Original Article

Calibrating Minimum Counts and Catch-per-Unit-Effort as Indices of Moose Population Trend

NICHOLAS J. DeCESARE,[1] *Montana Fish, Wildlife and Parks, Missoula, MT 59804, USA*

JESSE R. NEWBY, *Montana Fish, Wildlife and Parks, Kalispell, MT 59901, USA*

VANNA J. BOCCADORI, *Montana Fish, Wildlife and Parks, Butte, MT 59701, USA*

TONYA CHILTON-RADANDT, *Montana Fish, Wildlife and Parks, Libby, MT 59923, USA*

TIM THIER, *Montana Fish, Wildlife and Parks, Trego, MT 59934, USA*

DEAN WALTEE, *Montana Fish, Wildlife and Parks, Sheridan, MT 59749, USA*

KEVIN PODRUZNY, *Montana Fish, Wildlife and Parks, Helena, MT 59620, USA*

JUSTIN A. GUDE, *Montana Fish, Wildlife and Parks, Helena, MT 59620, USA*

ABSTRACT Monitoring wildlife population trends often involves indices assumed to correlate in proportion to abundance. We used aerial count data and harvest statistics for moose (*Alces alces*) populations in 16 hunting districts of Montana, USA, spanning 32 years (1983–2014) to assess population trends, drivers of uncertainty about those trends, and the relationship between aerial counts and hunter catch-per-unit-effort (CPUE). We found a great deal of statistical uncertainty surrounding population trends of moose measured with aerial minimum-count data, despite time series averaging >15 annual counts/district. State-space models of count-based trends suggested declining populations in 11 of 16 districts, yet 95% credible intervals overlapped 0 in all cases. The precision of count-based trends improved with increases in the number of years spanned by the time series ($\beta = -0.003$, $P < 0.001$) and average number of moose counted per survey ($\beta = -0.0006$, $P = 0.002$). Calibration of CPUE with count data showed positive correlations in only 5 of 16 (31%) districts and a catchability exponent (β) significantly <1. This indicated a generally poor level of agreement between these 2 indices, and evidence of "hyperstability," wherein declines measured by aerial counts were not reflected by proportionate declines in CPUE. Additionally, long-term trends measured with CPUE were not correlated to those in aerial counts ($P = 0.61$). We encourage explicit attention to the precision of trend estimates and local calibration of population indices to ensure both positive and proportionate relationships to underlying patterns of abundance. © 2016 The Wildlife Society.

KEY WORDS *Alces alces*, catchability, catch-per-unit-effort, CPUE, hunter harvest statistics, moose, population growth, state-space models.

Monitoring change in the abundance of wildlife populations is a core activity of wildlife management. Agencies employ a variety of techniques to monitor populations, including those that estimate abundance directly, as well as those that estimate trends in abundance indirectly using indices assumed to change proportionately with true abundance (Mason et al. 2006). Monitoring of local ungulate populations by Montana Fish, Wildlife & Parks, USA, biologists is conducted through a combination of annual aerial-survey counts, aerial age–sex composition ratios, and hunter harvest statistics from phone surveys. Comparability of aerial counts among years is sought by limiting surveys to specified areas and under specified conditions (e.g., spring green-up counts of elk [*Cervus canadensis*]), and these counts represent a minimum number of animals rather than statistical estimates of population size.

Time series of minimum count data can still allow unbiased estimates of population trend as long as both the survey effort and the mean sightability (proportion of animals in the population seen during a survey) remain constant over time (Harris 1986, Eberhardt and Simmons 1992, Humbert et al. 2009, Kéry and Schaub 2012). Sightability of animals is a complex product of 3 conditional probabilities: 1) the probability of an animal being present within the survey area, 2) the probability of an animal being available to be seen (e.g., not completely hidden by vegetation) given that it is within the survey area, and 3) the probability of a given observer actually seeing the animal given that it is available to seen (Peters et al. 2014). Several aerial-survey methodologies have

been developed to directly model portions of these imperfections in visibility for estimating abundance, but the minimum count approach instead assumes that the overall visibility varies about some constant mean. As such, imperfect sightability induces observation error into minimum count data, which in turn decreases statistical power to detect trends in abundance.

Aerial surveys often represent the "gold standard" for monitoring ungulate populations, but such methods are costly; and, in some scenarios, hunter statistics may provide a cost-effective means of monitoring population trends (Ericsson and Wallin 1999, Bontaities et al. 2000, Ueno et al. 2014). Boyce et al. (2012) simulated moose (*Alces alces*) monitoring data and resulting harvest prescriptions under different scenarios and showed that switching from aerial survey-based monitoring to monitoring solely with harvest statistics (e.g., hunter success and catch-per-unit-effort) would achieve a >600-fold reduction in costs and 30% decrease in yield, or hunter opportunity. Here, we pay specific attention to the role of catch-per-unit-effort (CPUE), a commonly collected harvest statistic, as a monitoring tool for hunted species. Despite the common management practice of interpreting CPUE and other hunter statistics as indices of game population abundances or trends, there are relatively few statistical evaluations of their reliability (Hatter 2001, Choate et al. 2006, Ueno et al. 2014). The relationship between population abundance and CPUE has received much more attention in the arena of fisheries monitoring and research, where CPUE plays a major role in global monitoring of fisheries (Harley et al. 2001, Hampton et al. 2005, Erisman et al. 2011).

Calibration between CPUE and abundance (*N*) is often characterized with the function,

$$CPUE_t = \alpha N_t^\beta \tag{1}$$

where α is the standardized coefficient of catchability and β is a shape parameter characterizing the shape of the relationship (Harley et al. 2001). It is often inherently assumed that the shape parameter, β, is equal to 1, which parameterizes a positive and linear correlation between CPUE and abundance; in other words, changes in CPUE are assumed to be proportionate to those in *N* (Fig. 1). However, allowing β to vary concedes the potential for nonlinear or threshold relationships between CPUE and abundance, which may be more common in reality. Situations where $\beta < 1$ (i.e., "hyperstability") can occur when harvest is particularly efficient or concentrated in areas of great abundance. This would result in CPUE estimates optimistically conveying stability in a population when it is in fact declining (Fig. 1). Conversely, $\beta > 1$ (i.e., "hyperdepletion") might convey pessimistic trends if a portion of the population is functionally unavailable for harvest.

For terrestrial species, there are some indications that CPUE may not be proportional to abundance across multiple taxa. Hatter (2001) found evidence of hyperstability when using CPUE to monitor declining populations of both black-tailed deer (*Odocoileus hemionus columbianus*; $\beta = 0.58$) and
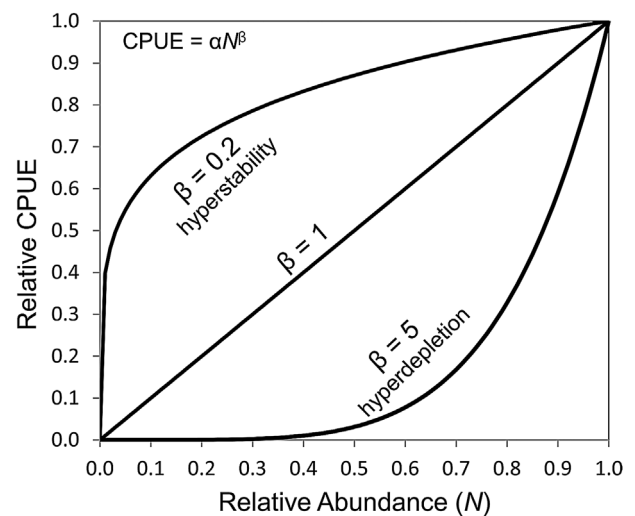

Figure 1. Theoretical relationships between relative values of catch-per-unit-effort (CPUE) and abundance as parameterized by different values of the shape parameter β, adapted from Harley et al. (2001).

moose ($\beta = 0.46$). Similarly, Ueno et al. (2014) found a less than proportional relationship (i.e., $\beta < 1$) between an index of moose seen-per-unit-effort (SPUE) and moose density. Most other studies evaluating CPUE have tested only for linear relationships, inherently assuming $\beta = 1$. For example, Lancia et al. (1996) and Fryxell et al. (1988) found evidence of linear relationships in white-tailed deer (*Odocoileus virginianus*) and moose, respectively, whereas Choate et al. (2006) found that CPUE was a poor predictor of abundance in cougars (*Puma concolor*). Lastly, Bowyer et al. (1999) found that CPUE actually increased with declines in abundance in a low-density moose population ($\beta < 0$). The shape of this relationship between CPUE and abundance for any particular wildlife population would have strong implications for managers interpreting population trends with CPUE (or SPUE) alone.

We used aerial count data and harvest statistics for moose populations in 16 hunting districts of Montana, USA, to assess: 1) the count-based population trends of moose in these districts, 2) drivers of uncertainty about those trends, and 3) the shape and strength of both short- and long-term relationships between CPUE and aerial counts. We used state-space models to estimate population trends with associated 95% confident intervals (Kéry and Schaub 2012) and assessed the relative effects of population size and the sample size of surveys upon the precision of these estimates. We then used a 2-tiered approach to test for both short- and long-term relationships between minimum counts and CPUE. First, we tested if annual estimates of CPUE were proportionally related to annual count data by estimating the shape parameter, β, which describes this relationship (Hatter 2001). Second, we compared long-term mean trends measured using both counts and CPUE to test whether trends measured with one population index were generally corroborated by the other.

STUDY AREA

We compiled moose monitoring data spanning the years of 1983–2014 for 16 moose hunting districts (HDs) in Montana (Fig. 2), including the Purcell Mountains (HD 100), Cabinet Mountains (HD 105), Salish Mountains (HDs 106 and 111), Whitefish Range (HD 110), Fleecer Mountains (HDs 319 and 341), upper Big Hole Valley (HDs 323, 326, and 327), Gravelly Range (HD 330), Ruby and Snowcrest Ranges (HD 331), Blacktail Mountains (HD 332), and Centennial Valley (HDs 333 and 334). Hunting districts averaged 1,304 km² in area (range = 333–3,057 km²), and were historically delineated with intent to include both winter and summer ranges for local moose populations. Moose harvest occurred in all 16 HDs over the study period and extended back to 1945 in some areas. The number of licenses offered varied considerably during the study period, averaging 13.5 licenses/HD-year, and ranging from 1 to 45 licenses in a given HD-year (Table 1). In general, the number of statewide licenses peaked in the mid-1990s and has since declined by >50% (DeCesare et al. 2014).

Hunt districts 100–111 were within Montana Fish, Wildlife & Parks administrative Region 1 in northwestern Montana. These HDs fell within the northern Rocky Mountain forest terrestrial ecoregion (Olson et al. 2001). This area was mountainous with extensive mesic conifer forest complexes including mixed stands of Engelmann spruce (*Picea engelmannii*), subalpine fir (*Abies lasiocarpa*), western red cedar (*Thuja plicata*), and mountain hemlock (*Tsuga mertensiana*), and lesser components of lodgepole pine (*Pinus contorta*), Douglas-fir (*Pseudotsuga menziesii*), and western larch (*Larix occidentalis*). In many areas, extensive wildfires during the 1900s–1920s and timber harvest during the 1950s–1980s affected large tracts that currently exist in various stages of succession; winter moose surveys typically show the greatest concentrations in these regenerating forest land-cover types. Hunt districts 319–341 were in Montana Fish, Wildlife & Parks's Region 3 in southwestern Montana. Each of these HDs included a combination of south-central Rocky Mountain forest and Montana valley and foothill grassland ecoregions. The area was characterized by forested mountain ranges separated by broad open valleys with a combination of sagebrush (*Artemisia* spp.) and grassland (including *Festuca* spp., *Poa* spp., and *Pseudoroegneria spicata*) communities, along with agriculture. Wet valley bottoms with extensive riparian communities made up of willows (*Salix* spp.) and sedges (*Carex* spp.) were common. Surrounding mountains were characteristic of somewhat xeric types relative to Region 1, dominated by lodgepole pine, with additional spruce–fir stands at higher elevations and Douglas-fir, western larch, and ponderosa pine (*Pinus ponderosa*) at lower elevations. Winter moose surveys in these regions typically find greatest abundances of moose in riparian willow complexes on valley floors.

METHODS

Data Collection

Aerial survey counts.—We conducted winter aerial surveys using either fixed- (HDs 319–341) or rotary-wing (HDs 100–111) aircraft, with the same platform always used within a given HD. The timing of surveys also targeted different periods according to HD, with most surveys in HDs 100–111 occurring in early winter (Dec) and surveys in HDs 319–341 occurring during mid- to late winter (Jan–Mar). We conducted surveys in standardized portions of HDs (i.e., trend areas) to maintain comparable survey effort across years. Boundaries of trend areas were comprised of both natural and anthropogenic landmarks including highways, streams, and ridgelines. Each trend area was historically selected to maximize the number of moose seen, both for generating minimum population counts and estimating sex and age classification ratios. Trend areas were estimated to represent an average of 70% of the moose winter range within a given HD (range = 20–90%) and expected to be the greatest-density winter ranges within each HD, as determined by initial scoping surveys and analyses of winter habitat selection and food habits.

In Region 3 of southwestern Montana, the fixed-wing surveys focused on riparian–willow complexes on valley floors. In this region, the winter diet of moose has been shown previously to be dominated (63–93%) by willow
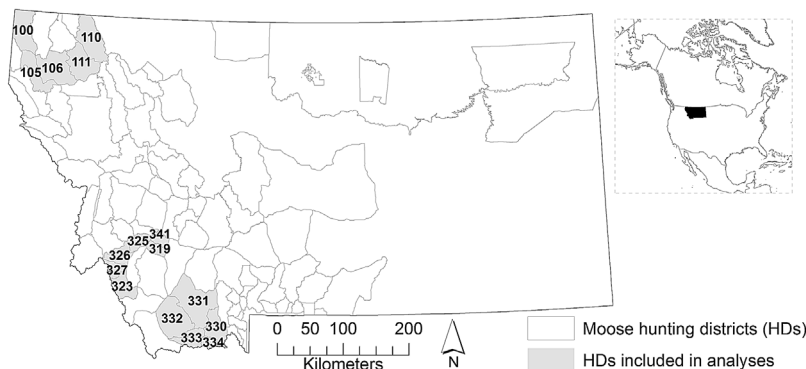

Figure 2. Moose hunting districts (HD) in Montana, USA, highlighting the 16 districts (grayed) for which aerial count data and estimates of catch-per-unit-effort were analyzed, 1983–2014.

Table 1. Descriptive statistics characterizing aerial minimum-count data for moose populations across 16 hunting districts (HDs) and intrinsic population growth rates (r), standard deviations and 95% credible intervals estimated with Bayesian state-space models, Montana, USA, 1983–2014.

HD	Time period	$\bar{x}$ no. of licenses	Range	N_{counts}	$\bar{x}$ count	Range	$\hat{r}$	SD_r	95% CI of $\hat{r}$
100	1985–2014	20	12–25	21	50	5–97	−0.050	0.070	−0.199 to 0.104
105	1985–2014	20	10–25	23	67	2–115	−0.040	0.060	−0.165 to 0.090
106	2001–2007	15	15–15	6	38	30–47	−0.047	0.117	−0.293 to 0.205
110	1994–2007	16	10–40	13	41	10–74	−0.014	0.116	−0.255 to 0.233
111	1986–2007	19	10–34	17	24	1–56	−0.035	0.105	−0.238 to 0.198
319	1994–2013	8	2–16	17	10	2–15	0.008	0.102	−0.205 to 0.226
323	1999–2014	23	11–45	13	67	25–112	0.038	0.101	−0.171 to 0.256
325	1983–2014	7	1–12	25	24	6–56	0.006	0.056	−0.099 to 0.140
326	1999–2014	14	2–22	11	31	3–54	−0.040	0.134	−0.308 to 0.241
327	1999–2014	16	3–25	12	23	6–59	−0.062	0.118	−0.311 to 0.190
330	1987–2013	9	5–10	14	18	4–47	−0.018	0.086	−0.192 to 0.174
331	1987–2013	14	3–20	13	55	20–135	−0.016	0.078	−0.180 to 0.153
332	1987–2013	7	2–13	15	33	5–92	0.058	0.072	−0.096 to 0.209
333	1987–2013	3	2–4	11	24	2–48	−0.025	0.109	−0.244 to 0.210
334	1985–2014	11	7–20	23	89	44–135	0.022	0.026	−0.031 to 0.079
341	2002–2012	5	2–11	9	11	3–20	−0.016	0.157	−0.325 to 0.318

species (Knowlton 1960, Dorn 1970) as well as local use of grasses and sedges within haystacks for forage (N. J. DeCesare, unpublished data). Winter aerial surveys in southwestern Montana involved a mixture of sinuous flight paths along riparian corridors and gridded paths in large willow patches with the goal of complete coverage within the trend areas. In Region 1 of northwestern Montana, the rotary-wing surveys primarily focused on regenerating forests within 15–30-year-old logging units and burns, which have been shown to be most selected by moose in this region (Matchett 1985). Flights in this region included the utilization of snow-tracks to locate moose and maximize the number of animals counted and classified. All flights in both regions were conducted from altitudes <150 m, from which both moose and their tracks could be most easily observed. Georeferenced survey track logs were collected with global positioning system units beginning in the early 1990s to document and maintain consistent coverage of trend areas across survey years. Surveys were conducted only when snow was present and funding and other logistics allowed; thus, flights were not conducted every year in each district. All detected moose were counted and classified; the summed count of detected moose within a given survey was treated as a minimum count, with no sightability corrections applied.

Hunting statistics.—During the study period, moose harvest statistics were estimated annually through efforts to contact all license holders following the hunting season. The moose hunting season began on 15 September and ended in late November of each year. Phone calls to moose hunters were then begun by a coordinated team of FWP staff in December and conducted through March following the hunting season. Montana FWP has long collected annual big-game harvest statistics by selecting a stratified random sample of hunters to ask predefined questions about harvest and hunting effort; this technique produces affordable and unbiased results (Lukacs et al. 2011). In the case of moose, the relatively small number of license holders (range = 362–769 annually permitted hunters during 1983–2014)

allowed for inclusion of all licensed hunters in the survey and for repeated attempts (up to 7) to contact all moose hunters statewide each year. If no contact was made with a given license-holder during the calling period, they were treated as nonrespondents. During and following the calling period, lead interviewers and a survey coordinator monitored the quality of data and corrected obvious mistakes. Following completion of the surveys, additional quality control was performed using structured query language within the database. Annual response rates to the moose harvest survey averaged >90%, and we assumed that the relatively small proportion of nonrespondents were missing at random from the sample (Lohr 2009).

We estimated harvest numbers and total hunter days (including both successful and unsuccessful hunters) for each HD from survey response data using estimators for simple random sampling within a given license type and stratified random sampling for districts with multiple license types (Thompson 2002). We corrected estimated variances and confidence intervals for large sampling fractions (i.e., high response rates) using a finite population correction factor (Thompson 2002). In cases when there was only one response in a given stratum, we took the stratum contribution to the total variance to be the average of all strata with a responding sample size of >1 individual. This adjustment results in a conservative estimate of the total variance (Lumley 2004). For the purposes of these analyses, we estimated CPUE as the total number of moose harvested divided by the total number of hunter days, combining harvest and hunter days across license types within districts with multiple types.

Data Analysis

Estimating mean growth rates from aerial-survey minimum counts.—Conventional models of population trends using log-linear regression of counts against time have been shown to consistently underestimate the variance of trend estimates and produce overly narrow confidence limits for trend estimates (Humbert et al. 2009). One explanation for this

underestimation involves a failure of serially autocorrelated time-series data to meet assumptions of independence for linear regression analysis (Ryding et al. 2007). A second explanation is that annual variation about the long-term mean growth rate is assumed to be the result solely of observation error, which discounts the potential for process variance, or biological variation in year-to-year growth rates when compared with the long-term average (Humbert et al. 2009, Hostetler and Chandler 2015). We used code developed by Kéry and Schaub (2012) to apply Bayesian state-space models for estimating growth rates and confidence intervals while accounting for both process variance and observation error contained within estimates of overall population trend. Following Kéry and Schaub (2012), we treated the state-process model for annual population growth rates, r_t, as $\log(N_{t+1}) = \log(N_t) + r_t$, with $r_t \sim \text{Normal}(\bar{r}, \sigma_r^2)$, where $\bar{r}$ was the overall mean growth rate and σ_r^2 was the process variance. The observation-process model for the annual data, y_t, was then $y_t \sim \text{Normal}\left[log\left(N_t, \sigma_y^2 \right) \right]$, where σ_y^2 was the observation variance.

To better understand the variability in precision of estimates across HDs, we then treated the standard deviation of growth-rate posterior estimates as a response variable for modeling which parameters were most responsible for improved precision. We used multiple linear regression to assess the effect of the number of annual surveys, the total duration of years spanned by surveys, the average count recorded during surveys, and the administrative region within Montana. We included administrative region to capture differences in both aerial survey platform and vegetation types among districts; in northwestern Montana (Region 1) surveys were conducted via helicopter over primarily forested habitats, whereas in southwestern Montana (Region 3) surveys were conducted via fixed-wing aircraft over primarily willow–riparian habitats.

Calibrating CPUE as an annual index of counts.—We used harvest and aerial survey data collected during 1983–2014 across 16 HDs to assess the relationship between CPUE during a given hunting season and aerial survey counts conducted within the subsequent winters of the same biological year, *i*. Monitoring of big-game populations with CPUE hinges initially upon a positive relationship between catch and effort. At the individual hunter level, increased effort is assumed to translate to increased cumulative probability of a "catch"; and at the population-level, increased cumulative hunter days are assumed to translate to increased total catch. To validate that this was the case in our own data under conditions when the maximum total catch was dictated by regulation, we conducted a preliminary analysis using a linear mixed-model to fit total harvest as a function of hunter days, with the intercept fixed at 0 and random slopes for each HD (see Supporting Information, S1). This did indeed reveal positive relationships between hunter days and harvest within all 16 HDs, with an overall conditional R^2 of 0.92. These results affirmed that catch and effort remained positively related across different levels of harvest regulation.

Typically, the relationship between CPUE and abundance is modeled by taking the natural log of both sides of Equation (1), such that,

$$\log_e (\text{CPUE}_i) = \log_e (\alpha) + \beta \left[\log_e (N_i) \right] \tag{2}$$

The intercept (exponentiated) and slope of this regression then represent estimates of the catchability coefficient (α) and shape parameter (β), respectively. This approach mirrors that of "classical calibration," where the direction of causation is such that CPUE is a dependent variable (y) to be calibrated to the independent variable (x), abundance (Krutchkoff 1967, Smith 2009). However, the underlying goal of calibration exercises is often to predict the independent variable when data are solely available for an index such as CPUE. For these reasons, "inverse calibration," where abundance would be on the left side of the equation and CPUE the right, is sometimes a more recommended approach (e.g., Centner et al. 1998), particularly because the results obtained by switching the x and y variables around are not symmetric using ordinary least-squares regression (OLS; Smith 2009). This asymmetry presents a conundrum in that one intuitively expects a single relationship between CPUE and abundance, yet will obtain different results depending on which variable is on which side of the OLS regression equation.

Fisheries researchers have addressed the issue of asymmetry between classical and inverse calibration approaches by applying reduced major axis (RMA) regression to models comparing CPUE and abundance (Erisman et al. 2011, Ward et al. 2013). Rather than the OLS approach of minimizing the vertical residuals between the fitted and observed y-values, RMA regression minimizes the product of both the dependent (x) and independent (y) deviations from the fitted line (Ward et al. 2013). This has 2 important benefits over OLS regression for our calibration of CPUE, particularly in our case of estimating using aerial minimum-count data, rather than estimates of abundance: 1) it accounts for the presence of measurement error in both the x- and y-variables, which is appropriate for these data; and 2) perhaps more importantly, it enforces a symmetric relationship between the 2 indices such that the slope of lines regressing each as the dependent variable will be reciprocal, as predictions made from either approach will be equivalent (Smith 2009). Although this approach may represent an improvement over previous applications of OLS regression to this question, RMA regressions have been shown to overestimate slopes and underestimate the widths of confidence intervals with uncorrelated variables and small sample sizes (Ricker 1984, Smith 2009). Thus, it has been recommended to limit interpretation of slopes estimated from RMA to only those data sets with significant correlations as determined by screening with the correlation coefficient (r) and its parametric P-value (Ricker 1984, Legendre 1998). To accommodate limited sample sizes of survey-years within the data for each HD, we used a

moderate criterion of $P < 0.20$ to screen correlation coefficients for entry into subsequent RMA regression analyses.

We used per-district RMA regressions to estimate α and β distinctly for each HD, following Equation (2). We used only those data for which estimates of both CPUE and aerial counts were available, and standardized and log-transformed both CPUE and count data within districts by first dividing the raw values of each index by the maximum value within the time-series of that HD alone. We exponentiated model intercepts to calculate α, and then estimated a weighted-average overall estimate of β across all districts following the methods of Murtaugh (2007), which adjusted the relative influence of individual β estimates according to their respective precision. For this weighting process, we approximated standard error estimates of βs by dividing the average span between β and 95% confidence interval limits by 1.96 (the critical z-value for estimating parametric 95% CIs).

Comparing long-term trends estimated from CPUE and counts.—Biologists working with indices of abundance often make harvest management recommendations based on trends measured over multiple years rather than on a given point estimate for a single year. In addition to calibrating CPUE and annual counts as described above, we also estimated long-term trends with paired data for each to assess if these were in agreement. Similar to sequential animal counts, time series of CPUE data include both process variation and observation error. Thus, we applied the same state-space modeling approaches as described above for count data to estimate mean rates of change per HD using raw CPUE data from each HD. We then used reduced major axis regression to compare paired mean population growth rates estimated with count and CPUE data across all 16 districts. We conducted all analyses using R 3.1.1 (R Core Team 2014), and conducted RMA regressions using the lmodel2 package (Legendre 2013; where RMA is called "standard major axis (SMA)" regression).

RESULTS

We estimated annual mean population growth-rates for 16 HDs using minimum counts from 243 annual aerial surveys, averaging 15.2 surveys/HD (range = 6–25; Fig. 3). Point estimates of mean annual growth-rates ($\hat{r}$) were negative for 11 of the 16 HDs. However, 95% credible intervals surrounding mean growth-rates overlapped 0 in all districts (Table 1; Fig. 4). The best model ($R^2 = 0.85$) explaining the differences in the magnitude of standard deviations among
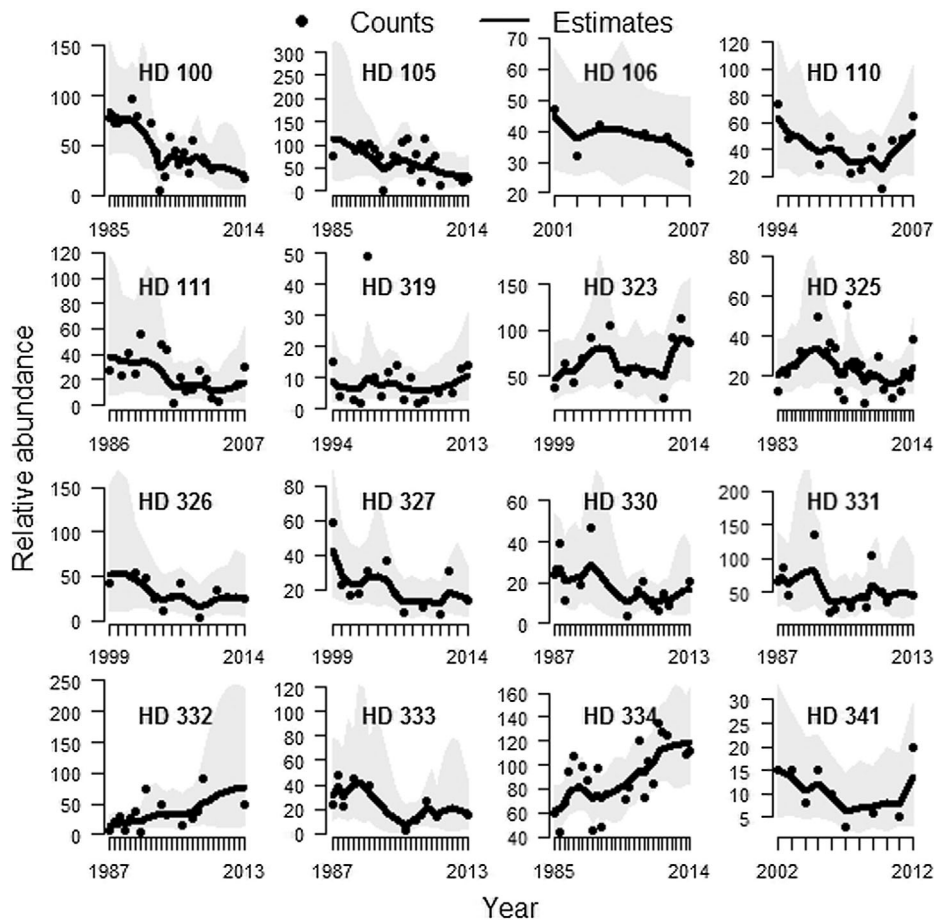

Figure 3. Time series of moose aerial counts (points), state-space model predicted counts (lines), and their respective 95% credible intervals (grayed area) in 16 hunting districts of Montana, USA, during portions or all of 1983–2014.
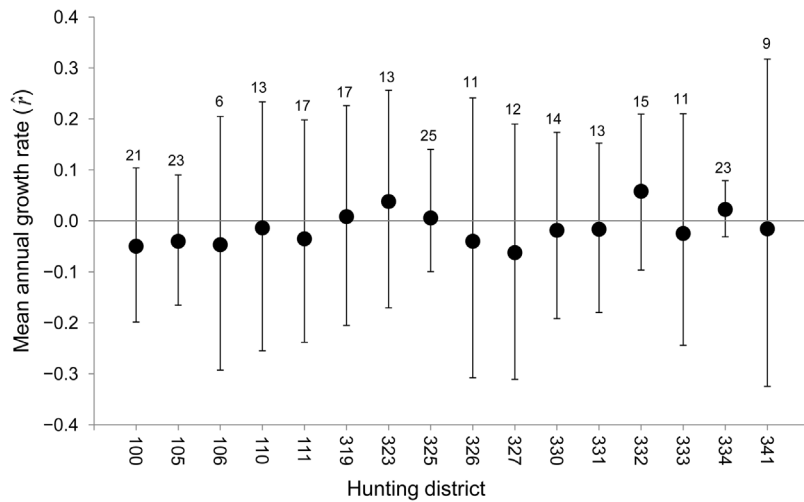
Figure 4. Intrinsic population growth rate estimates ($\hat{r}$), with 95% confidence intervals, and sample sizes of aerial counts (labels) for moose populations in 16 hunting districts of Montana, USA, during portions or all of 1983–2014.

growth-rate estimates showed improved precision with increases in the number of years spanned by the time series ($\beta = -0.003$, $P < 0.001$) and in the average number of moose counted per survey ($\beta = -0.0006$, $P = 0.002$). There was no difference among administrative regions in terms of the precision of surveys ($\beta = 0.0004$, $P = 0.96$).

We modeled the relationship between CPUE and aerial survey counts using 240 district-years of data for which both aerial survey and harvest statistic information were available. Values of the correlation coefficient (r) were positive for 14 of 16 HDs, suggesting that annual changes in CPUE were generally corroborated by changes in the same direction within aerial count data for most but not all districts (Table 2; Fig. 5). However, significance tests of r were moderately significant ($P < 0.05$) for only 2 HDs out of 16, and met our initial criteria for subsequent RMA regression analyses ($P < 0.2$) for 6 of 16 HDs (Table 2). Of these 6 HDs, credible intervals from RMA regression suggested a proportional relationship ($\beta = 1$) for 2 HDs, hyperstability in CPUE ($0 < \beta < 1$) for 3 HDs, and a negative proportional relationship ($\beta = -1$) for 1 HD (Fig. 5). A weighted mean estimate of β across these 6 HDs indicated significant hyperstability ($\beta = 0.39$, 95% CI $= 0.113$–0.669), though this was principally driven by strong evidence of hyperstability in relatively few HDs (Table 2; Fig. 5).

Lastly, to compare the long-term trends that would be detected with both population indices, we also estimated overall mean population growth-rates (r) for each HD using paired data sets of counts and CPUE for the same timespans. Regression analysis of these paired data revealed no significant relationship, as evidenced by the correlation coefficient ($r = -0.14$), its test of significance ($P = 0.61$), and the coefficient of determination ($R^2 = 0.019$; Fig. 6).

DISCUSSION

Our results reveal a great deal of statistical uncertainty surrounding population trends of moose in Montana when using aerial-based minimum-count data, despite time series

including an average of approximately 15 annual counts/ district. A recent review of the status of moose in Montana indicated concerns among management biologists over potential population declines since the 1990s (DeCesare et al. 2014). Our point estimates of r might corroborate concerns over widespread declines to some degree, but these data were generally insufficient to conclude with great confidence that populations have in fact declined. Because these data were minimum count data rather than statistical estimates of population size, estimates of trend also hinged upon an assumption of a constant mean sightability over

Table 2. Reduced major-axis regression estimates of the catchability coefficient (α) and shape parameter (β) describing the relationship between annual moose aerial-survey counts and moose hunter catch-per-unit-effort (CPUE) estimates across 16 hunting districts (HDs), as well as the correlation coefficient (r) and its parametric test of significance (P), Montana, USA, 1983–2014.

HD	α[a]	β[a]	SE(β)[a]	95% CI (β)[a]	Correlation coeff. r	P
100	−0.205	0.634	0.132	0.425–0.944	0.52	0.01
105	−0.346	0.399	0.087	0.263–0.605	0.33	0.13
106					0.002	1.00
110	−1.472	−0.779	0.226	−1.34 to −0.453	−0.51	0.07
111	−0.345	0.336	0.082	0.211–0.534	0.49	0.05
319					−0.33	0.23
323	0.036	1.005	0.291	0.585–1.726	0.52	0.07
325					0.20	0.37
326					0.18	0.60
327					0.15	0.63
330	−0.137	0.770	0.205	0.466–1.271	0.56	0.04
331					0.28	0.40
332					0.25	0.44
333					0.32	0.34
334					0.20	0.37
341					0.24	0.54

[a] Point estimates and CIs of α and β estimated using reduced major axis regression for data sets with insignificant ($P > 0.20$) correlation coeff. may be biased and are excluded.
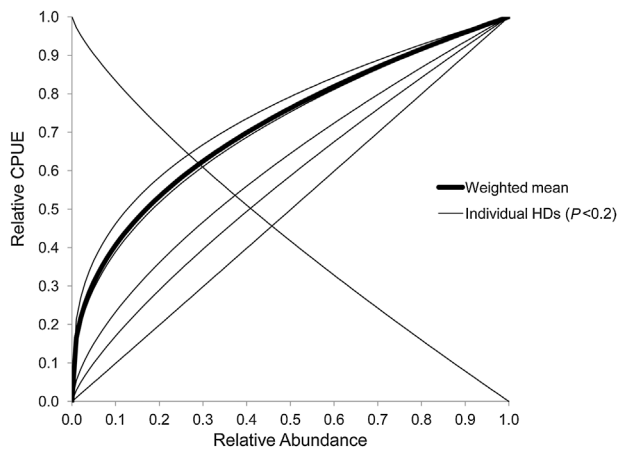
Figure 5. Empirical relationships between moose hunter catch-per-unit-effort (CPUE) and moose aerial counts estimated with reduced major axis (RMA) regression for 6 hunting districts with at least marginal correlations ($P < 0.20$), and a weighted average across all districts (bold line), Montana, USA, 1983–2014. Note an additional 10 districts that showed no relationship between CPUE and counts were withheld from this plot because RMA regressions are not recommended for data sets without significant correlation coefficients (Ricker 1984, Legendre 1998).

time. Although there were no data to suggest a change in sightability over the time period considered, it is conceivable that aging forest stands combined with reduced timber harvest (Spoelma et al. 2004), changes to riparian–willow habitats (whether degradation or restoration), or changes in rancher practices that alter the availability of alternative food sources (such as winter hay stacks) over this same time period may also be responsible for changes in counts by inducing gradual changes in mean sightability during surveys.

Much of the uncertainty about population trends likely stems from the inherent insufficiency of statistical power when surveying small populations with low sightability. The degree of sightability-induced error may be greater in small or low-density populations, making minimum counts of such populations particularly prone to large sampling variance
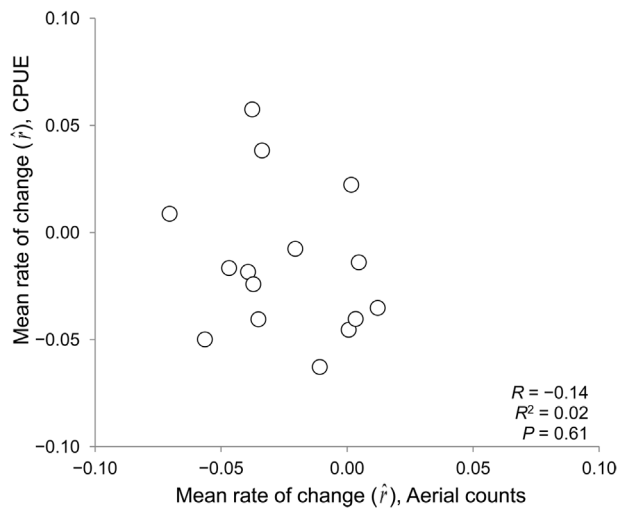

Figure 6. The relationship between long-term growth rates for moose populations in 16 hunting districts estimated using both aerial count data and hunter catch-per-unit-effort (CPUE) in Montana, USA, 1983–2014.

(MacKenzie et al. 2005). Supporting this, we found the magnitude of aerial counts to be an important driver of precision in growth estimates among HDs, in addition to the length of the time-series of the counts. In our comparison, we evaluated drivers of the posterior estimates of standard deviations about the population growth rates, which included combined effects of both observation and process error. Small numbers of animals counted (i.e., small populations) and short time-spans over which they are counted could presumably lead to increases in both observation and process error surrounding mean growth rates. These issues associated with error variance present a major challenge to those wishing to acquire reasonable precision when monitoring small populations (<100 animals counted/survey) even with relatively long (>20 yr) time series. We did not find evidence of regional differences in the precision of aerial surveys sufficient to suggest an effect of the habitat-based sightability differences among these regions. This lack of regional differences in precision may reflect confounding differences between habitats and aircraft types used during surveys across the 2 regions evaluated. It appears that using rotary aircraft in presumably low-sightability areas of Montana achieves comparable precision to the use of fixed-wing aircraft in greater sightability regions. During 1989–1993, Montana Fish, Wildlife & Parks used mark–resight methods to estimate average sightability as 0.34 and 0.53 in portions of HDs 111 and 105 (95% CIs of 0–0.77 and 0.39–0.67, respectively [J. Brown, Montana Fish, Wildlife & Parks, unpublished data]). Unfortunately, comparable studies of moose visibility are not available over other portions of the state or years of monitoring.

Our results, incorporating both process and observation error, uphold cautions that biologists should expect better estimates of trend in more abundant and visible populations and poor estimates in smaller and poorly visible populations. Harris (1986) recommended that multiple surveys per year be conducted in cases when the variability of any single count cannot be reduced enough to produce sufficiently precise trends. Simulations by Humbert et al. (2009) also suggest that the financial cost of increasing precision with multiple surveys per year might be offset by conducting such groups of surveys less frequently (e.g., every other year). Pilot data to evaluate the within-year variation of moose surveys in Montana would be valuable for future power analyses designed to assess these alternative survey strategies.

Calibrating another commonly used index of abundance, CPUE, to aerial counts revealed mixed results across HDs. These results suggest the assumption that CPUE will generally change proportionately with abundance is dubious. An overall weighted-average estimate of the shape parameter ($\beta = 0.39$) suggested hyperstability as the global average relationship, though attention to individual estimates per HD suggests some caution is needed when interpreting this result. In 10 of 16 HDs no correlation between CPUE and counts was found, and in one case a significant negative correlation existed. Thus, only in 5 of 16 (31%) HDs did CPUE correlate positively (using a very liberal cut-off of

$P < 0.20$) to aerial counts. Among those 5 HDs where a positive relationship was found, it was accompanied by evidence of hyperstability in 3 cases (60%). This hyperstability indicates areas where mean declines in abundance would not be reflected by proportionate declines in CPUE (Hatter 2001).

Although few studies have similarly assessed this relationship through direct estimation of the catchability exponent β, there is evidence to support both proportionate (Solberg and Sæther 1999, Kindberg et al. 2009, Cumberland 2012) and disproportionate relationships between trends measured via indices such as CPUE and trends in abundance (Bowyer et al. 1999, Hatter 2001, Ueno et al. 2014). Given the variety of results found both among and within (Hatter 2001) these studies, we interpret our results to generally uphold the recognized importance of site-specific, local conditions when interpreting index data such as CPUE (Lancia et al. 1996, Bowyer et al. 1999, Hatter 2001, Choate et al. 2006).

There are notable nuances to trophy big-game hunting and the regulation of such that might lead to additional complexities in CPUE data when compared with their traditional use in fisheries. First, as is the case with many big-game hunts across North America, moose licenses in Montana are limited spatially by regulation and allocated to hunters through a lottery. Thus, CPUE data are measured from a relatively small sample of hunters; in this Montana data set there were an average of 13.5 moose licenses allocated per HD-year. Additionally, the rarity of drawing a tag (e.g., 1.4% success rate/hunter in drawing 2014 license) combined with the idea of moose as a trophy species may cause hunter behavior to be less utilitarian and thus less efficient (Holsman and Petchenik 2006). Hunters in such tightly limited entry scenarios may be less likely to harvest an animal at the first opportunity, driving effort estimates upward. This highly limited-entry scenario may explain the questionable performance of CPUE as a population monitoring tool in our system relative to other systems with much larger quantities of permissible harvest (Ueno et al. 2014). Lastly, although the majority of licenses are for antlered moose, there are a subset of licenses issued for antlerless or either-sex harvest; an average of 29% of the animals harvested within these data were antlerless (adult females and calves). Differences in hunter behavior when hunting males versus females may also affect CPUE data, with potential to bias trends if the proportion of antlered versus antlerless harvest changes over time (Bhandari et al. 2006, Holsman and Petchenik 2006). Biologists monitoring populations with CPUE data may wish to restrict uses of this metric to within-sex comparisons, or consider working with other metrics such as animal sightings per unit effort (Ueno et al. 2014).

The rigor of our evaluation of CPUE as a source of monitoring data is also hampered by the relative lack of precision of aerial count data. Rather than assessing the relationship between CPUE and abundance, we assessed the relationship between one index of abundance (CPUE) and another (min. aerial counts). We used reduced major axis regression for explicitly accommodating potential for error in both data sets, and thus enforcing a symmetric relationship between each regardless of their treatment as dependent or independent variables. A lack of correlation and proportionality between CPUE and minimum aerial counts could be driven by the relationship between true abundance and either or both of these indices. However, a *post hoc* analysis assessing whether poor correlations between count and CPUE data were driven by the lack of precision in the count data themselves, showed no interactive effect ($P = 0.89$) of the relative precision of count data upon the resulting relationship between counts and CPUE (see Supporting Information, S2). Thus, we expect these results are not solely a reflection of the limitations in the count data.

Theoretical models for both moose (Boyce et al. 2012) and African lions (*Panthera leo*; Edwards et al. 2014) have shown that cost-effective and sustainable harvest management is achievable with collection of CPUE data alone. However, our test of this approach with empirical data suggests a note of caution. For small, harvested populations with limited data strings, multiple lines of evidence about population dynamics would also help make inferences supporting harvest management recommendations. Managers may also consider joint modeling of monitoring data using integrated population models, which can formally incorporate multiple types of data, such as vital rate monitoring, aerial counts, and harvest statistics (Schaub et al. 2007, Johnson et al. 2010). Incorporation of monitoring data into integrated population models would also permit an assessment of the relative value of data sources with respect to the precision of estimates needed to make management recommendations and allow a cost–benefit analysis comparing monitoring tactics (Lahoz-Monfort et al. 2014). This would allow not only for a statistically robust method but also one that may be affordable.

In general, we encourage managers to pay explicit attention to the precision of trend estimates when monitoring small or poorly visible populations with count data. It may be that repeated surveys within years (potentially at the allowable expense of conducting surveys every year; Humbert et al. 2009) are needed to provide sufficiently precise estimates of trend. Furthermore, we encourage local assessment of correlations between population indices and actual abundance. Our results showed a great degree of spatial variation in the sign, significance, and shape of correlations between CPUE and population counts, and a general lack of agreement in trends as measured by each. These results highlight a need for caution when monitoring populations with uncalibrated indices, which may not always be positively and linearly correlated to abundance as assumed.

ACKNOWLEDGMENTS

LITERATURE CITED

Bhandari, P., R. C. Stedman, A. E. Luloff, J. C. Finley, and D. R. Diefenbach. 2006. Effort versus motivation: factors affecting antlered and antlerless deer harvest success in Pennsylvania. Human Dimensions of Wildlife 11:423–436.

Bontaities, K. M., K. A. Gustafson, and R. Makin. 2000. A Gasaway-type moose survey in New Hampshire using infrared thermal imagery: preliminary results. Alces 36:69–75.

Bowyer, R. T., M. C. Nicholson, E. M. Molvar, and J. B. Faro. 1999. Moose on Kalgin Island: are density-dependent processes related to harvest? Alces 35:73–90.

Boyce, M. S., P. W. J. Baxter, and H. P. Possingham. 2012. Managing moose harvests by the seat of your pants. Theoretical Population Biology 82:340–347.

Centner, V., D. L. Massart, and S. de Jong. 1998. Inverse calibration predicts better than classical calibration. Fresenius' Journal of Analytical Chemistry 361:2–9.

Choate, D. M., M. L. Wolfe, and D. C. Stoner. 2006. Evaluation of cougar population estimators in Utah. Wildlife Society Bulletin 34:782–799.

Cumberland, R. E. 2012. Potvin double-count aerial surveys in New Brunswick: are results reliable for moose? Alces 48:67–77.

DeCesare, N. J., T. D. Smucker, R. A. Garrott, and J. A. Gude. 2014. Moose status and management in Montana. Alces 50:35–51.

Dorn, R. D. 1970. Moose and cattle food habits in southwest Montana. Journal of Wildlife Management 34:559–564.

Eberhardt, L. L., and M. A. Simmons. 1992. Assessing rates of increase from trend data. Journal of Wildlife Management 56:603–610.

Edwards, C. T., N. Bunnefeld, G. A. Balme, and E. J. Milner-Gulland. 2014. Data-poor management of African lion hunting using a relative index of abundance. Proceedings of the National Academy of Sciences 111:539–543.

Ericsson, G., and K. Wallin. 1999. Hunter observations as an index of moose Alces alces population parameters. Wildlife Biology 5:177–185.

Erisman, B. E., L. G. Allen, J. T. Claisse, D. J. Pondella, E. F. Miller, J. H. Murray, and C. Walters. 2011. The illusion of plenty: hyperstability masks collapses in two recreational fisheries that target fish spawning aggregations. Canadian Journal of Fisheries and Aquatic Sciences 68:1705–1716.

Fryxell, J. M., W. E. Mercer, and R. B. Gellately. 1988. Population dynamics of Newfoundland moose using cohort analysis. Journal of Wildlife Management 52:14–21.

Hampton, J., J. R. Sibert, P. Kleiber, M. N. Maunder, and S. J. Harley. 2005. Fisheries: decline of Pacific tuna populations exaggerated? Nature 434: E1–E2.

Harley, S. J., R. A. Myers, and A. Dunn. 2001. Is catch-per-unit-effort proportional to abundance? Canadian Journal of Fisheries and Aquatic Sciences 58:1760–1772.

Harris, R. B. 1986. Reliability of trend lines obtained from variable counts. Journal of Wildlife Management 50:165–171.

Hatter, I. W. 2001. An assessment of catch per unit effort to estimate rate of change in deer and moose populations. Alces 37:71–77.

Holsman, R. H., and J. Petchenik. 2006. Predicting deer hunter harvest behavior in Wisconsin's chronic wasting disease eradication zone. Human Dimensions of Wildlife 11:177–189.

Hostetler, J. A., and R. B. Chandler. 2015. Improved state-space models for inference about spatial and temporal variation in abundance from count data. Ecology 96:1713–1723.

Humbert, J.-Y., L. Scott Mills, J. S. Horne, and B. Dennis. 2009. A better way to estimate population trends. Oikos 118:1940–1946.

Johnson, H. E., L. Scott Mills, J. D. Wehausen, and T. R. Stephenson. 2010. Combining ground count, telemetry, and mark–resight data to infer population dynamics in an endangered species. Journal of Applied Ecology 47:1083–1093.

Kéry, M., and M. Schaub. 2012. Bayesian population analysis using WinBUGS: a hierarchical perspective. Elsevier/Academic Press, Waltham, Massachusetts, USA.

Kindberg, J., G. Ericsson, and J. E. Swenson. 2009. Monitoring rare or elusive large mammals using effort-corrected voluntary observers. Biological Conservation 142:159–165.

Knowlton, F. F. 1960. Food habits, movements and populations of moose in the Gravelly Mountains, Montana. Journal of Wildlife Management 24:162–170.

Krutchkoff, R. G. 1967. Classical and inverse regression methods of calibration. Technometrics 9:425–439.

Lahoz-Monfort, J. J., M. P. Harris, B. J. T. Morgan, S. N. Freeman, and S. Wanless. 2014. Exploring the consequences of reducing survey effort for detecting individual and temporal variability in survival. Journal of Applied Ecology 51:534–543.

Lancia, R. A., J. W. Bishir, M. C. Conner, and C. S. Rosenberry. 1996. Use of catch-effort to estimate population size. Wildlife Society Bulletin 24:731–737.

Legendre, P. 1998. Model II regression user's guide, R edition. R Vignette. https://cran.r-project.org/web/packages/lmodel2/vignettes/mod2user.pdf. Accessed 2 Jul 2015.

Legendre, P. 2013. lmodel2: Model II regression. R package. https://cran.r-project.org/web/packages/lmodel2/index.html. Accessed 2 Jul 2015.

Lohr, S. L. 2009. Sampling: design and analysis. Second edition. Brooks/Cole, Boston, Massachusetts, USA.

Lukacs, P. M., J. A. Gude, R. E. Russell, and B. B. Ackerman. 2011. Evaluating cost-efficiency and accuracy of hunter harvest survey designs. Wildlife Society Bulletin 35:430–437.

Lumley, T. 2004. Analysis of complex survey samples. Journal of Statistical Software 9:1–19.

MacKenzie, D. I., J. D. Nichols, N. Sutton, K. Kawanishi, and L. L. Bailey. 2005. Improving inferences in population studies of rare species that are detected imperfectly. Ecology 86:1101–1113.

Mason, R., L. H. Carpenter, M. Cox, J. C. DeVos, J. Fairchild, D. J. Freddy, J. R. Heffelfinger, R. H. Kahn, S. M. McCorquodale, D. F. Pac, D. Summers, G. C. White, and B. K. Williams. 2006. A case for standardized ungulate surveys and data management in the western United States. Wildlife Society Bulletin 34:1238–1242.

Matchett, M. R. 1985. Habitat selection by moose in the Yaak River drainage, northwestern Montana. Alces 21:161–190.

Murtaugh, P. A. 2007. Simplicity and complexity in ecological data analysis. Ecology 88:56–62.

Olson, D. M., E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. Powell, E. C. Underwood, J. A. D'amico, I. Itoua, H. E. Strand, J. C. Morrison, C. L. Loucks, T. F. Allnutt, T. H. Ricketts, Y. Kura, J. F. Lamoreux, W. W. Wettengel, P. Hedao, and K. R. Kassem. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. BioScience 51:933–938.

Peters, W., M. Hebblewhite, K. G. Smith, S. M. Webb, N. Webb, M. Russell, C. Stambaugh, and R. B. Anderson. 2014. Contrasting aerial moose population estimation methods and evaluating sightability in west-central Alberta, Canada. Wildlife Society Bulletin 38:639–649.

R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. Accessed 17 Sep 2014.

Ricker, W. E. 1984. Computation and uses of central trend lines. Canadian Journal of Zoology 62:1897–1905.

Ryding, K. E., J. J. Millspaugh, and J. R. Skalski. 2007. Using time series to estimate rates of population change from abundance data. Journal of Wildlife Management 71:202–207.

Schaub, M., O. Gimenez, A. Sierro, and R. Arlettaz. 2007. Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. Conservation Biology 21:945–955.

Smith, R. J. 2009. Use and misuse of the reduced major axis for line-fitting. American Journal of Physical Anthropology 140:476–486.

Solberg, E. J., and B. E. Sæther. 1999. Hunter observations of moose Alces alces as a management tool. Wildlife Biology 5:107–117.

Spoelma, T. P., T. A. Morgan, T. Dillon, A. L. Chase, C. E. Keegan, and L. T. DeBlander. 2004. Montana's forest products industry and timber harvest, 2004. U.S. Department of Agriculture Forest Service Resource Bulletin, Rocky Mountain Research Station, Fort Collins, Colorado, USA.

Thompson, S. K. 2002. Sampling. Second edition. Wiley, New York, New York, USA.

Ueno, M., E. J. Solberg, H. Iijima, C. M. Rolandsen, and L. E. Gangsei. 2014. Performance of hunting statistics as spatiotemporal density indices of moose (*Alces alces*) in Norway. Ecosphere 5:art13.

Ward, H. G., P. J. Askey, J. R. Post, and K. Rose. 2013. A mechanistic understanding of hyperstability in catch per unit effort and density-dependent catchability in a multistock recreational fishery. Canadian Journal of Fisheries and Aquatic Sciences 70:1542–1550.

Associate Editor: Thogmartin.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site. We include additional analyses upholding the positive relationship between catch and effort underlying these data (S1) and testing whether there was any indication that the precision of trends measured from aerial counts was related to the subsequent correlation of those counts to CPUE (S2).