esa ECOSPHERE

Comparing citizen science and professional data to evaluate extrapolated mountain goat distribution models

Elizabeth P. Flesch[1,†] and Jami J. Belt[2]

Crown of the Continent Research Learning Center, Glacier National Park, P.O. Box 128, West Glacier, Montana 59936 USA

Citation: Flesch, E. P., and J. J. Belt. 2017. Comparing citizen science and professional data to evaluate extrapolated mountain goat distribution models. Ecosphere 8(2):e01638. 10.1002/ecs2.1638

Abstract. Citizen science provides a prime opportunity for wildlife managers to obtain low-cost data recorded by volunteers to evaluate species distribution models and address research objectives. Using mountain goat (*Oreamnos americanus*) location data collected through aerial surveys by professionals, ground surveys by professionals, and ground surveys by volunteers, we evaluated two mountain goat distribution models extrapolated across Waterton-Glacier International Peace Park. In addition, we compared mountain goat location data by observer and survey type to determine whether there were differences that affected extrapolated model evaluation. We found that all dataset types compared similarly to both mountain goat models. A mountain goat occupancy model developed in the Greater Yellowstone Area (GYA) was the most informative in describing mountain goat locations. We compared Spearman-rank correlations (r_s) for occupancy probability bin ranks in the GYA model extrapolation and area-adjusted frequencies of mountain goat locations, and we found that all datasets had a positive correlation, indicating the model had useful predictive ability. Aerial observations had a slightly greater Spearman-rank correlation (r_s = 0.964), followed by the professional ground surveys (r_s = 0.946), and volunteer ground datasets (r_s = 0.898). These results suggest that with effective protocol development and volunteer training, biologists can use mountain goat location data collected by volunteers to evaluate extrapolated models. We recommend that future efforts should apply this approach to other wildlife species and explore development of wildlife distribution models using citizen science.

Key words: alpine; citizen science; mountain goats; national parks; *Oreamnos americanus;* population management; species distribution models; ungulate.

Received 31 August 2016; revised 10 November 2016; accepted 11 November 2016. Corresponding Editor: Debra P. C. Peters.

[1] Present address: Ecology Department, Montana State University, 310 Lewis Hall, Bozeman, Montana 59717 USA.

[2] Present address: Klondike Gold Rush National Historical Park, P.O. Box 517, Skagway, Alaska 99840 USA.

† **E-mail:** elizabeth.flesch@gmail.com

Introduction

Building models to predict how environmental factors influence habitat selection of wildlife is a critical step in development of monitoring strategies and assessment of the effects of habitat changes on species distribution, such as those caused by climate change or human impact (Stillman and Brown 1994, Guisan and Zimmermann 2000). Once a predictive model is developed using landscape characteristics acquired from remote sensing and GIS analysis, its accuracy and usefulness are often assessed by comparing the model with a subset of data or by collected data for comparison in the region of model development using different methods (Beard et al. 1999). Rather than building novel models for each discrete area, which can be expensive and time consuming, ecological models can ideally be extended to inform ecological studies in different areas and enable conservation questions to be addressed across a broader landscape scale. However, the extrapolated model must be evaluated, as landscape features in the new study area may affect species in different ways and at dissimilar spatial scales (Wiens 1989, Hobbs 2003, McAlpine et al. 2008, Mayor et al. 2009, Zanini et al. 2009). Evaluation of existing models can determine how well model predictions match actual species locations and can provide information for managers to decide whether an extrapolated model is useful or accurate (Gutzwiller and Barrow 2001).

Evaluation of wildlife distribution models necessitates the collection of species location information across the area of interest, which may require less intensive data collection than building a new model, but can still present a substantial logistical and budgetary challenge. To address this issue, citizen science could serve as an effective approach to gather data for evaluating wildlife distribution models. In general, citizen science programs typically involve research and monitoring where volunteers record data independently (Trumbull et al. 2000). As availability of research funding decreases (Pilz et al. 2006) and disturbances to ecosystems, such as climate change, impact wildlife (Morisette et al. 2008), the call for public involvement in ecological stewardship increases (Yung 2007) and many natural resource programs are employing citizen scientist research and monitoring to address management goals and research questions that encompass large spatial areas.

Citizen science programs with diligent volunteer training and strategic study design may develop datasets that are similar in quality to those collected by professionals (Hochachka et al. 2000, Gouveia et al. 2004). In addition, citizen science can allow research programs to collect data at greater spatial scales that may not be attainable through typical research budgets (Cooper et al. 2007, Greenwood 2007, Cohn 2008) and can be used to improve understanding of species abundance and distribution (Dickinson et al. 2010). Citizen science data were used successfully to build models that describe terrestrial wildlife distribution, such as that of bobcats, screech owls, and insects (Nagy et al. 2012, van Strien et al. 2013, Broman et al. 2014). To ensure the quality of volunteer data prior to its application to model evaluation, it can be helpful to compare citizen science observations with those recorded by professionals with greater research experience, to determine whether volunteer and professional observations were similar, using the same or different protocols (Broman et al. 2014).

Glacier National Park (Glacier) has coordinated mountain goat (*Oreamnos americanus*) surveys based on citizen science as a primary tool for population monitoring since 2008. Glacier hosts one of the largest intact native mountain goat populations in the continental United States, and the mountain goat is an iconic alpine species whose response to climate change remains uncertain (Pettorelli et al. 2007). Concern and lack of population trend data regarding the mountain goat population prompted the development of a citizen science monitoring program in Glacier. After 2 years of data collection, biologists compared the data collected by volunteers to data collected by professionals using the same methods and found that data collected by volunteers in ground surveys were statistically similar to those collected by biologists in ground and aerial surveys (Belt and Krausman 2012). They determined that citizen science could be used to monitor mountain goat population trends over time in Glacier (Belt and Krausman 2012). Mountain goat ground counts by volunteers were successful in Glacier in part because mountain goats are often highly visible and easily recognizable (Veitch et al. 2002) and high visitation provided a prime opportunity to engage the public for citizen science, resulting in large datasets collected by volunteers (Belt and Krausman 2012).

Glacier's past citizen science surveys focused on mountain goat density and population trends, but did not provide information on species distribution and habitat. A mountain goat occupancy model for Glacier and the adjacent Waterton Lakes National Park (Waterton-Glacier) would address that gap. Occupancy modeling provides the ability to account for variability of detection when evaluating habitat selection of rare wildlife species across a spatial area (MacKenzie et al. 2005, MacKenzie 2006). An evaluated mountain goat occupancy model developed in another area and extrapolated to the Waterton-Glacier region would provide increased understanding of mountain goat habitat use and could improve population density estimates derived from data collected by citizen scientists. Many mountain goat habitat selection models have been developed in areas where mountain goats are non-native, relatively new to the region, and increasing in number and distribution (Gross et al. 2002, Cobb et al. 2012, DeVoe et al. 2015). Extrapolating models developed in non-native range to assess whether they can accurately predict locations in native range, such as Waterton-Glacier, could also lead to ecological insight regarding mountain goat habitat selection. Researchers often caution against model extrapolation if there is a lack of species observations to support model predictions, but with the use of local species location data, evaluated species distribution model extrapolations can be informative if researchers rigorously assess their methods and results (Elith and Leathwick 2009).

Therefore, our first objective was to determine whether we could extrapolate mountain goat occupancy models developed in other areas and use them to predict mountain goat locations in Waterton-Glacier International Peace Park (Waterton-Glacier). We examined two models that were successful in predicting a very high percentage (72.1–87%) of mountain goat locations during model validation, one in the Greater Yellowstone Area (GYA; DeVoe et al. 2015) and the other in Mt. Evans, Colorado (Gross et al. 2002). We planned to evaluate these models using data collected by both volunteers and professionals. We selected the Colorado model because it evaluated similar primary habitat to Waterton-Glacier and emphasized proximity to "escape terrain," or a certain distance from areas of steep slope, a common variable chosen to develop mountain goat habitat models (Fox 1983, Gross et al. 2002, Cobb et al. 2012). The use of escape terrain as the only predictor of mountain goat occupancy has proven effective but limited. Terrain is typically deemed as either suitable or unsuitable using this approach, and the model does not provide a range of occupancy or relative probabilities across terrain. In addition, other environmental variables likely important for ungulate habitat selection, such as vegetation, are excluded. Therefore, we also considered an occupancy model developed in the GYA, which integrated multiple covariates and provided a range of probabilities for mountain goat occupancy (DeVoe et al. 2015).

Our second objective was to ask whether volunteers could record mountain goat locations as accurately as professional biologists, to contribute to evaluation of extrapolated models. To make this comparison, we needed to determine whether there was detection bias caused by either survey or observer type. To assess detection bias due to survey type, we compared mountain goat location data collected through professional aerial surveys to data collected during volunteer and professional ground surveys to determine whether the survey type would impact results and model evaluation. Aerial surveys are often used to count mountain goats, as they provide high animal detection, can cover large areas, and can be effective for population trend monitoring if detection rates can be estimated (Bender et al. 2003, Festa-Bianchet and Côté 2008). However, aerial surveys are often expensive and weather dependent, meaning that replication or thorough coverage of the study area may be cost-prohibitive. While potentially limited in visibility by surrounding terrain and tree cover, ground-based surveys can be advantageous in that observers have more time to search for mountain goats and record accurate locations. In addition, ground surveys are often less disruptive to mountain goats than aerial surveys, which can disturb mountain goats (Côté 1996). Thus, in national parks and wilderness areas that attempt to minimize overflights, ground surveys can serve as a less invasive alternative. To detect possible differences in results due to observer experience with surveying for mountain goats, we compared data collected by volunteers and professionals using the same ground survey protocol, to determine whether both types of observers with different levels of experience could provide comparable data for extrapolated model evaluation.

Methods

Study area

We collected mountain goat location data in Waterton-Glacier International Peace Park, which is composed of Glacier National Park and Waterton Lakes National Park. Glacier National Park encompasses over 400,000 ha in northwest Montana and is managed by the National Park Service. Waterton Lakes National Park consists of over 50,000 ha in southwest Alberta and is managed by

Parks Canada. The Continental Divide of the Rocky Mountains extends from north to south in the study area and serves as a boundary between Pacific-maritime climatic influences in the west and drier, continental climate and easterly winds on the east side. Elevations range from 960 m along the Flathead River to 3185 m at the summit of Mt. Cleveland. Pleistocene glaciations largely shaped landscape topography, resulting in many glacial lakes, moraines, cirques, and active glaciers throughout the study area. Glacier National Park hosts over two million visitors annually and mountain goats are highly visible and popular, providing an opportunity to involve the general public in wildlife research. In addition, mountain goat habitat is easily accessible from roads and a large number of hiking trails, which offer access to mountain goat viewing from close proximity.

Data collection

From 2013 to 2015, trained volunteers and park staff completed surveys and collected mountain goat location data from the ground from 9 May to 24 October, at 37 designated ground sites in Glacier National Park, selected according to the protocol by Belt and Krausman (2012). We defined the summer season as May through October to emulate the season used for model development by DeVoe et al. (2015) (June through October). Occupancy theory assumes that surveys will generally occur over a shorter timeframe; however, definition of a long, rather than narrowly defined, season to evaluate habitat use can be appropriate if congruent with study objectives (MacKenzie 2005). Our study objectives were to define all habitat used over the entire survey period available for ground-based observations of mountain goats in Waterton-Glacier to accommodate general evaluation of mountain goat distribution models and volunteer data collection. At survey sites, participants surveyed the surrounding terrain up to 3.2 km from the observer's position for 1 h, using magnifying optics (binoculars and spotting scopes). For every mountain goat observation, observers recorded total number, sex, age, predominant behavior upon detection, predominant landscape features, and location data (UTM coordinates). Observers also captured location data for opportunistic mountain goat sightings outside of surveys. Participants recorded locations using handheld GPS units, mobile mapping applications, digital aerial photos, and paper topographic maps of survey sites at approximately 1:45,000 scale. We recorded mountain goat location data that had fine spatial accuracy (recorded at 1:50,000 scale or finer, which was the level of accuracy of model validation data used by DeVoe et al. 2015). We asked all volunteers for information about their experience with using topographic maps or aerial photos to pinpoint locations of observations. The subset of volunteers who indicated that they were highly familiar with using topographic maps and/or aerial photos were asked to capture location data. Mountain goat locations collected by volunteers independently were classified as "volunteer ground" observations. Locations recorded by wildlife professionals independently or in a group with volunteers, such as during a training session, were classified as "professional ground" observations. Since mountain goats often occur in groups, such that individual locations within a social group are likely not independent, we collected and analyzed locations of mountain goat groups rather than locations of each individual within groups. Locations of single mountain goats were also recorded and weighted equally during analysis as locations of groups. The Glacier National Park Citizen Science Program coordinated volunteer training, field surveys, and data management.

Park staff conducted aerial surveys for mountain goats in Glacier National Park during 14–17 August 2008 and 18–19 August 2009 and in Waterton Lakes National Park on 25 July 2011. In 2008, three different park staff observers surveyed over four days from a fixed-wing airplane in mountainous areas the observers subjectively deemed as probable mountain goat habitat. In 2009, two observers simultaneously conducted surveys over two days by helicopter and recorded mountain goats found in viewsheds visible from citizen science ground survey sites. In 2011, Alberta province staff surveyed Waterton Lakes National Park from a helicopter, with two observers in the rear seat continually watching for mountain goats and one navigator in the left front passenger seat maintaining flight course and recording locations. During all aerial surveys, observers recorded mountain goat locations using a GPS unit in the aircraft during flight. We classified these recorded mountain goat locations as "professional aerial" observations.

We estimated the mountain goat population size of Waterton Lakes National Park using the

Table 1. Values and equations used to modify covariate raster spatial layers to extrapolate a predictive mountain goat occupancy model from the Greater Yellowstone Area in Waterton-Glacier International Peace Park.

Covariate	Raster description	Transformation	Mean ($\bar{X}$)	SD
SLP500	Slope (degrees) at 500-m scale	$(X - \bar{X})/\mathrm{SD}$	20.37329	12.46569
SLPv500	Slope variance at 500-m scale	$\ln((X/100) + 0.005)$	…†	…†
COV100	Forest canopy cover (percent) at 100-m scale	$(X - \bar{X})/\mathrm{SD}$	39.56991	34.91701
GRAD100	Global radiation (W/m^2) at 100-m scale	$(X - \bar{X})/\mathrm{SD}$	505,802.7	89,188.45
NDVI100	Normalized difference vegetation index (July 2001 Landsat derivative) at 100-m scale	$(X - \bar{X})/\mathrm{SD}$	0.099172	0.209397

† Mean and standard deviation were not required to transform slope variance at 500-m scale.

most recent aerial survey count from 25 July 2016 and a range of potential detection probabilities for mountain goat aerial surveys in the literature, including 0.55 (low), 0.70 (moderate), and 0.90 (high; Gonzalez-Voyer et al. 2001, Rice et al. 2009, Flesch et al. 2016). We report an estimated population size of Glacier National Park from the study completed by Belt and Krausman (2012).

Model extrapolation

We extrapolated a mountain goat occupancy model developed for the GYA (hereafter referred to as the Yellowstone model) to produce a layer of predicted mountain goat occupancy probability across Waterton-Glacier (DeVoe et al. 2015). This model was developed in the GYA using detection–nondetection data collected through ground-based occupancy surveys of viewsheds consisting of sampling units 100 m × 100 m in dimension (DeVoe et al. 2015). Habitat selection was associated with mean slope, slope variance, canopy cover, heat load, and normalized difference vegetation index (NDVI); 72.1% of model validation data were contained in areas classified as low suitability or better (DeVoe et al. 2015). We built spatial layers for environmental covariates at the specified scales of the top model for mountain goat occupancy in the GYA, including mean slope (500 m), slope variance (500 m), forest canopy cover (100 m), heat load (100 m), and NDVI (at 100 m; Table 1; DeVoe et al. 2015). All data processing was completed in the ArcMap module of ArcGIS 10.1 or 10.3.1, using NAD 1983 datum UTM Zone 12N (ESRI 2010). To develop the mean slope and slope variance layers at 500 m resolution, we calculated slope using the 1-arc-second resolution National Elevation Dataset (NED) raster and the ArcMap Spatial Analyst tool. We resampled 30-m grid cells describing slope to 100 m resolution using cubic interpolation and then calculated mean and SD focal statistics from a neighborhood of 5 × 5 cells to obtain 500 m resolution. We calculated heat load at 100 m resolution by using the NED raster and the ArcMap Solar Radiation tool. Latitude was set for 48.71° N to make radiation estimates accurate for the entire Waterton-Glacier study area. We resampled the solar radiation grid cells from 30 m resolution to 100 m resolution using cubic interpolation.

We calculated forest canopy cover by resampling a 30 m resolution canopy closure layer (Homer et al. 2007, Crown Managers Partnership 2012) to 100 m resolution using cubic interpolation. Finally, we developed the NDVI layer at 100 m resolution using a Landsat 7 SLC-on scene from 7 July 2001 with 0.02% cloud cover. High-elevation areas were covered by both perennial and seasonal snowpack in the summer, which may affect NDVI calculations (Appendix S1: Fig. S1). The image served as a single snapshot in time of NDVI, which was assumed to be generally representative of forage available to mountain goats during the summer. This image was processed to calculate NDVI values and resampled to 100 m resolution (Homer et al. 2007). The NDVI values of grid cells in areas with greater than 15% canopy cover were adjusted to 0.0609 to mask the effect of canopy cover on NDVI (DeVoe et al. 2015). The mean and SD for each covariate layer were comparable to those used in the occupancy model for the GYA (Table 1; Hirzel and Le Lay 2008, DeVoe 2015). Thus, the terrain of the northern GYA was sufficiently similar to that of Waterton-Glacier such that we extrapolated the unmodified Yellowstone model for evaluation in our study area (Hirzel and Le Lay 2008). Covariate rasters were centered and scaled or pseudo-threshold transformed (Table 1) and implemented in the logistic regression equation of the predictive occupancy model (DeVoe et al. 2015). Inverse logistic transformation converted the predictors from logit to relative probability scale (DeVoe et al. 2015). Since many lakes cover the terrain of Waterton-Glacier, we clipped the habitat suitability layer by lakes, to remove any grid cells with predicted occupancy that were covered by water and thus unavailable to mountain goats. The resulting product was a map of Waterton-Glacier habitat suitability, which displayed modeled mountain goat occupancy probabilities at 100 m resolution, because the same resolution was used for the Yellowstone model in the GYA (DeVoe et al. 2015). We defined occupancy probability (Ψ) in the same manner as DeVoe et al. (2015): "the probability of a group being in a sampling unit at the time of the survey" (DeVoe et al. 2015). We used the minimum occupancy probability definition used in the GYA study for suitable habitat, which was $\Psi \geq 0.0058$.

To consider an alternative model that emphasized distance to escape terrain, we replicated a model used to describe mountain goat locations of a non-native population in Colorado, where areas ≤258 m from ≥33° slope were classified as suitable mountain goat habitat (hereafter referred to as the Colorado model; Gross et al. 2002). Gross et al. (2002) collected location data on a map at 1:24,000 scale via a ground survey route to classify used and unused areas during summer and winter for 6 years. The simple escape terrain model correctly predicted 87% of model validation observations (Gross et al. 2002). We extrapolated this model across Waterton-Glacier using the 1-arc-second resolution NED raster and the ArcMap Spatial Analyst tool. We resampled the layer to 100 m resolution using cubic interpolation, to make the grid cell resolution directly comparable to the Yellowstone model. A buffer of 258 m was generated around areas of ≥33° slope, and the predicted occupancy layer was clipped to exclude lakes. The final layer predicted areas as either suitable or unsuitable mountain goat habitat.

Evaluation of extrapolated models

We used a percentage overall accuracy approach with presence data to assess and compare the number of locations within suitable areas predicted by the Yellowstone and Colorado model extrapolations, as we did not have unused data. We selected this approach because it is commonly used when unused data are not available (Fielding and Bell 1997), a similar approach was used to validate the original Yellowstone model (DeVoe et al. 2015), and presence-only evaluators are correlated with presence–absence evaluators (Hirzel et al. 2006). We intersected both Yellowstone and Colorado models with mountain goat location data collected through volunteer ground, professional ground, and professional aerial survey platforms. Since the Yellowstone model extrapolation produced a range of occupancy probabilities, we further evaluated this model by sampling availability of resources by generating 10,000 random locations within the Waterton-Glacier boundary, excluding lakes, using the ArcMap module of ArcGIS 10.3.1. We intersected these points with the Yellowstone predictive occupancy layer and classified 10 bins of increasing occupancy probability that each contained an equal number of random points (Boyce et al. 2002). Next, we calculated a Spearman-rank correlation between bin ranks of increasing occupancy probability and the area-adjusted frequency of mountain goat observations within each bin for each data type (professional aerial, professional ground, and volunteer ground; Boyce et al. 2002). If the model had useful predictive ability, we would expect a positive correlation between used locations and bins of increasing occupancy probabilities. We produced all graphs using R version 3.2.3 and the ggplot2 package (Wickham 2009, R Core Team 2015).

Comparison of volunteer and professional data

Simultaneous double-observer surveys were not possible within the scope of this study, presenting a potential challenge for assessing detection bias caused by observer error or survey type. To assess data collected by each survey platform (volunteer ground, professional ground, and professional aerial) for observer error, we examined the distribution of mountain goat group sizes recorded by each survey platform to determine whether certain group sizes were detected differently, due to observer classification (volunteer or professional) or survey type (aerial or ground). In addition, to assess whether data for any of the survey platforms compared differently to the alternative model extrapolations, we compared the location data of each survey platform independently to the Colorado and Yellowstone models to assess differences among survey platform types.

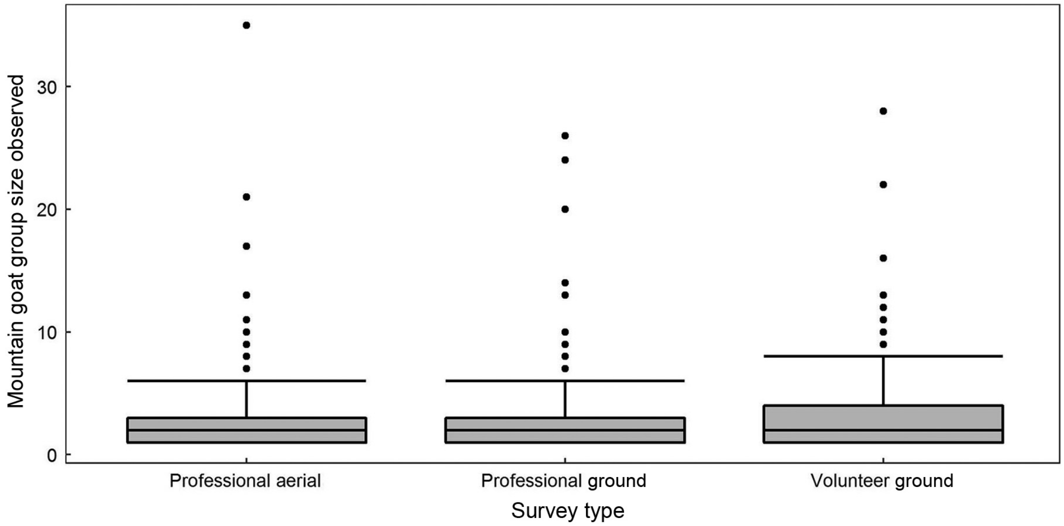

Fig. 1. Number of mountain goats observed within a group for mountain goat locations detected during aerial surveys by professionals (n = 229), ground surveys by professionals (n = 184), and ground surveys by volunteers (n = 200) in Waterton-Glacier International Peace Park.

Results

Data collection

Citizen scientists recorded 619 mountain goats at 200 locations, during 57 ground surveys and through incidental observations, from the second week of May to the third week of October, with observed group size ranging from 1 to 28 and averaging three mountain goats, with a SD of 3.4. Similarly, professionals recorded 549 mountain goats at 184 locations, during 45 ground surveys and through incidental observations, from the third week of May to the third week of September, with group size ranging from 1 to 26 and averaging three mountain goats, with a SD of 3.4. The Yellowstone model used data spanning from the third week of June to the second week of October for model development (DeVoe et al. 2015), but we retained all recorded locations for extrapolated model evaluation due to limited data availability. During summer aerial surveys by professionals over three different years, observers recorded 669 mountain goats at 229 locations, with group size ranging from 1 to 35 and averaging three mountain goats, with a SD of 3.7 (Fig. 1). All survey platforms observed a median group size of two mountain goats.

The most recent count data for Waterton Lakes National Park provided a minimum population estimate of 102 mountain goats in 2016. Using a range of detection probabilities reported in the literature for aerial surveys of mountain goats, the number of mountain goats in this area ranged from approximately 118, using a high detection probability of 0.9, to 193, using a low detection probability of 0.55. Applying a moderate detection of 0.7 provided an estimate of 151 mountain goats in Waterton Lakes National Park. Belt and Krausman (2012) extrapolated a population estimate for Glacier National Park of 1705–2349 using volunteer data and 1885–3269 using professional data.

Model extrapolation

The Yellowstone model extrapolated to Waterton-Glacier resulted in predicted mountain goat occupancy probabilities across the study area that ranged from 0.0 to 0.431 (Fig. 2). Using a suitability cutoff defined in the GYA of $\Psi \geq 0.0058$ (DeVoe et al. 2015), the Yellowstone model estimated that 2273 km^2 (50%) of Waterton-Glacier was suitable mountain goat habitat. The Colorado model estimated that 2551 km^2 (56%) of the study area was suitable mountain goat habitat. Thus, the extrapolated Yellowstone model predicted 278 km^2 (6%) less total area as suitable mountain goat habitat than the Colorado model. Approximately 82.81% of the Colorado model predicted

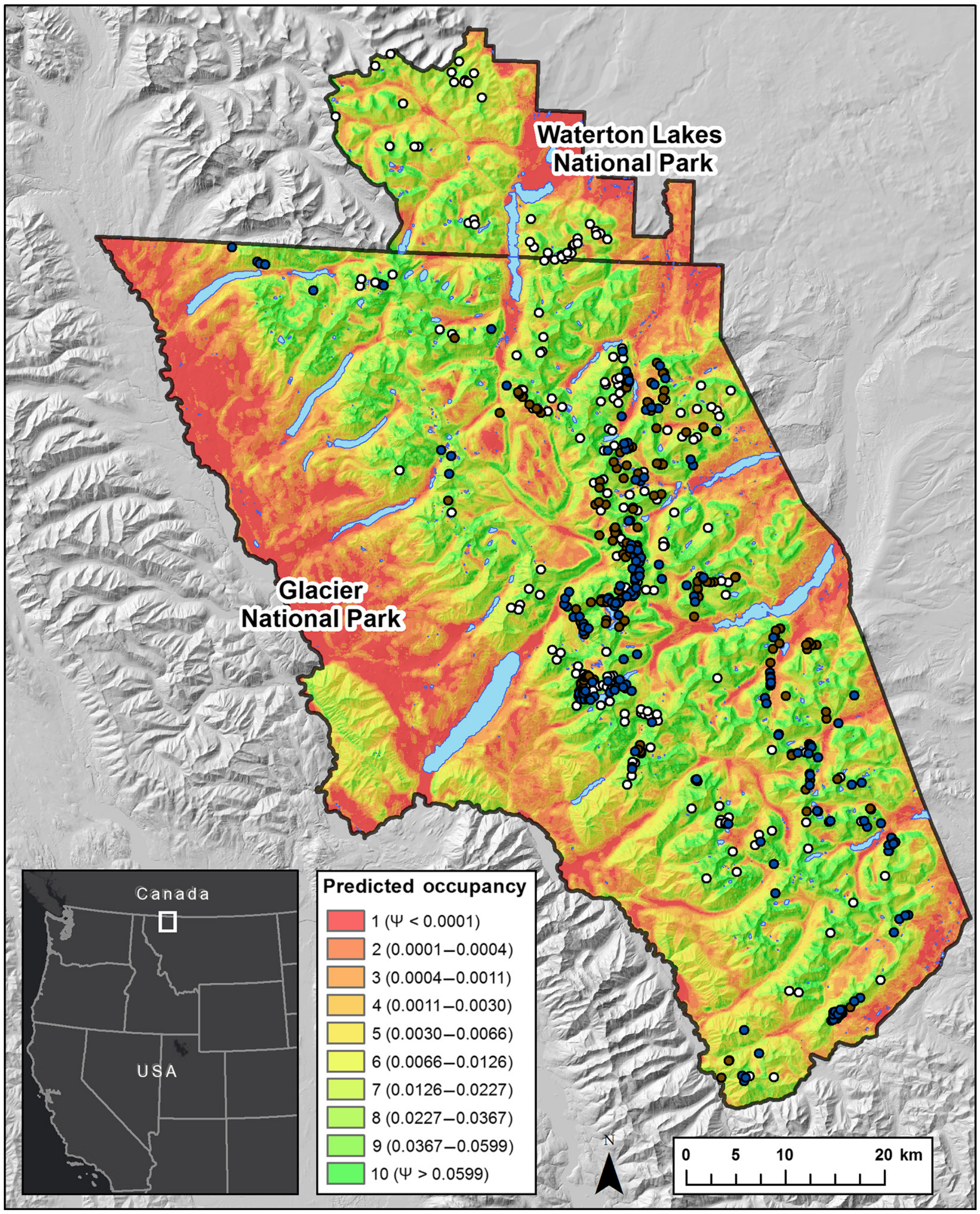

Fig. 2. Binned predicted mountain goat occupancy across Waterton-Glacier International Peace Park, based on an extrapolated model developed in the Greater Yellowstone Area (park boundaries shown as black lines). Mountain goat locations recorded by volunteers during ground surveys are blue points (n = 200), professionals during ground surveys are brown points (n = 184), and professionals during aerial surveys are white points (n = 229).

suitable area contained areas also predicted as suitable by the Yellowstone model across Waterton-Glacier. The area of the viewsheds surveyed from the ground by professionals and volunteers covered 8% of the predicted suitable area by the Colorado model and 9% of the predicted suitable area by the Yellowstone model. The total area of the viewsheds surveyed by professional and volunteer observers from the ground was predicted to be composed of 83% suitable mountain goat habitat, based on the Yellowstone model, and 89% suitable area, based on the Colorado model.

Table 2. Comparison of the number and percentage of observed mountain goat locations during aerial surveys by professionals (n = 229), ground surveys by professionals (n = 184), and ground surveys by volunteers (n = 200), located within suitable mountain goat areas predicted by two extrapolated predictive models developed in the Greater Yellowstone Area and Mt. Evans, Colorado.

Dataset	Yellowstone model		Colorado model	
	Number of locations	Percentage observed	Number of locations	Percentage observed
Professional aerial	226	98.7	227	99.1
Professional ground	176	95.7	177	96.2
Volunteer ground	194	97.0	197	98.5

Evaluation of extrapolated models

There was high consistency of observed mountain goat locations within suitable habitat predicted by both the Yellowstone and Colorado model extrapolations. When we compared observed mountain goat locations with suitable habitat area predicted by the Yellowstone model, 97.0% (n = 194) of volunteer ground, 95.7% (n = 176) of professional ground, and 98.7% (n = 226) of professional aerial locations were within predicted suitable habitat. When we compared observed mountain goat locations with suitable habitat area predicted by the Colorado model, 98.5% (n = 197) of volunteer ground, 96.2% (n = 177) of professional ground, and 99.1% (n = 227) of professional aerial locations were within predicted suitable habitat. Thus, a similar number of observations were located within predicted suitable areas for each extrapolated model, even though the Yellowstone model predicted a smaller area as suitable habitat (Table 2). For the Yellowstone model, we also compared the distribution of observed locations across binned predicted occupancy categories (Figs. 2 and 3). Observed locations for all data types were nonlinear and skewed right toward high-probability occupancy bins, and Spearman-rank correlations were significant and ranged from 0.898 to 0.964 (Fig. 3, Table 3).

Comparison of volunteer and professional data

The center and spread of the distribution of mountain goat group size observed per location were similar for all survey platforms (Fig. 1). The distribution of volunteer ground data was slightly more skewed to the right in comparison with the observations of the other survey platforms (Fig. 1). However, observer experience and survey platform did not impact extrapolated model evaluation results when locations were

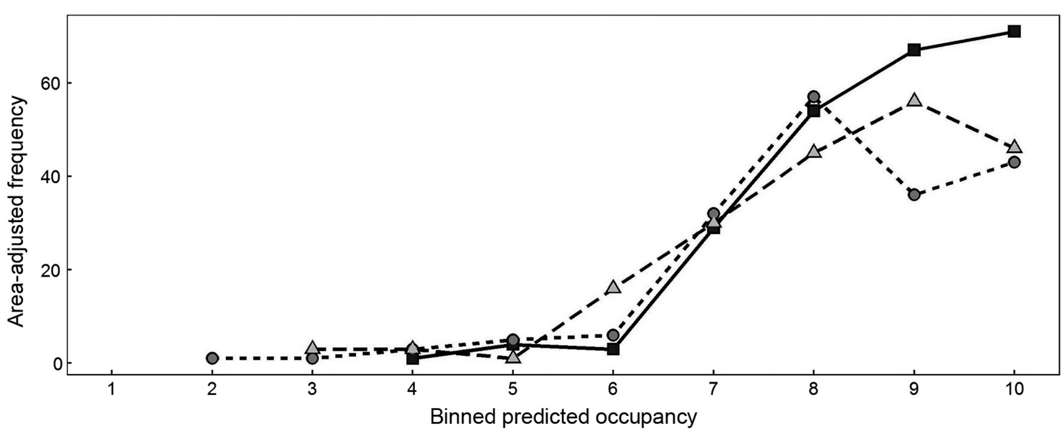

Fig. 3. Distribution of locations within predicted occupancy bins of sampling units in Waterton-Glacier International Peace Park for data detected during aerial surveys by professionals (n = 229), ground surveys by professionals (n = 184), and ground surveys by volunteers (n = 200). Solid squares for professional aerial observations, gray circles for professional ground, and gray triangles for volunteer ground.

Table 3. Spearman-rank correlations (r_s) between occupancy probability bin ranks and area-adjusted frequencies of mountain goat locations recorded during aerial surveys by professionals, ground surveys by professionals, and ground surveys by volunteers.

Dataset	r_s	P
Professional aerial	0.964	<0.003
Professional ground	0.946	<0.0002
Volunteer ground	0.898	<0.003

compared to the predicted suitable areas of the extrapolated Colorado and Yellowstone models. For all datasets, greater than 95% of observations were located within the predicted suitable areas for both models (Table 2). In regard to the Yellowstone model, Spearman-rank correlation and *P*-values for binned occupancy probabilities of the Yellowstone model were also similar across the observer categories (Table 3, Fig. 3). Professional aerial observations had a slightly greater Spearman-rank correlation (r_s = 0.964), followed by the professional ground (r_s = 0.946), and volunteer ground (r_s = 0.898).

Discussion

Both extrapolations of the Yellowstone and Colorado models predicted similar areas to be suitable mountain goat habitat across Waterton-Glacier. Over 95% of the observed mountain goat locations from volunteer ground, professional ground, and professional aerial survey platforms fell within the area predicted as suitable mountain goat habitat by both extrapolated models. When the Yellowstone model was validated in the GYA using aerial survey locations, 72.1% of points were within suitable habitat, 85.8% of points were within 100 m of predicted suitable areas, and 92.5% of points were within 200 m of predicted suitable areas (DeVoe et al. 2015). The Colorado model predicted 87% of mountain goat locations in the Colorado study area during model validation (Gross et al. 2002). Therefore, both extrapolated models were more effective in predicting mountain goat locations in Waterton-Glacier than in the areas where the models were originally developed. This observation may be because the mountain goat population in Waterton-Glacier is native, whereas non-native mountain goat populations were relatively recently introduced to the GYA and Colorado study areas. Habitat selection in non-native areas tends to be similar but not identical to that found in native range, as introduced vertebrates tend to occupy a subset of potential habitat in early stages of dispersal across non-native range (Strubbe et al. 2015). Mountain goat range expansion across the GYA has been relatively modest and limited over time, which suggests limited dispersal tendencies in contrast to wide range expansion achieved by another recently reintroduced mammal to the GYA, the wolf (*Canis lupus*; Smith et al. 2001, Flesch et al. 2016). In contrast, the mountain goat population in Waterton-Glacier likely occupies the vast majority of suitable habitat. In addition, the data used to validate the Yellowstone model in the GYA were collected over a period of 45 years throughout the species' range expansion history, which may have negatively affected the accuracy of validation for the Yellowstone model that was developed using occupancy data from 2011 to 2013 (DeVoe et al. 2015).

In general, the Yellowstone model was more informative than the Colorado model in predicting mountain goat locations in Waterton-Glacier. First, the suitable area predicted by the Yellowstone model captured a similar number of mountain goat locations as the Colorado model, even though suitable area predicted by the Yellowstone model covered 6% less total area. Without recording locations of mountain goat absence, we were not able to directly assess the proportion of predicted suitable areas that were incorrectly classified inactive sites, but this result suggests that the Yellowstone model was more effective in minimizing over-prediction of suitable area than the Colorado model. This was likely because the Yellowstone model accounted for multiple environmental covariates, whereas the Colorado model was based solely on slope. Secondly, the ability of the Yellowstone model to predict a range of occupancy probability values was an additional advantage for evaluating and deriving insight from the extrapolated model. Mountain goat locations recorded by all observation types (professional ground, volunteer ground, and professional aerial) in Waterton-Glacier demonstrated significant Spearman-rank correlations and were skewed toward sampling units that were predicted by the Yellowstone model to have high or very high relative occupancy probability

(Fig. 3, Table 3). These Spearman-rank correlation values were comparable to those used for evaluation of an extrapolated grizzly bear model, with Spearman-rank correlation values that ranged from 0.782 to 0.951 (Boyce et al. 2002). By using randomly generated locations to define occupancy probability bins, it was unlikely that the overlap of most mountain goat locations with high-probability areas was impacted by a disproportionate number of grid cells classified as high probability. The effectiveness of the Yellowstone model in Waterton-Glacier suggests that its application may be useful for predicting areas with high potential for mountain goat occupancy in other mountainous regions that are similar to the northern GYA. We recognize that our use of percentage overall accuracy may serve as a limitation in accuracy assessment of our model extrapolations, as very wide coverage would inflate the accuracy score. However, the area predicted as suitable by the Yellowstone model covered 50% of the study area and 6% less total area than the Colorado slope model, and thus, we suggest that its accuracy was likely not due to excessively wide coverage. Used-available data can be effective in modeling species distribution (Elith et al. 2006), and typical classification success metrics used to evaluate used–unused datasets, such as sensitivity analysis, are not applicable for the used-available data collected during our study (Boyce et al. 2002). This limited our evaluation options. Derivation of an independent suitability cutoff and use of a sensitivity analysis for model evaluation may be a more advantageous approach for model assessment if sufficient data are available for both model training and testing. We did not have sufficient data available to derive a suitability cutoff independent of the Yellowstone model in this study.

An additional advantage of the Yellowstone model was that DeVoe et al. (2015) accounted for imperfect detection of mountain goat groups through estimation of detection probability (MacKenzie et al. 2002, DeVoe et al. 2015). We expect that detection bias of our location data used to evaluate the Yellowstone model in Waterton-Glacier was similar to that used to build the Yellowstone model because our survey protocols were similar and based on visual observations. Terrain complexity and canopy cover can impact visual detection probability of mountain goats during aerial surveys, but DeVoe et al. (2015) found little evidence that these environmental variables impacted detection probability during ground surveys (Poole 2007, Rice et al. 2009). However, DeVoe et al. (2015) suspected that estimated detection probabilities and consequent occupancy probabilities of the Yellowstone model may have been biased, due to multiple influences on estimated detection probability (MacKenzie 2006, Rota et al. 2009). Possible influences included non-independent observations of the same mountain goat group in different grid cells due to mountain goat group movement during a survey, double observers not simultaneously viewing the same moving mountain goat group, and detection probability not accounted for due to unintentional exclusion of related covariates (MacKenzie 2006, DeVoe et al. 2015). Thus, occupancy probability values predicted by the Yellowstone model across Waterton-Glacier may be somewhat biased due to possible detection probability bias during original model development. DeVoe et al. (2015) concluded that while modeling of detection probability was a limitation of the study, the occupancy model was validated by independent data and thus likely a reasonable prediction of occupancy probability. By extrapolating the Yellowstone model to Waterton-Glacier rather than building a novel local occupancy model, we were not able to independently estimate detection probability in our study area. However, we expect that our evaluation of the extrapolated model was informative for predicting mountain goat habitat selection, due to similar survey protocols for data collection and use of an occupancy model that accounted for imperfect detection.

To determine whether there were significant differences between results from aerial vs. ground surveys, we compared volunteer and professional ground observations with aerial observations. Spatially, both volunteer and professional ground observations were concentrated near citizen science survey sites, whereas aerial observations were more dispersed. Aerial data may have included more locations in high occupancy probability grid cells due to protocol for some flights in which observers largely searched areas subjectively deemed to be high-probability observation areas, whereas surveys from the ground involved a visual search of all visible terrain near the survey site. Potential error in location data may have been introduced to aerial surveys, due to lag time in recording a GPS point for observations while the aircraft was in motion, especially in fixed-wing aircraft. In addition, aerial surveys were conducted in different years than ground surveys, but we expect that the distribution of mountain goats was likely comparable among different years, as the species is native to Waterton-Glacier. In addition, DeVoe et al. (2015) also validated their model using locations from aerial surveys largely collected in different years than the years the occupancy surveys were conducted. Our ground surveys may also have been impacted by inaccuracies in recorded point locations on maps, which could vary by observer skill and experience. However, we did not find any indication that spatial differences in sampling or these potential types of observer error impacted overall extrapolated model evaluation results, as the number of locations within suitable areas and the distributions of locations across occupancy probabilities were similar among professional aerial, volunteer ground, and professional ground survey platforms (Table 2, Fig. 3).

To assess whether location data collected by volunteers could be used to evaluate the extrapolated models, we compared results from volunteer ground and professional aerial/ground surveys. As expected, volunteers were able to collect more mountain goat locations during ground surveys than professionals, as a greater number of volunteers could dedicate more total time toward traveling to survey sites and conducting surveys. In regard to mountain goat group size observed per location, the mean, median, and spread of group size observed were similar for all datasets (Fig. 1). Thus, volunteer ground observations were not skewed toward larger group sizes in comparison with other datasets. Lack of volunteer bias toward observation of larger animal groups is encouraging for population estimate and trend research using citizen science, as well as extrapolated model evaluation. The lack of bias was contrary to our expectation that volunteers may be less likely than professionals to detect small groups, as reported by Belt and Krausman (2012). That study used a double-observer approach and found that mean detection probability for volunteers was lower and more variable than that of professionals and was positively influenced by increasing group size. In our study, we were unable to quantify the percentage of small groups of goats that were not detected by volunteers because a double-observer approach was not used. Additionally, we suspect that volunteer detection of small groups may have been higher because the subset of volunteers who self-selected to report location data due to their familiarity with topographic maps were mostly volunteers who had greater experience with the program. An extension of this work would be to use a double-observer approach to assess whether observer experience, measured by the number of surveys an observer conducted, would influence ability to detect smaller groups of mountain goats. Such an assessment would allow program managers to estimate the amount of observer experience necessary to maximize detection probability of mountain goats.

Volunteer ground observations were similarly captured in areas classified as suitable habitat predicted by both models in comparison with the professional datasets (Table 2). When we compared the datasets to the Yellowstone model range of occupancy probabilities, volunteer ground observations had the lowest Spearman-rank correlation value (Table 3). This may indicate that experience slightly impacted accuracy of observations or that volunteers were more likely to conduct surveys in more accessible and thus, lower probability areas. Nevertheless, volunteer ground observations produced a significant Spearman-rank correlation ($r_s = 0.898$; $P < 0.003$) to effectively evaluate the extrapolated Yellowstone model. Even if additional sources of volunteer bias were undetected, the similarity of volunteer and professional results implied that the level of data quality necessary for extrapolated model evaluation was comparable between professional and volunteer observers.

Our results suggest that citizen science ground data can be used to evaluate extrapolated wildlife distribution models, because results from observations by citizen scientists were similar to those by both professional ground and aerial surveyors. Therefore, in this case, collecting citizen science data was an effective method to evaluate extrapolated models with limited resources and served as an alternative to building a novel occupancy model using professionally collected datasets only. In regard to data collection efficiency, three multi-day aerial surveys were able to collect a similar number of locations as citizen scientists over three years. While aerial surveys could be seen as a preferred alternative to the considerable amount of time required for citizen science data collection methods, aerial surveys are generally more expensive, require more permitting, and are heavily reliant on favorable weather. Perhaps most importantly, citizen science efforts provide an opportunity for public engagement and stewardship during the scientific process. Consequently, a citizen science approach may be useful in other study areas with interested members of the general public and research questions that involve highly visible and easily identified species in regions accessible by road, foot, or water transportation. The approach used for model assessment should be determined by the objective of the research and planned use of the model, such that prediction error is within acceptable limits for the research goals (Schamberger and O'Neil 1986, Fielding and Bell 1997). Once an extrapolated predictive occupancy or habitat model is evaluated, the product could be used for species and ecosystem monitoring, management, and planning.

Researchers should also consider potential variation in environmental variables integrated into wildlife distribution models. Some environmental variables likely varied slightly from year to year during our survey effort, such as vegetation available for mountain goat use due to variation in timing and levels of precipitation. However, if a researcher's objective is not concerned with short-term fluctuations in occupancy, collection of a smaller number of observations per season over multiple years can serve to decrease the coefficient of variation for the estimated trend in occupancy and thus provide more accurate occupancy estimates (MacKenzie 2005). Since we collected a limited number of observations per year and our study objective was concerned with general assessment of mountain goat distribution models, conducting surveys over multiple years was a reasonable method to provide accurate information to evaluate the extrapolated models. In addition, we expected that there was variation in use of areas by mountain goats throughout the timeframe that we defined as the summer season (May through October). In general, it is acceptable to collect data over a timeframe where the species is present at random points in time throughout the season, if researchers acknowledge that the observed used area may be greater in extent than the area likely occupied over a shorter timeframe (MacKenzie et al. 2004, MacKenzie 2005).

Our results also suggest that citizen science data may be useful for building occupancy models if they do not already exist from other study systems. If we had sufficient local data to produce a site-specific model, we expect the local model would decrease potential over-prediction of suitable habitat, which we were not able to evaluate in this study without unused data. Local data can provide more information regarding habitat associations than an extrapolated model, because local heterogeneity in species distribution, caused by factors such as differing biogeographic history, may not be captured by environmental variables in other locations (Osborne et al. 2007, Hothorn et al. 2011). We recommend that other citizen science programs use citizen science data to build occupancy models and evaluate this in other study areas with different wildlife species. However, there are limitations to this approach. Despite similar results among citizen scientist and professional observations, citizen science data collection can be significantly less efficient than professional data collection. DeVoe et al. (2015) collected 505 mountain goat locations from professional ground surveys over three years to develop a novel occupancy model, whereas volunteers in Waterton-Glacier collected only 200 locations over three years of effort. Our dataset was more limited than that of DeVoe et al. (2015) in part because not all volunteers felt comfortable using topographic maps or aerial photos to record locations, resulting in specific mountain goat UTM locations recorded for only 37% of all volunteer ground surveys where mountain goats were observed. Therefore, developing a model using citizen science data may require more time to acquire additional points from proficient observers or more intensive training sessions to increase participant proficiency and participation in accurate spatial data collection. Alternatively, technology such as GPS, geotagging, or apps that record spatial data and allow volunteers to directly record locations at the exact location of observation could be used to develop models involving species observed at close range, rather than species observed from a distance, such as mountain goats that are often on inaccessible cliffs. Use of these technologies would reduce the potential for observer error and the need for additional training on use of maps for spatial data collection. We were not able to compare the accuracy of estimated locations with known locations in our study, due to logistical and safety concerns that often prevent access to mountain goat habitat. However, during professional and volunteer ground surveys we sought to emulate the level of accuracy achieved by DeVoe et al. (2015), by using a comparable field protocol for survey sessions and the same grid cell size in model development. In general, researchers should ensure that species location data are accurate at spatial scales similar to the grid cell size used for model development and analysis.

When using citizen science data for either model development or extrapolated model evaluation, it is important to minimize potential sources of bias associated with volunteer data, through careful attention to protocol development and training. While most observations in Waterton-Glacier obtained by volunteers were during systematic surveys, many programs may choose to use opportunistically collected observations. Opportunistic observations recorded by volunteers must be interpreted with caution, as these data can be biased toward more easily accessible areas, such as roadways (Broman et al. 2014). Volunteers may also bias their survey efforts and observations to habitat types that they subjectively deemed high-probability habitat. For example, volunteers often associate mountain goats with cliffs in Waterton-Glacier. To ensure that volunteers captured locations across the predicted range of occupancy probabilities, instructors at volunteer training sessions in Waterton-Glacier emphasized that citizen scientists must survey all visible terrain at survey sites, not just areas subjectively considered to be high-probability terrain.

Models developed or evaluated using low-cost citizen science data can have significant value for wildlife professional and natural resource managers. Managers could use an evaluated habitat suitability layer for planning and decision-making regarding issues that may impact natural resources, such as high visitation, roadway construction, and airspace use. For example, some conservation areas experience frequent air traffic tourism, and some species, including mountain goats, are impacted by and have energetic cost associated with air traffic noise disturbance (Côté 1996). Models evaluated with citizen science can also be used to determine how the distribution of different species potentially overlap, which can be informative in management of non-native species (Gormley et al. 2011). However, habitat suitability can change over time due to disturbances, such as human development and climate change (Walther et al. 2002, Araújo et al. 2005). These issues can reduce the accuracy of previously developed or evaluated models, necessitating efforts to build new models that account for these impacts. The advantage of citizen science in this context is that the low-cost approach can allow data to be collected over longer timeframes than traditional research studies, enabling changes in habitat suitability to be tracked over time and contribute toward development or evaluation of new models. Additionally, citizen science coordinators can consider developing survey protocols that allow for inferences regarding multiple population metrics, such as demography and population trends, to increase the types of information available regarding species of interest. To meet this need with limited resources, researchers can use carefully designed citizen science projects to achieve multiple population monitoring objectives, as the Glacier National Park Citizen Science Program used volunteer data to both estimate population trends and evaluate extrapolated spatial models.

Acknowledgments

Funding was provided by the Roundtable on the Crown of the Continent and the Glacier National Park Conservancy through the Crown of the Continent Research Learning Center and Glacier National Park. We would like to thank all of the citizen scientists who volunteered their time to collect mountain goat location data, including Bob Chinn, Susan Clothier, Steve Gniadek, Alex Guppy, Heather Jameson, Brian Johnson, Clayton Mauritzen, Robina Moyer, Cyndi Mullings, Todd Myse, Evan O'Connor, Shayla Paradeis, Kathy Ross, Jeannie Siegler, Barbara Summer, Kelly Van Frankenhuyzen, Dave Wollner, and many others. Glacier National Park staff Lisa Bate, Tara Carolin, Laura Luther, Josh Martin, Dennis Osborne, and Andrew Lahr contributed to this project. Barb Johnston and Robin Steenweg provided aerial survey data from Waterton Lakes National Park. Jesse DeVoe, Blake Lowrey, Robert Garrott, and Carson Butler of the Greater Yellowstone Area Mountain Ungulate Project, Tabitha Graves, and Barb Johnston provided technical expertise and comments to improve this manuscript. Richard Menicke of Glacier National Park and Shannon Blackadder of the Crown Managers Partnership provided GIS technical support. National Elevation Dataset (NED) raster and Landsat 7 SLC-on scene courtesy of the U.S. Geological Survey. We would like to thank two anonymous reviewers who provided helpful comments to improve this manuscript. Montana State University Library provided funding for open access publication of this manuscript.

Literature Cited

Araújo, M. B., R. G. Pearson, W. Thuiller, and M. Erhard. 2005. Validation of species–climate impact models under climate change. Global Change Biology 11:1504–1513.

Beard, K. H., N. Hengartner, and D. K. Skelly. 1999. Effectiveness of predicting breeding bird distributions using probabilistic models. Conservation Biology 13:1108–1116.

Belt, J. J., and P. R. Krausman. 2012. Evaluating population estimates of mountain goats based on citizen science. Wildlife Society Bulletin 36:264–276.

Bender, L. C., W. L. Myers, and W. R. Gould. 2003. Comparison of helicopter and ground surveys for North American elk *Cervus elaphus* and mule deer *Odocoileus hemionus* population composition. Wildlife Biology 9:199–205.

Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. Schmiegelow. 2002. Evaluating resource selection functions. Ecological Modelling 157:281–300.

Broman, D. J. A., J. A. Litvaitis, M. Ellingwood, P. Tate, and G. C. Reed. 2014. Modeling bobcat (*Lynx rufus*) habitat associations using telemetry locations and citizen-scientist observations: Are the results comparable? Wildlife Biology 20:229–237.

Cobb, M. A., H. Helling, and B. Pyle. 2012. Summer diet and feeding location selection patterns of an irrupting mountain goat population on Kodiak Island, Alaska. Biennial Symposium of the Northern Wild Sheep and Goat Council 18:122–135.

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. 1996. Mountain goat responses to helicopter disturbance. Wildlife Society Bulletin 24:681–685.

Crown Managers Partnership. 2012. Crown Closure in the Crown of the Continent Ecosystem. https://www.sciencebase.gov/catalog/item/565f398ee4b071e7ea544533

DeVoe, J. 2015. Occupancy modeling of non-native mountain goats in the greater Yellowstone area. Master of Science in Fish and Wildlife Management. Montana State University, Bozeman, Montana, USA.

DeVoe, J. D., R. A. Garrott, J. J. Rotella, S. R. Challender, P. J. White, M. O'Reilly, and C. J. Butler. 2015. Summer range occupancy modeling of non-native mountain goats in the greater Yellowstone area. Ecosphere 6:1–20.

Dickinson, J. L., B. Zuckerberg, and D. N. Bonter. 2010. Citizen science as an ecological research tool: challenges and benefits. Annual Review of Ecology, Evolution, and Systematics 41:149–172.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40:677–697.

Elith, J., et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Environmental Systems Resource Institute (ESRI). 2010. ArcGIS 10.1. ESRI, Redlands, California, USA.

Festa-Bianchet, M., and S. D. Côté. 2008. Mountain goats. Island Press, Washington, D.C., USA.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38–49.

Flesch, E. P., et al. 2016. Range expansion and population growth of non-native mountain goats in the greater Yellowstone area: challenges for management. Wildlife Society Bulletin 40:241–250.

Fox, J. L. 1983. Constraints on winter habitat selection by the mountain goat (*Oreamnos americanus*) in Alaska, dissertation, University of Washington, Seattle, Washington, USA.

Gonzalez-Voyer, A., M. Festa-Bianchet, and K. G. Smith. 2001. Efficiency of aerial surveys of mountain goats. Wildlife Society Bulletin (1973–2006) 29:140–144.

Gormley, A. M., D. M. Forsyth, P. Griffioen, M. Lindeman, D. S. L. Ramsey, M. P. Scroggie, and L. Woodford. 2011. Using presence-only and presence–absence data to estimate the current and potential distributions of established invasive species. Journal of Applied Ecology 48:25–34.

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. 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.
Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135:147–186.
Gutzwiller, K. J., and W. C. Barrow Jr. 2001. Bird-landscape relations in the Chihuahuan Desert: coping with uncertainties about predictive models. Ecological Applications 11:1517–1532.
Hirzel, A. H., and G. Le Lay. 2008. Habitat suitability modelling and niche theory. Journal of Applied Ecology 45:1372–1381.
Hirzel, A. H., G. Le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199:142–152.
Hobbs, N. T. 2003. Challenges and opportunities in integrating ecological knowledge across scales. Forest Ecology and Management 181:223–238.
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.
Homer, C., et al. 2007. Completion of the 2001 national land cover database for the conterminous United States. Photogrammetric Engineering and Remote Sensing 73:337.
Hothorn, T., J. Müller, B. Schröder, T. Kneib, and R. Brandl. 2011. Decomposing environmental, spatial, and spatiotemporal components of species distributions. Ecological Monographs 81:329–347.
MacKenzie, D. I. 2005. What are the issues with presence-absence data for wildlife managers? Journal of Wildlife Management 69:849–860.
MacKenzie, D. I. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Academic Press, Burlington, Massachusetts, USA.
MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.
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.
MacKenzie, D. I., J. A. Royle, J. A. Brown, J. D. Nichols, and W. L. Thompson. 2004. Occupancy estimation and modeling for rare and elusive populations. Pages 149–171 *in* W. Thompson, editor. Sampling rare or elusive species: concepts, designs, and techniques for estimating population parameters. Island Press, Washington, D.C., USA.
Mayor, S. J., D. C. Schneider, J. A. Schaefer, and S. P. Mahoney. 2009. Habitat selection at multiple scales. Ecoscience 16:238–247.
McAlpine, C. A., J. R. Rhodes, M. E. Bowen, D. Lunney, J. G. Callaghan, D. L. Mitchell, and H. P. Possingham. 2008. Can multiscale models of species' distribution be generalized from region to region? A case study of the koala. Journal of Applied Ecology 45:558–567.
Morisette, J. T., et al. 2008. 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.
Nagy, C., K. Bardwell, R. F. Rockwell, R. Christie, and M. Weckel. 2012. Validation of a citizen science-based model of site occupancy for eastern screech owls with systematic data in suburban New York and Connecticut. Northeastern Naturalist 19: 143–158.
Osborne, P. E., G. M. Foody, and S. Suárez-Seoane. 2007. Non-stationarity and local approaches to modelling the distributions of wildlife. Diversity and Distributions 13:313–323.
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. L. Ballard, and E. T. Jones. 2006. Broadening participation in biological monitoring: handbook for scientists and managers. General Technical Report, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA.
Poole, K. G. 2007. Does survey effort influence sightability of mountain goats *Oreamnos americanus* during aerial surveys? Wildlife Biology 13:113–119.
R Core Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org
Rice, C. G., K. J. Jenkins, and W.-Y. Chang. 2009. A sightability model for mountain goats. Journal of Wildlife Management 73:468–478.
Rota, C. T., R. J. Fletcher, R. M. Dorazio, and M. G. Betts. 2009. Occupancy estimation and the closure assumption. Journal of Applied Ecology 46: 1173–1181.
Schamberger, M., and J. O'Neil. 1986. Concepts and constraints of habitat model testing. Wildlife 2000: Modeling Habitat Relationships of Terrestrial

Vertebrates. University of Wisconsin Press, Madison, Wisconsin, USA.
Smith, D. W., K. M. Murphy, and D. S. Guernsey. 2001. Yellowstone wolf project: annual report, 2000. National Park Service, Yellowstone Center for Resources, Yellowstone National Park, Mammoth, Wyoming, USA.
Stillman, R. A., and A. F. Brown. 1994. Population sizes and habitat associations of upland breeding birds in the south Pennines, England. Biological Conservation 69:307–314.
Strubbe, D., O. Beauchard, and E. Matthysen. 2015. Niche conservatism among non-native vertebrates in Europe and North America. Ecography 38:321–329.
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.
van Strien, A. J., C. A. M. van Swaay, and T. Termaat. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. Journal of Applied Ecology 50:1450–1458.
Veitch, A., E. Simmons, M. Promislow, D. Tate, M. Swallow, and R. Popko. 2002. The status of mountain goats in Canada's Northwest Territories. Northern Wild Sheep and Goat Council 13:49–62.
Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. Nature 416: 389–395.
Wickham, H. 2009. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, New York, USA.
Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385–397.
Yung, L. 2007. Citizen monitoring and restoration: volunteers and community involvement in wilderness stewardship. US Department of Agriculture Forest Service Proceedings P-49, Rocky Mountain Research Station, Fort Collins, Colorado, USA.
Zanini, F., J. Pellet, and B. R. Schmidt. 2009. The transferability of distribution models across regions: an amphibian case study. Diversity and Distributions 15:469–480.

Supporting Information

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1638/full