Field validation of a distribution model for the yellow-bellied marmot (*Marmota flaviventris)*

Student: Erin Voss
Mentors: Tiffany Armenta and Dan Blumstein
Independent Research & Course
Rocky Mountain Biological Laboratory
Summer 2014

Abstract

Species distribution modeling is a powerful ecological tool that combines remotely sensed data with known presence locations to yield habitat suitability predictions that can be applied to a variety of hypotheses and conservation planning efforts, and can also be used on a smaller scale to examine individual species' habitat preferences. Here, species distribution modeling techniques are applied to the yellow-bellied marmot, *Marmota flaviventris*, a semi-fossorial, facultatively social rodent that is found across high elevations in much of western North America. *M. flaviventris* is known to prefer microhabitat patches with characteristics related to predator detection and avoidance; good visibility and the number of large rocks are the best predictors of marmot presence and persistence. To examine the habitat preferences of *M. flaviventris*, Maximum Entropy was used with environmental variables including slope, visibility, elevation, local geology, and vegetation, as well as 75 presence locations, to produce and verify a model of yellow-bellied marmot habitat preference in the East River Valley surrounding Gothic, Colorado. Statistical verification demonstrated that the model is a significant predictor of marmot presence, and field validation of the model yielded moderate success (72.9% correct predictions, or 97/133 sites). The importance of visibility and slope to the model indicate that predator-related variables are indeed the most important predictors of marmot presence, although the number of large rocks, as well as other potentially relevant biotic factors, were not included in the model. While MaxEnt modeling is generally undertaken on a larger scale, this model's success indicates that fine resolution modeling with the goal of field validation and examination of small-scale environmental variations is plausible, especially if environmental predictors are carefully selected and informed by a thorough consideration of the ecology of the focal species.

Introduction

Variation in environmental, geographic, and climatic factors determine species distribution, diversity, and abundance across various landscapes (Ball 1974, Cox & Moore 2010). Early ecologists recognized and attempted to describe the correlations they saw between species distributions and environmental variables (e.g. Hittinka 1963, cited in Guisan & Thuiller 2005), but their lack of powerful statistical techniques and accurate spatial data made such efforts difficult (Elith & Leathwick 2009). The rise of geographic information systems (GIS), in combination with advances in statistics and computing power, have enabled the creation of sophisticated species distribution models (SDMs), which produce habitat suitability maps based on species' presence/absence data and remotely-sensed environmental information (Elith & Leathwick 2009). SDMs have widespread applications across ecology and conservation biology (Guisan & Thuiller 2005), and have become vital tools for biologists. They are frequently used to test biogeographical, ecological, and evolutionary hypotheses (e.g. Anderson *et al.* 2002, Graham *et al.* 2004), and can also be helpful in examining the ecological effects of specific environmental alterations such as invasive species, land use, and climate change (e.g. Gormley *et al.* 2011, Peterson 2003, Araújo *et al.* 2006b). For example, Thomas *et al.* (2004) used species distribution modeling to predict that climate change will cause 15-37% of sampled species to be "committed to extinction" by 2050. Additionally, SDMs are used to support conservation planning and species management, and have been utilized successfully to identify areas of high biodiversity and to aid in the selection of biological reserves (Ferrier 2002, Moilanen 2005, Costa *et al.* 2009). Conservation biologists also employ species distribution models to identify potential habitat locations for rare or threatened species, as in the case of Raxworthy *et al.* (2003), who combined locality data from museums with GIS layers to create species distribution

models for 11 poorly-known chameleon species in Madagascar. Furthermore, SDMs can be valuable as practical tools for field ecologists, because such models can efficiently predict the presence of target organisms in specific locations. Here, I employed data from a long-term study at the Rocky Mountain Biological Laboratory in combination with species distribution modeling techniques to create and test the accuracy of a localized habitat suitability model for the yellow-bellied marmot, *Marmota flaviventris*.

Yellow-bellied marmots (YBMs) are facultatively social, semi-fossorial rodents that dwell in the montane meadows of the western United States and Canada (Frase and Hoffmann, 1980). They inhabit burrows within a home range of 0.13-1.02 ha (Armitage 1975), and are active for only 4-5 months annually. While yellow-bellied marmots are efficient hibernators (Armitage *et al.* 2003), hibernation-related mortality is high, and over-winter survival is dependent in part on individuals' ability to consume large quantities of food and store fat during their brief active season (Armitage 1976). Predation by a range of terrestrial and aerial predators during the summer season is a second major source of mortality (Van Vuren 2001). As microhabitats often vary in quality and safety (e.g. Lyons 2005, Prins & Iason 1989), habitat selection can have consequences for both food availability and predation risk, and thus affects not only individuals' future survival and reproductive success, but also population regulation on a larger scale (Morris 2003). While differing food availability is linked to variation in local demography for many organisms (e.g. Frey-Roos *et al.* 1995, Bergallo & Magnusson 1999), Blumstein *et al.* (2006) have shown that for yellow-bellied marmots, habitat characteristics related to predator detection and avoidance are better predictors of the presence and persistence of colonies. They discovered that yellow-bellied marmots require areas with high visibility (>50

meters) and large rocks for their safety, while variables related to food availability were not significant predictors of habitat suitability.

To test Blumstein *et al.*'s (2006) conclusions regarding the habitat preferences of yellow-bellied marmots, I constructed a predictive species distribution model of yellow-bellied marmot habitat preferences using marmot presence data and remotely sensed GIS layers to quantify the environmental parameters of marmot microhabitats. I hypothesized that this model is an accurate and effective tool for use in forecasting marmot occurrence when field-tested in the East River Valley near Gothic, Colorado.

Methods

Study area and population

M. flaviventris have been studied since 1962 at the Rocky Mountain Biological Laboratory in Gothic, Colorado; thus, we have extensive historic records on the species' presence in the surrounding East River Valley, as well as a detailed understanding of its habitat preferences. I used historic marmot presence data from this long-term study, in combination with my own observations of current marmot colony locations, to produce a model of YBM habitat suitability.

Data analysis and model production

To construct a model of YBM habitat preferences that would function as a predictive tool, I utilized the popular geographic distribution modeling software, MaxEnt (Phillips *et al*. 2006), which uses a machine learning algorithm to predict species occurrence based on known occurrences and environmental variables. I provided MaxEnt with 75 sites yielded from current and historic YBM presence-only data; 60 were utilized as training data for MaxEnt, while the

remaining 15 were set aside for testing the model. Additionally, I supplied MaxEnt with potentially relevant environmental predictors of marmot presence (per Blumstein *et al*. 2006), including precipitation, viewshed, geology, slope, vegetation type, and elevation values taken from 10-meter digital elevation maps (DEMs) and ArcGIS software (ESRI). Precipitation was later excluded due to its unrealistically strong impact on the model, as it was not expected to have a strong impact on YBM distribution at such a small scale. Viewshed, which is a GIS tool used for determining general visibility from specific locations, was calculated in ArcGIS using Shuttle Radar Topography Mission (SRTM) 90m elevation maps and 30 m DEMs obtained from RMBL to determine which areas were visible from each 90 m point. Slope was obtained from the 10m DEMs, and vegetation was obtained from a Bureau of Land Management aerial survey of the area, provided by RMBL. All environmental variables had 10m x 10m cells. These input files served as training data for the algorithm; MaxEnt then produced a logistic output map depicting the probability of YBM presence in each 10-meter pixel throughout the East River Valley.

Model validation and verification

Field-testing the predictive performance of the MaxEnt model took place in the East River Valley during July and August 2014. Specifically, I identified the pixels that produced the strongest prediction of presence or absence, as measured by the highest and lowest logistic output values yielded by MaxEnt, that were known to occur in relevant habitat types. Areas of dense shrub growth or tree cover, as well as those covered with bare rock, water, snow, or those used for agricultural purposes, were excluded because of the low likelihood of locating a burrow (Carey 1985, pers. obs.) From the pixels that were located in areas of grass/forb, alpine

grass/forb, alpine meadow, and sagebrush/low shrub cover, the pixels that fell within the highest and lowest 10% logistic output values were identified. 100 presence and 100 absence sites were randomly selected from among these pixels for field validation. Pixels that were inaccessible or dangerous to reach were excluded. I then navigated to each pixel of interest using GPS coordinates, and inspected a 100 m radius around the site thoroughly for burrows, feces, or other visual indications of YBM presence. I examined rocky areas within each site closely, because YBM burrows are often located in or around sloped, rocky outcrops in alpine meadows (Blumstein *et al.* 2006). If I did not find any signs of marmot occurrence within a 100-meter radius of the site, it was counted as an observed absence. To analyze the efficacy of the model, I calculated the percentages at which it correctly predicted the presence and absence of marmots using this field-tested independent data set. To verify the model statistically, MaxEnt was used to calculate AUC, or area under the receiver-operator curve, and TSS, the true skill statistic (Allouche *et al.* 2006), values were calculated, and a bootstrapping test was performed with 100 replicates, in which 15 of the 75 known presence sites were randomly subsampled for use as test locations for each iteration.

Results

Statistical verification

MaxEnt yielded a model of the East River Valley (Fig. 1, Appendix 1 contains variable coefficients and maximum and minimum values) with an AUC (area under the receiver-operator

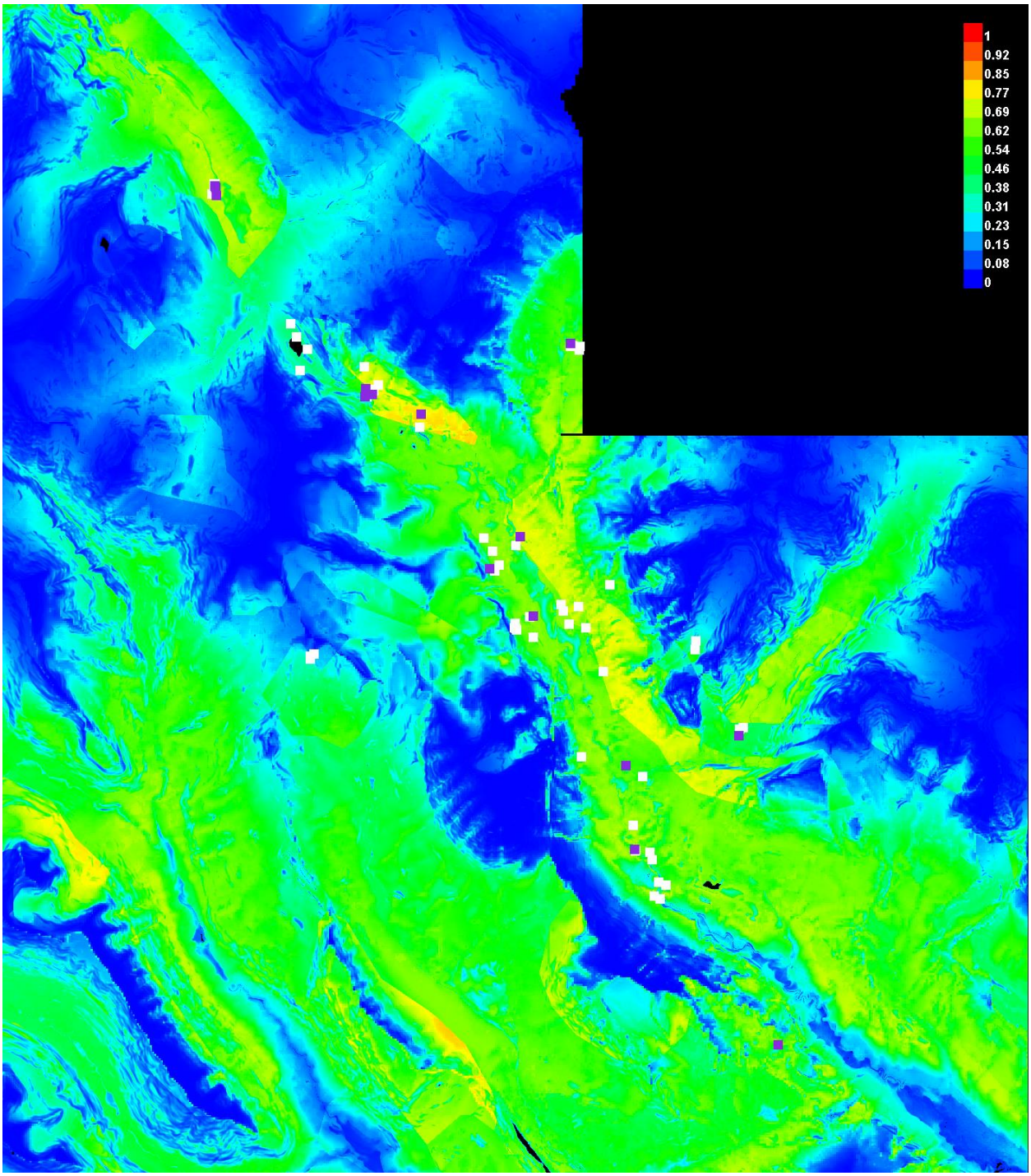

Figure 1. Maximum Entropy model of *M. flaviventris* in the East River Valley and environs. Cooler areas indicated low probability of YBM presence, while warmer colors indicate especially high probability of marmot presence. The black area was excluded due to a lack of data regarding vegetative cover for that region. White points indicate locations used as the 60 training points for the model (known current, historically occupied, or permanent marmot colony locations). 20% of known colony locations were set aside to use as test data (violet points).

curve) of 0.814 for training data (60 locations), and 0.841 for test data (15 locations), with a standard deviation of 0.035, and a regularized gain of 0.477. An AUC of 0.7-0.9 is considered good (Swets 1988), in comparison, a uniformly distributed, random model would have an AUC of 0.5. Analysis of the test and background probabilities (Appendix 2 for raw, cumulative, and logistic output values for training and test locations), when combined with a presence/absence threshold of the tenth percentile of the training presences, yielded a TSS (Allouche *et al.* 2006) of .5482, which is considered good (Monserud & Leemans 1992). TSS is perhaps a better representation of model performance, as AUC can be heavily influenced by the prevalence of presence points and the total geographic extent of the model (Lobo *et al.* 2008). MaxEnt also calculated several common thresholds for presence vs. absence using cumulative and logistic output values, and examined corresponding omission rates for both training and test data. It then used these values to calculate binomial probabilities and 1-sided p-values (Table 1), which varied based on the choice of presence vs. absence threshold, but all were significant ($p<0.05$). Values ranged from 0.02895 for the fixed cumulative value 1 threshold, to 9.935×10^{-7} for the maximum test sensitivity plus specificity threshold, indicating that the model has a significantly better ability to predict marmot presence than a random model would. The 100-replicate bootstrapping test yielded a mean training AUC of 0.836, indicating that the model chosen for field-validation is a good representative of mean model performance. Elevation, viewshed, and slope were the most important environmental predictors, while geology and vegetation contributed less than 5% to model function (Table 2). Response curves and jackknife tests of training and test gains support these results (Figures 2 and 3), with elevation performing the best when used alone.

Cumulative threshold	Logistic threshold	Description	Fractional predicted area	Training omission rate	Test omission rate	P-value
1.000	0.073	Fixed cumulative value 1	0.790	0.000	0.000	2.895E-2
5.000	0.190	Fixed cumulative value 5	0.622	0.017	0.000	8.141E-4
10.000	0.274	Fixed cumulative value 10	0.519	0.017	0.000	5.399E-5
4.728	0.183	Minimum training presence	0.630	0.000	0.000	9.71E-4
16.175	0.356	10 percentile training presence	0.436	0.100	0.067	7.928E-5
36.763	0.487	Equal training sensitivity and specificity	0.267	0.267	0.133	2.054E-6
42.672	0.514	Maximum training sensitivity plus specificity	0.230	0.267	0.200	4.908E-6
43.163	0.516	Equal test sensitivity and specificity	0.227	0.300	0.200	4.283E-6
39.147	0.499	Maximum test sensitivity plus specificity	0.251	0.267	0.133	9.935E-7
4.728	0.183	Balance training omission, predicted area and threshold value	0.630	0.000	0.000	9.71E-4
5.098	0.192	Equate entropy of thresholded and original distributions	0.620	0.017	0.000	7.667E-4

Table 1. Common presence vs. absence thresholds and corresponding omission rates of training and test locations from presence values. As the number of test values was <25, binomial probabilities were calculated exactly, and 1-sided p-values indicated in the final column are for for the null hypothesis that test points are predicted no better by the model than would a uniform distribution with the same fractional predicted area. The "balance" threshold minimizes 6 * training omission rate + 0.4*cumulative threshold +1.6 * fractional predicted area (MaxEnt 3.3.3k). All p-values are significant (<0.05), indicating that by any choice of threshold, this model predicts YBM presence better than the null (random) model.

Variable	Percent contribution	Permutation importance
Elevation	70.9	54.9
Viewshed	13.7	18.7
Slope	10.8	16.5
Geology	4.4	9.9
Vegetation	0.2	0

Table 2. Estimates of relative contributions of environmental variables to the Maximum Entropy model.

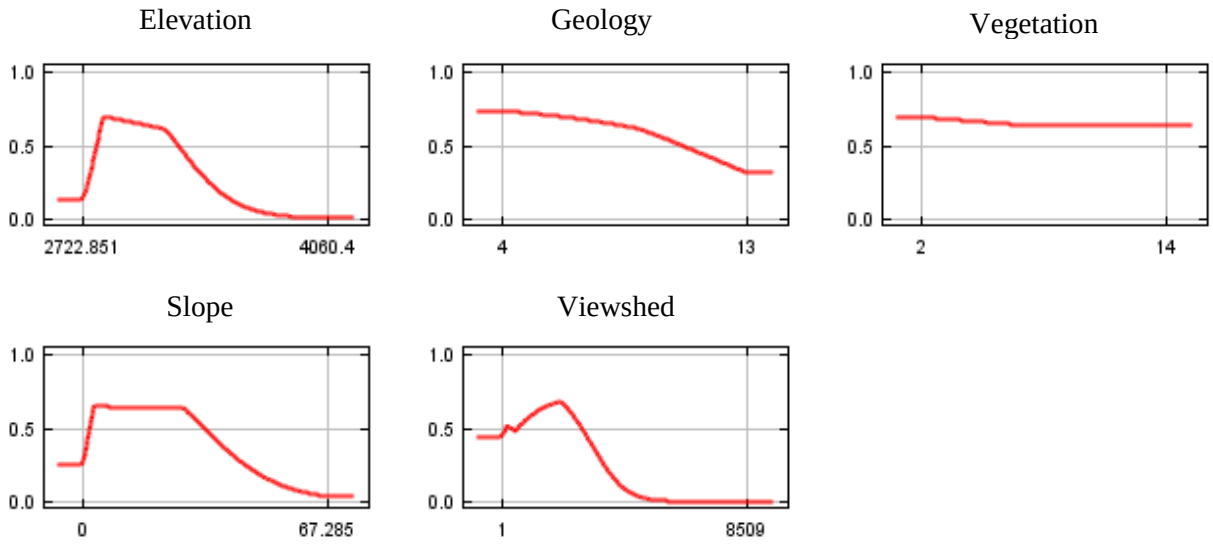

Figure 2. Response curves for each environmental variable. The curves show how logistic prediction values change as each environmental covariate is altered, when all other variables are kept constant.

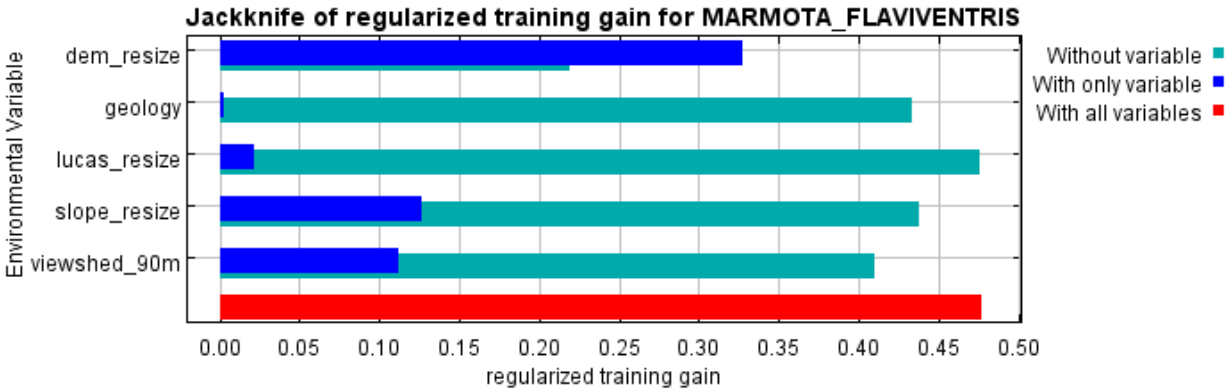

Figure 3. Jackknife test of variable importance, measured using regularized training gain.

Field validation

Field-testing the model showed a moderate amount of success at prediction of marmot presence and absence. I evaluated 133 sites throughout the East River Valley. Out of 72 high probability sites tested in the field, 38 showed signs of marmot presence (52.8 % success). 59 of 61 low probability sites visited in the field showed no signs of marmot presence (96.7 % success). The

combined prediction accuracy of the model for the East River Valley was 97/133, or 72.9% correct (Figure 4). Of the 40 locations where I detected signs of marmot presence, 38 were predicted presence sites (95.0%). Conversely, of the 93 sites were no signs of marmots were discovered, 59 were predicted absences (63.4%).

	Success	Failure	Total	% Correct
Predicted Presences	38	34	72	52.8 %
Predicted Absences	59	2	61	96.7%
Overall	97	36	133	72.9%

Table 3. Observed successes and failures at presence vs. absence sites test locations.

In sum, the model produced using MaxEnt (Figure 1) was significant according to all presence/absence threshold tests (Table 3), and correctly predicted marmot occurrence for 72.9% of field-tested locations (Figure 4). Response curves and jackknife tests performed on the environmental predictors included in the model indicate that elevation, viewshed, and slope were the most significant covariates in predicting marmot presence (Table 2, Figures 2 and 3), while vegetation and local geology were less important (Table 2).

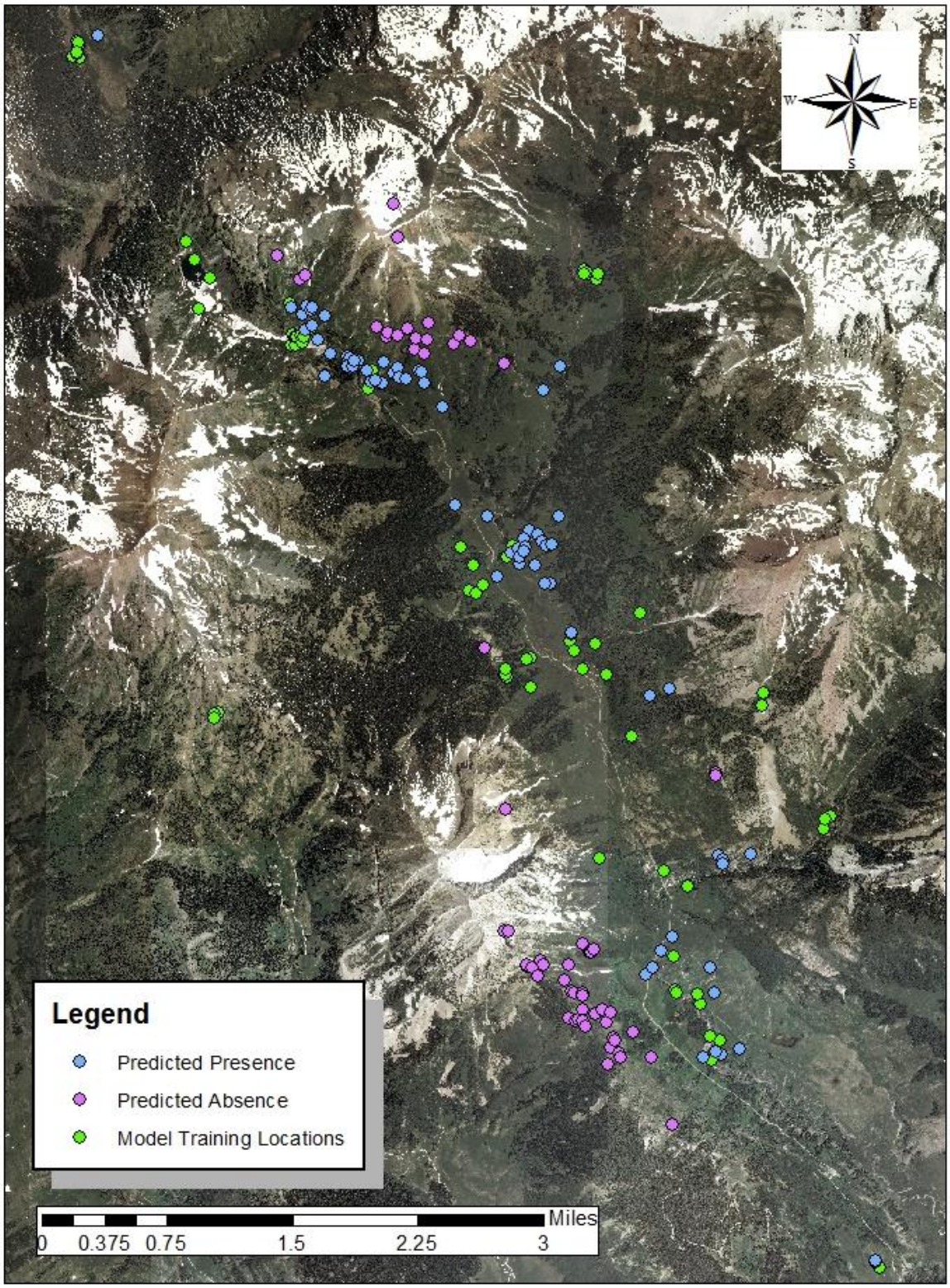

Figure 4. Map of GPS locations of training sites, as well as field validated high and low probability sites, as predicted by the model.

Discussion

Statistical verification of the model indicated that it was very successful, and predicted testing points significantly better than a null distribution would, but field validation yielded only moderate success. While marmots were rarely seen in unsuitable habitats, and the model appeared to be very effective at predicting absence locations, they were frequently absent from areas the model defined as suitable habitat space, with only 33/64 high probability locations yielding observed presences. Part of this high rate of commission is potentially due to the fact that during the summer of 2014 when I was conducting field validation, the marmot population was much smaller than it had been in recent years. Snow pack from the winter of 2013-2014 lasted longer than usual, resulting in a drop in the study population from approximately 170 to 70 individuals due to over-winter mortality (D.T. Blumstein, unpubl. data). Due to their decreased population size, this summer's YBM population likely does not occupy all of the valley's potentially suitable habitat space. The model may have captured some of these appropriate but unoccupied locations, which yielded observed absences when field-tested and registered as failures to predict marmot presence.

The model's relative success at predicting YBM presence and absence in specific areas within the valley indicates that marmot occurrence differs across the variety of available habitats. While marmots disperse widely (Van Vuren & Armitage 1994b), the model's success at predicting YBM absence indicates that it has captured some aspects of the landscape that are unappealing or dangerous for *M. flaviventris*. If YBMs do settle in these undesirable habitat patches, they may experience local extinction at higher rates than other marmot colonies, as Blumstein *et al.* (2006) posit that some habitat patches, if colonized, are more likely to go extinct than others due to their lower habitat quality. In order to interpret the model's results, the

importance of specific variables to the model was assessed to determine which local environmental differences separated undesirable locations from the areas where *M. flaviventris* persist.

The importance of viewshed to the model is consistent with Blumstein *et al.*'s (2006) conclusions regarding the importance of predation-related characteristics to marmot presence and persistence. They found that visibility was the most important environmental variable associated with permanent marmot colonies. For this model, response curves indicate that logistic output values, and thus probability of YBM presence, improved as viewshed increased, although areas with especially high viewshed values had extremely low probabilities of marmot presence (Figure 2). This pattern most likely occurred because especially high viewshed areas tend to be exposed mountaintops, which are often unsuitable for long-term inhabitance. As habitat selection can play a major role in individual survival and larger-scale population regulation (Morris 2003), the importance of higher viewshed values to the model indicates that visibility, and by extension, predator detection, is a critical to *M. flaviventris* survival and prevention of local extinctions. Yellow-bellied marmot behavior also indicates that YBMs prioritize visibility and vigilance against predation; they spend 40-60% of the day aboveground, sitting and looking (Armitage *et al*. 1996), and they remain watchful during 33% of the time spent in a foraging bout (D. T. Blumstein, unpubl. data, cited in Blumstein *et al.* 2006).

Response curves indicate that slope was also important to the model, as areas with a slight to moderate slope (approximately 5-30˚) have the highest probability of YBM presence (Figure 2), providing further evidence that environmental factors related to predator detection are important to the local persistence of marmots. Blumstein *et al.* (2006) suggest that areas with a slight slope often have good visibility in multiple directions, while steeper or flatter sites may

have poorer visibility, and may therefore be less attractive to marmots looking to avoid predation.

The importance of viewshed to the model indicates not only that predator detection and avoidance is central to local YBM survival, but also that predation risk varies across the landscape, and that it may be lower in areas with good visibility. However, some aspects of the "landscape of fear," which Laundré *et al.* (2010) define as the variation in predation risk across time and space, can be difficult to quantify using remotely sensed data. For example, Blumstein *et al.* (2006) found the number of large rocks at each site to be a significant predictor of marmot presence. Large rocks are useful to YBMs in predator avoidance and detection, as marmots spend much of their time being vigilant from the tops of rocks and dig their burrows underneath large rocks to discourage fossorial predators (Svendsen 1976) such as badgers (*Taxidea taxus*). Unfortunately, the GIS layers available in the study area can not quantify the number and/or size of rocks across the landscape, so I could not include this variable in the model. While the model did include a variable with information regarding the underlying geology of the region, it contributed less than one percent to model function, indicating that it did not capture any information relevant to YBM habitat preferences.

The model's omission of the number of large rocks at each site indirectly addresses one of the fundamental challenges of SDMs and ecological niche modeling: incorporating the effects of biotic interactions on species into their distribution models (e.g. Araújo *et al.* 2014). Guisan & Thuiller (2005) recognize a need for distributional analyses that include environmental predictors relating to the presence and absence of known competitors (e.g. Anderson *et al.* 2002) or other participants in biotic interactions with the focal species. However, as obtaining accurate and comprehensive biotic data often requires labor-intensive collection efforts, they are often

excluded from SDMs in favor of climatic and spatial variables. Nevertheless, through the use of remotely sensed data, this model was able to capture the marmots' preference for high-visibility microhabitat patches, which demonstrates that detection and avoidance of predation risk structures YBM habitat selection. Inclusion of data directly related to the yellow-bellied marmot's predators and competitors could further enhance the model's performance, and could help researchers examine linkages between the marmot's interspecific interactions and its distribution across the landscape. In addition to predation, other biotic interactions thought to be potentially significant for marmot habitat suitability include the presence of previously established burrow sites, as well as proximity to other marmot colonies, as females at satellite sites have been shown to experience lower fitness than those at well-established colonies (Van Vuren & Armitage 1994a).

The irrelevance of vegetation to the model also supported Blumstein *et al.*'s (2006) findings, as they concluded that food-related variables, including vegetative biomass and ground cover, are not significantly related to habitat suitability for YBMs. Changes in the vegetative type only caused changes in logistic output of less than 5%, indicating that vegetation did not play a large part in determining model function (Figure 2). While specific vegetation types may contain nutrients that are vital to marmot survival, *M. flaviventris* generally does not lack for food; in comparison with golden marmots (*M. caudata aurea*), yellow-bellied marmots have access to an order of magnitude more aboveground vegetation at the peak of the growing season (Blumstein & Foggin, 1997), which may indicate why variation in food availability does not correspond to variation in local YBM demography. Rather, weather-related events such as late summer droughts and long winters have been associated with local extinction events and changes in local demography (Armitage 2003), and over-winter mortality is a frequent source of local

extinctions. YBMs are efficient hibernators (Armitage *et al.* 2003), however, young marmots tend to lose more mass at colder temperatures, leading to increased mortality among pups. As such, Blumstein *et al.* (2006) suggest that hibernacula structure may contribute to local extinction, but were not able to include measures of hibernation suitability in their study, as characteristics related to burrow structure are difficult to quantify, and their null sites lacked hibernacula for comparison. In an attempt to consider some of the larger-scale environmental variables that could affect hibernation-related mortality, local geology was included in this model. However, the insignificance of local geology to the model indicates that the variable did not capture any meaningful information about environmental factors that could affect the construction of effective and safe hibernacula. This model was unable to demonstrate whether any environmental factors related to hibernacula structure are relevant to YBM habitat selection.

The critical importance of elevation to the model was surprising, as *M. flaviventris* inhabits a wide elevation range above 2000 m across North America (Frase & Hoffmann 1980). Variable jackknife tests of training gain indicate that when each variable was taken individually and used as the sole predictor of marmot presence, elevation performed best. Similarly, the model performed most poorly when elevation was removed (Figure 3). The extremely large contribution of elevation may be an artifact of sampling bias, as many of the sites included in the long-term RMBL marmot study fall within similar elevations around Gothic because they are easy to access and observe. Efforts were made to explore the valley and surrounding mountains in order to discover marmot colonies in other areas for inclusion in the model, but areas farther away from the road and the main expanse of the valley may still be more poorly represented. While sampling bias is a common problem in species distribution modeling (Syfert *et al*. 2013, e.g. Leitão *et al.* 2011), methods for correcting sampling bias are complex, and require further

development (Araújo *et al*. 2006a, Fourcade *et al.* 2013). As such, no efforts to correct sampling bias were undertaken in this study.

Other difficulties with using MaxEnt may have affected the efficacy of the model at predicting marmot habitat sites. MaxEnt modeling is generally undertaken on a larger scale, often encompassing the entire range of the species under examination (e.g. Abba *et al.* 2012). At such scales, remotely sensed environmental data is often available only at much larger operational scales, even at the scope of 50x50 km grids (e.g. Bakkenes *et al* 2006). Meanwhile, 1x1 km output grids are often considered fine-scale (Seo *et al.* 2009). At these scales, MaxEnt is capable of capturing the entire niche of the focal species. In contrast, this model utilized environmental data of 10x10 m resolution, and the entire study area was approximately 15x15 km; the use of very fine-scale, accurate environmental layers allowed for the creation of a more detailed model. Wiens (2002, cited in Guisan & Thuiller 2005) suggests that identifying the appropriate scale on which to conduct ecological modeling is a frequent problem; for this project, the use of a smaller study extent and finer resolution data allowed us to target microhabitat-level differences across the landscape. At the scale of the East River Valley, the model was constructed with a different goal in mind, one with a bottom-up, species' viewpoint approach: to examine *M. flaviventris'* microhabitat preferences rather than its realized niche. Additionally, small-scale modeling made field validation of the model viable, allowing for the examination of the model's efficacy through an independent data set in addition to use of statistical tools such as AUC, TSS and bootstrapping. Many models do not attempt field validation (e.g. Abba *et al.* 2012, Bergallo & Magnusson 1999), as it would be too time consuming and laborious on large scales, and instead rely on test statistics that may be influenced by a variety of inaccuracies and errors (Lobo *et al.* 2008) to assess model function.

Species distribution modeling has many practical uses in the fields of ecology and wildlife biology, but as discussed above, many argue that SDMs lack linkages to biological theory (Guisan & Thuiller 2005). In this study, in an effort to utilize the wide range of research on YBM life-history traits and habitat preferences, environmental variable selection was informed by previous studies of yellow-bellied marmot microhabitat patch features. Selection of environmental variables for inclusion is a critical part of the modeling process, and rigorous and careful selection of ecologically relevant predictors can increase the degree of explanation and insight that models are capable of (Elith and Leathwick 2009, Mac Nally 2000). As such, predictors thought to be irrelevant at such a small scale, including climatic variables related to temperature and precipitation, were omitted from the model, while potentially significant variables such as viewshed, vegetation, and geology were included. However, as discussed above, remotely sensed data fails to capture some aspects of the landscape that are most applicable to YBM habitat selection, including the number of large rocks and the potential for good hibernacula. Perhaps this is where some of the difficulties of linkages with ecological theory emerge; while usage of remotely sensed data and geographic information systems is both informative and practical, such data may lack the capability to express information that would be biologically relevant.

In conclusion, the model was a moderate success at predicting marmot presence and absence locations when field-tested, and significantly successful when tested using presence/absence thresholds and known presence locations. The importance of variables such as viewshed and slope indicate that predator detection and avoidance structure marmot habitat selection, while food-related variables do not, in support of Blumstein *et al.*'s (2006) conclusions. While SDMs are generally used for larger, niche-based examinations of species

distribution, it is clear that smaller -scale modeling can be useful for testing hypotheses surrounding habitat preferences, especially because field-validation of the model is possible in addition to statistical verification. While no single variable is a significant predictor of YBM presence and persistence, the combination of a variety of environmental factors allows examination of the range of habitats available to marmots, and identification of the ones they find most appealing. Future studies could focus on incorporating more potentially relevant factors, such as the number of large rocks and the distribution of predators and competitors, as well as examining a wider range of elevations to determine whether habitat selection patterns hold across a larger distribution of YBM-occupied patches. In general, application of ecological knowledge and hypotheses to distributional modeling, along with careful selection and consideration of environmental variables and rigorous model testing, can be effective tools for exploring species' habitat preferences, as well as the environmental and geographic variables that structure territory selection.

Acknowledgements

I would like to thank Tiffany Armenta and Dan Blumstein for their guidance throughout the model construction and testing processes, Lucas Moyer-Horner for providing vegetation data and sharing information about his modeling efforts, and the Rocky Mountain Biological Laboratory for making this project possible.

Literature Cited

Abba, AM, MF Tognelli, VP Seitz, JB Bender, & SF Vizcaíno. (2012). Distribution of extant xenarthans (Mammalia: Xenartha) in Argentina using species distribution modeling. *Mammalia* 76: 123-136.

Allouche, O, A Tsoar, & R Kadmon. (2006). Assessing the accuracy of species distribution models: prevalence, kappa, and the true skill statistic (TSS). *Journal of Appl. Ecol.* 43: 1223-1232.

Anderson, RP, AT Peterson, & M Gómez-Laverde. (2002). Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos* 98: 3-16.

Araújo, MB, & A Guisan. (2006a). Five (or so) challenges for species distribution modeling. *Journal of Biogeography* 33: 1677-1688.

Araújo, MB, RG Pearson, W Thuiller, & M Erhard. (2006b). Validation of species-climate impact models under climate change. *Global Change Biology* 11: 1-10.

Araújo, CB, LO Marcondes-Machado, & GC Costa. (2014). The importance of biotic interactions in species distribution models: a test of the Eltonian noise hypothesis using parrots. *Journal of Biogeography* 41: 513-523.

Armitage, KB. (1975). Social behavior and population dynamics of marmots. *Oikos* 26: 341-354.

Armitage, KB. (1976). Seasonal changes in weights of marmots. *American Midland Naturalist* 96: 36-51.

Armitage, KB. (2004). Badger predation on yellow-bellied marmots. *American Midland Naturalist* 151: 378-387.

Armitage, KB, DT Blumstein, & BC Woods. (2003). Energetics of hibernating yellow-bellied marmots (*Marmota flaviventris*). *Comparative Biochemistry and Physiology* 134: 101-114.

Armitage, KB, CM Salsbury, EL Barthelmess, RC Gray, & A Kovach. (1996). Population time budget for the yellow-bellied marmot. *Ethol., Ecol., & Evol.* 8: 7-95.

Bakkenes, M, B Eickhout, & R Alkemade. (2006). Impacts of different climate stabilization scenarios on plant species in Europe. *Global Environmental Change* 16: 19-28.

Ball, IR. (1974). Nature and formulation of biogeographical hypotheses. *Systematic Zoology* 24: 407-430.

Bergallo, HG, & WE Magnusson. (1999). Effects of climate and food availability on four rodent species in southeastern Brazil. *Journal of Mammology* 80: 472-486.

Blumstein, DT, & JM Foggin. (1997). Effects of vegetative variation on weaning success, overwinter survival, and social group density in golden marmots (*Marmota caudata aurea*). *Journal of Zoology* 243: 57-69.

Blumstein, DT, A Ozgul, V Yovovich, DH Van Vuren, & KB Armitage. (2006). Effect of predation risk on the presence and persistence of yellow-bellied marmot (*Marmota flaviventris*) colonies. *Journal of Zoology* 270: 132-138.

Carey HV. (1985). The use of foraging areas by yellow-bellied marmots. *Oikos* 44: 273-279.

Costa, GC, C Nogueira, RB Machado, & GR Colli. (2009). Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hot spot. *Biodivers. Conserv*. 2009: 1-17.

Cox, CB, & PD Moore. (2010). *Biogeography: an ecological and evolutionary approach*. 8th ed. Hoboken, New Jersey: John Wiley & Sons, Inc.

Elith, J, & JK Leathwick. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst*. 40: 677-697.

Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Syst. Biol.* 51: 331-363.

Fourcade, Y, JO Engler, D Rödder, & J Secondi. (2013). Mapping species distributions with MAXENT using a geographically bias sample of presence data: performance assessment of methods for correcting sampling bias. *PLoS* 9: n.p.

Frase, BA, & RS Hoffmann. (1980). *Marmota flaviventris*. *Mammalian Species* 135: 1-8.

Frey-Roos, F, PA Brodmann, & HU Reyer. (1995). Relationships between food resources, foraging patterns, and reproductive success in the water pipit, *Anthus* sp. *spinoletta. Behavioral Ecology and Sociobiology* 6: 287-295.

Gormley, AM, DM Forsyth, P Griffioen, M Lindeman, DSL Ramsey, MP Scroggie, & L Woodford. (2011). Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species. *Journal of Appl. Ecol.* 48: 25-34.

Graham, CH, SR Ron, JC Santos, CJ Schneider, & C Moritz. (2004). Integrating phylogenetic and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58: 1781-1793.

Guisan, A, & W Thuiller. (2005). Predicting species distribution: offering more than simple habitat models. *Ecol. Letters* 8: 993-1009.

Laundré, JW, L Hernández, & WJ Ripple, (2010). The landscape of fear: the ecological implications of being afraid. *The Open Ecology Journal* 3: 1-7.

Leitão, PJ, F Moreira, & PE Osborne. (2011). Effects of geographical data sampling bias on habitat models of species distributions: a case study with steppe birds in southern Portugal. *International Journal of Geographical Information Science* 25: 439-453.

Lyons, JE. (2005). Habitat-specific foraging of prothonotary warblers: deducing habitat quality. *The Condor* 107: 41-49.

Mac Nally, R. (2000). Regression and model-building in conservation biology, biogeography, and ecology: the distinction between—and reconciliation of—"predictive" and "explanatory" models. *Biodiversity and Conservation* 9: 655-671.

Moilanen, A. (2005). Reserve selection using nonlinear species distribution models. *The American Naturalist* 165: 695-706.

Monserud, RA, & R Leemans. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling* 62: 275-293.

Morris, DW. (2003). Toward an ecological synthesis: a case for habitat selection. *Oecologia* 136: 1-13.

Peterson, AT. (2003). Predicting the geography of species' invasions via ecological niche modeling. *Q. Rev. Biol.* 78: 419-433.

Phillips, SJ, RP Anderson, & RE Schapire. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231-259.

Prins, HHT, & GR Iason. (1989). Dangerous lions and nonchalant buffalo. *Behaviour* 108: 262 296.

Ramirez-Villegas, J, F Cuesta, C Devenish, M Peralvo, A Jarvis, & CA Arnillas. (2014). Using species distribution models for designing conservation strategies of Tropical Andean biodiversity under climate change. *Journal for Nature Conservation*, in press.

Raxworthy, CJ, E Martinez-Meyer, N Horning, RA Nussbaum, GE Schneider, MA Ortega Huerta, & AT Peterson. (2003). Predicting distributions of known and unknown reptile species in Madagascar. *Nature* 426: 837-841.

Seo, C, JH Thorne, L Hannah, & Wilfried Thuiller. (2009). Scale effects in species distribution models: implications for conservation planning under climate change. *Biology Letters* 5: 39-43.

Swets, JA. (1988). Measuring the accuracy of diagnostic systems. *Science* 240: 1285-1293.

Svendsen, GE. (1976). Structure and location of burrows of yellow-bellied marmot. *The Southwestern Naturalist* 20: 487-494.

Syfert, MM, MJ Smith, & DA Coomes. (2013). The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS ONE* 8: np.

Thomas, CD, A Cameron, RE Green, M Bakkenes, LJ Beaumont, YC Collingham, BFN Erasmus, M Ferreira de Siqueirea, A Grainger, L Hannah, L Hughes, B Huntley, AS van Jaarsveld, GF Midgley, L Miles, MA Ortega-Huerta, AT Peterson, OL Phillips, & SE Williams. (2004). Extinction risk from climate change. *Nature* 427: 145-148.

Van Vuren, DH, & KB Armitage. (1991). Duration of snow cover and its influence on life history variation. *Canadian Journal of Zoology* 69: 1755-1758.

Van Vuren, DH, & KB Armitage. (1994a). Reproductive success of colonial and noncolonial female yellow-bellied marmot (*Marmota flaviventris*). Journal of Mammology 75: 950-955.

Van Vuren, DH, & KB Armitage. (1994b). Survival of dispersing and philopatric yellow-bellied marmots: what is the cost of dispersal? *OIKOS* 69: 179-181.

Appendix 1. Model coefficients (1st value), minimum values (2nd value), and maximum values (3rd value).

```
dem_resize, 0.0, 2722.85107421875, 4060.39990234375,
geology, 0.0, 4.0, 13.0,
lucas_resize, 0.0, 2.0, 14.0,
slope_resize, 0.0, 2.47279298491776E-4, 67.28468322753906,
viewshed_90m, 0.0, 1.0, 8509.0,
dem_resize^2, -1.7071417768886954, 7413917.972374201, 1.6486847366953135E7,
geology^2, -1.314532577168307, 16.0, 169.0,
viewshed_90m^2, -16.372009279773653, 1.0, 7.2403081E7,
'dem_resize, -2.007990156355097, 3181.4415283203125, 4060.39990234375,
`dem_resize, -2.5208639932538386, 2722.85107421875, 2836.45947265625,
'dem_resize, -0.2930425768131503, 3170.113525390625, 4060.39990234375,
'slope_resize, -1.4819584866430995, 27.213879585266113, 67.28468322753906,
'dem_resize, -1.1147335909271079, 3169.964111328125, 4060.39990234375,
`slope_resize, -1.7360577556157684, 2.47279298491776E-4, 3.200950026512146,
'dem_resize, -1.3280612346902567, 3166.5845947265625, 4060.39990234375,
'slope_resize, -2.4414818749680287, 27.74373435974121, 67.28468322753906,
`lucas_resize, 0.08766625185177224, 2.0, 6.5,
`slope_resize, 0.03573056324340886, 2.47279298491776E-4, 13.431439876556396,
`viewshed_90m, -0.13663580951388424, 1.0, 2036.5,
'geology, -0.5044911769921115, 9.0, 13.0,
`viewshed_90m, -0.716729970616207, 1.0, 2037.5,
`viewshed_90m, 0.08642913108056736, 1.0, 465.5,
`viewshed_90m, -0.5820614821591441, 1.0, 2050.5,
`viewshed_90m, 0.27237747135384915, 1.0, 470.5,
`viewshed_90m, -0.4371269575105955, 1.0, 2051.5,
`viewshed_90m, 0.2912924126167372, 1.0, 471.5,
`viewshed_90m, -0.3416450756022805, 1.0, 193.5,
`dem_resize, -0.2830287944736194, 2722.85107421875, 2837.2999267578125,
`lucas_resize, 0.166514974796529, 2.0, 7.5,
`slope_resize, 0.01386725855285356, 2.47279298491776E-4, 13.383665084838867,
`viewshed_90m, -0.39134345302861634, 1.0, 2244.5,
`slope_resize, 0.034476189017905105, 2.47279298491776E-4, 13.385924816131592,
linearPredictorNormalizer, -1.1409828935907134,
densityNormalizer, 1466.4992866646126,
numBackgroundPoints, 10060,
entropy, 8.73825109693661,
```

Appendix 2. Model predictions for test and training values.

X	Y	Test or train	Raw prediction	Cumulative prediction	Logistic prediction
324816	4318770	train	5.45E-04	99.30112065	0.772659271
324195	4319261	train	4.89E-04	98.10238117	0.753092025
324253	4319341	train	4.73E-04	97.76474609	0.746734856
324157	4319281	train	4.07E-04	95.46869236	0.717445034
321998	4322000	train	3.75E-04	93.98440693	0.700523919
328097	4312530	train	3.49E-04	92.51409243	0.685228383
322016	4322102	train	3.45E-04	92.02480595	0.682541341
321998	4321988	train	3.40E-04	91.37955153	0.679632832
321969	4321969	train	3.36E-04	91.00159329	0.676838822
321998	4322108	train	3.25E-04	89.77656345	0.669307105
328041	4312337	train	3.09E-04	88.03653392	0.658641123
327751	4313308	train	3.07E-04	87.64106472	0.656816677
327341	4315415	train	2.89E-04	85.25729944	0.643517268
325816	4317061	train	2.85E-04	83.96936276	0.639732623
326998	4316311	train	2.75E-04	81.28512695	0.631765717
326140	4317149	train	2.75E-04	81.22312018	0.631601883
325913	4316878	train	2.70E-04	79.75599999	0.627504026
326334	4316162	train	2.69E-04	79.48229025	0.626922237
326872	4316063	train	2.69E-04	79.27093046	0.62646797
326793	4316243	train	2.67E-04	78.84783766	0.62507114
326758	4316336	train	2.53E-04	74.01026238	0.612307559
327783	4312954	train	2.41E-04	69.51807225	0.600049868
327025	4319879	train	2.32E-04	66.51658492	0.591483019
326890	4319872	train	2.30E-04	65.89142847	0.589624168
326918	4319890	train	2.30E-04	65.37727714	0.588770005
327007	4319823	train	2.29E-04	65.24725109	0.588431089
326931	4319875	train	2.27E-04	64.51790839	0.586197027
327009	4319847	train	2.25E-04	63.50039309	0.58380909
325697	4317249	train	2.20E-04	61.63152978	0.578571497
325846	4316797	train	2.11E-04	57.71687116	0.568765187
326127	4316068	train	2.11E-04	57.41921265	0.568052267
326126	4316074	train	2.10E-04	56.97115063	0.566910393
326380	4315888	train	2.07E-04	55.42546591	0.563132596
327042	4314247	train	2.03E-04	54.02381854	0.558243376
327882	4313975	train	1.98E-04	52.15979659	0.55245396
326137	4316010	train	1.95E-04	51.08629474	0.54817406
326149	4315988	train	1.92E-04	49.9338937	0.544406325
326125	4316015	train	1.88E-04	48.61272342	0.539587861
327428	4316604	train	1.80E-04	46.34055887	0.529051312

324057	4319595	train	1.79E-04	46.04230941	0.528012032
324115	4319265	train	1.74E-04	44.34156638	0.520920945
324111	4319271	train	1.74E-04	44.16033977	0.520396739
327980	4312932	train	1.71E-04	43.00113874	0.515616915
328195	4312492	train	1.70E-04	42.71726368	0.514161576
328013	4312837	train	1.42E-04	33.10258423	0.469989371
329205	4314630	train	1.35E-04	30.11409721	0.456774134
329261	4314645	train	1.30E-04	28.50344818	0.448688999
328597	4315712	train	1.30E-04	28.1420543	0.44701407
324105	4319190	train	1.27E-04	26.94529124	0.441347092
323134	4320010	train	1.07E-04	20.87124292	0.401152119
327101	4316019	train	1.07E-04	20.67479705	0.399761772
328117	4312298	train	9.81E-05	18.20328132	0.379687952
323279	4319837	train	9.57E-05	17.58166596	0.373700107
323183	4319541	train	8.85E-05	16.13947804	0.355550621
323056	4320183	train	8.78E-05	16.03151302	0.353864772
328611	4315837	train	8.25E-05	14.89009902	0.33978676
323320	4315591	train	6.98E-05	12.13195668	0.303393565
323322	4315599	train	6.89E-05	11.92163698	0.300482144
323321	4315631	train	6.40E-05	10.7559963	0.285377019
323367	4315655	train	3.60E-05	4.731550831	0.183430728
324844	4318943	test	5.01E-04	98.4006726	0.757435196
322032	4321947	test	3.24E-04	89.7173338	0.669069648
322019	4322056	test	3.14E-04	88.60736058	0.662018402
322010	4322079	test	3.01E-04	86.87887281	0.652219085
324171	4319213	test	2.89E-04	85.00424923	0.642927323
326379	4316179	test	2.60E-04	76.87403521	0.618995904
326893	4319920	test	2.24E-04	63.15551954	0.58272806
326199	4317262	test	2.18E-04	60.79484761	0.576084753
327653	4314123	test	1.99E-04	52.74825908	0.554286548
329741	4310301	test	1.98E-04	52.15827177	0.552448618
327767	4312981	test	1.82E-04	47.10824699	0.532209999
324078	4319291	test	1.71E-04	43.1727336	0.516369249
325777	4316826	test	1.60E-04	39.16019628	0.499369264
329198	4314536	test	1.10E-04	21.58505909	0.407230754
324073	4319186	test	6.55E-05	11.18917128	0.290133367