Where in Washington State is the pygmy saxifrage? The predicted distribution of *Saxifraga hyperborea*

by

Jaileen C. Merced Hoyos

A thesis
submitted in partial fulfillment
of the requirements for the degree
Master of Environmental Studies
The Evergreen State College
December 2021

©2021 by Jaileen C. Merced Hoyos. All rights reserved.

This Thesis for the Master of Environmental Studies Degree

by

Jaileen C. Merced Hoyos

has been approved for

The Evergreen State College

by

____

John Withey, Ph.D.

Member of the Faculty

December 31, 2021

Date

ABSTRACT

Where in Washington State is the pygmy saxifrage? The predicted distribution of *Saxifraga hyperborea*

Jaileen C. Merced Hoyos

Alpine ecosystems face some of the highest risks from climate change around the world. Conservation efforts are necessary to prevent the extinction of many alpine species. In order to develop reliable conservation policies and practices, more information about species' distributions, habitats, and threats in alpine areas is necessary. This project built a species distribution model (SDM) for *Saxifraga hyperborea*, a Washington state alpine sensitive plant species, and examined its state status. *Saxifraga hyperborea* is ranked as sensitive in Washington but bordering countries and states have ranked the species as secured. To build the model, I used 6 climatic layers; 1) continentality, 2) mean annual precipitation, 3) growing degrees day, 4) number of frost-free days, 5) precipitation as snow, and 6) climatic moisture deficit, along with 31 presence and 93 pseudo-absence points. I used the ArcGIS Forest Based Classification tool, which is based on the Random Forest algorithm. The result of this model was a map of potential habitat for *Saxifraga hyperborea* in Washington State. The model predicted promising areas in the Okanagan and Canadian Rockies, where the species has not been found to date in Washington State. This study is the beginning of a potential journey to resolve the species' confusing state status and to identify areas to potentially find more populations of *Saxifraga hyperborea*.

Table of Contents

Table of Figures

List of Tables

Acknowledgements

I would like to extend my deepest gratitude to my thesis advisor John Withey for his patience, guidance, and his continued support throughout my thesis journey. Thank you for always lifting my spirit when I most needed it. This project would not have been possible without Washington Natural Heritage program and Walter Fertig who provided access to the baseline data. I also want to thank the Rare Care program and Wendy Gibble for providing valuable advice and the opportunity to work and explore the wonderful alpine ecosystem of Washington State.

I would like to recognize the invaluable support from my partner Chris Taber. Thank you for holding my hand through this long and enriching graduate experience, for making me dinner and snacks on the long nights and, for the celebratory beers at the end of a productive night. Mom, dad and Nata, thank you for the last minutes edits and for providing me with an example to follow. I would have not been able to do this without your love and support. Judy and Rick thank you for always cheering me on. My friends Shohei, Madison, Friscka and the quails. Thank you for believing in me; our friendship is one of the best things that has happened on this thesis ride. Finally, but not least, Coquito, thank you for the licks, the snuggles and the forced walks; I love you.

Introduction

Approximately 15-24% of the Earth is covered by arctic and alpine tundra together (Charles et al., 2020). In the alpine zones in Europe, Inouye (2019) estimated that more than 80% of suitable habitat will be lost, for 36-55% of alpine species, and 31-51% subalpine species, by 2070-2100. Among the many effects on the world's flora due to climate change, especially warming temperatures, alpine plants face some of the highest risks (Verrall et al., 2020; Wershow et al., 2018). For example, one of the biggest threats from climate warming to alpine species is the migration of subalpine trees and woody shrubs into the alpine ecosystem (Charles et al., 2020; Rumpf et al., 2017). In the Tibetan Plateau, Wang et al. (2014) found that vegetation coverage distribution has increased by 18% from 1972-2009, triggered by higher temperatures and precipitation. The arrival of new plant and animal species is predicted to introduce competitors that could potentially create changes in the plant and pollinator interactions (Inouye, 2019). More importantly, the change in range of species has impacts on biodiversity, biome integrity and ecosystem services (Oldfather, 2018).

Additional threats to alpine ecosystems include changes in weather patterns, melting of permafrost, changes in trophic interactions, and loss of relict species with disjunct distributions (Charles et al. 2020). Also, changes in precipitation, snowmelt and snowpack will affect the growing season, reproductive behavior, and activity of pollinators (Korner, 2021; Inouye, 2019). The rapid change in climate is "winning the race" and alpine plants have not typically been able to adapt or respond—and so we have reached the point where assisted species migration and hybridization has been proposed to be a conservation strategy for some alpine plants (Charles et al., 2020).

Saxifraga hyperborea is an alpine plant that grows in moist rock crevasse, talus on mountain peaks and can be found in Canada, Russia, Greenland, and the western United States (Herbaria, 2020). *Saxifraga hyperborea*'s unique habitat requirements make searching and monitoring this species complex. Only those willing to climb and backpack can ensure an encounter with this plant. Additionally, the plant is small and hard to notice. It could grow up to 10 cm tall, it forms patches of up to 8 cm wide, have mostly basal leaves that are 3-6 mm broad, and its population distribution is sparse (Camp et al., 2011).

I was fortunate to encounter *Saxifraga hyperborea* twice last summer. My job with the Rare Care program at the University of Washington brought to my attention the alarming state of alpine ecosystems. Part of my job was to set permanent plots for the National Park Service to continue the monitoring of endangered, threaten and sensitives alpine plants in the National Parks areas. At the end of the summer, I spoke with Wendy Gibble, Rare Care program coordinator, she mentioned multiple alpine plant species that could benefit from Species Distribution Model study. Based on present data availability I decided to study *Saxifraga hyperborea*.

The plant is categorized as secure and widespread in most of these locations. In contrast, in the state of Washington, it is listed as sensitive. For example, in Idaho, the plant is listed as SNR, which means "conservation not yet assessed," in Oregon, it is not listed as a species of concern, and in Canada, it is listed as secure (Washington Natural Heritage Program, 2019; Klinkenberg et al., 2020; Idaho Fish and Game, n.d.). There is a concerning discrepancy in the status of this plant in the Pacific Northwest, particularly in Washington, which is the only state in the west that has it listed as sensitive.

After a conversation with two rare plant experts in Washington State, both agreed on the enigma behind the status of this plant. According to the Washington Natural Heritage Program database, there are 46 known locations of this plant in Washington State. For instance, the distribution of the species is concentrated in the Northeast Olympic Peninsula and the North Cascades. To better understand the distribution of the species in Washington State, this project intended to build a species distribution model (SDM) for *Saxifraga hyperborea*. The SDM was made using the Forest-Based Classification tool in ArcGIS, six environmental layers and presence and pseudo-absence points. The reason for *Saxifraga hyperborea* to be listed as sensitive could be due to the lack of information on population locations. Therefore, this SDM aimed to detect areas of potentially suitable habitat for *Saxifraga hyperborea*. The suitable habitat detected by the model could also serve as potential areas for the reintroduction of *Saxifraga hyperborea* in the future.

Literature Review

<u>Plant Conservation in the United States</u>

"*The Earth's ecosystems continue to be destroyed and degraded at unprecedented rates, and the capacity to restore them is extremely limited. But we must begin to address this need if we want to maintain biodiversity and ecosystem services into the future. Ecosystem services are the benefits people obtain from ecosystems, including provisioning of food and water; regulation of atmosphere, floods, drought, land degradation, and disease; support for soil formation and nutrient cycling; pollution filtering; and cultural services such as recreational, spiritual, religious, and other nonmaterial benefits (Millennium Ecosystem Assessment 2005), and plants serve as the foundation of ecosystems. Plants are not optional; they are essential to life and central to the future of human well-being. Plants provide habitat, food, cover, nesting areas, and more for the planet's wildlife. This rich legacy of biodiversity is an invaluable and irreplaceable component of our natural heritage and deserves our protection."* (Havens et al., 2014, p. 10)

A reflection on Havens' quote allows us to realize that we need to remember the significance of ecosystem services for our survival. More importantly, it reflects on how plants are necessary to maintain balance. Humans cannot allow or afford the extinction of more plant species or any species. Preserving and sustaining biodiversity must be a priority since we depend on these ecosystems (Brandt et al., 2014). Therefore, the protection and preservation of endangered species play an essential role in our survival (Mahoney, 2009). Mass extinction has happened before on Earth, driven by natural events. Today, mass extinction is 99% due to human behavior. (Primack, 2014). According to the United Nations report, more than 30% of amphibians, 33% of coral, 10% insects and, 33% of all mammals are threatened (Msuya et al., 2019). Angiosperms are disappearing at a rate of 3 species per year, 500 times faster than if it had happened by natural causes (Humphreys et al.,2019). To reduce the stress these statistics signal, we need to focus on what has survived and protect these species. According to O.G. Wilson, ecosystems may be reduced and degraded, but if a species survives, there is still hope to recover an ecosystem (Wilson, 2010).

In addition to the fast degradation of ecosystems, conservationists must accommodate political interests. Available funding and budgets play a significant role in this race against time. The Endangered Species Act (ESA) passed in 1973 required preserving and restoring endangered species populations. (Mahoney, 2009). The ESA and the Fish and Wildlife Service developed a recovery guideline to identify the most at-risk species to allocate funding. Species are categorized according to the 1) degree of threat, 2) potential for recovery, and 3) taxonomic distinctiveness (Negron-Ortiz,2014; Fish and Wildlife Service, Interior, 1983). Species classified as endangered are the ones who get ESA recovery funding (Mahoney, 2009). Since the listing of endangered species is a political decision influenced by external factors, this brings a disproportionate allocation of funding that focuses money on only a few species (Ando, 1999; Mahoney, 2009). According to Mahoney (2009), the size of the species affects the amount of funds it receives; the larger the species, the greater the fund.

Approximately one-third of the United States flora is considered threatened, and plants account for over one-half of the federal listed species (Havens et al.,2014; NatureServe, 2012). In addition, plant conservation research receives approximately 4.1% of the available government funding (Negron-Ortiz, 2014). Yet, plant conservation still lacks support compared to wildlife conservation (Simon et al., 1995). As conservationists, we must accept that data on the status of species of concern is the most powerful tool to guide our effort to provide an adequate assessment for the recovery of the species. Data on species may also include biological and ecological limiting factors for the populations, threats, and possible management needs (Fish and Wildlife Service, Interior, 1983). But before we assess those issues, information on the geographic distribution of the species is essential. This knowledge gap is called the Wallace shortfall. This shortfall refers to the lack of data on the geographic distribution of species (Ladle et al., 2011). The Wallace shortfall of *Saxifraga hyperborea* is the type of knowledge gap that my study means to tighten.

Conservation Biogeography

The distribution of species is an essential puzzle piece to understanding species habitat and threats. Fields like conservation biogeography depend on this type of data. Conservation biogeography is a relatively new field in conservation. It seeks to predict the effects of humans on biodiversity and inform management decisions on behalf of the conservation of biodiversity (Serra-Diaz, 2019; Lomolino, 2004; Whittaker, 2005). Lomolino (2004) argues that this discipline has two focuses: 1) conservation of the biological diversity and 2) the preservation of the geographic, ecological, and evolutionary context. There are three main scopes within the conservation biogeography: 1) population scale, 2) landscape scale, and 3) geographical scale. The population scale evaluates population decline, viability, competition, and degradation of small populations. The landscape-scale refers to the study on habitat corridors and metapopulation. Finally, the geographical scale maps and models biogeographical patterns and diversity (Ladle et al., 2011). This research study falls within the geographical scale, modeling the distribution of a species.

Conservation organizations, government agencies, and consultants play an essential role in collecting demographic data on rare species in Washington State. The Washington Natural Heritage Program (WNHP) manages and stores this data. The WNHP role is to categorize species, ecosystems, and their conservation needs. They currently have around ~7,000 records of rare species in their database. This data is essential to allocate conservation funds and determine well-informed conservation priorities (Natural Heritage Program, 2021). Programs like WNHP are fundamental to expanding our understanding of species of concern. My study will use

WNHP data to predict the potential distribution of the Washington state rare plant species *Saxifraga hyperborea*.

Species Distribution Models

I will attempt to predict the current distribution of *Saxifraga hyperborea* by building a species distribution model (SDM). SDMs have become one of the most important tools for conservation biology, ecology, and evolution (Zurell et al., 2020; Mi et al., 2017). According to Elith and Leathwick (2009), SDMs is a tool that combines the known location of species with environmental estimates. According to Sierra-Diaz et al. (2016), we have benefited from SDMs by projecting in space and time the distribution of species. SDMs have been critical instruments to 1) identify the conservation problem, 2) define possible conservation actions, 3) predict consequences of actions, and 4) detect habitat suitability at a local and global scale (Guisan et al., 2013; Maguire, 2016). There are many applications for SDMs, according to an Araujo et al. (2019) study that examined >600 studies utilizing SDMs. The most common applications of SDM were focused on 1) predicting the effects of climate change on biodiversity, 2) selecting places for protected areas, 3) selecting areas for habitat restoration, and 4) selecting species translocation (Araujo et al.,2019). My study's focus falls within the second and fourth most common applications mentioned above. Predicting potential distribution areas could also provide information on potential areas where this plant species could be relocated or reintroduced. SDMs have various techniques; Random Forest (RF) is the machine learning algorithm applied in this project.

Random Forest

Today we can find different modeling algorithms to predict a species distribution. The most common are the Generalized Linear Model (GLM), BioClim, Regression Trees, Domain, Maxent, GARP, and Random Forest (Duan et al., 2014). I will be using the Random Forest model, which has two primary frameworks, classification or regression. Since my response variable is binary (presence/absence) I used the classification model. Random Forest works by creating decision trees based on the selected independent variables (in this case, climatic variables). Based on those variables and response data (presence and absence points), the model (once created and trained) can predict outcomes. In this case, the model identifies areas as potential presence and absence. The Random Forest's main structure is to create multiple numbers of decisions trees. Every tree function as a flow chart. Depending on the random climatic variables and decisions, the results on every tree act as a vote. When the same data goes through multiple randomized trees, the finding that repeats the most is the "winner." This type of machine learning can then predict based on the trained data. The model is trained with approximately 2/3 of the data, and the other 1/3 is used to validate your model.

The output of Random Forest classification is a confusion matrix. This matrix provides information on the accuracy and sensitivity of your model. Sensitivity explains when one area was correctly assigned as a suitable habitat. But accuracy describes both 1) when the model correctly assigned an area that was not suitable as a non-suitable habitat and 2) when one area was correctly assigned as suitable habitat (ESRI, 2019). In the final output, the model will predict potential areas of suitable habitat (Breiman, 2001; Biau et al., 2016). One item of note is that one of the most significant constraints of dealing with rare species is the small size of the data. The main driver in using Random Forest was that this model had demonstrated a good performance even with a small data size (Mi et al., 2016).

Saxifraga hyperborea

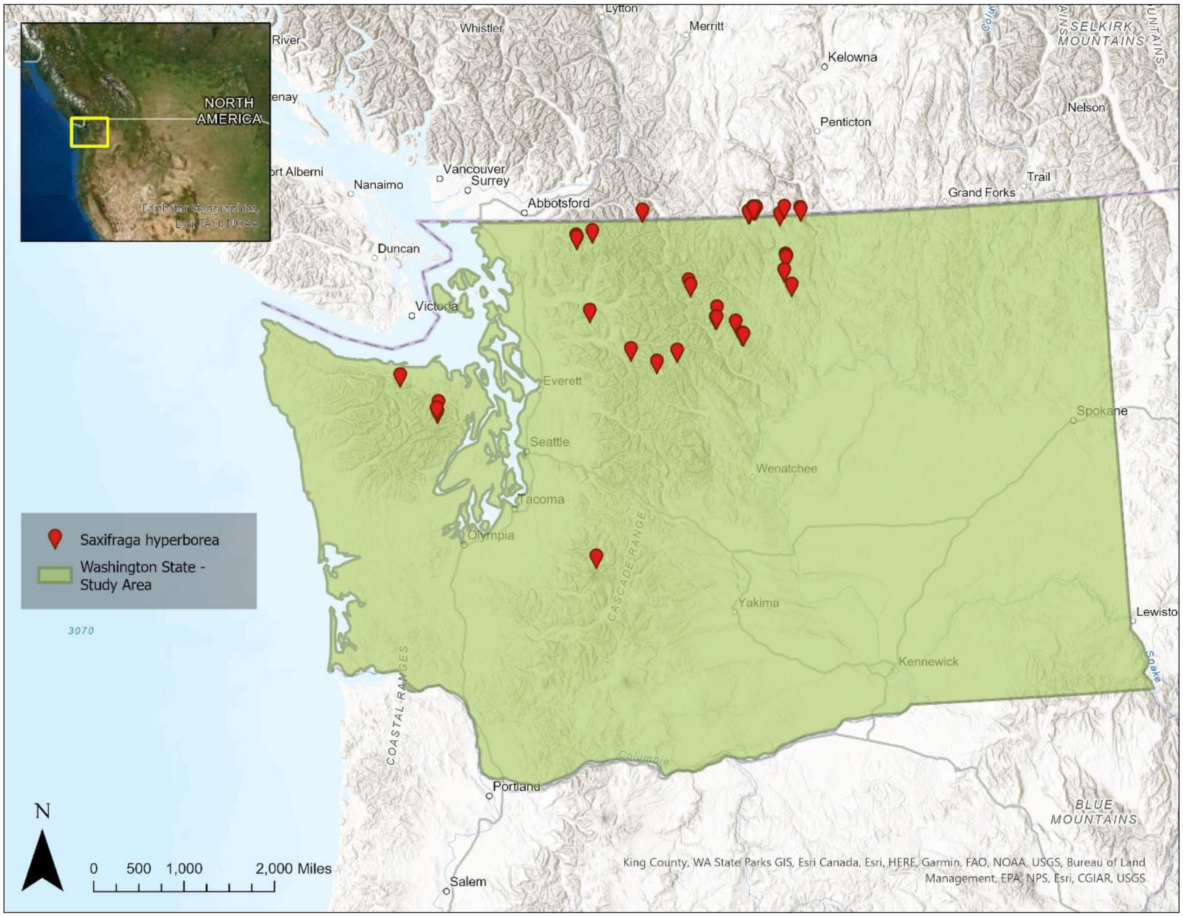

Figure 1:Saxifraga hyperborea occurrence in Washington State. Data from the Washington Natural Heritage Program.

The baseline data I used had 46 presence points for *Saxifraga hyperborea*, provided by the WNHP, which are concentrated in the North Cascades and the Olympics (Figure 1). This plant tends to grow in subalpine and alpine ecosystems; it grows at an elevation of ~5,200-8,800 feet (WTU Herbarium, 2004). The plant is found in shaded cliffs, talus, and rock crevices (Camp et al., 2011). As shown in Figure 2, *Saxifraga hyperborea* can also be found in Russia, Greenland, and Canada. In the United States, it can be found in California, Montana, Wyoming, Idaho, New Mexico, Arizona, Oregon, and Alaska (Herbaria, 2020; WTU Herbarium, 2004). *Saxifraga hyperborea* is a tufted perennial, forming patches up to 8 cm broad and 1-10 cm tall (Figure 3). Leaves are mostly basal, round to kidney shape, with 3-9 round lobes. This plant blooms from June to August; terminal flowers have 5 white to purplish petals, 5 erect sepals, 10 statements, and half inferior ovaries. The fruit is a 4-6 mm long capsule (WTU Herbarium, 2004; Camp et al., 2011; Saxifraga hyperborea-FNA, 2020). *Saxifraga hyperborea* is a rare species in Washington State (Camp et al., 2011).

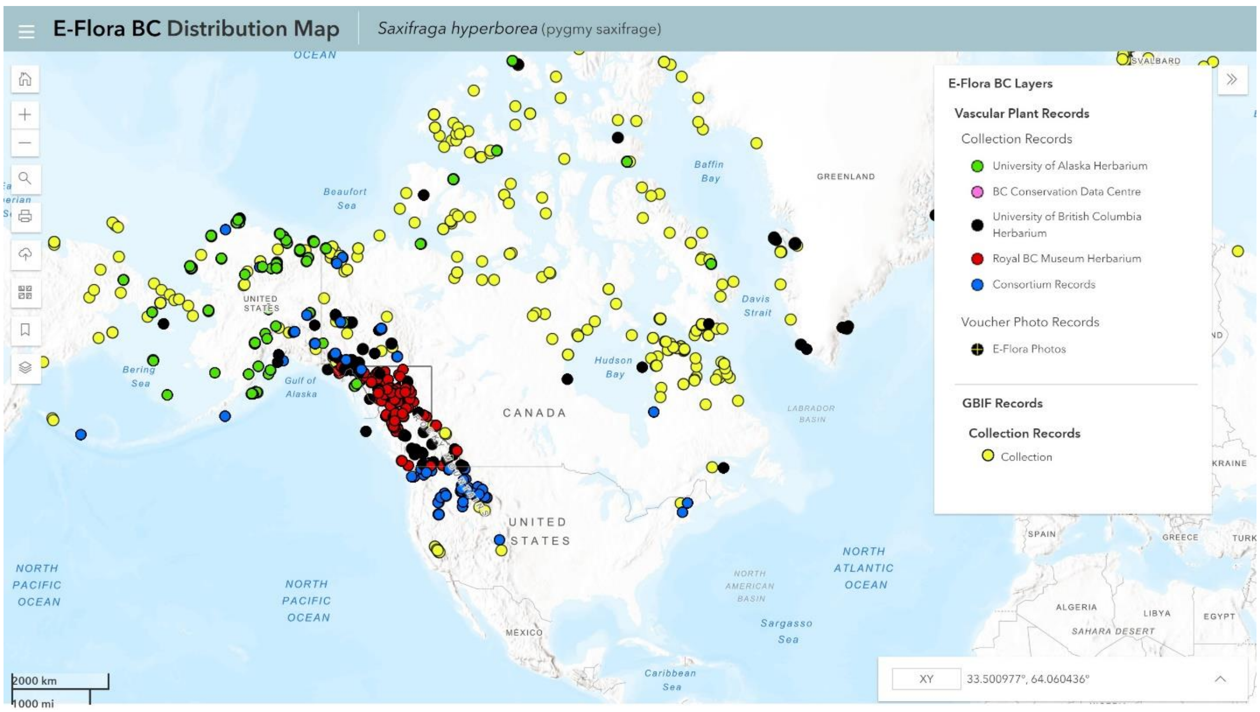

Figure 2.Distribution of Saxifraga hyperborea *based on herbarium specimens and photos. Map retrieved from E-Flora BC: Electronic Atlas of the Flora of British Columbia. (https://linnet.geog.ubc.ca/eflora_NewFullMap/index.html?sciname=Saxifraga%20hyperborea&BCStat)*

*Figure 3.*Saxifraga hyperborea*. Copyright 2020 Gary Brill.*

Conservation Status of *Saxifraga hyperborea*

Saxifraga hyperborea is classified as a G5/S3 on the Heritage Network Ranking System. The Heritage Network Ranking System is a classification system for species and ecosystems developed by Nature Serve (Nature Serve, 2012). The "G" refers to the Global conservation status, and the 5 is the level. In this case, a "G5" species is described by the Nature Serve (2012) as "Secure- At a very low risk of extinction or elimination due to a very extensive range, abundant populations or occurrences, and little to no concern from declines or threats." The "S" refers to the subnational conservation status, and the 3 is the level. The "S3" is defined by the Nature Serve (2012) as "vulnerable- moderate risk of extirpation in the jurisdiction due to a fairly restricted range, relatively few populations or occurrences, recent and widespread declines, threats, or other factors." WNHP then uses the Heritage Network Ranking System to categorize the state status of species of conservation concern in Washington State (Washington Natural Heritage Program, 2019). Species like *Saxifraga hyperborea* categorized as "G5/S3" are then classified in Washington as sensitive based on the Washington matrix (Washington Natural Heritage Program, 2019). According to the Washington Natural Heritage Program (2019), sensitive is defined as "vulnerable or declining and could become Threatened or Endangered."

However, *Saxifraga hyperborea* is not listed as a species of concern in other states and countries. For example, in British Columbia, the heritage rank is "G5/S5". The "S5" is defined by Nature Serve (2012) as "demonstrably widespread, abundant and secure." In the British Columbia list status, this plant is classified as "yellow." According to the BC Conservation Data Center (2021) "yellow" is defined as "species or ecological communities that are apparently secure and not at risk of extinction. Yellow listed species may have red- or blue-listed subspecies". In addition, the Committee on the Status of Endangered Species in Canada and the Species at Risk Act haven't ranked *Saxifraga hyperborea*.

In the same way, in the United States, the situation is a little unclear. In California, Oregon, Montana, Colorado, New Mexico, Arizona, and Alaska are not listed as species of concern. On the other hand, in Wyoming, *Saxifraga hyperborea* is listed as "G5/S3". Finally, in Idaho, the plant is listed as SNR, which Nature Serve (2012) defines as "State conservation status not yet assessed."

Again, the difference of *Saxifraga hyperborea* status between Washington and the adjacent areas is unusual. Why would *Saxifraga hyperborea* be sensitive in Washington but not listed as a species of concern across the border and states around it? In a conversation with Wendy Gibble, program manager for the Washington Rare Plant Care and Conservation program, she mentioned her curiosity about the distribution of *Saxifraga hyperborea* in Washington. Specifically, she thought it was interesting that there is not enough information on the presence of this rare plant. She also pointed out that the *Saxifraga hyperborea* growth habit complicates the search for this species. As mentioned before, this plant species tends to grow in alpine areas, particularly in rock crevices on top of mountains. From two populations I had the chance to monitor this summer, we found less than 200 individuals in an area of 5 to 3meter radius on both occasions. The population arrangement of this plant is very patchy.

I had a similar conversation with Walter Fertig, the botanist for the WNHP, where he agreed with Wendy's concern. Both experts express that *Saxifraga hyperborea* could be classified as sensitive in Washington State because there is insufficient information on the current distribution. Both conversations brought me to the conclusion that WNHP needed more information on the presence of *Saxifraga hyperborea.* This project aims to inform the potential distribution of *Saxifraga hyperborea* in Washington State.

Methods

Data Management

I obtained 46 presence location points of *Saxifraga hyperborea* from WNHP. To avoid pseudo-replication, I removed 15 presence points located less than 0.5 miles from each other at random, which resulted in a final *n* of 31 presence points. Data provided by WNHP did not include absence points. Therefore, pseudo-absence points were created in ArcGIS Pro. To make the pseudo-absence points, I built a layer with suitable habitat for *Saxifraga hyperborea* based on the known ecology (Camp et al., 2011; WTU Herbarium, 2004). The suitable habitat layer was developed by eliminating unsuitable habitat based on the land cover data layer from the National Land Cover Dataset (National Land Cover, 2009). I used the reclassify tool to remove the following classifications: “Open Water,” “Developed, Open Space,” “Developed, Low Intensity,” “Developed, Medium Intensity,” “Developed, High Intensity,” “Dwarf Scrub,” “Shrub/Scrub,” “Grassland/Herbaceous,” “Sedge/Herbaceous,” “Lichens,” “Moss,” “Pasture/Hay,” “Cultivated Crops,” “Woody Wetlands” and “Emergent Herbaceous Wetlands.” Then I created a buffer area of 138 miles (~2 degrees) around each presence point based on the recommendation of Barbet-Massin et al. (2012). To ensure random points were within the state of Washington, I cropped the buffer area with a Washington State outline polygon. I created ten pseudo-absence points per presence point within the buffer area using Create Random Points tool. All the pseudo-absence points outside the suitable area range were deleted utilizing the Clip tool. The final pseudo-absence layer had 93 pseudoabsence points and 31 presence points (Figure 4). All layers created were then projected to NAD 1983 UTM zone 10N.

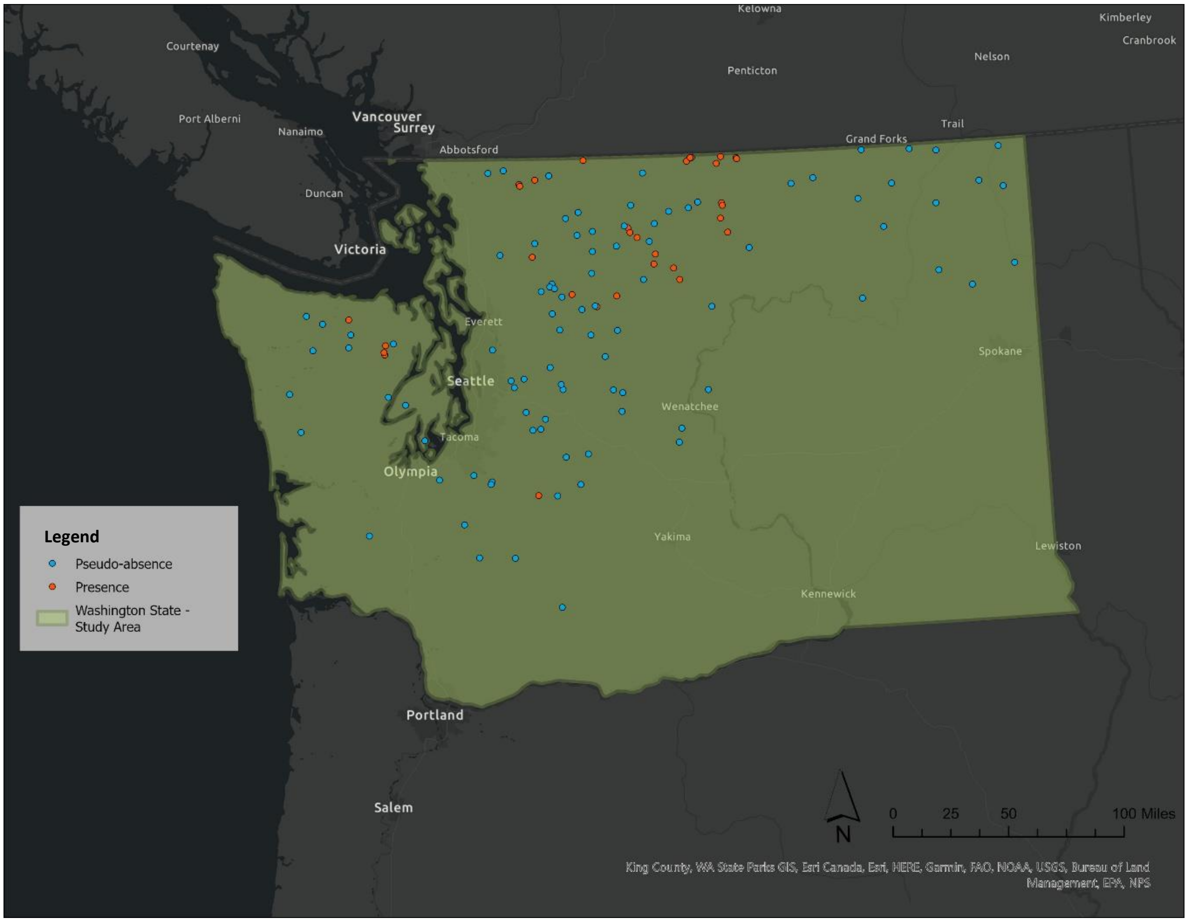

Figure 4. Final locations of layer presence and pseudo-absence points within Washington used in the study.

Environmental Variables

I selected twelve potential climatic and topographic variables based on the Wershow and DeChaine (2017) study of five endemic alpine plant species in the Olympics. These included: 1) mean annual temperature (MAP), 2) mean warmest month temperature (MWMT), 3) mean coldest month temperature (MCMT), 4) continentality (TD), 5) mean annual precipitation (MAP), 6) growing degrees days (DD5), 7) number of frost-free days (NFFD), 8) precipitation as snow (PAS) and 9) climatic moisture deficit (CMD), 10) slope, 11) aspect and 12) elevation. I tested all variables for collinearity and when any pairwise correlation was high (r >0.8, Table 3) one of the variables was taken out to avoid model inaccuracy (Wershow and Dechaine, 2017). The final six variables used in our model were: 1) continentality, 2) mean annual precipitation, 3) growing degrees day, 4) the number of frost-free days, 5) precipitation as snow, and 6) climatic moisture deficit (Figure 5). Climatic variables were obtained from Adapt West (Wang et al., 2016) from a 1991-2020 time period at 1km resolution. Elevation was obtained from the U.S. Geological Survey (U.S. Geological Survey, 2020) at 30m resolution, as a digital elevation model (DEM). Slope and Aspect layer was created from the DEM using the Slope tool and Aspect tool in ArcGIS pro.

Table 1. Results of the correlation test for the environmental variables. Values with * r >0.8.

Layer	MAT	MWMT	MCMT	TD	MAP	DD5	NFFD	PAS	CMD	Elevation	Slope	Aspect
MAT	1	0.76258	0.71201	0.04863	-0.26414	0.94423*	0.82794	-0.80467	0.51442	-0.91114*	-0.54544	0.06146
MWMT	0.76258	1	0.09532	0.68067	-0.6965	0.92102*	0.28485	-0.72662	0.90057*	-0.56493	-0.51519	0.02389
MCMT	0.71201	0.09532	1	-0.66427	0.35327	0.45206	0.9615*	-0.43297	-0.18246	-0.76899	-0.27619	0.06936
TD	0.04863	0.68067	-0.66427	1	-0.78298	0.3589	-0.4937	-0.22697	0.8105	0.14172	-0.18364	-0.03317
MAP	-0.26414	-0.6965	0.35327	-0.78298	1	-0.51445	0.20428	0.50243	-0.84869	0.10822	0.38959	0.03421
DD5	0.94423	0.92102*	0.45206	0.3589	-0.51445	1	0.61495	-0.80816	0.74151	-0.80448	-0.58011	0.04633
NFFD	0.82794*	0.28485	0.9615*	-0.4937	0.20428	0.61495	1	-0.58034	-0.02624	-0.85412	-0.35999	0.06472
PAS	-0.80467*	-0.72662	-0.43297	-0.22697	0.50243	-0.80816	-0.58034	1	-0.55104	0.74478	0.55387	-0.03569
CMD	0.51442	0.90057*	-0.18246	0.8105	-0.84869	0.74151	-0.02624	-0.55104	1	-0.31306	-0.45174	-0.01741
Elevation	-0.91114*	-0.56493	-0.76899	0.14172	0.10822	-0.80448	-0.85412	0.74478	-0.31306	1	0.49956	-0.05316
Slope	-0.54544	-0.51519	-0.27619	-0.18364	0.38959	-0.58011	-0.35999	0.55387	-0.45174	0.49956	1	-0.03787
Aspect	0.06146	0.02389	0.06936	-0.03317	0.03421	0.04633	0.06472	-0.03569	-0.01741	-0.05316	-0.03787	1

Random Forest Model

Random Forest is a supervised learning algorithm that creates decision trees (Donges, 2019). Decision trees split categorical data in this study between presence (1) and absence (0) based on the environmental variables. Multiple trees are made and merged to ensure the most accurate prediction. The prediction that repeats the most or has the most votes is the chosen one. If multiple decision trees predicted potential presence more than absence in a particular area, the model would then mark that area as presence based on the environmental variables.

The model was built in ArcGIS Pro with the Forest-based Classification tool. I 'trained' several different models with different combinations of the original environmental variables. I evaluated accuracy and sensitivity as a primary indicator for a "good" model. First, I tried the combination of the nine environmental variables that were not highly correlated ($r > 0.8$, Table 2) with each other. Accuracy for the training data was high (0.95; Table 2), but accuracy for the validation was lower (0.86). Another attempt to find the most precise model was to lower correlation coefficient ($r > 0.7$) to evaluate collinearity. From the previous 9 layers I ended up with only six layers. In this version of the model, accuracy and sensitivity stayed the same for the training data, but for the validation data, the sensitivity for presence was significantly lower (0.75) compared to the previous attempt (Table 3). Various combinations of environmental variable were trained, but the previous attempts mentioned were the second highest best scores I obtained. The final six environmental variables (TD, MAP, DD5, NFFD, PAS, CMD) selected had the best results for accuracy and sensitivity.

After selecting our final environmental variables, a total of 100 decision trees was selected for this model. I tested our model using 50 decision trees, 100 trees and 200 trees. Increasing the number of trees helps evaluate the performance of the model (ESRI, 2019). While testing for the different trees I evaluated the score in the mean square error and model accuracy. The 200-decision tree model had a low mean square error, but it also had a very low accuracy score (Table 4). Results for the 50 decision trees model had similar result as the 100 decision trees model. The 50-descision trees model overall mean square error and the absence mean square score was lower than the 100-desions trees model. But the presence mean square error was lower for the 100 decision trees model. I decided to select the 100-decision model since it is recommended to use as many trees as possible to assure a more stable result and a less susceptible model to noise in the data (ESRI, 2019). The model was trained with 90% of the data, and the remaining 10% was used to validate our model. Since our presence and absence data was not balanced (31 presence and 93 absences) I checked the Advanced Forest Option to "Compensate for Sparse Categories". This option ensures that each category was represented in each tree (ESRI, 2019).

Finally, another metric that I used to select the final model was the Out of Bag errors (OOB). The Out of Bag errors for the final model were fairly low (Table 5). These errors help us understand the accuracy of the model by indicating the percentage of incorrect classifications. The OOB error is calculated using the subset of training data that was not used in the model analysis (Bhatia, 2019). The lower the OOB error, the higher the accuracy and success of our model (ESRI, 2019). The percentage OOB error for presence (1) was 0.296, while the percentage OOB error for absence (0) was 8.634. The mean square error is the overall percentage of the incorrect OOB classifications (ESRI, 2019). I obtained a 6.569 mean square error. All the OOB errors were calculated based on 100 decision trees.

Table 2. Accuracy and sensitivity results for the first potential 9 layers after the first correlation test (r>0.8). Accuracy and Sensitivity results for this model had lower scores compared to our final model.

Prediction with 9 layers (TD, MAP, DD5, NFFD, PAS, CMD, Elevation, Slope and Aspect)		
Training Data		
Category	**Sensitivity**	**Accuracy**
0	0.94	0.95
1	1.0	0.95
Validation Data		
Category	**Sensitivity**	**Accuracy**
0	0.80	0.86
1	1.0	0.86

Table 3. Accuracy and sensitivity results for the second correlation test (r>0.7). Accuracy and Sensitivity results for this model had lower scores compared to our final model. scores compared to our final model.

Prediction with 6 layers (TD, MAP, NFFD, PAS, Slope and Aspect) after correlation test r>0.7		
Training Data		
Category	**Sensitivity**	**Accuracy**
0	0.94	0.95
1	1.0	0.95
Validation Data		
Category	**Sensitivity**	**Accuracy**
0	0.90	0.86
1	0.75	0.86

Table 4. Results for different number of trees. These results were used to select the right number of trees for our final model.

Number of Trees	50	100	200
Mean Square Error	6.550	6.569	5.532
0	8.463	8.634	7.280
1	0.741	0.296	0222
Training Data:			
Category	**Accuracy**	**Accuracy**	**Accuracy**
0	0.96	0.96	0.94
1	0.96	0.96	0.94
Validation Data:			
Category	**Accuracy**	**Accuracy**	**Accuracy**
0	0.93	0.93	0.57
1	0.93	0.93	0.57

Table 5. Model Out of Bag (OOB) mean square errors results. The mean square error is the overall percentage of the incorrect OOB classifications.

Model Out of Bag Errors	
Number of Trees	**100**
Mean Square Error	6.569
0	8.634
1	0.296

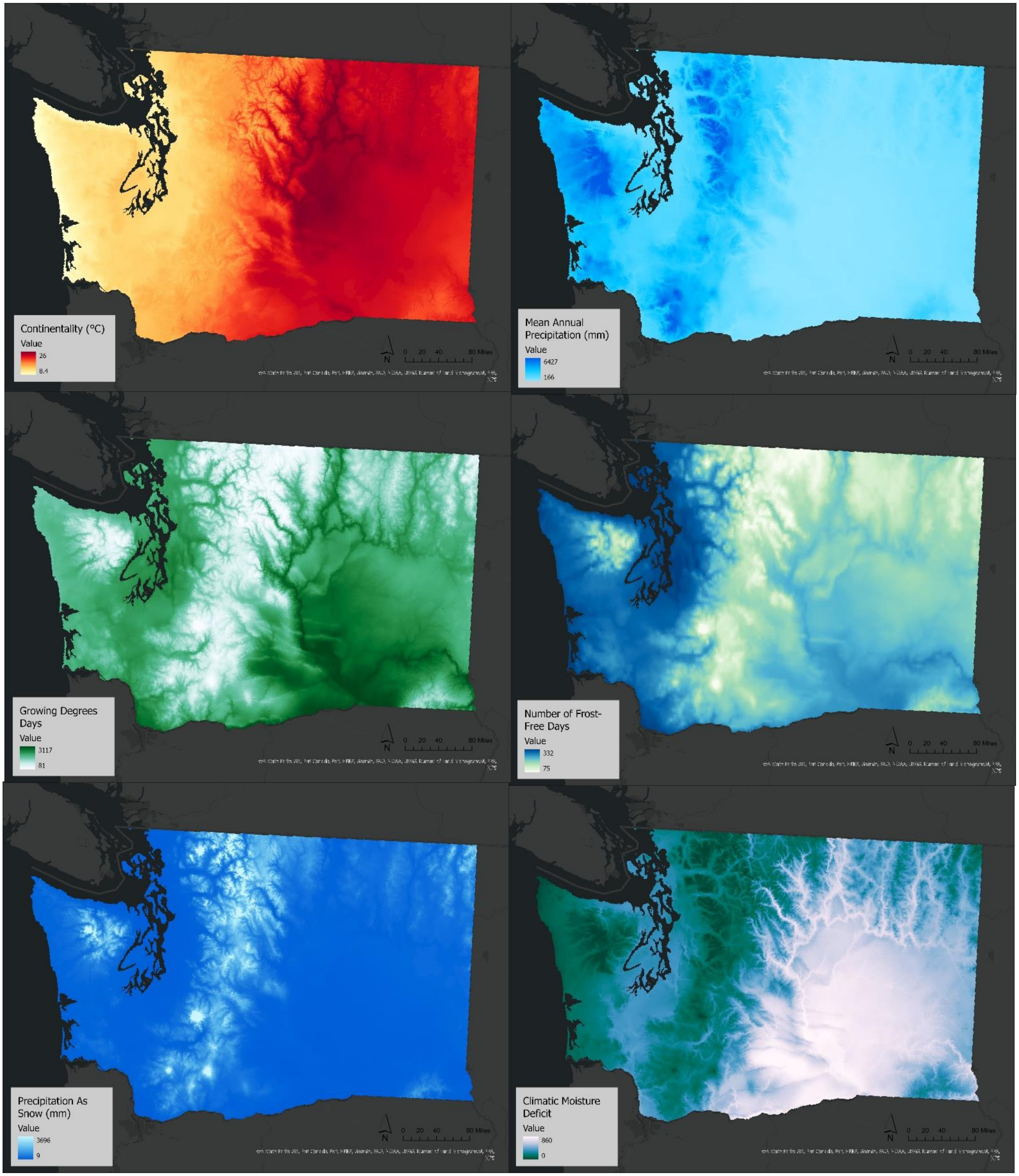

Figure 5.Climatic variables selected for the final model. Layers retrieved from AdaptWest at 1km resolution. Climatic layers are dated from 1991-2020 time period.

Results

Model

The top variable of importance from the final Random Forest model of *Saxifraga hyperborea* in Washington was precipitation as snow (Table 6). Importance is calculated using the Gini coefficient. The Gini coefficient is the sum of the number of times a variable creates a split and the impact of the split divided by the number of trees (ESRI, 2019). The importance value of precipitation as snow was 0.39, followed by the mean annual precipitation (0.32), with the rest of the variables receiving similar importance values (0.28-0.30, see Table 6).

Table 6. Final model top *variable of importance results. Importance is the sum of the Gini Coefficient. The % is the percentage of the total sum of the Gini Coefficient.*

Top Variable Importance		
Variable	**Importance**	**%**
Precipitation as Snow	0.39	21
Mean Annual Precipitation	0.32	17
Number of Frost-Free Days	0.30	16
Continentality	0.30	16
Growing Degree Days	0.29	15
Climatic Moisture Deficit	0.28	15

The second output of the model is the sensitivity and accuracy score for the training and validation data (Table 7). Higher scores (closer to 1.0) connote better prediction. In the category section, one represents presence, and zero represents absence. Sensitivity for the training and validation data in the presence category was 1.0. For the absence category on sensitivity, both the training and the validation data received scores higher than 0.90. For the accuracy results, the training data score a 0.96, and the validation data score a 0.93. Again, on both occasions, accuracy was higher than 0.90, representing a good result for the model.

Table 7. Final model sensitivity and accuracy *results. Sensitivity is the percentage a category (presence or absence) was correctly predicted. Accuracy is the percentage that takes in account when 1) a category is correctly predicted and when 2) a category was not correctly predicted.*

Training Data		
Category	Sensitivity	Accuracy
0	0.95	0.96
1	1.0	0.96
Validation Data		
Category	Sensitivity	Accuracy
0	0.90	0.93
1	1.0	0.93

The third output of the model is the explanatory variable range diagnostics. This diagnostic evaluates the training, validation and prediction values, and assesses if they are adequate to produce a precise model by looking at the overlap of the training and prediction data (ESRI, 2019). The training column is the percentage of overlap for the training values and all the climatic variables values. The validation column is the percentage of overlap for the training values and the validation data. The prediction column is the percentage of overlap for the training values and prediction data. A value of one indicates that the training data and the prediction value are equivalent. Values less than zero indicate the model intends to make predictions on data that was not trained. Values higher than one indicates that the training data range is bigger than the range used for the prediction (ESRI, 2019). Values for the training share were very close to 1.00, i.e. a good result (Table 8). This comparably occurs in the validation share, where values are higher than 0.75. On the other hand, for the prediction shares all values are over 1.00. This means that the training value range was bigger than the prediction, and the model prediction is attempting to extrapolate.

Table 8. Explanatory range diagnostic results. Training column is the percentage of overlap between the ranges of the training data and the input explanatory variable. The validation column is the percentage of overlap between ranges of the validation data and the training data. The prediction column is the percentage of overlap between the ranges of the training data and the prediction data.

Variable	Share		
	Training	Validation	Prediction
Continentality	1.00	0.73	1.54
Mean Annual Precipitation	1.00	0.95	1.51
Growing-degree days	0.98	0.95	1.52
Number of Frost-Free Days	1.00	0.89	1.32
Precipitation as Snow	0.84	1.00	2.00
Climatic Moisture Deficit	1.00	0.76	1.26

Prediction

The final output of the model is the prediction map. As shown in Figure 6, in red you can see the predicted potential habitat for *Saxifraga hyperborea*. Most of the predicted suitable areas are concentrated in the North Cascades and the Northeast corner of the Olympics. In addition, there is some dispersed potential habitat in the Okanogan and Canadian Rockies. Finally, in the south of Washington, the three main volcanos Mt. Rainier, Mt. St. Helens, and Mt. Adams, also appear to be a potential habitat for *Saxifraga hyperborea*.

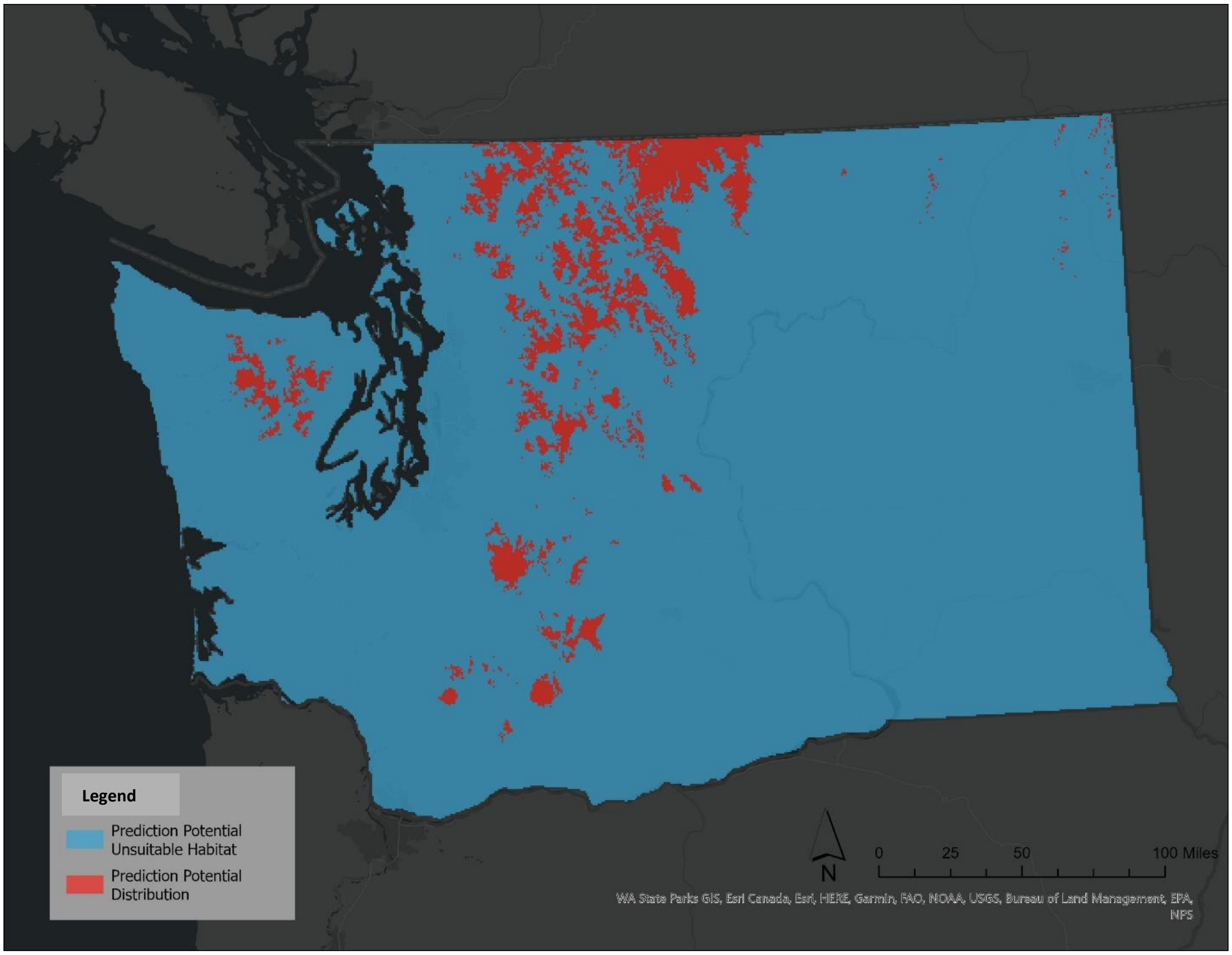

Figure 6.Areas in red are the model prediction for potential distribution of Saxifraga hyperborea.

Discussion

The Random Forest-based output in ArcGIS predicted areas in Washington State of suitable habitat for *Saxifraga hyperborea*. The model accuracy and sensitivity was high (Table 9) and the mean squared error was low (Table 6). Even with a small sample of known presences in Washington, I am confident with the model performance based on the results. As shown in Figure 6, most of the prediction for suitable habitat occurred in the North Cascades, followed by the Olympic peninsula and the Southern Cascades. Areas without previous records were predicted as suitable by the model in the Okanagan and the Canadian Rockies (i.e., in Northeastern Washington). The model also identified snow precipitation as the main climatic variable driver for predicting *Saxifraga hyperborea* habitat. These results may serve as a preliminary approach to better reflect the potential discovery of new populations of *Saxifraga hyperborea* in Washington State. As discussed below, the model has areas where it can be improved.

Okanagan and the Canadian Rockies

A very interesting aspect of the model results was the prediction of potential habitat of *Saxifraga hyperborea* in the Okanagan and Canadian Rockies. As shown in Figure 7 all the different shades of green represent the Selkirk Mountain which is part of the Canadian Rockies. Most of the Selkirk Mountain in Washington State sit on the Colville National Forest. This area has been described as some of the wildest country in Washington State (N.S., 2022) and is the home of the mountain caribou (*Rangifer tarandus caribou*), a federal endangered and state endangered species (W.D.F.W., 2022). From the botanical perspective this ecoregion also holds the Halliday Fen Research Natural Area. In this research area, one can find thirteen rare plant species, a wetland community, and three terrestrial plant communities (P.N.I.N.A.N., 2013). This corner of Washington State has been protected because it has an important ecological value.

There is no *Saxifraga hyperborea* population known in this ecoregion in Washington State. Most of the known populations are in the North Cascades and the Northeast corner of the Olympics (Figure 1). However, the model predicted potential habitat in this region are very close to herbarium specimens of *Saxifraga hyperborea* from Idaho (Figure 7). These herbarium specimen data was obtained from the Consortium of Pacific Northwest Herbaria. *Saxifraga hyperborea* has therefore been identified in the same ecoregion (the Selkirk Mountains) as the model predictions as recently as 2 years ago (In Figure 7 the herbarium specimen labeled as 1 was collected in 1986, specimen number 2 in 1982, specimen number 3 in 2019, specimen number 4 in 2019, specimen number 5 in 2020, specimen number 6 in 2004 and specimen number 7 in 1989). The model is predicting potential habitat in Washington, very close to where there are known populations in Idaho, lending confidence to the prediction. This area has been protected for its natural value and provides a sense of urgency for potential exploration for the search of *Saxifraga hyperborea*.

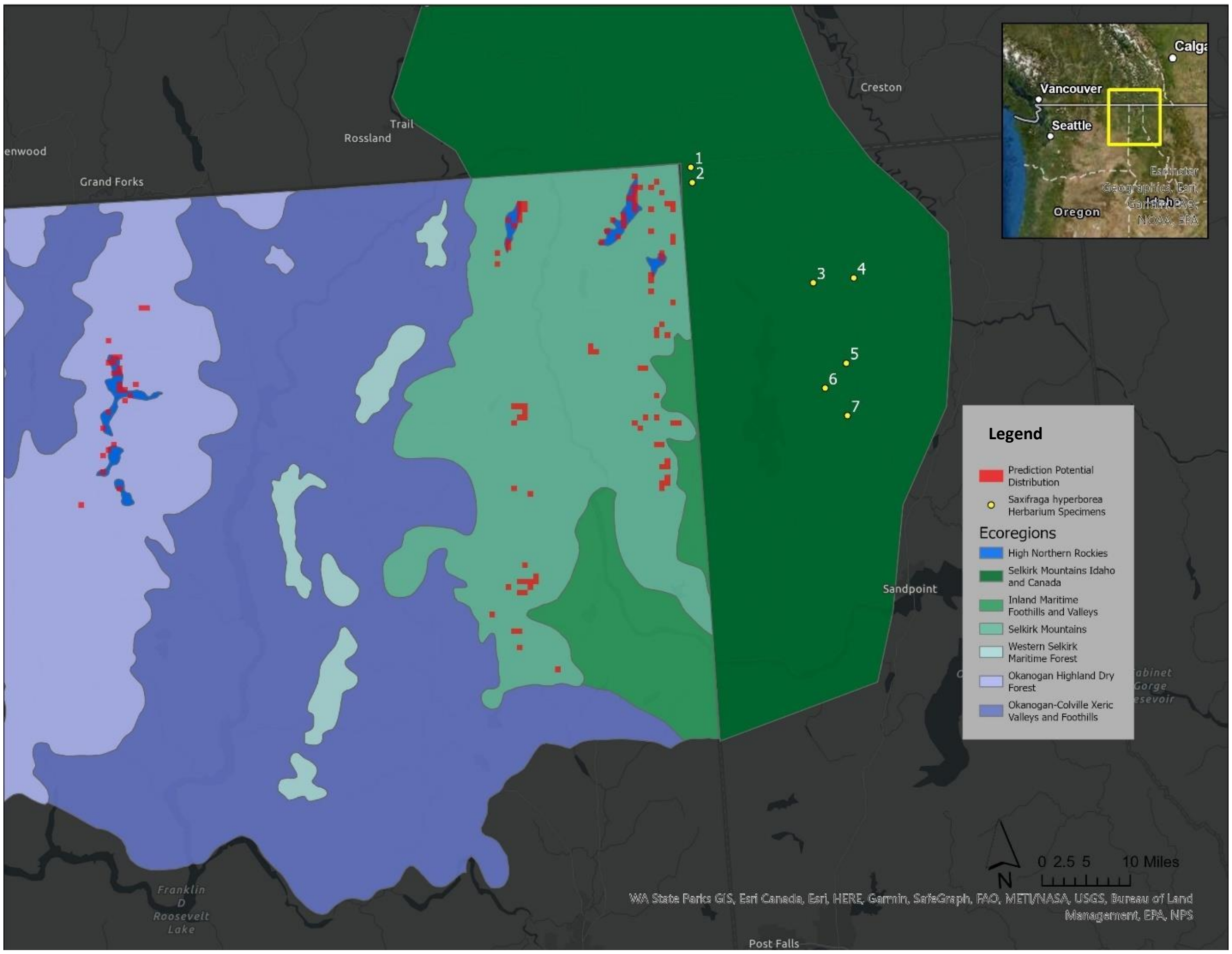

Figure 7. Image zoomed in to Northeastern Washington. Areas in red represent the potential distribution of Saxifraga hyperborea*. Yellow dots are herbarium specimens of* Saxifraga hyperborea *in Idaho State. The specimen labeled as 1 was collected in 1986, specimen 2 in 1982, specimen 3 in 2019, specimen 4 in 2019, specimen 5 in 2020, specimen 6 in 2004 and specimen 7 in 1989.*

Resolution of Climate layers

Building an SDM requires two types of data layers, 1) presence data of the species and 2) environmental data. This data is then transformed into a grid. The size of the cells in the grid determines the resolution of the data. Resolution can also be referred or called grain size. It might appear that the smaller the grain size, the better the model accuracy, since a small grain size can be able to represent more accurately the local environment and conditions (Manzoor et al., 2017). Nevertheless, some studies (Guisan et al., 2007; Manzoor et al., 2017) have found that a fine grain size data does not always ensure a more accurate prediction.

The climatic data used in this study had a resolution of 1 km (~0.6 mi). This data was obtained from Adapt West (Wang et al., 2016). According to the monitoring site description data from the WNHP, *Saxifraga hyperborea* populations found in Washington State do not typically exceed the 6 square meter size. Surveyors often describe the population growth habit as “scattered-patchy” or with a “patchy distribution”. The large grain climatic data in this study may have had an impact on the accuracy of the model I used. In the presence data, each 1 km cell that was labeled as presence was overestimating the real species habitat extent. *Saxifraga hyperborea* is also a habitat specialist as it grows in moist cool rock crevasses, completely shaded areas, and with little to no soil (WTU Herbarium, 2004; Camp et al., 2011). According to Manzoor et al (2017) using a small grain data, for habitat specialist species, will increase the distinction between different climatic units, proving a more precise species-habitat relationship. Finally, Lauzeral et al. (2013) suggested that plant distribution should be modeled using a small grain size, this is because they are restricted to the location they exist. Species should be modeled at a spatial grain size based on the environment they depend on (Lauzeral et al., 2013).

On the other hand, a too small of a grain size data made also generate limitations in the model. Species occurrence data at a small grain size needs to be highly accurate (Manzoor et al., 2017). The presence data used in this study have population locations that were documented in 1912. Unfortunately, these points are not accurate enough to be certain of the actual location where the population was found. Additionally, this study area happened to be the entire Washington State, which is considered a large geographical extent for these types of studies. Climatic data at a small grain size at this large extent is not available (Connor et al., 2017). Commonly used open-source spatial data sites such as AdaptWest, PRIMS, Climate Toolbox, and WorldClim offer, as their smallest grain size, environmental data at 1 km resolution. Finally, some studies (Hanberry, 2013; Lauzeral et al., 2013) suggested the use of large grain size climatic data for large geographical extents to avoid failed predictions at the environmental habitat edges.

There is no rule of thumb to determine the ideal grain size for environmental data when building an SDM. For future studies the following should be considered: a) the quality of the data, b) the ecology of the species and a) the extent of the study area (Lauzeral et al., 2013). For this study a smaller climatic grain size data, that its more representative of the species habit growth, and a smaller geographical extent may provide a more accurate model outcome.

Pseudo-absence implications

Another potential limitation of this study could have been the method to generate pseudo-absence and the ratio of pseudo-absence and presence points. There has been a debate whether pseudo-absence ratio could affect model accuracy (Lauzeral et al., 2013). Pseudo-absence is described by Phillips et al., (2009) as a point that provides "information on the available environment in the study region." According to Stokland et al., (2011) models that used pseudo-absence points, rather than presence only models, tend to perform better because it increases species occurrence and environmental factor relationships. In the final model I used 31 presence points and 93 pseudo absence points. I decided to use an unbalanced data set, even though, Barbet-Massin (2012) study suggested that the same number of presence and absence points will generate a more reliable model. However, Lobo et al. (2011) recommended a highly unbalanced design when working with poorly surveyed species to increase the sample size. Additionally, Valverde et al. (2009) suggested that unbalanced data enables restriction in the model predictions when dealing with species that occupy a small fraction of the study area.

I generated pseudo-absence points based on the Barbet-Massin et al. (2012) study "2 degrees far method". This method advised that pseudo-absence should be selected at random in a radius of 2 degrees (~138 mi) from each presence point. This was also recommended specifically for the Random Forest algorithm. For the selection of pseudo-absence points I used the random sampling with background extend limitations method (Iturbe et al., 2015).   This method involves a two-step process: 1) create a buffer around the presence point of ~0.5 mile (avoid pseudo-absence close to presence points) and 2) removing unsuitable habitat (refer to Methods, page 14). Pseudo absence points were then selected at random within the 2-degree radius, outside of the 0.5-mile buffer and within the suitable habitat. The suitable habitat included the following classifications "Ice and Snow", "Barren Land", "Deciduous Forest", "Evergreen Forest" and "Mixed Forest" (National Land Cover, 2009).

One of the possible limitations of this study could have been that I did not constrain the random selection of pseudo-absences to the high elevations or more-specific land cover types where *Saxifraga hyperborea* are typically found. I made this choice as a "let the chips fall where they may" (i.e. more relaxed) approach with the model. Most of the known locations in Washington State of Saxifraga hyperborea are in open alpine rocky slopes. Mixed forest, deciduous forest and evergreen forest does not align with the habitat requirements of *Saxifraga hyperborea,* particularly because some of these forests grow at a lower elevation that I have found *Saxifraga hyperborea*. Assuming that these areas were potentially suitable habitat might have increased the overestimation in the model prediction.

For future studies I suggest a more highly unbalanced presence absence ratio. I could potentially allow 5 times more pseudo-absence points (instead of 3), this will make available a bigger sample size. One could also constrain more the suitable habitat. I recommend removing areas that were below ~1,500 m (~5,000 ft) of elevation. The addition of elevation would create a more limited, but potentially more realistic, suitable habitat of *Saxifraga hyperborea*.

Variables of importance

Alpine ecosystem is characteristically recognized by its weather, topography, and isolation. These conditions have created a highly biodiverse and a high endemism zone (Verrall et al., 2020). This area is described as the zone above the tree line (Verrall et al., 2020). In this region sun, wind and snow can change the weather very quickly. The cold weather and short growing season make it challenging for life to thrive in this environment. Growing seasons are described as days with temperatures above 5 Celsius (41 Fahrenheit) (Ghaberr et al., 2010). It is not a coincidence that precipitation as snow was the model's top variable of importance (Table 6) for the determination of suitable habitat for *Saxifraga hyperborea*.

The top variable of importance is calculated by using the Gini Coefficient (refer to page 20). Table 6 shows the top variables of importance in the model, precipitation as snow got the highest rating, with a 21% of importance. As shown in Table 1, precipitation as snow, and elevation were (not surprisingly) highly correlated (r=0.75). *Saxifraga hyperborea* is found mainly at high elevations. This may be one of the explanations of why precipitation as snow is the top variable of importance. On the other hand, Korner (2021) argues that elevation is not necessarily a predictor for plant distribution in the alpine zone. He emphasizes that snow distribution determines plant distribution but, the distribution of snow is determined by wind and relief. For this reason, I can agree that snow, wind, and relief are the main drivers of plant distribution in the alpine ecosystems. Finally, snow also determines the reproductive behavior (Korner, 2021). Some of these plants, like *Saxifraga hyperborea,* wait until snow melts to start flowering.

Snow plays a very important role in the survival of alpine plants. As climate change continues its destructive course, the alpine ecosystem's climate warming is the biggest threat. The snowmelt and snowpack will be altered by the change in precipitation and temperature (Inouye, 2019). Snow also provides protection from extreme temperature, winter desiccation, ice blast and solar radiation (Korner, 2021). The result of precipitation as snow as the top variable of importance was a reassurance that the model was detecting an important climatic variable for *Saxifraga hyperborea* and other alpine species.

Conclusion

The main goal of this study was to detect potential habitat for the sensitive alpine species, *Saxifraga hyperborea*. The model predicted promising areas in the Okanagan and Canadian Rockies, where the species has not been found in Washington State. The high accuracy, low mean squared error and recent collection of herbarium specimens near the model predictions gave confidence to the model. There are multiple recommendations for the improvement of the model. The first improvement may be the selection of a bigger sample size by increasing the pseudo-absences (choosing 155 instead of 93). The second improvement would be to constrain pseudo-absence selection on areas above 1,600 m (5,200 ft) to simulate a more realistic habitat for *Saxifraga hyperborea*. The third improvement may be to find a higher resolution climatic layer that is more representative of the species habitat growth. This study is the beginning of a potential exploration for *Saxifraga hyperborea* in Washington State in areas where it has not been recorded. This is the first approach to potentially start resolving the enigma on the species sensitive state status. Finally, in the case that the species is reconfirmed as sensitive, the prediction map could serve as candidate areas for the reintroduction of *Saxifraga hyperborea.*

Bibliography

Ando, A. (1999). Waiting to Be Protected Under the Endangered Species Act: The Political Economy of Regulatory Delay. *The Journal of Law and Economics*, *42*(1), 29–60. https://doi.org/10.1086/467417

Andrade, A. F. A., Velazco, S. J. E., & de Marco Júnior, P. (2020). ENMTML: An R package for a straightforward construction of complex ecological niche models. *Environmental Modelling & Software*, *125*, 2–35. https://doi.org/10.1016/j.envsoft.2019.104615

Araújo, M., Anderson, R., Márcia Barbosa, A., Beale, C., Dormann, C., Early, R., Garcia, R., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R., Zimmermann, N., & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, *5*(1), 1–10. https://doi.org/10.1126/sciadv.aat4858

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, *3*(2), 327–338. https://doi.org/10.1111/j.2041-210x.2011.00172.x

Biau, G., & Scornet, E. (2016). A Random Forest guided tour. *TEST*, *25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7

B.C. Conservation Data Center (2021). BC Species and Ecosystems Explorer. B.C. Min. of Environ. Victoria, B.C. Available: https://a100.gov.bc.ca/pub/eswp/ (accessed Nov 6, 2021).

Bhatia, N. (2019, June 26). *What is Out of Bag (OOB) score in Random Forest?* Towards Data Science. Retrieved November 6, 2021, from https://towardsdatascience.com/what-is-out-of-bag-oob-score-in-random-forest-a7fa23d710

Brandt, P., Abson, D. J., DellaSala, D. A., Feller, R., & von Wehrden, H. (2014). Multifunctionality and biodiversity: Ecosystem services in temperate rainforests of the Pacific Northwest, USA. *Biological Conservation*, *169*, 362–371. https://doi.org/10.1016/j.biocon.2013.12.003

Breiman, L. (2001). Random Forest. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Camp, P., & Gamon, J. G. (2011). *Field Guide to the Rare Plants of Washington* (Illustrated ed.). University of Washington Press.

Charles, K. M., & Stehlik, I. (2020). Assisted species migration and hybridization to conserve cold-adapted plants under climate change. *Conservation Biology*, *35*(2), 559–566. https://doi.org/10.1111/cobi.13583

Clark, J., Gelfand, A., Woodall, C., & Zhu, K. (2014). More than the sum of the parts: forest climate response from joint species distribution models. *Ecological Applications*, *24*(5), 990–999. https://doi.org/10.1890/13-1015.1

Connor, T., Hull, V., Viña, A., Shortridge, A., Tang, Y., Zhang, J., Wang, F., & Liu, J. (2017). Effects of grain size and niche breadth on species distribution modeling. *Ecography*, *41*(8), 1270–1282. https://doi.org/10.1111/ecog.03416

Donges, N. (2021, September 17). *A Complete Guide to the Random Forest Algorithm*. Built In. Retrieved November 6, 2021, from https://builtin.com/data-science/random-forest-algorithm

Duan, R. Y., Kong, X. Q., Huang, M. Y., Fan, W. Y., & Wang, Z. G. (2014). The Predictive Performance and Stability of Six Species Distribution Models. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0112764

ESRI. (2019). *Forest-based Classification and Regression (Spatial Statistics)*. Spatial Statistic Toolbox. Retrieved November 6, 2021, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/forestbasedclassificationregression.htm

Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Farnsworth, E. J., & Ogurcak, D. E. (2006). Biogeography And Decline of Rare Plants in New England: Historical Evidence And Contemporary Monitoring. *Ecological Applications*, *16*(4), 1327–1337. https://doi.org/10.1890/1051-0761

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fish and Wildlife Service, Interior. (1983, September). *Endangered and Threatened Species Listing and Recovery Priority Guidelines* (No. 48). https://www.fws.gov/endangered/esa-library/pdf/1983_LPN_Policy_FR_pub.pdf

Forest, F., Grenyer, R., Rouget, M., Davies, T. J., Cowling, R., Faith, D., Balmford, A., Manning, J., Procheş, Ş., van der Bank, M., Reeves, G., Hedderson, T., & Savolainen, V. (2007). Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature*, *445*(7129), 757–760. https://doi.org/10.1038/nature05587

General Land Use Final Dataset. (2018). Washington Geospatial Open Data Portal. https://geo.wa.gov/datasets/a0ddbd4e0e2141b3841a6a42ff5aff46_0

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., . . . Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424–1435. https://doi.org/10.1111/ele.12189

Guisan, A., Graham, C. H., Elith, J., & Huettmann, F. (2007). Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, *13*(3), 332–340. https://doi.org/10.1111/j.1472-4642.2007.00342.x

Havens, K., Kramer, A. T., & Guerrant, E. O. (2014). Getting Plant Conservation Right (or Not): The Case of the United States. *International Journal of Plant Sciences*, *175*(1), 3–10. https://doi.org/10.1086/674103

Havens, K., Vitt, P., & Masi, S. (2012). Citizen science on a local scale: the Plants of Concern program. *Frontiers in Ecology and the Environment*, *10*(6), 321–323. https://doi.org/10.1890/110258

Herbaria, C. O. P. N. (2020). *Consortium of Pacific Northwest Herbaria*. Copyright (c) 2007–2018 Consortium of Pacific Northwest Herbaria. https://www.pnwherbaria.org/index.php

Humphreys, A. M., Govaerts, R., Ficinski, S. Z., Nic Lughadha, E., & Vorontsova, M. S. (2019). Global dataset shows geography and life form predict modern plant extinction and rediscovery. *Nature Ecology & Evolution*, *3*(7), 1043–1047. https://doi.org/10.1038/s41559-019-0906-2

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., . . . Wang, Y. P. (2011). Harmonization of land-use scenarios for the period 1500–2100: 600

years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, *109*(1–2), 117–161. https://doi.org/10.1007/s10584-011-0153-2

Idaho Fish and Game. (n.d.). *Pygmy Saxifrage (*Saxifraga hyperborea*) | Idaho Fish and Game*. Retrieved December 11, 2021, from https://idfg.idaho.gov/species/taxa/56222

Inouye, D. W. (2019). Effects of climate change on alpine plants and their pollinators. *Annals of the New York Academy of Sciences*, *1469*(1), 26–37. https://doi.org/10.1111/nyas.14104

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., & Gutiérrez, J. M. (2015). A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling*, *312*, 166–174. https://doi.org/10.1016/j.ecolmodel.2015.05.018

Klingkenberg, B., (2020) *E-Flora BC: Electronic Atlas of the Plants of British Columbia [eflora.bc.ca]. Lab for Advanced Spatial Analysis, Department of Geography, University of British Columbia, Vancouver.* Retrieved November 6, 2021, from https://linnet.geog.ubc.ca/eflora_NewFullMap/index.html?sciname=Saxifraga%20hyperborea&BCStatus=yellow&synonyms=%27Saxifraga%20rivularis%27&commonname=pygmy%20saxifrage&PhotoID=66468&mapservice=Vascular

Körner, C. (2021). *Alpine Plant Life: Functional Plant Ecology of High Mountain Ecosystems* (3rd ed.). Springer.

Ladle, R. J., & Whittaker, R. J. (2011). *Conservation Biogeography* (1st ed.). Wiley-Blackwell.

Lauzeral, C., Grenouillet, G., & Brosse, S. (2013). Spatial range shape drives the grain size effects in species distribution models. *Ecography*, *36*(7), 778–787. https://doi.org/10.1111/j.1600-0587.2013.07696.x

Lehtomäki, J., Kusumoto, B., Shiono, T., Tanaka, T., Kubota, Y., & Moilanen, A. (2018). Spatial conservation prioritization for the East Asian islands: A balanced representation of multitaxon biogeography in a protected area network. *Diversity and Distributions*, 414–429. https://doi.org/10.1111/ddi.12869

Lobo, J. M., & Tognelli, M. F. (2011). Exploring the effects of quantity and location of pseudo-absences and sampling biases on the performance of distribution models with limited point occurrence data. *Journal for Nature Conservation*, *19*(1), 1–7. https://doi.org/10.1016/j.jnc.2010.03.002

Lomolino, M. V., & Heaney, L. R. (2004). *Frontiers of Biogeography* (1st ed.). Sinauer Associates.

Lorini, M. L., Paese, A., & Uezu, A. (2011). GIS and Spatial Analysis Meet Conservation: a Promising Synergy to Address Biodiversity Issues. *Natureza & Conservação*, *9*(2), 129–144. https://doi.org/10.4322/natcon.2011.019

Maguire, K. C., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C., Williams, J. W., Ferrier, S., & Lorenz, D. J. (2016). Correction to 'Controlled comparison of species- and community-level models across novel climates and communities.' *Proceedings of the Royal Society B: Biological Sciences*, *283*(1837), 20161705. https://doi.org/10.1098/rspb.2016.1705

Mahoney, J. (2009). What Determines the Level of Funding for an Endangered Species? *Major Themes in Economic*, *11*(4), 1–17. https://scholarworks.uni.edu/mtie/vol11/iss1/4

Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size – finer may not always be better. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-25437-1

Mi, C., Huettmann, F., Guo, Y., Han, X., & Wen, L. (2017). Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ*, *5*, 1–22. https://doi.org/10.7717/peerj.2849

Msuya, J., Azoulay, A., Graziano Da Silva, J., Steiner, A., & Pasca Palmer, C. (2019, May 6). *UN Report: Nature's Dangerous Decline "Unprecedented"; Species Extinction Rates "Accelerating."* United Nations Sustainable Development. https://www.un.org/sustainabledevelopment/blog/2019/05/nature-decline-unprecedented-report/

Multi-Resolution Land Characteristics. (2019). *Missouri Botanical Garden*. National Land Cover Database. https://www.missouribotanicalgarden.org/

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858. https://doi.org/10.1038/35002501

National Land Cover Database. (2019). *Data*. The National Land Cover Database. https://www.mrlc.gov/data

Natural Heritage Program. (2021). Washington State Department of Natural Resources. https://www.dnr.wa.gov/natural-heritage-program

Nature Serve. (2012, June). *NatureServe Conservation Status Assessments: Methodology for Assigning Ranks*. https://www.natureserve.org/sites/default/files/natureserveconservationstatusmethodology_jun12.pdf

Negrón-Ortiz, V. (2014). Pattern of expenditures for plant conservation under the Endangered Species Act. *Biological Conservation*, *171*, 36–43. https://doi.org/10.1016/j.biocon.2014.01.018

N.S. (2022). *Canadian Rocky Mountains ecoregion // LandScope America*. LandScope America. Retrieved January 24, 2022, from http://www.landscope.org/washington/natural_geography/ecoregions/canadian_rockies/

Oldfather, M. (2018). *Population and Community Dynamics of Alpine Plants in a Changing Climate Across Topographically Heterogeneous Landscapes*. Ph.D. Dissertation, University of California, Berkeley. https://escholarship.org/uc/item/0vq6v6vs

Ovaskainen, O. (2020). *Joint Species Distribution Modelling (With Applications in R)* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781108591720

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, *19*(1), 181–197. https://doi.org/10.1890/07-2153.1

Primack, R. B. (2014). *Essentials of Conservation Biology* (6th ed.). Sinauer Associates is an imprint of Oxford University Press.

P.N.I.N.A.N. (2013). *Pacific Northwest Interagency Natural Areas Network- Halliday Fen RNA*. Pacific Northwest Interagency Natural Areas Network. Retrieved January 24, 2022, from http://www.fsl.orst.edu/rna/sites/Halliday_Fen.html

Rare Care. (2019). *Washington Rare Plant Care and Conservation | University of Washington Botanic Gardens*. Rare Plant Monitoring. https://botanicgardens.uw.edu/science-conservation/rarecare/

Raven, P. H., & Wilson, E. O. (1992). A Fifty-Year Plan for Biodiversity Surveys. *Science*, *258*(5085), 1099–1100. https://doi.org/10.1126/science.258.5085.1099

Rumpf, S. B., Hülber, K., Klonner, G., Moser, D., Schütz, M., Wessely, J., Willner, W., Zimmermann, N. E., & Dullinger, S. (2018). Range dynamics of mountain plants decrease with elevation. *Proceedings of the National Academy of Sciences*, *115*(8), 1848–1853. https://doi.org/10.1073/pnas.1713936115

Saxifraga hyperborea - *FNA*. (2020, November). Flora of North America. Retrieved November 6, 2021, from http://beta.floranorthamerica.org/Saxifraga_hyperborea

Serra-Diaz, J. M., & Franklin, J. (2019). What's hot in conservation biogeography in a changing climate? Going beyond species range dynamics. *Diversity and Distributions*, *25*(4), 492–498. https://doi.org/10.1111/ddi.12917

Simon, B. M., Leff, C. S., & Doerksen, H. (1995). Allocating Scarce Resources for Endangered Species Recovery. *Journal of Policy Analysis and Management*, *14*(3), 415. https://doi.org/10.2307/3325033

Smyth, R. (0000). *Habitat Suitability Modeling | NatureServe*. Nature Serve. https://www.natureserve.org/conservation-tools/habitat-suitability-modeling

Soil Databases | NRCS Soils. (2020). USDA. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/tools/?cid=nrcseprd1407024

Sosa, V., & De-Nova, J. A. (2012). Linajes de angiospermas endémicas en México: zonas de alto endemismo para la conservación. *Acta Botanica Mexicana*, *100*, 293. https://doi.org/10.21829/abm100.2012.38

Stokland, J. N., Halvorsen, R., & Støa, B. (2011). Species distribution modelling—Effect of design and sample size of pseudo-absence observations. *Ecological Modelling*, *222*(11), 1800–1809. https://doi.org/10.1016/j.ecolmodel.2011.02.025

Tikhonov, G., Opedal, Ø. H., Abrego, N., Lehikoinen, A., Jonge, M. M. J., Oksanen, J., & Ovaskainen, O. (2020). Joint species distribution modelling with the r-package Hmsc. *Methods in Ecology and Evolution*, *11*(3), 442–447. https://doi.org/10.1111/2041-210x.13345

U.S. Geological Survey. (2020). The National Elevation Dataset. *Fact Sheet*, *148*(99). https://doi.org/10.3133/fs14899

Velazco, S. J. E., Villalobos, F., Galvão, F., & de Marco Júnior, P. (2019). A dark scenario for Cerrado plant species: Effects of future climate, land use and protected areas ineffectiveness. *Diversity and Distributions*, *25*(4), 660–673. https://doi.org/10.1111/ddi.12886

Verrall, B., & Pickering, C. M. (2020). Alpine vegetation in the context of climate change: A global review of past research and future directions. *Science of The Total Environment*, *748*, 141344. https://doi.org/10.1016/j.scitotenv.2020.141344

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally Downscaled and Spatially Customizable Climate Data for Historical and Future Periods for North America. *PLOS ONE*, *11*(6). https://doi.org/10.1371/journal.pone.0156720

Wang, X., Sun, Z., & Zhou, A. G. (2014). Alpine Cold Vegetation Response to Climate Change in the Western Nyainqentanglha Range in 1972–2009. *The Scientific World Journal*, *2014*, 1–9. https://doi.org/10.1155/2014/514736

Washington Natural Heritage Program. (2019, July). *2019 Washington Vascular Plant Species of Special Concern List* (No. 2019–04). Washington Department of Natural Resources. https://www.dnr.wa.gov/publications/amp_nh_vascular_ets.pdf?1z6b4u

Washington Rare Plant Care and Conservation | University of Washington Botanic Gardens. (2020). Botanic Gardens UW. https://botanicgardens.uw.edu/science-conservation/rarecare/

Wershow, S. T., & DeChaine, E. G. (2018). Retreat to refugia: Severe habitat contraction projected for endemic alpine plants of the Olympic Peninsula. *American Journal of Botany*, *105*(4), 760–778. https://doi.org/10.1002/ajb2.1042

Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E. M., & Willis, K. J. (2005). Conservation Biogeography: assessment and prospect. *Diversity and Distributions*, *11*(1), 3–23. https://doi.org/10.1111/j.1366-9516.2005.00143.x

Wilson, E. O. (2010). *The Diversity of Life: With a New Preface (Questions of Science)* (2nd ed.). Belknap Press: An Imprint of Harvard University Press.

WTU Herbarium, Burke Museum, University of Washington. (2004). *Burke Herbarium Image Collection*. Copyright (c) 2004–2021 WTU Herbarium, Burke Museum, University of Washington. Retrieved November 6, 2021, from https://biology.burke.washington.edu/herbarium/imagecollection/taxon.php?Taxon=Saxifraga%20hyperborea

W.D.F.W. (2022). *Woodland caribou*. Washington Department of Fish & Wildlife. Retrieved January 24, 2024, from https://wdfw.wa.gov/species-habitats/species/rangifer-tarandus#:%7E:text=The%20southern%20mountain%20caribou%20distinct,is%20listed%20as%20state%20endangered.

Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., . . . Merow, C. (2020). A standard protocol for reporting species distribution models. *A Standard Protocol for Reporting Species Distribution Models*, *43*(9), 1261–1277. https://doi.org/10.1111/ecog.04960