Wildlife Society Bulletin 36(2):264–276; 2012; DOI: 10.1002/wsb.139

Original Article

Evaluating Population Estimates of Mountain Goats Based on Citizen Science

JAMI J. BELT,[1,2] *Boone and Crockett Wildlife Conservation Program, University of Montana, Missoula, MT 59812, USA*
PAUL R. KRAUSMAN, *Boone and Crockett Wildlife Conservation Program, University of Montana, Missoula, MT 59812, USA*

ABSTRACT Citizen science programs that use trained volunteers may be a cost-effective method for monitoring wildlife at large scales. However, few studies have compared data collected by volunteers versus biologists. In Glacier National Park (GNP), Montana, USA, we assessed whether citizen science is a useful method to monitor mountain goat (*Oreamnos americanus*) populations. We compared estimates of mountain goat abundance by volunteers at 32 sites throughout GNP with estimates by biologists and aerial surveys at a subset of 25 and 11 sites, respectively. We used multiple-observer surveys to calibrate the indices of abundance at each site for observer variation between volunteers and biologists. We used N-mixture models to obtain estimates of abundance across all sites. Population estimates by citizen scientists overlapped estimates by biologists. Density estimates from aerial surveys were lower than ground estimates. Mean detection probability from multiple-observer surveys for biologists was significantly higher and less variable than that of volunteers. More frequent site visits balanced out lower detection probability by volunteers and resulted in abundance estimates that were less variable than those of biologists. When large spatial and temporal coverage can be achieved, citizen science can provide mountain goat population estimates that are statistically similar to those of biologists.

KEY WORDS citizen science, Glacier National Park, N-mixture models, *Oreamnos americanus*, population monitoring, volunteers.

Research and monitoring that relies on volunteers to collect data without direct supervision is called citizen science (Trumbull et al. 2000). The use of citizen science for long-term ecological data collection is increasing (Newman et al. 2003, Danielsen et al. 2005, Greenwood 2007, Cohn 2008). As ecosystem-level disturbances (e.g., global climate change; Morisette et al. 2009) proliferate, as funding for ecological monitoring declines (Pilz et al. 2005), and as demand for public participation in resource management increases (Yung 2007), many organizations are developing citizen science programs to address the need for cost-effective monitoring that covers large geographic areas. Many grant-providing organizations (e.g., National Science Foundation) now require grant holders to incorporate public participation in research and monitoring (Silvertown 2009).

Careful training and sampling design may allow citizen science programs to yield results that are as reliable as those from professional monitoring programs (Hochachka et al. 2000, Yoccoz et al. 2003, Gouveia et al. 2004). Citizen science data are often collected on spatial scales beyond the reach of most research budgets (Cooper et al. 2007, Greenwood 2007, Cohn 2008), and they may also be conducted over a longer term than professional monitoring programs with larger funding needs (Danielsen et al. 2005). For example, the National Audubon Society's Christmas Bird Count, a citizen science effort that began in 1900, has yielded the longest unbroken record of bird diversity and distribution data (Root and Alpert 1994) with the broadest temporal and geographic coverage of North America's avian fauna (Dunn et al. 2005).

However, the question of whether citizen science data can yield useful scientific inferences remains unanswered. Many of the assumptions underpinning statistically rigorous ecological surveys (e.g., skill levels of observers, consistent survey effort, and homogeneity of temporal variables) are violated with citizen science (Danielsen et al. 2005). The impact of these departures from traditional survey design on the validity of the inferences that can be drawn from the data needs to be assessed.

A major obstacle to understanding the effectiveness of citizen science is the paucity of studies that have compared data and results from volunteers with professionally collected data (Fitzpatrick et al. 2009). A few studies have reported that volunteers with varied skill level often underestimate species abundance, but their estimates may be consistent (Newman et al. 2003, Delaney et al. 2008) and have a linear relationship to density estimates obtained from other sources (Kindberg et al. 2009). Other studies have found that volunteers tend to reliably detect species presence, but that

Received: 18 November 2010; Accepted: 3 February 2012;
Published: 11 May 2012

Additional supporting information may be found in the online version of this article.

[1]*E-mail: jami_belt@nps.gov*
[2]*Present address: Glacier National Park, P.O. Box 128, West Glacier, MT 59936, USA.*

they often underestimate species richness (e.g., birds [Sauer et al. 1994], amphibians [Genet and Sargent 2003], and invertebrates [Delaney et al. 2008, Kremen et al. 2010]). Some studies have found experience to be a significant factor (Genet and Sargent 2003, Fitzpatrick et al. 2009), while others found it had little to no effect on population estimates (Hochachka et al. 2000, Newman et al. 2003). Two studies that used double-sampling by biologists to assess population estimates based on data collected by volunteers have reported that the large sample sizes (i.e., >1,000 samples) outweighed the disadvantage of variation among observers (Hochachka et al. 2000, Kindberg et al. 2009).

In 2008, managers at Glacier National Park (GNP), Montana, USA, established a citizen science program to conduct baseline monitoring for mountain goats (*Oreamnos americanus*). Declines in goat numbers at a prominent mineral lick in GNP (S. Gniadek, Glacier National Park, unpublished data) and throughout the neighboring Bob Marshall Wilderness (Koeth 2008), and uncertainty about mountain goat response to climate change (Pettorelli et al. 2007), generated concern about the stability of mountain goat populations in GNP. Broad-scale abundance estimates of mountain goats in GNP are needed for future monitoring. Due to the sensitivity of mountain goats to trapping (Côté et al. 1998) and the costs, safety concerns, and potential impacts to visitor experience associated with traditional mark–recapture methods, the National Park Service suggested citizen science as a noninvasive monitoring approach.

Aerial survey by helicopter is the primary noninvasive technique used to census mountain goats (Shackleton 1997), but the high cost of this method means that coverage or replication is often sacrificed. A combination of aerial surveys and observational ground counts may improve precision of population estimates (Festa-Bianchet and Côté 2008). Mountain goats are an ideal candidate species for ground counts due to the high visibility of their exposed habitats (Veitch et al. 2002). Ground counts are rarely conducted, however, because they are not typically cost-efficient due to the rugged and remote places inhabited by mountain goats (Festa-Bianchet and Côté 2008). Studies using the number of ungulates seen per day by hunters and outfitters as an estimate of ungulate density (Ericsson and Wallin 1999, Veitch et al. 2002, Pettorelli et al. 2007) have reported promise as viable long-term monitoring techniques.

The large number of visitors to GNP (approx. 2 million/yr) similarly provides a potentially useful resource for monitoring mountain goat populations. In spring 2008, GNP created the High Country Citizen Science (HCCS) program to train volunteers to conduct observational surveys of mountain goats. These surveys take place on a park-wide scale and focus on data collection at backcountry locations. The goals of the program are to estimate the distribution and abundance of mountain goats in GNP and to establish protocols for long-term trend monitoring.

Our objective was to evaluate whether citizen science is a viable method for long-term population monitoring of mountain goats. We compared mountain goat population estimates from ground counts by volunteers with estimates from ground counts by biologists at a smaller sample of sites and to the earlier density estimates of Chadwick (1977) and Singer and Doherty (1985). We also compared estimates from volunteers with aerial survey counts (by helicopter) at a smaller sample of sites, because this is the most commonly used mountain goat survey technique and represents the "gold standard." Because we did not have knowledge of the true number of mountain goats, we could not measure the accuracy of each method. Our aim, rather, was to determine whether estimates based on citizen science would yield statistically similar population estimates to more traditional approaches.

We made comparisons at the level of individual surveys and across all surveys to answer the following questions:

1. Do volunteers have similar detection probabilities to biologists and aerial surveyors? To answer this question, we used multiple-observer methods to analyze the differences between estimates of relative abundance attained during individual surveys by volunteers, biologists, and aerial surveys. We predicted volunteers would detect fewer mountain goats than would biologists and, due to differences in skill level among volunteers, would have more variable detection rates. As a result, abundance estimates of mountain goats by volunteers from individual surveys would be slightly more variable and biased low compared with counts by biologists. If detection probability could be adequately measured and was consistent (i.e., relatively low variability), average detection probability could be useful as a correction factor for counts from volunteers.
2. Do the benefits of large spatial and temporal sample sizes attainable using citizen science balance the limitations that result from varied skill levels and heterogeneous survey effort? To answer this question, we analyzed differences between volunteers, biologists, and aerial surveyors in estimates of average abundance across all sites based on N-mixture models (MacKenzie et al. 2002). Due to budget constraints, biologists and aerial surveyors were only able to count mountain goats at approximately one-third of all sites. On an individual survey basis, professionals (i.e., biologists and aerial surveyors) may attain higher detection probability and less variable estimates, whereas a team of volunteers can often survey more sites and visit sites more often. Our research aimed to provide a quantitative basis for managers (with a given amount of funding for mountain goat monitoring) to evaluate the trade-off between hiring a small team of skilled biologists versus coordinating a larger team of volunteers.

We predicted that, despite higher variation in detection probability during individual surveys by volunteers, the larger sample size would stabilize estimates across all surveys. Additionally, movement of mountain goats within their home ranges may cause some animals to be out of view and unavailable to be counted during each survey. Therefore, volunteers may be more likely than biologists to capture a higher minimum count of mountain goats as the number of site visits increases. Overall, we expected that abundance estimates across all survey sites would be similar to estimates from biologists due to larger spatial and temporal coverage.

STUDY AREA

Glacier NP contained 4,081 km^2 of federally protected land and was situated in the northern Rocky Mountains, Montana, USA. For a description of the study area, see Chadwick (1977) and Hop et al. (2007). Glacier NP provided habitat to an unhunted native population of mountain goats of undetermined size. Earlier studies yielded density estimates of 1.16 mountain goats/km^2 in a 310-km^2 central portion of GNP (Chadwick 1977) and 2.9 mountain goats/km^2 in a 32.7-km^2 area (Singer and Doherty 1985). Glacier NP maintained an extensive network of >1,127 km of hiking trails. Our study area included all portions of the Livingston and Lewis mountain ranges that had slope angles $\geq 25^\circ$ and were within 3.2 km of a trail (Fig. 1). Surveys were conducted at 32 sites with an area of 1,311 km^2 (32.1% of GNP).

METHODS

Site Selection

Proximity to escape terrain (i.e., precipitous terrain used to evade predators, with cliffs, rocky ledges, and slope angles $\geq 25^\circ$ (Chadwick 1976, Varley 1994, Gross et al. 2002)) is a strong determinant of habitat use by mountain goats (Brandborg 1955, McFetridge 1977, Haynes 1992, Hamel and Côté 2007). A model based solely on distance to escape terrain with slope angles $\geq 33^\circ$ correctly predicted occurrence in 87% of mountain goat observations from a study in Colorado, USA (Gross et al. 2002). Therefore, we identified escape terrain from a 30-m digital-elevation model using slope-angle classifications of 25–32°, 33–39°, and 40–90° (Gross et al. 2002). The area of GNP with slope angles $\geq 25^\circ$ was 1,653 km^2 (40.5% of GNP). We then created 8-km $\times$ 11-km grid cells over the escape terrain to systematically locate observation sites. The grid-cell size was large enough to encompass the maximum home-range size of a mountain goat (Rideout 1977), to minimize the likelihood that mountain goats would move from one grid cell to another.

Most (76.6%) of the escape terrain used by mountain goats in GNP exceeded 40° (Chadwick 1977). Therefore, we identified all grid cells with $\geq 40^\circ$ escape terrain ≤ 3.2 km from a hiking trail for observation sites. The distance of 3.2 km is the maximum line-of-sight distance for reliable detection and identification of mountain goats (S. Gniadek, personal communication; D. Chadwick, National Geographic, personal communication), which we verified through field tests. We then divided each trail into 2-km segments, randomly selected 1 segment/cell, hiked each segment, and recorded Universal Transverse Mercator coordinates of all points from which slope features

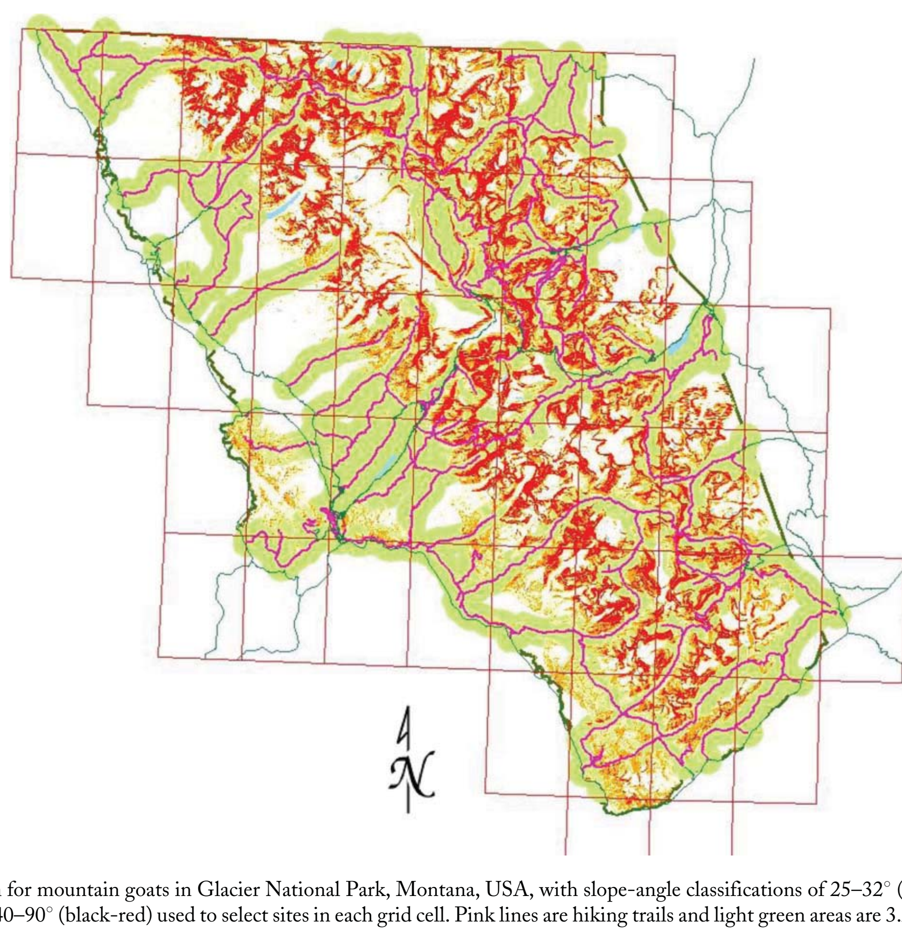

Figure 1. Escape terrain for mountain goats in Glacier National Park, Montana, USA, with slope-angle classifications of 25–32° (light gray-yellow), 33–39° (dark gray-orange), and 40–90° (black-red) used to select sites in each grid cell. Pink lines are hiking trails and light green areas are 3.2-km buffers around trails.

were visible. We then selected in each grid cell, from all points, the observation point (i.e., site) from which the largest area of escape terrain $\geq 25^\circ$ was visible.

We selected 32 sites, but topographic features blocked portions of each site. We calculated the portion of each site that was visible from each observation point using the Spatial Analyst Viewshed function in ArcGIS. We used the Spatial Analyst Extract by Mask function in ArcGIS to calculate the area of escape terrain in the viewshed of each site.

Data Collection

We recruited volunteers for the HCCS program using press releases, newspaper articles, public presentations, and flyers. In 2008 and 2009, selected volunteers attended a standardized 6-hour training session where they learned ecology of mountain goats, field identification, and classification of age and sex following criteria developed by Smith (1988). Volunteers then worked in the field with HCCS staff to learn survey protocols and the use of survey equipment (e.g., binoculars, spotting scopes, Global Positioning System [GPS], and compasses).

All volunteers in 2009 completed a participant information sheet (Supplementary Materials, Appendix 1) detailing their experience with spotting scopes and viewing wildlife. We used this information as a self-evaluation measure to assess experience level (Martin 1997, Scott et al. 2005). We also assessed the experience level of biologists who conducted mountain goat surveys in GNP using the same questionnaire (scores could range from 9 to 42). We pooled separate scores for volunteers and biologists, calculated the quartiles for the pooled data, and ranked each participant as follows: novice (min. to first quartile, 9–20.75 points), some relevant experience (first quartile to median, 20.76–26 points), moderate (median to third quartile, 26.01–31 points), and skilled (third quartile and above, 31.01–42 points).

Once trained, volunteers conducted surveys at selected sites based on their schedule, hiking ability, and preference. Volunteers navigated to each site using a GPS unit and site map. Photos of the observation point and the views due north and due south were provided to ensure that volunteers could locate the correct site despite GPS error (± 10 m). Volunteers conducted a 1-hour survey, located groups of goats, and recorded the number, age, and sex of mountain goats detected, time of initial detection, and group size, and the power and field of view of their optical equipment. Volunteers were also asked to take photos of each group of mountain goats with a digital camera through a spotting scope, and to submit photos for verification purposes.

Volunteers also recorded temperature, cloud cover, weather, time of day, and behavior that may affect detection probability. Volunteers recorded behavior of individual mountain goats upon detection (e.g., bedded, standing, and foraging, walking), which we converted into the percentage of mountain goats that were moving upon detection. Beginning in 2009, volunteers also documented visibility (as a proxy for distance estimation) and habitat use. We used a 2-step process to estimate visibility. Volunteers first recorded how they detected each group of mountain goats (visible with the naked eye, visible with binoculars, or visible only with spotting scope). We then scaled this visibility information into a single value for each survey by weighting the percentage of mountain goats seen in each category (naked eye by 1, binoculars by 2, and spotting scope by 3), and then summed the total. Habitat use was estimated by recording landscape features where mountain goats were detected. Landscape features recorded were those that may influence the distribution of mountain goats in GNP (Chadwick 1976) and included permanent snow or ice-fields, ledges, talus–scree–moraine, meadows, shrubs–krummholz, forests, roads, and trails. We identified the dominant landscape feature in which the majority of mountain goats were seen from these data. We chose the covariates listed in the previous 2 paragraphs because they were the factors most likely to influence detection probability that could also easily be recorded by volunteers.

Surveys were conducted between the second week of June and the last week of October 2008 and 2009, after parturition and before the rut. During this time, mountain goats are more likely to remain within their home ranges (Festa-Bianchet and Côté 2008) and we could assume that the number of mountain goats present within each home range would not change due to birth, immigration, or emigration. Our use of sites that were larger than maximum home-range size estimates for mountain goats also increased the likelihood of maintaining population closure during sampling periods (MacKenzie et al. 2003). If population closure assumptions were violated by movement of mountain goats between sites, it could result in inaccurate estimates. However, closure violations would similarly affect estimates by volunteers and biologists and would therefore not inhibit relative comparisons to determine whether estimates were statistically similar.

The goal was for volunteers to conduct ≥ 3 surveys at each site. We sent periodic e-mails to volunteers to inform them of sites that needed to be surveyed. Due to the voluntary nature of the program we could not assign survey locations. Because volunteers chose their own survey locations and schedule, and individual volunteers rarely surveyed the same site ≥ 1 time, potential sources of heterogeneity from observer and time of day effects were minimized (MacKenzie and Royle 2005).

Biologists in GNP who had >1 year experience in monitoring mountain goats conducted observational surveys at a subset of sites to compare with data from volunteers. Biologists conducted ≥ 3 surveys, following the same protocols as volunteers, at each of 14 sites that were randomly selected from all sites accessible within 1 day of travel (1-way distance). The order of site visits was rotated to avoid the introduction of systematic variation (MacKenzie and Royle 2005). Although mountain goats are most active in the morning (0700–1000 hours) and late afternoon (1500–2000 hours; Rideout 1977, Singer and Doherty 1985), biologists conducted surveys during times of day that volunteers most commonly conducted surveys. We also asked biologists to conduct surveys at additional sites whenever possible.

Estimating Detection Probability and Abundance at Sites Using Multiple-Observer Methods

Mean detection probability ($\hat{P}_i$) is rarely constant at all sites and times and not all covariates can be measured, so direct estimation of detection probability is an important part of monitoring (Alldredge et al. 2006). Experience-level differences among observers can also influence detection probability, which biased abundance estimates (Nichols et al. 2000, Genet and Sargent 2003). Multiple-observer methods account for detection bias due to animals that are present but missed due to observer error (Nichols et al. 2000). In this case, $\hat{P}_i$ refers to the probability that mountain goats were detected at a site given they were present, rather than the probability that a mountain goat was present, or available to be counted, at the site.

We conducted independent multiple-observer surveys (Nichols et al. 2000) to directly estimate differences in $\hat{P}_i$ between biologists ($\hat{P}_{\text{biol}}$) and volunteers ($\hat{P}_{\text{vol}}$). Incorporating detection probability measured by multiple-observer approaches has been proposed as a correction factor for data with high inter-observer variation (Nichols et al. 2000, Alldredge et al. 2006). Such a correction factor could be useful in calibrating data from citizen science programs to make it comparable to data from biologists. Multiple-observer approaches also enable use of mark–recapture methodology to move point counts from indices to estimates of abundance (Nichols et al. 2000, Johnson 2008).

One biologist conducted 76 multiple-observer surveys simultaneously with volunteers. Constant use of spotting scopes and binoculars during the 1-hour survey inhibited the ability of observers to cue off detections of others and ensured that observers maintained independence during multiple-observer surveys (Nichols et al. 2000). We selected a sample of volunteers from each experience rank (i.e., novice, some relevant experience, moderate, and skilled) to represent a proportional sample of the number of volunteers in the whole program within that experience rank to ensure that $\hat{P}_i$ was measured for volunteers with all levels of experience. Each biologist also conducted $\geq$2 multiple-observer surveys on 2 separate occasions with each other ($n = 53$) to measure differences in $\hat{P}_i$ among biologists. We used the Lincoln–Petersen estimator to obtain an estimate of abundance for each survey, and then divided the observer's count by that estimate to determine $\hat{P}_i$ for each observer (Nichols et al. 2000).

We used Welch's t-approximation (Welch 1947), which assumes unequal variances, to test for differences in detection probabilities between volunteers and biologists ($n = 50$), and between biologists ($n = 56$). We divided high counts (i.e., the largest observed count) at each site from volunteers and from biologists by mean $\hat{P}_i$ across all sites for each group ($\hat{P}_{\text{biol}}$ for biologists and $\hat{P}_{\text{vol}}$ for volunteers) to get corrected counts. Although the use of high counts rather than mean counts may introduce a bias, we chose to do this because high counts are commonly used for aerial surveys of mountain goats and other ungulates (Shackleton 1997).

Density for each site was estimated by dividing the high count of mountain goats by the area of escape terrain at each site. To correct for detection bias, we divided density estimates by $\hat{P}_i$ for each observer class (i.e., volunteer or biologist). We then divided density estimates by the average density to obtain relative density estimates for each observer class (i.e., volunteer, biologist, and aerial observers). Relative densities provided an understanding of whether distribution estimates were similar despite differences in actual densities. We determined quartiles of relative density estimates for each observer class and assigned each site a density rank of no mountain goats, low (first quartile), moderate (second quartile), high (third quartile), and very high (fourth quartile). We considered relative density estimates between observer classes to be in agreement if they were within 1 quartile of one-another.

Estimating Average Abundance Across All Sites Using N-Mixture Models

Because all or nearly all members of a mountain goat group occasionally travel together (Chadwick 1977, Singer and Doherty 1985), these estimates represent the highest density of mountain goats observed within the viewshed of each site. It is unlikely, however, that all mountain goats occupying a site will be available for detection in the viewshed simultaneously. Density estimates based on high counts in the viewshed account for detection bias, but do not account for availability bias. Availability bias occurs when animals are not available to be detected due, in our study, to not being present in the viewshed during surveys (Nichols et al. 2000). Therefore, these estimates were likely biased low and do not accurately reflect the number of mountain goats inhabiting the site (including the area surrounding the viewshed) that were not detected.

To enable estimation of abundance beyond the viewshed at each site, we needed estimates that accounted for availability bias by incorporating the probability that a mountain goat was present at the site but not present in the viewshed during surveys. Patterns of detection and nondetection during spatially replicated counts can be used to adjust for biases in counts that are caused by false absences (MacKenzie and Kendall 2002, MacKenzie et al. 2002). N-mixture models account for false absences due to detection bias and availability bias (MacKenzie and Kendall 2002, MacKenzie et al. 2002). We used N-mixture models to derive 2 estimates of average abundance (λ) and occupancy (ψ) across all survey areas (Royle 2004), one from volunteer data and the second from biologist data. N-mixture models assume that site-specific abundance influences detection (or nondetection) of animals at a site (P), that distribution of animals across survey sites is random, and that it can be described by a Poisson distribution (Royle and Nichols 2003). An estimate of λ is derived by integrating the binomial probabilities of detecting a certain count of animals at each site over the possible values of abundance for that site (Royle 2004).

We developed a series of models of covariates with the potential to influence estimates of λ and P but used only 2009 data; the number of visits to each site in 2008 by biologists was too low (≤ 2) to obtain adequate estimates of P. Covariates tested in relation to P included observer experience, size of largest group of mountain goats detected, total number of mountain goats, temperature, binocular power, binocular field of view, scope power, scope field of view, start time of survey, wind speed, weather, the percentage of mountain goats that were moving when detected, visibility of mountain goats, and the dominant landscape feature. If covariates that could be controlled to some degree were included in models that explained a large portion of variation in detection probability, changes to survey design could potentially aim to remove these effects. Covariates that could not be controlled could be accounted for to adjust detection probabilities.

Covariates with potential to influence estimates of λ included viewshed area (km^2), area of escape terrain within viewshed (km^2), area of escape terrain at site (km^2), and number of site visits. N-mixture estimates of λ use observed counts to determine a minimum abundance (Royle 2004). Because mountain goats needed to be within viewsheds to be counted, a larger viewshed size or a larger amount of escape terrain within viewsheds could lead to higher observed counts. Similarly, the observed count may be higher at sites that were visited more frequently due to the increased likelihood of observing mountain goats within each viewshed as they moved around within the surrounding site. By modeling the potential effects of these covariates on estimates of λ, we could determine whether variation was based more on habitat availability (i.e., area of escape terrain) as opposed to factors that affected availability of mountain goats to be counted.

We then used Akaike's Information Criterion (AIC; Burnham and Anderson 2002) to select the top model and evaluated the goodness-of-fit of our fitted model using parametric bootstrapping. We estimated average density of mountain goats in GNP using the following equation:

$$\overline{\text{density}}_{\text{GNP}} = \frac{\lambda \times \text{no. of sites surveyed}}{\text{area of escape terrain in all sites}} \tag{1}$$

To extrapolate abundance to all areas with escape terrain $\geq 25°$, we used the following equation:

$$\hat{N}_{\text{GNP}} = \overline{\text{density}}_{\text{GNP}} \times \text{area of escape terrain in GNP} \tag{2}$$

We conducted N-mixture analysis, model development, and model selection using the "Unmarked" package (http://r-forge.r-project.org/projects/unmarked) in R.

To evaluate the viability of citizen science for detecting long-term trends in mountain goat populations using N-mixture model estimates, we also conducted a power analysis using methods proposed by Field et al. (2005) (Supplementary Materials, Appendix 2).

Aerial Surveys

We conducted aerial surveys of 11 sites, including area outside of the viewshed (450.56 km^2; 11% of GNP) to obtain estimates of the number of mountain goats at each site against which to compare our ground-count estimates. We had no estimate of detection probability to use for correcting aerial survey counts, but again were most interested in determining whether volunteer estimates were statistically similar to this commonly used method for estimating mountain goat populations. Montana Fish, Wildlife & Parks personnel, who had extensive experience with aerial mountain goat surveys, conducted aerial surveys by helicopter during 2 days in August 2009 at minimum above-ground elevations of 150 m. Locations were recorded using GPS for all mountain goats observed. We overlaid all goat locations onto our site viewshed maps to determine the count of mountain goats at each site that were within viewsheds. We developed a density estimate for viewsheds by dividing this count by the summed area of escape terrain in all viewsheds surveyed during aerial surveys. We used regression analysis to compare raw counts within site viewsheds from aerial surveys with raw and corrected high counts from volunteers and biologists. We estimated density from aerial surveys for the entire survey area by dividing the sum of counts at each site by the area of escape terrain at all sites surveyed during aerial surveys. We compared this density estimate with N-mixture model density estimates from volunteers and biologists to determine whether the aerial survey estimate fell within the confidence intervals of either estimate.

RESULTS

The mean area of escape terrain in the viewshed of each site was 4.7 km^2 $\pm$ 2.3 (SD). The area of escape terrain within viewsheds was 149 km^2 (9.0% of escape terrain in GNP) and the area of escape terrain at sites was 727 km^2 (43.9% of escape terrain in GNP). Most escape terrain was above tree-line.

During 2008 and 2009, 140 volunteers were trained and 104 volunteers conducted ≥ 1 survey. Volunteers and biologists spent 4,401.3 and 1,219.1 hours, respectively, conducting mountain goat surveys (Table 1). Experience ranks of biologists ($\bar{x} = 34.9 \pm 1.64$ [SD], range = 33–37) were higher and less varied than those of volunteers

Table 1. Site visits and surveys conducted by volunteers and biologists in 2008–2009 in Glacier National Park, Montana, USA, for mountain goat population estimates.

Year	Observer	Sites visited	No. of surveys	$\bar{x}$ no. of site visits	SD	Sites surveyed $\geq$3 times
2008	Volunteers	30	132	5.95	2.60	22
2008	Biologists	21	33	1.84	0.62	0
2009	Volunteers	32	197	7.24	2.85	31
2009	Biologists	25	76	4.78	2.94	14

($\bar{x} = 24.5 \pm 7.09$, range = 10–39). All biologists were in the skilled experience rank. The proportion of volunteers in each experience rank (including volunteers who conducted ≥ 1 survey) varied: novice = 27%; some relevant experience = 31%; moderately skilled = 27%; and skilled = 15%. . The proportion of volunteers in each experience rank that accepted our invitation to conduct multiple-observer surveys was similar but volunteers in the novice and skilled ranks were slightly underrepresented: novice = 20%; some relevant experience = 38%; moderately skilled = 35%; and skilled = 7%.

Mean $\hat{P}_i$ for GNP biologists and the GNP biologist who conducted multiple-observer surveys with volunteers was not significantly different ($t = 1.534$, $df = 35.2$, $P = 0.93$). Therefore, we combined $\hat{P}_i$ from all multiple-observer surveys by biologists into mean $\hat{P}_{\text{biol}}$ for all biologists. Mean $\hat{P}_{\text{biol}}$ (0.809 ± 0.249) was significantly higher ($t = 3.161$, $df = 81.5$, $P = 0.001$) than mean $\hat{P}_{\text{vol}}$ (0.647 ± 0.317). No misidentifications of other species as mountain goats or other false positives were reported during multiple-observer surveys. Verification photos made during single-observer surveys by volunteers and biologists showed no evidence of misidentifications or false positives. Verification photos were submitted for 15% of groups of mountain goats detected by volunteers and for 74% of groups of mountain goats detected by biologists.

Raw high counts (uncorrected for $\hat{P}_i$) from volunteers had a stronger statistical relationship with counts from biologists in 2008 ($R^2_{\text{adj}} = 0.66$, $df = 19$, $P < 0.001$) than in 2009 ($R^2_{\text{adj}} = 0.38$, $df = 23$, $P < 0.001$; Fig. 2), but the correlation in 2008 may have been influenced by a single point. Density estimates followed the same pattern (Fig. 3). In 2008, 64% of the variation in density estimates by volunteers was explained by variation in density estimates by biologists ($R^2_{\text{adj}} = 0.64$, $df = 19$, $P < 0.001$), but variation was not correlated in 2009 ($R^2_{\text{adj}} = 0.49$, $df = 23$, $P < 0.001$). However, when 2009 density estimates for sites surveyed <3 times by volunteers and biologists were excluded, and one strong leverage point (Cook's distance >1.0) removed, regression results had high explanatory power ($R^2_{\text{adj}} = 0.84$, $df = 11$, $P < 0.001$). Raw, uncorrected counts from aerial surveys were not correlated with estimates from 2009 corrected high counts by volunteers ($R^2_{\text{adj}} = 0.20$, $df = 8$, $P < 0.06$), nor with estimates by biologists ($R^2_{\text{adj}} = 0.07$, $df = 6$, $P = 0.26$; Fig. 4). Density estimates for aerial surveys were also poorly correlated with 2009 density estimates from volunteers ($R^2_{\text{adj}} = 0.47$, $df = 8$, $P < 0.001$) and biologists ($R^2_{\text{adj}} = 0.45$, $df = 6$, $P < 0.04$).

Aerial counts in survey viewsheds and 2009 density estimates from uncorrected high counts by volunteers and biologists in survey viewsheds were similar (1.99 mountain goats/km^2, 1.91 mountain goats/km^2, and 1.87 mountain goats/km^2, respectively). Mean density estimates for all escape terrain at sites, based on corrected high counts by volunteers (0.54–0.72 mountain goats/km^2) and biologists (0.48–0.55 mountain goats/km^2) were lower than density estimates from the aerial survey counts (0.95 mountain goats/km^2). Density estimates by volunteers were higher and more variable than density estimates by biologists, but confidence intervals overlapped. When calculated only for sites visited ≥ 3 times in 2009, however, density estimates by volunteers remained nearly the same (0.56–0.74 mountain goats/km^2) and density estimates by biologists were higher (0.71–0.8 mountain goats/km^2). Relative density estimates (Fig. 5) were in agreement between volunteers and biologists at 18 of 25 sites, between volunteers and aerial surveys at

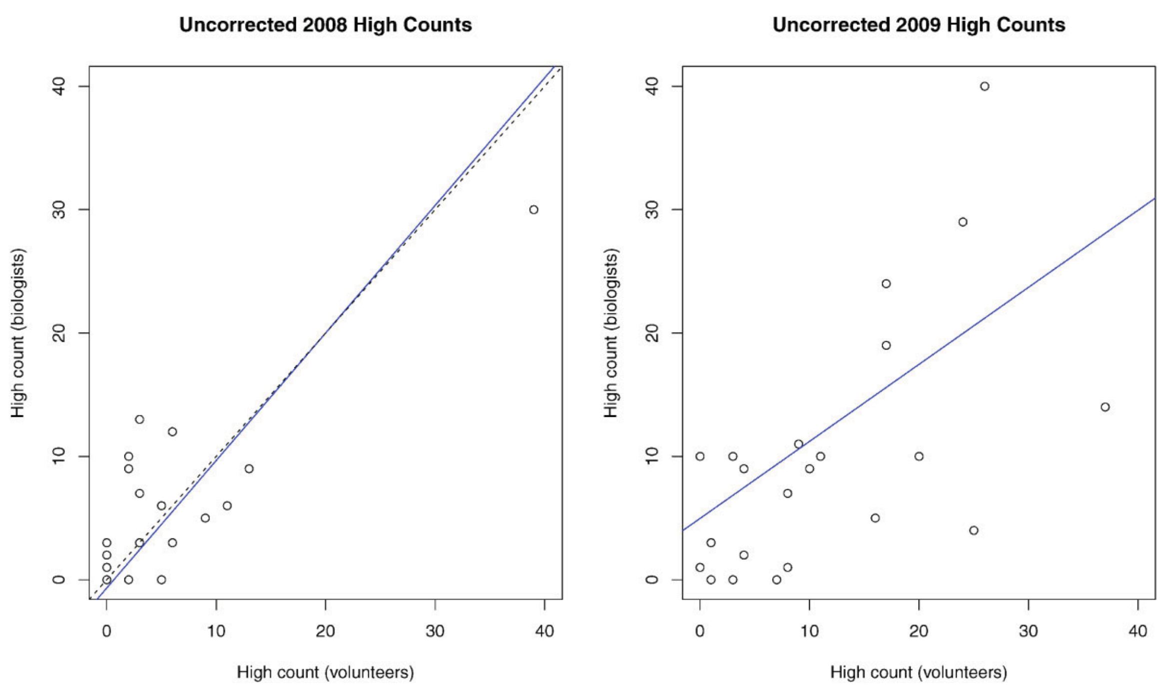

Figure 2. Regression of high counts of mountain goats by volunteers with high counts by biologists for 2008 (left) and 2009 (right) in Glacier National Park, Montana, USA. The solid line is the regression line. The dashed line is 1:1 line.

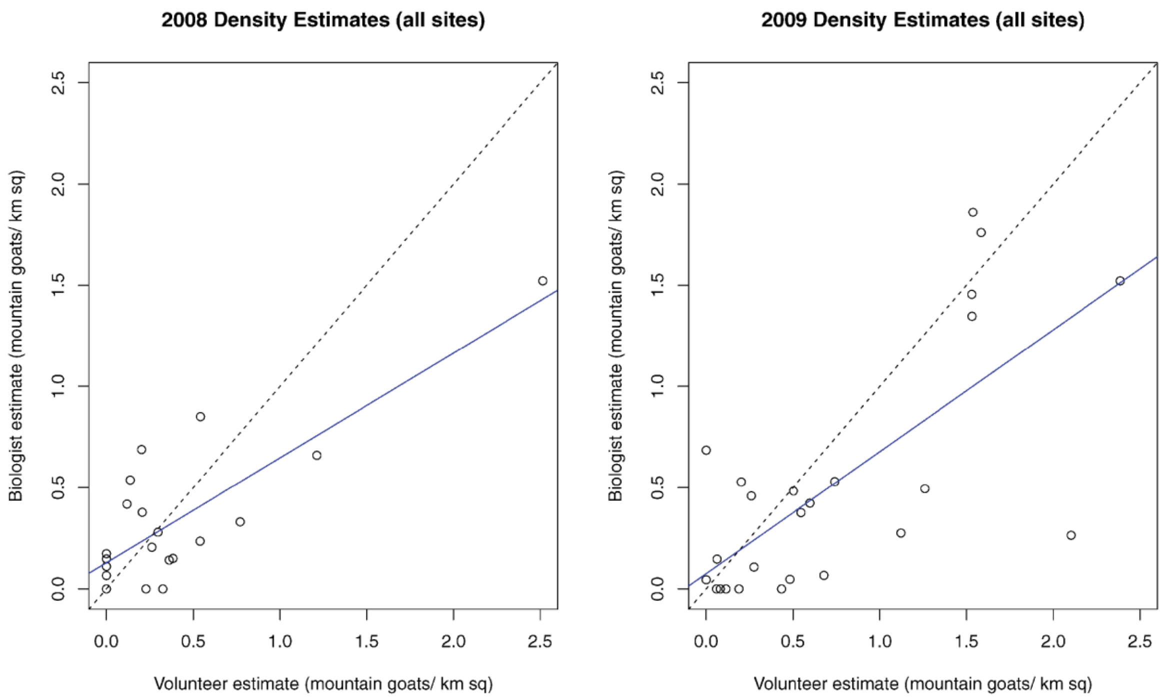

Figure 3. Regression of density estimates of mountain goats by volunteers with density estimates by biologists at all survey sites for 2008 (left) and 2009 (right), Glacier National Park, Montana, USA. The solid line is the regression line. The dashed line showing perfect correlation ($r^2 = 1.0$) is included for comparison.

10 of 11 sites, between biologists and aerial surveys at 6 of 9 sites, and between all three (volunteers, biologists, and aerial) at 5 of 9 sites.

The best supported N-mixture model for 2009 volunteer data included number of site visits (SiteVisits) as a predictor of mean abundance at sites (λ), and group size of largest group of mountain goats (GroupSize) and landscape feature in which the majority of mountain goats were seen (DomFeat) as predictors of detection of mountain goats at sites (*P*). In the top model for 2009 data from biologists, area of escape terrain within viewshed (ViewshedEscape) was the most influential predictor of λ. GroupSize and DomFeat were also influential predictors, but the addition of visibility of mountain goats (visibility) as an additional predictor of

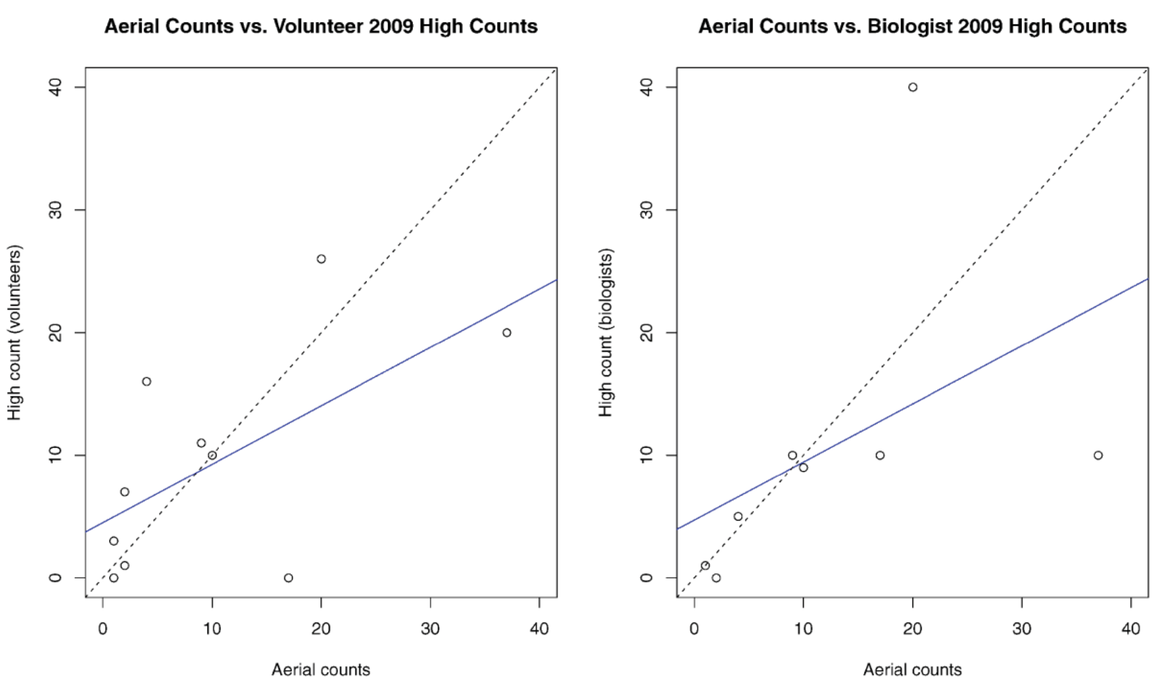

Figure 4. Regression of 2009 corrected high counts of mountain goats by volunteers (left) and corrected high counts by biologists (right) with raw counts from 2009 aerial surveys in Glacier National Park, Montana, USA. High counts for volunteers and biologists are corrected by mean detection probability. The solid line is the regression line. The dashed line showing perfect correlation ($r^2 = 1.0$) is included for comparison.

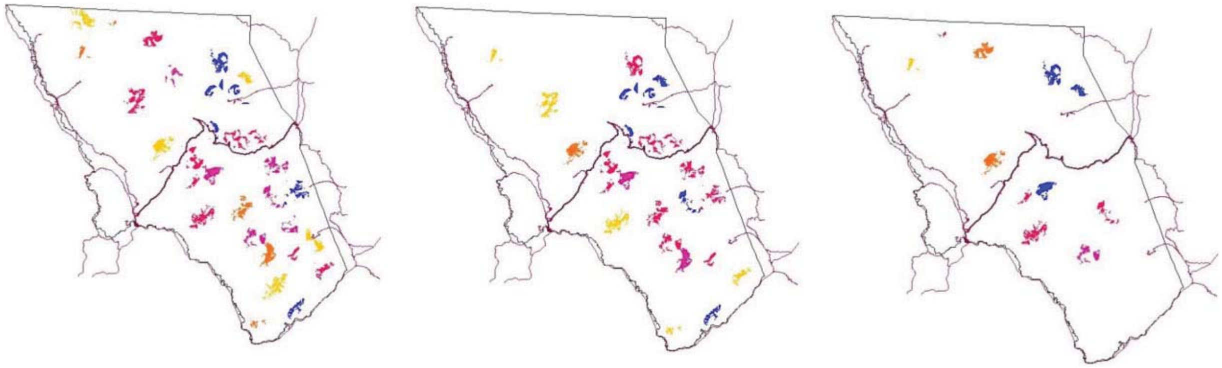

Figure 5. Maps of relative densities of mountain goats estimated from corrected high counts by volunteers (left) and biologists (center), and raw counts from aerial surveys (right) in 2009 Glacier National Park, Montana, USA. Legend: yellow = no mountain goats, orange = low density, red = moderate density, pink = high density, and blue = very high density.

P improved the model performance. Goodness-of-fit tests for the selected models for volunteer data ($P = 0.96$) and biologist data ($P = 0.455$) yielded small differences between observed residual deviance and expected residual deviance, indicating that the N-mixture models fit the data. N-mixture models were not developed for 2008 data due to the low number of site visits by biologists.

The best N-mixture model for 2009 volunteer data (Table 2) estimated abundance at 23.44–32.3 mountain goats/site ($\lambda = 27.52 \pm 2.25$ [SE], $P = 0.06 \pm 0.41$, $\psi = 0.96$). Multiplying λ by the number of sites surveyed ($N = 32$) and dividing by escape terrain at sites (km^2) yielded a density estimate of 1.23 ($\pm$0.195) mountain goats/km^2. The best N-mixture model for 2009 data from biologists estimated abundance at 26.13–45.51 mountain goats/site ($\lambda = 32.95 \pm 3.89$, $P = 0.09 \pm 0.17$, $\psi = 0.97$), yielding a density estimate of 1.56 ($\pm$0.42) mountain goats/km^2 ($N = 25$). Extrapolating these density estimates from volunteer and biologist models to all escape terrain in GNP yielded an estimate of 1,705–2,349 mountain

Table 2. N-mixture models for mountain goat abundance from 2009 volunteer survey data and 2009 biologist survey data from Glacier National Park, Montana, USA. Detection probability of mountain goats is denoted in the column labeled P.

	K	AIC	Δ_i	w_i	λ	SE (λ)	P
Models from 2009 volunteer data[a]							
~GroupSize + DomFeat ~ SiteVisits	10	462.20	0.00	0.40	27.52	2.25	0.06
~GroupSize + DomFeat ~ SiteVisits + Viewshed	11	464.18	1.98	0.15	27.54	2.26	0.06
~GroupSize + DomFeat ~ SiteVisits + Viewshed + ViewshedEscape	12	465.90	3.70	0.06	27.39	2.27	0.06
~GroupSize + DomFeat + StartTime + WindSpeed + ExpRank + SkyCover + Visibility + Temp + PercentMov + Visibility ~ SiteVisits + Viewshed + ViewshedEscape + SiteEscape	26	469.92	7.72	0.01	24.48	2.08	0.05
~GroupSize + DomFeat + StartTime + WindSpeed + ExpRank + SkyCover + Visibility + Temp + PercentMov + BinocPower + BinocView + ScopePower + ScopeView ~ SiteVisits + Viewshed + ViewshedEscape + SiteEscape	30	473.79	11.59	0.00	23.64	2.04	0.06
Null	2	1687.85	1225.65	0.00	27.39	2.28	0.12
Models from 2009 biologist data[a]							
~GroupSize + DomFeat + Visibility ~ ViewshedEscape	9	264.13	0.20	0.26	32.95	3.89	0.09
~GroupSize + DomFeat + Visibility ~ ViewshedEscape + SiteVisits + SiteEscape + Viewshed	12	267.21	3.28	0.06	31.58	3.78	0.09
~GroupSize + DomFeat ~ ViewshedEscape	8	272.43	8.50	0.00	34.36	3.99	0.14
~GroupSize + DomFeat + Visibility + Temp + SkyCover + StartTime + WindSpeed + ExpRank + SkyCover + PercentMov ~ ViewshedEscape + SiteVisits + SiteEscape + Viewshed	24	279.60	15.67	0.00	32.52	4.12	0.09
~GroupSize + DomFeat + Visibility + Temp + SkyCover + StartTime + WindSpeed + ExpRank + SkyCover + PercentMov + BinocPower + BinocView + ScopePower + ScopeView ~ ViewshedEscape + SiteVisits + SiteEscape + Viewshed	28	284.48	20.55	0.00	31.56	4.09	0.09
Null	2	739.94	476.01	0.00	22.87	2.84	0.26

[a] Detection probability covariates for N-mixture models were observer experience (ExpRank), size of largest group of mountain goats detected (GroupSize), temperature (Temp), binocular power (BinocPower), binocular field of view (BinocView), scope power (ScopePower), scope field of view (ScopeView), start time of survey (StartTime), wind speed (WindSpeed), weather (SkyCover), percentage of mountain goats seen that moving when detected (PercentMov), visibility of mountain goats (Visibility), and landscape feature in which the majority of mountain goats were seen (DomFeat). Abundance covariates with potential to influence λ were viewshed area in km^2 (Viewshed), area of escape terrain $\geq 25°$ within viewshed (ViewshedEscape), area of escape terrain $\geq 25°$ within survey site (SiteEscape), and no. of site visits (SiteVisits).

goats by volunteers and 1,885–3,269 mountain goats by biologists.

The N-mixture density estimates by biologists and volunteers were higher than the aerial survey estimate (0.95 mountain goats/km^2). Estimates by volunteers and biologists overlapped the estimate of 1.16 mountain goats/km^2 by Chadwick (1977). Volunteer and biologist estimates were lower than the estimate of 2.9 mountain goats/km^2 by Singer and Doherty (1985).

DISCUSSION

Population estimates from citizen science data were similar to those from data collected by biologists. Uncorrected counts in survey viewsheds by volunteers and biologists were similar and were close to counts from aerial surveys. Confidence intervals of our density estimates from high counts by volunteers and biologists overlapped, despite a lower mean detection probability by volunteers. Density estimates by volunteers and biologists from N-mixture models also overlapped each other and those from earlier research (Chadwick 1977). Volunteer estimates provide similar baseline information, comparable to biologists, for planning future monitoring and research.

A few discrepancies in the estimates raise important considerations. The average density and relative abundance estimates were considerably lower than estimates by Singer and Doherty (1985; i.e., 1.03–1.42 mountain goats/km^2 vs. 2.9 mountain goats/km^2). This may be partially explained by the small area encompassed by their study area and its proximity to a heavily used mineral lick. Conversely, our N-mixture model estimates were higher than aerial survey counts. We found no evidence of misidentification of mountain goats to suggest that our estimates were falsely inflated. A simple explanation may be that detection probability during aerial surveys was less than perfect. We did not measure detection probability during aerial surveys but other studies have reported detection probabilities from 0.55 to 0.84 (Gonzalez-Voyer et al. 2001) and 0.75 to 0.91 (Rice et al. 2009). Using this range (i.e., 0.55–0.91) as a hypothetical correction factor for our aerial survey count yields an estimate of 1.03–1.38 mountain goats/km^2, which overlaps our N-mixture density estimates by volunteers and biologists (Fig. 6).

An alternative explanation for the discrepancy between our estimates and aerial counts is that that our assumption of closure may have been violated. Mountain goat home-range sizes vary widely throughout their range, from 6.3 km^2 (Singer and Doherty 1985) to 24 km^2 (Rideout 1977) for 2 populations studied in Montana. Due to a lack of specific information about mountain goat home-range sizes and locations in GNP, we assumed a rectangular home range. Home ranges may, in fact, be more linear, reflecting fidelity to escape terrain (Brandborg 1955, Hamel and Côté 2007). If home ranges are in fact linear, mountain goats may have moved from one survey site to another, potentially inflating population estimates from ground counts. Ensuring closure using linear home ranges would require a more terrain-specific approach with specific knowledge about movement

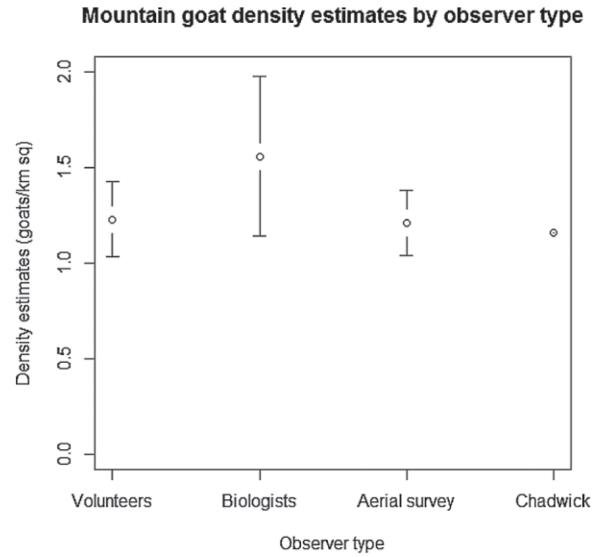

Figure 6. Density estimates of mountain goats by volunteers, biologists, and aerial surveys in 2009 and estimates from previous research (Chadwick 1977) in Glacier National Park, Montana, USA, derived from N-mixture models with confidence intervals where available. Confidence intervals for aerial-survey estimates are derived using a range of detection probabilities (0.55–0.91) from other mountain goat aerial surveys (Gonzalez-Voyer et al. 2001, Rice et al. 2009) as a hypothetical correction factor for our aerial-survey count.

patterns. Alternatively, in future replication of this study, we could adjust the survey design to remove the need for population closure, using an extension of occupancy modeling that allows for temporary emigration (MacKenzie et al. 2003).

Observer experience for volunteers was lower and more varied than for biologists, as we expected, but did not correlate with detection probability or directly influence estimates. The proportion of volunteers in the novice and skilled experience ranks who accepted our invitations to conduct multiple-observer surveys was lower than the proportion of overall volunteers in these experience ranks. This potential underrepresentation of volunteers at the lower and upper ends of observer experience in our multiple-observer sample may have negatively affected the correlation between experience and detection probability.

Other citizen science programs use scores from observer experience to weight data from volunteers (Silvertown 2009). This approach may be effective if skill level is correlated with experience, but our results suggest that weighting data in this manner may not be valid and we caution against this practice. The adequacy of the chosen tool for measuring observer experience is also an important consideration for other programs. In our study, low correlation between experience and estimates of detection probability and abundance may have resulted from failure of our participant survey to accurately measure experience level.

A few relevant metrics of experience were not included because they are difficult to quantify (e.g., the degree to which volunteers have a search image, the amount of investment volunteers have in surveying). Observer bias, such as lower detection probability, generally decreases as observers become more experienced (Delaney et al. 2008) and factoring in number of surveys conducted by each observer may provide a promising avenue for exploring the relationship between experience and observer bias. If number of surveys was strongly correlated to skill level, programs could require a minimum number of surveys before incorporating data from volunteers. An alternative means of assessing skill level in the future could include asking citizen scientists to find goats in photographs with known counts, similar to the "Frog Quiz" used by the North American Amphibian Monitoring Program (Genet and Sargent 2003). Participants who performed poorly could be excluded or trained more intensively.

We used multiple-observer surveys to correct for observer variation between volunteers and biologists and to calibrate our indices of abundance. Our results suggest that detection probabilities from multiple-observer surveys are not sufficient to correct counts because they are not consistent across all levels of mountain goat abundance. This may explain the low correlation between density estimates by volunteers and biologists, but fails to explain the low correlation between high counts by volunteers and biologists that were not corrected by detection probability.

To obtain estimates of true abundance we used N-mixture models, which incorporated the probability that a mountain goat was present at the site, but was not detected because it was either outside of the viewshed during the survey period or missed due to observer error. Similar to multiple-observer surveys, N-mixture models resulted in lower detection probabilities and more variation among volunteers than among biologists. Detection probability in N-mixture models for volunteer and biologist data was again influenced by mountain goat group size. Habitat use also influenced detection probability for volunteers and biologists, while visibility was influential only for biologist data. Habitat use and visibility may have similarly influenced multiple-observer models but these parameters were not tested because data were not available for 2008.

N-mixture model estimates were higher than estimates from high counts in viewsheds that were corrected for detection probability, most likely because not all mountain goats were available for detection in viewsheds during survey periods. The high probability of occurrence ($\geq$0.96) and low detection probability ($\leq$0.094) estimated by N-mixture models confirm this explanation and suggest that mountain goats inhabiting sites were frequently absent from survey viewsheds. Abundance estimates by biologists were influenced by area of escape terrain in viewsheds, a result supported by previous research that identified escape terrain as the best predictor of mountain goat occurrence (Gross et al. 2002, Hamel and Côté 2007). While the area of escape terrain in viewsheds also influenced abundance estimates by volunteers in several of the highest ranking models, the number of site visits alone best explained the variation and had the largest influence on abundance estimates by volunteers. Other studies have reported that volunteers underestimate abundance due to lower detection probability (Newman et al. 2003, Delaney et al. 2008). In our study, however, the effect of lower detection probability on abundance estimates by volunteers was balanced out by the larger number of site visits. Variation in abundance estimates was lower for volunteers than for biologists because volunteers surveyed sites more frequently.

The difference in number of site visits by volunteers and biologists may explain why uncorrected high counts and density estimates by volunteers and biologists were not more highly correlated. By surveying sites more often, volunteers captured a higher minimum count of mountain goats at >50% of sites. The larger variation in detection probability among volunteers, however, led to volunteers reporting a larger proportion of counts that were lower than biologist counts. Therefore, higher N-mixture model estimates by biologists also may be explained by greater consistency in counts despite higher minimum counts by volunteers.

Alternately, higher estimates by biologists may be a result of instability in the N-mixture models caused by the smaller sample size of surveys by biologists combined with the low detection probability and high rate of occupancy for mountain goats. This explanation highlights a potential advantage of using citizen science to attain large enough sample sizes to attain stable N-mixture estimates for species that are abundant and easy to identify, but difficult to detect. Conducting surveys by biologists and aerial surveys at the same spatial and temporal scale as surveys by volunteers would be an ideal way to assess the impact of sample size on stability of N-mixture estimates, but would be cost-prohibitive. A future investigation could conduct simulations to develop an optimal sample size for biologists and volunteers, given the detection probability and occupancy of mountain goats we reported.

MANAGEMENT IMPLICATIONS

Wildlife managers faced with limited funding to meet their monitoring needs are increasingly turning to the free labor source provided by the public (Silvertown 2009), but establishing and coordinating a citizen science program requires financial commitment and effort (Yung 2007). Managers must determine which will better meet their conservation objectives: hiring a citizen science project manager to coordinate volunteers to cover a larger sample area, or enlisting a small number of biologists to cover a smaller sample area. Our results suggest that the 2 methods may yield statistically similar population estimates if enough data are collected by volunteers.

Citizen science programs involved with long-term monitoring should incorporate some measure of data quality. The cost of employing biologists or using other methods (e.g., mark–recapture) limits comparisons of data on a similar scale to the data that can be collected by volunteers. The use of multiple-observer surveys to correct volunteer data may not be viable for citizen science data due to the high variability in detection probability. However, double-sampling using data collected by biologists or data from aerial surveys over a smaller subsample offers a useful comparison provided that enough data are collected to measure detection probability. Data quality comparisons will likely be most effective once the program has been established (e.g., $\geq$1 yr after initiation), because managers can then determine how much data needs to be collected by volunteers to yield meaningful inferences.

Because citizen science population estimates from small sample sizes are not comparable to biologist estimates, we do not recommend citizen science as a direct substitute for professional monitoring. Citizen science will only produce similar population estimates to those of biologists when sample sizes are larger than those attainable by biologists. We reported that abundance estimates were positively influenced by number of site visits and that variation was negatively influenced by number of sites surveyed. Future research on the number of site visits and number of surveys at each site that maximize precision of citizen science estimates would contribute toward increasing the power of detecting trends.

ACKNOWLEDGMENTS

We thank GNP (S. J. Hejl, T. Carolin, M. Sladek, S. Sindt, B. Thomas, P. Ollig, J. Potter, J. Waller, L. Bate, T. Shepherd, D. Kotter, and R. Menicke), the citizen science volunteers and S. Gniadek, D. Chadwick, and Montana Fish, Wildlife & Parks biologists. The comments of J. Berger, M. Mitchell, and J. Derbridge improved this manuscript. Funding was provided by The GNP Fund, the Crown of the Continent Research Learning Center at GNP, the Boone & Crockett Program in Wildlife Conservation at the University of Montana, and the International Order of Rocky Mountain Goats.

LITERATURE CITED

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

Brandborg, S. M. 1955. Life history and management of the mountain goat in Idaho. Idaho Wildlife Bulletin 2:1–142.

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

Chadwick, D. H. 1976. Ecological relationships of mountain goats in Glacier National Park. American Institute of Biological Sciences First Conference on Scientific Research in the National Parks, New Orleans, Louisiana, USA.

Chadwick, D. H. 1977. The influence of mountain goat social relationships on population size and distribution. Proceedings of the International Mountain Goat Symposium 1:74–91.

Cohn, J. P. 2008. Citizen science: can volunteers do real research? BioScience 58:192–197.

Cooper, C. B., J. Dickinson, T. Phillips, and R. Bonney. 2007. Citizen science as a tool for conservation in residential ecosystems. Ecology and Society 12:11.

Côté, S. D., M. Festa-Bianchet, and F. Fournier. 1998. Life-history effects of chemical immobilization and radiocollars on mountain goats. Journal of Wildlife Management 62:745–752.

Danielsen, F., N. D. Burgess, and A. Balmford. 2005. Monitoring matters: examining the potential of locally-based approaches. Biodiversity and Conservation 14:2507–2542.

Delaney, D. G., C. D. Sperling, C. S. Adams, and B. Leung. 2008. Marine invasive species: validation of citizen science and implications for national monitoring networks. Biological Invasions 10:117–128.

Dunn, E. H., C. M. Francis, P. J. Blancher, S. R. Drennan, M. A. Howe, D. Lepage, C. S. Robbins, K. V. Rosenberg, J. R. Sauer, and K. G. Smith. 2005. Enhancing the scientific value of the Christmas bird count. The Auk 122:338–346.

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

Festa-Bianchet, M., and S. D. Côté. 2008. Mountain goats: ecology, behavior and conservation of an alpine ungulate. Island Press, Washington, D.C., USA.

Field, S. A., A. J. Tyre, and H. P. Possingham. 2005. Optimizing allocation of monitoring effort under economic and observation constraints. Journal of Wildlife Management 69:473–482.

Fitzpatrick, M. C., E. L. Preisser, A. M. Ellison, and J. S. Elkington. 2009. Observer bias and the detection of low-density populations. Ecological Applications 19:1673–1679.

Genet, K. S., and L. G. Sargent. 2003. Evaluation of methods and data quality from a volunteer-based amphibian call survey. Wildlife Society Bulletin 31:703–714.

Gonzalez-Voyer, A., M. Festa-Bianchet, and K. G. Smith. 2001. Efficiency of aerial surveys of mountain goats. Wildlife Society Bulletin 29:140–144.

Gouveia, C., A. Fonseca, A. Câmara, and F. Ferreira. 2004. Promoting the use of environmental data collected by concerned citizens through information and communication technologies. Journal of Environmental Management 71:135–154.

Greenwood, J. J. D. 2007. Citizens, science and bird conservation. Journal of Ornithology 148:77–124.

Gross, J. E., M. C. Kneeland, D. F. Reed, and R. M. Reich. 2002. GIS-based habitat models for mountain goats. Journal of Mammalogy 83:218–228.

Hamel, S., and S. D. Côté. 2007. Habitat use patterns in relation to escape terrain: are alpine ungulate females trading off better foraging sites for safety? Canadian Journal of Zoology 85:933–943.

Haynes, L. A. 1992. Mountain goat habitat of Wyoming's Beartooth Plateau: implications for management. Biennial Symposium of the Northern Wild Sheep and Goat Council 8:325–339.

Hochachka, W. M., K. Martin, F. Doyle, and C. J. Krebs. 2000. Monitoring vertebrate populations using observational data. Canadian Journal of Zoology 78:521–529.

Hop, K., M. Reid, J. Dieck, S. Lubinski, and S. Cooper. 2007. U. S. Geological Survey-National Park Service Vegetation Mapping Program: Waterton-Glacier International Peace Park. U. S. Geological Survey, Upper Midwest Environmental Sciences Center, La Crosse, Wisconsin, USA.

Johnson, D. H. 2008. In defense of indices: the case of bird surveys. Journal of Wildlife Management 72:857–868.

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.

Koeth, C. 2008. Clinging to existence. Montana Outdoors (Mar–Apr): 23–27.

Kremen, C., K. S. Ullman, and R. W. Thorp. 2010. Evaluating the quality of citizen-scientist data on pollinator communities. Conservation Biology 25:607–617.

MacKenzie, D. I., and W. L. Kendall. 2002. How should detection probability be incorporated into estimates of relative abundance? Ecology 83:2387–2393.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84:2200–2207.

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

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

Martin, S. R. 1997. Specialization and differences in setting preferences among wildlife viewers. Human Dimensions of Wildlife 2:1–18.

McFetridge, R. J. 1977. Strategy of resource use by mountain goat nursery groups. Proceedings of the International Mountain Goat Symposium 1:169–173.

Morisette, J. T., A. D. Richardson, A. K. Knapp, J. I. Fisher, E. A. Graham, J. Abatzoglou, B. E. Wilson, D. D. Breshears, G. M. Henebry, J. M. Hanes, and L. Liang. 2009. Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. Frontiers in Ecology and the Environment 7:253–260.

Newman, C., C. D. Buesching, and D. W. Macdonald. 2003. Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation—"*Sed quis custodiet ipsos custodies*?" Biological Conservation 113:189–197.

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

Pettorelli, N., F. Pelletier, A. Von Hardenberg, M. Festa-Bianchet, and S. D. Côté. 2007. Early onset of vegetation growth vs. rapid green-up: impacts on juvenile mountain ungulates. Ecology 88:381–390.

Pilz, D., H. Ballard, and E. T. Jones. 2005. Broadening participation in biological monitoring: guidelines for scientists and managers. Institute for Culture and Ecology, Portland, Oregon, USA.

Rice, C. G., K. J. Jenkins, and W. Chang. 2009. A sightability model for mountain goats. Journal of Wildlife Management 73:469–478.

Rideout, C. B. 1977. Mountain goat home ranges in the Sapphire Mountains of Montana. Proceedings of the International Mountain Goat Symposium 1:201–211.

Root, T. L., and P. Alpert. 1994. Volunteers and the NBS. Science 263:1205.

Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. Biometrics 60:108–115.

Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence–absence data or point counts. Ecology 84:777–790.

Sauer, J. R., B. G. Peterjohn, and W. A. Link. 1994. Observer differences in the North American Breeding Bird Survey. The Auk 111:50–62.

Scott, D., R. B. Ditton, J. R. Stoll, and T. L. Eubanks, Jr. 2005. Measuring specialization among birders: utility of a self-classification measure. Human Dimensions of Wildlife 10:53–74.

Shackleton D. M., editor. 1997. Wild sheep and goats and their relatives. Status survey and conservation action plan for Caprinae. International Union for Conservation of Nature/Status Survey and Conservation Caprinae Specialist Group, Gland, Switzerland and Cambridge, United Kingdom.

Silvertown, J. 2009. A new dawn for citizen science. Trends in Ecology and Evolution 24:467–471.

Singer, F. J., and J. L. Doherty. 1985. Movements and habitat use in an unhunted population of mountain goats, *Oreamnos americanus*. Canadian Field-Naturalist 99:205–217.

Smith, B. L. 1988. Criteria for determining age and sex of American mountain goats in the field. Journal of Mammalogy 69:395–402.

Trumbull, D. J., R. Bonney, D. Bascom, and A. Cabral. 2000. Thinking scientifically during participation in a citizen-science project. Science Education 84:265–275.

Varley, N. C. 1994. Summer–fall habitat use and fall diets of mountain goats and bighorn sheep in the Absaroka Range, Montana. Biennial Symposium of the Northern Wild Sheep and Goat Council 9:131–138.

Veitch, A., E. Simmons, M. Promislow, D. Tate, M. Sallow, and R. Popko. 2002. The status of mountain goats in Canada's Northwest Territories. Northern Wild Sheep and Goat Council 13:49–62.

Welch, B. L. 1947. The generalization of "Student's" problem when several different population variances are involved. Biometrika 34: 28–35.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2003. Monitoring of biological diversity—a response to Danielsen et al. Oryx 37:412.

Yung, L. 2007. Citizen monitoring and restoration: volunteers and community involvement in wilderness stewardship. U.S. Department of Agriculture Forest Service Proceedings P-49, Rocky Mountain Research Station, Fort Collins, Colorado, USA.

Associate Editor: McDonald.