Estimating sightability of greater sage-grouse at leks using an aerial infrared system and N-mixture models

Authors: Coates, Peter S., Wann, Gregory T., Gillette, Gifford L., Ricca, Mark A., Prochazka, Brian G., et al.

Estimating sightability of greater sage-grouse at leks using an aerial infrared system and *N*-mixture models

Peter S. Coates, Gregory T. Wann, Gifford L. Gillette, Mark A. Ricca, Brian G. Prochazka, John P. Severson, Katie M. Andrle, Shawn P. Espinosa, Michael L. Casazza and David J. Delehanty

P. S. Coates (https://orcid.org/0000-0003-2672-9994) ✉ (pcoates@usgs.gov), G. T. Wann, M. A. Ricca, B. G. Prochazka, J. P. Severson and M. L. Casazza, U.S. Geological Survey, Western Ecological Research Center, Dixon Field Station, Dixon, CA 95620, USA. – G. L. Gillette and D. J. Delehanty, Dept of Biological Sciences, Idaho State Univ., Pocatello, ID, USA. – K. M. Andrle, Nevada Dept of Wildlife, Sagebrush Ecosystem Program, Carson City, NV, USA. – S. P. Espinosa, Nevada Dept of Wildlife, Reno, NV, USA.

Counts of grouse present at leks (breeding grounds) during spring are widely used to monitor population numbers and assess trends. However, only a proportion of birds available to count are detected resulting in a biased population index. We designed a study using an aerial integrated infrared imaging system (AIRIS) and experimental pseudo-leks to quantify sightability (proportion of birds detected) of conventional ground-based visual (GBV) surveys for greater sage-grouse *Centrocercus urophasianus*. Specifically, we calibrated AIRIS at pseudo-leks composed of known numbers of captively-raised birds, primarily ring-necked pheasant *Phasianus colchicus*. We then carried out AIRIS and GBV surveys, simultaneously, on nearby sage-grouse leks, allowing us to model AIRIS and GBV sightability. AIRIS detected ~93% of birds on pseudo-leks while GBV detected ~86% of sage-grouse on leks. Thus, the ground count observation error was −14% from the 'true' number of male sage-grouse attending the leks. We also found sagebrush cover decreased sightability for GBV counts but did not influence sightability by AIRIS. Because standard GBV protocols typically make repeated counts of sage-grouse in a single morning, we also modeled repeated GBV counts using *N*-mixture models and found an 88% sightability, which was nearly the same as GBV sightability from the AIRIS analysis. This suggests that the use of repeated morning counts can potentially account for imperfect detection in the standard GBV surveys currently implemented. We also provide generalized correction values that could be employed by resource managers using either GBV or AIRIS to better estimate 'true' numbers of sage-grouse attending leks within similar environments to this study. The findings and interpretation presented can help guide effective monitoring protocols that account for observation error and improve accuracy of data used for population trend and abundance estimation.

Keywords: aerial survey, *Centrocercus urophasianus*, detection, greater sage-grouse, infrared, lek counts, *N*-mixture model, observation error, sightability

A primary goal in designing wildlife monitoring surveys is to develop data collection protocols capable of informing managers of changes in population abundance over time (Nichols 1991, Williams et al. 2002). Count data obtained from leks (traditional breeding grounds) of greater sage-grouse *Centrocercus urophasianus* (hereafter sage-grouse) have been a primary source of information used to assess population trends since the 1940s when lek monitoring first began (Patterson 1952, Connelly and Schroeder 2007). Sage-grouse numbers have declined throughout their range since the 1950s averaging an annual decrease of approximately 0.85% per

year (Garton et al. 2015, WAFWA 2015). The species currently occupies roughly one half of its historic distribution (Schroeder et al. 2004), and with further habitat losses in the sagebrush biome likely in coming decades (Coates et al. 2016, Smith et al. 2016, Green et al. 2017) and upcoming consideration for protection under the Endangered Species Act, improved information on populations from lek count data will be central to sage-grouse conservation.

Accordingly, accounting for intrinsic biases in count data due to observation error would improve estimation of true demographic patterns resulting from environmental change. Like many types of survey data used as population indices, lek-count data is often scrutinized as a biased representation of true population numbers (Beck and Braun 1980, Applegate 2000). Part of the uncertainty in lek-count data results from imperfect observation rates of individual grouse during conventional ground-based visual (GBV) surveys lead-

ing to variable and biased estimates. This long-recognized problem has led to several evaluations of lek counts and the factors that affect their accuracy (Fremgen et al. 2016, Baumgardt et al. 2017). Nichols et al. (2009) described four distinct components to detectability in count surveys. The first component is the probability (p) that an individual's home range overlaps the sampling unit (p_s). Because the sampling unit for sage-grouse is the lek site, home ranges of all sage-grouse within those populations are assumed to intersect the lek. However, not all leks on the landscape are known and counted (Shyvers et al. 2018), so individuals associated with those leks will not be counted. A second component is the probability an individual is present at the sampling unit during the time of survey (p_p). For example, individual attendance on leks can vary within mornings (Monroe et al. 2016), throughout the breeding season, and among years (Blomberg et al. 2013, Fremgen et al. 2019, Wann et al. 2019). Given the individual is present at the sampling unit, the third and fourth components are the probability of being available for detection during the count (p_a, e.g. not obscured by vegetation) and probability of detection conditional on availability (p_d; e.g. accurate count of unobscured individuals), respectively, and are collectively referred to as sightability ($p_a p_d$; Fremgen et al. 2016, Baumgardt et al. 2017). Thus, sightability can also be interpreted as the proportion of available individuals which are observed.

Infrared video-surveying is an emerging technology that is particularly useful for sensing endothermic animals (Havens and Sharp 1998) and shows promise as a tool for monitoring upland gamebird populations. The spectral signatures of these animals in the infrared wavelength (i.e. heat) is generally distinct from their environment. Infrared technology has been applied in wildlife studies for decades but has been limited primarily to uncooled infrared sensors (Gillette et al. 2013). Current systems include a single, gyroscopically-stabilized unit (to reduce motion blur in a moving aircraft) which contains both an infrared camera that is cryogenically-cooled (to improve measurement precision of spectral intensity) and a high-resolution camera in the visible spectra allowing for high-magnification zooming (e.g. to distinguish between male and female sage-grouse). We refer to this combination of technologies into one device operated from fixed-wing aircraft as an aerial integrated infrared imaging system (AIRIS).

Several state agencies have initiated lek surveys using a combination of AIRIS and GBV counts within their monitoring programs. However, without accounting for differences in sightability among survey types, population trend estimates can be confounded by mixed survey methodology which may misinform population performance and ultimately management actions. Rigorous measurement of sightability differences between AIRIS and GBV surveys may provide appropriate adjustment to lek counts and improve accuracy of trend estimates. Furthermore, AIRIS can also be used to assess the accuracy of GBV counts, as we demonstrate in this study. Although past studies have compared similarities between counts recorded with infrared cameras to those collected on the ground for sharp-tailed grouse *Tympanuchus phasianellus* and sage-grouse

(Gillette et al. 2013, 2015), sightability of newer AIRIS has not been formally estimated.

Conversely, AIRIS surveys can be relatively expensive and may not be feasible for extensive surveying (Gillette et al. 2015). Therefore, agencies charged with monitoring sage-grouse populations over large areas may be interested in cost-effective alternatives to AIRIS to account for imperfect detection in lek counts. The *N*-mixture model developed by Royle (2004) offers one promising alternative because it only requires repeated counts during a period of population closure (i.e. no movement in and out of survey site during time counts occur), which is a crucial assumption. *N*-mixture models estimate sightability and true population abundance (*N*; i.e. animals available at the survey site for observation). The *N*-mixture model has been used to estimate the male population of sage-grouse at leks using repeated surveys conducted throughout the breeding season (McCaffery and Lukacs 2016). However, those estimates may be difficult to interpret given that the closure assumption is likely violated due to variation in attendance rates across survey days (Fremgen et al. 2019, Wann et al. 2019, Monroe et al. 2019). In contrast, repeated counts that occur across a relatively short period in morning hours should satisfy the closure assumption, although the estimated lek abundance will be specific to the day the counts occurred and will change by survey day given the variability in lek attendance. Nonetheless, estimating the day-specific abundance is analogous to the conventional survey estimate of using the maximum daily count and is precisely of interest in our study.

Sightability can vary considerably among lek surveys (Fremgen et al. 2016, Baumgardt et al. 2017), meaning the error in raw lek counts (i.e. the proportion of the true number of birds missed) may include substantial bias, and there is a lack of consensus on how to account for these errors. In this study, we quantify sightability error through an experimental approach that combined emerging technology with traditional methods, and then offer multiple options to account for error to managers assessing and collecting lek data. Our first objective was to estimate AIRIS sightability by quantifying the proportion of a known number of captively-raised galliform birds serving as proxies for sage-grouse on pseudo-leks. Our second objective was to estimate overall GBV error by combining the sightability of GBV counts relative to simultaneous AIRIS counts with the AIRIS sightability from objective one. We also assessed the effects of environmental factors such as sagebrush cover (serving as an index visual obstruction) and the time since sunrise (serving as an index of degree of daylight) on sightability for GBV and AIRIS counts. We were particularly interested in differences in the effects of environmental predictors and how sightability varied between the two types of survey counts. Our third objective was to derive an alternative estimate of sightability using *N*-mixture models from repeated within-morning GBV counts. We compared the *N*-mixture estimate to the result of objective two and discuss the practical potential of all three methods (i.e. GBV, AIRIS, *N*-mixture GBV) for integrated sage-grouse monitoring designs.

Methods

Study area

We surveyed sage-grouse leks and pseudo-leks located in northeastern California (Lassen County; lattitude: 40°58′N, 120°27′W), eastern Idaho (Clark, Fremont and Jefferson Counties; 43°99′N, 111°96′W), southwestern Idaho (Owyhee County; 42°98'N, 116°50'W), northeastern Nevada (Elko County; 41°38′N, 115°68′W), and north-central Nevada (Eureka and Lander counties; 40°08′N, −116°36′W) over three breeding seasons during April and May, 2015–2017 (Fig. 1). Vegetation communities in our study areas were typical of the sagebrush ecosystem of the northern Great Basin. Dominant shrubs included several species of sagebrush (primarily *Artemisia arbuscula*, *A. nova* and *A. tridentata*), rabbitbrush *Ericameria nauseosa* and *Chrysothamnus viscidiflorus*, snowberry *Symphoricarpos* spp., western serviceberry *Amelanchier alnifolia*, and antelope bitterbrush *Purshia tridentata*. Forbs and grasses were largely dormant when our surveys occurred, but leks were generally snow free.

Study design

We surveyed birds at two different location types: 1) active leks consisting of wild sage-grouse and 2) pseudo-leks consisting of captively-raised ring-necked pheasant *Phasianus colchicus* (hereafter, pheasant) or chukar partridges *Alectoris chukar* (hereafter, chukar) which were tethered to the ground. Pseudo-leks contained known numbers of birds which provided a true population size for deriving AIRIS sightability. Pseudo-lek locations were randomly generated (given the following constraints) between 500 and 600 m from a real sage-grouse lek. We chose 500 m as a minimum because leks were clearly distinct from the aircraft at this distance. We chose 600 m as a maximum so general habitat characteristics were similar and flight time for the aircraft between locations was minimized, allowing for similar levels of ambient infrared radiation between both survey types. We targeted areas with percent shrub cover ≤20% within the boundaries of the pseudo-lek, which was similar to our real sage-grouse leks. Additionally, pseudo-leks had to be relatively close to an unimproved or two-track road (≤100 m) to facilitate transportation and placement of pheasant and chukar.

We placed a known number of captively-raised pheasant or chukar at pseudo-leks as a proxy for sage-grouse to estimate a proportion of birds observed by AIRIS. We chose these morphologically different galliform species to create size variation and to avoid unknown idiosyncrasies of a single species. We rationalized that lack of difference in AIRIS sightability between pheasant and chukar would indicate that sage-grouse share similarities in sightability. AIRIS at sage-grouse leks and their paired pseudo-leks occurred on the same mornings so weather and visibility conditions
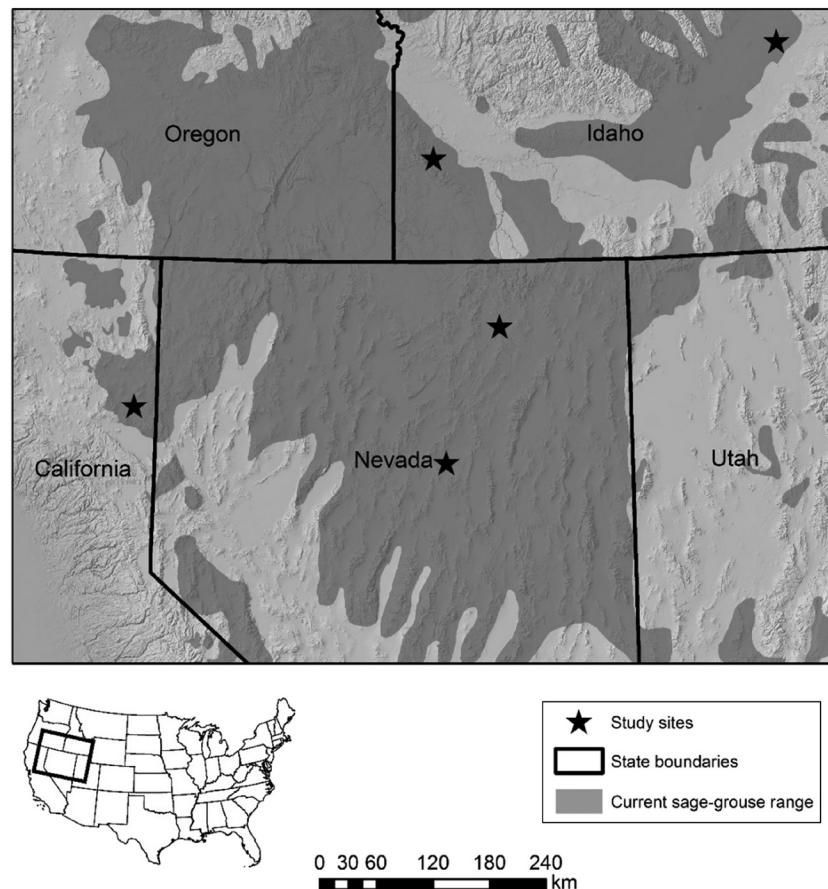

Figure 1. Greater sage-grouse *Centrocercus urophasianus* study areas where aerial integrated infrared imaging system (AIRIS) and ground-based visual (GBV) surveys were assessed for sightability at real and pseudo-leks in California, Nevada and Idaho from 2015 to 2017.

were similar. Chukar and pheasant were not mixed on the same pseudo-lek (i.e. only one species occurred on a given pseudo-lek).

Placement of individual pheasant and chukar on pseudo-leks (i.e. distance from center point) followed observed patterns of sage-grouse locations on real leks digitized in a geographic information system (GIS) from infrared images recorded in Nevada and Idaho. We then measured distances of each individual sage-grouse to the geometric mean of all sage-grouse present in the image (mean = 16.3 m; standard deviation = 9.4 m) using Euclidian distance tool in ArcMap 10.3. We used these measured distances to estimate Gamma distribution parameters (shape $\alpha = 2.4$; rate $\theta = 0.038$) using the 'MASS' package (Venables and Ripley 2002) in program R (<www.r-project.org>). We chose Gamma distribution because sage-grouse locations were clustered on leks and this distribution was skewed (Kéry 2010). We determined the pseudo-lek size by sampling from a normal distribution based on sage-grouse lek counts from across Nevada reported to the Nevada Dept of Wildlife in 2015. To determine bird placement relative to the pseudo-lek center, we sampled a distance from the Gamma distribution and randomly selected a directional azimuth for each bird. Each pseudo-lek bird was tethered to a stake at the pre-determined location prior to the surveys using paracord attached to its tibiotarsus (see Supplementary material Appendix 1 more information).

Lek counts

AIRIS counts were recorded at pseudo-leks and nearby real sage-grouse leks sequentially during the morning on survey days. A detailed description of the AIRIS technology and methods used in this study is provided in Supplementary material Appendix 2. We carried out double-blind GBV counts at real leks using two independent observers on the ground simultaneous with AIRIS surveys (Supplementary material Appendix 3).

GBV counts were conducted over two periods: 1) before the AIRIS plane arrived, and 2) during the plane visit. Within each of these primary periods, observers counted and recorded the number of male, female and unknown (i.e. sex could not be determined) grouse three times over a period of 10–15 min. Conducting three successive counts is consistent with most state agency lek count protocols (Connelly et al. 2003). For all double-blind ground counts, we randomly selected either the first or second observer's maximum count of males recorded simultaneous to the AIRIS survey with two exceptions. First, if grouse were visibly disturbed by the plane (e.g. stopped displaying and hid behind or under shrubs), we used the count that occurred prior to the AIRIS survey. Counts obtained before and during the AIRIS survey were highly correlated (r = 0.97), indicating the choice was likely to have minimal bias. In addition, we checked the influence of plane activity on our estimates by re-running models using only the GBV count data recoded prior to the AIRIS survey. Second, at Idaho sites in 2016, we used the total number of grouse from our GBV survey, because the AIRIS survey did not distinguish between males and females. We accounted for this difference in count type in the model.

Sightability modeling using AIRIS and GBV

We used a Bayesian modeling framework to simultaneously estimate AIRIS and GBV sightability from 1) true numbers of birds deployed to pseudo-leks, 2) AIRIS counts of pseudo-leks, 3) AIRIS counts of sage-grouse leks, and 4) GBV counts of sage-grouse leks. This framework allowed for parameter-sharing across multiple models, which provided a unique opportunity to estimate GBV sightability. The first two equations of our model formulated a calibration for AIRIS surveys using pseudo-leks (P). A Poisson distribution was specified to model counts as:

$$n_{AP,i} \sim \text{Poisson}\left(\lambda_{P,i}\right) \tag{1}$$

$$\lambda_{P,i} = \omega_{P,i} n_{TP,i} \tag{2}$$

Here, $n_{AP,i}$ is the number of birds on pseudo-lek i counted from the plane, and the rate $\lambda_{P,i}$ is a function of the product of the true number of birds located on each pseudo-lek ($n_{TP,i}$) and a proportional variable ($\omega_{P,i}$). The proportional variable allowed for proportions >1 (overcounting). Thus, ω_P represented the AIRIS sightability parameter and was modeled as a deterministic function of lek-level covariates (J) using a log-link function as:

$$\log\left(\omega_{P,i}\right) = \alpha_P + \sum_{j=1}^{J} \beta_{P,j} X_{P,ij} \tag{3}$$

Therefore, ω_P represented the proportion of birds on pseudo-lek i recorded by AIRIS cameras (e.g. an estimate <1 suggests AIRIS cameras on average undercounted the true number of birds), and this proportion was determined by covariates and associated model parameters including the intercept (α_P) and slope coefficients ($\beta_{P,j}$).

The fourth and fifth equations established the relationship between GBV counts of sage-grouse on real leks and predicted 'true' numbers of sage-grouse. We derived separate posterior distributions of ω_P based on the conditions observed at real leks (R), which we refer to as $\hat{\omega}_{P,i}$ and divided that value into the number of sage-grouse observed from the air for each real lek ($n_{AR,i}$). We added a constant (C) of 0.01 to $n_{AR,i}$ to avoid taking the log of 0 when AIRIS counts failed to observe any birds. We assigned a Poisson distribution to the number of sage-grouse observed on the ground $n_{GR,i}$ as:

$$n_{GR,i} \sim \text{Poisson}\left(\lambda_{R,i}\right) \tag{4}$$

$$\lambda_{R,i} = \omega_{R,i} \left(\frac{n_{AR,i} + C}{\hat{\omega}_{P,i}} \right) \tag{5}$$

Thus, $\omega_{R,i}$ represented the proportion of sage-grouse recorded on the ground relative to the predicted 'true' number, providing a GBV sightability parameter. In parallel, we modeled $\omega_{R,i}$ as a deterministic function of lek-level covariates (J) using a log link function as:

$$\log\left(\omega_{R,i}\right) = \alpha_R + \sum_{j=1}^{J} \beta_{R,j} X_{R,ij} \tag{6}$$

4

We specified vague priors in terms of mean and precision (i.e. inverse-variance) for all model coefficients (intercepts and slopes) including $\alpha_T \sim$ normal(0, 0.0001) and $\beta_T \sim$ normal(0, 0.0001), and subscripts denote coefficients estimated for real or pseudo-leks ($T = \{R, P\}$).

Covariates

We considered several covariates as potentially influencing the accuracy of AIRIS and GBV counts. However, prior to fitting all covariates, we tested the assumption that captively-raised pheasant and chukar were equally detectable by AIRIS using a model that included only an intercept and coefficient for the two-level species effect (1 = pheasant, 0 = chukar). An estimated coefficient for the species effect with 95% credible interval (CI) overlapping 0 supported similar detectability between the species by AIRIS.

A covariate for count type (i.e. males only or both males and females) was considered for the GBV sightability model. Additionally, we considered temporal effects (minutes before or after sunrise at which count occurred, i.e. 'time since sunrise'), and concealment effects (topographic roughness and shrub canopy cover) for both the GBV sightability and AIRIS sightability models. We calculated time since sunrise for each lek location and date that a count occurred using the spatial package 'sp' (Bivand et al. 2013) in program R. We also calculated average shrub canopy cover from 30-m resolution National Land Cover Database Shrubland Products (NLCD; Xian et al. 2015) and topographic roughness as the variance in elevation from a 30-m digital elevation model (Riley et al. 1999) within 100 m of leks using the zonal statistics tool in ArcMap 10.3.

We first estimated AIRIS and GBV sightability without environmental effects but accounting for count type (see above) and reported estimates of GBV sightability of males attending leks. We then estimated sightability accounting for covariate effects and predicted the average sightability while holding the habitat characteristics at the mean values for real leks. Sightability was not constrained between 0 and 1 because, although rare, overcounting sometimes occurred in AIRIS surveys at pseudo-leks.

Our full AIRIS sightability model included an intercept and four covariates (species, time since sunrise, shrub cover and topographic roughness), and our full GBV sightability model included an intercept and four covariates (count type, time since sunrise, shrub cover and topographic roughness). Covariates were considered supported by data if 95% CI of estimated coefficient (β) did not overlap 0. We also evaluated support based on the posterior probability of nonzeroness derived from a stochastic search variable selection (SSVS) method (George and McCulloch 1996). Specifically, we assigned a Bernoulli prior with probability of inclusion of 0.5 and derived a posterior probability of β being included in the model. This value represents how likely $\beta \neq 0$ given the data. We considered evidence substantial for values >0.6, marginal for 0.5–0.6, and deficient for <0.5.

Sightability N-mixture modeling

In addition to the sightability model using AIRIS data in conjunction with GBV data, we analyzed repeated counts from GBV data only collected at real leks as an alternative

approach using a basic binomial N-mixture model (Royle 2004). The purpose of this analysis was to compare sightability estimates between the two approaches and provide wildlife managers with alternative methods in accounting for observation error using repeated count designs. As previously described, during sage-grouse lek surveys, GBV observers recorded three repeated ground counts simultaneously with the AIRIS counts. N-mixture models were fit to the repeated GBV counts during single morning surveys. Thus, unlike the GBV-AIRIS sightability analysis which used the maximum GBV count, the N-mixture model analysis used all three GBV counts recorded during a survey.

For each real lek, we randomly selected one of the two observers and used their repeated GBV counts. We modeled counts at real lek i during count period j as arising from a binomial distribution as $y_{ij} \sim$ binomial(N_i, p), where N_i is abundance at lek i, which is a latent state estimated from the repeated counts. The parameter p in the N-mixture model can also be thought of as the probability of detecting an individual conditional on availability ($p_a p_d$; i.e. sightability) on a given count. Because these surveys were conducted in a single morning over a relatively short period of time, the component p_p (i.e. probability of presence) was not included because bird movement into and out of leks was not expected. Thus, sightability and abundance were conditional on the set of birds on lek during this time frame (Nichols et al. 2009). We fit simple intercept structures for both N and p using a log and logit link, respectively, and specified vague priors for both intercepts as $\beta_0 \sim$ normal(0, 0.0001). Because state agencies generally report the maximum count when multiple counts occur in a morning, and p by itself is not informative for datasets which only report the maximum from repeated counts, we also calculated a derived maximum sightability, p_D, as:

$$\frac{\sum_{i=1}^{I} \max\left(y_{ij}\right) N_i^{-1}}{I}$$

In this equation, every lek has the maximum of its repeated counts divided by its estimated abundance, and the total summation of this value is divided by the total number of leks (I) to obtain an average, represented as p_D. Only lek counts with >1 displaying male were used in the N-mixture analysis (i.e. the maximum of repeated counts had to have 2 or more males recorded).

Model implementation

All models were fit using the package 'R2jags' (Su and Yajima 2015) in Program R, which interfaced with the MCMC sampler program JAGS (ver. 4.2.0; Plummer 2003). We monitored three posterior chains over 20 000 MCMC iterations, the first 5000 of which were discarded as burn-in. Convergence of the marginal posterior distributions were assessed using the Brooks–Gelman statistic, $\hat{R}$ (Brooks and Gelman 1998). Values of $\hat{R} > 1.1$ suggest lack of convergence. We ran the AIRIS and GBV sightability models simultaneously and saved output from the three MCMC chains for parameter inference. We summarized statistics (i.e. median and 95% CI) from the posterior marginal

Table 1. Number of pseudo-leks (ring-necked pheasant *Phasianus colchicus* and chukar *Alectoris chukar*) and real leks (greater sage-grouse *Centrocercus urophasianus*) surveyed by site and year to quantify sightability using ground visual surveys and aerial infrared surveys within 6 field sites in the Great Basin during 2015–2017. Pseudo-lek counts were not conducted in 2015.

Site	No. of psuedo-leks	No. of real leks	Pseudo-lek counts		Real lek counts		
			2016	2017	2015	2016	2017
Northeastern California	10	5	0	20	0	0	9
Eastern Idaho	9	13	9	0	0	13	0
Southwestern Idaho	5	8	5	0	0	8	0
Northeastern Nevada	24	18	20	15	0	17	10
Northcentral Nevada	0	11	0	0	11	0	0
Total	48	55	34	35	11	38	19

distributions for parameters monitored in our models. Derived parameters were calculated from the saved MCMC output from both the pseudo-lek and real lek models.

Results

We conducted surveys at 48 pseudo-leks and 55 real leks (Table 1) and used 69 maximum counts at pseudo-leks and 68 maximum counts at real leks in our analysis (Table 1). Thus, some leks consisted of >1 maximum count based on sampling across years. Pheasant were used during 62 pseudo-lek counts and chukar were used during 7. Most field data were collected in Nevada and Idaho. Sampling effort varied by year and site and we did not conduct pseudo-lek counts in 2015 at any sites (Table 1). On rare occasions sage-grouse were observed to stop displaying or crouch low to the ground during aerial counts.

All parameter estimates from our models converged ($\hat{R}<1.1$). Based on model parameters, the AIRIS sightability model produced an average sightability (ω_P) of 0.93 (95% CI: 0.87, 0.99), the GBV sightability model produced an average sightability (ω_R) of 0.86 (95% CI: 0.78, 0.95). Thus, the estimated ground count observation error was −14% from the 'true' number of male sage-grouse attending the lek (Fig. 2). Replacing GBV counts with those recorded before the plane arrived also produced an average GBV sightability of 0.86 (95% CI: 0.77, 0.95). Overall, the average counts were similar between paired AIRIS (19.3; SE = 2.1) and GBV (17.9; SE = 2.1). The strong correlation between AIRIS counts and pseudo-lek numbers (i.e. truth; r = 0.94) was similar to the correlation between AIRIS and GBV counts (r = 0.94; Fig. 3a–b). We also found no differences in sightability between pheasant and chukar at pseudo-leks based on counts collected in AIRIS surveys ($\bar{\beta}_{species}$ = 0.02, 95% CI: −0.20, 0.23; Supplementary material Appendix 3 Fig. A1). Correlations between the double-blind GBV counts at sage-grouse leks were high (r = 0.99, indicating agreement in counts obtained between observers. However, residuals between the paired counts increased with lek size, suggesting decreasing precision as a function of lek size (Fig. 3c).

Shrub cover reduced sightability for GBV surveys but did not affect AIRIS surveys (Fig. 4) based on non-overlap of 95% CI for β and SSVS analysis (Table 2). We found marginal evidence that sightability increased as time elapsed from sunrise (linear) for GBV surveys (Fig. 4c) but not for AIRIS (Table 2, Fig. 4d). Although weaker, evidence suggests differences in sightability associated with topographic roughness (Table 2). The type of count (males only versus

combined males and females) showed some evidence of influencing AIRIS sightability, but 95% CIs overlapped 0. Using a model that included count type and covariates fixed at their median values for sagebrush, roughness, and time since sunrise, we estimated average GBV sightability ($\bar{\omega}_R$) to be 0.85 (95% CI: 0.76, 0.95).

The binomial *N*-mixture model was fit to repeated GBV counts recorded at 31 leks. All parameters converged (all $\hat{R}<1.1$). The estimated sightability using repeated counts for any given GBV count was 0.82 (95% CI: 0.78, 0.86), whereas maximum sightability derived from the maximum count (p_D) among these counts was 0.88 (95% CI: 0.83, 0.93).

Discussion

Our empirical calibration of AIRIS allowed a novel and robust assessment of effectiveness of GBV counts for sage-grouse population monitoring. Our findings corroborate
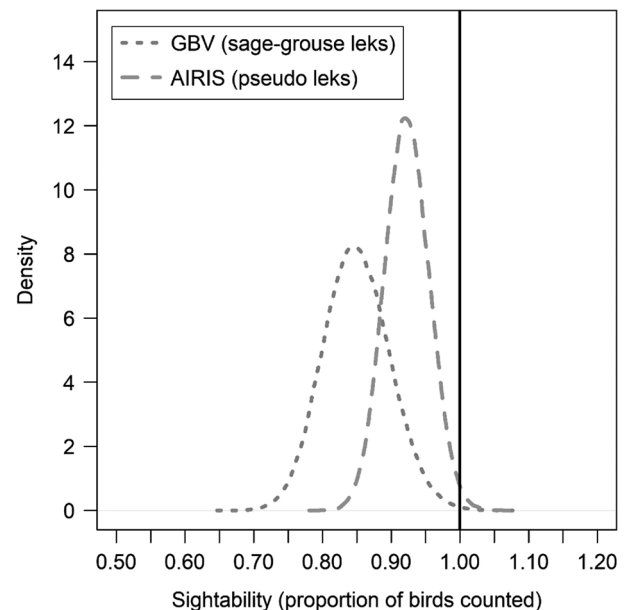

Figure 2. Posterior distributions of the proportion of birds detected at multiple sites in the Great Basin during 2015–2017 using ground-based visual (GBV) surveys and aerial integrated infrared imaging system (AIRIS). Estimates of GBV were derived from real-leks attended by unknown numbers of greater sage-grouse *Centrocercus urophasianus*. Estimates of AIRIS were derived from pseudo-leks with known numbers of ring-necked pheasant *Phasianus colchicus* and chukar *Alectoris chukar*. Perfect detection is denoted by the solid black line.
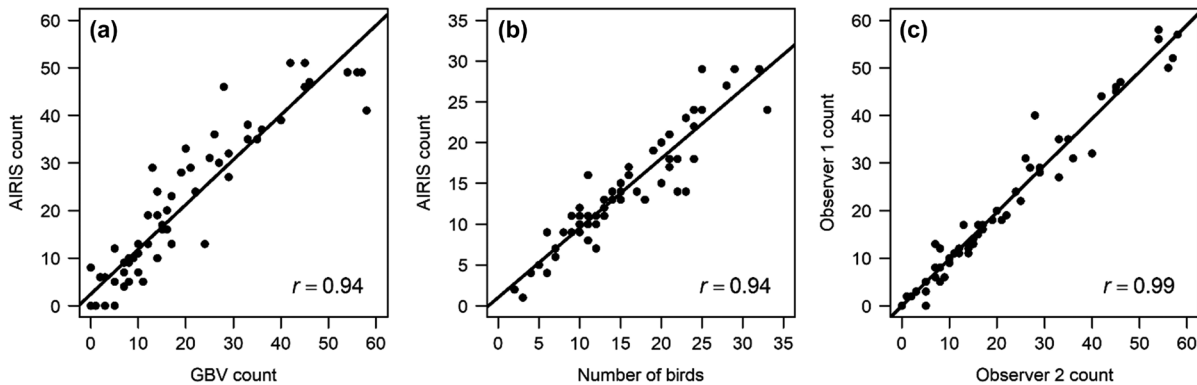
Figure 3. Correlation between counts obtained from (a) aerial infrared imaging system (AIRIS) and ground-based visual (GBV) surveys at real leks with unknown numbers of greater sage-grouse, (b) AIRIS obtained counts of pseudo-leks with known numbers of ring-necked pheasant *Phasianus colchicus* and chukar *Alectoris chukar*, and (c) paired counts of real leks obtained by double-blind observers. Data were collected from multiple sites in the Great Basin from 2015 to 2017.

previous aerial infrared studies of lekking grouse that also reported correlations between infrared cameras and ground counts (Gillette et al. 2013, 2015). However, we extend this study to provide robust estimates based on actual proportion of grouse counted from the aircraft relative to a 'true' number on the ground. Although GBV and AIRIS counts were highly correlated, AIRIS sightability was greater on sage-grouse leks than GBV sightability. Additionally, calibrating AIRIS with known numbers at pseudo-leks and calibrating GBV surveys with AIRIS at sage-grouse leks provided an experimental approach to robustly estimate that GBV surveys observed ~86% of male sage-grouse attending a lek during the survey period. Because our double-blind surveys indicated agreement between observers, especially for smaller leks, failure to detect all males by GBV was not driven intrinsically by individual observer effects.

Aerial infrared technology for wildlife surveys has advanced rapidly and use on sage-grouse lek counts has increased substantially across the western US (J. Romero, Owyhee Air Research, pers. comm.), largely because more leks can be counted per morning especially in remote areas (Gillette et al. 2013, 2015). Because AIRIS and GBV surveys can vary across years at individual leks, population trend estimates may be confounded without appropriately adjusting count data based on methodology.

One option to improve precision and decrease bias of population estimates, is for managers to apply a published estimate of sightability, such as ours, to their maximum lek count data. While this may be a coarse correction for different regions, it can readily be applied to existing lek databases. We therefore provide sightability estimates that may serve as adjustment factors for single maximum lek count data from GBV and AIRIS. These adjustment factors are intended to better approximate true numbers of sage-grouse attending leks and reduce confounding effects of survey type. For example, simply dividing observed counts by the median GBV sightability value reported here, as well as upper and lower 95% credible limits, will provide more accurate estimates of the numbers of males attending a lek during the survey. Additionally, counts obtained with AIRIS can be divided by AIRIS sightability to be comparable with adjusted GBV counts. Adjusted values can then be used to improve

accuracy in estimates of population trends and factors influencing population changes by accounting for detection.

A second option to improve population estimates is for managers to develop their own detection probabilities and or corrected population sizes specific to their leks, regions, and survey times. Our use of an *N*-mixture model provided a relatively simple modeling framework to estimate sightability and lek abundance that can be carried out readily by wildlife managers. Most state agency lek databases currently consist only of maximum counts derived from a series of repeated counts conducted in one morning, while the lower counts are discarded. Applying *N*-mixture models to estimate sightability would only require recording and retaining all the repeated count data within each morning in the lek database rather than just the daily or annual maximum count. Single morning successive repeated counts also allow the closure assumption to be met (Royle 2004). The correspondence between our GBV–AIRIS and *N*-mixture results increased our confidence in the reliability of this method for lek counts, and the reduced cost compared to AIRIS surveys makes it an attractive alternative. Future research that critically evaluates the use of *N*-mixture models on repeated counts during single morning lek surveys would be highly beneficial.

We found that the effects of environmental factors on sightability varied among GBV and AIRIS surveys. The most influential factor that decreased sightability for GBV surveys was increased shrub cover at the lek, which was consistent with findings elsewhere (Fremgen et al. 2016). Fremgen et al. (2016) observed a negative effect of shrub height, and both height and cover likely affect visual obstruction. Thus, GBV observers are seemingly limited by visual screening from shrubs when counting sage-grouse from the ground. AIRIS methods overcome this issue to some extent owing to the plane's ability to circle sage-grouse and observe them from multiple angles, as well as infrared camera's ability to detect partially obstructed birds. We found some evidence that time since sunrise influenced sightability of birds using GBV but not AIRIS, which may be explained by increased ambient lighting. However, another recent sightability study (Baumgardt et al. 2017) observed a negative relationship with time since sunrise, which they
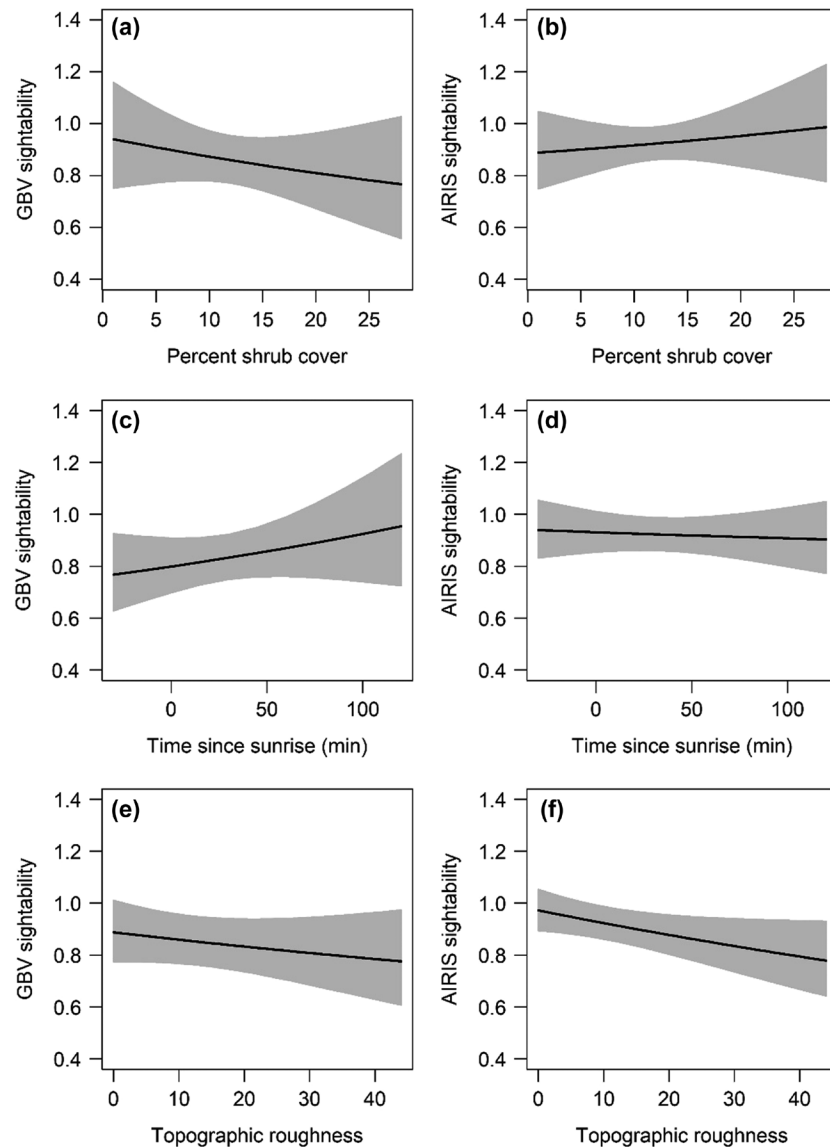
7

Figure 4. Effects of shrub cover (a and b), time since sunrise (c and d), and topographic roughness (e and f) on sightability estimates for ground-based visual (GBV) surveys (left column) at real leks with unknown numbers of greater sage-grouse *Centrocercus urophasianus* and aerial integrated infrared imaging system (AIRIS) surveys (right column) at pseudo-leks with known numbers of ring-necked pheasant *Phasianus colchicus* or chukar *Alectoris chukar*. GBV (2015–2017) and AIRIS (2016–2017) surveys were conducted at multiple study sites within the Great Basin.

Table 2. Parameter estimates from sightability models fit to pseudo-lek data (AIRIS sightability model) collected from the air (ring-necked pheasant *Phasianus colchicus* and chukar *Alectoris chukar*) and real lek data (GBV sightability model) collected in the air and on the ground (greater sage-grouse *Centrocercus urophasianus*) at study areas in the Great Basin. Coefficients (β) are reported for different covariates with their associated median and 95% credible intervals (CI). Subscripts indicate covariates which included shrub cover ('shrub'), time since sunrise ('tsr'), terrain roughness ('rough'), and intercepts.

Lek type	Parameter	Median β	95% CI	P ($\beta=1$)*
Pseudo-lek	$\beta_{intercept}$	−0.060	−0.253 to 0.135	na
	β_{shrub}	0.004	−0.010 to 0.017	0.435
	β_{tsr}	0.000	−0.002 to 0.001	0.488
	β_{rough}	−0.005	−0.010 to 0.000	0.505
Real lek	$\beta_{intercept}$	−0.145	−0.466 to 0.171	na
	β_{shrub}	−0.008	−0.025 to −0.010	0.640
	β_{tsr}	0.001	−0.001 to 0.003	0.511
	β_{rough}	−0.003	−0.010 to 0.003	0.502

* Indicator function representing whether β is included in the model using stochastic search variable selection (George and McCulloch 1996). Evidence was considered substantial for values >0.6, marginal for 0.5–0.6, and deficient for <0.5.

8

attributed to decreased strutting activity. That study also observed that cloud cover and presence of females influenced sightability (Baumgardt et al. 2017). While we did not record cloud cover, given the absence of time since sunrise effects in AIRIS surveys, we suspect cloud cover is unlikely to influence the ability of the AIRIS to detect male sage-grouse but may be expected to influence sightability in ground counts. Weak evidence suggested that terrain roughness reduced sightability for AIRIS and GBV. This might be explained by fragmentation of field of view for both survey methods. Importantly, our assessment of using variables from relatively high-resolution GIS layers allows managers to use readily-available spatial data to adjust their estimates based on measurements associated with leks (e.g. shrub cover) remotely following lek monitoring, as opposed to conducting field measurements (Fremgen et al. 2016).

The overall lack of covariate effects on AIRIS sightability provide support for the hypothesis that most environmental predictors should not be as concerning in AIRIS surveys as in GBV surveys (Fremgen et al. 2016, Baumgardt et al. 2017). One explanation is that sightability associated with AIRIS is almost entirely comprised of probability of detection (p_d) and not probability of being available (p_a). This is because the p_a is likely very close to 1.0 given that factors that influence p_a for detection are not influential. In contrast, sightabilty in GBV surveys is likely driven by factors that influence p_a, such as visual obstruction by shrubs.

Several features of our study may have influenced our results and are important to consider for application of our methods to other systems. First, our proxy birds varied from sage-grouse in characteristics such as size and plumage. Mean sage-grouse mass across sexes at breeding (2323 g; Beck and Braun 1978) are substantially larger then female pheasant (954 g; Giudice and Ratti 2001) or chukar (680 g; Nagel 1945), suggesting the larger sage-grouse should be at least as detectable by infrared cameras as our proxy birds. Additionally, we did not observe a difference between pheasant and chukar which may imply that size did not have an effect. Furthermore, while the plumage coloration was different among all three species, infrared imagery does not use the visible spectrum so is unaffected by color, and the high-resolution color camera was only used to distinguish sex in the real lek surveys. Second, while sex was not distinguished in some of the surveys, and we found slight evidence of sightability differences among sexes, we accounted for the different survey types in the model and only report sightability estimates for males because male sage-grouse are the primary interest for wildlife agencies and land managers (WAFWA 2015). In using AIRIS for lek surveys, we stress the importance of separating males and females in the counts and having trained technicians capable of accurately identifying the sex of sage-grouse. Third, while our study assessed sightability (i.e. $p_a p_d$), it did not account for other components important for true population abundance estimates previously described by Nichols et al. (2009). One such component was the probability of sage-grouse being present on lek during the time of sampling (p_p) which can vary substantially throughout the season and among years (Wann et al. 2019). Additionally, not all lek locations are known (Sedinger 2007), which can lead to underestimation

of population abundance, and sage-grouse leks that are easily accessible (e.g. near roads) are more likely to be surveyed which can introduce sampling bias (Applegate 2000, Anderson 2001, Walsh et al. 2004). Our study was not designed to address these issues, but they should be considered when assessing populations. We note, however, the potential application of AIRIS in locating unknown leks and surveying inaccessible leks, and we urge assessments of such uses.

Although AIRIS is a promising tool for lek surveys, this method has advantages and disadvantages compared to GBV surveys. First, AIRIS can be costly, currently averaging approximately $800 per hour (Gillette et al. 2015), whereas GBV methods to survey the same number of leks has approximately one third the cost. However, under time constraints, substantially more leks can be counted in a single morning using AIRIS than conventional methods of GBV surveys (Gillette et al. 2015). Second, flight time often must be scheduled well in advance because of limited availability of suitable aircraft and pilots, potentially constraining the use of AIRIS, whereas GBV surveys can be implemented more readily. Lastly, small aircraft surveys can carry increased safety risk for personnel (Sasse 2003), and weather conditions can additionally limit survey windows (Gillette et al. 2015). An economic assessment contrasting these survey techniques was beyond the scope of our study but see Gillette et al. (2015) for thorough cost-comparison between AIRIS and GBV.

In conclusion, our study provides information that can be used to improve inference to population sizes and trends and can help advance lek survey methods. While decisions of using AIRIS techniques over those of conventional GBV might be based on multiple factors including differences in costs, lek access, etc., likely a combination of both techniques will allow for most effective surveying for population assessments. We found GBV methods captured approximately 86% of males attending leks, while AIRIS increased the proportion of sage-grouse detected to about 93%, on average. Thus, for management application, our estimates may serve as general baseline adjustments on single lek counts for AIRIS and GBV (i.e. maximum counts) methods aimed at standardizing databases and accounting for detection uncertainty. Moreover, we provide adjustments in sightability for specific sites as a function of environmental covariates derived from widely available GIS layers. Lastly, *N*-mixture models using repeated within-morning ground counts from a single observer are useful to estimate sightability and lek abundance, which should be useful in accounting for spatial and temporal trends in observation error. These methods provide multiple options for managers to improve previously collected data and refine their monitoring programs to make better use of lek data for population studies.

References

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. – Wildl. Soc. Bull. 29: 1294–1297.

Applegate, R. D. 2000. Use and misuse of prairie chicken lek surveys. – Wildl. Soc. Bull. 28: 457–459.

Baumgardt, J. A. et al. 2017. Visibility bias for sage-grouse lek counts. – Wildl. Soc. Bull. 41: 461–470.

Beck, T. D. I. and Braun, C. E. 1978. Weights of Colorado sage grouse. – Condor 80: 241–243.

Beck, T. D. I. and Braun, C. E. 1980. The strutting ground count: variation, traditionalism, management needs. – Proc. W. Assoc. Fish Wildl. Agen. 60: 558–556.

Bivand, R. S. et al. 2013. Applied spatial data analysis with R, 2nd edn. – Springer.

Blomberg, E. J. et al. 2013. Annual male lek attendance influences count-based population indices of greater sage-grouse. – J. Wildl. Manage. 77: 1583–1592.

Brooks, S. P. and Gelman, A. 1998. General methods for monitoring convergence of iterative simulations. – J. Comput. Graph. Stat. 7: 434–455.

Coates, P. S. et al. 2016. Wildfire, climate and invasive grass interactions negatively impact an indicator species by reshaping sagebrush ecosystems. – Proc. Natl Acad. Sci. USA 113: 12745–12750.

Connelly, J. W. and Schroeder, M. A. 2007. Historical and current approaches to monitoring greater sage-grouse. – In: Reese, K. P. and Bowyer, R. T. (eds), Monitoring populations of sage-grouse. College of Natural Resources Experiment Station Bulletin 88. Univ. of Idaho, Moscow, USA, pp. 3–9.

Connelly, J. W. et al. 2003. Monitoring of greater sage-grouse habitats and populations. Station Bulletin 80. – College of Natural Resources Experiment Station, Univ. of Idaho, Moscow, Idaho, USA.

Fremgen, A. L. et al. 2016. Male greater sage-grouse detectability on leks. – J. Wildl. Manage. 80: 266–274.

Fremgen, A. L. et al. 2019. Weather conditions and date influence male sage grouse attendance rates at leks. – Ibis 161: 35–49.

Garton, E. O. et al. 2015. Greater sage-grouse population dynamics and probability of persistence. Final Report to Pew Charitable Trusts. – <http://www.pewtrusts.org/~/media/assets/2015/04/garton-et-al-2015-greater-sagegrouse-population-dynamics-and-persistence-31815.pdf>.

George, E. I. and McCulloch, R. E. 1996. Stochastic search variable selection. – In: Markov chain Monte Carlo in practice. Springer US.

Gillette, G. L. et al. 2013. Can reliable sage-grouse lek counts be obtained using aerial infrared technology? – J. Fish Wildl. Manage. 4: 386–394.

Gillette, G. L. et al. 2015. Evaluating the potential of aerial infrared as a lek count method for prairie grouse. – J. Fish Wildl. Manage. 6: 486–497.

Giudice, J. H. and Ratti, J. T. 2001. Ring-necked pheasant (*Phasianus colchicus*), ver. 2.0. – In: Rodewald, P. G. (ed.), The birds of North America. Cornell Lab of Ornithology, Ithica, NY, USA.

Green, A. W. et al. 2017. Investigating impacts of oil and gas development on greater sage-grouse. – J. Wildl. Manage. 81: 46–57.

Havens, K. J. and Sharp, E. J. 1998. Using thermal imagery in the aerial survey of animals. – Wildl. Soc. Bull. 26: 17–23.

Kéry, M. 2010. Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses. – Academic Press.

McCaffery, R. and Lukacs, P. M. 2016. A generalized integrated population model to estimate greater sage-grouse population dynamics. – Ecosphere 7: e01585.

Monroe, A. P. et al. 2016. Effect of lek count protocols on greater sage-grouse population trend estimates. – J. Wildl. Manage. 80: 667–678.

Monroe, A. P. et al. 2019. The importance of simulation assumptions when evaluating detectability in population models. – Ecosphere 10: e02791.

Nagel, W. O. 1945. Adaptability of the chukar partridge to Missouri conditions. – J. Wildl. Manage. 9: 207–216.

Nichols, J. D. 1991. Extensive monitoring programmes viewed as long-term population studies: the case of North American waterfowl. – Ibis 133: 89–98.

Nichols, J. D. et al. 2009. Inferences about landbird abundance from count data: recent advances and future directions. – In: Thomson, D. L. et al. (eds), Modeling demographic processes in marked populations. Springer, pp. 201–235.

Patterson, R. L. 1952. The sage grouse in Wyoming. – Wyoming Game and Fish Commission and Sage Books, Inc., Denver, CO, USA.

Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. – In: Hornik, K. et al. (eds), Proc. 3rd Int. workshop on distributed statistical computing (DSC 2003), March 2003, Vienna, Austria, pp. 1–10. <www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf>

Riley, S. J. et al. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. – Intermount. J. Sci. 5: 23–27.

Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. – Biometrics 60: 108–115.

Sasse, D. B. 2003. Job-related mortality of wildlife workers in the United States, 1937–2000. – Wildl. Soc. Bull. 31: 1015–1020.

Schroeder, M. A. et al. 2004. Distribution of sage-grouse in North America. – Condor 106: 363–376.

Sedinger, J. S. 2007. Improving understanding and assessment of greater sage-grouse populations. – Univ. Idaho Coll. Nat. Resour. Exp. Stat. Bull. 88: 43–56.

Shyvers, J. E. et al. 2018. Dual-frame lek surveys for estimating greater sage-grouse populations. – J. Wildl. Manage. 82: 1689–1700.

Smith, J. T. et al. 2016. Reducing cultivation risk for at-risk species: predicting outcomes of conservation easements for sage-grouse. – Biol. Conserv. 201: 10–19.

Su, Y. S. and Yajima, M. 2015. R2jags: using R to run 'JAGS'. – R package ver. 0.5-7. <https://CRAN.R-project.org/package=R2jags>.

Venables, W. N. and Ripley, B. P. 2002. Modern applied statistics with S, 4th edn. – Springer.

WAFWA. 2015. Greater sage-grouse population trends: an analysis of lek count databases 1965–2015. – Western Association of

Fish and Wildlife Agencies, Cheyenne, WY. <www.wafwa.org/Documents%20and%20Settings/37/Site%20Documents/News/Lek%20Trend%20Analysis%20final%208-14-15.pdf>.

Walsh, D. P. et al. 2004. Evaluation of the lek-count index for greater sage-grouse. – Wildl. Soc. Bull. 32: 56–68.

Wann, G. T. et al. 2019. Assessing lek attendance of male greater sage-grouse using fine-resolution GPS data: implications for population monitoring of lek mating grouse. – Popul. Ecol. 61: 183–197.

Williams, B. K. et al. 2002. Analysis and management of animal populations, 1st edn. – Academic Press.

Xian, G. et al. 2015. Characterization of shrubland ecosystem components as continuous fields in the northwest United States. – Remote Sens. Environ. 168: 286–300.

Supplementary material (available online as Appendix wlb-00552 at <www.wildlifebiology.org/appendix/wlb-00552>). Appendix 1–3.