The Influence of Climate Change On The Ecology of the Arctic Ground Squirrel in Denali National Park, AK.

Nigel A. Golden
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

Part of the Genetics Commons, and the Population Biology Commons

Recommended Citation
https://doi.org/10.7275/26908472 https://scholarworks.umass.edu/dissertations_2/2461

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
The Influence of Climate Change On The Ecology of the Arctic Ground Squirrel in Denali National Park, AK.

A Dissertation Presented
by
NIGEL A. GOLDEN

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2022

Department of Environmental Conservation
THE INFLUENCE OF CLIMATE CHANGE ON THE ECOLOGY OF THE ARCTIC GROUND SQUIRREL IN DENALI NATIONAL PARK, AK

A Dissertation Presented
By
NIGEL A. GOLDEN

Approved as to style and content by:

Toni Lyn Morelli, Chair

Bethany Bradley, Member

Tammy Wilson, Member

Susan Natali, Outside Member

Timothy Randhir, Graduate Program Coordinator
Environmental Conservation
ACKNOWLEDGMENTS

I would like to give my thanks to my advisor Dr. Toni Lyn Morelli for giving me the opportunity to pursue my dream – a complete dissertation. Thank you for the guidance, support, and resources necessary to complete this work. Thank you for the balanced mixture of gentle prodding and pressure which helped me to bring this research to the finish line.

Dr. Paige Warren, Dr. Forrest Bowlick, Dr. Susannah Lehrman, Dr. Curtice Griffin, and Dr. Tim Randhir; thank you for your generous support, mentoring and contributions to my graduate school experience. Thank you, Heidi Bauer-Clapp and the team of professionals, in the Office of Professional Development, Dr. Denise Pope for helping to sharpen my academic toolset and be the best educator, mentor, and writer that I could be. Thank you to my dissertation committee members Dr. Bethany Bradley, Dr. Tammy Wilson and Dr. Sue Natali for all of your input and enriching my research. Dr. Lisa Komoroske, thank you for providing me with advice, guidance, and a wealth of knowledge and insight to this project.

The Department of Environmental Conservation, W.E.B. Dubois Department of Afro-American Studies, Northeast Climate Adaptation Center, BRiDGE, the Morelli Lab and the wider UMass Community, thank you for your support and friendship. Thank you to the 2015 Eco Graduate Student Cohort who started with me – especially Alexej Siren, Ana Quinonez, Nereyda Falconi Lopez, Kate Iaquinto, Juan Carlos Cruz, Evan Kuras and Victor Montalvo. I did not anticipate your presence making graduate school to be so fun. Thank you to my chosen family, especially Raquel Bryant, Cecile Yezou, Benjamin Kiesling, Edwin Murenzi, Marsha Allen, Sanket Sabnis, Laura Fattaruso, Mariela Garcia Arredondo, Nadia Fernandez, and Kadambari Devarajan, for feeding me, listening to me rant and rave, the Octavia Butler book clubs, Star Trek watch clubs, for believing me and being there for me.

I could have not completed my fieldwork without my talented field techs and volunteers, especially Andrea Corradini, Jailene Hidalgo, and Erika Smith. A special shout out to Ashley Morocco Powell. Thank you to the passionate educators at the Murie Science Learning Center, Sierra McLane, Stephanie Ford, David Tomeo, Dana Lindauer, Sophie May, Katie McClellan for your logistical support, opportunities to interact with park visitors and friendship. Thank you to the park rangers, especially Eric Filardi, Anastasia Brease, Anna Kirk, Nick Virgil, and Tina and Bob King for your incredible insights to the local ecology, logistical support, and friendships.
Funding support for my research was provided by the Northeast Climate and Adaptation Science Center (NECASC), the Explorers Club Foundation, the UMass Amherst Graduate School Dissertation Fieldwork Grant, and Early Career Research Grant from National Geographic Society.
ABSTRACT

THE INFLUENCE OF CLIMATE CHANGE ON THE ECOLOGY AND MONITORING OF THE ARCTIC GROUND SQUIRREL IN DENALI NATIONAL PARK, AK.

FEBRUARY 2022

NIGEL A. GOLDEN, B.S. UNIVERSITY OF WISCONSIN-STEVENS POINT
PH.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Toni Lyn Morelli

This dissertation research focuses on the ecology of the Arctic ground squirrel (Urocitellus parryii) in Denali National Park and Preserve, AK. The Arctic ground squirrels is a species of interest for monitoring efforts under the National Park Services’ Vital Signs Monitoring Program under the Vital Signs Monitoring program. The focus of this program is to monitor what is considered to be the most significant indicators of ecological conditions of the specific park resources that are of the greatest concern. The CAKN designated the Arctic ground squirrel (Urocitellus parryii) as one indicator species of park ecosystems. Despite being easy to observe and having a geographic range larger than most other ground-dwelling species, the ecology of Arctic ground squirrels and their potential vulnerability to climate change is not well studied.

This study aimed at collecting baseline information about the species status in the park by estimating the spatial distribution and density of Arctic ground squirrels and their environmental correlates, assessing the genetic diversity and population structure, and predict their occurrence and abundance in the future with respect to climate change to fill in critical information gaps relative to future management concerns.

In Chapter 2, I analyzed data collected on Arctic ground squirrels during three seasons of monitoring to generate estimates of the probability of occupancy, density, and other demographic parameters of interest of Arctic ground squirrels in DENA using hierarchical occupancy and distance transect models. The distance-transect surveys results suggest an estimated density of 19.5 individuals/ha. For Arctic ground squirrels, slope and mean precipitation was an important positive predictor, while sampling year was also predictive.
Occupancy models predicted an 0.54 occupancy probability across surveyed sites in the park. The five-year temperature mean was a negative predictor of occupancy. These relationships are uncertain as there was moderate to high overdispersion.

In Chapter 3, I used single nucleotide polymorphisms (SNPs) to delineate genetic structure and assess genetic diversity to guide future monitoring efforts. I identified three major genetic clusters across the five populations, with low genetic diversity and genetic differentiation among the populations and clusters. The low genetic diversity is possibly attributed to genetic drift caused by population declines and contractions in the species range caused by climate fluctuations. The low genetic differentiation detected could be due to occupying a narrow band of suitable habitat in Denali National Park and Preserve.

In Chapter 4, I make use of presence-absence, abundance, and genetic data previously collected in this study to assess potential conservation implications of climate change on the distribution of Arctic ground squirrels using the Random Forest algorithm. The results indicate that regions of predicted high occurrence and abundance showed similar patterns where high concentrations are mainly in the mid to southern portion of the administrative boundaries of the park. The maps of predicted spatial distribution projected with future climate predictions here represent an initial attempt to capture the geographical and temporal range of the species and may be useful for identifying potential areas of conservation concern and improving sampling for future studies.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td><strong>1</strong> ARCTIC GROUND SQUIRRELS: A USEFUL SPECIES FOR ASSESSING AND MONITORING THE IMPACTS OF CLIMATE CHANGE</td>
<td>1</td>
</tr>
<tr>
<td><strong>2</strong> ABUNDANCE AND OCCUPANCY OF ARCTIC GROUND SQUIRRELS (UROCITELLUS PARRYII) IN DENALI NATIONAL PARK, AK: IMPlications for Monitoring</td>
<td>7</td>
</tr>
<tr>
<td>2.1 ABSTRACT</td>
<td>8</td>
</tr>
<tr>
<td>2.1.1 Keywords</td>
<td>8</td>
</tr>
<tr>
<td>2.2 INTRODUCTION</td>
<td>8</td>
</tr>
<tr>
<td>2.3 METHODS</td>
<td>10</td>
</tr>
<tr>
<td>2.3.1 STUDY AREA</td>
<td>10</td>
</tr>
<tr>
<td>2.3.2 DATA COLLECTION</td>
<td>10</td>
</tr>
<tr>
<td>2.3.2.1 Vegetation and abiotic data</td>
<td>10</td>
</tr>
<tr>
<td>2.3.2.2 Survey</td>
<td>11</td>
</tr>
<tr>
<td>2.3.2.3 Remote-sensed spatial and climatic data</td>
<td>12</td>
</tr>
<tr>
<td>2.3.3 STATISTICAL ANALYSIS</td>
<td>12</td>
</tr>
<tr>
<td>2.3.3.1 Occupancy modeling</td>
<td>12</td>
</tr>
<tr>
<td>2.3.3.2 Distance sampling modeling</td>
<td>13</td>
</tr>
<tr>
<td>2.3.3.3 Simulations</td>
<td>15</td>
</tr>
<tr>
<td>2.4 RESULTS</td>
<td>16</td>
</tr>
<tr>
<td>2.4.1 Occupancy Modeling</td>
<td>16</td>
</tr>
<tr>
<td>2.4.2 Distance Sampling Modeling</td>
<td>17</td>
</tr>
<tr>
<td>2.4.3 Spatial capture-recapture</td>
<td>18</td>
</tr>
<tr>
<td>2.4.4 Simulations and Bayesian Estimates</td>
<td>18</td>
</tr>
<tr>
<td>2.5 DISCUSSION</td>
<td>19</td>
</tr>
<tr>
<td>2.6 CONCLUSIONS</td>
<td>22</td>
</tr>
<tr>
<td><strong>3</strong> GENETIC DIVERSITY AND POPULATION OF UROCITELLUS PARRYII IN DENALI NATIONAL PARK AND PRESERVE</td>
<td>36</td>
</tr>
<tr>
<td>3.1 ABSTRACT</td>
<td>36</td>
</tr>
<tr>
<td>3.1.1 Keywords</td>
<td>36</td>
</tr>
</tbody>
</table>
3.2 INTRODUCTION ................................................................................. 37
3.3 METHODS ......................................................................................... 39
  3.3.1 Study area and sampling ............................................................... 40
  3.3.2 DNA extraction, library construction, and SNP calling ............... 40
  3.3.3 Genomic analyses ................................................................. 41
    3.3.3.1 Population structure ......................................................... 41
    3.3.3.2 Genetic diversity ............................................................. 42
    3.3.3.3 Genetic differentiation ..................................................... 42
3.4 RESULTS ............................................................................................... 42
  3.4.1 Genotyping ................................................................................. 43
  3.4.2 Population structure ................................................................. 43
  3.4.3 Genetic diversity ........................................................................ 44
  3.4.4 Genetic differentiation ............................................................. 44
3.5 DISCUSSION ......................................................................................... 45
3.6 CONCLUSION ....................................................................................... 46

4 BUILDING A BETTER MODEL: USING REPEAT SURVEY DATA AND GENETIC MARKERS TO ESTIMATE THE ARCTIC GROUND SQUIRREL (UROCITELLUS PARRYII) SPECIES DISTRIBUTION .................................. 58
4.1 ABSTRACT ......................................................................................... 58
  4.1.1 Keywords ..................................................................................... 59
4.2 INTRODUCTION ................................................................................... 59
4.3 METHODS ......................................................................................... 61
  4.3.1 Study area and sampling ............................................................... 61
  4.3.2 Climate projections ...................................................................... 62
  4.3.3 Classification and regression trees ................................................. 63
4.4 RESULTS ............................................................................................... 63
  4.4.1 Model performance ..................................................................... 63
  4.4.2 Variable importance .................................................................... 63
  4.4.3 Partial dependence plots ............................................................. 64
  4.4.4 Occurrence and abundance distribution patterns ........................... 65
4.5 DISCUSSION ......................................................................................... 65

BIBLIOGRAPHY ......................................................................................... 79
LIST OF TABLES

2.1 Description of covariates considered during exploratory analysis for predicting occurrence and density of Arctic ground squirrels in Denali National Park and Preserve ................................................................. 32

2.2 Model-averaged parameter estimates, unconditional standard errors and 95% confidence limits of coefficients of covariates in occupancy models. Parameter estimates whose model-averaged 95% CI did not overlap zero are displayed in bold ........................................................................................................ 33

2.3 Results of QAICc-based model selection for covariates influencing probability of detection (p) and occupancy (ψ) of Arctic ground squirrels. Year was included as a term in all models whereas all other covariates were allowed to vary. K = number of parameters .............................................................. 34

2.4 Results of AICc-based model selection for covariates influencing probability of detection (p) and density (λ) of Arctic ground squirrels. Year was included as a term in all models whereas all other covariates were allowed to vary. K = number of parameters. The φ parameter was held at it’s intercept in all models ....... 35

2.5 Results of QAICc-based model selection of probability of detection (p) and density (λ) for Arctic ground squirrels in Denali National Park and Preserve, AK ........................................................................................................ 36

3.1 Location information and the number of genotyped Arctic ground squirrels that were sampled (# Tissue Samples), failed to pass quality checks (# Failed Tissue Samples), genotyped for downstream analysis (Total Individuals) in Denali National Park and Preserve, AK ................................................................................. 54

3.2 Mean observed heterozygosity Hobs and mean expected heterozygosity Hex, and allelic richness Ar estimated by rarefaction is reported per site ......................................................................................................................... 55

3.3 Analysis of molecular variance (AMOVA) for 46 individuals in 5 populations of Arctic ground squirrels in Denali National Park and Preserve, AK ......................................................................................... 56

3.4 Pairwise FST values between Arctic ground squirrel sampling locations (aka putative populations) are shown below the diagonal ................................................................................................................. 57

4.1 Description of climatic covariates considered during exploratory analysis for predicting occurrence and density of Arctic ground squirrels in Denali National Park and Preserve .......................................................................................................................... 75

4.2 Top 4 variables importance ranking of Random Forest occurrence, abundance, and genetic models ....................................................................................................................................................... 76
LIST OF FIGURES

1.1 An Arctic ground squirrel standing upright alert for predators. Photo Credit: Anna Kirk, National Park Service ................................................................. 6

2.1 Map of entire study extent. Locations for study sites (nodes) are represented by black boxes. The Denali Park Road that runs lengthwise through the park is represented by the black line................................................................. 25

2.2 Example mini-grid design of the permanent vegetation monitoring plots in Denali National Park and Preserve. The sampling frame is used for the collocation and integration of monitoring efforts occurring across the park. The survey sites are represented by red dots arranged in five rows of five points, spaced 500 meters apart.................................................. 25

2.3 Model-averaged predicted relationships between occupancy of Arctic ground squirrels and top supported covariates in the top <2 Δ QAICc model set. Solid lines represent model-average predicted occupancy and the shaded region and error bars delineate upper and lower 95% CI. .............................................. 27

2.4 Model-averaged predicted relationships between detection probability of Arctic ground squirrel and supported covariates in the top <2 Δ QAICc model set. Solid lines represent model-averaged predicted detection and shaded regions delineate upper and lower 95% CI.................................................. 28

2.5 Detection probability as a function of distance from hazard line transect model fitted to Arctic ground squirrels in Denali National Park and Preserve....................................... 29

2.6 Boxplots showing mean square error (MSE) of occupancy probability (A) and detection probability (B) across 9 simulated sampling designs. Shaded panels are included to assist with comparisons between sampling designs. The red diamonds are the group means.................................................................. 30

2.7 Contains confidence intervals for the difference in the means for all 36 pairs of groups. This is the graphical display of the multiple (pair-wise) comparisons using Tukey’s HSD for the scenario data. Any confidence intervals that do not overlap with zero provide evidence of a difference in group means.......................... 31

3.1 Sites of sampled and analyzed individuals. Sites are grouped and colored by location name. Inset shows location in Denali National Park and Preserve, AK......................................................................................... 48

3.2 Principal components analysis (PCA) of 46 Arctic ground squirrel samples with 17,460 unlinked SNPs........................................................................ 49

3.3 Stratified cross-validation of DAPC using varying numbers of PCs to ensure the appropriate number is retained. 5 PCs give the highest percent of correctly predicted subsamples with the lowest error............................................. 50

3.4 Discriminant analysis of Principal Components (DAPC) scatterplot drawn against 17,466 SNPs across 46 individuals. Dots represent individuals, with cluster denoting sampling origin and inclusion of 95% inertia ellipses.................................................................................. 52

3.5 Hierarchical fastStructure barplots of 46 individuals using 17,460 unlinked SNPs.
Each column is an individual. Rows are plots of $K = 1$, $K = 2$, and $K = 3$, respectively

3.6 AMOVA permutations test using 1000 permutations for the number of subpopulation $K = 5$. The three histograms represent the distribution of the randomized strata and the black line represents the observed data.

4.1 Partial dependence plots for the influential variables in the occurrence distribution models for Arctic ground squirrels. Covariates include bio11, bio14, bio15, bio8, bio9, and bio1 in that order.

4.2 Partial dependence plots for the influential variables in the abundance distribution models for Arctic ground squirrels. Covariates include bio1, bio3, bio8, bio9, bio11, and bio14, and bio15 in that order.

4.3 Partial dependence plots for the influential variables for genetic cluster 1 in the binary models for Arctic ground squirrels. Covariates include bio11, bio8, bio9, and bio6 in that order.

4.4 Partial dependence plots for the influential variables for genetic cluster 2 in the binary models for Arctic ground squirrels. Covariates include bio13, bio18, bio19, and bio8 in that order.

4.5 Partial dependence plots for the influential variables for genetic cluster 3 in the binary models for Arctic ground squirrels. Covariates include bio13, bio18, bio19, and bio8 in that order.

4.6 Spatial distribution map of relative occurrence with near-current and future climate data. Values scale from high (green) to low (pink) (green) predicted occurrence.

4.7 Spatial distribution map of relative abundance with near-current and future climate data. Values scale from high (green) to low (pink) (green) predicted occurrence.
CHAPTER 1

ARCTIC GROUND SQUIRRELS: A USEFUL SPECIES FOR ASSESSING AND MONITORING THE IMPACTS OF CLIMATE CHANGE

There are few global challenges more urgent than the current climate change crisis. The five warmest years globally have occurred since 2010 (GISTEMP, 2018), with the last three years the warmest on record (Hansen et al., 2010; GISTEMP, 2018). The Arctic is one of the most highly vulnerable biomes to global climate change; it is warming at 2 times the global rate (Hu et al., 2015; Iwahana et al., 2016; Overland et al., 2018). Since the observational record began in 1900, there has been a 3.5 °C increase in annual surface air temperature north of 60° N (Overland et al., 2018). Although there is high year-to-year variability, air temperatures have exceeded previous records for the past five years (2014-2018). The persistent, and rapidly intensifying atmospheric warming continues to drive broad, long-term trends: sea ice extent is at its lowest in the 38-year (1979-2018) instrumental record (Perovich et al., 2018); spring snow cover extent over the Arctic hit a new record low in 2016 (Derksen et al., 2016); and dramatic changes to the Arctic landscape are being driven by increases in the rate of formation of thermokarst collapses in the land surface due to thawing permafrost and the degradation of ice-wedges, large ice deposits formed in ground cracks caused by freeze thaw cycles (Lara et al., 2016; Liljedahl et al., 2016). While there is debate over the extent, rate and nature of change that will occur, what is certain is that a new ’normal’ is settling over the Arctic, characterized as a warmer, less frozen, and biologically changed region.

Climate change in the Arctic will cause dramatic ecological transformations that cannot easily be reversed with changes in management. There is mounting evidence that global climate change has caused species distributions to shift northward and to higher elevations (Lipton et al., 2018). Arctic biodiversity will thus potentially face simultaneous challenges of shrinking suitable habitat in the North (Marcot et al., 2015) and displacement to the South of their range (Macdonald et al., 2010). Marcot et al. (2015) projected a significant shift in land cover types in the Arctic and concluded that most wildlife will experience habitat declines. As the Arctic becomes warmer and wetter, wildlife distributions are increasingly shifting poleward (Hof et
al., 2012). Because of warming, terrestrial Arctic wildlife, delimited in the north by the Arctic Ocean, will be squeezed by receding suitable climate at their southern edge (Meltofte et al., 2013). Effects on species will likely alter food-web structures and complexity, increase risk of extinction, and threaten the cultural integrity and subsistence traditions of Alaska Natives (Marcot et al., 2015; Anisimov et al., 2007).

Thus, these rapidly changing ecosystems are likely to present significant philosophical and practical challenges for natural resource managers. For example, the range of management options for the National Park Service (hereafter NPS), accustomed to maintaining natural conditions, is shrinking as 96% of NPS land are in areas of observed 20th-century warming (Gonzalez, 2011; Lynch et al. 2021). NPS areas in Alaska are particularly impacted by global climate change (Gonzalez et al., 2018). In Denali National Park and Preserve in interior, AK, (hereafter DENA) the summer of 2019 was the wettest on record – 16.61 total inches of rain, 206% of normal and the 6th warmest 1.11 °C warmer than normal (Sousanes and Hill, 2020). It was preceded by the second warmest December on record (7.5 °C warmer than normal) in DENA, with snowfall at <50 % of normal (Sousanes and Hill, 2020). Swanson et al. (2021) show that from 2014-2019, the mean annual air temperature in DENA increased nearly by 2 °C. Models estimating changes to the spatial distribution and composition of vegetation indicates a spread of mesic forest and shrubland into the park, primarily at the expense of vulnerable Arctic tundra habitat (Holsinger et al., 2019). These changes and new, unexpected transformations to DENA landscapes in the future are likely to persist for decades to come. With over 6 million acres of parklands to administer, resource managers in DENA, need to understand possible climate trends and their impact on species within the park to better adopt different adaptation options under the unpredictable and novel trajectory of ecosystem changes.

The NPS Inventory and Monitoring Division is tasked with monitoring changes in overall ecosystem health using indicators called “vital signs.” The US national park units are grouped into 34 networks based on common geography and ecosystem function. Denali is part of the Central Alaska Network (CAKN) which represents over 25% of total land in the National Park System. The purpose of the CAKN vital signs monitoring program is to help properly manage park resources (https://science.nature.nps.gov/IM/units/cakn/index.cfm). With input from its member parks, the CAKN established a combination of physical, chemical, and biological components of the Arctic ecosystem to serve as “vital signs” of the overall condition of network parks. Focusing on habitat change, vegetation structure and composition, and wildlife is part of
meeting this goal in understanding the baseline condition of habitat in ecosystems so that any impending threats can be detected and addressed.

Representatives from DENA selected the Arctic ground squirrel *Urocitellus parryii* (Figure 1.1) as a Vital Sign for vulnerable Arctic tundra habitat. Arctic ground squirrels are small, conspicuous, diurnal sciurids. Their geographic range extends farther north than any other species of squirrel, and they are also the only hibernating terrestrial animal species with a Holarctic distribution. The species is colonial and lives in shallow, subterranean burrow systems. Densities of breeding adults vary with latitude and habitat, with reports of between 1-3/ha in boreal forest habitats (Werner, 2015; Donker and Krebs, 2011; Hik et al., 2001; Byrom et al., 2000) and 3-16/ha in tundra or meadow habitats (Werner, 2015; Donker and Krebs, 2011; Hik et al., 2001; Lacey et al., 1997). These estimates are likely to vary substantially across its geographic range as turnover in colony size and type fluctuates widely with habitat and latitude (Lacey and Wieczorek, 2001; Boonstra et al., 2001; Batzli and Sobaski, 1980).

Arctic ground squirrels are distinct from other small Arctic mammals such as snowshoe hare because their populations are noncyclical; thus, they can act as keystone species. They have a variety of predators, many of who are dependent on them when other prey species are at low abundance (Barker and Derocher, 2010). Karels and Boonstra (2000) documented that 96% of active season Arctic ground squirrel mortalities were due to predation during the low phase of the ten-year cycle in snowshoe hares. As a soil-disturbing vertebrate, the Arctic ground squirrels acts as an ecosystem engineer; its burrowing can have significant direct and indirect geomorphic impact by sediment displacement (Price, 1971), can assist plants in recolonizing the disturbed surfaces through nutrient addition by deposition of waste (McKendrigk et al., 1980; Gough et al., 2008; Eriksson, 2011) and can create habitat for secondary animals.

Arctic ground squirrels are theorized to be highly vulnerable to climate change (Wheeler and Hik, 2013). Burrowing mammals like the Arctic ground squirrel are thought to be extremely vulnerable to landscape changes associated with permafrost thaw (Wheeler and Hik, 2013). Permafrost thaw due to climate change can impact population distribution of burrowing animals like Arctic ground squirrels through two key processes. First, as the active layer (seasonally thawed surface soils) thickens and depth to the top of the permafrost layer increases, new habitat will be available (Wheeler and Hik, 2013). Second, as permafrost thaws, surface water will drain to deeper soil layers, increasing access to the well-drained areas that burrowing animals prefer in some soil types and in upland regions (Avis et al., 2011; Barker and Derocher, 2010).
However, soil saturation may increase in lowland or poorly drained areas, or in regions with high ice content, where permafrost thaw leads to ground slumping and inundation (Jorgenson et al., 2001; Wrona et al., 2016). It is not clear yet from the literature the impact that these shifts in the landscape will have on Arctic ground squirrel persistence.

Changes in the timing and amount of precipitation (including snow cover), as well as changes in tundra plant communities, will also likely impact Arctic ground squirrel foraging ecology and population abundance (Flower et al., 2019; Wheeler and Hik, 2013). Earlier springs due to rising atmospheric temperatures are also known to affect hibernation behavior by causing potential phenological mismatches with important plant or arthropod food sources (Wheeler and Hik, 2013). Additionally, low snowpack from higher atmospheric temperatures can decrease overwintering success by reducing snow insulation. Increases in minimum temperatures are theorized to push the Arctic ground squirrel both northwards and up to higher elevations. Additionally, changes in vegetation structure and composition may substantially affect their abundance. Hibernators such as Arctic ground squirrels could respond positively to the higher abundance of food in the longer summer and a higher survival rate due to shorter winters. Conversely, studies have documented Arctic ground squirrel populations with sink characteristics in study sites dominated by boreal and shrub vegetation, and with tall non-forage functional vegetation groups (Wheeler et al., 2015). As an integral part of subarctic and Arctic ecosystems and food webs, determining the effects of climate change on Arctic ground squirrels is a significant ecological challenge. It is critical to determine the stability and sustainability of the species through long-term monitoring.

The purpose of this study was to provide managers with baseline information on Arctic ground squirrels ecological conditions in DENA, gather knowledge about the efficacy of different monitoring strategies, and gather information on the type of effects that climate changes may have, to reduce critical uncertainties and consequently lead to better decisions for enabling species to adapt to ecological change.

This dissertation contains three chapters in addition to the introductory chapter. Chapter 2 estimates the spatial distribution and population size of Arctic ground squirrels in DENA and then uses the results to address the technical and logistical challenges of how to survey climate-change-vulnerable wildlife species. Chapter 3 presents the results of research using genomic techniques to assess the population structure and gene flow of Arctic ground squirrels. Chapter 4 presents results testing how detection/non-detection, abundance, and population genomic
structure data can predict Arctic ground squirrel distribution in response to climate change.
Figure 1.1: An Arctic ground squirrel standing upright on alert for predators. Photo Credit: Anna Kirk, National Park Service
CHAPTER 2

COMPARISON OF METHODS FOR SURVEYING CLIMATE CHANGE VULNERABILITY OF ARCTIC GROUND SQUIRRELS (UROCITELLUS PARRYII) IN DENALI NATIONAL PARK AND PRESERVE, AK

2.1 ABSTRACT

Occupancy and density estimates are often considered necessary to predict the impact of climate change on species and communities. Thus, identifying methods to assess the environmental conditions that are tied to a species’ highest likelihood of occurrence and density are top priorities in monitoring programs. However, limited budgets and staff time require informative yet low cost methods. Density estimates produced by traditional distance-transect sampling and spatial capture-recapture can be time-consuming and expensive over large areas, while estimating the proportion of sites occupied can be hampered by imperfect detection. Despite its status as a National Park Service indicator species, population-level data for the Arctic ground squirrel (Urocitellus parryii) are incomplete. In this study, I collected data during the summers of 2017, 2018, and 2019 in Denali National Park, and used a simulation framework to assess the use of distance sampling, spatial capture-recapture, and presence-absence surveys for monitoring Arctic ground squirrels. Distance-transect surveys resulted in an estimated density of 19.5 individuals/ha while presence-absence surveys resulted in a 0.54 occupancy probability across surveyed sites in the park. The estimated densities and occupancy across the surveyed area, and the properties for conducting distance transect sampling and spatial capture-recapture methods suggest presence-absence surveys as the measure of choice for monitoring Arctic ground squirrels. The results of this study provide the first empirical estimates of the spatial distribution and density of Arctic ground squirrels in Denali National Park and can be used to guide monitoring efforts of Arctic ground squirrels and inform the design of future studies.
2.1.1 Keywords:

*Urocitellus parryii*; climate change; conservation; indicator species; monitoring; occupancy modeling; abundance modeling; capture-recapture; survey

### 2.2 INTRODUCTION

Conservation planners spend a substantial amount of time and resources monitoring wildlife populations (Martin et al. 2007). Ecological monitoring is a common tool for management strategies where the purpose is to assess the condition or status of a species at a particular point in time to implement actions based on the current state. The knowledge gained from monitoring can generate ecological inferences that can be used to test theories and make predictions of future population dynamics and distribution (Marolla et al., 2021). More fundamentally, monitoring provides information on baseline states as well as changes that are required for most management decisions (Fancy et al. 2009).

Depending on the management objective, monitoring efforts focus on understanding a species’ distribution and estimating its population size across that distribution. Measures of density (number of individual/unit area) and/or occupancy (proportion of area occupied) are key to informing conservation planning (Guisan et al., 2013), monitoring (Randin et al., 2020), and understanding the processes that constrain ecological niches (Dawson et al., 2011) and drive population size (Ehrlén and Morris, 2015). Despite potentially improving conservation decisions, practitioners and scholars alike recognize that there are multiple limitations and daunting technical challenges that ecological monitoring presents (Stem et al., 2005). Monitoring is (1) expensive as it usually requires human resources, and the commitment needs to be sufficiently long term to explain data trends; (2) logistically difficult as a sufficient sample size needs to be obtained while adhering to rigorous sample design; and (3) can produce results that are difficult to analyze (Field et al., 2007). In addition to some of these limitations/challenges, there are potential trade-offs when operating under finite budgets that require the consideration of balancing multiple management activities.

Monitoring the status of the Arctic ground squirrel (*Urocitellus parryii*) in Denali National Park and Preserve in interior Alaska poses several challenges. First, the relatively large size of the park (approximately the size of the state of Maine) and limited road system presents constraints on collection of data which may limit dataset size and representativeness. Second, extrapolating ecological relationships identified from previous studies of Arctic ground
squirrels is challenging because most were conducted at their southern range limit in the Southwest Yukon where they may experience environmental conditions different than those in DENA (Thogmartin et al., 2006).

In order to address these monitoring challenges for the Arctic ground squirrel and for other priorities that are lacking sufficient resources, I compared three common survey methods to determine the most reliable way to estimate their patterns of distribution and density: spatial capture-recapture, distance transect sampling, and presence-absence surveys. Spatial capture-recapture (SCR), a recent development of mark-recapture (Borchers and Marques, 2017), attempts to capture (using an array of traps) or identify with detectors (using cameras to identify natural marks, scars, leg bands, etc.) individuals over multiple sampling occasions to create an encounter history, denoting the occasions where an individual was detected or not (Royle and Dorazio, 2008). The encounter histories can then be used to estimate detection probability, and by extension, the proportion of the total population that was captured. Thus, SCR requires that individuals in the population of interest be individually identifiable.

Distance transect sampling and presence-absence surveys are noninvasive and don’t rely on individuals being identifiable. With occupancy modeling (MacKenzie et al., 2002; Tyre et al., 2003; Ahumada et al., 2013; MacKenzie et al., 2003), the detection/non-detection of individuals in a given site collected and each site is surveyed on multiple occasions. Occupancy models are used to determine the true presence or absence of a species at a site through the joint modeling of site occupancy and detection probabilities. N-mixture models, are used to estimate species abundance on the basis of species counts. Sharing similar properties of the occupancy model, the data are collected using standard counting techniques, and sites are revisited on multiple occasions. The repeated counts are then used to jointly estimate the detectability of the species of interest and the population size (Royle, 2004; Kery et al., 2009; Dail and Madsen, 2011). N-mixture models are increasingly used for monitoring of small vertebrates (Romano et al., 2017).

Comparing across these three methods, I analyzed data collected on Arctic ground squirrels during three field seasons to: 1) generate estimates of the probability of occurrence, density, and other demographic parameters of interest of Arctic ground squirrels in DENA, 2) identify the factors that most influence density and occurrence, and 3) determine whether different survey designs could obtain the same reliable estimates using similar or reduced effort. Finally, I considered the implications for continued monitoring of Arctic ground
squirrels, including how the information from this study could contribute to assessments of the vulnerability of Arctic ground squirrels to climate change.

2.3 METHODS

2.3.1 STUDY AREA

I collected data on Arctic ground squirrels from June to August in 2017-2019 in DENA (63°7N, 151°11W; datum = WGS84) (Figure 2.1), which makes up approximately 24,524 km2 of interior South-Central Alaska. DENA is bounded primarily by state land, including Denali State Park on its southern border. Fieldwork occurred on either side of the 145 km park road. DENA Terrain varies dramatically with habitat composed of coniferous forest, ranging upward into dense and open shrubland, and then to open alpine habitat. Roland et al. (2019) provides additional details about this study area. DENA has a continental climate characterized by long, cold, dry winters with short, warm summers (Shulski and Wendler, 2007). The mean annual temperature is -2.2 °C, while the summer mean temperature is 11.6 °C, and winter mean temperature is -14.6 °C.

2.3.2 DATA COLLECTION

2.3.2.1 Vegetation and abiotic data

All ground squirrel sampling in this study was co-located with long-term vegetation monitoring plots (details in Roland (2004)). Within this design, a systematic grid with random start spans a 12,800-km2 area of the park, with nodes spaced every 20 km and additional nodes added at 10-km intervals. At each node, a nested sampling grid was located with sampling stations in a 5x5 arrangement, and each station separated by 500 m (25 points) (Figure 2.2). This study sampled 15 of the possible different nodes established across the park. Within each plot at each of the 25 points within a mini grid, was a circular plot 16 m in diameter (area of about 200 m2), and four separate 4-m2 quadrat placements that were sampled measuring a suite of topographic and edaphic variables and recorded evidence of fire history, flooding, or geomorphic activity. Data and methodology associated with this study is described in detail by (Stehn and Roland, 2018).
2.3.2.2 Survey

Presence-absence surveys (1-3 visits/site) were used to estimate the occupancy probability in DENA. Surveys consisted of 10 minutes of listening for alarm calls and scanning with binoculars. The recorded data were classified in two categories: species was detected, 1; and species was not detected, 0. Repeated distance line transect surveys (1-3 visits/site) were conducted to estimate the density of Arctic ground squirrels in DENA (Figure 2.1). Selected sites consisted of pre-determined transects comprising a 500-m straight line from one plot to the next of the 5x5 mini-grid nested within the larger sampling grid. I counted and recorded the amount of Arctic ground squirrels along the line. I recorded Arctic ground squirrel detections (by sight or sound) during the walk. When identified, the sighting distance and sighting angle were measured. I used a rangefinder to record the distance of Arctic ground squirrels detected from the line and used a combination of GPS and compass for measuring sighting angles. Survey-level covariates were measured after each count and included: wind speed, cloud cover, Julian date, and time of day.

In the 2018-2019 field seasons, I conducted mark-recapture surveys. At each site, I placed Tomahawk live traps (26cm * 6cm * 6cm) (Tomahawk Live Trap Co., Tomahawk, Wisconsin, USA) baited with slices of carrot, at burrows, and checked every hour for captures. Individual trap locations were inventoried with a GPS unit during the field season. To minimize thermal stress and predation risk to the animals, wind and sun guards were placed along the edge of each trap. A 10-second scan with a tag reader of each trap holding an individual determined whether it was a new individual or a recapture. If new, pre-sterilized PIT tags were inserted subcutaneously between the shoulder blades using a syringe and a 12-gauge needle. If an Arctic ground squirrel was a recapture, it was released without being sampled again.

2.3.2.3 Gridded spatial and climatic data

I considered remotely sensed data from ArcticDEM (Porter et al., 2018) and climatic data from Daymet (Thornton et al., 2017) for probability of occupancy (ψ) and density. In addition, I extracted landscape data (e.g., elevation, slope, azimuth) from the ArcticDEM Mosaic dataset using ArcGIS 10.7.1 for each site. I used the “biovars” function in the package dismo (Hijmans et al., 2017) to summarize monthly bioclimatic data calculated from Daymet datasets using the package daymetr (Hufkens et al., 2018).
2.3.3 STATISTICAL ANALYSIS

2.3.3.1 Occupancy modeling

To estimate the probability an Arctic ground squirrel occupied a site, I modeled Arctic ground squirrel presence and absence using single season occupancy models in the R package *unmarked* (Fiske et al., 2011). Occupancy models are hierarchical state-space models: a site-specific logit-linear model for occupancy ($\phi$) and the observation process describing imperfect detection ($p$). I did not use the multi-season occupancy modeling (MacKenzie et al., 2003) due to unequal sampling effort across seasons which was insufficient to support the greater model complexity design. Instead, I “stacked” the site detection histories across years, treating each unique site-year combination as independent sampling units ($N = 152$), and included a categorical year covariate ($0 = 2017$, $1 = 2018$, $2 = 2019$) in calculating $\psi$ to avoid pseudoreplication (MacKenzie et al., 2017). Additionally, because I was more interested in site-use and spatial patterns than rates of colonization and extinction, I used the “implicit dynamics” modeling approach (MacKenzie et al., 2017), which provides a great deal of flexibility to model unequal survey effort to estimate $\psi$ and $p$ probabilities from repeated detection/nondetection data. Sites that had missing surveys were kept in the data set as “NA,” because occupancy modeling allows for missing sampling occasions. I used the package *unmarked*, which allows fitted linear models within a maximum likelihood framework that can be combined with an information theoretic approach (Anderson and Burnham, 2002) for model selection.

To choose the appropriate environmental predictors, I calculated the variation inflation factor (VIF) with a threshold of VIF=3, to reduce multicollinearity in covariates. Remaining variables were retained based on their relevance to Arctic ground squirrel ecology (Zuur et al. 2010). Prior to constructing models, all continuous covariates were standardized to a mean of zero and standard deviation of 1. With the reduced covariate set, I considered the following potential effects on occupancy: slope, vector rugged measure (vrm; quantifies terrain ruggedness more independently of slope), percent of plot covered in vegetation stature (low: $<50$ cm; medium: 50-200 cm; and high: over 200 cm), maximum soil depth (cm), maximum temperature of warmest month (bioclimatic variable 5; bio5), minimum temperature of coldest month (bioclimatic variable 6; bio6), annual mean precipitation (pmean), five year temperature mean (fivetm), year of sampling (year) and all combinations of interactions were considered (Table 2.1). For the detection probability I considered cloud cover (%), high vegetation stature
To model candidate model space reliably and efficiently, I followed the modeling approach suggested by Morin et al. (2020) by fitting the sub-models of detection and occupancy independently and combining the top set of models from each sub-model for selection in a final stage. I used quasi-Akaike’s Information Criterion (QAIC) to compare balanced variable model sets, and calculated parameter estimates, unconditional standard errors, and 95% confidence intervals via model averaging (Anderson and Burnham, 2002). I treated all competing models (e.g., ΔQAICc < 2.0) as plausible and included them in the final candidate set. I assessed the most general model fit through a $\chi^2$ goodness-of-fit test using the function “mb.gof.test” with 10,000 bootstrapping events in the AICcmodavg package (Mazerolle, 2020). QAIC values were used to compare models when overdispersion and a lack of fit were indicated ($\hat{c}$ > 1). When data were overdispersed, I inflated unconditional standard errors by the square root of the value (MacKenzie and Bailey, 2004). I then averaged candidate models (QAIC < 2) using the package MUMin (Barton and Barton, 2015). This procedure determines the Akaike weights and averages the estimates of the parameters of interest among the set of candidate models. All analyses were conducted in R Version 4.1.2 (R Core Team, 2021).

### 2.3.3.2 Distance sampling modeling

In order to get a population size estimate, I modeled density with hierarchical distance models (HDM) using the “gdistsamp” function in the R package unmarked (Fiske et al., 2011; R Core Team, 2021). Using the model of Royle (2004), based on three conditional and partially observable random processes, I applied a coupled logistic regression where the detection process is conditional on the density of animals within each transect at each site(s). I grouped continuous values of distance that were measured in the field, in distance classes (h) of 100 m such that I had four distance bins to model observations and stacked data as with the occupancy analyses (Buckland et al., 2001, 2005). I then assigned each observation ($y_{sh}$) along each transect (t) to its specific distance class (h). I stacked data as with the occupancy analysis. I fit models using the package unmarked in R (Fiske et al., 2011) in program R 3.6.2 (R Core Team, 2021) with the function “gdistsamp” (Fiske et al., 2011). I modeled local density ($X_t$) at each transect, assuming a Poisson distribution: ($X_t$ Poisson ($\lambda_t$); $t = 1, ..., n$) with $\lambda$ representing the expected value of $X(\lambda = E(x))$. I also modeled detection frequencies (i.e., the observed count of individuals in each h), assuming for these a multinomial distribution, conditional on the...
population size $X_t: (y_{t1}, \ldots, y_{tH}) \sim \text{Multinomial}(X_t, \pi)$, where $\pi_{th}$ is the multinomial probability for distance class $h$ and transect $t$, that depend on the parameter of the detection function, which I assumed a hazard-rate distribution, as $g(y) = \exp - (y^2/2\sigma^2)$. Prior to each analysis, I reduced the full environmental dataset I hypothesized to be associated with Arctic ground squirrel habitat (Table 2.1). I calculated the VIF and removed variables at the VIF = 3 threshold and used biological relevance to decide which variables to retain (Zuur et al. 2010).

To model candidate space reliably and efficiently, I followed the modeling approach suggested by Morin et al. (2020) by fitting the sub-models of detection (“detection” model set) and density independently and combining the top set of models from each sub-model for selection in a final stage. In the model, availability was assumed to be constant because I did not have enough observation data to accommodate more covariates. I first independently modeled factors that influenced detection probability using three survey covariates: (1) time of sampling (tos), (2) Julian date (yday), and (3) cloud cover (%). I created all possible combinations of the survey covariates to model the detection probability. The top sub-models from this candidate set were used in the final stage candidate selection. As with occupancy, I tested covariates that assessed broad patterns of density: covariates here include a range of climatic, topographic and habitat variables using all possible combinations of additive covariates and quadratic relationships with elevation and year as a site covariate. Finally, using the supported models from both the detection and density sub-model set, I created all possible combinations of the top models and compared them using AICc. All models were generated using the hazard-rate key functions. Using the top ranked model, I plotted functional relationships using the predict function in R.

To evaluate how well a model fits the data collected, I used QAIC to compare candidate models, and calculated parameter estimates, unconditional standard errors, and 95% confidence intervals via model averaging (Anderson and Burnham, 2002). I treated all competing models (e.g. $\Delta$QAICc < 2.0) as plausible and included them in the final candidate set. I assessed the most general model fit through a $\chi^2$ goodness-of-fit test using the function “mb.gof.test” with 10,000 bootstrapping events. QAIC values were used to compare models when overdispersion and a lack of fit were indicated ($\hat{c}$ > 1). When data were overdispersed, I inflated unconditional standard errors by the square root of the value (MacKenzie and Bailey, 2004). I then averaged candidate models ($\Delta$QAIC < 2) using the package MUMin (Barton and Barton, 2015). This procedure determines the Akaike weights and averages the estimates of the parameters of
interest among the set of candidate models.

### 2.3.3.3 Simulations of effect of survey effort

I used a simulation framework to evaluate the effects of survey effort on occupancy estimator quality. I simulated data for 9 scenarios representing a range of sites and survey replicates that resource managers at DENA could potentially implement. These scenarios represented the monitoring data collected on the Arctic ground squirrel, and so I used the mean detection probabilities estimated from the null model (detection probability was modeled as constant) for each scenario. The following levels were being investigated:

- **n (sites)** = 50 (n1), 75 (n2), or 100 (n3);
- **s (surveys)** = 1 (s1), 2 (s2), or 3 (s3)
- **p** = 0.8
- **ψ** = 0.5

For each scenario, I simulated data and compared the average estimation error for a given value of the detection probability and density and occupancy probability. While it would be desirable to model more combinations, the computational resources required were prohibitive. I ran 1000 simulations on the models, of which we ran three chains with 80,000 MCMC iterations each, of which 40000 were discarded as burn-in samples. Samples were thinned at a rate of 30 to obtain the posterior distributions for the occupancy and detection probabilities.

I used the “simocc” function in the *AHMBook* R package (Kéry et al., 2017) as a basis to create simulated datasets for each scenario. The function “simOcc” simulate data for static occupancy models under a wide range of conditions. It generates a matrix of M sites (rows) during J surveys (columns) with each element containing a value of 0 (non-detection) or 1 (detection/non-detection). In the simulations, the number of sites and number of survey replicates matched the scenarios combinations stated above.

I used JAGS v3.4.0 (Plummer et al., 2003) called through the *JAGSUI* R (Kellner et al., 2021) package, launched from R to implement Bayesian estimation of model parameters via Markov chain Monte Carlo (MCMC) samples from posterior distributions. For the Bayesian MCMC approach, I ran 1000 simulations for each of the 9 scenarios. For each simulation I ran 3 chains with 80000 MCMC iterations each, of which 40000 iterations were discarded as burn-in, where samples were thinned at a rate of 20 to obtain posterior distributions for all model
parameters. Vague priors were used after assessing their effect on model results because there was no published occupancy or density study using distance transect sampling of Arctic ground squirrels to provide prior information. I determined whether the models converged using traceplots and the Gelman-Rubin statistic where values $\hat{r}<1.1$ indicated adequate sampling of the posterior distribution (Gelman and Rubin, 1992). Boxplots were used to visualize and compare the distribution of mean standard error (MSE), bias and variance among detection and occupancy/distance sampling parameters and sampling scenarios. I used ANOVAs and post hoc Tukey tests to assess whether apparent differences in MSE estimated parameters were statistically significant. For each parameter I extracted the posterior mean and 95% credible intervals. As part of this analysis, I also compared maximum-likelihood estimates with the Bayesian MCMC approach here.

2.4 RESULTS

2.4.1 Occupancy Modeling

I detected Arctic ground squirrels on 60 of 150 visits (naive occupancy or the proportion of sites at which target species was detected = 0.40) across the three years of the study. They were detected on 33 of 150 (22%) sites surveyed in 2017, 27 of 53 (51%) surveyed in 2018, and 24 of 48 (50%) surveyed in 2019. They were detected in more than one year at 27 of 52 (52%) of sites that were resampled (>1 site visit).

The goodness-of-fit test indicated high overdispersion ($\hat{c}=2.77$) and showed evidence of lack-of-fit for the global model containing all covariates. I calculated QAIC scores using this value for subsequent model selection of the full generated model set ($n=44$; including the null model). It was clear that there was no single best model; 7 models in the set differed by only small amounts as defined by the information criterion ($\Delta$QAIC <2). The results of the model averaging are summarized in Table 2.2. In the detection sub-model, model-averaged results suggested little evidence that any of the variables predicted to affect detection probability were significant as all confidence intervals for the parameter estimates included zero. For the occupancy sub-model, only the five-year mean temperature (fivetm) had confidence intervals that excluded 0. The model-averaged estimate with shrinkage was -1.68, indicating a negative effect.

However, it could still be relevant to use the model-averaged estimates to predict point
estimates for occupancy and detection probability. Mean Arctic ground squirrel occupancy decreased with an increase in max soil depth and the five year mean temperature (fivetm) and increased with terrain ruggedness (vrm) and mean precipitation (pmean; Figure 2.3). The relationship between year and occupancy changed across the years, suggesting annual differences in factors that varied within sampling periods. Occupancy was highest at the beginning of the sampling period (year 1) and was at its lowest during the middle of the sampling period (year 2). Average site-occupancy ranged from 0.25-0.90 depending on these covariates. Mean Arctic ground squirrel detection probability increased with cloud cover (cloud) and decreased with percent site of high vegetation stature (high; Figure 2.4).

The candidate set of the most probable detection sub-models (QAIC weights summing to 0.93) included the null model, percent cloud cover, and high vegetation stature (Table 2.3). The last model included an interaction term between cloud and high and received a negligible weight. The confidence set of the most probable occupancy sub-models included soilmax, fivetm, minimum temperature of the coldest month (bio6), slope, vrm, pmean, and the year that the site was sampled (year; Table 2.3). The top QAICc-ranked covariates in the occupancy sub-model were soilmax, fivetm, and year. Mean occupancy was estimated as 0.54 and mean detection probability was estimated as 0.83.

2.4.2 Distance Sampling Modeling

I counted Arctic ground squirrels on 68 of 157 sites (43%) where transects were conducted across the three years of the study. The maximum number of observations per site was 12, and the mean number of observations per site was 6.8 individuals. The goodness-of-fit test indicated moderate overdispersion (ĉ=1.22) and showed evidence of lack-of-fit for the most global model containing all covariates. I calculated QAIC scores using this value for subsequent model selection of the full generated model set. Of the full generated model set (n=20), there was no single best model; several models (n=5) in the set differed by only small amounts as defined by the information criterion (Δ QAIC <2) suggesting model uncertainty (Table 2.4). The model-averaged predictions in the detection sub-model suggests that the only predictor that included evidence to support an effect was time of sampling (tos). The model-averaged estimate with shrinkage was 0.12, indicating a small positive effect for the density sub-model. For the density (λ) submodel, the model-averaged predictors, slope, mean precipitation, and sampling year 2-3 included evidence to support an effect on density. The following covariates model-
averaged estimates with shrinkage indicated a positive effect: slope, and mean precipitation. The covariates with model-averaged estimates indicating a negative effect included the second and third year of sampling.

The candidate set of the most probable detection submodel (QAICc weights summing to 0.8) included models with tos, interaction between tos and cloud cover, an interaction between tos and date of sampling, and all three covariates in one model (Table 2.5). The confidence set of the most probable density submodel (QAIC weights summing to .86) included a model with the covariates fivetm, slope, vrm, bio6, mean temperature of coldest month (bio11), pmean and the second and third year of sampling (Table 2.5). Arctic ground squirrel detection along the transect was considered ‘perfect’ (probability of detection is 100%) up to about 140 meters (Figure 2.5). The estimated detection probability for the entire hypothetical population was 0.82. The estimated population density for the sampling area was 19.5 individuals/hectare.

2.4.3 Spatial capture-recapture

Of the 70 individuals trapped over the two field seasons (year 1 = 40, year 2 = 39), only 4 were recaptures. Because of a low number of recaptures, it was impossible to conduct a spatial capture-recapture analysis.

2.4.4 Simulations of survey effort effects

Overall, the distribution of the MSE estimated p and ψ parameters decreased as the number of replicate surveys and sites increased (Figure 2.6). The MSE of the occupancy component was higher than the detection component for each sampling scenario (occupancy; Figure 2.6). Within the occupancy component, the most noticeable decreases in MSE occurred between the lowest sampling scenarios n1 (25 sites) and the highest site number scenario, n3 (150 sites). The MSE is noticeably lower for all scenarios that included sampling 150 sites (n3) (Figure 2.6). For the detection component, this difference is less pronounced as decrease in MSE across each sampling scenario was less pronounced indicating less improvement in model performance (detection; Figure 2.6). The most notable decrease in MSE happens when you increase the number of survey visits from one to two, even when only the number of sites is at the lowest setting (n = 25). Parameter estimates met the convergence criteria. Based on Ř values (all <1.1) and traceplots, parameters were deemed reliable. ANOVAs revealed significant differences between the sampling scenario groups (Table 2.6). Within the occupancy
component, Tukey’s test showed that the change in MSE was not significant for scenarios where the number of surveys increased by one but number of sites decreased by one (e.g., n2s1-n1s3) and vice-versa (e.g., n3s1-n2s3) (Figure 2.7). This pattern also held true within the detection component (Figure 2.7).

2.5 DISCUSSION

With the consent of the National Park Service, I developed and tested a monitoring protocol using occupancy, N-mixture models, and SCR to address the unmet objectives of the CAKN inventory and monitoring program for the Arctic ground squirrel. I used the protocol to provide needed information to assess the status of the species, and for understanding the underlying mechanisms that drive its habitat relationships and population size estimates. This study compared noninvasive and invasive methods to estimate the probability of occupancy and density of the Arctic ground squirrel in Denali National Park and Preserve. This analysis also highlighted the habitat attributes that most impact occupancy and density.

Neither cloud cover nor the percent of the site with high vegetation stature (\(>200\) cm), had an important effect on the detection probability of Arctic ground. This outcome is surprising given the behavior and ecology of the species. As an Arctic species, the Arctic ground squirrel is especially vulnerable to heat stress caused by hot weather and a warming climate. Exposing Arctic ground squirrels to temperatures above 30 °C for 55 minutes is lethal (Sullivan and Mullen, 1954). Basic mechanisms for cooling off are lacking (Gilg et al., 2012), and they avoid heat through behavioral thermoregulation by retreating into their burrows (Watton and Keenleyside, 1974; Long et al., 2005). I used cloud cover as a proxy for solar radiation during the survey, predicting higher cloud cover would increase foraging time above ground and thus improve availability for species to be detected. It might be that visual inspection of cloud cover once per day was insufficiently fine-grained as a predictor variable. Or cloud cover might not be a good enough proxy for heat.

I also tested vegetation covers effect on detection. Climate change is leading to substantial vegetation changes in the Arctic. Woody shrub and boreal forest species, which provide poor habitat for Arctic ground squirrels (Eddingsaas et al., 2004; Gillis et al., 2005a; Donker and Krebs, 2011), are increasing in DENA (Roland, 2004). I predicted that taller vegetation would inhibit the ability of Arctic ground squirrels to detect predators and thus
increase predator avoidance behavior (e.g., more time in the burrow) in these habitats and ultimately decrease species detection. However, I found that vegetation cover had no effect on detection probability. This may be a result of Arctic ground squirrel alarm-calling behavior (Melchior, 1971) increasing my ability to detect the species when I initially didn’t detect them. In most sampling scenarios, Arctic ground squirrels were detected because they emitted an alarm call due to our presence. Future research could incorporate instances when alarm calls led to detection, or experiment with alarm calls and their impact on detection probability.

In Denali National Park, Arctic ground squirrels occupied about half of the sites that I surveyed (0.54). Their occupancy was best predicted by the five-year temperature mean. This may mean that Arctic ground squirrels, which exhibit physiological traits that are plastic (Sheriff et al., 2011, 2013; Williams et al., 2012), are impacted by temperatures in aggregate rather than during the duration of one season (mean temperature of warmest and coldest quarter) or at the extreme (i.e., max temperature of warmest month and min temperature of the coldest). This result suggests that increasing warming may create unsuitable habitat conditions to support a long-term population of this vital Sign, and likely other Arctic species, in Denali National Park and Preserve.

Arctic ground squirrel detection probability in the distance transect models was positively associated with the time of sampling. This outcome was not surprising. The Arctic ground squirrel is a small mammal, and likely is more susceptible to rapid changes in body temperature due to high body surface to volume ratios and less insulation. Arctic ground squirrels are thus required to rapidly make adjustments to their body temperature depending on the thermal environment, by behaviors such as retreating to their burrows or using shade. Arctic ground squirrel greatest percent of time spent above their burrow has been reported to increase during the morning until about noon, and then decreased as the day progressed (Long et al., 2005). This is correlated with how temperature varies during the day and night due to incoming sunlight.

My study estimated the density of arctic ground squirrels as 19.5/ha. This is at the high end compared to previous studies that reported population size estimates of the Arctic ground squirrel (eWerner, 2015; Donker and Krebs, 2011; Hik et al., 2001; Lacey et al., 1997). Density was positively associated with slope and mean precipitation. Previous studies have reported that greater numbers of burrows and burrow systems were associated with steeper slopes (Wheeler et al., 2015) which should support a higher number of individuals. Steeper slopes may
be associated with better visibility (Barker and Derocher, 2010), and an increase in burrow activity due to drainage that reduces the waterlogging of soils (Wheeler and Hik, 2013). Precipitation may be associated with forbs which are important forage for Arctic ground squirrels (Batzli and Sobaski, 1980; McLean, 1985; Gillis et al., 2005b). Altered precipitation patterns and shifts in seasonality may significantly impact plants (Ernakovich et al., 2014). While the potential growing season is increasing, plants will need to advance the timing of phenological events which could be constrained by water availability thus impacting habitat quality.

Estimating site use and population size accurately and precisely is critical to wildlife monitoring. Using simulations, I aimed to determine whether different survey designs would need to increase or could reduce the amount of sites/replicate surveys to provide reliable estimates based on the results of this study. Simulations from the occupancy model revealed that models using a single survey replicate always provided unreliable estimates of detection and occupancy, however, results could be improved by increasing the number of sites surveyed. Survey designs with more sites and replicates always outperformed those with fewer. However, the models only improved marginally along the upper ends of survey replicates and sites. This is especially seen with the detection sub-model perhaps due to high detection probabilities in this species.

The large uncertainty in the density estimators may be a result of failing to meet assumptions underlying the models. Distance sampling is difficult to implement in mountainous terrain, especially where visual coverage of the entire survey area is insufficient. One of the key assumptions is that animals are randomly and evenly distributed throughout the survey area. Arctic ground squirrels were seldomly found distributed independent of one another. Methods moving forward could accommodate group structure by redefining the analysis to be on the group rather than individuals and including a covariate for group size (Kery and Royle, 2015). Additionally, there may be missing covariates that are more functionally related to both parameters of interest. Due to constraints on sample size, I was not able to include all the same covariates into both models to conserve degrees of freedom for subsequent model selection. Thus, I opted to include covariates that I thought reflected the most important plausible hypothesis for each parameter of interest.

Additionally, I was unable to estimate Arctic ground squirrel density with spatial capture-recapture due to insufficient recaptures. While sampling design for SCR can be optimized, it is
also unclear whether trapping arrays can be accommodated in the park. This has been an issue in other sciurid studies (TL Morelli, pers comm). Moreover, trapping in national parks, especially wilderness, is complicated logistically and ethically. For a species that is “trap shy”, the cost in time and equipment, not to mention the invasive nature of this technique, makes this monitoring method less preferable.

2.6 CONCLUSIONS

With over 150 sites sampled over three years, this study on spatial distributions and density of Arctic ground squirrels through occupancy modeling and distance transect sampling is one of the first conducted in Denali National Park and Preserve. Due to their patchy distribution and difficulty to trap, I recommend the collection of presence-absence data to estimate occupancy as the proper measure to monitor Arctic ground squirrels in Denali National Park and Preserve. The properties for estimating population size through spatial capture-recapture methods may not be met as demonstrated through an inability to recapture individuals. Additionally, the transect sampling may fail to account for the species availability during the survey.

Addressing imperfect detection in distance sampling is complicated both by species availability and perceptibility (Marsh and Sinclair, 1989; Amundson et al., 2014). While Arctic ground squirrels may not be rare or elusive, as hibernating burrowers they are certainly not always available for detection. Time spent above ground for Arctic ground squirrels may be significantly influenced by time of day, habitat, and weather conditions all of which I was not able to accommodate in the distance sampling availability submodels (Nichols et al., 2009). Failure to take this into account can negatively bias population size estimates (Chandler et al., 2015).

While population size estimates like density are more data rich and can be used to make inferences about presence and absence, extreme caution should be taken to not draw presumed population size patterns from predictions of the probability of occupancy (Tempel and Gutierrez, 2013; Wilson and Schmidt, 2015; Casner et al., 2014; Ellis et al., 2014). As evidenced by the data, the hierarchical occupancy and distance sampling models neither shared equally the same predictors or relationships. It is likely that occupancy and density for Arctic ground squirrels are determined by different processes. Previous studies on Arctic ground squirrels, revealed concurrent density-dependent and density-independent (Karels and
Boonstra, 2000), and food availability effects on abundance (Hubbs and Boonstra, 1997). There is yet to be an analysis comparing how reliable occupancy analyses are for characterizing abundance patterns and vice-versa for Arctic ground squirrels. Future analyses should focus on formulating competing hypotheses based on the predictors from this study and assess potential mismatch between predictors of occupancy and density.
FIGURES

Figure 2.1: Map of entire study extent. Locations for study sites (nodes) are represented by black boxes. The Denali Park Road that runs lengthwise through part of the park is represented by the black line.
Figure 2.2: Example mini-grid design of the permanent vegetation monitoring plots in Denali National Park and Preserve. The sampling frame is used for the collocation and integration of monitoring efforts occurring across the park. The survey sites are represented by red dots arranged in five rows of five points, spaced 500 meters apart.
Figure 2.3: Model-averaged predicted relationships between occupancy of Arctic ground squirrels and top supported covariates in the top $<2 \Delta$ QAICc model set. Solid lines represent model-average predicted occupancy and the shaded region and error bars delineate upper and lower 95% CI.
Figure 2.4: Model-averaged predicted relationships between detection probability of Arctic ground squirrel and supported covariates in the top $\Delta Q$Ak model set. Solid lines represent model-averaged predicted detection and shaded regions delineate upper and lower 95% CI.
Figure 2.5: Detection probability as a function of distance from hazard line transect model fitted to Arctic ground squirrels in Denali National Park and Preserve.
Figure 2.6: Boxplots showing mean square error (MSE) of occupancy probability (A) and detection probability (B) across 9 simulated sampling designs. Shaded panels are included to assist with comparisons between sampling designs. The red diamonds are the group means.
Figure 2.7: Contains confidence intervals for the difference in the means for all 36 pairs of groups. This is the graphical display of the multiple (pairwise) comparisons using Tukey’s HSD for the scenario data. Any confidence intervals that do not overlap with zero provide evidence of a difference in group means.
Table 2.1: Description of covariates considered during exploratory analysis for predicting occurrence and density of Arctic ground squirrels in Denali National Park and Preserve.

<table>
<thead>
<tr>
<th>Functional Group Covariate</th>
<th>Abbrev.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Climate</strong></td>
<td></td>
</tr>
<tr>
<td>Max Temperature of Warmest Month</td>
<td>bio4</td>
</tr>
<tr>
<td>Min Temperature of Coldest Month</td>
<td>bio6</td>
</tr>
<tr>
<td>Mean Temperature of Warmest Quarter</td>
<td>bio10</td>
</tr>
<tr>
<td>Mean Temperature of Coldest Quarter</td>
<td>bio11</td>
</tr>
<tr>
<td>Five Year Precipitation Mean</td>
<td>pmean</td>
</tr>
<tr>
<td>Five Year Temperature Mean</td>
<td>fivetm</td>
</tr>
<tr>
<td><strong>Topography</strong></td>
<td></td>
</tr>
<tr>
<td>Elevation</td>
<td>elev</td>
</tr>
<tr>
<td>Slope</td>
<td>slope</td>
</tr>
<tr>
<td>Vector Ruggedness Measure</td>
<td>vrm</td>
</tr>
<tr>
<td>Max Soil Depth</td>
<td>soilmax</td>
</tr>
<tr>
<td><strong>Vegetation</strong></td>
<td></td>
</tr>
<tr>
<td>Percent of Plot Covered in Low Vegetation Stature (&lt; 50 cm)</td>
<td>low</td>
</tr>
<tr>
<td>Percent of Plot Covered in Medium Vegetation Stature (50 – 200 cm)</td>
<td>med</td>
</tr>
<tr>
<td>Percent of Plot Covered in High Vegetation Stature (over 200 cm)</td>
<td>high</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
</tr>
<tr>
<td>Sampling Year</td>
<td>year</td>
</tr>
<tr>
<td>Julian Day</td>
<td>yday</td>
</tr>
<tr>
<td>Time of Sampling</td>
<td>tos</td>
</tr>
<tr>
<td>Cloud Cover</td>
<td>cloud</td>
</tr>
</tbody>
</table>
Table 2.2: Model-averaged parameter estimates, unconditional standard errors, and 95% confidence limits of coefficients of covariates in occupancy models. Parameter estimates whose model-averaged 95% CI did not overlap zero are displayed in bold.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>β</th>
<th>Unconditional SE</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Detection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cloud</td>
<td>0.05</td>
<td>0.21</td>
<td>-0.37</td>
<td>0.46</td>
</tr>
<tr>
<td>high</td>
<td>-0.04</td>
<td>0.19</td>
<td>-0.41</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>Occupancy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pmean</td>
<td>0.04</td>
<td>0.22</td>
<td>-0.4</td>
<td>0.48</td>
</tr>
<tr>
<td>fivetm</td>
<td>-1.68</td>
<td>0.6</td>
<td>-2.85</td>
<td>-0.51</td>
</tr>
<tr>
<td>vrm</td>
<td>0.04</td>
<td>0.18</td>
<td>-0.32</td>
<td>0.41</td>
</tr>
<tr>
<td>soilmax</td>
<td>-0.77</td>
<td>0.51</td>
<td>-1.77</td>
<td>0.23</td>
</tr>
<tr>
<td>high</td>
<td>-0.03</td>
<td>0.17</td>
<td>-0.36</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1.5</td>
<td>1.13</td>
<td>-3.72</td>
<td>0.72</td>
</tr>
<tr>
<td>3</td>
<td>-0.44</td>
<td>0.96</td>
<td>-2.32</td>
<td>1.45</td>
</tr>
</tbody>
</table>
Table 2.3: Results of QAICc-based model selection for covariates influencing probability of detection (p) and occupancy (ψ) of Arctic ground squirrels. Year was included as a term in all models whereas all other covariates were allowed to vary. K = number of parameters.

<table>
<thead>
<tr>
<th>Model</th>
<th>K</th>
<th>QAICc</th>
<th>ΔQAICc</th>
<th>Wt</th>
<th>Quasi L. L.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Detection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>~ 1 ~ Ψ</td>
<td>15</td>
<td>107.84</td>
<td>0</td>
<td>0.54</td>
<td>-37.13</td>
</tr>
<tr>
<td>~ cloud ~ Ψ</td>
<td>16</td>
<td>109.68</td>
<td>1.84</td>
<td>0.22</td>
<td>-36.79</td>
</tr>
<tr>
<td>~ high ~ Ψ</td>
<td>16</td>
<td>110.13</td>
<td>2.29</td>
<td>0.17</td>
<td>-37.02</td>
</tr>
<tr>
<td>~ cloud + high ~ Ψ</td>
<td>17</td>
<td>112.09</td>
<td>4.25</td>
<td>0.07</td>
<td>-36.32</td>
</tr>
<tr>
<td><strong>Occupancy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>~ p ~ soilmax + fivetm + year</td>
<td>9</td>
<td>95.45</td>
<td>0</td>
<td>0.12</td>
<td>-38.08</td>
</tr>
<tr>
<td>~ p ~ soilmax + vrm + fivetm + year</td>
<td>10</td>
<td>97.07</td>
<td>1.62</td>
<td>0.05</td>
<td>-37.75</td>
</tr>
<tr>
<td>~ p ~ fivetm + year</td>
<td>8</td>
<td>97.34</td>
<td>1.89</td>
<td>0.05</td>
<td>-40.16</td>
</tr>
<tr>
<td>~ p ~ soilmax + fivetm + pmean + year</td>
<td>10</td>
<td>97.38</td>
<td>1.93</td>
<td>0.05</td>
<td>-37.9</td>
</tr>
<tr>
<td>~ p ~ soilmax + high + fivetm + year</td>
<td>10</td>
<td>97.57</td>
<td>2.11</td>
<td>0.04</td>
<td>-37.99</td>
</tr>
<tr>
<td>~ p ~ soilmax + low + fivetm + year</td>
<td>10</td>
<td>97.62</td>
<td>2.17</td>
<td>0.04</td>
<td>-38.02</td>
</tr>
<tr>
<td>~ p ~ soilmax + bio6 + fivetm + year</td>
<td>10</td>
<td>97.63</td>
<td>2.18</td>
<td>0.04</td>
<td>-38.02</td>
</tr>
<tr>
<td>~ p ~ soilmax + slope + fivetm + year</td>
<td>10</td>
<td>97.67</td>
<td>2.22</td>
<td>0.04</td>
<td>-38.05</td>
</tr>
<tr>
<td>~ p ~ soilmax + medium + fivetm + year</td>
<td>10</td>
<td>97.68</td>
<td>2.23</td>
<td>0.04</td>
<td>-38.05</td>
</tr>
<tr>
<td>~ p ~ soilmax + bio5 + fivetm + year</td>
<td>10</td>
<td>97.75</td>
<td>2.29</td>
<td>0.04</td>
<td>-38.08</td>
</tr>
<tr>
<td>~ p ~ low + fivetm + year</td>
<td>9</td>
<td>98.43</td>
<td>2.98</td>
<td>0.03</td>
<td>-39.57</td>
</tr>
<tr>
<td>~ p ~ vrm + fivetm + year</td>
<td>9</td>
<td>99.02</td>
<td>3.57</td>
<td>0.02</td>
<td>-39.87</td>
</tr>
<tr>
<td>~ p ~ pmean + bio5 + year</td>
<td>9</td>
<td>99.1</td>
<td>3.64</td>
<td>0.02</td>
<td>-39.91</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>~ p ~ vrm + bio6 + high + year</td>
<td>10</td>
<td>112.2</td>
<td>16.75</td>
<td>0</td>
<td>45.31</td>
</tr>
</tbody>
</table>
Table 2.4: Results of AICc-based model selection for covariates influencing probability of detection ($p$) and density ($\lambda$) of Arctic ground squirrels. Year was included as a term in all models whereas all other covariates were allowed to vary. $K =$ number of parameters. The $\phi$ parameter was held at it’s intercept in all models.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>$\beta$</th>
<th>Unconditional SE</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Detection Probability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tos</td>
<td>0.12</td>
<td>0.04</td>
<td>0.03</td>
<td>0.2</td>
</tr>
<tr>
<td>yday</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.05</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Density</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fivetm</td>
<td>-0.69</td>
<td>0.4</td>
<td>-1.46</td>
<td>0.09</td>
</tr>
<tr>
<td>slope</td>
<td>-0.36</td>
<td>0.08</td>
<td>0.21</td>
<td>0.51</td>
</tr>
<tr>
<td>vrm</td>
<td>0.05</td>
<td>0.05</td>
<td>-0.05</td>
<td>0.16</td>
</tr>
<tr>
<td>bio6</td>
<td>0.05</td>
<td>0.1</td>
<td>-0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>bio11</td>
<td>0.08</td>
<td>0.15</td>
<td>-0.21</td>
<td>0.37</td>
</tr>
<tr>
<td>bio10</td>
<td>-0.22</td>
<td>0.29</td>
<td>-0.79</td>
<td>0.36</td>
</tr>
<tr>
<td>elev</td>
<td>0.15</td>
<td>0.28</td>
<td>-0.04</td>
<td>0.71</td>
</tr>
<tr>
<td>pmean</td>
<td>0.62</td>
<td>0.19</td>
<td>0.26</td>
<td>0.99</td>
</tr>
<tr>
<td>year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1.34</td>
<td>0.58</td>
<td>-2.47</td>
<td>-0.21</td>
</tr>
<tr>
<td>3</td>
<td>-1.29</td>
<td>0.29</td>
<td>-1.85</td>
<td>-0.73</td>
</tr>
</tbody>
</table>
Table 2.5: Results of QAICc-based model selection of probability of detection (p) and density (λ) for Arctic ground squirrels in Denali National Park and Preserve, AK.

<table>
<thead>
<tr>
<th>Model</th>
<th>K</th>
<th>QAICc</th>
<th>ΔQAICc</th>
<th>Wt</th>
<th>Qau_L.L.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda \sim 1 \sim \text{tos}$</td>
<td>15</td>
<td>502.92</td>
<td>0</td>
<td>0.49</td>
<td>-235.98</td>
</tr>
<tr>
<td>$\lambda \sim 1 \sim \text{cloud + tos}$</td>
<td>16</td>
<td>504.99</td>
<td>2.06</td>
<td>0.17</td>
<td>-235.79</td>
</tr>
<tr>
<td>$\lambda \sim 1 \sim \text{tos + yday}$</td>
<td>16</td>
<td>505.37</td>
<td>2.44</td>
<td>0.14</td>
<td>-235.98</td>
</tr>
<tr>
<td>$\lambda \sim 1 \sim \text{cloud + tos + yday}$</td>
<td>17</td>
<td>507.44</td>
<td>4.52</td>
<td>0.05</td>
<td>-235.78</td>
</tr>
<tr>
<td>Occupancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fivetz + slope + vrm + bio6 + year $\sim 1 \sim p$</td>
<td>9</td>
<td>95.45</td>
<td>0</td>
<td>0.29</td>
<td>-238.44</td>
</tr>
<tr>
<td>elev + vrm + bio6 + bio10 + year $\sim 1 \sim p$</td>
<td>10</td>
<td>97.07</td>
<td>1.62</td>
<td>0.22</td>
<td>-241.18</td>
</tr>
<tr>
<td>fivetz + slope + bio6 + bio10 + year $\sim 1 \sim p$</td>
<td>8</td>
<td>97.34</td>
<td>1.89</td>
<td>0.21</td>
<td>-241.2</td>
</tr>
<tr>
<td>fivetz + slope + vrm + bio6 + bio10 + year $\sim 1 \sim p$</td>
<td>10</td>
<td>97.38</td>
<td>1.93</td>
<td>0.15</td>
<td>-239.12</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>elev + slope + vrm + bio6 + bio11 + year $\sim 1 \sim p$</td>
<td>10</td>
<td>97.62</td>
<td>2.17</td>
<td>0</td>
<td>-613.49</td>
</tr>
</tbody>
</table>
CHAPTER 3

GENOMIC DIVERSITY AND POPULATION STRUCTURE OF ARCTIC GROUND SQUIRRELS IN DENALI NATIONAL PARK AND PRESERVE

3.1 ABSTRACT

Understanding the impact of climate change on Arctic species will require comprehensive assessments of how species interact with their environment. While important insights have been gained related to how Arctic biodiversity will experience shifts in their distributions and population size, little is known regarding contemporary genetic diversity of most species which could inform the management of wildlife populations, including identifying populations and units for conservation, assessing the potential of populations to adapt, and the factors that influence such potential. Here, I provide an example of this approach for Arctic ground squirrels (Urocitellus parryii) within Denali National Park and Preserve. Ongoing monitoring efforts have specifically focused on estimating their abundance and describing patterns of their spatial distribution, but have not provided insights into demographic history of the sampled populations and potential consequences to fitness. Here I report on the use of nuclear DNA (17,460 SNPs) to delineate genetic structure and assess genetic diversity to guide future monitoring efforts. I identified three major genetic clusters across the five populations, with medium genetic diversity and clear genetic differentiation among the populations and clusters. Monitoring efforts in the future should focus on expanding genetic sampling of different populations and inclusion of analyses that estimate potential for long-term persistence in the face of climate change.

3.1.1 Keywords:

Urocitellus parryii; Arctic; mammal; connectivity; genomics; genetic differentiation; genomic diversity; SNPs
3.2 INTRODUCTION

While protecting species via habitat conservation remains a key objective under climate change, other objectives such as maintaining genetic variability are emerging. Understanding how landscape characteristics influence gene flow and genetic differentiation can be critical to improving our knowledge of how declines in loss of suitable habitat may impact species vulnerability (Wasserman et al., 2012; Cushman et al., 2006; Castillo et al., 2014). This is especially true for Arctic systems, where the emergence of a new climate system is rapidly changing the landscape (Duncan et al., 2020; Landrum and Holland, 2020). From the perspective of conservation practitioners in the Arctic, climate change is the most serious threat to long-term ecosystem functioning and biodiversity (Meltofte et al., 2013). However, when addressing the vulnerability of species to climate change, most assessments treat the species of conservation concern as a single unit, ignore intraspecific genetic variation, and/or ignore the impacts of climate change on intraspecific genetic variation (Shafer et al., 2015; Orgeret et al., 2021; van Beest et al., 2021).

An individual species’ ability to respond to climate change will depend in part on standing genetic variation and heterogeneity in the distribution of adaptive alleles (Parmesan, 2006; Hoffmann and Gröning, 2011; Chhatre et al., 2019). These aspects are in turn driven by a species’ dispersal capacity (Lenormand, 2002; Kawecki and Ebert, 2004) and the effect of distribution and connectivity of populations on gene flow (Waterhouse et al. 2018). Maintaining genetic variation and connectivity of individuals and populations is important as it allows species to adapt to environmental change (Sexton et al., 2014). In many cases, climate change is expected to reduce genetic diversity, through shifts in a species range and changes in the local selection regime. Simulation and empirical studies demonstrate the loss of genetic diversity through range expansions, contractions, and shifts (Arenas et al., 2012; Alsos et al., 2012; Rubidge et al., 2012). Measurements of genetic diversity and connectivity among habitat patches via gene flow are essential particularly to highlight where adaptive capacity may be lacking (Nicotra et al., 2015; Beever et al., 2016). The application of genomic approaches can inform all these issues and provide critical information for designing management strategies to address them (Allendorf et al. 2010).

To develop appropriate conservation and management strategies, it will be critical to estimate species genetic diversity, genetic structure of a population, and the role of landscape
features in shaping those patterns. Despite an increasing number of studies investigating declining populations, contracting ranges, and local extirpations, many species lack any measurements of genetic diversity which would help identify their likelihood of future extinction (Shafer et al., 2015). Thus, the evolutionary potential of many species to respond to climate change, as well as the usefulness of current species distribution models, is uncertain (Scoble and Lowe, 2010; Araujo and Peterson, 2012).

Populations of the Arctic ground squirrel (Urocitellus parryii) living in Denali National Park and Preserve, AK (hereafter DENA), provides an opportunity to use genetic diversity as a metric for identifying species of conservation concern in the Arctic. The Arctic ground squirrel has a Holarctic geographic range and is distributed across alpine, tundra, meadows, and boreal forest habitats. Being both Arctic and alpine distributed with specific habitat requirements, it serves as an indicator species for climate and other global change (Wheeler and Hik, 2013), is useful for understanding fine-scale responses to environmental change, and appears particularly vulnerable to climate change (Wheeler and Hik, 2013). As a burrowing hibernator, the individuals of this species are highly specialized in occupying well-drained soils. In the center-most range in Alaska, populations are patchily distributed in open tundra and meadows in subalpine areas.

While the physiology and population biology of Arctic ground squirrels have been extensively studied (Sheriff et al. 2012; Frare et al., 2018, 2019), major gaps remain in our knowledge on their dispersal capability, and how the connectivity among habitat patches via gene flow affects their genetic diversity and evolutionary potential. The historical landscape, characterized by formation of glacial refugia through Quaternary climatic oscillations, is thought to have significantly affected the demographic history, population genetic structure, and gene flow of Arctic ground squirrels in North America (Eddingsaas et al., 2004; Galbreath et al., 2011).

Based upon analyses of phylogeographic research on how Pleistocene glacial cycles influenced the spatial distribution of genetic variation for alpine and Arctic-adapted species, Lanier et al. (2015a) suggested that Arctic ground squirrels are already experiencing population and/or range declines and exhibit a long history of isolation in glacial refugium. The genetic consequences of refugial survival may have resulted in reduced genetic diversity but increased divergence in populations within Alaska. Fragmentation of habitat between these glacial cycles may have resulted in reductions in effective population sizes and disruptions in gene flow,
leading to an acceleration of genetic drift. The contraction and fragmentation of the habitat of this species by the encroachment of shrubs and forests moving rapidly northward and upslope (Wheeler et al. 2013) presents a significant conservation threat that can create barriers to movement and gene flow.

Furthermore, Arctic ground squirrels exhibit ecological characteristics which may have important consequences for expectations of genetic diversity and structure. In DENA, the habitat occupied by these species sits within a landscape matrix dominated by forests at elevations less than 760 meters, scrub vegetation and spruce woodland in the subalpine zone (760-1070 meters) and low tundra in the alpine zone (above 1070 m) and is fragmented by natural barriers such as glacially fed rivers and anthropogenic barriers such as the park road. In DENA, Arctic ground squirrels can be commonly found patchily distributed in the subalpine zone and low tundra in the alpine zone. Because of their subterranean habits, the Arctic ground squirrel needs well drained soils. Thus, their specialized habitat requirements can create patchy distributions. These patchy populations in turn can be connected or isolated depending on the distance or intervening matrix. Indeed, species with narrow habitat requirements may display genetic diversity patterns in line with theories of isolation by barriers or resistance (McRae and Beier, 2007). Even those populations in suitable habitats may be susceptible to density fluctuations and turnover of individuals due to increases in predation pressure after cyclical crashes in regional hare populations. In addition, Arctic ground squirrels exhibit a high degree of philopatry with low dispersal capacities and colonization capacities. Therefore, they may exhibit low genetic variation and show strong genetic structure sustained by low rates of migration and gene flow.

My goal in this study was to gain first insights into the population genomics of Arctic ground squirrels located at the center of its geographical range in Alaska using single nucleotide polymorphisms (SNPs) that will inform the basic landscape and molecular ecology of the species. Given the patchy distribution of suitable habitat and the inferred limited individual movement and dispersal of Arctic ground squirrels, I hypothesize that: (1) Arctic ground squirrels in DENA have low levels of genetic diversity; and 2) Arctic ground squirrels populations are highly differentiated.

3.3 METHODS:
3.3.1 Study area and sampling

The study occurred in Denali National Park and Preserve, AK (Figure 3.1.), an area encompassing roughly 6.1 million acres, in interior Alaska. During the summers of 2018-2019, Arctic ground squirrels were live trapped at 5 sampling grids (Figure 3.1). The sampling units at Primrose, Tributary Creek, Divide Mountain, Polychrome and Gorge Creek ranged from 8 km to 28 km from each other. Arctic ground squirrels were sampled using live collapsible Tomahawk traps (Tomahawk Live-trap Co., WI, USA; 19X6X6 cm) placed on top of known occupied burrows. Trapping locations were targeted through a combination of spatial modeling that included distance from each other and for representing environmental variations related to their ecological habitat. Trapped Arctic ground squirrels were weighed, sexed and individuals were marked with pre-sterilized PIT tags, and a DNA sample was collected in the form of roughly 2 mm metal ear punch. Squirrels were subsequently released at the same location that they were caught. Tissues were stored in 95% ethanol in the field, and then in -20 C freezers until DNA extraction. All handling procedures were approved by UMass Institutional Animal Care and Use Committee under protocol #1278 19-0023, the Alaskan Fish and Game under protocol 18-108, and the National Park Service DENA-00934.

3.3.2 DNA extraction, library construction, and SNP calling

Genomic DNA was extracted from preserved tissue samples using the BestRAD illumina library prep protocol and the NEBnext Ultra 11 Dna library prep kit for Illumina (E7645) to isolate genome-wide fragments of DNA for sequencing. Arctic ground squirrel genomic DNA samples were examined on a NanoDrop 2000 Spectrophotometer (ThermoScientific) to identify samples with appropriate concentration and purity for next-gen sequencing. Genomic DNA was plated and sent to Novogene in Sacramento, CA. for sequencing using restriction site-associated DNA sequencing (RADseq; (Davey and Blaxter, 2010)). Using a reference-guided bioinformatic pipeline, Arctic ground squirrel sequences were aligned to a reference genome and then genotyped using Freebayes creating a VCF SNP file output. Calling quality was assured through post processing exploratory data analysis and various filtering strategies. VCF files were then converted into the class objects genind, genpop and genlight for downstream analysis. All subsequent genomic analyses were completed in R 4.1.2 (R Core Team, 2021) unless otherwise noted. The molecular marker of choice in this analysis was SNPS. SNPS are
variations in DNA sequences that occur through substitutions at a single base-pair, and can differ at the same position between individuals, populations, and species.

### 3.3.3 Genomic analyses

#### 3.3.3.1 Population structure

I explored the population structure of Arctic ground squirrels first with a two-stage nonparametric approach using principal components analysis (PCA) (Bryc et al., 2010) and discriminant analysis of principal components (DAPC) (Jombart et al., 2010). Parametric approaches use statistical models to infer population structure and assign individuals into subpopulations through estimating population parameters, based on genetic assumptions about the data (Pritchard et al. 2000). In contrast, nonparametric methods have no modelling assumption requirements and group together individuals with similar genetic profiles (Limpti et al. 2000).

To visualize genetic structure, I used the function “glPCA” in package adegenet (Jombart, 2008; Jombart and Ahmed, 2011) along five principal components to create scatter plots that indicate where the most genetically isolated subpopulations appear as distinct cluster of individuals. To summarize genetic differentiation between groups, along with the assignment of individuals to clusters, I used the function “find.clusters” to first infer genetic clusters. I then implemented the discriminant analysis of principal components on the inferred individual group memberships using the function “dapc” to summarize the genetic differentiation between groups. Because the assignment of group membership can be unstable depending on the number of principal components retained, and the discriminant functions employed, I used a cross-validation method to assess how many should be retained to be informative enough to accurately discriminate between groups using the “xvalDapc” function. I then visualized group membership as a graphic output using the “scatter.dapc” function, and showed the probabilities of group membership of individuals to different clusters using the “assignplot” function.

To compare the non-parametric methods of PCA and DAPC with a parametric approach that uses statistical models to infer population structure, I used the Bayesian assignment analysis implemented in the FastStructure software (Raj et al., 2013). I ran FastStructure using the simple prior with K values from 1 to 10. I used the utility tool chooseK.py to choose the appropriate number of model components that best explained structure in the genomic dataset.
I then visualized the ancestry proportions inferred by fastStructure using the “plotq” function in the *pophelper* package (Francis, 2017).

### 3.3.3.2 Genetic diversity

Observed heterozygosity ($H_{\text{obs}}$), expected heterozygosity ($H_{\text{exp}}$), and allelic richness ($A_r$) estimated by rarefaction were calculated for all individuals and across the populations to gain insight into the genomic diversity of Arctic ground squirrels using the function “basic.stats” in the *hierfstat* package (Goudet, 2005). I also estimated mean observed $H_{\text{obs}}$, mean gene diversities for all samples. $H_{\text{obs}}$ and $H_{\text{exp}}$ were calculated with the package *dartR* using the function “gl.Ho” and “gl.He” respectively (Gruber et al., 2018). The inbreeding coefficient was also calculated for each location ($F_{\text{IS}} = (H_{\text{exp}}-H_{\text{obs}})/H_{\text{exp}}$), and the average $F_{\text{IS}}$ and $F_{\text{ST}}$ were calculated for the overall region.

### 3.3.3.3 Genetic differentiation

To examine population differentiation of Arctic ground squirrels I first calculated differentiation statistics including Nei’s GST and Jost’s D using the function “diff_stats” in the *mmmod* package (Winter, 2012). I also used genet.dist to calculate pairwise Nei’s Ds (Nei et al., 1983; Takezaki and Nei, 1996) between individuals, and Weir and Cockerham’s $F_{\text{ST}}$ (Weir and Cockerham, 1984) between populations. To determine if there is a pattern of isolation by distance among populations in our sampled area, I used the genetic distances and geographic distances in a Mantel test using the “mantel.randtest” and “mantel” function in the *ade4* and *vegan* packages (Dray et al., 2007; Oksanen et al., 2007). Additionally, I used an analysis of molecular variance (AMOVA) to test differences among the clusters indicated by fastStructure populations using the function “stammpAmova” based on genetic distance estimates estimated using the function “stemmpNeisD” from the *StAMPP* package (Pembleton et al., 2013). To test if populations were significantly different, I performed a randomization test using the function “randtest” from the *ade4* package. Pairwise $F_{\text{ST}}$ values were generated with the function “gl.fst.pop” in the package *dartR* and 100 bootstraps were performed across loci to generate the corresponding p-values.

### 3.4 RESULTS
3.4.1 Genotyping

I collected ear tissue samples from 69 Arctic ground squirrel individuals (Table 3.1; Figure 3.1). Of the 69 Arctic ground squirrels sequenced, during the final filtering stage, 23 tissue samples were dropped due to not meeting quality control standards and were excluded from the analysis. The number of dropped samples were similar to each other except for the Gorge Creek site, where over half of the samples were dropped. Stringent filtering in the final VCF file of 46 individuals contained 17,460 SNPs with up to 6% missing data per site. This dataset was used in subsequent analyses.

3.4.2 Population structure

The PCA (Figure 3.2) revealed three clusters on PC1, and additional PCs indicated separation among individual sites. PC1 (8.06% of variance explained) separated Tributary Creek and Primrose from Polychrome with Gorge Creek and Divide Mountain falling in between the two. PC2 (6.85 of variance explained) indicates more separation of Divide Mountain and Gorge Creek from Tributary, Primrose and Polychrome. PC3 (5.48 variance explained) further characterized genetic variation by grouping together Gorge Creek, Polychrome and Divide Mountain corresponding to their geographic proximity to each other and potentially describing a southern cluster from the other two sites. Ultimately, most sites, even those closest by straight-line distance (8 km), showed genetic isolation, suggesting restricted movement among sites.

DaPC confirmed these results by indicating anywhere between one to four clusters could provide useful summaries of the data. Depending on the number of principal components retained, and the discriminant functions used, there was instability with regards to group membership. Using a cross-validation method to assess the tradeoff between retaining too many and too few PCs in the model to accurately discriminate between groups indicated five to be optimum achieving the lowest MSE and having the highest mean of success (Figure 3.3). DAPC, with five PCs retained and two discriminant functions (Figure 3.4), indicated separation among two clusters along DF1. DF2 showed a further split with three clusters.

FastStructure revealed additional structure when run on a simple prior with K values from 1 to 10 using the same seed. FastStructure returned a K = 3-5 as an appropriate number of model components that explain structure in the dataset. At K = 3, individuals were clustered
into groups that aligned into: an eastern cluster containing samples from Primrose and Tributary Creek (Cluster 1), a southern cluster containing samples from Polychrome (Cluster 2), and a western cluster containing samples from Divide Mountain and Gorge Creek (Cluster 3) (Figure 3.5). There was little to no admixture at K=3 at most sites; however, individuals at Polychrome and Divide Mountain showed a rough mix from Cluster 1. At K=4, all clusters remained like splits in sites shown at K=3 except at Primrose. Primrose was split into two clusters, with most individuals retained in Cluster 1 with Tributary Creek. At K=5, all clusters retained their composition except for Tributary Creek where all but two individuals remained in the same cluster. Upon visualizing the K = 3 fastStructure admixture plot, it was clear that Arctic ground squirrels were divided into three clusters consistent with the PCA and DAPC results. Admixture was not present at all sampling locations.

### 3.4.3 Genetic diversity

The \( H_{obs} \) averaged overall for the region was estimated at 0.29. Within population genetic diversity (HS) was 0.26. \( H_{exp} \) was compared to \( H_{obs} \) at each site to understand the levels of Arctic ground squirrel genetic diversity (Figure 3.6; Table 3.2). \( H_{exp} \) was lower than observed at every single location and did not vary significantly. \( A_r \) also did not vary by population and was comparable to each other. I found no significant differences in \( H_{exp} \) and \( H_{obs} \) between each sampling site.

### 3.4.4 Genetic differentiation

An AMOVA examined the genetic variation of DENA Arctic ground squirrel populations. The results showed that 100% of the genetic variation was contained within the Arctic ground squirrel samples (Table 3.3). This result held true when run on the three clusters of the sampled populations determined by both non-parametric and parametric models. A randomization test to determine if populations were significantly different revealed that there was no significant population structure considering all levels of the population strata used in the AMOVA (Figure 3.7). Similar to other results, a Mantel test confirmed these populations exhibited genetic isolation by distance across the study area (\( r = .81, p < .01 \)). Hierfstat returned a global \( F_{ST} \) of 0.067, and examining the inbreeding coefficients shows little to no evidence of inbreeding (mean \( F_{IS} = -0.10 \)). The Weir and Cockerham estimates of Fstatistics were similar
(FST = 0.0745, FIS = -0.11). I compared genetic differentiation in the form of pairwise FST (Table 3.4). Statistically significant pairwise FST values ranged from 0.04 (Polychrome and Tributary) to 0.11 (Primrose and Gorge Creek), These results are in line with the global FST value.

3.5 DISCUSSION

Climate change combined with additional stressors that are likely to shift patterns of distribution, abundance, and phenology of Arctic species will also impact genetic diversity (Pauls et al., 2013). Genetic diversity has long been recognized as important for providing the raw material for adaptation to climate change (Hohenlohe et al., 2021; DeWoody et al., 2021), is positively correlated with population fitness and species persistence (Reed and Frankham, 2003; Scott et al., 2020) and is thought to be a good indicator for predicting adaptive potential (Harrisson et al., 2014; Sella and Barton, 2019). Thus, wildlife management would benefit from population-level monitoring that also incorporates measures of adaptation to climate change – specifically genomic information.

A critical source of information for conservation action is to identify populations and the relationships among them through evolutionary processes. Current approaches to biodiversity conservation rests on discrete categories, such as the species, which runs counter to the biological reality that most species are not genetically uniform (Edwards et al. 2016). Many studies have revealed cryptic species; these are highly divergent genetic populations within species (Struck et al. 2017). The importance of identifying populations is highlighted by how previous failure to recognize cryptic species have resulted in the failure to protect geographically restricted species (Gippoliti et al. 2017). The first step in delineating populations is to characterize population genetic structure.

In our assessment of population structure, I found evidence for three genetic clusters within the sampled populations of Arctic ground squirrels in DENA. These three clusters could comprise management units based on differentiation at neutral loci (Moritz et al. 1994), with a distinction between Arctic ground squirrels in Tributary Creek and Primose Ridge (Cluster 1), Polychrome (Cluster 2) and Gorge Creek and Divide Mountain (Cluster 3). Consistent with our analysis, not many individuals from these clusters showed mixed ancestry. This is consistent with the pattern of isolation I found through the Mantel test. In contrast, I found notably low population differentiation among the samples; AMOVA tests showed that none of the total
differentiations were partitioned among populations indicating a high gene exchange (or low differentiation) between the populations and the fastStructure clusters.

In addition to patchy habitat availability, topography and waterway barriers could explain the moderate genetic differentiation. This may be the product of moderate to high gene flow and/or signal that an incomplete biogeographic barrier exists in DENA. Rivers and streams that are fed by the snow melt off the mountains and glaciers in DENA may represent one level of an incomplete barrier. As streams dry up from reduced runoff at the end of the summer season, Arctic ground squirrels may be able to cross these otherwise “hard” barriers into new habitat, not accessible previously. Humans have also provided a pathway, the park road, by which Arctic ground squirrels in DENA can circumvent potential biogeographic barriers. Arctic ground squirrel burrows are found along the park road and it may facilitate gene flow amongst the populations sampled, which were within 5 km of the road. Although sites were differentiated, gene flow did appear higher than in some ground-dwelling sciurids (e.g. the black-tailed prairie dogs; Magle et al. 2010), indicating a lack of habitat fragmentation in this generally well-protected area.

Population genetic theory predicts that low genetic diversity in species could be a symptom of past genetic drift and inbreeding caused by a low effective population size (N_e; Charlsworth, 2009). In this analysis we were not able to include estimates of N_e which warrant future investigation to understand the demographic history of Arctic ground squirrels in DENA. It’s inclusion, along with estimates of genetic variation would provide tools to potentially understand the adaptive capacity of the Arctic ground squirrel to potential environmental changes. In this study, I there was some indication of low genetic diversity, in line with our expectations based on the geological history of the region (Alsos et al., 2005) and the evolutionary and biogeographic past of this species (Hewitt, 2004). Recent range contraction and population declines may have led to eroded genetic diversity in the population sampled in this study (Lanier et al., 2015b). Range contractions have important consequences for expectations of the genetic diversity of species. Studies have demonstrated that overall, it should lead to decreases in genetic diversity of populations (O’Brien, 1994; Luikart and Cornuet, 1998). Also, genetic diversity did not vary strongly among the sites sampled.

3.6 Conclusion
Maintaining genetic diversity of Arctic ground squirrels, and other Arctic wildlife will be critical for preserving biodiversity in the face of climate change. This will be particularly important for small, isolated populations occupying patchy habitat that historically have experienced range contractions. The Arctic ground squirrel in Denali National Park seems to be such a species. The population genomic approaches applied here in this study provided detailed information on the structuring of, differentiation of, and genetic diversity of Arctic ground squirrels in DENA. These results suggest that the potential for evolutionary adaptation in these populations may be low. Results could be improved through expanded monitoring in the park that samples more populations.
FIGURES

Figure 3.1: Sites of sampled and analyzed individuals. Sites are grouped and colored by location name. Inset shows location in Denali National Park and Preserve, AK.
Figure 3.2: Principal components analysis (PCA) of 46 Arctic ground squirrel samples with 17,460 unlinked SNPs.
Figure 3.3: Stratified cross-validation of DAPC using varying numbers of PCs to ensure the appropriate number is retained. 5 PCs give the highest percent of correctly predicted subsamples with the lowest error.
Figure 3.4: 4 Discriminant analysis of Principal Components (DAPC) scatterplot drawn against 17,466 SNPs across 46 individuals. Dots represent individuals, with cluster denoting sampling origin and inclusion of 95% inertia ellipses.
Figure 3.5: Hierarchical fastStructure barplots of 46 individuals using 17,460 unlinked SNPs. Each column is an individual. Rows are plots of $K=1$, $K=2$, and $K=3$, respectively.
Figure 3.6: AMOVA permutations test using 1000 permutations for the number of subpopulation $K = 5$. The three histograms represent the distribution of the randomized strata and the black line represents the observed data.
Table 3.1: Location information and the number of genotyped Arctic ground squirrels that were sampled (# Tissue Samples), failed to pass quality checks (# Failed Tissue Samples), genotyped for downstream analysis (Total Individuals) in Denali National Park and Preserve, AK.

<table>
<thead>
<tr>
<th>Site</th>
<th># Tissue Samples</th>
<th># Failed Tissue Samples</th>
<th>Total Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primrose</td>
<td>8</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Tributary</td>
<td>17</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>Gorge Creek</td>
<td>14</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Polychrome</td>
<td>20</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Divide Mountain</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>23</td>
<td>46</td>
</tr>
</tbody>
</table>
Table 3.2: Mean observed heterozygosity (Hobs) and mean expected heterozygosity (Hexp), and allelic richness (Ar) estimated by rarefaction is reported per site.

<table>
<thead>
<tr>
<th>Pop</th>
<th>n</th>
<th>H_{obs}</th>
<th>H_{exp}</th>
<th>Ar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tributary Creek</td>
<td>15</td>
<td>0.297</td>
<td>0.259</td>
<td>1.26</td>
</tr>
<tr>
<td>Polychrome</td>
<td>14</td>
<td>0.290</td>
<td>0.253</td>
<td>1.25</td>
</tr>
<tr>
<td>Divide Mountain</td>
<td>5</td>
<td>0.276</td>
<td>0.234</td>
<td>1.24</td>
</tr>
<tr>
<td>Gorge Creek</td>
<td>6</td>
<td>0.270</td>
<td>0.229</td>
<td>1.23</td>
</tr>
<tr>
<td>Primrose</td>
<td>6</td>
<td>0.299</td>
<td>0.247</td>
<td>1.25</td>
</tr>
</tbody>
</table>
Table 3.3: Analysis of molecular variance (AMOVA) for 46 individuals in five populations of Arctic ground squirrels in Denali National Park and Preserve, AK.

<table>
<thead>
<tr>
<th></th>
<th>Sigma</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variations Between pop</td>
<td>-10.79</td>
<td>-0.22</td>
</tr>
<tr>
<td>Variations Between samples within pop</td>
<td>-194.19</td>
<td>-3.99</td>
</tr>
<tr>
<td>Variations Within samples</td>
<td>5065.48</td>
<td>104.21</td>
</tr>
<tr>
<td>Total Variations</td>
<td>4860.49</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 3.4: Pairwise FST values between Arctic ground squirrel sampling locations (i.e. putative populations) are shown below the diagonal.

<table>
<thead>
<tr>
<th></th>
<th>Tributary Creek</th>
<th>Polychrome</th>
<th>Divide Mountain</th>
<th>Gorge Creek</th>
<th>Primrose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tributary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polychrome</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divide</td>
<td>0.07</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gorge</td>
<td>0.09</td>
<td>0.10</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primrose</td>
<td>0.06</td>
<td>0.08</td>
<td>0.09</td>
<td>0.12</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 4

BUILDING A BETTER MODEL: USING REPEAT SURVEY DATA AND GENETIC MARKERS TO ESTIMATE THE ARCTIC GROUND SQUIRREL (*UROCITELLUS PARRYII*) SPECIES DISTRIBUTION

4.1 ABSTRACT:

Species distribution models (SDMs) are essential tools in the conservation toolbox, for forecasting the potential impacts of climate change. However, most SDMs are built with presence-only data, potentially leading to inaccurate predictions that can reduce the impact of conservation actions. While debate remains over the modeling procedures that make the best use of presence-only data, another suggested improvement is the use of more detailed data to increase our ability to discriminate range boundaries and improve conservation efforts. Here, I make use of true presence-absence, abundance, and genetic data to assess potential impacts of climate change on the distribution of Arctic ground squirrels (*Urocitellus parryii*) in Denali National Park and Preserve (DENA). We applied the Random Forest algorithm with the WorldClim bioclimatic predictor variables to forecast the spatial occurrence as well as the abundance distribution, and to incorporate intra-specific variation. I split the data into training and test sets for validation and to assess model performance. The Random Forest model trained on presence-absence data performed well according to AUC (0.89), while models built on the abundance data performed poorly RMSE (0.19). Model performance was high for the genetic dataset potentially highlighting adaptive variation although there was evidence of model overfitting. The regions of high occurrence and abundance showed similar patterns with high concentrations mainly in the mid to southern portion of the administrative boundaries of DENA. Maps of predicted spatial distribution projected with future climate projections represent an initial attempt to capture the geographical and temporal range of the species and can be used for identifying potential areas of conservation concern and improving sampling for future studies.
4.1.1 Keywords:

_Urocitellus parryii_; species distribution model; species abundance model; machine learning *
Random Forest * climate change

4.2 INTRODUCTION:

A major concern and challenge for conservation practitioners and managers is to understand the consequences of climate change for species (Hughes, 2000; Change et al., 2014). Climate change is predicted to be the major cause of species extinctions (Urban, 2015) as novel climates are projected to develop and impact the behavior and fitness of species (Williams et al., 2007; Cahill et al., 2013). Ongoing research continues to demonstrate that as climates change, so do species distributions (Pecl et al., 2017). Changes in habitat suitability which exceed the physiological tolerance of species (RuizAravena et al., 2014) is likely to elevate the risk of extinction, especially for species of management concern. Thus, forecasting how species ranges will shift under climate change is critical to adapt conservation strategies and useful for exploratory analysis to identify potential study sites or areas where additional survey effort is needed.

Recent decades have seen the increased use of GIS and technical tools that predict the spatial distribution of species, including species distribution models (SDMs) (Elith and Leathwick, 2009; Guisan and Thuiller, 2005; Guillera-Arroita, 2017). SDMs relate georeferenced biodiversity observations (e.g., species detection) with a set of (spatially-explicit) environmental variables under the basic ecological assumption that biodiversity distribution is determined by the environment (Hutchinson, 1957; Elith and Leathwick, 2009; Zurell and Engler, 2019). Once the biodiversity-environment relationship has been estimated, SDMs can be used to delineate the potential distribution of biodiversity under current and future climate/environments. SDMs are a popular tool because they are comparably easy to use (Thuiller et al., 2009; Guisan et al., 2017), can be constructed in a variety of ways (presence, presence/absence) with a complex set of environmental variables (e.g., climate, topographic, landscape cover), and result in a wide range of outputs (e.g., potential of invasive species, distribution of cryptic species, suitable sites for reintroduction of threatened species).

Nevertheless, there are a suite of methodological challenges with the building of SDMs and their application to conservation decisions. While other studies have highlighted these
methodological and statistical limitations (Elith and Leathwick, 2009; Mateo et al., 2011), of particular concern to this chapter is that SDMs have mostly relied on presence-only data with the assumption of perfect detection which constrains the interpretation and use of SDMs (Royle and Dorazio, 2008), confounds information about habitat preferences and availability (Phillips and Elith, 2013) and creates a strong spatial bias (Guillera-Arroita, 2017). In recognition of these shortcomings, other data and methodologies, including those that control for detection probability through occupancy, are needed to improve our capacity for predicting the impacts of climate change on species (Urban et al. 2016).

Additionally, abundance data are thought to produce more accurate species abundance models (SAM) (Mi et al., 2017) than occurrence data. Abundance might be more indicative of species extinction risk (O’Grady et al., 2004) by reflecting factors such as habitat quality (Pearce and Ferrier, 2001), whereas suitability models generated from presence-absence data may reflect only the potential to colonize a site and overstating where potential management efforts are needed (Jarnevich et al. 2021; O’Neil et al. 2021). Abundance may also be more useful in management contexts where species site extinction is better predicted by a species’ local abundance (Shulz et al. 2020), occurrence patterns may not change until after population depletion (Hastings et al. 2020), and in some cases the most productive areas only occupy a subset of the total geographic space (Kallasvuo et al. 2017). In many management contexts, the goal is often to maintain abundance rather than just presence (Hutchings and Reynolds, 2014).

Another recent advancement in species distribution modeling is a focus on genetic markers (e.g., single nucleotide polymorphism (SNP), restriction fragment length polymorphism (RFLP), and microsatellites) to identify spatial priorities in conservation planning to promote the persistence of species under climate change scenarios (Sunnuck and Taylor, 2008). The ability of species to track climate conditions not only depends on the speed at which change occurs (Berestycki et al. 2008), but also on the ability of the species to reach new areas or shift their ranges away from climatically unsuitable habitat (Legrand et al. 2017). Additionally, a species population genetic patterns affects its ability to respond to climate change (Marcer et al. 2016). For example, modeling approaches that incorporate genetic markers that signal demography can identify how different genetic lineages within a species or subspecies may differ in their response to environmental change (Hällfors et al., 2015; Maguire et al. 2018). Landscape genetic methods which quantify landscape permeability and identify source/sink population dynamics have much insight to provide in projecting species distribution...
ranges and improving the predictions of the effects of climate change (Manel and Holderegger et al. 2013; Murphy et al. 2010).

In this chapter, I seek to demonstrate how SDMs derived from true presence/absence, abundance, and genetic data may be combined to offer useful information for predicting the impacts of climate change on vulnerable species, specifically the Arctic ground squirrel (*Urocitellus parryii*) in Denali National Park and Preserve, AK (DENA). I will 1) assess and predict patterns of Arctic ground squirrel occurrence and abundance; 2) incorporate molecular markers to identify if potential genetic differences among clusters reflect different responses to climate conditions, 3) identify the climatic variables that influence occurrence, abundance, and genetic clustering; and 4) investigate the relationship between predicted occurrence and abundance. This research demonstrates the advantage of incorporating detailed survey data into SDMs and SAMs for conservation in the face of climate change.

### 4.3 METHODS

#### 4.3.1 Study area and sampling

The study occurred in DENA in interior Alaska (see Chapter 1), an area encompassing roughly 24,524 km². Point count surveys (1-3 visits per site) were conducted from 2017-2019 to estimate the occupancy probability of Arctic ground squirrels in DENA. Surveys consisted of 10 min of listening for alarm calls and scanning with binoculars. I recorded Arctic ground squirrel detection and distance using a range finder. The recorded data were classified in two categories: species was detected, 1; and species was not detected, 0. More information on surveys is available in Chapter 2. To develop the species distribution model, these data were collapsed into simple presence/absence data if an Arctic ground squirrel was detected at any point during the three-year survey, that site was considered occupied (N = 62 total sites).

Repeated distance line transect surveys (1-3 visits per site) were conducted to estimate the density of Arctic ground squirrels in DENA. Selected sites consisted of pre-determined transects comprising a 500-m straight line from one plot to the next of the 5x5 mini-grid nested within the larger sampling grid. I counted and recorded the amount of Arctic ground squirrels along the line. I recorded Arctic ground squirrel detections (by sight or sound) during the walk. I observed Arctic ground squirrels with binoculars and a rangefinder for distance measurements, and a combination of GPS and compass for measuring sighting angles. More
details are available in Chapter 2. To develop the SAM, this dataset was collapsed by using the highest recorded number of individuals during either any of the 1-3 site visits, and then averaged over the three years of the study (N = 69 total sites) to produce a minimum count.

Arctic ground squirrels were live trapped to collect tissue samples at five sampling grids. Arctic ground squirrels were sampled using live collapsible Tomahawk traps (Tomahawk Live-trap Co., WI, USA; 19X6X6 cm) placed near burrow entrances. Trapping locations were targeted through a combination of spatial modeling that included distance from each other and for representing environmental variations related to their ecological habitat. I identified three major genetic clusters using fastStructure through SNPs (Chapter 3). The data frame was structured such that each row represents a different individual, whereas columns hold the proportion ancestry for each of the three clusters, the binary presence/absence vector that would be used in the model (N = 46 total individuals across 5 sampling sites). This produced a presence-absence data set, which represents sites where a particular genetic cluster has been detected, and in analysis, can be predicted by a set of environmental covariates.

4.3.2 Climate projections

I characterized the environmental conditions in DENA based on 19 biologically relevant proxy climatic (bioclimatic) variables (Table 4.1) drawn from the WorldClim dataset (climate data for 1970 to 2020; or neat current) at a resolution of 30 arc seconds (http://www.worldclim.org). I also used the CMIP6 downscaled future bioclimatic projections using the 126 shared socio-economic pathway (SSP) for the years 2041-2060 from the WorldClim dataset. Using a shapefile map of DENA, I used the “extract by mask” tool in ArcGIS Pro 2.9 (ESRI Inc., http://www.esri.com/) to delineate the current and projected climatic data within the boundaries of the park. Projected climatic data averaged over the period 2041-2060, were rasterized to match the cell size of the baseline climatic data using the “resample” tool in ArcGis Pro 2.9. Prior to model development, I reduced potential collinearity (or multicollinearity) using a priori variable selection. Prior to each analysis, I reduced our full bioclimatic dataset by calculating the variation inflation factor (VIF) and removing variables at the VIF = 3 threshold and used biological relevance to decide which variables to retain (Zuur et al. 2010). For the genetic data, we ran RF on all the available predictors to examine if different clusters may be best fit with different set of predictor variables.
4.3.3 Classification and regression trees

I employed a machine learning technique, Random Forest (RF), to model the occurrence, abundance, and occurrence probability for each cluster fit with different sets of predictor variables. Random Forest is an ensemble classifier which builds many classifications or regression trees from random subsets of the data and aggregates the results (Breiman, 2001; Liaw et al., 2002). I used the package ‘randomForest’ (Liaw et al., 2002) implemented in R 4.1.2 (R Core Team, 2021) to perform the classification tree algorithm (CT) on each genetic cluster separately, and the regression tree algorithm (RT) on the occurrence (to compute a relative index of occurrence) and abundance data. I split the dataset into a training set (75% of the data) and test set (25%) and assessed the RF model performance for each dataset using a set of ‘out-of-bag’ training points and the independent test data, I used the area under receiver-operating curve (AUC) as a discrimination performance criteria and Kappa statistics (Cohen, 1960)(AUC ranges from 0 to 1, with low values below 0.6 indicating a performance no better than random, values between 0.7-0.9 considered as useful, and values > 0.9 as excellent. For the the root-mean-square deviation (RMSE), which measures how far from the regression line data points are, values from 0.2-0.5 show that the model predicts the data relatively accurately. Random Forest was also applied to rank the relative importance of the bioclimatic variables. I then created partial response plots for each variable included in the occurrence, abundance, and genetic models.

4.4 RESULTS

4.4.1 Model performance

The occurrence model composed of presence/absence data obtained a high measure of performance (AUC: 0.89), whereas the abundance model performed poorly (RMSE: 0.19). For the genetic data, the RF model was able to moderately discriminate between areas where genetic clusters were present and absent (Accuracy: 0.5; Kappa: 0.23); however, the accuracy of the training dataset was higher than the test set for each cluster.

4.4.2 Variable importance

The variable importance ranking for the occurrence, abundance, and genetic models obtained from the Random Forest method retained some similarities (Table 4.2). For the
occurrence model, after accounting for collinearity, the following covariates were used: annual mean temperature (bio1), mean temperature of wettest quarter (bio8), mean temperature of driest quarter (bio9), mean temperature of coldest quarter (bio11), precipitation of driest month (bio14), and precipitation seasonality (bio15). Of these covariates bio15, bio14, bio11 and bio1 were the top important variables (in descending order of importance). For the abundance model, after accounting for collinearity, the following covariates were used to construct the model: annual mean temperature (bio1), isothermality (bio3), mean temperature of wettest quarter (bio8), bio9, mean temperature of coldest quarter (bio11), precipitation of driest month (bio14), and precipitation seasonality (bio15). Of these covariates, bio11, bio1, bio3 and bio9 were the top important variables (in descending order of importance). For the genetic model, variable importance was different for each cluster. For Cluster 1, the top four important variables consisted of bio11, mean temperature of wettest quarter (bio8), bio9, and minimum temperature of coldest month (bio6). In contrast, for Cluster 2, the top four important variables included temperature seasonality (bio4), mean diurnal range(bio2), temperature annual range (bio7) and mean temperature of coldest quarter (bio10), while for Cluster 3 the top four variables consisted of precipitation of wettest month (bio13), precipitation of wettest quarter (bio16) precipitation of coldest quarter (bio19), and bio8.

4.4.3 Partial dependence plots

Based on model results, the occurrence preference of Arctic ground squirrels with respect to the bioclimatic variables was for bio14 between 14-22 mm, for precipitation variability (bio15) between 49-53%, bio1 below -3.6 °C (Figure 4.1). For predicting abundance, more individuals occurred when winter temperatures were above -15.5 °C (bio11), when annual mean temperature (bio1) was below -3.7 °C, isothermality (bio3) was between 26-27%, and summer mean (bio9) was above -6 °C (Figure 4.2). Genetic cluster 1, was associated with high values for mean temperature of wettest quarter (bio8; 9.3-12 °C), and low values of mean temperature of coldest quarter (bio11; -15.2–14.9 °C), mean temperature of driest quarter (bio9; 15 14 °C), and min temperature of coldest month (bio6; -20.7-20.5 °C) (Figure 4.3). Genetic cluster 2 is associated with low values for all of the top 4 important variables: temperature seasonality (bio4; 10.0 10.3 °C), annual mean diurnal range (bio2; 9.5 9.7 °C), annual temperature range (bio7; 36.0 37.0 °C), and mean temperature of warmest quarter (bio10; 9.0-9.6 °C) (Figure 4.4). Genetic cluster 3 was associated with low values of bio 8 (8.0-9.0 °C), and with high values for precipitation of warmest quarter (bio18; 243 260 mm), precipitation of coldest quarter (bio19;
86-98 mm) and precipitation of wettest month (bio13; 93-100 mm) (Figure 4.5). For each genetic cluster, at least half of the range is near zero for the range of the variables suggesting little to no actual relationship between them.

### 4.4.4 Occurrence and abundance distribution patterns

The region of high possibility of Arctic ground squirrel occurrence based on the bioclimatic variables is in the mid to southern portion of the administrative boundaries of DENA. The area habitat seems to be continuous (Figure 4.6). Climate change is expected to shift the areas with high probability of occurrence to the southern portion of the park while decreasing occurrence in the northern portion of the administrative boundaries. The abundance models showed similar patterns, with low abundance concentrated in the mid to southern portions of the park administrative boundaries (Figure 4.7). Overall, climate change is expected to confine the distribution mainly in the southern portion of the administrative boundaries of the park (Figure 4.6).

### 4.5 DISCUSSION

Our use of the Random Forest algorithm was designed to predict the current and future distribution of Arctic ground squirrels. Our results show that the occurrence models predicted with high accuracy the area inhabited by Arctic ground squirrels, whereas the abundance models did not. The reduction in occurrence with projected climate is inline with predictions made in an earlier study (Marcot et al. 2015) that saw low-shrub-associated species such as lemmings, voles, shrews, and the Arctic ground squirrel will lose suitable habitat with climate change. The reduction in occurrence closely follows changes in elevation; where new habitat opened up in the sections of the Alaskan Range in DENA, it decreased in the lower elevational areas. This is in line with predictions of how small mammals climb up in elevation in response to projected climate (McCain et al. 2021).

The weak performance of the abundance Random Forest model is potentially due to containing many absences (sites where we did not count Arctic ground squirrels) that overwhelmed the dataset. While the literature is still developing for the use of the Random Forest algorithm on SAM models, it has been reported that this behavior is well known when modeling presence-absence data (Evans and Cushman, 2009; Freeman et al., 2012; Robinson et al., 2018). Additionally, while AUC may provide information about the model’s fit, it doesn’t
show the biological significance, thus a cautious interpretation of variable importance for the abundance model should be followed.

Finding effective ways to model abundance will remain an important area of research for monitoring Arctic ground squirrels. Earlier studies suggest that different analytical assumptions of models combined with species specific ecology influence the predictive power of machine learning methods (Oppel et al., 2012; Kosicki, 2020). Future work should focus on combining and comparing the performance of multiple analytical approaches to predictive modeling of species distribution, i.e., the Generalized Linear Model, Generalized Additive Model, Random Forest, Boosted Regression Trees to predict the species relative abundance (Oppel et al., 2012; Kosicki, 2020). Also, abundance data for this species was difficult to collect (see Chapter 2), potentially influencing model quality (Pearce and Ferrier, 2000; Kadmon et al., 2003; Hernandez et al., 2006). Additionally, species with low frequency of occurrence and low mean abundance are thought to be more challenging to predict as they may be more likely to be constrained by effects such as availability of resources (Venter et al. 2014) rather than the coarser scale climatic variables like those used in this analysis.

The inclusion of genetic data in this analysis represents a significant step forward by incorporating adaptive capacity into SDMs. Standard SDMs built from presence data assume that the modeled species has a single environmental niche rather than different populations potentially responding to different local conditions. This is equivalent to the assumption of a single panmictic population, even though there is significant evidence suggesting the opposite (e.g. fragmentation of habitat by humans isolating populations). Thus, the product of the SDM is the species response encompassing the entire geographical range with little or no genetic variation, rather than modeling how different populations (or clusters) may respond differently. Conservation actions with this consideration may look very different. For example, if two management units with significant differentiation from each other in the same administrative area have opposite predicted range shifts in the future to projected climate, the focus could be on translocating individuals with adaptations for the extreme end of environmental change to populations that are less adapted. While the Random Forest model for the genetically distinct clusters from the fastStructure results indicated an ability to moderately differentiate between clusters, and highlighted a subset of projected climate predictors that each cluster may respond to, I caution interpretation of this as there is evidence of overfitting.

The occurrence and abundance models of Arctic ground squirrels developed here
illustrate the current distribution, the potential impacts of climate change, and relevant locations for where to prioritize monitoring efforts. For occurrence, annual mean temperature, mean temperature of the coldest quarter, precipitation of the driest month, and precipitation seasonality were the top most important predictors. In contrast the top important covariates for the abundance model consisted of annual mean temperature, isothermality, mean temperature of the coldest quarter, and precipitation of the coldest quarter. While providing an avenue to better understand how Arctic ground squirrels may respond to climate change, it is important to note that the climate projections used in this study represents only a snapshot of possible predictions. Future work could focus on (1) predicting over a range of time periods, as species distributions are likely to change over time; (2) a range of climate scenarios that considers a variety of possible futures; and (3) different climate data sources to assess model performance.

In this study, occurrence and abundance displayed identical spatial distribution patterns. These similarities may represent a mixture of effects that reflect similar differences between the underlying biological processes that gave rise to the specific occurrence and abundance estimated in this study. How to interpret these similarities given the differences in model performance is a problem waiting to be further resolved. Whereas some predictors were shared across the models, caution is needed in uniting these predictions as the importance of the effects varied. Results from the RF genetic models raised the possibility that different clusters may be best fit with different sets of predictor variables indicating that genetic differences among the clusters potentially reflect adaptive variation to different abiotic conditions.
Figure 4.1: Partial dependence plots for the influential variables in the occurrence distribution models for Arctic ground squirrels. Covariates include mean temperature of coldest quarter (bio11), precipitation of driest month (bio14), precipitation seasonality (bio15), mean temperature of wettest quarter (bio8), mean temperature of driest quarter (bio9), and annual mean temperature (bio1) in that order.
Figure 4.2: Partial dependence plots for the influential variables in the abundance distribution models for Arctic ground squirrels. Covariates include annual mean temperature (bio1), isothermality (bio3), mean temperature of wettest quarter (bio8), mean temperature of driest quarter (bio9), bio8, bio9, mean temperature of coldest quarter (bio11), bio11, precipitation of driest month (bio14), and precipitation seasonality (bio15), and bio14, and bio15 in that order.
Figure 4.3: Partial dependence plots for the influential variables for genetic cluster 1 in the binary models for Arctic ground squirrels. Covariates include mean temperature of coldest quarter (bio11), mean temperature of wettest quarter (bio8), mean temperature of driest quarter (bio9), and minimum temperature of coldest month, in that order.
Figure 4.4: Partial dependence plots for the influential variables for genetic cluster 2 in the binary models for Arctic ground squirrels. Covariates include precipitation of wettest month (bio13), precipitation of warmest quarter (bio18), precipitation of coldest quarter (bio19), and mean temperature of wettest quarter (bio8),bio8 in that order.
Figure 4.5: Partial dependence plots for the influential variables for genetic cluster 3 in the binary models for Arctic ground squirrels. Covariates include precipitation of wettest month (bio13), precipitation of warmest quarter (bio18), precipitation of coldest quarter (bio19), and mean temperature of wettest quarter (bio8), in that order. Covariates include bio13, bio18, bio19, and bio8 in that order.
Figure 4.6: Spatial distribution map of relative occurrence with near-current and future climate data. Values scale from high (green) to low (pink) predicted occurrence. Note that the scales differ between the two distribution maps.
Figure 4.7: Spatial distribution map of relative abundance with near-current and future climate data. Values scale from high (green) to low (pink) (green) predicted abundance. Note that the scales differ between the two distribution maps.
Table 4.1: Description of climatic covariates considered during exploratory analysis for predicting occurrence and density of Arctic ground squirrels in Denali National Park and Preserve.

<table>
<thead>
<tr>
<th>Bioclimatic Variable</th>
<th>Units</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual Mean Temperature</td>
<td>°C</td>
<td>bio1</td>
</tr>
<tr>
<td>Annual Mean Diurnal Range</td>
<td>°C</td>
<td>bio2</td>
</tr>
<tr>
<td>Isothermality</td>
<td>%</td>
<td>bio3</td>
</tr>
<tr>
<td>Temperature Seasonality</td>
<td>°C</td>
<td>bio4</td>
</tr>
<tr>
<td>Max Temperature of Warmest Month</td>
<td>°C</td>
<td>bio5</td>
</tr>
<tr>
<td>Min Temperature of Coldest Month</td>
<td>°C</td>
<td>bio6</td>
</tr>
<tr>
<td>Annual Temperature Range</td>
<td>°C</td>
<td>bio7</td>
</tr>
<tr>
<td><strong>Mean Temperature of Wettest Quarter</strong></td>
<td>°C</td>
<td>bio8</td>
</tr>
<tr>
<td>Mean Temperature of Driest Quarter</td>
<td>°C</td>
<td>bio9</td>
</tr>
<tr>
<td>Mean Temperature of Warmest Quarter</td>
<td>°C</td>
<td>bio10</td>
</tr>
<tr>
<td>Mean Temperature of Coldest Quarter</td>
<td>°C</td>
<td>bio11</td>
</tr>
<tr>
<td>Annual Precipitation</td>
<td>mm</td>
<td>bio12</td>
</tr>
<tr>
<td><strong>Precipitation of Wettest Month</strong></td>
<td>mm</td>
<td>bio13</td>
</tr>
<tr>
<td>Precipitation of Driest Month</td>
<td>mm</td>
<td>bio14</td>
</tr>
<tr>
<td>Precipitation Seasonality</td>
<td>%</td>
<td>bio15</td>
</tr>
<tr>
<td>Precipitation of Wettest Quarter</td>
<td>mm</td>
<td>bio16</td>
</tr>
<tr>
<td><strong>Precipitation of Driest Quarter</strong></td>
<td>mm</td>
<td>bio17</td>
</tr>
<tr>
<td>Precipitation of Warmest Quarter</td>
<td>mm</td>
<td>bio18</td>
</tr>
<tr>
<td>Precipitation of Coldest Quarter</td>
<td>mm</td>
<td>bio19</td>
</tr>
</tbody>
</table>
Table 4.2: Variable importance ranking for Random Forest occurrence, abundance, and genetic models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Ranking</th>
<th>Predictor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occurrence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Precipitation Seasonality (bio15)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Precipitation of Driest Month (bio14)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Mean Temperature of Coldest Quarter (bio11)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Annual Mean Temperature (bio1)</td>
</tr>
<tr>
<td><strong>Abundance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Mean Temperature of Coldest Quarter (bio11)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Annual Mean Temperature (bio1)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Isothermality (bio3)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Mean Temperature of Driest Quarter (bio9)</td>
</tr>
<tr>
<td><strong>Genetic Cluster 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Mean Temperature of Coldest Quarter (bio11)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Mean Temperature of Wettest Quarter (bio8)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Mean Temperature of Driest Quarter (bio9)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Mean Temperature of Coldest Month (bio6)</td>
</tr>
<tr>
<td><strong>Genetic Cluster 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Temperature Seasonality (bio4)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Annual Mean Diurnal Range (bio2)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Annual Temperature Range (bio7)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Mean Temperature of Warmest Quarter (bio10)</td>
</tr>
<tr>
<td><strong>Genetic Cluster 3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>Precipitation of Wettest Month (bio13)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Precipitation of Warmest Quarter (bio18)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Precipitation of Coldest Quarter (bio19)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Mean Temperature of Wettest Quarter (bio8)</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


Rapid shifts in Arctic tundra species' distributions and inter-specific range overlap under future climate change. *Diversity and Distributions.*


