Using patch occupancy models to estimate area of crevice-nesting seabird colonies

Authors: Reynolds, Joel H., and Renner, Heather M.

COMMENTARY

Using patch occupancy models to estimate area of crevice-nesting seabird colonies

Joel H. Reynolds[1,a]* and Heather M. Renner[2]

[1] Division of Natural Resources, National Wildlife Refuge System, Anchorage, Alaska, USA
[2] Alaska Maritime National Wildlife Refuge, Homer, Alaska, USA
[a] Current address: Western Alaska Landscape Conservation Cooperative, Anchorage, Alaska, USA
* Corresponding author: joel_reynolds@fws.gov

ABSTRACT

Crevice-nesting seabirds are notoriously difficult to monitor. We present a survey design and analysis that estimates both colony area and geographic extent, using indirect evidence to determine whether a cell is "occupied." The approach is to define a grid of cells across potential habitat and randomly sample small plots within each cell, surveying for signs of occupancy. Visiting ≥ 1 plot cell^{-1} provides a basis for mapping geographic extent. Occupancy models are used to estimate colony area, probability of detection for an occupied cell, and standard errors for all estimated parameters (allowing for statistical comparisons across surveys or colonies). We estimated the area of a colony of Least Auklets (*Aethia pusilla*) and Crested Auklets (*A. cristatella*) on Segula Island, Aleutian Archipelago, Alaska, in 2006, and use this as an example of how to adapt the survey design to the logistical constraints common in seabird colony surveys. Surveying only a handful of sample plots of ~ 16 m^2 in each $\sim 2,500$-m^2 cell in the grid was adequate to estimate the detection bias from spatial subsampling, correcting a $>50\%$ underestimate of colony area due to plots without evidence having been interpreted as unoccupied cells.

Keywords: Aethia cristatella, *Aethia pusilla*, area estimation, Crested Auklet, detection bias, Least Auklet, monitoring

Uso de modelos de ocupación de parches para estimar el área de las colonias de aves marinas que anidan en grietas

RESUMEN

Las aves marinas que anidan en grietas son muy difíciles de monitorear. Presentamos un diseño de estudio y análisis que estima tanto el área de la colonia como la extensión geográfica, usando evidencia indirecta para determinar si una celda esta "ocupada." El enfoque es definir una grilla de celdas a través del hábitat potencial y muestrear al azar pequeñas parcelas dentro de cada celda, evaluando signos de ocupación. La visita de ≥ 1 parcela por celda^{-1} brinda una base para mapear la extensión geográfica. Los modelos de ocupación son usados para estimar el área de la colonia, la probabilidad de detección de una celda ocupada y los errores estándar para todos los parámetros estimados (permitiendo comparaciones estadísticas a través de los muestreos o las colonias). Estimamos el área de una colonia de *Aethia pusilla* y de *A. cristatella* en la Isla Segula, Archipiélago Aleutiano, Alaska, en 2006, y usamos esto como un ejemplo de cómo adaptar el diseño de muestreo a las limitantes lógicas comunes en los muestreos de colonias de aves marinas. El estudio de solo un puñado de parcelas de muestreo de ~ 16 m^2 en cada celda de $\sim 2,500$ m^2 de la grilla resultó adecuado para estimar el sesgo de detección del submuestreo espacial, corrigiendo una subestimación de $>50\%$ del área de la colonia debido a parcelas sin evidencia que habían sido interpretadas como celdas desocupadas.

Palabras clave: Aethia cristatella, *Aethia pusilla*, estimación de área, monitoreo, sesgo de detección

INTRODUCTION

Least Auklets (*Aethia pusilla*) and Crested Auklets (*A. cristatella*) are an integral part of the Bering Sea ecosystem; together they are likely the most abundant seabirds breeding in the ecosystem (Springer and Roseneau 1985, Stephensen and Irons 2003). Efforts to improve monitoring of these auklets have been under way since the 1960s (summarized in Renner et al. 2006, 2011). As specialist predators of small zooplankton, in particular copepods and euphausiids (Hunt et al. 1998), Least and Crested auklets have long been viewed as valuable indicators of changes in Bering Sea trophic systems (Springer et al. 2007). This ecosystem changes dramatically

as a result of decadal-scale climate oscillations (Hunt et al. 2002) and is expected to be heavily affected by changes in global climate (Niebauer 1998) that are already being observed, such as loss of sea ice and increasing ocean acidification (Perovich et al. 2012, Mathis 2011). Crevice-nesting auklets may also be particularly susceptible to threats such as oil spills (Piatt et al. 1990), introduction of mammalian predators (Williams et al. 2003), and volcanic explosions (Williams et al. 2010), given that the entire population is concentrated in a few large colonies for breeding.

Unfortunately, as with all crevice-nesting seabirds, monitoring populations of Least and Crested auklets is challenging. Although these auklets annually aggregate into a few large breeding colonies (Jones 1993a, 1993b), the colonies occur on remote, isolated islands in the Aleutians and the Bering Sea, where access is difficult and expensive. Population estimates for any auklet breeding colony have been few and largely conjectural, sometimes varying by an order of magnitude among observers (Shuntov 1999). Furthermore, their nests are underground, which prevents direct observation, so counts are of birds socializing on the surface (see Renner et al. 2011). In the case of auklets, their sheer abundance and variable surface attendance have foiled efforts using photography, video, and various double-count methods (for a summary, see Renner et al. 2006).

These challenges have stymied traditional population monitoring techniques, so in 2006 we proposed documenting changes in colony extent and density instead of bird abundance (Renner et al. 2006). Unlike the numbers of auklets arriving, departing, or socializing on the colony surface at any given time, colony area is more temporally stable across the course of the nesting season and, thus, much simpler to measure and/or estimate with limited monitoring resources. Developing a monitoring program around measures of colony area requires assuming that changes in colony area over the time scales of interest (e.g., decades) reflect changes in the breeding population. We used three metrics for monitoring colony change: (1) a statistical estimate of total area occupied by the colony; (2) a map of the colony's geographic distribution and extent; and (3) relative nesting density information at known locations in the colony. The estimate of total area provides a basis for statistical estimation of the magnitude of changes through time, tracking net changes in colony size. The map of known occupied sites allows for tracking gross changes in spatial occupancy (e.g., colony distribution and extent), thus providing a basis for insight into the underlying mechanisms of change.

A useful way to measure colony area is by documenting locations where indirect evidence of nesting is present, such as belly feathers, guano, or worn vegetation near a crevice entrance (Renner et al. 2006). This allows for simpler measurement protocols than traditional repeated counts of numbers of auklets socializing on the colony surface (for an assessment of that metric, see Renner et al. 2011) and, thus, provides sample coverage across the full colony in substantially less time. Importantly, the density of this indirect evidence is strongly correlated with counts of socializing auklets (Renner et al. 2006).

Because auklet nesting colonies cover large areas (up to multiple square kilometers) in relation to the limited personnel and time available for monitoring, a colony area survey includes defining a regular grid of cells over a region encompassing the colony (grid cells are usually about 50–100 m on a side; see below), randomly selecting small sample plots within each grid cell (on the order of $\sim$16 m^2), and thoroughly searching them for indirect physical evidence of auklet occupancy (i.e. it relies on spatial subsampling; e.g., Renner et al. 2006). Searching just 1 plot cell^{-1} results in "unoccupied" classifications for both (1) grid cells unoccupied by nesting auklets and (2) those occupied by nesting auklets but for which no indirect evidence of nesting was detected in the small sample plot, thus underestimating the total number of occupied cells. Searching multiple plots in each grid cell, or at least in a randomly selected subset of grid cells, allows application of occupancy models, which account for detection probablility (MacKenzie et al. 2006).

In their usual application, occupancy models are fitted to data derived from multiple visits in time to a subset of sites (hereafter "cells"). Here, we adapt the concept to data derived from surveying multiple plots in a cell, substituting multiple locations in space for multiple visits in time. An occupancy model allows one to survey just a small portion of the cell (e.g., a plot) and derive estimates of both the percentage of cells occupied by nesting birds and the probability of detecting evidence of occupancy in an occupied cell; the models thus account for cells that were occupied by nesting birds but for which evidence was not detected.

We present a survey design and analyses for estimating crevice-nesting colony area while explicitly accounting for imperfect detection. The method is illustrated in application to mapping the colony of Least and Crested auklets on Segula Island, Aleutian Archipelago, Alaska, USA.

METHODS

The survey fulfills two goals, providing (1) a statistical estimate of colony area and (2) a map, with known scale, of cells known to be occupied. The former allows for assessments of net change through time, the latter for assessments of gross change in colony occupancy and insight into the types of changes that have occurred. Both are important monitoring objectives and, like all successful monitoring programs, must be approached using carefully

prescribed methods that are repeatable and reliable (Fuller 1999). The survey design balances the requirements of these two objectives, with a systematic-sampling aspect for meeting the mapping objective and a repeated-spatial-subsampling aspect for meeting the occupancy modeling objective. The result is a "double sampling" design (e.g., MacKenzie and Royle 2005). Stages 1 to 3 (below) were developed from the initial work of Renner et al. (2006).

Survey Design Overview

The survey design relies on developing a uniform grid of cells (ignoring surface topology) over a region encompassing the colony, then dividing each cell into a uniform grid of smaller sample units. The set of sample units available for selection and measurement defines the "sample frame" (de Gruijter et al. 2006). Usually the sample frame will not be fully determined until the field crew arrives at the colony site and establishes the current colony perimeter. Therefore, the survey design must include stages for establishing and refining the sample frame.

Developing the colony map requires sampling and measuring ≥ 1 plot cell^{-1}. Fitting the occupancy models requires sampling and measuring ≥ 2 plots in at least a subset of grid cells (MacKenzie et al. 2006), although generally the precision of the results is greatly improved with ≥ 3 plots (MacKenzie and Royle 2005). Usually the time required to locate and measure a sample unit at the colony will not be known, because of topography, vegetation type and density, and other factors, until the field crew establishes those rates through initial work at the colony site during the current visit (see "Stage 3" below). Therefore, the survey design must provide the flexibility to adjust sample sizes in the field to the logistical realities encountered. This is achieved by conducting two rounds of measurements: (1) measuring a randomly selected plot from each cell and (2) measuring an additional random selection of multiple plots from at least a random sample of cells (ideally, if time permits, all cells, as a means of improving the final colony map). The survey stages summarized below incorporate this flexibility and assume that observers have only general information on the colony location prior to arriving at the field site. They also allow for use of a protocol whereby only a portion of the sample unit (the "observational unit") is actually measured for indirect evidence of nesting, as illustrated in the application.

Stage 1: Develop preliminary sample frame. Before traveling to the field site, review aerial photographs and/or previous colony mapping efforts to identify a region likely to encompass all the potential nesting habitat in and surrounding the colony. Nonhabitat areas do not provide any crevice openings and consist of some combination of standing water, permanent snow or ice fields, or grass-covered terrain that lacks rock crevices (Renner et al.

2006). Overlay a uniform grid of cells across a map of this region. Define the size and shape of a cell and a sample unit so that a cell is evenly divisible into a two-dimensional array of sample units.

Stage 2: Refine the sample frame to encompass current colony perimeter. Survey the apparent edges of the colony on foot and identify the approximate colony perimeter using auditory and visual cues of occupancy and marking waypoints on a global positioning system. Cues include surface and subsurface vocalizations, birds flying in and out of crevices or socializing on the surface, or indirect evidence of occupancy (i.e. guano, feathers, and/or worn vegetation near crevice openings). In field camp, refine the sample frame to include only cells crossed by or contained within the approximate perimeter or within a boundary (1–2 cells thick) of potential habitat around the approximate perimeter.

Stage 3: Measurement, round 1. Randomly select 1 sample unit cell^{-1} in the refined sample frame. Search the observational unit centered on each selected sample unit for any evidence of auklet presence: belly feathers, droppings, vocalizations, vegetation wear, or birds attending on the surface of the plot (Renner et al. 2006). A circular observational unit of radius 2.25 m centered on the sample unit is easily implemented in the field and is a reasonable size to thoroughly search (quarter by quarter) in a brief period (Renner et al. 2006). Classify each searched sample unit as "present" (occupied; evidence detected); "absent" (no evidence detected, possibly because the sample unit was not potential habitat but rather a snow field, water, a grass plain with no visible crevices, etc.); or "missing" (inaccessible because located on a cliff face or inaccessible beach segment). Small patches of nonhabitat can occur within a cell classified as potential habitat.

Locate the cell's center for orientation and decide whether the whole cell is nonhabitat or contains at least some potential habitat. We defined "nonhabitat cells" as those either without any crevice openings or both (1) composed of <1% (surface area) of crevice-containing habitat and (2) isolated or discontinuous with the main colony. Potential habitat cells contain crevice openings. Record whether the whole cell is nonhabitat. Record the time required to complete this stage, which will be used to determine final sample sizes in stage 5.

Stage 4: Further refine the sample frame. Remove from the sample frame all cells classified as nonhabitat in stage 3.

Stage 5: Plan and conduct measurement, round 2. There is a tradeoff between allocating the remaining time in the field and searching more sample units in fewer cells or fewer sample units in more cells. Simulations under simple scenarios reveal that, in general, statistical efficiency requires ≥ 3 sample units cell^{-1} (for those cells with

TABLE 1. The assumptions of the standard occupancy model (MacKenzie et al. 2006).

1. A cell's occupancy status does not change during the period of sampling.
2. Detection probability is constant across all cells (see below for extensions).
3. Detection of evidence at a cell is independent of detection at any other cell.
4. Detection of evidence in a sample unit is independent of detection at any other sample unit within the cell.
5. Probability of occupancy is constant across the cells.
6. T randomly selected sample units are observed in each cell.

multiple sample units searched) when probability of detection is >0.5, and that the optimal number of sample units per cell increases as detection probability decreases and probability of occupancy increases (MacKenzie and Royle 2005).

Using the time required to complete stage 3, calculate how many cells, randomly selected from the revised sample frame, can be visited in the remaining field time, such that a simple random sample of 2 to 4 additional sample units can be measured within each cell. Randomly select the sample of cells and their sample units for measurement in round 2. Search the observational unit centered on each selected sample unit for any evidence of auklet presence as described above, classifying each searched sample unit as present, absent, or missing. Note that although spatial subsampling without replacement may generate bias in the occupancy estimate (Kendall and White 2009), the practical impact will be negligible in the application under consideration because the percentage of the sample units in the cell that are actually surveyed is expected to be much less than 10% (Guillera-Arroita 2011).

Statistical Analysis

Occupancy models are thoroughly described by Mac-Kenzie et al. (2006). Under the standard assumptions (Table 1), the model's likelihood is

$$L(\Psi, p \mid \{n_i\}) = \left[\Psi^{n.} \prod_{sites\ i\ with\ detects} \binom{T}{n_i} p^{n_i} (1-p)^{T-n_i} \right] \times \left(\Psi(1-p)^T + (1-\Psi) \right)^{N-n.}$$

where Ψ is the probability that the species occupies a randomly chosen cell; p is the probability of detecting evidence in a sample unit, given that the cell is occupied; N is the total number of surveyed cells; T is the number of independently randomly selected sample units searched in each cell; n_i is the total number of sample units in cell i containing evidence ($n_i = 0, 1, \ldots, T$); and $n.$ is the total number of cells in which evidence was detected. Mac-

Kenzie et al. (2006) extended this model in a variety of ways, including using differing numbers of sample units searched per cell and using mixture models to accommodate unmodeled heterogeneity among cells in the probability of detection (Pledger 2000).

We fit the basic model and at least the two-point mixture model using maximum likelihood methods and select the best-fitting model using Akaike's Information Criterion (AIC; Burnham and Anderson 2002). Goodness-of-fit for the selected model is assessed using a chi-square goodness-of-fit statistic with the null reference distribution estimated via Monte Carlo simulation (1,000 simulated samples) from the fitted model (MacKenzie and Bailey 2004). Standard errors and 95% confidence intervals (CI) are estimated from 1,000 nonparametric bootstrap resamples (Lunneborg 2000). Analyses were conducted using the program PRESENCE (see Acknowledgments) and in R using code from the first author (R Development Core Team 2010).

RESULTS

Study Site

The Gula Point auklet colony on the north end of Segula Island (52°02′00′′N, 178°09′00′′E) was surveyed (using the methods described above) between May 24 and June 6, 2006. This was assumed to be the early incubation period, on the basis of the timing at nearby colonies where nests were directly monitored. Weather in this period was dry with no rain.

Stage 1. The observational unit was a circular plot with a 2.25-m radius (area 16 m^2); the sample unit was a hexagon with a 2.26-m radius (area 17.58 m^2) circumscribing the observational unit (note that the observation unit is slightly smaller than the sample unit; Figure 1). An array of sample units was overlain across the region encompassing the best understanding of the colony's location prior to arrival at the colony, then partitioned into a grid of roughly square cells, each a 26 row × 24 column array of sample units, approximately 100 × 100 m (area = 10,969.92 m^2; Figure 1). A grid of ~400 such cells was defined as the preliminary sample frame before arriving at the field site (Renner and Reynolds 2006).

Stage 2. The preliminary sample frame was refined to a contiguous set of 308 cells of potential habitat (1 day for a team of 2 people).

Stages 3 and 4. Of the 308 cells, the stage-3 search identified 92 as nonhabitat, leaving 216 cells in the final sample frame (stage 3: 6 days, 2 people working independently; stage 4: 1 day). In 6 of the cells, the randomly selected sample unit was inaccessible because of cliffs or other barriers; a randomly selected sample unit was searched in each of the other 210 cells.
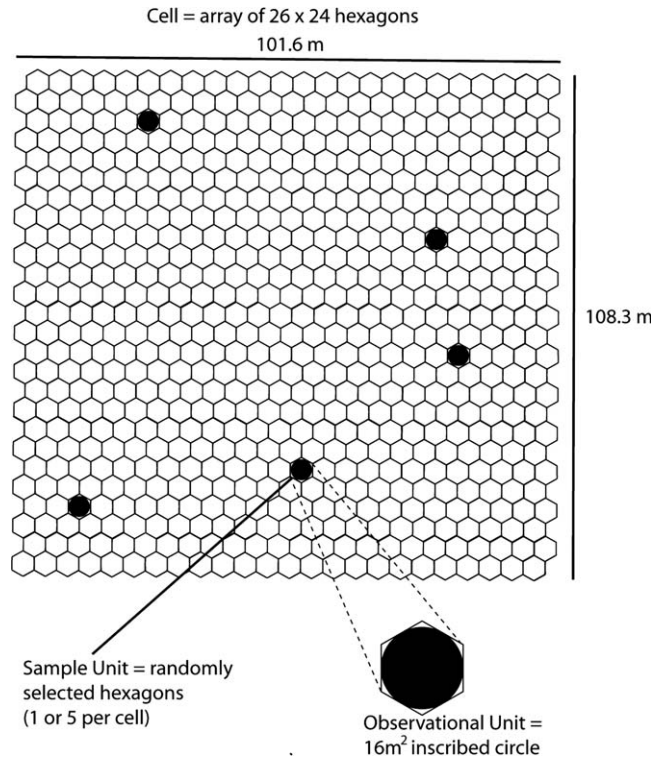
Cell = array of 26 × 24 hexagons
101.6 m

FIGURE 1. Diagram of sampling design used for documenting occupancy by auklets at Segula Island, Alaska, in 2006.

Stage 5. We chose a goal of searching 4 additional randomly selected sample units in a random sample of cells and estimated, on the basis of time records from implementing stage 3, that we could visit ~82 cells in the remaining 4 days of the field visit. We randomly selected 80 cells from the 216 in the sample frame and surveyed the additional sample units in each.

Colony Survey Results

Searching just a single sample unit per cell (i.e. the stage-3 survey) identified signs of auklet occupancy in 76 cells (Table 2). Following the full method, stages 3 and 5 identified a total of 18 more cells with signs of auklet occupancy, further refining the map (Table 2). The 52 cells in which multiple sample units were searched and in which

TABLE 2. Summary data from surveying the region of the Segula Island auklet colony for evidence of occupancy as described in the text.

	Sample units searched per cell		
	1	5	Subtotal
Inaccessible	6		6
No evidence	88	28	116
Evidence	42	52	94
Total cells	136	80	216

TABLE 3. Model selection results from fitting occupancy models for the 216 potential habitat cells; inaccessible sample units were treated as missing data. Detection probabilities were modeled as constant (model 1) or as heterogeneous and represented a mixture of 2 or 3 components (models 2 and 3, respectively). The data are best described by the model allowing for a two-component mixture model of detection probabilities.

Model of probability of detection	Delta AIC [a]	Model weight	Goodness-of-fit
1 (constant)	23.2	0.00	0.00
2	0.0	0.88	0.22
3	4.0	0.12	0.03

[a] Lowest value of AIC = 600.83.

evidence of occupancy was found (Table 2) consisted of 50 cells in which 5 sample units were searched in each cell (18 cells exhibited evidence in 1 sample unit, 10 in 2 sample units, 7 in 3 sample units, 8 in 4 sample units, and 7 in all 5 sample units); 1 cell in which only 4 sample units were searched (1 was inaccessible), 3 of which exhibited evidence; and 1 cell in which only 3 sample units were searched (2 were inaccessible), 1 of which exhibited evidence. A map of the study colony depicting the 216 cells in the refined sample frame and the results of the searches is available online in Renner and Reynolds (2006).

Analysis Results

The basic, two-point mixture, and three-point mixture occupancy models were fitted to the summary results from the revised survey method; data strongly supported the two-point mixture model of heterogeneity in probability of detection among cells (Table 3), likely reflecting habitat quality and auklet abundance within a cell. The goodness-of-fit assessment did not detect any significant departure of the observations from the selected model (Table 3). The majority of the occupied colony cells were attributed to the mixture component with lower probability of detection ($\pi_{Low} = 68\%$ of occupied colony; Table 4). The basic model, which did not allow for any heterogeneity in detection probabilities, estimated occupancy as 68% rather than 83% from the final model (Table 4), implying an unaccounted-for heterogeneity in detection probabilities among sample units.

TABLE 4. Parameter estimates from the occupancy model with probability of detection modeled as a two-component mixture (Table 3, model 2), with bootstrap SE in parentheses. Prob(Occupied), or Ψ, is the probability that a randomly selected cell in the sample frame is occupied. π_{Low} is the probability that a randomly selected occupied cell has low detection probability.

Prob(Occupied)	0.83 (0.11)
π_{Low}	0.68 (0.11)
Prob(Detect$_{Low}$)	0.20 (0.07)
Prob(Detect$_{High}$)	0.79 (0.07)

Of the sample frame's 216 cells, 83% were estimated to be occupied (Table 3), for an estimated colony area of 1.96 km² (95% CI: 1.47–2.37 km²). Ignoring imperfect detection and stopping the survey at the end of stage 3 (following Renner et al. 2006) would give naive estimates of 35% occupancy and a colony area of 0.83 km², with no estimate of the detection probability or each of the associated uncertainties.

DISCUSSION

The survey method and analyses presented here provide a standardized, repeatable survey to both statistically estimate colony area (using occupancy models) and produce a colony map with an explicit scale, thereby advancing our ability to understand the status and trends of breeding colonies of crevice-nesting seabirds. The method meets the logistical constraints of surveying (often quite large) colonies on remote islands by surveying only a small portion (1 or a few sample units) of each cell for evidence of auklet occupancy. In the context of the Segula Island colony, this means that a cell of ~11,000 m² is determined to be occupied, or not, on the basis of whether evidence is detected in the handful of 16-m² sample units searched in the cell (Table 2). The high potential for failure to detect that a cell is occupied when only 0.15% of it (1 sample unit) is searched creates a high potential for severe underestimation of the number of occupied cells and, thus, of colony area. This is eliminated, as demonstrated here, by surveying a random sample of additional plots in at least a random subset of the cells, so as to directly estimate both the colony area occupied and the detection probability.

Colony area can be severely underestimated if detection bias is ignored: The estimated colony area for Segula Island that resulted from surveying just 1 plot cell⁻¹ (the method of Renner et al. 2006) was only 44% of the estimate when detection bias was accounted for. The magnitude of the probability of detection will vary among colonies, with colonies qualitatively judged to be "low-density" on the basis of surface attendance, like Segula, presumably having much lower detection probabilities than "higher-density" colonies like that at Ulakaia Ridge, St. George Island, the Pribilofs. Yet detection bias will still greatly affect a colony area estimate even in such "high-density" colonies as Ulakaia Ridge. For example, a 2004 survey of the Ulakaia Ridge colony surveyed just 1 sample unit cell⁻¹ (Renner et al. 2006), effectively stopping after stage 3 of the occupancy method; so the colony was revisited in 2006 and stages 4 and 5 of the method were applied, resulting in 4 additional sample units in approximately one-third of the cells (for details of the methods and results, see Appendix). Accounting for detection bias provided a colony area estimate that was 77% larger than when the bias was ignored (Appendix, Table A3). Clearly, detection bias is a substantial issue even in this relatively high-density colony. Logistics and time available for fieldwork constrain our ability to eliminate the subsampling bias by using smaller cells (more to visit) or larger sample units (more to search) to guarantee detection of occupied cells.

Bias from imperfect detection will vary among surveys of the same colony at different points of the season (as evidence accumulates or is lost), in different years, and among colonies, owing to differences in observers, survey effort, field conditions, colony breeding activity, habitat features, colony density, and so on. Timing in relation to colony occupancy should have little effect on the colony area estimates if (1) surveys are conducted after the population has been at the colony for ≥2 wk (so that there is some non-negligible probability of detection) and (2) the field surveys are conducted over a short enough period that probability of detection remains relatively constant within that period (we recommend 2 wk, though this has not been directly tested).

As seen at Ulakaia, detection bias cannot be ignored even at a high-density colony. Further, the large potential for changes in density to accompany changes in colony area implies that failing to account for detection at every survey can severely confound estimates of temporal change at a colony or of differences between colonies, undermining any potential insight from colony monitoring. Eliminating such confounding is essential to achieving the goal, articulated in Renner et al. (2006), of informative regional analyses of an auklet colony monitoring network using colony area as a metric. Occupancy models allow for assessment of net change in colony area through fitting a multiple-season patch occupancy model with a trend in the occupancy parameter (MacKenzie et al. 2006).

By accounting for imperfect detection, occupancy models also suggest two potentially informative metrics for monitoring colony change: the magnitude and structure of the detection probability. Colony-wide changes in breeding density of sufficient magnitude should result in changes in the magnitude of the detection probability, all other factors being held constant. Similarly, a shift from a fairly spatially uniform colony density to more heterogeneous density should result in a shift from a simple detection model (e.g., colony-wide average probability of detection) to a more heterogeneous detection model (e.g., ≥2 mixture models), such as seen at both Segula (Table 3) and St. George (Appendix, Table A2).

Although occupancy models estimate the total number of occupied cells, they do not distinguish which cells were occupied among those in which no evidence of occupancy was seen. Thus, the mapping goal of the survey requires the systematic coverage provided by visiting ≥1 plot in each cell. Probability of occupancy could be estimated for those cells whose occupancy status remains unknown if covariates related to occupancy status were collected at the cell scale;

this could also improve the precision of the occupancy (and, hence, area) estimates. For the Segula colony, data layers such as dominant gradient, substrate, and cover classes might prove useful (if they were available).

Some of the occupied cells could be identified by spending a few additional minutes in each cell during stage 3 of the survey to quickly assess occupancy of the most likely habitat within the cell (not just in the randomly selected sample unit). This "expert search" would suffer from unknown probability of detection and could not contribute to the statistical estimate of the colony area or the probability of detection, but the results could inform the colony map of occupied cells for those cells where no evidence of occupancy is detected in the randomly selected sample units. A combination of both approaches would best satisfy the dual objectives of developing an explicit georeferenced map of known occupied cells and an unbiased colony area estimate with associated standard error. However, this additional search effort should be incorporated only if the additional time required does not jeopardize completion of all stages of the survey.

By also estimating the uncertainty of the colony area estimate, occupancy models provide a basis for statistical assessments of differences in colony area; uncertainty estimates are required to determine statistical differences in observed area estimates. This is a fundamental limitation of past estimates of colony area. Although the area of the Segula auklet colony was estimated by Early et al. (1980) as 0.78 km^2, with "60% unused habitat," and by Thomson (1995) as 1.28 km^2, no formal quantitative assessment of change in colony area is possible because neither of those surveys employed a well-documented, repeatable, statistically valid survey method. Unfortunately, almost no insight can be gained about changes in colony area over the intervening 30 yr, let alone causes of changes, because the lack of documentation and rigorous methods prevents even a qualitative assessment of bias and uncertainty for the earlier estimates. For example, how do the surveys differ in their underlying definitions of occupancy, in detection probabilities, in measurement scale or counting unit (Renner et al. 2006), or in the treatment of uninhabited habitat within the colony perimeter?

Survey Design Considerations

The occupancy method requires planning and forethought before visiting the field site. Surveying multiple sample units from at least some cells requires that one carefully define a shape for the sample unit that allows the cell to be completely partitioned into identically shaped sample units that fit evenly into the cell. The easiest way to do this, ignoring surface topography, is to define square or rectilinear cells and then partition them into an array of square sample units. Alternatively, one can construct cells "from the bottom up" by starting with square or hexagonal

sample units, then defining the cell as an array of sample units. The Segula survey used hexagonal sample units to maximize the spatial correspondence between the sample unit (hexagon) and observational unit (circle) and, thus, minimize this aspect of spatial subsampling in the probability of detection (Figure 1). Although the circular observation unit was slightly smaller than the hexagonal sample unit (Figure 1), we believe that any bias related to coverage was negligible because evidence of occupancy in just the border region was unlikely; Renner et al. (2006) found that evidence was fairly uniform at the scale of the observation unit. Given that patch occupancy models simply incorporate such subsampling into the overall probability of detection, eliminating any gain from using hexagonal sample units, we recommend that future surveys use square sample units to simplify the survey design and its description and retain the circular observational unit for ease of implementation in the field. Incorporating a digital elevation model into the survey design would improve the accuracy of the area estimate as well as improve comparability across colonies with substantial differences in surface topography (Renner et al. 2006).

This spatial subsampling design relies on "sampling without replacement" because occupancy is determined on the basis of physical evidence that, barring a storm during the survey, it is reasonable to assume is temporally stable over the duration of the field session. Thus, there is no gain from revisiting the same plot ("temporal replication"), in contrast to more common applications of occupancy models.

In this application, the basic occupancy model's assumption that probability of detection is constant across occupied cells (MacKenzie et al. 2006) will seldom be tenable. Auklet colonies are patchy at the scale of this survey's cells ($\sim$11,000 m^2 cell^{-1}), hence the strong support for the two-point mixture model to capture the heterogeneity across cells in probability of detection (Table 2). This suggests the value of incorporating survey plot-scale covariates associated with detection, such as land cover, in the design.

By eliminating detection bias, the occupancy method survey and analyses achieve the dual objectives of having a statistically rigorous estimate of colony area and a map of known occupied colony cells. Although future work will undoubtedly improve upon these foundations, they provide a means for statistically sound monitoring of a network of colonies of crevice-nesting seabirds.

LITERATURE CITED

Burnham, K. P., and D. R. Anderson (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, second edition. Springer, New York, NY, USA.

de Gruijter, J. J., D. J. Brus, M. F. P. Bierkens, and M. Knotters (2006). Sampling for Natural Resource Monitoring. Springer, New York, NY, USA.

Early, T., J. Beall, W. Henry, and A. Taber (1980). Results of bird and mammal surveys of the western Aleutian Islands, summer 1979. U.S. Fish and Wildlife Service, Adak, AK, USA.

Fuller, W. A. (1999). Environmental surveys over time. Journal of Agricultural, Biological, and Environmental Statistics 4:331–345.

Guillera-Arroita, G. (2011). Impact of sampling with replacement in occupancy studies with spatial replication. Methods in Ecology and Evolution 2:401–406.

Hunt, G. L., Jr., R. W. Russell, K. O. Coyle, and T. Weingartner (1998). Comparative foraging ecology of planktivorous auklets in relation to ocean physics and prey availability. Marine Ecology Progress Series 167:241–259.

Hunt, G. L., Jr., P. Stabeno, G. Walters, E. Sinclair, R. D. Brodeur, J. M. Napp, and N. A. Bond (2002). Climate change and control of the southeastern Bering Sea pelagic ecosystem. Deep-Sea Research Part II: Tropical Studies in Oceanography 49:5821–5853.

Jones, I. L. (1993a). Crested Auklet (*Aethia cristatella*). In The Birds of North America, no. 70 (A. Poole and F. Gill, Editors). Academy of Natural Sciences, Philadelphia, PA, and American Ornithologists' Union, Washington, DC, USA.

Jones, I. L. (1993b). Least Auklet (*Aethia pusilla*). In The Birds of North America, no. 69 (A. Poole and F. Gill, Editors). Academy of Natural Sciences, Philadelphia, PA, and American Ornithologists' Union, Washington, DC, USA.

Kendall, W. L., and G. C. White (2009). A cautionary note on substituting spatial subunits for repeated temporal sampling in studies of site occupancy. Journal of Applied Ecology 46:1182–1188.

Klostermann, M. R., and B. Drummond (2012). Biological monitoring at St. George Island, Alaska 2012. U.S. Fish and Wildlife Service Report AMNWR 2012/08.

Lunneborg, C. E. (2000). Data Analysis by Resampling: Concepts and Applications. Duxbury Press, Pacific Grove, CA, USA.

MacKenzie, D. I., and L. L. Bailey (2004). Assessing the fit of site-occupancy models. Journal of Agricultural, Biological, and Environmental Statistics 9:300–318.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines (2006). Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence. Elsevier, San Diego, CA, USA.

MacKenzie, D. I., and J. A. Royle (2005). Designing occupancy studies: General advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

Mathis, J. T. (2011). The extent and controls on ocean acidification in the western Arctic Ocean and adjacent continental shelf areas. Arctic Report Card update for 2011. http://www.arctic.noaa.gov/report11/ocean_acidification.html

Niebauer, H. J. (1998). Variability in Bering Sea ice cover as affected by a regime shift in the North Pacific in the period 1947–1996. Journal of Geophysical Research 103:27717–27737.

Perovich, D., W. Meier, M. Tschudi, S. Gerland, and J. Richter-Menge (2012). Sea Ice. Arctic Report Card Update for 2012. http://www.arctic.noaa.gov/report12/sea_ice.html

Piatt, J. F., C. J. Lensink, W. Butler, M. Kendziorek, and D. R. Nysewander (1990). Immediate impact of the 'Exxon Valdez' oil spill on marine birds. The Auk 107:387–397.

Pledger, S. (2000). Unified maximum likelihood estimates for closed capture–recapture models using mixtures. Biometrics 56:434–442.

R Development Core Team (2010). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Renner, H. M., and J. H. Reynolds (2006). Mapping distribution and relative density of nesting Least and Crested auklets at Segula Island, Alaska, May–June 2006. U.S. Fish and Wildlife Service Report AMNWR 06/08.

Renner, H. M., M. Renner, J. H. Reynolds, A. M. A. Harding, I. L. Jones, D. B. Irons, and G. V. Byrd (2006). Colony mapping: A new technique for monitoring crevice-nesting seabirds. The Condor 108:423–434.

Renner, H. M., J. H. Reynolds, M. Sims, and M. Renner (2011). Evaluating the power of surface attendance counts to detect long-term trends in populations of crevice-nesting auklets. Environmental Monitoring and Assessment 177:665–679.

Shuntov, V. P. (1999). Seabirds of the western Bering Sea. In Dynamics of the Bering Sea (T. R. Loughlin and K. Ohtani, Editors). University of Alaska Sea Grant, Fairbanks, AK, USA. pp. 651–682.

Springer, A. M., G. V. Byrd, and S. J. Iverson (2007). Hot oceanography: Planktivorous seabirds reveal ecosystem responses to warming of the Bering Sea. Marine Ecology Progress Series 352:289–297.

Springer, A. M., and D. G. Roseneau (1985). Copepod-based food webs: Auklets and oceanography in the Bering Sea. Marine Ecology Progress Series 21:229–237.

Stephensen, S. W., and D. B. Irons (2003). Comparison of colonial breeding seabirds in the eastern Bering Sea and Gulf of Alaska. Marine Ornithology 31:167–173.

Thomson, G. (1995). Counts of Least and Crested auklets following fox removal at Segula Island, Alaska, in 1995. U.S. Fish and Wildlife Service Report AMNWR 95/05.

Williams, J. C., G. V. Byrd, and N. B. Konyukhov (2003). Whiskered auklets *Aethia pygmaea*, foxes, humans and how to right a wrong. Marine Ornithology 31:175–180.

Williams, J. C., B. A. Drummond, and R. T. Buxton (2010). Initial effects of the August 2008 volcanic eruption on breeding birds and marine mammals at Kasatochi Island, Alaska. Arctic, Antarctic and Alpine Research 42:306–314.

APPENDIX

Accounting for Imperfect Detection in the Colony Area Estimate of the Ulakaia Ridge Auklet Colony, St. George Island, the Pribilofs

The use of occupancy models to estimate auklet colony area grew from a 2004 survey of Ulakaia Ridge, St. George Island, the Pribilofs, detailed in Renner et al. (2006). Because that effort did not sample multiple sample units in each cell, it did not allow proper accounting for imperfect detection. The colony was revisited during the 2006 field season and stages 4 and 5 of the revised method were implemented. The combined 2004–2006 dataset was analyzed using the methods described in the main text. Distinct features of this effort and the results are summarized below.

Survey methods. The sample frame of feasible auklet habitat, defined from the 2004 survey, consisted of 147 square cells, each 2,500 m^2. In each cell, in 2004, a single circular observation unit (radius 2.25 m) was searched for evidence of auklet occupancy (Renner et al. 2006). In 2006, 59 of these cells were randomly selected and, in each, 4 more sample units were randomly selected and searched for evidence of auklet occupancy.

Analysis methods. The 2004 and 2006 results were combined for analysis using the models, model fitting, model selection, and colony area, standard error, and confidence interval estimation methods described in the main text.

Results. In total, the surveys revealed evidence of auklet occupancy in 77 of the 147 cells (Table A1). The 40 cells in which 5 sample units were searched and in which evidence of occupancy was found (Table A1) consisted of 8 cells where evidence was detected in 1 sample unit, 7 where evidence was detected in 2 sample units, 7 where evidence was detected in 3 sample units, 11 where evidence was detected in 4 sample units, and 7 where evidence was detected in all 5 sample units.

The two-component mixture model for probability of detection was identified as the best description of the colony data (Table A2), with approximately half the colony

TABLE A2. Model selection results from fitting occupancy models for the 147 potential habitat cells. Detection probabilities were modeled as constant (model 1) or as heterogeneous and represented a mixture of 2 or 3 components (models 2 and 3, respectively). The data are best described by the model allowing for a two-component mixture model of detection probabilities.

Model of probability of detection	ΔAIC [a]	Model weight	Goodness-of-fit
1	9.69	0.01	0.00
2	0.0	0.87	0.22
3	4.00	0.12	0.03

[a] Smallest AIC = 454.97.

estimated as having "high probability of detection" (mean ± SE = 0.78 ± 0.07; Table A3). The number of cells occupied in the colony was estimated as 108 (95% CI: 96–147), corresponding to a colony area estimate of 0.27 km^2 (95% CI: 0.24–0.37 km^2). The estimate from the method of Renner et al. (2006) was 61 cells occupied and a colony area of 0.15 km^2; the method does not provide standard error estimates.

Discussion. Even at what is generally considered a "high-density" auklet colony, detection of evidence of occupancy is less than perfect and, thus, will cause underestimation of colony area unless directly accounted for (Table A3). Accounting for it led to a colony area estimate >75% larger than that originally reported by Renner et al. (2006). The revised method also provides an estimate of the uncertainty of the colony area estimate, providing a basis for future quantitative assessments of change in colony area. Combining the 2 datasets for analysis required assuming no substantial changes in colony occupancy and no probability of detection between the 2004 and 2006 surveys. The former assumption is tenable given the brief span of time and the lack of differences in the surface attendance counts during these 2 seasons (Klostermann and Drummond 2012). The choice of a small observation unit and season-long access to the field site helped minimize differences among observers in ability to detect auklet evidence.

TABLE A1. Summary data from surveying the region of the auklet colony at Ulakaia Hills, St. George, Pribilof Islands, for evidence of occupancy as described in the text.

	Sample units searched per cell		
	1	5	Subtotal
No evidence	51	19	70
Evidence	37	40	77
Total cells	88	59	147

TABLE A3. Parameter estimates from the occupancy model with two-point mixture model of probability of detection (Table A2, model 2), with bootstrap SE in parentheses. π_{Low} is the probability that a randomly selected occupied cell has low probability of detection.

Prob(Occupied)	0.76 (0.11)
π_{Low}	0.44 (0.11)
Prob(Detect$_{Low}$)	0.25 (0.12)
Prob(Detect$_{High}$)	0.78 (0.07)