Building a Geodatabase Design for American Pika Presence and Absence Data

by

Kyle Krueger Burke

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

May 2018

To My Wife:
For without her support this
would not have been possible

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would especially like to say thank you to Dr. Eric A. Beever whose research and enthusiastic support for this project I owe everything. Special thanks to my father for the seemingly endless discussions which eventually led to this paper and to my wife and son for allowing me hours of seclusion when necessary. Thank you to my USC thesis committee for the guidance and support to complete this project.

List of Abbreviations

ERD	Entity Relationship Diagram
GIS	Geographic information system
GISci	Geographic information science
SSI	Spatial Sciences Institute
USC	University of Southern California
USGS	United State Geological Survey

Abstract

A United States Geological Survey (USGS) researcher has been studying impacts of climate change on American Pika (*Ochotona princeps)* from the mid 1990's through 2017. This project aims to contribute to research on the American Pika by building a geodatabase to store and provide access to data on pika populations throughout the Great Basin region of the Western United States. The geodatabase contains pika presence and absence data for locations of talus, which includes habitat areas that have been previously surveyed or may be potentially surveyed in the future. The project used formatted data provided from field surveyed talus that have been digitized on www.caltopo.com, digitized new talus that have been more recently surveyed, and imported GPS points for presence/absence captured in Excel spreadsheets.

The end result of this project was a geodatabase that housed presence/absence points, talus polygons, site locations, temperature sensor locations, and temperature/relative humidity data. Several queries were completed that show proper importation and relationships of all data. Working closely with project researchers, this study allows for database expansion as needed for future research needs.

Studying presence/absence of American Pika allows for further understanding of climatic impacts in niche habitats that are especially susceptible to environmental change. This project also provides the opportunity for improved analysis and long term data storage relating to these presence/absence locations throughout the Great Basin region. The end result supports expansion of the database structure for future field seasons and data inclusion.

Chapter 1 Introduction

This project aims to create a geodatabase of a longitudinal study on American Pika (*Ochotona princeps*) habitat areas from previously surveyed and potential survey locations in the Great Basin region of the Western United States. The end result aims to support speedy analysis of habitat changes of small montane animals that have been studied for more than 20 years. Data housed in this geodatabase is specifically related to the presence and absence of American Pika at surveyed talus (rocky slopes habitat) locations and potential future survey talus. These data have been previously compiled and, therefore, this effort augments work performed by USGS scientist, Dr. Erik Beever, by creating a single database to contain several years of data collected in the field. The creation of a database to house the field collection data also allows for additions of future field seasons. Beever's project work focuses on the climate change impact on pika and how they may be used as an indicator species for future climate impacts (Beever et al. 2003; Beever et al. 2010; Beever et al. 2011; Beever et al. 2013). The American Pika has a very limited range of habitat in which it has the ability to survive. Because of that niche habitat, pika have developed specific adaptations to allow for survival. Dr. Beever's research aims to investigate how changes in climate impact on a species like the American Pika and what that might tell us about how other species and environments react longer term.

The creation of a geodatabase compiled collected data and added the ability to swiftly import future data from every field season. The goal is to allow for swifter analysis of these collected data, as well as storage of all data in one singular location. The geodatabase allows for faster exploration and analysis of collected data to see any gaps that may exist for upcoming field seasons or modeling. The ability to continuously build upon previously collected work is an advantage to researchers who may want to incorporate these data into future work. While the ability to swiftly and accurately share data is a great advantage, the main goal of this study is to allow researchers to have spatially referenced talus with the associated presence/absence point data and attributes all housed in one location.

Section 1.1 provides further background on the project research. Section 1.2 provides details of the motivation behind this project. Section 1.3 discusses the objectives that this project looks to accomplish. Section 1.4 looks at an overview of the project methods and Section 1.5 breaks down the structure of the thesis.

1.1. Background

The Background section is broken into three subsections. 1.1.1 discusses the background on the research area, 1.1.2 discusses the American Pika (*Ochotona princeps*) biology and ecology and, 1.1.3 discusses the field collection techniques.

1.1.1. Study Area

The study area for this project focuses primarily on the Great Basin hydrographic region located in the western United States. The Great Basin includes an area of roughly 500,000km² that stretches between the Sierra Nevada and Rocky Mountains (see Figure 1). The general geography of the region consists of north-south mountain ranges with a cold-temperate semidesert climate. Within the 20th century, the Great Basin has warmed an average of 0.3 to 0.6°C and there are expectations that it will rise a further 2.5 to 4°C by 2100 (Rowe and Terry 2014). The change in temperature along with the great biodiversity of the region, especially among small mammals, make it an exceptional study area in regards to climatic impact.

1.1.2. The American Pika (Ochotona princeps)

American Pika (*Ochotona princeps*) (hereafter referred to as pika) are small mammals that have limited, discontinuous, and isolated range on the high rocky slopes of the western

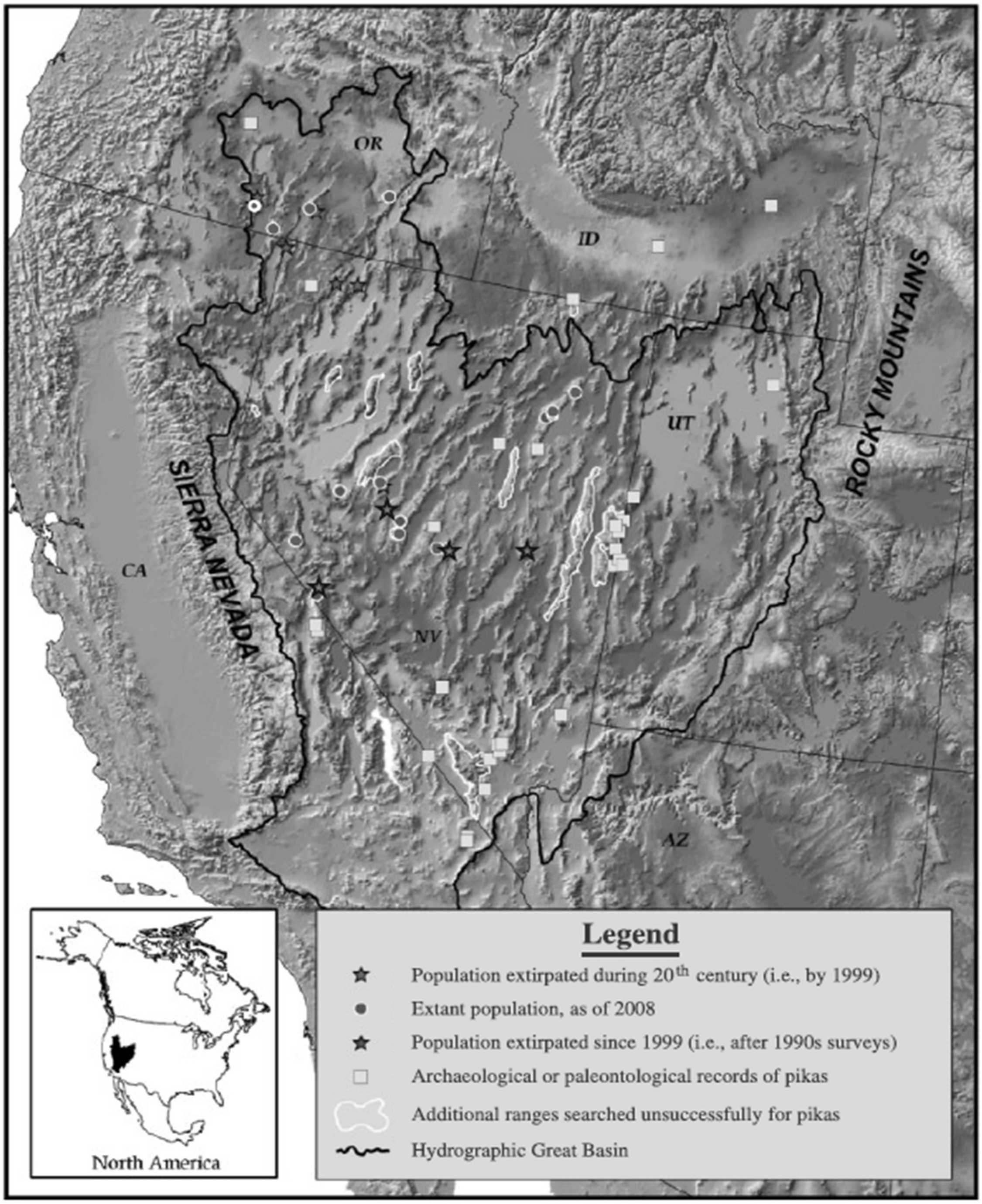


Figure 1. Original Map from Dr. Beever's Research of the Great Basin Region with locations of pika (Beever et al. 2011).

North American mountains (Millar and Westfall 2010). Pika require a generally cool climate with higher than average snowfall during winter to insulate them from extreme cold, as they do not hibernate (Beever et al. 2010). While typically found at elevations anywhere from sea level to 3000m along the northern edge of their range, at the southern edge most are found at 2500m and higher due to their cool weather climate requirement. These high montane requirements have create niche or island habitats that isolate population groups in specific rocky areas surrounded by suitable vegetation (Smith and Weston 1990).

With mean temperatures in North America expected to rise >2°C by the middle of the 21st century and >4°C by the late 21st century, the impact of climate change will be felt across the continent. Further, North America is expected to have daily extremes in excess of 5°C warmer than current temperatures and >80% of future years are predicted to have snow falls less than the middle of the 20th century levels (IPCC 2015). These factors along with the understanding that alpine areas will be significantly impacted because of climate change (Villers-Ruiz and Castaneda-Aguado 2013) make the pika population and range an important part of identifying how temperature increase will impact montane species.

Another major drawback for pika adaptability is their relatively small dispersal area. Found to be 0.8-3.0km (maximum and under cool conditions), the limited range of dispersal poses severe restrictions on the pika's ability to diffuse to new habitat should the current become inhospitable (Beever et al. 2003; Hafner 1993; Rodhouse et al 2010).

1.1.3. Field Collection Techniques

The field collection techniques used to collect the presence/absence data are detailed below. During multiple fields seasons talus locations that where known to have pika populations where surveyed for continuing populations. Each location was surveyed for 8 hours along with all talus locations within 3km as designated by the maximum known dispersal range. Field surveys included walking 50 m long transect lines across talus slopes around 15 vertical meters apart, using handheld GPS units to record the location of any pika sign, such has haypiles, feces, sighting, etc.

During these surveys Celsius degree and relative humidity sensor readings were recorded. These readings were obtained by placing temperature sensors roughly 80 cm under the surface of the talus at ideal living areas for pika. The sensors used were DS1921G Thermochron i-Buttons manufactured by Maxim Integrated Products of Sunnyvale, California, USA. The sensors were placed at locations throughout the Great Basin research region and recorded temperature to the nearest 0.1°C every two hours for the first five months after placement and every four hours after the initial period of time (Beever et al. 2010). During original field seasons all sensors collected data using the previously mentioned collection techniques. This has changed in more recent field seasons to only collect a reading every four hours. All sensors continue in operation until they no longer function or are replaced. The sensors are often replaced due to loss of function or inability to find the old sensor. The data that have been collected in previous years, and will continue to be collected in future years, are the main focus of this study and were used to create the geodatabase.

1.2. Motivation

According to the Intergovernmental Panel on Climate Change (IPCC) the results of numerous studies have led to the general consensus that humans have caused a rapid change in climate (IPCC 2015). In the coming years, the ability to further research and model the impacts of climate change will become more important. Researching specific species allows for scientists to attempt to understand the future impacts of climate change on the world. Pika are one such species that can be used as a predictor of climate impact because of its susceptibility to dramatic shifts in temperature. The overall goal of Dr. Beever's study is to further the understanding of how climate impacts the biology and ecology of specific areas such as montane or coastal zones. In Dr. Beever's project, the focus is montane zones. This project will allow researchers to further analyze and study climate change by building a geodatabase that allows for queries and spatial analysis to provide insight into these goals.

1.3. Objectives

In this study there are two primary objectives. The first objective is to build a geodatabase to house the majority of Dr. Beever's collected research which allows for the ability to quickly import, analyze, and present results of the presence/absence analysis. The completed geodatabase permits numerous years of data to be housed together and will create a system of simple data importation for future data collections efforts. This eases the current data storage issues surrounding multiple Excel spreadsheets. The geodatabase also allows for almost instant importation of field data into a permanent structure. The second objective is to build a framework for a geodatabase which other researchers will be able to use as a template for future research projects of a similar nature.

1.4. Overview of Project Methods

The following is a brief overview of the methods for this project and process involved. This segment is broken into multiple sections starting with 1.3.1 which discusses the geodatabase design. Section 1.3.2 describes the cleanup of data that was required for importation. Section 1.3.3 describes the digitization of talus, which was needed to expand the previous data. Finally, 1.3.4 discusses the data importation and testing of the geodatabase.

1.4.1. Geodatabase Design

The first step taken in this project was the design of the geodatabase. This was completed by first designing an entity relationship diagram (ERD) and getting feedback on the initial structure. Revisions and changes were then made based on that feedback. The geodatabase structure was then built in ArcCatalog 10.5 and importation of an initial sample dataset was completed to test functionality.

1.4.2. Data Cleaning

The second step that was taken in this project was cleaning up data provided. Data were previously housed in numerous Excel spreadsheets that resulted in several different versions and years of data. These spreadsheets were formatted in numerous different ways by several different people and needed a lot of effort to assign proper naming conventions for all attributes. Once completed this greatly eased importation into the geodatabase.

1.4.3. Digitization of Talus

The next step for this project involved digitizing the remaining talus via www.Caltopo.com. This part of the project was previously started by Dr. Beever and other researchers in an attempt to begin to assign point data to specific talus/polygons. Although consideration was given to continuing to use www.Caltopo.com in an effort to be consistent with previous research due to attributes not being assigned properly, it was necessary to complete this effort in ArcMap 10.5 using an ArcGIS Online base layer.

1.4.4. Importation and Testing of Data

The final step was a combination of importing all data into the completed geodatabase and then testing functionality. The data importation was much faster during this step due to the numerous measures taken in early areas of the project to ensure proper formatting. Queries were also written based on several research objectives discussed with Dr. Beever (i.e. number of continuously extant sites between specific years). Once importation was completed, the queries were run which demonstrated proper database functionality.

1.5. Structure of this Paper

The next chapter discusses research related to this project providing background on geodatabase design, pika ecology, and climate change studies using wildlife. The third chapter outlines the methods used to complete this project. The fourth and fifth chapters discuss the results and conclusions respectively. Following the fifth chapter is a list of references.

Chapter 2 Related Work

This chapter looks at the research and background for geodatabase design, the ecology and biology of the American Pika (*Ochotona princeps*), the impact of climate on their habitat, and the reasons for studying this particular species. Section 2.1 discusses the research for building and populating wildlife biology data in a geodatabase. Section 2.2 explores the biology, ecology, and climate change studies of pika.

2.1. Geodatabase Design

For this project, using examples of geodatabase design for wildlife tracking (Urbano and Cagnacci 2014), as well as multiple sources on geodatabase design for archaeological sites, was invaluable (Mocanu and Velicanu 2011; Gonzalez-Tennant 2009; Breunig et al. 2016). While using the geodatabase design for wildlife tracking may make perfect sense, the use of archaeological geodatabase design may be less obvious. For example, with the island nature of habitable sites for pika (i.e. talus) the model of an archaeological site geodatabase becomes more suitable. Both this geodatabase and archaeological geodatabases are usually bounded by a specific area. Archaeological geodatabases use the excavation site, this project used small talus areas. Outside of these areas, limited factors tend to be included in the database giving both designs island like results. Both archaeological and talus sites are required to have point data within the polygon specifying specific locations which data and attributes are recorded for each. As an example, for archaeological sites, excavators, artifacts, provenance, etc. and for talus, data collectors, pika sign, temperature, etc. These correlations allow for design techniques to be used from archaeological geodatabases and applied to this project.

This project also aims to fill a void within current research for creating geodatabases for niche species. Wildlife tracking databases, such as Urbano and Cagnacci (2014) use thousands of individual data points and then are able to extrapolate a range for the tracked species. This geodatabase model differs and was developed for species that have a specific, known set of criteria in which they persist (i.e. niche species). Instead of tracking animal movement and analyzing range and area based on movement, this database houses survey data for specific sites know to contain the attributes for pika to survive.

Another key consideration with any wildlife tracking is there tends to be very large datasets that need to be "securely, consistently and efficiently managed" as to allow for any number of changes and people to use the dataset (Urbano and Cagnacci 2014). Wildlife spatial data also tends to be housed locally and does not set any standards for interoperability which can then require several separate procedures when analyzing the same datasets (Urbano et al. 2010). These considerations were taken into account when deciding to develop a geodatabase to house all previous information together for faster importation/exportation and analysis along with the idea that because of the cost associated with surveying and collecting wildlife data it is of great importance to share this research with many other researchers (Urbano and Cagnacci 2014).

The geodatabase offers other benefits to housing large amounts of spatial data. Rather than having many different files that need to be accessed for large projects, like wildlife tracking or sensor records, the geodatabase contains all of the files required in one location which frees up system resources. The geodatabase also provides several key elements that allow for better functionality when dealing with spatial data. The creation of feature datasets and the relationships between different feature classes within the datasets provides a perfect layout for the data used in this project. Also, the ability to create domains that limit the inputs for a particular field will help with data entry and error reduction. Finally, the geodatabase can allow for quick updates based on the collected data using a data dictionary which aligns with GPS receivers (Gonzalez-Tennant 2009).

The archaeological geodatabases give us a great basis for developing a presence/absence geodatabase. Building a geodatabase for archaeological sites use temporal, spatial, archaeological, and document data. For this presence/absence wildlife geodatabase there is also temporal, spatial, and instead of archaeological the presence/absence data. The use of the Entity Relationship Diagram (ERD) is vitally important to the creation of the geodatabase. Special attention must also be given to geometry of the data because of its spatial nature and the type of points, lines, or polygons input in a particular field. (Mocanu and Velicanu 2011; Breunig et al. 2016).

The design of the geodatabase must also take into account what the geodatabase will be used for, as well as who will be using it. This project deals with researchers at the USGS which is a federal agency. While not specifically designed for the entire agency some needs based assessment can be used. Some federal agencies have strained to change from simple data collection and storage to “a more collaborative system of data management” (Smith et al. 2015). Many discussions where had with Dr. Beever on what the geodatabase design and usage would encompass. The use of Smith et al. (2015) to better understand what questions to ask allowed for a better understanding of how to construct the geodatabase for long term use.

2.2. Pika Habitat and Climate Change

North American temperature is expected to rise by more than 2°C by the middle of the 21^{st} century and more than 4° by the end of the century, which will lead to changes in climate on a micro and macro level (IPCC 2015). Some of the initial areas expected to be impacted by these temperature changes are mountainous, or montane, regions due to the fact that temperature can change rapidly with elevation. These montane habitats support a wide variety of species of plants and animals, including the pika (Beniston 2003).

Species, including pika, that subsist exclusively in these montane areas have specially adapted traits that allow for them to survive in these select environments (e.g. high body temperature and low thermal conductivity). These traits have evolved previously by climatic pressure forcing species to higher elevations and into these island habitats (Smith and Weston 1990; Beever et al. 2010). However, these selector traits were generally produced by long term adaptation to climate factors rather than short rapid changes. Selector traits normally are formed over the course of thousands of years and with the much more rapid change of climate anticipated in the coming years, those selector traits will not change fast enough for some animals to adapt. For pika, and other high montane species, this rapid and changing climate in their habitable areas is unlikely to allow for the similar dispersal or adaptations seen in previous shifts of habitat range (Barnosky and Kraatz 2007).

Talus are described as "piles of broken rock fringed by suitable vegetation" with rock pieces ranging in size from 0.2m to 1m in diameter (Smith and Weston 1990). Along with selecting talus locations, pika have adapted specific traits to live in these island habitats at high elevation. A high natural body temperature (mean=40.1°C), with a limited range of thermoregulation, and a low thermal conductance allow them to survive in these talus (MacArthur and Wang 1973; Smith and Weston 1990). Although these traits have provided the pika with the ability to survive the climate in high montane areas, they have put limitations on pika's ability to adapt to sudden changes in their climate. The natural high body temperature is close to their high lethal limit and makes the pika susceptible to extreme temperatures, especially higher than normal temperatures during summer months (Otto et al. 2015; Stewart et al. 2015; Beever et al. 2003; 2010; Wilkening et al. 2011).

As a result of expected rapid change in temperatures at high elevations, the range and suitable habitats for pika are likely to greatly diminish. There have been numerous studies of climate related change on pika habitats using many different models (Mathewson et al. 2017; Beever et al. 2003; Beever et al. 2010; Beever et al. 2011; Beever et al. 2013; Wilkening et al. 2011; Millar and Westfall 2010; Calkins et al. 2012; Calkins 2010). These all shows that with expected temperature gains, suitable pika habitats will decrease in total area as temperatures rise. Calkins (2010) showed that with a temperature increase of 7°C, expected habitable areas would decrease by 95% in the Rocky Mountain lineages. Pika habitat loss is expected, and in some cases already observed, to be especially prevalent in the lower elevation habitats, as well as toward the southern boundary of their North American range (Beever et al. 2003; Beever et al. 2010).

While climate related extirpations are on the rise, other pika extirpations of historically habitable sites were seen on the northern latitude of the Great Basin. These extirpations had strong correlations with habitable talus located at lower elevations, in livestock grazing areas, near roadways, and limited nearby habitable talus for dispersion (Beever et al. 2003). In addition, surveys conducted in the 2000's revealed extirpations exclusively located on the southern edge of the pika Great Basin range. These extirpations were much more strongly correlated with rises in overall temperature at those surveyed locations than with the previously seen extirpations at the historical sites (Beever et al. 2010). So, while in recent years there has been strong evidence to support temperature being the cause of extirpations, there are other factors to consider when solely analyzing presence/absence at a given talus location.

Due to their restricted habitat (i.e. high mountain talus), as well as their relative ease of field documentation, pika are a prime candidate for climate impact studies. As Beever et al. (2010) note, the ability to document pika presence, by experienced observers, is as high as 95.9%. This allows for a highly effective field study in which it is easy to detect a change in presence, absence, or extirpation from a particular given site or talus. This project allowed for previously collected pika presence, absence, and extirpation data to be housed and associated with specific talus.

The data that has been collected in previous field seasons also includes temperature data at several pika presence sites collected over multiple years. The inclusion of this data allows for analysis of pika habitats and persistence, but also allows for a collection of temperature data to be accumulated for montane/alpine conditions that may not be easily tracked otherwise. As temperature is a vital determinate of pika presence, having this data available for multiple field sites will dramatically increase the ability to predict extirpation (Beever et al. 2003; Beever et al. 2010; Beever et al. 2011; Beever et al. 2013).

Chapter 3 Methods

This chapter describes the methods used to complete the geodatabase project presented in this paper. Section 3.1 is an overview of the design of the geodatabase and an overview of the methods. Section 3.2 looks at the data, software, and hardware required for the completion of the geodatabase, as well as future field work. Section 3.3 discusses the timeline for completion of this project. Section 3.4 discusses the step by step process for completing the geodatabase and data importation.

3.1 Methods Discussion

The design for this project is laid out in the following section. Section 3.1.1 discusses the geodatabase design. Section 3.1.2., provides an overview of the data used in this project and Section 3.1.3 discusses the detail of the relationships that were built in the database.

After completion of the database, data were formatted from previously housed Excel spreadsheets and exported KML files. These data were then imported into the proper feature classes within the database. Upon completion of the data import, digitization of numerous PATCH polygons was completed to fill in some of the missing information for that feature class. Finally, some queries were run to test for proper importation and housing of imported data.

3.1.1. Geodatabase Design

Initial geodatabase design was done via several personal communications with project researchers to understand what data were already collected and what would be needed for future research. While initial design sessions resulted in the ERD seen in Appendix A. This ERD was developed through several discussions with Dr. Beever and before his field data were turned over to the author (due to field season scheduling conflicts). As such, it was initially much broader and the geodatabase design attempted to contain many more attributes in each Feature Class than was finally deemed necessary. Once the data were procured and visual inspection of the data were completed, some of these attributes changed within their respective feature classes to accommodate formatting of the actual data. The Final ERD can be seen in Figure 2. Many of these changes were made due to data which were not always consistent between field seasons. Also, several of the feature classes, Survey, Mountain range, County, and State, were excluded after further discussions and during examination of the data due to the fact that they could be housed within an attribute of an alternate feature class and were not required to be separate classes at this time.

The design of the initial structure is crucial to the success and usability of a database. The building of a geodatabase requires knowing your audience and knowing your data (Smith et al. 2015). This project's main audience are current USGS researchers. While audience is an important factor so are the data. The data for this project involves years of presence/absence points for pika at specific talus among the Great Basin region of North America. Included in the data are presence/absence points and temperature/relative humidity from several of the field study talus. Knowing that this initial project seeks to answer questions and queries about specific talus and sites of extirpation makes the database design much easier.

This project requires a geodatabase that houses the spatial data from the digitized polygons (i.e. the talus locations), the spatial points taken during site surveys for presence/absence of pika, and temperature records taken using sensors buried at specific locations in chosen talus. Among these data there are different types of attributes for each set of surveys. Attributes of the digitized talus include: surveyed, possible habitable talus for survey,

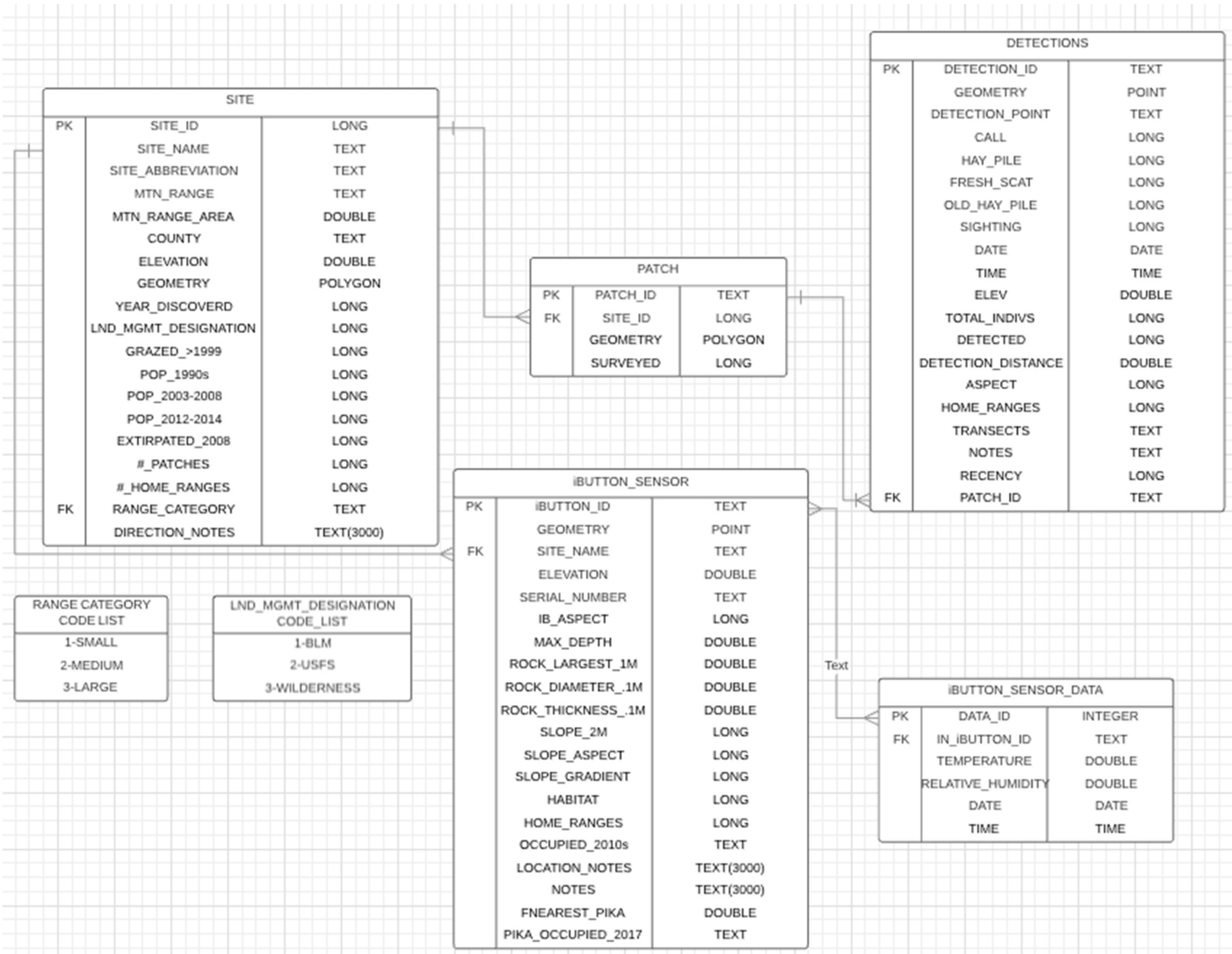


Figure 2. Final ERD with changes to Feature Classes and Attributes

deemed unsuitable, persistent, extirpated, and functionally extirpated. Point data associated with each polygon include: temperature recording points, which include year around temperature data, and presence points with specific attributes the Dr. Beever's research has shown to indicate high probability of extant pika (i.e.) hay piles, calls, etc.) (Beever et al. 2003). The development of feature classes for geodatabases, allows for specific relationships to be built between the different existing features (Gonzoles-Tennant 2009). The ERD in Figure 2 shows feature classes for talus (polygon), temperature (point), presence (point) data, and their relationships.

3.1.2. Data Overview

The data used in this project was all provided by Dr. Beever. It includes point data for 40 individual sites. These sites are several historically known pika extant locations along with sites that Dr. Beever has added to the field seasons through the years. They are provided as points and then are buffered to 3000 meters, which allows Dr. Beever to know which patches need to be field surveyed.

The next data set are patches. These areas encompass an area of talus that is continuously connected. The patches are polygons and are linked to the site data based on an attribute and a relationship class (which is further discussed in Section 3.1.3 and in Figure 4). The patches house many attributes that are recorded during field season surveys which includes surveyed status, pika population, mountain range, land manager, etc.

The final two datasets are point data which included iButton temperature/relative humidity sensors and the presence/absence points recorded from the field surveys. These data contain specific point locations of iButton sensors and pika presence/absence points within each patch polygon. They are also nested in the geodatabase using a relationship class (Figure 4).

3.1.3. Detail of Relationships

Construction of the database occurred by creating a Feature Dataset with multiple Feature classes within the dataset (see Figure 3). The feature classes included point feature classes DETECTIONS (the presence/absence data), the iBUTTON (temperature/relative humidity data), and SITES (overall site data). Also included was the polygon feature class PATCH (patch/talus

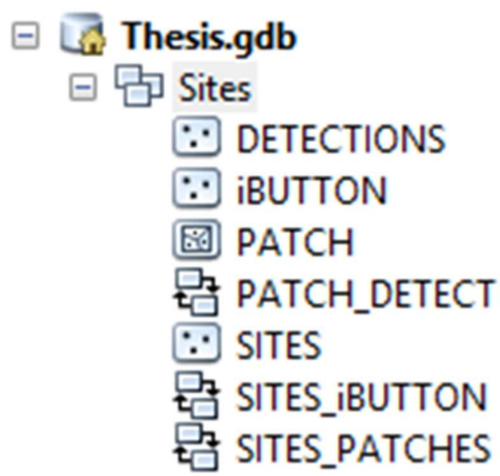


Figure 3. Feature Dataset, Feature Classes, and Joins

data). These were also given relationship classes between PATCH and DETECTIONS, SITES and iBUTTON, and SITES and PATCH to allow for understanding the nested behavior of each different type of feature class in relation to the others. Using Figure 4 as an example, Arc Dome is a Site in the SITES feature class and ARDO1f, ARDO1c, and ARDO1h are talus patches within the PATCH feature class. Within the ARDO1f patch there is a PRESENCE/ABSENCE point 1447. These relationship classes were then used to create the corresponding joins between the feature classes for querying purposes. Initially all database construction was completed within ArcCatalog 10.5, with some subsequent editing done within ArcMap 10.5 and the Catalog extension housed within.

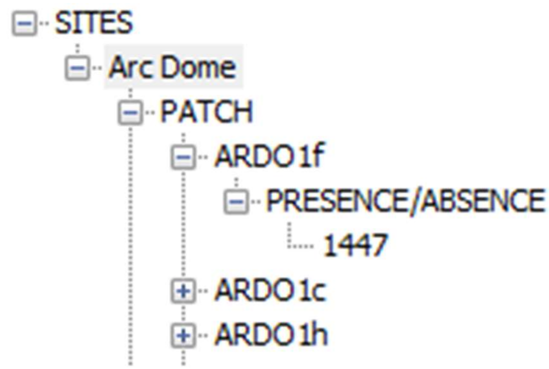


Figure 4. Depiction of Relationship Classes in ArcMap

3.2. Data, Software, and Hardware Requirements

The data, software, and hardware requirements are listed in the following subsections 3.2.1, 3.2.2, and 3.2.3 respectively. 3.2.4 discusses timeframe for project completion.

3.2.1. Data Requirements

The data requirements for this project rely almost exclusively on data provided by Dr. Erik Beever. These data are mostly point features that include presence/absence GPS points taken over numerous field seasons. Other data that was included for this project are polygons and alternative point data that have been housed on www.Caltopo.com. These data are a compilation of previously digitized talus and point data representing both sites and presence/absence points (see Figure 5).

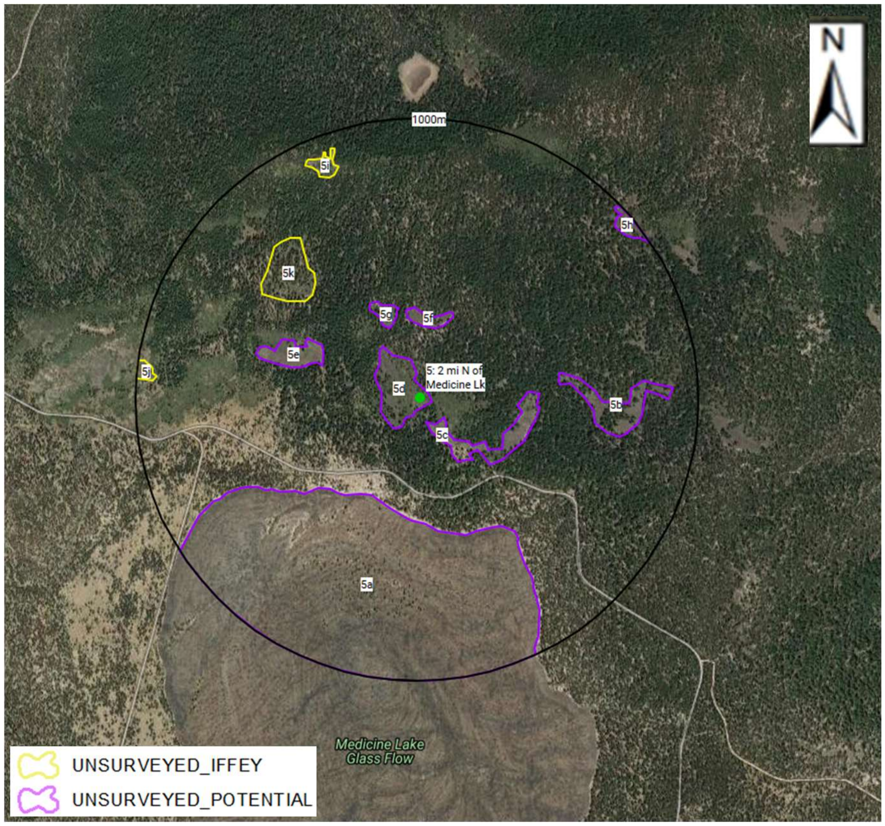


Figure 5. Example of digitized talus on www.caltopo.com

3.2.2. Software Requirements

This project used several different pieces of software to complete. Initial data were housed in Microsoft Excel and www.Caltopo.com. Microsoft Excel was also used to format the field data into the appropriate fields for proper importation into the geodatabase. The polygon and some point data were housed on www.Caltopo.com before exportation into the geodatabase via KML files. The geodatabase was built in ArcCatalog 10.5 and ArcMap 10.5 was used to run queries and verify all data were imported and house correctly.

3.2.3. Hardware Requirements

The software for this project was all used on Windows PCs and several different handheld GPS units were also used during the field collecting process.

3.3. Timeline

The timeline for the completion of this project was structured roughly as seen in Table 1 below. Overall from start to finish the project took roughly 7 weeks to complete starting at data collection and creation of the database structure. This was an easier than expected task, due to all data relevant to the project being accessible. However, there was some work needed to develop the database structure and multiple conversations were had to complete this process.

The formatting/standardization of the field data were a process that was much more involved than originally planned. Due to the fact that data collection has been happening over the span of roughly 20 years, much of the data were housed in different formats. This caused a significant amount of formatting to allow for proper importation into the database. The formatting of these data pushed back the original timetable considerably.

The formatting of data were followed by data exportation from www.Caltopo.com and importation into the geodatabase. There were differences between the importation methods used for the data from Excel versus the data from www.Caltopo.com in KML format and is discussed in Section 3.4.

Finally, many remaining talus patches were digitized to allow for better proof of concept of the geodatabase. Much of the PATCH data were not completed and was required to be added at this point to fully allow for querying of more than one Site. By no means were all of the remaining talus areas digitized at this point and further time was required to complete this section.

Table 1. Project Timeline

	Week 1-2	Weeks 3-6	Weeks 6-7	Week 8
Task 1	Creating database structure			
Task 2		Formatting field data		
Task 3			Exporting/importing datasets	
Task 4				Digitizing polygons

3.4. Detailed Methods

This section discusses in detail the different processes taken after design and build of the geodatabase. Section 3.4.1 discusses the process used for data standardization and importation into the geodatabase. Section 3.4.2 discusses the digitization of additional talus polygons and 3.4.3 discusses the queries used to test the geodatabase once completed.

3.4.1. Data Standardization and Importing

Field detection data for Dr. Beever's Sites have been collected on and off since the mid 1990's and were all previously housed in Excel spreadsheets with some years combined and others housed separately. These data were collected as point data and attributes were assigned based on field collection (i.e. hay piles, call, sighing, etc.). One issue with these data were that they have been collected and formatted previously by different people and as such not all data were collected or housed with the same attributes. The inconsistency with these data coupled with the upwards of 2000 records made for many manual corrections, as well as many null values for records in which the data were not captured. One example of some Detection data is seen in Figure 6. The spreadsheet housed attributes for each point including Site, Date, Time, Pika detected, etc.

wpt ID	ele	magvar	name	cmt	desc	SiteName	Abbrev'n	
78	2493.526367	CM132PK1C	20:21	7/19/2013	20:21	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
96	2527.893311	CM13PK2CT	20:40	7/19/2013	20:40	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
79	2492.80542	CM132PK3C	21:03	7/19/2013	21:03	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
97	2510.829834	CM13PK4C	21:17	7/19/2013	21:17	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
80	2520.202881	CM132PK5C	5:43	7/20/2013	5:43	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
98	2487.278076	CM13PK6CT	6:05	7/20/2013	6:05	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
99	2487.278076	CM13PK7HC	6:12	7/20/2013	6:12	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
100	2487.518311	CM13PK8C	7:01	7/20/2013	7:01	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
84	2482.951904	CM13PK10S	6:33	7/20/2013	6:33	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
101	2469.493652	CM13PK9C	6:40	7/20/2013	6:40	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
85	2483.913086	CM13PK11H	6:48	7/20/2013	6:48	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
86	2515.636475	CM13PK12C	7:16	7/20/2013	7:16	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
87	2522.365723	CM13PK13C	8:00	7/20/2013	8:00	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
81	2530.056152	CM133PK14CS	8:08	7/20/2013	8:08	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
88	2392.348633	CM13PK15C	8:45	7/20/2013	8:45	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
89	2390.906494	CM13PK16C	8:52	7/20/2013	8:52	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
90	2401.480957	CM13PK17C	9:00	7/20/2013	9:00	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
82	2396.434082	CM13NOPK1	9:09	7/20/2013	9:09	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
91	2402.682617	CM13PK19C	9:17	7/20/2013	9:17	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
92	2406.047119	CM13PK19H	9:33	7/20/2013	9:33	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
93	2384.177246	CM13PK20H	10:48	7/20/2013	10:48	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
94	2371.920654	CM13PK21H	11:39	7/20/2013	11:39	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx
95	2445.701172	CM13PK22C	12:25	7/20/2013	12:25	Crane Mtn., Warner Mtns., se OR	CRMO	eTrex VistaHCx

Figure 6. Unformatted Detection Data (Latitude and Longitude Redacted)

Beginning with one spreadsheet that represented the most complete record of attributes, the initial formatting was constructed. The easiest and most consistent process was found to format/copy all records into one single CSV file for importation rather than attempting to import multiple CSV files into the feature class. This was due to formulas and other formatting within the original Excel spreadsheets. Each column was titled appropriately based on the naming convention in the DETECTIONS feature class and formatted for the correct type of data (i.e. Double, Integer, Text, etc.). These names were all based on Dr. Beever's collection records and can be seen below in Figure 7. Once all records were copied into the proper columns the spreadsheet was saved as a CSV.

Field Name	Data Type
OBJECTID	Object ID
DETECTION_POINT	Text
CALL	Long Integer
HAY_PILE	Long Integer
FRESH_SCAT	Text
OLD_HAY_PILE	Long Integer
SIGHTING	Long Integer
Date_	Date
Time_	Text
Elev	Double
Total_indivs	Long Integer
Detected	Text
Detection_distance	Double
Aspect	Long Integer
HOME_RANGES	Long Integer
TRANSECTS	Text
NOTES	Text
LONGITUDE	Double
LATITUDE	Double
Location_Notes	Text
RECENCY	Text
Shape	Geometry

Figure 7. Detection Feature Class Formatted Fields

In one instance of the previously saved detection data, the XY coordinates were saved in a separate location. To create the necessary CSV file, the data from the two separate spreadsheets was merged into one workbook. Once the data were within the same workbook a VLOOKUP was used to find the matching data based on the location name. The matching data were copied to the proper location within a single spreadsheet and then formatted correctly to allow for proper importation into the database.

Importing the CSV files into the geodatabase was done in a couple different steps. In ArcMap, the Add Data tool was used to add the CSV files to a blank map. The CSV files initially come in as tables with no geographic data displayed. Right clicking and selecting view XY data allows for selection of the columns associated with Longitude and Latitude in decimal degrees, as well as elevation data, and the proper datum (in this case WGS84). Once the XY data is displayed the whole table can be exported as a feature class. This feature class can then be imported into the premade feature dataset already housed in the geodatabase, with each column creating an attribute corresponding to the attribute desired in the geodatabase feature class. This technique was used for importing the Presence/Absence, Site, and iBUTTON data.

Data for the talus polygons was housed two separate ways. The polygons of the talus themselves were housed via www.CALTOPO.com. This website has high resolution imagery that allows for digitization of polygons and uploading points to have the visual data associated with the detections and surveyed talus. This data is available to export via KML files which was done for all previously digitized talus. The polygons were then imported into ArcMap via the KML to Layer tool, which creates a feature class from a KML file. Upon completion of this step this feature class was imported into the PATCH feature class in the feature data set only importing the polygon data and the color data, as this was how the surveyed/unsurveyed

PATCH

PATCH_NAME	SITE_ID *	SURVEYED	OBJECTID *	HALL LOCN *	Site ID *	Site ab	Mtn Range	County	Year Discov'd	Elevation	LandMgmt Design	Grazed since 1999	Mtn-range Area
bmP1	31	UNSURVEYED	31	Blowout Mountain	31	BLMO	<Null>	Humboldt	2011	2045.4	Wilderness	<Null>	<Null>
bmY1	31	UNSURVEYED	31	Blowout Mountain	31	BLMO	<Null>	Humboldt	2011	2045.4	Wilderness	<Null>	<Null>
bmY2	31	UNSURVEYED	31	Blowout Mountain	31	BLMO	<Null>	Humboldt	2011	2045.4	Wilderness	<Null>	<Null>
bmY3	31	UNSURVEYED	31	Blowout Mountain	31	BLMO	<Null>	Humboldt	2011	2045.4	Wilderness	<Null>	<Null>
bmY4	31	UNSURVEYED	31	Blowout Mountain	31	BLMO	<Null>	Humboldt	2011	2045.4	Wilderness	<Null>	<Null>
bmY5	31	UNSURVEYED	31	Blowout Mountain	31	BLMO	<Null>	Humboldt	2011	2045.4	Wilderness	<Null>	<Null>
COPEb1	23	SURVEYED	23	Cougar Pk.	23	COPE	<Null>	Lake	1925	1692.2	USFS	1	<Null>
COPEb2	23	SURVEYED	23	Cougar Pk.	23	COPE	<Null>	Lake	1925	1692.2	USFS	1	<Null>
COPEp1	23	UNSURVEYED	23	Cougar Pk.	23	COPE	<Null>	Lake	1925	1692.2	USFS	1	<Null>
COPEp2	23	UNSURVEYED	23	Cougar Pk.	23	COPE	<Null>	Lake	1925	1692.2	USFS	1	<Null>
COPEp3	23	UNSURVEYED	23	Cougar Pk.	23	COPE	<Null>	Lake	1925	1692.2	USFS	1	<Null>
COPEp4	23	UNSURVEYED	23	Cougar Pk.	23	COPE	<Null>	Lake	1925	1692.2	USFS	1	<Null>
COPEp5	23	UNSURVEYED	23	Cougar Pk.	23	COPE	<Null>	Lake	1925	1692.2	USFS	1	<Null>
DUPE	1	SURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE1	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE10	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE11	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE12	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE13	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE2	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE3	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE4	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE5	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE6	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE7	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78
DUPE8	1	UNSURVEYED	1	Duffer Pk.	1	DUPE	Pine Forest	Humboldt	1935	1885.7	BLM	1	1425.78

Figure 8. Example of Attributes for PATCH Feature Class joined with SITES Feature Class

polygons were distinguished. Once imported, the CSV file housing the remaining data were imported with the proper attributes for each talus polygon (see Figure 8). These attributes coupled with the Joins between the different feature classes allow for substantial querying of collected field data.

A table for the collected data by each iButton sensor was added in the final step. This data is a simple Excel spreadsheet containing temperature and relative humidity that is collected every four hours from placement of the sensor until retrieval or end of life. This data were included in the geodatabase in a table simply to show proof of concept that the data could be housed if necessary and for querying purposes for this project. In the future it may be more beneficial to house this data externally due to the sheer number of records that each sensor generates (upwards of 2000 records per sensor per year). The table was then joined with the iButton point feature

3.4.2. Digitizing Polygons

The final step to completing the geodatabase was the continuation of digitizing the talus that had been surveyed or may be surveyed in the future. This process was relatively simple once the feature class for PATCH was complete. Using the create feature tool, an Esri Basemap with satellite imagery, additional talus patches were digitized. This process included choosing what type of patch was to be digitized, i.e. Surveyed, Unsurveyed – Potential, and Unsurveyed – Iffy (this designation was included from previously digitized data for consistency). Patches were selected in discussions with Dr. Beever and examining previously digitized areas for similar categorized areas. The selected patch was then digitized by simply encompassing the patch area in the polygon on top of the imagery (see Figure 9). The patch data does come with a caveat, the patch data included for this project in no way encompasses all patch data that is needed for complete querying of presence/absence points. It was required to digitize more patch areas because of how few had been previously completed. However, with the already completed talus patches, along with the patches digitized for this project, there is ample data for proof of concept/design.

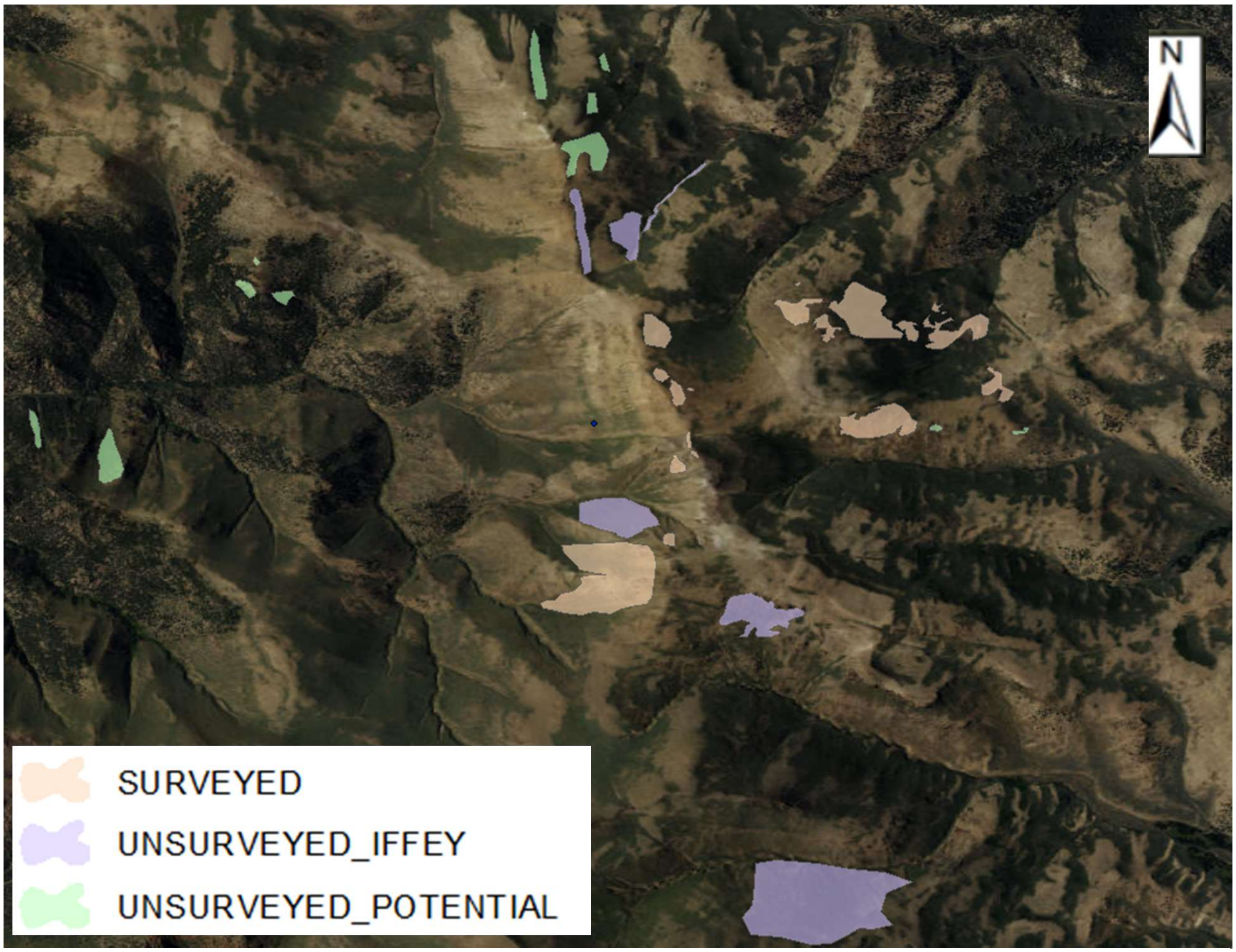


Figure 9. Example of Digitization of patches in ArcMap

One difficulty that was uncovered in digitization and importation of the talus was some attributes that had to be manually altered to allow for proper querying. Due to www.Caltopo.com only being able to record polygons and a color (i.e. Surveyed, Unsurveyed, etc.) and not being able to house attributes such as the site relationship for the PATCH polygons or the patch relationship to the PRESENCE point data, much of this was required to be added manually in the database. This required inspecting each site area and adding the proper Site number to the proper attribute field in the PATCH feature. The same was required for each Presence/Absence and iButton point feature that fell within a digitized PATCH. This required some effort because some sites have many different patches and each needed to be done individually. Because of this limitation it was also decided to digitize additional talus patches in ArcMap rather than on www.Caltopo.com and export them into the geodatabase.

3.4.3. Queries

Finally, once all data were imported, data verification occurred by testing several queries (see Table 2). These queries were built based on conversations with Dr. Beever regarding his future research questions. The results of these queries are discussed in full in the following chapter.

Table 2. Queries for Testing Geodatabase Functionality

Queries	Questions Asked	Question Drivers
Query 1	What records were compiled during the 2016 field season?	Will allow for better planning of future field seasons.
Query 2	What are the minimum and maximum elevations of detected pika?	Elevation is an important factor for ongoing pika extirpation research.
Query 3	How many potential patches need to be surveyed?	Will focus field season surveys towards patches in need of survey.
Query 4	What iButton locations have pika occupancy from 2010 and 2017?	Example of basic analysis of extant populations.
Query 5	What is the highest temperature recorded by an iButton Sensor?	Allows for understanding of where highest temperature was recorded.
Query 6	What sites had a decrease in pika population from the 1990’s to the 2000’s?	Example of basic analysis of population decreases between surveyed patches.

Chapter 4 Results

This chapter discusses the results seen in the completed database. Section 4.1 describes all final data and relationships used, as well as numbers of records included in the final geodatabase build. Section 4.2 provides the answers to the results of the queries. Finally, section 4.3 summarizes all results for this project.

4.1. Final Design and Data

Section 4.1.1 looks at the final geodatabase design. Section 4.1.2 looks at the final data and records housed in the geodatabase after importation and digitization.

4.1.1. Final Design/Schema

The final design took the initial ERD and changed certain feature classes and attributes to account for collected data and ease of querying. The final ERD for this project is shown in Figure 3 above and includes the feature classes and relationships deemed necessary for database design. As discussed in the previous chapter the end design for the project required several changes to allow for proper importation of the data. The original schema was designed before taking possession of any data and was based on initial understanding of database needs. Once the data were in hand, the understanding of how to properly house it in the geodatabase was better realized, leading to the changes previously seen.

4.1.2. Geodatabase Data and Records

At the completion of this project, the feature classes and tables did not house all of Dr. Beever's data. There are still many talus polygons to digitize, detection and temperature data to import, and possible future survey sites to be identified and imported. However, for this project 40 records were imported for the SITES feature class, 292 records for the iBUTTON feature class, 405 records that combine previously digitized and current digitization for the PATCH feature class, 1602 records for the DETECTIONS feature class, and 4096 records for the temperature and relative humidity data for one iBUTTON location (see Figure 10 and Table 3). At this time, the SITES and iBUTTON feature classes house all locations and data known, while the PATCH and DETECTIONS feature classes and the SENSORS table house some, but not all, of the data collected.

Table 3. Final Count of Imported Records

Features	Type	# of Records
Sites	Point	40
Detections	Point	1602
iButton	Point	292
Patch	Polygon	405
Temperature/Relative Humidity	Table	4096

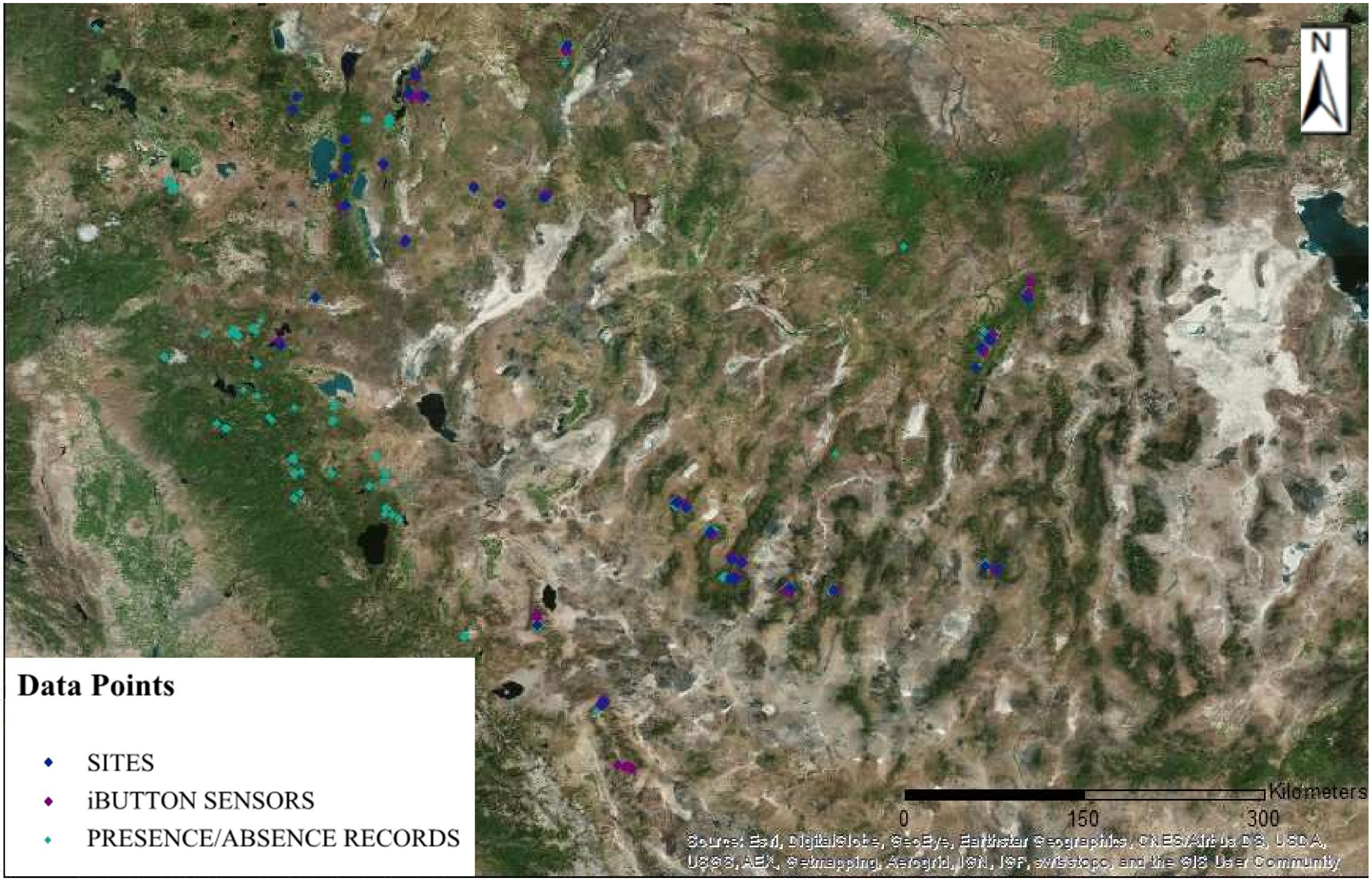


Figure 10. Visualization of all data showing Pika habitat in the study area of the Great Basin Region (Patch/Polygon data is not visible at this scale)

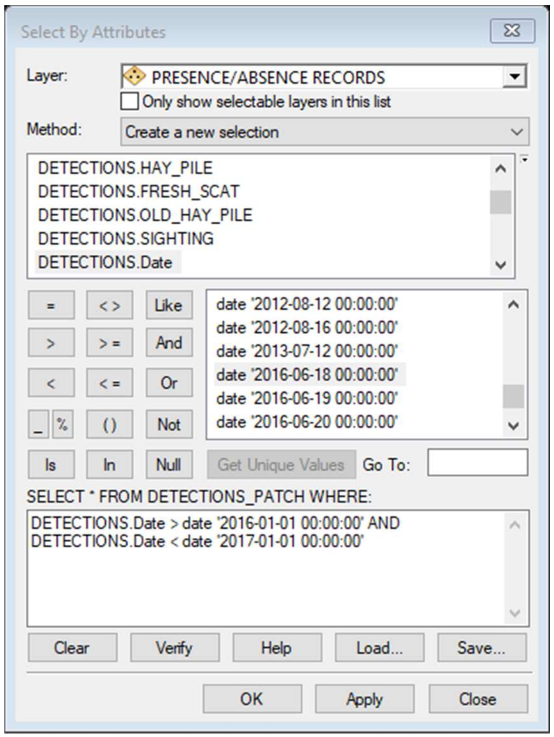
4.2. Query Results

Upon completion of all data importation, the previously listed queries (seen in Table 2) were done using ArcMap and the Selector tool to confirm proper importation and containment. The Tables below show the queries with their results both graphically and in table format. The individual queries showed all results as expected and allowed for nesting queries as needed. These were used to accommodate some of the queries that involved calculating the maximum or minimum values. Once imported into the geodatabase the spatial nature of the data becomes more apparent. The ability to spatial analyze these data along with the other attributes provides an enhanced understanding of how all the pieces of the field data interconnect.

Query 1 returned all values in the Detections feature class that were recorded in 2016. This was 120 results and they are visualized in Figure 11. Queries such as this will allow better understanding of field collection seasons and will allow for a better understanding what sites and surveys were recorded in a given year. Seeing these spatially also allows for a specific and concise plan for future field seasons.

Query 2 shows how researchers could better understand what elevations pika are currently inhabiting and where they have extirpated from previously known habitat. The ability to look at how far pika have retreated upwards or extirpated from specific locations, based on elevation, will be valuable in future research. These queries returned both older records from the mid 1990s with a minimum value of 1258.8m and a maximum of 3557m (see Figure 12 and Figure 13). Further analysis regarding where these detections were recorded, as well as if any have been recorded more recently, would allow for better understanding of pika persistence today.

Query Number	Query	Results
Query 1	What records were surveyed during field season 2016?	120



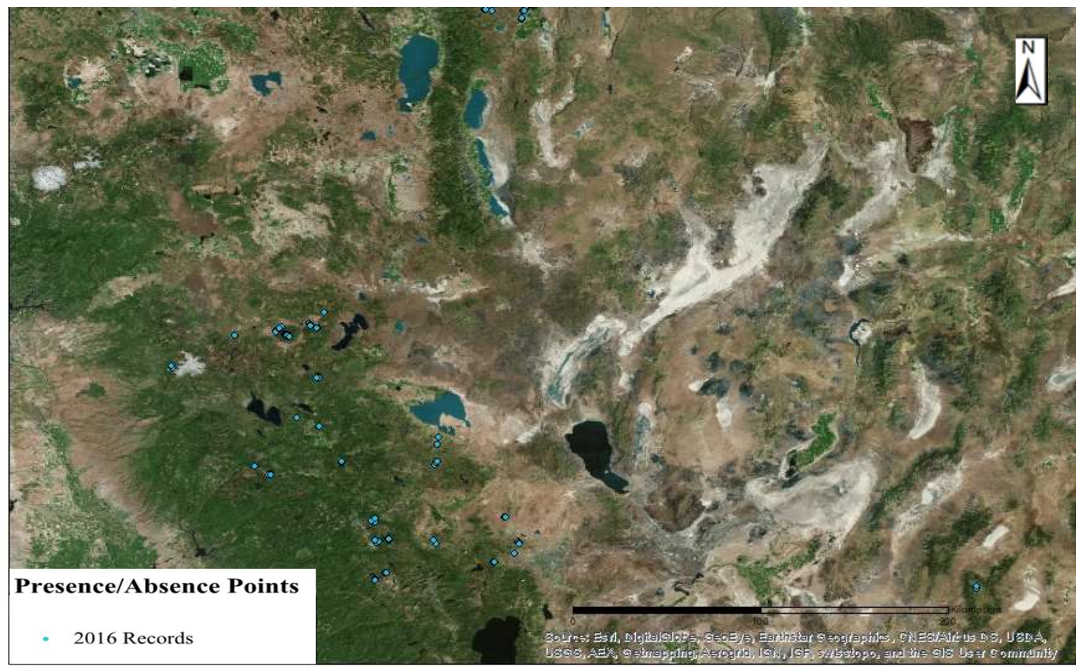
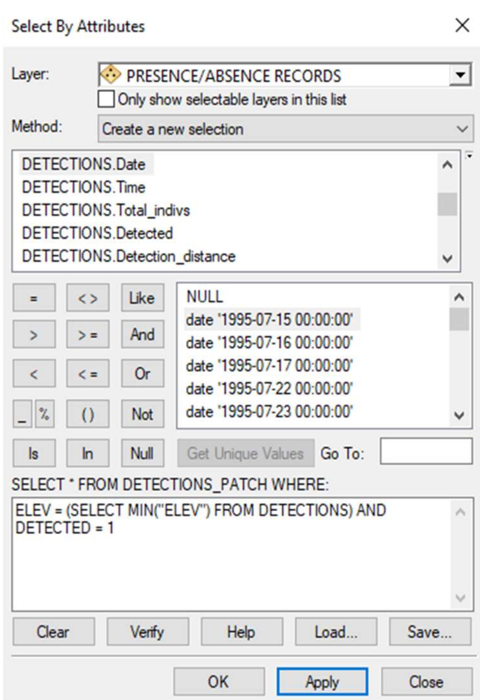
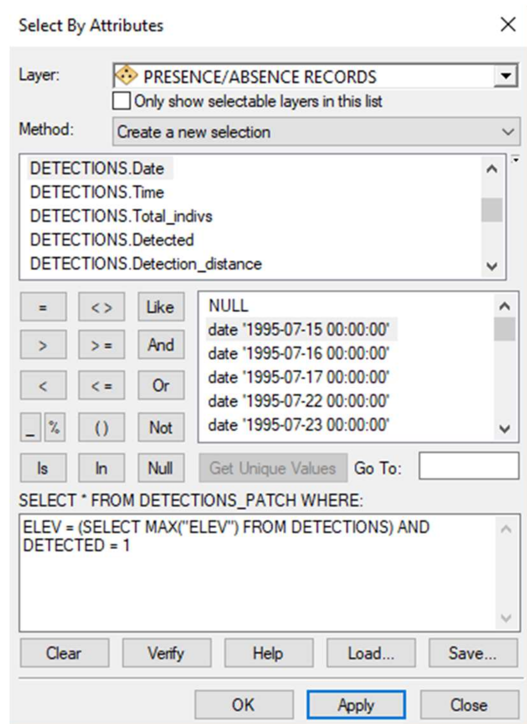


Figure 11. Query and Results showing locations of presence/absence points taken during the 2016 field season.

Query 2	What are the minimum and maximum elevations of detected pika?	Min: 1258.8m Max: 3557m



Query of Minimum Elevation with Detected Pika



Query of Maximum Elevation with Detected Pika

Figure 12. Queries for finding the maximum and minimum elevation of pika detections

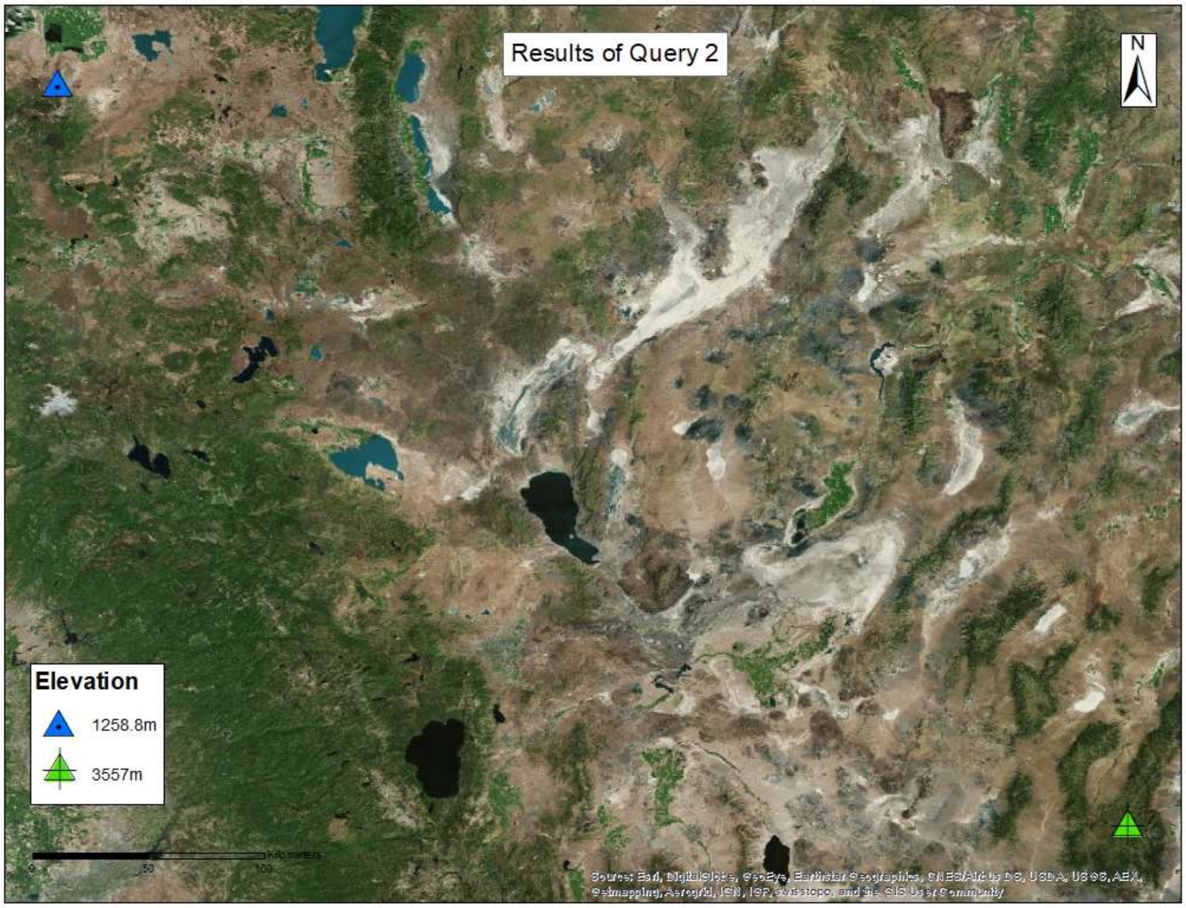
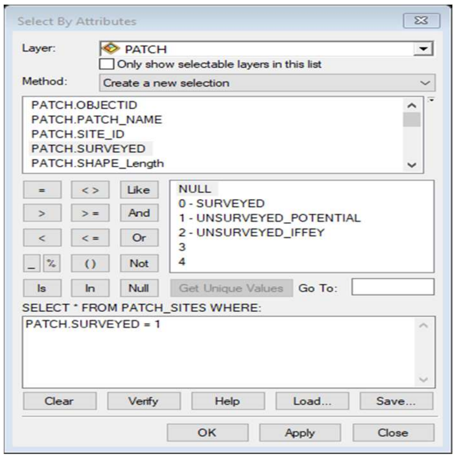
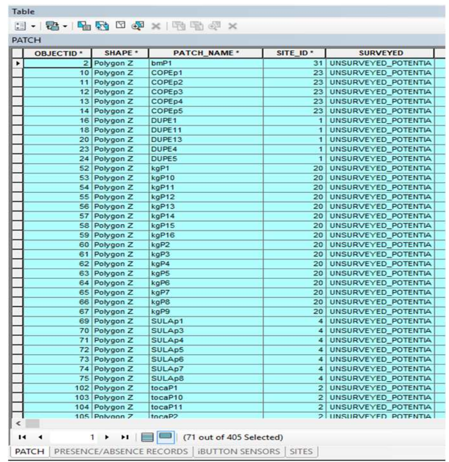


Figure 13. Results for finding the maximum and minimum elevation of pika detections

In Query 3, the patches that still need to be surveyed in a given site were queried. This will allow users to plan future field seasons more efficiently. By better understanding and spatially visualizing when and what areas have been surveyed recently, new survey areas can be more effectively planned for a given year. Currently the database showed 71 unsurveyed potential patch locations that need to be surveyed (see Figure 14). As discussed previously, not all talus areas have been digitized. Once this is completed this query will generate increased results.

Query 3	How many potential patches need to be surveyed?	71





Table

PATCH

OBJECTID *	SHAPE *	PATCH_NAME *	SITE_ID *	SURVEYED
2	Polygon Z	bmP1	31	UNSURVEYED_POTENTIA
10	Polygon Z	COPEp1	23	UNSURVEYED_POTENTIA
11	Polygon Z	COPEp2	23	UNSURVEYED_POTENTIA
12	Polygon Z	COPEp3	23	UNSURVEYED_POTENTIA
13	Polygon Z	COPEp4	23	UNSURVEYED_POTENTIA
14	Polygon Z	COPEp5	23	UNSURVEYED_POTENTIA
16	Polygon Z	DUPE1	1	UNSURVEYED_POTENTIA
18	Polygon Z	DUPE11	1	UNSURVEYED_POTENTIA
20	Polygon Z	DUPE13	1	UNSURVEYED_POTENTIA
23	Polygon Z	DUPE4	1	UNSURVEYED_POTENTIA
24	Polygon Z	DUPE5	1	UNSURVEYED_POTENTIA
52	Polygon Z	kgP1	20	UNSURVEYED_POTENTIA
53	Polygon Z	kgP10	20	UNSURVEYED_POTENTIA
54	Polygon Z	kgP11	20	UNSURVEYED_POTENTIA
55	Polygon Z	kgP12	20	UNSURVEYED_POTENTIA
56	Polygon Z	kgP13	20	UNSURVEYED_POTENTIA
57	Polygon Z	kgP14	20	UNSURVEYED_POTENTIA
58	Polygon Z	kgP15	20	UNSURVEYED_POTENTIA
59	Polygon Z	kgP16	20	UNSURVEYED_POTENTIA
60	Polygon Z	kgP2	20	UNSURVEYED_POTENTIA
61	Polygon Z	kgP3	20	UNSURVEYED_POTENTIA
62	Polygon Z	kgP4	20	UNSURVEYED_POTENTIA
63	Polygon Z	kgP5	20	UNSURVEYED_POTENTIA
64	Polygon Z	kgP6	20	UNSURVEYED_POTENTIA
65	Polygon Z	kgP7	20	UNSURVEYED_POTENTIA
66	Polygon Z	kgP8	20	UNSURVEYED_POTENTIA
67	Polygon Z	kgP9	20	UNSURVEYED_POTENTIA
69	Polygon Z	SULAp1	4	UNSURVEYED_POTENTIA
70	Polygon Z	SULAp3	4	UNSURVEYED_POTENTIA
71	Polygon Z	SULAp4	4	UNSURVEYED_POTENTIA
72	Polygon Z	SULAp5	4	UNSURVEYED_POTENTIA
73	Polygon Z	SULAp6	4	UNSURVEYED_POTENTIA
74	Polygon Z	SULAp7	4	UNSURVEYED_POTENTIA
75	Polygon Z	SULAp8	4	UNSURVEYED_POTENTIA
102	Polygon Z	tocaP1	2	UNSURVEYED_POTENTIA
103	Polygon Z	tocaP10	2	UNSURVEYED_POTENTIA
104	Polygon Z	tocaP11	2	UNSURVEYED_POTENTIA
105	Polygon Z	tocaP2	2	UNSURVEYED_POTENTIA

(71 out of 405 Selected)

PATCH | PRESENCE/ABSENCE RECORDS | iBUTTON SENSORS | SITES

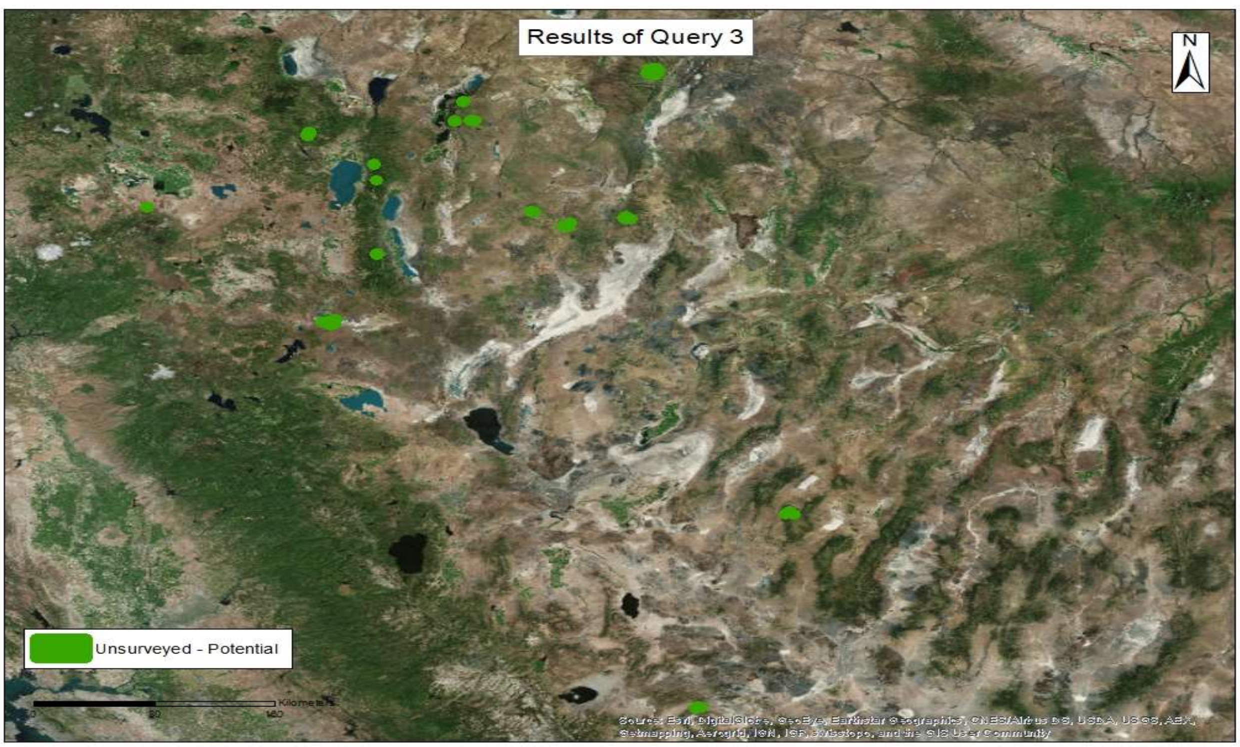
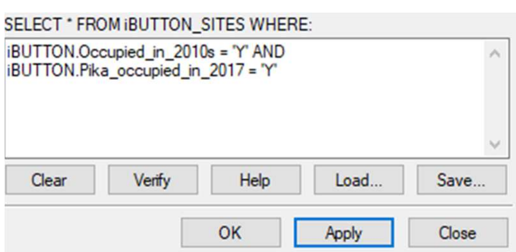
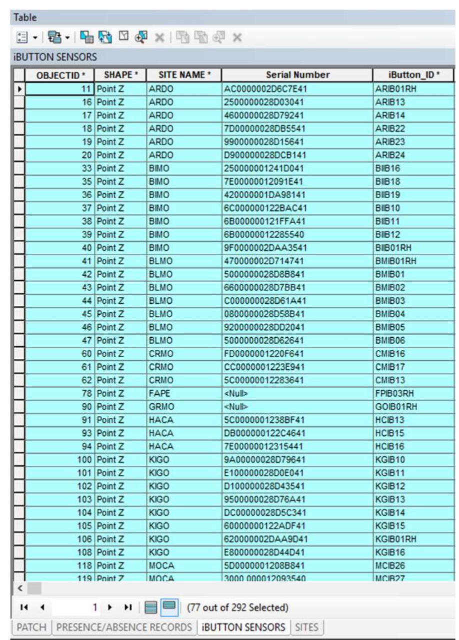


Figure 14. Query and Results of Unsurveyed Potential Patches

The data for the iButton locations is particularly interesting as it adds relative humidity and temperature data for each talus patch where the sensors are placed. Also, all talus data (i.e. presense/absence, etc.) is collected with each placing and collection of the sensor data, which provides additional valuable information. Query 4 finds the results of pika persistence from iButton sensor locations from 2010 survey to 2017 survey. The query returned 77 of the iButton locations as having persistent pika (see Figure 15).

Query 4	What iButton locations have pika occupancy from 2010 and 2017?	77





Table

iBUTTON SENSORS

OBJECTID *	SHAPE *	SITE NAME *	Serial Number	iButton_ID *
11	Point Z	ARDO	AC0000002D6C7E41	ARIB01RH
16	Point Z	ARDO	2500000028D03041	ARIB13
17	Point Z	ARDO	4600000028D79241	ARIB14
18	Point Z	ARDO	7D00000028DB5541	ARIB22
19	Point Z	ARDO	9900000028D15641	ARIB23
20	Point Z	ARDO	D900000028DCB141	ARIB24
33	Point Z	BIMO	250000001241D041	BIIB16
35	Point Z	BIMO	7E00000012091E41	BIIB18
36	Point Z	BIMO	420000001DA98141	BIIB19
37	Point Z	BIMO	6C000000122BAC41	BIIB10
38	Point Z	BIMO	6B000000121FFA41	BIIB11
39	Point Z	BIMO	6B00000012285540	BIIB12
40	Point Z	BIMO	9F0000002DAA3541	BIIB01RH
41	Point Z	BLMO	470000002D714741	BMIB01RH
42	Point Z	BLMO	5000000028D8B841	BMIB01
43	Point Z	BLMO	6600000028D7BB41	BMIB02
44	Point Z	BLMO	C000000028D61A41	BMIB03
45	Point Z	BLMO	0800000028D58B41	BMIB04
46	Point Z	BLMO	9200000028DD2041	BMIB05
47	Point Z	BLMO	5000000028D62641	BMIB06
60	Point Z	CRMO	FD0000001220F641	CMIB16
61	Point Z	CRMO	CC0000001223E941	CMIB17
62	Point Z	CRMO	5C00000012283641	CMIB13
78	Point Z	FAPE	<Null>	FPIB03RH
90	Point Z	GRMO	<Null>	GOIB01RH
91	Point Z	HACA	5C0000001238BF41	HCIB13
93	Point Z	HACA	DB000000122C4641	HCIB15
94	Point Z	HACA	7E00000012315441	HCIB16
100	Point Z	KIGO	9A00000028D79641	KGIB10
101	Point Z	KIGO	E100000028D0E041	KGIB11
102	Point Z	KIGO	D100000028D43541	KGIB12
103	Point Z	KIGO	9500000028D76A41	KGIB13
104	Point Z	KIGO	DC00000028D5C341	KGIB14
105	Point Z	KIGO	6000000122ADF41	KGIB15
106	Point Z	KIGO	620000002DAA9D41	KGIB01RH
108	Point Z	KIGO	E800000028D44D41	KGIB16
118	Point Z	MOCA	5D0000001208B841	MCIB26
119	Point Z	MOCA	3000 000012093540	MCIB27

(77 out of 292 Selected)

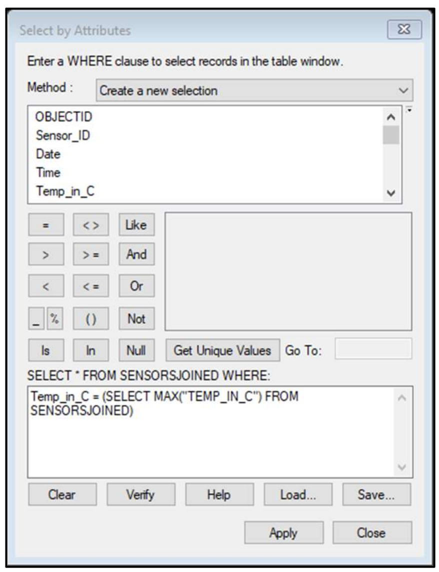
PATCH | PRESENCE/ABSENCE RECORDS | iBUTTON SENSORS | SITES

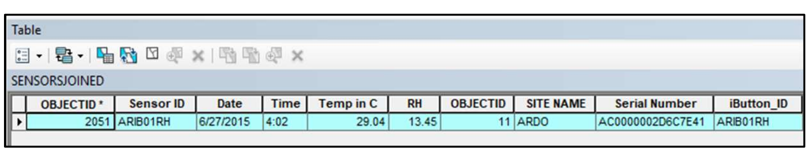
Figure 15. Query and Results of iButton Locations with Pika Occupancy in both 2010 and 2017

Continuing with iButton queries, but this time focusing on the sensor table, Query 5 asks what the highest temperature recorded in the sensor table was. This resulted in a record for

29.04°C (see Figure 16). A caveat with this query is that only one set of sensor data were imported from only one sensor. This was due to the large number of records that are stored by each sensor. As stated above, one sensor for a 2 year period generated over 4000 records. This data is also not readily available to the author at this time and so one sensor was included to show proof of concept.

Query 5	What is the highest temperature recorded by an iButton Sensor?	29.04°C





Table

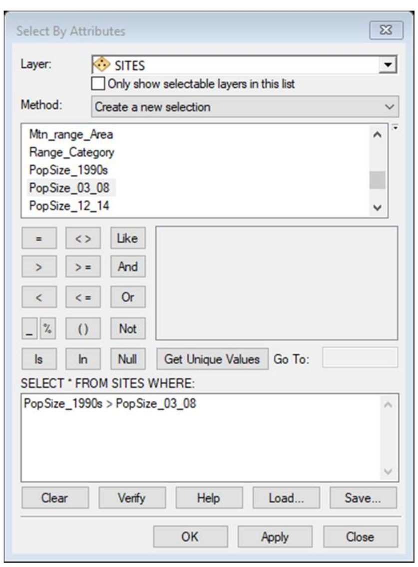
SENSORSJOINED

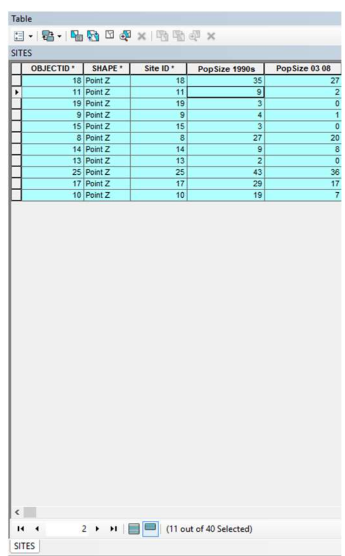
OBJECTID *	Sensor ID	Date	Time	Temp in C	RH	OBJECTID	SITE NAME	Serial Number	iButton_ID
2051	ARIB01RH	6/27/2015	4:02	29.04	13.45	11	ARDO	AC0000002D6C7E41	ARIB01RH

Figure 16. Query and Results of Highest Temperature Recorded

The final query was run to search how many sites saw a decrease in pika population from the 1990s levels to the 2000s levels. The results returned 11 sites that have seen pika population decreases between the two time periods (see Figure 17 and 18). This final query was run as a more personal exercise to see if pika populations were in fact reducing during the study timeframe. While some of the sites did see reductions in pika populations, it should be noted that others saw increases. Some of these variations could simply be due to factors surrounding the survey, but the number of sites that saw an increase was far outweighed by the number that decreased.

Query 6	What sites had a decrease in pika population from 1990's to the 2000's?	11





Table

SITES

OBJECTID *	SHAPE *	Site ID *	PopSize 1990s	PopSize 03 08
18	Point Z	18	35	27
11	Point Z	11	9	2
19	Point Z	19	3	0
9	Point Z	9	4	1
15	Point Z	15	3	0
8	Point Z	8	27	20
14	Point Z	14	9	8
13	Point Z	13	2	0
25	Point Z	25	43	36
17	Point Z	17	29	17
10	Point Z	10	19	7

(11 out of 40 Selected)

SITES

Figure 17. Query and Results of Sites that Lost Pika Population between the 1990's field seasons to the 2003-2008 field seasons

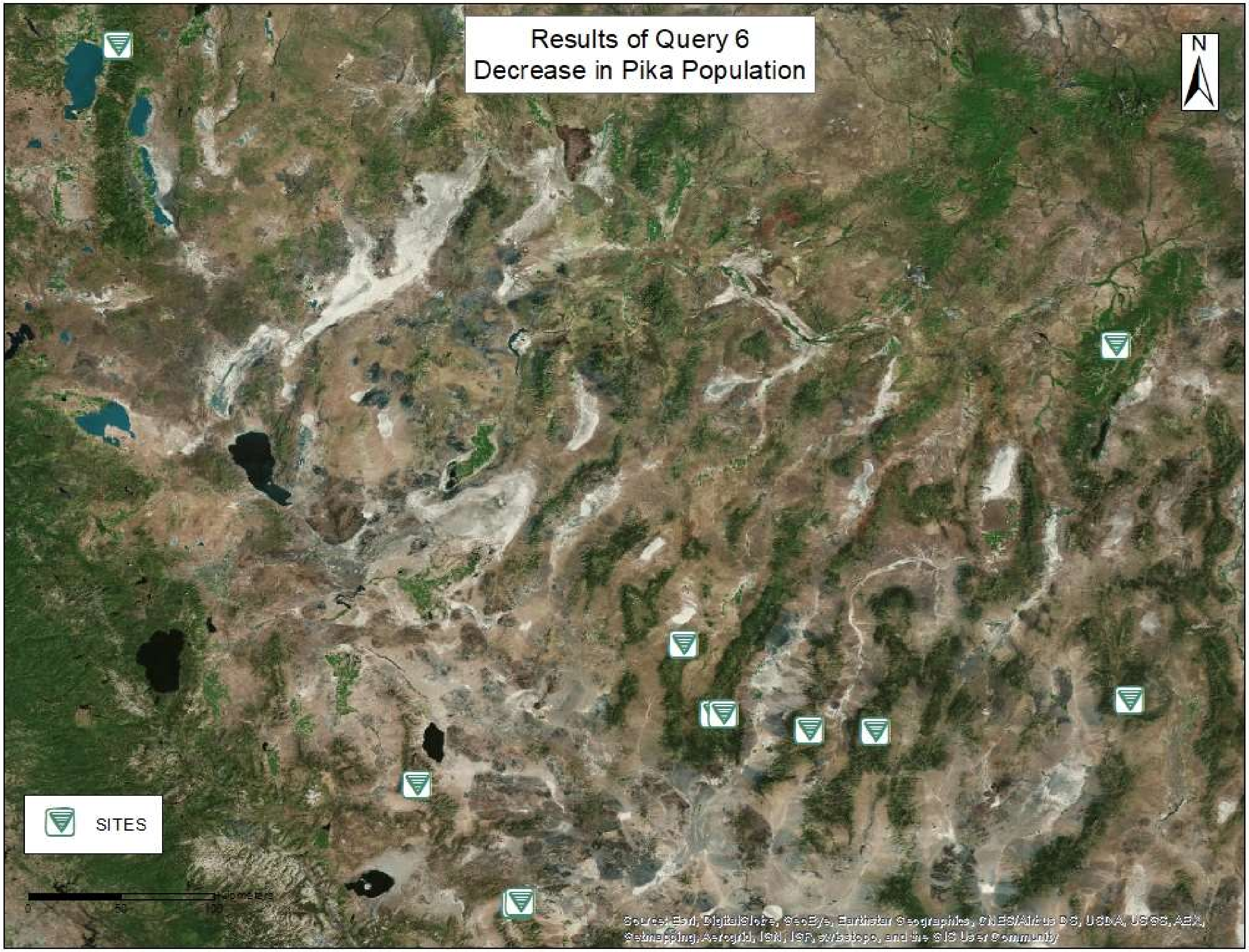


Figure 18. Results of Query 6

4.3. Summarizing of Results and Errors

As can be seen by the previous sections, the database was successfully populated to allow for querying of the data. The importation of many of the Detection records and creating the relationships with the attributes and the sites will greatly benefit future research. The ability for users to visualize the relationships between the Detections, the Patch, and the Site will allow them to greatly increase the speed of analysis and presentation of results.

The building of the relationship classes between Sites, Patches, Detections, and iButton sensors will allow for future collaboration in this research and understand the nesting concept of this work (see Figure 19). The hope is that future field seasons can be easily planned using this geodatabase and the contained records. The visualization of all data from the macro (Site level) to the micro (Presence/Absence points) allows the users to intuitively see the relationships.

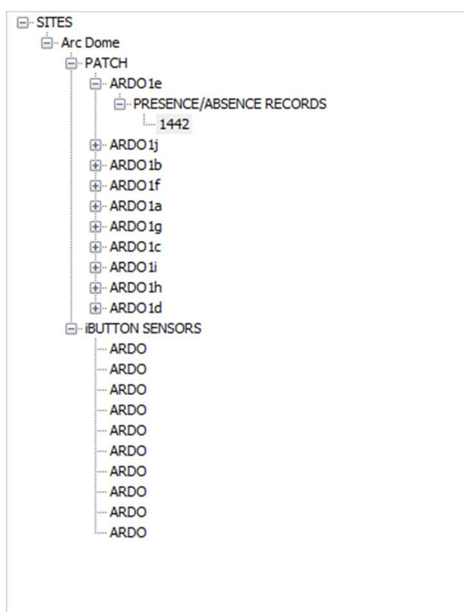


Figure 19. Relationship of nested features example showing the Arc Dome Site with the ARDO1e Talus Patch related to it with the Presence/Absence Point 1442 related to both.

A final example, which can be seen below in Figure 20, shows a complete view of Site 11 with all associated feature classes. The relationships for this site are shown above in Figure 19, which represent Site 11, the Patches associated with Site 11, the Detection points for each Patch, and the iButton sensors placed in each Patch. This was the goal at the beginning of this project and so the belief is that this will ease the analysis of collected data and greatly increase the research capabilities. With the goal of housing all data in one database met, the ability to continue with further spatial analysis on all data will be a huge benefit to current and future research.

There were several errors that were found when importing the data. These revolved around the formatting of each attribute in Excel and how they were saved in the CSV file. One example was several of the attributes were initially built in the database as a Short or a Long type but had to be adjusted once the data were imported. This required several iterations of importation and reformatting the database to accommodate how the data were originally formatted during the field collection.

One feature in ArcCatalog that the author found particularly helpful was the ability to build a geodatabase schema based on the formatting of the import data. Gathering all the data into the CVS file and formatting each attribute accordingly allows for ArcCatalog to build a feature class schema based on the formatting from the CVS file. This allows for quick database construction and eliminates the struggles of possibly building your feature class attributes with the wrong type code.

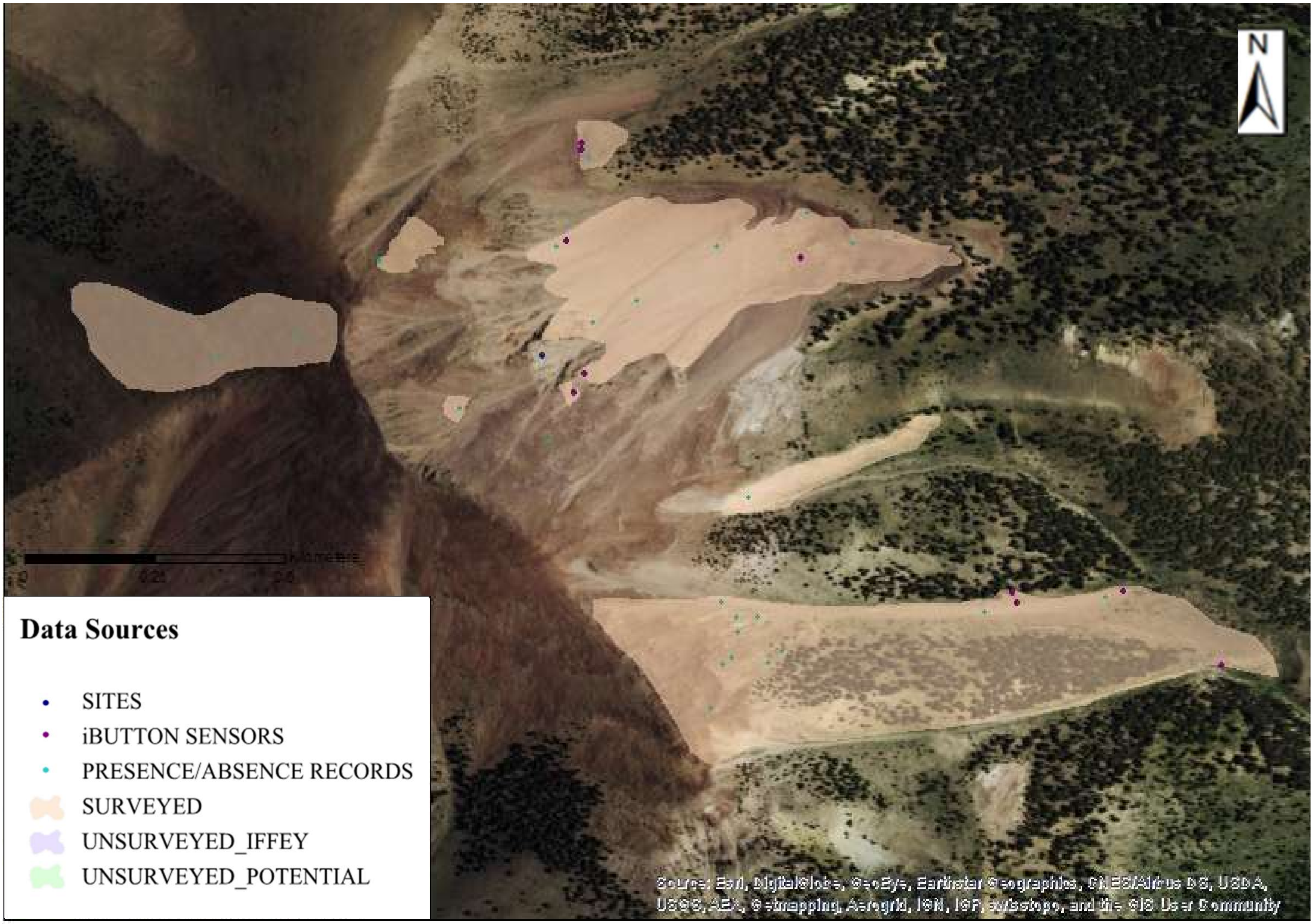


Figure 20. Example of Site with all Feature Classes

Chapter 5 Conclusion

In this chapter the conclusions drawn from this project are discussed. Section 5.1 discusses the lessons learned while completing this project. Section 5.2 discusses the uses and current completion state of housing all the data in the geodatabase. Section 5.3 discusses the future work related to this project.

5.1. Project Lessons

This study aimed to build a geodatabase to house the research collected by Dr. Beever's research project on pika and climate change. This project was originally much more complex than the final design. Due to many of the time commitments to formatting and importing the data, less time was spent on including all surveyed data and more on formatting the base data. The underlying structure that was completed during this project provides a framework that allows for expansion should others continue with this work.

The original complexity of the project was mostly due to limited understanding on how the data were surveyed. This then led to the initial design of the database which proved much more complicated than necessary, as well as created some more formatting issues to allow for proper importation. The design portion would have been much better served to have allowed time to gather all datasets and review them at once, rather than slowly piece them together over several weeks. This probably could have been avoided but, due to overlapping deadlines for this project and field season requirements, it was necessary.

Another challenge of this project was understanding exactly how all pieces of research were connected. After several discussions with Dr. Beever and reviewing of all data this became clearer and allowed for much faster geodatabase design. The ability for laymen to quickly visualize and understand how the data fit together will allow for a greater collaboration on research in the future. The visualization of the geographically nested data will benefit the project, in the ability to quickly find points associated with each Site or Patch with relative ease. The hope is that future field seasons can be easily planned using this geodatabase and the contained records.

Additionally, much more time than was previously thought was needed to format the data for proper importation. This did involve some querying within Excel to allow for the proper attributes to be imported to the proper feature classes. In the future, this may be able to be avoided by field collecting the data in slightly different format, or translating field work to the Excel spreadsheets in a more efficient way. Much time was spent simply testing data importation because of errors in proper formatting of the columns in the CSV files. Also as previously discussed, it was also found to be much more efficient to simply format the CSV file and then use that to create the formatting for each attribute in the feature class by importing the schema (e.g. Double, Text, Float, etc.) which alleviated much of the formatting issues encountered early in the project.

5.2. Current State of Geodatabase

Due to the ongoing nature of the pika research project, this geodatabase will more than likely need to be updated frequently. The included records and schema for this project were an attempt to show that using a geodatabase will greatly increase querying ability to quickly understand which sites have been surveyed, when they were surveyed, and what the results of those surveys. This process will also be enhanced due to the visual results provided by the geodatabase platform. Previously it was required to comb through, query separate spreadsheets, or rely on institutional knowledge, to understand what was needed to be accomplished each field season.

While this project does not include all records that have been collected. The basis has been provided to show how the data may be housed that will allow for speedy importation in the future. With the addition of a data dictionary, the data collection to importation process would speed up immensely and allow for much faster results each field season. Having a data dictionary would be a great asset if implemented in future field seasons. The construction of a data dictionary will eliminate all the need to format data and allow for concise and accurate collection of the data. However, the data dictionary construction was not attempted during this project due to time constraints.

5.3. Future Work

The future work for this project could be extensive. The data collection is currently scheduled to continue for the foreseeable future and because of this, the need for this database should be ongoing for years to come. With continuation of the data collection comes the opportunity to ease the data collection methods. It would greatly speed up collection and post-processing of the field data if some type of data dictionary could be used to collect the data in the field. This could be set up on iPad's using Collector for ArcGIS or via a Trimble GPS unit and GPS Pathfinder. These hardware and software options would provide a valuable source of collecting that would all but eliminate any formatting needs once the data were collected.

The use of a data dictionary would allow for coded values and other such methods to limit certain inputs that are being used in the database. This would make for a simple easy download and import into the geodatabase at the end of each field day or field session. Once this system is implemented, the geodatabase would be much more up to date and alleviate the formatting issues that plagued this project.

Another aspect that should be expanded on is the inclusion of more temperature data. This data encompasses over 200 locations with accurate temperature and relative humidity for an extended area of the western United States. The inclusion of all of these records will allow for a valuable understanding of how climate is being affected at higher elevations over the course of many years. These data can also be compared to data from temperature models to determine how accurate they are at predicting temperature in montane environments.

Collected alongside the temperature data were also relative humidity data. These data will also allow for further analysis of climate at higher elevations and allow for a more complete picture of change, should it be shown to have occurred. The inclusion of relative humidity data can greatly increase the spatial analysis that can be performed and correlated to burn areas and further Dr. Beever's research on climate change and impacts. The links between relative humidity and fire danger have been shown to be particularly correlated with mean relative humidity as the best overall predictor (Holsten et al. 2013). Many important analyses could be completed using these two different data sets.

The inclusion of temperature and relative humidity data will cause some issues because of the number of records generated. The sensors collect roughly 2000 records a year per sensor which can become unwieldy. As seen in other environmental studies, other software such as *Loggernet* may be better equipped to handle data volumes that are seen for these sensors (Gries et al. 2016).

There is also an opportunity to house more data that has been collected at some, but not all, of the Sites (i.e. survey specific data). This data includes, but is not limited to, wind speeds, percent of area recently burned, percent of area with snow, animal sightings, and types of trees and shrubs. While these attributes were not included for this project, it would be beneficial in the long term to add this data, which would provide a more complete understanding of the surveyed locations. The inclusion of landscape and other variables are often the best predictors when trying to understand wildlife habitat (Carroll et al. 1999).

Finally the complete housing of all of research in one location will allow for a much more thorough analysis of these data. With the ability to see all work in one location, and with a streamlined data collection process, research will be able to continue in a much more efficient way. It will also allow for the data to be presented and analyzed geographically. This will help to understand and visualize where specific changes have occurred and provide a palette from which to work from for future research.

References

Beever, Erik A., Peter F. Brussard, and Joel Berger. 2003. "Patterns of Apparent Extirpation among Isolated Populations of Pikas (Ochotona princeps) in the Great Basin. *Journal of Mammalogy.* 84, No. 1: 37-54. http://www.jstor.org/stable/1383633.

Beever, Erik A. Chris Ray, Philip w. Mote, and Jennifer Wilkening. 2010. "Testing alternative models of climate-mediate extirpations." *Ecological Applications.* 20, No. 1:164-178. http://www.jstor.org/stable/27797796.

Beever, Erik A., Chris Ray, Jenifer L. Wilkening, Peter F. Brussard, and Philip W. Mote. 2011. "Contemporary climate change alters the pace and drivers of extinction." *Global Change Biology.* 17: 2054-2070. Doi: 10.1111/j/1365-2486.2010.02389.x.

Beever, Erik A., S.Z. Dobrowski, J. Long, A.R. Mynsberge, and N.B. Piekielek. 2013. "Understanding relationships among abundance, extirpation, and climate at ecoregional scales." *Ecology.* 94, No. 7: 1563-1571. http://www.jstor.org/stable/23596945.

Bruenig, Martin, Paul V. Kuper, Edgar Butwilowski, Andreas Thomsen, Markus Jahn, Andre Dittrich, Mulhim Al-Doori, Darya Golovko, and Mathia Menninghaus. 2016. "The story of DG4GeO – A service-bassed geo-database architecture to support multi-dimensional data analysis and visualization." *ISPRS Journal of Photogrammetry and Remote Sensing.* 117:187-205. Doi: 10.1016/j.isprsjprs.2015.12.006

Calkins, Michael T., Erik A. Beever, Kenneth G. Boykin, Jennifer K. Frey, and Mark C. Andersen. 2012. "Not-so-splendid isolation: modeling climate-mediated range collapse of a montane mammal *Ochotona* princeps across numerous ecoregions." *Ecography.* 35 No. 9: 780-791. Doi: 10/1111/j.1600-0587.2011.07227.

Calkins, Michael T. 2010. "Modeling Climate Change Impacts on the American Pika Over the Western United States." Master's Thesis. New Mexico State University.

Carroll, Carlos, William J. Zielinski, and Reed F. Reed F. 1999. "Using Presence-Absence Data to Build and Test Spatial Habitat Models for the Fisher in the Klamath Region, U.S.A." *Conservation Biology.* Vol. 13, No. 6:1344-1359. http://www.jstor.org/stable/2641958.

Gonzalez-Tennant, Edward. 2009. "Using Geodatabases to Generate 'Living Documents' for Archaeology: A Case Study from Otago Goldfields, New Zeland." In "Current Research in Australia and New Zeland.," *Historical Archaeology* 43, No. 3: 20-37. Accessed July 11, 2016. http://www.jstor.org/stable/25617568.

Gries, Corinna, Mark R. Gahler, Paul C. Hanson, Timpoth K. Kratz, and Emily H. Stanley. 2016. "Information management at the North Temperate Lakes Long-term Ecological Research site – Successful support of research in a large, diverse, and long running project." *Ecological Informatics* 36: 201-208. http://dx.doi.org/10.1016/j.ecoinf.2016.08.007

Hafner, David J. 1993. “North American Pika (*Ochotona princeps*) as a Late Quaternary Biogeographic Indicator Species.” *Quaternary Research* 39: 373-380. DOI:10.1006/qres.1993.1044

Holsen, Anne, Anto Raja Dominic, Luis Costa, and Jurgen P. Kropp. 2013. “Evaluation of the performance of meteorological forest fire indices for German federal states.” *Forest Ecology and Management.* 287:123-131. Doi: 10.1016/j.foreco.2012.08.035.

MacArther, R.A. and L.C.H. Wang. (1974). “Physiology of thermoregulation in pika, *Ochotona princeps.” Canadian Journal of Zoology* 51:11-16. https://doi-org.libproxy1.usc.edu/10.1139/z73-002.

Mathewson, Paul D., Lucas Moyer-Horner, Erik A. Beever, Natalie J. Briscoe, Michael Kearney, Jeremiah M. Yahn, and Warren P. Porter. 2017. “Mechanistic variables can enhance predictive models of endotherm distributions: the American pika under current, past, and future climates.” *Global Change Biology.* 23:1048-1064. DOI:10/1111/gcb.13454.

Millar, Constance I., and Robert D. Westfall. 2010. “Distribution and Climatic Relationship of the American Pika (*Ochotona princeps*) in the Sierra Nevada and Western Great Basin, U.S.A.; Periglacial Landforms as Refugia in Warming Climates.” *Arctic, Antarctic, and Alpine Research* 42, No. 1: 76-88. http://www.jstor.org/stable/25642094.

Mocanu, Aura-Mihaela and Manole Velicanu. 2011. “Building a Spatial Database for Romanian Archaeological Sites.” *Database Systems Journal* 2 No. 1 (2011): 3-12. http://www.dbjournal.ro/archive/3/3.pdf

Moyer-Horner, Lucas, Paul D. Mathewson, Gavin M. Jones, Michael R. Kearney and Warren P. Porter. 2015. “Modeling behavioral thermoregulation in a climate change sentinel.” *Ecology and Evolution* 5. No. 24:5810-5822. DOI:10.1002/ece2.1848.

Otto, Hans W., James A. Wilson, and Erik A. Beever. 2015. “Facing a Changing World: Thermal Physiology of American Pikas.” Western North American Naturalist. 75, No. 4:429-445. http://libproxy.usc.edu/login?url=https://search-proquest-com.libproxy1.usc.edu/docview/1770937946?accountid=14749.

Rodhouse, Thomas J. Erik A. Beever, Lisa K. Garrett, Kathryn M. Irvine, Mackenzie R. Jeffress, Michael Munts, and Chris Ray. 2010. “Distribution of American pikas in a low-elevation lava landscape: conservation implications from the range periphery. *Journal of Mammalogy.* 97, No. 5: 1287-1299. American Society of Mammalogists. http://www.jstor.org/stable/40925736.

Romero-Lankao, P., J.B. Smith, D.J. Davidson, N.S. Diffenbaugh, P.L. Kinney, P. Kirshen, P. Kovacs, and L. Villers Ruiz. 2014: North America. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*

[Barros, V.R., C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1439-1498. http://www.ipcc.ch/pdf/assessment-report/ar5/wg2/WGIIAR5-Chap26_FINAL.pdf

Rowe, Rebecca J. and Rebecca C. Terry. 2014. "Small mammal responses to environmental change: integrating past and present dynamics." *Journal of Mammalogy.* 95, No. 6: 1157-11174. American Society of Mammalogists. DOI: 10.1644/13-MAMM-S-079.

Smith, Jordan W., William S. Slocumb, Charlynne Smith, and Jason Matney. 2015. "A Needs-Assessment Process for Designing Geospatial Data Management Systems within Federal Agencies." *Journal of Map & Geography Libraries* 11, No. 2: 226-244. DOI:10.1080/15420353.2015.1048035.

Stewart, Joseph A. E., John D. Perrine, Lyle B. Nichols, James H. Thorne, Constance I. Millar, Kenneth E. Goehring, Cody P. Massing, and David H. Wright. 2015. "Revisiting the past to foretell the future: summer temperature and habitat area predict pika extirpations in California." *Journal of Biogeography* 42:880-890. http://wileyonlinelibrary.com/journal/jbi DOI:10.1111.12466.

Urbano, Ferdinando and Francesca Cagnacci, (eds.) *Spatial Database for GPS Wildlife Tracking Data.* Springer International Publishing Switzerland, 2014. DOI: 10.1007/978-3-319-03743-1_1

Wilkening, Jennifer L., Chris Ray, Erik A. Beever, and Peter F. Brussard. 2011. "Modeling contemporary range retraction in Great Basin pikas(Ochotona princeps) using data on microclimate and microhabitat." *Quaternary International.* 235: 77-88. Elsevier. DOI: 10.116/j.quaint.2010.05.004.

Appendices

Appendix A – Initial Entity Relationship Diagram