SAMPLING METHODS FOR LEK AND BROOD COUNTS OF GREATER SAGE-GROUSE: ACCOUNTING FOR IMPERFECT DETECTION

A Thesis
Presented in Partial Fulfillment of the Requirements for the
Degree of Master of Science
with a
Major in Natural Resources
in the
College of Graduate Studies
University of Idaho
by
Ian P. Riley

Major Professor: Courtney Conway, Ph.D.
Committee Members: Ryan Long, Ph.D.; Shane Roberts, Ph.D.; Kerri Vierling, Ph.D.
Department Administrator: Lisette Waits, Ph.D.

May 2019

Authorization to Submit Thesis

This thesis of Ian P. Riley, submitted for the degree of Master of Science with a Major in Natural Resources and titled "SAMPLING METHODS FOR LEK AND BROOD COUNTS OF GREATER SAGE GROUSE: ACCOUNTING FOR IMPERFECT DETECTION," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: ______________________________ Date: __________
Courtney Conway, Ph.D.

Committee Members: ______________________________ Date: __________
Ryan Long, Ph.D.

______________________________ Date: __________
Shane Roberts, Ph.D.

______________________________ Date: __________
Kerri Vierling, Ph.D.

Department
Administrator: ______________________________ Date: __________
Lisette Waits, Ph.D.

Abstract

A central component of wildlife conservation and harvest management is accurately estimating population abundance and parameters that influence breeding productivity. Inferences about wildlife population abundance are only accurate if study designs adequately reflect temporal and spatial changes in vital rates. Greater sage-grouse (*Centrocercus urophasianus*; hereafter, sage-grouse) are a species of conservation concern due to continued population decline since the 1940's and ongoing threats to their habitat. Currently, investigators use counts of male greater sage-grouse at leks and brood counts of radio-marked hens to indicate annual change in sage-grouse abundance and breeding productivity. Lek and brood counts are essential to inform harvest management, Endangered Species Act listing decisions, and land management policies. However, inferences based on these uncorrected counts may be biased if investigators fail to account for spatial and temporal variation in detection probability.

In chapter 1, we used generalized linear mixed models to estimate the probability of visually and aurally detecting active leks during mock point-count surveys along lek routes. We also examined whether 15 factors influenced aural lek detection and we used a Huggins model to compare aural lek detection probabilities between audio-recording equipment and surveyors. Our results demonstrated that surveyors can aurally detect active leks at greater distances (300% further) compared to visual-only surveys. Our results suggest aural detection probabilities of an active lek were highest near 8 April, within 30 mins of sunrise, on relatively calm and cold days, when surveyors were at higher elevations relative to the lek, in areas of low topographic roughness, during times with no background noise, and on mornings following moonless nights. Audio-equipment had higher lek detection rates compared to real-time surveyors. Our results suggest that 1-min point-count surveys on lek routes would greatly increase detection rates (and reduce bias). To implement 1-min point-count surveys along lek routes, some of the existing lek routes may need to be shortened or split to ensure that the lek routes can be completed during the optimal times when grouse attend leks. By incorporating 1-min point-count surveys at regular intervals along lek routes, surveyors will be more likely to locate all leks along a survey route (including leks that move from one year to another). Therefore, analysts will be able to use the data to produce more accurate estimates of population trends for male sage-grouse.

In chapter 2, we used Cormack-Jolly-Seber (CJS) and Huggins models to estimate survival and detection probabilities of radio-marked hens with broods. Furthermore, we reviewed the sage-grouse literature to summarize brood survey methods used in past studies, including variance, bias, and detection probability associated with each method. We used CJS models to compare brood survival and detection probabilities with data generated from daytime surveys (flush and visual) and nighttime spotlight surveys against a novel brood fecal survey method whereby brood status is assessed indirectly via the presence of brood feces at nighttime roost-sites. A Huggins model was used to estimate differences in detection probabilities among daytime flush surveys, fecal surveys, and nighttime spotlight surveys for broods at 42 days after hatch. We demonstrated that detection probability of daytime surveys (visual and flush surveys) changes with brood age, whereas brood fecal surveys had comparably high detection rates that did not vary with brood age. Our novel brood survey method was as effective as daytime surveys but had a lower probability of flushing hens and their broods (a drawback of daytime brood surveys). Based on our literature review, we need more standardization and rigor in sage-grouse brood survey methods and design. For example, we found only 2 studies that accounted for imperfect detection when using daytime brood surveys to estimate brood or chick survival. We also found wide disparity among studies regarding the metrics they used to determine brood fate, statistical analyses used, and the frequency and duration of monitoring periods used to estimate habitat selection, hen productivity metrics, and estimate brood or chick survival.

Greater sage-grouse are of substantial management concern. Unbiased estimates of sage-grouse population trends, abundance, and breeding productivity metrics are important to prevent inaccurate conclusions and inappropriate management actions. Our results provide managers and future researchers with some potential ways to improve future monitoring efforts and management decisions.

Acknowledgements

If this thesis is considered a success, I first owe it to my advisor Courtney Conway, my committee, and lab mates. Courtney Conway's seemingly eternal patience in focusing my budding ideas into fruition is only topped by his generosity and encouragement. I wish to express my gratitude to my committee: Ryan Long, Shane Roberts, and Kerri Vierling, for their continued professional guidance, constructive feedback, and time spent on my written work that vastly improved my thesis' final form. I further thank my lab mates: Austin Allison, Emilia Breitenbach, Amanda Goldberg, Dave Gotsch, Eamon Harrity, Carl Lundblad, Deo Lachman, and Bryan Stevens for helping and supporting me with their respective expertise and answering my itinerant thoughts and questions. I especially thank Bryan Stevens for his statistical and programming knowledge that was invaluable with simulations and bootstrapping efforts in my 1st chapter.

Secondly, I would like to thank the Grouse & Grazing project leaders, crew leaders, technicians, members of the Idaho Department of Fish and Game (IDFG), and ranchers that provided technical support for my project. I am personally indebted to David Musil, Anthony Locatelli, Andrew Meyers, Emilia Breitenbach, Dylan O'Leary, Emily Prosser, Jackie Dougherty, and Matt Modlin for their untiring, difficult work managing the oft-overlooked logistics of this project, trapping sage-grouse, and following broods. Also, I would like to thank all the IDFG lek survey personnel that generously worked with me to coordinate sampling at leks to reduce stepping on each other's toes. We worked with numerous private ranchers in the past 2 years throughout southern Idaho that were easy to work with, told a good story or two, and occasionally pulled a wayward graduate student from the mud.

My thesis would not have been possible without funding from the Veterans Affairs program, Idaho Department of Fish and Game, Bureau of Land Management in Idaho, and the College of Natural Resources at the University of Idaho.

Dedication

To my family and friends. Without them, I am but a snake that cannot cast its skin. My Uncle Paul that first showed me the elegant, yet capricious ways of nature. And to the loved ones that showed me the bright, dancing stars in the chaotic abyss.

Table of Contents

List of Tables

Chapter 1.

Chapter 2.

List of Figures

Chapter 1.

Chapter 2.

Chapter 1: Factors that Affect Detection Probability of Greater Sage-Grouse Leks and Implications for Lek Count Methods

Abstract

Counts of greater sage-grouse (*Centrocercus urophasianus*) at leks have been used to inform harvest management, Endangered Species Act listing decisions, and land management policies. Current lek count sampling methods have several sources of sampling error. We focus on imperfect detection of leks due to males that are temporarily unavailable for counting and leks that moved, therefore, the relationship between lek counts and true abundance of male sage-grouse is currently unknowable. We assessed the potential benefits of a revised sampling design whereby observers conduct multiple point-count surveys annually along defined survey routes which provides a standardized approach to locate unknown leks. To evaluate the validity of this approach, we used binomial generalized linear fixed and hierarchical models to estimate the likelihood that observers would detect active leks during 1-min point-count surveys. Furthermore, we examined whether 15 factors influenced aural detection probability of leks. Finally, we used a Huggins model to compare lek detection probabilities between audio-recording equipment and field-observers. Our results suggest that surveyors can aurally detect active leks at distances nearly 300% greater than the distances at which they can visually detect leks. Numerous factors influenced aural detection probability of leks. The probability of hearing an active lek was highest: near 8 April, within an hour of sunrise, on relatively calm and cold days, when the surveyor was at higher elevation relative to the lek, when moon illumination was low, and when there was no background noise. Detection probability declined with distance and the probability of aural detection was 59% at 1km from a lek. Hence, ≥3 surveys along a lek route (as is typical for current lek count protocols in many states) would yield ≥ 93% detection of leks within 1 km. Detection probability of active leks was higher with audio-equipment compared to a field observer. Our results suggest that 1-min point-count surveys along lek routes would greatly improve detection probability of sage-grouse leks and provide more precise estimates of population trends. Moreover, our results provide insights about how environmental factors influence detection probability (and communication behavior of grouse) during surveys.

Introduction

A common objective in wildlife management is to estimate population trends. Population trends are useful metrics of population abundance if there is a known relationship between these quantities. Several sources of error may make this relationship unknowable including spatial and temporal variation in detection probability. To make viable inferences, investigators must ensure that the study design accounts for spatial and temporal variation in detection of individuals so that sampling detects temporal changes in abundance for the population of interest (Thompson 2002; Buckland et al. 2015). However, many survey efforts commonly used by wildlife biologists do not allow for estimates of detection probability (i.e., they assume perfect detection). Unbiased estimates of population trend are crucial, especially for harvested species of great management concern such as the greater sage-grouse (*Centrocercus urophasianus*; hereafter sage-grouse). Analyses of lek count data suggest that sage-grouse occupy <60% of their historic range (Schroeder et al. 2004) and that abundance has decreased by ~0.83% per year between 1965-2015 (Garton 2015). These analyses of lek count data have influenced policy decisions. Sage-grouse were listed as endangered under the Canadian Species at Risk Act in 1998 (COSEWIC 2017) and have been petitioned for listing 8 times under the U.S. Endangered Species Act (ESA). The U.S. Fish and Wildlife Service (USFWS) determined that sage-grouse were warranted for ESA listing in 2010 but revised their decision and determined that a sage-grouse listing was no longer warranted (Department of the Interior 2015). Data from sage-grouse lek counts are also used to guide harvest regulations in 7 states (e.g., Idaho Sage-grouse Advisory Committee 2006), measure the effects of energy development on sage-grouse populations (LeBeau et al. 2017), and classify and rank habitat for conservation purposes (Doherty et al. 2010; Coates et al. 2013). Hence, unbiased estimates of sage-grouse population trends are important to prevent inaccurate conclusions and inappropriate management actions.

Lek counts are frequently used to estimate sage-grouse population trends, but methodological biases in those counts have prompted numerous authors to question the reliability of inferences based on lek counts (Anderson 2001, Walsh et al. 2004, Sedinger 2007, Johnson and Rowland 2007). Observers typically count grouse at leks along survey routes that were established explicitly to conduct male counts at a series of known leks within a standardized timeframe (within 1 hr of sunrise; Connelly et al. 2003). With most current lek route survey methods, surveyors typically expend little or no effort searching for new leks (or leks that have moved) while conducting lek routes. Not counting grouse at new or moved leks creates 3 main issues that make the accuracy of population trend estimates unknowable (Johnson and Rowland 2007). First, non-detection at a known lek along a lek route does not necessarily mean true absence from the surveyed area but instead may result from males being temporarily absent (e.g., not attend leks to avoid predators) or short-distance male movements or emigration (Sedinger 2007). Second, leks that are not easily visible from the road are not likely to be detected unless they are close to a historical lek. Third, most lek counts are not based on a defined area surrounding the route so the counts may not represent overall population trends (Walsh et al. 2004). In other words, current lek route survey methods cannot distinguish between temporal changes in detection and temporal changes in abundance. However, annual lek counts are still widely used to estimate sage-grouse population trends and make decisions regarding management, policy, and harvest regulations (Garton 2015).

One approach that could potentially improve accuracy of trend estimates from lek counts is if observers stopped periodically at points along lek route surveys and conducted a short 1-min point-count survey. Periodic point-count surveys along lek routes could potentially allow observers to detect new or moved sage-grouse leks (and account for spatial and temporal variation in detection probability of leks) thereby improving the data generated from sage-grouse lek counts. However, the probability of hearing or seeing a sage-grouse lek on standardized surveys has not been estimated nor have the factors that influence the detection probability of leks. Periodic point-count surveys along lek routes would help overcome some of the current biases if they substantially increased detection probability and allow analysts to account for spatial and temporal variation in detection probability. How far away can a surveyor hear an active sage-grouse lek and how far away can a surveyor see an active leks? The distance beyond which detection probability is zero (i.e., distance thresholds; Dejong and Emlen 1985) likely varies with environmental features (Brewster and Simon 2009; Anderson et al. 2015) and likely differs between aural and visual detection. For example, visual distance thresholds of a sage-grouse lek vary with vegetation density, obstructions, and sage-grouse behavior (Fremgen et al. 2016; Baumgardt et al. 2017). Meanwhile, aural distance thresholds and the factors that influence them have not been studied with sage-grouse. Evidence from lesser (*Tympanuchus pallidicinctus*, Butler et al. 2010) and greater prairie-chickens (*T. cupido pinnatus*, Raynor et al. 2017), northern bobwhites (*Colinus virginianus*, Rusk et al. 2009), and ruffed grouse (*Bonasa umbellus,* Zimmerman et al. 2003) suggest many factors (e.g., wind speed, terrain, etc.) may influence aural detection of sage-grouse leks.

Documenting the factors that influence detection probability will help inform debates about how best to improve survey methods and potentially increase the accuracy of sage-grouse population trend estimates. To address this need, we quantified how visual and aural detection probability varies with distance from surveyor to lek for sage-grouse leks in Idaho. We also examined the relationship between aural detection probability and 15 environmental variables.

Methods

Study Area

We conducted lek detection trials from early March until early May in 2016 and 2017 in the Big Southern Butte area of the upper Snake River plain in Bingham, Butte, and Power counties, southeastern Idaho (43.24°N, 113.07°W). We coordinated our lek detection trials to minimize conflicts with two other efforts in the area: a long-term sage-grouse study (Conway et al. 2017) and annual lek counts coordinated by Idaho Department of Fish and Game (IDFG). Topography consists of relatively flat, rolling hills interspersed by shallow valleys, playas, lava fields, and buttes (elev. 1536-2304m). Monthly precipitation at the nearby Idaho National Laboratory (INL) averages 1.5, 2.0, and 3.0 cm, respectively, for March, April, and May (National Oceanic and Atmospheric Administration's INL Weather Center. Undated). The landscape is dominated by patches of sagebrush (*Artemisia* ssp.) interspersed with grass or open ground. Common shrubs include Wyoming big sagebrush (*A. tridentata wyomingensis*), three-tip sagebrush (*A. t. tripartita*), basin big sagebrush (*A. t. tridentata*), and rabbitbrush (*Chrysothamnus visciidflorus*). Common grasses include squirreltail (*Elymus elymoides*), bluebunch wheatgrass (*Pseudoroegneria spicata*), Sandberg bluegrass (*Poa secunda*), needle-and-thread grass (*Hesperostipa comata*), Indian rice grass (*Achnatherum hymenoides*), and crested wheatgrass (*Agropyron cristatum*). Potential predators of sage-grouse seen during the lekking season include: coyote (*Canis latrans*), short-eared owl (*Asio flammeus*), great horned owl (*Bubo* *virginianus*), golden eagles (*Aquila chrysaetos*), bald eagles (*Haliaeetus leucocephalus*), rough-legged hawks (*Buteo lagopus*) and ferruginous hawks (*Buteo regalis*). While conducting our sampling, we observed coyotes, golden-eagles, and a northern harrier (*Circus hudsonius*) attack sage-grouse on leks. Bureau of Land Management (BLM) manages most of the land within our study area for cattle and sheep grazing (Idaho Sage-grouse Advisory Committee 2006). Spring grazing is common in the study area and most sage-grouse leks in the study area are in locations that appear to have been modified or created for water catchments and salt licks for cattle. Recreational vehicle activity is popular in the area.

Detection trials

We used detection trials at active leks to examine the relationship between distance and aural and visual detection probability and factors that influence detection probability. We considered a lek as a location where ≥ 1 male congregates for courtship on ≥ 2 days for ≥1 year. All leks that we sampled were on lek routes established by IDFG or the author. We conducted all detection trials near sunrise to coincide with the times when male sage-grouse attend leks (~1.5 hr. before sunrise to ~2 hr. after sunrise; Jenni and Hartzler 1978). A detection trial included a sequential series of 1-min aural surveys at various distances from a focal lek known to be active. We began most detection trials with a 1-min survey at 1.5km from the focal lek. We then moved away from or towards the focal lek depending on whether we aurally detected grouse on the lek at the initial 1-min survey point. We initially moved towards or away from the focal lek in 500m increments between sequential 1-min surveys until we recorded a change in outcome (e.g., detected then not detected, or vice versa), and then we moved in the opposite direction in progressively smaller increments (i.e., 250m, 100m, 50m) between sequential 1-min surveys to narrow down the detection distance threshold (between where we could and could not detect the lek; Dejong and Emlen 1985). Therefore, the number and location of 1-min survey points varied among detection trials, and we did not conduct surveys along each trial progressively further from (or closer to) the focal lek. Although we typically conducted the first survey at 1.5km from the focal lek, the location and order of all subsequent surveys were selected based on outcome from previous 1-min surveys on that detection trial, road conditions, and time constraints, with the explicit objective of efficiently locating the detection distance threshold. We occasionally conducted a second detection trial (73 of the 155 trials) on the same focal lek on the same morning to increase our sampling size. At leks where we conducted a second detection trial on the same morning, the first 1-min survey started 75-975m away from the opposite side of the lek boundary from where the first detection trial ended (e.g., we conducted the first trial north of the lek and the second trial south of the lek). We conducted a second trial only in instances where the observer had no noticeable effect on displaying male grouse behavior on the focal lek during the first trial. The final 1-min survey on each detection trial was at the lek count observation point (the designated point where observers typically conduct annual lek counts). All 1-min surveys were performed while standing 10m away from the vehicle (vehicle was turned off) perpendicular to the road (except at the lek count observation point which were conducted from within the vehicle, as per the current lek count survey protocols in Idaho). In 2017, we also recorded the distance beyond which we could no longer see lekking birds during all 1-min surveys. One person (I. Riley) with normal hearing ability (based on a hearing test administered prior to the first field season) and with experience listening to sage-grouse calls and counting sage-grouse on leks performed all 1-min surveys on the 155 detection trials.

For each 1-min survey, we calculated the distance from the surveyor to the geometric center of the lek. We returned to each lek ≥2 hr after sunrise when no sage-grouse were present to determine the geometric center of the lek. To do so, we mapped the maximum spatial extent of males present on the lek (i.e., the lek boundary). Aided by the map, we recorded the lek boundary via the "tracks" features on a global positioning system (GPS) receiver (Garmin rino 120 and etrex 30x). The geometric center of the lek was estimated from the recorded boundary. In cases where the center of the lek moved substantially between repeated visits, we recalculated distances from the surveyor to the new lek center coordinates with ArcGIS 10.3 distance tools (Environmental Systems Research Institute, Redlands, CA).

We recorded a suite of weather, temporal, noise-related, and topographic factors that we thought *a priori* could potentially influence probability of lek detection. We used a wind meter (Kestrel 1000) to measure mean wind speed (km/hr) and temperature (Celsius). We recorded sky conditions as: (1) ≤50 % cloud cover; (2) > 51% cloud cover; or (3) precipitation. We recorded the number of the following terrain features between the surveyor and the focal lek: (1) draw; (2) knoll; (3) knoll side; (4) flat ground (defined as ~0 slope for ≥ 200 m). We also recorded the prominent terrain feature of the observation point (i.e., where the surveyor was standing during the 1-min survey). We visually estimated the dominant vegetation height in the 100m surrounding the surveyor as: (1) 0-0.5m; (2) 0.5-0.99m; (3) 1-1.49m; or (4) 1.5-2m. We recorded elevation (m) where each 1-min survey was conducted. In 2017, we recorded whether any of the following types of background noise were present during surveys: (0) no noise, (1), airplane, (2) vehicle(s), (3) mechanical wells, (4) cows, (5) coyotes, and (6) sage-grouse from non-focal leks.

We conducted a formal lek count at the end of each detection trial from the lek count observation point (the designated point where observers typically conduct the lek count at that lek during their annual lek route survey). We conducted the lek count while inside the vehicle ≥75 m from the lek boundary, as per state like count protocols (D. Musil, M. Commons-Kemner, pers. comm). Surveyors on lek routes are instructed to remain inside the vehicle for the formal lek count to minimize the probability that the grouse would flush from the lek during the count. We counted sage-grouse with binoculars if the observation point was ≤100 m from the lek boundary and we used a 20*x*60 mm spotting scope if the observation point was ≧100 m from the lek boundary. At each lek count observation point, we recorded the number of males, displaying males, and females present. We also recorded the 15 covariates at the lek observation point (i.e., those recorded at the other 1-min surveys during the detection trials).

Audio-recordings

At 50 of the 52 detection trials in 2017, we used portable audio-equipment to help us estimate lek detection probabilities and identify covariates that may affect aural detection probability. We used a recorder (ZOOM H4nPro) and omnidirectional microphone (Sennheiser, ME 62/K6) covered by a windscreen (Rycote, 5 cm Classic Softie-19/22) that was stabilized on a telescopic pole (K-Tek, Avalon KE-89CC) and shock-mount (Auray, DUSM-1). We extended the assembly ~2.44m above the ground to minimize noise caused by turbulence near the observer. Sound was recorded in 96-kHz, WAV format at 24-bit resolution. We transferred recordings to a computer and renamed audio files with a unique code.

A technician with field experience identifying sage-grouse calls and sagebrush-obligate songbird vocalizations listened to all the 1-min recordings in a quiet office with headphones. Each audio-file was randomly assigned to prevent predictive associations with recording order. We used spectrograms in the program Audacity (window type: Blackman-Harris; window length: 2048) to see or hear recorded sounds. The technician recorded whether she heard sage-grouse calls (presence/absence) and recorded the primary source and the intensity of any background noise (i.e., wind, songbirds, vehicles, etc.). We categorized noise intensity as: (0) no background noise during most of the 1-min survey, (1) faint background noise during >50% of the 1-min survey, (2) sufficient background noise to likely prevent hearing sage-grouse during <25% of the 1-min survey, (3) sufficient background noise to likely prevent hearing sage-grouse during 25-50% of the 1-min survey, and (4) sufficient background noise to likely prevent hearing sage-grouse during >50% of the 1-min survey (modified from Conway 2009). We standardized headphone volume to maximize noise identification without causing hearing damage. The audio-technician could replay files or refer to reference recordings and spectrograms as needed to verify noise sources. After analysis, we matched analyzed files to their corresponding observation point to create detection histories.

Model construction

Detection distances. - One of our objectives was to compare the relationship between distance and detection probability between aural and visual modes of detection. Given the limited time window available to detect displaying males at leks (~ 2½ hrs each morning), logistical challenges, and the potential risk of disturbing grouse at leks, we did not conduct further observations after recording the aural and visual detection distance thresholds on a detection trial. Detection of sage-grouse is a threshold phenomenon (Dejong and Emlen 1985) and so we assumed that we would see or hear sage-grouse at any distance between the lek and the detection distance threshold, and that we would not hear or see sage-grouse beyond that threshold. This assumption was undoubtedly not always correct, but we do not think it biased our ability to accurately model the relationship between distance and detection or the factors that affect detection.

Factors that influence lek detection. – We constructed models to identify the factors that influence aural detection of a sage-grouse lek. To find the most parsimonious model, we developed and sequentially analyzed 4 model sets of hypothesized covariates. Hypothesized sets included covariates that are known to influence lek-specific behavior and sound attenuation.

Detection of lekking sage-grouse is potentially affected by lek-specific behavior. For example, larger groups of birds are typically detected more frequently than smaller groups (Quinn 1981, but see Hayward et al. 1991). Hence, we expected that leks with more males would be easier to detect compared to those with fewer males. The probability that a bird is available for aural detection varies daily and seasonally (Robbins 1981). Male sage-grouse attend leks most frequently within 1-hr of sunrise (Jenni and Hartzler 1978; Connelly et al. 2003; Walsh et al. 2004; Fremgen et al. 2016; Baumgardt et al. 2017), possibly as an evolved strategy to avoid predators (Boyko et al. 2003). We assumed that the probability of hearing grouse on a lek would follow a negative quadratic relationship with time relative to sunrise. We standardized the starting time for all 1-min surveys based off the difference between the earliest starting time that we ever conducted a survey (i.e., 88 min prior to sunrise) to sunrise. For example, the earliest 1-min survey that we conducted (88 min prior to sunrise) had a value of zero for time relative to sunrise and the values for all other 1-min surveys were positive values that reflected how much later they were relative to that earliest survey. Seasonal variation in male attendance at leks may result from energetic constraints or female presence (Beck and Braun 1978; Jenni and Hartzler 1978; Walsh et al. 2004). Therefore, we expected that aural detection of leks would follow a negative quadratic relationship with date. Grouse, like other birds (York et al. 2014), display during moonlit nights (Hjorth 1968; Archibald 1976; Johnson 1989) which potentially affects lek attendance and grouse behavior in the morning following moonlit nights. Indeed, moon phase influenced the proportion of males and females seen at leks in southern Idaho (Garton et al. 2007). Because moon phase influences lek attendance, we expected a corresponding increase in aural detection probabilities by surveyors. Percent daily moon illumination on the day of each survey was calculated with getMoonIllumination function in the R package suncalc (Agafonkin and Thieurmel 2018). We assumed that increasing temperatures would decrease lek detection because male display rates are constrained by thermal conditions (Vehrencamp et al. 1989) and because higher temperatures increase absorption of sound into the atmosphere (Ingård 1953). Wind speed and sky conditions could also impact sound quality, observer hearing ability, and male lek activity (Ingård 1953; Sherfy and Pekins 1995; Simon et al. 2007; Baumgardt 2017). We treated the sky condition values (1-3; see above) as a continuous variable in this analysis.

Aural detection probability is likely influenced by any factor that influences sound attenuation. Sound intensity level typically decreases rapidly with distance because sound is scattered, deflected, or absorbed as it passes through air (Brumm and Slabbekoorn 2005). Areas with greater vegetation height may cause sound to scatter or be absorbed more readily and thus may lower aural detection probabilities (Morton 1975; Wiley and Richards 1978; Pacifici et al. 2008; Yip et al. 2017). Hence, we assumed that increasing sagebrush height between the observation point and the lek may decrease aural detection probability of leks. Sound is absorbed by the ground (Aylor 1972) and so we assumed that areas with more topographic roughness between the lek and the observation point would likely decrease aural detection probability of leks. We calculated topographic roughness as the total number of hills between each observation point and the lek. We also considered additive and interactive effects of distance and topographic roughness. We assumed that aural detection probability would be higher at survey points on terrain features that are unimpeded by other terrain features (e.g., leks are more likely heard on survey points on hills rather than within valleys).

Background noise can reduce detection probability during bird surveys and detection probability often varies among observers (Link and Sauer 1998; Yip et al. 2017). Low-frequency chronic anthropogenic noise such as mechanical wells can potentially mask avian calls or influence birds' behavior (Blickley and Patricelli 2010; Ortega and Francis 2012; Koper et al. 2016). Similarly, cows (*Bos taurus*), coyotes, and non-focal sage-grouse leks can interfere with a surveyor's ability to hear a focal lek. We assumed that aural detection probability of leks would be lower during surveys with higher levels of background noise. We included two metrics associated with background noise: if cows were bellowing (a binary covariate) and the number of other types of background noise present during the survey.

Audio-technicians. – One of our objectives was to quantify if lek detection probability differed between field surveyors versus lab technicians who listened to digital recordings of the 1-min surveys. We examined the same lek-specific and sound attenuation covariates described above (see Model construction*: Factors that influenced lek detection*. We suspected that the difference in the time necessary to detect sage-grouse calls (i.e., effort) between a field-observer (always 1-min) and an audio-technician (unlimited) may influence lek detection probabilities. Hence, we calculated the effort expended per survey to detect the focal lek for the field-observer (1-min) and the audio-technician (range = 1-10 mins). We used the difference in effort between methods as a potential covariate. We also assessed the effect of the type and intensity (see *Audio-recordings*) of background noise on lek detection probabilities. We considered 4 types of background noise: (1) avian calls or songs, (2) anthropogenic noises, (3) mammal noises, and (4) wind.

Model analysis

Detection distances. - We used data from 2017 to examine whether the detection distance functions differed between aural and visual modes of detection (i.e., whether observers could detect leks from further away via visual or auditory cues). We separately modeled the visual and aural detection functions: the response variable for these models was detection or non-detection (visually or aurally, respectively) of the focal lek and distance was the predictor variable. We facilitated parameter comparison by standardizing distance values to mean 0 and standard deviation of 1. We used binomial generalized linear mixed model (GLMM) in R package lme4 (Bates et al. 2015). We accounted for autocorrelation among point-surveys by including lek as a random factor. We compared β coefficients and 95% CI between aural and visual detection distances. We back-transformed unstandardized estimates from the logit to real scale and we calculated 95% confidence intervals by non-parametric bootstrap sampling procedures.

Factors that influence aural detection probability. - To test for multicollinearity, we used generalized variance inflation factors adjusted for the confidence ellipsoid ($GVIF^{1/(2*df)}$; Fox and Monette 1992) on our global model that included only fixed effects and contained unstandardized and untransformed variables (car package; Fox and Weisberg 2011). We considered variables collinear if $GVIF^{1/(2*df)}$ was ≥ 2.0. Prior to further exploratory analyses and model selection, we standardized continuous covariates to mean 0 and standard deviation of 1 to facilitate comparisons among parameter coefficients given the large numerical differences between parameters. We compared fully parameterized binomial GLM and GLMM regression models, with and without leks as a random factor (Zuur et al. 2009). We stabilized model convergence with a bound optimization by quadratic approximation algorithm (BOBYQA; Powell 2009) set to 10^5 model iterations. A GLMM model structure that included lek as a random factor was favored. Prior to model selection, we compared linear and quadratic effects for time of day and Julian date in separate global models. We used 2 sequential model selection efforts to explore which covariates influenced lek detection probability. In the 1st model selection effort, we explored all covariates that were measured in both 2016 and 2017. In the 2nd model selection effort, we included only 2017 data which allowed us to include all the covariates in the 1st model selection effort (except year) plus cow bellowing and background noise (which we had not recorded in 2016). We assessed influential factors in the 2nd model selection effort with a GLM due to limited sample size. In each of the two model selection efforts, we used 2 successive model selection steps: the 1st step included candidate models with all combinations of lek-specific variables, and the 2nd step included the top-ranked model from the 1st step plus sound attenuation variables. We calculated 95% confidence intervals by non-parametric bootstrap sampling procedures. We used generalized conditional ($GLMM_c$) and marginal ($GLMM_m$)R^2 scores to assess model adequacy for the top-ranked GLMM model (Nakagawa and Schielzeth 2013). We also assessed model performance of the fixed-effect component of all top-ranked models by calculating the area under the curve (AUC) with the R package pROC (Robin et al. 2011).

Effects of multiple visits on aural detection probability. - We assessed the impact of the number of visits per year on aural detection probability and the detection distance function. We used back-transformed distance estimates and bootstrapped 95% confidence intervals from the top-ranked model from the 2nd step of our model selection effort to calculate cumulative probabilities of hearing leks with 1-4 repeated surveys (i.e., 1-4 visits during the lekking season) based on a binomial density function (R base function: dbinom).

Audio-recordings. - We used a Huggins model implemented in Program MARK (Huggins 1989, 1991, White and Burnham 1999) via RMark v3.1.1 (Laake and Rexstad 2012) to compare aural detection probability of leks between a lab technician who listened to digital recordings from portable audio-equipment (recordings examined by a technician in the office months later) and a field surveyor. We created sage-grouse lek detection histories whereby i = heard by field observer and j = detected by a lab technician based on audio recordings made by the field observer. Hence, i and j were 0 if the lek was not detected or 1 if the lek was detected. Our data consisted of numerous (k) detection histories per trial (x^{k}_{ij}). Model convergence was facilitated by a simulated annealing algorithm (Kirkpatrick et al. 1983; Bowker 2008). We tested whether linear or quadratic effects of date and time relative to sunrise were more informative prior to further analyses. We did not include individual lek as a random factor in any of our candidate models because preliminary analysis suggested that models that included it were not competitive ($\Delta AIC_c > 4$). After optimizing our model structures, we evaluated the inclusion of covariates with 2 successive model selection steps similar to the process described previously (see Model Analysis: *Factors that influence aural detection probability*). We used the delta method (Seber 1982) to calculate lek detection probability and associated variance estimates.

Model selection. - We used Akaike's Information Criterion corrected for small sample size (ΔAIC_c) and normalized Akaike model weights (ω_i; Burnham and Anderson 2002) to compare and rank candidate models. We conducted all analyses in Program R (Version 3.1.1, www.r-project.org, accessed 21 Jun 2016). We modeled lek detection as a binary response variable (lek detected or not detected) as a function of the covariates that we assumed influenced detection. We only analyzed data from detection trials if we learned that males were not present on the lek at the end of the trial). We considered models to be competitive if ΔAIC_c was <2. If we found evidence of model uncertainty (i.e., $\Delta AIC_c \leq 2$ for >1 models), we compared the 95% confidence intervals (CI) for the β coefficients among the top-ranked models. We considered a β coefficient meaningful if its associated 95% CI did not overlap 0.

Results

We did not include data from 21 detection trials because either: no males were present on the lek while conducting the 1-min surveys (which we were unable to determine until we conducted the formal lek count at the end of the trial) (*n* = 19). On 24 of the 155 detection trials, we did not detect the lek aurally on any of the 1-min surveys (non-displaying males were present on the lek on 10 of these trials and displaying males were present, but not heard, on 14 of these trials). We conducted trials from 11 Mar to 30 Apr ($\bar{x}$ = 7 Apr) in 2016 and 28 Mar to 4 May ($\bar{x}$ = 17 Apr) in 2017. Minutes relative to sunrise at the start of a 1-min survey varied from -88 to 124 mins in 2016 (0 - 212 mins, after standardizing) and from -48 to 127 mins (40 - 215 mins after standardizing) in 2017. The average number of grouse detected on the formal lek counts (at the end of our detection trials) was 16.1 (SE = 1.2) males and 1.3 females (SE = 0.2).

Detection distances. - We conducted 52 detection trials at 18 leks in 2017 where we recorded aural and visual detections separately. A random-effect structure for lek was strongly supported compared to models without a random factor for lek for both the visual and aural detection models ($\Delta AIC_c \geq 4.47$; $\Delta AIC_c \geq 16.86$; respectively). The probability of visually detecting a lek sharply declined with distance whereas the probability of aurally detecting a lek declined much more gradually with distance (Table 1.1, Fig. 1.1). Visual detection of a lek was slightly higher than aural detection when observers were close to the lek (until approximately 250 m) but vice versa beyond 250 m (Fig. 1.1). Aural detection distance thresholds were substantially greater than visual detection thresholds. For example, we could aurally detect a lek at 300m with a 90% detection probability whereas we could not visually detect a lek at that distance (i.e., visual detection probability was 0) (Fig. 1.1). Our 2017 aural detection model performed moderately well based on AUC (0.72) and R^2 values (0.09, 0.16 for $GLMM_{(c)}$ and $GLMM_{(m)}$ respectively). Visual detection probability models performed well based on AUC (0.65) and R^2 values (0.99, 0.98 for $GLMM_{(c)}$ and $GLMM_{(m)}$ respectively).

Factors that influence aural detection probability of leks. - We conducted 899 1-min surveys during 155 detection trials at 31 leks in 2016-17. None of our explanatory variables were highly correlated ($GVIF^{1/(2*df)} \geq 2.0$). A random-intercept structure (i.e., accounting for among-lek variance) had greater support than a fixed-effects structure ($\Delta AIC_c > 15.1$). Models with quadratic effects for both Julian date and time relative to sunrise had greater support ($\Delta AIC_c > 3.4$) than models with linear effects. In the 1st step, we compared 256 candidate models and 8 of those models had $\Delta AIC_c < 2.0$. Four of the 8 potential explanatory variables considered in the 1st step were included in the top model and all 8 variables were included in at least 1 of the 8 top-ranked models (Table 1.2), but several models were uncompetitive (Arnold 2010) because they included only one additional variable (4 of the 8 possible variables) that had β coefficients with 95% CIs that included zero (Table 1.3). Accordingly, we used the covariates from our 3rd best model as the base for model construction in our 2nd-step. In the 2nd-step, we compared 39 candidate models and 2 of them had $\Delta AIC_c < 2.0$. The top-ranked models in the 2nd step explained 91% of variation in aural detection probability (Table 1.2). Vegetation height received support in the top model but had β coefficients with 95% CIs that included zero, hence we considered it as not meaningful (Table 1.3) and we therefore used our 2nd best model for predictions. Aural detection probability varied by date and time relative to sunrise with a peak near 8 April (Julian date = 109) and near sunrise (Fig. 1.2a-b). Aural detection probability was negatively associated with wind speed (Fig. 1.2c) and temperature (Fig. 1.2d). Increased topographic roughness had a negative effect on the distance an observer could aurally detect a lek (Fig. 1.2e). Aural detection probability was highest if the surveyor was higher in elevation than the lek (Fig. 1.2f). We found no evidence that the terrain feature where the surveyor was standing during the 1-min survey or the sky conditions affected aural detection probability (Table 1.4). Our top model assessing factors that influence aural detection probabilities performed well based on AUC (0.77) and moderately well based on R^2 values (0.25, 0.36 for $GLMM_{(m)}$ and $GLMM_{(c)}$ respectively).

In the 2nd model selection effort (which used only 1-min surveys from 2017), we examined whether cow bellows and other background noise affected aural detection probability of leks. This analysis included 360 1-min surveys on 52 detection trials at 18 leks. Most of the background noise was from non-focal sage-grouse leks (17 %; *n* = 61). We rarely detected anthropogenic noise (5%; *n* = 18), cows (3.6 %; *n* = 10), or coyote yelps (3 %, *n* = 9). First, we compared 127 lek-specific candidate models. Within top-ranked models ($\Delta AIC_c < 2.0$), wind speed, moon luminance, and quadratic effects of both date and time relative to sunrise appeared to influence aural detection probability of leks (Table 1.5). The 95% confidence interval for temperature coefficient overlapped zero suggesting that it is unmeaningful (Table 1.6). Date was included in 66% of our top-ranked models and was found to be biologically relevant in our prior analysis (1st modeling effort). We used our top-ranked model as the base for model construction in our 2nd-step. For the 2nd step of the modeling process, we compared 159 sound attenuation candidate models. Similar to the prior analysis (1st modeling effort), wind speed was negatively associated with aural detection, elevation was positively associated with aural detection, and both date and time relative to sunrise influenced aural detection probability (Table 1.6). Also, vegetation height and sky conditions did not influence aural detection probability of leks. Moon illumination, distance, and background noise were also negatively associated with detection (Fig. 1.3a-b); factors that were not included in the 1st modeling effort. Cows bellowing did not influence detection probability (Table 1.6). Our 2nd phase models performed well based on AUC (0.77) and moderately well based on R^2 values (0.30, 0.30 for $GLMM_{(c)}$ and $GLMM_{(m)}$, respectively).

Effects of survey replication. - Aural detection probability of an active lek increased with repeated surveys (i.e., some of the leks not heard on the first visit will be heard on subsequent visits) such that aural detection probability was near 100% out to 1500 m from the lek (Fig. 1.4). An observer on a single survey had a 65% probability of aurally detecting a lek with grouse present when standing near the lek and detection declined steadily with distance. Grouse did not always make noise on the lek, but the detection probability increased up to 98.5% by the 4th survey when standing near the lek (Fig. 1.4).

Audio-recordings. - We compared lek detection probabilities between a lab technician who listened to audio-recordings made during the 1-min surveys and a field observer present during the actual 1-min surveys in 2017. We conducted 233 audio-surveys with a field observer present on 50 detection trials at 16 leks in 2017. We excluded 5 detection trials (55 surveys) because of one of the following issues: heavy fog, males were not present, or operators failed to record audio data. On 44 surveys, both the lab technician and the field observer did not detect the focal lek and these surveys were not included in the analysis (pursuant to Huggins model specifications). We excluded an additional 32 1-min surveys because the field observer detected ≥ 1 non-focal leks and the lab technician did as well. We removed these 32 surveys because the lab technician had no way to differentiate between focal and non-focal leks based on the audio recordings. The audio-technician detected most (95%) of the grouse aurally from the recordings and only detected ~5% visually (based on the spectrogram). Our analysis included data from 125 1-min surveys during 45 detection trials at 13 leks. Mean difference between audio-technician and field observer survey effort was 1.47 mins. (SE = 0.01, n = 125). Both observers detected the lek aurally at 63.2% (n =79) of the 125 surveys. The lab technician aurally detected the lek (from the audio-recording) but the field observer did not aurally detect it on 30.4 % (n = 38), and the field observer aurally detected the lek, but the lab technician did not detect it on 6.4% (n = 8). Preliminary analysis suggested quadratic effects for both date and time relative to sunrise had greater support (ΔAIC_c = 2.4 - 2.5), so we used quadratic forms of these covariates. On the 1st step, we compared 254 candidate models and 10 of those models had $\Delta AIC_c < 2.0$. All 10 models included moon and survey method (lab versus field technician; Table 1.7). Only moon and survey method had coefficients that were informative in ≥ 1 model (Table 1.8). In the 2nd step, we compared 321 candidate models. Models with $\Delta AIC_c < 2.0$ included distance, terrain, and our main effects from the 1st step (Table 1.7). Most variables except distance and terrain had 95% CIs that overlapped zero (Table 1.9). Our results suggest that aural detection of leks is lower during waxing moon phases and is negatively correlated with distance. Also, the lab technician listening to digital recordings of the 1-min surveys had higher aural detection probability compared to a field observer present during the survey (Table 1.9). Aural detection probability was higher when the surveyor was standing on a hill or knoll, but other terrain categories did not influence aural detection (Table 1.9). When the surveyor was on a hill, the field technician and lab technician had similar detection probabilities (Fig. 1.5).

Discussion

To our knowledge, this is the first study to document the relationship between distance and detection probability of sage-grouse leks and, hence, our results may inform managers in designing more efficacious lek route surveys. Our results document (and control for) the many intuitive factors that influence aural detection probability of leks including distance from observer to lek, time of day, date, weather, topography, and background noise. Conducting repeated 1-min surveys on lek routes would greatly increase the surveyor's probability of detecting a new or a moved lek during lek route surveys and, hence, would reduce the potential for bias in trend estimates (because of less room for spatial and temporal variation in detection rates). Current lek route survey protocols in many states implicitly rely on visual detection of new or moved leks while driving between known leks. With current lek route protocols (that do not include periodic stops), repeated visits are much less likely to improve visual detection probability of a new or moved lek because if a surveyor cannot see a new lek from the survey route on one visit (e.g., because of a knoll or vegetation) they will likely not see it during subsequent visits either. If surveyors exit the vehicle for a 1-min survey along lek routes they obviously would use both senses to detect leks but the improved aural detection (compared to current methods) would reduce spatial and temporal variation in detection probability and, hence, improve the rigor of the count data. Aural detection probability can be further improved if surveyors create audio recordings during 1-min surveys along lek routes and process those audio recordings in the lab later to help detect new or moved leks. This approach would allow surveyors with minimal experience to collect data, would eliminate observer variation in hearing ability, and could potentially be automated for survey routes in remote areas. Once an observer detects a lek at a 1-min survey point (either in-person or in the lab later), a subsequent visit would validate that a lek is indeed present and document the exact location of that new lek. Subsequent lek route surveys would then include that lek in the standard count. Moreover, distance detection functions like those reported in our study can be used to generate density estimates and further improve population trend estimates for sage-grouse.

Our results demonstrate that hearing leks is more effective than seeing leks during surveys, which is typical for avian survey efforts (Brewster et al. 2009; Anderson et al. 2015). Sage-grouse often lek in depressions surrounded by hills or sagebrush, making visual detection of leks difficult in some landscapes (Fremgen et al. 2016; Baumgardt et al. 2017). The magnitude of difference between visual and aural distance detection functions could result from acoustic properties of male sage-grouse advertisement calls. Male sage-grouse calls have low tonal frequency and call repetition (Dantzker et al. 1999) which often equates to low absorption (thus their calls are heard at greater distances) in vegetation, topography, and other environmental factors compared to other high-frequency bird calls (Morton 1975; Wiley and Richards 1978). Despite this, most current sage-grouse lek survey protocols rely almost exclusively on visual detection of new or moved leks.

Surprisingly, male numbers at leks did not affect the probability of an observer aurally detecting leks in our GLMM regression models. The range of variation in male numbers was modest in our study (1-64) and so perhaps we would have detected a relationship between male numbers and detection probability if we had included leks with >100 males. However, 98% of sage-grouse leks in Idaho have <64 males (IDFG, unpubl. data) so our results are relevant to most sage-grouse lek routes. The lack of a relationship between male abundance and lek detection is helpful for monitoring because methods where detection is correlated with abundance make estimates of population trend more challenging. We found no evidence that sky conditions affected aural detection probability; variables which others have found to affect male attendance at leks (Johnson et al. 1989; Garton et al. 2007; Baumgardt et al. 2017). Nevertheless, we did find some evidence that moon stage matters; aural detection probability decreased as the moon waxed. One possible explanation for why moon illumination may influence aural detection is that males may have limited energy after lekking on moonlit nights, which could affect their tonal quality or display rates at sunrise. Appropriate behavioral studies that clarify why sage-grouse display on leks at night and its impact on lekking behavior would potentially help explain these results. Our results suggest that aural detection probability of sage-grouse leks peaks at sunrise and drops off very sharply after sunrise, and that aural detection has a distinct seasonal peak (near 8

April in our study area; Fig. 1.2a-b). Seasonal and diurnal patterns in aural detection are typical in songbirds (Robbins 1981) and ruffed grouse (*Bonasa umbellus*) (Zimmerman et al. 2009) and these temporal patterns corroborate lek attendance patterns reported in other sage-grouse populations (Jenni and Hartlzer 1978; Walsh et al. 2004; Fremgen et al. 2016; Baumgardt et al. 2017). The similarity between diurnal and seasonal peaks in lek attendance rates in other studies and peaks in estimates of aural detection probability (if males are present) may reflect one or more of the following: 1) diurnal and seasonal variation may reflect selection pressures on sage-grouse communication (Ey and Fischer 2009); and 2) increases in the male's vocal or display intensity when female sage-grouse are on leks (Wiley 1973; Patricelli and Krakauer 2009). We detected few females at leks and, hence, did not assess the potential impact of female presence on aural detection probability of leks. Our results also demonstrate that high wind and high ambient temperature reduce aural detection probability. Increasing temperature may increase sound attenuation or males may display less due to thermal constraints (Vehrencamp et al. 1989). Temperature also did not influence detection probability of ruffed grouse (Zimmerman et al. 2009), but wind is well-known to affect aural detection during most avian surveys (Alldredge et al. 2007; Buckland et al. 2015) due to sound scattering and wind turbulence near the human ear. Wind can also decrease overall lek activity in sage-grouse (Sherfy and Pekins 1995). We found evidence that topography roughness influenced the relationship between distance and detection probability; aural detection probabilities did not vary predictably with distance in areas with high topographic roughness, but detection probability decreased with distance in a stereotypical manner in areas with low topographic roughness. Our results suggest that aural detection thresholds of sage-grouse are greater compared to those of northern bobwhite (x = 944 m; Rusk et al. 2009), lesser prairie-chickens (x = 645 m; Butler et al. 2010), and greater prairie chickens (x = 310 m; Raynor et al. 2017). Survey points in elevated areas likely increase aural detection probability due to less sound interference.

Not surprisingly, background noise impacted a surveyor's ability to aurally detect a lek. Background noise could interfere with detection by masking focal sounds, distracting observers, reducing display frequency or intensity, or disrupting lek activity (Blickley and Patricelli 2010; Ortega and Francis 2012; Koper et al. 2016). Our study area does not contain low-frequency chronic noise emitters such as wind turbines (Whalen et al. 2018) and appears to contain only a few mechanical wells that may disrupt lek activity or display behavior (Blickley and Patricelli 2010; Ortega and Francis 2012). We know little about the effects of wind and energy extraction equipment on long-distance communication in sage-grouse. However, wind turbines negatively affect lekking male greater prairie-chickens (Raynor et al. 2017; Whalen 2018). Background noises that reduce a surveyor's ability to detect new or moved leks can create bias in population trend estimates derived from lek counts if the frequency or intensity of background noises changes over time and those changes are not accounted for in the analyses. Future lek counts should record the intensity of background noise and wind speed so that future analysts can control for long-term changes in background noise and wind speed (if any occur) and thereby reduce the likelihood of systematic bias in trend estimates (Pacifici et al. 2008; Conway 2009, 2011; Goodwin and Shriver 2011).

Our results suggest that audio-recordings on 1-min point counts could be used to further increase the probability of detecting new or moved leks during lek route surveys. Our results corroborate past studies on other species that also reported higher detection probabilities for audio-recordings compared to field observers (Hobson et al. 2003; Cellis-Murillo et al. 2012; Tegeler et al. 2012). Multiple observers (Nichols et al. 2000; Alldredge et al. 2006) during surveys allow analysts to measure (and possibly account for) detection error and the use of audio-recordings paired with single field observers can provide estimates of detection probability (as we did here) without the added cost of two observers conducting field surveys. Personnel could potentially deploy automated recorders in remote locations to obtain information on lek presence in remote areas or in areas where leks aren't currently known to occur. We did not use >1 field observer, but our field observer had substantial prior experience and still failed to detect a surprising number of leks that were audible on the recording equipment, which suggests that aural detection probability of sage-grouse leks likely varies among field surveyors. Set-up of audio-equipment in the field and downloading data took little effort but listening to audio clips in the lab/office requires a moderate amount of additional effort (recordings from 125 surveys required 5hrs to process, transcribe, and analyze). However, processing the audio clips afterwards in the lab or office would allow lek routes to be conducted by field staff with less (and more variable) experience and hearing abilities. Moreover, future efforts could potentially automate processing of audio recordings to substantially reduce post-survey processing time. Excluding vehicle maintenance, gas cost, and travel to the field sites, the total cost of 1 person ($13/hr) conducting 125 surveys over 21 days in the field with audio-equipment and analyzing the recordings was ~$1823 (audio-equipment = ~$1,147; pay = ~$546 (survey time: 2 hrs); and analysis = ~$65 (survey time: 5 hrs). The cost of using two people to estimate lek detection probabilities (Nichols et al. 2000; Alldredge et al. 2006) would be greater than the cost of audio-equipment after ~475 surveys.

Future studies in other areas would help document the extent to which our results regarding factors that affect lek detection probabilities differ among sage-grouse populations. We assumed that the number of males generating noise during a traditional lek count was the same as when we conducted a 1-min survey. We could not verify this assumption given our sampling design, but future researchers could compare our 1-min point-count surveys to independent surveys conducted in a blind near the lek. We may have overestimated lek detection probabilities because our aural detection distances may have been influenced by prior knowledge (Riddle et al. 2010); that is, the field observer knew there was a lek present (and the direction and distance to it) during all our detection trials. However, our primary goal was to compare aural and visual detection distance functions; prior knowledge affected both equally and, hence, did not affect our main conclusion that periodic stops along lek routes would increase detection probability and reduce spatial and temporal variation in detection probability. Furthermore, we compared aural detection during a 1-min point count to visual detection during a 1-min point count (with a surveyor standing outside his truck scanning the landscape) whereas most of the current sage-grouse lek route protocols rely on observers driving along a route without stopping to detect new or moved leks. Hence, our estimates of visual detection probabilities are undoubtedly higher than what current survey methods attain. The next step would be to compare naïve observers that conduct 1-min point-count surveys along lek routes to naïve observers that use current survey methods without periodic stops.

Management Recommendations

We recommend that short duration stops be added to existing lek route survey protocols to allow for point-count surveys. To maximize survey efficiency, we recommend 1-min point-counts at intervals of 1km between known lek sites (see Conway et al. 2018) and at known inactive leks. Adding 1-min stops along lek routes would increase the time to conduct lek route surveys but would provide substantial improvements in rigor. Where necessary, splitting or shortening preexisting lek routes to accommodate the 1-min stops would be one way to ensure that lek route surveys stay within the desired time window (Conway et al. 2018).

Our results can be used to improve sage-grouse lek route survey protocols so that lek count data can provide more robust estimates of population trends. Adding 1-min point counts along lek routes would help ameliorate one source of bias. Inferences from lek route surveys would be improved even further if the selection of lek routes were based on a probabilistic method (Johnson and Rowland 2007). Robust sampling approaches have been proposed (Garton et al. 2007) that provide a framework for probabilistic lek route sampling. If the proposed lek route design by Garton et al. (2007) is combined with 1-min point counts along routes, analysts could produce more accurate and more defensible estimates of abundance and population trends for male sage-grouse.

Literature Cited

Agafonkin, V., and B. Thieurmel. 2018. suncalc: Compute sun position, sunlight phases, moon position and lunar phase. R package version 0.4. <https://CRAN.R-project.org/package=suncalc>. Assessed 5 Nov 2018

Alldredge, M. W., K. H. Pollock, and T. R. Simons. 2006. Estimating detection probabilities from multiple-observer point counts. Auk 123:1172-1182.

Alldredge, M. W., T. R. Simons, and K. H. Pollock. 2007. Factors affecting aural detection of songbirds. Ecological Applications 17:948–955.

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29:1294-1297.

Anderson, A. S., T. A. Marques, L. P. Shoo, and S. E. Williams. 2015. Detectability in audio-visual surveys of tropical rainforest birds: the influence of species, weather and habitat characteristics. <http://PLoS ONE 10: e0128464. doi:10.1371/journal.pone.0128464>

Archibald, H. L. 1976. Spring drumming patterns of ruffed grouse. Auk 93:808-829.

Arnold, T. W. 2010. Uninformative parameters and model selection using Akaike's information criterion. Journal of Wildlife Management 74:1175-1178.

Aylor, D. 1972. Sound transmission through vegetation in relation to leaf area density, leaf width, and breadth of canopy. Journal of the Acoustical Society of America 51:411-414.

Baumgardt, J. A., K. P. Reese, J. W. Connelly, and E. O. Garton. 2017. Visibility bias for sage-grouse lek counts. Wildlife Society Bulletin 41:461-470.

Bates, D., M. Maechler, B. M. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1-48.

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

Blickley, J. L., and G. L. Patricelli. 2010. Impacts of anthropogenic noise on wildlife: research priorities for the development of standards and mitigation. Journal of International Wildlife Law and Policy 13:274-292.

Bowker, B. 2008. Ecological models and data in R. Princeton University Press. Princeton, NJ. USA.

Boyko, A. R., R. M. Gibson, and J. R. Lucas. 2004. How predation risk affects the temporal dynamics of avian leks: greater sage-grouse versus golden eagles. American Naturalist 163:154-165.

Brewster, J. P., and T. R. Simons. 2009. Testing the importance of auditory detections in avian point counts. Journal of Field Ornithology 80:178-182.

Brumm, H., and H. Slabbekoorn. 2005. Acoustic communication in noise. Advances in the Study of Behavior 35:151–209.

Buckland, S. T., E. A. Rexstad, T. A. Marques, and C. S. Oedekoven. 2015. Distance Sampling: Methods and Applications. Springer International. Switzerland

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. Springer-Verlag, New York, NY, USA

Butler, M. J., W. B. Ballard, R. D. Holt, and H. A. Whitlaw. 2009. Sound intensity of booming in lesser prairie-chickens. Journal of Wildlife Management 74:1160-1162.

Celis-Murillo, A., J. L. Deppe, and M. P. Ward. 2012. Effectiveness and utility of acoustic recordings for surveying tropical birds. Journal of Field Ornithology 83:166-179.

Coates, P. S., M. L. Casazza, E. J. Blomberg, S. C. Gardner, S. P. Espinosa, J. L. Yee, L. Wiechman, and B. J. Halstead. 2013. Evaluating greater sage-grouse seasonal space use relative to leks: Implications for surface use designations in sagebrush ecosystems. Journal of Wildlife Management 77: 1598-1609

Connelly, J. W., K. P. Reese, and M. A. Schroeder. 2003. Monitoring of greater sage-grouse habitats and populations. University of Idaho, College of Natural Resources Experiment Station Bulletin 80. Moscow, ID, USA

Conway, C. J. 2009. Standardized North American marsh bird monitoring protocols, version 2009-2. Wildlife Research Report #2009-02. U.S. Geological Survey. Arizona Cooperative Fish and Wildlife Research Unit, Tucson, AZ, USA.

Conway, C. J. 2011. Standardized North American marsh bird monitoring protocol. Waterbirds 34: 319-346.

Conway, C. J., A. Meyers, K. Launchbaugh, D. Musil, P. Makela, and S. Roberts. 2017. The Grouse & Grazing Project: Effects of cattle grazing on sage-grouse demographic traits – 2017 Annual Report. College of Natural Resources, University of Idaho. <https://idahogrousegrazing.files.wordpress.com/2018/11/2017_gg_annualreport_final.pdf. > Accessed 4 Feb 2018.

Conway, C. J., E. O. Garton, G. H. Dicus, and J. J. Lonneker. 2018. Monitoring protocol for greater sage-grouse (*Centrocercus urophasianus*) at Craters of the Moon National Monument and Preserve and City of Rocks National Reserve. Natural Resource Report NPS/UCBN/NRR—2018/1789. National Park Service, Fort Collins, Colorado.

COSEWIC. 2017. COSEWIC assessment and update status report on the greater sage-grouse *Centrocercus urophasianus*, *Phaios* subspecies and *Urophasianus* subspecies, *Centrocercus urophasianus urophasianus*, in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa, CA.

Dantzker, M. S., G. B. Deane, and J. W. Bradbury. 1999. Directional acoustic radiation in the strut display of male sage grouse *Centrocercus urophasianus*. Journal of Experimental Biology 20:2893-2909.

Dejong, M. T., and J. T. Emlen. 1985. The shape of the auditory detection function and its implications for songbird censusing. Journal of Field Ornithology 56:213-223.

Department of the Interior. 2015. Endangered and threatened wildlife and plants; 12-month finding on a petition to list greater sage-grouse (*Centrocercus urophasianus*) as an endangered or threatened species; Proposed Rule. 80 (191)

Doherty, K. E., J. D. Tack, J. S. Evans, and D. E. Naugle. 2010. Mapping breeding densities of greater sage-grouse: a tool for range-wide conservation planning. Bureau of Land Management Completion Report, Washington, D.C., USA.

Ey, E., and J. Fischer. 2009. The 'acoustic adaptation hypothesis'- a review of the evidence from birds, anurans and mammals. Bioacoustics 19:21-48.

Fremgen, A. L., C. P. Hansen, M. A. Rumble, R. S. Gamo, and J. J. Millspaugh. 2016. Male greater sage-grouse detectability on leks. Journal of Wildlife Management 80:266-274.

Fox, J., and G. Monette. 1992. Generalized collinearity diagnostics. Journal of American Statistical Association 87:178-183.

Fox, J., and S. Weisberg. 2011. An R companion to applied regression. 2nd Ed. Sage Publications. Thousand Oaks, CA. USA

Garton, E. O., D. D. Musil, K. P. Reese, J. W. Connelly, and C. L. Anderson. 2007. Sentinel lek-routes: an integrated sampling approach to estimate greater sage-grouse population characteristics. University of Idaho, College of Natural Resources Experiment Station Bulletin 88, Moscow, ID, USA

Garton, E. O., A. G. Wells, J. A. Baumgardt, and J. W. Connelly. 2015. Greater sage-grouse population dynamics and probability of persistence. Final Report to Pew Charitable Trusts

Hartzler, J. E. 1974. Predation and the daily timing of sage grouse leks. Auk 91:532-536.

Hayward, G. D., C. B. Kepler, and J. M. Scott. 1991. Point counts from clustered populations: lessons from an experiment with Hawaiian crows. Condor 93:676-682.

Hjorth, I. 1968. Significance of light in the initiation of morning display of the black grouse (*Lyrurus tetrix*). Viltrevy 5:39-94.

Hobson, K. A., R. S. Rempel, H. Greenwood, B. Turnbull, and S. L. Van Wilgenburg. 2002. Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system. Wildlife Society Bulletin 30:709-720.

Huggins, R. M. 1989. On the statistical analysis of capture experiments. Biometrika 76:133-140

Huggins, R. M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. Biometrics 47:725-732.

Hutto, R. L. and R. J. Stutzman. 2009. Humans versus autonomous recording units: a comparison of point-count results. Journal of Field Ornithology 80:387-398.

Idaho Sage-grouse Advisory Committee. 2006. Conservation plan for the greater sage-grouse in Idaho.

Ingård, U. 1953. A review of the influence of meteorological conditions on sound propagation. Journal of the Acoustical Society of America 25: 405-411.

Jenni, D. A. and J. E. Hartzler. 1978. Attendance at a sage grouse lek: implication for spring census. Journal of Wildlife Management 42:46-52.

Johnson, D. H. and M. M. Rowland. 2007. The utility of lek counts for monitoring greater sage-grouse. University of Idaho College of Natural Resources Experiment Station Bulletin 88:15-23.

Johnson, L. L. 1989. Field investigations of lek mating system in sage grouse. Thesis. University of Wyoming, Laramie, WY, USA

Kirkpatrick, S., C. Gelatt, and M. Vecchi. 1983. Optimization by simulated annealing. Science 220:671-680.

Koper, N., L. Leston, T. M. Baker, C. Curry, and P. Rosa. 2016. Effects of ambient noise on detectability and localization of avian songs and tones by observers in grasslands. Ecology and Evolution 6:245-255.

Laake, J. L., and E. Rexstad. 2012. RMark: R Code for MARK Analysis. R Package Version 2.1.1.

LeBeau, C. W., J. L. Beck, G. D. Johnson, R. M. Nielson, M. J. Holloran, K. G. Gerow, and T. L. McDonald. 2017. Greater sage-grouse male lek counts relative to a wind energy development. Wildlife Society Bulletin 41:17-26.

Link, W. A., and J. R. Sauer. 1998. Estimating population change from count data: application to the North American Breeding Bird Survey. Ecological Applications 8:258-268.

Morton, E. S. 1975. Ecological sources of selection on avian sounds. American Naturalist 109:17-34.

Nakagawa, S. and H. Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-Effects models. Methods in Ecology and Evolution 4:133-142

National Oceanic and Atmospheric Administration's INL Weather Center. Undated. <http://niwc.noaa.inel.gov/climate.htm> Accessed 29 May 2018.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393-408.

Ortega, C. P. and C. D. Francis. 2012. Effects of gas-well compressor noise on the ability to detect birds during surveys in northwest New Mexico. Ornithological Monographs 74:78-90.

Pacifici, K., T. R. Simons, and K. H. Pollock. 2008. Effects of vegetation and background noise on the detection process in auditory avian point-count surveys. Auk 125:600-607.

Patricelli, G. L. and A. H. Krakauer. 2010. Tactical allocation of effort among multiple signals in sage grouse: an experiment with a robotic female. Behavioral Ecology 21:97-106.

Powell, M. J. D. 2009. The BOBYQA algorithm for bound constrained optimization without derivatives. Department of Applied Mathematics and Theoretical Physics. Cambridge, England. technical reportNA2009/06

Quinn, T. J. 1981. The effect of group size on line transect estimators of abundance. Pages 502-508 in C. J. Ralph and J. M. Scott, editors. Estimating the numbers of terrestrial birds. Studies in Avian Biology. No. 6

Raynor, E. J., C. E. Whalen, M. Bomberger Brown, and L. A. Powell. 2017. Location matters: evaluating greater prairie-chicken (*Tympanuchus cupido*) boom chorus propagation. Avian Conservation and Ecology 12:17.

Robbins, C. S. 1981. Effect of time of day on bird activity. Pages 275-286 in C. J. Ralph and J. M. Scott, editors. Estimating the numbers of terrestrial birds. Studies in Avian Biology. No. 6

Robin, X. N., Turck, A. Hainard, N. Tiberti, F. Lisacek, J. Sanchez, and M. Müller (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12:77.

Rusk, J. P., J. L. Scott, F. Hernandez, and F. C. Bryant. 2009. Refining the morning covey-call survey to estimate northern bobwhite abundance. Pages 38-45 in S. B. Cederbaum, B. C. Faircloth, T. M. Terhune, J. J. Thompson, and J. P. Carroll, editors. Gamebird 2006: quail VI and perdix XII. 31 May-4 June 2006. D. B. Warnell School of Forestry and Natural Resources, University of Georgia, Athens, USA.

Schroeder, M. A., C. L. Aldridge, A. D. Apa, J. R. Bohne, C. L. Braun, S. D. Bunnell, J. W. Connelly, P. A. Deibert, S. C. Gardner, M. A.Hilliard, , G. D. Kobridger, S. M. McAdam, C. W. McCarthy, J. J. McCarthy, D. L. Mitchell, E. V. Rickerson, and S. J. Stiver. 2004. Distribution of sage-grouse in North America. Condor 106:363-376.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Second edition. Macmillan Book, New York, NY, USA

Sedinger, J. S. 2007. Improving understanding and assessment of greater sage-grouse populations. Pages 43-56 in Reese, K.P. and R.T. Bowyer, editors. Monitoring Populations of Sage-Grouse: Proceedings of a Symposium at Idaho State University Hosted by University of Idaho and Idaho State University, Pocatello, ID. College of Natural Resources Experimental Station Bulletin 88

Sherfy, M. H., and P. J. Pekins. 1995. Influence of wind speed on sage grouse metabolism. Canadian Journal of Zoology 73:749-754.

Tegeler, A. K., M. L. Morrison, and J. M. Szewczak. 2012. Using extended-duration audio recordings to survey avian species. Wildlife Society Bulletin 36:21-29.

Thompson, W. L. 2002. Towards reliable bird surveys: accounting for individuals present but not detected. Auk 119:18-25.

Vehrencamp, S. L., J. B. Bradbury, and R. M. Gibson. 1989. The energetic cost of display in male sage-grouse. Animal Behavior 38:885-896.

Walsh, D. P., G. C. White, T. E. Remington, and D. C. Bowden. 2004. Evaluation of the lek-count index for greater sage-grouse. Wildlife Society Bulletin 32:56-68.

Whalen, C. E., M. B. Brown, J. McGee, L. A. Powell, and E. J. Walsh 2018. Male greater prairie-chickens adjust their vocalizations in the presence of wind turbine noise. Condor 120:137-148.

Wiley, R. H. 1973. Territoriality and non-random mating in sage grouse (*Centrocercus urophasianus*). Animal Behavior Monographs 6:85-169.

Wiley, R. H., and D. G. Richards. 1978. Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. Behavioral Ecology and Sociobiology 3:69-94.

White, G. C. and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46:120-139.

Yip, D. A., E. M. Bayne, P. Sólymos, J. Campbell, and D. Proppe. 2017. Sound attenuation in forest and roadside environments: implications for avian point-count surveys. Condor 119:73-84.

York, J. E., A. J. Young, and A. N. Radford. 2014. Singing in the moonlight: dawn song performance of a diurnal bird varies with lunar phase. Biology Letters 10:20130970. http://dx.doi.org/10.1098/rsbl.2013.0970

Zimmerman, G. S., and R. J. Gutiérrez. 2007. The influence of ecological factors on detecting drumming ruffed grouse. Journal of Wildlife Management 71:1765-1772.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed Effects Models and Extensions in Ecology with R. Springer, NY,

Table 1.1. Distance from the surveyor to the lek affected both aural and visual detection probability of sage-grouse leks; lower (LCI) and upper (UCI) 95% confidence intervals of coefficients for distance did not overlap zero. Standardized β coefficients and standard error (SE) for visual and aural detection distances of greater sage-grouse leks in the Southern Big Butte region, Idaho, USA were based on a General Linear Mixed Model with 52 detection trials at 18 leks in 2017.

Mode of detection	Variable	β	SE	LCI	UCI
visual	distance	-15.50	3.14	-27.71	-9.30
aural	distance	-0.59	0.14	-0.87	-0.30

Table 1.2. Top 2 models that describe covariates that influence aural detection probability of greater sage-grouse leks in the Southern Big Butte region, Idaho, in 2016-17. Models were ranked with Akaike's Information Criterion corrected for small sample sizes (AIC_c) and Akaike's normalized weights (ω_i). We only show the models with $\Delta AIC_c < 2$ and the null model for each step in our analysis. K = Number of parameters, *DD* = quadratic Julian date, *TT* = quadratic time scaled relative to sunrise (mins.), *Sky* = sky conditions, *Temp* = temperature (C°), *Year* = 2016 or 2017, *Wind* = mean wind speed, *Males* = number of males at focal lek, *Moon* = moon luminance (%), *Veg* = vegetation height (m), *Elev* = elevation difference (m) between the observation point and lek center, *Dist* = distance from surveyor to lek; *Hills* = topographic roughness metric based on the total number of hills between the survey point and the lek.

Step 1

Candidate models	K	AIC_c	ΔAIC_c	ω_i
DD+TT+Sky+Temp+Wind	9	1159.28	0.00	0.12
DD+TT+Sky+Temp+Wind+Moon	10	1159.90	0.62	0.09
DD+TT+Temp+Wind	8	1160.23	0.95	0.07
DD+TT+Sky+Temp+Wind+Males	10	1160.25	0.97	0.07
DD+TT+Sky+Temp+Wind+Males+Moon	11	1160.67	1.39	0.06
DD+TT+Sky+Temp+Year+Wind	10	1160.91	1.63	0.05
DD+TT+Temp+Wind+Moon	9	1161.09	1.81	0.05
DD+TT+Temp+Wind+Males	9	1161.26	1.98	0.04
null	2	1218.12	58.84	≤0.01

Step 2

Candidate models	K	AIC_c	ΔAIC_c	ω_i
*DD+TT+Sky+Temp+Wind+Veg+Dist+Hills+Elev+Dist*Hills*	13	1095.57	0.00	0.48
*DD+TT+Sky+Temp+Wind+Dist+Hills+Elev+Dist*Hills*	12	1095.78	0.21	0.43
null	2	1218.12	120.68	<0.01

Table 1.3. Eight top-ranked General Linear Mixed Models ($\Delta AIC_c < 2$) from step 1 of our analysis and their standardized β coefficients, standard errors (SE), and 95% lower and upper confidence intervals (LCI, UCI) from analysis that was designed to explore the factors that affect aural detection probability of sage-grouse leks in the Southern Big Butte region, Idaho in 2016-2017. *DD* = quadratic Julian date, *TT* = quadratic time scaled relative to sunrise, *Sky* = sky conditions, *Temp* = temperature (C°), *Year* = 2016 or 2017, *Wind* = mean wind speed, *Males* = number of males at focal lek, and *Moon* = moon luminance (%).

Model	Est.	*DD*	*TT*	*Sky*	*Temp*	*Year*	*Wind*	*Male*	*Moon*
1	β	-0.22	-0.21	-0.15	-0.23		-0.45		
	SE	0.08	0.07	0.08	0.08		0.09		
	LCL	-0.38	-0.35	-0.32	-0.39		-0.63		
	UCL	-0.05	-0.07	0.02	-0.06		-0.27		
2	β	-0.22	-0.20	-0.15	-0.24		-0.44		-0.09
	SE	0.08	0.07	0.08	0.08		0.09		0.08
	LCL	-0.38	-0.34	-0.32	-0.40		-0.62		-0.25
	UCL	-0.06	-0.05	0.01	-0.08		-0.25		0.06
3	β	-0.26	-0.21		-0.25		-0.46		
	SE	0.08	0.07		0.08		0.09		
	LCL	-0.42	-0.36		-0.42		-0.64		
	UCL	-0.10	-0.06		-0.09		-0.27		
4	β	-0.21	-0.20	-0.15	-0.22		-0.44	0.11	
	SE	0.08	0.07	0.08	0.08		0.09	0.10	
	LCL	-0.37	-0.34	-0.31	-0.38		-0.62	-0.09	
	UCL	-0.05	-0.05	0.02	-0.06		-0.26	0.31	
5	β	-0.22	-0.18	-0.15	-0.23		-0.43	0.12	-0.10
	SE	0.08	0.07	0.08	0.08		0.09	0.10	0.08
	LCL	-0.38	-0.33	-0.31	-0.39		-0.61	-0.08	-0.25
	UCL	-0.06	-0.04	0.01	-0.07		-0.25	0.31	0.05
6	β	-0.21	-0.20	-0.15	-0.20	0.12	-0.45		
	SE	0.08	0.07	0.08	0.09	0.18	0.09		
	LCL	-0.37	-0.35	-0.32	-0.38	-0.24	-0.64		
	UCL	-0.04	-0.06	0.01	-0.02	0.48	-0.27		
7	β	-0.26	-0.20		-0.27		-0.45		-0.09
	SE	0.08	0.07		0.08		0.09		0.08
	LCL	-0.42	-0.34		-0.43		-0.63		-0.25
	UCL	-0.11	-0.05		-0.11		-0.26		0.07
8	β	-0.26	-0.20		-0.25		-0.45	0.12	
	SE	0.08	0.08		0.08		0.09	0.11	
	LCL	-0.41	-0.35		-0.41		-0.64	-0.10	
	UCL	-0.10	-0.05		-0.09		-0.27	0.34	

Table 1.4. Two top-ranked General Linear Mixed Models ($\Delta AIC_c < 2$) from step 2 in our initial model selection effort and their standardized β coefficients, standard errors (SE), and 95% lower and upper confidence intervals (LCI, UCI) from analysis that was designed to explore the factors that influence aural detection probability of sage-grouse leks in the Southern Big Butte region, Idaho in 2016-2017. All candidate models included 4 variables from the top-ranked model in step 1 (Table 1.3) but varied in whether they included 5 additional covariates that we predicted *a priori* might influence aural detection probability. Covariates include: *DD* = quadratic Julian date, *TT* = quadratic time scaled relative to sunrise (mins.), *Sky* = sky conditions (<50 % cloud cover; > 51% cloud cover; precipitation), *Temp* = temperature (C°), *Wind* = wind speed, *Veg* = vegetation height (m), *Dist* = distance from surveyor to lek center (m), *Hills* = topographic roughness metric based on the total number of hills between the survey point and the lek, and *Elev* = elevation difference (m) between the observation point and lek center.

Model	Est.	DD	TT	Temp	Wind	Veg	Dist	Hills	Elev	Dist*Hills
1	β	-0.33	-0.20	-0.31	-0.57	0.14	-0.39	-0.58	0.38	0.33
	SE	0.10	0.08	0.09	0.10	0.09	0.13	0.14	0.13	0.08
	LCL	-0.53	-0.35	-0.50	-0.78	-0.04	-0.63	-0.85	0.12	0.17
	UCL	-0.14	-0.04	-0.13	-0.37	0.31	-0.14	-0.30	0.64	0.50
2	β	-0.32	-0.20	-0.31	-0.58		-0.38	-0.55	0.39	0.32
	SE	0.10	0.08	0.09	0.10		0.13	0.14	0.13	0.08
	LCL	-0.51	-0.35	-0.49	-0.78		-0.63	-0.82	0.13	0.16
	UCL	-0.13	-0.04	-0.13	-0.37		-0.13	-0.28	0.65	0.48

Table 1.5. Top-ranked models relating the effect of covariates on aural detection probability of greater sage-grouse leks in the Southern Big Butte region, Idaho, USA, in 2017. Models were ranked based on Akaike's Information Criterion corrected for small sample sizes (AIC_c) and Akaike's normalized weights (ω_i). We only show models with $\Delta AIC_c < 2.0$ and null models. K = number of parameters, *DD* = quadratic Julian date, *TT* = quadratic time scaled relative to sunrise (mins.), *Wind* = mean wind speed, *Moon* = moon luminance (%), *Temp* = temperature (C°) *Veg* = vegetation height (m), *Dist* = distance from surveyor to lek center (m), *Elev* = elevation difference (m) between the observation point and lek center, *BG* = background noise, and *Hills* = topographic roughness metric based on the total number of hills between the observer point and the lek.

Step 3

Candidate models	K	AIC_c	ΔAIC_c	ω_i
DD+TT+Wind+Moon	8	452.46	0.00	0.20
TT+Wind+Moon	6	454.03	1.56	0.09
DD+TT+Temp+Wind+Moon	9	454.23	1.77	0.08
null	2	478.22	25.75	<0.01

Step 4

Candidate models	K	AIC_c	ΔAIC_c	ω_i
DD+TT+Wind+Moon+Dist+Elev+BG	11	413.55	0.00	0.21
DD+TT+Wind+Moon+Veg+Dist+Elev+BG	12	415.31	1.76	0.09
DD+TT+Wind+Moon+Dist+Hills+Elev+BG+Dist:Hills	13	415.52	1.97	0.08
DD+TT+Wind+Moon+Dist+Terrain+Elev+BG	14	415.55	1.99	0.08
null	2	478.22	64.6	<0.01

Table 1.6. Three top-ranked General Linear Mixed Models ($\Delta AIC_c < 2.0$) from steps 1 and 2 of our second model selection effort describing their standardized β coefficients, standard errors (SE), and 95% lower and upper confidence intervals (LCI, UCI). Each model examined *a priori* predicted effects on aural detection probability of sage-grouse leks in the Southern Big Butte region, Idaho based only on trials in 2017 (which allowed us to examine the effects of 2 covariates that we didn't measure in 2016). *DD* = quadratic Julian date, *TT* = quadratic time scaled relative to sunrise, *Temp* = temperature (C°), *Wind* = mean wind speed, *Moon* = moon luminance (%), *Veg* = vegetation height (m), *Dist* = distance from surveyor to lek center (m), *Elev* = elevation difference (m) between the observation point and lek center, *Hills* = topographic roughness metric based on the sum of hills between the observer point and the lek, and *BG* = background noise.

Step 1

Model	Est.	*DD*	*TT*	*Temp*	*Wind*	*Moon*
1	β	-0.38	-0.45		-0.36	-0.40
	SE	0.17	0.11		0.13	0.14
	LCI	-0.72	-0.66		-0.61	-0.68
	UCI	-0.04	-0.24		-0.12	-0.13
2	β		-0.43		-0.28	-0.23
	SE		0.10		0.12	0.12
	LCI		-0.63		-0.51	-0.46
	UCI		-0.23		-0.05	0.00
3	β	-0.36	-0.45	-0.07	-0.34	-0.42
	SE	0.18	0.11	0.13	0.13	0.14
	LCI	-0.71	-0.66	-0.33	-0.60	-0.69
	UCI	-0.02	-0.24	0.18	-0.08	-0.14

Step 2

Model	Est.	*DD*	*TT*	*Wind*	*Moon*	*Veg*	*Dist*	*Hills*	*Elev*	*BG*	*Dist*Hills*
1	β	-0.51	-0.40	-0.61	-0.53		-0.79		0.59	-0.43	
	SE	0.19	0.11	0.15	0.15		0.15		0.15	0.13	
	LCI	-0.89	-0.61	-0.90	-0.83		-1.09		0.29	-0.68	
	UCI	-0.14	-0.18	-0.32	-0.22		-0.49		0.89	-0.17	
2	β	-0.59	-0.41	-0.61	-0.60	-0.16	-0.81		0.57	-0.40	
	SE	0.25	0.12	0.15	0.22	0.18	0.19		0.16	0.14	
	LCI	-1.07	-0.64	-0.91	-1.02	-0.51	-1.18		0.26	-0.68	
	UCI	-0.10	-0.18	-0.32	-0.17	0.19	-0.43		0.88	-0.13	
3	β	-0.51	-0.38	-0.61	-0.51		-0.82	-0.06	0.58	-0.43	0.24
	SE	0.20	0.11	0.15	0.16		0.19	0.20	0.15	0.13	0.16
	LCI	-0.89	-0.60	-0.91	-0.81		-1.20	-0.45	0.28	-0.68	-0.06
	UCI	-0.12	-0.16	-0.32	-0.20		-0.44	0.32	0.88	-0.17	0.55

Table 1.7. We used a Huggins model to evaluate factors that influenced aural detection probability of greater sage-grouse leks in the Southern Big Butte region, Idaho, USA, in 2017. We compared models based on Akaike's Information Criterion corrected for small sample sizes (AIC_c) and Akaike's normalized weights (ω_i). We only show the models with $\Delta AIC_c < 2.0$ and null models. K = Number of parameters, *DD* = quadratic Julian date, *TT* = quadratic time relative to sunrise (mins.), *Wind* = mean wind speed, *Males* = number of males at focal lek, *Moon* = moon luminance (%), *Dist* = distance from surveyor to lek center (m), *Hills* = topographic roughness metric, *Elev* = elevation difference (m) between the observation point and lek center, *Background* = background noise, *Method* = inherent differences between field observer and lab audio-technician methods, *Terrain* = observation point terrain types: (1) draw; (2) knoll; (3) knoll side; (4) flat ground, *Int* = continuous noise intensity levels; (0) no background noise during most of the survey, (1) faint background noise during > half of the survey, (2) sufficient background noise to likely prevent hearing sage-grouse during < 25% of the survey, (3) sufficient background noise to likely prevent hearing sage-grouse 25-50% of the survey, and (4) sufficient background noise to likely prevent hearing sage-grouse >50% of the survey (modified from Conway 2009), and *Eff* = time difference in effort to detect sage-grouse between audio-technician and field observer surveys.

step 1

Candidate models	K	AIC_c	ΔAIC_c	ω_i
Moon + Method	3	196.49	0.00	0.06
Wind + Moon + Method	4	196.77	0.28	0.05
Temp + Moon + Method	4	197.42	0.93	0.04
DD + Moon + Method	4	197.54	1.05	0.04
Males + Moon + Method	4	197.87	1.39	0.03
DD + Wind + Moon + Method	5	198.19	1.70	0.03
TT + Moon + Method	4	198.29	1.81	0.03
Temp + Wind + Moon + Method	5	198.37	1.88	0.02
DD + Temp + Moon + Method	5	198.43	1.94	0.02
TT + Wind + Moon + Method	5	198.46	1.97	0.02
null	1	223.43	26.94	≤0.01

step 2

Candidate models	K	AIC_c	ΔAIC_c	ω_i
Moon + Method + Terrain + Dist + Elev + Int	9	185.31	0.00	0.07
Moon + Method + Terrain + Dist + Elev + Int + Eff	10	185.59	0.28	0.06
Moon + Method + Terrain + Dist + Int + Eff	9	186.26	0.95	0.05
Moon + Method + Terrain + Dist + Elev + Eff	9	186.52	1.21	0.04
Moon + Method + Terrain + Dist + Elev	8	186.64	1.34	0.04
Moon + Method+ Terrain + Dist + Int	8	186.88	1.57	0.03
Moon + Method+ Terrain + Dist + Elev + Hills + Int	9	187.15	1.84	0.03
null	1	223.43	38.12	<0.01

Table 1.8. Unstandardized β coefficients, standard errors (SE), and 95% lower and upper confidence intervals (LCI, UCI) based on detection probabilities estimated with a Huggins model to compare a field observer and an audio-technician in the Southern Big Butte region, Idaho, USA, in 2017. We ranked models based on Akaike's Information Criterion corrected for small sample sizes (AIC_c) and Akaike's normalized weights (ω_i). We only show models with $\Delta\text{AIC}_c < 2.0$. *DD* = quadratic Julian date, *TT* = quadratic time scaled relative to sunrise, *Temp* = temperature (C°), *Wind* = mean wind speed, *Males* = number of males at focal lek, *Moon* = moon luminance (%), *Method* = detection probabilities differ between field observers and audio-technicians.

Model	Est	*DD*	*TT*	*Temp*	*Wind*	*Males*	*Moon*	*Method*
1	β						-1.92	1.50
	SE						0.58	0.39
	LCI						-3.07	0.74
	UCI						-0.78	2.27
2	β				-0.08		-1.68	1.50
	SE				0.06		0.61	0.39
	LCI				-0.21		-2.87	0.74
	UCI				0.04		-0.48	2.27
3	β			-0.07			-2.05	1.50
	SE			0.07			0.60	0.39
	LCI			-0.21			-3.23	0.74
	UCI			0.06			-0.87	2.27
4	β	-0.41					-1.44	1.50
	SE	0.00					0.74	0.39
	LCI	-0.41					-2.89	0.74
	UCI	-0.41					0.00	2.27
5	β					0.01	-1.91	1.50
	SE					0.01	0.58	0.39
	LCI					-0.01	-3.05	0.74
	UCI					0.03	-0.77	2.27
6	β	-0.78			-0.08		-1.31	1.50
	SE	577.41			0.06		0.75	0.39
	LCI	-1132.51			-0.20		-2.77	0.74
	UCI	1130.96			0.05		0.15	2.27
7	β		2.35				-1.96	1.50
	SE		0.00				0.59	0.39
	LCI		2.35				-3.11	0.74
	UCI		2.35				-0.80	2.27
8	β			-0.05	-0.07		-1.81	1.50
	SE			0.07	0.07		0.64	0.39
	LCI			-0.20	-0.20		-3.06	0.74
	UCI			0.09	0.06		-0.55	2.27
9	β	-1.88		-0.08			-1.57	1.50
	SE	0.00		0.07			0.75	0.39
	LCI	-1.88		-0.21			-3.04	0.74
	UCI	-1.88		0.06			-0.09	2.27
10	β		3.66		-0.09		-1.71	1.50
	SE		948.00		0.06		0.61	0.39
	LCI		-1854.42		-0.21		-2.92	0.74
	UCI		1861.75		0.04		-0.51	2.27

Table 1.9. Variation in unstandardized β coefficients, standard errors (SE), and 95% lower and upper confidence intervals (LCI, UCI) based on comparison between aural detection probability of a field observer and that of a lab technician (who listened to audio files recorded during the surveys) in the Southern Big Butte region, Idaho, USA, in 2017. We used Huggins closed-capture models and ranked models based on Akaike's Information Criterion corrected for small sample sizes (AIC_c) and Akaike's normalized weights (ω_i). We only show models with $\Delta AIC_c < 2.0$. *Moon* = moon luminance (%), *Terrain* = observation point terrain types: (1) draw; (2) knoll; (3) knoll side; (4) flat ground, *Dist* = distance, *Elev* = elevation difference between observation point and lek center, *Int* = background noise intensity, *Method* = difference between observer and audio-equipment, and *Eff* = difference in effort between methods.

Model	Est	*Moon*	*Method*	*Terrain*	*2*	*3*	*4*	*Dist*	*Hills*	*Elev*	*Int*	*Eff*
1	β	-2.16	1.50		2.84	0.37	0.91	≤-0.01		0.07	-0.56	
	SE	0.70	0.39		1.15	0.54	0.64	≤-0.01		0.04	0.31	
	LCI	-3.54	0.74		0.59	-0.69	-0.34	≤-0.01		≤-0.01	-1.16	
	UCI	-0.79	2.27		5.10	1.43	2.15	≤-0.01		0.14	0.04	
2	β	-2.13	1.50		2.79	0.36	0.75	≤-0.01		0.06	-0.53	-0.26
	SE	0.71	0.39		1.16	0.55	0.65	≤-0.01		0.04	0.31	0.19
	LCI	-3.52	0.74		0.52	-0.72	-0.52	≤-0.01		-0.01	-1.13	-0.63
	UCI	-0.74	2.27		5.06	1.44	2.03	≤-0.01		0.14	0.07	0.12
3	β	-2.01	1.50		3.00	0.32	0.72	≤-0.01			-0.63	-0.29
	SE	0.69	0.39		1.17	0.54	0.65	≤-0.01			0.30	0.18
	LCI	-3.36	0.74		0.71	-0.74	-0.55	≤-0.01			-1.21	-0.66
	UCI	-0.66	2.27		5.28	1.37	1.99	≤-0.01			-0.04	0.07
4	β	-2.22	1.50		2.62	0.25	0.69	≤-0.01		0.08		-0.28
	SE	0.70	0.39		1.13	0.54	0.65	≤-0.01		0.04		0.19
	LCI	-3.59	0.74		0.41	-0.80	-0.58	≤-0.01		≤0.01		-0.66
	UCI	-0.85	2.27		4.84	1.30	1.95	≤-0.01		0.15		0.10
5	β	-2.25	1.50		2.67	0.24	0.84	≤-0.01		0.08		
	SE	0.69	0.39		1.12	0.52	0.63	≤-0.01		0.04		
	LCI	-3.60	0.74		0.46	-0.78	-0.39	≤-0.01		0.01		
	UCI	-0.91	2.27		4.87	1.27	2.08	≤-0.01		0.16		
6	β	-2.08	1.50		3.13	0.34	0.94	≤-0.01			-0.67	
	SE	0.68	0.39		1.16	0.53	0.63	≤-0.01			0.30	
	LCI	-3.42	0.74		0.86	-0.69	-0.30	≤-0.01			-1.25	
	UCI	-0.74	2.27		5.41	1.37	2.17	≤-0.01			-0.09	
7	β	-2.08	1.50		2.84	0.37	1.01	0.00	0.07	0.08	-0.57	-0.22
	SE	0.71	0.39		1.15	0.54	0.66	0.00	0.12	0.04	0.31	0.51
	LCI	-3.48	0.74		0.58	-0.69	-0.29	0.00	-0.17	0.00	-1.18	-1.21
	UCI	-0.67	2.27		5.10	1.43	2.31	0.00	0.31	0.15	0.03	0.77

Figure 1.1. Gradual decay of aural (black line) and sharp decline of visual (blue dotted line) lek detection probabilities for greater sage-grouse in Idaho in 2017.

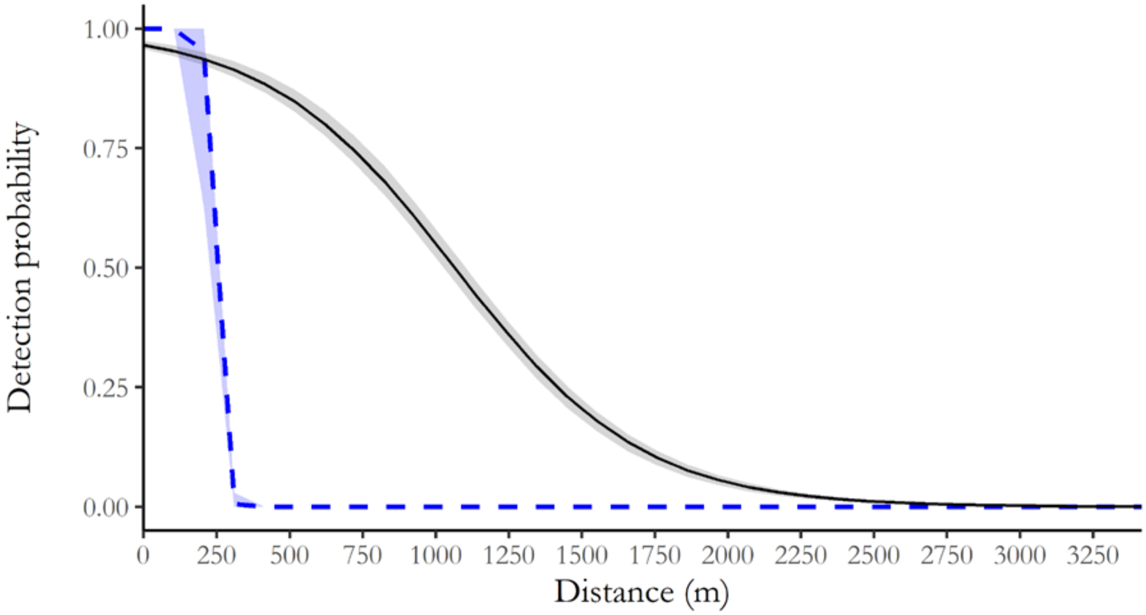
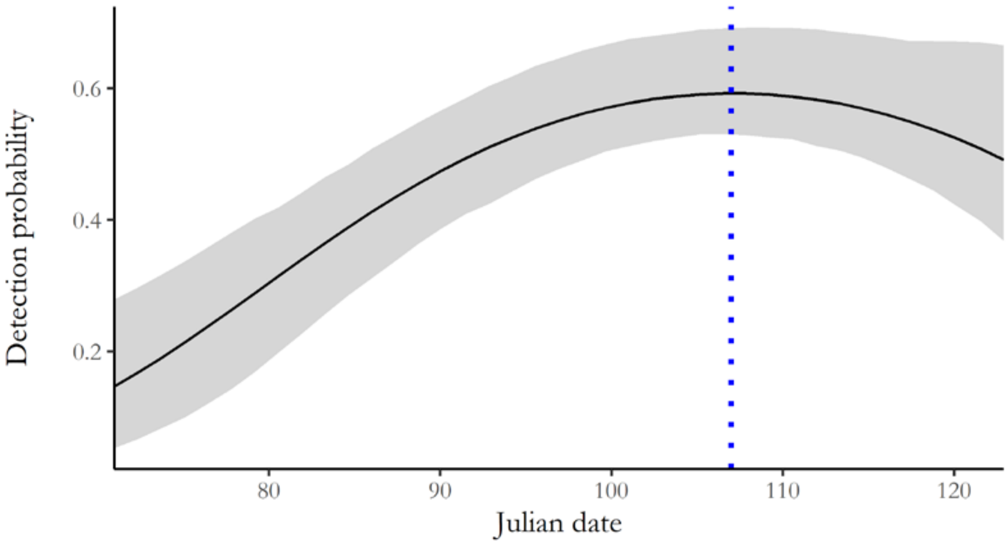
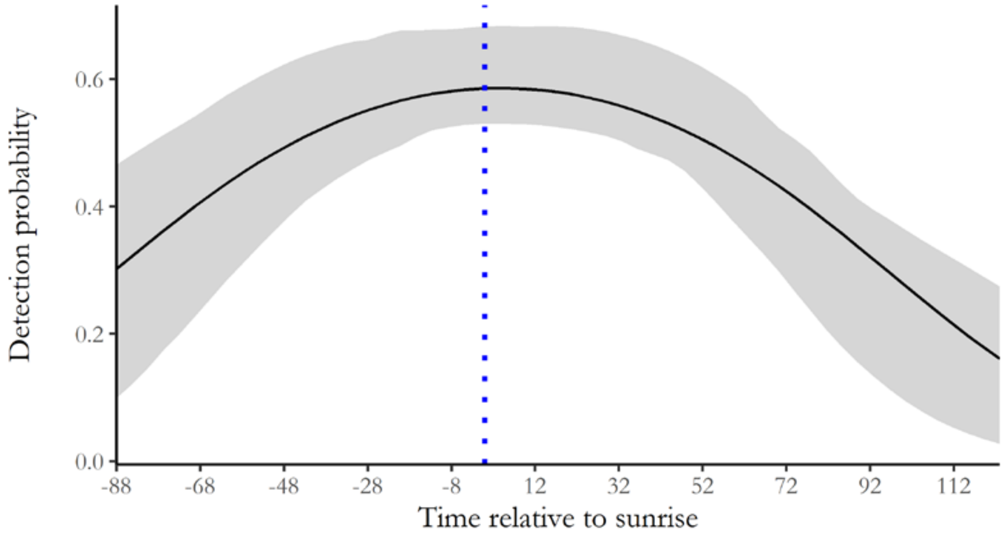


Figure 1.2. Aural detection probability of greater sage-grouse leks was highest in early April (dotted line) (a), highest at sunrise (dotted line) (b), negatively associated with wind speed (c), negatively associated with ambient temperature (d), increased with distance in areas of low topographic roughness (dotted), but did not vary with distance in areas of high topographic roughness (dash) (e), and was positively associated with the elevation difference between the surveyor and the lek (f). Results were based on detection trials at active leks in Southern Big Butte Region, Idaho in 2016-2017.

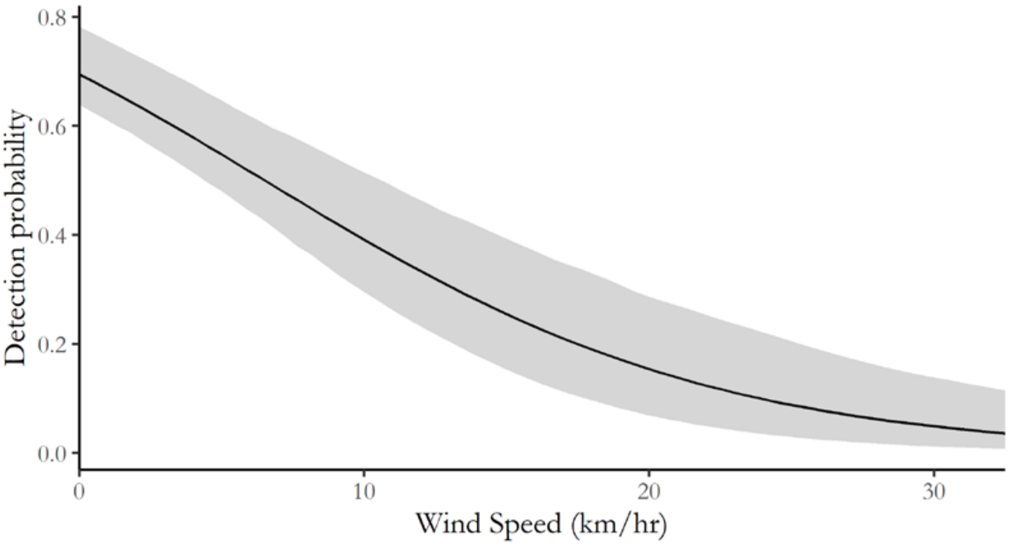
a.



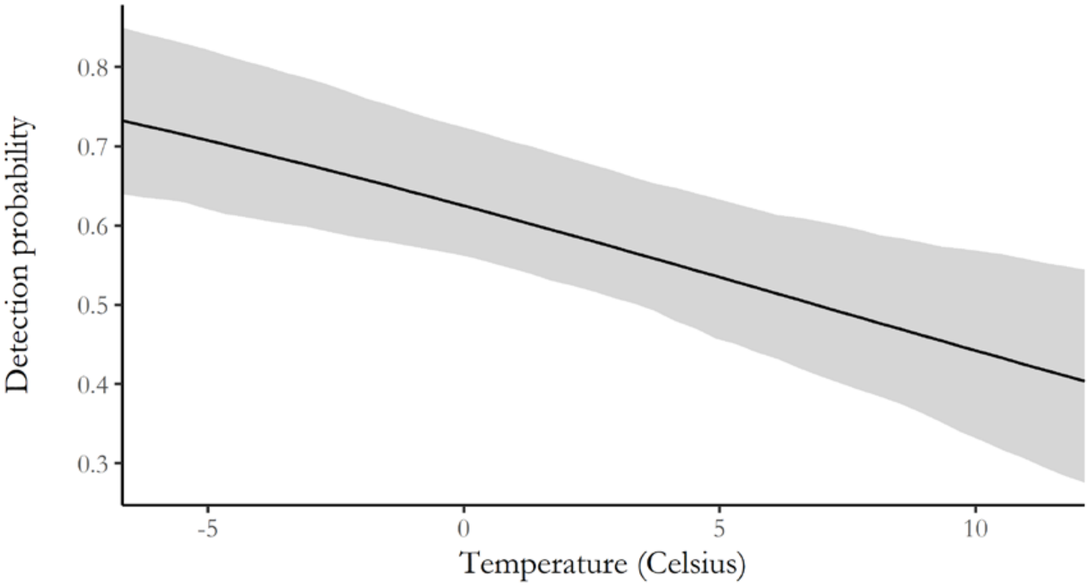
b.



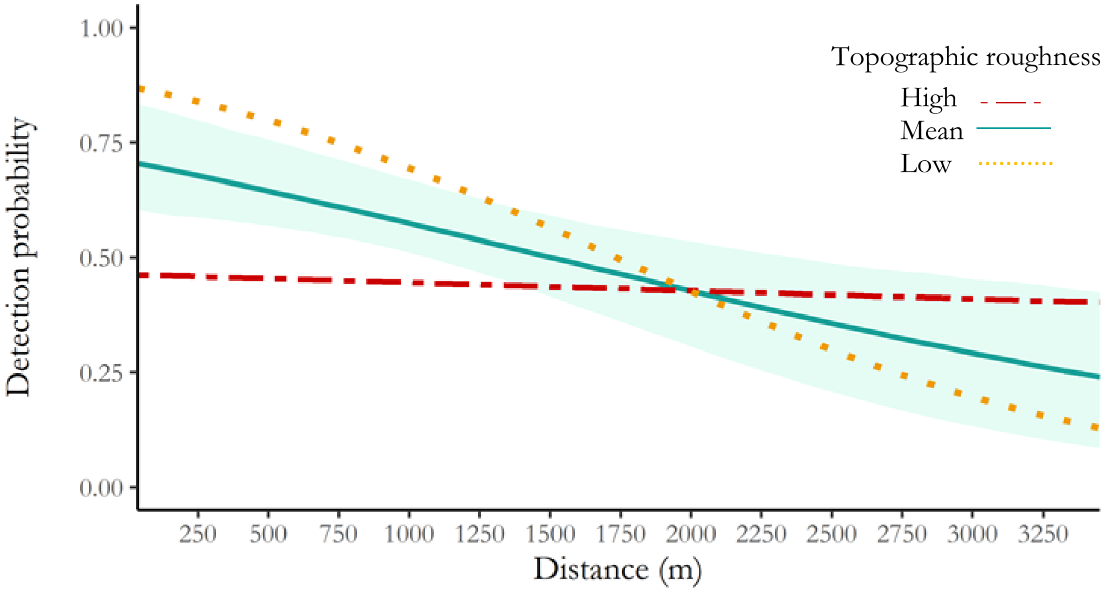
c.


0.8
0.6
0.4
0.2
0.0
Detection probability
0
10
20
30
Wind Speed (km/hr)

d.


0.8
0.7
0.6
0.5
0.4
0.3
Detection probability
-5
0
5
10
Temperature (Celsius)

e.


Detection probability
1.00
0.75
0.50
0.25
0.00
250
500
750
1000
1250
1500
1750
2000
2250
2500
2750
3000
3250
Distance (m)
Topographic roughness
High
Mean
Low

f.

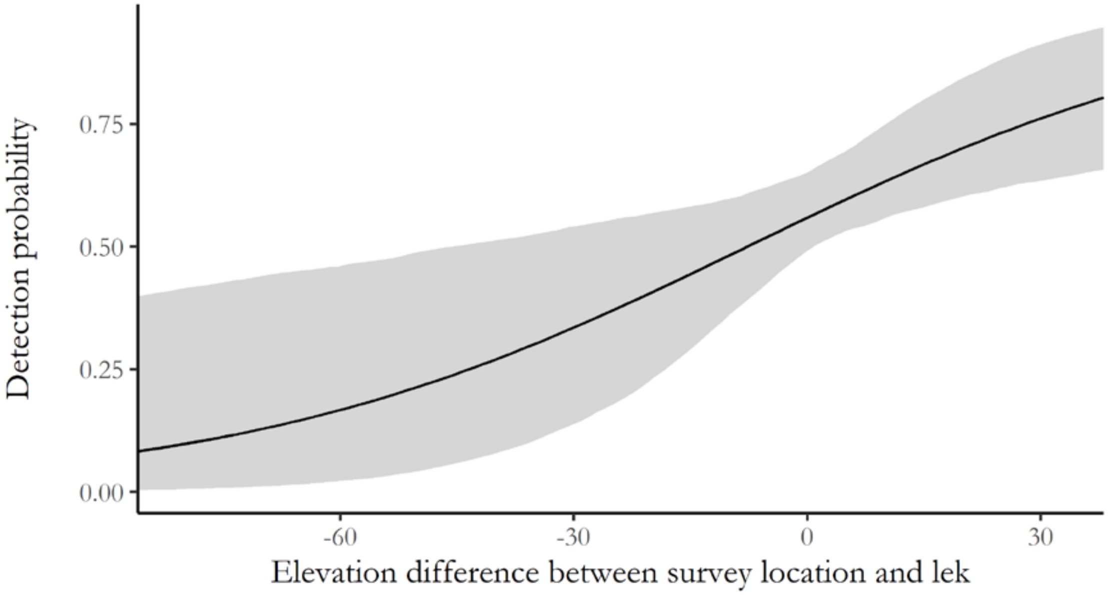
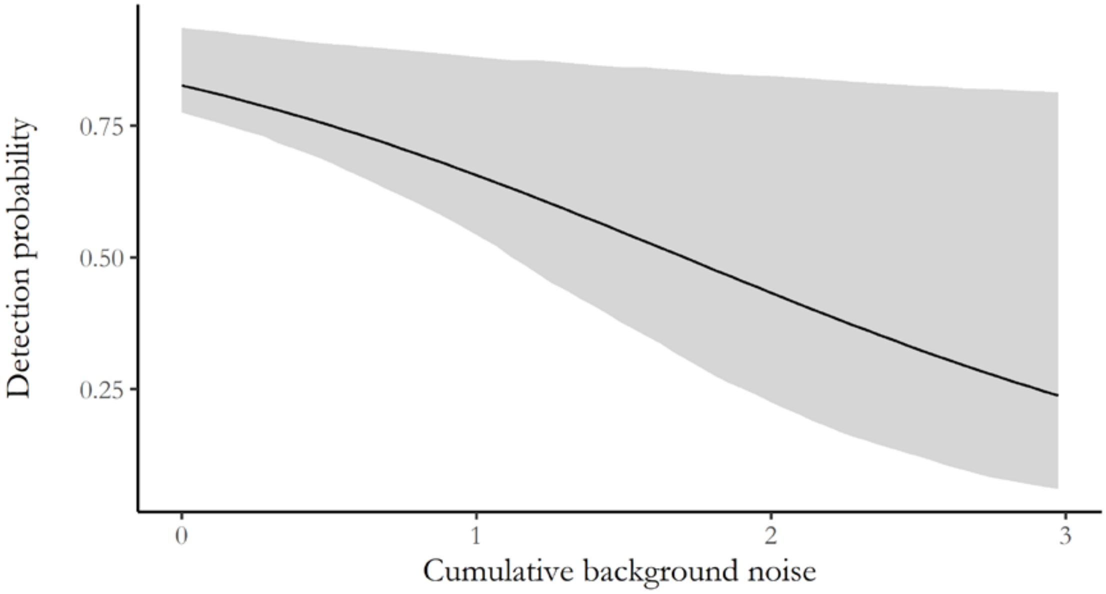

Detection probability
0.75
0.50
0.25
0.00
-60
-30
0
30
Elevation difference between survey location and lek

Figure 1.3. Aural detection probability of greater sage-grouse leks was negatively associated with background noise and moon brightness in Idaho in 2017.

a.



b.

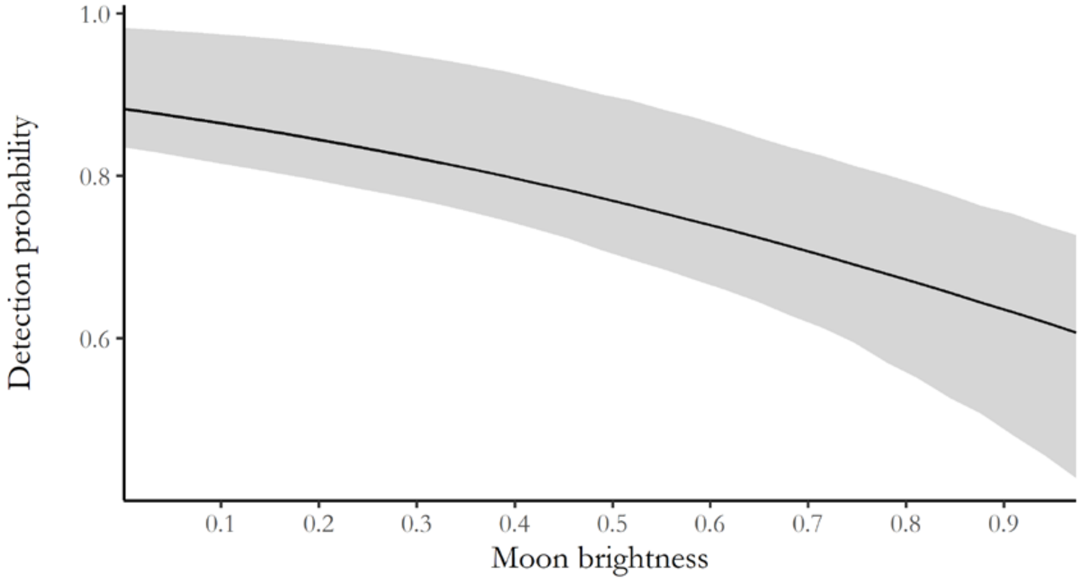


Figure 1.4. Cumulative aural detection probabilities of sage-grouse leks based on 1-4 repeated lek counts performed per year and the interaction between distance and topographic roughness from the top-ranked model (i.e., 2016-2017 data). Darker and thicker lines indicate that lek detection increases when the number of surveys increases from 1 to 4. Aural detection probabilities increase in areas of low topographic roughness (yellow dotted) but do not vary predictably with distance in areas of high (red dash) topographic roughness.

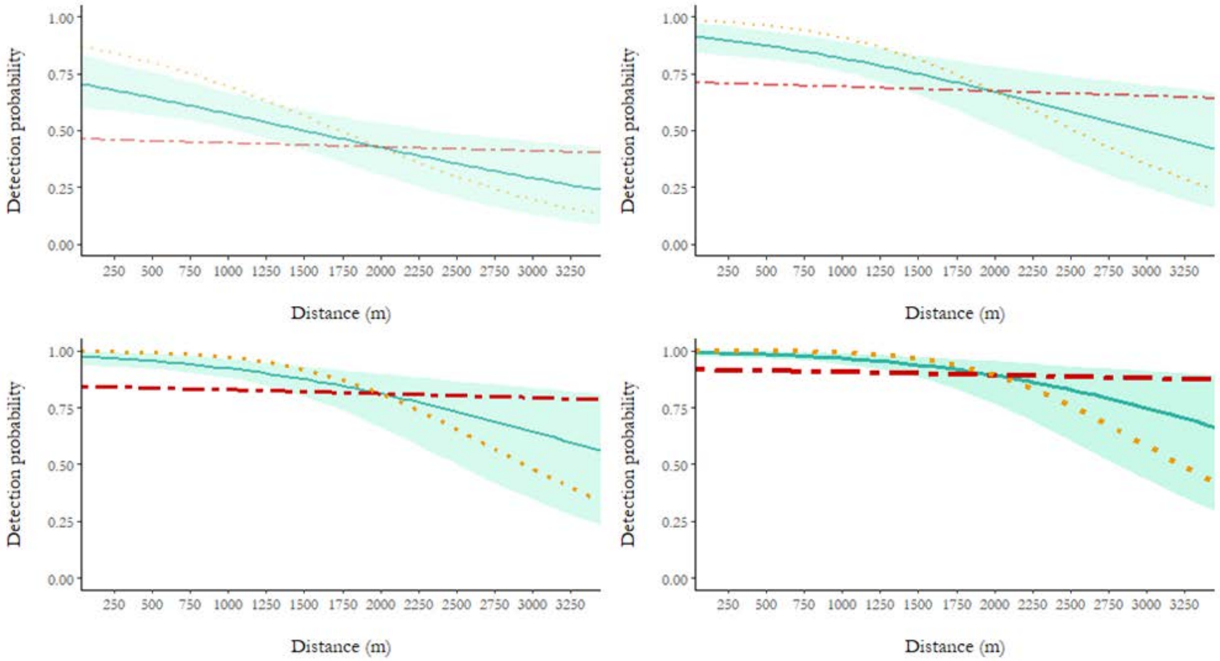
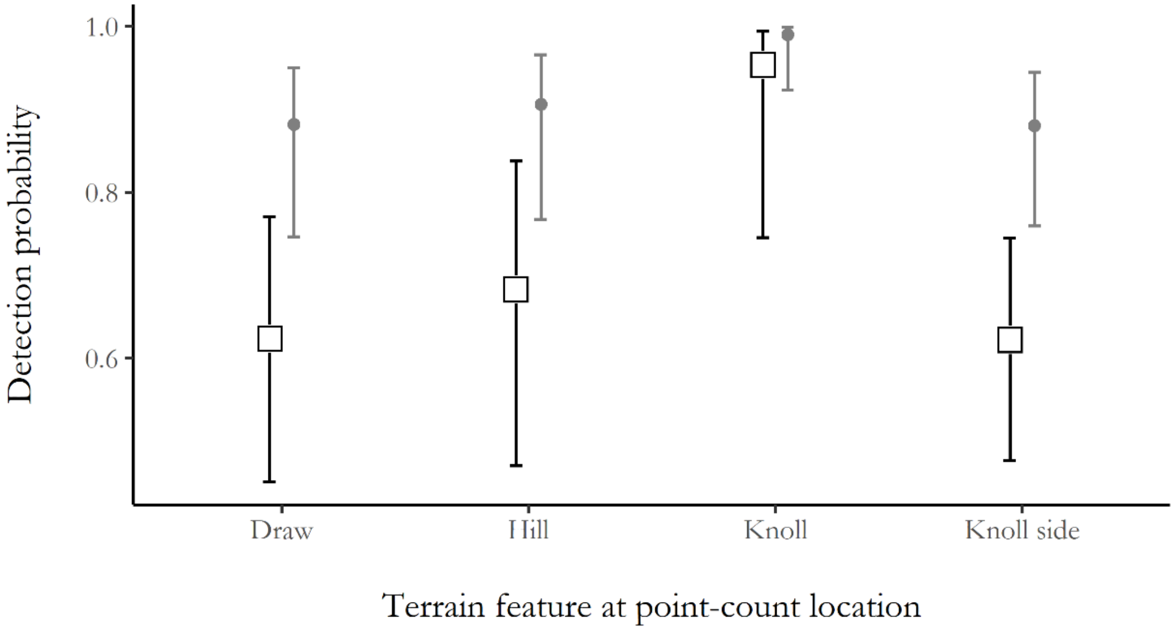


Figure 1.5. Aural detection probability of greater sage-grouse leks was higher for a lab technician (listening to audio files recorded during surveys) compared to a field observer who conducted the surveys, but the effect differed by terrain. Results are based on 125 1-min aural surveys during detection trials at leks in Southern Big Butte, Idaho, 2017. Open-square and black lines represent field observer detection probabilities whereas solid circles and gray lines represent lab technician aural detection probabilities.



Chapter 2: Estimating Detection and Survival Probabilities of Sage-Grouse Broods: A Review and Comparison of Field Methods

Abstract

Research and monitoring data are used to make inferences that shape conservation efforts and policy and such inferences are weakened if investigators incorrectly assume detection of individuals is perfect or that detection does not vary spatially or temporally. Grouse researchers commonly use brood survey methods to estimate brood productivity, occupancy, and habitat selection. However, these methods likely have imperfect detection probabilities due to variation in observer ability, vegetation, and brood behavior. We reviewed 50 greater sage-grouse (*Centrocercus urophasianus*) papers from 1998-2017, documented the brood survey methods used, and identified whether the authors estimated or accounted for detection probability. We also used radio-marked hens with broods to estimate and compare detection probability among three brood survey methods: daytime surveys (flush and visual), nighttime spotlight surveys, and nighttime fecal surveys at roost sites. We used 2 approaches to compare these 3 survey methods: 1) Cormack-Jolly-Seber (CJS) models to estimate brood detection and survival probabilities from 45 radio-marked hens with broods up to 42 days after hatch (DAH), and 2) a double-observer method to compare detection probabilities at 42 DAH. Of the 50 reviewed papers, the most common methods included brood or chick daytime visual surveys (36.2%; *n* = 25), daytime flush surveys (27.5%; *n* = 19), marked chicks (18.8 %; *n* = 13), nocturnal spotlight surveys (14.5%; *n* = 10), or pointing dogs (2.90%; *n* = 2). Nineteen of the 50 papers used >1 method and only 2 of the 50 papers measured or reported detection probability. Studies varied widely regarding the brood age used for estimates of habitat selection, hen productivity, and brood size or brood survival ($\overline{x}$ = 43 days post-hatch, range 14 to 84 days). We found it difficult to compare estimates of brood and chick success or survival because the frequency of consecutive brood sampling visits also varied greatly among studies (range = 1-70 surveys per brood). Furthermore, 22 papers used one or more maternal behaviors as indicators of brood fate but did not validate the utility of these behaviors. From our field trials, daytime visual surveys had brood detection probabilities that varied from 57.2% immediately after hatch to 87.4% at 42 days post-hatch. In contrast, detection probability of fecal surveys did not vary with brood age (75.0%). Our estimates of daily brood survival from fecal surveys suggested an age-dependent pattern of survival where daily survival probability of newly

hatched broods was very low and increases rapidly with age, whereas we failed to detect age-related variation in daily survival from daytime visual surveys. At 42 days after hatch, brood detection probabilities were relatively high (>85% but <100%) for daytime flush surveys, nighttime spotlight surveys, and brood fecal surveys. Our results suggest that brood fecal surveys are a novel survey method that has comparable detection probability compared to daytime flush surveys and minimizes the potential detrimental effects of daytime flushing.

Introduction

In many wildlife studies, counts of individuals are frequently used to infer population trends, occupancy, reproductive output, and habitat selection. However, individuals are inevitably missed during most counts (Nichols et al. 2000; MacKenzie et al. 2002; Gu and Swihart 2004; Gimenez and Gaillard 2017). To improve inference, several methods have been developed to estimate detection probability or account for imperfect detection during counts, including repeated counts (double sampling or double-observer surveys), capture-recapture, and distance sampling (MacKenzie et al. 2002; Williams et al. 2002; Buckland et al. 2015). Nevertheless, many wildlife monitoring programs (Schmidt et al. 2013) and research studies still do not account for imperfect detection with either statistical modeling (Kellner and Swihart 2014) or appropriate sampling methods and designs (Hutto 2016).

Numerous species of grouse (subfamily Tetraoninae) in North America have declined in abundance and are of increasing conservation concern (Storch 2007; Hovick et al. 2014). To effectively manage populations, wildlife managers need accurate information on grouse vital rates. Several authors have emphasized the importance of chick or brood survival in population growth of grouse (Wisdom and Mills 1997; Taylor et al. 2012). Numerous methods have been used to estimate chick and brood survival, but few studies have estimated or compared detection probability associated with these methods. Daytime flush surveys and daytime visual surveys are the most commonly used methods for estimating chick or brood survival of grouse. Daytime flush survey data has been used in the U.S. for >80 yr (Leopold, 1933) to estimate brood survival and hen productivity for ptarmigan (*Lagopus* sp.; Wong et al. 2009), prairie grouse (*Tympanuchus* sp.; Hagen et al. 2005; Goddard and Dawson 2009; McNew et al. 2011), ruffed grouse (*Bonasa umbellus*; Tirpak et al. 2008); and sage-grouse (*Centrocercus urophasianus*; Dahlgren et al. 2010a). During a typical daytime flush survey, observers intentionally flush a hen and subsequently count chicks or record the presence of ≥1 chick (i.e., brood status). Only one study has estimated the detection probability associated with daytime flush count surveys: 72% of greater sage-grouse (*Centrocercus urophasianus*; hereafter sage-grouse) chicks (5-8 weeks old) were detected during daytime flush surveys in Utah (Dahlgren et al. 2010a). Numerous authors have expressed uncertainty in the accuracy of daytime flush surveys for detecting brood status of grouse because grouse chicks have cryptic coloration and behavior (Wing et al. 1944; Godfrey 1975; Kubisiak 1978; Schroeder et al. 1997; Aldridge and Brigham 2001; Walker 2008, Dahlgren et al. 2010a). Additionally, flushing hens with broods may break-up broods, induce partial or full brood abandonment, or increase the risk of chick predation (Schroeder 1997). Therefore, the method used to estimate productivity metrics may affect the very metrics they are intended to measure. To reduce the potential negative effects of flushing chicks and brooding hens during daytime flush surveys, some authors have used daytime visual surveys whereby field personnel attempt to obtain visual observations of the brood from a short distance without flushing the hen or the brood (e.g., Casazza et al. 2011; Lebeau et al. 2017; this paper). However, daytime visual surveys may have lower detection rates than daytime flush surveys, but we are not aware of any estimates of detection probabilities for daytime visual surveys. Moreover, observers may unintentionally flush the hen and brood during daytime visual surveys and we are not aware of estimates of the percentage of visits that observers inadvertently flush hens and broods while conducting daytime visual surveys. Nighttime spotlight surveys are a third brood survey method that has been used for estimating brood survival and hen productivity in grouse (Walker 2008; Dahlgren et al. 2010a). Nighttime spotlight surveys on 42-day-old broods had 100% chick detection probability in Utah (*n* = 21 broods; Dahlgren et al. 2010a), but the study area was dominated by very short (≤ 0.6 m; Fryer 2009) black sagebrush (*Artemisia nova*) and herbaceous cover (Dahlgren et al. 2010a). Further assessment of nighttime spotlight surveys in other regions, at other brood ages, and in other sagebrush systems are warranted because: 1) chicks are likely harder to detect in areas with taller vegetation, 2) hens must sometimes be flushed to see brooding chicks when chicks are young (i.e., its effectiveness might vary with chick age), 3) chicks often do not roost next to the hen when chicks are older, and 4) hen brooding behavior and proximity of chicks to the hen varies with ambient temperature (IR, pers. obs.).

A fourth potential survey method for assessing brood survival and hen productivity that has not previously been used or tested is brood fecal surveys at nighttime roosts. Grouse chicks will often huddle near their mother at nighttime roost sites and defecate distinctive fecal pellets directly underneath or nearby the hen (Popper et al. 1996; Tirpak et al. 2005; pers. obs.). If valid, brood fecal surveys may provide an alternative method that is less invasive and does not have some of the drawbacks of the other 3 brood survey methods.

Past studies that have measured breeding productivity metrics for grouse have used various daytime and nighttime brood survey methods but few have included estimates of detection probability and whether detection varies spatially and temporally. We reviewed the sage-grouse literature to summarize and provide insights regarding field methods used, including: variance, bias, and detection probability. We also used field data to compare detection probabilities and daily survival probabilities among 3 brood survey methods based on detection and non-detection data from repeat surveys.

Study area

Our study was part of a larger sage-grouse project that included estimates of breeding productivity (Conway et al. 2017). We conducted field work at 5 study sites in southern Idaho, including: Big Butte, Brown's Bench, Jim Sage, Sheep Creek, and Pahsimeroi). The study sites fall within Butte (43.37°N, 113.21°W), Twin Falls (42.13°N, -114.77°W), Cassia (42.18°N, -113.44°W), Owyhee (42.01°N, -115.82°W), and Custer counties (44.39°N, -113.75°W), respectively. We conducted brood surveys from May-Jul in 2016 and May-Aug in 2017. The mean and range (min.-max.) precipitation (cm) at each study site during our brood surveys were: Big Butte (2016 = 12.1, 3.1-29.2; 2017 = 15.0, 1.8-25.8), Brown's Bench (2016 = 17.0, 3.6-39.7; 2017 = 3.05, 0.0-8.4), Jim Sage (2016 = 24.2, 5.1-49.5; 2017 = 14.4, 9.7-18.3), Sheep Creek (2017 = 13.1, 2.3-23.4), and Pahsimeroi (2017 = 24.4, 6.6-42.7). The minimum and maximum mean ambient temperature (Celsius) at each study site during our brood surveys were: Big Butte (2016 = 2.5-31.4; 2017 =2.1-33.6), Brown's Bench (2016 = 2.4-31.4; 2017 = 0.7-33.9), Jim Sage (2016 = 3.6-31.1; 2017 = 3.1-32.6), Sheep Creek (2017 = 2.1-32.6), and Pahsimeroi (2017 = 1.3-32.7). All 5 study sites are remote areas with little anthropogenic development, are ≥ 100 km from each other, and contain relatively low abundance of invasive annual grasses. Elevation ranges from 1300-2300m at the 5 study sites. The Bureau of Land Management (BLM) manages cattle production at all 5 study sites. Pastures within the study sites were grazed by cattle only in the spring, only in the fall, or rotated yearly between seasons. Common shrubs at our sites include Wyoming big sagebrush (*Artemisia tridentata wyomingensis*), black sagebrush (*Artemisia nova*), little sagebrush (*A. arbuscula*), three-tip sagebrush (*A. t. tripartita*), mountain big sagebrush (*A. t. vaseyana*), basin big sagebrush (*A. t. tridentata*), and rabbitbrush (*Chrysothamnus visciidflorus*). Common grasses include bluebunch wheatgrass (*Pseudoroegneria spicata*), Sandberg bluegrass (*Poa secunda*), squirreltail (*Elymus elymoides*), crested wheatgrass (*Agropyron cristatum*), Western wheatgrass (*Pascopyrum smithii*), needle-and-thread grass (*Hesperostipa comata*), and Indian rice grass (*Achnatherum hymenoides*). We observed the following potential grouse predators on our sites: American badger (*Taxidea taxus*), coyote (*Canis latrans*), short-eared owl (*Asio flammeus*), great horned owl (*Bubo virginianus*), golden eagles (*Aquila chrysaetos*), rough-legged hawk (*Buteo lagopus*), ferruginous hawks (*Buteo regalis*), Northern harrier (*Circus hudsonius*), American kestrel (*Falco sparverius*), common raven (*Corvus corax*), and black-billed magpie (*Pica hudsonia*). We observed antler-hunting, ungulate and upland bird hunting, and recreational off-road activities at all sites.

Methods

Review

We reviewed articles that quantified metrics associated with sage-grouse broods, including: brood success, brood survival, hen productivity, brood habitat selection, and brood occupancy. We systematically searched Google Scholar for papers irrespective of date with the following keywords: sage-grouse, *Centrocercus*, *Centrocercus urophasianus*, chick, brood, and survival. The author (I.Riley) carefully read the objectives, methods, and results of 264 peer-reviewed journal articles and graduate theses. We limited our search to articles that explicitly used sage-grouse brood survey methods to estimate some metric related to sage-grouse broods. We recorded study duration, number of broods sampled, number of surveys, survey intervals, if broods were flushed or detected visually (without flushing), time of flushes (day or night), the proportion of hens or broods that were flushed, whether the study examined the effects of survey method on hen or brood survival, how brood fate was determined, and the response variables estimated (chick or brood survival, brood success, brood presence, brood habitat use, etc.). We assumed that researchers used a daytime visual survey if the text said that researchers explicitly “located”, “observed”, or “confirmed” chicks or broods. We assumed that researchers used a daytime flush survey if researchers explicitly reported that chicks or broods were flushed. To avoid pseudo-replication, we thoroughly examined the author's publication history for evidence that data were reused among multiple papers. We only selected papers that reused data if that paper also included additional, previously unexamined data.

Field Comparison of Brood Survey Methods

Grouse capture. – We captured sage-grouse hens at night near leks between Feb. and Apr. of 2016-17. We used binoculars and high-powered spotlights to locate night-roosting sage-grouse hens while traveling by foot, ATV, or truck (Giesen et al. 1982; Wakkinen et al. 1992). Once a hen was detected, we used engine noise or played loud music to mask our approach. We captured hens with a large handheld net while the spotlight disoriented the roosting hen. We estimated the age (i.e., adult, yearling, or unknown) of captured sage-grouse based on plumage characteristics (Braun and Schroeder 2015). We attached a metal leg band (size 14) and a 25-g necklace-style radio-transmitter (Advanced Telemetry Systems, Isanti, MN, USA) to captured hens. We obtained locations on radio-marked hens every 2-3 days by triangulation with a handheld Yagi antenna and a VHF radio-telemetry receiver.

Identifying hatch date. – We monitored radio-collared sage-grouse hens to determine nest fate and hatch or fail date. We determined a hen was nesting when she was in the same location (within a 100 m) on 2 consecutive daytime visits. When an observer first assumed that a hen was at a nest (after the 2 consecutive locations at the same location), they attempted to observe the nesting hen from 15-30 m away with binoculars to precisely locate the nest site. If the hen was not detected visually without the risk of flushing her, we used triangulation of the radio signal to approximate the nest location. We established 2 monitoring points 100 m away from the nest, 90° - 150° apart, to safely monitor the nest with radio-telemetry gear. We also created a near-nest monitoring point 30-50 m from the nest to confirm nest occupancy visually when our far monitoring points could not confirm nest occupancy by radio-signal. We then created a detailed map of the nest location relative to the monitoring points. We monitored each nest every 2-3 days. If a hen was not at her nest location during a daytime telemetry visit, we walked into the nest location to determine whether the nest had failed or the eggs had hatched. We examined any remaining eggshells for patterns that suggested hatching or predation (Girard 1937). We estimated hatch date as the date midway between the last date the hen was documented on or at the nest and the first date the hen was detected off or away from the nest. For hatched nests, we obtained locations of hens with broods at 3, 7, 14, 21, 28, 35, and 42 days after hatch (DAH) by using ≥1 of 4 brood survey methods (see below).

Brood fecal surveys. – We conducted brood fecal surveys for radio-collared hens with broods at 3, 7, 14, 21, 28, 35, and 42 DAH. Within 30 mins of dawn, we used the hen's radio-signal to encircle the hen at 20-30 m to find the nocturnal roost-site. While encircling the hen, we kept track of our position and the potential roost-site with the "tracks" feature on a Global Positioning System (GPS). We infrequently determined the exact position of the roosting hens or chicks with binoculars. More often, the hen's radio-signal was used to triangulate the approximate location of the nocturnal roost-site. Once the roost-site was approximated within a confined area, we used a low-lying rock cairn to create a reference point 30-40 m away from the brooding hen. We then created a detailed map depicting major landmarks (unique shrubs, rocks, etc.) and estimated the distance and bearing from the reference point to the presumed roost-site. We returned to the reference point 2 hrs after sunrise. Aided by the map and GPS, we searched the area within 5-m of the presumed roost-site for hen and chick pellets. We identified roost-sites by characteristic scrapings, feathers, and chick fecal pellets (Fig. 2.1). Chick pellets were discernably smaller than the hen's "clocker" pellets or other breeding-age sage-grouse pellets (Fig. 2.2). Observers counted all hen (i.e., normal and clocker) and chick pellets at each roost-site. We did not conduct brood fecal surveys when vegetation was wet or during inclement weather (e.g., rain or snow).

Daytime surveys. – We conducted daytime visual surveys at 7, 14, and 28 DAH and a daytime flush survey at 42 DAH. Observers used the hen's radio-signal to locate hens with broods during the daytime between 2 hours after sunrise and 2 hours before sunset. For daytime visual surveys, we encircled the radio-collared hen and attempted to visually confirm brood presence (i.e., observe ≥1 chick near the hen) without flushing any grouse. If we saw a chick without flushing the hen, we left the area to avoid further disturbance. If a chick was not seen, we continued to search until we either saw a chick or flushed the hen and then searched for chicks within a 15 - 20 m radius. For daytime flush surveys (all at 42 DAH), we located and always flushed the brooding hen and recorded the number of chicks observed (i.e., we always flushed the hen regardless of whether we detected $\geq$1 chick prior to flushing her).

Double surveys. – At 42 DAH, we used an independent double survey approach (Marsh and Sinclair 1989; Fletcher and Hutto 2006; Collier et al. 2013) to estimate method-specific brood detection probabilities for 3 brood survey methods: daytime flush surveys, nighttime spotlight surveys, and roost-site fecal surveys. We attempted to conduct all 3 of these paired surveys on the same day (within 12 hours of exactly 42 DAH), but weather, technician days off, and road conditions created slight variation in brood age when these paired surveys were conducted (i.e., 41-47 DAH). We conducted 59.7% (n = 67) of all 3 paired surveys for the same brood within 24 hrs (range = 0.5-4 days) of each other. The daytime flush survey and nighttime fecal survey methods were described above. For nighttime spotlight surveys, we used a spotlight to locate and count all chicks at night roost-sites between >1 hour after sunset and >1 hour before sunrise (Walker 2008; Dahlgren et al. 2010a). For the brood daytime flush survey and both nighttime surveys, we randomly chose which were conducted first (we always conducted these surveys >6 hrs apart to avoid additional stress). We always conducted the fecal survey the morning after the nighttime spotlight survey.

Model construction. – Many factors potentially influence daily brood survival (φ) in grouse. Most brood mortality in grouse (including sage-grouse) typically occurs within 2-weeks after hatch (Burkepile et al. 2002; Fields et al. 2006; Hannon and Martin 2006). We therefore expected brood survival to increase non-linearly with age. Annual brood survival can vary with moisture availability which likely constrains vital food sources such as forbs and insects (Blomberg et al. 2012; Guttery et al. 2013). Brood survival if often positively correlated with hen age (Caizergues and Ellison 2000; Fields 2006; Guttery et al. 2013, but see Schroeder 1997; Dahlgren et al. 2010b; Ludwig et al. 2010) and hatch date could influence brood survival because differences in weather, resource availability, protective cover, or changes in predator foraging decisions may impact brood survival (Fields et al. 2006; Ludwig et al. 2006; Thompson 2012; Guttery et al. 2013). Hence, we included brood age, hen age, and hatch date as covariates.

Sage-grouse brood detection probability (p) likely differs among survey methods. Brood detection probabilities may also vary with brood age, especially for daytime visual surveys because an observer's ability to visually detect chicks may increase with chick body mass and age-related changes in escape behavior and coloration. Detection probability of brood fecal pellets at roost-sites may be negatively correlated with brood age because brood size tends to decline with brood age (Burkepile et al. 2002; Fields et al. 2006; Hannon and Martin 2006) and young grouse chicks lack the capacity to thermoregulate on cold nights (Aulie 1976). Therefore, younger chicks are more likely to huddle in a group and defecate in tight, highly visible piles underneath the hen (Fig. 2.1; pers. obs.). In contrast, older chicks can thermoregulate and therefore tend to spread out and defecate farther away from the hen at night which may make fecal detection more difficult (pers. obs.). Hence, we explored the effects of brood age on brood detection probability and how these relationships differed among brood survey methods.

Data handling. – Three of our radio-marked hens died during the brood-rearing season and we could not ascertain whether those broods were also killed (brood adoption and brood amalgamation are common in sage-grouse; Dahlgren et al. 2010b). Therefore, we recorded the brood as non-detected after a hen died. The radio signal for 16 brood-rearing hens disappeared prior to their brood reaching 42 DAH (2016 = 8, 2017 = 8) and we were unable to confirm their fate, so we also marked those broods as a "missed" survey (see below) after the hen's signal disappeared. In some instances, we could not conduct a brood survey on the desired date (e.g., 42 DAH) due to road conditions, poor weather, lack of available field personnel, or because the brood was temporally unavailable (e.g., on private property). We created daily detection histories for all broods and marked all days between sampling periods and days when sampling was not conducted as "missed" ("." notation by convention). For example, one detection history for a brood from hatch to 42 DAH was: 1.....1......1.............0...........1, where 1 indicates a day when the brood was surveyed and found alive (i.e., detected), 0 indicates a day when the brood was surveyed and not detected, and "." indicates a day when the brood was not surveyed. Our goal was to conduct the final brood survey at 42 DAH (i.e., cease surveys after 42 DAH), but our final survey was not exactly at 42 DAH (range = 41-47 DAH) for 26 of the 45 fecal surveys and 30 of the 49 daytime visual surveys because of logistical constraints discussed above. For our CJS analysis, we coded all of the final surveys at 42 DAH to simplify the detection histories and ensure model convergence. We only included detection histories for known broods (nests that we knew had hatched); we could not find two nests even after an extensive search and did not detect a brood on any of the subsequent surveys. Two surveys were marked as missed (".") because the focal hen with a brood was flocking with other hens and we could not determine brood ownership. We classified 1 hen's age to adult whose age was unknown at capture.

CJS analysis. – For both fecal surveys and daytime visual surveys, we used a maximum likelihood-based Cormack-Jolly-Seber (CJS) model (Cormack 1964; Jolly 1965; Seber 1965; Lebreton et al. 1992) to estimate brood detection probability and brood survival within a capture-recapture framework. We conducted analyses in the RMark v3.1.1 interface (Laake and Rexstad 2012) of program MARK (White and Burnham 1999). The response variable for both CJS models was φ and p was the detection (or non-detection) of >1 chick fecal pellet or a brood (≥ 1 chick) for fecal surveys and daytime visual surveys, respectively. Like most studies of sage-grouse brood survival, we were unable to account for brood-mixing. Brood-mixing occurs at all ages but seems to occur frequently during the first 21 DAH (Gregg et al. 2007; Dahlgren et al. 2010b; Thompson 2012). Our primary goal was to compare detection probability among brood survey methods, and we assumed that brood-mixing would affect the validity of these methods equally and, therefore, not cause bias in our methods comparison. We developed competing model sets to estimate the effects of explanatory covariates on brood φ and p. For brood φ, we considered models with all logical combinations of the additive effects of hen age (h), Julian hatch date (j), year (y), non-linear age effects log(A + 0.01), or no effect (null). We added 0.01 to our linear effects of age prior to logarithmic transformation to ensure non-zero values. Hen age was treated as a continuous variable because this approximates the progression of a hen's age. Exploratory analysis indicated that hatch date (j) was approximately normal and, therefore, no transformation was needed. For brood detection probability (p), we considered models with all combinations of (A), (A + 0.01), or (null). We used a simulated annealing algorithm (Kirkpatrick et al. 1983) to aid in model convergence (Bowker 2008). Goodness-of-fit tests (GOF) for capture-recapture models (Burnham 1987) are currently unavailable for data sets with missing survey periods. As such, robust estimates of overdispersion ($\hat{c}$; Anderson and Burnham 1994) are also not derivable. Therefore, we assessed goodness-of-fit graphically with deviance residual plots and compared model sensitivity when $\hat{c}$ is adjusted in 0.5 increments from 1-2. We used an information theoretic approach (Burnham and Anderson

2002) to evaluate support for competing models. We ranked models based on quasi-Akaike Information Criterion corrected for small sample size ($QAIC_c$; Burnham and Anderson 2002). Our criteria for assuming models were competitive was $\Delta QAIC_c \leq 2$ (Burnham and Anderson 2002) and normalized Akaike model weights (ω_i). Furthermore, we considered parameters uninformative if the β coefficient's 95% confidence intervals (CI) overlapped zero (Arnold 2010). We used the delta method to estimate the variance of the mean marginal effects (Seber 1982).

Double-survey. – We used a maximum likelihood-based Huggins model (Huggins 1989, 1991) to compare detection probability among 3 brood survey methods on the intended 42 DAH survey (i.e., 41 - 47 DAH). The method provides model parameter estimates of individual covariates (i.e., method-specific detection probabilities) conditional that ≥ 1 chick is detected (Huggins 1989, 1991). We created a detection summary for each brood at 42 DAH whereby h = detection or non-detection from the fecal survey, i = detection or non-detection on the daytime flush survey, and j = detection or non-detection on the nighttime spotlight survey. Therefore, our 42 DAH detection summaries consisted of k^{th} observations per survey combination (i.e., $p_1 = x^{k}_{hij}$; $p_2 = x^{k}_{ihj}$; or $p_3 = x^{k}_{jih}$). To ensure the 3 'observers' (3 methods in our case) are independent, we forced the initial capture probabilities to equal recapture probabilities (Fletcher and Hutto 2006). Given the small sample sizes ($n = 16$), we pooled data among study sites and observers and did not consider these as covariates in the model. Standard closed-population capture-recapture assumptions are defined in Otis (1978). We estimated detection probabilities in program MARK (White and Burnham 1999) via the RMark v3.1.1 interface (Laake and Rexstad 2012). We used the delta method (Seber 1982) to estimate the variance of $p_{1\text{-}3}$.

Results

Review

We reviewed 264 sage-grouse papers that met our search criteria. Of those, 50 papers included brood survey methods (Table 2.1): 32 peer-review papers and 18 graduate theses or dissertations. Most of the study sites (36 of 57) in the 50 papers were within the core sage-grouse distribution (i.e., Idaho, Nevada, Oregon, Utah, Wyoming) and fewer were in states on the periphery of the sage-grouse range (e.g., Alberta, California, Colorado, North Dakota, Saskatchewan, South Dakota, Washington; Schroeder et al. 2004). Studies that have reported sage-grouse brood metrics have increased since 1998. Studies used brood survey methods to: estimate chick or brood success (41.0%; *n* = 32), estimate chick or brood survival (29.5%; *n* = 23), document brood habitat selection (25.6%; *n* = 20), and address life history questions related to hen's fitness and brood-rearing (*n* = 3). Most studies (62%, *n* = 31) used one survey method and 38% (*n* = 19) used 2 or 3 different survey methods. The most common methods (Fig. 2.3) were brood or chick daytime visual surveys (36.2%; *n* = 25), daytime flush surveys (27.5%; *n* = 19), marked chick monitoring (18.8 %; *n* = 13), nighttime spotlight surveys (14.5%; *n* = 10), and pointing dog surveys (2.9%; *n* = 2). Three papers included results from a daytime survey, but the text included insufficient information to classify whether they conducted a daytime flush or daytime visual survey. Nighttime survey methods were primarily used to determine final brood fate (*n* = 7) or validate the results of daytime surveys (*n* = 3). Only 1 paper (Schreiber et al. 2016) used a nighttime survey method throughout the sampling period as their primary brood survey method.

Studies varied widely in study design, sample size of brood or chicks included in the study, how they determined brood or chick fate, chick or brood survival, and statistical methodology (Table 2.1). Sample size of broods monitored varied among studies from 2 to 272 ($\bar{x}$ = 57.8) and sample size of radio-marked chicks monitored varied from 25 to 518 ($\bar{x}$ = 185.7; Table 2.1). All studies sought to estimate some measure of productivity but varied widely in the brood age relative to hatch when productivity was estimated (i.e., when they recorded brood fate; 14- 84 DAH; $\bar{x}$ = 43). The frequency with which the same broods or chicks were resurveyed varied among survey methods from 1-50 times for daytime surveys, 14-78 for radio-marked chick monitoring, and 1-7 times for nighttime surveys. The study that included the most frequent revisits to broods included 648 daytime visual surveys on broods to estimate human-imprinted, released (rather than wild) chick survival rates (Huwer et al. 2008). No study recorded the proportion of hens or broods that were flushed during their brood surveys. Six studies that used radio-marked chicks examined whether their survey method influenced brood survival. Twenty-one studies used the following indirect or ancillary clues to classify brood detection, non-detection, or fate: hen distracting behavior (29.8 %; *n* = 14), hens observed without chicks on $\geq$1 consecutive daytime (31.9%; *n* = 15) or nocturnal (21.3%; *n* = 10) survey, the presence of other yearling or adult sage-grouse (10.6% *n* = 5), or large hen movements (≥ 1-3 km) between previous surveys (6.4%; *n* = 3).

The clues used to document brood fate were not mentioned in 6 of the 50 studies. Mean chick survival ranged from 0-78% and mean brood survival ranged from 10-100% (Table 2.1). Statistical methods varied by paper (Fig. 2.4) and only 1 study provided estimates of detection probability associated with their methods (Table 2.1). Furthermore, 1 paper (Gibson et al. 2017) accounted for chick detection with a Lukacs young survival model (Lukacs et al. 2004) but did not report their estimates of chick detection.

Brood Analysis

We conducted all 3 brood survey methods between 8 May-25 Jul 2016 and 10 May-12 Aug 2017. The range of dates differed slightly between years because the mean hatch date in 2016 (24 May, range = 1 May-5 Jun) was earlier than in 2017 (31 May, range = 8 May-28 Jun).

Fecal surveys. – We explored covariates that influenced sage-grouse brood φ and p based on fecal survey data. Our capture histories consisted of 100 fecal surveys in 2016 and 90 fecal surveys in 2017, representing 45 radio-marked sage-grouse hens and their broods. These hens included 13 yearlings (2016 = 7, 2017 = 6) and 32 adults (2016 = 18, 2017 = 14). We detected chick feces at 67.4% (n = 128) of 190 nocturnal roost-sites (some unknown percent of those non-detections reflect brood loss versus imperfect detection). In 51% of surveys, we detected adult or yearling pellets (i.e., normal or clocker) and the mean number of detected adult or yearling pellets per roost-site was 4.6 (range = 1-57). Of the 67.4% of surveys where we detected chick pellets, the mean number of chick pellets per roost-site was 28 (range = 1-154). We inadvertently flushed the hen on 25% of the fecal surveys (n = 47) and flushed ≥1 chick (n = 12) on 6% of the 190 fecal surveys while triangulating the roost-site location in the early morning.

Our fully parameterized model that was not adjusted for $\hat{c}$ showed good model fit in the deviance plot but some evidence of overdispersion (Fig. 2.5). When we incrementally adjusted $\hat{c}$ (Table 2.2), the models $\varphi\{y + \log(A+0.01)\}$ $p\{c\}$ and $\varphi\{y + h + \log(A+0.01)\}$ $p\{c\}$ were frequently favored (80%) as the top-ranked models, again suggesting good model fit. We conservatively corrected for model $\hat{c}$ to its mean value (i.e., 1.5) due to inherent uncertainty seen in the deviance plot and model sensitivity. CJS modeling suggested that brood survival (φ) differed between years and varied non-linearly with brood age (Table 2.2).

We also found support in our 2nd best model that adults may produce broods that have lower brood survival compared to yearlings. Our 3rd best model included hatch date indicating that broods hatched later in the year may have lower survival (Table 2.2). Our best-supported model suggests daily brood survival increases nonlinearly with brood age ranging from 96.8% just after hatch to 99.9% by 42 DAH in 2016 and from 99.8% just after hatch to 99.9% by 42 DAH in 2017. Detection probability of broods during fecal surveys did not differ between years and was not affected by chick age (Tables 2.3). Brood detection probability during fecal surveys was constant at 75% (95 % CI = 65.1 - 82.9).

Daytime visual surveys. – For daytime visual surveys, our capture histories consisted of 83 visual surveys in 2016 and 48 visual surveys in 2017 representing 49 hens: 14 yearlings (2016 = 8, 2017 = 6) and 35 adults (2016 = 22, 2017 = 13). We visually detected >1 chick on 56.5% (*n* = 74) of the 131 daytime visual surveys. The 57 non-detections reflect an unknown percentage of brood mortality versus imperfect detection. We inadvertently flushed the radio-marked hen on 93.1% (*n* = 122) of the 131 surveys and inadvertently flushed ≥1 chick on 35.9% (*n* = 47) of the 131 surveys. Deviance plots suggested good model fit but some evidence of overdispersion (Fig. 2.6). Our sensitivity analysis also suggested good overall fit (Table 2.4). We corrected overdispersion by the mean adjustment value ($\hat{c}$ = 1.50) to conservatively accommodate the slight overdispersion that we observed in the deviance plot and small lack of model fit. Brood survival was influenced by hatch date and detection probability varied by brood age (Table 2.4). All models containing hatch date and brood age had coefficients whose 95% C.I.'s overlapped zero (Table 2.5). Our results suggest that none of our covariates influence brood survival or detection probability according to our criterion (Table 2.4; 2.5). However, models containing detection probability parameters for brood age had coefficients whose 95 C.I.'s just barely overlapped zero (Table 2.5). If we accept that constant-only model is best supported, then daily brood survival was 99.2 % (95% CI = 97.8 - 99.7) and brood detection probability was 71.2 % (95% CI = 58.1 - 81.5). If we accept that brood age is meaningful, the top-ranked model suggests that daily brood survival is high (99.1% ± 95 CI = 97.9 - 99.6; Table 2.5) and brood detection probability increased linearly from 57.2 % (95% CI = 36.8 - 75.4) just after hatch to 87.4% at 42 DAH (95% CI = 63.7 - 96.5).

Nighttime spotlight surveys. – We conducted 26 nighttime spotlight surveys at 42 DAH. We detected ≥1 chick on 65.4% (*n* = 17) of the 26 nighttime spotlight surveys. We inadvertently flushed the hen on 30.8% (*n* = 8) of the 26 surveys and inadvertently flushed ≥1 chick on 11.5% (*n* = 3) of the 26 surveys.

Daytime flush surveys. – We conducted 31 daytime flush surveys. We flushed the hen on 90.3% (*n* = 28) of the 31 surveys and flushed ≥1 chick on 71.0% (*n* = 22) of the 31 surveys. We flushed both the hen and chicks on 64.5% (*n* = 20) of the 31 surveys. The other hens and chicks ran away and did not flush despite our efforts to flush them during daytime flush surveys.

Double-surveys. – Detection probability at 42 DAH was high for all 3 survey methods (87.5% - 93.7%; Table 2.6).

Discussion

Our results corroborate and build upon recent studies (Taylor et al. 2012) that have mentioned the lack of standardization in sage-grouse brood survey methods and sampling design. Based on our review of 50 papers, investigators have primarily used daytime visual surveys and daytime flush count surveys to estimate brood productivity metrics of sage-grouse and most have used these methods without estimating or accounting for variation in detection probability. Indeed, we found only 2 studies (Dahlgren et al. 2010a; Gibson et al. 2017) that estimated or accounted for detection probability when using daytime brood survey data to estimate brood or chick survival. We found wide disparity among studies in the cues or triggers used to infer brood fate, and substantial variation among studies in the frequency and duration of monitoring visits. This variation makes comparisons among studies difficult and limits an investigator's ability to put their study results into proper context. Variation in detection probability among brood survey methods is most pronounced at younger brood ages and detection probability varies with brood age for some methods and not others. Better inferences would be possible if investigators, grant reviewers, and journal editors implemented or required age-specific estimates of brood detection probability in future grouse studies. Moreover, we lack estimates regarding the effects of most brood survey methods (except the radio-marked chick method) on chick survival or body condition in sage-grouse. And, we found no studies that reported the proportion of times that hens and broods were flushed during their brood surveys. Such metrics should be reported in future studies so that investigators can make more informed decisions regarding which brood survey methods to use and how many repeated visits to use.

Past studies have often inferred sage-grouse brood fate based on untested assumptions about hen behavior. For example, hens with broods will sometimes act "broody" or use protective or distractive behaviors when approached by humans. But the extent to which the following behaviors serve as reliable cues of brood status (alive or dead) need verification: staying close to the flush site, feigning injury (e.g., broken-wing or wing-drag display), rushing towards the observer (Atamian et al. 2010; Lebeau et al. 2017), etc. Willow ptarmigan (*Lagopus lagopus*) use some of these defensive behaviors when they have broods (Sandercock 1994), but the validity of this behavior for inferring brood fate has not been quantified with any grouse species to our knowledge. Moreover, some studies have inferred brood fate based on hen movements (e.g., assumed brood mortality when a hen moves >1km between subsequent telemetry locations or in response to a flush count survey; Moynahan 2004; Dzialak et al. 2011) or when a hen flocks with other hens (e.g., Sandford et al. 2017). However, we only found unquantified or unsubstantiated evidence that large movements by a hen indicate recent brood mortality (Thompson 2012) and hens with intact broods (even those with very young broods) will move 1-3 km in a day (unpubl. data). Use of hen behaviors to infer brood presence, absence, or fate is unwise until future research provides evidence that such behaviors are reliable indicators of brood fate.

Use of radio-marked chicks or pointing dogs are alternative methods to sample hen productivity, chick, or brood success but few studies have reported detection rates with these methods. Pointing dogs are used extensively in Europe and less commonly in North America to locate chicks or broods (reviewed by Dahlgren et al. 2012). Pointing dogs located 96% of 21 5-8-week-old sage-grouse chicks (Dahlgren et al. 2010a), but the accuracy of this method has not been tested on younger sage-grouse broods and detection probability likely varies among dogs (Orange et al. 2017), vegetative communities, trainers, weather, etc. Detection probability of radio-marked chicks (Larson et al. 2001; Burkepile et al. 2002; Gregg et al. 2007; Dahlgren et al. 2010b) is often assumed to be 100%, but adverse reactions to handling or marking may confound survival estimates (Amundson and Arnold 2010; Taylor et al. 2012) and lost signals are a form of imperfect detection (and right-censoring of missing signals can produce biased estimates). Lastly, radio-marked chicks may not accurately estimate brood survival unless researchers are certain that all chicks within a brood are radio-marked.

Appropriate statistical methodology and sampling methods are needed to increase the accuracy of important brood productivity parameters. Brood productivity estimates are likely underestimated if studies fail to explicitly account for imperfect detection. Biased estimation likely varies among studies based on brood age when surveys were conducted and by differing survey methods. Such biases hinder comparisons of brood survival across studies based on treatments, land-uses, or regions. Common statistical methods used to analyze daytime surveys in our review include Kaplan-Meier, nest survival models, and known-fate models (Fig. 2.4) that require clear binary outcomes (Andersen and Gill 1982; Dinsmore et al. 2002; Williams et al. 2002: 343). It is unclear whether surveys without marked chicks or without complete observations can achieve this outcome; although some authors report that repeated surveys or the use of nighttime surveys is enough to validate brood fate (e.g., LeBeau et al. 2017). Given that imperfect detection likely occurs in most brood survey methods, future studies can reduce variance in survival estimates by using CJS models (Lebreton et al. 1992; Lukacs et al. 2004; Schaub and Royle 2014), conducting >1 visit, or using >1 method (Williams et al. 2002).

This study is one of the first (also see Gibson et al. 2017) to estimate and incorporate imperfect detection with a capture-recapture CJS model to estimate brood survival in sage-grouse or to estimate differences in method-specific detection probabilities based on a double-survey approach. We demonstrated that daytime surveys (daytime visual and daytime flush surveys) have <100% detection probability and, importantly, detection probability varies with brood age. In contrast, a novel, less-invasive survey method (brood fecal surveys) had high detection rates that did not vary with brood age. Variation in detection probability based on age of chicks is particularly problematic for sage-grouse because brood-mixing is relatively common (>45% of broods; Dahlgren et al. 2010b) and amalgamated broods often include chicks of different ages. Moreover, comparisons across studies, years, or management treatments will be biased if daytime brood surveys are conducted at different ages and investigators used survey methods that includes age-specific variation in detection probability. Our review of past studies indicated that studies varied greatly in this regard.

Our results corroborate past studies (Dahlgren et al. 2010a) indicating that nighttime spotlight surveys have relatively high brood detection probabilities at 42 DAH. However, nighttime spotlight surveys were much less effective for broods <28 DAH and we stopped conducting them on younger broods due to difficulties observing chicks without pushing the hen off her chicks prior to 42 DAH (unpubl. data). Fecal surveys and daytime flush surveys had high and comparable detection probabilities at 42 DAH. None of the methods had 100% detection, even at 42 DAH, so 2 visits are needed if investigators want to determine brood fate definitively (without error). To our knowledge, our results are the first to document that brood detection probability varies with brood age with daytime visual surveys and one of the first studies to document the relationship between daily brood survival and brood age.

Our results revealed important differences in brood detection between survey methods. Fecal surveys detected broods better than daytime visual surveys for younger-aged broods and performed equally well for broods 28-42 DAH. Fecal surveys had relatively high and constant probability of detecting broods despite the wide range of chick fecal pellets detected at nighttime roost sites (Fig. 2.7). Fresh chick feces, even if only a few pellets, are visible because of their high contrast against the ground (Fig. 2.2). In contrast, daytime surveys had detection probability that increased with brood age. Smaller, younger chicks likely rely more on camouflage for defense whereas older, larger chicks rely on flight to escape danger making older chicks more detectable during daytime visual or flush surveys. Age-specific detection probability makes it more difficult to compare estimates across studies if those studies conducted brood surveys at different (or unknown) brood ages and failed to explicitly estimate detection probability even if those studies used the same method.

We found important differences in daily brood survival between fecal surveys and daytime visual surveys. Fecal surveys are less invasive than daytime visual or flush surveys and, hence, we were comfortable conducting them more frequently on broods and conducting them on young broods (3-7 DAH). The added sampling frequency with fecal surveys allowed us to document that brood survival was lowest within the first week after hatch and increased rapidly thereafter. In contrast, we failed to detect any age-related differences in daily brood survival with daytime visual surveys (because we did not want to disturb broods <7 DAH with daytime surveys). Furthermore, we found some evidence that year, hen age, and hatch date influence brood survival based on fecal surveys (Table 2.3), but we found no such evidence with daytime visual surveys. Our estimates of daily brood survival based on fecal surveys (96.8 % - 99.9 %) was comparable to daily chick survival in Utah (0.987-0.994%, *n* = 150 chicks; 21 broods; Dahlgren et al. 2010b). However, our daily survival estimates were generally higher compared to estimates of brood survival from many other studies (Table 2.1). We found that brood survival increased quickly within the 1st week after hatch, a result that corroborates results from other grouse species (Zwickel and Bendell 1966; Moss et al. 1980; Larson et al. 2001; Ludwig et al. 2010). With fecal survey data, we found some evidence (3rd best model in Table 2.3) that late-hatched broods have lower survival than earlier hatched broods; a pattern also reported in Colorado (Thompson 2012). We failed to detect differences in brood survival between yearling and adult hens as others have (Schroeder 1997; Fields 2006; Dahlgren et al. 2010b).

Our results suggest that fecal surveys are an alternative brood survey method that has several benefits and one drawback compared to other methods. Compared to daytime visual or flush surveys, fecal surveys had greater or equal detection probabilities than daytime visual surveys to 42 DAH and were less likely to flush the hen or brood. Detection probability for fecal surveys might be even higher if the sampling timeframe at nighttime roost-sites (i.e., within 30 mins of dawn) was earlier (e.g., 90-30 min prior to dawn). On several occasions, we failed to detect the nighttime roost-site because hens had already begun moving when we arrived just before dawn. Future researchers can alleviate this problem by finding the radio-marked hen earlier (e.g., 100-60 min prior to dawn) or by using high-powered spotlights or thermal imaging cameras (Mitchell et al. 2019) to find the roost-site. Or, fecal surveys and spotlight surveys could be combined into a single method to maximize detection under all conditions and at a variety of ages to take advantage of the benefits of both methods. Moreover, investigators could estimate chick survival (in addition to brood survival) if genetic analysis is used to identify individual chicks via the fecal pellets.

All brood survey methods have imperfect detection, and detection rates of each method may differ based on many extrinsic factors, including observer, vegetation, temperature, wind, etc. Temporal and spatial variation in detection probability causes bias when making comparisons in survival among groups when those groups vary in any of those environmental conditions. We recommend that future studies document the effects of all methods on survival (we could not address this issue because we conducted all methods on the same broods), how far the hen and brood flush, and how quickly they reunite during both daytime flushes and nighttime flushes. Nighttime spotlight surveys and fecal surveys are particularly valuable because of their high rates of brood detection and they are less invasive compared to traditional brood sampling methods like daytime flush or visual surveys.

Literature Cited

Aldridge, C. L. and R. M. Brigham. 2001. Nesting and reproductive activities of greater sage-grouse in a declining northern fringe population. Condor 103: 537-543.

Aldridge, C. L. and M. S. Boyce. 2007. Linking occurrence and fitness to persistence: habitat-based approach for endangered greater sage-grouse. Ecological Applications. 508-526.

Amundson, C. L., and T. W. Arnold. 2010. Effects of radiotransmitters and plasticine bands on mallard duckling survival. Journal of Field Ornithology 81: 310-316.

Andersen, P. K. and R. D. Gill 1982. Cox's regression model for counting processes: a large sample study. Annals of Statistics 10: 1100-1120.

Anderson, D. R., K. P. Burnham, and G. C. White. 1994. AIC model selection in over-dispersed capture-recapture data. Ecology 75: 1780-1793.

Arnold, T. W. 2010. Uninformative parameters and model selection using Akaike's information criterion. Journal of Wildlife Management 74: 1175-1178.

Atamian, M. T., J. S. Sedinger, J. S. Heaton, and E. J. Blomberg. 2010. Landscape-level assessment of brood rearing habitat for sage-grouse in Nevada. Journal of Wildlife Management 74: 1533-1543.

Aulie, A. 1976. The pectoral muscles and the development of thermoregulation in chicks of willow ptarmigan (*Lagopus lagopus*). Comparative Biochemistry and Physiology 53: 343-346.

Baxter, J. J., J. P. Hennefer, R. J. Baxter, R. T. Larsen, and J. R. Flinders. 2013. Evaluating survival of greater sage-grouse chicks in Strawberry Valley, Utah, by use of microtransmiters: does handling time negatively influence survival rates? Western North American Naturalist 73: 419-425.

Blomberg, E. J., J. S. Sedinger, M. T. Atamian, and D.V. Nonne. 2012. Characteristics of climate and landscape disturbance influence the dynamics of greater sage-grouse populations. Ecosphere <http://doi: 10.1890/ES11-00304.1>. Assessed 20 Jun 2018.

Blomberg, E J., J. S. Sedinger, D.V. Nonne, and M.T. Atamian. 2013. Seasonal Reproductive Costs Contribute to Reduced Survival of Female Greater Sage-Grouse. Journal of Avian Biology 44: 149-158.

Bowker, B. 2008. Ecological Models and Data in R. Princeton University Press. Princeton, NJ, USA.

Braun, C.E. and M.A. Schroeder. Age and sex identification from wings of sage grouse. 2015. Wildlife Society Bulletin 39: 182-187.

Byrne, M.W. 2002. Habitat use by female greater sage-grouse in relation to fire at Hart Mountain National Antelope Refuge, Oregon. Thesis. Oregon State University. Corvalis, OR, USA

Buckland, S. T., E. A. Rexstad, T. A. Marques, and C. S. Oedekoven. 2015. Distance sampling: methods and applications. Methods in Statistical Ecology. Springer International Publishing. Switzerland.

Bui, T., B. Bedrosian, and J. Marzluff. 2010. Common raven activity in relation to land use in western Wyoming: implications for greater sage-grouse reproductive success. Condor 112: 65-78.

Burkepile, N. A, J. W. Connelly, D. W. Stanley, and K. P. Reese. 2002. Attachment of radiotransmitters to one-day-old sage grouse chicks. Wildlife Society Bulletin 30: 93-96.

Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. Design and analysis methods for fish survival experiments based on release-recapture. American Fisheries Society Monograph. No. 5. Bethesda, MD, USA.

Burnham, K. P. and D. R. Anderson. 2002. Second Edition. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag, New York, NY. USA.

Caizergues, A. and L. N Ellison. 2000. Age-specific reproductive performance of black grouse *Tetrao tetrix* females. Bird Study 47: 344-351.

Casazza, M. L., P. S. Coates, and C. T. Overton. 2011. Linking habitat selection and brood success in greater sage-grouse. Pages 151–167 in B. K. Sandercock, K. Martin, and G. Segelbacher, editors. Ecology, Conservation, and Management of Grouse. Studies in Avian Biology. 39 University of California Press, Berkeley, CA, USA.

Caudill, D., M.R. Guttery, B. Bibles, T.A. Messmer, G. Caudill, E. Leone, D.K. Dahlgren, and R. Chi. 2014. Effects of climatic variation and reproductive trade-offs vary by measure of reproductive effort in greater sage-grouse. Ecosphere 5: 154; http://dx.doi.org/10.1890/ES14-00124.1

Chi, R. Y. 2004. Greater sage-grouse reproductive ecology and Tebuthiuron manipulation of dense big sagebrush on Parker Mountain. Thesis. Utah State University. Logan, UT, USA

Collier, B. A., S. S. Ditchkoff, C. R. Ruth Jr., and J. B. Raglin. 2013. Spotlight surveys for white-tailed deer: monitoring panacea or exercise in futility? Journal of Wildlife Management 77: 165-171.

Conway, C. J., A. Meyers, K. Launchbaugh, D. Musil, P. Makela, and S. Roberts. 2017. The Grouse & Grazing Project: Effects of cattle grazing on sage-grouse demographic traits – 2017 Annual Report. College of Natural Resources, University of Idaho. <https://idahogrousegrazing.files.wordpress.com/2018/11/2017_gg_annualreport_final.pdf. > Accessed 4 Feb 2018.

Cook, A. Greater sage-grouse seasonal habitat models, response to juniper reduction and effects of capture behavior on vital rates, in northwest Utah. 2015. Thesis. Utah State University. Logan, UT, USA

Cormack, R. M. 1964. Estimates of survival from the sighting of marked animals. Biometrika 51: 429-438.

Dahlgren, D. K., T. A. Messmer, E. T. Thacker, and M. R. Guttery. 2010a. Evaluation of brood detection techniques: recommendations for estimating greater sage-grouse productivity. Western North American Naturalist 70: 233-237.

Dahlgren, D. K., T. A. Messmer, and D. N. Koons. 2010b. Achieving better estimates of greater sage-grouse chick survival in Utah. Journal of Wildlife Management 74: 1286-1294.

Dahlgren, D. K., R. D. Elmore, D. A. Smith, A. Hurt, E. B. Arnett, and J. W. Connelly. 2012. Use of dogs in wildlife research and management. Pages 140-153 in N.J. Silvy, editor. The Wildlife Techniques Manual: Research. Vol. 1. John Hopkins University Press. Baltimore, MD, USA.

Davis, D. M. Breeding season habitat use and response to management activities by greater sage-grouse on Sheldon National Wildlife Refuge, Nevada. Thesis. Oregon State University. Corvallis, OR, USA.

Davis, D. M., K. P. Reese, and S. Gardner. 2014. Demography, reproductive ecology, and variation in survival of greater sage-grouse in northeastern California. Journal of Wildlife Management 78: 1343-1355.

Dinkins, J. B., M. R. Conover, C. P. Kirol, and J. L. Beck. Greater sage-grouse (*Centrocercus urophasianus*) select nest sites and brood sites away from avian predators. Auk 129: 600-610.

Dinsmore, S. J., G. C. White, and F. L. Knopf. 2002. Advanced techniques for modeling avian nest survival. Ecology 83: 3476-3488.

Drut, M. S. Habitat use and selection by sage-grouse broods in southeastern Oregon. 1992. Thesis. Oregon State University. Corvallis, OR. USA.

Dunbar, M. R., M. A. Gregg, J. A. Crawford, M. R. Giordano, and S. J. Tornquist. 2005. Normal hematologic and biochemical values for prelaying greater sage-grouse (*Centrocercus urophasianus*) and their influence on chick survival. Journal of Zoo and Wildlife Medicine 36: 422-429.

Duvuvuei, O. V., N. W. Gruber-Hadden, T. A. Messmer, M. R. Guttery, B. D. Maxfield. 2017. Contribution of translocated greater sage-grouse to population vital rates. The Journal of Wildlife Management 81: 1033-1041.

Dzialak, M. R., C. V. Olson, S. M. Harju, S. L. Webb, J. P. Mudd, J. B. Winstead, and L. D. Hayden-Wing. 2011. Identifying and prioritizing greater sage-grouse nesting and brood-rearing habitat for conservation in human-modified landscapes. PLoS ONE. 6: e26273. <doi:10.1371/journal.pone.0026273>. Assessed 20 Jun 2018.

Fields, T. L., G. C. White, W. C. Gilgert, and R. D. Rodgers. 2006. Nest and brood survival of lesser prairie-chickens in west central Kansas. Journal of Wildlife Management 70: 931-938.

Fletcher, Jr., R. J., and R. L. Hutto. 2006. Estimating detection probabilities of river birds using double surveys. Auk 123: 695-707.

Fryer, J. L. 2009. *Artemisia nova*. Fire effects information system. U.S. Department of Agriculture. Forest Service. Rocky Mountain Research Station, Fire Sciences Laboratory. <https://www.fs.fed.us/database/feis/plants/shrub/artnov/all.html>. Accessed 26 Nov 2018.

Gibson, G., E. J. Blomberg, M. T. Atamian, and J. S. Sedinger. 2017. Weather, habitat composition, and female behavior interact to modify offspring survival in greater sage-grouse. Ecological Applications 27: 168-181.

Giesen, K. M., T. J. Schoenberg, and C. E. Braun. 1982. Methods for trapping sage grouse in Colorado. Wildlife Society Bulletin 10: 224-231.

Gimenez, O. and J. Gaillard. 2017. Estimating individual fitness in the wild using capture–recapture data. Population Ecology 60: 101-109.

Girard, G. L. 1937. Life history, habits and food of the sage grouse, *Centrocercus urophasianus* Bonaparte. University of Wyoming Publications, Publication 3. University of Wyoming, Laramie, WY, USA.

Goddard, A. D. and R. D. Dawson. 2009. Factors influencing the survival of neonate sharp-tailed grouse *Tympanuchus phasianellus*. Wildlife Biology 15: 60-67.

Godfrey, G. A. 1975. Underestimation experienced in determining ruffed grouse brood size. Journal of Wildlife Management 39: 191-193.

Gregg, M. A., M. R. Dunbar, and J. A. Crawford. 2007. Use of implanted radio-signal transmitters to estimate survival of greater sage-grouse chicks. Journal of Wildlife Management 71: 646-651.

Gregg, M. A. and J. A. Crawford. 2009. Survival of greater sage-grouse chicks and broods in the Northern Great Basin. Journal of Wildlife Management 73: 904-913.

Gruber, N. W. 2012. Population dynamics and movements of translocated and resident sage-grouse on Anthro Mountain, Utah. Thesis. Utah State University. Logan, UT. USA.

Gu, W. and R. K. Swihart. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. Biological Conservation 116: 195-203.

Guttery, M. R., D. K. Dahlgren, T. A. Messmer, J. W. Connelly, K. P. Reese, P. A. Terletzky, N. Burkepile, and D. N. Koons. 2013. Effects of landscape-scale environmental variation on greater sage-grouse chick survival. PLoS ONE. 8: e65582. <http://doi:10.1371/journal.pone.0065582>. Assessed 20 Jun 2018

Hagen, C. A., G. C. Salter, J. C. Pitman, R. J. Robel, and R. D. Applegate. 2005. Lesser prairie-chicken brood habitat in sand sagebrush: invertebrate biomass and vegetation. Wildlife Society Bulletin 33: 1080-1091.

Hannon, S. J., and K. Martin. 2006. Ecology of juvenile grouse during the transition to adulthood. Journal of Zoology 269: 422-433.

Harju, S. M., C. V. Olson, L. Foy-Martin, S. L. Webb, M. R. Dzialak, J. B. Winstead, and L.D Hayden-Wing. 2013. Occurrence and success of greater sage-grouse broods in relation to insect-vegetation community gradients. Human-Wildlife Interactions 7: 214-229.

Herman-Brunson, K. E. 2007. Nesting and brood-Rearing habitat selection of greater sage-grouse and associated survival of hens and broods at the edge of their historic distribution. Thesis. South Dakota State University. Brookings, SD, USA.

Hollaran, M.J. 2005. Greater sage-grouse (*Centrocercus urophasianus*) population response to natural gas field development in western Wyoming. Dissertation. University of Wyoming. Laramie, WY, USA

Hovick, T. J., R. D. Elmore, D. K. Dahlgren, S. D. Fuhlendorf, and D. M. Engle. 2014. Evidence of negative effects of anthropogenic structures on wildlife: a review of grouse survival and behaviour. Journal of Applied Ecology 51: 1680-1689.

Huggins, R. M. 1989. On the statistical analysis of capture experiments. Biometrika 76: 133-140.

Huggins, R. M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. Biometrics 47: 725-732.

Hutto, R. L. 2016. Should scientists be required to use a model-based solution to adjust for possible distance-based detectability bias? Ecological Applications 26: 1287-1294.

Huwer, S. L., D. R. Anderson, T. E. Remington, and G. C. White. 2008. Using human-imprinted chicks to evaluate the importance of forbs to sage-grouse. Journal of Wildlife Management. 72: 1622-1627.

Idaho Sage-grouse Advisory Committee. 2006. Conservation plan for the greater sage-grouse in Idaho.

Johnson, G. D., and M. S. Boyce. 1991. Survival, growth, and reproduction of captive-reared sage grouse. Wildlife Society Bulletin 19: 88-93.

Jolly, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52: 225-247.

Kaczor, N. W., K. M. Herman-Brunson, K. C. Jensen, M. A. Rumble, R. W. Klaver, and C. C. Swanson. 2011. Resource selection during brood-rearing by greater sage-grouse. Pp. 169–177 in B. K. Sandercock, K. Martin, and G. Segelbacher (eds). Ecology, conservation, and management of grouse. Studies in Avian Biology (no. 39), University of California Press, Berkeley, CA, USA.

Kellner, K. F., and R. K. Swihart. 2014. Accounting for imperfect detection in ecology: a quantitative review. PLoS ONE 9(10): e111436. <http://doi:10.1371/journal.pone.0111436>. Assessed 5 Nov 2018.

Kirkpatrick, S., C. Gelatt, and M. Vecchi. 1983. Optimization by simulated annealing. Science 220: 671-680.

Kirol, C. P., J. L. Beck, S. V. Huzurbazar, M. J. Hollaran, and S. N. Miller. 2015. Identifying greater sage-grouse source and sink habitats for conservation planning in an energy development landscape. Ecological Applications 25: 968-990.

Knerr, J. S. Greater sage-grouse ecology in western Box Elder County, Utah. 2007. Thesis. Utah State University. Logan, UT. USA.

Kubisiak, J. F. 1978. Brood characteristics and summer habitats of ruffed grouse in central Wisconsin. Wisconsin Department of Natural Resources. Technical Bulletin 108.

Laake, J. L., and E. Rexstad. 2012. RMark: R Code for MARK Analysis. R Package Version 3.1.1.

Larson, M. A., M. E. Clark, and S. R. Winterstein. 2001. Survival of ruffed grouse chicks in northern Michigan. Journal of Wildlife Management 65: 880-886.

LeBeau, C.W., J. L. Beck, G. D. Johnson, and M. J. Holloran. 2014. Short-term impacts of wind energy development on greater sage-grouse fitness. Journal of Wildlife Management 78: 522-530.

LeBeau, C. W., G. D. Johnson, M. J. Holloran, J. L. Beck, R. M. Nielson, M. E. Kauffman, E. J. Rodemaker, and T. L. McDonald. 2017.Greater sage-grouse habitat selection and survival, and wind energy infrastructure. Journal of Wildlife Management 81: 690-711.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62: 67-118.

Leopold, A. 1933. Game Management. University of Wisconsin Press, Madison, WI, USA.

Lukacs, P. M., V. J. Dreitz, F. L. Knopf, and K. P. Burnham. 2004. Estimating survival probabilities of unmarked dependent young when detection is imperfect. Condor 106: 927-932.

Ludwig, G. X., R. V. Alatalo, P. Helle, H. Lindén, J. Lindström, and H. Siitari. 2006. Short- and long-term population dynamical consequences of asymmetric climate change in black grouse. Proceedings of the Royal Society B: Biological Sciences 273: 2009-2016.

Ludwig, G. X., R. V. Alatalo, P. Helle, and H. Siitari. 2010. Individual and environmental determinants of early brood survival in black grouse *Tetrao tetrix*. Wildlife Biology 16: 367-378.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. L. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248-2255.

Marsh, H. and D. F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management 53: 1017-1024.

Mabray, S. T. Microhabitat selection by greater sage-grouse hens in southern Wyoming. Thesis. Utah State University. Logan, UT. USA.

McNew, L. B., A. J. Gregory, S. M. Wisely, and B. K. Sandercock. 2011. Reproductive Biology of a Southern Population of Greater Prairie-Chickens. Pages 209-221 in B. K. Sandercock, K. Martin, and G. Segelbacher, editors. Ecology, conservation, and management of grouse. Studies in Avian Biology. 39. University of California Press, Berkeley, CA, USA.

Moss, R., A. Watson, P. Rothery, and W. W. Glennie. 1980. Clutch size, egg size, hatch weight, and laying date in relation to early mortality in red grouse *Lagopus Lagopus Scoticus* chicks. Ibis 123: 450-462.

Moynahan, B.J. 2004. Landscape-scale factors affecting population dynamics of greater sage-grouse (*Centrocercus urophasianus*) in north-central Montana 2001-2004. Dissertation. University of Montana. Missoula, MT. USA.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117: 393-408.

Orange, J. P., J. M. Yeiser, D. L. Baxley, J. J. Morgan, and B. A. Robinson. 2017. Evaluating hunting success of pen-reared and wild Northern bobwhite in a reclaimed Kentucky mineland. National Quail Symposium Proceedings: Vol. 8, Article 75.

Orning, E. K. 2014. Effect of predator removal on greater sage-grouse (*Centrocercus urophasianus*) ecology in the Bighorn Basin Conservation Area of Wyoming. Thesis. Utah State University. Logan, UT. USA.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62.

Popper, K. J., E. C. Pelren, and J. A. Crawford. 1996. Summer nocturnal roost sites of blue grouse in northeastern Oregon. Great Basin Naturalist 56: 177-179.

Rebholz, J. L. 2007. Influence of habitat characteristics on greater sage-grouse reproductive success in the Montana Mountains, Nevada. Thesis. Oregon State University, Corvallis, OR, USA.

Robinson, J. D. and T. A. Messmer. 2013. Vital rates and seasonal movements of two isolated greater sage-grouse populations in Utah's west desert. Human-Wildlife Interaction 7: 182-194.

Sandercock, B. K. 1994. The effect of manipulated brood size on parental defence in a precocial bird, the Willow Ptarmigan. Journal of Avian Biology 25: 281-286.

Sandford, C. P., M. T. Kohl, T. A. Messmer, D. K. Dahlgren, A. Cook, and B. R. Wing. 2017. Greater sage-grouse resource selection drive reproductive fitness under a conifer removal strategy. Rangeland Ecology and Management 70: 59-67.

Schaub, M., and J. A. Royle. 2014. Estimating true instead of apparent survival using spatial Cormack-Jolly-Seber models. Methods in Ecology and Evolution 5: 1316-1326.

Schmidt, J. H., C. L. McIntyre, and M. C. MacCluskie. 2013. Accounting for incomplete detection: what are we estimating and how might it affect long-term passerine monitoring programs? Biological Conservation 160: 130-139.

Schreiber, L. A., C. P. Hansen, M. A. Rumble, J. J. Millspaugh, F. R. Thompson III, R. S. Gamo, J. W. Kehmeier, and N. Wojik. 2016. Greater sage-grouse apparent nest productivity and chick survival in Carbon County, Wyoming. Wildlife Biology 22: 37-44.

Schroeder, M. A. 1997. Unusually high reproductive effort by sage grouse in a fragmented habitat in north-central Washington. Condor 99: 933-941.

Seber, G. A. F. 1965. A note on the multiple-recapture census. Biometrika 52: 249-259.

Seber, G. A. F. 1982. Second edition. The estimation of animal abundance and related parameters. Macmillan Book, New York, NY, USA.

Sika, J. 2006. Breeding ecology, survival rates, and causes of mortality of hunted and nonhunted greater sage-grouse in central Montana. Montana State University. Bozeman, MT. USA.

Smith, L. S. 2009. Greater sage-grouse and energy development in northeastern Utah: implications for management. Thesis. Utah State University. Logan, UT. USA.

Sveum, C. M. Habitat selection by sage grouse hens during the breeding season in south-central Washington. Thesis. Oregon State University. Corvallis, OR. USA.

Tack, J. D. 2009. Sage-grouse and the human footprint: implications for conservation of small and declining populations. Thesis. University of Montana. Missoula, MT. USA.

Taylor, R. L., B. L. Walker, D. E. Naugle, and L. S. Mills. 2012. Managing multiple vital rates to maximize greater sage-grouse population growth. Journal of Wildlife Management 76: 336-347.

Thompson, T. R. 2012. Dispersal ecology of greater sage-grouse in northwestern Colorado: evidence from demographic and genetic methods. Dissertation. University of Moscow, ID. USA.

Tirpak, J. M., W. M. Guiliano, and C. A. Miller. 2005. Nocturnal roost habitat selection by ruffed grouse broods. Journal of Field Ornithology 76: 168-174.

Tirpak, J. M., W. M. Guiliano, and C. A. Miller. 2008. Ruffed grouse brood habitat selection at multiple scales in Pennsylvania: implications for survival. Canadian Journal of Zoology 86: 329-337.

Storch, I. 2007. Conservation status of grouse worldwide: an update. Wildlife Biology 13: 5-12.

Walker, B. L. 2008. Greater sage-grouse response to coal-bed natural gas development and West Nile virus in the Powder River Basin, Montana and Wyoming, USA. Dissertation. University of Montana. Missoula, MT. USA.

Wakkinen, W. L., K. P. Reese, J. W. Connelly, and R. A. Fischer. 1992. An improved spotlighting technique for capturing sage grouse. Wildlife Society Bulletin 20: 425-426.

Williams, B. K., J. D. Nichols, and M. J. Conroy, 2002. Analysis and management of animal populations. Academic Press, San Diego, CA, USA.

Wing, L., J. Beer, and W. Tidyman. 1944. Brood habits and growth of 'blue grouse'. Auk 61: 426-440.

Wing, B. R. 2014. The role of vegetation structure, composition, and nutrition in greater sage-grouse ecology in northwestern Utah. Thesis. Utah State University. Logan, UT. USA.

Wisdom, M. J. and L. S. Mills. 1997. Sensitivity analysis to guide population recovery: prairie-chickens as an example. Journal of Wildlife Management 61: 302-312.

White, G. C. and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study 46: 120-139.

Wong, M. M. L., B. C. Fedy, S. Wilson, and K. M. Martin. 2009. Adoption in rock and white-tailed ptarmigan. Wilson Journal of Ornithology 121: 638-664.

Zwickel, F. C. and J. F. Bendell. 1966. Early mortality and the regulation of numbers in blue grouse. Canadian Journal of Zoology 45: 817-851.

Table 2.1. Summary of 50 greater sage-grouse (*Centrocercus urophasianus*) papers from 1998-2017 that used brood surveys methods to estimate productivity or habitat use, including: state, years of study, method used, number of chicks or broods studied, final date to assess survival (survival date), how the researchers appraised brood fate, which vital rates were calculated, which survival analysis was used, and chick or brood survival estimates and variances. Papers that accounted for or estimated chick or brood detection probabilities are in bold.

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
Aldridge & Brigham 2001	AB	1998-1999	DF		22	1	50		C1	AP	0.14 - 0.23	
Aldridge & Boyce 2007		2001-2004	R	41	35	23	56		C2, HS	COX	0.12	
Atamian et al. 2010	NV	2003-2006	DF		29	7	12	HD, RD	B1, HS	RAW		1.00 [29/29][1]
		2003-2006					45			RAW		0.83 [24/29][1]
Baxter et al. 2013	UT	2005-2006	R	40	19	14	49		C2	KF	0.25 (0.10)[f]	
Blomberg et al. 2013	NV	2003-2011	DF		272	6	45		B1	RAW		0.52 [50/96][1]
Bryne 2002	OR	1998-2000	UNK		58	UNK	1 Aug		B1, HS	NA		UNK
Bui et al. 2010	WY	2007-2008	DF		UNK	2	35		B1, HS	NA		UNK
Burkepile et al. 2002	ID	1999	R	75	28	70	70		C1	AP (1g)	0.32	
		2000								AP (1g)	0.21	
		2000								AP (1.4g)	0.50	
Casazza et al. 2011	CA	2003-2005	DV		38	16-50	50	RD	B1, HS	NA		UNK
Caudill et al. 2014	UT	1998-2010	DV, R	UNK	142	21-25	42-50		B1, HF	NA		UNK

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
Chi 2004	UT	2000	DV		30	13	40		B1	AP		0.73
		2001										0.88
		2002										0.55
Cook 2015	UT	2012-2013	DV		43	2-14	50		B2	NS (TRT)		0.74 (0.53-0.96)[h]
	UT	2012-2013								NS (CO)		0.42 (0.23-0.61)[h]
	UT (WB)	2012-2013								NS (TRT)		0.81 (0.63-0.99)[h]
	UT (WB)	2012-2013								NS (CO)		0.61 (0.37-0.86)[h]
	UT (RC)	2012-2013								NS (TRT)		0.24 (0.02-0.45)[h]
	UT (RC)	2012-2013								NS (CO)		0.54 (0.16-0.91)[h]
Dahlgren et al. 2010a	UT	2006-2007	DF, DOG, R	25	21	1	56		C1	ANOVA		
Dahlgren et al. 2010b	UT	2005-2006	R	150	42	21-42	42		C2	NS	0.60 (0.51-0.72)[h]	
Davis 2002	NV, OR	1998-2000	DV		14	8-38	1 Aug		B1, HS	AP		0.68
		1998										0.50
		1999										0.88
		2000										0.62
Davis et al. 2014	CA	2007-2009	DF, DV		25	4	7	HD, FL	B1	AP		0.92
		2007-2009					14					0.80

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
		2007-2009					30					0.60
		2007-2009					60					0.44
		2007					60					0.33
		2008										0.71
		2009										0.33
Dinkin et al. 2012	WY	2008-2010	DV		83	3	21	HD	B1, HS	NA		UNK
Drut et al. 1992	OR (JA)	1989-1991	UNK		18	≤12	42		B1, HS	AP		0.43
	OR (HM)	1989-1991										0.36
	OR (JA)	1989-1991					84					0.43
	OR (HM)	1989-1991										0.36
Dunbar et al. 2005	NV, OR	1999-2001	DF		UNK	1	15 Jul-1 Aug		B1, HF	NA		UNK
Duvuvuei et al. 2017	UT	2009-2012	DF, DV, R, NS	UNK	47	2	50	HD, NSD	B2	NS		0.55 (0.41-0.69)[h]
Gibson et al. 2017	NV	2005-2012	DF, NS		120	6	42	HD, RD, NSD	C2, HF, HS	LYS	0.26 (0.02)[f]	
		2005									0.48 (0.37-0.62)[2h]	
		2006									0.22 (0.18-0.30)[2h]	

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
		2007									0.09 (0.04-0.26)[2h]	
		2008									0.28 (0.14-0.42)[2h]	
		2009									0.16 (0.10-0.22)[2h]	
		2010									0.30 (0.22-0.38)[2h]	
		2011									0.34 (0.26-0.42)[2h]	
		2012									0.18 (0.12-0.24)[2h]	
Gregg et al. 2007	NV, OR	2001-2002	R	288	52	28	28		C2	KM	0.22 (0.03)[f]	
Gregg et al. 2009	NV, OR	2000-2003	DOG, R	506	94	28	28		C2, B2	KM	0.39 (0.02)[f]	0.67 (0.06)[f]
		2000										0.89 (0.11)[f]
		2001										0.52 (0.10)[f]
		2002										0.44 (0.10)[f]
		2003										0.92 (0.06)[f]
Gruber 2012	UT	2009	R	99	24	21	20		C2	MS (RH)	0.14 (0.08-0.30)[h]	
		2009								MS (TH)	0.08 (0.04-0.15)[h]	

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
		2009					50		C2	MS (RH)	0.08 (0.04-0.15)[h]	
		2009								MS (TH)	(0.002, 0.00-0.04)[h]	
		2010					20		C2	MS (RH)	0.17 (0.10-0.34)[h]	
		2010								MS (TH)	0.11 (0.05-0.28)[h]	
		2010					50		C2	MS (RH)	0.16 (0.08 -0.32)[h]	
		2010								MS (TH)	0.08 (0.02-0.20)[h]	
Guttery et al. 2013	ID, UT	1999-2009	R	518	142	26-42	42		C2	MS	0.48 (0.38-0.57)[h]	
Harju et al. 2013	WY	2008-NA	DV		11	5	35	RD	B1, HS	RAW		0.45 [5/11][1]
Herman-Brunson 2007	ND	2005	DF, R	UNK	13	6-9	21		B1, HS	AP	0.34	
		2006									0.42	
		2005				52-78	35-42-1 Jan		C2	KM	0.50 (0.23-0.58)[h]	
		2006									0.32 (0.14-0.49)[h]	
		2005				UNK	1 Aug		B1	RAW		0.86 [6/7][1]
		2006										0.50 [3/6][1]

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
Hollaran 2005	WY	1998-2004	DF, DV		123	UNK	1-15 Aug	HD, RD	C2	KF	0.32 (0.17)[f]	
Huwer et al. 2008**	CO	2002	DV	UNK	UNK	648	27		C2	KF	0.35 (0.17-0.59)[h]	
		2003									0.78 (0.63-0.88)[h]	
Kaczor et al. 2011	ND, SD	2005-2006	DF, DV		43	3-5	35	RD	B1, HS	NA		UNK
Kirol et al. 2015	WY	2008-2009	DV, NS		35	6	36	HD, RD, NSD	B2, HS	KM		0.76 (0.08)[f]
Knerr et al. 2007	UT	2005-2006	DF, DV		9	21-22	50		C1, B1, HS	AP	0.24	0.44
		2005					50		C1	AP	0.56	
		2006									0.16	
		2005					50		B1	AP		1.00
		2006										0.29
LeBeau et al. 2014	WY	2009	DV, NS		31	5	14	HD, RD, NSD	B2	AG		0.95 (0.78-1.00)[g]
		2010										0.73 (0.44-0.92)[g]
		2009					35-37		B2	AG		0.61 (0.39-0.80)[g]
		2010										0.80 (0.49-0.96)[g]
LeBeau et al. 2017	WY	2011	DV, NS		123	5	35	HD, RD, NSD	B2	AG (CO)		0.80 (0.62-1.00)[2g]
		2012										0.92 (0.78-1.00)[2g]

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
		2013										0.68 (0.50-0.90)[2g]
		2014										1.00[2g]
		2011					35		B2	AG (TRT)		0.66 (0.45-0.92)[2g]
		2012										0.64 (0.40-1.00)[2g]
		2013										0.64 (0.45-0.92)[2g]
Mabray 2015	WY	2008-2011	DV		8	UNK	UNK	HD	B1, HS	NA		UNK
Moynahan 2004	MT	2001	DF, DV		115	7-8	30	HD, DI, RD, FL	B2	NS		0.21 (0.12)[f]
		2002										0.69 (0.07)[f]
		2003										0.76 (0.06)[f]
Orning 2014	WY	2011-2012	UNK, NS		8	UNK	35	NSD	C1	AP (CO)	0.4	
		2011-2012								AP (TRT)	0.44	
Rebholz 2007	NV	2004-2005	R	115	21	18	18		C2	KM	0.44 (0.05)[f]	
		2004-2005							B1	AP		0.71
Robinson et al. 2013	UT (SR)	2005	DV		UNK	25	50		B1	AP		0.29
	UT (DC)	2005										0.50
	UT (SR)	2005										0.30
	UT (DC)	2006										0.67

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
Sandford et al. 2017	UT	2012-2015	DF, DV, NS		56	14-21	50	RD, FL, NSV	B1, HS	RAW		0.86 [43/50][1]
Schrieber et al. 2016	WY	2011	NS		37	3-7	70	RD, NSD	C2	LE	0.19 (0.06-0.37)[h]	
		2012									0.04 (0.01-0.12)[h]	
		2013									0.16 (0.08-0.27)[h]	
Schroeder 1997	WA	1992-1996	UNK		99	1	50		C1, B1	AP	0.33	0.50
Sika et al. 2006	MT	2003-2005	DF		73	2-3	30	FL	B2	NS		0.79 (0.002)[f]
Smith 2012	UT	2007	UNK		2	14-21	50		C1, HS	AP	0.17	
		2008									0.00	
Sveum 1996	WA	1992	DV		38	<16	1 Aug		B1, HS	AP		0.10
		1993										0.50
Tack 2009	MT, SK	2007	DF, DV		39	10-16	50		C1	AP	0.33	
		2008									0.38	
Thompson et al. 2006	WY	1999-2003	DV		82	1-2	14	HD, RD	B1, HS	NA		UNK
Walker et al. 2008	WY, MT (DE)	2003	DF, DV, NS		246	7-11	35	HD, DI, FL, NSV	B1, C2	AP	0.43	0.67
	WY, MT (DE)	2004									0.38	0.91
	WY, MT (DE)	2005									0.50	0.80
	WY, MT (DE)	2006									0.40	0.76

Paper	State/ Province[a]	Years	Method[b]	Chicks	Broods	Survey number	Survival date	Fate appraisal[c]	Vital Rate[d]	Survival Analysis[e]	Chick survival (var)[f-h]	Brood survival (var)[f-h]
	WY (BUF)	2004									0.33	0.71
	WY (BUF)	2005									0.55	0.96
	WY (BUF)	2006									0.48	0.91
	WY (SH)	2003									0.54	0.83
	WY, MT (DE)	2003					35		B2	GLM		0.63
	WY, MT (DE)	2004										0.84
	WY, MT (DE)	2005										0.67
	WY, MT (DE)	2006										0.52
	WY (BUF)	2004										0.68
	WY (BUF)	2005										0.92
	WY (BUF)	2006										0.93
	WY (SH)	2003										0.83
Wing et al. 2014	UT	2012-2013	DF, NS		28	7	50	RD, NSV	B1	AP		0.71

[a] Letters in parentheses indicate different study sites: BUF = Buffalo, DC = Deep Creek, DE = Decker, HM = Hart Mtn., JA = Jackass Creek, RC = Rich County, SH= Spotted Horse, SR = Sheep Rock, WB = West Box Elder

[b] DF = daytime flush, DV = daytime visual, R = radio-marked chicks, NS = nighttime spotlight, UNK = unknown

[c] DI = distance between consecutive surveys, HD = hen distractive behavior, NSD = nighttime spotlight survey used to determine fate, NSV = nighttime survey used to validate other survey's fate determination, RD = repeated detection/non-detection

[d] B1 = brood success, B2 = brood survival, C1 = chick success, C2 = chick survival, HF = hen's fitness or productivity, HS = habitat selection

[e] AP = apparent, AG = Anderson-Gill, ANOVA = analysis of variance, COX = Cox proportional, GLM = generalized linear model, KF = Known-fate, KM = Kaplan-Meier, LE = logistic exposure, LYS = Lukacs young-survival, NA = brood or chick survival or success not estimated, NS = nest survival, RAW = calculated apparent values of brood or chick success

[f] Estimated standard error variance

[g] Estimated 90 % confidence interval

[h] Estimated 95 % confidence internal

[1] Papers didn't calculate chick or brood success, but there was enough information for (IR) to do so.

[2] Survival or success rates are approximate due to values being depicted on graphs.

[**] The author used human-imprinted chicks to estimate diet selection and chick survival. Authors surveyed chicks 24 times per day thus the extraordinary survey number

Table 2.2. Sensitivity of Cormack-Jolly-Seber joint survival (φ) and detection probability (p) model rankings when adjusted for overdispersion ($\hat{c}$) to qualitatively assess model goodness-of-fit based on brood fecal surveys. We collected fecal survey data from 45 sage-grouse hens and their broods in 2016-2017. Models are ranked by quasi-likelihood Akaike's information criterion corrected for sample size ($\Delta QAIC_c$) and normalized Akaike model weights (ω_i). We show the top-ranked models with $\Delta QAIC_c \le 2$. Explanatory variables include year (y), hen age (h), Julian hatch date (j), linear (A) or non-linear effects {log(A + 0.01)} of brood age, or no effect (null).

$\hat{c}$	φ	p	K	$QAIC_c$	$\Delta QAIC_c$	ω_i
1	y + j + h + log(A + 0.01)	null	6	244.14	0.00	0.13
	y + h + log(A + 0.01)	null	5	244.41	0.27	0.11
	y + log(A + 0.01)	null	4	245.10	0.96	0.08
	y + j + log(A + 0.01)	null	5	245.89	1.75	0.05
	y + j + h + log(A + 0.01)	log(A + 0.01)	7	246.11	1.97	0.05
1.25	y + h + log(A + 0.01)	null	5	197.61	0.00	0.09
	y + log(A + 0.01)	null	4	197.73	0.12	0.09
	y + j + h + log(A + 0.01)	null	6	197.83	0.22	0.08
	y + j + log(A + 0.01)	null	5	198.79	1.18	0.05
	y + j + h	null	5	199.44	1.83	0.04
	y + h + log(A + 0.01)	log(A + 0.01)	6	199.55	1.94	0.04
1.5	y + log(A + 0.01)	null	4	166.15	0.00	0.08
	y + h + log(A + 0.01)	null	5	166.41	0.25	0.07
	y + j + h + log(A + 0.01)	null	6	166.95	0.80	0.06
	y + j + log(A + 0.01)	null	5	167.39	1.24	0.05
	log(A + 0.01)	null	3	167.83	1.68	0.04
	y + j + h	null	5	167.93	1.78	0.03
	y + log(A + 0.01)	log(A + 0.01)	5	168.14	1.99	0.03
1.75	y + log(A + 0.01)	null	4	143.60	0.00	0.08
	y + h + log(A + 0.01)	null	5	144.12	0.52	0.06
	log(A + 0.01)	null	3	144.73	1.14	0.04
	y + j + h + log(A + 0.01)	null	6	144.90	1.30	0.04
	y + j + log(A + 0.01)	null	5	144.97	1.37	0.04
	j + log(A + 0.01)	null	4	145.33	1.73	0.03
	j + h	null	4	145.41	1.81	0.03
	y + j + h	null	5	145.43	1.83	0.03
2	y + log(A + 0.01)	null	4	126.68	0.00	0.07
	y + h + log(A + 0.01)	null	5	127.41	0.72	0.05
	log(A + 0.01)	null	3	127.41	0.73	0.05
	y + j + log(A + 0.01)	null	5	128.15	1.46	0.04
	j + log(A + 0.01)	null	4	128.19	1.51	0.03
	j + h	null	4	128.26	1.58	0.03
	y + j + h + log(A + 0.01)	null	6	128.36	1.67	0.03
	h + log(A + 0.01)	null	4	128.51	1.82	0.03
	y + j + h	null	5	128.55	1.87	0.03

Table 2.3. Standardized β coefficients, standard errors (SE), and 95% lower and upper confidence intervals (LCI, UCI) from top-ranked models ($\Delta QAIC_c \leq 2$) of factors that affect survival and detection probabilities of sage-grouse broods via fecal surveys in southern Idaho from 2016-2017. Explanatory variables include year (y), hen age (h), Julian hatch date (j), linear (A) or non-linear {log(A + 0.01)} effects of brood age, or no effect (null). Covariates that are in bold are considered meaningful (i.e., 95 % C.I. does not overlap zero).

Model	Parameter	β	SE	LCI	UCI
1	φ{y}	-2.19	1.64	-5.41	1.03
	φ{log(A + 0.01)}	**0.48**	**0.15**	**0.18**	**0.78**
	p{null}	**1.10**	**0.24**	**0.62**	**1.58**
2	φ{y}	-5.72	13.88	-32.91	21.48
	φ{h}	-2.52	3.04	-8.48	3.45
	φ{log(A + 0.01)}	**0.46**	**0.16**	**0.14**	**0.77**
	p{null}	**1.01**	**0.23**	**0.57**	**1.46**
3	φ{y}	-4.10	7.91	-19.61	11.40
	φ{j}	-0.04	0.03	-0.11	0.03
	φ{h}	-3.41	5.05	-13.31	6.49
	φ{log(A + 0.01)}	**0.42**	**0.18**	**0.06**	**0.77**
	p{null}	**1.01**	**0.23**	**0.56**	**1.45**
4	φ{y}	-2.15	2.21	-6.48	2.17
	φ{j}	-0.03	0.04	-0.11	0.04
	φ{log(A + 0.01)}	**0.46**	**0.16**	**0.14**	**0.78**
	p{null}	**1.09**	**0.26**	**0.57**	**1.60**
5	**φ{log(A + 0.01)}**	**0.43**	**0.14**	**0.15**	**0.72**
	p{null}	**1.16**	**0.26**	**0.65**	**1.68**
6	φ{y}	-1.84	1.64	-5.05	1.37
	φ{j}	-0.05	0.03	-0.12	0.01
	φ{h}	-3.13	2.95	-8.92	2.65
	p{null}	**0.99**	**0.24**	**0.52**	**1.45**
7	φ{y}	-2.26	1.68	-5.55	1.02
	φ{log(A + 0.01)}	**0.50**	**0.15**	**0.20**	**0.81**
	p{log(A + 0.01)}	-0.08	0.22	-0.51	0.34

Table 2.4. The qualitative assessment of goodness-of-fit for joint survival (φ) and detection probability (p) from Cormack-Jolly-Seber models based on sensitivity to adjustment for overdispersion ($\hat{c}$). Brood capture histories came from 45 radio-collared hens with broods in southern Idaho. Models are ranked by quasi-likelihood Akaike's information criterion corrected for sample size ($\Delta QAIC_c$) and normalized Akaike model weights (ω_i). Top-ranked models are shown ($\Delta QAIC_c \leq 2$). Explanatory variables include hen age (h), Julian hatch date (j), linear (A) or non-linear {log(A + 0.01)} effects of brood age, or no effect (null).

	φ	p	K	$QAIC_c$	$\Delta QAIC_c$	ω_i
1	log(A + 0.01)	A	4	212.12	0.00	0.09
	null	A	3	212.22	0.10	0.08
	j	A	4	212.67	0.55	0.07
	j + log(A + 0.01)	A	5	212.86	0.74	0.06
	log(A + 0.01)	null	3	213.31	1.19	0.05
	h	A	4	213.71	1.59	0.04
	h + log(A + 0.01)	A	5	213.79	1.67	0.04
	j + h	A	5	213.92	1.80	0.04
1.25	null	A	3	171.01	0.00	0.09
	log(A + 0.01)	A	4	171.36	0.35	0.08
	j	A	4	171.81	0.79	0.06
	log(A + 0.01)	null	3	171.89	0.87	0.06
	j + log(A + 0.01)	A	5	172.39	1.38	0.04
	h	A	4	172.63	1.62	0.04
	j + log(A + 0.01)	null	4	172.97	1.96	0.03
1.5	null	A	3	143.54	0.00	0.09
	log(A + 0.01)	A	4	144.19	0.65	0.07
	log(A + 0.01)	null	3	144.27	0.73	0.06
	j	A	4	144.56	1.02	0.05
	null	null	2	144.96	1.41	0.04
	h	A	4	145.25	1.70	0.04
	null	log(A + 0.01)	3	145.41	1.86	0.04
	j + log(A + 0.01)	A	5	145.41	1.87	0.04
	j + log(A + 0.01)	null	4	145.53	1.99	0.03
1.75	null	A	3	123.92	0.00	0.09
	log(A + 0.01)	null	3	124.55	0.62	0.07
	log(A + 0.01)	A	4	124.78	0.86	0.06
	null	null	2	124.83	0.91	0.06
	j	A	4	125.10	1.18	0.05
	null	log(A + 0.01)	3	125.52	1.60	0.04
	h	A	4	125.69	1.77	0.04
	j	null	3	125.75	1.83	0.04

2	null	A	3	109.21	0.00	0.09
	null	null	2	109.74	0.53	0.07
	log(A + 0.01)	null	3	109.75	0.55	0.07
	log(A + 0.01)	A	4	110.23	1.02	0.05
	j	A	4	110.50	1.30	0.05
	null	log(A + 0.01)	3	110.61	1.40	0.04
	j	null	3	110.81	1.60	0.04
	h	A	4	111.02	1.81	0.04

Table 2.5. Standardized β coefficients, standard errors (SE), and 95% lower and upper confidence intervals (LCI, UCI) from top-ranked models ($\Delta QAIC_c \leq 2$) designed to identify factors that affect survival and detection probabilities of sage-grouse broods based on daytime visual surveys in southern Idaho from 2016-2017. Explanatory variables include hen age (h), Julian hatch date (j), linear (A) or non-linear {log(A + 0.01)} effects of brood age, or no effect (null). Covariates that are in bold are considered meaningful (95% C.I. does not overlap zero).

Model	Parameter	β	SE	LCI	UCI
1	**φ{null}**	**4.66**	**0.42**	**3.84**	**5.48**
	p{A}	0.04	0.02	-0.01	0.09
2	φ{log(A+0.01)}	0.31	0.19	-0.06	0.68
	p{A}	0.03	0.02	-0.01	0.08
3	**φ{log(A+0.01)}**	**0.39**	**0.18**	**0.05**	**0.74**
	p{null}	**0.99**	**0.30**	**0.40**	**1.59**
4	φ{j}	-0.04	0.04	-0.11	0.04
	p{A}	0.04	0.02	-0.01	0.09
5	**φ{null}**	**4.77**	**0.49**	**3.81**	**5.73**
	p{null}	**0.91**	**0.30**	**0.33**	**1.48**
6	φ{h}	-0.59	0.96	-2.47	1.29
	p{A}	0.04	0.02	-0.01	0.09
7	**φ{null}**	**4.68**	**0.44**	**3.82**	**5.55**
	p{log(A+0.01)	0.41	0.32	-0.23	1.04
8	φ{j}	-0.03	0.03	-0.10	0.04
	φ{log(A+0.01)}	0.30	0.19	-0.08	0.67
	p{A}	0.03	0.02	-0.01	0.08
9	φ{j}	-0.04	0.04	-0.12	0.05
	φ{log(A+0.01)}	**0.37**	**0.18**	**0.02**	**0.72**
	p{null}	**0.99**	**0.31**	**0.38**	**1.60**

Table 2.6. Proportion of sage-grouse broods detected or not detected independently by 3 methods within 41-47 days after hatch in southern Idaho, 2016-2017. A Huggins model was used to calculate detection probabilities (*p*) per method.

Method	*p*	LCI	UCI
nighttime fecal	87.5%	61.3%	96.8%
daytime flush	93.7%	66.3%	99.1%
nighttime spotlight	93.7%	66.3%	99.1%

Figure 2.1. A nighttime roost-site of a radio-marked greater sage-grouse (*Centrocercus urophasianus*) hen in southern Idaho. Camera cap is 58 mm diam. Note the disturbed soil and white uric caps (circled) of chick fecal pellets.

Figure 2.2. Size differences between fecal pellets of greater sage-grouse (*Centrocercus urophasianus*): 15-day-old chick (top; in black circle), hen's *clockers* (middle), and other breeding-age sage-grouse fecal pellets (bottom) at nighttime roost-sites in southern Idaho. Camera cap is 58 mm diam and knife is 36 mm x 125 mm. Note the white uric cap and green to dark brown coloration.

Figure 2.3. Most sage-grouse (*Centrocercus urophasianus*) studies that used brood survey methods (*n* = 50) used daytime visual surveys or daytime flush surveys. Unk = daytime survey but it is unknown whether the author performed a daytime visual or daytime flush survey.

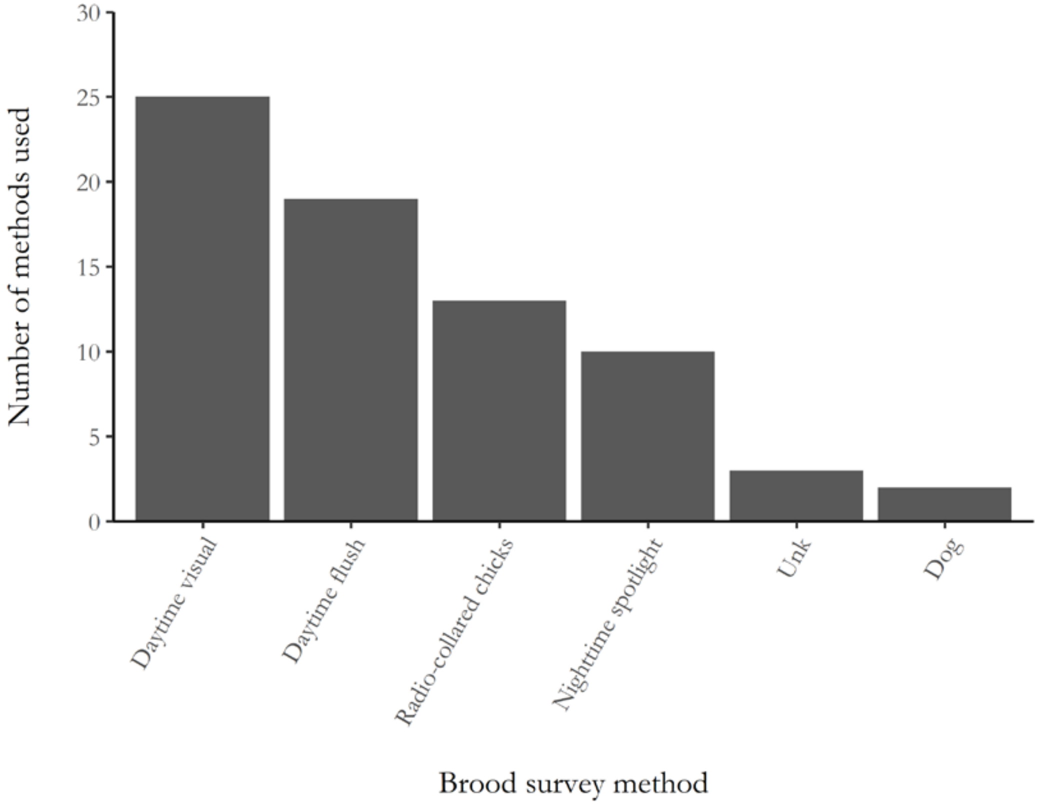


Figure 2.4. Most of the sage-grouse studies that included brood sampling used an analytical method that did not address detection probability. Only 2 of 50 papers that we reviewed accounted for brood detection probabilities by using an ANOVA and Lukacs young survival model (Dahlgren et al. 2010a; Gibson et al. 2017).

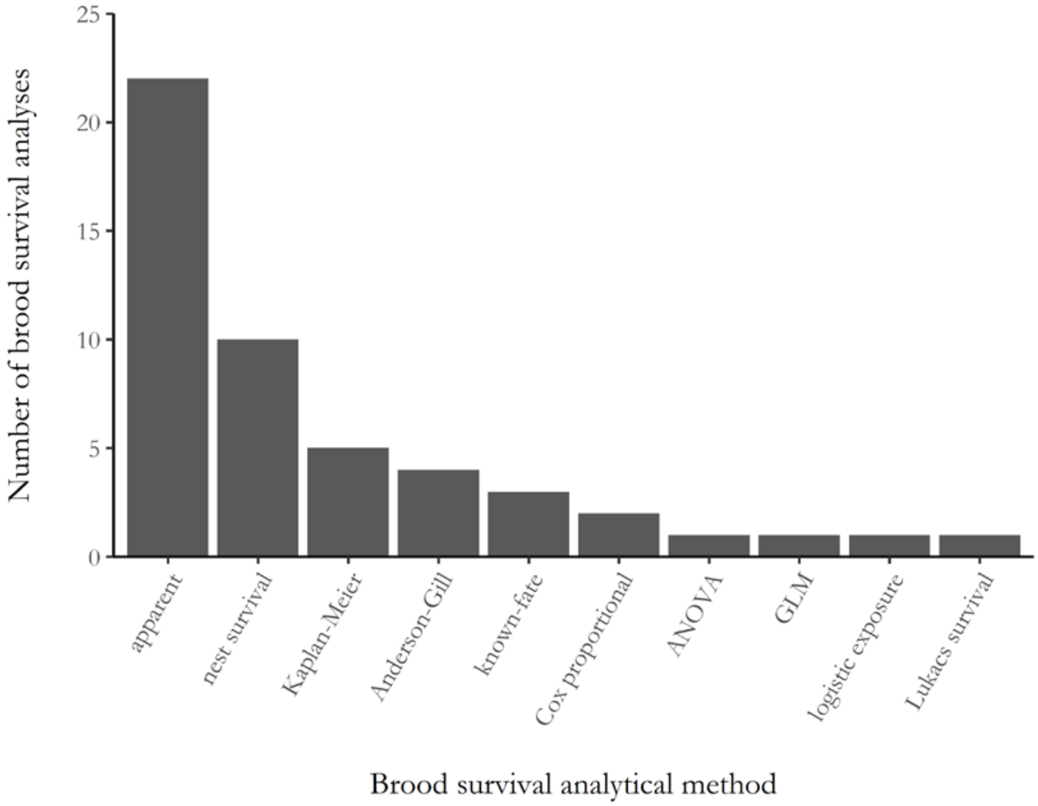


Figure 2.5. The deviance residual plot from a full-parameterized Cormack-Jolly-Seber model based on fecal pellet surveys showing no apparent trends in the residual values but some evidence of overdispersion. Overdispersion (i.e., lack of fit) is indicated when model residuals (solid dots) are outside the range of ± 2.5 deviance residuals (dotted line).

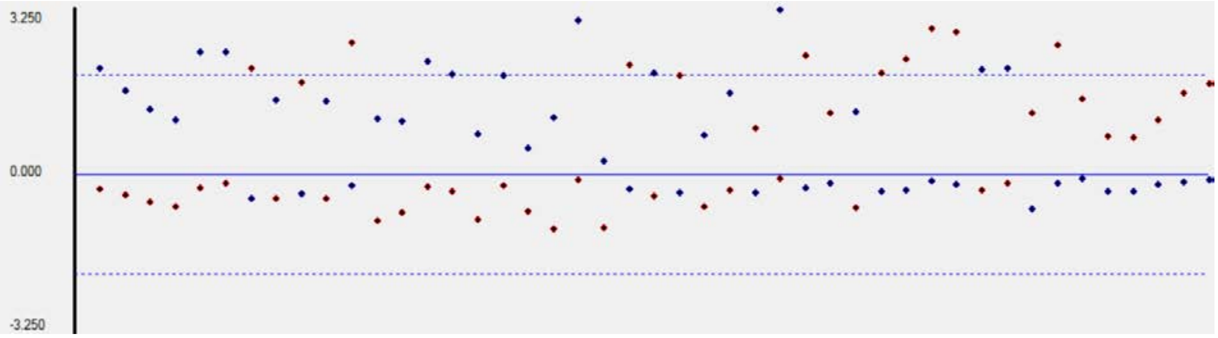


Figure 2.6. The deviance residual plot from a full-parameterized Cormack-Jolly-Seber model based on daytime visual surveys showing no apparent trends in the residual values but some evidence of overdispersion. Overdispersion (i.e., lack of fit) is indicated when model residuals (solid dots) are outside the range of ± 2.5 deviance residuals (dotted line).

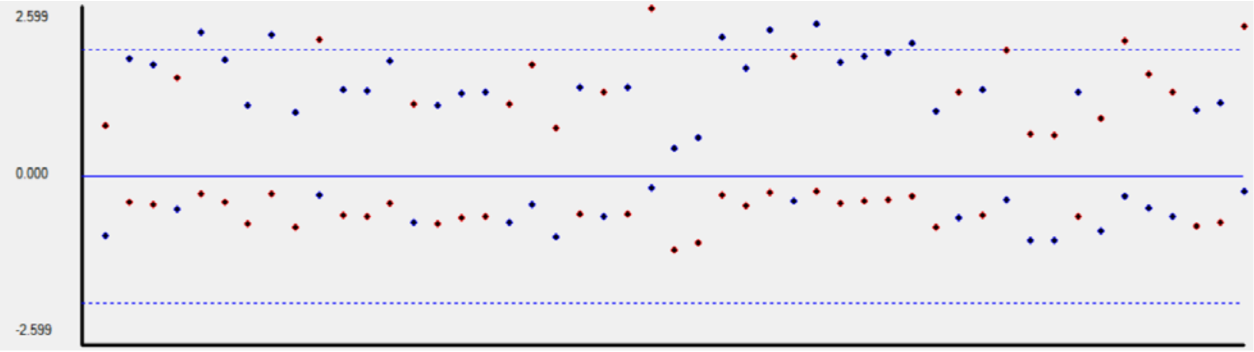


Figure 2.7. Number of greater sage-grouse chick pellets counted per roost-site in southern Idaho in 2016-2017.