Supporting Information

8 February 2022

Forsyth, D. M., S. Comte, N. E. Davis, A. J. Bengsen, S. D. Côté, D. G. Hewitt, N. Morellet, and A. Mysterud. Methodology matters when estimating deer abundance: a global systematic review and recommendations for improvements. Journal of Wildlife Management

**S7. Bias of Deer Abundance and Density Estimation Methods**

**Definition of Accuracy and Bias**

We defined accuracy as the absolute difference between estimated abundance (*N̂*) and true abundance (*N*), and bias as the signed difference (i.e., negative or positive; Caughley 1974, Pollock and Kendall 1987, Hone 2008). Accuracy has also been defined as squared deviations of estimated abundance from true abundance (Hone 2008), but that definition has seldom been used in the deer abundance estimation literature and is not used here. Hence, accuracy can be considered the absolute value of bias, such that methods can be considered more or less accurate without referring to the direction of the bias. For this review, however, it was important to understand the direction (either an underestimate or overestimate of *N*) and magnitude of the bias of deer abundance estimation methods.

**Criteria for study inclusion**

The key requirement for a study to provide a robust assessment of accuracy and bias was to know the true abundance of the deer population. This meant that 2 criteria needed to be met (Pollock and Kendall 1987, McCullough and Hirth 1988). First, the deer population had to be physically closed to immigration, emigration, and temporary movements; in practice, this usually required populations to be physically enclosed by a deer-proof fence. Second, the number of deer in the population had to be known with complete, or near-complete, certainty. In practice, this meant that the population was either introduced to the enclosure immediately prior to the assessment being undertaken, or that the population was subject to capture and marking with or without some other intensive monitoring (e.g., direct observations or population reconstruction) immediately prior to the assessment.

These criteria excluded simulation studies of accuracy and bias, and also studies of real deer populations that used one method to estimate *N* (with potentially substantial uncertainty) and then compared the result from another method with that estimate. Both approaches have merit, but neither provides a robust assessment of accuracy and bias as per our definitions. Simulation studies can identify important assumptions of sampling designs and analyses but do not account for the multitude of variables potentially affecting field assessments of deer abundance. The second approach potentially involves a tautology because it assumes that one abundance estimate is more accurate than another. In our review of the literature, it proved straightforward to include or exclude studies within deer-proof enclosures in which deer abundance was known with complete, or near complete, certainty (i.e., include) or with potentially substantial uncertainty (i.e., exclude).

**Results**

Our literature searches (described in the main text) identified only 15 peer-reviewed journal articles and 1 book chapter that met our criteria for robustly assessing the bias of deer abundance estimation methods. One of these studies was identified in the first search (Hewison et al. 2007), 3 studies were identified in the second search (Vincent et al. 1996, Zabransky et al. 2016, Beaver et al. 2020) and 12 studies were identified in the third search (LeResche and Rausch 1974, Rice and Harder 1974, Floyd et al. 1979, McCullough 1982, Gasaway et al. 1985, Bartmann et al. 1986, Beasom et al. 1986, Bartmann et al. 1987, McCullough and Hirth 1988, White et al. 1989, Cogan and Diefenbach 1998, Maillard et al. 2010).

Since there were insufficient studies for a quantitative analysis of bias, we qualitatively summarized the bias reported in those publications (Table 1 of the main text) as follows. The magnitude of the bias was summarized as either minor (i.e., [*N̂* – *N*] / *N* is <10%), moderate (i.e., [*N̂* – *N*] / *N* is 10–25%), or major (i.e., [*N̂* – *N*] / *N* is >25%).

Below we briefly describe those publications in terms of how absolute deer abundance (*N*) was known, the deer species, the region, and the abundance estimation method that was assessed. Studies are grouped by first-order method; there were no studies of pedestrian sign counts, motion-sensitive cameras, or harvest data. Readers wanting further information can examine the relevant publication(s).

*Pedestrian direct counts.*—McCullough and Hirth (1988) used a 464-ha deer-proof enclosure stocked with white-tailed deer (*Odocoileus virginianus*) to test 1 pedestrian direct count method (capture–recapture). For the 31 monthly estimates, the mean bias was 46% (i.e., there was major positive bias). The monthly bias ranged from –26% to 407%.

Vincent et al. (1996) used a 130-ha deer-proof enclosure stocked with fallow deer (*Dama dama*) to evaluate 2 pedestrian direct count methods (capture–recapture and strip count) in Europe. Capture–recapture provided estimates of between −8.1 and +8.2% (i.e., minor negative and positive bias). Strip counts provided estimates of between −19.5% and −3.4% bias (i.e., moderate and minor negative bias).

Hewison et al. (2007) evaluated the accuracy and bias of the drive count (conducted in fragmented habitat) and walked count (forest habitat) methods to estimate roe deer (*Capreolus capreolus*) abundance in Europe. A sample of deer was radio-collared so that the number of marked individuals available to be counted was known with near certainty. A mean 75% of collared deer was counted by drive count (i.e., there was moderate negative bias). The walked count (with a capture–mark–recapture approach) detected 22% of deer (i.e., there was major negative bias).

One book chapter briefly summarized an unpublished study that apparently meets our criteria for inclusion. Maillard et al. (2010: 465–466) reported that “only 30% of the marked roe deer were counted during a total drive” of an enclosed roe deer population in Europe, but we did not find any further information on this assessment (i.e., there was major negative bias).

*Vehicle direct counts.*—McCullough (1982) used a 464-ha deer-proof enclosure stocked with white-tailed deer to test 1 vehicular direct count method (spotlight counts). Less than 50% of deer were counted (i.e., there was major negative bias).

*Aerial direct counts.*—LeResche and Rausch (1974) used deer-proof enclosures stocked with moose (*Alces alces*) to evaluate one aerial direct count method (fixed-wing strip count) in North America. In the best conditions, experienced and inexperienced observers counted 68% and 43% of moose, respectively (i.e., there was major negative bias).

Rice and Harder (1977) used a 122-ha deer-proof enclosure to evaluate helicopter capture–recapture of white-tailed deer in North America. The true population size was determined by a drive count with a maximum distance of 3 m between drivers. The mean estimate had minor positive bias (+3%), but if they had ceased sampling after 4 rather than 5 samples then the mean estimate would have had moderate negative bias (−17%).

Floyd et al. (1979) used a population of radio-collared white-tailed deer to evaluate one aerial direct count method (total counts by fixed-wing) in North America. Between 50% and 56% of the radio-collared deer were observed (i.e., there was major negative bias).

Gasaway et al. (1985) used individually marked moose (ear-tags and branding) to evaluate one aerial direct count method (total count from fixed-wing) in North America. Thirty-six percent and 26% of moose were observed in total counts before and after leafout (i.e., there was major negative bias).

Beasom et al. (1986) used individually marked white-tailed deer (numbered collars) to evaluate 1 aerial direct count method (total count from helicopter) in 2 enclosures in North America. The mean percentage of marked deer seen in the total counts (i.e., 100% of area surveyed in their Table 1) in the 2 enclosures was 26% and 30%.

Cogan and Diefenbach (1998) used known group sizes of elk (*Cervus canadensis*) in North America to assess the accuracy of one aerial direct count method (sightability models). The group sizes were determined in open conditions by fixed-wing aircraft or ground crews prior to or during the surveys, respectively. It was found that “the 36 observations of elk groups of known size later observed by helicopter crews indicated that undercounting occurred, but in no instance did overcounting occur” (Cogan and Diefenbach 1998: 274). The magnitude of the bias was not reported.

Zabransky et al. (2016) used individually marked mule deer (*Odocoileus hemionus*) (fitted with global positioning system collars) to evaluate the accuracy of one aerial direct count method (total count from helicopter) in North America. The percentage of marked deer detected averaged 42 ± 16% and ranged among surveys from 19% to 77%.

Three journal articles used deer-proof enclosures stocked with mule deer to evaluate 3 aerial direct count methods (helicopter quadrat counts [Bartmann et al. 1986], helicopter capture–recapture [Bartmann et al. 1987], and helicopter distance-sampling methods [White et al. 1989]) in North America. The helicopter quadrat count method counted, on average, 60–68% of deer (i.e., there was major negative bias; Bartmann et al. 1986). When the percentage of marked deer was >30% of the total population, the 3 helicopter capture–recapture analysis approaches gave mostly negatively biased estimates; when the percentage of marked deer was <30% of the total population, the estimates had positive bias (Bartmann et al. 1987). Because the true abundance was not always known with certainty (rather, a range was given), simple summaries of bias cannot be given for this study. Interested readers should consult Table 1 in Bartmann et al. (1987) for further information. The helicopter distance-sampling method counted 90% (exponential polynomial model) and 102% (negative exponential model) of the deer (i.e., there was minor negative or no bias; White et al. 1989).

Beaver et al. (2020) used a 174-ha deer-proof enclosure containing white-tailed deer to evaluate one aerial direct count method (an unmanned aerial vehicle equipped with a thermal sensor conducting strip counts along parallel transects). The deer population was intensively monitored, by annual capturing and unique marking of individuals, and by motion-sensitive cameras. The population during the study was 151–163 deer. When morning and evening survey data were both used, the estimate was 78% of the mean known population. Thermal contrast between animals and background was better during evening flights, and when only evening survey data were used the estimate was 92% of the mean known value.

**Literature Cited**

Bartmann, R. M., L. H. Carpenter, R. A. Garrott, and D. C. Bowden. 1986. Accuracy of helicopter counts of mule deer in pinyon-juniper woodland. Wildlife Society Bulletin 14:356–363.

Bartmann, R. M., G. C. White, L. H. Carpenter, and R. A. Garrott. 1987. Aerial mark–recapture estimates of confined mule deer in pinyon-juniper woodland. Journal of Wildlife Management 51:41–46.

Beasom, S. L., F. G. Leon III, and D. R. Synatzske. 1986. Accuracy and precision of counting white-tailed deer with helicopters at different sampling intensities. Wildlife Society Bulletin 14:364–368.

Beaver, J. T., R.W. Baldwin, M. Messinger, C. H. Newbolt, S. S. Ditchkoff, and M. R. Silman. 2020. Evaluating the use of drones equipped with thermal sensors as an effective method for estimating wildlife. Wildlife Society Bulletin 44:434–443.

Caughley, G. 1974. Bias in aerial survey. Journal of Wildlife Management38:921–933.

Cogan, R. D., and D. R. Diefenbach. 1998. Effect of undercounting and model selection on a sightability-adjustment estimator for elk. Journal of Wildlife Management 62:269–279.

Floyd, T. J., L. D. Mech, and M. E. Nelson. 1979. An improved method for censusing deer in deciduous-coniferous forest. Journal of Wildlife Management 43:258–261.

Gasaway, W. C., S. D. Dubois, and S. Harbo. 1985. Biases in aerial transect surveys for moose during May and June. Journal of Wildlife Management49:777–784.

Hewison, A. J. M., J. M. Angibault, B. Cargnelutti, A. Coulon, J.-L. Rames, E. Serrano, H. Verheyden, and N. Morellet. 2007. Using radio-tracking and direct observation to estimate roe deer *Capreolus* *capreolus* density in a fragmented landscape: a pilot study. Wildlife Biology 13:313–320.

Hone, J. 2008. On bias, precision and accuracy in wildlife aerial surveys. Wildlife Research 35:253–257.

LeResche, R. E., and R. A. Rausch. 1974. Accuracy and precision of aerial moose censusing. Journal of Wildlife Management 38:175–182.

Maillard, D., J.-M. Gaillard, A. J. M. Hewison, P. Ballon, P. Duncan, A. Loison, C. Toïgo, E. Baubet, C. Bonenfant, M. Garel, et al. 2010. Ungulates and their management in France. Pages 441–474 *in* M. Apollonio, R. Andersen, and R. Putman, editors. European ungulates and their management in the 21st century. Cambridge University Press, Cambridge, United Kingdom.

McCullough, D. R. 1982. Evaluation of night spotlighting as a deer study technique. Journal of Wildlife Management 46:963–973.

McCullough, D. R., and D. H. Hirth. 1988. Evaluation of the Petersen–Lincoln estimator for a white-tailed deer population. Journal of Wildlife Management 52:534–544.

Pollock, K. H., and W. L. Kendall. 1987. Visibility bias in aerial surveys: a review of estimation procedures. Journal of Wildlife Management 51:502–510.

Rice, W. R., and J. D. Harder. 1977. Application of multiple aerial sampling to a mark–recapture census of white-tailed deer. Journal of Wildlife Management 41:197–206.

Vincent, J.-P., A. J. M. Hewison, J.-M. Angibault, and B. Cargnelutti. 1996. Testing density estimators on a fallow deer population of known size. Journal of Wildlife Management 60:18–28.

White, G. C., R. M. Bartmann, L. H. Carpenter, and R. A. Garrott. 1989. Evaluation of aerial line transects for estimating mule deer densities. Journal of Wildlife Management 53:625–635.

Zabransky, C. J., D. G. Hewitt, R. W. DeYoung, S. S. Gray, C. Richardson, A. R. Litt, and C. A. DeYoung. 2016. A detection probability model for aerial surveys of mule deer. Journal of Wildlife Management 80:1379–1389.