2020

OPTIMIZING THE USE OF WILDLIFE MONITORING RESOURCES

Charles Raymond Henderson Jr.

Recommended Citation

Henderson, Charles Raymond Jr., "OPTIMIZING THE USE OF WILDLIFE MONITORING RESOURCES" (2020). *Graduate Student Theses, Dissertations, & Professional Papers*. 11604.
https://scholarworks.umt.edu/etd/11604

OPTIMIZING THE USE OF WILDLIFE MONITORING RESOURCES

By

CHARLES RAYMOND HENDERSON JR.

Bachelor of Arts, University of Tennessee, Knoxville, Tennessee, 1998
Master of Science, University of Montana, Missoula, Montana, 2014

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Fish and Wildlife Biology

The University of Montana
Missoula, MT

May 2020

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Paul M. Lukacs, Chair
Wildlife Biology Program

Chad J. Bishop
Wildlife Biology Program

Mark A. Hurley
Idaho Department of Fish and Game

Angela D. Luis
Wildlife Biology Program

Alexander L. Metcalf
Department of Society and Conservation

Michael S. Mitchell
Montana Cooperative Wildlife Research Unit

Henderson, Charles, Ph. D., Spring 2019 *Fish and Wildlife Biology*

Optimizing the use of wildlife monitoring resources

Chairperson: Paul M. Lukacs

ABSTRACT

Accurate knowledge regarding trends in the abundance of wildlife populations provides a foundation for the understanding of wildlife ecology and effective wildlife management. Abundance estimates enable managers and researchers to track the status of wildlife populations, supply information on which to base wildlife management decisions, and provide a metric to assess the outcome of specific management actions. Consequently, accurate estimates of abundance provide essential knowledge for managing wildlife populations effectively. However, the amount of resources available for monitoring wildlife populations is limited and often fluctuates in response to changes in annual budgets. Therefore, agencies responsible for the management of wildlife populations would benefit from knowing the most efficient manner in which to allocate their limited monitoring resources.

In order to determine the most efficient manner for allocating monitoring resources, I investigated several methods for improving upon and expanding current wildlife monitoring techniques. I used data collected by the Idaho Department of Fish and Game on mule deer (*Odocoileus hemionus*) populations throughout the state of Idaho. I compared survival rate estimates generated from global positioning system and very high frequency radio collars to determine if there are significant differences in estimates from the two data collection technologies. I developed a method for incorporating classification error into age and sex ratio estimates from data collected via aerial surveys. I assessed which types of commonly collected data about mule deer populations have the greatest influence on estimates of abundance generated using integrated population models. Finally, I developed a method to quantify the amount of information that can be gained about the abundance of a wildlife population under various levels of monitoring resource availability. The predicted gains in information are then used to suggest data collection scenarios that result in the optimal allocation of wildlife monitoring resources for mule deer. While all methods are developed in relation to the mule deer population of Idaho they can be extended to other species in a variety of geographical settings.

ACKNOWLEDGEMENTS

I would like to extend my sincere thanks to the Idaho Department of Fish and Game for providing the funding and support for this research and my education. I also wish to thank the Shikar Safari Club for their support through the Susan and John Monson Hunting Heritage Scholarship.

I would like to extend my deepest gratitude to my advisor Dr. Paul Lukacs, who gave me a chance to pursue my education and whose patience, insight, unparalleled knowledge, and willingness to explain the same thing over and over on the subject matter was instrumental in the completion of this research. I'm also deeply indebted to Dr. Mike Mitchell whose friendship, perspective, and passion inspired me to punch way above my weight class and achieve something I had not thought possible. I'm extremely grateful to Dr. Mark Hurley for keeping the heady ideas for research grounded in what is useful and practical for biologists on the ground and whose support was critical to the inception and completion of this research. I would like to express my deepest appreciation to the rest of my committee whose mentorship and input improved the quality of this research and facilitated my development into a scientist and wildlife biologist. I especially appreciate the theoretical understanding I developed under Dr. Angie Luis, the understanding of the underlying complexity of personal relationships and their impact on achieving conservation goals as explained by Dr. Chad Bishop, and the larger role of human dimensions in natural resource management highlighted by Dr. Alex Metcalf as well as his invaluable advice on techniques for getting through the written comps exam. I would like to acknowledge Gretchen McCaffery from the UM writing center for the development of my writing and editing skills which were vital for the completion of this dissertation.

I cannot begin to express my thanks to all of my friends and lab mates whose support and feedback were essential to my success. I especially want to thank Dr. Josh Nowak, Dr. Sarah Sells, Dr. Dave Ausband, Dr. Sara Williams, Anna Moeller, Jenny Helm, Molly McDevitt, Ben Jimenez, Sarah Bassing, Kristin Barker, Dr. Ally Keever, Kenneth Loonam and all other Lukacs and Mitchell lab members past and present. I cannot leave the U. of Montana without thanking the other faculty, staff, and graduate students of the Wildlife Biology Program, who make this such an amazing and special place to learn and develop both personally and professionally.

Finally, my biggest thanks to my family for all the support you have shown me through this process, the culmination of years of schooling and living on student wages. For my son, Forrest, thank you for making me smile every day and always being a source of inspiration to keep on going even when things weren't always that great. To my wife Stacey, thank you for all your support, for being there to hear me vent, for dealing with odd hours and an ever changing schedule because without you I would have stopped this insanity a long time ago. You have been amazing, and I promise I'm done with school and will now get a real job!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

2016 by mule deer population management units (PMU) designated by the Idaho Department of Fish and Game and Bio_Year. Bio_Year represents the year when a capture period ended. Numbers located in the grey box on the left hand side of the figure correspond to the following PMUs: Weiser-McCall (2), Central Mountains (4), Boise River (5), Smoky-Bennet (6), South Hills (8), Bannock (9), Palisades (11), Island Park (12), and Mountain Valley (13).

Figure 2-1. Differences in the observed fawn:doe ratios and simulated fawn:doe ratios when the probabilities of correctly classifying an individual within any age or sex class are less than 1. The probability of correctly classifying a fawn is held constant at 0.8, the probability of correctly classifying a buck is held constant at 0.9, and the probability of correctly classifying a doe ranges from 0.95 to 0.75. DD is the probability of classifying an individual as a doe given the individual is a doe. Red line indicates no difference between the observed and true ratio.

Figure 2-2. Differences in the observed buck:doe ratios and simulated buck:doe ratios when the probabilities of correctly classifying an individual within any age or sex class are less than 1. The probability of correctly classifying a fawn is held constant at 0.8, the probability of correctly classifying a buck is held constant at 0.9, and the probability of correctly classifying a doe ranges from 0.95 to 0.75. DD is the probability of classifying an individual as a doe given the individual is a doe. Red line indicates no difference between the observed and true ratio.

Figure 2-3. Estimates of abundance for the Boise River mule deer population management unit (PMU) over a 5 year period when fawn:doe ratios are increased

doe, BB is the probability of classifying an individual as a buck given the individual is a buck, and FF is the probability of classifying an individual as a fawn given the individual is a fawn.

Figure 2-8. Plot of the increase in the precision of the fawn:doe ratio estimate as the number of groups resampled is increased. The probability of correctly classifying a fawn as a fawn is held constant at 0.8. The values in the grey box on the right of the graph represent the probability of correctly classifying a buck as a buck. DD is the probability of classifying an individual as a doe given the individual is a doe, BB is the probability of classifying an individual as a buck given the individual is a buck, and FF is the probability of classifying an individual as a fawn given the individual is a fawn.

Figure 2-9. Plot of the increase in the precision of the buck:doe ratio estimate as the number of groups resampled is increased. The probability of correctly classifying a fawn as a fawn is held constant at 0.8. The values in the grey box on the right of the graph represent the probability of correctly classifying a buck as a buck. DD is the probability of classifying an individual as a doe given the individual is a doe, BB is the probability of classifying an individual as a buck given the individual is a buck, and FF is the probability of classifying an individual as a fawn given the individual is a fawn.

Figure 2-10. Elasticity values for survival rates from the population matrix as age and sex ratios change.

Figure 2-11. Elasticity values for reproductive rates from the population matrix as age and sex ratios change.

DISSERTATION OVERVIEW AND INTRODUCTION

Accurate knowledge regarding trends in the abundance of wildlife populations provides a foundation for the understanding of wildlife ecology and effective wildlife management. Abundance estimates enable managers and researchers to track the status of wildlife populations, supply information on which to base wildlife management decisions, and provide a metric to assess the outcome of specific management actions (Terletzky and Koons 2016). Consequently, accurate estimates of abundance provide essential knowledge for managing wildlife populations effectively. However, the amount of resources available for monitoring wildlife populations is limited and often fluctuates in response to changes in annual budgets (Chades et al. 2008). Therefore, agencies responsible for the management of wildlife populations would benefit from knowing the most efficient manner in which to allocate their limited monitoring resources. In order to determine the most efficient manner for allocating monitoring resources, we began several converging lines of research to gather information regarding how and where to use monitoring resources most effectively. The information resulting from these lines of research are then combined to produce a set of data collection scenarios, which simultaneously maintain or increase the accuracy of abundance estimates while reducing the cost of monitoring. The resulting data collection scenarios provide a set of potential options to management agencies to decide from when allocating monitoring resources when monitoring resource availability and management needs change periodically.

The first element in the converging research lines is an assessment of the impacts of using different radio collar technologies to estimate survival. This assessment provides information regarding the how inherent differences in the function of global positioning system

(GPS) and very high frequency (VHF) radio collar technologies impact the estimation of a vital rate crucial to the estimation of wildlife abundance. The second element of the converging lines of research develops a multinomial likelihood to account for the classification error present in most surveys of wildlife populations. This tool will allow managers and researchers to deal explicitly with the presence of classification error within a modeling framework and eliminate the common assumption that this type of error can be ignored or is too small to make a difference in the estimation of abundance. The third line of research explores and quantifies the impact of data with varying levels of quality on the estimation of abundance using integrated population models (IPM). The quantification of the behavior of data types of different relative qualities within the IPM provides information suggesting which data types contribute the most to the estimation of abundance. The information on the contribution of each of the different data types can inform decisions regarding the allocation of wildlife monitoring resources for data collection. The final research thread brings elements of the previous lines of research together to determine the optimal data collections scenarios for collecting data for estimating the abundance of wildlife populations. This research line develops a new metric for describing the relationship between the costs of gathering different data types and the expected gains in information about abundance from varying amounts and types of data.

Chapter 1 provides an analysis of the differences in the estimates of survival generated from survival data collected via GPS and VHF radio collars. Survival data is expected to have a large influence on the estimate of abundance and therefore any biases related to how this information is collected should be quantified. Our initial analysis focused on the estimates of survival for female adults. We expect the survival of this age and sex class to have the least amount of variability due to the evolutionary canalization of this vital rate (Gaillard and Yoccoz

2003). Therefore, when we analyze individuals that overlap both temporally and spatially, differences in the estimate of survival of this age and sex class are more likely to be related to the technology used to collect survival information about them. Analysis of these two survival data gathering technologies can serve to guide choices concerning the appropriate type of monitoring technology to use, account for differences between survival estimates from data collected by the different technologies, and improve the quality of data used in the IPM for this highly influential data source.

Ratio data provides information on potential or realized population growth and is an important component for models that estimate the abundance of populations. However, there are potential problems with ratio data that have been recognized for decades (Caughley 1974, McCullough 1994). One potential problem with ratio data is the presence of classification errors, assigning an individual to the wrong age or sex class. Misclassification can result in inaccurate age and sex ratios, which then influence the accuracy of abundance estimates generated using these ratios. Due to the potential for problems caused by inaccurate ratio data, it is important to address the presence of classification error in wildlife surveys to eliminate one potential source of error that can lead to poor quality ratio data. Chapter 2 provides a method using a multinomial classification likelihood to account for this type of error when estimating age and sex ratios.

Chapter 3 explores the ways in which IPMs incorporate data of different relative qualities. The influence that different quality data has on estimates of abundance are of interest when determining which data types to collect and how much of each data type is necessary for accurate estimates of abundance. By testing the performance of the IPM, by measuring the bias and precision of abundance estimates generated by the model, while changing the bias and precision of the data used by the model we can assess the influence that data of different relative qualities have on final estimate of abundance. Using this methodology we can also track the weights given to each data type within the IPM and follow how these weights change as the quality of the data changes. Tracking weights is one way to determine whether the IPM is performing in a manner consistent with expectations, poorer quality data should receive less weight within the model than higher quality data. This is important information to quantify because it provides further evidence that integrated models are capable of accommodating different quality data types within the hierarchical model structure. More importantly, tracking weights also provides information that can be used to prioritize which data types are most important to the accurate estimate of abundance and suggests that the quality of certain data types can be changed without influencing the accuracy of abundance estimates. This analysis provides information regarding both the quality and quantity needed of different data types to generate accurate estimates of abundance.

In Chapter 4 we quantify how the amount of information gained, in the form of the precision of the abundance estimate, is related to the cost of gathering different quantities and qualities of data. We accomplish this by combining the methods and results from previous analyses and by developing a new metric, the information gain (IG) ratio, which quantifies the relationship between data collections costs and abundance estimate precision. The IG ratio is then used to assess various data collection scenarios to determine which scenarios optimally allocate available monitoring resources. The results of this research provide a method which wildlife management agencies can use to ascertain the most efficient use of the limited and changing amounts of monitoring resources available to them. The technique developed here also provides a defensible method for making decisions regarding the use of public resources.

LITERATURE CITED

Caughley, G., 1974. Interpretation of age ratios. The Journal of Wildlife Management, pp.557-562.

Chadès, I., McDonald-Madden, E., McCarthy, M.A., Wintle, B., Linkie, M. and Possingham, H.P., 2008. When to stop managing or surveying cryptic threatened species. Proceedings of the National Academy of Sciences.

Gaillard, J.M. and Yoccoz, N.G., 2003. Temporal variation in survival of mammals: a case of environmental canalization?. Ecology, 84(12), pp.3294-3306.

McCullough, D.R., 1994. In My Experience: What Do Herd Composition Counts Tell Us?. Wildlife Society Bulletin (1973-2006), 22(2), pp.295-300.

Terletzky, P.A. and Koons, D.N., 2016. Estimating ungulate abundance while accounting for multiple sources of observation error. Wildlife Society Bulletin, 40(3), pp.525-536.

CHAPTER 1: A COMPARISON OF GLOBAL POSITION SYSTEM (GPS) AND VERY HIGH FREQUENCY (VHF) SURVIVAL RATE ESTIMATES

INTRODUCTION

Accurate estimates of population abundance play a central role in the effective management and study of wildlife populations, and rely on the quality and quantity of the data available. Data commonly used for abundance estimation include counts of the population, data regarding the number of births and deaths occurring in the population during the period of interest, and the number of individuals moving in and out of the population during the same time period. Since counting the number of individuals that died in a large population is difficult to accomplish and often cost prohibitive, many models rely on estimates of survival to provide information regarding mortality in the population. Therefore, the accuracy of estimates of survival for the various age and sex classes in a population have a large impact on the accuracy of abundance estimates.

Decomposing one of the basic models used to describe population growth demonstrates the importance of survival rates to the estimation of abundance. The following equation represents this basic abundance model:

$$N_{t+1} = N_t + B_t + I_t - D_t - E_t$$

where N_{t+1} is the estimate of abundance one time step in the future, N_t is the abundance at time *t*, B_t is the number of births in the population at time *t*, I_t is the number of immigrants into the population at time *t*, D_t is the number of deaths in the population at time *t*, and E_t is the number of emigrants out of the population at time *t*. Assuming a closed population further simplifies this equation. The closure assumption allows the dropping of the terms for immigration and emigration, I_t and E_t respectively, leaving the three data types essential for all estimates of abundance. Of the three remaining terms, D_t, is the one that is directly related to survival since it is a count of the number of individuals that died in a population. Breaking down D_t into the equation, $D_t = N_t * d_t$, where N_t is the number of individuals in the population at time t and d_t is the rate at which the individuals in the population die provides a way to calculate survival (s_t) via the equation, $s_t = 1 - d_t$. Thus demonstrating that estimates of survival can provide one of the three critical pieces of information needed to estimate abundance.

Estimates of survival rates come from data gathered using various types of technology and one of the most common technologies used are tracking collars. Previously, tracking collars had been limited to the use of very high frequency (VHF) radio signals to monitor individuals outfitted with these devices. However, in the past two decades the emergence, increase in availability, and reduction in price of global positioning system (GPS) tracking collars have led to a wider use of this technology for monitoring wildlife. Both of these monitoring technologies have attributes that could potentially bias survival data in a systematic way.

The attributes of each technology system that have the potential to bias the data gathered by them relate to the inherent differences between how each system operates. GPS collars rely on access to satellites to gather and transmit data, whereas VHF collars rely on the broadcast range of the radio frequencies transmitted by the collar to collect data. These innate differences in each technology affect how environmental conditions can influence the data each type of collar is capable of collecting. Weather conditions and canopy coverage can influence the accessibility of satellite signals to GPS collars, thus interfering with their ability to collect data. Weather and topography can impact the range of VHF signals and weather can directly impact the ability to conduct surveys for VHF signals. Inherent differences in the technologies can also result in a large difference in the amount of geographic area each technology can monitor, with GPS researchers are able to monitor a much larger geographic area relative to VHF. GPS collars generally have a shorter battery life relative to VHF collars due to the increased energetic costs of searching for and connecting to satellites, storing data, and transmitting collected data. Finally, monitoring VHF collared individuals is much more labor intensive relative to monitoring GPS collared individuals, which often results in a difference in the frequency that individuals with different type collars are "observed". Any one of these differences have the potential to systematically influence the data collected by GPS and VHF collars resulting in differences in the accuracy, defined as bias and precision, of the survival rates estimated from the data collected by the different types of tracking collar.

It is important to determine whether data collected by different technologies result in different estimates of survival because small differences in the estimates of survival have the potential to influence estimates of abundance. Depending on the value of the survival estimate, very different conclusions about the population of interest may result. For example, Figure 1-1 shows how different conclusions regarding the trajectory of a population can be made if the population of interest is projected into the future using two survival rates (0.89, SD = 0.01; 0.86, SD = 0.01) that differ by a relatively small amount while all other demographic values are held constant. In this example, the trajectory of the population with a survival estimate for adult females of 0.89 is positive over time. Conversely, the trajectory of the population with a survival estimate for adult females of 0.86 is negative over the same time period. The difference in the point estimates of survival for the hypothetical survival rate values is only 0.03 but the population trajectory associated with each of the survival estimates result in opposite population trajectories. Figures 1-2 and 1-3 show two other hypothetical situations where the survival estimates remain 0.03 apart but do not contain the value of adult female survival which results in a stable population. In these two hypothetical examples, the population trajectories are similar but the abundance estimates are increasingly different over time. These hypothetical examples highlight that while differences in the survival estimates generated from data collected by two different technologies may be relatively small, they have the potential to influence the inferences made about the trajectory and size of the population and consequently management decisions made regarding a population.

Due to the important role that survival data play in the estimation of abundance, determining if there are differences between the estimates of survival generated by the two data-collecting systems is essential to effective wildlife management and study. Our analysis focuses on the estimates of survival for female adults in the mule deer (*Odocoileus hemionus*) population of Idaho generated from data collected using both GPS and VHF tracking collars for two reasons. First, small changes in the survival rate of adult females in populations of mule deer have large impacts on population growth rates and consequently the abundance of the population due to the high elasticity values associated with this vital rate. The second reason for focusing on the survival of adult females is the expectation that the survival of this age and sex class will have the least amount of variability due to the evolutionary canalization of this vital rate (Gaillard and Yoccoz 2003). Therefore, if we analyze data only collected from individuals in the adult female class, which overlap both temporally and spatially, any differences in the estimate of the survival are more likely to be due to the technology used to collect the survival data rather than differences between individuals. Analysis of these two survival data-gathering technologies serves to improve the quality of data used to estimate abundance from this highly influential data source.

METHODS

We used data collected between 2007-2018 by Idaho Department of Fish and Game (IDFG) on the survival of adult female mule deer for this analysis. We analyzed survival data from the following mule deer population management units (PMU): Bannock (9), Boise River (5), Central Mountains (4), Island Park (12), Mountain Valley (13), Palisades (11), Smoky-Bennet (6), South Hills (8), and Weiser-McCall (2). The number in the parentheses represents the PMU ID number associated with each PMU in the data.

We split the data into three periods to compare estimates of survival rates. We defined each period by the relative distribution of collar types and demarcated each period by the Bio Year, or year when the capture period ended. For example, Bio Year 2012 includes the capture period beginning in November of 2011 and ending in April of 2012. The first time period, referred to as VHF dominant, includes individuals captured between December 2007 and January 2011 (Bio Years 2008 – 2011) and VHF collars make up the largest proportion of the collars deployed during this time period (Figure 1-4). The second time period in which the distribution of deployed collars are more evenly split between GPS and VHF collar types, is referred to as the transitional period and includes individuals captured between December 2011 and January 2014 (Bio Years 2012 – 2014, Figure 1-4). The third time period, referred to as GPS dominant, includes individuals captured between December 2014 and January 2016 (Bio Years 2015 – 2016) and GPS collars make up the largest proportion of the collars deployed during this time period (Figure 1-4).

The deployment distribution of GPS and VHF collars is also broken down by the number of collars deployed over the entire data set and in each time period by PMU and year (Figures 1-5, 1-6, 1-7, and 1-8). We assumed that if survival estimates were similar within each of the three periods of time for both VHF and GPS collars, any differences in the estimates of survival relate to differences in the technological capabilities of the collars and not due to other factors, such as environmental or demographic stochasticity.

We created encounter histories for each collared deer to estimate annual and monthly survival rates. Due to conflicting dates within the data regarding last date observed, fate dates, and censor dates for multiple individuals, the capture date and the last date observed were used to begin and end each individual's encounter history. We did not include any individuals captured after January 2016 to eliminate any bias in the estimate of survival rates due to a relatively shorter monitoring period. The encounter histories we created only included observations until January 2018 to ensure that all individuals had the possibility of being monitored for a minimum of two years. We also estimated survival for each time period with and without harvested individuals being included in the sample. This was to ensure that differences in hunting regulations across PMUs did not influence our comparison of survival rates between GPS and VHF collared individuals. This also allowed us to determine if there were different harvest rates potentially linked to collar type.

We estimated survival using a known fate model for adult female mule deer for each of the 3 time periods by individual PMU, by combining all data from all PMUs in a single time period, and by combining all data from all PMUs across all 3 time periods (White and Burnham 1999). The linear regression model used for this analysis was

$$y_t = \beta_0 + \beta_1 * collartype$$

where y_t is observation of alive or dead, β_0 is the y-intercept, and $\beta_1 * collartype$ represents whether the individual was outfitted with a GPS or VHF tracking collar. We employed a Bayesian statistical framework to conduct the survival analysis in program JAGS. We used

JAGS 4.2.0 (JAGS 4.2.0, www.mcmc-jags.sourceforge.net, accessed 28 Nov 2016) and R package R2jags (Su and Yajima 2015, www.r-project.org, accessed 22 Aug 2015). The R code used for the survival model was adapted from the PopR software (Nowak et al. 2018). We employed uninformative priors drawn from a uniform distribution (-6, 6). We assessed convergence using the "Rhat" value from the BUGS output. We ran models for 50,000 iterations and discarded the initial 5,000 iterations for the burn-in period. We used 5 chains with different initial values to ensure that initial values did not influence estimates.

RESULTS

We used the point estimates and 95% credible intervals (CRI) of the survival estimates for adult female mule deer from data collected via GPS and VHF tracking collars to determine whether there were differences between survival estimates generated from the data collected by each tracking collar type. The CRIs of the survival estimates overlapped to some degree for all estimates however the amount of overlap varied. We used the amount of overlap between the CRIs of GPS and VHF based survival estimates to quantify the probability or strength of the difference between estimates made from GPS and VHF collected data. For instance, if the GPS survival estimate was 0.85, the VHF estimate was 0.78, and the CRIs for the two estimates overlapped by 30% then we concluded that the probability that the GPS estimate is higher than the VHF estimate is 70% (Pr[GPS>VHF] = 0.70).

Over all three time periods the survival estimates from GPS collar data were higher than the survival estimates from the VHF collar data when harvested individuals were included (Table 1). The amount of overlap in the CRIs of the survival estimate when harvested individuals were included for the three time periods was between 0.007% – 70.8%. In the VHF dominant period, the probability that the GPS estimate of survival was greater than the VHF estimate of survival was 0.804 (n_{GPS} = 17, n_{VHF} = 365). During the transitional period, the probability of the GPS estimate of survival being greater than the VHF estimate of survival was 0.993 (n_{GPS} = 89, n_{VHF} = 131). In the GPS dominant period, the probability that the GPS estimate of survival was greater than the VHF estimate of survival was 0.292 (n_{GPS} = 147, n_{VHF} = 6). Results from the analysis of survival estimates for adult female mule deer from the two tracking collar technologies over the three time periods show that estimates from GPS collars are always higher when harvested individuals are included. These results also demonstrate when the distribution of collars is more even, during the transitional period, there is an increase in the probability that the survival estimate generated from GPS data will be greater than the survival estimate generated from data collected by VHF collars.

When we removed harvested individuals, the GPS based estimates of survival were greater than the VHF estimates of survival for two of the three time periods. Over the VHF dominant and transitional periods the survival estimates from GPS collar data were higher than the survival estimates from the VHF collar data but the VHF survival estimate was higher than the GPS survival estimate for the GPS dominant period (Table 1-1). The amount of overlap in the CRIs of the survival estimate when harvested individuals were not included for the three time periods was between 26.5% – 63.9%. In the VHF dominant period, the probability that the GPS estimate of survival was greater than the VHF estimate of survival was 0.674 (n_{GPS} = 15, n_{VHF} = 190). During the transitional period, the probability of the GPS estimate of survival being greater than the VHF estimate of survival was 0.735 (n_{GPS} = 74, n_{VHF} = 90). However in the GPS dominant period the pattern reversed and the VHF estimate of survival was greater than the GPS estimate of survival with a probability of 0.391 (n_{GPS} = 133, n_{VHF} = 3). During the GPS dominant period, the strength of difference between GPS and VHF estimates of survival was the lowest no matter if harvested individuals were included (Pr[GPS>VHF] = 0.292) or removed (Pr[VHF>GPS] = 0.361). These results again demonstrate when the distribution of collars is more even, during the transitional period, there is an increase in the probability that the survival estimate generated from GPS data will be greater than the survival estimate generated from data collected by VHF collars.

We estimated the survival rates for adult females in individual PMUs during the transitional period, when the distribution of the two collar types was more even, to determine if survival rates in individual PMUs followed a similar pattern as the results from the combined PMU analysis during the transitional period. GPS based survival estimates were higher in 88.9% of the PMUs than VHF based survival estimates with and without harvested individuals (Tables 1-2 and 1-3). PMU 4 (Central Mountains) was the only PMU in the transitional period where estimates of survival based on data collected by VHF collars was higher than survival estimates based on data collected from GPS collars. VHF survival estimates were higher for PMU 4 both with and without harvested individuals.

The strength of the difference between estimates of survival by PMU during the transitional period from data collected by GPS and VHF collars was different when harvested individuals were included versus when harvested individuals were removed from the data. When harvested individuals were included and the GPS based survival estimate was greater than the VHF based estimate, the probability of GPS estimates being higher ranged from 0.344 to 0.977 (Table 1-2), with a mean probability of GPS being greater than VHF of 0.676. In the single PMU where the VHF survival estimate was greater than the GPS estimate of survival, the probability of VHF being greater than the GPS was 0.347. When harvested individuals were removed and the GPS based survival estimate was greater than the VHF based estimate, the probability of GPS estimates being higher ranged from 0.140 to 0.838 (Table 1-3), with a mean probability of GPS being greater than VHF of 0.527. In the single PMU where the VHF survival estimate was greater than the GPS estimate of survival, the probability of VHF being greater than the GPS was 0.514. The estimates generated from GPS collars were uniformly higher across all PMUs, except for PMU 4, during the transitional period however the strength of the difference decreased in 63% of the PMUs when harvested individuals were not present in the sample. The reduction in strength of difference is likely related to the decrease in sample sizes when harvested individuals were removed.

Survival estimates from GPS collars were higher in the majority of each of the time periods when we combined all PMUs and for all individual PMUs, except PMU 4, during the transitional period. Annual survival rates for individuals outfitted with both collar types typically increased when we removed harvested individuals. The mean magnitude of the increase was greater for VHF collared individuals (0.074) than GPS collared individuals (0.022).

DISCUSSION

Survival estimates play a critical role in the estimation of wildlife abundance, so determining if there are systematic biases introduced into the estimates of survival related to the technologies used to collect data for these estimates is critical for wildlife management and study. The technology used to collect data for estimating survival has shifted during the past two decades from VHF to GPS tracking collars. We compared survival estimates from data gathered by the two tracking collars types which overlapped both spatially and temporally. This analysis was to determine if there are consistent differences that researchers and managers should consider when deciding which monitoring technology to deploy. Our results demonstrate that in the majority of cases, data collected using GPS tracking collars result in survival estimates that are higher than survival estimates from data collected via VHF tracking collars.

The analyses with the greatest support for the observed pattern of higher GPS based estimates occur during the transitional period when the number of collars deployed for each type of data collecting technology are most similar. Further support for the observed pattern reflecting a true difference in survival estimates from this time period comes from the relatively small and similar standard deviations associated with the survival estimates from both technology types for the transitional period (Table 1-1). The similarity of standard deviation values during the transitional period suggest both tracking collar types are capturing the variability in the population yet the survival estimates are different by 0.112 when harvested individuals are included and 0.048 when harvested individuals are removed. These results demonstrate that when deployed in similar numbers and under similar conditions survival estimates from data collected by GPS tracking collars will result in higher estimates of survival than survival estimates based on data collected by VHF tracking collars.

Analyses from the other time periods, except when we remove harvested individuals from the GPS dominant period, also result in a pattern of higher GPS based survival estimates but the probability of the GPS and VHF based survival estimates being different is smaller. The reduction in the strength of the difference between survival estimates relates to the reduced sample sizes during these analysis periods (Table 1-1). Reduced sample sizes for one or the other tracking collar types increase the standard deviations and consequently results in greater overlap of the CRIs of the survival estimate. Greater overlap of the CRIs of the survival estimates reduce the strength of the evidence for the difference between GPS and VHF based survival estimates. However, the pattern of higher survival estimates remains consistent across time periods and PMUs with one exception. The exception to the pattern of higher GPS based estimates of survival occur when we remove harvested individuals from the GPS dominant period. This exception is the likely a result of the small number of VHF collars ($n_{VHF} = 3$) included in this analysis and is also associated with the second smallest value for strength of difference. Therefore, we attribute the exception to the pattern of higher GPS based estimates of survival to the VHF sample size and not a change in the capabilities of the technology.

There are several potential explanations for the estimates of survival from GPS collars being higher than survival estimates from VHF collars in the majority of the analyses performed. First, there may have been an unintentional bias when deploying GPS collars, especially during the VHF dominant and transitional time periods. During these first two time periods, GPS collars were more expensive and less common and therefore may have only been placed on individuals that were in the best physical condition and thus more likely to survive. A second potential explanation for the generally higher estimate of survival from GPS collars may indicate that VHF collared individuals were harvested at a higher rate. This has some support from the magnitude of the change in VHF survival estimates relative to the magnitude of the change in GPS survival estimates when we removed harvested individuals from the data. Survival rate estimates for both collar types increased when we removed harvested individuals in each time period. However, the increase in the mean estimate of survival for VHF collared individuals is 0.074 compared to 0.022 for GPS collared individuals yet GPS based estimates of survival remained the higher relative to VHF based estimates of survival in the majority of analyses.

A third potential explanation for the higher estimates of survival from data collected via GPS relates to an assumption related to the frequency of observations associated with the known fate model used to estimate survival. The assumption regarding the frequency of observations in the known fate model is that we observe each individual on every occasion. Differences between GPS and VHF technological capabilities result in more violations of the observation assumption in data collected via VHF than data collected via GPS. Violation of this assumption in the known fate model results in an estimate of survival which is biased low. This trait of known fate survival models suggests the pattern of higher estimates of survival from data collected by GPS tracking collars may be the result of the increased number of violations of the observation assumption in the data collected via VHF.

We suggest differences in the technological attributes of the two tracking collars which directly influence the amount of violations of the observation assumption of the known fate model are responsible for the consistent pattern of higher estimates of survival from data collected via GPS across time periods and PMUs. Factors impacting the frequency of data collection for each collar type include the environmental attributes that impact access to satellites versus broadcast range of the radio frequencies, differences in the amount of geographic area each technology can monitor, and the frequency that individuals with different type collars are "observed". Due to the design of this study we were not able to isolate a single technology attribute or set of technology attributes that are wholly responsible for the frequency of observations and the consequent increase in estimates of survival from data collected via GPS relative to data collected by VHF. Discerning the particular technological attributes impacting the difference is one potential line of further inquiry which requires a study design in which control of a variety of factors is necessary. This type of study may be better suited to a captive setting or more controlled environment but was outside the scope of this research.

The results of this analysis demonstrate that researchers and managers can typically expect higher estimates of survival from data collected via GPS tracking collars relative to estimates of survival from data collected via VHF tracking collars. We suggest that a reduction in the amount of model assumption violations also results in less biased estimates in the estimates of survival from data collected via GPS. Therefore, the choice of which technology to use should depend on the goal of an individual analysis and the amount of precision required from the survival estimates generated from the study. There may be instances when a less expensive and survival estimate that is biased low is adequate for management or research purposes. These results suggest when non-biased estimates of survival are required, data from GPS tracking collars should provide less biased survival estimates that will typically be greater than survival estimates from data collected via VHF. Therefore, the choice of which technology to use should depend on the goal of an individual analysis and the resources available.

LITERATURE CITED

Gaillard, J.M. and Yoccoz, N.G., 2003. Temporal variation in survival of mammals: a case of environmental canalization?. Ecology, 84(12), pp.3294-3306.

Nowak, J.J., Lukacs, P.M., Hurley, M.A., Lindbloom, A.J., Robling, K.A., Gude, J.A. and Robinson, H., 2018. Customized software to streamline routine analyses for wildlife management.Wildlife Society Bulletin,42(1), pp.144-149.

Plummer, M., 2003, March. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. InProceedings of the 3rd international workshop on distributed statistical computing(Vol. 124, No. 125.10).

Su, Y. and Yajima, M. 2015. R2jags: Using R to Run 'JAGS'. R package version 0.5-7. https://CRAN.R-project.org/package=R2jags

White, G.C. and Burnham, K.P., 1999. Program MARK: survival estimation from populations of marked animals.Bird study,46(sup1), pp.S120-S139.

TABLES

Table 1-1. Comparison of estimates of annual Survival of global positioning system (GPS) and very high frequency (VHF) collared adult females captured between 12/2007 - 01/2016 with and without harvested individuals including the following mule deer population management units (PMU) designated by the Idaho Department of Fish and Game: Weiser-McCall, Central Mountains, Boise River, Smoky-Bennet, South Hills, Bannock, Palisades, Island Park, and Mountain Valley. GPSann is the annual survival estimate for adult females from data collected by GPS tracking collars. VHFann is the annual survival estimate for adult females from data collected by VHF tracking collars. SD represents the standard deviation of the estimate of survival, Percent CRI overlap represents the amount of overlap of the credible intervals for the estimates of survival, and PR> represents the probability that the highest survival estimate is actually greater than the other.

Year	Harvest Included	Parameter	Mean	SD	Number of Collars	Percent CRI Overlap	PR >
2007-2011	Yes	GPSann	0.897	0.067	17	0.196	0.804
		VHFann	0.81	0.012	356		
2011-2014		GPSann	0.943	0.021	89	0.007	0.993
		VHFann	0.831	0.021	131		
2014-2016		GPSann	0.808	0.027	147	0.708	0.292
		VHFann	0.787	0.148	6		
2007-2011	No	GPSann	0.946	0.051	15	0.326	0.674
		VHFann	0.897	0.012	190		

2011-2014	GPSann	0.941	0.023	74	0.265	0.735
	VHFann	0.893	0.02	90		
2014-2016	GPSann	0.83	0.027	133	0.639	0.361
	VHFann	0.86	0.13	3		

Table 1-2. Annual survival by mule deer population management units (PMU) designated by the Idaho Department of Fish and Game for adult female mule deer captured between 12/2011-01/2014 with harvested individuals included. PMU numbers are associated with the following PMU: Bannock (9), Boise River (5), Central Mountains (4), Island Park (12), Mountain Valley (13), Palisades (11), Smoky-Bennet (6), South Hills (8), and Weiser-McCall (2). SD represents the standard deviation of the estimate of survival, Percent CRI overlap represents the amount of overlap of the credible intervals for the estimates of survival, and PR> represents the probability that the highest survival estimate is actually greater than the other.

PMU	parameter	Mean	SD	LCI	UCI	Number of collars	Percent Overlap	PR >
pmu2	GPSann	0.940	0.056	0.790	0.998	11	0.095	0.905
	VHFann	0.740	0.064	0.607	0.854	16		
pmu4	GPSann	0.820	0.091	0.610	0.958	16	0.653	0.347
	VHFann	0.873	0.052	0.755	0.954	16		
pmu5	GPSann	0.925	0.069	0.742	0.997	9	0.437	0.563
	VHFann	0.784	0.125	0.493	0.957	7		
pmu6	GPSann	0.927	0.067	0.750	0.998	7	0.631	0.369
	VHFann	0.881	0.042	0.788	0.950	21		
pmu8	GPSann	0.965	0.067	0.762	1.000	2	0.247	0.753
	VHFann	0.830	0.051	0.719	0.918	19		
pmu9	GPSann	0.952	0.046	0.831	0.998	11	0.202	0.798
	VHFann	0.846	0.038	0.766	0.913	31		
pmu11	GPSann	0.946	0.107	0.605	1.000	3	0.656	0.344
	VHFann	0.838	0.161	0.362	0.969	1		
pmu12	GPSann	0.959	0.068	0.755	1.000	5	0.301	0.699
	VHFann	0.616	0.352	0.000	0.968	0		
pmu13	GPSann	0.991	0.014	0.949	1.000	25	0.023	0.977
	VHFann	0.808	0.070	0.654	0.925	20		

Table 1-3. Annual survival by mule deer population management units (PMU) designated by the Idaho Department of Fish and Game for adult female mule deer captured between 12/2011-01/2014 without harvested individuals included. PMU numbers are associated with the following PMU: Bannock (9), Boise River (5), Central Mountains (4), Island Park (12), Mountain Valley (13), Palisades (11), Smoky-Bennet (6), South Hills (8), and Weiser-McCall (2). SD represents the standard deviation of the estimate of survival, LCI represents the lower limit of the 95% credible interval of the posterior distribution of the estimate of adult female survival, and UCI represents the upper limit of the 95% credible interval of the posterior distribution of the estimate of adult female survival.

PMU	parameter	Mean	SD	LCI	UCI	Number of collars	Percent Overlap	PR >
pmu2	GPSann	0.935	0.061	0.773	0.998	9	0.405	0.595
	VHFann	0.827	0.069	0.672	0.939	9		
pmu4	GPSann	0.810	0.096	0.591	0.956	14	0.486	0.514
	VHFann	0.900	0.049	0.782	0.966	13		
pmu5	GPSann	0.925	0.069	0.743	0.998	9	0.860	0.140
	VHFann	0.901	0.082	0.667	0.970	5		
pmu6	GPSann	0.986	0.030	0.897	1.000	5	0.464	0.536
	VHFann	0.948	0.023	0.886	0.970	12		
pmu8	GPSann	0.966	0.067	0.759	1.000	2	0.318	0.682
	VHFann	0.848	0.053	0.731	0.936	14		
pmu9	GPSann	0.941	0.055	0.795	0.998	10	0.625	0.375
	VHFann	0.903	0.034	0.826	0.958	21		

pmu11	GPSann	0.946	0.105	0.611	1.000	3	0.649	0.351
	VHFann	0.838	0.162	0.361	0.969	1		
pmu12	GPSann	0.959	0.067	0.757	1.000	5	0.299	0.701
	VHFann	0.615	0.352	0.000	0.968	0		
pmu13	GPSann	0.987	0.024	0.918	1.000	17	0.162	0.838
	VHFann	0.866	0.069	0.702	0.962	15		

FIGURES

Figure 1-1. Differences in the estimate of abundance over time when hypothetical survival estimates for adult females from global positioning system (GPS) and very high frequency (VHF) collared collar types that are not statistically different straddle the value where the population growth rate changes from increasing to decreasing.

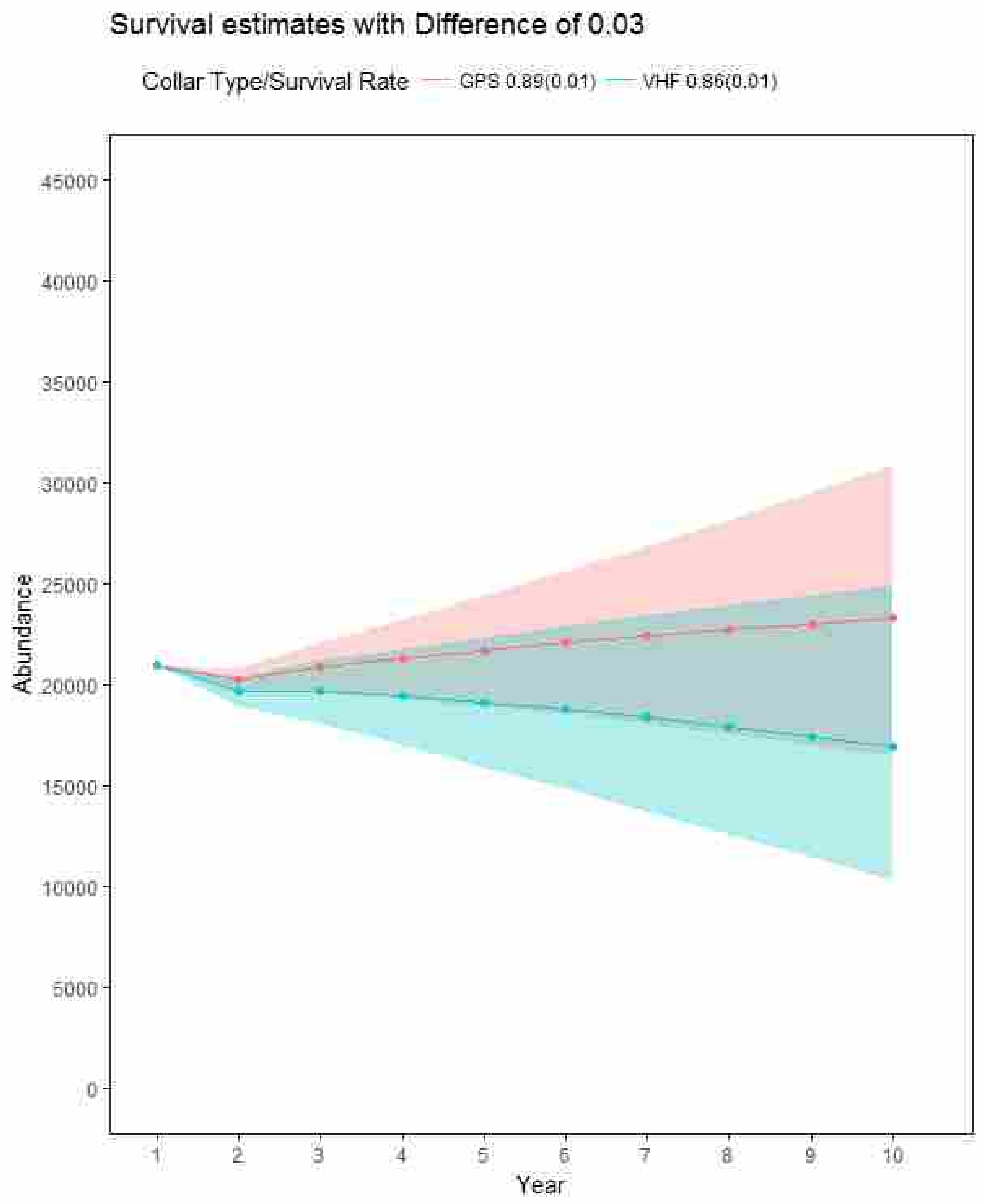


Figure 1-2. Differences in the estimate of abundance over time when hypothetical survival estimates for adult females from global positioning system (GPS) and very high frequency (VHF) collared collar types that are not statistically different are both below the value where the population growth rate is decreasing.

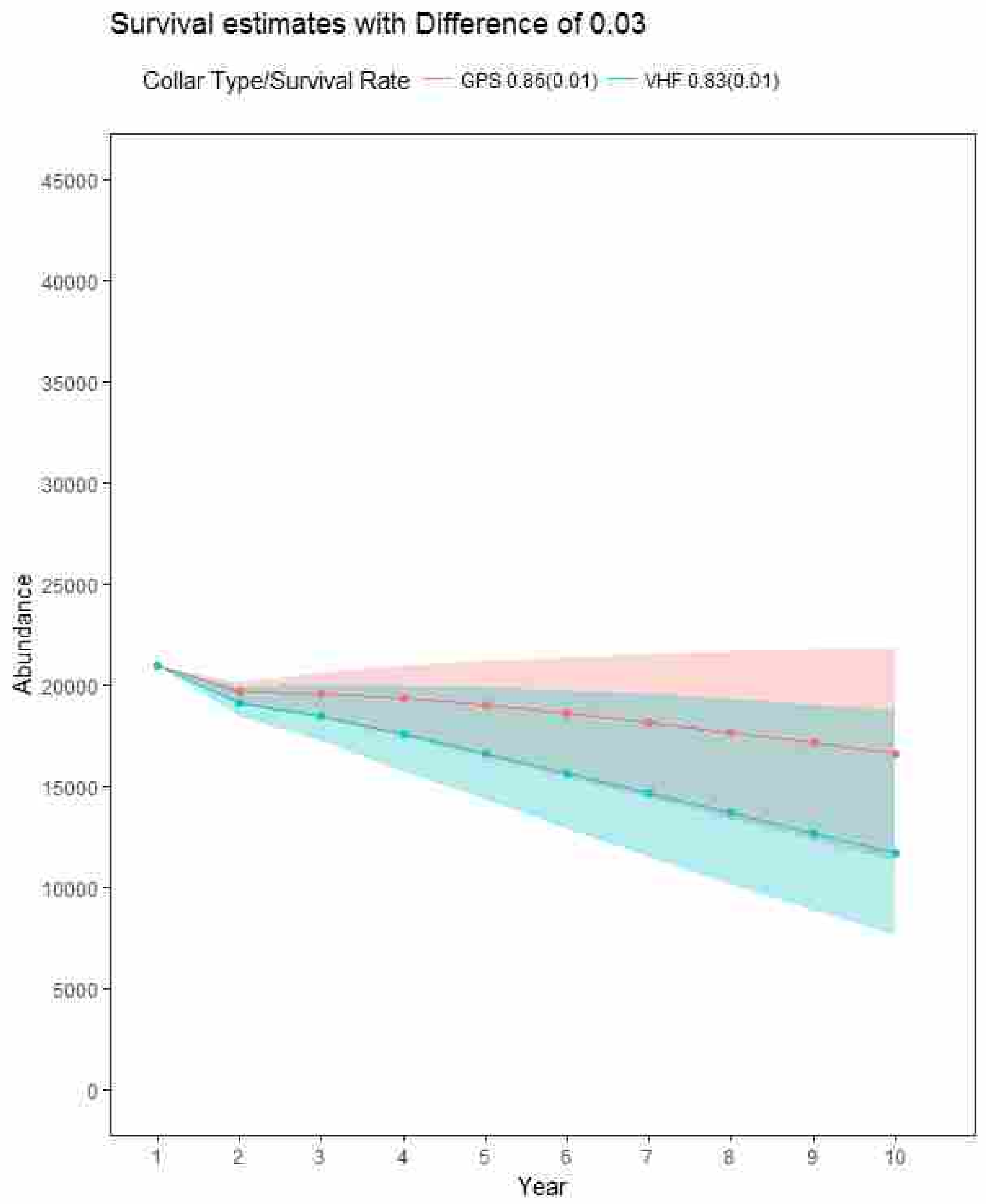


Figure 1-3. Differences in the estimate of abundance over time when hypothetical survival estimates for adult females from global positioning system (GPS) and very high frequency (VHF) collared collar types that are not statistically different are both above the value where the population growth rate is increasing.

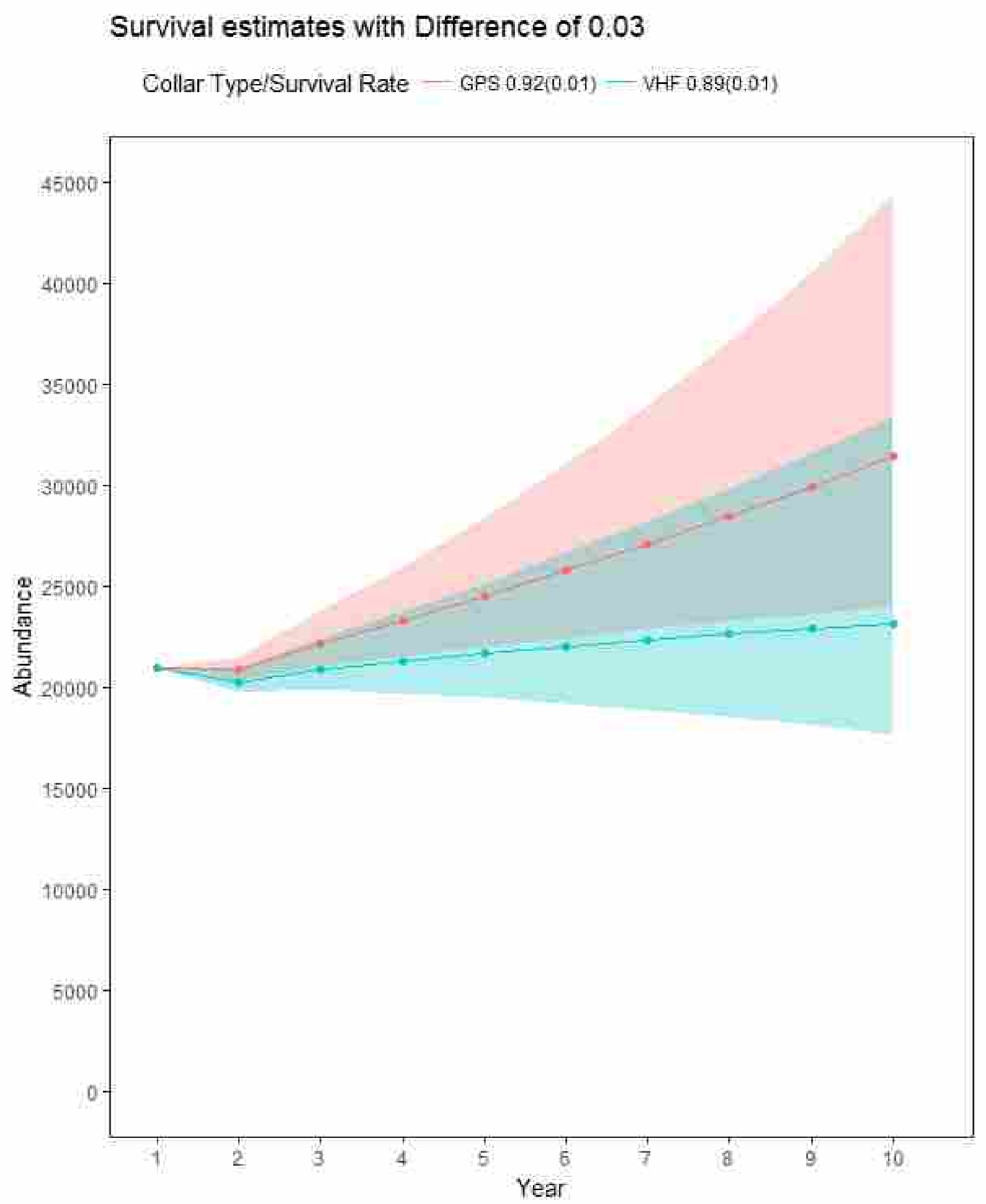


Figure 1-4. Distribution of global positioning system (GPS) and very high frequency (VHF) tracking collars deployed on individuals captured between December 2007 and January 2016 for mule deer population management units (PMU) designated by the Idaho Department of Fish and Game: Weiser-McCall, Central Mountains, Boise River, Smoky-Bennet, South Hills, Bannock, Palisades, Island Park, and Mountain Valley. Bio_Year represents the year when a capture period ended.

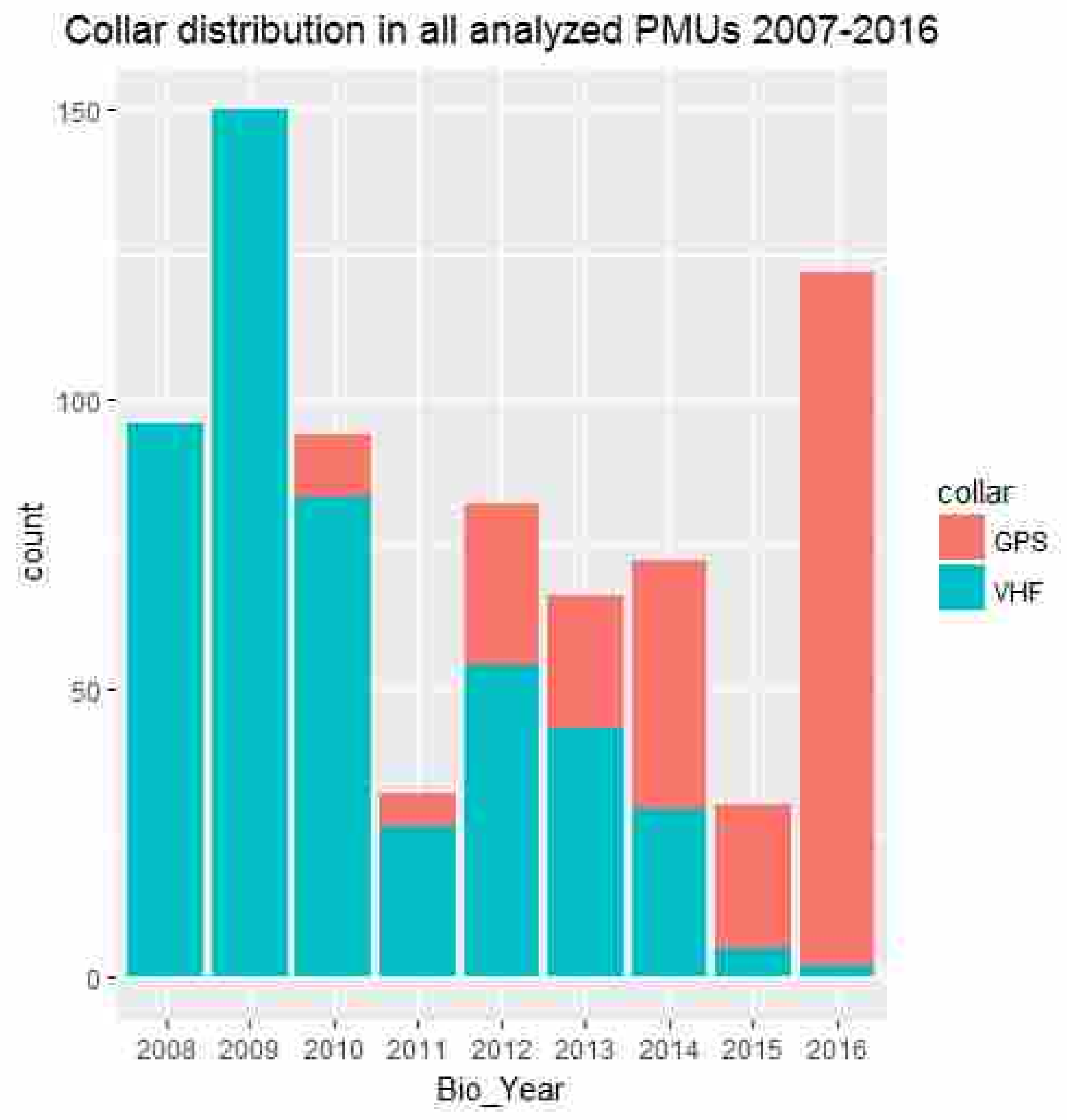


Figure 1-5. Distribution of global positioning system (GPS) and very high frequency (VHF) tracking collars deployed on individuals captured between December 2007 and January 2016 by mule deer population management units (PMU) designated by the Idaho Department of Fish and Game and Bio_Year. Bio_Year represents the year when a capture period ended. Numbers located in the grey box on the left hand side of the figure correspond to the following PMUs: Weiser-McCall (2), Central Mountains (4), Boise River (5), Smoky-Bennet (6), South Hills (8), Bannock (9), Palisades (11), Island Park (12), and Mountain Valley (13).

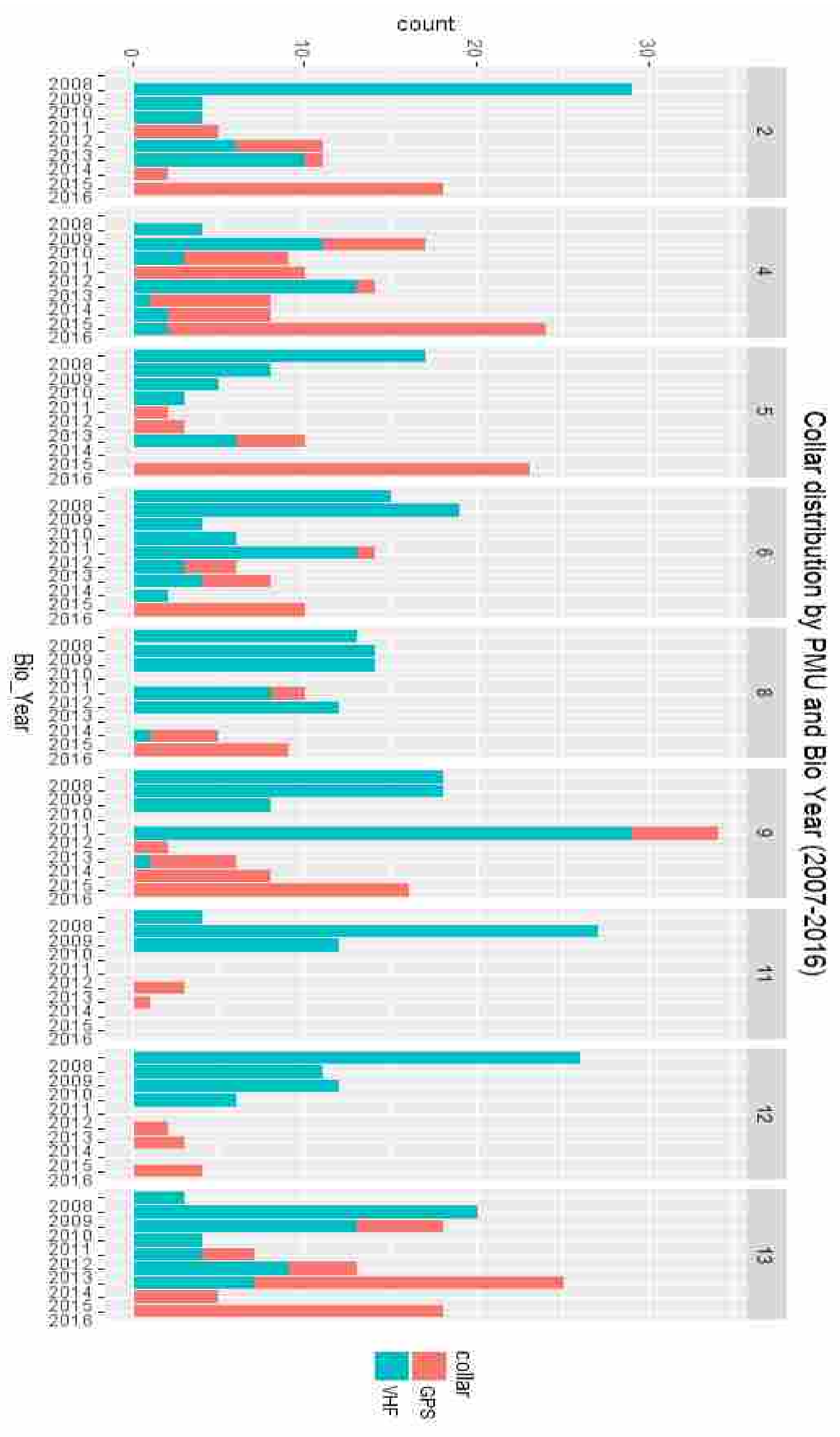

Collar distribution by PMU and Bio Year (2007-2016)

Figure 1-6. Distribution of global positioning system (GPS) and very high frequency (VHF) tracking collars deployed on individuals captured in Bio_Years 2008 - 2011 by mule deer population management units (PMU) designated by the Idaho Department of Fish and Game and Bio_Year. Bio_Year represents the year when a capture period ended. Numbers located in the grey box on the left hand side of the figure correspond to the following PMUs: Weiser-McCall (2), Central Mountains (4), Boise River (5), Smoky-Bennet (6), South Hills (8), Bannock (9), Palisades (11), Island Park (12), and Mountain Valley (13).

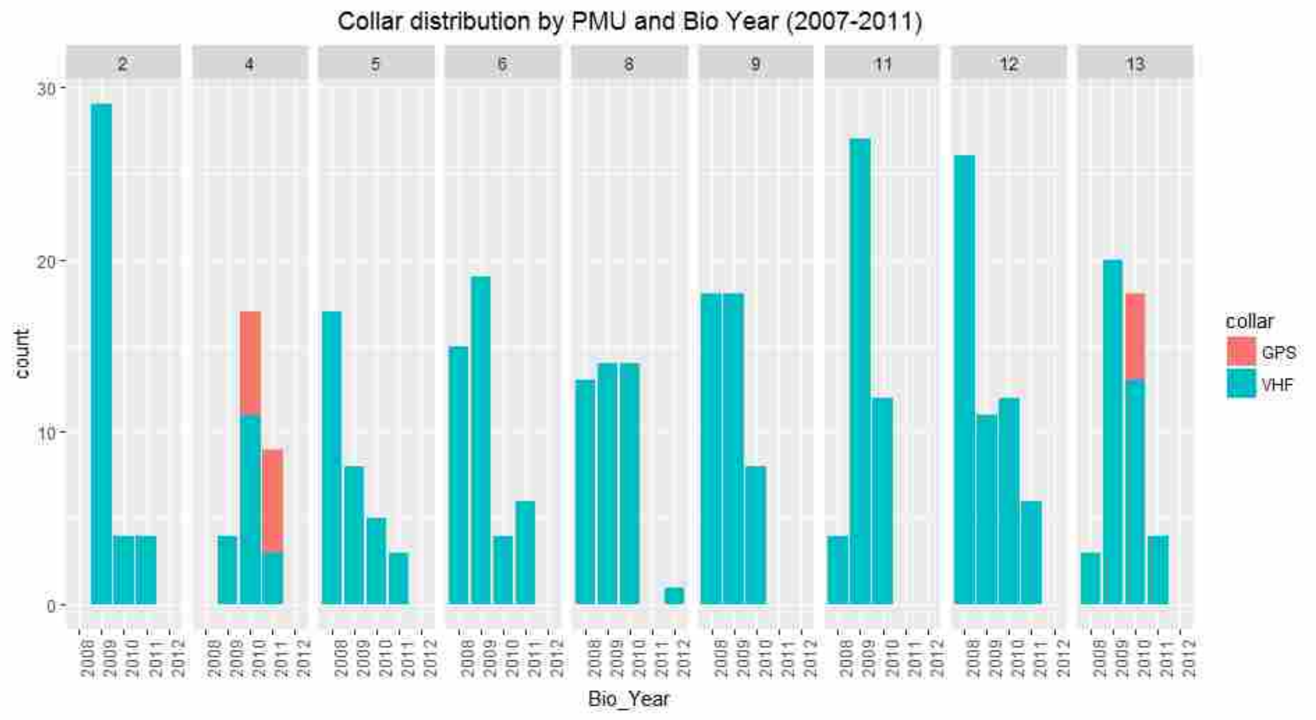


Figure 1-7. Distribution of global positioning system (GPS) and very high frequency (VHF) tracking collars deployed on individuals captured in Bio_Years 2012 - 2014 by mule deer population management units (PMU) designated by the Idaho Department of Fish and Game and Bio_Year. Bio_Year represents the year when a capture period ended. Numbers located in the grey box on the left hand side of the figure correspond to the following PMUs: Weiser-McCall (2), Central Mountains (4), Boise River (5), Smoky-Bennet (6), South Hills (8), Bannock (9), Palisades (11), Island Park (12), and Mountain Valley (13).

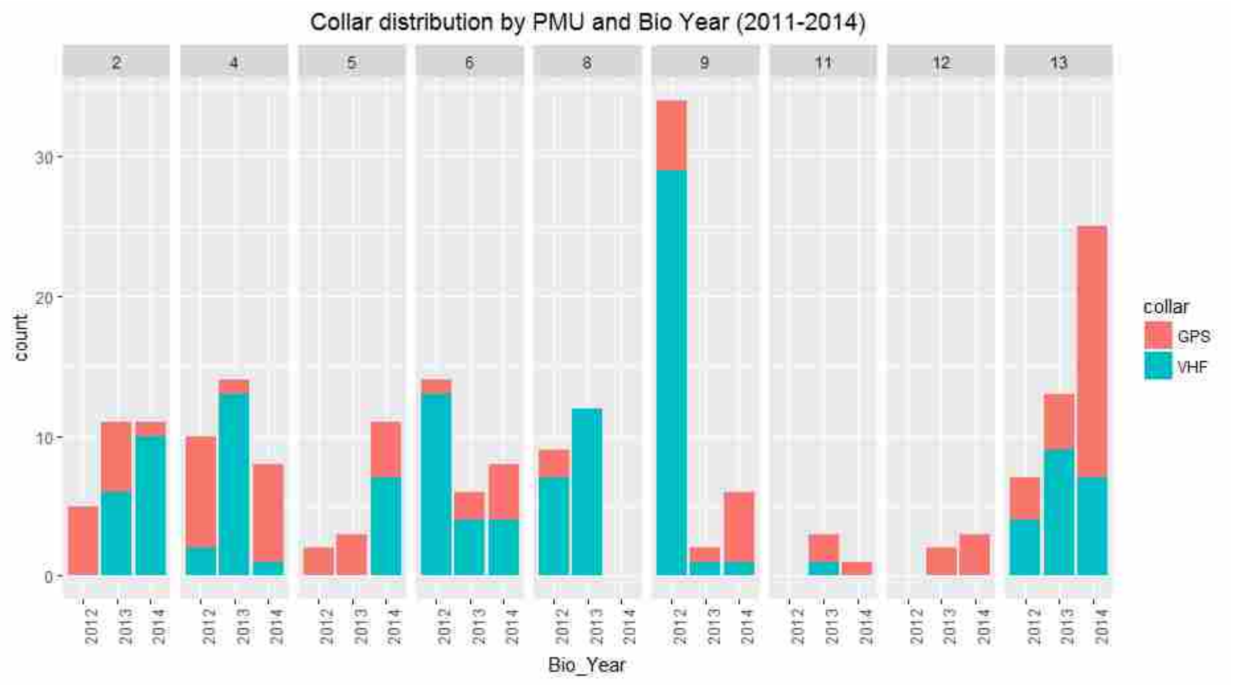
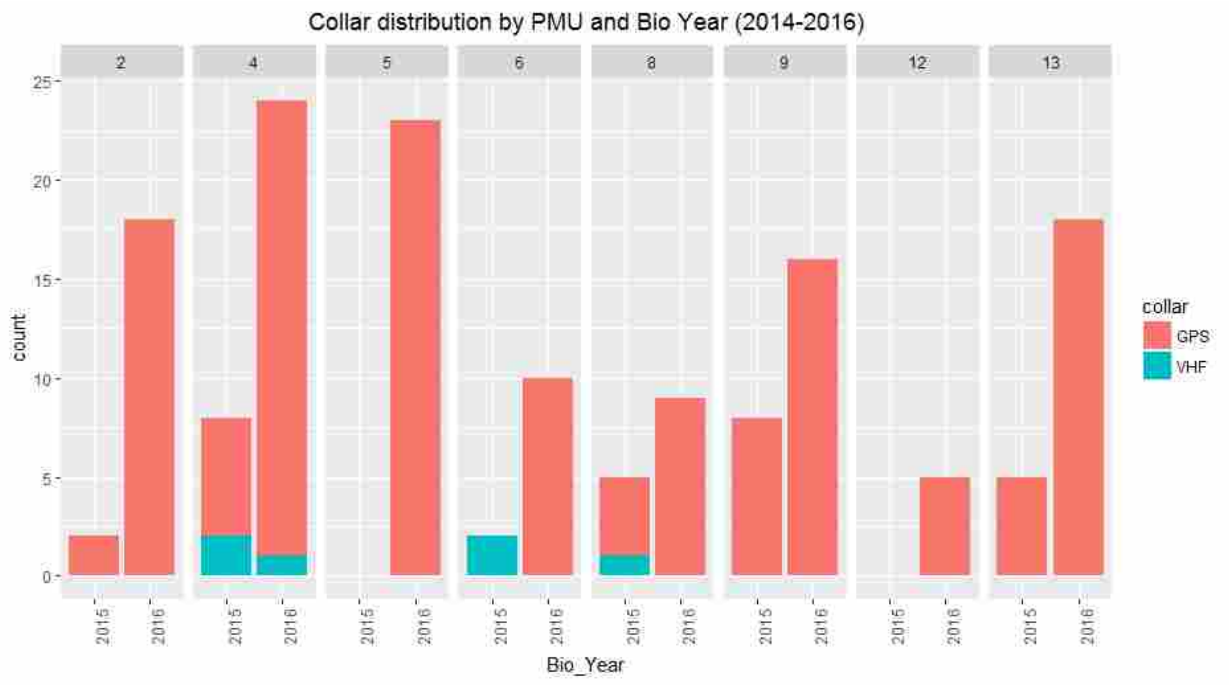


Figure 1-8. Distribution of global positioning system (GPS) and very high frequency (VHF) tracking collars deployed on individuals captured in Bio_Years 2015 - 2016 by mule deer population management units (PMU) designated by the Idaho Department of Fish and Game and Bio_Year. Bio_Year represents the year when a capture period ended. Numbers located in the grey box on the left hand side of the figure correspond to the following PMUs: Weiser-McCall (2), Central Mountains (4), Boise River (5), Smoky-Bennet (6), South Hills (8), Bannock (9), Palisades (11), Island Park (12), and Mountain Valley (13).



CHAPTER 2: ACCOUNTING FOR CLASSIFICATION ERRORS IN WILDLIFE SURVEYS WITHIN A MODELING FRAMEWORK

INTRODUCTION

Accurate knowledge regarding trends in the abundance of wildlife populations provides a foundation for the understanding of wildlife ecology and effective wildlife management. Abundance estimates enable managers and researchers to track the status of a population, supply information on which to base wildlife management decisions, provide a metric to assess the outcome of specific management actions, and quantify the impacts of natural or anthropogenic caused variations in the environment (Terletzky and Koons 2016). Therefore, accurate estimates of abundance are critical information for managing wildlife populations.

Accurate abundance estimates are especially important to the management of harvested populations of game species, such as mule deer (*Odocoileus hemionus*). Common types of data used to estimate the abundance of mule deer include survival, reproduction, harvest, counts of individuals, and herd composition. Herd composition data, used to estimate age and sex ratios, are one of the most commonly collected types of data and are relatively easy to obtain (Bowden et al. 1984, Samuel et al. 1992). Age and sex ratios estimated from these data provide information pertaining to survival, potential reproduction, and recruitment and are a commoncomponent for models that estimate abundance (Bowden et al. 1984, Samuel et al. 1992). However, small changes in the bias and precision of age and sex ratio estimates have the potential to have a large impact on estimates of abundance (Downing et al. 1977).

A variety of factors can influence the bias and precision associated with estimates of age and sex ratios. These include aspects of both mule deer biology and the observation or data collection process. The first way in which the attributes of mule deer biology and data collection can potentially impact the quality of herd composition data is through the group formation dynamics of mule deer and survey timing. Herd composition data is gathered post breeding season for the populations in this study and immediately following the breeding season adult males split off from the mixed sex groups of the breeding season to inhabit the peripheries of the seasonal range (Geist 1981). Therefore, adult males are typically not associated with larger, more easily detectable groups of females, juveniles, and sub-adult males making them more difficult to find and observe during this period (Samuel et al. 1987, Samuel et al. 1992). A survey conducted during this period to collect herd composition data is likely biased low in the adult male category due to the difficulty in detecting this age and sex class. However, by mid- to late winter adult males rejoin the larger mixed sex groups (Geist 1981, Lingle 2003) thereby increasing their detection probability (Samuel et al. 1987, Samuel et al. 1992). Unfortunately, the annual casting of antlers by adult males presents a new issue impacting the quality of herd composition data during the winter.

A second interaction between mule deer biology and data collection that can potentially impact the quality of herd composition data involves the correct classification of individuals according to their age and sex. As an extreme example, if aerial surveys take place after the casting of antlers begins, then the lack of an easily visible indicator of sex and age can result in the misclassification of individuals according to their age and sex. However, classification errors can happen at any time. Classification error, the assignment of an individual to the wrong age or sex class, potentially skews the numbers of individuals in all age and sex classes leading to erroneous estimates of age and sex ratios and consequently inaccurate estimates of abundance.

The problem of misclassification is widespread throughout wildlife surveys and is likely present in most surveys to varying degrees (Smith and McDonald 2002, Simons et al. 2007, McClintock et al. 2015). There are multiple studies showing that misclassification occurs even when the observer is an experienced professional biologist (Alldredge et al. 2008, McClintock et al. 2010, McClintock et al. 2015). This is likely due to the classification process involving the simultaneous assessment of multiple traits in a short period of time (McClintock et al. 2015). Since the classification of individuals into the wrong age or sex class will always result in some amount of bias in the estimate of age and sex ratios it is important to account for classification error to accurately estimate the precision of age and sex ratio estimates. Bias in the age and sex ratio estimates have the potential to influence the accuracy of abundance estimates generated using these ratios especially when the precision associated with the ratio estimate is high. Therefore, it is important to examine how the violation of common assumptions made about ratio data influence the accurate estimates of the precision associated with age and sex ratios so that small amounts of bias present due to classification error will not have an increased impact of abundance estimates generated using these data.

There are several assumptions commonly made about the occurrence of classification errors during the collection of wildlife survey data. A frequent assumption is that no misclassification occurs or that it occurs at such small rates that it has little influence on the data collected. A second assumption in many wildlife surveys is that misclassification rates are uniform within and across categories and space (Conn et al. 2013). A third commonly made assumption is that observer training and strict sampling protocols can eliminate classification error (Miller et al. 2012, McClintock et al. 2015). Based on one or more of these assumptions, a typical way to deal with classification error is to ignore its presence or censor data associated with classification errors (Conn and Diefenbach 2007). When not censored, addressing classification error through the introduction of an “unknown” category when individuals are difficult to classify is a common method. Both censoring and the addition of an unknown category can bias or reduce the precision of the abundance estimates made using this data and influence the inferences made from these results (Conn and Diefenbach 2007, Conn et al. 2013, McClintock et al. 2015). Therefore another approach to addressing the violation of assumptions associated with classification error when collecting ratio data is necessary to improve the abundance information that many wildlife management decisions are based upon.

This project develops a method that addresses the violation of some of the assumptions associated with classification errors made during the collection of herd composition data that can be incorporated in the estimation of age and sex ratios in order to accurately calculate the precision associated with age and sex ratio estimates. In addition, we quantify the amount of sampling effort needed to increase the precision of age and sex ratio estimates. The motivation for both aspects of this study is to improve the precision associated with age and sex ratio estimates which are common data sources used in population models for estimating abundance. Through improving the estimates of precision associated with age and sex ratios we anticipate an overall increase in the quality of the estimates of abundance generated by these data and consequently an increase in the quality of information used by management agencies when making decisions regarding the management and conservation of wildlife species.

METHODS

We developed a joint likelihood to link standard herd composition data to a resampled set of observations to directly estimate misclassification error and composition ratios. The joint likelihood consisted of combining the likelihood of observing a particular combination of classes in a group (Observation likelihood = $log\mathcal{L}_{Obs}$) and the likelihood of classifying an individual into a particular group (Classification likelihood = $log\mathcal{L}_{Class}$). We used the following observation likelihood:

$$log\mathcal{L}_{Obs} = \sum_{i=1}^{n_D}\Big((x_{FD} * log(\phi_{FD})) + (x_{BD} * log(\phi_{BD})) + \big(x_{DD} * log(1-\phi_{FD}-\phi_{BD})\big)\Big) +$$
$$\sum_{i=1}^{n_F}\Big((x_{DF} * log(\phi_{DF})) + (x_{BF} * log(\phi_{BF})) + \big(x_{FF} * log(1-\phi_{DF}-\phi_{BF})\big)\Big) +$$
$$\sum_{i=1}^{n_B}\Big((x_{DB} * log(\phi_{DB})) + (x_{FB} * log(\phi_{FB})) + \big(x_{BB} * log(1-\phi_{DB}-\phi_{FB})\big)\Big) \quad \text{(Equ. 1)}$$

The classification likelihood was as follows:

$$log\mathcal{L}_{Class} = \sum_{i=1}^{N_D} d_D * log(((1-\phi_{FD}-\phi_{BD}) * \theta_D) + (\phi_{DF} * \theta_F) + (\phi_{DB} * \theta_B)) +$$
$$\sum_{i=1}^{N_F} d_F * log\Big(\big((1-\phi_{DF}-\phi_{BF}) * \theta_F\big) + (\phi_{FD} * \theta_D) + (\phi_{FB} * \theta_B)\Big) + \sum_{i=1}^{N_B} d_B *$$
$$log(((1-\phi_{DB}-\phi_{FB}) * \theta_B) + (\phi_{BD} * \theta_D) + (\phi_{BF} * \theta_F)) \quad \text{(Equ. 2)}$$

Where n is the number of individuals observed in a particular class, N is the upper limit of the summation of the number of individuals observed in a particular class, x is count data from resampling, d are the number of individuals observed from a particular class, D represents the adult female class (Does), F represents the juvenile class (Fawns), B represents the adult male class (Bucks), θ proportion of a particular age or sex class in the total population, ϕ represents the probability of making a particular observation, subscripts with 2 letters represent the same probabilities as in the classification matrix. We estimated the maximum likelihood using the "L-BFGS-B" method in the optim function from the base R package stats (R Core Team 2013).

We tested our proposed model through simulation based on observed data. We used data collected by Idaho Department of Fish and Game (IDFG) on the mule deer population in Idaho to simulate populations to investigate the effects of classification error on the precision of age and sex ratio estimates. The elements used to simulate the mule deer population incorporate average values based on data collected by IDFG when conducting aerial surveys in winter and included the average size of all observed groups of deer, the average number of groups observed in a population management unit (PMU), and the average proportion of each of the 3 population classes observed. The 3 population classes in which an individual could be classified included juvenile (fawn), adult female (doe), or adult male (buck). The average group size observed during the aerial survey was 5 individuals and the average number of groups observed within a PMU was 200. The average population consisted of 50% adult females, 30% juveniles, and 20% adult males. The simulated population had group sizes drawn randomly from a Poisson distribution with a mean of 5. We assigned the number of individuals in each age and sex class within a single group using a random draw from a multinomial distribution where the probability of being in a specific class reflected the proportion of that class in the population.

Observations of the simulated population where created using a random draw from a multinomial distribution where the probability of being classified reflected all possible ways in which an individual from a single population class could be classified. All possible classifications are included in the following matrix where the first letter indicates the observed class of an individual and the second letter indicates the true class of the individual.

$$\begin{bmatrix} DD & DF & DB \\ FD & FF & FB \\ BD & BF & BB \end{bmatrix}$$

DD is the probability that an individual is classified as an adult female given it is an adult female, *FD* is the probability that an individual is classified as a juvenile given it is an adult female, *BD* is the probability that an individual is classified as an adult male given it is an adult female, *FF* is the probability that an individual is classified as a juvenile given it is a juvenile, *DF* is the probability that an individual is classified as an adult female given it is a juvenile, *BF* is the probability that an individual is classified as an adult male given it is a juvenile, *BB* is the probability that an individual is classified as an adult male given it is an adult male, *DB* is the probability that an individual is classified as an adult female given it is an adult male, and *FB* is the probability that an individual is classified as a juvenile given it is an adult male. We assigned a probability to each of these possible classification events which represent how physical attributes, survey timing, and observer experience could impact the correct classification of an individual. We created multiple matrices that reflected a range changes in the classification probability which represents various combinations of changes in physical attributes, survey timing, and observer experience that could occur during an aerial survey.

When classification error is present, there are differences between simulated or known age and sex ratios and observed age and sex ratios. We calculated the mean differences between ratios from the known and observed data sets using 100 simulated populations and the observations from the simulated populations. The classification probabilities used to demonstrate the differences in the known and observed age and sex ratios were as follows: the probability of correctly classifying an adult male as an adult male (BB) was 0.9, the probability of correctly classifying a juvenile as a juvenile (FF) was 0.8, and the probability of correctly classifying an adult female as an adult female (DD) ranged from 0.75 to 0.95 in 0.05 increments. We calculated the difference between the known and observed ratios by subtracting the observed ratio from the known ratio and taking the mean of the differences from the 100 simulated data sets.

Both positive and negative errors in the age and sex ratios influence estimates of abundance. We selected a 5 year data set from the Boise River mule deer PMU to demonstrate the impact of these errors on the estimate abundance for each year in the 5 year period. We varied the levels of ratio error from +20 to -20 individuals in increments of 10 to determine the impact of ratio error on estimates of abundance. . Ratio errors were included singly (age ratio error or sex ratio error) or in combination (age ratio error and sex ratio error) for the different abundance estimate scenarios. We did not change the measures of precision for the ratio estimates for the different abundance estimation scenarios. We implemented the current integrated population model used by IDFG from PopR to estimate abundance (Nowak et al. 2018). This IPM uses both age and sex ratio estimates as data to estimate abundance. Recruitment values used in the IPM to estimate the number of female and male fawns are the estimates of the annual age ratio. Sex ratio estimates provide information on adult male survival to the IPM and allow the IPM to estimate the number of males and females from the total count of abundance data which contains no information on the sex of the individual counted. The resulting estimates of abundance show the impact of the magnitude of the ratio error on the estimate of abundance and the differences in the impact of age ratio error versus sex ratio error.

Finally, we investigated the amount of effort required to change the precision of age and sex ratios through repeated observations of groups of mule deer. In this analysis, the group was the unit of observation as suggested for a social species like mule deer (Samuel et al. 1992). Therefore, we used the number of groups observed more than once as the measure of additional sampling effort through which it should be possible to increase the precision of age and sex ratio estimates in the presence of classification error. The number of groups resampled ranged from 25 to 200 groups representing a 12.5% to 100% increase in the amount of additional sampling effort. We tested this range of resampling effort under several different classification probability scenarios. Every scenario analyzed included the probability of correctly classifying an adult female as an adult female (DD) which ranged from 0.75 to 0.95 in 0.05 increments. Scenarios were tested in which the probability of correctly classifying a juvenile as a juvenile (FF) was held constant at 0.8 and the probability of correctly classifying an adult male as an adult male (BB) ranged from 0.95 to 0.85 in 0.05 increments to investigate how changes to the probability of correctly classifying an adult male would influence the amount of effort needed to increase age and sex ratio precision. The other set of scenarios tested how changes to the probability of correctly classifying a juvenile would influence the amount of effort needed to increase age and sex ratio precision. In these scenarios the probability of correctly classifying an adult male as an adult male (BB) was held constant at 0.9 and the probability of correctly classifying a juvenile as a juvenile (FF) ranged from 0.9 to 0.7 in 0.1 increments. We employed the observation/classification likelihood previously described to estimate the age and sex ratios and their precision. .

RESULTS

Figure 2-1 shows the differences in the observed age (fawn:doe) ratio and the true fawn:doe ratio represented by the horizontal red line. When the probability of observing a doe correctly was equal to or greater than 0.9 the observed fawn:doe ratio was higher than the true fawn:doe ratio. When the probability of observing a doe correctly was equal to or less than 0.85 the observed fawn:doe ratio was lower than the true fawn:doe ratio. Figure 2-2 shows the differences in the observed sex (buck:doe) ratio and the true buck:doe ratio represented by the horizontal red line. When the probability of observing a doe correctly was equal to or greater than 0.95 the observed buck:doe ratio was higher than the true buck:doe ratio. When the probability of observing a doe correctly was equal to or less than 0.9 the observed buck:doe ratio was lower than the true buck:doe ratio. The interaction between the probabilities of correctly classifying each class relative to one another and the proportion of each class in the population influenced the magnitude and direction of the error in the observed age and sex ratios.

Age and sex ratios can influence estimates of abundance generated by population models. Figures 2-3, 2-4, and 2-5 demonstrate the amount of influence age and sex ratios can have on abundance estimates. Abundance estimates did not change significantly for the first year in any of the scenarios. Figure 3 shows abundance estimates when only the fawn:doe ratio was manipulated. Estimates for years 1 and 2 remained similar but differences in abundances estimates appeared in the following 3 years, especially years 3 and 5. Figure 4 shows abundance estimates when only the buck:doe ratio was manipulated. There were significant differences in the abundance estimates for years 2 to 5. 9000 to 10000 individuals separate the largest and smallest abundance estimates for years 2, 4, and 5. Figure 2-5 shows abundance estimates when both the fawn:doe ratio and the buck:doe ratio was manipulated. There were significant differences in the abundance estimates for years 2 to 5. The differences between the highest and lowest abundance estimates for years 2 through 5 ranged from 6000 to 12000 individuals. These results indicate that errors in age and sex ratios influence the estimates of abundance generated by population models and can compound over time.

Resampling of more groups increase in the precision of age and sex ratio estimates. The amount of groups that need to be resampled to increase the precision of age and sex ratio data under different classification probability scenarios are shown in Figures 2-6, 2-7, 2-8, and 2-9. We quantified resampling effort by the number of groups resampled and which ranged from 25 to 200 groups. Overall, precision increased quickly as the number of groups re-sampled increased from 25 to 100 groups but was not significantly different when more than 100 groups were re-sampled (Figures 2-6, 2-7, 2-8, and 2-9).

Figures 2-6 and 2-7 show changes to the precision of the fawn:doe ratio (Figure 2-6) and the buck:doe ratio (Figure 2-7) when the probability of classifying a buck as a buck is held constant at 0.9 and the probabilities of classifying a doe as a doe range from 0.75 to 0.95 and the probabilities of classifying a fawn as a fawn range from 0.7 to 0.9. In the scenarios testing the influence of changes to the classification error associated with classifying fawns on the amount resampling effort needed to increase precision, the precision of the fawn:doe ratio ranged from 3.9 to 24.11. The fawn:doe ratio was the most precise (SD = 3.9) when DD = 0.9, BB = 0.9, and FF = 0.9 and 200 groups were resampled. The fawn:doe ratio was the least precise (SD = 24.11) when DD = 0.75, BB = 0.9, and FF = 0.7 and 25 groups were resampled. The precision of the buck:doe ratio ranged from 2.88 to 13.11. The buck:doe ratio was the most precise (SD = 2.88) when DD = 0.75, BB = 0.9, and FF = 0.8 and 200 groups were resampled. The buck:doe ratio was the least precise (SD = 13.11) when DD = 0.75, BB = 0.9, and FF = 0.9 and 25 groups were resampled.

Figures 2-8 and 2-9 show changes to the precision of the fawn:doe ratio (Figure 2-8) and the buck:doe ratio (Figure 2-9) when the probability of classifying a fawn as a fawn is held constant at 0.8 and the probabilities of classifying a doe as a doe and a buck as a buck change. In the scenarios testing the influence of changes to the classification error associated with classifying bucks on the amount resampling effort needed to increase precision, the precision of the fawn:doe ratio ranged from 3.96 to 17.33. The fawn:doe ratio was the most precise (SD = 3.96) when DD = 0.85, BB = 0.9, and FF = 0.8 and 200 groups were resampled. The fawn:doe ratio was the least precise (SD = 17.33) when DD = 0.8, BB = 0.95, and FF = 0.8 and 25 groups were resampled. The precision of the buck:doe ratio ranged from 2.96 to 10.73. The buck:doe ratio was the most precise (SD = 2.96) when DD = 0.95, BB = 0.85, and FF = 0.8 and 200 groups were resampled. The buck:doe ratio was the least precise (SD = 10.73) when DD = 0.75, BB = 0.85, and FF = 0.8 and 25 groups were resampled.

DISCUSSION

Estimates of age and sex ratios provide information about wildlife populations regarding survival, reproduction, and population structure. Information about survival, reproduction, and population structure influence the accuracy, bias and precision, of estimates of abundance generated by population models. Since some level of classification error is likely present in all animal surveys it must be accounted for when estimating abundance (Smith and McDonald 2002, Simons et al. 2007) or it can bias or reduce the precision of abundance estimates (Conn et al. 2013, McClintock et al. 2015). The amount of bias and level of precision in age and sex ratio estimates influence the accuracy of abundance estimates especially over multiple time steps (Figures 2-3, 2-4, and 2-5). This study develops a multinomial likelihood for classification probabilities to account for the presence of classification error in herd composition data. Using this likelihood when estimating age and sex ratios should lead to more accurate estimates of the precision of the age and sex ratio estimates and consequently more accurate abundance estimates on which to base wildlife management decisions. We also quantify the amount of resampling effort needed to improve the estimate of precision associated with age and sex ratio estimates during data collection.

Multiple studies have shown that despite high levels of training or experience it is rare that all individuals are classified correctly with regards to age and sex during a wildlife survey (Alldredge et al. 2008, McClintock et al. 2010, McClintock et al. 2015). Therefore the probability of correctly classifying an individual in any age or sex class is typically less than 1 and consequently the resulting estimates of age and sex ratios contain some amount of bias.

However, the type of ratio estimate and the magnitude of the bias in relation to the precision associated with that ratio estimate influence the magnitude of the impact that classification error has on the accuracy of the subsequent estimate of abundance.

Similar amounts of bias in the estimates of age and sex ratios from classification error varied in the magnitude of their impact on abundance estimates. Typically, bias in the age ratio had a smaller influence on the estimate of abundance (Figure 2-3) relative to the impact of bias in the sex ratio (Figure 2-4). The smaller impact of the age ratio is likely due to the variability in juvenile survival and recruitment in mule deer and other ungulate species that have a similar life history (Gaillard et al. 1998). The relatively low elasticity values of reproductive rates as compared to the elasticity of the survival rates of female adults and the elasticity of the survival rates of female fawns when the sex ratio is less than 0.7, provide another reason why bias in age ratios have a reduced impact on the estimate of abundance (Table 2-1). Due to the use of age ratios as an index for reproduction, the smaller influence of bias in the age ratio, as compared to the sex ratio, on the estimate of abundance is not surprising. However, the increased magnitude of the impact of bias in the estimate of sex ratios on estimates of abundance is unexpected (Figures 2-4 and 2-5).

The larger relative impact of bias in the estimate of the sex ratio on the estimate of abundance is likely the result of the elasticity values associated with vital rates that sex ratio estimates contain information about. The sex ratio contains information regarding the survival of all age and sex classes in the population. Consequently, bias in the sex ratio inaccurately represents information regarding the survival rates of either or both sexes of adults and juveniles in the population. In every population age and sex ratio scenario investigated using the population matrix from this study, the survival of adult females has a higher elasticity value than the elasticity value of any reproductive rate in the same scenario (Table 2-1). When we added the elasticity values across all survival rates in a single scenario the cumulative elasticity (0.92 – 0.68) is always greater than the cumulative elasticity of the reproductive rates (0.32 – 0.08; Table 2-1). Depending on the combination of the age and sex ratios, there are occasions when the elasticity of a reproductive rate is greater than the elasticity of a particular survival rate but the occurrence of this phenomenon fluctuated as the age and sex ratios in the population changed and did not result in a consistent pattern (Table 2-1, Figures 2-10 and 2-11). Overall our results demonstrate that sex ratio estimates, which are most closely associated with the vital rates having the highest elasticity values, have the greatest impact on the estimate of abundance. Consequently, bias in the sex ratio estimate has a larger impact on abundance estimates than bias in age ratio estimates.

The differences in the relative magnitude of the impact of bias in age and sex ratios on the estimate of abundance further emphasizes the need for accurate estimates of precision. Accounting for classification error using the multinomial likelihood within the process of estimating age and sex ratios increases the accuracy of the estimates of the precision associated with the age and sex ratios. Better estimates of precision consequently allow the IPM to better compensate for the small levels of bias inherent in ratio estimates due to the presence of classification error during the observation process. The ability of the population model to better compensate for bias as a consequence of better estimates of precision results in improved accuracy of the abundances estimates produced by the IPM.

Ignoring classification error can lead to biased and poorly estimated measures of precision related to age and sex ratio estimates (Figures 2-3, 2-4, and 2-5). Observer training and experience has been shown to improve the probability of correctly classifying individuals during wildlife surveys but even the most highly trained and experienced observers still have a probability of misclassify individuals during the course of a survey (Miller et al. 2012, McClintock et al. 2015). The multinomial likelihood we developed helps account for the classification error that increased observer training and experience does not eliminate. We did this by constructing a multinomial likelihood for classification probabilities that incorporates all of the possible ways to classify an individual during a single observation event. We assigned a probability to each possible type of observation (a female classified as a female, a female classified as a male, etc.). The probability of classification for a particular observation is flexible and based on the identifiability of a specific age or sex class which is determined via experimentation or solicited from expert opinion. For example in mule deer, an adult male with antlers has a higher probability of correct identification than a juvenile male. Adjustment of classification probabilities according to the attributes of the environment in which the survey takes place or the level of experience and training of the observer is another example of the flexibility of the multinomial likelihood. Thus the multinomial likelihood for classification probabilities is capable of incorporating both the attributes that make each age and sex class more or less identifiable and the capabilities of the observer.

The multinomial likelihood for classification probabilities developed by this study addresses the common assumption that classification error does not exist or that is small enough to ignore. This method allows the user to adjust the probability of making each type of observation even at minute levels thereby accounting for small changes in the probability of committing a classification error. In addition, the ability to adjust the probability of making each type of observation eliminates the need for an unknown category. The elimination of the unknown category reduces bias in ratio estimates especially if the individuals assigned to the unknown category are not distributed randomly across age and sex classes as well as increasing the precision of ratio estimates by including all observations (Conn and Cooch 2009). Through the flexible nature of the multinomial likelihood classification method we remove the violation of these assumptions from consideration and decrease their impact on the precision of the age and sex ratio estimates.

The flexibility of the multinomial likelihood classification method also addresses a second common assumption made concerning classification error, that classification error rates are uniform across space and age and sex categories. The ability to adjust the probability of making each type of observation allows the user to assign variable classification probabilities to different age and sex classes. These adjustments reflect not only differences in the probability of correctly identifying a specific age and sex class but can also reflect changes to the attributes of the environment in which the collection of the composition data occurred.

Another benefit of the flexibility of the multinomial likelihood for classification probabilities developed here is the possibility of adapting this methodology to other types of wildlife surveys that are subject to the presence of classification error. Most wildlife surveys whether visually based, like aerial surveys and camera traps, or auditory based, such as bird or amphibian vocalization counts, struggle with the presence of error in classifying all individuals within the appropriate category (Conn et al. 2013). Other methods used to improve the accuracy of these surveys include repeated surveys over a short time period and the use of multiple observers (Pollock and Kendall 1987). We suggest our method using flexible classification probabilities based on observer, environmental, and species attributes can supplement or replace the aforementioned survey methodologies. While this application of the multinomial likelihood for classification probabilities was not investigated during this study it is a potential avenue for the further study and development of wildlife survey methods and techniques.

We also investigated another way to improve the precision of ratio data, through quantifying the amount of resampling effort needed to increase the precision of ratio estimates. Resampling in this case is analogous to using a second observer to resample groups during surveys. The second observer used to resample could be present during the original survey or contribute observations through an analysis of photographic data gathered during the original survey. However, this requires an increase in the amount of resources needed to conduct the resampling effort. Due to the increase in resources needed to resample, it is important to quantify the expected return in increased precision relative to the amount of effort required. This analysis provides information regarding the amount of resampling effort needed to increase the precision of age and sex ratios when surveying mule deer populations. The number of groups of deer that must be resampled is the unit we used to measure resampling effort.

We focused on observed groups as the unit of observation because previous work has shown that it is the smallest sampling unit for social species (Samuel et al. 1992). In addition, group size influences both the probability of observation (Samuel and Pollock 1981, Samuel et al. 1992) and the probability of observing and correctly classifying all members of a group (Fleming and Tracey 2008, Graham and Bell 1989, Clement et al. 2017). Our results suggest that a noticeable increase in the precision of age and sex ratio estimates occurs when the number of groups resampled increases from 25 to 100 groups (Figures 2-6, 2-7, 2-8, and 2-9). Increases in the precision of the age and sex ratios are negligible after resampling greater than 100 groups. Changes in the classification probabilities have less influence on the precision of the age and sex ratio estimates than the number of groups resampled for the scenarios tested. There is one exception to this pattern. The increase in the classification probability of identifying a fawn as a fawn had a noticeable impact on the fawn:doe ratio when the probability of classifying a doe as a doe is low (0.75) and the number of groups resampled is small (25). This exception is likely the product of a combination of factors including both classes that make up the fawn:doe ratio having classification error, the reference class (does) having a largest amount of classification error tested, and a small amount of resampling effort associated with these scenarios. We used the classification probabilities associated with different age and sex classes in the population to estimate the amount error in observed age and sex ratios as the number of groups resampled increased. The quantification of the increase in the precision of the ratio data estimates as the number of groups to be resampled changes under different classification probability scenarios will allow managers and researchers to weigh the cost of increasing sampling effort against the increases in the precision of ratio data when determining how to best allocate monitoring resources based on their specific precision needs.

Age and sex ratio estimates from population composition data is a common type of data gathered about wildlife populations but its accuracy is often subject to question due to the presence of classification error. The probabilities of correctly classifying an individual in the appropriate age or sex class impacts the amount of bias in ratio estimates and the bias in ratio estimates often leads to biased estimates of abundance. This is especially true if the precision of the ratio estimates is not carefully calculated. We developed a multinomial likelihood for classification probabilities that takes into account the ubiquitous presence of classification error in wildlife surveys and specifically accounts for classification error. The likelihood is flexible enough to be adapted to incorporate the various factors which influence classification error. The estimates of various population classes using the multinomial likelihood are less biased and have a more accurate measure of precision in the presence of classification error. We also quantified the amount resampling effort, based on the number of groups of mule deer resampled, needed to improve precision in ratio estimates to a level acceptable for different management and conservation purposes. While the results of the resampling portion of this analysis are specific to mule deer the sampling quantification methodology employed here is applicable to other species surveyed in a similar manner and having similar group living dynamics. Finally, classification error is likely present in all wildlife survey data and this study develops methods that address this source of error. This study provides tools to address classification error and improve the quality of the estimate of precision associated with age and sex ratio estimates. Through the use of these methods we expect the quality of information used to make management and conservations decisions regarding wildlife populations to improve.

LITERATURE CITED

Alldredge, M.W., Pacifici, K., Simons, T.R. and Pollock, K.H., 2008. A novel field evaluation of the effectiveness of distance and independent observer sampling to estimate aural avian detection probabilities. Journal of Applied Ecology, 45(5), pp.1349-1356.

Anderson, A. E. 1981. Morphological and physiological characteristics. Pp. 27-97, in Mule and black-tailed deer of North America (O. C. Wallmo, ed.). Univ. Nebraska Press, Lincoln, xvii + 605 pp.

Anderson, A.E., Bowden, D.C. and Medin, E., 1990. Indexing the annual fat cycle in a mule deer population. *The Journal of Wildlife Management*, pp.550-556.

Bandy, P.J., Cowan, I.M. and Wood, A.J., 1970. Comparative growth in four races of black-tailed deer (Odocoileus hemionus). Part I. Growth in body weight. *Canadian Journal of Zoology*, *48*(6), pp.1401-1410.

Bowden, D.C., Anderson, A.E. and Medin, E., 1984. Sampling plans for mule deer sex and age ratios. *The Journal of wildlife management*, pp.500-509.

Caughley, G., 1974. Interpretation of age ratios. *The Journal of Wildlife Management*, pp.557-562.

Clement, M.J., Converse, S.J. and Royle, J.A., 2017. Accounting for imperfect detection of groups and individuals when estimating abundance. *Ecology and evolution*, *7*(18), pp.7304-7310.

Conn, P.B., McClintock, B.T., Cameron, M.F., Johnson, D.S., Moreland, E.E. and Boveng, P.L., 2013. Accommodating species identification errors in transect surveys. *Ecology*, *94*(11), pp.2607-2618.

Conn, P.B. and Cooch, E.G., 2009. Multistate capture–recapture analysis under imperfect state observation: an application to disease models. *Journal of Applied Ecology*, *46*(2), pp.486-492.

Conn, P.B. and Diefenbach, D.R., 2007. Adjusting age and stage distributions for misclassification errors. *Ecology*, *88*(8), pp.1977-1983.

Downing, R.L., Michael, E.D. and Poux Jr, R.J., 1977. Accuracy of sex and age ratio counts of white-tailed deer. *The Journal of Wildlife Management*, pp.709-714.

Fleming, P.J. and Tracey, J.P., 2008. Some human, aircraft and animal factors affecting aerial surveys: how to enumerate animals from the air. Wildlife Research, 35(4), pp.258-267.

Gaillard, J.M., Festa-Bianchet, M. and Yoccoz, N.G., 1998. Population dynamics of large herbivores: variable recruitment with constant adult survival. *Trends in Ecology & Evolution*, *13*(2), pp.58-63.

Geist, V. 1981. Behavior: adaptive strategies in mule deer. Pp. 157-224, in Mule and black-tailed deer of North America (O. C. Wallmo, ed.). Univ. Nebraska Press, Lincoln, xvii + 605 pp.

Goss, R.J., 1983. *Deer antlers: regeneration, function and evolution*. Academic Press.

Graham, A. and Bell, R., 1989. Investigating observer bias in aerial survey by simultaneous double-counts. The Journal of wildlife management, 53(4), pp.1009-1016.

Lingle, S., 2003. Group composition and cohesion in sympatric white-tailed deer and mule deer. Canadian Journal of Zoology, 81(7), pp.1119-1130.

Mackie, R.J., Pac, D. F., Hamlin, K. L., and Dusek, G. L., 1998. *Ecology and management of mule deer and white-tailed deer in Montana*. Montana Fish, Wildlife, and Parks, Wildlife Division.

Mcclintock, B.T., Bailey, L.L., Pollock, K.H. and Simons, T.R., 2010. Experimental investigation of observation error in anuran call surveys. *The Journal of Wildlife Management*, *74*(8), pp.1882-1893.

McClintock, B.T., Moreland, E.E., London, J.M., Dahle, S.P., Brady, G.M., Richmond, E.L., Yano, K.M. and Boveng, P.L., 2015. Quantitative assessment of species identification in aerial transect surveys for ice-associated seals. *Marine Mammal Science*, *31*(3), pp.1057-1076.

McCullough, D.R., 1994. In My Experience: What Do Herd Composition Counts Tell Us?. *Wildlife Society Bulletin (1973-2006)*, *22*(2), pp.295-300.

Miller, D.A., Talley, B.L., Lips, K.R. and Grant, E.H.C., 2012. Estimating patterns and drivers of infection prevalence and intensity when detection is imperfect and sampling error occurs. *Methods in Ecology and Evolution*, *3*(5), pp.850-859.

Nowak, J.J., Lukacs, P.M., Hurley, M.A., Lindbloom, A.J., Robling, K.A., Gude, J.A. and Robinson, H., 2018. Customized software to streamline routine analyses for wildlife management. *Wildlife Society Bulletin*, *42*(1), pp.144-149.

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

Robinette, W.L., Baer, C.H., Pillmore, R.E. and Knittle, C.E., 1973. Effects of nutritional change on captive mule deer. *The Journal of Wildlife Management*, pp.312-326.

Samuel, M.D. and Pollock, K.H., 1981. Correction of visibility bias in aerial surveys where animals occur in groups. The Journal of Wildlife Management, 45(4), pp.993-997.

Samuel, M.D., Garton, E.O., Schlegel, M.W. and Carson, R.G., 1987. Visibility bias during aerial surveys of elk in northcentral Idaho. The Journal of wildlife management, pp.622-630.

Samuel, M.D., Steinhorst, R.K., Garton, E.O. and Unsworth, J.W., 1992. Estimation of wildlife population ratios incorporating survey design and visibility bias. *The Journal of wildlife management*, pp.718-725.

Simons, T.R., Pollock, K.H., Wettroth, J.M., Alldredge, M.W., Pacifici, K. and Brewster, J., 2009. Sources of measurement error, misclassification error, and bias in auditory avian point count data. In *Modeling demographic processes in marked populations* (pp. 237-254). Springer, Boston, MA.

Smith, B.L. and McDonald, T.L., 2002. Criteria to improve age classification of antlerless elk. *Wildlife Society Bulletin*, pp.200-207.

Sorin, A.B., 2004. Paternity assignment for white-tailed deer (Odocoileus virginianus): mating across age classes and multiple paternity. Journal of Mammalogy, 85(2), pp.356-362.

Terletzky, P.A. and Koons, D.N., 2016. Estimating ungulate abundance while accounting for multiple sources of observation error. Wildlife Society Bulletin, 40(3), pp.525-536.

TABLES

Table 2-1. Elasticities of vital rates used in population matrix.

Age Ratio	Sex Ratio	Adult Female Reproduction, Fawn Female	Adult Female Reproduction, Fawn Male	Adult Male Reproduction, Fawn Female	Adult Male Reproduction, Fawn Male	Adult Female Survival	Fawn Female Survival	Adult Male Survival	Fawn Male Survival
0.2	0.1	0.0768	0.0017	0.0017	0.0000	0.8395	0.0784	0.0002	0.0017
0.4	0.1	0.1244	0.0047	0.0047	0.0002	0.7316	0.1291	0.0005	0.0049
0.6	0.1	0.1566	0.0079	0.0079	0.0004	0.6535	0.1646	0.0007	0.0083
0.8	0.1	0.1799	0.0110	0.0110	0.0007	0.5941	0.1908	0.0010	0.0116
1	0.1	0.1974	0.0137	0.0137	0.0010	0.5472	0.2111	0.0011	0.0147
0.2	0.3	0.0747	0.0064	0.0064	0.0006	0.8211	0.0812	0.0026	0.0070
0.4	0.3	0.1154	0.0173	0.0173	0.0026	0.6881	0.1327	0.0067	0.0199
0.6	0.3	0.1385	0.0276	0.0276	0.0055	0.5915	0.1661	0.0101	0.0331
0.8	0.3	0.1526	0.0362	0.0362	0.0086	0.5200	0.1888	0.0127	0.0448
1	0.3	0.1616	0.0434	0.0434	0.0116	0.4655	0.2050	0.0145	0.0550
0.2	0.5	0.0704	0.0146	0.0146	0.0030	0.7804	0.0850	0.0145	0.0176
0.4	0.5	0.0988	0.0352	0.0352	0.0125	0.6035	0.1340	0.0331	0.0477
0.6	0.5	0.1099	0.0506	0.0506	0.0233	0.4869	0.1604	0.0445	0.0739
0.8	0.5	0.1145	0.0613	0.0613	0.0328	0.4093	0.1758	0.0507	0.0942
1	0.5	0.1166	0.0691	0.0691	0.0409	0.3551	0.1856	0.0538	0.1100
0.2	0.7	0.0588	0.0292	0.0292	0.0145	0.6660	0.0880	0.0707	0.0437
0.4	0.7	0.0693	0.0542	0.0542	0.0423	0.4433	0.1235	0.1167	0.0965
0.6	0.7	0.0714	0.0670	0.0670	0.0629	0.3359	0.1384	0.1277	0.1299
0.8	0.7	0.0719	0.0746	0.0746	0.0774	0.2753	0.1465	0.1278	0.1519
1	0.7	0.0721	0.0796	0.0796	0.0880	0.2365	0.1517	0.1247	0.1677
0.2	0.9	0.0300	0.0416	0.0416	0.0578	0.3592	0.0716	0.2989	0.0994
0.4	0.9	0.0343	0.0567	0.0567	0.0938	0.2371	0.0910	0.2800	0.1505
0.6	0.9	0.0368	0.0647	0.0647	0.1138	0.1885	0.1014	0.2517	0.1784
0.8	0.9	0.0385	0.0699	0.0699	0.1269	0.1609	0.1083	0.2290	0.1967
1	0.9	0.0397	0.0736	0.0736	0.1363	0.1425	0.1133	0.2111	0.2099

FIGURES

Figure 2-1. Differences in the observed fawn:doe ratios and simulated fawn:doe ratios when the probabilities of correctly classifying an individual within any age or sex class are less than 1. The probability of correctly classifying a fawn is held constant at 0.8, the probability of correctly classifying a buck is held constant at 0.9, and the probability of correctly classifying a doe ranges from 0.95 to 0.75. DD is the probability of classifying an individual as a doe given the individual is a doe.

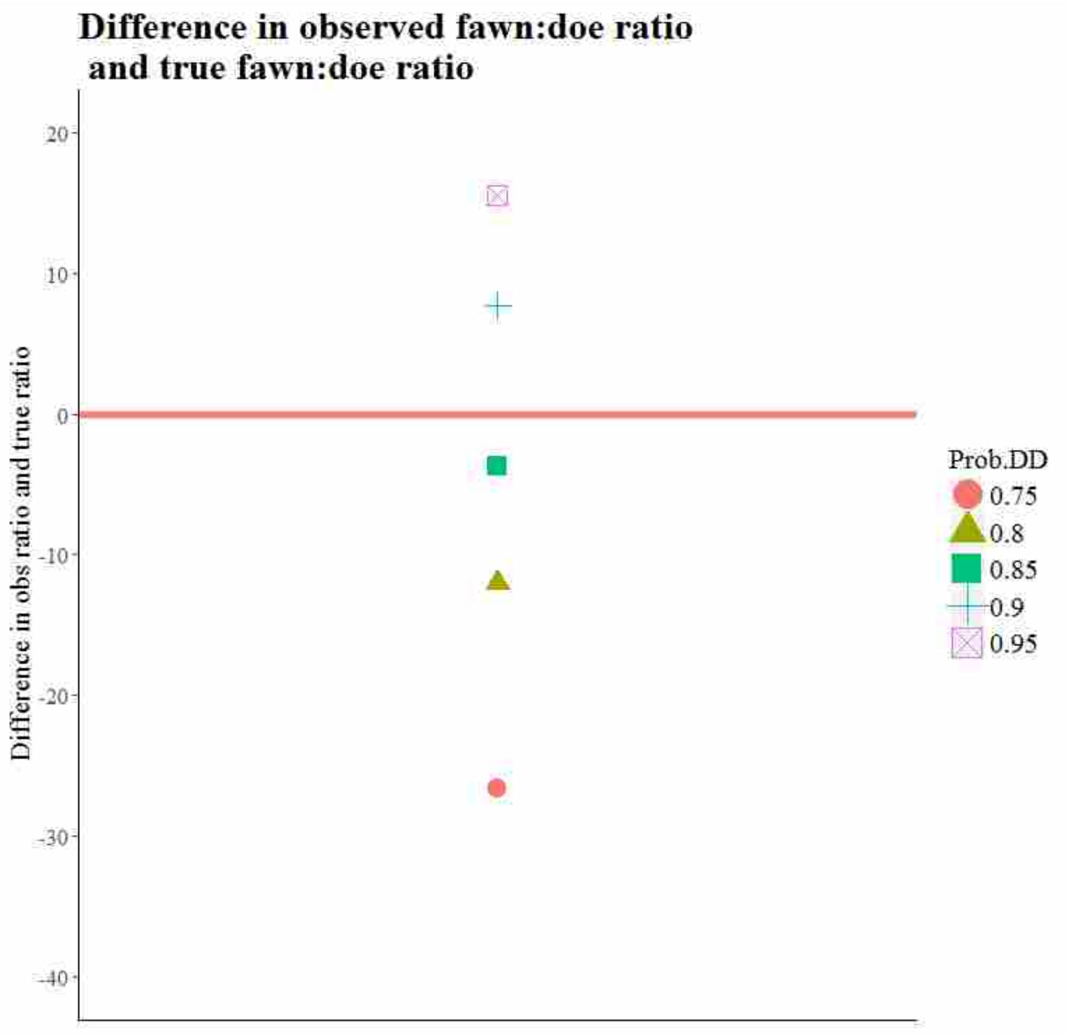


Figure 2-2. Differences in the observed buck:doe ratios and simulated buck:doe ratios when the probabilities of correctly classifying an individual within any age or sex class are less than 1. The probability of correctly classifying a fawn is held constant at 0.8, the probability of correctly classifying a buck is held constant at 0.9, and the probability of correctly classifying a doe ranges from 0.95 to 0.75. DD is the probability of classifying an individual as a doe given the individual is a doe.

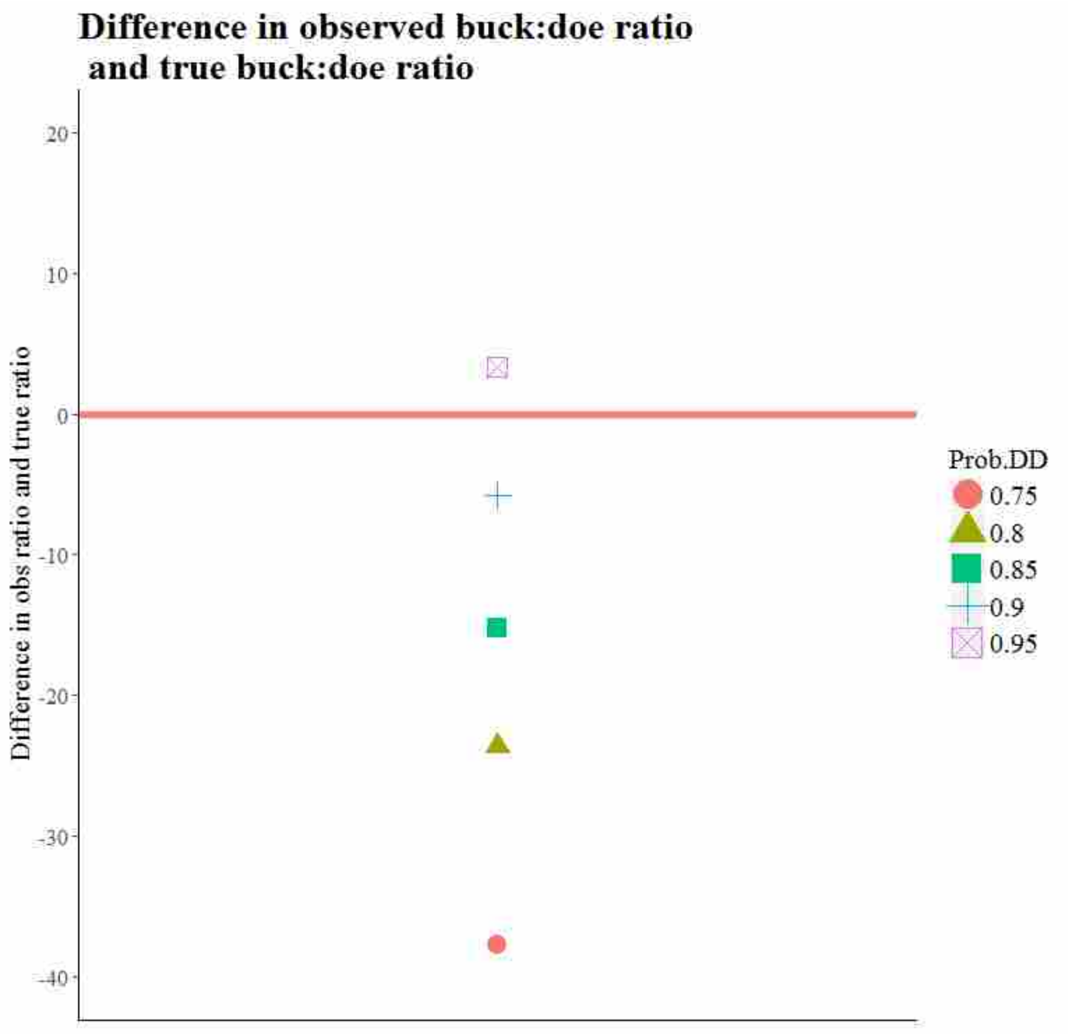


Figure 2-3. Estimates of abundance for the Boise River mule deer population management unit (PMU) over a 5 year period when fawn:doe ratios are adjusted over or under the observed ratio in increments of 10 individuals.

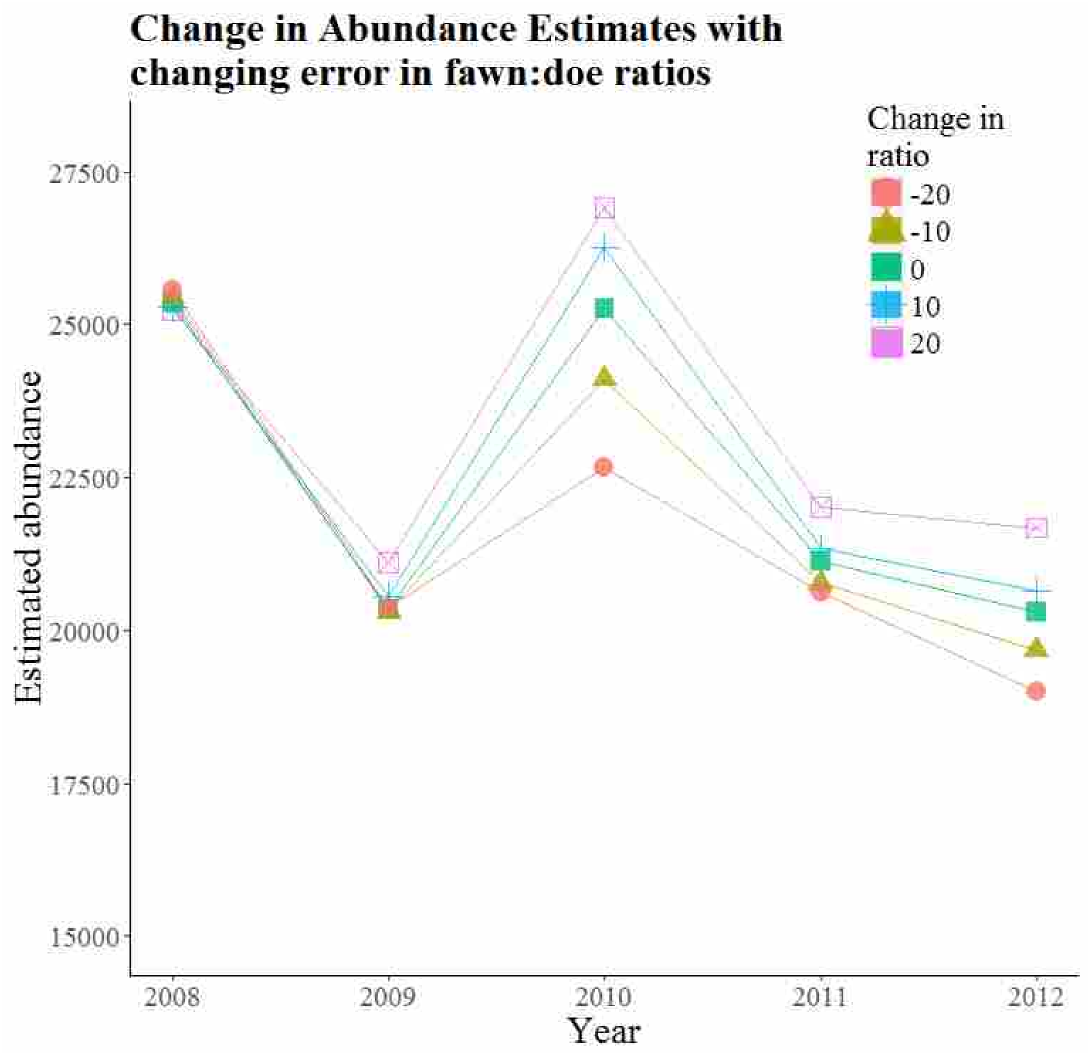


Figure 2-4. Estimates of abundance for the Boise River mule deer population management unit (PMU) over a 5 year period when buck:doe ratios are adjusted over or under the observed ratio in increments of 10 individuals.

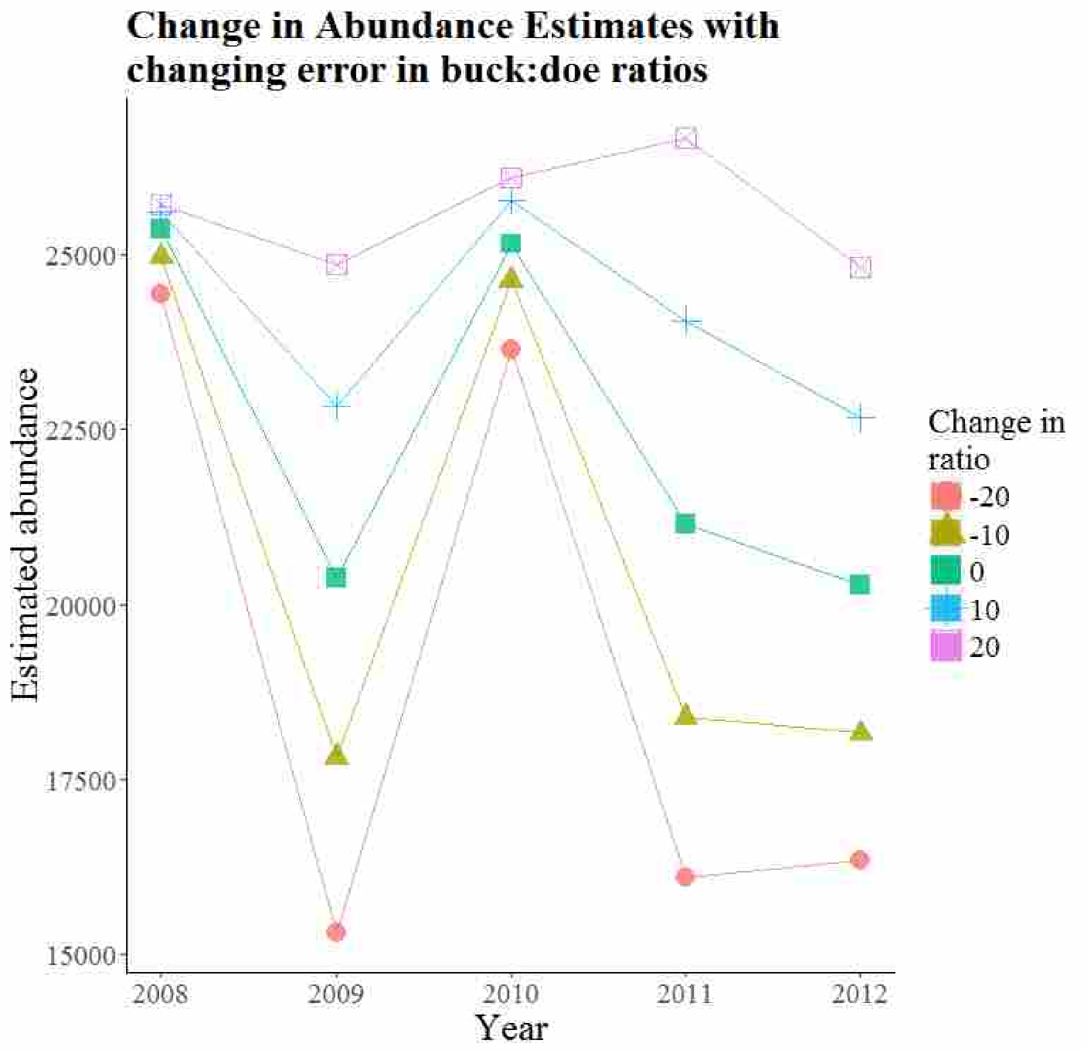


Figure 2-5. Estimates of abundance for the Boise River mule deer population management unit (PMU) over a 5 year period when fawn:doe and buck:doe ratios are adjusted over or under the observed ratio in increments of 10 individuals.

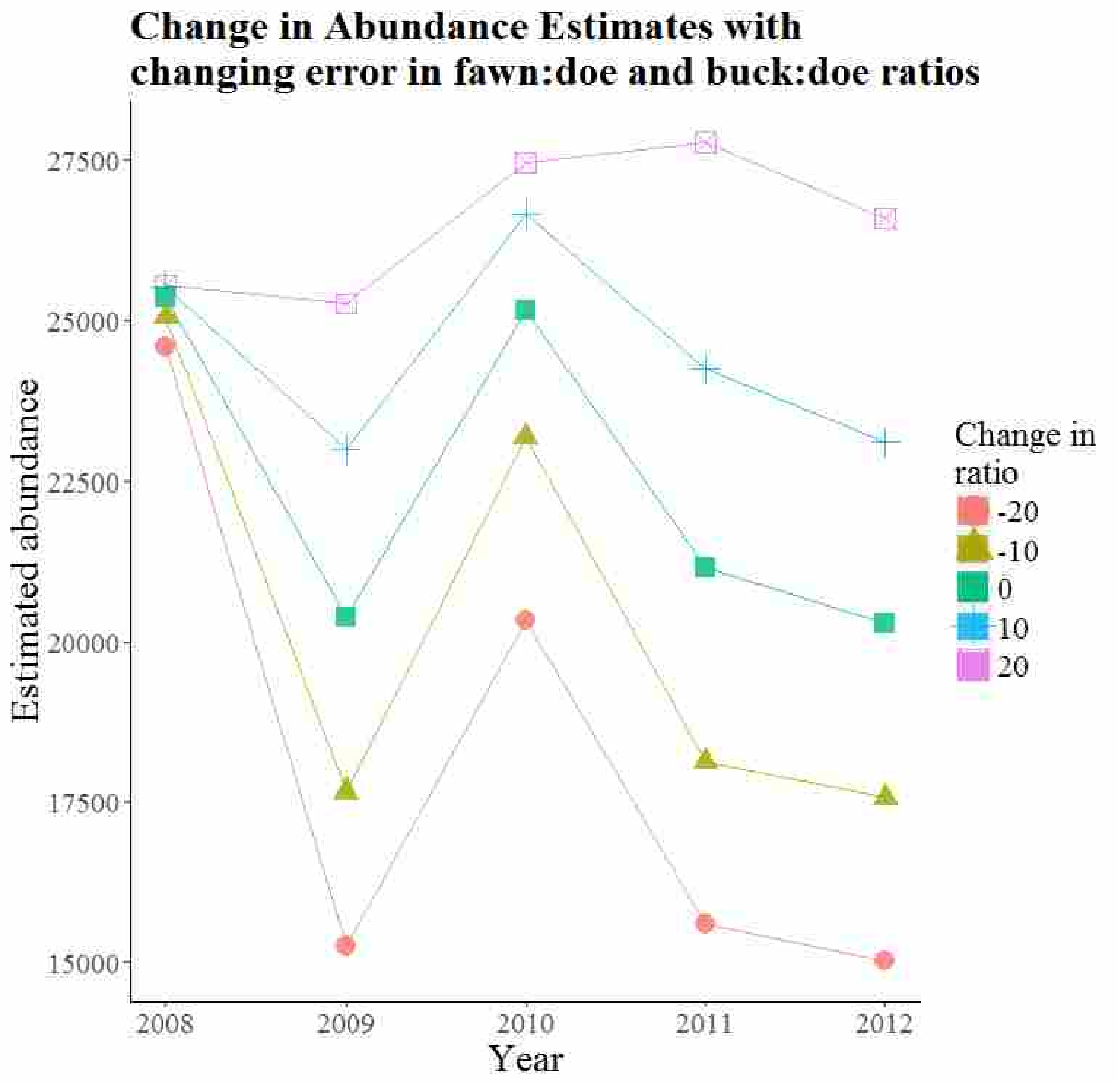


Figure 2-6. Plot of the increase in the precision of the fawn:doe ratio estimate as the number of groups resampled is increased. The probability of correctly classifying a buck as a buck is held constant at 0.9. The values in the grey box on the right of the graph represent the probability of correctly classifying a fawn as a fawn. DD is the probability of classifying an individual as a doe given the individual is a doe, BB is the probability of classifying an individual as a buck given the individual is a buck, and FF is the probability of classifying an individual as a fawn given the individual is a fawn.

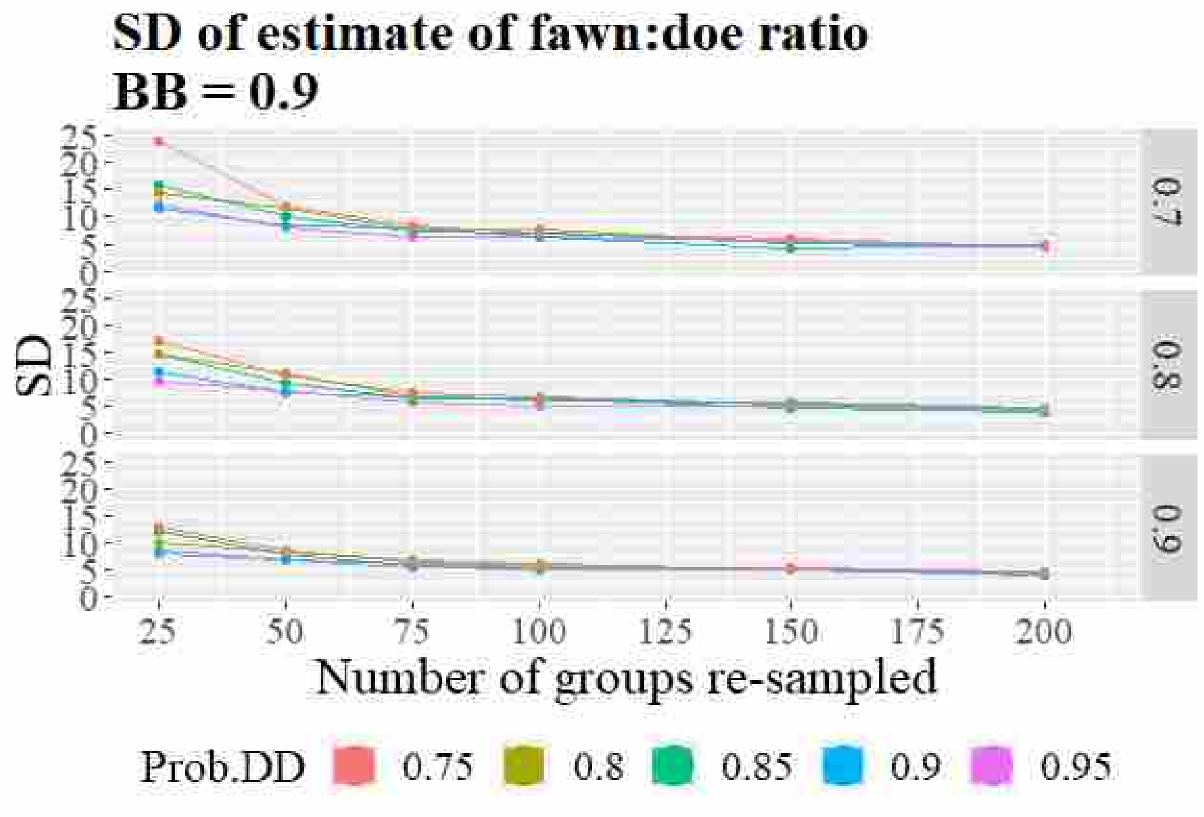


Figure 2-7. Plot of the increase in the precision of the buck:doe ratio estimate as the number of groups resampled is increased. The probability of correctly classifying a buck as a buck is held constant at 0.9. The values in the grey box on the right of the graph represent the probability of correctly classifying a fawn as a fawn. DD is the probability of classifying an individual as a doe given the individual is a doe, BB is the probability of classifying an individual as a buck given the individual is a buck, and FF is the probability of classifying an individual as a fawn given the individual is a fawn.

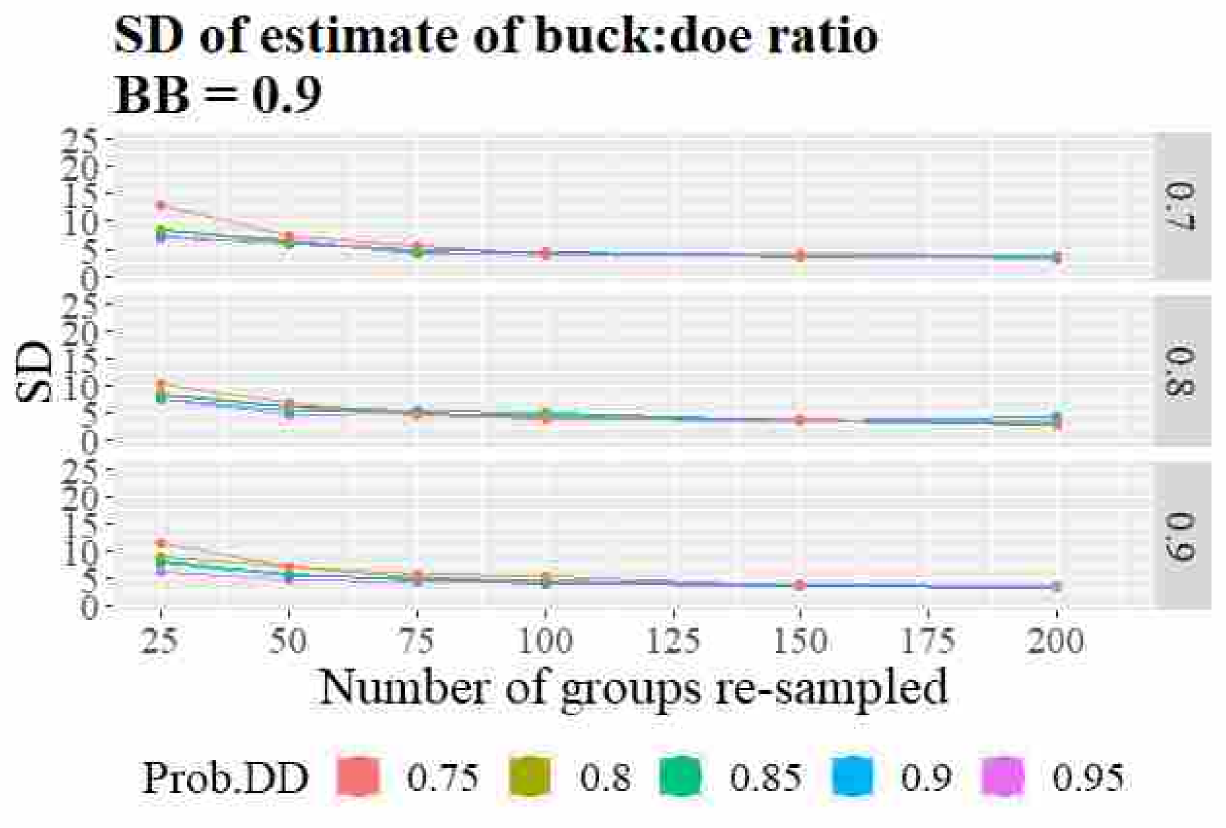


Figure 2-8. Plot of the increase in the precision of the fawn:doe ratio estimate as the number of groups resampled is increased. The probability of correctly classifying a fawn as a fawn is held constant at 0.8. The values in the grey box on the right of the graph represent the probability of correctly classifying a buck as a buck. DD is the probability of classifying an individual as a doe given the individual is a doe, BB is the probability of classifying an individual as a buck given the individual is a buck, and FF is the probability of classifying an individual as a fawn given the individual is a fawn.

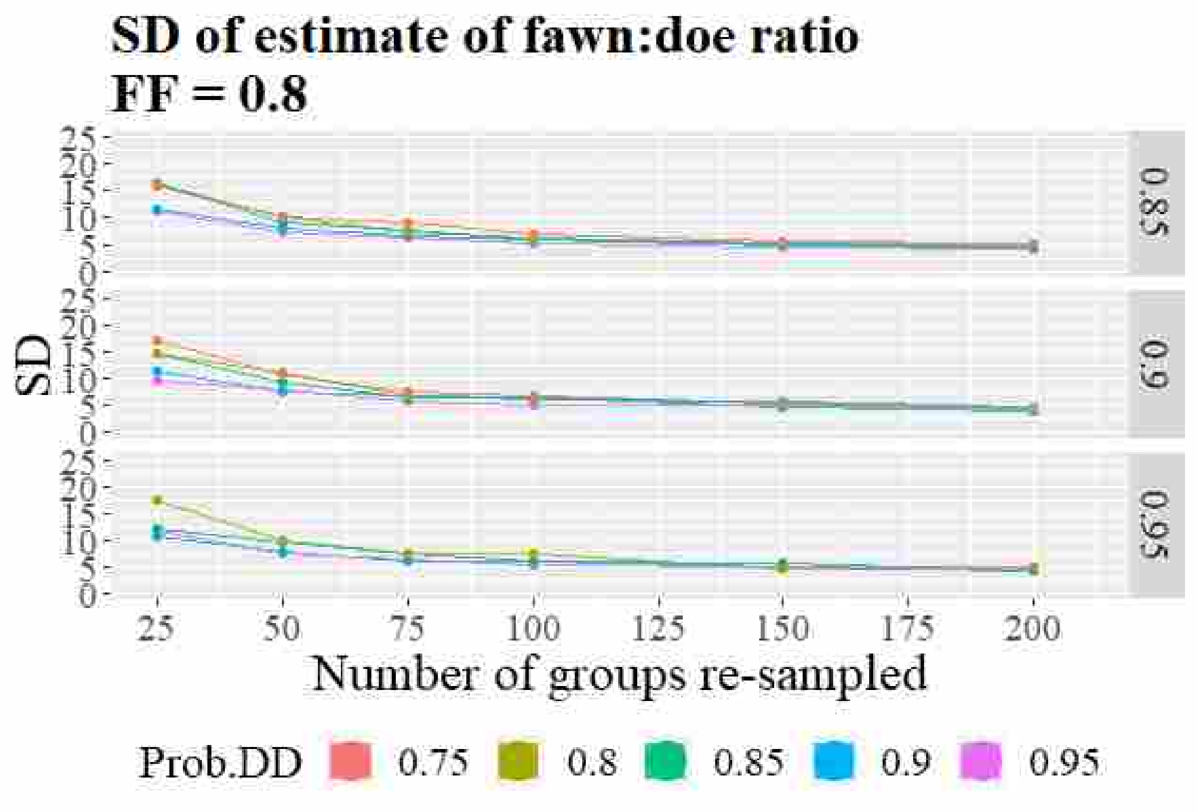


Figure 2-9. Plot of the increase in the precision of the buck:doe ratio estimate as the number of groups resampled is increased. The probability of correctly classifying a fawn as a fawn is held constant at 0.8. The values in the grey box on the right of the graph represent the probability of correctly classifying a buck as a buck. DD is the probability of classifying an individual as a doe given the individual is a doe, BB is the probability of classifying an individual as a buck given the individual is a buck, and FF is the probability of classifying an individual as a fawn given the individual is a fawn.

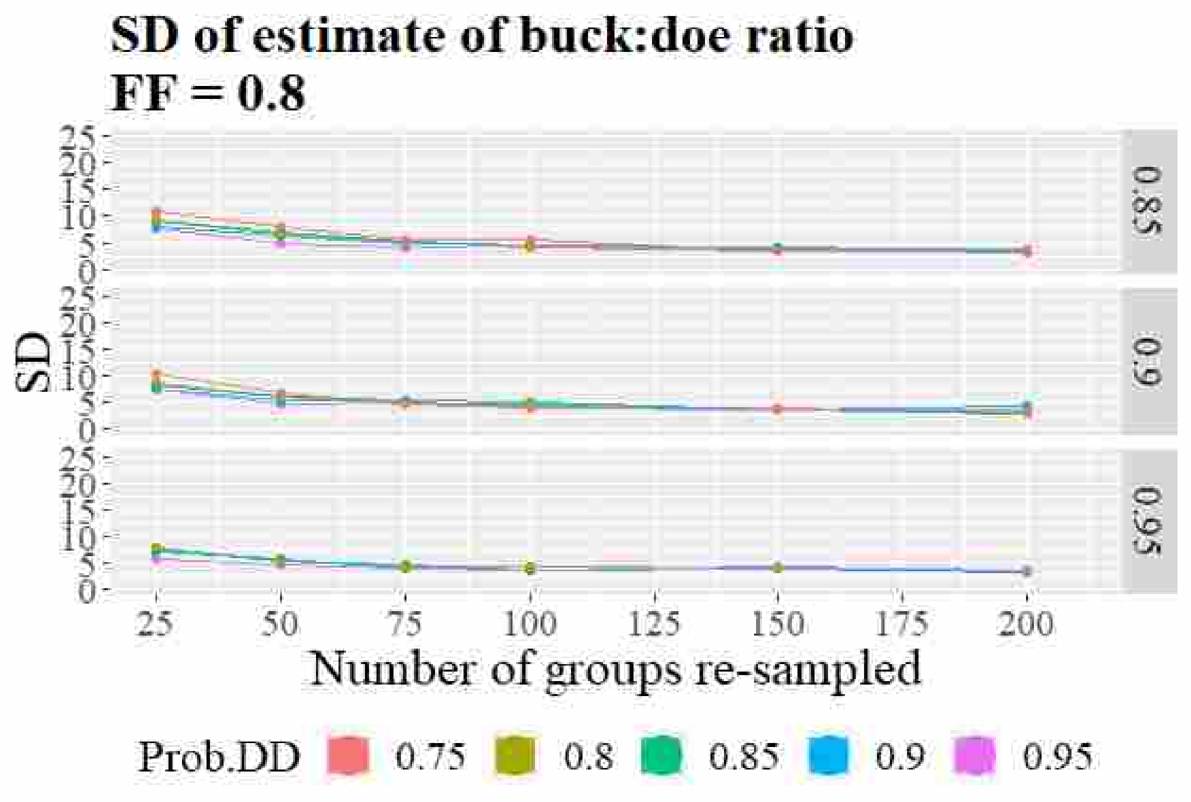


Figure 2-10. Elasticity values for survival rates from the population matrix as age and sex ratios change.

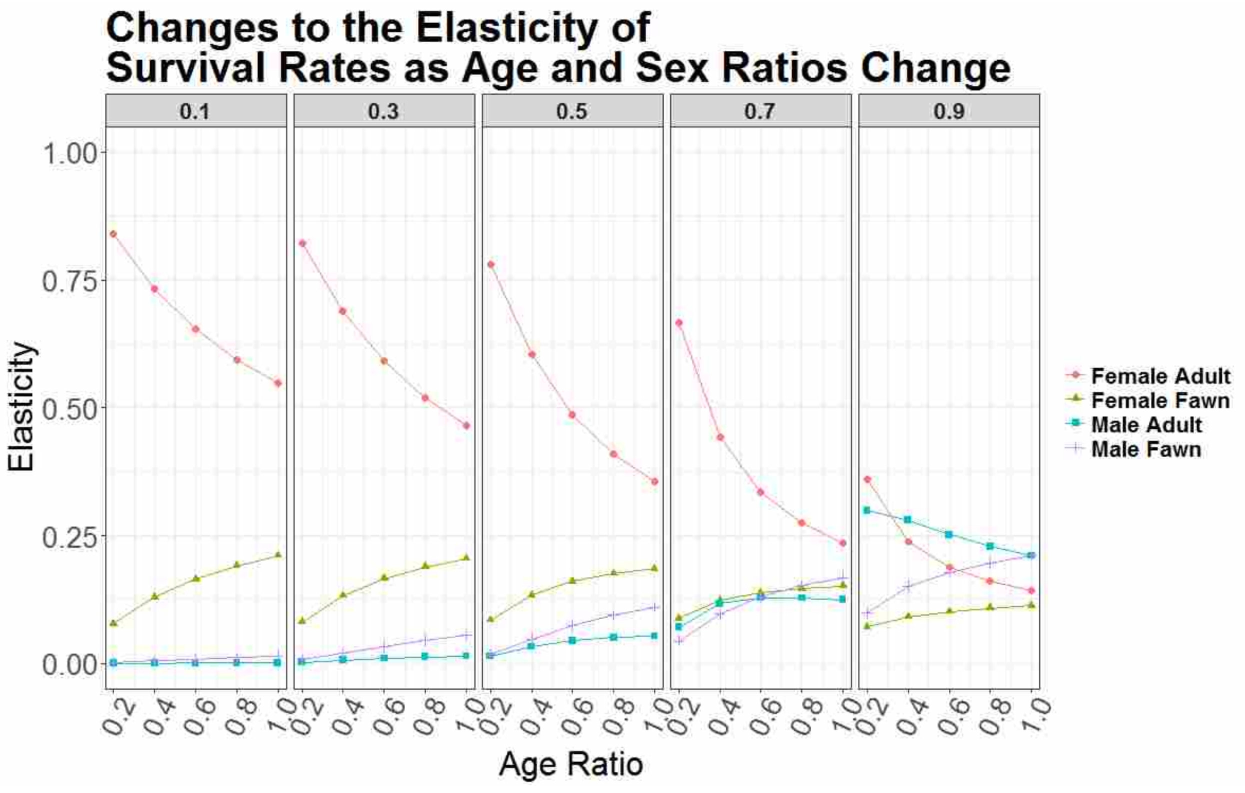
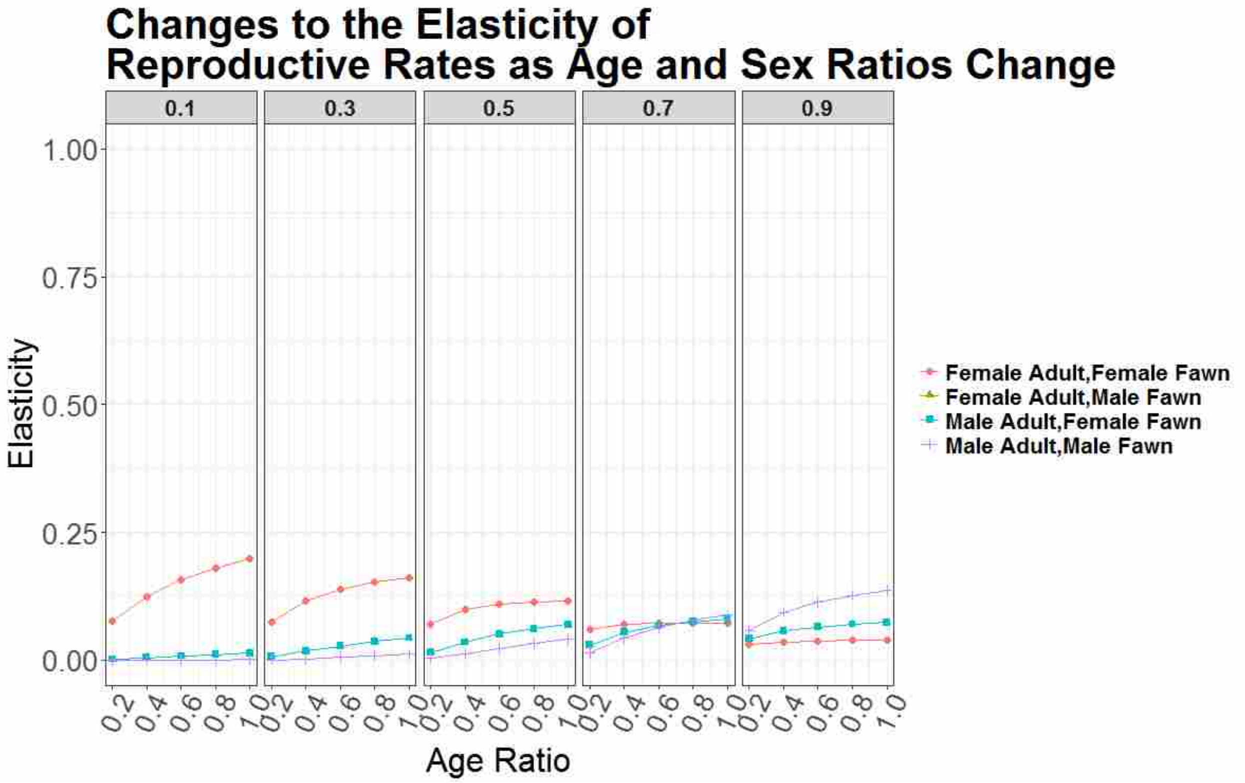


Figure 2-11. Elasticity values for reproductive rates from the population matrix as age and sex ratios change.



CHAPTER 3: THE ROLE OF DATA TYPES WITH VARYING LEVELS OF QUALITY IN GENERATING ESTIMATES OF ABUNDANCE WITH INTEGRATED POPULATION MODELS

Introduction

Accurately estimating the abundance of wildlife populations provides a foundation for understanding their ecology and for making effective management decisions. Managers and researchers use population models to predict trends, conceptualize dynamics, and test hypotheses regarding population dynamics (White 2000). Population models also enable managers to assess the outcome of specific management actions and the effects of natural or anthropogenic caused variations in the environment. However, long term, broadly distributed datasets often result in the gathering of a variety of data types. Traditionally, there is an independent analysis of each different type of data. Separate analyses risk producing parameter estimates with increased bias and uncertainty, especially if sample sizes are small because there is no sharing of information among the different data types (Doak et al. 2005, Abadi et al. 2010). Advances in population modeling have resulted in the development of integrated population models (IPM), a method that combines separate data types for analysis within a single hierarchical model (Besbeas et al. 2002, Brooks et al. 2004).

IPMs are a modeling framework which allows the incorporation of different types of data into a single hierarchical model. Hierarchical models facilitate the study of complex systems by breaking these systems into smaller more easily studied components. Hierarchical models consist of a set of sub-models that describe ecological and observation processes related to one another through conditional probabilities (Royle and Dorazio 2008, Kéry and Royle 2016). A hierarchical IPM typically consists of a state process model and an observation model for each type of data incorporated into the state process model. The state process model describes variation in the ecological process and includes the latent variable of interest such as the true population size in a model used to estimate abundance (Royle and Dorazio 2008). The observation model describes the data collection processes that produce the observable data. Observation models often include information related to detection bias such as environmental factors, animal behavior, or observer ability (Royle and Dorazio, 2008). Benefits of using a single hierarchical model framework for IPMs include an easier way to account for different sources of variance, the sharing of information between different data types, and a better accommodation of data observed at different scales (Royle and Dorazio, 2008). These attributes result in more precise estimates compared to other population modeling methods that analyzed data types individually and allows for the estimation of parameters for which little or no data were collected (Besbeas et al. 2003, Abadi et al. 2010). Another attribute of hierarchical IPMs is that Bayesian statistical methods can be employed for the analysis of these models by employing MCMC (Markov chain Monte Carlo) algorithms which allow for latent variables to remain in the model rather than being integrated out as is the case in frequentist statistical methodology (Kéry and Royle 2016).

Bayesian inference is a statistical paradigm that characterizes uncertainty about the parameters of a model by using probability directly in the calculation of parameter values. In this mode of statistical analysis, parameters are variables in a statistical distribution that the observed data provides more information about as opposed to frequentist statistical analysis that considers parameters to be fixed values that are unknown but estimated over multiple hypothetical sets of data (Kéry and Royle 2016). Bayesian inference updates prior knowledge about a parameter with data collected about the parameter using Bayes theorem represented by the following equation

$$[\theta|y] = \frac{[y|\theta][\theta]}{[y]}$$

where y is the data, θ is the parameter, $[\theta]$ represents the prior knowledge about the parameter (marginal distribution of θ), $[y]$ is the marginal distribution of the data, $[y|\theta]$ is the likelihood of the data given the parameter (conditional probability distribution), and $[\theta|y]$ is updated knowledge regarding the parameter given the data (posterior distribution). The computation of the numerator is relatively straightforward but the intensive computation required to integrate $[y]$ has prevented the widespread use of Bayes' rule until the recent development of methods to bypass this integration problem which are referred to as MCMC algorithms (Kéry and Royle 2016). MCMC algorithms allow the characterization of the posterior distribution without having to calculate [y] thereby giving access to this probability rule that is capable of updating prior knowledge about a parameter. The resulting posterior distribution produces a distribution of possible values for the parameter of interest and their associated probability of occurring given the prior knowledge concerning the parameter and the current data. The 95% credible interval (CRI) of the posterior distribution includes the range of possible parameter values located between the 2.5 and 97.5 quantiles of the posterior distribution.

Wildlife population data are notoriously difficult to gather and include both process error, the variation in population size due to biotic and abiotic processes, and observation error, variation in the estimate of population size due to the methods used to collect data. IPMs allow the separate analysis of these two types of error by lower level models within a single IPM. The benefit of accounting for these two types of error in the same hierarchical model is that by explicitly accounting for errors made during the observation process it is easier to recognize and quantify the influence of natural processes (Ahrestani et al. 2013). These attributes of an IPM represent an improvement over the individual analysis of single data types. However, we need a method to determine which data types have the greatest influence on the output of an IPM. In addition, this method should also provide a better understanding of how changes in the bias and precision of the estimates made from each data impact IPM-generated estimates of abundance.

One of the challenges of including multiple data types within a single model is to ensure that data of relatively low quality do not have greater influence on the model output than data of higher relative quality (Francis 2011, Maunder and Punt 2013). Data from different sources vary in their quality relative to one another due to the challenges encountered during the collection process for each data type. Two methods previously used to determine the influence of a particular data type on the model output are to remove a single data type from the model or change a pre-specified weight associated with the data type and then quantify the changes in the model outcome to determine the amount of influence associated with each data type (Maunder and Punt 2013). We define weight as the contribution a particular data type makes to the IPM-generated estimate of abundance. We chose to explore the second method, manipulation pre-specified weights, to track how changes in the bias and precision attributes of each data type impact the weight given to a particular data type by the IPM during the estimation process. We also assess the impact that changes to the bias and precision attributes of each data type have on the abundance estimate. These two assessments provide a methodology for determining data collection priorities and the allocation of population monitoring resources.

We use data about the mule deer (*Odocoileus hemionus*) population in Idaho to test how the relative quality and quantity of each data type influences how the IPM uses different data types and how changes to the attributes of the estimates generated from each data type impact the estimates of abundance generated by the IPM. We use data collected by the Idaho Department of Fish and Game (IDFG) to estimate the size of regional and statewide mule deer populations from the past decade (2008-2018). The types of data used in this IPM are estimates made from data collected about the vital rates associated with the mule deer population in Idaho. The types of estimates used as data by the IPM include estimates of adult and juvenile survival, estimates of abundance generated using population count data and a sightability model, age and sex ratios estimated from herd composition data and the sightability model, and harvest. The collection process and models used to generate estimates from these data influence their relative quality. Most of the problems affecting the relative quality of the estimates from different data types have potential solutions that have been suggested elsewhere (Fieberg et al. 2013, Henderson Chapter 2). Therefore, this research focuses on developing a methodology for assessing how the bias and precision of the data incorporated in the IPM impacts the estimate of abundance from the IPM and determines which types of data have the most influence on the IPM. The development of this methodology will allow for the assessment of multiple data types across a range of species specific IPMs that will assist in the prioritization of data collection and help in determining how to use monitoring resources most efficiently.

Survival data come from a statewide radio collaring and monitoring program and are subject to several assumptions. The first assumption is the sample of radio-collared individuals is representative of the overall population. However, the logistical limitations of capturing individuals in a completely random manner with regards to space may violate this assumption to some extent. A second assumption is survival estimates for both sexes in the adult age class are equal because of the lack of data for adult males. This is likely a poor assumption due to the differences in male and female survival related to the differences in risk from human harvest as well as differences in behavior that likely changes the amount of predation risk each sex experiences throughout the year (Bowyer 2004). Equal survival for adult males and females may also be a poor assumption due to behavioral differences which influence access to resources due to differences in space use by the different sexes throughout the year (Mackie et al. 1998). The final assumption relates to the use of a known fate model to estimate survival for this population (White and Burham 1999). While we know the fates of the majority of individuals, there remains some portion of the sample whose fate is unknown. The assumption is that the number of individuals whose fate remains unknown is small enough not to bias the estimates of survival made by the known fate model for different age and sex classes within the population.

The abundance data come from counts of animals collected by aerial surveys conducted when mule deer populations are on their winter range. An adjustment of the raw counts occurs using a sightability model that incorporates information on variables that influence detectability (Unsworth et al. 1994). The variables in the sightability model include group size, group visibility, activity, snow coverage, and vegetation. The relative quality of count data is affected by the under estimation of variance due to violating assumptions associated with the sightability model. The first assumption is that the survey stratums in each PMU accurately reflect the true probability of mule deer occupancy within that PMU. Violation of this assumption could result in the allocation of survey/sampling effort being different than the actual spatial distribution of mule deer in the PMU. Changes in sampling effort for various strata have the potential to bias estimates of abundance high or low according to how the sampling effort is misallocated (Thompson 2002). The second assumption made in the sightability model is that there is no difference in the detectability of different age and sex classes within the mule deer population. Differences in the behavior between different ages and sexes of mule deer during at the time when the survey occurs may result in the violation of this assumption (Mackie et al. 1998). Not accounting for different rates of detection for different age and sex classes has the potential to bias estimates of abundance by overestimating the number of individuals in a group composed of individuals that are more difficult to detect. A third assumption made by the sightability model is that all observers are of equal ability in detecting and accurately counting the number of individuals in a group. Differences in observer ability has been shown by multiple studies as an important factor in determining the number of individuals in a group during aerial surveys (Caughley et al. 1976, Kendall et al. 1996) and has the potential to bias estimates of abundance made from the sightability model.

Estimates of age and sex ratios come from herd composition data also collected by aerial surveys when mule deer are located on their winter range. The ratio estimates made using the sightability model share the same assumptions as the abundance estimate generated from count data with one addition. The additional assumption made when estimating age and sex ratios is the correct classification of all individuals according to their true age and sex. Violation of this assumption could occur as a result of the timing of the survey, low visibility conditions, or observer experience. Other work has shown that even small amounts of classification error (probability of correctly classifying an individual < 1) can influence the estimation of the age and sex ratios estimated from this data (Henderson Chapter 2).

The various collection methods employed during this period influence the relative quality of the mule deer harvest data used in this IPM. Methods for collecting harvest data include self-reporting online or via phone, game check stations, mail surveys, and phone surveys. There are several assumptions associated with the collection of this data that may influence the level of accuracy of the harvest counts. One basic assumption is that the number of individuals reported harvested reflect the true number of individuals harvested. Second, that there is no reporting bias associated with the success or failure of a hunter to harvest a deer. In other words, a hunter that is successful is not more likely to report than a hunter that is not successful. The third assumption associated with harvest data is that the location of the harvest event is accurate at least to the PMU level. Violation of this assumption is more likely to happen in areas where harvest occurs near the boundary of two or more PMUs.

We seek to provide managers and researchers with information regarding the influence that different data types have on the estimates of abundance generated by an IPM. We will track the weights associated with different data types within the IPM under a variety of scenarios which change the bias and precision attributes of the estimates made using each data type. Tracking the changes in weights associated with changes in the relative bias and precision of each data type and the resulting impact on the bias and precision of the IPM-generated abundance estimates will allow the quantification of how data types of different relative qualities influence abundance estimation. Given that resources for monitoring wildlife populations are limited and fluctuate on an annual basis, the results of this analysis provide two important pieces of information for wildlife management and conservation. First, these results provide information that will improve the overall quality of population abundance estimates on which to base management and conservation decisions. Second, the methodology used to assess the influence of different types of data will produce a template that managers and researchers can use to help prioritize data collection. This template provides managers and researchers a defensible basis for their decisions regarding data collection priorities, and facilitates a more efficient and cost effective manner to monitor wildlife populations. While we develop and test this methodology using a mule deer IPM, we expect it to have the flexibility to apply to IPMs used to estimate population parameters and abundance for many other species.

METHODS

We use the following definitions for bias, precision, and accuracy throughout the rest of the manuscript. We define bias as the distance from the true value of a parameter to the expected value of that parameter. When bias was used in relation to data, we an assumed a direct correlation between data quality and the amount of bias in the estimator associated with that data type. Lower quality data suggests the estimate from that data contains more bias. We define precision as how close a set of estimated values are to one another. When using precision in relation to data, we assume a direct correlation between data quantity and the amount of precision associated with the estimate made from that data type. We assumed that as the quantity of data increased the precision of the estimate from that data increased. We define accuracy as the combination of bias and precision. Also in an effort to provide further clarity, in the rest of this manuscript the term ‘raw data’ refers to the data collected to estimate a population parameter while ‘data’ refers to the estimate itself when used by the IPM for estimating abundance.

Data

IDFG collected raw data used for this analysis from the mule deer population in the Boise River PMU from 2008-12 (Figure 3-1). The types of raw data included are: 1) encounter histories to estimate survival rates for fawns and juveniles of both sexes and adult females, 2) counts of abundance to be used in sightability models to estimate total abundance in a specific PMU, 3) counts of individuals in different age and sex classes to be used by the sightability model to estimate age and sex ratios, and 4) the number of individuals harvested by sex in each PMU.

The raw survival data collected from adult females and juveniles of both sexes are the result of capturing individuals from these age and sex classes using a combination of drive nets, net gunning from helicopters, and chemical immobilization when mule deer are present on their winter range. Captured individuals were radio collared using both GPS and VHF devices and monitored at regular intervals. The number of individuals monitored fluctuated across PMUs. IDFG collected raw survival data at regular intervals via aerial and ground telemetry as well as remote transmission of data from GPS devices. The observation of the state of each individual at each monitoring occasion fell into one of three categories: alive, dead, or unknown. We created encounter histories for each individual over time from the raw data collected through monitoring. Due to the monitoring intensity, we used a known fate model to estimate the survival rate of different age and sex classes. Known fate models do not include a term for probability of detection, based on the assumption that the probability of detection is equal to 1.0, and is typical of studies when individuals are outfitted with radio monitoring devices (White and Burnham 1999).

IDFG collected raw count data during aerial surveys conducted in each PMU every 5 to 10 years for use in sightability models to predict total abundance by PMU. IDFG divides each PMU into high, medium, and low strata according to the probability of mule deer occupancy. Sampling effort is allocated according to strata with high occupancy strata being sampled the most followed by medium and then low occupancy strata. During the survey, raw data were collected regarding the activity of the group when first observed (moving, standing, bedded), visibility of the group (in 15% intervals based on the amount of each deer that is screened from view by vegetation), vegetation type (categories are specific to each PMU), the amount of ground covered by snow within an specific distance of the observed group divided into 3 categories (0-20%, 21-79%, and 80-100%), and the size of each group observed (number of individuals; Unsworth et al. 1994). The estimate of the true number of deer present comes from the processing of the raw data through the sightability model. Based on a linear regression of the sightability model predictive variables were valued as follows: the y-intercept was -0.254, activity coefficient values were 0 if bedded, 1.56 if standing, or 4.43 if moving, vegetation coefficients values were 0 if grass/open/agriculture, -0.88 if sagebrush, -2.383 if juniper/mtn.mahogany, -0.60 if aspen/mtn. brush, or -0.63 if conifer, snow cover coefficients values were 0 if < 20% snow, -1.37 if > 20% and < 79% snow, or -0.60 if > 80% snow, and the group size coefficient was 0.047 (Unsworth et al. 1994).

IDFG collected raw data to estimate age and sex ratios during aerial surveys conducted every 1 to 3 years in most PMUs. Similar to the aerial surveys made to gather count data for abundance estimates each PMU was stratified into high, medium, and low strata according to the probability of mule deer occupancy and sampling effort was divided accordingly. These surveys also collect raw data regarding the deer activity, visibility, vegetation type, and snow coverage. In addition, observers classify the members of each group of observed deer into one of the following six categories: fawn, adult female, yearling male, sub-adult male, adult male, and unknown. The sightability model estimated the number of individuals in each age and sex class using the raw data. IDFG uses these estimated numbers to calculate age and sex ratios.

IDFG collected raw harvest data using several methods during this time period. These methods included self-reporting online or via phone, game check stations, mail surveys, and phone surveys. Raw harvest data include information on the sex and location of each individual harvested. During the time period used in this analysis raw harvest data consisted only of counts of the number of each sex reported harvested in each PMU.

Integrated Population Model

The IPM used for this analysis is one of the current formulations of the model used by IDFG to estimate mule deer populations (Nowak et al. 2018). The following equations describe the expected abundance values for a single year for each age and sex class in the population from the stochastic process models used throughout this analysis:

$$N_{fawn,F} = N_{adult,F} * R_{year} * 0.5$$

$$N_{fawn,M} = N_{fawn,F}$$

$$N_{juv,F} = N_{fawn,F} * S_{juv,F}$$

$$N_{juv,M} = N_{fawn,M} * S_{juv,F}$$

$$N_{adult,F} = \left(N_{juv,F} * S_{juv,F}\right) + N_{adult,F} - H_{adult,F}$$

$$N_{adult,M} = \left(N_{juv,M} * S_{juv,M}\right) + N_{adult,M} - H_{adult,M}$$

where *N* is the number of individuals in a specific age and sex class, *S* is the survival rate for a specific age and sex class, *R* is the recruitment rate based on the age ratio, and *H* is the number of individuals harvested by sex. The fawn age class consists of individuals below 1.5 years of age, juveniles are individuals between the ages of 1.5 to 2 years of age, and adults are individuals older than 2 years of age. This version of the IPM holds adult survival constant over time, allows juvenile survival to vary over time, and allows reproduction to vary over time. The observation models used to estimate demographic parameters for the data types included in stochastic process models in this IPM are as follows:

$$n \sim Normal(N_{total}, tau\ N_{total})$$

$$YF \sim Normal\left(R_{year}, tau\ YF\right)$$

$$MF \sim Normal(mf, tau\ MF)$$

$$S \sim Normal(S_{age,F}, tau\ S_{age,F})$$

where *n* is the number of individuals, *YF* is the age ratio (fawns per adult females), *R* is the recruitment rate based on the age ratio (fawns per adult females), *MF* is the sex ratio (adults males per adult females), and *S* is the survival rate for a specific age and sex class. In this model, we assumed the observed data for each type of data are normally distributed with a mean and a measure of precision *tau*. We assumed a normal distribution because we are approximating the error using the uncertainty associated with estimated values from the specific data type sub-models, such as the sightability or survival models. The measure of precision is given as *tau* because these models were being run in the program JAGS (Plummer 2003) and is the reciprocal of the variance $\left(\frac{1}{\delta^2}\right)$. *R* and *mf* are values derived within the model using the following equations

$$R_{year} = yN/N_{adult,F}$$

$$mf_{year} = N_{adult,M}/N_{adult,F}$$

where *R* is the age ratio, *mf* is the sex ratio, *N* is the number of individuals estimated to be in a specific age and sex class, and *yN* is the estimate of the total number of fawns in a specific year.

In order to calculate the influence that each data type was having on the IPM-generated estimate of abundance in each scenario a weighting term was included for each data type. We created a set of possible weight combinations to track the influence of each data type in the various bias and precision scenarios. The weight combinations resulted in a matrix that included all possible combination of weights ranging from 0.3 to 1.0 for each data type. This resulted in 4096 possible combinations of weights for each bias and precision scenario. We limited the lower cutoff point for possible weights to 0.3 for two reasons. First, we assumed that any data type receiving less 0.3 weight in the IPM would be a likely candidate for not being included in the data types used by the IPM for generating abundance estimates. The second reason for limiting the lower range of possible weights was to decrease computational time.

Population Matrix

We used a matrix model populated with the estimates from each data type's specific sub-model to simulate the "true" population size through time. The 'true' abundance value from the population matrix provided a way to measure the amount of bias associated with the estimate of abundance generated by the IPM for different bias and precision scenarios for each data type. We employed the following population matrix

$$\begin{bmatrix} 0 & 0 & R_{year} * S_{adult,F} * 0.5 & 0 \\ 0 & 0 & R_{year} * S_{adult,F} * 0.5 & 0 \\ S_{fawn,F} & 0 & S_{adult,F} & 0 \\ 0 & S_{fawn,M} & 0 & S_{adult,M} \end{bmatrix} * \begin{bmatrix} N_{fawn,F,t} \\ N_{fawn,M,t} \\ N_{adult,F,t} \\ N_{adult,M,t} \end{bmatrix}$$

where *N* is the number of individuals in a specific age and sex class at time *t*, *S* is the survival rate for a specific age and sex class, and *R* is the recruitment rate based on the age ratio.

Scenarios Tested

We tested multiple scenarios with varying levels of bias and precision for each data type to determine how changes to these attributes influenced the weight associated with each data type within the IPM and the bias and precision of the IPM-generated estimate of abundance. The scenarios tested for survival estimate data included a relatively more biased estimate of survival with more precision (more bias/more precision), a relatively less biased estimate of survival with less precision (less bias/less precision), and an estimate of survival that was both more biased and less precise (more bias/less precision). The amount of bias associated with each survival scenario reflected the values from the low end of the range of recorded adult female and fawn survival estimates (Forrester and Wittmer 2013). Each scenario represents a potential way that real world conditions influence survival estimate data. The more bias/more precision scenario could arise from a large but biased sample due to the difficulties involved with randomly capturing individuals in space. The less bias/less precision scenario could arise from a sample that is relatively small but very representative of the population with the small size potentially related to the difficultly in capturing individuals from a specific age or sex class. The more bias/less precision scenario could occur when the sample size is small and biased due to lack of random spatial sampling and difficulty in capturing a sufficient number of individuals.

The scenarios tested on the abundance estimate data from the sightability model included abundance estimates with relatively less bias and less precision (less bias/less precision), abundance estimates with relatively more bias and more precision (more bias/more precision), and abundance estimates with relatively less bias and more precision (less bias/more precision). The less bias/less precision scenario represents a situation when the groups observed during the data collection are representative of the population as a whole but factors such as topography, canopy cover, or weather make obtaining a large sample difficult. The more bias/ more precision scenario could arise from making observations during a time period when the detection of groups in the population is easier but are scattered over a larger geographical area due to season behaviors such as migration to winter range or the stage of the reproductive cycle. The less bias/more precision scenario represents the best case scenario. This situation could arise when the population is spatially compact and groups are easily detectable. We tested this scenario because we assume that having the perfect conditions for gathering raw count data is the exception rather than the rule. Therefore, we wanted to test if the IPM would rely most heavily on the abundance estimate data when it contained the highest quality information pertaining to abundance.

We analyzed three scenarios to determine how the IPM incorporates age ratios estimated from raw data collected during aerial surveys for herd composition. These scenarios included age ratio estimates that had relatively less bias and more precision (less bias/more precision), age ratio estimates that had relatively more bias and less precision (more bias/less precision), and age ratios that were relatively less biased with less precision (less bias/less precision). We based the amount of bias on the range of age ratio estimates from previous estimates made by IDFG (IDFG 2008). The less bias/more precision scenario for age ratios represented the situation when there is little or no classification error present when assigning individuals to an age class and that the number of groups observed is large enough to represent the population in the area of interest. The more bias/less precision scenario could arise when classification error is present and the number of groups observed is limited. The less bias/less precision scenario represents a case when the age ratio estimate contains no classification error but the number of groups sampled is small.

We analyzed three scenarios to determine how the IPM incorporates sex ratios estimated from raw data collected during aerial surveys for herd composition. The scenarios used were the same as the scenarios tested for age ratio data due to the similarity in collection methodology, assumptions made regarding these data types, and the use of the same model for estimation. The sex ratio estimate scenarios included sex ratio estimates that had relatively less bias and more precision (less bias/more precision), sex ratio estimates that had relatively more bias and less precision (more bias/less precision), and sex ratios that were relatively less biased with less precision (less bias/less precision). The less bias/more precision scenario for sex ratios represented a scenario when there is no classification errors made when assigning individuals to a sex class and that the number of groups observed is large enough to represent the population in the area of interest. The more bias/less precision scenario represents a situation when classification error is present and the number of groups observed is not sufficient. The less bias/less precision scenario represents a case when the sex ratio estimate when there is no classification error but the number of groups sampled is small.

We tested male and female harvest values with a single scenario for each one due to these values being a single count and lacking any measure of precision. The scenarios tested for these data were a harvest value for each that was relatively less biased. The less biased values used to test the impact of changes to the came from the matrix population model and were of a smaller magnitude than the values used in the current formulation of the IPM.

Assessment Values

Each bias and precision scenario had 4096 unique combinations of weights assigned across the data types. For example, the first weight combination for each scenario assigned weights to each data type as follows: abundance weight = 0.3, age ratio weight = 0.3, sex ratio weight = 0.3, survival weight = 0.3, female harvest weight = 0.3, male harvest weight = 0.3. The second weight combination for each scenario assigned weights to each data type as follows: abundance weight = 0.6, age ratio weight = 0.3, sex ratio weight = 0.3, survival weight = 0.3, female harvest weight = 0.3, male harvest weight = 0.3. We estimated abundance using the IPM 30,000 times for each unique weight combination for each bias and precision scenario tested. We then calculated the mean value of the weight assigned to each data type, the estimate of abundance, and the precision associated with each IPM estimate of abundance from the 30,000 iterations. This resulted in 4096 sets of mean values for each bias and precision scenario tested.

We assessed the influence of each type of data under the different data bias and precision scenarios on the estimate of abundance from the IPM using three criteria. We calculated the criteria values using the mean values for the weight assigned to each data type, the IPM estimate of abundance, and the precision of the IPM estimate of abundance from the 5% of the sets of mean values from a single bias and precision scenario with the least biased estimates of abundance. We focused on the top 5% with the least bias because we did not manipulate the bias and precision of any data types except the one being tested in that specific scenario. Therefore, we assumed that the bias contained in the IPM-generated estimates of abundance are most likely the result of the bias contained in the data type assessed in that particular scenario. The first assessment criterion was the average weight associated with each data type under a particular data bias and precision scenario. We used the average weight criterion to assess how much influence each data type had on the IPM estimate of abundance when only changing the bias and precision attributes of a single data type. The second assessment criterion was the amount of bias in the IPM-generated estimate of abundance compared to the estimate of abundance from the population matrix. We calculated bias by subtracting the average IPM estimate of abundance from the abundance estimate generated using a population matrix. The third assessment criterion was the precision of the abundance estimate as measured by the average width of the 95% CRI of the posterior distribution of the IPM abundance estimate. Based on discussions with agency personnel, an acceptable range of bias in the estimate of abundance for management purposes was an estimate of abundance that is within 15% of the true abundance (pers. comm. Hurley). Consequently, an acceptable CRI would incorporate values within 15% of the true abundance in either a positive or negative direction. We used values based on these criteria to assess the acceptability of different levels of bias and precision for the estimates of abundance generated by the IPM under various data bias and precision scenarios.

RESULTS

Data weights

The results of the analysis of survival estimate data show that changes made to the bias and precision of this data type, have little impact on the weights associated with each of the other data types across all three scenarios, with one exception (Table 3-1). As the attributes of the survival estimate used as data by the IPM changed, the average weight remained relatively high for abundance estimate data ($\hat{N}$, 0.97 – 0.99), age ratio data (YF, 0.86 – 0.89), and sex ratio data (MF, 0.89 – 0.91). Weights associated with harvest data were also consistent but much lower. Female harvest data weights ranged from 0.58 – 0.59 and weights for male harvest data ranged from 0.52 – 0.58. The weights associated with survival estimate data had the most variability across scenarios. The weights associated with survival estimate data ranged from 0.42 to 0.68. The lowest weight associated with survival estimate data, 0.42, was from the more bias/more precision scenario. In the less bias/less precision scenario the weight associated with survival was 0.67. The weight associated with survival in the more bias/less precision scenario was 0.68. These results demonstrate that when the survival estimate data was both biased and overly precise, the impact of this data type was reduced by approximately 26% on the IPM-generated estimate of abundance.

The results of the analysis of survival estimate data show that changes made to the bias and precision of this data type, have little impact on the weights associated with each of the other data types across all three scenarios, with one exception (Table 3-1).

The results of the analysis of abundance estimate data from the sightability model show that changes made to the bias and precision of this data type have a large impact on the weights associated with other data types (Table 3-2). When abundance estimate data was less biased and less precise, the abundance estimate data was associated with the highest weight (0.99) followed by sex ratio data (0.95), age ratio data (0.72), survival data (0.62), female harvest data (0.57), and male harvest data (0.53). When abundance estimate data was more biased and more precise, abundance estimate data was associated with a substantially lower weight of 0.65. Age ratio data and survival data remained consistent with associated weights of 0.73 and 0.64, respectively. The weights associated with female and male harvest data both increased, with female harvest data increasing by 14% and male harvest data increasing by 25%. Similar to abundance estimate data, the weight associated with sex ratio data decreased substantially (-28%). The distribution of weights associated with all of the data types in the more bias/more precision scenario was the second smallest range of weights (0.65 - 0.78) for all scenarios tested. The less bias/more precision scenario for abundance estimate data was associated with an increase in the weight of the abundance estimate data to 0.77. This scenario also resulted in marked increases in the weights associated with age and sex ratio data, 0.92 for age ratio data and 0.97 for sex ratio data. The weight associated with survival data remained similar as the weights from the other two scenarios. Data weights associated with female and male harvest dropped to the lowest values in this set of scenarios, 0.52 and 0.41 respectively. The large fluctuations in the weights associated with most of the data types in the scenarios associated with changes to the attributes of the abundance estimate data, excluding survival data, demonstrate the large amount of influence that the bias and precision attributes of the abundance estimate data have on the impact of each data type in the generation of abundance estimates by the IPM.

The results of the analysis of age ratio data show that changes to the bias and precision of age ratio data results in consistent weights for each data type across all three scenarios with one exception (Table 3-3). As the attributes of the age ratio data used by the IPM changed, the weights of all of the other data types remained stable with abundance estimate data ranging from 0.99 to 1.0, sex ratio data ranging from 0.9 to 0.92, survival data ranging from 0.62 to 0.67, female harvest data ranging from 0.53 to 0.59, and male harvest data ranging from 0.49 to 0.51. The weights associated with age ratio data were consistent for the more bias/less precision scenario and the less bias/less precision scenarios at 0.9 and 0.88, respectively. The single exception to the consistent weights across data types and scenarios was the weight associated with the age ratio data in the less bias/more precision scenario. In the less bias/more precision scenario, the weight associated with age ratio data was 0.71, a decrease of 17% to 19% as compared to the other two scenarios involving changes to the bias and precision attributes of this data type. The consistency of the weights associated with all data types across all scenarios suggest that fluctuations in the bias and precision of age ratio data does not have a large influence on how the IPM incorporates any of the data types into the estimation of abundance nor how the IPM uses other data types in the presence of changes to the bias and precision of age ratio data. The single exception to the pattern of consistency of weights occurred when the age ratio data was at its highest levels of quality and quantity and resulted in a decreased weight being associated with age ratio data.

The results of the analysis of sex ratio data show that changes to the bias and precision attributes of this data type have increased influence on the weights associated with other data types used by the IPM (Table 3-4). The weights associated with the various data types fluctuated as much as 46% between the three scenarios involving changes to the attributes of sex ratio data. When sex ratio data was less biased and more precise, abundance estimate data was associated with the highest weight (1.0), followed by sex ratio data (0.76), age ratio data (0.74), female and male harvest data which both had weight of 0.59 associated with them, and survival estimate data (0.49). When sex ratio data more biased and less precise, the weight associated with abundance estimate data remained the high at 0.91. The weight associated with survival estimate data increased substantially to 0.81 as did the weight associated with female and male harvest data, 0.66 and 0.65 respectively. Under the more bias/less precision sex ratio data scenario the weights associated with both age ratio and sex ratio data decreased. The weight associated with age ratio data decreased to 0.66 and the weight associated with sex ratio data decreased to 0.3. The decrease in the weight associated with sex ratio data was the lowest weight associated with any data type across all scenarios. The less bias/less precision sex ratio data scenario resulted in a more even distribution of weights across all data types. The weights associated with the different data types under this scenario were as follows: abundance estimate data = 0.73, age ratio data = 0.67, sex ratio data = 0.76, survival estimate data = 0.64, female harvest data = 0.67, and male harvest data = 0.64. These results demonstrate the large influence that changes to the bias and precision attributes of sex ratio data have on how the IPM incorporates different data types in the process of estimating abundance.

The results of the analysis of female and male harvest estimate data show that changes to the bias associated with this data type have some effect on the weights associated with other data types used by the IPM. However, there was a limit to the full investigation of the influence of the bias and precision attributes of these two data types due to the lack of a measure of precision for harvest data (Table 3-5). Due to the lack of a measure of precision associated with these data, this analysis was limited to a single scenario for each harvest estimate data type that assessed the impact of a less biased estimate of harvest. When female harvest estimate data was less biased the weights associated with each data type were within a relatively small range of 17%. In this scenario, abundance estimate data was associated with the highest weight (0.83) followed by sex ratio data (0.73), male harvest estimate data (0.7), survival estimate data and female harvest estimate data which both were associated with a weight of 0.67, and age ratio data (0.66). When male harvest estimate data was less biased the distribution of weights associated with each data type encompassed a much larger range of 52%. Abundance estimate data and sex ratio data remained the two data types associated with the highest weights but with substantial increases to the weights associated with these data types. The weight associated with abundance estimate data increased to 1.0 and the weight associated with sex ratio data increased to 0.85. The weights associated with survival estimate data and age ratio data also increased to 0.77 and 0.76, respectively. Conversely, the weights associated with female and male harvest estimate data decreased substantially by 19% for female harvest data and 22% for male harvest data. While these analyses showed that changes in the bias of harvest estimation data can influence the weights associated with different data types the lack of a precision component limits the amount of inference possible from the results of this analysis.

IPM bias and precision

The impact of different scenarios, on the bias and precision of the IPM-generated estimates of abundance, varied by data type. Changes to the bias and precision of the survival data and age ratio data had a consistent impact on the bias and precision of the IPM-generated estimate of abundance (Figure 3-2). The bias and precision of the IPM-generated estimate of abundance for the six scenarios tested for these two data type clustered between bias values of 3193 to 4485 and precision values of 2775 to 3465 (Figure 3-2). These results demonstrate that changes to the bias and precision of survival data and age ratio data have a limited impact on the bias and precision of the IPM-generated estimate of abundance. The impact of changes to the bias and precision of the abundance data and sex ratio data have a much larger impact of the bias and precision of the IPM-generated estimate of abundance. Changes to the bias and precision of the abundance data had a relatively large impact on the bias of the IPM-generated estimate of abundance with the amount of bias ranging from 2 to 4958 (Figure 3-2). However, changes to the bias and precision of the abundance data had minimal impact on the precision of the IPM-generated estimate of abundance with precision measurements ranging from 2369 to 2838 (Figure 3-2). We observed the opposite pattern of impact for changes to the bias and precision of sex ratio data. Changes to the bias and precision of the sex ratio data had a relatively large impact on the precision of the IPM-generated estimate of abundance with the precision measurements from 3465 to 6209 (Figure 3-2). However, changes to the bias and precision of the sex ratio data had minimal impact on the amount of bias associated with the IPM-generated estimate of abundance. Bias in the IPM-generated estimate of abundance ranged from 0.7 to 33 for the three sex ratio scenarios tested (Figure 3-2).

DISCUSSION

Our goal is to provide information to managers and researchers so they can make informed decisions when prioritizing which data to collect and how to allocate monitoring resources. The way in which IPMs incorporate data types of different relative qualities and the influence that these data have on the estimate of abundance is of great interest when deciding which data to collect and how much of each data type is required for accurate estimates of abundance. We assessed the influence of four common data types (abundance estimates from count data, age and sex ratios, survival, and harvest) on the performance of an IPM used to estimate mule deer abundance to determine how the model uses different data types and how the quality and quantity of each data type can influence the output of the model. We demonstrated that some data types have more influence on the IPM-generated estimates of abundance than others and that the relative quality and quantity of certain data types can influence these IPM-generated estimates of abundance. The weights associated with abundance estimate data and sex ratio data fluctuated greatly with changes in the bias and precision of each data type. Conversely, we found that the weights associated with age ratio data and survival data remained predominantly stable when we made changes in their bias and precision in the scenarios we tested. Changes in the abundance data and sex ratio data also produced five of the six IPM-generated estimates of abundance with levels of bias and precision that are useful for informing management decisions. The differences between the bias and precision of estimates of abundance from the IPM that were the result of changes to the bias and precision of age ratio data and survival data were all larger than what we considered useful for informing management decisions. Our results suggest that changes in the bias and precision of abundance data and sex ratio data have the most influence on how the IPM uses different data types. Changes to the bias and precision attributes of these two data types also appeared to have the most impact on the overall accuracy of abundance estimates generated by the IPM. These results highlight the role that the bias and precision of abundance data and sex ratio data play in the generation of abundance estimates by this IPM.

Survival data

The effect of survival data on the weights associated with other data types used in the IPM is minimal. The consistent values for the weights associated with all data types in the different survival data bias and precision scenarios demonstrate that changes to survival data do not affect the influence of the other data types used by the IPM for generating estimates of abundance. The single exception to this pattern of consistency occurred in the scenario when the survival data was more biased and more precise. The more bias/more precision scenario represents one of the worst possible states that the survival data could be in for inclusion in the IPM because it suggests that the survival is much different than its true value and that there is a high degree of confidence in this estimate of survival. The consequent large reduction in the weight associated with survival data in the more bias/more precision scenario suggests that the IPM was able to ‘recognize’ that the survival data in this case was a poor representation of the actual survival rates in the population. The consistency of weights associated with different data types within the rest of the more bias/more precision scenario and across the other survival data bias and precision scenarios along with the ability to adjust the weight associated with survival data in the more bias/more precision scenario is likely due to the IPM’s ability to share information across data types. The sharing of information is possible due to the hierarchical structure of the IPM (Kéry and Royle, 2016) and because there is information related to survival contained within all other data types used in the IPM. Survival rates directly influence the number of individuals available to count in the population (abundance data) which consequently influence the estimates of age and sex ratios (age and sex ratio data). Survival also contains information related to the estimates of male and female harvest by influencing the number of individuals available for harvest. Therefore, it is possible to gain information regarding survival from all other data types included in the IPM which likely reduces the effect of changes in survival data bias and precision on the influence of other data types within the model and on survival data itself except in a worst case scenario.

The influence of changes to the bias and precision of survival data did have a more pronounced and consistent effect on the bias and precision of abundance estimates generated by the IPM. The IPM-generated estimates of abundance from all three scenarios involving changes to the bias and precision of survival estimate data were highly biased in a negative direction by a similar magnitude and had a relatively small CRI width (Figure 3-2). This suggests that changing the amount of bias and precision of the survival data, by the amounts we tested, does not influence the precision of the of the IPM abundance estimate but does influence the amount of bias in this abundance estimate.

Increased bias in the survival estimate data did appear to increase the amount bias in the IPM abundance estimate slightly more than decreasing the precision of the survival data. However, the amount of change in the bias of the IPM abundance estimate was less than 320 across all three testing scenarios suggesting that a larger set of scenarios need to be tested to find the point at which changes to the bias and precision attributes of the survival data begin to reduce the amount of bias in the IPM estimate of abundance. One possible explanation for the increased bias in the IPM estimate of abundance for the three survival data scenarios is the stable nature of the survival rates for adult female ungulates, which make up the largest proportion of the study population. Gaillard et al. (1998) showed adult female survival to be the most stable vital rate in ungulate populations as well as the most sensitive. Due to the sensitivity associated with this vital rate, we would expect small changes in adult survival to have a large impact on the estimate of abundance as shown in this analysis. In the scenarios when the survival data had more bias, we biased adult survival data by 0.15 and by 0.2 for juveniles. The value used for adult survival represents an extreme change in this vital rate with high sensitivity. Therefore the resulting large amount of bias in the abundance estimate of the two scenarios when the value of survival estimate data was more biased is not surprising. The more interesting result was the similar amount of bias in the IPM estimate of abundance when the survival data was less precise and less biased. This could be the result of the large drop in precision from the original SD (0.06) and the SD (0.45) for the less precise scenarios. Another potential source of the large amount of bias in the IPM abundance estimates may relate to the structure of the IPM which holds survival for both sexes within each age class as equal and constant. While holding survival constant for adult female survival is a reasonable assumption making adult male survival equal to females and constant as well as holding juvenile survival equal across sexes and constant may have unforeseen consequences on the bias of the abundance estimate. Finally, the large amount of bias associated with the IPM estimates of abundance from these three scenarios may be the product of the presence of unaccounted for bias in one of the other data sources.

Our results suggest that changes to the bias and precision of survival data can bias estimates of abundance generated from the IPM but have a smaller impact on the precision of the estimate. The results also demonstrate that the IPM is sharing information between different data sources and that the IPM is capable of reducing the influence of a particular data type when the information provided by that source does not have support from information contained about the same population parameter in the other data types. We suggest further testing to determine the exact magnitude of change in the bias and precision of survival data that is allowable before changes in this data source start to impact the IPM estimate of abundance to point where it is not useful for informing management.

Abundance data

The effect of the abundance data on the weights associated with data types used in the IPM varied according to data type and testing scenario. The variability in weights associated with most data types across and within the testing scenarios associated with changes to the bias and precision of the abundance data demonstrate the large influence that this data type has on the IPM. Over all three scenarios testing the influence of the bias and precision of the abundance data, the weights associated with survival data varied relatively little (0.62 – 0.7). This provides further support to the previous conclusion about the stability of the influence of survival data on the weights associated with other data types within the IPM. When the abundance data had less bias and less precision the weight associated with this data type was high (1.0). We expect the increased weight given to the abundance data because this data type contains the greatest amount of information about the population abundance. A more interesting result from the less bias/less precision scenario is the high weight associated with the sex ratio data (0.95) relative to the other data types. While the abundance data contains information related to sex ratios it also contains information about age ratios which was not associated with a high weight value. It is unclear why sex ratio estimate data would be associated with such a high weight relative to the other data types whose weights ranged from 0.53 to 0.72 in this scenario.

The distribution of weights associated with all data types were more even in the scenario when the abundance data was more biased and more precise. The weights associated with the abundance data and sex ratio data both dropped to 0.66 and 0.67, respectively. The drop in the weight associated with the abundance data in this scenario is likely due to the abundance data being further from the true population abundance but having a higher measure of precision associated with it. This provides further evidence that the IPM is capable of ‘recognizing’ when the information from a single data source is not supported by the other data included in the IPM.

In the third scenario when the abundance data is less biased and more precise, the weight associated with the abundance data increases to 0.77 but surprisingly the weights associated with both age and sex ratio data increase to 0.92 and 0.97, respectively. In this scenario, we expect the abundance data to have the highest weight associated with it because it contains the least biased and most precise information about population abundance of all data testing scenarios. However, the high weights associated with age and sex ratios in this testing scenario are unexpected. One possible reason for the high weights associated with the two types of ratio estimate data is the high measure of precision associated with the abundance data. The high precision of the abundance data may influence the weights associated with the ratios due to the interval at which the raw data for estimating abundance from the sightability model is collected. Raw abundance data is typically collected intervals in from 5 to 10 years. Therefore, the IPM must infer abundance for the years that occur in between data collection. If high measures of precision are associated with the abundance data in years where no raw data is collected the IPM recognizes this mismatch in precision and consequently reduces the weight associated with the abundance data. If the abundance data is associated with a lower weight then the IPM likely compensates by increasing the weights associated with the two ratio data types because the two ratio data types contain the most current information related to the population abundance in years when no abundance estimate data collected.

The influence of changes to the bias and precision of the abundance data had a varied effect on the bias and precision of abundance estimates generated by the IPM. The changes made to increase the precision of the abundance data in the more bias/more precision and less bias/more precision scenarios produced IPM abundance estimates with relatively small amounts of bias and precision. Across the three scenarios tested to assess the impacts of changes to the bias and precision attributes of the abundance data, the CRI width of the IPM abundance estimate remained relatively stable while the amount of bias in the IPM abundance estimate varied greatly.

The IPM-generated estimate with the least bias occurs when the abundance data was more biased and more precise. This was an unexpected result because it suggests that when abundance data is most misleading the resulting IPM estimate of abundance has the least bias and relatively high precision. This seemingly contradictory result likely relates the distribution of the weights associated with each of the data types in this scenario. The range of this particular combination of weights was the second smallest and the abundance data is associated with the smallest weight of all scenarios tested. This suggests that the IPM reduced the weight associated with the abundance data due to the increased amount of bias associated with it and that information from other data types had greater influence on the IPM-generated estimate of abundance when abundance estimate data type is more biased. It is also surprising that in this scenario the highest weights are associated with the data types that we assume to be of the lowest relative quality, female and male harvest estimate data, and the lowest weight is associated with the data type that we assume to be of the highest relative quality, survival data.

We expected the abundance data scenario with less bias and more precision to produce the IPM estimate of abundance with the smallest bias and highest precision. The expectation was that this bias and precision scenario would provide the most information regarding the population to the IPM. However under the less bias/more precision scenario it is the age and sex ratios that are associated with the highest weights and the greatest influence on the IPM estimate of abundance. The higher weights associated with the ratio data likely explains the increased bias in the IPM-generated estimate of abundance because we assume these two data types to contain increased an increased amount of bias. In this scenario abundance data is associated with one of the lowest weights it received across all scenarios tested. The low weight suggests that the increased precision associated with the abundance data in the less bias/more precise scenario does not have support from the other sources of information available to the IPM.

When the abundance data was less biased and less precise the resulting IPM-generated estimate of abundance is associated with the second largest amount of bias from all scenarios tested. This suggests that these changes in the bias and precision attributes of the abundance data result in IPM-generated estimates of abundance that are not informative for management. In the less bias/less precise scenario, the weight associated with the abundance data is one of the highest values of all scenarios tested. However, the sex ratio estimate data is also associated with the highest weight value it received over all scenarios tested. This suggests that the IPM supplements the lack of precision in the less biased/less precise abundance data with information from the sex ratio data. Since, sex ratio estimate data is typically of a lower relative quality then the bias contained in the sex ratio data may be responsible for the increased amount of bias in the IPM-generated estimate of abundance. This outcome suggests a possible interaction between the information contained in the sex ratio data and the abundance data within the IPM. The interaction may be the result of using the same model, sightability, to estimate both values and could represent the compounding of an error contained within the estimation process of the sightability model itself.

Age ratio data

Similar to survival data, the effect of changes to the attributes of age ratio data on the weights associated with data types used in the IPM is minimal. Under all three scenarios tested, the weights associated with each data type remained stable with the exception of age ratio data in the less biased/more precise scenario. Under the less biased/more precise scenario the weight associated with age ratio estimate data decreased by 17% to 19%. The decrease in the weight associated with age ratio data is unexpected because of the bias and precision attributes of the age ratio in this scenario. The decrease in weight may be due to the increased measure of precision associated with the age ratio data in this scenario not having support from the other data types. The relative consistency of the weights associated with all data types under all age ratio scenarios tested may relate to the inherent variability associated with the age ratio data type.

The variability associated with the age ratio data type reflects both the biology of this species and the observation process. The vital rate that most influences age ratios is the survival rate of the youngest age class (Harris et al. 2010).   Survival of this age class in ungulate species is expected to be highly variable (Galliard et al. 1998) due to a combination of biotic and abiotic factors including fall NDVI, snow depth, the timing of major weather events, and potentially higher rates of predation on the juvenile age class (Hurley et al. 2011). Therefore, we expect high annual variation in age ratio estimates as it represents the recruitment of approximately 6 month old individuals into the population. The observation process may also introduce variation into the age ratio data. The timing of the collection of data for estimating age ratios, in this study, is early to mid- winter which likely increases the probability of committing some level of classification error when assigning individuals to specific age classes due to the similarity in body sizes of older fawns and juvenile (yearling) females which are already ‘recruited’ into the population. The amount of classification error is likely dependent on observer experience and environmental conditions.

The influence of changes to the bias and precision of age ratio estimate data did have a consistent effect on the bias and precision of abundance estimates generated by the IPM. Abundance estimates from all three scenarios involving changes to the bias and precision of age ratio estimate data were highly biased in a negative direction by a similar magnitude and had a similar CRI width. The CRI width for all three scenarios that included changes to the bias and precision of age ratio estimate data are in the acceptable range for management but the amount of bias associated with the IPM estimates of abundance are larger than acceptable for management purposes. The IPM abundance estimates with a bias greater than 3000 suggests that these results may not be useful for informing management. The large amount of bias associated with the IPM estimates of abundance when making changes to the attributes of age ratio estimate data are likely related to this data type containing the most information about population level reproductive rates. The other data types used in this IPM contain little reproductive information therefore changes to the bias and precision of age ratio data are likely to have an increased impact on the bias of estimates of abundance generated by the IPM. However, the presence of unaccounted for bias in other data types cannot be completely ruled out as a factor impacting the amount of bias associated with the IPM abundance estimates from these scenarios.

Sex ratio data

Changes made to the bias and precision attributes of sex ratio data have the largest impact on the weights associated with all data types used in the IPM. The weights associated with each of the data types had the widest range of values of any scenarios tested across all data types. In the first scenario, less biased/more precision, the abundance data is associated with the highest weight (1.0) and survival data is associated with the lowest weight (0.49). While the high weight associated with the abundance data is expected, the low weight associated with survival data is not. We expect survival data to have a higher weight because it is the data type considered to be of the highest relative quality. Under this scenario survival data was associated with the lowest weight across all scenarios tested, except for the survival scenario when survival data was more biased and more precise. The reduction in the weight associated with survival data may be due to sex ratio data containing a relatively large amount of information pertaining to adult survival of both sexes. The decrease in bias and increase in precision of sex ratio data in this scenario may have reduced the weight associated with survival data in the IPM by placing greater emphasis on the information regarding survival contained in the sex ratio data.

When sex ratio data was more biased and less precise, the weight associated with sex ratio data was the lowest of any weight associated with any data type across all data types and testing scenarios. Conversely, survival data is associated with the highest weight for survival data across all data types and testing scenarios. This provides further support for our suggestion that the IPM is adjusting the influence that each data type has when two different data types contain information about the same population rate, adult survival, and the adjustments appear to relate to the bias and precision attributes of the type of data. When sex ratio data is more biased and less precise the IPM shifts to a heavier reliance on survival data, conversely, when sex ratio data is less biased and more precise the IPM relies more heavily on the sex ratio data for information related to adult survival.

In the final scenario when sex ratio data is less biased and less precise the weight distribution across all data types is more even. The weights of different data types in this scenario encompass a range of only 12%. As the sex ratio is less biased and less precise the weight associated with survival increases to a value that is approximately half way between its weight in the two previous bias and precision scenarios for sex ratio data. The changes in the weights associated with survival data, which track changes to the bias and precision of the sex ratio data, provide further evidence of the impact of each of these data types on one another in the IPM with regards to where the IPM is getting its information pertaining to adult survival.

The influence of changes to the bias and precision of sex ratio estimate data resulted in very low and consistent levels of bias in the abundance estimates generated by the IPM. However, the precision of IPM estimates of abundance under the sex ratio data scenarios is more variable compared to the impact that changes to the attributes of other data have on the precision of the IPM estimates of abundance. The pattern of bias and precision related to the attributes of the IPM abundance estimates in the sex ratio data testing scenarios are the reverse of the patterns that we observed when we made changes to the same attributes of other data types and their impact on the bias and precision of IPM-generated estimates of abundance.

All changes made to the bias and precision of the sex ratio estimate data resulted in IPM estimates of abundance with no or very small amounts of bias. The lack of bias in the estimates of abundance generated by the IPM when we made changes to the amount of bias in the sex ratio data suggests that this data source has a relatively small impact on the amount of bias present in IPM-generated abundance estimates. Conversely, when we made changes to the amount of precision associated with the sex ratio data the precision of the IPM-generated estimates of abundance fluctuated over a larger range. Specifically, when the precision of the sex ratio data is low the resulting estimate of abundance has less precision. These results suggest that the bias and precision attributes of sex ratio data impact the precision of IPM-generated estimates of abundance more than the bias of the IPM-generated estimates of abundance.

One potential reason for the large influence of sex ratio data on the precision of IPM estimates of abundance may be the difficulty in accurately estimating the precision of this data type in the sightability model. The precision of sex ratio data can stem from issues related to the collection of this data type. Similar to age ratio data, IDFG collects sex ratio data when mule deer are present on their winter range. One factor potentially influencing the amount of precision associated with estimates of sex ratios may be the failure to account for the presence of classification error when collecting raw sex ratio data. Classification error can stem from various sources including the similarity in body size of mature females and the younger age classes of adult males, the timing of the shedding of antlers by males during this period of the year, or observer experience. Small errors in the classification of a sample of individuals into the correct sex class can have potentially large impacts on the estimation of sex ratios (Henderson Chapter 2). The lack of a specific mechanism within the sightability model to account for classification error may result in the estimates of precision of sex ratio data containing influential amounts of error and consequently influencing the precision of the IPM abundance estimate.

Harvest data

Changes made to the bias of male and female harvest data have varied impacts on the weights associated with all data types used in the IPM. When male harvest data is less biased the weights associated with each data type are highly variable. The weights associated with both harvest data types in the scenario when male harvest has less bias are two of the lowest weights associated with harvest data types across all testing scenarios. This suggests that the less biased male harvest data did not agree with the information about harvest provided by other data types. Specifically, the weight associated with the abundance data increases to the highest possible value (1.0) and the weight associated with sex ratio estimate is also increases to a relatively high value (0.85). The increase in the weights associated with the abundance data and sex ratio data suggest that the IPM relies more heavily on these data when the male harvest data does not agree with the other data types containing information about harvest. The shift to reliance on other data in this scenario may be the result of the lack of a measure of precision for the male harvest data.

Conversely, in the scenario when female harvest data was less biased the range of weights associated with different data types is relatively small. The more even distribution of weights across data types suggests an increased amount of information about females available in other data types. Abundance data, ratio data, and survival data all contain information about the number of adult females in the population. Therefore, the impact of a change in the bias of female harvest data may be diffused across other data sources used in the IPM resulting in the more even distribution of weights associated with each data type. It is unclear why the weight associated with male harvest data increases in this scenario. However, the weights of both harvest data types are approximately equal and increase or decrease at approximately the same magnitude in almost every scenario. The equivalent weights associated with these two harvest data types across testing scenarios may be the result of the attributes of these data such as the lack of a measure of precision and a single count value.

IPM bias and precision

Changes made to the bias and precision of the male and female harvest data resulted in two opposite impacts on the bias of the estimates of abundance. When male harvest data was less biased, the amount of bias associated with the estimate of abundance is the largest and the CRI width is the smallest of all scenarios tested. These results suggest that the IPM abundance estimate from this scenario contains the least amount of information for management decisions. Male harvest data is also associated with a relatively low weight in this scenario and the weights associated with abundance estimate data and sex ratio data are relatively high. This suggests that either or both abundance data and sex ratio data may have a substantial amount of bias associated with them that is unaccounted for in this scenario thus leading to the large amount of bias in the abundance estimate.

In the scenario when female harvest data was less biased the estimate of abundance is one of the most informative for management purposes. This result may be linked to the relatively small range of weights associated the data types in the model which suggests a greater sharing of information across data in the IPM thus resulting in a relatively higher quality estimate of abundance.

Issues and further research

The results of these analyses provide evidence confirming how different data types interact within the IPM and the impact that different data types have on the estimates of abundance generated by this model. However, the lack of clarity in some of the results of the testing scenarios is likely due to the use of estimates made from real data for all but the data type analyzed in the specific scenario. For example, when we decreased the bias of the abundance data the estimates from all other data types are the result of the analysis of real data. We expect that there are unaccounted for sources of error affecting the bias and precision of the estimates from other data types. We did this on purpose to provide an opportunity to assess how the IPM functions under conditions that better represent typical conditions encountered by researchers and managers. Quite often elements out of the control of the researcher or manager impact the quality and quantity of data available to estimate population level attributes. Therefore, the incorporation of estimates based on the most realistic data are important to assess the how the IPM functions under realistic rather than idealistic conditions. We suggest testing a larger variety of scenarios which encompass more and finer gradations in the changes made to bias and precision of the estimate from each data type will improve the inferences made from this type of analysis. Deeper analysis in this manner can provide a more detailed picture of the bias and precision values at which the influence of each data type increases or decreases the bias and precision of the IPM abundance estimate appropriate to the management of the species under consideration. The gradations used for this analysis are specific to mule deer and potentially similar ungulate species but may not be appropriate for species with other life history characteristics.

Conclusions

Arguably the most important data types to improve the accuracy of IPM abundance estimates are abundance data and sex ratio data when substantial prior information exists for survival. Improvements in the bias and precision of these data types are associated with the IPM-generated estimates of abundance that have the least amounts of bias and the highest precision. The results are not surprising for abundance data since it contains the most information about the parameter of interest, abundance. Specifically, increases in the precision of the abundance reduced the bias in the IPM estimate of abundance.

Abundance estimate data are expensive to collect and the intervals between collection ranges from 5 to 10 years, therefore the IPM must infer between widely spaced data points to fill in abundance data for years when raw data is not collected. The infrequency of the collection of raw abundance data makes improving the bias and precision related to the abundance estimate used as data by the IPM even more important. There are several approaches that could help improve the bias and precision attributes of this influential data type. These approaches include increasing sampling intensity and frequency, improve upon the sightability model used to generate abundance data from raw abundance observations, and alter the timing when raw abundance data is collected. While increasing sampling intensity and frequency is an obvious way to improve the bias and precision associated with any estimate, it is also a much more expensive way to improve the attributes of this data type. Improving upon the existing sightability model may also be cost prohibitive given the limited resources available to monitor populations. The third option of changing the timing of when count data is collected is likely the least expensive option to improve the attributes of abundance estimate data (Keegan et al. 2011).

Collecting raw count data via aerial surveys during the period in which there is a higher probability of observing a greater portion of the population more easily is typically hampered by the overlap of this time period with hunting season. Recent improvements in technology and estimation methods provide alternatives to collecting raw abundance data from aerial surveys during hunting season. Two alternative methods are the estimation of the abundance of unmarked populations using data collected by camera trapping (Moeller et al. 2018) and DNA sampling from fecal matter (Goode et al. 2014, Ebert et al. 2012). Both of these methods have attributes that may make them preferable to aerial surveys and include: 1) a general reduction in the need for aerial surveys which reduces risk to personnel, 2) an increase in the frequency at which raw count data can be collected, 3) neither method interferes directly with hunting so these methods can be implemented during the time period during which this data type may be most representative of the entire population, 4) both have the potential to be more cost effective over time relative to aerial surveys. The results of this analysis demonstrate the importance of abundance data to the generation of low biased and highly precise estimates of abundance by the IPM. Therefore, any improvement to the current abundance data or the addition of other supplemental sources of data related to abundance can only improve the accuracy of the IPM-generated estimates of abundance and consequently the information available for managing wildlife populations.

Sex ratio data is also a highly influential data type on the IPM-generated estimate of abundance. Changes in the bias and precision attributes had a large influence of the IPM estimate of abundance. Changes to the bias and precision of sex ratio data resulted in both the largest and smallest ranges of weights in all of the scenarios tested, three of the IPM estimates of abundance with the least amount of bias, and the largest range of changes in the precision of the IPM estimate of abundance related to changes in the attributes of a single data type. The results of the analysis of sex ratio data suggest that it contains a considerable amount of bias. Over all of the scenarios the amount of bias in the IPM estimate of abundance decreases to less than 33 individuals when the weight associated with the sex ratio is less than 0.76. Conversely, when weight of the sex ratio estimate data is larger than 0.85 the bias of the IPM estimate of abundance is always greater than 3193. The large fluctuations in the weights associated with other data types as the bias and precision attributes of sex ratio data change demonstrate that the IPM is adjusting the influence of other data types to compensate for the bias in sex ratio data by using information about the sex ratio contained in other data types, specifically abundance data and harvest data. The precision associated with sex ratio estimate data appears to directly influence the precision of the IPM estimates of abundance. As sex ratio data precision increases the precision of the IPM abundance estimate increases and vice versa. Therefore improving the precision of the sex ratio data should help to improve the precision of the IPM estimate of abundance.

Potential ways to improve the upon the bias and precision of the sex ratio estimate include the same methods as suggested for improving the abundance data, accounting for unequal detection probabilities between sexes, and one additional component specific to classification. Sex ratio data depends on the accurate classification of individuals to estimate sex ratios and previous work shows that even small differences in the probability of correctly classifying a single age and sex class can introduce a relatively large amount of bias (Henderson Chapter 2). Therefore, we suggest the inclusion of a term relating to classification error in the sightability model or the implementation of the ratio estimation method developed by Henderson (Chapter 2). Either approach would help to increase the accuracy of the sex ratio data used by the IPM. Improvements to the precision of the sex ratio data will result in improving both the precision of the IPM estimate of abundance which sex ratio data has a large influence upon.

The conclusions regarding the impact of changes to the bias and precision of survival data and age ratio data are similar. The influence of changes to the bias and precision attributes of age ratio and survival data have little impact on the weights associated with other data types within the IPM. The single instance of substantial change to occurs in the scenarios that age ratio and survival data are more precise. In both scenarios the weights associated with age ratio and survival data decrease by 18 to 26%, respectively, while the weights associated with all other data types remain relatively constant. The manipulation of the bias and precision attributes of age ratio and survival data resulted in IPM-generated estimates of abundance that are consistently biased low by several thousand individuals. The precision of the IPM abundance estimates from these scenarios remained fairly constant and within an acceptable range for management. We suggest the large amounts of bias in the IPM estimates of abundance from the age ratio and survival data scenarios are the result of the large weights associated with the sex ratio data in these scenarios. The large influence of the sex ratio data in these scenarios may introduce unaccounted for bias via the sex ratio data.

The results of the analysis of male and female harvest data are mixed. Less biased female harvest data results in a more even distribution of weights associated with data types in the IPM. The effect of female harvest data that is less biased led to an IPM estimate of abundance with virtually no bias and relatively high precision. The relatively small range of weights associated with different data types coupled with the relatively high accuracy of the IPM estimate of abundance from this analysis may be a reflection of the potentially large influence of the female portion of the population on the dynamics of this deer population. Conversely, male harvest data that is less biased results in a large range of weights associated with different data types in the IPM, a low weight associated with male harvest data, and an estimate of abundance from the IPM that had the largest bias and highest precision of any scenario tested. The low weight associated with male harvest data and the wide range of weights associated with the other data types suggest the less biased value for male harvest data did not agree with information regarding males contained in other data types. The large impact on the accuracy of the IPM estimate of abundance from male harvest data may be due to how the IPM incorporates harvest data. This version of the IPM holds harvest data constant from year to year which may be an acceptable assumption for female harvest but is likely a poor assumption for male harvest. Therefore a constant rate for male harvest data may result in increased bias in the estimate of population abundance. Changes to the quality of harvest data for both males and females appears to have the potential to influence the amount of bias in the IPM estimate of abundance therefore increasing the accuracy of this data type could be important to improving the accuracy of IPM estimates of abundance. However, due to the lack of any measure of precision associated with the harvest data any inference based on the results of this analysis are to be made cautiously. Ongoing improvements in the estimation of harvest data will make a more meaningful analysis of the role of these data in the generation of IPM estimates of abundance possible in the future.

The methods used in this analysis have the potential to improve IPMs used for other species and for estimating other population parameters. The assessment of the weights associated with each data type under various scenarios gives insight into how the IPM shares information between data types and suggests which types of data have the most influence on the bias and precision of the estimates from the IPM. Using these methods as a template, managers and researcher can assess the influence of the data types currently available for their species of interest on IPM parameter estimates, determine how changes to the bias and precision attributes of various data types influence the results generated from the IPM, and explore the potential impacts of additional data types under various collection effort scenarios on IPM-generated estimates. Currently, the statistical methods used to assess IPM goodness of fit are limited to a single formal method that uses calibrated simulation (Besbeas and Morgan 2014). The methodology and results of this analysis have the potential to provide a starting point for the development of another formal goodness of fit test for IPMs. Finally, this methodology for assessing how IPMs use data and the impacts of data with varying levels of bias and precision on the estimates generated by an IPM provide a transparent and defensible source of information on which to base decisions regarding the allocation of monitoring resources for wildlife populations.

LITERATURE CITED

Abadi, F., O. Gimenez, R. Arlettaz, and M. Schaub. 2010a. An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. Ecology 91:7-14.

Ahrestani, F.S., Hebblewhite, M. and Post, E., 2013. The importance of observation versus process error in analyses of global ungulate populations. Scientific reports, 3, p.3125.

Besbeas, P., S. N. Freeman, B. J. T. Morgan, and E. A. Catchpole. 2002. Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. Biometrics 58:540-547.

Besbeas, P., Lebreton, J.D. and Morgan, B.J., 2003. The efficient integration of abundance and demographic data. Journal of the royal statistical society: series C (applied statistics), 52(1), pp.95-102.

Besbeas, P. and B. J. T. Morgan. 2014. Goodness-of-fit of integrated population models using calibrated simulation. Methods in Ecology and Evolution 5, 1373-1382.

Bowyer, R.T., 2004. Sexual segregation in ruminants: definitions, hypotheses, and implications for conservation and management. Journal of Mammalogy, 85(6), pp.1039-1052.

Brooks et al. 2004 Brooks, S. P., R. King, and B. J. T. Morgan. 2004. A Bayesian approach to combining animal abundance and demographic data. Animal Biodiversity and Conservation 27:515-529.

Caughley, G., Sinclair, R. and Scott-Kemmis, D., 1976. Experiments in aerial survey. The Journal of Wildlife Management, pp.290-300.

Doak, D. F., K. Gross, and W. F. Morris. 2005. Understanding and predicting the effects of sparse data on demographic analyses. Ecology 86:1154-1163.

Ebert, C., Sandrini, J., Spielberger, B., Thiele, B. and Hohmann, U., 2012. Non-invasive genetic approaches for estimation of ungulate population size: a study on roe deer (Capreolus capreolus) based on faeces. Animal Biodiversity and Conservation, 35(2), pp.267-275.

Fieberg, J., Alexander, M., Tse, S. and St. Clair, K., 2013. Abundance estimation with sightability data: a Bayesian data augmentation approach. Methods in Ecology and Evolution, 4(9), pp.854-864.

Forrester, T.D. and Wittmer, H.U., 2013. A review of the population dynamics of mule deer and black-tailed deer Odocoileus hemionus in North America. *Mammal Review*, *43*(4), pp.292-308.

Francis, R.C., 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences, 68(6), pp.1124-1138.

Gaillard, J.M., Festa-Bianchet, M. and Yoccoz, N.G., 1998. Population dynamics of large herbivores: variable recruitment with constant adult survival. Trends in Ecology & Evolution, 13(2), pp.58-63.

Geist, V. 1981. Behavior: adaptive strategies in mule deer. Pp. 157-224, in Mule and black-tailed deer of North America (O. C. Wallmo, ed.). Univ. Nebraska Press, Lincoln, xvii + 605 pp. Harris et al. 2010

Goode, M.J., Beaver, J.T., Muller, L.I., Clark, J.D., Manen, F.T.V., Harper, C.A. and Basinger, P.S., 2014. Capture—recapture of white-tailed deer using DNA from fecal pellet groups. Wildlife biology, 20(5), pp.270-278.

Hurley, M.A., Unsworth, J.W., Zager, P., Hebblewhite, M., Garton, E.O., Montgomery, D.M., Skalski, J.R. and Maycock, C.L., 2011. Demographic response of mule deer to experimental reduction of coyotes and mountain lions in southeastern Idaho: Réponse Démographique du Cerf Mulet à la Réduction Expérimentale des Populations de Coyotes et de Pumas dans le Sud de l'Idaho. *Wildlife Monographs*, *178*(1), pp.1-33.

Idaho Department of Fish and Game. 2008. Mule Deer Management Plan 2008-2017. Idaho Department of Fish and Game, Boise, USA.

Keegan T. W., B. B. Ackerman, A. N. Aoude, L. C. Bender, T. Boudreau, L. H. Carpenter, B. B. Compton, M. Elmer, J. R. Heffelfinger, D. W. Lutz, B. D. Trindle, B. F. Wakeling, and

B. E. Watkins. 2011. Methods for monitoring mule deer populations. Mule Deer Working Group, Western Association of Fish and Wildlife Agencies, USA.

Kendall, W.L., Peterjohn, B.G. and Sauer, J.R., 1996. First-time observer effects in the North American breeding bird survey. The Auk, pp.823-829.

Kéry, M. and Royle, J.A., 2016. Applied hierarchical modeling in ecology. Academic, Waltham.

Mackie, R.J., Pac, D. F., Hamlin, K. L., and Dusek, G. L., 1998. *Ecology and management of mule deer and white-tailed deer in Montana*. Montana Fish, Wildlife, and Parks, Wildlife Division.

Maunder, M.N. and Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research, 142, pp.61-74.

Moeller, A.K., Lukacs, P.M. and Horne, J.S., 2018. Three novel methods to estimate abundance of unmarked animals using remote cameras. Ecosphere, 9(8).

Nowak, J.J., Lukacs, P.M., Hurley, M.A., Lindbloom, A.J., Robling, K.A., Gude, J.A. and Robinson, H., 2018. Customized software to streamline routine analyses for wildlife management. *Wildlife Society Bulletin*, *42*(1), pp.144-149.

Plummer, M., 2003, March. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing (Vol. 124, No. 125.10).

Royle, J.A. and Dorazio, R.M., 2008. Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Elsevier.

Samuel, M. D., E. O. Garton, M. W. Schlegel, and R. G. Carson. 1987. Visibility bias during aerial surveys of elk in northcentral Idaho. Journal of Wildlife Management 51:622-630.

Thompson, S. K. 2002. Sampling. Second edition. John Wiley and Sons, Inc., New York, New York, USA.

Unsworth, J. W. 1994. Mule Deer Sightability: Project W-160-R-21. Idaho Department of Fish and Game, Boise, USA.

White, G. C. 2000. Modeling population dynamics. Pages 84-107 in S. Demarais and P. R. Krausman, editors. Ecology and Management of Large Mammals in North America.

White, G.C. and Burnham, K.P., 1999. Program MARK: survival estimation from populations of marked animals. Bird study, 46(sup1), pp.S120-S139.Prentice-Hall, New Jersey, USA.

TABLES

Table 3-1. Results of the bias and precision scenarios assessing the impact of changes to the bias and precision attributes of survival estimate data (S) on the weights associated with all data types and the precision and bias of the IPM-generated estimate of abundance. Definitions of abbreviations used in the table are as follows: N = abundance data, YF = age ratio estimate data, MF = sex ratio estimate data, S = survival estimate data, HARVF = female harvest data, HARVM = male harvest data, BIAS = bias associated with the IPM-generated estimate of abundance, and CRI = 95% credible interval width of the posterior distribution of the IPM-generated estimate of abundance.

Scenario	N	YF	MF	S	HARVF	HARVM	BIAS	CRI
more bias/more precision	0.97	0.86	0.89	0.42	0.59	0.58	-4485	3438
less bias/less precision	0.99	0.90	0.91	0.68	0.58	0.54	-4167	3461
more bias/less precision	0.99	0.90	0.91	0.68	0.58	0.52	-4179	3453

Table 3-2. Results of the bias and precision scenarios assessing the impact of changes to the bias and precision attributes of abundance data (N) on the weights associated with all data types and the precision and bias of the IPM-generated estimate of abundance. Definitions of abbreviations used in the table are as follows: N = abundance data, YF = age ratio estimate data, MF = sex ratio estimate data, S = survival estimate data, HARVF = female harvest data, HARVM = male harvest data, BIAS = bias associated with the IPM-generated estimate of abundance, and CRI = 95% credible interval width of the posterior distribution of the IPM-generated estimate of abundance.

Scenario	N	YF	MF	S	HARVF	HARVM	BIAS	CRI
less bias/less precision	1.00	0.72	0.95	0.62	0.57	0.53	-4658	2651
more bias/more precision	0.65	0.73	0.67	0.64	0.71	0.78	2	2838
less bias/more precision	0.77	0.92	0.97	0.70	0.52	0.41	-2326	2369

Table 3-3. Results of the bias and precision scenarios assessing the impact of changes to the bias and precision attributes of age ratio estimate data (YF) on the weights associated with all data types and the precision and bias of the IPM-generated estimate of abundance. Definitions of abbreviations used in the table are as follows: N = abundance data, YF = age ratio estimate data, MF = sex ratio estimate data, S = survival estimate data, HARVF = female harvest data, HARVM = male harvest data, BIAS = bias associated with the IPM-generated estimate of abundance, and CRI = 95% credible interval width of the posterior distribution of the IPM-generated estimate of abundance.

Scenario	N	YF	MF	S	HARVF	HARVM	BIAS	CRI
less bias/more precision	1.00	0.71	0.92	0.62	0.53	0.49	-3918	2775
more bias/less precision	0.99	0.90	0.90	0.67	0.59	0.51	-3193	3363
less bias/less precision	0.99	0.88	0.90	0.66	0.57	0.51	-4129	3465

Table 3-4. Results of the bias and precision scenarios assessing the impact of changes to the bias and precision attributes of sex ratio estimate data (MF) on the weights associated with all data types and the precision and bias of the IPM-generated estimate of abundance. Definitions of abbreviations used in the table are as follows: N = abundance data, YF = age ratio estimate data, MF = sex ratio estimate data, S = survival estimate data, HARVF = female harvest data, HARVM = male harvest data, BIAS = bias associated with the IPM-generated estimate of abundance, and CRI = 95% credible interval width of the posterior distribution of the IPM-generated estimate of abundance.

Scenario	N	YF	MF	S	HARVF	HARVM	BIAS	CRI
less bias/more precision	1.00	0.74	0.76	0.49	0.59	0.59	-13	3465
more bias/less precision	0.91	0.66	0.30	0.81	0.66	0.65	-33	6209
less bias/less precision	0.73	0.67	0.76	0.64	0.67	0.64	0.7	5925

Table 3-5. Results of the bias and precision scenarios assessing the impact of changes to the bias and precision attributes of male and female harvest data (HARVM, HARVF) on the weights associated with all data types and the precision and bias of the IPM-generated estimate of abundance. Definitions of abbreviations used in the table are as follows: N = abundance data, YF = age ratio estimate data, MF = sex ratio estimate data, S = survival estimate data, HARVF = female harvest data, HARVM = male harvest data, BIAS = bias associated with the IPM-generated estimate of abundance, and CRI = 95% credible interval width of the posterior distribution of the IPM-generated estimate of abundance.

Scenario	N	YF	MF	S	HARVF	HARVM	BIAS	CRI
less bias	0.83	0.66	0.73	0.67	0.67	0.70	-0.37	3164
less bias	1.00	0.76	0.85	0.77	0.48	0.48	-8273	2245

FIGURES

Figure 3-1. Map of Mule deer Population Management Units (PMU, Idaho Department of Fish and Game 2008).

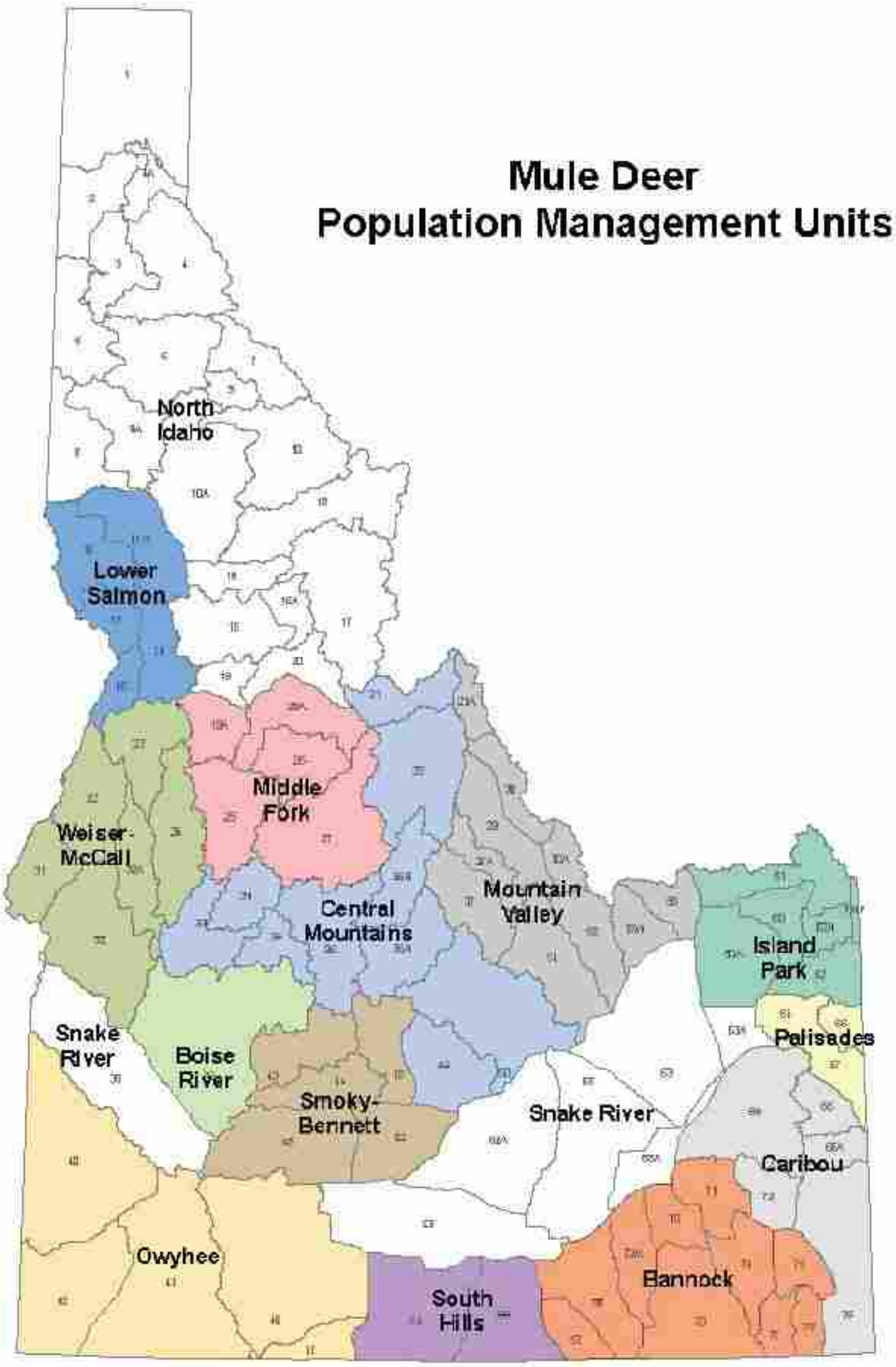
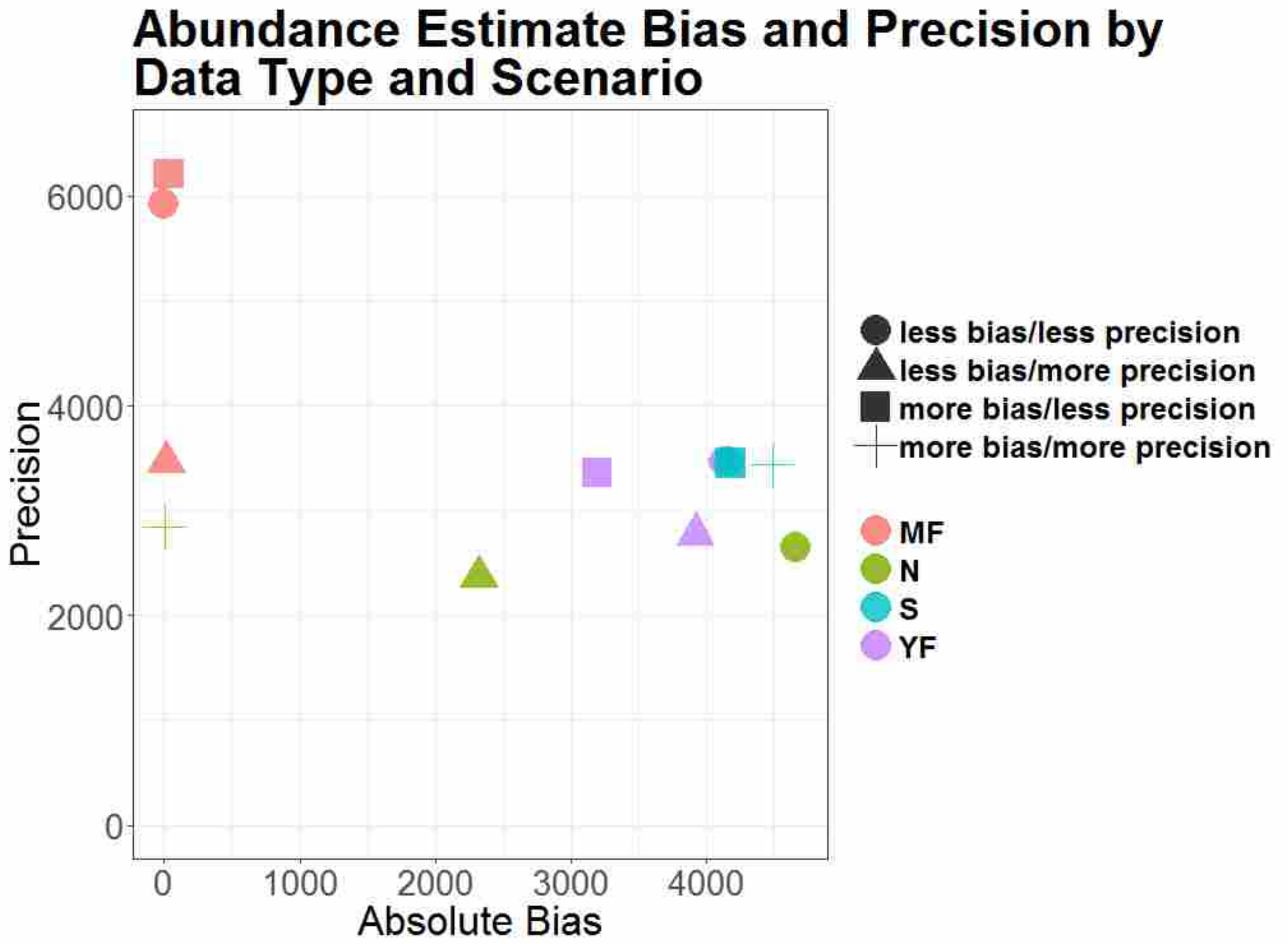


Figure 3-2. IPM generated abundance estimates from the bias and precision scenarios tested for each data type. Shapes indicate which combination of bias and precision for each scenario tested and color indicates the data type manipulated for the specific bias and precision scenario. Data types shown in this graph are estimates generated by sub-models and used as data by the IPM. Absolute bias indicates the distance between the IPM-generated estimate of abundance and the estimate of abundance from the population matrix model. The 86% of these IPM-generated estimates of abundance are biased in a negative direction. Precision indicates the width of the 95% credible interval from the posterior distribution of each IPM-generated estimate of abundance. MF = Sex ratio data, N = abundance data, S = survival data, and YF = age ratio data.



CHAPTER 4: OPTIMAL ALLOCATION OF MONITORING RESOURCES FOR COLLECTING DATA TO ESTIMATE POPULATION SIZE

INTRODUCTION

The ability of wildlife agencies to accurately monitor population trends is of the utmost importance (Field et al. 2005). We define monitoring here as the recurrent recording of certain attributes of a system of interest, in order to determine if the system attributes match previously defined norms or lie within predetermined limits (Hellawell 1991). Through population monitoring wildlife agencies determine whether the population of interest is in the desired state, detect the impacts of disturbances and changes to the ecological system on the population, and assess the impact of management actions and decisions (Legg and Nagy 2005). Monitoring programs provide information to wildlife agencies on which to base their decision making regarding conservation and management actions even when perfect knowledge of system dynamics in which these populations reside is lacking (Chades et al. 2008). As well as being a tool to inform the management of wildlife, monitoring programs also play a critical role in fulfilling the legal mandate of many of the agencies responsible for the conservation and management of wildlife populations.

State laws establish many state wildlife agencies and thus the agencies are required to fulfill the legal mandate under which they were established. One example of the legal requirement of wildlife management agencies to monitor wildlife population is found in the section of Idaho state code which articulates that as an agency the Idaho Department of Fish and Game (IDFG) shall "…preserve, protect, and perpetuate […] wildlife, [to] provide for the citizens of this state and, as by law permitted to others, continued supplies of […] wildlife for hunting, fishing and trapping" (Idaho Code Section 36-103). In order to fulfill this and similar legal mandates, wildlife management agencies rely on population monitoring to determine what if any conservation or management actions are required to maintain and sustain sufficient population sizes of wildlife. While many methods exist for monitoring wildlife populations, implementing these methods on a landscape scale can be prohibitively expensive. Therefore, it is paramount to optimize the efficiency of implementing these methods to insure the largest return on the investment of the limited public resources available for wildlife monitoring (Field et al. 2005).

Resources to manage and monitor wildlife species are limited worldwide (Chades et al. 2008) and are becoming less available in the U.S. The U.S. Fish and Wildlife Service (2018) reports a decline in hunting license holders from a peak number of 16,748,541 in 1982 to 15,620,578 in 2018. The decline in license holder numbers is important because the majority of funding (58.8%) for state wildlife agencies comes from hunting and fishing related activities, either directly through licenses sales or indirectly through excise taxes on hunting and fishing equipment (Assoc. of Fish and Wildlife Agencies 2017). The decrease in available resources for wildlife monitoring is further exacerbated by the increasing cost of monitoring. In 1998, the total estimated cost to conduct game surveys in nine western states including: Arizona, Colorado, Idaho, Montana, New Mexico, Oregon, Utah, Washington, and Wyoming in was approximately $7M (Rabe et al. 2002). The cost in 2018 for these same surveys accounting only for inflation was >$11M (US Dept. of Labor 2018). Therefore, one of the largest problems facing wildlife management agencies is how to optimally allocate limited monitoring resources and how to balance the trade-off between managing, monitoring and doing nothing (Chades et al. 2008).

We address the problem of optimal allocation of monitoring resources for wildlife populations. We use data previously collected on the mule deer (*Odocoileus hemionus*) population in Idaho as an example. IDFG collected these data over the past decade (2008-2018) and used them to estimate the size of regional and statewide mule deer populations. The types of data collected include adult and juvenile survival, population count data, and herd composition data. Survival data comes from a statewide radio-collaring and monitoring program and is collected for adult females and juveniles of both sexes. IDFG collects count data and herd composition data via aerial surveys conducted when mule deer populations are on their winter range. A sightability model that incorporates information on variables that influence the detectability of mule deer uses these raw data to estimate abundance and age and sex ratios (Unsworth et al. 1994). IDFG currently combines abundance, ratio, and survival estimates to estimate abundance using an integrated population model (IPM; Nowak et al. 2018).

We strive to develop a methodology to determine the most efficient use of monitoring resources for collecting data to estimate the abundance of wildlife populations. In order to achieve this goal, we devise a metric that describes the relationship between resource use and the precision of abundance estimates. We focus on the precision of the estimate of abundance because we assume the quality of the sampling design drives the bias in an estimate of abundance. Based on this assumption, we expect the relative cost per observation to reflect the potential for bias in the data. We develop a predictive model to quantify the expected results of different data collection scenarios and rank the data collection scenarios according to their efficiency at producing precise estimates of abundance. The expected result of the methodology developed here is the production of a set of data collection scenarios which simultaneously minimize resource use and maximize the precision of estimates of population abundance. We expect this methodology to provide decision makers a set of options for consideration when making decisions regarding the allocation of monitoring resources for tracking mule deer abundance. In addition, we expect this methodology to be adaptable to a wide range of species and able to function at several different scales. Finally, the results of this study should provide wildlife management agencies with a method for allocating limited monitoring resources that is defensible and adjustable according to the changing levels of resources available for wildlife monitoring.

METHODS

The data for this analysis comes from the following mule deer population management units (PMU) located in the central and southern regions of Idaho: Boise River, Island Park, Middle Fork, Palisades, Smoky-Bennett, and South Hills (Figure 4-1). These PMUs are located within three of the predominant ecotypes found in the state of Idaho where mule deer are present: conifer, aspen, and shrub-steppe (Hurley et al. 2017). We used these data in a multistep process to determine what data collection scenarios represent the optimal allocation of monitoring resources for each PMU. The optimal scenarios are specific to each PMU to accommodate differences in the cost of collecting each data type across with changing environmental attributes and on the varying amounts of monitoring resources available at the PMU scale (Figure 4-2).

The first step in determining the optimal scenarios for resource allocation was to generate estimates from each type of raw data using the model appropriate for that data type and varying the amount of raw data available for each estimate. For example, we used a known fate model to generate estimates of adult and juvenile survival (White and Burnham 1999). We used a known fate model because the raw survival data collected was from radio collars which allowed for the fate of the majority of individuals to be determined without error. Next, we repeatedly estimated survival for different age classes in the population while changing the amount of raw data, individual encounter histories, available to the known fate model. We varied the number of individual encounter histories for each age class from 10 to 150 in increments of 10 encounter histories. Encounter histories were drawn from the available survival data for adult females and juveniles in a given PMU and year. When the number of encounter histories needed for a specific data amount scenario exceeded the number of encounter histories available from a PMU in a given year, we added additional encounter histories to the sample by randomly drawing from the existing encounter histories for that PMU and year. The results of these survival analyses provided data for use in the IPM which included both an estimate of survival and a measure of precision that was unique to the PMU and number of encounter histories used to estimate survival.

We then estimated abundance for each PMU using raw count data and the current formulation of the sightability model used by IDFG (Unsworth et. al. 1994). We created a set of abundance estimates for each PMU for every year that IDFG collected count data in that PMU. We repeatedly estimated abundance for each PMU and year using a portion of the raw count data collected for that specific year and PMU in intervals of 10% using from 10% to 100% of the raw count data available. We estimated age and sex ratios from the raw herd composition data using the same sightability model and methodology as was employed to estimate abundance from the raw count data (Unsworth et. al. 1994).

We gathered the results of the individual PMU estimation processes using varying amounts of data for age and sex ratios, survival, and abundance into a single database to estimate abundance for each PMU for every year from 2012 to 2018. We used the individual estimates from each data type as data for estimating abundance using IDFG's current IPM developed by

Nowak et al. (2018). There were 3000 possible available data-scenario combinations for each PMU per year. We passed the results from each data-scenario combination to the IPM to estimate abundance by PMU and year. The specific formulation of the IPM used in this analysis held adult survival constant, allowed juvenile survival to vary with time, and allowed recruitment to vary with time (Nowak et al. 2018). The estimates of abundance were generated using the statistical computing language and environment R (R Core Team 2017) and the JAGS program for Bayesian analysis (Plummer 2003). We ran each IPM model for 40,000 iterations with a burn in period of 10,000 iterations. Convergence was assessed using the R-hat metric or Gelman-Rubin statistic (Gelman and Rubin 1992).

We paired IPM estimates of abundance from the different data-collection scenarios with cost data to quantify the relationship between the cost of gathering different types and quantities and the resulting changes to the precision of the mule deer abundance estimate generated by the IPM. IDFG provided information regarding the total cost to conduct both count and herd composition surveys using the Bell 47 helicopter and IDFG personnel (pers. comm. Mark Hurley). The wide variety of radio-collar types and fluctuating cost of radio-collar technologies used during the time period in which IDFG collected the data for this analysis made it difficult to assign a specific monetary value to each encounter history used for the survival analysis. Therefore, we used a consistent, mid-range value of \$1200 as the cost for deploying a single radio collar to collect survival data for each encounter history regardless of PMU and year. We also compared how changes to the price of survival impacted the set of optimal data collection scenarios. We conducted this analysis for a single PMU, the Middle Fork, and used prices for a single encounter history ranging from \$200 to \$4000. We used the cost data and number of observations to calculate the average cost per observation by PMU and data type (Table 4-1).

We calculated the total amount of money spent collecting data for each IPM-generated estimate of abundance for each data-collection scenario using the cost per observation of each data type. The total cost of data collection, the IPM-generated estimate of abundance, and the 95% credible interval width (CRI) of the posterior distribution of the estimate of abundance were combined to estimate the information gain (IG) ratio for each data availability scenario.

We created the information gain ratio (IG) to describe the relationship between the amount of resources used collecting data and the resulting precision of the IPM-generated estimate of abundance. The formula we developed to calculate the IG ratio was

$$IG = \frac{1}{\left(cost_{total} * \frac{CRI}{N}\right)} * 10000$$

where *IG* is the information gain ratio, $cost_{total}$ is the total cost of collecting all data types for a given data-collection scenario, *CRI* is the 95% credible interval width of the posterior distribution of the estimate of abundance, and *N* is the point estimate of abundance generated by the IPM. We multiplied the ratio by 10,000 for ease of interpretation. Lower $\frac{CRI}{N}$ values indicate more precise estimates of abundance. High IG ratio values indicate an increase in the amount of information gained regarding abundance, relative to the cost of collecting the data needed to produce an estimate of abundance with a particular level of precision. For example, if $\frac{CRI}{N}$ has a value of 0.01 and the cost of collecting the data to produce an estimate of abundance with that precision is $100,000 then, the IG ratio value is 10. If it costs $200,000 to produce an estimate of abundance with a $\frac{CRI}{N}$ value of 0.01 then the value of the IG ratio is 5. Figure 4-3 demonstrates the behavior of the IG ratio value as the $\frac{CRI}{N}$ increases from 0.01 to 1 (precision of the abundance estimate decreases) and the amount of money available for data collection increases in $100,000 intervals from $100,000 to $500,000. The shape, negative exponential, of the relationship between the $\frac{CRI}{N}$ remains stable but the magnitude of change in the IG ratio decreases as the amount of money used to collect data increases. The IG ratio was calculated for each data-collection scenario using: 1) the IPM generated point estimate of abundance and the 95% CRI of the IPM estimate of abundance generated using the specific combination of data amounts and types for each data collection scenario and 2) the combined PMU specific cost of collecting the type and amount of data for that data-collection scenario.

The next step in determining the optimal allocation of resources for collecting data to estimate abundance for each PMU was to use the IG ratio values calculated from the different data-collection scenarios to develop a model for predicting IG ratio values under different resource availability conditions. We developed a generalized additive model (GAM) for each PMU to predict IG ratio values under changing resource availability scenarios. We selected the GAM for use due to this model's ability to accommodate the non-linear relationship between the IG ratio value and the varying amounts of each data type without including polynomial terms in the regression equation (Wood 2017). While lacking the typical output coefficients for explanatory variables that are the product of linear regression models, GAMs allow for an improved interpretation of the non-linear relationship between the dependent variable and the explanatory variables. Graphing the expected values of the dependent variable over a large range of explanatory variable values facilitates the interpretation of the influence of explanatory variables on the dependent variable (Larson 2015). Graphing highlights the changing, non-linear relationship between the explanatory and dependent variables. We used the mgcv R package (Wood 2011, Wood 2017) to develop and predict the relationship between the IG ratio value and the varying amounts of survival, count, and herd composition data. The smoothing method used was the cubic regression spline and the number of knots for each smoothing function was determined by comparing Akaike information criteria (AIC) values and generalized cross validation (GCV) values between models with different knot values for each explanatory variable. We configured all models compared via AIC and GCV in the following format

$$IG = \alpha + s_1(survival) + s_2(count) + s_3(composition)$$

where IG is the information gain ratio, α is the y-intercept, $survival$ is the number of encounter histories for both adult females and juveniles of either sex that were included as data, $count$ is the number of groups observed during abundance surveys, $composition$ is number of groups observed during herd composition surveys, s_x is the smoothing function for a specific explanatory variable. We selected a separate GAM for each PMU due to differences in the environmental attributes of each PMU impacting the cost of collecting data, the changes in deer population density across PMUs, and the collection frequency of different data types across PMUs.

After we selected a single GAM for each PMU based on AIC and GCV scores, we used the specific GAM to predict the IG ratio value for all possible data collection scenarios for that PMU. We created the pool of possible data collection scenarios using the PMU specific cost per observation of each data type and the amount of resources available for data collection. We created data collection scenarios at 3 different resource availability levels: \$100,000, \$200,000, and \$300,000. As the amount of resources available for data collection increased the number of possible data-collection scenarios increased. We curtailed the amount of data allowed in each data-collection scenario differently for survival data than for count and herd composition data. We allowed the number of possible encounter histories for survival data to fluctuate between 0 and 150 for each age class (adults and juveniles). This allowed for the collection of a maximum of 300 encounter histories for any single year in a single PMU. We selected 300 as the maximum value because it represents the greatest amount of effort that would likely be expended collecting survival data in a single PMU during one year. We allowed count and herd-composition data to fluctuate between 0 and the maximum number of observations that could be collected with the given level of available resources. This allowed the amount of monitoring resources available to dictate the maximum number of count and herd composition observations. We ranked the top 25 optimal resource-allocation scenarios for collecting data to estimate abundance for each PMU based on the IG ratio values predicted for these data-collection scenarios. Figure 4-2 provides a diagram of how each data type, model type, and results from the various models contributed to the identification of the optimal resource allocation scenarios for each PMU.

RESULTS

The cost per observation for count and herd-composition data varied between both PMUs and ecotypes (Table 4-1). The cost per count observation was consistently higher than the cost per herd composition observation for all PMUs. The Middle Fork PMU had the most expensive per-observation costs for both count and herd-composition data. Island Park had the least expensive cost per observation for both count and herd-composition data (Table 4-1).

Estimates generated for survival, abundance (count data only), and age and sex ratios from the known fate and sightability models had relatively consistent point estimates but the precision fluctuated widely as the number of observations for each data type changed. Abundance estimates from the IPM were also relatively consistent and were more precise when the number of observations increased and were less precise as the number of observations decreased. There was a large range of IG ratio values across PMUs from the direct calculations made using the differing amounts of each data type collected by IDFG. The maximum IG ratio values range from 1002.44 in the Smoky-Bennett PMU to 91.27 in the Middle Fork PMU while minimum values across all PMUs are small and have a range from 0.03 to 0.09 (Table 4-2).

The number of knots for each data type's smoothing function and the R^2 value of the GAM with the lowest AIC and GCV for each PMU are located in Table 4-3. The number of knots for the survival data smoothing function and the composition data smoothing had the most variability, ranging from 5 to 11 for survival and 10 to 16 for composition. The number of knots for the count data smoothing function remained consistent at 5 across PMUs with Island Park having the sole exception with 4 knots. All of the selected GAMs had a relatively high R^2 value and ranged from 0.698 to 1.0. Figures 4-3 to 4-8 show the influence of the number of observations of each data type on the predicted IG ratio for each PMU. These figures show the contribution to the IG ratio value as the number of observations for each data type increase when the number of observations for the remaining two data types are constant. As the number of survival observations increase the predicted IG ratio generally decreases. When the number of count data observations is low, count data increases the IG ratio value rapidly but as the number count observations continues to increase the rate of increase of the IG ratio value slows but generally remains positive. Only in the Boise River PMU does the number of count observations reach a point where they have a negative impact on the IG ratio value (Figure 4-3). The influence of the amount of composition data on the IG ratio value appears relatively flat and consistent over the first 500 observations but then begins to fluctuate in both the positive and negative direction for all PMUs except Boise River. The influence of increasing the number of composition observations on the IG ratio value remains consistent for the Boise River PMU.

We used the predicted IG ratio values for each data-collection scenario to determine the top 25 optimal data-collection scenarios to estimate mule deer abundance for each PMU.

Figures 4-9 to 4-14 show the optimal allocation of monitoring resources, for the top 25 data-collection scenarios for each PMU, when the resources available for monitoring are equal to $100,000, $200,000, or $300,000. Count data makes up the highest percentage of data in the top allocation scenarios for most of the PMUs with 2 exceptions. There is a more even distribution in the results from the Smoky-Bennett PMU between the collection of count and composition data especially as the amount of available resources increase (Figure 4-13). The other exception is the top allocation scenarios from the Boise River PMU. The collection of composition data dominates the optimal allocation of resources for the Boise River PMU except when the available resources are $100,000. The collection of survival data is not prioritized in any of the top allocation scenarios except in the Middle Fork PMU when available resources are $100,000.

We used a consistent value of $1200 per encounter history for all of the previously data-collection scenarios but the lack emphasis on this important data source was surprising. Therefore, we wanted to determine the impact of the changing the cost of collecting survival data on the predicted relationship between the IG ratio and the number of encounter histories. We selected the Middle Fork PMU to test how changes to the cost of collecting survival impacted our prediction of the most optimal data-collection scenarios. Figures 4-16 – 4-18 show the relationship between the cost of collecting survival data and the predicted influence of each data type on the IG ratio. As the cost of collecting survival data increases from $200 to $4000 several patterns emerge. First, the shape of the relationship between the number of encounter histories and the IG ratio remains fairly stable but becomes smoother as cost increases. Second, the relationship between the amount of count data collected and the IG ratio remains virtually unchanged. Finally, there is large change in the relationship between the amount of herd composition collected and the IG ratio. When the cost of collecting survival data is low ($200) the impact of the number of herd composition observations on the IG ratio fluctuates widely between 30 and 60. However, as survival data becomes more expensive to collect, $2,000 to $4,000, the impact of herd composition data stabilizes to have a relatively low and almost constant effect on the IG ratio. Despite the impact of the changes in the cost of collecting survival data on the relationships between data types and the IG ratio, we observed very little change in the overall patterns of the top 25 optimal data-collection scenarios (Figure 4-20). Specifically, the collection of survival data continued to be the least prioritized type of data to collect.

DISCUSSION

The goal of this research is to provide information to wildlife management agencies concerning the most efficient use of monitoring resources to estimate wildlife abundance. In order to accomplish this goal, we develop a multistep process to determine a set of scenarios for optimally allocating resources for data collection that is capable of adapting to the changes in the availability of wildlife monitoring resources. The process to go from the original data collected in the field to a set of optimal resource allocation scenarios requires the following steps: 1) the estimation of age and sex ratios, survival rates, and abundance with varying amounts of data, 2) use the results from the single data type analyses in the current population model to generate abundance estimates whose precision reflects the differing amounts of each type of data that are available, 3) calculate the area specific cost per observation for each data type, 4) develop a metric, the IG ratio, to quantify the relationship between the amount of resources used and the expected gain in the precision of the estimate of abundance, 5) create a model to predict the IG ratio under changing data and resource availability conditions, and 6) generate a set of data collection scenarios which simultaneously reduce cost and increase the precision of abundance estimates which can serve as a to guide management agencies when deciding how to best use the limited resources available for wildlife monitoring. The cost per observation for each data type is a critical component of determining the most efficient allocation of resources for estimating wildlife abundance. While the calculation of these rates is straightforward once cost and observation information is gathered, there are some interesting patterns to emerge from this analysis. The first interesting pattern is the consistently higher cost per count observation relative to the cost per composition observation. During the time period which this analysis covers, count data is typically available for a single year for each PMU whereas composition data is often available for multiple years. However, the number of count observations range from several hundred to several thousand per year while the number of composition observations are typically lower than two hundred per year in a single PMU. Given the number of observations per year and the equal costs in labor and equipment for all aerial surveys, the expectation is that count observations should be less costly per observation than composition observations, but this is not true for any of the PMUs analyzed. This demonstrates that the methodology in use to collect count data, the ‘censusing’ of entire population units, results in a greater cost per observation than the random sampling methods used to collect composition data. This suggests that a method for collecting count data which resembles the method used to collect of composition data may use monitoring resources more efficiently. Further support for modifying count data collection methods comes from the graphs showing the influence of count data on the IG ratio values (Figures 4-3 – 4-8). These graphs clearly show that in most PMUs after collecting between 500 to 1000 count observations the rate of information gain begins to decrease, thus negating the need to collect thousands of count observations.

A second result to emerge from the cost per observation analysis is the lack of a direct relationship between the cost per observation and the predominant ecotype in each PMU. The expectation is that in ecotypes with higher visibility collecting data should be less costly due to the ease of detecting the species of interest (Unsworth et al. 1994). However, the results of this analysis do not support these expectations. There are several potential explanations for the lack of the expected patterns in the relationship between cost per observation and ecotype. First, the difference between visibility and detection rates across the three ecotypes present in this study may not be large enough in magnitude to impact the cost per observation. Second, the topography of each PMU may impact detection in ways that the visibility scores of the ecotypes do not reflect. A third potential reason for the lack of an ecotype related cost pattern may be the influence of deer density on the cost per observation. Deer density in a given PMU at the time of the survey may have a greater impact on cost than the visibility of deer in a particular ecotype.

The information gain (IG) ratio is the metric developed to measure the relationship between the amount of resources used to collect a specific amount of data and the resulting gain in precision of the estimate of abundance. This metric incorporates both the cost of data collection and the resulting ratio of the precision of the abundance estimate to the abundance estimate and is useful in determining the most efficient data collection scenarios. Both cost and precision impact the behavior of the IG ratio as seen in Figure 4-15. This figure shows that as the resources used to collect certain amounts of data increases, the IG ratio value decreases rapidly if the $\frac{CRI}{N}$ ratio is less than 0.125 but remains stable if $\frac{CRI}{N}$ is 0.5 or greater. This demonstrates that the IG ratio rewards more precise estimates of abundance at a greater rate than increases in spending. For example, if the amount of resources spent to gather data increases from \$100,000 to \$200,000 and the $\frac{CRI}{N}$ remains at 0.01 the resulting IG ratio decreases from 10 to 5 thus penalizing the doubling of resources used to collect data for abundance estimates of the same precision. Since, the goal of this analysis is to use wildlife monitoring resources most efficiently we suggest that the IG ratio metric captures the relationship between data collection costs and the precision of abundance estimates in such a way that it can be the basis for determining which scenarios represent the most efficient use of monitoring resources.

The GAMs in use for predicting the IG ratio across all PMUs share several characteristics. First, there is a rapid positive gain in the value of the IG ratio as the number of count data observations increase from 0 (Figures 4-3 – 4-8). We expect a large positive influence of count data observations since the count data contains the greatest amount of information concerning abundance. Second, there is generally a negative impact on the value of the IG ratio when the number of survival observations increase. In the Island Park, Palisades, and South Hills PMUs, there is a small positive impact on the IG ratio value when the number of survival observations is less than 50, but as the number of survival observations increase over 50 the value of the IG ratio declines quickly (Figures 4-3 – 4-8). In the other PMUs, IG ratio values are negatively impact by any number of survival observations. The general pattern of shrinking IG ratio values as the number of survival observations increase is unexpected.

Survival data contains a large amount of information related to factors contributing to the size of a wildlife population. Therefore, we expect collecting more survival data to increase the precision of abundance estimates. There are two potential explanations for the generally negative trend associated with increases in the number of survival observations and the decreased contribution of this data type to the IG ratio. First, the large amount of survival data collected by IDFG over the past several decades provides the IPM with prior distributions for survival that have low bias and high precision. This results in a situation where the collection of additional survival data contributes very little to increasing the amount of high quality information about survival already available to the IPM. Thus collecting relatively smaller amounts of survival data do not contribute to improving the precision of the IPM-generated estimates of abundance. Another analysis that provides support for this explanation shows that increasing the information about survival when there is already a large amount of available information regarding survival requires sample sizes greater than 150 individuals (pers. comm. Paul Lukacs).

A second potential explanation for this pattern of a generally negative influence of increasing amounts of survival data on the IG ratio value relates to the relatively high and constant cost value assigned to survival observations. A more precise characterization of the costs associated with collecting survival data from the different age classes which reflects the changing costs of radio collar technologies will potentially change the direction and magnitude of the impact of survival observation numbers on the value of the predicted IG ratio. We investigated this explanation by re-analyzing the predictive GAM models and data-collection scenarios using a set of survival data collection costs that ranged from $200 to $4000 for the Middle Fork PMU. The results from this further analysis showed that changes to the cost of collecting survival data had the most impact on the expected contribution of herd composition data to the IG ratio (Figures 4-16 – 4-18). However, when we predicted the most efficient data-collection scenarios using the GAM results from the variable survival price scenarios the general trends of a focus on collecting count and composition data remained even when survival data collection costs are low (4-20). This provides further support for our hypothesis that the large amount of survival data already available in this study system de-emphasizes the continued collection of large amounts of survival data. We expect survival data to be more important to estimating abundance for populations that do not already have such a rich source of survival data but did not test this expectation during this analysis.

The impact of increases in amount of composition data collected on the IG ratio value is more variable relative to the other two data types. In the Boise River PMU the impact of increased composition observations is consistent and relatively stable but in the rest of the PMUs the influence of composition data on the IG ratio switches between positive and negative at irregular intervals. The ‘wavy’ shape of the relationship between the amount of composition data collected and the predicted IG ratio value is potentially the result of over-fitting the composition data smoothing function. Over-fitting is the result of having too many knots assigned to the smoothing function of an explanatory variable in the GAM but less smooth curves indicate less bias (Larson 2015, Wood 2017). However, for this analysis the top models based on the AIC and GCV values all include a relatively large number of knots for the composition data smoothing function.

Composition data is the only data type where increased observations result in a negative value for the IG ratio. A possible explanation for the presence of negative IG ratio values occurring as the number of composition observations increase is the related to the precision of the estimate of abundance. Previous work (Henderson Chapter 2 and 3) shows that composition data often contains significant amounts of error and this error impacts the precision of the age and sex ratios estimated from composition data. Furthermore, as imprecision in age and sex ratios enter the IPM to generate estimates of abundance, the resulting IPM-generated abundance estimates decrease in precision. In the calculation of the IG ratio, a decrease in the precision of the abundance estimate typically results in a lower IG ratio value unless there is a significant decrease in the amount of resources used to collect that data type. Thus, the negative IG ratio values associated with increases in amount of composition data collected are most likely the result of introducing larger amounts of a relatively low quality (imprecise) data source into the abundance estimation process.

Using the PMU specific predictive models, we determine the most efficient data collection scenarios for precisely estimating mule deer abundance under three different levels of resource availability. In most PMUs, the majority of scenarios for efficient data collection suggest devoting more resources to the collection of count data followed by composition data and survival data (Figures 4-9 – 4-15). The single exception to this trend is the data-collection scenarios for the Boise River PMU in which the collection of composition data is most favored.

The top scenarios for optimal data collection in the Boise River PMU are intriguing for two reasons. First, Boise River data collection scenarios result in the second smallest increase in the predicted IG ratio values (+16.45) as the amount of resources devoted to data collection increase from \$100,000 to \$300,000. The lack of increase in the IG ratio value relates to the absence of a significant gain in the precision of the IPM generated estimate of abundance as the amount of available monitoring resources increases. The lack of gain in precision of the abundance estimate may be connected to the uncommonly large R^2 value (1.0) associated with the predictive IG model for this PMU, but why the R^2 value could result in a relatively stable IG ratio value across resource availability levels is unclear. Second interesting feature of the top optimal data collections scenarios for the Boise River PMU is the prioritization of the collection of composition data in the majority of the data collection scenarios. As the level of available monitoring resources increases to \$200,000 and above the emphasis on collecting composition data grows (Figure 4-9). The stable contribution of composition data to the IG ratio value

(Figure 4-3) as the number of composition observations increase is most likely responsible for the prioritization of the collection of composition data in this PMU.

The top data-collection scenarios for the Island Park, Palisades, Smoky-Bennett, and the South Hills PMUs are similar to one another. The most efficient scenarios for all of these PMUs generally prioritize the collection of count data over the other two types of data. However, there are several exceptions to the general pattern when the collection of composition data is favored over the collection of count data. The prioritization of collecting composition data occurs in 19% of the data-collection scenarios in the $100,000 resource availability level, in 4% of the data-collection scenarios in the $200,000 resource availability level, and in 1% of the data-collection scenarios in the $300,000 resource availability level in these four PMUs. The difference in cost between count data observations and composition data observations for these PMUs increase the number of data-collection scenarios which prioritize composition data at the $100,000 resource level. The cost per count observation is higher than the cost per composition observation across all PMUs (Table 4-1). Therefore when the available resources for monitoring are $100,000 or less, there may not be enough money available to collect a sufficient amount of count data for precise estimates of abundance. Since the contribution of composition data to the IG ratio value is relatively stable over the first 500 to 1,000 observations (Figures 4-3 – 4-8) the optimal allocation scenarios likely shift to the collection of the less expensive data type to take advantage of this stable relationship between the influence of composition data and the IG ratio value. However, it is important to point out that the top ranked scenarios for optimal allocation of monitoring resources across all of these PMUs always prioritizes the collection of count data over the collection of composition and survival data.

The data collection scenarios for the optimal allocation of monitoring resources have the most variability in the Middle Fork PMU (Figure 4-11). Moreover, the greatest variability in data collection scenarios for the Middle Fork PMU occurs when the available resource level is \$100,000. Within the \$100,000 resource availability level, the model prioritizes the gathering of count data in the 1st and 2nd ranked scenarios, the collection of composition data in the 3rd ranked scenario, 4th ranked collection scenario implies that collecting no data is the optimal strategy, and the 5th ranked scenario prioritizes the collection of survival data. The increased cost of collecting data in this PMU is likely responsible for the switching of collection priorities in the Middle Fork PMU at the \$100,000 level. The cost per observation of count and composition data is the highest relative to all other PMUs considered in this analysis. The cost per observation of composition data is also higher than the cost of collecting count data in 67% of all PMUs. The Middle Fork is also the PMU with the lowest average number of count observations and has the 2nd lowest average number of composition observations. These low numbers of observations decrease the precision of the estimates of abundance generated by the IPM. The combination of high data collection costs and low precision estimates of abundance combine to give the Middle Fork PMU the lowest mean predicted IG ratio values (10.45) for the top 25 data collection scenarios for all resources levels across all PMUs. Since the predicted IG ratio value for all scenarios are the consistently low, it is likely that a wider variety of data collection scenarios are capable of producing IG ratio values in this lower range and therefore result in the wide variety of optimal data collection scenarios suggested for the Middle Fork PMU.

The optimal scenarios for data collection for the Middle Fork PMU at the \$200,000 and \$300,000 levels of resource availability more closely resemble the results from the other PMUs, excluding Boise River. At the \$200,000 resource availability level, the collection of count data receives the highest priority in 72% of the data collection scenarios and at the $300,000 resource availability level the collection of count data receives the highest priority in 84% of the data collection scenarios. However, even when data collection priorities change to the collection of count data at the $200,000 and $300,000 resource availability levels, the average predicted values for the Middle Fork PMU IG ratio remain low at 10.0 and 16.94, respectively. The low predicted IG ratio values suggest that in this PMU estimate abundance values that are equivalent in precision to the other PMUs require a greater amount of resources. A post-hoc analysis shows that an increase in available resources to at least a $600,000 level is required to raise the predicted IG ratio value for the top ranked data collection scenario for the Middle Fork PMU to the same IG ratio value predicted for the top ranked data collection scenario at the $100,000 resource availability level for the PMU with the next smallest IG ratio value, the Island Park PMU.

The results of this analysis and the methods developed during this study will be useful to wildlife management agencies when deciding how to efficiently allocate monitoring resources and prioritize data collection. However, the usefulness of these results can be improved through the addition of more precise information regarding the costs of collecting each data type. Specifically, a better estimate of costs for collecting survival data will improve the ability of this method to determine which data collection scenarios represent the most efficient allocation of monitoring resources. The constant value assigned to the cost of securing survival data likely results in the underutilization of this data type in some data collection scenarios. A potential avenue for further research using the method developed by this study is the inclusion of a greater number of PMUs across more ecotypes. A greater number of PMUs representing more ecotypes will allow for the investigation of more general resource allocation patterns that are applicable across larger landscapes. The investigation of larger landscape scale patterns of optimal resource allocation should lead to better understanding of the role environmental variables play in the efficient employment of monitoring resources.

A disadvantage to this methodology is the amount of data needed to perform this type of analysis. In order to accurately determine the most efficient allocation of resources for collecting data, researchers and managers would have access to sets of data collected over multiple time steps in order to include the variety of stochastic influences that impact data collection. However, the substitution of simulated data could help determine efficient data collection scenarios using the methodology developed in this study, when long term sets of data are not available.

The goal of this study is to provide a defensible methodology for making decisions regarding the use of public resources for the monitoring of wildlife abundance. The methods described here highlight a process in which wildlife management agencies can determine which data types provide the most return on their monitoring resource investment. We develop this method using data collected on mule deer populations in Idaho but, it is flexible enough for implementation over a variety of species across a larger geographic range. The data types commonly collected to estimate the abundance of other species and taxa, such as nest survival or clutch size, can easily replace the data types used in this analysis. Any population models that provide a measure of precision along with the estimated population size can replace the IPM employed in this analysis. The user could also replace the GAM used here with a model better suited to predicting the IG ratio for the particular data types available to them. We suggest the most important contributions of this research are the development of the IG ratio, a metric to determine the relationship between available resources and the precision of the abundance estimate, and the set of optimal allocation scenarios generated using the IG ratio. The generation of multiple scenarios that optimize the use of monitoring resources allow decision makers to see the full range of options at their disposal and tailor their decisions to fit the changing availability of monitoring resources that are available for wildlife management.

LITERATURE CITED

Association of Fish & Wildlife Agencies and the Arizona Game and Fish Department. 2017. The State Conservation Machine [PDF file]. Retreived from https://www.fishwildlife.org/application/files/3615/1853/8699/The_State_Conservation_Machine-FINAL.pdf.

Chadès, I., McDonald-Madden, E., McCarthy, M.A., Wintle, B., Linkie, M. and Possingham, H.P., 2008. When to stop managing or surveying cryptic threatened species. Proceedings of the National Academy of Sciences.

Field, S.A., Tyre, A.J. and Possingham, H.P., 2005. Optimizing allocation of monitoring effort under economic and observational constraints. The Journal of Wildlife Management, 69(2), pp.473-482.

Gelman, A. and Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Statistical science*, *7*(4), pp.457-472.

Hellawell, J.M., 1991. Development of a rationale for monitoring. In Monitoring for conservation and ecology (pp. 1-14). Springer, Dordrecht.

Hurley, M.A., Hebblewhite, M., Lukacs, P.M., Nowak, J.J., Gaillard, J.M. and Bonenfant, C., 2017. Regional-scale models for predicting overwinter survival of juvenile ungulates. *The Journal of Wildlife Management*, *81*(3), pp.364-378.

Idaho Department of Fish and Game. (2008). Mule Deer management plan 2008-2017. Idaho Department of Fish and Game, Boise, USA.

Idaho Department of Fish and Game. 2015. Directions: FY 2015-2018, Annual Strategic Plan. Idaho Department of Fish and Game, Boise, ID, USA.

Larson, K. 2015. GAM: The Predictive Modeling Silver Bullet[PDF file]. Retrieved from https://multithreaded.stitchfix.com/blog/2015/07/30/gam/.

Legg, C.J. and Nagy, L., 2006. Why most conservation monitoring is, but need not be, a waste of time. Journal of environmental management, 78(2), pp.194-199.

Nowak, J.J., Lukacs, P.M., Hurley, M.A., Lindbloom, A.J., Robling, K.A., Gude, J.A. and Robinson, H., 2018. Customized software to streamline routine analyses for wildlife management. *Wildlife Society Bulletin*, *42*(1), pp.144-149.

Plummer, M., 2003, March. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, No. 125.10).

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rabe, M.J., Rosenstock, S.S. and deVos Jr, J.C., 2002. Review of big-game survey methods used by wildlife agencies of the western United States. Wildlife Society Bulletin, pp.46-52.

United States Department of Labor, Bureau of Labor Statistics. 2018. US Inflation Calculator. Accessed January, 2019, https://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/.

United States Fish and Wildlife Service. 2018. National Hunting License data [PDF file].Retrieved from https://wsfrprograms.fws.gov/Subpages/LicenseInfo/Hunting.htm.

Unsworth, J. W. 1994. Mule Deer Sightability: Project W-160-R-21. Idaho Department of Fish and Game, Boise, USA.

White, G.C. and Burnham, K.P., 1999. Program MARK: survival estimation from populations of marked animals. *Bird study*, *46*(sup1), pp.S120-S139.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36.

Wood, S.N. (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC. Version 1.8-26.

TABLES

Table 4-1. The cost per observation by Population Management Unit (PMU), data type, and predominant ecotype as classified by Hurley et al. (2017).

PMU	Cost per Survival Observation	Cost per Count Observation	Cost per Composition Observation	Ecotype
Boise River	$1,200	$89	$49	Conifer
Island Park	$1,200	$65	$46	Conifer
Middle Fork	$1,200	$429	$154	Conifer
Palisades	$1,200	$104	$81	Aspen
Smoky-Bennett	$1,200	$97	$48	Conifer
South Hills	$1,200	$172	$84	Shrub-Steppe

Table 4-2. The maximum and minimum information gain (IG) ratio values by Population Management Unit (PMU) calculated from data collected by the Idaho Department of Fish and Game between 2008-2018.

	IG ratio	
PMU	maximum	minimum
Boise River	167.45	0.03
Island Park	106.14	0.05
Middle Fork	91.27	0.09
Palisades	218.36	0.04
Smoky-Bennett	1002.44	0.05
South Hills	442.96	0.04

Table 4-3. Attributes of the Generalized Additive Model (GAM) with the lowest Akaike information criteria (AIC) values and generalized cross validation (GCV) for each Population Management Unit (PMU). These models are used to predict the information gain (IG) ratio values to determine the scenarios representing the optimal allocation of monitoring resources for collecting data to estimate mule deer abundance.

	Knots per smoothing function by data type			
PMU	Survival	Count	Herd Composition	Adjusted R-squared
Boise River	5	5	12	1
Island Park	10	4	13	0.698
Middle Fork	7	5	12	0.871
Palisades	11	5	10	0.813
Smoky-Bennett	11	5	16	0.816
South Hills	5	5	15	0.795

FIGURES

FIGURE 4-1. Map of Mule deer Population Management Units (PMU, Idaho Department of Fish and Game 2008).

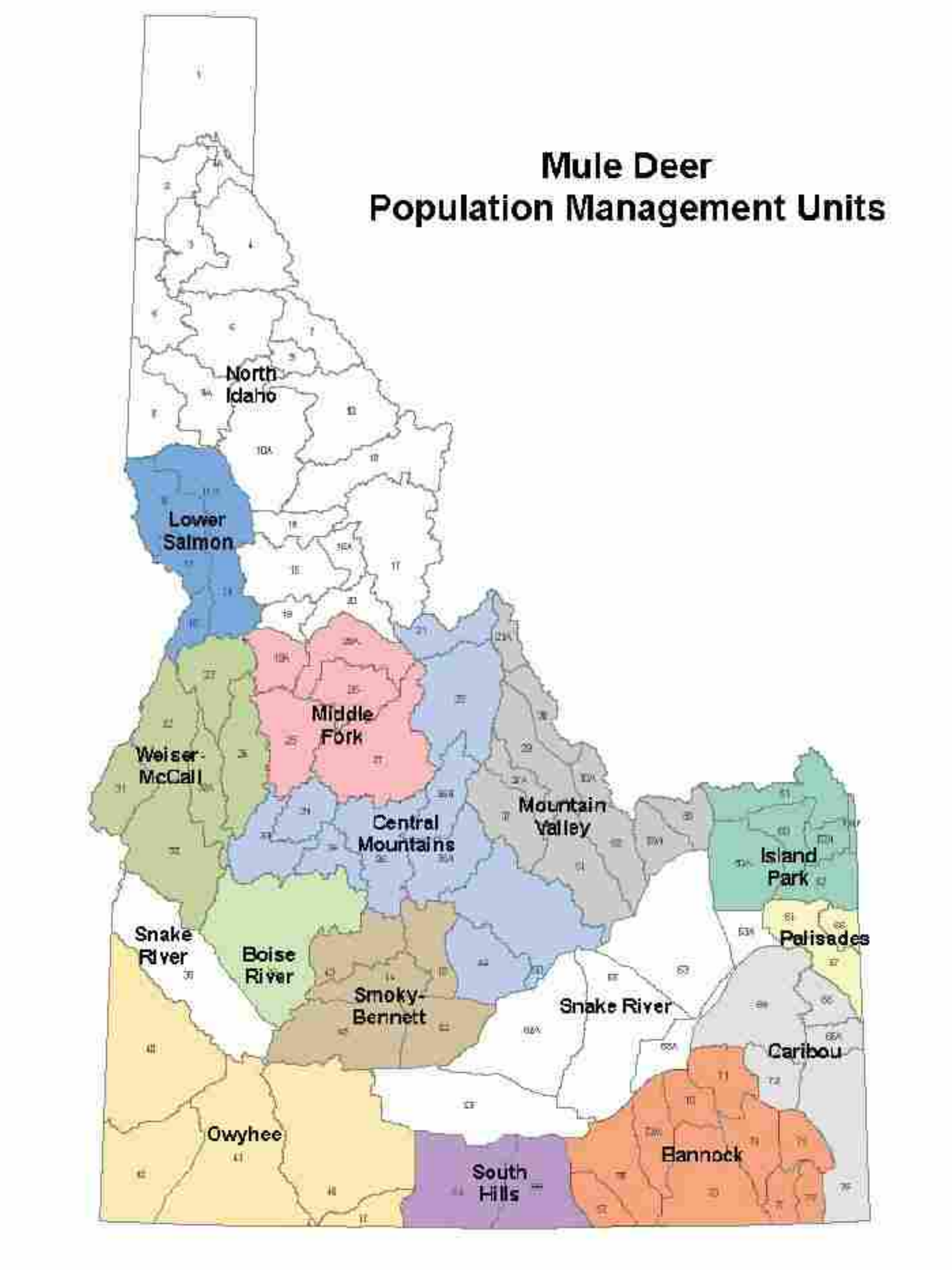


Figure 4-2. Conceptual diagram of the multistep process for determining the optimal allocation for the collection of data to estimate mule deer abundance.

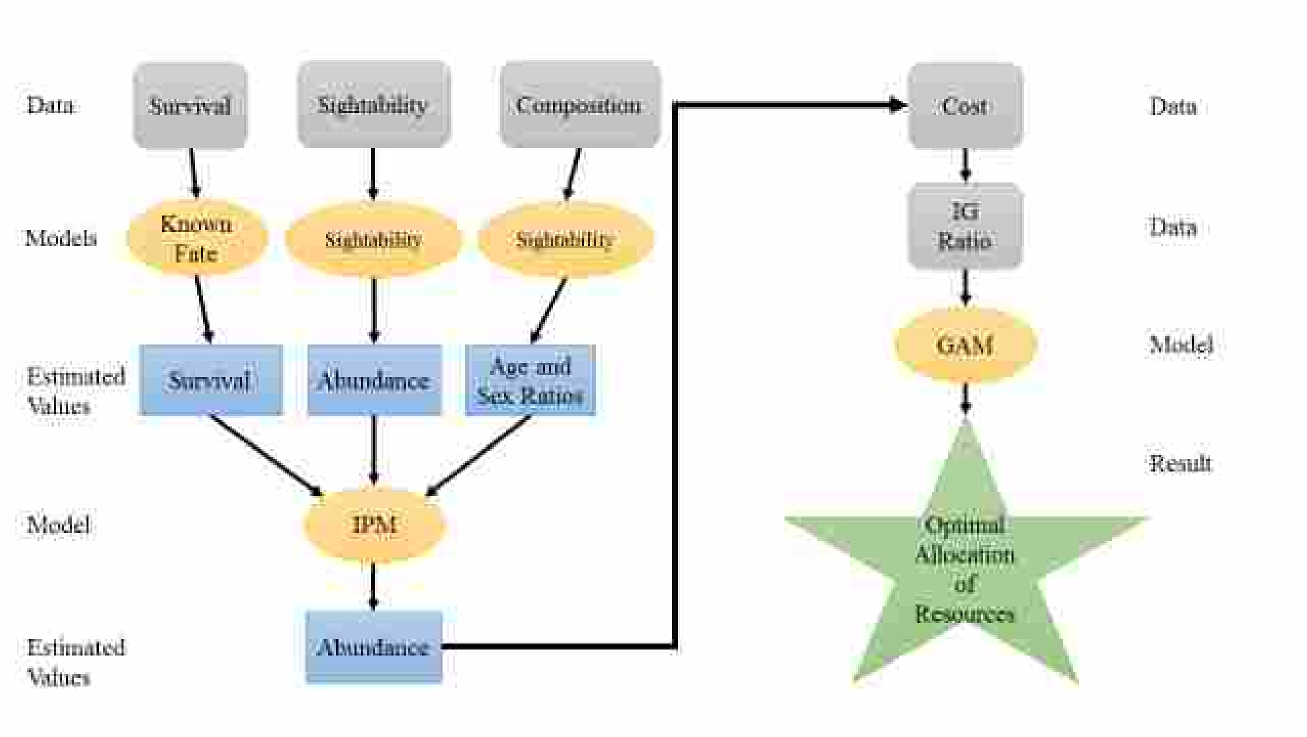


Figure 4-3. Graph displaying the relationship between the information gain (IG) ratio values and the ratio ($\frac{CRI}{N}$) of the 95% credible interval width (CRI) of the posterior distribution of the integrated population model (IPM) generated estimate of abundance and the point estimate of abundance from the IPM (N) as the amount of available resources changes from \$100,000 to \$500,000 in \$100,000 increments.

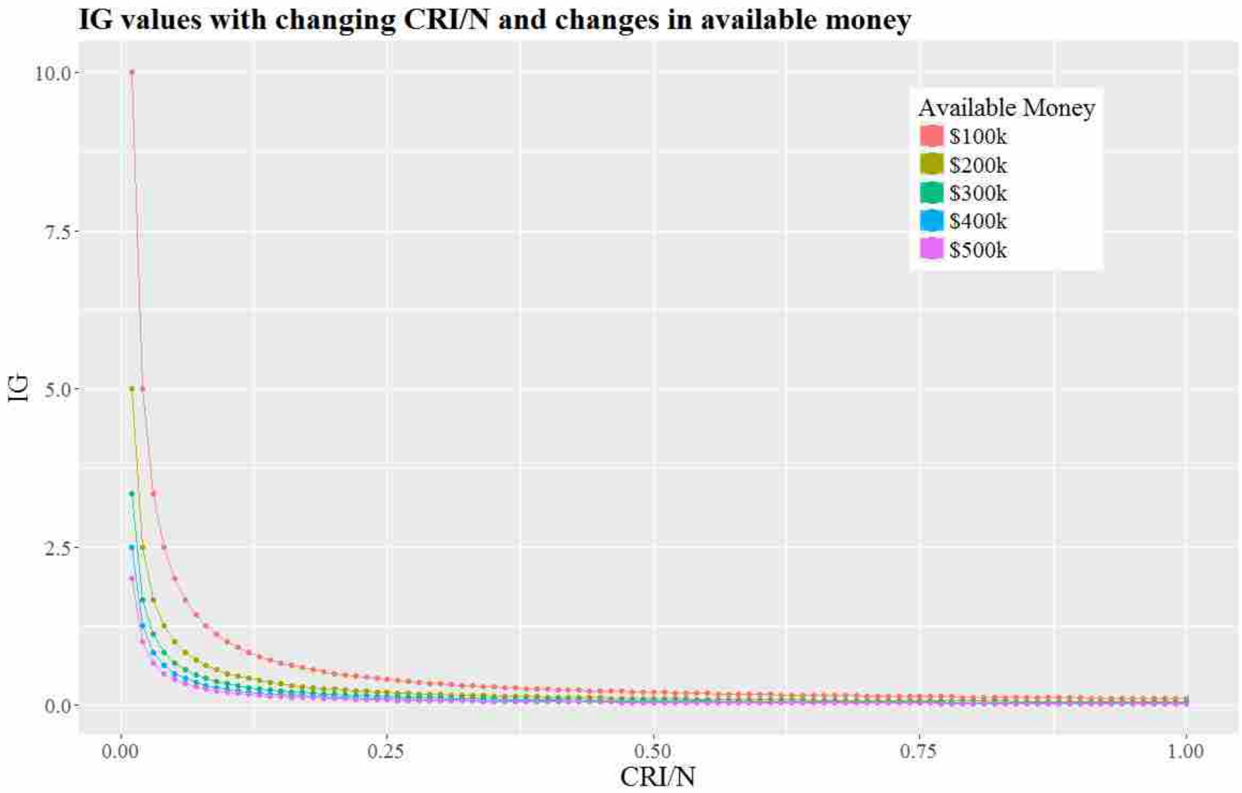


Figure 4-4. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Boise River Population Management Unit (PMU).

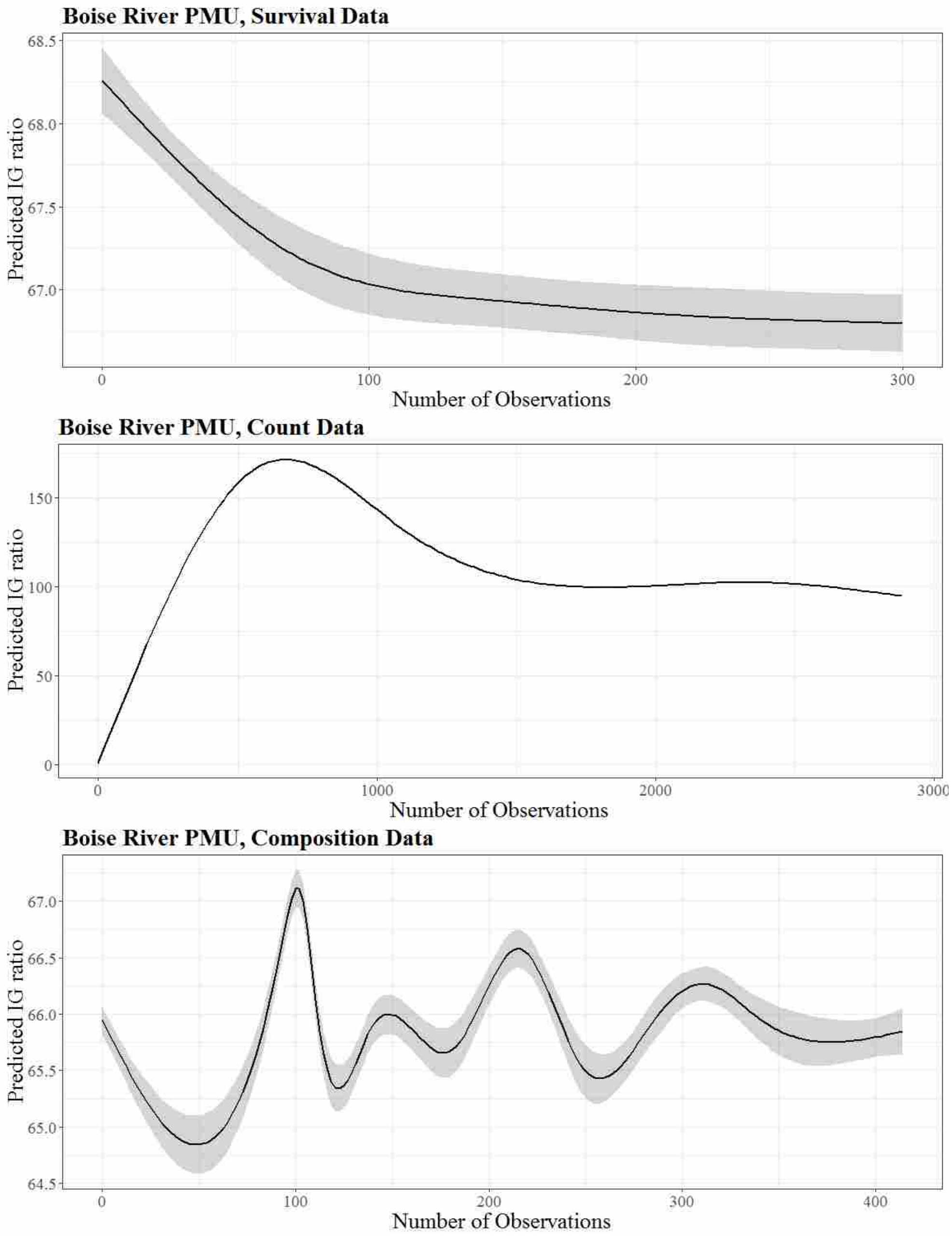


Figure 4-5. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Island Park Population Management Unit (PMU).

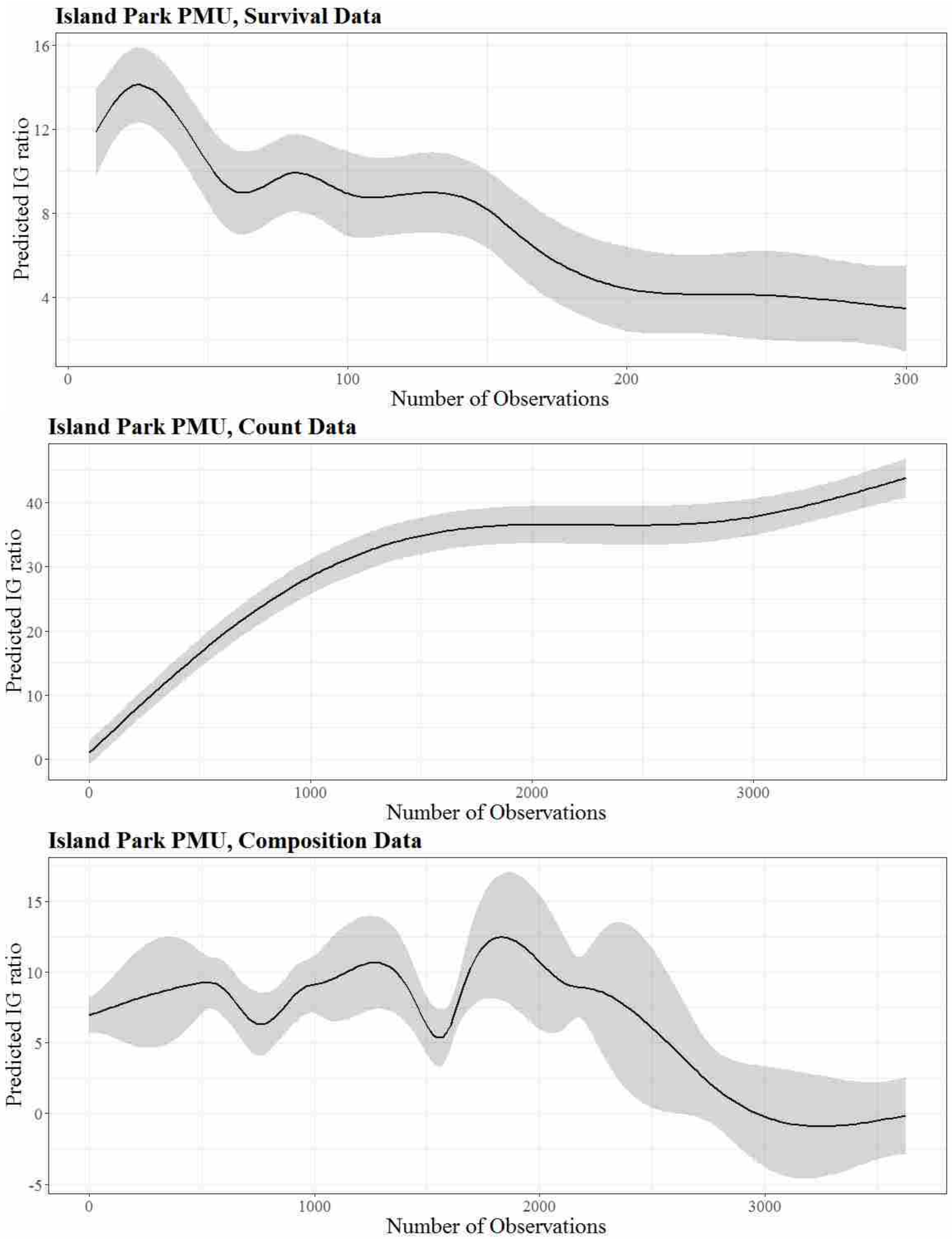


Figure 4-6. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Middle Fork Population Management Unit (PMU).

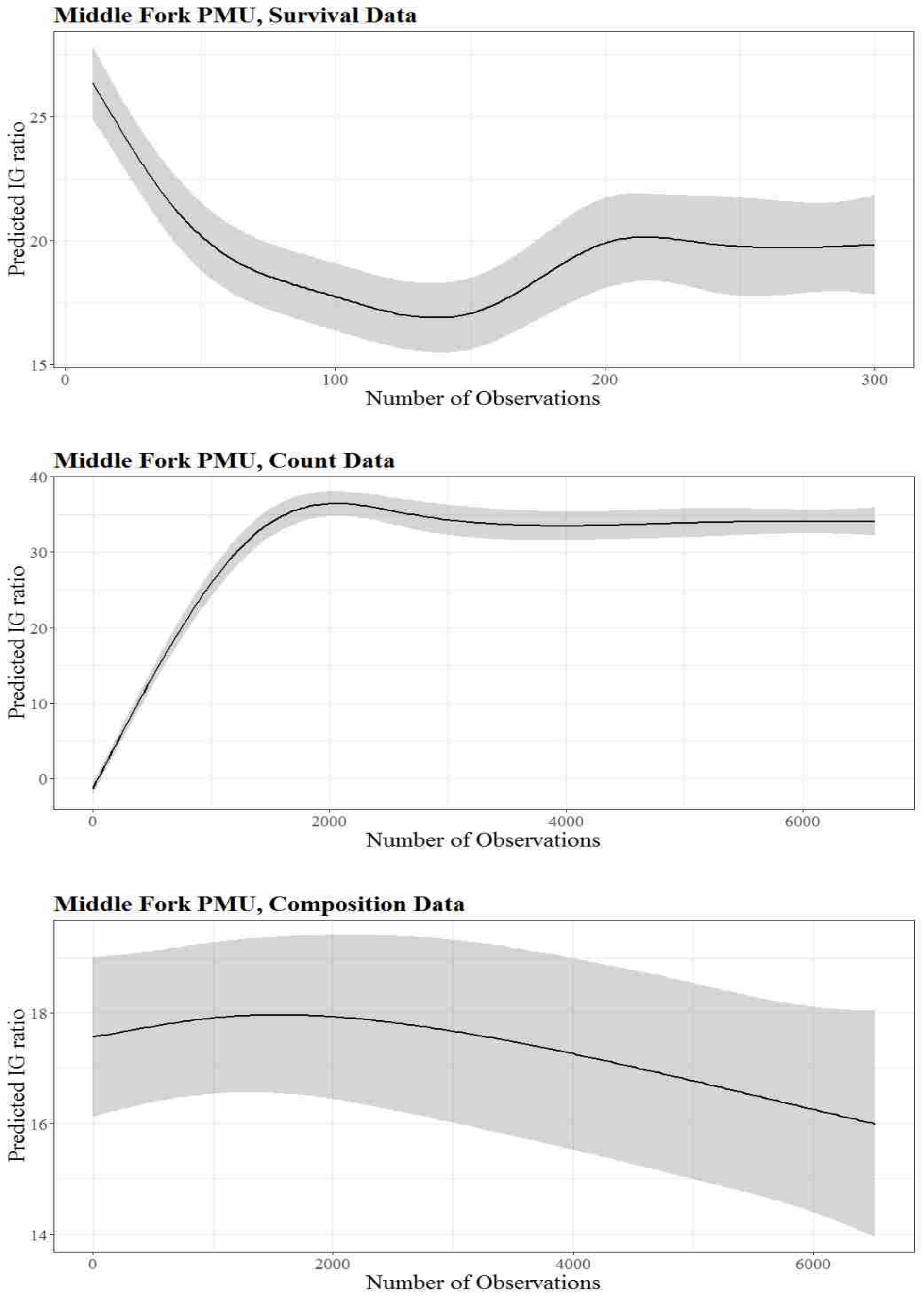


Figure 4-7. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Palisades Population Management Unit (PMU).

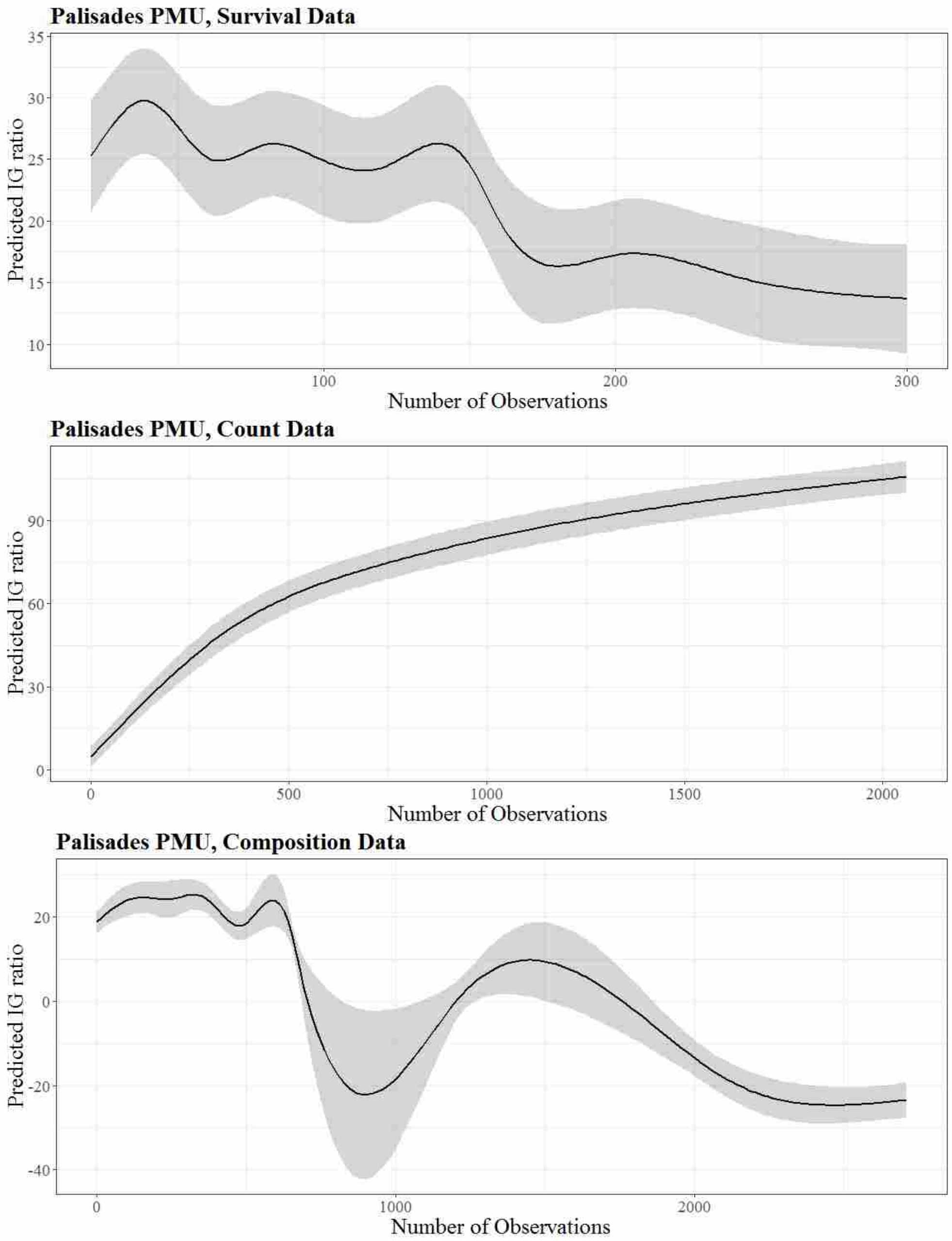


Figure 4-8. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Smoky-Bennett Population Management Unit (PMU).

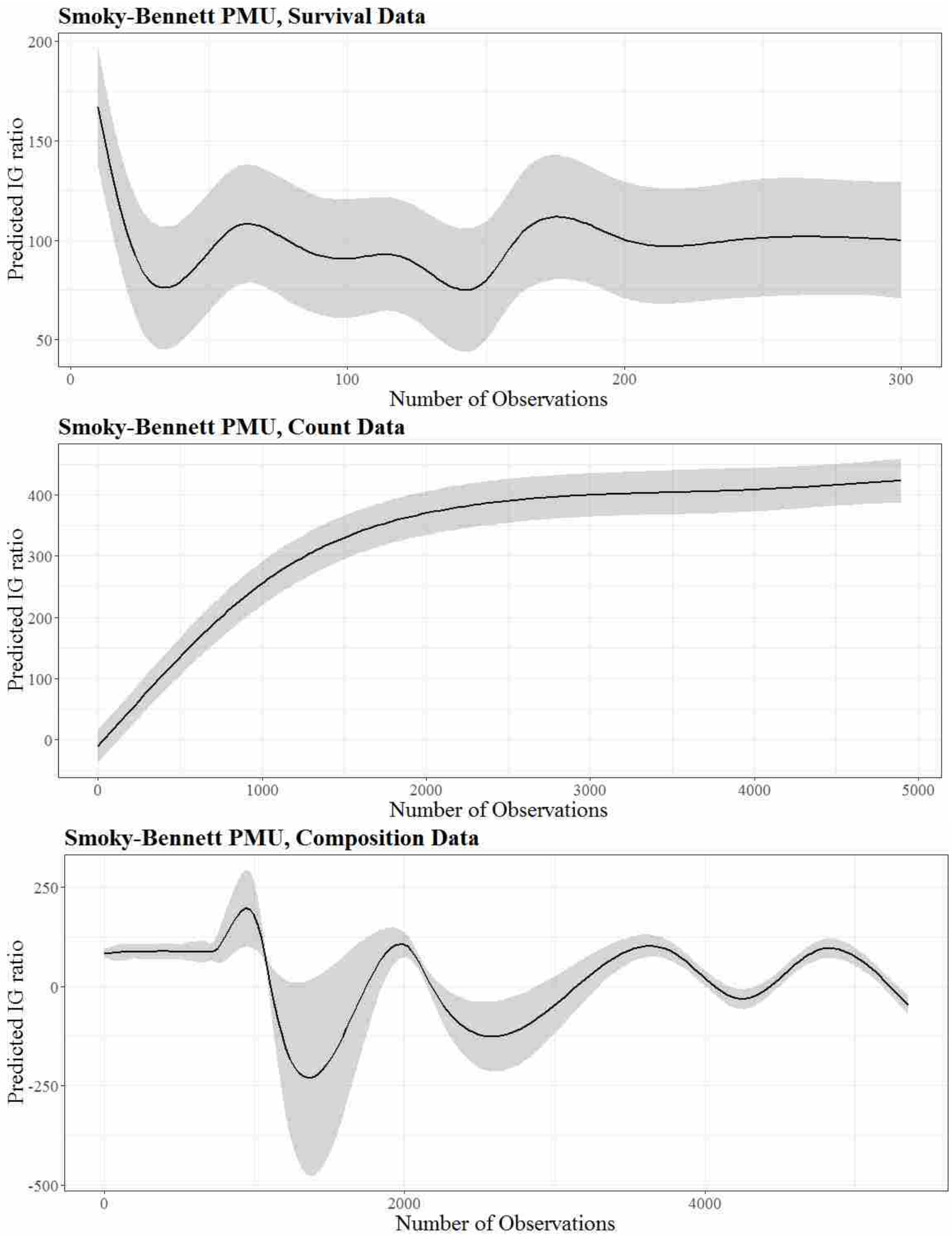


Figure 4-9. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the South Hills Population Management Unit (PMU).

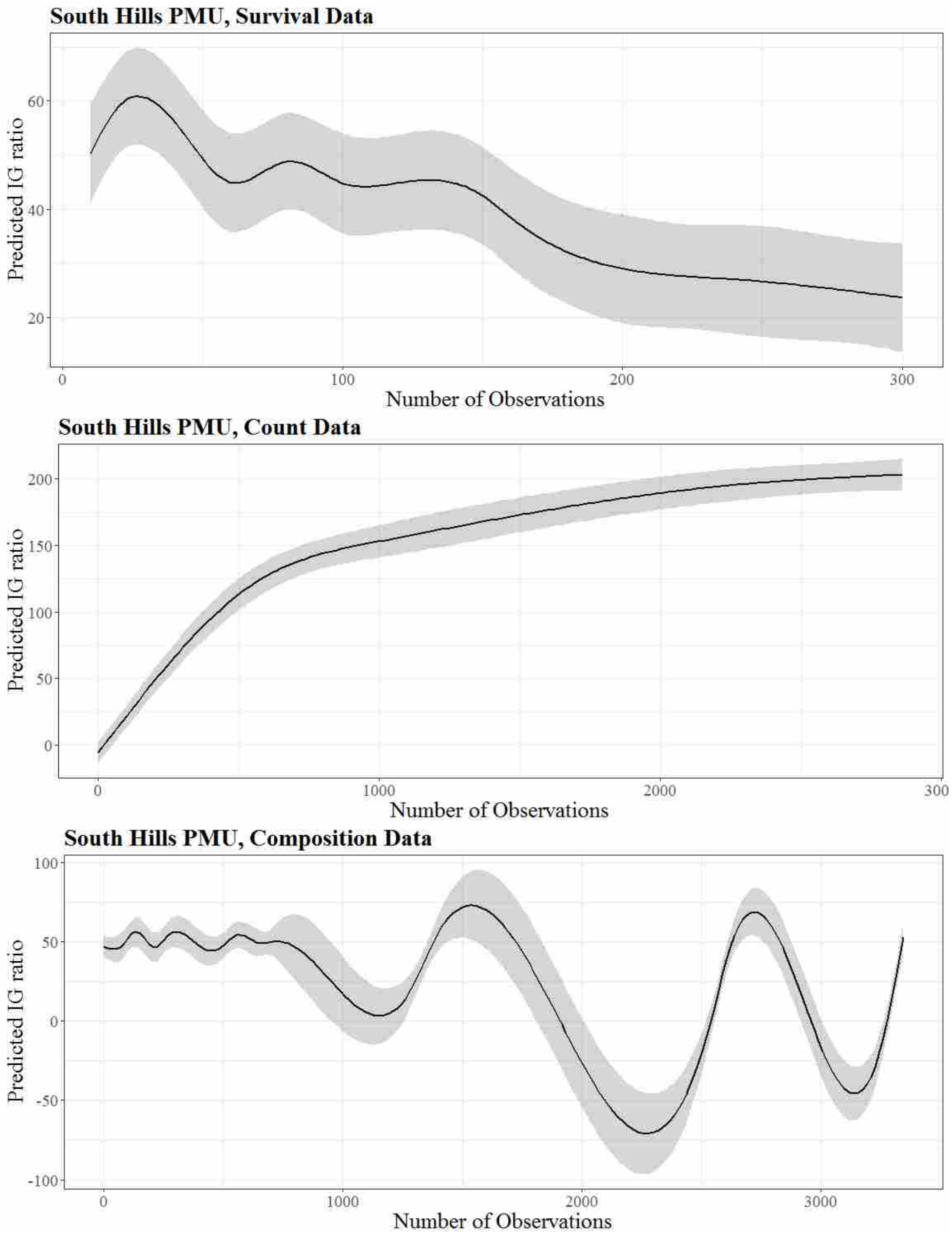


Figure 4-10. The top 25 scenarios showing the optimal allocation of monitoring resources, for collecting data to estimate abundance in the Boise River Population Management Unit (PMU). IG represents the information gain ratio. Data collection scenarios are ranked from 1 to 25 based on the highest IG ratio value and each colored portion of the bar represents the percentage of each data type collected in that scenario. The values on the right hand side of each graph represent the total amount of available resources in increments of $100000. $100k represents $100000 in available resources, $200k represents $200000 in available resources, and $300k represents $300000 in available resources.

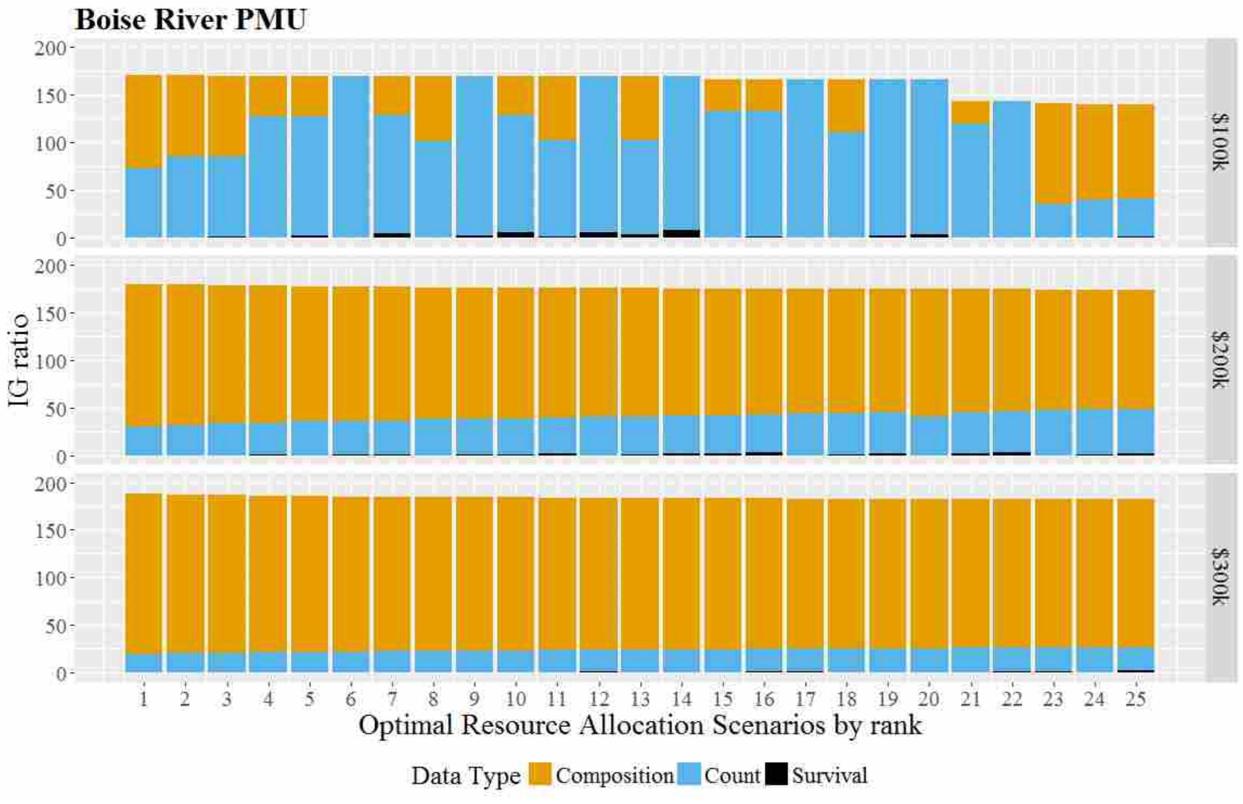


Figure 4-11. The top 25 scenarios showing the optimal allocation of monitoring resources for collecting data to estimate abundance for the Island Park Population Management Unit (PMU). IG represents the information gain ratio. Data collection scenarios are ranked from 1 to 25 based on the highest IG ratio value and each colored portion of the bar represents the percentage of each data type collected in that scenario. The values on the right hand side of each graph represent the total amount of available resources in increments of $100000. $100k represents $100000 in available resources, $200k represents $200000 in available resources, and $300k represents $300000 in available resources.

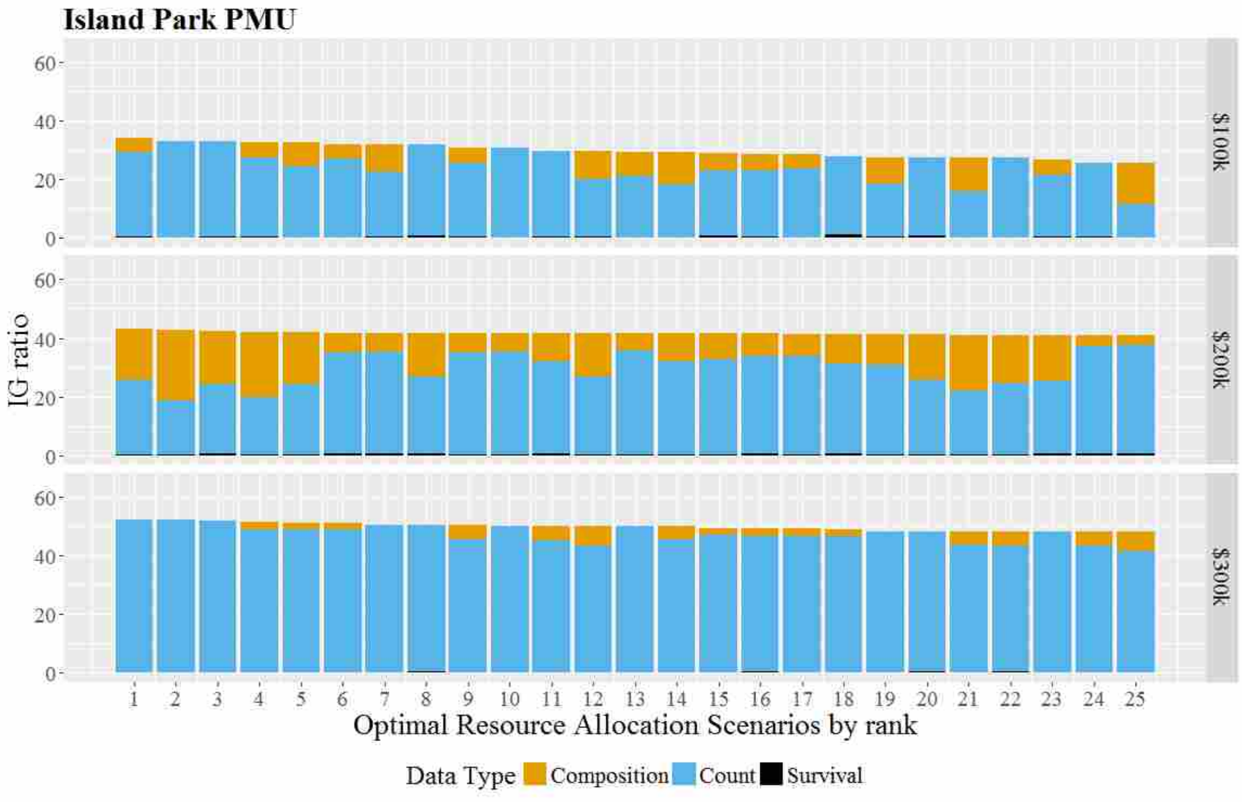


Figure 4-12. The top 25 scenarios showing the optimal allocation of monitoring resources for collecting data to estimate abundance for the Middle Fork Population Management Unit (PMU). IG represents the information gain ratio. Data collection scenarios are ranked from 1 to 25 based on the highest IG ratio value and each colored portion of the bar represents the percentage of each data type collected in that scenario. The values on the right hand side of each graph represent the total amount of available resources in increments of $100000. $100k represents $100000 in available resources, $200k represents $200000 in available resources, and $300k represents $300000 in available resources.

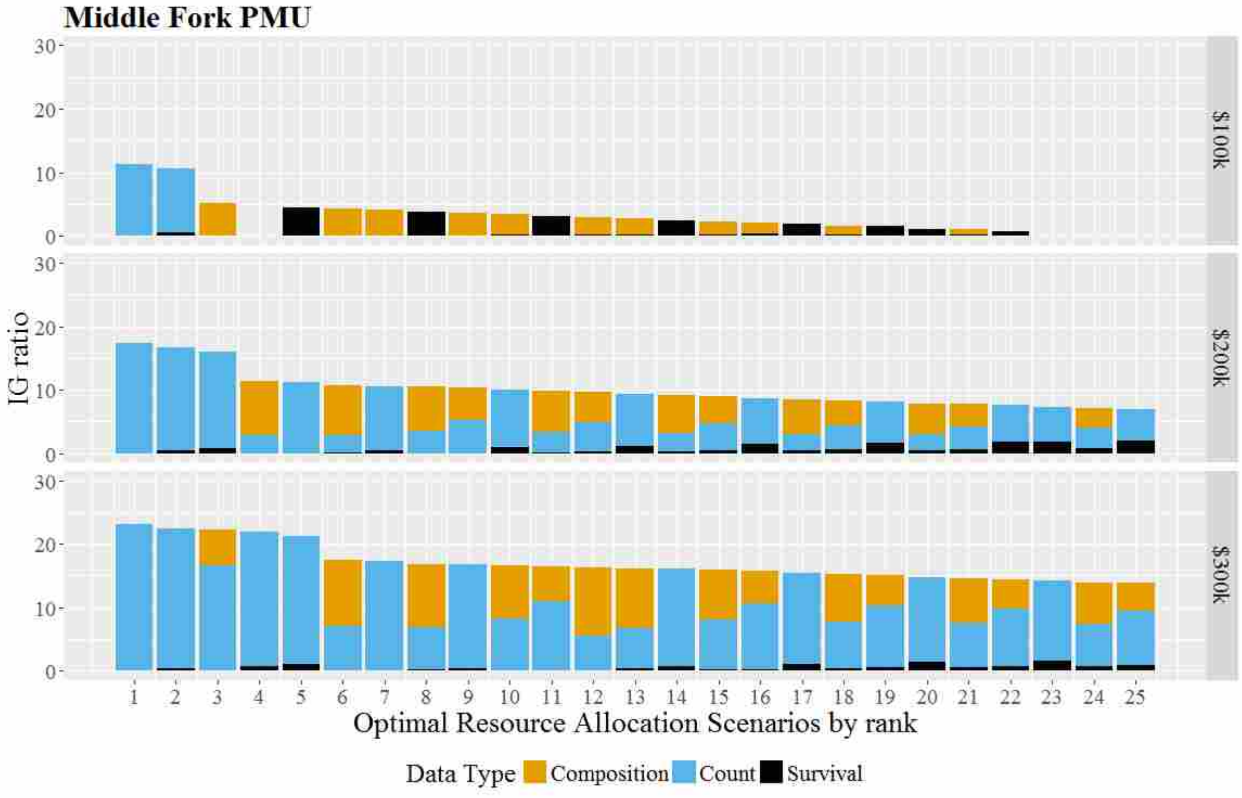


Figure 4-13. The top 25 scenarios showing the optimal allocation of monitoring resources for collecting data to estimate abundance for the Palisades Population Management Unit (PMU). IG represents the information gain ratio. Data collection scenarios are ranked from 1 to 25 based on the highest IG ratio value and each colored portion of the bar represents the percentage of each data type collected in that scenario. The values on the right hand side of each graph represent the total amount of available resources in increments of $100000. $100k represents $100000 in available resources, $200k represents $200000 in available resources, and $300k represents $300000 in available resources.

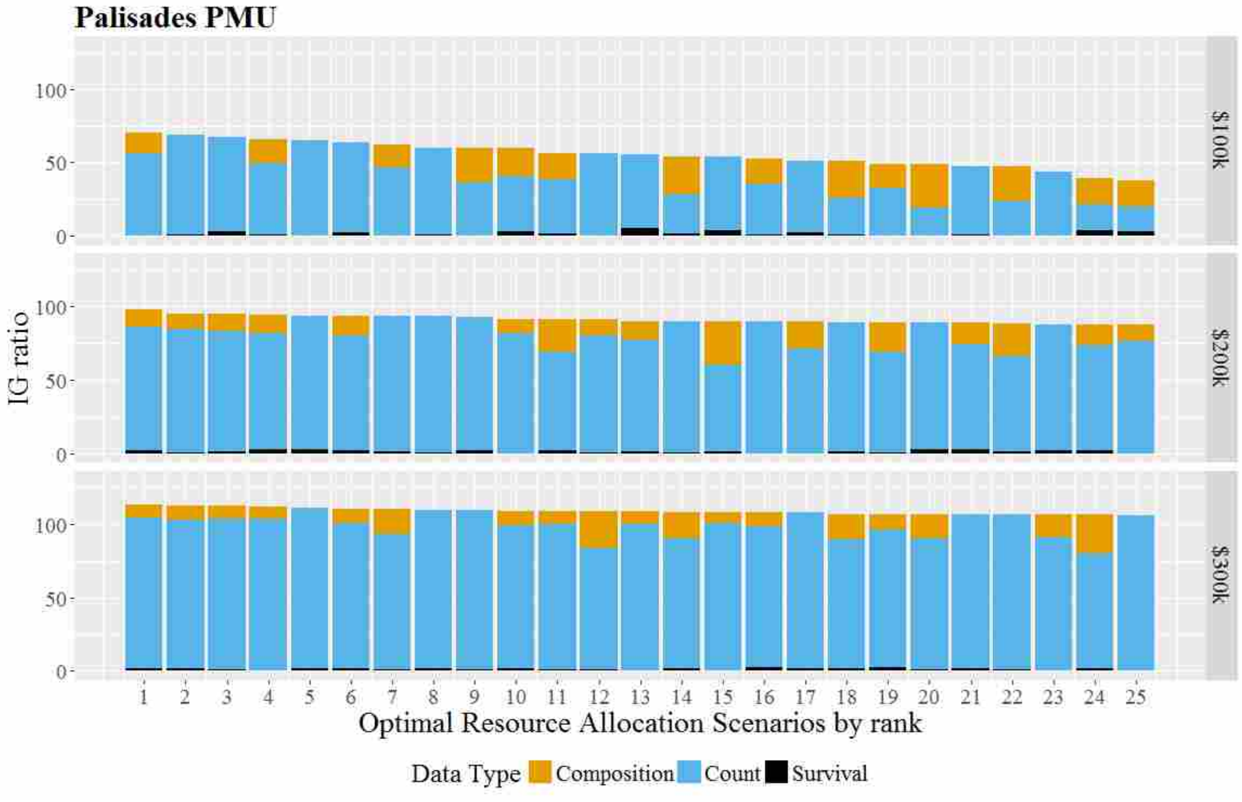


Figure 4-14. The top 25 scenarios showing the optimal allocation of monitoring resources for collecting data to estimate abundance for the Smoky-Bennett Population Management Unit (PMU). IG represents the information gain ratio. Data collection scenarios are ranked from 1 to 25 based on the highest IG ratio value and each colored portion of the bar represents the percentage of each data type collected in that scenario. The values on the right hand side of each graph represent the total amount of available resources in increments of $100000. $100k represents $100000 in available resources, $200k represents $200000 in available resources, and $300k represents $300000 in available resources.

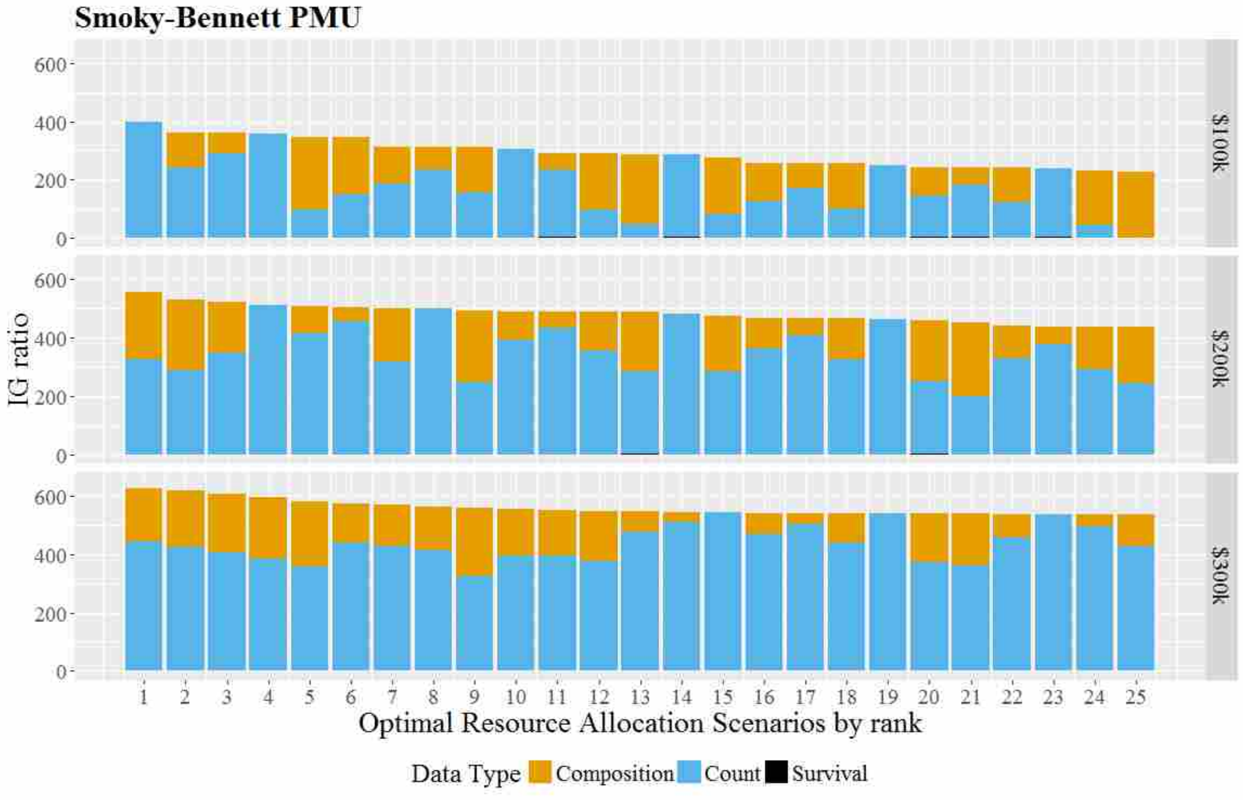


Figure 4-15. The top 25 scenarios showing the optimal allocation of monitoring resources for collecting data to estimate abundance for the South Hills Population Management Unit (PMU). IG represents the information gain ratio. Data collection scenarios are ranked from 1 to 25 based on the highest IG ratio value and each colored portion of the bar represents the percentage of each data type collected in that scenario. The values on the right hand side of each graph represent the total amount of available resources in increments of $100000. $100k represents $100000 in available resources, $200k represents $200000 in available resources, and $300k represents $300000 in available resources.

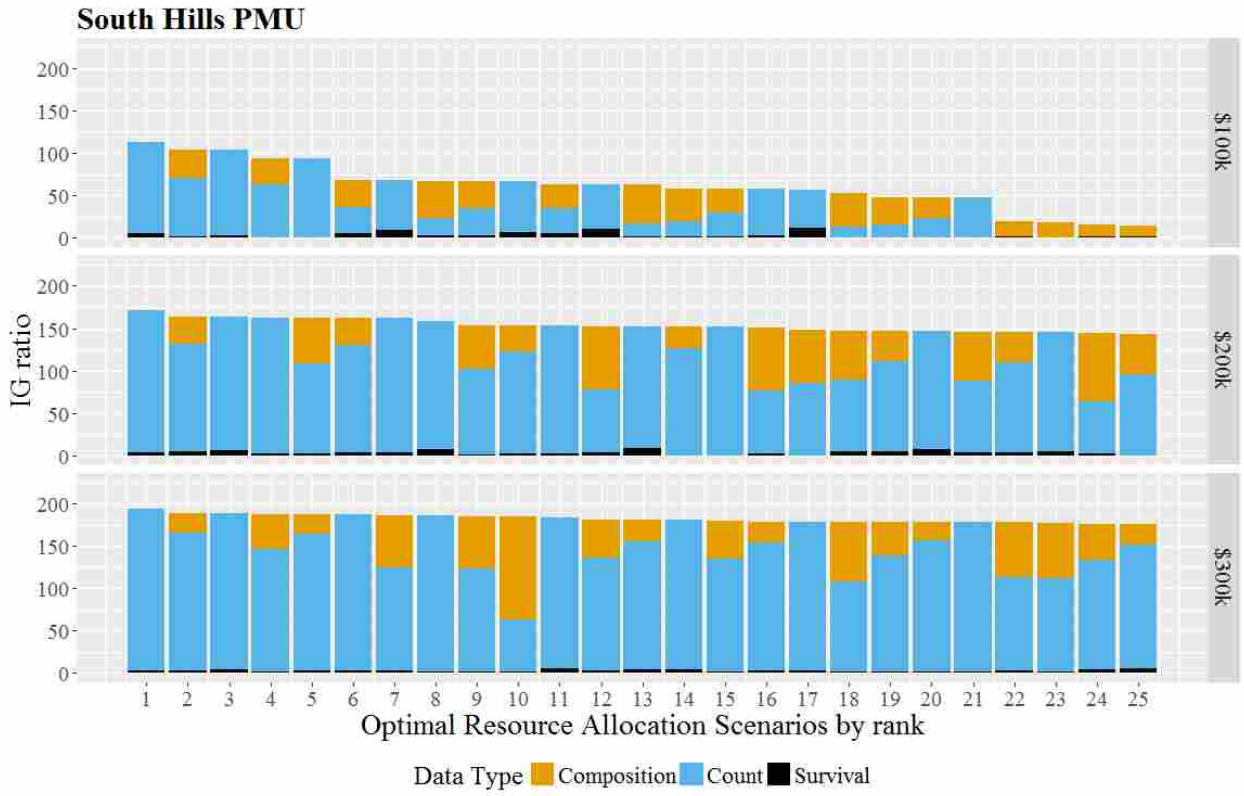
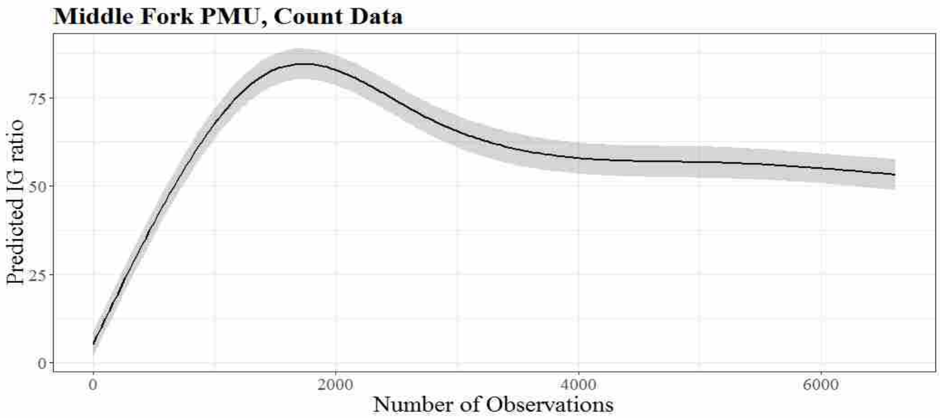
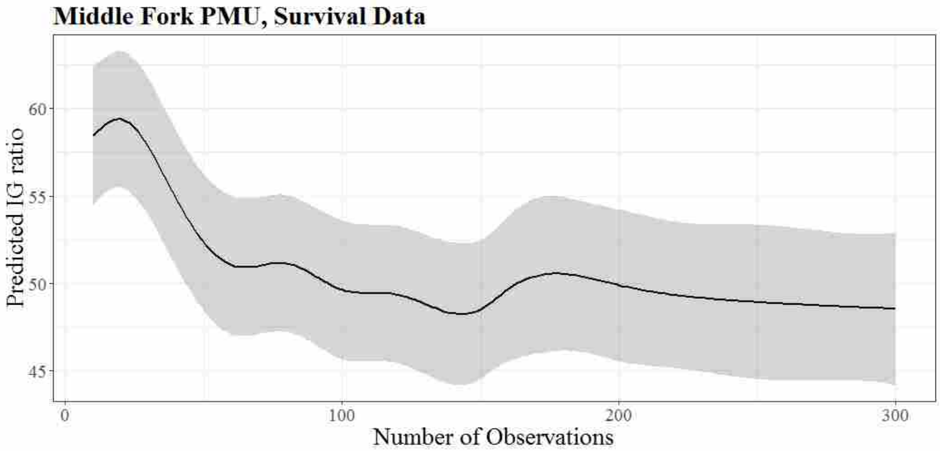


Figure 4-16. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Middle Fork Population Management Unit (PMU) when the cost of a survival observation is $200.





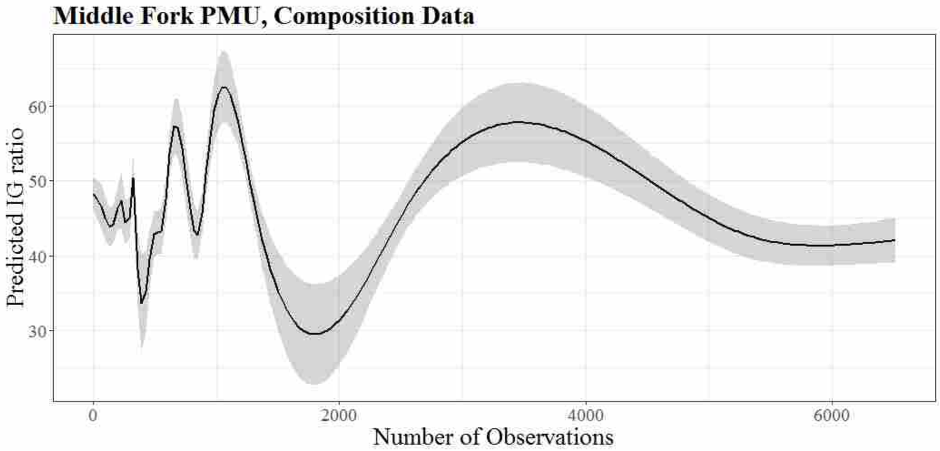
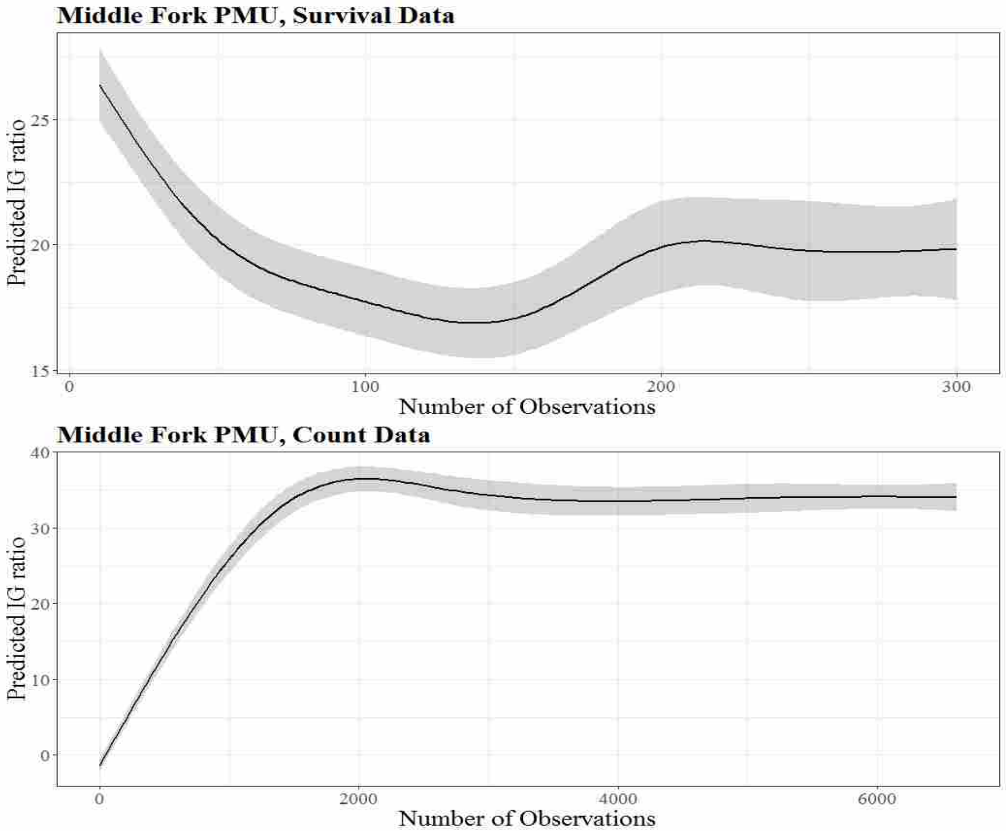


Figure 4-17. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Middle Fork Population Management Unit (PMU) when the cost of a survival observation is $2000.



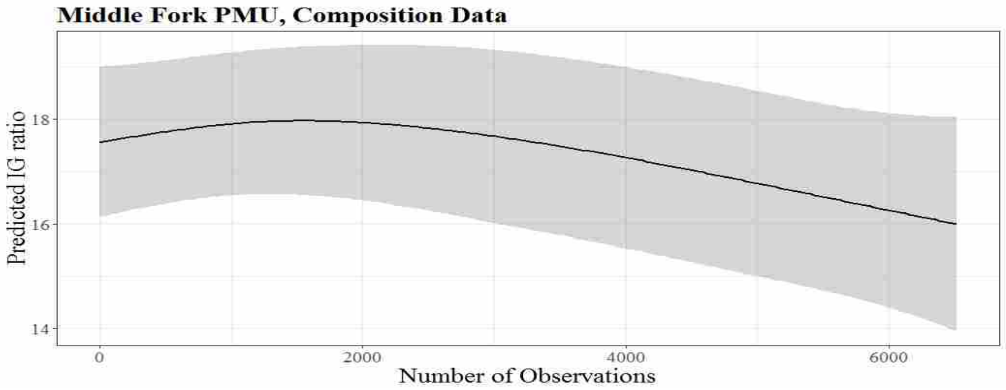
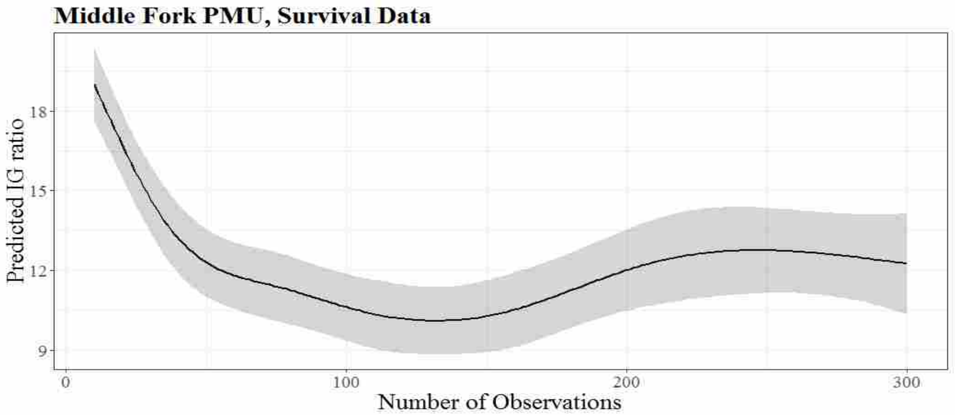
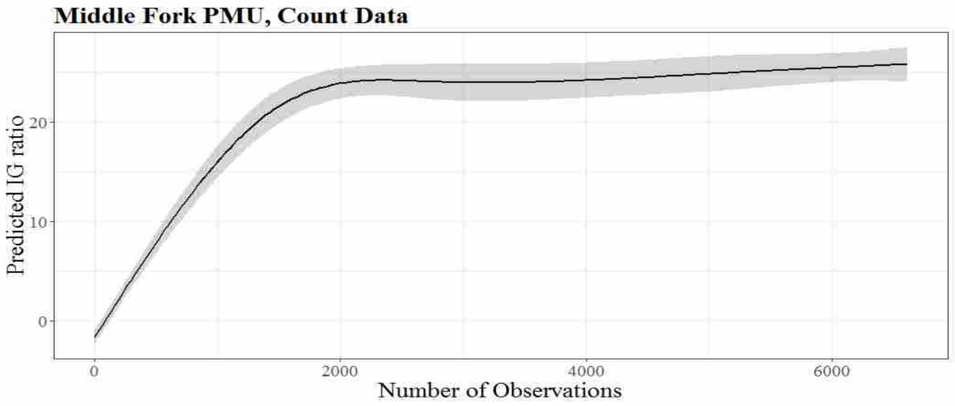


Figure 4-18. Graphical representations of the smoothing functions for survival data, count data, and herd composition data showing the predicted value of the information gain (IG) ratio as the number of observations for each data type increases. These are the results from the Generalized Additive Model (GAM) used to predict the IG ratio value for the Middle Fork Population Management Unit (PMU) when the cost of a survival observation is $4000.





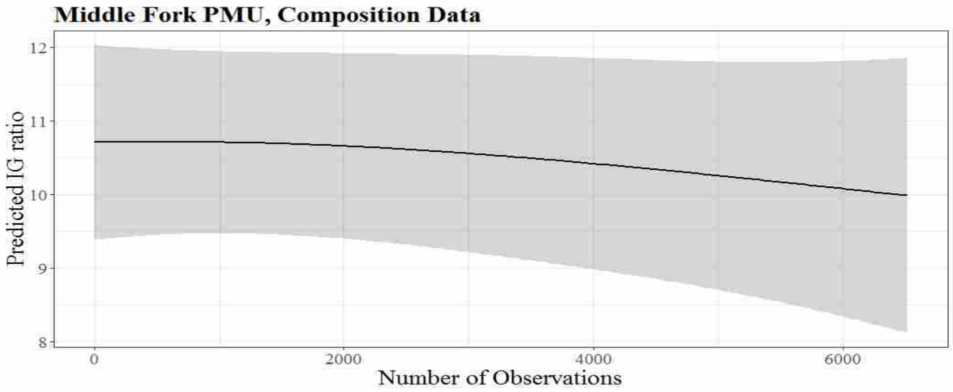
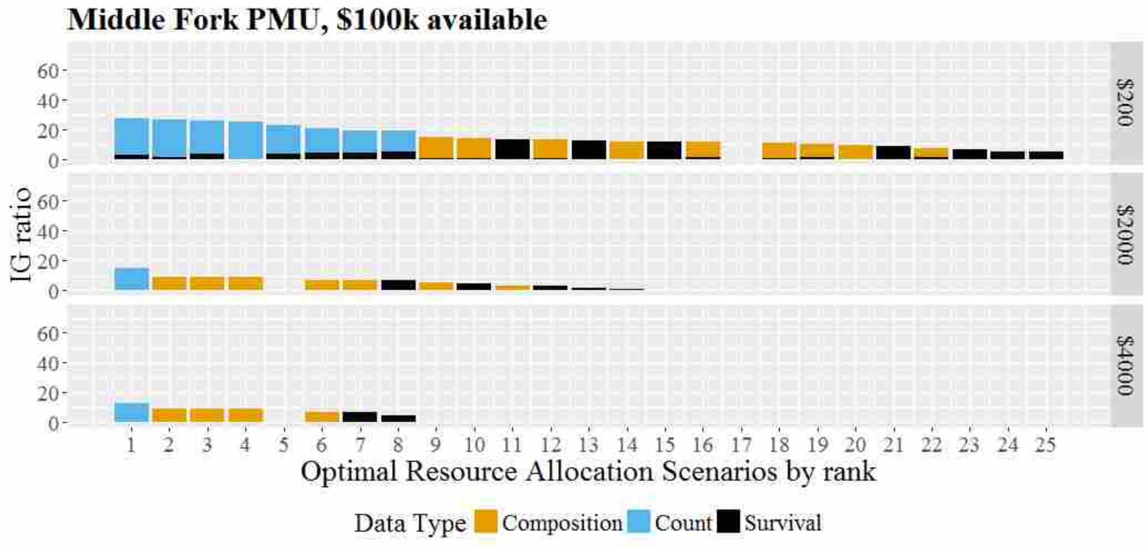
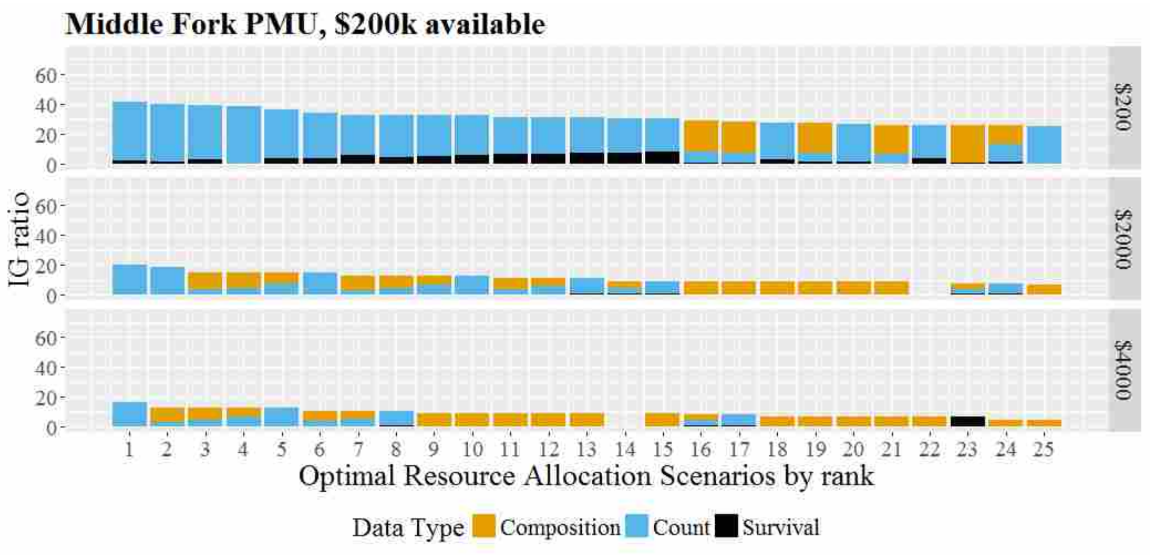
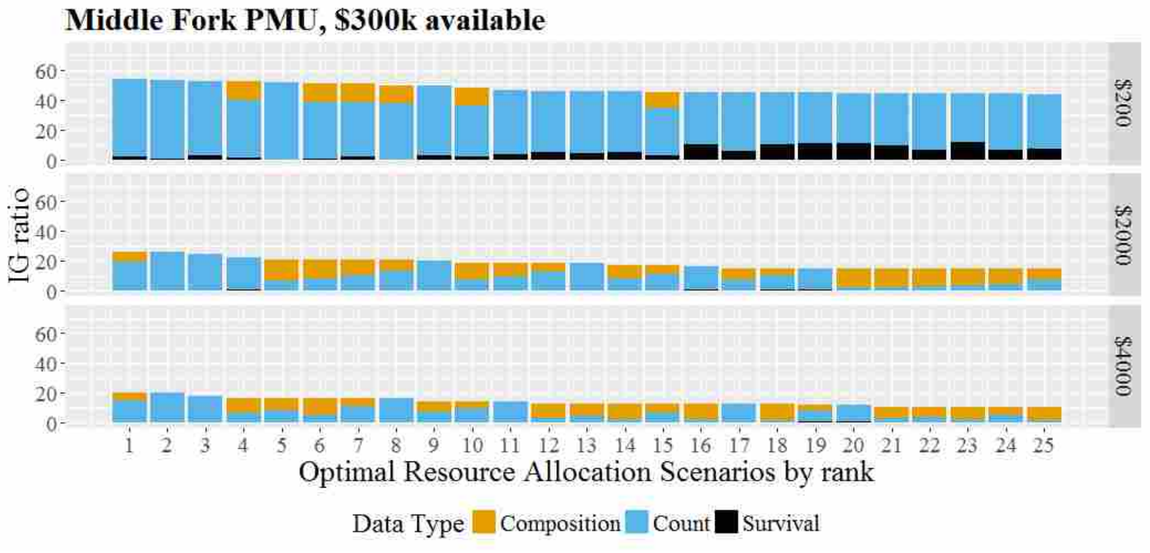


Figure 4-19. The top 25 scenarios showing the optimal allocation of monitoring resources for collecting data to estimate abundance for the Middle Fork Population Management Unit (PMU). IG represents the information gain ratio. Data collection scenarios are ranked from 1 to 25 based on the highest IG ratio value and each colored portion of the bar represents the percentage of each data type collected in that scenario. The values on the right hand side of each graph represent the cost of collecting a single survival observation. $100k represents $100000 in available resources, $200k represents $200000 in available resources, and $300k represents $300000 in available resources.




Middle Fork PMU, $200k available


Middle Fork PMU, $300k available