

RESEARCH ARTICLE



Integrating distance sampling survey data with population indices to separate trends in abundance and temporary immigration

Joshua H. Schmidt¹ | Tammy L. Wilson² |
William L. Thompson³ | Buck A. Mangipane⁴

¹U.S. National Park Service, Central Alaska Network, 4175 Geist Road, Fairbanks, AK 99709, USA

²U.S. Geological Survey, Massachusetts Cooperative Fish and Wildlife Research Unit, Department of Environmental Conservation, University of Massachusetts, 160 Holdsworth Way, Amherst, MA 01003, USA

³U.S. National Park Service, 215 South Ferry Road, Narragansett, RI 02882, USA

⁴U.S. National Park Service, Lake Clark National Park and Preserve, General Delivery, Port Alsworth, AK 99653, USA

Correspondence

Joshua H. Schmidt, U.S. National Park Service, 4175 Geist Road, Fairbanks, AK 99709, USA.
Email: joshua_schmidt@nps.gov

Funding information

Funding for this project was provided by the Southwest Alaska Network, the National Park Service Natural Resource Preservation Program, Lake Clark National Park and Preserve, and the Alaska Department of Fish and Game.

Abstract

Managers rely on accurate estimators of wildlife abundance and trends for management decisions. Despite the focus of contemporary wildlife science on developing methods to improve inference from wildlife surveys, legacy datasets often rely on index counts that lack information about the detection process. Data integration can be a useful tool for combining index counts with data collected under more rigorous designs (i.e., designs that account for the detection process), but care is required when datasets represent different population processes or are mismatched in space and time. This can be particularly problematic in cases where animals aggregate in response to a spatially or temporally limited resource because individuals may temporarily immigrate from outside the study area and be included in the abundance index. Abundance indices based on brown bear (*Ursus arctos*) feeding aggregations within coastal meadows in early summer in Lake Clark National Park and Preserve, Alaska, USA, are one such example. These indices reflect the target population (brown bears residing within the park) and temporary immigrants (i.e., bears drawn from outside the park boundary). To properly account for the effects of temporary immigration, we integrated the index data with abundance data collected via park-wide distance sampling surveys, the latter of which properly addressed the detection process. By assuming that the distance data provide inference



on abundance and the index counts represent some combination of abundance and temporary immigration processes, we were able to decompose the relative contribution of each to overall trend. We estimated that the density of brown bears within our study area was 38–54 adults/1,000 km² during 2003–2019 and that abundance increased at a rate of approximately 1.4%/year. The contribution of temporary immigrants to overall trend in the index was low, so we created 3 hypothetical scenarios to more fully demonstrate how the integrated approach could be useful in situations where the composite trend in meadow counts may obscure trends in abundance (e.g., opposing trends in abundance and temporary immigration). Our work represents a conceptual advance supporting the integration of legacy index data with more rigorous data streams and is broadly applicable in cases where trends in index values may represent a mixture of population processes.

KEYWORDS

Alaska, brown bear, distance sampling, index, integrated model, resource use, superpopulation

Effective wildlife monitoring and management programs rely on accurate estimators of population abundance and trend (Pollock et al. 2002, Williams et al. 2002). As a result, developing methods to overcome bias caused by imperfect detection has been the focus of extensive methodological developments in wildlife science (e.g., capture-recapture, double observer, distance sampling, occupancy). These advancements can be used to provide strong inference to the target population (i.e., the group of individuals for which inference is desired), but collecting the data that these approaches require is monetarily and logistically expensive. As a result of this extra cost, these rigorous approaches may be infrequently employed (Reynolds et al. 2011).

Indices of abundance (i.e., field measures containing information on the relative size of the population; Williams et al. 2002) are commonly used to monitor trends in populations, in part because they tend to be cheaper and easier to implement than more rigorous approaches that account for imperfect detection of individuals (Anderson 2001, Williams et al. 2002, Johnson 2008). As is the case for any survey type, however, inference is dependent, in part, on how well a given approach addresses the detection process (Nichols et al. 2009, Schmidt et al. 2013). Given that detection information is generally lacking for index data, inference is reliant on strong and often untestable assumptions about the relative constancy of the detection process and the relationship between the index value and target population abundance (Anderson 2001, Pollock et al. 2002, Thompson 2002b, Williams et al. 2002, Sollmann et al. 2013). Natural resource agencies commonly have extensive collections of index data, perhaps gathered at different spatial or temporal scales, providing weak inference to the population or process of interest.

Indices may also suffer from *ad hoc* sampling designs where the spatial and temporal extent of the sampled population is poorly defined. A common example is unadjusted counts of individuals that are aggregated in response to a limited resource (e.g., bears [*Ursus* spp.] at salmon [*Oncorhynchus* spp.] spawning streams [Deacy et al. 2019], breeding aggregations of sea birds [Byrd et al. 2005], raptor migration counts [Farmer et al. 2007]). In such cases, counts represent some unknown combination of 2 distinct processes: target population abundance (i.e., abundance) and temporary immigration from the surrounding area (i.e., the superpopulation; Schwarz and Arnason 1996, Kendall 1999,



Efford and Dawson 2012, Kendall et al. 2013). Proper inference to the process of interest (i.e., abundance and trend) requires information about how temporary immigrants contribute to the index values (e.g., source-sink dynamics); however, index data alone do not contain sufficient information to resolve this confounding.

Data integration provides a potential solution to some of the problems associated with indices when more rigorous data are also available. Integrated modeling approaches formally link datasets to improve estimators of the parameters of interest and can be used to increase the amount of information available in a variety of settings (Pacifi et al. 2017, Zipkin et al. 2017, Arnold et al. 2018, Zipkin and Saunders 2018, Hostetter et al. 2019). Much of the recent development has been focused on integrated population models (Besbeas et al. 2002, Schaub and Abadi 2011, Saunders et al. 2019, Zipkin et al. 2021), but the same principles apply when formally linking datasets collected using different survey protocols (Hanks et al. 2011, Steger et al. 2017, Schmidt et al. 2019). In doing so, parameters are jointly estimated using >1 data source, thereby increasing estimator precision (Sanderlin et al. 2019, Schmidt et al. 2019, Farr et al. 2021). The process is analogous to extensions in a capture-recapture framework that allow additional parameters to be estimated with additional information (Otis et al. 1978, Pollock et al. 1990, Kendall et al. 1997). Complications arise when combined data measure different processes or occur at different spatial or temporal scales, leading to model misspecification or apparent discrepancies between datasets (Maunder and Piner 2017). Therefore, although data integration offers a potential solution for linking index data with more rigorous approaches, properly accounting for spatial or temporal scale misalignment is necessary (Saunders et al. 2019).

Brown bear (*Ursus arctos*) abundance is commonly monitored using a variety of rigorous survey approaches (Miller et al. 1997, Becker and Quang 2009, Thompson et al. 2010, Walsh et al. 2010), although cost often limits survey frequency. In Lake Clark National Park and Preserve (Lake Clark NPP), Alaska, USA, biologists conduct distance sampling surveys every 5–10 years to obtain estimators of abundance for the population of adult brown bears occurring within the Lake Clark NPP study area during early spring—the target population. Managers also collect annual index counts of brown bears in coastal sedge meadows during early summer when bears congregate to feed on seasonally important food resources (Schoen et al. 1994, Schwartz et al. 2003, Smith and Partridge 2004, Rode et al. 2006). Because the index represents a mixture of 2 population processes, direct inference to brown bear abundance is not possible. Conversely, the long survey return interval for the distance sampling efforts results in relatively few data points for trend assessment. Given the limitations of both approaches, we expected that an integrated modeling approach would improve estimators of abundance and trend for the brown bear population in Lake Clark NPP.

We developed an approach to formally integrate the Lake Clark NPP brown bear distance sampling survey data with the meadow count indices to increase the precision of estimators of population abundance and trend and decompose the relative contribution of the abundance and temporary immigration processes to the overall trend in the meadow count indices. We also investigated 3 hypothetical scenarios to further explore how data integration can be used to improve inference when trends in individual population components differ.

STUDY AREA

Our 2003–2019 study encompassed a 2,204-km² area that included all watersheds draining the coastal areas of Lake Clark NPP (Figure 1). Elevations ranged from sea level to 3,053 m, but areas >914 m did not provide habitat for brown bears (Olson and Putera 2007). In early summer (mid-Jun through mid-Jul), brown bears congregate in 5 coastal meadow complexes (Figure 1) along the Lake Clark NPP coast to feed on forbs, sedges, and adjacent coastal marine resources, important food sources prior to the period when migrating salmon become accessible. The coastal meadows were located along the north and south shores of Tuxedni Bay (49.3 km² and 36.5 km², respectively), at the mouths of Silver Salmon and Shelter Creeks (9.5 km² and 5.2 km², respectively), and along Chinitna Bay (39 km²). We defined the boundaries of each meadow complex by the presence of salt marsh vegetation and absence of trees. A portion of the Chinitna Bay meadow extends beyond Lake Clark NPP and outside the study area boundary. The area of all 5 meadows combined represents approximately 6.3% of the area covered by the distance sampling surveys. Terrain varies from flat coastal meadows and

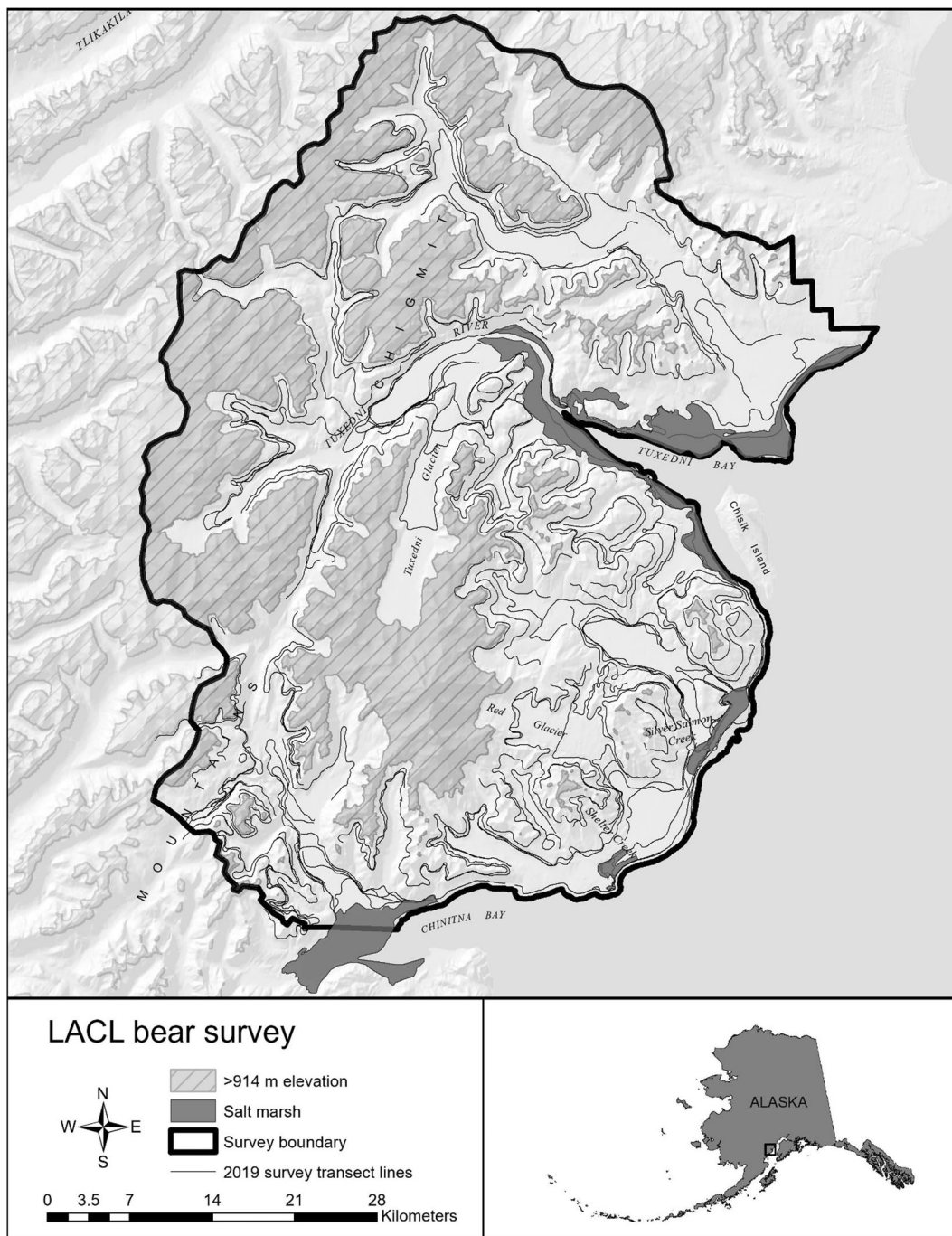


FIGURE 1 Lake Clark National Park and Preserve (LACL), Alaska, USA, study area boundary (heavy line) showing transect layout (thin lines) for the 2019 distance sampling survey for brown bears. We did not generate transects in areas >914 m elevation (hatched) where we assumed brown bears did not occur. The 5 meadow complexes where we counted brown bears annually during the mid-June to late-July period between 2003 and 2019 are also shown (dark gray)



mudflats to steep glaciated mountains, including Mount Iliamna, an active volcano. Woody vegetation consists of shrub patches (willows [*Salix* spp.], alders [*Alnus* spp.]) and forest (spruce [*Picea* spp.], balsam poplar [*Populus balsamifera*], birch [*Betula neolaskana*]). Primary food resources for brown bears in spring and early summer include sedges in coastal meadows, razor clams (*Siliqua patula*), marine mammal carcasses, migrating salmon, and moose (*Alces alces*) calves. Moose are present at low densities, and black bears (*Ursus americanus*) commonly co-occur with brown bears in coastal forests.

METHODS

Data collection

We conducted aerial distance sampling surveys in Lake Clark NPP in 2003, 2010, and 2019 to estimate brown bear abundance. We randomly placed 20-km transects along elevational contours and used straight lines with random orientation (2003 and 2010 only) in areas with little topographical relief (Figure 1). We surveyed transects in late May after most bears had emerged from dens and prior to leaf-out of woody vegetation that may completely obscure bears. We conducted surveys from a tandem fixed-wing aircraft flown at approximately 90 m above ground, searching for bear groups either on the uphill side of the line (contour transects) or on a randomly selected side (straight-line transects). Upon detecting a bear group, we recorded the spatial location, size, and composition (i.e., adults, cubs) of each group and the distance from the transect line to the group (Olson and Putera 2007). We restricted our analysis to the number of adults in each group, thereby providing inference to the population of adult bears.

We conducted the 2003 and 2010 surveys using mark-recapture distance sampling methods whereby the pilot and observer each recorded group detections independently to provide a means to estimate the detection probability on the line, $g(0)$, given that the $g(0) = 1$ assumption (i.e., complete detection on the transect line; $p_d = 1$) of conventional distance sampling (CDS) approaches was potentially violated (Olson and Putera 2007, Becker and Quang 2009, Thompson et al. 2010). Schmidt et al. (2017) reported that bias due to $p_d < 1$ at $g(0) = 1$ was low for distance sampling applications for bears; therefore, we conducted the 2019 survey using CDS (Buckland et al. 2001) wherein the pilot and observer worked as a team to detect bear groups. We converted the 2003 and 2010 data to simple CDS detection data for equivalency with the 2019 surveys (Schmidt et al. 2017). As in prior applications, we assumed that all adult brown bears were present and available for sampling during the survey period.

In 2003–2019, we used tandem fixed-wing aircraft to count bears in each coastal meadow. We visited meadows 1–4 times between mid-June and mid-July, attempting to cover the peak bear aggregation period. The lack of obscuring cover and simplicity of the habitats in these areas meant that bear groups were highly visible during these surveys. Therefore, we assumed that our counts reflected the true number of bears present in each area during each replicate survey. Various conditions, such as phenological timing (e.g., timing of green-up) or the availability of alternative foods, however, can result in some proportion of bears being located outside of the meadows on a given survey date.

Analysis

The overall trend in an index is dependent on the relative contributions of the abundance and temporary immigration processes. This relationship can be described as

$$T = \frac{(T' \times N' + T'' \times N'')}{N' + N''},$$

where overall trend (T) is a function of trend in abundance (T') and the number of individuals from the target population that are present (N') and the trend in temporary immigration (T'') and the number of additional individuals that temporarily immigrated into the study area (N''). We can estimate the overall trend directly from the series of index counts, but without

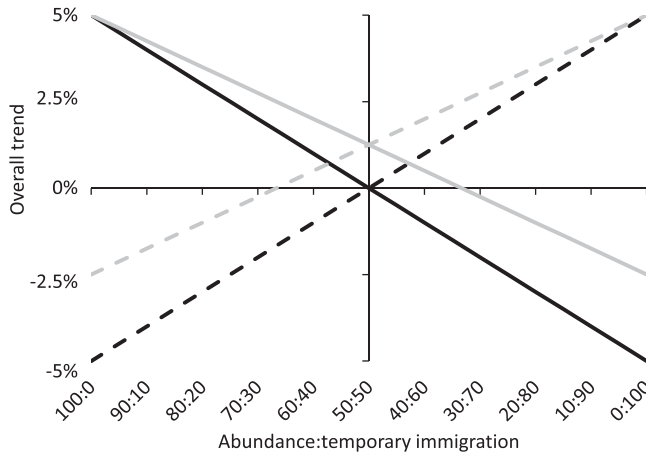


FIGURE 2 The relative impact of abundance and temporary immigration on overall trend (y-axis) based on the relative proportion of individuals in each category (abundance:temporary immigration; x-axis) for 4 scenarios with opposing trends: abundance increasing at 5%/year and temporary immigration decreasing at -5%/year (solid black line), temporary immigration increasing at 5%/year and abundance decreasing at -5%/year (dashed black line), abundance increasing at 5%/year and temporary immigration decreasing at -2.5%/year (solid gray line), and temporary immigration increasing at 5%/year and abundance decreasing at -2.5%/year (dashed gray line). Lines of the same color intersect at $x = 50:50$ and the overall trend is equal to the mean of the 2 trends at that point (i.e., $y = 0$ when opposing trends are of equal value)

additional information, the individual trend components are confounded. In cases where N'' is small relative to N' , T will provide a reasonable estimator of trend in abundance and any confounding will have little impact (Figure 2). As N'' increases, inference to abundance becomes weaker because the overall trend reflects some unknown combination of the abundance and temporary immigration processes (Figure 2). Given that we have an additional source of information on abundance (i.e., CDS data), we can resolve the confounding in the above equation through data integration.

We jointly analyzed the 2 datasets by integrating a distance sampling model with a log-linear mixed effects model for the meadow counts. We linked the 2 datasets by sharing the abundance and trend terms based on the CDS data. We followed the hierarchical group-based CDS model of Schmidt et al. (2012), which uses data augmentation to account for unobserved groups (Royle et al. 2007). We modeled the observations (y) on each transect (i) for each potential group (j) as

$$y_{ij} \sim \text{Bernoulli}(\mu_{ij}),$$

and

$$\mu_{ij} = p_{dij} \Psi_{ij}$$

where μ_{ij} is a function of the probability of detecting an individual given that it is present and available for detection (p_{dij}) and the probability of group occurrence (Ψ_{ij}). The half-normal detection function for p_{dij}

$$p_{dij} = \exp\left(-\frac{x_{ij}^2}{2\sigma_{ij}^2}\right),$$

is a function of the distance from the transect to each group (x_{ij}), and the scale parameter (σ_{ij}) that is modeled as

$$\log(\sigma_{ij}) = \alpha,$$



where α is the intercept. We assumed α was constant across surveys, although this could be relaxed through the addition of covariates. With a slight modification to the group-based model of Schmidt et al. (2012), we expressed Ψ_{ij} as a function of the expected number of groups on transect (λ_i) and the size of the augmented population (M). This can be written as

$$\Psi_{ij} = \sum_{i=1}^I \lambda_i / M,$$

where

$$\log(\lambda_i) = \mathbf{X}_i' \boldsymbol{\beta}$$

and $\mathbf{X}_i'$ is a vector of known covariate values (e.g., yr) and $\boldsymbol{\beta}$ are the associated coefficients. We included an intercept term with year as the only covariate (i.e., annual trend in abundance) and assumed group size (s_{ij}) comes from a Poisson distribution:

$$s_{ij} \sim \text{Poisson}(\delta),$$

where δ is mean group size. We derived annual abundance at the level of the study area by summing across all potential groups on the sampled transects within a given year, multiplying by mean group size, and extrapolating to the area of interest.

We assumed the observed index counts (y') in each year (k) come from a Poisson distribution with mean N_k :

$$y'_k \sim \text{Poisson}(N_k)$$

$$\log(N_k) = \mathbf{Y}'_k \boldsymbol{\mu} + \varepsilon_k,$$

where $\mathbf{Y}'_k$ is a vector of known covariate values (e.g., yr), $\boldsymbol{\mu}$ are the associated coefficients, and ε_k is a random error term representing unexplained variation. We included date and date² as covariates in the index count model to account for not all bears being present due to survey timing, and assumed different intercept values for each of the 5 meadows. To properly account for the relative contribution of trends in abundance and temporary immigration, 2 intercept terms corresponding to N' and N'' are required for each meadow. The inclusion of the year covariate in $\mathbf{X}'_i$ and $\mathbf{Y}'_k$ allowed the coefficient T' to be jointly estimated using the information contained in both datasets (i.e., T' was a component of both $\boldsymbol{\beta}$ and $\boldsymbol{\mu}$). In effect, by using the distance sampling data to inform the target population abundance and trend terms, T'' and N'' become estimable. See the JAGS code in the online supplement for more details (Supplement S1, available in Supporting Information).

We right-truncated the CDS data at 600 m based on prior work (Olson and Putera 2007), and left-truncated at 70 m to account for the incompletely observed strip beneath the aircraft (i.e., $p_d < 1$ at $x < 70$ m). To account for the fact that the actual area surveyed for each transect had been previously calculated using various left-truncation values (i.e., 2003 = 22 m; 2010 and 2019 = 90 m), we adjusted the surveyed area of each transect by the proportional change of the strip-width relative to the 70-m left-truncation point. We used an informative prior for the intercept term for N' , $\text{Norm}(85,25)\text{T}(25,85)$, based on the CDS estimate of target population abundance for the initial year. This prior reflected our presumption that most of the bears in the study area used the meadows during late summer but allowed us to acknowledge some uncertainty in that assumption. Alternative priors for N' reflecting different hypotheses or additional data sources may also be valid, depending on the survey context. We specified weakly informative priors for all remaining parameters. We fitted the integrated model in JAGS 4.3.0 (Plummer 2003), using the rjags package (Plummer 2019) and R 3.6.1 (R Core Team 2019) as an interface, running 2 chains for 20,000 iterations and discarding the first 5,000 as burn-in. We used the Gelman-Rubin diagnostic ($\text{Rhat} < 1.1$;



Brooks and Gelman 1998) and visual inspection of the chains to verify convergence had been reached. We also analyzed each dataset independently (i.e., no data integration) to compare estimator precision between the independent versus integrated approaches. We present estimates as means with 95% Bayesian credible intervals (95% CrI).

Exploration of hypothetical scenarios

Given that the results from the Lake Clark NPP data suggested that the overall observed trend in the meadow count indices largely reflected changes in abundance (see Results), we created 3 hypothetical datasets to more fully illustrate how an integrated approach may be useful when temporary immigration represents a larger component of the index. For each scenario, we generated a new dataset imitating the structure of the meadow count data but with an increased contribution of temporary immigrants (as compared to our actual data). Our intent was to mimic our original dataset, while including a large proportion of bears drawn from outside the park to the meadows. We could then demonstrate how inference from the index values changes when trends in the 2 population processes differ (Figure 2), which could profoundly affect their interpretation.

We began by creating a hypothetical population of 85 bears representing a target population that was increasing at a rate of 1.5%/year (i.e., T'). We chose the initial size of this population based on the CDS estimate for 2003. Next we assumed that 100 temporary immigrants (i.e., N'') used the meadows in the initial year and that this number changed through time at a rate of -2.5%/year, -5%/year, or 5%/year (i.e., T'') under the 3 scenarios. We used the date of each original count and the estimated coefficients from the analysis of the original data to adjust each index count for survey timing within the season and distributed the total counts across the 5 meadows in proportion to their occurrence in the original data. In doing so, we created a hypothetical series of paired values for abundance and temporary immigrants that together represented the bears measured by the meadow count index. We then fit the integrated model to the 3 new datasets. Through these scenarios, we were then able to explore the impacts on inference when large proportions of the meadow counts represent temporary immigrants drawn to this limited resource.

RESULTS

We surveyed 349, 229, and 147 transects during the distance sampling surveys in 2003, 2010, and 2019, respectively, and used 163 detected bear groups for analysis. Total survey coverage based on the estimated observable strip was 3,497 km², 2,168 km², and 1,356 km² in 2003, 2010, and 2019, respectively. We estimated adult bear abundance within Lake Clark NPP to be 85, 83, and 119 for the 3 surveys (Table 1) with an annual population trend of 1.014/year (95% CrI = 0.990–1.038; Figure 3).

TABLE 1 Estimated abundance of adult bears in the population ($\hat{N}$), with the associated 95% Bayesian credible intervals (95% CrI) and estimated coefficient of variation ($\widehat{CV}$), and the estimated bear density ($\hat{D}$; adult bears/1,000 km²) in the Lake Clark National Park and Preserve, Alaska, USA, study area during late May for each of 3 survey years based on the conventional distance sampling estimator

Survey year	$\hat{N}$	95% CrI	$\widehat{CV}$	$\hat{D}$
2003	85	69–104	11%	38.6
2010	83	62–107	13.5%	37.7
2019	119	91–153	13%	54.0

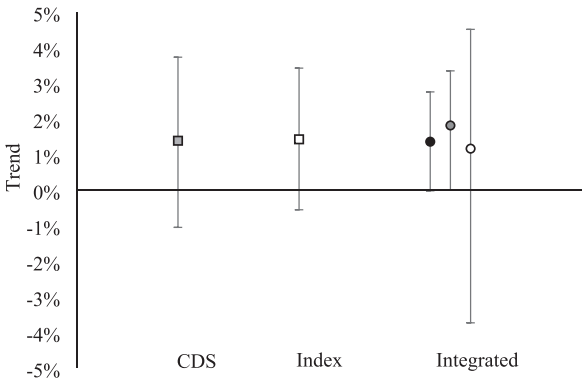


FIGURE 3 Estimates of trend in abundance (%/yr) based on the conventional distance sampling data alone (CDS), the index data alone (index), and the integrated model: overall trend (black), trend in abundance (gray), and trend in temporary immigration (white) for brown bears in Lake Clark National Park and Preserve, Alaska, USA, 2003–2019

We collected 157 meadow counts across 5 count areas from 2003–2019 (≤ 4 counts/area/yr) with the exception of 2006 when we were unable to complete surveys. There was a curvilinear relationship between date and the observed counts with peak counts generally occurring around 26 June and declining rapidly by mid-July in the average year. Annually, the maximum number of adult bears recorded across all 5 units combined varied between 97 and 225 bears. The estimated annual rate of change in the meadow count index was 1.014/year (95% CrI = 0.995–1.035), similar to the trend estimated with the CDS data (Figure 3).

When we jointly estimated overall trend using both the distance sampling and index data, estimated population growth rate was 1.014/year (95% CrI = 1.00–1.028). Integration of the 2 datasets increased precision, indicating that the trend in abundance was likely >0 (Figure 3). The integrated model, separately estimating trends in temporary immigration and abundance, indicated that the observed trend in the index was largely due to population trend (1.018/yr [95% CrI = 1.00–1.033]) with a smaller but imprecisely estimated trend in temporary immigration (1.01/yr [95% CrI = –1.037–1.045]; Figure 3).

The hypothetical scenarios demonstrated how the overall trend in the index values is dependent on the relative proportion of the abundance and temporary immigration processes captured by the meadow counts and their respective trends (Figure 4). The estimates of T were positively skewed towards the abundance trend when temporary immigration was declining and towards temporary immigration when temporary immigration was increasing faster than abundance. This occurs because as one component declines, that component represents a declining proportion of the overall dataset (Figure 2).

DISCUSSION

Our work represents an important conceptual advance that formally links legacy index datasets with more rigorous survey approaches to increase the amount of information available for populations of interest to managers. In our specific application, we integrated data from less frequent but more rigorous CDS surveys with an annual population index based on feeding aggregations of brown bears. By doing so, we were able to decompose the index values into 2 components (i.e., abundance, temporary immigration) and provide more precise estimators of the trend in brown bear abundance in Lake Clark NPP. This approach thereby provided direct information on the target population and the role of temporary immigration of bears into the study area during early summer, improving our understanding of brown bear dynamics in the area. Given the

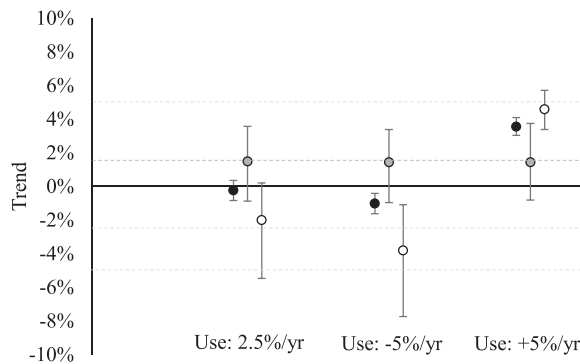


FIGURE 4 Estimates of the overall trend (black), trend in abundance (gray), and trend in temporary immigration (white) under 3 hypothetical scenarios consisting of a target population of 85 bears increasing at 1.5%/year and an additional 100 temporary immigrants from outside Lake Clark National Park and Preserve, Alaska, USA, changing at rates of -2.5%/year, -5%/year, or +5%/year. Horizontal dotted lines represent the true values for trend in abundance (gray) and for each of the 3 values of temporary immigration (light gray)

pervasiveness of legacy index datasets, our work will serve as a useful guide for practitioners in determining how to extract more information from these data by integrating them with those collected under more rigorous survey designs, or to plan for such integration by supplementing index counts with these surveys in the future.

Our main concern in the brown bear application was the mismatch in the population processes reflected in 2 survey approaches. In other cases, an index and a designed survey might sample the same target population (assuming the population subject to sampling matches the target population; Thompson 2002a), but the index may not address the issue of incomplete detection (Schmidt et al. 2019). Alternatively, available datasets may represent opposing trends, which in turn requires a clear understanding of the different processes and spatial extents measured by each data source (Maunder and Piner 2017). The task for the practitioner is to determine where 2 datasets intersect to ensure proper inference to the population of interest because resolving differences in sampling approaches through data integration can help correct for poor study design or violations of key assumptions (Hanks et al. 2011, Steger et al. 2017). Therefore, properly accounting for the spatial and temporal extents of individual data sources is important (Saunders et al. 2019, Zipkin et al. 2021).

In most settings, one assumes that a population index is directly proportional to changes in abundance in the target population (Williams et al. 2002). When that assumption is violated, however, inference is compromised. Without identifying the relative contributions of each process affecting the index value, the potential effectiveness of management actions can be limited (Roy et al. 2019). For example, declines in abundance caused by human factors (e.g., overharvest, disturbance) may be masked in an index if temporary immigration increases in the area covered by the index (i.e., possible source-sink dynamics). Without decomposing the 2 trends, the best course of action is uncertain because the overall trend in the index (T) may remain approximately stable through time. Understanding the interplay between the abundance and temporary immigration processes allows managers to make more effective decisions.

In addition to providing a conceptual advance in the application of data integration, our work represents the first comprehensive assessment of brown bear abundance and trend since Lake Clark NPP was created in 1980. Although Olson and Putera (2007) reported on the 2003 survey data, their analysis was based on the gamma detection function approach of Becker and Quang (2009), which is biased (Becker and Christ 2015, Becker and Crowley 2021). An analysis of 2003 distance sampling data from Lake Clark NPP and surrounding areas by Schmidt et al. (2017) suggested the density of brown bears of all ages was at the lower end



of the range reported for other coastal populations in Alaska (Miller et al. 1997). This is consistent with our current estimates of adult densities of 38–54 adults/1,000 km² from 2003–2019. Our finding of a slow increasing trend in abundance (i.e., ~1.4%/yr) is also consistent with a productive and lightly harvested population, which is similar to recent findings in northwestern Alaska (Schmidt et al. 2021). Together, this information suggests that current management actions are likely not limiting the Lake Clark NPP brown bear population.

An important consideration for ongoing index surveys in Lake Clark NPP and elsewhere is that changes in phenology and quality of the coastal meadows may affect habitat use by brown bears if meadow availability becomes temporally misaligned with alternative resources (Deacy et al. 2017). This hypothesis could be addressed through the inclusion of covariates representing phenological processes, or alternatively, the abundance or quality of coastal resources. Changing patterns in the use of important but spatially limited resources may occur because of the impacts of density dependence, variation in resource quality or quantity, harvest rates, disturbance by bear viewers, or other factors governing the abundance and distribution of individuals on the landscape. Similar problems may arise in areas where counts of feeding bears along salmon streams are used to index bear populations (Deacy et al. 2019) because the size and temporal distribution of the salmon run could affect the number and characteristics (e.g., sex, age) of bears counted. We conduct index surveys in Lake Clark NPP when herbaceous vegetation in coastal meadows provides bears with a spatially limited but abundant source of highly digestible protein (Rode et al. 2001), leading to our assumption that the target population of bears would largely be present in the meadow counts. Relaxing this assumption would require auxiliary data on meadow use during the survey period (e.g., collared bears). Resource use, in particular has important spatial and temporal factors that could be explicitly modeled to improve understanding of the factors that affect animal abundance in management units. Explicitly modeling the process that draws bears to these resources will improve their usefulness as population indices, especially when management regimes (e.g., harvest, land use) differ between 2 adjacent management units that are not closed to animal movement.

Although we present an application specific to brown bear surveys in Alaska, the basic concept of integrating datasets to separate the role of individual population processes is broadly applicable. Management agencies often have more index data than any other type (Anderson 2001). Simply discarding what may represent decades of effort and potential information because the data were not collected using a rigorous survey approach is simply not reasonable. Our conceptual approach to data integration provides a much needed bridge between historical approaches and those that have been more recently developed, and in some cases may be useful in justifying a switch to a more rigorous survey approach. Recent interest in integrating nonstandardized datasets (e.g., citizen science) highlights the importance of properly accounting for the intersection of individual data sources (Steger et al. 2017, Miller et al. 2019, Zipkin et al. 2019, Zulian et al. 2021). Our work contributes to the development of integrated approaches for use with datasets of varying qualities, and highlights the need to ensure that spatial and temporal mismatches representing different population processes are addressed.

RESEARCH IMPLICATIONS

Our results suggest that growth of the Lake Clark NPP brown bear population is consistent with current management practices. Several oil and mineral exploration and development activities have been proposed within and adjacent to the boundaries of Lake Clark NPP, and if implemented, these projects have the potential to affect bear populations and their use of certain habitats. In addition, increasing visitation on the Lake Clark NPP coast could alter spatial and temporal patterns of bear use of meadows. Annual meadow index counts, in combination with periodic distance sampling surveys, provide a practical solution for long-term population monitoring in this area. The ongoing collection of both data types would mitigate the risk that changes in temporary immigration would bias



interpretation of trends in abundance and ensure that reliable trend information was available for managers. In a broader context, we suggest that biologists and managers consider current and future monitoring programs in a multi-metric framework rather than focusing on projects or surveys in isolation. Data integration provides a means to formally combine multiple data types collected on a single species or system, providing more information than would be possible when treating projects individually. We encourage researchers to explicitly plan for data integration, thereby increasing the quality and quantity of data available for management, and reducing the cost of ongoing programs.

ACKNOWLEDGMENTS

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the United States Government. We thank all the pilots for safely surveying hundreds of transects over the years, and all the observers for meticulous data collection. We thank C. N. Miller and D. T. Schertz for technical assistance with survey preparation, data processing, and figure design. M. C. MacCluskie, A. E. Miller, and J. P. Lawler provided helpful comments on an earlier draft of this manuscript. Funding for this project was provided by the Southwest Alaska Network, the National Park Service Natural Resource Preservation Program, Lake Clark National Park and Preserve, and the Alaska Department of Fish and Game. The Dena'ina people have called the land around Qizhjah Vena (Lake Clark) home for thousands of years and we respect their deep connection to the land and animals.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICS STATEMENT

We adhered to the guidelines provided by the American Society of Mammologists for the conduct of aerial surveys of vertebrate animals.

DATA AVAILABILITY STATEMENT

The brown bear data are available at: <https://irma.nps.gov/DataStore/>.

ORCID

Joshua H. Schmidt  <http://orcid.org/0000-0002-5803-6431>

Tammy L. Wilson  <http://orcid.org/0000-0002-3672-8277>

REFERENCES

- Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. *Wildlife Society Bulletin* 29:1294–1297.
- Arnold, T. W., R. G. Clark, D. N. Koons, and M. Schaub. 2018. Integrated population models facilitate ecological understanding and improved management decisions. *Journal of Wildlife Management* 82:266–274.
- Becker, E. F., and A. M. Christ. 2015. A unimodal model for double observer distance sampling surveys. *PLoS One* 10: e0136403.
- Becker, E. F., and D. W. Crowley. 2021. Estimating brown bear abundance and harvest rate on the southern Alaska Peninsula. *PLoS One* 16:e0245367.
- Becker, E. F., and P. X. Quang. 2009. A gamma-shaped detection function for line-transect surveys with mark-recapture and covariate data. *Journal of Agricultural, Biological, and Environmental Statistics* 14:207–223.
- Besbeas, P., S. N. Freeman, B. J. T. Morgan, and E. A. Catchpole. 2002. Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics* 58:540–547.
- Brooks, S. P., and A. Gelman. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7:434–455.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. *Introduction to distance sampling*. Oxford University Press, Oxford, United Kingdom.



- Byrd, G. V., H. M. Renner, and M. Renner. 2005. Distribution patterns and population trends of breeding seabirds in the Aleutian Islands. *Fisheries Oceanography* 14:139–159.
- Deacy, W. W., J. B. Armstrong, W. B. Leacock, C. T. Robbins, D. D. Gustine, E. J. Ward, J. A. Erlenbach, and J. A. Stanford. 2017. Phenological synchronization disrupts trophic interactions between Kodiak brown bears and salmon. *Proceedings of the National Academy of Sciences* 114:10432–10437.
- Deacy, W. W., W. B. Leacock, E. J. Ward, and J. B. Armstrong. 2019. Aerial surveys cause large but ephemeral decreases in bear presence at salmon streams in Kodiak, Alaska. *PloS One* 14:e0222085.
- Efford, M. G., and D. K. Dawson. 2012. Occupancy in continuous habitat. *Ecosphere* 3:32. <https://doi.org/10.1890/ES11-00308.1>
- Farmer, C. J., D. J. T. Hussell, and D. Mizrahi. 2007. Detecting population trends in migratory birds of prey. *Auk* 124: 1047–1062.
- Farr, M. T., D. S. Green, K. E. Holekamp, and E. F. Zipkin. 2021. Integrating distance sampling and presence-only data to estimate species abundance. *Ecology* 102:e03204.
- Hanks, E. M., M. B. Hooten, and F. A. Baker. 2011. Reconciling multiple data sources to improve accuracy of large-scale prediction of forest disease incidence. *Ecological Applications* 21:1173–1188.
- Hostetter, N. J., B. Gardner, T. S. Sillett, K. H. Pollock, and T. R. Simons. 2019. An integrated model decomposing the components of detection probability and abundance in unmarked populations. *Ecosphere* 10:e02586.
- Johnson, D. H. 2008. In defense of indices: the case of bird surveys. *Journal of Wildlife Management* 72:857–868.
- Kendall, W. L. 1999. Robustness of closed capture-recapture methods to violations of the closure assumption. *Ecology* 80: 2517–2525.
- Kendall, W. L., J. E. Hines, J. D. Nichols, and E. H. Campbell Grant. 2013. Relaxing the closure assumption in occupancy models: staggered arrival and departure times. *Ecology* 94:610–617.
- Kendall, W. L., J. D. Nichols, and J. E. Hines. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology* 78:563–578.
- Maunder, M. N., and K. R. Piner. 2017. Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. *Fisheries Research* 192:16–27.
- Miller, D. A. W., K. Pacifici, J. S. Sanderlin, and B. J. Reich. 2019. The recent past and promising future for data integration methods to estimate species' distributions. *Methods in Ecology and Evolution* 10:22–37.
- Miller, S. D., G. C. White, R. A. Sellers, H. V. Reynolds, J. W. Schoen, K. Titus, V. G. Barnes, Jr., R. B. Smith, R. R. Nelson, W. B. Ballard, and C. C. Schwartz. 1997. Brown and black bear density estimation in Alaska using radiotelemetry and replicated mark-resight techniques. *Wildlife Monographs* 133:3–55.
- Nichols, J. D., L. Thomas, and P. B. Conn. 2009. Inferences about landbird abundance from count data: recent advances and future directions. Pages 201–235 in D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. *Modeling demographic processes in marked populations*. Environmental and Ecological Statistics, volume 3. Springer, Boston, Massachusetts, USA.
- Olson, T. L., and J. A. Putera. 2007. Refining monitoring protocols to survey brown bear populations in Katmai National Park and Preserve and Lake Clark National Park and Preserve. Alaska Region Natural Resources Technical Report NPS/AR/NRTR-2007-66. National Park Service, Alaska Regional Office, Anchorage, Alaska, USA.
- Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monographs* 62:3–135.
- Pacifici, K., B. J. Reich, D. A. W. Miller, B. Gardner, G. Stauffer, S. Singh, A. McKerrow, and J. A. Collazo. 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. *Ecology* 98: 840–850.
- Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20–22, Vienna, Austria.
- Plummer, M. 2019. rjags: Bayesian graphical models using MCMC. R-package version 4-10. <https://CRAN.R-project.org/package=rjags>
- Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture-recapture experiments. *Wildlife Monographs* 107:3–97.
- Pollock, K. H., J. D. Nichols, T. R. Simons, G. L. Farnsworth, L. L. Bailey, and J. R. Sauer. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics* 13:105–119.
- R CoreTeam. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reynolds, J. H., W. L. Thompson, and B. Russell. 2011. Planning for success: identifying effective and efficient survey designs for monitoring. *Biological Conservation* 144:1278–1284.
- Rode, K. D., S. D. Farley, and C. T. Robbins. 2006. Behavioral responses of brown bears mediate nutritional effects of experimentally introduced tourism. *Biological Conservation* 133:70–80.



- Rode, K. D., C. T. Robbins, and L. A. Shipley. 2001. Constraints on herbivory by grizzly bears. *Oecologia* 128:62–71.
- Roy, C., N. L. Michel, C. M. Handel, S. L. Van Wilgenburg, J. C. Burkhalter, K. E. B. Gurney, D. J. Messmer, K. Princé, C. S. Rushing, J. F. Saracco, R. Schuster, A. C. Smith, P. A. Smith, P. Sólymos, L. A. Venier, and B. Zuckenberg. 2019. Monitoring boreal avian populations: how can we estimate trends and trajectories from noisy data? *Avian Conservation and Ecology* 14(2):8.
- Royle, J. A., R. M. Dorazio, and W. A. Link. 2007. Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics* 16:67–85.
- Sanderlin, J. S., W. M. Block, B. E. Strohmeier, V. A. Saab, and J. L. Ganey. 2019. Precision gain versus effort with joint models using detection/non-detection and banding data. *Ecology and Evolution* 9:804–817.
- Saunders, S. P., M. T. Farr, A. D. Wright, C. A. Bahlai, J. W. Ribeiro, Jr., S. Rossman, A. L. Sussman, T. W. Arnold, and E. F. Zipkin. 2019. Disentangling data discrepancies with integrated population models. *Ecology* 100:e02714.
- Schaub, M., and F. Abadi. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology* 152:S227–S237.
- Schmidt, J. H., C. L. McIntyre, and M. C. MacCluskie. 2013. Accounting for incomplete detection: What are we estimating and how might it affect long-term passerine monitoring programs? *Biological Conservation* 160:130–139.
- Schmidt, J. H., K. L. Rattenbury, J. P. Lawler, and M. C. MacCluskie. 2012. Using distance sampling and hierarchical models to improve estimates of Dall's sheep abundance. *Journal of Wildlife Management* 76:317–327.
- Schmidt, J. H., J. H. Reynolds, K. L. Rattenbury, L. M. Phillips, K. S. White, D. Schertz, J. M. Morton, and H. S. Kim. 2019. Integrating distance sampling with minimum counts to improve monitoring. *Journal of Wildlife Management* 83: 1454–1465.
- Schmidt, J. H., H. L. Robison, L. S. Parrett, T. S. Gorn, and B. S. Shults. 2021. Brown bear density and estimated harvest rates in northwestern Alaska. *Journal of Wildlife Management* 85:202–214.
- Schmidt, J. H., T. L. Wilson, W. L. Thompson, and J. H. Reynolds. 2017. Improving inference for aerial surveys for bears: the importance of underlying assumptions and the cost of unnecessary complexity. *Ecology and Evolution* 7:4812–4821.
- Schoen, J. W., R. W. Flynn, L. H. Suring, K. Titus, and L. R. Beier. 1994. Habitat-capability model for brown bear in southeast Alaska. *Bears: Their Biology and Management* 9:327–337.
- Schwartz, C. C., S. D. Miller, and M. A. Haroldson. 2003. Grizzly bear. Pages 556–586 in G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, editors. *Wild mammals of North America: biology, management, and conservation*, 2nd edition. Johns Hopkins University Press, Baltimore, Maryland, USA.
- Schwarz, C. J., and A. N. Arnason. 1996. A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics* 52:860–873.
- Smith, T. S., and S. T. Partridge. 2004. Dynamics of intertidal foraging by coastal brown bears in southwestern Alaska. *Journal of Wildlife Management* 68:233–240.
- Sollmann, R., A. Mohamed, H. Samejima, and A. Wilting. 2013. Risky business or simple solution – relative abundance indices from camera-trapping. *Biological Conservation* 159:405–412.
- Steger, C., B. Butt, and M. B. Hooten. 2017. Safari science: assessing the reliability of citizen science data for wildlife surveys. *Journal of Applied Ecology* 54:2053–2062.
- Thompson, S. K. 2002a. *Sampling*. Wiley, New York, New York, USA.
- Thompson, W. L. 2002b. Towards reliable bird surveys: accounting for individuals present but not detected. *Auk* 119: 18–25.
- Thompson, W. L., K. Peirce, and B. A. Mangipane. 2010. Protocol for monitoring brown bears—version 1.0: Southwest Alaska Inventory and Monitoring Network. Natural Resource Report NPS/SWAN/NRR–2010/275. National Park Service, Fort Collins, Colorado, USA.
- Walsh, P., J. Reynolds, G. Collins, B. Russell, M. Winfree, and J. Denton. 2010. Application of a double-observer aerial line-transect method to estimate brown bear population density in southwestern Alaska. *Journal of Fish and Wildlife Management* 1:47–58.
- Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. *Analysis and management of animal populations*. Academic Press, San Diego, California, USA.
- Zipkin, E. F., B. D. Inouye, and S. R. Beissinger. 2019. Innovations in data integration for modeling populations. *Ecology* 100: e02713.
- Zipkin, E. F., S. Rossman, C. B. Yackulic, J. D. Wiens, J. T. Thorson, R. J. Davis, and E. H. Campbell Grant. 2017. Integrating count and detection-nondetection data to model population dynamics. *Ecology* 98:1640–1650.
- Zipkin, E. F., and S. P. Saunders. 2018. Synthesizing multiple data types for biological conservation using integrated population models. *Biological Conservation* 21:240–250.



- Zipkin, E. F., E. R. Zylstra, A. D. Wright, S. P. Saunders, A. O. Finley, M. C. Dietze, M. S. Itter, and M. W. Tingley. 2021. Addressing data integration challenges to link ecological processes across scales. *Frontiers in Ecology and the Environment* 19:30–38.
- Zulian, V., D. A. W. Miller, and G. Ferraz. 2021. Integrating citizen-science and planned-survey data improves species distribution estimates. *Diversity and Distributions* 27:2498–2509.

Associate Editor: Quresh Latif.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Schmidt, J. H., T. L. Wilson, W. L. Thompson, and B. A. Mangipane. 2022. Integrating distance sampling survey data with population indices to separate trends in abundance and temporary immigration. *Journal of Wildlife Management* 86:e22185. <https://doi.org/10.1002/jwmg.22185>