COUNTABILITY OF SANDHILL CRANES IN AERIAL SURVEYS

DOUGLAS H. JOHNSON
USGS Northern Prairie Wildlife Research Center, Douglas_H_Johnson@usgs.gov

JOHN W. SOLBERG
U.S. Fish and Wildlife Service

COURTNEY L. AMUNDSON
University of Minnesota

Follow this and additional works at: http://digitalcommons.unl.edu/nacwgproc

Part of the Behavior and Ethology Commons, Biodiversity Commons, Ornithology Commons, Population Biology Commons, and the Terrestrial and Aquatic Ecology Commons

http://digitalcommons.unl.edu/nacwproc/134

This Article is brought to you for free and open access by the North American Crane Working Group at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in North American Crane Workshop Proceedings by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
COUNTABILITY OF SANDHILL CRANES IN AERIAL SURVEYS

DOUGLAS H. JOHNSON,¹ USGS Northern Prairie Wildlife Research Center, 200 Hodson Hall, 1980 Folwell Avenue, Saint Paul, MN 55108, USA
JOHN W. SOLBERG, U.S. Fish and Wildlife Service, 3425 Miriam Avenue, Bismarck, ND 58501, USA
COURTNEY L. AMUNDSON, Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, 200 Hodson Hall, 1980 Folwell Avenue, Saint Paul, MN 55108, USA

Abstract: Aerial surveys are used to monitor populations of many wildlife species, including sandhill cranes (Grus canadensis). In addition to the usual problems of detectability (involving both availability and perceptibility), aerial surveys of concentrated animals are subject to countability issues; from a rapidly moving aircraft, observers cannot count or accurately estimate the number of animals in a large group. Calibration is sometimes performed in an effort to adjust aerial counts for incomplete detectability and countability by calculating the ratio of animals actually in a group to the number in the group estimated from the aircraft. Here we explore alternative, model-based approaches to the analysis of those adjustment ratios using aerial survey data of sandhill crane concentrations during 1978-2007 in the Platte River Valley of Nebraska. Ratios varied by year and by observer. In addition, the ratio varied with the actual size of the concentration. Modeling can be used to develop improved estimates of the ratio.

PROCEEDINGS OF THE NORTH AMERICAN CRANE WORKSHOP 11:89-97

Key words: aerial survey, countability, detectability, Platte River, sandhill crane, sightability.

Aerial surveys are often used to estimate or track populations of wildlife, particularly large mammals and birds (Caughley 1974, 1977). Large and conspicuous animals that are widely spaced over a known area (e.g., wintering whooping cranes, Grus americana) pose few if any special problems in such surveys. Typically, aerial surveys must overcome two primary concerns: lack of availability—animals are hidden or emit no detectable cues, and incomplete detectability—animals are available for detection, but the observer misses them (e.g., Johnson 2008). A variety of design and analytic methods have been devised to address some of these issues, including distance sampling (Buckland et al. 1993), multiple-observer sampling (Cook and Jacobson 1979), and time-to-detection sampling (Farnsworth et al. 2002, Allredge et al. 2007).

In some situations a third issue arises, which we term countability. If the animals are numerous, observers may find it difficult or impossible to enumerate all the animals during the brief pass in an aircraft. Just as availability and detectability can be confounded in many surveys of birds (Johnson 2008), countability can be confounded with either of them in an aerial survey. Problems of reduced accuracy are likely exacerbated when groups are large, animals are inconspicuous, or surveying conditions are suboptimal.

These conditions commonly occur during surveys of the midcontinental population of sandhill cranes (Grus canadensis). Accuracy of aerial surveys of birds is highly variable, and causes of such variability are poorly understood (Frederick et al. 2003). The objective of this paper is to explore the variation in countability of this surveyed population and to determine if there are methods to improve the analysis of survey results. The ultimate goals would be to improve accuracy of estimates by better accounting for countability and possibly to reduce survey effort. Although we focus on the single situation involving sandhill cranes along the Platte River, our results are of greater generality.

METHODS

Of the several populations of sandhill cranes (Meine and Archibald 1996), the mid-continental population is by far the largest, encompassing about a half-million birds (Sharp and Vogel 1992). Although birds in this population breed from Siberia, through Alaska and northern Canada, and east to Ontario and Quebec (G. L. Krapu et al., unpublished data), their defining characteristic is that they migrate through the Central Plains of North America. In particular, during their spring migration most of them stop for a period of time in mid to late March along the north or central Platte River valleys of central Nebraska (USFWS 1981, 1981).
Krapu et al. 1987).

Operational surveys of sandhill cranes along the Platte River have been conducted since 1972 (Ferguson et al. 1979, Benning and Johnson 1987). The method first used involved flying an aircraft along the river before dawn, when cranes are still on their riverine roosts. Two or 3 days were needed to survey the designated portion of the Platte River. Problems with this survey method included frequent poor weather and poor visibility during the early mornings, difficulties in seeing and counting all cranes present, and the possibility that some cranes already may have left their roosts to feed.

A different approach was evaluated in 1978 (Ferguson et al. 1979), which involved conducting surveys later in the day, when most of the cranes are in fields in the valley, foraging or resting; the method is still in use. The 2,174-km² area of the Platte River and North Platte River Valleys in which most of the cranes were thought to occur in was designated the survey area. That area was divided into 10 strata, based on similarities in crane densities in roosts along adjacent stretches of river during previous years.

Each stratum was divided into a number of north-south transects, each 0.8 km wide and of various lengths (range 4.4-25.6 km). A systematic sampling plan was used. About a fourth of the transects could be surveyed operationally; thus, a random number between 1 and 4 was chosen, and that-numbered transect and every fourth one thereafter was included in the sample. Most strata contained 4 or 5 surveyed transects (range 3-10).

A pilot accompanied by an observer flew an aircraft along the center line of each sampled transect. The pilot and observer each attempted to count all sandhill cranes on his or her side of the aircraft within 0.4 km of the center line. Densities of cranes were computed for each transect and projected to the entire stratum. After multiplying the average density by the area of a stratum, resulting estimates of population size were added across all strata.

Observed counts cannot be expected to be exact, for 2 reasons. First, cranes were rather cryptic, especially under certain ground and light conditions. Second, cranes often occurred in large flocks, making it impossible for observers to count them from rapidly moving aircraft. So any error reflected both imperfect detectability and inaccuracy in counting or estimating sizes of flocks (Fig. 1).

In an attempt to compensate mathematically for such errors, independent data on countability were obtained. For this endeavor, a third person in the aircraft photographed certain flocks of cranes. Efforts were made to photograph flocks in a range of sizes that were representative of those observed in the operational survey. Subsequently cranes in each flock were carefully counted on large-format photographs. Results were compared to the number of cranes that the pilot or observer estimated in the flock. A goal of 50 flocks per observer each year was established, but often not quite met due to a variety of logistic constraints (mean per year per observer was 41 flocks, range 23-56).

The current adjustment for countability is design-based; that is, only data from the photographed flocks in a particular year are used. We explore model-based adjustment, incorporating explanatory variables such as observer, year, and flock size. Although distance to object is a major influence on detectability in many surveys, it was not recorded in this survey and was not considered important, due to the large size of objects (sandhill cranes) and limited maximum range (half-width of transect, 0.4 km).

Under the current protocol, the ratio of total cranes counted on photographs to the total estimated in the same flocks (hereafter, Ratio) by each observer is computed. Recorded counts are adjusted by multiplying them by this Ratio, separately for each observer in each year. During the 1982-2007 period, Ratios averaged 1.27 and ranged from 0.80 to 2.32.

In general, observers tended to underestimate the size of flocks, although errors in estimating large concentrations can be appreciable and in either direction (Fig. 2). Many Ratios were near 1, with a few somewhat below 1 and a few even exceeding 2 (Fig. 3). Because of this asymmetry, we based our analyses on the natural logarithm of the Ratio. Here we explore 3 new approaches to the analysis of these countability data.

**A Mixed-model Approach**

Our first approach involved the notion of fixed and random effects in linear models (Littel et al. 2006). Fixed effects are explanatory variables for which the study includes all values of interest. Random effects are variables for which the study includes only a portion (ostensibly randomly chosen) of the values that the variable can assume. For example, year typically is treated as a random effect, because inference is desired...
Figure 1. Examples of concentrations of sandhill cranes in the Platte River Valley of Nebraska as photographed for comparison with counts estimated from aircraft, (a) a small flock of 57 fairly detectable sandhill cranes, estimated as 55 from the aerial survey, (b) concentration of 123 less-detectable sandhill cranes, estimated as only 75 from the aerial survey, (c) large flock of 934 readily detectable sandhill cranes, estimated as only 400 from the aerial survey. Pen lines and check marks were used to facilitate counting cranes on the photographs.
for more years than just the ones under study (even though years certainly are not randomly selected).

The response variable in our models was the logarithm of Ratio, the ratio of number of cranes counted on a photograph (Count) to the number estimated from the aircraft. We modeled the log of Ratio as a function of the number of cranes counted by the observer (log Count), year of the survey (Year, a class variable), observer (Observer, a class variable), and the experience of the observer (log Experience). Experience was defined as the number of years an observer had conducted this particular aerial survey of cranes (e.g., Experience = 1 for an observer’s first year surveying cranes). We further tested for quadratic effects of log of the Count (log^2 Count) to examine whether the log of the Ratio changed nonlinearly with the log of crane abundance. The 6 observers who surveyed cranes in only a single year confounded Experience and Observer effects, so we did not consider models containing both variables; rather, we treated those models as competing hypotheses in an exploratory analysis. We compared 2 models that included Year, log Count, log^2 Count, and either Observer or log Experience effects. We selected the variable (Observer or Experience) for which the model including that variable explained more of the variation in the Ratios.

Observer could justifiably be treated either as a random effect (i.e., if observers are representative of a population of potential observers) or a fixed effect (i.e., if interest is in only particular observers; note that the use of the same observers in several years suggests they were not selected randomly). Because neither assumption is

Figure 2. Counts of sandhill cranes recorded from aerial photographs of selected flocks against the size of the flock as estimated by observers, (a) pilot and (b) non-pilot, in 2007. Diagonal line represents perfect accord.
strictly correct, we addressed both possibilities by conducting 2 parallel analyses (F: Observer fixed, R: Observer random) and comparing selected models and coefficient estimates for each. Year was treated as a random effect; all other variables were treated as fixed effects.

An exploratory analysis revealed that Observer effects were stronger than any Experience effect, suggesting that the amount of experience with this crane survey did not completely account for the variability among observers. We next proposed 13 biologically plausible model structures \textit{a priori}; these formed our candidate model set (Table 1). We compared models using Akaike's Information Criterion (AIC; Burnham and Anderson 2002), so that the model with the lowest AIC value is the one best supported by the data. Also, $\Delta$AIC measures the difference between AIC values for a model under consideration and the model with lowest AIC value. Our sample size to parameter ratio exceeded 40, so we did not require the AIC small-sample-size correction in our analysis (Anderson et al. 2001, Burnham and Anderson 2002). We assessed the importance of each explanatory variable by examining its coefficient estimate along with its associated 95% confidence interval, and also by summing AIC weights of all models that included each variable. We model-averaged parameter estimates unless overwhelming

\begin{table}[h]
\centering
\begin{tabular}{lll}
\hline
Model number & Explanatory variables included\textsuperscript{a} & $K$\textsuperscript{b} \\
\hline
1 & Year + Obs + log C + log\textsuperscript{2} C + Obs$ \cdot$ log C + Obs$ \cdot$ log\textsuperscript{2} C & 64 \\
2 & Year + Obs + log C + log\textsuperscript{2} C + Obs$ \cdot$ log C & 52 \\
3 & Year + Obs + log C + Obs$ \cdot$ log C & 52 \\
4 & Year + Obs + log C + log\textsuperscript{2} C & 41 \\
5 & Year + Obs + log C & 40 \\
6 & Obs + log C + log\textsuperscript{2} C + Obs$ \cdot$ log C + Obs$ \cdot$ log\textsuperscript{2} C & 39 \\
7 & Year + log C + log\textsuperscript{2} C & 29 \\
8 & Year + log C & 28 \\
9 & Obs + log C + log\textsuperscript{2} C + Obs$ \cdot$ log C & 27 \\
10 & Obs + log C + log C & 15 \\
11 & Obs + log C & 14 \\
12 & log C + log\textsuperscript{2} C & 3 \\
13 & log C & 2 \\
\hline
\end{tabular}
\caption{Candidate model set for estimating count bias in aerial surveys of sandhill cranes.}
\end{table}

\textsuperscript{a} Obs represents observer; C is Count, the number of cranes estimated from aerial surveys.

\textsuperscript{b} Number of parameters in the model.
support (defined as AIC weight > 0.9) existed for a single model (Anderson et al. 2001).

A Bayesian Approach

Our second approach involved Bayesian hierarchical modeling. We modeled the log of Ratio as a linear function of the log Count, the squared log Count, Year effects, and Observer effects. The major difference between this Bayesian approach and the mixed-modeling approach is that the Bayesian paradigm treats all parameters, not just the “random” components, as having statistical distributions. Those distributions can reflect either uncertainty about parameters that are assumed to be fixed, or actual variation in parameters that are not assumed fixed.

We used “standard” objective Bayesian modeling in which we assumed no prior knowledge of any parameter. We used WinBUGS (Thomas 1994) to perform the analysis. We ran 50,000 simulations on our model to obtain parameter estimates, after discarding the initial 60,000 simulations to ensure convergence. We checked density functions and autocorrelation plots to assess model convergence.

An Empirical Bayes Approach

Empirical Bayes can be viewed as 1 way of incorporating prior knowledge or information from similar situations into an inference problem. For our purposes, we considered a simple empirical Bayes estimator of the form

$$\text{Ratio}_{EB}(t) = w \cdot \text{Ratio}(t) + (1 - w) \cdot \text{Mean Ratio}$$

where Ratio$_{EB}(t)$ is the empirical Bayes estimator for a particular observer in year $t$, Ratio$(t)$ is the observed value for that observer and year, Mean Ratio is the overall mean (or mean from prior years), and $w$ reflects the relative precision of Ratio$(t)$ versus Mean Ratio:

$$w = \frac{\text{Var(Mean Ratio)}}{\text{Var(Mean Ratio)} + \text{Var(Ratio}(t))}$$

Accordingly, if Ratio$(t)$ is a good estimator for a particular observer and year (i.e., Var(Ratio$(t)$) is small), $w$ will be near 1 and the empirical Bayes estimator will be nearly identical to the observed value for that observer and year. Conversely, if Ratio$(t)$ is poorly estimated, the empirical Bayes estimator will be close to the long-term mean, Mean Ratio, reflecting a lack of confidence in the observed Ratio.

RESULTS

Observers generally tended to underestimate flock size (all expected Ratios >1; Fig. 4). Treating observer as random or fixed produced only slightly different best-fitting models. The most-supported model in both analyses indicated that the Ratio varied by Year, Observer, and a quadratic function of log Count. In addition, treating Observer as a random effect supported interactions between Observer and linear ($\Delta$AIC = 0) and quadratic ($\Delta$AIC = 1.20) effects of Count, whereas there was less support for those interactions when Observer was treated as a fixed effect ($\Delta$AIC = 4.3 and 5.6, respectively).

Parameter estimates were similar between the best-supported Fixed and Random models (Table 2). The Year coefficients ranged from -0.13 (Fixed-effects model: 1985) to 0.24 (Random-effects model: 1995). Observer coefficients ranged from -0.21 (Observer W) to 0.25 (Observer U). Confidence intervals (95%) for coefficients of log Count and log$^2$ Count did not include 0. Predicted Ratios followed similar trends as mixed models (Fig. 4) with observers, on average, underestimating a representative flock size of 200 birds (mean Ratio = 1.24), ranging from 0.92 (Observer W in

Figure 4. Predicted values of Ratio from various models, with model-averaged values for fixed-effects and random-effects models.
Table 2. Estimates of intercept and regression coefficients (standard errors in parentheses) for the best-supported models of sandhill crane countability ratio. Years and Observers are arbitrarily selected examples.

<table>
<thead>
<tr>
<th></th>
<th>Observer random</th>
<th>Observer fixed</th>
<th>Bayesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.227 (0.14)</td>
<td>-0.436 (0.13)</td>
<td>-0.226 (0.13)</td>
</tr>
<tr>
<td>log Count</td>
<td>0.245 (0.05)</td>
<td>0.244 (0.05)</td>
<td>0.243 (0.05)</td>
</tr>
<tr>
<td>log’ Count</td>
<td>-0.026 (0.01)</td>
<td>-0.025 (0.01)</td>
<td>-0.026 (0.01)</td>
</tr>
<tr>
<td>Year 1982</td>
<td>0.011 (0.05)</td>
<td>0.016 (0.05)</td>
<td>0.013 (0.06)</td>
</tr>
<tr>
<td>1990</td>
<td>-0.020 (0.05)</td>
<td>-0.023 (0.05)</td>
<td>-0.025 (0.06)</td>
</tr>
<tr>
<td>2000</td>
<td>-0.092 (0.04)</td>
<td>-0.093 (0.04)</td>
<td>-0.110 (0.05)</td>
</tr>
<tr>
<td>2007</td>
<td>-0.011 (0.04)</td>
<td>-0.010 (0.04)</td>
<td>-0.020 (0.05)</td>
</tr>
<tr>
<td>Observer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>-0.150 (0.08)</td>
<td>-0.378 (0.12)</td>
<td>-0.131 (0.06)</td>
</tr>
<tr>
<td>S</td>
<td>-0.211 (0.08)</td>
<td>-0.439 (0.12)</td>
<td>-0.201 (0.05)</td>
</tr>
<tr>
<td>E</td>
<td>-0.042 (0.10)</td>
<td>-0.274 (0.14)</td>
<td>-0.035 (0.08)</td>
</tr>
<tr>
<td>B</td>
<td>-0.015 (0.08)</td>
<td>-0.218 (0.13)</td>
<td>-0.023 (0.05)</td>
</tr>
</tbody>
</table>

For most years and observers, empirical Bayes estimates were very similar to the observed Ratios (Fig. 5). Substantial differences occurred only when the observed ratios were rather extreme, greater than about 1.5; in those instances the empirical Bayes estimates were shifted noticeably toward the overall average of 1.27. The weights reflected the fairly high precision of the observed ratios (average variance was 0.014) in contrast to the substantial variation among years and observers (variance = 0.108).

**DISCUSSION**

The spring count of sandhill cranes along the Platte River is a valuable monitoring tool for an important population of birds. That such a large fraction of the entire population is present at the same time enables managers to estimate its size much more accurately than most populations, game or nongame. Yet, the aerial counts of cranes in the Platte River survey are imperfect, generally lower than actual numbers (as also found by Frederick et al. [2003] for simulated wading birds and by Pearse et al. [2008] for waterfowl decoys). Errors were especially large, and unpredictable, for large flocks of cranes. Errors, as measured by the difference between numbers of birds in a flock counted from aerial photographs and the numbers estimated from an aerial survey, varied by observer, year, and size of the flock.

![Figure 5. Differences between empirical Bayes estimates and observed Ratios in relation to observed Ratio.](image-url)
This variability reduces the value of the survey counts even as an index of population size. Fortunately, concurrent photography of selected flocks permits an adjustment to the observed counts to enhance their accuracy through the use of these ratios. Current analytic methods involve estimating this ratio separately for each observer in each year, under conditions similar to those occurring in the survey year. The current protocol does not incorporate any adjustment based on flock size, but efforts are made to sample flock sizes that are representative of those found during the survey.

The current method is totally design-based; that is, it is based only on data collected in a particular year by a particular observer under a prescribed sampling design. The major assumption is that the design is adequate. Alternatives to design-based estimators are model-based estimators, which rely on assumptions about the variable of interest itself, such as how it relates to other variables. We illustrated two approaches, mixed models and Bayesian, to developing such models. The empirical Bayes approach effectively combines a design-based estimate with a model-based estimate, essentially getting the “best of both worlds.”

The mixed-model and Bayesian approaches generated similar results. Both Year and Observer effects were included in the best-supported mixed models, and 95% confidence intervals for coefficient estimates in our Bayesian analysis did not include zero, suggesting that both variables had some influence on countability. The ability of observers to count cranes also varied by the size of the concentration. These results demonstrate the value of annual countability estimates, obtaining photographs of crane groups that vary in size and are representative of those surveyed, and accounting for observer differences. Coefficient estimates and predicted values were remarkably similar for both analyses, and either mixed or Bayesian models could be used in future analyses. The benefit of the Bayesian approach is its ability to incorporate prior information into the analysis, which for long-term datasets such as this could improve precision and reduce the number of surveys needed.

The empirical Bayes estimates were little different from the observed Ratios, primarily because sample sizes on which observed Ratios were based were fairly large, averaging 41 flocks per observer and year. The only appreciable differences between empirical Bayes estimates and observed Ratios occurred when the latter were extreme; the empirical Bayes procedure shifted them toward the overall mean. A benefit of empirical Bayes estimators is that extreme values are somewhat discounted. Another advantage is that empirical Bayes estimates could be generated in years in which the number of photographed flocks was small, or even zero.

The simple empirical Bayes estimator we examined weighted the observed Ratio, for a particular observer in a particular year, with the overall average Ratio. Somewhat more sophisticated estimators could be developed by weighting the observed Ratio with a different value. In particular, estimators from the mixed-model or Bayesian approaches could be used to replace the overall average Ratio. In a detailed examination of empirical Bayes estimators for waterfowl populations based on aerial surveys with some on-the-ground double sampling, Johnson (1986) found that simple estimates weighted with the mean performed virtually as well as estimates weighted with values based on covariates. Those results, however, may not be general.

Our objective was not to propose a particular alternative to the estimation method currently in use. Rather, it was to examine factors that influence countability and to explore possibilities for improved estimation of the mid-continental sandhill crane population, but with potential generalization to other situations.

In the sandhill crane example, current methods appear satisfactory, primarily because large numbers (averaging 41) of photographed flocks were available for each observer in each year. If photography was unavailable or samples were small, in any particular year, methods discussed here would provide viable alternatives.

With such a long-term data base available (1978-2007), it is reasonable to ask if there is sufficient structure in the data to improve the accuracy of adjustments or even to permit a reduction of data collection in the future. These changes do not seem warranted until improved model-based estimators become available. The alternative models we examined incorporated only non-mechanistic explanatory variables, such as year and observer. Without collecting data in a particular year, determining an effect for that year would be very difficult.

Developing mechanistic models of countability bears investigation. Variables that affect countability in meaningful ways should be incorporated into the models. For example, effects due to year probably reflect annual
differences in ground conditions (partial snow cover versus open ground, for example) and possibly sunlight and cloud cover. It may be worthwhile to gather such information during the course of future crane surveys. Mechanistic models, which are based on realistic processes and causal relationships, can be more useful and general than phenomenological models, which rely on observed associations among variables. For now, however, it appears that accurate estimation of this population of sandhill cranes will require some photography to help calibrate countability.

ACKNOWLEDGMENTS

D. S. Benning and E. L. Ferguson of the U.S. Fish and Wildlife Service developed the surveys of sandhill cranes in the Platte River Valley. We are grateful for comments on the manuscript from D. S. Benning, T. L. Shaffer, an anonymous referee, and the editor.

LITERATURE CITED


