Comparison of helicopter and ground surveys for North American elk Cervus elaphus and mule deer Odocoileus hemionus population composition

Authors: Louis C. Bender, Woodrow L. Myers, and William R. Gould
Source: Wildlife Biology, 9(4) : 199-205
Published By: Nordic Board for Wildlife Research
URL: https://doi.org/10.2981/wlb.2003.051
Comparison of helicopter and ground surveys for North American elk *Cervus elaphus* and mule deer *Odocoileus hemionus* population composition

Louis C. Bender, Woodrow L. Myers & William R. Gould


Both ground and helicopter surveys are commonly used to collect sex and age composition data for ungulates. Little attention has been paid, however, to whether data collected by each technique are similar. We compared helicopter and ground composition data for both elk *Cervus elaphus* and mule deer *Odocoileus hemionus* across a variety of habitats in the state of Washington, USA. We found that ground and helicopter counts differed (P’s < 0.002) consistently in male age structure estimates for elk, and that the two survey methods differed in estimates of adult sex ratios for mule deer (P = 0.023). Counts from helicopters provided larger sample sizes, tended to be more consistent annually in their results, and were corroborated by other demographic studies of the test populations. We conclude that helicopter and ground surveys differ for male age structure and perhaps male:female ratios, but are similar for young:female ratios. Managers should maintain a standardized technique using the same survey vehicle for trend analysis of composition data.

Key words: age structure, elk, ground surveys, helicopter, mule deer, population composition, sex ratios, surveys

Louis C. Bender*, Washington Department of Fish and Wildlife, 2950 Ley Road, Cle Elum, Washington 98922, USA - e-mail: lbender@nmsu.edu
Woodrow L. Myers, Washington Department of Fish and Wildlife, 8702 North Division, Spokane, Washington 99216, USA - e-mail: myerswlmln@dfw.wa.gov
William R. Gould, University Statistics Center, Department of Economics, New Mexico State University, Las Cruces, New Mexico 88003, USA - e-mail: wgould@nmsu.edu

*Present address: U.S. Geological Survey, New Mexico Cooperative Fish and Wildlife Research Unit, PO Box 30003 MSC 4901, Las Cruces, New Mexico 88003, USA

Corresponding author: Louis C. Bender

Received 19 December 2001, accepted 5 December 2002

Associate Editor: John P. Ball

Ground, helicopter and fixed-wing surveys are the primary techniques used to collect sex and age composition data and other population data for ungulates (Lovaas, Egan & Knight 1966, DeYoung 1985, Beasom, Leon & Synatzske 1986, Ericsson & Wallin 1999). Each of these techniques is frequently used in combination and/or interchangeably in the collection of composition data. However, each technique includes numerous potential biases that may affect the accuracy of the data collected (Caughley 1974, McCullough, Weckerly, Gar-
cia & Evett 1994). These differences may result in the techniques being neither complimentary nor interchangeable. For example, visibility of surveyed animals differs greatly between ground and aerial surveys (Gilbert & Grieb 1957, Caughley 1974, Samuel, Garton, Schlegel & Carson 1987) as well as between helicopter and fixed-wing surveys (R. Spencer & L. Bender, Washington Department of Fish and Wildlife, unpubl. data).

Previous studies comparing survey techniques have contrasted ground and fixed-wing surveys (Gilbert & Grieb 1957, Wolfe & Kimball 1989, Woolley & Lindzey 1997). The consistent conclusion from these studies is that herd composition as found using the two techniques differs, and that whichever technique is used should be used consistently.

Helicopter surveys are anecdotally viewed as highly accurate relative to either fixed-wing or ground surveys (Hess 1997, Smith & Anderson 1998). Reasons for this include sampling a larger number of animals, classifying a greater proportion of animals located, ability to survey broader geographic areas (e.g. surveys are not limited to areas near roads), and better visibility of surveyed animals. However, comparative studies between helicopter and other survey methods are rare. Tsukamoto (1977) found that spring fawn:adult ratios of mule deer Odocoileus hemionus did not differ between helicopter and ground samples. However, Tsukamoto (1977) looked at only one demographic characteristic (fawn: adult ratio), rather than the range of commonly collected ratios (e.g. male:female and male age structure).

We compared composition data from helicopter and ground surveys for both North American elk Cervus elaphus and mule deer in the state of Washington, USA. Our goal was to test similarity of survey data from these two methods of assessing sex and age ratios. Specific objectives included: 1) comparing ground and helicopter surveys in terms of resultant male:female, young:female and male age structure, and 2) recommending appropriate application of both survey types.

Material and methods

Study area
We conducted elk composition surveys in both heavily-forested western Washington and the more open grassland/forest mix of southeastern Washington. Western Washington elk composition surveys were in the western Cascade Mountains of southwest Washington, specifically Game Management Unit (GMU) 524 (Margaret unit; approximately 46°10'N, 122°40'W). The site was >80% forested and was dominated by tree farms of Douglas-fir Pseudotsuga menziesii and western hemlock Tsuga heterophylla. Characteristic species of riparian areas were red alder Alnus rubra and big-leaf maple Acer macrophyllum.

Eastern Washington elk composition surveys were conducted in the northern Blue Mountains over the entire Blue Mountains herd area, which included six GMUs (Blue Creek, Watershed, Dayton, Tucannon, Mountain View and Lick Creek; approximately 46°15'N, 117°30'W). Vegetation in the Blue Mountains varied considerably with elevation and aspect. Bunch grass communities including bluebunch wheatgrass Agropyron spicatum/Idaho fescue Festuca idahoensis or bluebunch wheatgrass/blue grass Poa secunda were common at lower and intermediate elevations and south-facing slopes. Dense mallow ninebark Physocarpus malvaceous stands were found on steep north and east-facing slopes of low or intermediate elevations. Ponderosa pine Pinus ponderosa, Douglas-fir, and grand fir Abies grandis were found in both open and dense stands at intermediate and high elevations.

We conducted mule deer surveys in the Methow valley of north-central Washington, located on the east slope of the northern Cascade Mountains in western Okanogan County, including five GMUs (Chewuch, Pearlygin, Gardner, Chiliwist and Alta; approximately 48°20'N, 121°10'W). This study site was covered by a mix of shrub-steppe and forested communities. Mid and lower elevation south-facing slopes were dominated by bluebunch wheatgrass, bluebunch wheatgrass/bitterbrush Purshia tridentata, bluebunch wheatgrass/big sagebrush Artemisia tridentata, or ponderosa pine/bitterbrush communities. Douglas-fir communities were found on north slopes at low and intermediate elevations.

Elk composition surveys
In western Washington, we flew helicopter surveys on 22 September 1995 and 24 September 1996 for pre-hunting season elk composition, using a Bell 206B Jet Ranger flown by a pilot highly experienced in composition work and three experienced observers. The entire Margaret site was flown in a single flight lasting from three hours before sunset to sunset. The lead observer in the helicopter surveys conducted ground surveys < 5 days after the completion of the helicopter survey. For ground surveys, a permanent route was established along ridgetop roads that allowed visual coverage of >80% of open habitat types (mostly clearcuts and other early successional communities) in the site. We drove the ground route from three hours before sunset until sunset on two consecutive days to allow coverage of the entire site, and recorded all elk encountered along the
route. We classified all observed elk as bulls, cows or calves. We further segregated bulls by antler points and placed them into one of two categories, yearling or adult (Bender & Miller 1999).

In eastern Washington, we flew early-summer elk composition surveys during 26-30 June 1996-1997; surveys were flown during the first two hours after sunrise and late afternoon-early evening prior to sunset to avoid hot mid-day temperatures and periods of lower elk activity. We surveyed high and mid-elevation portions of the GMUs in the Blue Mountains identified as summer use areas based upon observed movements of radio-marked elk (Myers 1999). Fall composition surveys were flown the third week of September over the same search areas as the mid-summer surveys and using the same protocol. We flew all helicopter surveys of elk over eastern Washington in a Hiller 12E helicopter with an experienced pilot and two experienced observers. During the first two weeks of July 1996-1997, we conducted summer ground counts daily during early morning and late evening hours by driving roads which followed ridges through summer use areas and stopping periodically to glass meadows and other openings with binoculars and spotting scopes. Fall ground counts occurred between late August through September 1996-1997, following the same driving routes and protocol as the mid-summer ground counts. During all surveys, we classified observed elk as bulls, cows or calves; bulls were further divided into yearlings or adults based on antler characteristics (Bender & Miller 1999). We pooled count data within methods (ground or helicopter) so that survey data provided a single representation of population composition for the entire Blue Mountains site. To minimize the potential for double counting among flights, adjacent areas were separated by ≥8 km from their common border unless we surveyed these units on the same day. This separation distance corresponds to twice the diameter of the average home range of elk in Washington (Bender & Miller 1999).

Mule deer composition surveys
We flew surveys over mule deer winter ranges in the Methow Valley (Zeigler 1978) to measure mule deer composition during the first three weeks of December each year during 1984-1986. We surveyed ≥80% of the winter range using a Hughes 500C or 500D model helicopter with a crew of three experienced observers and pilot. We classified deer as bucks, does or fawns. Ground counts were conducted in December following completion of helicopter surveys each year. We counted deer from the ground by hiking four traditional (established >10 years) routes that varied in length from 5 to 8 km; deer classifications were similar to those of the helicopter surveys. We pooled data within methods to be representative of the entire study site as described under elk composition surveys. No snow cover was present during any mule deer surveys.

Data analysis
Our design involved paired surveys conducted by experienced observers on the same sites during the same time period, with the same observer doing both helicopter and ground surveys to minimize observer biases known to be present in composition surveys (Caughley 1974). We conducted each survey pair within a time frame that was short enough to be considered biologically simultaneous, especially since hunting harvest, the major factor influencing elk and deer demographics in Washington (Bender & Miller 1999), was either absent or trivial during the time between survey pairs. Each site was examined separately.

For a given site, our design could be considered a repeated-repeated measures, with the helicopter vs ground counts considered as one repeat and years or seasons as another temporal component. However, we have treated years independently based on our knowledge of substantial population changes (i.e. high turnover of males annually at all sites and transfer of young from a previous year into ‘adult’ categories the following year). For the Blue Mountains site, seasons were also treated separately due to differential sightability and social grouping patterns of elk between early summer and the fall rut (Franklin & Lieb 1979, Myers 1999). Given the discrete nature of our data (counts of numbers of males, females and young, and numbers of males in each age category) we used the Cochran-Man- tel-Haenszel Test (Cochran 1954, Mantel & Haenszel 1959) to identify any general association between survey method and the elk or deer counts. This test determines if there is an association between survey method and age/gender ratios while controlling for survey year, i.e. years are treated as strata. When the directionality of the association is reversed from one stratum (year) to another, the test is likely to indicate a non-detectable result.

A statistically significant result indicates that for at least one stratum, there is an association between survey method and age/gender ratios. When an association was detected, individual examinations of 2×2 tables were used in conjunction with odds ratios for describing the relationship(s) underlying the result (Agresti 1990). The odds ratio describes the ratio of the odds that an observation within a row is classified in the first column, which in our case corresponds to the odds that an animal is classified as a male in a male:female composition.
Table 1. Male age structure (SE) and male:female:young ratios (SE) for helicopter and ground composition survey pairs for elk and mule deer in Washington. For elk, the composition surveys were collected in western Washington (Margaret, during fall (F) 1995 and 1996) and eastern Washington (Blue Mountains, during summer (S) and fall (F), 1996 and 1997), and for mule deer in eastern Washington (Methow, during 1984-1986). Survey results shown in italics within a site year/season contrast consistently differ (P < 0.05).

<table>
<thead>
<tr>
<th>Species</th>
<th>Area (Time)</th>
<th>Type</th>
<th>Male age classes</th>
<th>Male:female:young</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yearling</td>
<td>Mature</td>
<td></td>
</tr>
<tr>
<td>Elk</td>
<td>Margaret (F95)</td>
<td>Air</td>
<td>0.34(0.03)</td>
<td>0.66(0.06)</td>
<td>57(4):100:55:4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>0.38(0.04)</td>
<td>0.82(0.14)</td>
<td>90(1):100:46:6</td>
</tr>
<tr>
<td></td>
<td>Margaret (F96)</td>
<td>Air</td>
<td>0.38(0.01)</td>
<td>0.62(0.02)</td>
<td>54(3):100:45:2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>0.21(0.03)</td>
<td>0.79(0.11)</td>
<td>37(3):100:48:3</td>
</tr>
<tr>
<td></td>
<td>Blues (S96)</td>
<td>Air</td>
<td>0.63(0.05)</td>
<td>0.37(0.03)</td>
<td>9(1):100:49:3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>0.91(0.08)</td>
<td>0.09(0.01)</td>
<td>9(3):100:52:8</td>
</tr>
<tr>
<td></td>
<td>Blues (S97)</td>
<td>Air</td>
<td>0.24(0.01)</td>
<td>0.76(0.03)</td>
<td>18(1):100:52:3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>0.86(0.11)</td>
<td>0.14(0.02)</td>
<td>9(3):100:48:8</td>
</tr>
<tr>
<td></td>
<td>Blues (F96)</td>
<td>Air</td>
<td>0.36(0.05)</td>
<td>0.64(0.09)</td>
<td>20(3):100:41:4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>0.64(0.10)</td>
<td>0.36(0.06)</td>
<td>21(4):100:63:9</td>
</tr>
<tr>
<td></td>
<td>Blues (F97)</td>
<td>Air</td>
<td>0.33(0.08)</td>
<td>0.67(0.16)</td>
<td>17(3):100:36:5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>0.93(0.07)</td>
<td>0.07(0.01)</td>
<td>12(3):100:29:5</td>
</tr>
<tr>
<td>Mule deer</td>
<td>Methow (84)</td>
<td>Air</td>
<td>-</td>
<td>-</td>
<td>3(0.5):100:75:5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>-</td>
<td>-</td>
<td>4(0.5):100:75:5</td>
</tr>
<tr>
<td></td>
<td>Methow (85)</td>
<td>Air</td>
<td>-</td>
<td>-</td>
<td>3(0.1):100:76:1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>-</td>
<td>-</td>
<td>2(0.5):100:75:5</td>
</tr>
<tr>
<td></td>
<td>Methow 86</td>
<td>Air</td>
<td>-</td>
<td>-</td>
<td>10(0.5):100:64:2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ground</td>
<td>-</td>
<td>-</td>
<td>6(0.5):100:58:4</td>
</tr>
</tbody>
</table>

Results

Standard errors for surveys were small relative to means because of large sample sizes and the fact that a large proportion of each population was generally counted. There was a detectable association between survey method and age ratios for male elk at both the Margaret ($\chi^2_1 = 9.56; P = 0.002$) and Blue Mountain site (Spring: $\chi^2_1 = 13.20, P = 0.001$; Fall: $\chi^2_1 = 14.28, P = 0.001$). The odds of classifying a male elk as mature at Margaret by helicopter survey were 0.42 times that by ground counts (95% confidence interval: 0.24-0.73). In the Blue Mountains, odds of classifying a male elk as mature by helicopter survey were 10.49 times that by ground counts during summer surveys (95% confidence interval: 2.30-47.79). Fall surveys were similar in that the odds of classification as an adult by helicopter survey were greater (3.06 times) than by ground counts in 1996. In 1997, the odds ratio increased to 26.0, meaning the odds of classifying a male elk as an adult were 26 times higher for helicopter than for ground surveys. The large differences between years prevented presentation of a common odds ratio for fall surveys in the Blue Mountains.

Bull:cow:calf ratios were not associated with survey method in a consistent manner for any of the elk site-season combinations (Margaret: $\chi^2_2 = 0.027, P = 0.987$; Blue Mountains summer: $\chi^2_2 = 1.52, P = 0.466$; and Blue Mountains fall: $\chi^2_2 = 0.98, P = 0.613$). In general, there were differences between paired ratios, but directionality of differences was inconsistent between years for elk.

We found an association between survey method and mule deer buck:doe:fawn ratios ($\chi^2_2 = 7.56; P = 0.023$). This association was significant due to differences in buck:doe ratios, which were higher for helicopter surveys than for ground surveys (odds ratio = 1.66 in 1985 and 1986) except for 1984, when the odds ratio was 0.81. There was virtually no difference between doe:fawn ratios between methods (odds ratio = 1.0 in 1984 and 1985), with the exception of 1986 in which the odds ratio equaled 0.91 for ground survey classification of females.

Discussion

Helicopter surveys are anecdotally considered more accurate than ground surveys, due to several factors associated with survey procedures (Hess 1997, Smith & An...
Comparative tests of data obtained from helicopter and ground surveys are rare for ungulates, however (Tsukamoto 1977). We found that ground and helicopter counts differed consistently in male age structure estimates for elk, and that the two methods differed in estimates of adult sex ratios for mule deer. Counts from helicopters tended to be more consistent annually in their results (see Table 1), which matched biologists’ subjective views on population parameter trends and were more consistent with other data from the study sites (see below). Moreover, sample sizes from helicopter counts were always larger, which resulted in helicopter counts having smaller or equal standard errors in all comparisons of male:female and young:female ratios, and in 7 of 12 comparisons of male age structure (the exceptions being when proportions from ground counts approached 0 or 1, which resulted in very small standard errors despite smaller sample sizes due to the binomial data; see Table 1). Thus, our data supported suggestions by Hess (1997) and Smith & Anderson (1998) that helicopter surveys may be superior to ground surveys for ungulate composition.

Differences in age structure estimates were consistent between ground and helicopter counts at a given site, but not between sites. Helicopter counts were less likely to classify a bull elk as an adult in Margaret (odds ratios = 0.42), but more likely to classify a bull as an adult in the Blue Mountains for either summer (odds ratios = 5.8-19.4) or fall (odds ratios = 3.1-26.0). Differing odds of classifying bulls between the sites was likely due to cover characteristics associated with the areas. The Margaret site was comprised of dense forested habitat, where observation distances were limited and most elk were consequently observed at close distances. This likely facilitated identification of yearling bulls with small spike antlers, which are often difficult to see in cover or at long distances. Further, adult bull elk occur in smaller social groups, which are less likely to be observed in heavy cover (Franklin & Lieb 1979, Samuel et al. 1987). Conversely, elk in the more open grassland-woodland mixture of the Blue Mountains were more likely to be detected and observed at greater distances (Samuel et al. 1987). At longer distances, yearling bulls with smaller spike antlers are more likely to be misidentified, especially since they are more frequently associated with larger cow groups than with bull groups (Franklin & Lieb 1979). Although the directionality of differences varied between study sites, ground and helicopter surveys consistently differed in bull age structure estimates, indicating that the two methods were not interchangeable nor complimentary in collection or interpretation of age structure data.

Differences between ground and helicopter surveys were detectable for mule deer composition, due to differing buck:doe ratios. The odds of classifying a mule deer as a buck were generally higher from helicopter surveys (odds ratios = 0.81, 1.66, and 1.66), probably because helicopter counts allowed more area to be covered and thus greater chance of observing bucks, which occur in smaller groups and are segregated from does during this time (Zeigler 1978, Geist 1998). Despite often large differences between survey pairs, especially for bull:cow ratios, inconsistent directionality between years in bull:cow:calf ratios for our elk surveys likely prevented statistical significance since our testing procedure looked for a consistent effect between techniques. For example, ground count estimates of bull:cow ratios were much higher than helicopter estimates in 1995 for Margaret, but were much lower in 1996 (see Table 1). Because of this variability, we caution against using ground and helicopter counts interchangeably for male:female ratios for either mule deer or elk. Similar to Tsukamoto (1977), we found no differences between techniques in classifying either elk or mule deer in terms of young:female ratios.

We found that ground and helicopter composition surveys did not produce similar results for male age structure estimates, nor buck:doe deer ratios. Which method, if either, was accurate? Based on our experience with the study populations, including long-term population dynamics studies (Myers, Naney, Lloyd, Quinn & Dixon 1990, Myers 1999, Bender & Miller 1999), we feel that helicopter counts were more representative than ground counts. First, mortality rates derived from bull elk age structure collected from a helicopter did not differ from telemetry-based estimates for the Margaret site (Bender & Miller 1999). Similarly, bull:cow ratios derived from comparisons of bull and cow elk age structure (Lang & Wood 1976) agreed with helicopter composition estimates for both eastern and western Washington elk herds, including the Margaret and Blue Mountains populations (L. Bender, Washington Department of Fish and Wildlife, unpubl. data). Further, parameters estimated from helicopter counts fit better with predicted demographic trends from simulation models using either mark–resight, aerial sight-bias, or reconstruction population estimates and known harvests (Bender & Spencer 1999, Myers 1999). Thus, although we were unable to compare each technique with true values due to the inability to census large free-ranging populations (Bender & Spencer 1999), we feel that helicopter counts provide a much more accurate accounting of sex and age characteristics for species such as elk and mule deer for which such characteristics can be differentiated from the air.
This conclusion may not hold for species such as bison *Bison bison*, which are difficult to identify by sex from the air (Wolfe & Kimball 1989).

We agree with the conclusions of Woolley & Lindzey (1997) in their analysis of fixed-wing vs ground surveys for pronghorn *Antilocapra americana*, and reiterate the same cautions here. Herd composition data collected using differing techniques (e.g. helicopter vs ground, fixed-wing vs ground, and likely helicopter vs fixed-wing) may differ substantially, and whatever technique is used should be used consistently. Managers should not mix survey techniques, since they may give different results (Wolfe & Kimball 1989, Woolley & Lindzey 1997; see Table 1), and it is either difficult or impossible to determine which technique, if any, is accurate. This is especially problematic with male age structure, for which we believe representative data cannot be collected from ground surveys. We caution that this may be true for male:female ratios as well, but note that the data is less definitive. Lastly, young:female or young:adult ratios have not been shown to consistently differ between methods, and may be collected by either ground or helicopter surveys (Tsukamoto 1977; see Table 1). Managers should maintain a standardized technique using the same survey vehicle for trend analysis of population data. If composition data is used to characterize the population rather than for relative trends, helicopters are likely more accurate than the other vehicles and should be used exclusively.

Acknowledgements - funding for this project was provided by the Washington Department of Fish and Wildlife through Pittman-Robertson contracts. We thank P. Fowler, J. King and P. Miller for field assistance.

References

Smith, B.L. & Anderson, S.H. 1998: Juvenile survival and pop-