Sea Duck Joint Venture
Final Project Summary
FY23 (October 1, 2022 – September 30, 2023)

Project Title: *Evaluating Sea Duck Detectability in the Puget Sound Winter Ambient Monitoring Program – SDJV #136*

Principal Investigators:

- *Sarah J. Converse, USGS Washington Cooperative Fish and Wildlife Research Unit, University of Washington, sconver@uw.edu*
- *Kyle A. Spragens, Waterfowl Section, Washington Department of Fish and Wildlife, Kyle.Spragens@dfw.wa.gov*
- *Postdoctoral scientist: Jamie L. Brusa, School of Environmental and Forest Sciences, University of Washington, jbrusa@uw.edu*
- *Postdoctoral scientist: Matthew T. Farr, School of Aquatic and Fishery Sciences, University of Washington, farrm@uw.edu*

Partners:

- *Emily Silverman, U.S. Fish and Wildlife Service, Division of Migratory Bird Management, emily_silverman@fws.gov*
- *Joseph Evenson, Washington Department of Fish and Wildlife, Joseph.Evenson@dfw.wa.gov*

Project Description: *See attached manuscript.*

Project Objectives: *See attached manuscript.*

Preliminary Results:

(The following abstract is from the attached manuscript): Long-term, large-scale monitoring of wildlife populations is an integral part of conservation research and management. However, some traditional monitoring protocols lack the information needed to account for sources of measurement error in data analyses. Ignoring measurement error, such as partial availability, imperfect detection, and species misidentification, can lead to mischaracterizations of population states and processes. Accounting for measurement error is key to robust monitoring of populations, which can inform a wide variety of decisions, including harvest, habitat restoration, and determination of the legal status of species. We undertook an effort to retroactively minimize bias in a large-scale, long-term monitoring program for marine birds in the Salish Sea, Washington, USA by conducting an auxiliary study to jointly estimate components of measurement error. We built a novel model in a Bayesian framework that simultaneously harnessed human observer and photographic data types to produce estimates necessary to correct for the effects of partial availability, imperfect detection, and species misidentification. Across all 31 species identified in photographs, both of the two participating observers had instances of undercounting and overcounting birds but tended to undercount (observers undercounted totals across all species on 69.3% – 78.9% of transects). We estimated species-specific correction factors that can be used to correct both historical and future counts from the Salish Sea survey, which has been running since 1992. Our novel modeling framework can be applied in other multi-species monitoring contexts where minimal

photographic data can be collected for the purposes of correcting for measurement error in large-scale, long-term datasets.

Project Status: *This project is complete. We have submitted the attached manuscript for publication. This serves as our final report to SDJV.*

Project Funding Sources (US$): *NA*

Total Expenditures by Category (SDJV plus all partner contributions; US$): *NA*

This draft manuscript is distributed solely for purposes of scientific peer review. Its content is deliberative and pre-decisional, and it must not be disclosed or released by reviewers. Because the manuscript has not yet been approved for publication by the U.S. Geological Survey (USGS), it does not represent any official USGS finding or policy. The findings and conclusions in this article are those of the authors and do not necessarily represent the view of the U.S. Fish and Wildlife Service.

Ecological Applications

Article

Correcting for measurement errors in a long-term aerial survey with auxiliary photographic data

Jamie L. Brusa[1a*], Matthew T. Farr[2], Joseph Evenson[3], Emily Silverman[4], Bryan Murphie[5], Thomas A. Cyra[5], Heather J. Tschaekofske[6] Kyle A. Spragens[6], and Sarah J. Converse[7]

1. School of Environmental and Forest Sciences, University of Washington, Seattle, Washington, USA
2. Washington Cooperative Fish and Wildlife Research Unit, School of Aquatic and Fishery Sciences, University of Washington, Seattle, Washington, USA
3. Washington Department of Fish and Wildlife, Bellingham, Washington, USA
4. U.S. Fish and Wildlife Service, Division of Migratory Bird Management, Laurel, Maryland, USA
5. Washington Department of Fish and Wildlife, Port Townsend, Washington, USA
6. Washington Department of Fish and Wildlife, Olympia, Washington, USA

[a] current affiliation: Integrated Statistics in support of Northeast Fisheries Science Center, National Marine Fisheries Service, NOAA, Woods Hole, Massachusetts, USA

7. U.S. Geological Survey, Washington Cooperative Fish and Wildlife Research Unit, School of Aquatic Fishery Sciences & School of Environmental and Forest Sciences, University of Washington, Seattle, Washington, USA

*Corresponding author: jamie.brusa@noaa.gov

Open research statement: Code and data for analyses will be available at https://github.com/Quantitative-Conservation-lab/seaduck_detection.

Keywords: aerial survey analysis; Bayesian statistics; imperfect detection; species misidentification; marine birds; multi-species monitoring

Abstract

Long-term, large-scale monitoring of wildlife populations is an integral part of conservation research and management. However, some traditional monitoring protocols lack the information needed to account for sources of measurement error in data analyses. Ignoring measurement error, such as partial availability, imperfect detection, and species misidentification, can lead to mischaracterizations of population states and processes. Accounting for measurement error is key to robust monitoring of populations, which can inform a wide variety of decisions, including harvest, habitat restoration, and determination of the legal status of species. We undertook an effort to retroactively minimize bias in a large-scale, long-term monitoring program for marine birds in the Salish Sea, Washington, USA by conducting an auxiliary study to jointly estimate components of measurement error. We built a novel model in a Bayesian framework that simultaneously harnessed human observer and photographic data types to produce estimates necessary to correct for the effects of partial availability, imperfect detection, and species misidentification. Across all 31 species identified in photographs, both of the two participating observers had instances of undercounting and overcounting birds but tended to undercount (observers undercounted totals across all species on 69.3% – 78.9% of transects). We estimated species-specific correction factors that can be used to correct both historical and future counts from the Salish Sea survey, which has been running since 1992. Our novel modeling framework can be applied in other multi-species monitoring contexts where minimal photographic data can be collected for the purposes of correcting for measurement error in large-scale, long-term datasets.

Introduction

Robust estimates of population parameters obtained from monitoring data are valuable for wildlife research and management. However, robust estimates depend on monitoring protocols and analytical approaches designed to account for measurement errors. Raw count data from population surveys frequently result in biased measures of abundance as a result of such errors (Caughley 1974, Samuel and Pollock 1981, Pollock and Kendall 1987, Russell et al. 1996, Davis et al. 2022). Ignoring measurement error may lead to either over- or underestimation of parameters, such as population abundance. Measurement error can arise from multiple sources, including partial availability, imperfect detection, or species misidentification. Partial availability occurs when animals flee from observers before they are detected or exhibit other cryptic behaviors (e.g., diving under water). Imperfect detection occurs when observers fail to detect animals that are available. Finally, even when individuals of a species are counted accurately, they may be misidentified as members of another species (Miller et al. 2011). Often, these processes occur simultaneously during data collection and may not be readily apparent to observers. A variety of survey methods and corresponding statistical models, such as detection/non-detection data and occupancy modeling (MacKenzie et al. 2002, Miller et al. 2011) or repeated count data and N-mixture models (Royle and Nichols 2003) may be adopted to account for such measurement errors.

Monitoring across large geographic areas is often conducted through aerial surveys (Siniff and Skoog 1964, Briggs et al. 1985a, Buckley and Buckley 2000, Chabot et al. 2018). Established methods for estimating observation error in aerial surveys include distance sampling (Buckland et al. 2001), double-observer methods (Cook and Jacobson 1979), simultaneous observations from multiple platforms, such as plane and ship (Briggs et al. 1985b), or estimation

78 of correction factors using comparisons of observed aerial counts to a known number of decoys
79 (Frederick et al. 2003, Strobel and Butler 2014). However, each of these methods has potential
80 drawbacks, including additional assumptions or higher resource (e.g., time, money) demands.
81 Distance sampling is a useful method but can be challenging to implement in aerial surveys
82 because observers must record exact distances or distance bins in addition to counting and
83 identifying species, which is often not feasible for large groups composed of multiple species
84 (e.g., Davis et al. 2022); further, distance sampling cannot directly accommodate partial
85 availability. Double observer methods and counts from multiple platforms are typically more
86 expensive, do not address partial availability and can, for some species, be difficult to implement
87 in practice (Samuel and Pollock 1981, Briggs et al. 1985b, Pollock and Kendall 1987). Using
88 decoys to develop correction factors can have limited applicability if decoys do not reasonably
89 mimic the behaviors of live animals.
90 Photographs taken from aerial platforms are a promising approach, especially as
91 autonomous aerial vehicles become more readily available and affordable. As photographs allow
92 for identifying and counting species without a time limit, these methods may eliminate or
93 substantially reduce measurement error associated with imperfect detection and species
94 misidentification. However, long-term collection and analysis of aerial photographs currently
95 tends to be more costly and time-intensive than observer-led surveys and can be logistically
96 prohibitive for large landscapes (Watson 1969, Bayliss and Yeomans 1990, Béchet et al. 2004).
97 Using autonomous aerial systems tends to lower costs and risks to observer safety, but it can also
98 introduce challenges that existing technology has not yet overcome, such as additional sound
99 disturbance, privacy issues, limitations on the duration of flights (Wang et al. 2019), time
100 consuming processing procedures, and high sensitivity to inclement weather (e.g., Weiser et al.

2022). Additionally, photographs with poor contrast between animals and their background or including heavy cover that conceals animals can create difficulties in identifying species or detecting individual animals, which can lead to large biases in abundance estimates (Siniff and Skoog 1964, Brack et al. 2018). Finally, photographs provide only a single snapshot whereas human observers usually have several seconds to observe animals, which may improve species identification. However, whereas photography alone may have weaknesses as a monitoring method, coupling photographs with observer counts in aerial surveys can support estimation of reliable correction factors for observer counts and can serve as a cost-effective approach to improve observer-led wildlife monitoring (Bayliss and Yeomans 1990, Lamprey et al. 2020).

Numerous marine bird species are either residents of or overwinter in the Salish Sea, an ecologically, economically, and culturally important ecosystem in the North American portion of the Pacific flyway (Gaydos and Pearson 2011, Crewe et al. 2012). The Washington Department of Fish and Wildlife has monitored wintering marine birds in the United States portion of the Salish Sea annually since 1992 using strip-transect aerial surveys. These surveys have provided consistent evidence for declining abundances of multiple marine bird species (Anderson et al. 2009, Bower 2009, Vilchis et al. 2014). Accurate abundance estimates are valuable for guiding decision-making processes, e.g., regarding harvest, habitat restoration, and legal protection for declining species. Additionally, marine birds have been identified as important indicators of ecosystem health in the Salish Sea (Pearson and Hamel 2013, Blight et al. 2015, Miller et al. 2015, Bishop et al. 2016), increasing the importance placed on monitoring these species. Aerial surveys are a particularly important method for monitoring marine birds in this region because some sections of the Salish Sea are difficult to access by other means (Vilchis et al. 2014).

123 Here, we present a novel model designed to correct for measurement error in aerial
124 survey counts of marine birds. Our model integrates observer and photographic data collected
125 simultaneously during a one-time aerial survey and is designed to account for several sources of
126 measurement error, including partial availability, imperfect detection, and species
127 misidentification. We applied the model to develop correction factors that can be used to correct
128 counts from both past and future aerial surveys. Our model can be applied to any monitoring
129 situation in which limited photographic data can be collected, simultaneously with observer
130 counts, for the purpose of calculating correction factors that can then be used to correct a larger
131 dataset of observer counts.
132
133 **Methods**
134 *Data collection*
135 Data collection from a high-wing de Havilland DHC-2 aircraft on floats occurred over 5
136 days in March 2012 in a portion of the Salish Sea, Washington, USA (48° N, 123° W) known to
137 contain a high diversity of overwintering marine bird species. Surveys were designed using a
138 strip transect method with a strip width of 50 m on the left side of the plane. A 0.64-cm poly line
139 tied to the wing strut at 33° and the edge of the floats at 58° created visual boundaries for the
140 transect. The aircraft flew at a speed of 157-167 km/hour at an altitude of about 61 m; the plane
141 flew directly into the wind to maintain a forward orientation. Each transect was about 2 km long
142 and took, on average, 44 seconds to complete. A total of 625 transects were flown over the 5-day
143 study, capturing 175,680 photographs.
144 Two experienced observers (17 and 15 years of experience for observers 1 and 2,
145 respectively) sat in the middle and rear seats on the left side of the aircraft, and the aircraft

146 landed once per day for observers to switch seats. Observers recorded the number of birds
147 detected and identified each bird to the lowest taxonomic group possible (usually species).
148 Observers did not communicate with each other during data collection and were visually
149 separated by an opaque divider. The observer in the middle seat had a slightly larger window
150 than the observer in the rear seat; only the middle seat is used during standard surveys. Observers
151 recorded an index of glare and the Beaufort sea-state during surveys. However, we did not
152 include either of these variables in our analyses because glare had negligible variation across
153 transects, and Beaufort sea-state varied within transects, which were our unit of analysis.
154 Meanwhile, a Canon EOS 5D Mark II equipped with a Canon EF 70-200 mm f/2.8 L IS
155 USM lens set to 200 mm, attached high on the wing strut to eliminate vibrations from the
156 propeller, captured photographs of birds. This forward-facing camera took continuous
157 photographs at 3.9 frames per second imaging the transect strip from 250-270 m ahead of the
158 aircraft to 500-540 m ahead, often capturing the same birds in multiple images. The camera was
159 calibrated to the transect strip at the start of each survey day. An additional camera, the point-of-
160 view camera, was designed to capture the area seen by observers simultaneous with observers.
161 However, the field of view from this camera proved to be different from that of observers (the
162 camera was aimed abeam of the aircraft while the observers could see abeam as well as behind
163 and ahead), limiting its usefulness. We do not refer to the point-of-view camera further;
164 hereafter, "camera" refers to the forward-facing camera. The in-flight observers synchronized
165 their watches with the clocks on the cameras and Global Positioning System (GPS) to allow for
166 accurate image matching to the observer records.
167 An independent observer, not present on any of the flights, counted and identified to the
168 lowest taxonomic group possible (usually species) each bird in each photograph from the

camera. We used data from 321 of the 625 completed transects for analyses. As the focus of the study was to derive correction factors for sea ducks (Tribe Mergini),transects were randomly selected from a group where the observers recorded sea ducks or where the transects were over habitats used by sea ducks. In addition, randomly selected transects from this group were prioritized such that all Beaufort sea states and glare categories were represented. See Evenson et al. (2013) for full details on data collection and processing.

Analytical approach

We assume that the camera bird species composition and abundance data represent the population of interest. Some small differences between the true composition and abundance and what was observed in the photographs may occur because foraging marine birds dive (independent of a response to the plane) and, thus, are not always available and because 100% of all birds in the photographs could not be identified to species. By contrast, the observer data potentially contained multiple sources of measurement error. We identified three sources of measurement error: 1) movement of birds in response to the plane, 2) misidentification of species, and 3) imperfect detection of birds. Movement in response to the plane may include diving or flying out of the transect but may also include flying into the transect or surfacing after the camera passed over their location. Species misidentification occurs if an observer detects an individual but incorrectly identifies it. Imperfect detection of birds occurs if an observer misses an entire group of one or more individuals or if they under- or overcount the number of individuals in a detected group. Many marine birds form large and mixed-species groups during the winter, and some species have quite similar physical characteristics, e.g., Common Goldeneyes (*Bucephala clangula*) and Barrow's Goldeneyes (*B. islandica*), which may contribute to imperfect detection or misidentification.

192 Birds in photographs could not always be identified to species (~7% of marine birds,
193 excluding gulls and scaup, could not be identified to species). Because the photographic data
194 were used as truth in the analysis, this limited our ability to fully account for species
195 misidentification. To develop species-specific abundances despite this limitation, we allocated
196 the individuals in species groups identified in photographs to appropriate species in proportion to
197 their occurrence in the photographic detections that were identified to species (Table 1, Conn et
198 al. 2012). For example, if 450 Common Goldeneyes, 50 Barrow's Goldeneyes, and 200
199 unclassified goldeneyes were recorded in photographs, $\frac{450}{50+450}$ = 90% (180) of the unclassified
200 goldeneyes would be allocated to the count of Common Goldeneyes and the remainder to
201 Barrow's Goldeneyes. In this way, we eliminated groups recorded as "unclassified goldeneyes,"
202 in the dataset. For the observer data, we left counts in species groups unchanged. The model then
203 dealt with these as misidentifications using the methods described below. The assumption
204 implicit in our approach – that the relative abundance of unidentifiable individuals within species
205 groups is proportional to the relative abundance of identifiable individuals within species groups
206 in photographs – may not be perfectly met, and, therefore, we ran two analyses: a taxonomically
207 fine-filtered analysis and a taxonomically coarse-filtered analysis. In the taxonomically coarse-
208 filtered analysis, we defined species groups to a higher taxonomic level (e.g., "goldeneyes")
209 rather than assigning them to species as described above. Additionally, given challenges with
210 identification and the goals of the monitoring program, gulls (likely *Larus spp.* and
211 *Chroicocephalus spp.* in the Salish Sea in winter) were not identified beyond species group even
212 in the fine-filtered analysis. We also treated scaup, including Greater Scaup (*Aythya marila*) and
213 Lesser Scaup (*Aytha affinis*), as a single group given the challenges in distinguishing them.
214 *Camera*

We summarized observations from the camera including birds flying and birds on the surface of the water, giving us a count, FF_{ji}, for transect *j* and species *i*. We assumed the camera contained no measurement error and captured the true abundance of birds available within the field of view. To estimate species composition in each transect, *j*, we described FF_{ji}, using a multinomial distribution as shown in equation 1 (Figure 1):

$$FF_{ji} \sim multinomial(\pi_i, \sum_i FF_{ji}), \quad (1)$$

where π_i is the bird species composition pre-aircraft contact (i.e., the proportion of birds on the transect that are members of each species *i*), and $\sum_i FF_{ji}$ is the total bird abundance (i.e., the summation of the data across species, *i*, within each transect, *j*). We described the species-wide bird abundance on transect *j* in equation 2 using a Poisson distribution with mean species-wide bird abundance, Λ:

$$\sum_i FF_{ji} \sim Poisson(\Lambda) \quad (2)$$

We used the forward-facing camera data to estimate π_i and Λ. We parameterize species composition as: $\pi_i = \frac{\lambda_i}{\Lambda}$, where λ_i is the expected species-specific mean abundance. We constrained $\Lambda = \Sigma_i \lambda_i$, which implies equation 3:

$$FF_{ji} \sim Poisson(\lambda_i). \quad (3)$$

The following subsections describe a method to estimate the three parameters associated with each type of measurement error (i.e., movement, species misidentification, and imperfect detection). However, we found that disentangling these three parameters can prove difficult in practice. Therefore, we follow the presentation of that analytical method with a condensed method that estimates those three parameters as one parameter to capture the total measurement error generated from movement, imperfect detection, and species misidentification.

Movement

238 The observer's field of view is on the side of the aircraft; thus, observers count birds after
239 they come into contact with the aircraft (i.e., are nearly or directly below the aircraft). Individual
240 birds may move into or out of view at random or due to a behavioral response to the aircraft. The
241 available latent abundance for observers is, therefore, an outcome of species-wide abundance and
242 a movement process. We used equation 4 to describe available latent species-wide abundance,
243 $\Sigma_i N_{ji}$, as the summation of individuals captured by the camera for each species *i* and transect *j*
244 after aircraft contact:
245 $$\Sigma_i N_{ji} \sim Poisson(\Lambda \times \bar{\alpha}), \quad (4)$$
246 where $\bar{\alpha}$ is the mean of the species-specific movement, α_i, which captures movement rates of
247 individual species. The support for the movement rate is 0 to ∞, which allows for individuals to
248 move both in and out of the observers' field of view, where $\bar{\alpha} < 1$ indicates more birds moving
249 out of the field of view than in, $\bar{\alpha} > 1$ indicates more birds moving into the field of view than out,
250 and $\bar{\alpha} = 1$ indicates the same magnitude of birds moving in and out of view.
251 *Species misidentification*
252 Observers recorded the counts of each species *i*, which required correctly identifying
253 individuals to species. Despite rigorous training and experience, species misidentification can
254 occur during the observation process (Johnston et al. 2015). For a given species *i*, the observed
255 number of individuals contains the number of correctly identified individuals of species *i* in
256 addition to false positives (i.e., individuals from species *k* misidentified by the observer as
257 species *i*). There may also be individuals of species *i* misidentified as species *k*. Again, we
258 considered the true bird species composition, π, and we defined the bird species composition
259 seen by the observers as ϕ. We can use the conditional probability $(\phi_k|\pi_i)$ to estimate how many
260 individuals of species *i* were misidentified as species *k*. These conditional probabilities are

contained within a square $i \times k$ matrix, where elements along the diagonal are the probabilities of correctly identifying species *i,* ($\phi_{k=i}|\pi_i$), and off-diagonal elements contain the misidentification probabilities ($\phi_{k\neq i}|\pi_i$). We use equation 5 to describe this observational process with a multinomial distribution:

$$C_{jik} \sim multinomial((\phi_k|\pi_i), N_{ji}), \quad (5)$$

where the species-specific available latent abundance, N_{ji}, for each species *i* in transect *j* is distributed into the elements C_{jik}, which contain the additional dimension *k*. We define this latent value, C_{jik}, as the scalar elements of the confusion matrix for each transect *j*, representing the number of individuals of species *i* that were correctly identified (diagonal elements, $k = i$) and incorrectly identified (off-diagonal elements, $k \neq i$). Using the $i \times k$ confusion matrix, the sum of column *k* (i.e., $\sum_i C_{jik}$) is the number of individuals of species *k* recorded under perfect detection, including both species correctly identified ($k = i$, diagonal element) and misidentified ($k \neq i$, off-diagonal elements), which might also include individuals counted by the aerial observers that were not captured by the camera. Then, we sum across columns $k = 1, \ldots, K$ species (i.e., $\sum_k \sum_i C_{jik}$), which collapses the species-specific abundances to the species-wide latent abundance for each transect *j*. Here, $\sum_k \sum_i C_{jik}$ is equivalent to $\sum_i N_{ji}$, as misidentification only changes the species-specific magnitudes but not the species-wide value.

Imperfect detection

In addition to misidentification, observers may also undercount or overcount the number of individuals in their field of view due to imperfect detection. We can describe this process using a binomial distribution as in equation 6; we use *i* to denote true species identification from the photographs and *k* to denote observed species identification from real-time observations:

$$\sum_k OBS_{jko} \sim binomial(\bar{p}_o \sum_i N_{ji}), \quad (6)$$

where $\sum_k OBS_{jko}$ are data containing the observed bird species-wide counts for each observer *o* and transect *j*. Detection probability, $\bar{p}_o$, is the mean observer- and species-specific detection probability; and $\sum_i N_{ji}$ (i.e., $\sum_k \sum_i C_{jik}$) is the species-wide number of individuals.

Model identifiability

As noted above, the data and the structure of the model did not allow for separately estimating movement rate and detection probability. However, the mathematical product of these parameters, $\bar{\alpha} \times \bar{p}_o = \bar{\varepsilon}_o$, is estimable, giving us the modified species-wide observation process illustrated in equation 7:

$$\sum_k OBS_{jko} \sim Poisson(\Lambda \times \bar{\varepsilon}_o). \quad (7)$$

Here, we summed the observer data across species and described the data with a Poisson distribution. The expected value of the distribution is the product of the expected bird species-wide abundance, movement rate, and detection probability. We generalize the combination of movement and detection to estimate $\bar{\varepsilon}_{jo}$ specific to each transect and observer, allowing us to account for potential differences in $\bar{\varepsilon}_o$ due to the seat assignment (middle or rear) of observer, *o*, during transect, *j*, as described by equation 8:

$$\log(\bar{\varepsilon}_{jo}) = \bar{\varepsilon}_o + \beta_1 \times seat_{jo}, \quad (8)$$

where $\bar{\varepsilon}_o$ is the estimate for observer *o* when in the rear seat, β_1 is the additive effect of being in the middle seat, and $seat_{jo}$ is a binary indicator (1 for middle seat, 0 for rear seat) for each transect *j* and observer *o*.

We cannot fully estimate the confusion matrix and associated conditional probabilities. However, we simply corrected for misidentification by specifying it as a rate, $\frac{\phi_{ok=oi}}{\pi_i}$, the rate at which species *i* is correctly identified, rather than a series of conditional probabilities. Thus, we

used equation 9 to estimate the observed bird species composition, ϕ_{ok}, with a multinomial distribution and observer data:

$$OBS_{jko} \sim multinomial(\phi_{ok}, \sum_k OBS_{jko}), \quad (9)$$

where ϕ_k is the proportion of species identified by an observer as species, *k*.

Out-of-sample correction

We can use the estimated parameters (i.e., $\bar{\varepsilon}_o, \pi_i, \phi_{ok}$) to correct for out-of-sample observations made when photographic data do not exist (Figure 1). We used equation 10 to model species-specific observation data with a Poisson distribution:

$$OBS_{jko} \sim Poisson\left(\lambda_i \times \frac{\phi_{o,k=i}}{\pi_i} \times \bar{\varepsilon}_o\right). \quad (10)$$

Therefore, by combining the new observation data with the estimated parameters, we can calculate the parameter of interest, λ_i. The estimated correction factor, *CF*, for each species and observer is $\widehat{CF}_{io} = \frac{\phi_{o,k=i}}{\pi_i} \times \bar{\varepsilon}_o$.

Parameter estimation

We fit our model using a Markov Chain Monte Carlo (MCMC) approach (Gelfand and Smith 1990, Casella and George 1992, Geman and Geman 1993). We fit the model in the R package NIMBLE version 0.6-10 (de Valpine et al. 2017) in R version 4.0 (R Core Team 2020) with 3 chains, a burn-in of 10,000 iterations, and a sampling period of 20,000 iterations. We used the R packages coda (Plummer et al. 2006), ggmcmc (Fernández-i-Marín 2016), and MCMCvis (Youngflesh 2018) to inspect model convergence using trace plots, density plots, and Gelman-Rubin statistic, $\hat{R}$, values (Gelman and Rubin 1992). We used vague priors, including the log of $\lambda_i \sim Normal(0, 100)$, $\phi_{ok} \sim Dirichlet(1)$, $\bar{\varepsilon}_o \sim Normal(0, 100)$, and $\beta_1 \sim Normal(0, 100)$.

Model testing

We used simulated data to evaluate the model's ability to correctly return parameter estimates for out-of-sample predictions (see Appendix SI for full description of model testing methods). For the simulations, we simplified the model to only include one observer and did not include a seat assignment covariate. We generated a parameter value for the true movement while also generating an offset (between 1 and 1.5) that allowed for a larger field of view for observers compared to the camera. Similarly, we generated probabilities of misidentification (pairwise values across all species) and detection (one value assigned for all species) and the true species composition for each simulated transect. We used the generated misidentification values to construct a species × species confusion matrix for each transect. Using the generated values for true movement and true composition, we simulated species-specific data for the camera. Data and model code for the simulations and empirical model are provided at https://github.com/Quantitative-Conservation-lab/seaduck_detection.

Results

Summed across all 321 transects, the camera captured 55,000 photographs containing 9029 individual marine birds, whereas observer 1 recorded 6434 birds, and observer 2 recorded 5021. Thirty-one marine bird species were identified across observers and photographs (Table 1). The total number of individuals per species captured on the camera across all transects ranged from 2 (black scoter, *Melanitta americana*) to 1644 (bufflehead, *Bucephala albeola*). The camera captured 1612 groups of marine birds across all transects, ranging from 1 to 518 birds per group with a mean of 5.60 birds per group (SD = 19.16). Group size was left skewed: 79% of groups had ≤5 birds and 91% had ≤10 birds. Including only birds that could be identified to species, 27% of groups captured by the camera included more than one species.

Both observers tended to undercount birds compared to the camera (Figure 2); observer 1 recorded a mean of 78.9% (SD = 1.2%) of the birds caught on the camera per transect, and observer 2 recorded a mean of 68.7% (SD = 1.0%). However, both observers counted more birds than caught on the camera on some transects (28.3% of the transects for observer 1 and 25.2% of the transects for observer 2). The species with the smallest difference in total counts between observer and camera was Barrow's Goldeneye (observer 1 total was 96.7% of the camera total) and Brant (*Branta bernicla.*; observer 2 total was 100.4% of the camera total). The species with the greatest difference in counts between observer and camera were Pelagic Cormorant (*Phalacrocorax pelagicus*; observer 1 total was 10.6% of the camera total) and Northern Pintail (*Anas acuta*; observer 2 total was 1.4% of the camera total).

The results from our simulations indicated that our model was able to capture the data-generating parameters accurately. The simulations resulted in a 100% convergence rate for $\bar{\varepsilon}$ with minimal bias (< -0.01%, interquartile range (IQR) = -0.05 – 0.04), a 96.2% convergence rate for λ with minimal bias (0.01%, IQR = -0.06 – 0.08), and a 98.2% convergence rate for Λ with minimal bias (<0.01%, IQR = -0.05 – 0.07).

Based on model diagnostics and general agreement between our corrected counts and our camera counts, our model appeared to perform well on the empirical dataset. Overall diagnostics indicated that our models converged; although, we had $\hat{R}$ estimates > 1.1 for $\widehat{CF}_{Ancient\ Murrelet}$ and $\widehat{CF}_{Black\ Scoter}$ (Appendix S2 Table S1), which had very few observations (Table 1). Correction factor estimates varied across species/species groups and between observers (Appendix S2 Table S1, Figure 3). For most species/species groups, our calculated corrected counts (calculated by multiplying the inverse of our correction estimates from our model by the observer counts) were very similar to the counts from the camera. As one would expect, the

model tended to produce correction estimates that yielded corrected counts proportionally closer to the counts of the camera for species with larger sample sizes. The seat position of the observer had a negligible effect on detection/species identification (taxonomically fine-grained analysis: $\beta_{1(fine)}$ = -0.001, 95% credible interval = -0.04 – 0.04; taxonomically coarse-grained analysis: $\beta_{1(coarse)}$ = -0.002, 95% credible interval = -0.04 – 0.03).

Discussion

Comparing detected animals to known numbers of animals can be an effective method for estimating measurement errors and adjusting abundance estimates to improve accuracy of wildlife monitoring data (Caughley et al. 1976, Bayliss and Yeomans 1990, Pearse et al. 2008). Our results suggest that uncorrected counts from aerial surveys of marine birds in the Salish Sea include notable measurement error resulting from some combination of animal movement in response to the plane, imperfect detection, and species misidentification. Comparing real-time observations of bird counts from aerial surveys, especially with multiple species, to counts from photographs taken of the area in front of the plane is a complicated process because animal movement, detection, and species identification cannot be separately estimated. Therefore, it is impossible to determine if an individual captured by the camera but not by an observer resulted from the bird not being available for detection, the observer not detecting the bird, or the observer detecting the bird but identifying it as the incorrect species. As is common in other studies investigating observation errors from aerial surveys (Caughley et al. 1976, Pearse et al. 2008, Alisauskas and Conn 2019), observers counted fewer individuals than were in the population of interest (i.e., in the photographs from the camera), and measurement error varied

by species and observer. Applying the correction factors that we estimated to the observer counts yielded accurate estimates based on comparison to the camera counts.

At its most basic level, our study demonstrates that the ability to detect and correctly identify marine bird species from aircraft varies both across observers and species, reinforcing that estimates correcting for biases must account for, at minimum, these two factors. Although multiple methods to account for imperfect detection of wildlife in aerial surveys have emerged, many of these approaches either do not mimic the actual field conditions under which monitoring data are collected, do not account for partial availability, or fail to account for species misidentification (Davis et al. 2022). Given the ability of our model to provide accurate correction factors for the species in the dataset, our model results can be used to adjust both future and historical counts of marine birds in the Salish Sea aerial survey by the same observers under the assumption that observer performance is constant over time. For surveys with the same observers, marine bird wintering counts can be retroactively adjusted to improve the ability of the survey to capture true abundance while integrating uncertainty in the observation process into the abundance estimates. However, observers will change over time, as will the abilities of individual observers (Sauer et al. 1994), underscoring the value of repeating the simultaneous collection of observer and photographic data as funding allows, when staff turnover occurs, or when conditions that may influence detection are altered (e.g., change of aircraft platform).

Although the method we have developed can provide accurate correction factors for most of the species recorded in the survey, the method has limited value when species are rare or highly sensitive to plane presence. Given that the correction is multiplicative, the models we present cannot provide an estimate for the true number of birds present when the observer count for a species is zero. In the data for the current study, observers did not record any observations

419 of Black Scoter or Eurasian Wigeon, but the camera captured both. From the surveys alone, one
420 could erroneously but reasonably conclude that neither of these species were present on the
421 transects. Black Scoter was only captured by the camera on one transect, and one observer may
422 have seen the same animal and identified it as an unknown scoter. Eurasian Wigeon were
423 identified in mixed groups with American Wigeon, which substantially outnumber the Eurasian
424 Wigeon. Others have also noted that less numerous species in a mixed groups may receive
425 disproportionately lower counts or may not be recognized as separate species (Gilbert et al.
426 2021).

427 Although our correction estimates are specific to the Salish Sea, specific characteristics
428 of the survey (e.g., aircraft used and altitude flown), observers, and species recorded during the
429 surveys, our approach has potential to be useful for other monitoring programs, especially those
430 using multiple observers and in study regions where the ability to account for imperfect detection
431 using double observer methods or multiple survey platforms is logistically prohibitive.
432 Researchers can mimic our field approach and apply our model to their own data to estimate
433 their own correction factors. Several expansions and refinements are also possible. Future
434 applications may benefit from using an array of point-of-view cameras that accurately capture
435 the same field of view as the observers. This source of data would allow for development of a
436 detection model that accounts for group size. As demonstrated in other studies investigating
437 detection probabilities in aerial surveys (Cook and Jacobson 1979, Pearse et al. 2008, Clement et
438 al. 2017, Gilbert et al. 2021), group size can affect the detection probabilities of observers.
439 However, achieving an exactly equivalent field of view between observers and cameras could be
440 difficult in practice because observers would need to maintain fixed head and eye positions
441 throughout each transect.

442 Advances in plane-based or autonomous aerial vehicle-based photography, coupled with
443 artificial intelligence-based analysis of images, have the potential to transform aerial monitoring
444 of wildlife (e.g., Gonzalez et al. 2016, Chabot et al. 2018). Removing the reliance on plane-based
445 observers with photographic images that can be saved and revisited over time as technology
446 improves could address many of the challenges of observer-based surveys. However, to be
447 viable, surveys of this type will have to be of sufficiently high quality and low cost, and the
448 timeline for development of the technology necessary to facilitate such surveys is uncertain. In
449 multi-species systems, this timeline is likely to be extended.
450 Many wildlife population and community analyses require accurate long-term monitoring
451 of abundance to inform species management or address scientific hypotheses (Tinkle 1979,
452 Nichols and Williams 2006). Inaccurate abundance estimates can lead to poor management
453 decisions or erroneous scientific conclusions (Ward-Paige et al. 2010, Elliot et al. 2020). By
454 means of a small supplemental study using cameras to compare observed counts to known counts
455 from photographs, we have shown that it is possible to account for multiple sources of
456 measurement error, including animal movement, imperfect detection, and species
457 misidentification. Additionally, we have demonstrated how to apply estimated correction factors
458 retroactively to existing data. Our results contribute to a growing recognition of the need for
459 tools to improve abundance estimates from aerial-based survey counts. The ability to maintain
460 long-term aerial monitoring efforts while simultaneously improving our confidence in the
461 estimates and trends obtained from them is imperative for the management of wildlife
462 populations.
463
464 Acknowledgements

Support for JLB was provided by the US Fish and Wildlife Service via a grant from the Sea Duck Joint Venture. Support for MTF was provided by the Washington Department of Fish and Wildlife. We thank the Washington Cooperative Fish and Wildlife Research Unit for facilitating funding. Chuck Perry with Kenmore Air served as the pilot on all flights. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author contributions: Conceptualization: K.A.S., J.E., E.S., Data curation: J.L.B., Data collection: J.E., B.M., T.C., H.J.T., Formal analysis: J.L.B., M.T.F., S.J.C., Funding acquisition: S.J.C., K.A.S., J.E., Investigation: J.L.B., M.T.F., K.A.S., J.E., E.S., S.J.C., H.J.T., Methodology: J.L.B., M.T.F., K.A.S., J.E., E.S., S.J.C., B.M., T.C., Project administration: K.A.S., Visualization: J.L.B., M.T.F., K.A.S., J.E., E.S., S.J.C., Writing – original draft: J.L.B., Writing – review and edition: all authors.

Conflict of Interest: We declare that we do not have any conflicts of interest

References

Alisauskas, R. T., and P. B. Conn. 2019. Effects of distance on detectability of Arctic waterfowl using double-observer sampling during helicopter surveys. Ecology and Evolution 9:859–867.

Anderson, E. M., J. L. Bower, D. R. Nysewander, J. R. Evenson, and J. R. Lovvorn. 2009. Changes in avifaunal abundance in a heavily used wintering and migration site in Puget Sound, Washington, during 1966-2007. Marine Ornithology 37:19–27.

Bayliss, P., and K. M. Yeomans. 1990. Use of low-level aerial photography to correct bias in aerial survey estimates of magpie goose and whistling duck density in the Northern Territory. Australian Wildlife Research 17:1–10.

Béchet, A., A. Reed, N. Plante, J.-F. Giroux, and G. Gauthier. 2004. Estimating the size of the greater snow goose population. The Journal of Wildlife Management 68:639–649.

Bishop, E., G. Rosling, P. Kind, and F. Wood. 2016. Pigeon guillemots on whidbey Island, washington: a six-year monitoring study. Northwestern Naturalist 97:237–245.

Blight, L. K., K. A. Hobson, T. K. Kyser, and P. Arcese. 2015. Changing gull diet in a changing world: A 150-year stable isotope (δ13C, δ15N) record from feathers collected in the Pacific Northwest of North America. Global Change Biology 21:1497–1507.

Bower, J. L. 2009. Changes in marine bird abundance in the Salish Sea: 1975 to 2007. Marine Ornithology 37:9–17.

Brack, I. V., A. Kindel, and L. F. B. Oliveira. 2018. Detection errors in wildlife abundance estimates from Unmanned Aerial Systems (UAS) surveys: Synthesis, solutions, and challenges. Methods in Ecology and Evolution 9:1864–1873.

Briggs, K. T., W. B. Tyler, and D. B. Lewis. 1985a. Aerial surveys for seabirds: methodological experiments. The Journal of Wildlife Management 49:412–417.

Briggs, K. T., W. B. Tyler, and D. B. Lewis. 1985b. Comparison of ship and aerial surveys of birds at sea. The Journal of Wildlife Management 49:405–411.

Buckland, S. T. 183326, D. R. 183327 Anderson, K. P. 183328 Burnham, J. L. 183329 Laake, D. L. 183330 Borchers, and L. 183331 Thomas. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, UK.

510 Buckley, P. A., and F. G. Buckley. 2000. The role of helicopters in seabird censusing. Proceedings
511 of the Society of Caribbean Ornithology, Special Publication 1:134–147.
512 Casella, G., and E. I. George. 1992. Explaining the Gibbs sampler. The American Statistician
513 46:167–174.
514 Caughley, G. 1974. Bias in aerial survey. The Journal of Wildlife Management 38:921–933.
515 Caughley, G., R. Sinclair, and D. Scott-Kemmis. 1976. Experiments in aerial survey. The Journal
516 of Wildlife Management 40:290–300.
517 Chabot, D., C. Dillon, and C. Francis. 2018. An approach for using off-the-shelf object-based
518 image analysis software to detect and count birds in large volumes of aerial imagery.
519 Avian Conservation and Ecology 13.
520 Clement, M. J., S. J. Converse, and J. A. Royle. 2017. Accounting for imperfect detection of
521 groups and individuals when estimating abundance. Ecology and Evolution 7:7304–
522 7310.
523 Conn, P. B., J. L. Laake, and D. S. Johnson. 2012. A hierarchical modeling framework for multiple
524 observer transect surveys. PLOS ONE 7:e42294.
525 Cook, R. D., and J. O. Jacobson. 1979. A design for estimating visibility bias in aerial surveys.
526 Biometrics 35:735–742.
527 Crewe, T., K. Barry, P. Davidson, and D. Lepage. 2012. Coastal waterbird population trends in
528 the Strait of Georgia 1999–2011: results from the first 12 years of the British Columbia
529 Coastal Waterbird Survey. British Columbia Birds 22:8–35.
530 Davis, K. L., E. D. Silverman, A. L. Sussman, R. R. Wilson, and E. F. Zipkin. 2022. Errors in aerial
531 survey count data: Identifying pitfalls and solutions. Ecology and Evolution 12:e8733.

532 Elliot, N. B., A. Bett, M. Chege, K. Sankan, N. de Souza, L. Kariuki, F. Broekhuis, P. Omondi, S.
533 Ngene, and A. M. Gopalaswamy. 2020. The importance of reliable monitoring methods
534 for the management of small, isolated populations. Conservation Science and Practice
535 2:e217.
536 Evenson, J. R., H. Tschaekofske, B. Murphie, T. Cyra, and D. Kraege. 2013. Status and summary
537 of the 2013 WDFW winter sea duck aerial survey detectability project - phase 3. Page
538 35. Washington Department of Fish and Wildlife, Olympia, Washington, USA.
539 Fernández-i-Marín, X. 2016. ggmcmc: Analysis of MCMC samples and Bayesian inference.
540 Journal of Statistical Software 070:1–20.
541 Frederick, P. C., B. Hylton, J. A. Heath, and M. Ruane. 2003. Accuracy and variation in estimates
542 of large numbers of birds by individual observers using an aerial survey simulator.
543 Journal of Field Ornithology 74:281–287.
544 Gaydos, J. K., and S. F. Pearson. 2011. Birds and mammals that depend on the Salish Sea: a
545 compilation. Northwestern Naturalist 92:79–94.
546 Gelfand, A. E., and A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal
547 densities. Journal of the American Statistical Association 85:398–409.
548 Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple
549 sequences. Statistical Science 7:457–472.
550 Geman, S., and D. Geman. 1993. Stochastic relaxation, Gibbs distributions and the Bayesian
551 restoration of images. Journal of Applied Statistics 20:25–62.

552 Gilbert, A. D., C. N. Jacques, J. D. Lancaster, A. P. Yetter, and H. M. Hagy. 2021. Visibility bias of
553 waterbirds during aerial surveys in the nonbreeding season. Wildlife Society Bulletin
554 45:6–15.
555 Gonzalez, L. F., G. A. Montes, E. Puig, S. Johnson, K. Mengersen, and K. J. Gaston. 2016.
556 Unmanned aerial vehicles (UAVs) and artificial intelligence revolutionizing wildlife
557 monitoring and conservation. Sensors 16:97.
558 Johnston, A., C. B. Thaxter, G. E. Austin, A. S. C. P. Cook, E. M. Humphreys, D. A. Still, A. Mackay,
559 R. Irvine, A. Webb, and N. H. K. Burton. 2015. Modelling the abundance and distribution
560 of marine birds accounting for uncertain species identification. Journal of Applied
561 Ecology 52:150–160.
562 Lamprey, R., F. Pope, S. Ngene, M. Norton-Griffiths, H. Frederick, B. Okita-Ouma, and I. Douglas-
563 Hamilton. 2020. Comparing an automated high-definition oblique camera system to
564 rear-seat-observers in a wildlife survey in Tsavo, Kenya: Taking multi-species aerial
565 counts to the next level. Biological Conservation 241:108243.
566 MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm.
567 2002. Estimating site occupancy rates when detection probabilities are less than one.
568 Ecology 83:2248–2255.
569 Miller, A., J. E. Elliott, K. H. Elliott, M. F. Guigueno, L. K. Wilson, S. Lee, and A. Idrissi. 2015.
570 Brominated flame retardant trends in aquatic birds from the Salish Sea region of the
571 west coast of North America, including a mini-review of recent trends in marine and
572 estuarine birds. Science of The Total Environment 502:60–69.

Miller, D. A., J. D. Nichols, B. T. McClintock, E. H. C. Grant, L. L. Bailey, and L. A. Weir. 2011. Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification. Ecology 92:1422–1428.

Nichols, J. D., and B. K. Williams. 2006. Monitoring for conservation. Trends in Ecology & Evolution 21:668–673.

Pearse, A. T., P. D. Gerard, S. J. Dinsmore, R. M. Kaminski, and K. J. Reinecke. 2008. Estimation and correction of visibility bias in aerial surveys of wintering ducks. The Journal of Wildlife Management 72:808–813.

Pearson, S. F., and N. J. Hamel. 2013. Marine and terrestrial bird indicators for Puget Sound. Page 55. Washington Department of Fish and Wildlife and Puget Sound Partnership, Olympia, Washington, USA.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11.

Pollock, K. H., and W. L. Kendall. 1987. Visibility bias in aerial surveys: a review of estimation procedures. The Journal of Wildlife Management 51:502–510.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence–absence data or point counts. Ecology 84:777–790.

Russell, J., S. Couturier, L. G. Sopuck, and K. Ovaska. 1996. Post-calving photo-census of the Rivière George caribou herd in July 1993. Rangifer:319–330.

Samuel, M. D., and K. H. Pollock. 1981. Correction of visibility bias in aerial surveys where animals occur in groups. The Journal of Wildlife Management 45:993–997.

Sauer, J. R., B. G. Peterjohn, and W. A. Link. 1994. Observer differences in the North American Breeding Bird Survey. The Auk 111:50–62.

Siniff, D. B., and R. O. Skoog. 1964. Aerial censusing of caribou using stratified random sampling. The Journal of Wildlife Management 28:391–401.

Strobel, B. N., and M. J. Butler. 2014. Monitoring whooping crane abundance using aerial surveys: Influences on detectability. Wildlife Society Bulletin 38:188–195.

Tinkle, D. W. 1979. Long-Term Field Studies. BioScience 29:717–717.

de Valpine, P., D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik. 2017. Programming with models: writing statistical algorithms for general model structures with NIMBLE. Journal of Computational and Graphical Statistics 26:403–413.

Vilchis, L. I., C. K. Johnson, J. R. Evenson, S. F. Pearson, K. L. Barry, P. Davidson, M. G. Raphael, and J. K. Gaydos. 2014. Assessing ecological correlates of marine bird declines to inform marine conservation. Conservation Biology 29:154–163.

Wang, D., Q. Shao, and H. Yue. 2019. Surveying wild animals from satellites, manned aircraft and unmanned aerial systems (UASs): a review. Remote Sensing 11:1308.

Ward-Paige, C., J. M. Flemming, and H. K. Lotze. 2010. Overestimating fish counts by non-instantaneous visual censuses: consequences for population and community descriptions. PLOS ONE 5:e11722.

Watson, R. M. 1969. Aerial photographic methods in censuses of animals. East African Agricultural and Forestry Journal 34:31–37.

Weiser, E. L., P. L. Flint, D. K. Marks, B. S. Shults, H. M. Wilson, S. J. Thompson, and J. B. Fischer. 2022. Optimizing surveys of fall-staging geese using aerial imagery and automated counting. Wildlife Society Bulletin:e1407.

Youngflesh, C. 2018. MCMCvis: tools to visualize, manipulate, and summarize MCMC output. The Journal of Open Source Software 3:640.

623 Table 1. Species and species groupings of marine birds identified from photographs taken during
624 aerial surveys in the Salish Sea, Washington, USA. To account for unclassified birds even in the
625 reference (i.e., camera) data, we conducted two analyses, a taxonomically fine-grained analysis
626 and coarse-grained analysis. The classes in the fine-grained and coarse-grained analyses are
627 shown below, along with scientific names for species.

Species (i.e., fine-grained)	Species group (i.e., coarse-grained)	Scientific name	Counts (observer 1, observer 2, **camera**)
SPECIES			
Ancient Murrelet	alcid[a]	*Synthliboramphus antiquus*	0, 6, **1**
Common Murre	alcid	*Uria aalge*	29, 40, **63**
Marbled Murrelet	alcid	*Brachyramphus marmoratus*	51, 9, **104**
Pigeon Guillemot	alcid	*Cepphus columba*	24, 34, **88**
Rhinoceros Auklet	alcid	*Cerohinca monocerata*	103, 89, **377**
unclassified alcid[b]	alcid	NA	6, 4, **4**
unclassified small alcid[b]	alcid	NA	0, 1, **2**
unclassified murrelet[b]	alcid	NA	0, 8, **26**
Brant	Brant	*Branta bernical*	11566, 1195, **1190**

Bufflehead	Bufflehead	*Bucephala albeola*	1196, 991, **1659**
Double-crested Cormorant	cormorant	*Nannopterum auritum*	6, 15, **22**
Pelagic Cormorant	cormorant	*Urile pelagicus*	7, 9, **66**
unclassified cormorant[b]	cormorant	NA	50, 36, **65**
American Wigeon	dabbling duck[c]	*Mareca americana*	496, 258, **1055**
Eurasian Wigeon	dabbling duck	*Marcea penelope*	0, 0, **10**
Mallard	dabbling duck	*Anas platyrhynchos*	22, 10, **66**
Northern Pintail	dabbling duck	*Anas acuta*	90, 2, **138**
Barrow's Goldeneye	goldeneye	*Bucephala islandica*	29, 19, **30**
Common Goldeneye	goldeneye	*Bucephala clangula*	90, 32, **174**
unclassified goldeneye[b]	goldeneye	NA	68, 44, **61**
Horned Grebe	grebe	*Podiceps auratus*	55, 41, **221**
Red-necked Grebe	grebe	*Podiceps grisegena*	46, 7, **52**
Western Grebe	grebe	*Aechmophorus occidentalis*	254, 194, **399**
unclassified grebe[b]	grebe	NA	0, 4, **9**

gull[d]	gull	NA	515, 422, **726**
Harlequin Duck	Harlequin Duck	*Histrionicus histrionicus*	36, 21, **74**
Long-tailed Duck	Long-tailed Duck	*Clangula hyemalis*	320, 260, **402**
Common Loon	loon	*Gavia immer*	59, 25, **89**
Pacific Loon	loon	*Gavia pacifica*	3, 8, **18**
Red-throated Loon	loon	*Gavia stellata*	51, 84, **98**
unclassified loon[b]	loon	NA	41, 22, **121**
Common Merganser	merganser	*Mergus merganser*	1, 2, **1**
Red-breasted Merganser	merganser	*Mergus serrator*	57, 19, **104**
unclassified merganser[b]	merganser	NA	4, 62, **20**
Ruddy Duck	Ruddy Duck	*Oxyura jamaicensis*	26, 5, **30**
scaup[e]	scaup	NA	90, 75, **124**
Black Scoter	scoter	*Melanitta americana*	0, 0, **2**
Surf Scoter	scoter	*Melanitta perspicillata*	691, 560, **993**
White-winged Scoter	scoter	*Melanitta deglandi*	129, 92, **450**
unclassified scoter[b]	scoter	NA	158, 171, **107**

unclassified duck[b]	[f]	NA	0, 75, **79**
unclassified seabird[b]	[f]	NA	11, 0, **79**

[a]alcids are members of the Family Alcidae

[b]In the taxonomically fine-grained analysis, counts of the following groups were allocated proportional to the counts for individuals within the group that were identified to species, as follows: "unclassified alcid" to all alcids; "unclassified small alcid" to all murrelets; "unclassified murrelet" to all murrelets; "unclassified cormorant" to all cormorants; "unclassified goldeneye" to all goldeneyes; "unclassified grebe" to all grebes; "unclassified loon" to all loons; and "unclassified merganser" to all mergansers. The species group "unclassified duck" was allocated proportionally across all ducks (i.e., Bufflehead, all dabbling ducks, all goldeneyes, Harlequin Duck, Ling-tailed Duck, all mergansers, Ruddy Duck, scaup, and all scoters). The species group "unclassified seabird" was allocated proportionally across all other species

[c]dabbling ducks are members of the Family Anatidae, Subfamily Anatinae

[d]Given challenges with identification and the objectives of the monitoring program, we grouped all gulls (Family Laridae), likely including *Larus spp.* and *Chroicocephalus philadelphia*, in both the fine-grained and coarse-grained analysis

[e]scaup are composed of two species, Greater Scaup (*Aythya marila*) and Lesser Scaup (*Aytha affinis*), but because of challenges in distinguishing them, they were never identified to species in the survey data

[f]In the taxonomically coarse-grained analysis, counts of the following groups were allocated proportionally to species groups as follows: "unclassified duck" to Bufflehead, dabbling ducks, goldeneyes, Harlequin Duck, Long-tailed Duck, mergansers, Ruddy Duck, scaup, and scoters; "unclassified seabird" to all species groups

649

Figure Captions

Figure 1. Directed acyclic graph showing relationship between in-sample data and out-of-sample correction.

Boxes are data, and dashed circles are parameter values. Solid arrows represent relationships between data and parameters. Dashed arrows indicate information used to correct the out-of-sample data.

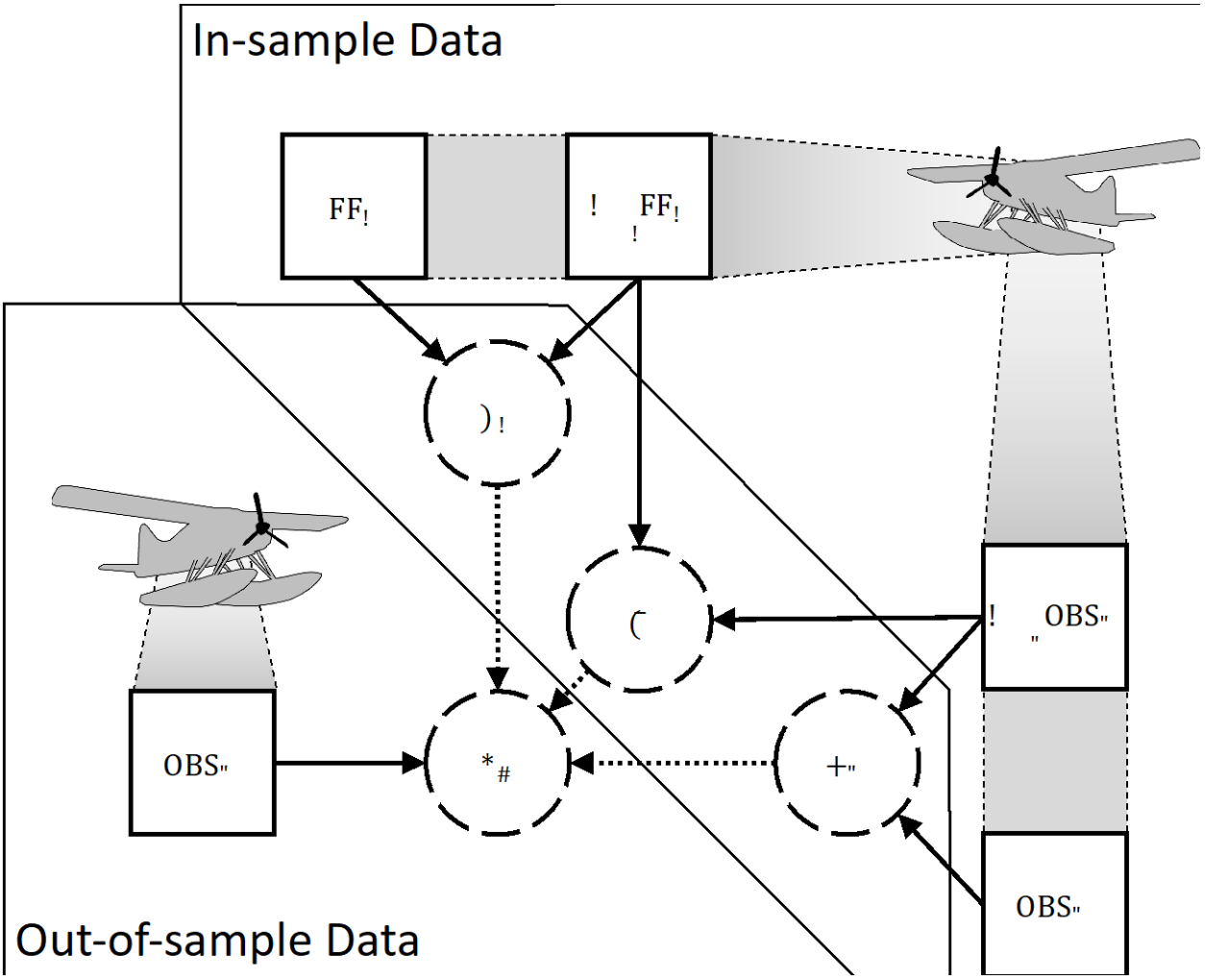
Figure 2. Observer count accuracy per species

Points display the count captured by the camera subtracted from each observer's count for each species for each group of birds observed. Only groups that were captured by camera were included.

Figure 3. Model performance for count corrections

Points represent raw counts from observers or the camera, and triangles represent estimated abundances produced by our model for select species. Error bars on estimates indicate 95% credible intervals. Note, y-axis is on logarithmic scale.

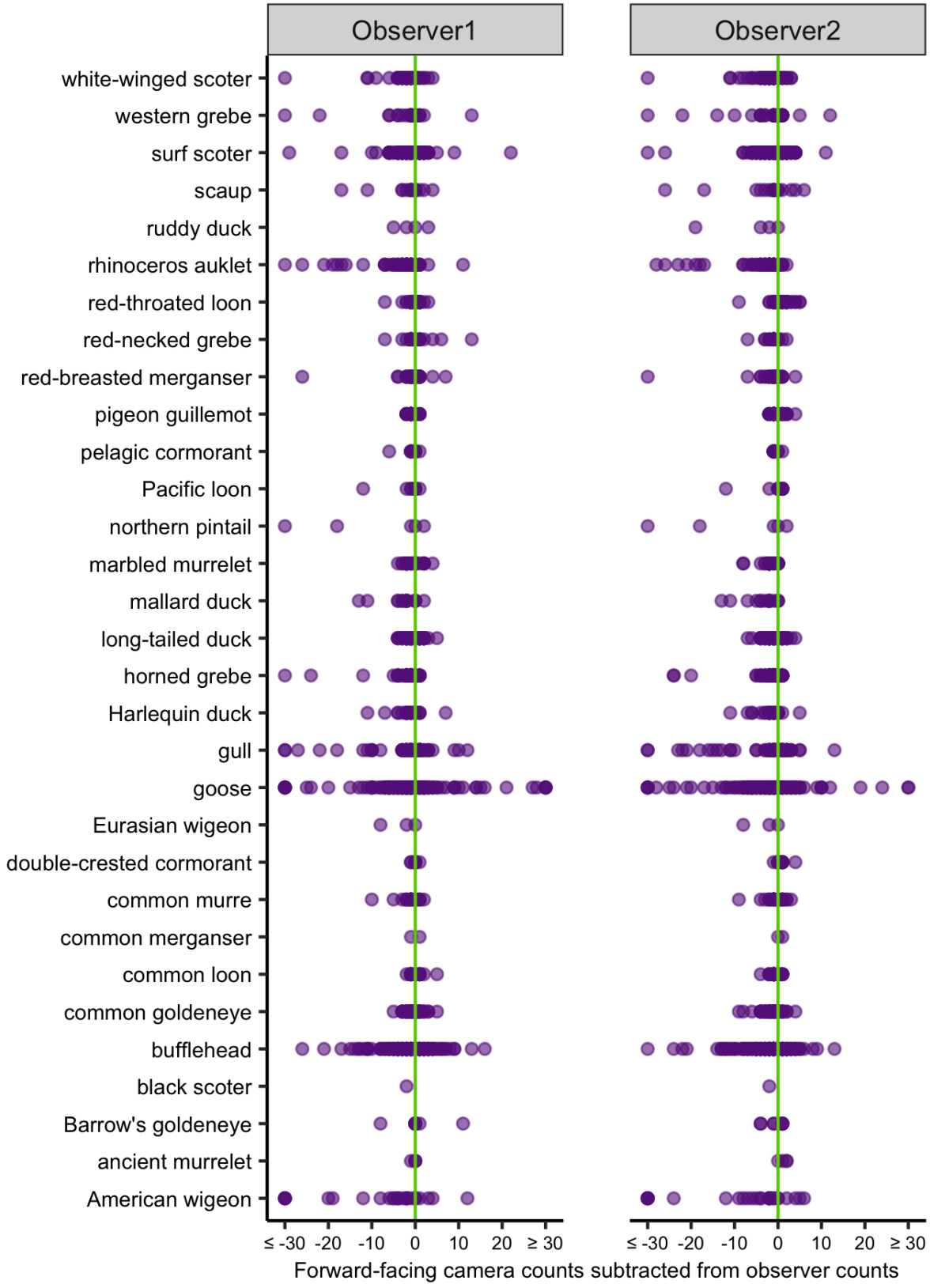
667 Figure 1
668



669 Figure 2

670

671



672 Figure 3

673