5-2020

The Utility of Environmental DNA and Species Distribution Models in Assessing the Habitat Requirements of Twelve Fish Species in Alaskan North Slope Rivers

James B. Eddings
Utah State University

THE UTILITY OF ENVIRONMENTAL DNA AND SPECIES DISTRIBUTION MODELS IN ASSESSING THE HABITAT REQUIREMENTS OF TWELVE FISH SPECIES IN ALASKAN NORTH SLOPE RIVERS

by

James B. Eddings

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Ecology

Approved:

Charles Hawkins, Ph.D. Major Professor	Phaedra Budy, Ph.D. Committee Member
Joseph Wheaton, Ph.D. Committee Member	Richard S. Inouye, Ph.D. Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2019

ABSTRACT

The Utility of Environmental DNA and Species Distribution Models in Assessing the Habitat Requirements of Twelve Fish Species in Alaskan North Slope Rivers

by

James B. Eddings, Master of Science

Utah State University, 2019

Major Professor: Dr. Charles Hawkins
Department: Watershed Sciences

Subsistence fishing is a vital component of Alaska's North Slope borough economy and culture that is being threatened by anthropogenic disturbances. These threats mean the fish must be protected, but the size of the region makes conservation planning difficult. Fortunately, advances in species distribution models (SDMs), environmental DNA (eDNA), and remote sensing technologies provide potential to better understand species' needs and guide management. The objectives of my study were to: (1) map the current habitat suitability for twelve fish species occurring in Alaska's North Slope, (2) determine if SDMs based on eDNA data performed similarly to, or improved, models based on traditional sampling data, and (3) predict how species distributions will shift in the future in response to climate change. I was able to produce robust models for 8 of 12 species that relate environmental characteristics to a species' presence or absence. I also produced maps from model predictions to identify stream reaches where species are likely, or not, to occur. Unfortunately, the use of eDNA data did not produce useful

models in Northern Alaskan rivers. However, we were able to use data obtained from traditional sampling methods to predict current and future species distributions that should help inform management.

(77 pages)

PUBLIC ABSTRACT

The Utility of Environmental DNA and Species Distribution Models in Assessing the Habitat Requirements of Twelve Fish Species in Alaskan North Slope Rivers

James B. Eddings

Subsistence fishing is a vital component of Alaska's North Slope borough economy and culture that is being threatened by human disturbance. These threats mean the fish must be protected, but the size of the region makes conservation planning difficult. Fortunately, advances in species distribution models (SDMs), environmental DNA (eDNA), and remote sensing technologies provide potential to better understand species' needs and guide management. The objectives of my study were to: (1) map the current habitat suitability for twelve fish species, occurring in Alaska's North Slope, (2) determine if SDMs based on eDNA data performed similarly to, or improved, models based on traditional sampling data, and (3) predict how species distributions will shift in the future in response to climate change. I was able to produce robust models for 8 of 12 species that relate environmental characteristics to a species' presence or absence and identify stream reaches where species are likely to occur. Unfortunately, the use of eDNA data did not produce useful models in Northern Alaskan rivers. However, I was able to generate predictions of species distributions into the future that should help inform management for years to come.

ACKNOWLEDGMENTS

I would first like to thank Dr. Charles Hawkins for taking me on as a student and mentoring me throughout this entire project. Similarly, I thank Dr. John Olson for his mentorship and assistance throughout. I would also like to thank my committee members, Drs. Phaedra Budy and Joe Wheaton for their invaluable support and input from proposal to final manuscript. Drs. Jiming Jin and Adele Cutler also contributed by providing useful opinions and data for which I am extremely grateful.

I give special thanks to my family, girlfriend, friends, and colleagues for their encouragement, moral support, input, questions, and patience over the last three years.

James B. Eddings

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Subsistence fishing is an important component of Alaska's North Slope borough economy and culture that is being threatened by human alteration of the landscape (Jones et al., 2009). These fish are important, not only as catch, but also for their roles in sustaining the aquatic ecosystems in which they are found (Earnst, 2004; R. Davies and Walker, 1986). Significant oil production and continuing exploration paired with a changing climate are likely to negatively affect many fish species throughout the region (Grunblatt and Atwood, 2014). There is a need for innovative methods of monitoring and managing fish species because of the increase in potential threats associated with National Petroleum Reserve development within the region. With ongoing development and a continually changing climate, there is a clear need to identify habitat requirements of important fish species. With information on important habitat requirements, resource managers can design mitigation strategies to best minimize habitat loss and protect populations (Rosenfeld, 2003).

Species distribution models (SDMs) can help us understand a species' habitat requirements and identify important areas for protection (Elith and Leathwick, 2009). These models create statistical relationships between the presence or absence of a species and the ambient environmental conditions at a particular site (Guisan and Zimmermann, 2000). However, SDMs require a systematic approach in which ecological theory is used in all steps of the model building process to ensure the predictions are useful (Austin, 2002). It is therefore important to consider ecological theory while selecting predictive environmental variables (Pearson et al., 2004), assessing model assumptions and their validity when making projections into the future (Araújo and Pearson, 2005), and

determining potential causes of prediction errors (Guisan and Thuiller, 2005). If implemented properly, SDMs can help resource managers identify specific areas or stream reaches that are most important to protect to maintain species (Guisan and Thuiller, 2005).

SDMs are capable of answering many biological or ecological questions. Many SDMs have been developed, but they have not been used frequently to address conservation problems (Guisan et al., 2013). In a few cases, SDMs have been used to identify critical habitat needs and inform decision making (e.g., Heinrichs et al., 2010; Mainali et al., 2015; Porfirio et al., 2014). Unfortunately, many models are built only for small regions or use predictors whose relationship to the species is not direct. Producing SDMs with interpretable results is key in making conservation decisions, as is ensuring models can be reproduced and updated as needed. Closer collaboration between researchers and decision makers could facilitate the use of SDMs in more decision-making (Guisan et al., 2013). SDMs are capable of answering questions about many potentially manageable habitat requirements if interpretable predictors are used and communication with resource managers is effective.

Many habitat features can influence where fish occur in streams, so predictor selection will affect the robustness of models (Guisan et al., 2013). Furthermore, environmental characteristics that can be observed remotely are vital because the sheer size and remoteness of Alaska’s North Slope makes physical sampling difficult. Factors such as temperature, waterbody connectivity, and hydrology can all influence one or more life stages of a given fish species (Radinger and Wolter, 2014). Many of these factors also have direct or indirect relationships with climate and can thus be used to make predictions of how species distributions may change in the future.

Robust calibration of SDMs requires comprehensive distribution data across regions, which is often difficult to obtain across large or inaccessible areas when using traditional sampling methods (Thuiller et al., 2004). This is especially challenging in Alaska's North Slope, whose size and remoteness make gathering distribution data time consuming, expensive, and impractical. Traditional methods are also often non-standardized and depend on taxonomic expertise, which is in decline (Bonar et al., 2009; Perez et al., 2017). Fortunately, a novel technique for detecting aquatic species from water samples has been developed in recent years—environmental DNA (eDNA). eDNA is genetic material shed by species into the environment that is used to estimate the presence of species (Ficetola et al., 2008; Takahara et al., 2012). This technique is relatively new, but it has already proved to be a useful tool in detecting aquatic species (Nathan et al., 2014; Rees et al., 2014). The use of eDNA data in SDMs could improve the quality of models, if the data it provides is as, or more, accurate than data collected traditionally. eDNA sampling can result in larger and more precise samples, while also reducing effort required in data acquisition, as it only requires a small water sample at each site (Minamoto et al., 2012).

Similar to distribution data, environmental data can be difficult, time consuming, and expensive to collect. Fortunately, recent advances in remote sensing technology provide a potential solution. Habitat features like temperature, connectivity, and hydrology can all influence one or more life stages of the different fish species. Remote sensing (e.g. satellites) provides the ability to quantify many environmental features across large spatial scales that may be predictive of aquatic species distributions. In an area as large as

Alaska's North Slope, remotely sensed predictor variables offer a more efficient and cost-effective means for building an SDM.

SDMs based on current habitat conditions can be useful in identifying important habitat characteristics, but models that can predict potential future distributions as well as current distributions are especially needed (Guisan and Thuiller, 2005). Resource managers and conservationists can make decisions about regions to protect or habitat characteristics to maintain based on current conditions. However, with land use and climate change, management strategies may need to change to remain effective. Data from global climate models (GCMs) provide the capability to calibrate models based on spatial variation in current climate conditions and then predict how distributions will shift with climate change (Huntley et al., 2004).

In this thesis, I focused on 12 fish species that are present in the North Slope and important to local ecosystems and communities. Eight of the species studied here are salmonids: Round Whitefish (*Prosopium cylindraceum)*, Humpback Whitefish (*Coregonus pidschian*), Broad Whitefish (*Coregonus nasus*), Bering Cisco (*Coregonus laurettae*), Least Cisco (*Coregonus sardinella*), Arctic Grayling (*Thymallus arcticus*), Arctic Char complex (*Salvelinus*), and Chum Salmon (*Oncorhynchus keta*). The other four species include Alaska Blackfish (*Dallia pectoralis*), Ninespine Stickleback (*Pungitius pungitius*), Slimy Sculpin (*Cottus cognatus*), and Burbot (*Lota lota*). The whitefish, cisco, char, salmon, and grayling are all known to be anadromous or semi-anadromous species in Alaska's North Slope (Harper et al., 2012; Lee David S. (David Stephen) et al., 1980; McPhail and Lindsey, 1970). Most of these fish inhabit river mouths and brackish waters near the coast other than when they move upstream in late summer or fall to spawn

(McPhail and Lindsey, 1970). The other four species have slightly different life histories. Most are not anadromous, but all species have been found in brackish or marine waters along Alaska's coast (Lee, David S. et al., 1980). Blackfish and stickleback are small species that prefer Alaska's lowlands. Burbot and sculpin also prefer lowland areas because streams in these areas have relatively low stream gradients and are close to lakes, where they spawn. However, they differ from many of the other species in that they are most active at night and spawn in spring and winter respectively (Hofmann and Fischer, 2002; McPhail and Lindsey, 1970; McPhail and Paragamian, 2000). Although we know basic aspects of the ecology and life histories of these twelve species, we lack critical information regarding what factors most strongly influence their distributions.

The objectives of my study were to: (1) map the current habitat suitability of Alaska's North Slope for the twelve target fish species, (2) determine if SDMs based on eDNA data performs similarly to, or improve, models based on traditional data, and (3) predict how species distributions will shift in the future in response to climate change. The ~245,000 km^2 study region, which also contains the National Petroleum Reserve and nearly 9,500 people, is a vast and remote area that is home to a unique ecosystem upon which many fish species, and the people who inhabit the region, depend (Davies and Walker, 1986; Wolfe and Walker, 1987). To address my goals, I built SDMs for each of the twelve fish species based on both eDNA data and data produced from traditional sampling methods throughout Alaska's North Slope and then applied these models to current and future climate scenarios.

METHODS

2.1 Overall approach

I used Random Forest models, remotely sensed predictors, and species occurrence data to create SDMs for the twelve fish species. I used several model performance metrics to compare the performance of models based on eDNA versus traditionally sampled occurrence data. Then I used recent climate data to calibrate models and forecast future distributions. Finally, I used the SDMs to predict how species' distributions will likely change in response to climate change.

2.2 Study area and data

The North Slope borough of Alaska (Fig. 1) is the northernmost human occupied region in both Alaska and North America, and according to the 2010 US census is home to nearly 9500, mostly indigenous, people. Despite its small population, Alaska's North Slope covers 245,521 square kilometers extending from the Brooks Range to the Arctic Ocean. The North Slope is home to a uniquely diverse landscape (Kittel et al., 2011). Its landscape includes the mountains of the Brooks Range in the south and coastal lowlands dotted with glacial lakes and braided rivers in the north. A major reason for such high interest in protecting species throughout this region is the ongoing development of the National Petroleum Reserve in Alaska (NPRA). At 95,506 square kilometers, the NPRA covers roughly 39% of the North Slope and includes the headwaters of the Colville River and other river systems (Fig. 1).

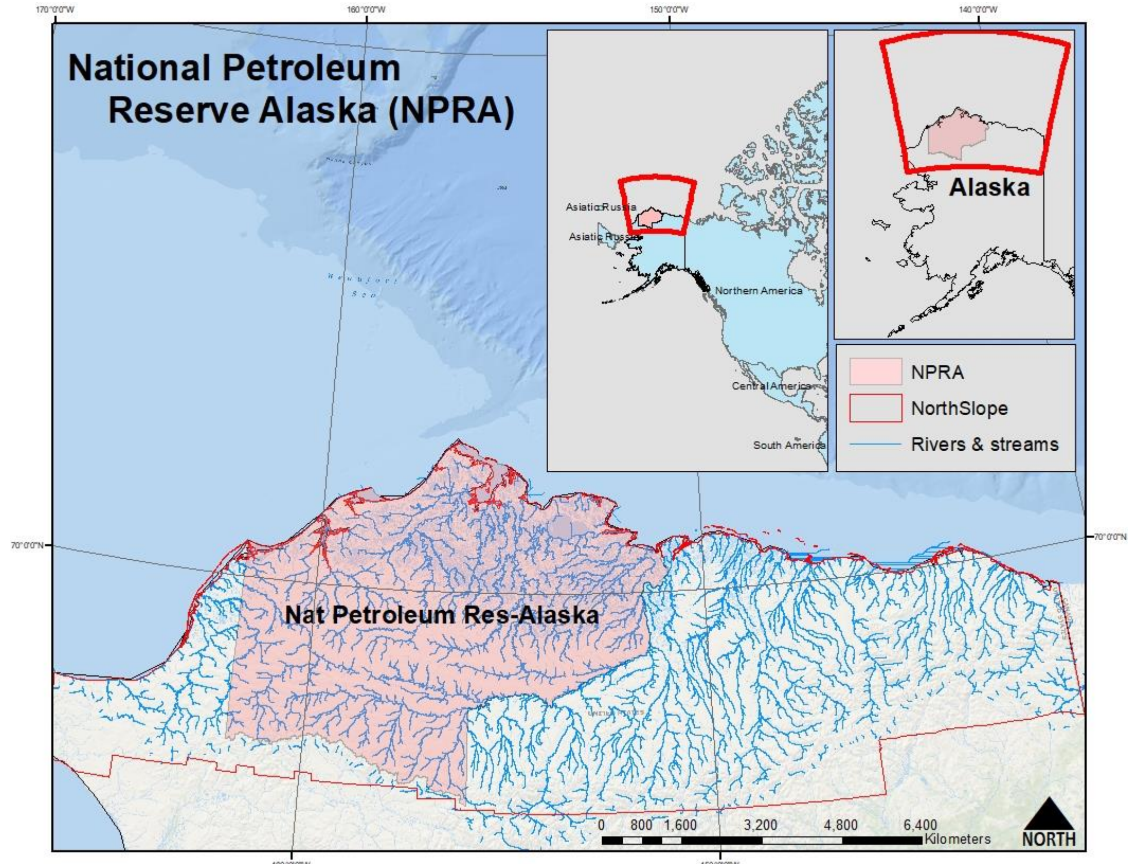


Figure 1: The North Slope borough of Alaska and the National Petroleum Reserve Alaska (NPRA).

The U.S. Bureau of Land Management and Alaska Department of Fish and Game collected presence-absence data for the 12 fish species through the summer and fall months of 2001-2016. Species presence or absence was determined with both traditional and eDNA sampling methods in stream segments from 1-5 km in length. Traditional sampling methods were similar to those described by Portt et al., (2006) and included electrofishing and seine netting. eDNA samples were collected, extracted, and identified following methods similar to those in Thomsen et al., (2012). One hundred ninety-eight river and stream sites were sampled throughout the North Slope. Of these, 68 sites were sampled for eDNA and 130 were sampled with traditional methods.

2.3 Random Forest model development and evaluation

Random Forest (RF) is a nonparametric modeling technique that takes advantage of bootstrapped resampling to produce and aggregate large numbers of different classification and regression trees (Breiman, 2001). There are many advantages to using RF models for species distribution modeling, which has led to their use in several ecological assessments (Pacifici et al., 2015; Terrado et al., 2016; Treglia et al., 2015). RF models are not only highly accurate relative to other methods, but they also are capable of handling large numbers of predictor variables and complex interactions (Cutler et al., 2007). I used the *randomForest* package in R (Liaw et al., 2002) to develop models.

2.3.1 Balancing presence-absence data

The presence/absence data were generally severely unbalanced for all species, with many species having much fewer presences than absences. Unbalanced data can be a problem when building RF models because a percentage of the data is pulled as test data and not used to train the model. If all or most of the presences were to get pulled as test data, the model algorithm has fewer presences to build the model from. To combat this problem, I used the *sampsize* function in *randomForest*, which allowed me to specify that the model be built based on a one to one ratio of presences to absences.

2.3.2 Predictor selection

I considered 83 candidate environmental predictor variables that were available for the entire North Slope borough (Table 1). Past studies have shown environmental characteristics such as stream temperature, evapotranspiration, and connectivity perform

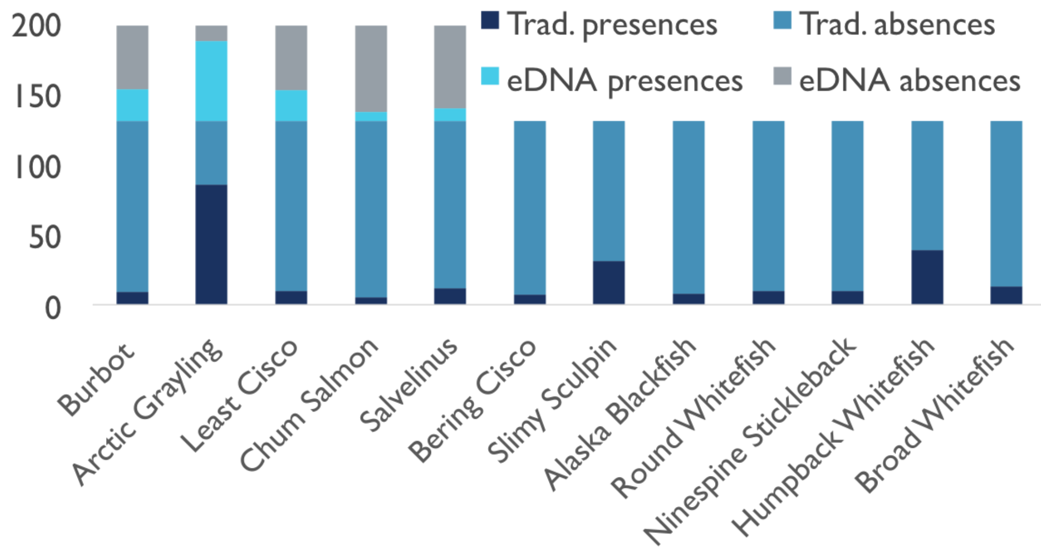


Figure 2: Numbers of presences and absences for each of the twelve fish species.

well in predicting fish distributions (Radinger and Wolter, 2014). However, stream temperature cannot be remotely sensed either easily or accurately, so I used air temperature as a surrogate for stream temperature (Bailing et al., 2018; Rodell et al., 2004). Sixty-two of the predictors were compiled from a variation of StreamCat, a stream-catchment dataset that includes both land use and natural landscape variables (Hill et al., 2016), built specifically for the North Slope by Dr. John Olson (California State University at Monterey Bay). These variables include estimates of evapotranspiration, enhanced vegetation index, gross primary productivity, snow cover, and fire derived from the Moderate Resolution Imaging Spectroradiometer (MODIS). Each of these predictors were available at both the 'local-catchment' (reach-scale watershed) and 'watershed' (upstream watershed) scale.

Table 1. All (83) of the candidate remotely sensed environmental predictor variables considered for inclusion in the species distribution models for 12 fish species in Alaska's North Slope. The 6 predictors used in final models are denoted with a '*.'

Predictor name	**Description**
AreaSqKM*	Watershed area, (km^2)
ET_Cat	Longterm average Evapotranspiration (2000-13) from MODIS--catchment scale ($kg/m^2/8day$)
ET_Ws	Longterm avg. Evapotranspiration (2000-13) from MODIS--watershed scale
EVI_Cat	Longterm avg. Enhanced Vegetation Index from MODIS (EVI)
EVI_Ws	Longterm avg. Enhanced Vegetation Index from MODIS
Fire_Cat	Longterm avg. Fire data from MODIS (km^2)
FireMax_Cat	Longterm max. Fire data from MODIS
FireSum_Cat	Longterm sum of Fire data from MODIS
Fire_Ws	Longterm avg. Fire data from MODIS
FireMax_Ws	Longterm max. Fire data from MODIS
FireSum_Ws	Longterm sum of Fire data from MODIS
GPP_Cat	Longterm avg. Gross Primary Productivity from MODIS ($kg\ C\ m^2$)
GPP_Ws	Longterm avg. Gross Primary Productivity from MODIS
WV_Cat	Longterm avg. atmospheric water vapor from MODIS (cm)
WV_WS	Longterm avg. atmospheric water vapor from MODIS
Alder_Cat	Proportion of the catchment that is Alder
Arcto_Cat	Proportion of the catchment that is Arctophila
Bare_Cat	Proportion of the catchment that is Bare
Birch_Cat	Proportion of the catchment that is Bircerial
Burned_Cat	Proportion of the catchment that is Burned
Carex_Cat	Proportion of the catchment that is Carex
CoastMarsh_Cat	Proportion of the catchment that is Coastal Marsh
Deciduous_Cat	Proportion of the catchment that is Deciduous
Dwarf_Cat	Proportion of the catchment that is Dwarfsdryas
Dwarfother_Cat	Proportion of the catchment that is Dwarfsdryas other
IceSnow_Cat	Proportion of the catchment that is Ice/Snow
Willow_Cat	Proportion of the catchment that is Low Tall Willow
Marine_Cat	Proportion of the catchment that is Marine Beach
Mesicsedge_Cat	Proportion of the catchment that is Mesic sedge
Mesiherb_Cat	Proportion of the catchment that is Mesiher
Mixleaf_Cat	Proportion of the catchment that is open mixed leaf
Needle_Cat	Proportion of the catchment that is open needle
Openwater_Cat	Proportion of the catchment that is Open water
Sparsveg_Cat	Proportion of the catchment that is Sparsveg
Tussock_Cat	Proportion of the catchment that is Tussocks
Unclassified Cat	Proportion of the catchment that is Unclassified

Table 1 (cont.)

Wetsedgesph_Cat	Proportion of the catchment that is Wet sedge_sph
Wetsedge_Cat	Proportion of the catchment that is Wet sedge
Woodneedle_Cat	Proportion of the catchment that is Wood needle leaf
Alder_Ws	Proportion of the watershed that is Alder
Arcto_Ws	Proportion of the watershed that is Arctophila
Bare_Ws	Proportion of the watershed that is Bare
Birch_Ws	Proportion of the watershed that is Bircerial
Burned_Ws	Proportion of the watershed that is Burned
Carex_Ws	Proportion of the watershedthat is Carex
CoastMarsh_Ws	Proportion of the watershed that is Coasalt Marsh
Deciduous_Ws	Proportion of the watershed that is Deciduous
Dwarf_Ws	Proportion of the watershed that is Dwarfsdryas
Dwarfother_Ws	Proportion of the watershed that is Dwarfsdryas other
IceSnow_Ws	Proportion of the watershed that is Ice/Snow
Willow_Ws	Proportion of the watershed that is Low Tall Willow
Marine_Ws	Proportion of the watershed that is Marine Beach
Mesicsedge_Ws	Proportion of the watershed that is Mesic sedge
Mesiherb_Ws	Proportion of the watershed that is Mesi herb
Mixleaf_Ws	Proportion of the watershed that is open mixed leaf
Needle_Ws	Proportion of the watershed that is open needle
Openwater_Ws	Proportion of the watershed that is Open water
Sparsveg_Ws	Proportion of the watershed that is Sparsveg
Tussock_Ws	Proportion of the watershed that is Tussocks
Unclassified_Ws	Proportion of the watershed that is Unclassified
Wetsedgesph_Ws	Proportion of the watershed that is Wet sedge_sph
Wetsedge_Ws	Proportion of the watershed that is Wet sedge
Woodneedle_Ws	Proportion of the watershed that is Wood needle leaf
PetCCC_Cat	Potential evapotranspiration for 2000-09 from Canadian Center for Climate Modeling (mm)
PetCCC_Ws	Potential evapotranspiration for 2000-09 from Canadian Center for Climate Modeling
PetMPI_Cat	Potential evapotranspiration for 2000-09 from Max Planck Institute Earth model
PetMPI_Ws	Potential evapotranspiration for 2000-09 from Max Planck Institute Earth model
MAGT_Cat	Decadal avg. Mean Annual Ground Temperature from Geophysical Institure Permafrost Lab (C)
MAGT_Ws	Decadal avg. Mean Annual Ground Temperature from Geophysical Institure Permafrost Lab
ALT_Cat	Decadal avg. Active Layer Thickness from Geophysical Institute Permafrost Lab (cm)
ALT_WS	Decadal avg. Active Layer Thickness from Geophysical Institute Permafrost Lab

Table 1 (cont.)

CoastMin*	Minimum distance to coast (km)
CoastMax	Maximum distance to coast
CoastRange	Range (max. - min.) distance to coast
CoastMean	Mean distance to coast
LakeMin*	Minimum distance to winter lake refugia (i.e., does not completely freeze through) (m)
LakeMax	Maximum distance to winter lake refugia
LakeRange	Range (max. - min.) distance to winter lake refugia
LakeMean	Mean distance to winter lake refugia
ElevDif_m	Elevation drop from upstream to downstream end of reach (m)
Slope*	Segment slope calculated from ElevDif and segment length
ObsTemp*	Observed air temperature, 1981-2010 (C)
ObsPrecip*	Observed precipitation, 1981-2010 (mm/month)

Random Forest can produce models with a large suite of predictors, but I selected six predictors to create final models that would be both parsimonious and interpretable. Of the 83 initial candidate predictors, I eliminated 21 immediately because they had little to no variation across the North Slope (Appendix D). I produced initial models based on the remaining 62 variables to identify the strongest predictors. I then eliminated 56 more variables, most of which described vegetation types and productivity (e.g., gross primary productivity [GPP], enhanced vegetation index [EVI], willow, etc.) because relationships between fish occurrence and these variables would be difficult to interpret. The final list of six predictor variables included mean annual air temperature from 1981-2010 (ObsTemp), mean monthly precipitation (ObsPrecip) from 1981-2010, minimum distance to the coast (CoastMin), minimum distance to an unfrozen lake (LakeMin), channel slope, and watershed area (AreaSqKM) (Table 1). I compared performance of these models with the full models to ensure I did not sacrifice performance for interpretability. Models built with the six predictors I selected performed very similarly to the models built with 62 predictors. I therefore report the results of models built with only the six habitat

characteristics. I also used a scatterplot matrix to identify any correlations among predictor variables (Albuquerque et al., 2009), which could confound interpretation. These six predictors had the potential to produce simple, reproducible, and interpretable models that could also be used in predicting how climate change would influence distributions.

2.3.3 Model performance evaluation

I evaluated the performance of models with measures of sensitivity, specificity, the true skill statistic (TSS), and AUC as well as receiver operating characteristic (ROC) plots. Sensitivity, also called the true positive rate, measures the rate at which the model correctly predicts presences, and specificity, or the true negative rate, measures the rate at which the model correctly classifies absences (Lalkhen and McCluskey, 2008). The ROC plot is a graphical display of the ability of a model to predict presences and absences. The AUC is the area underneath the receiver operating curve and equals the probability of a classifier ranking a randomly chosen presence higher than a randomly chosen absence (Fawcett, 2006). AUC is often used to compare model performance (Hand, 2009). However, AUC can be noisy in that it does not do a good job of showing whether a model successfully predicts presences or absences (Hanczar et al., 2010). The true skill statistic is another method for comparing model performance that yields less noise than AUC. The TSS formula (sensitivity + specificity -1) provides a simple balance between the ability of the model to predict presences and absences and offers a good representation of how well a model is calibrated (Vezza et al., 2015). I defined models as robust, or useful for management purposes, if the AUC was greater than 0.75 and the TSS was greater than 0.5 (Allouche et al., 2006; McNyset, 2005).

I also generated variable importance and partial dependence plots for each species to assess how habitat characteristics influenced each species. Variable importance plots show the degree to which each variable is associated with the presence or absence of a species. Partial dependence plots show the marginal effects of an individual environmental characteristic on a species – i.e., how the probability of observing a species varies with changes in the value of each predictor.

2.3.4 eDNA and traditional model comparison

To compare performance of models built with eDNA data with those built with traditionally sampled data, I modeled each of the five species for which I had both types of data in 4 different scenarios. I had access to eDNA data for five fish species at 68 sites: Burbot, Arctic Char complex, Least Cisco, Arctic Grayling, and Chum Salmon. I had occurrence data based only on traditional sampling methods from 130 sites for all twelve fish species. Models for those seven species I did not have eDNA data for were built only with traditional data. For the five species with both data types, I built models with all of the traditional data (n = 130). Then I built models with just the eDNA data (n = 68). Next, I built models with traditional data randomly selected to match the sample size of the eDNA data (n = 68). Finally, I built models with traditional and eDNA data combined (n = 198). Whichever data type produced the best performing models was the data type I used to produce final models.

2.4 Predicting and mapping recent (1981-2010) distributions

To map likely current species distributions, I used the *predict* function from the *randomForest* package for each model built with whichever data type performed best.

This function produced a likelihood of occurrence (0-1) of a species in each stream segment in the North Slope. Then, I appended those predicted values to the attribute table of a National Hydrological Dataset (NHD) shapefile, an ARCGIS shapefile delineating the streams in the region. I color-coded each stream reach based on the predicted probability of occurrence for a species.

2.5 Predicting and mapping potential future (2071-2100) distributions

To project how fish distributions are likely to shift by the end of the century, I substituted end-of-century temperature and precipitation projections from three different climate models and two different emission scenarios into SDMs for each species. The climate projections I used were produced from (1) the Geophysical Fluid Dynamics Laboratory Earth System Model (GFDL-ESM2), (2) the Institut Pierre-Simon Laplace (IPSL) model, and (3) the Model for Interdisciplinary Research on Climate (MIROC5) (Dunne et al., 2013; Hourdin et al., 2013; Watanabe et al., 2010). I used these three different climate projections to account for several different potential scenarios (Figures 2 & 3, Appendix B). Data from each climate model included predictions based on both a low and high emission scenario to bracket the likely range of possible outcomes. Projections from each model-emission scenario combination were downscaled (Bailing et al., 2018; Rodell et al., 2004) to 150-250 km^2 resolution by Dr. Jiming Jin (Utah State University) based on observed climate data for the region (Global Land Data Assimilation System Version 2 [GLDAS-2]). I also used the different climate scenarios to model temperature and precipitation for recent years to see which projection produced temperature and precipitation values most similar to actual conditions. I then built RF models with the same six predictors (ObsTemp, ObsPrecip, CoastMin, LakeMin, Slope,

and AreaSqKM) but with future temperature and precipitation values from the downscaled climate model that best predicted recent temperature and precipitation (Figures 3 & 4).

I then used the same procedures described above to map the predicted probabilities of occurrence at the end of the century and show potential shifts in distributions. To show how potential distributions might change over time, I used the *calculate field* tool in ARCGIS to create a new field in the NHD shapefile's attribute table. This field included the difference between the present and end of the century in predicted probabilities of occurrence for individual reaches. I also summed the predictions of occurrence for all species in each stream segment to estimate species richness (out of the 12 target species) in each reach. Summing the same predictions for the end of century allowed me to identify areas where species richness is likely to change over time.

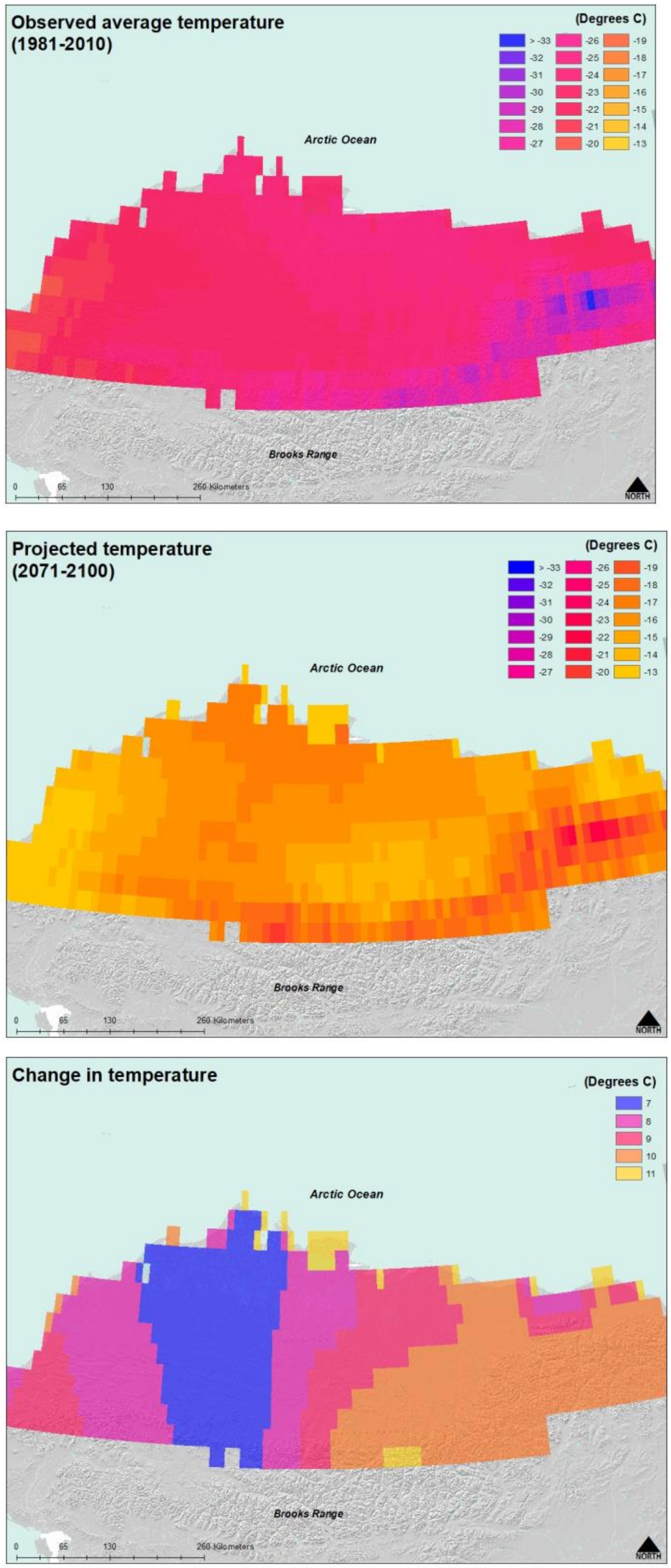


Figure 3: Current, projected, and change in temperature (GFDL 2 high emission scenario) in Alaska's North Slope.

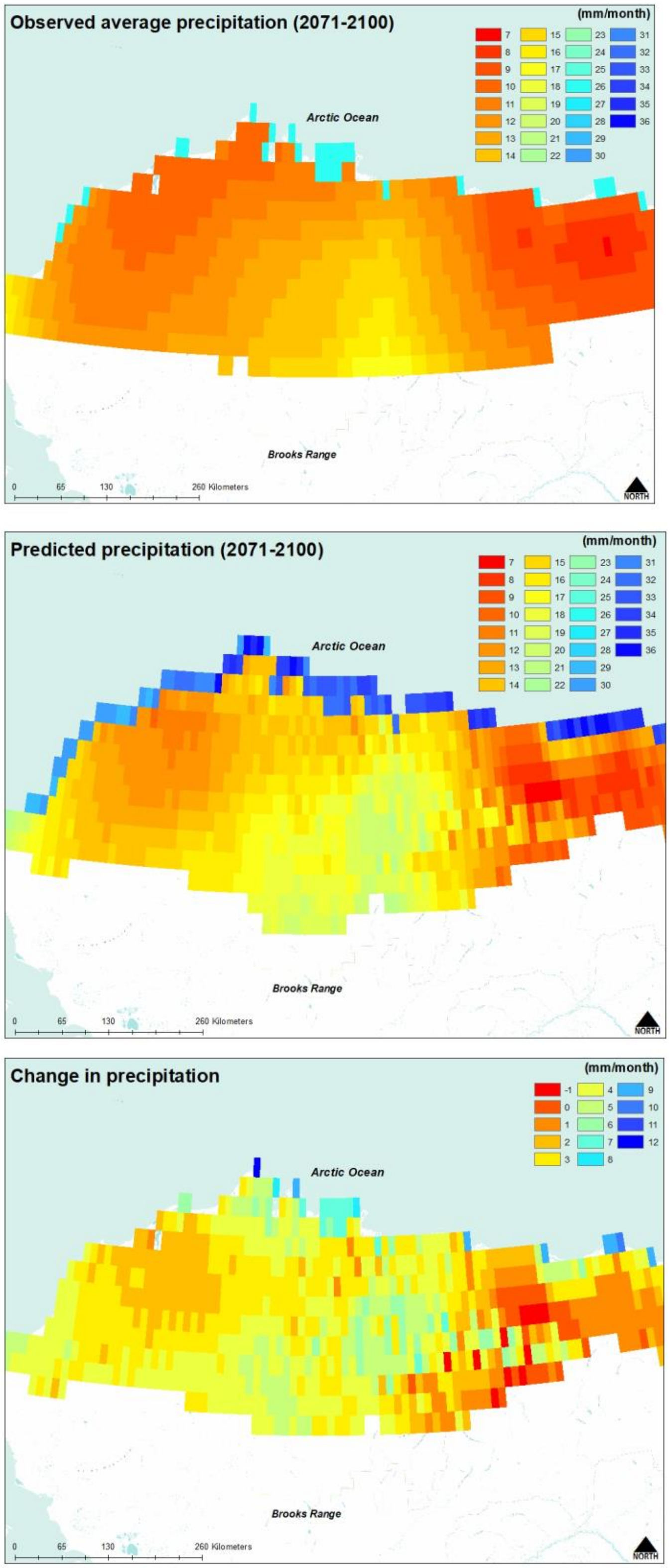


Figure 4: Current, projected, and change in precipitation (GFDL 2 high emission scenario) in Alaska's North Slope.

RESULTS

<u>3.1 eDNA and traditional model performance results</u>

In general, eDNA models performed markedly worse than traditional models (Table 2). Of the five species for which I compared model performance, four of the models built with traditionally sampled data far outperformed models built with eDNA data. Models built with traditional data for Burbot, Least Cisco, Chum Salmon, and *Salvelinus* had TSSs between 0.28 and 0.92, and AUCs between 0.56 and 0.97 (Table 2). Models built for the same species with eDNA data had TSSs of 0.01 to 0.39 and AUCs from 0.55 to 0.62 (Table 2). Sensitivity and specificity were also consistently higher for models of Burbot, Least Cisco, Chum Salmon, and *Salvelinus* built with traditional data (Table 2).

Models built with eDNA data for Arctic Grayling performed better than models built with traditional data. The eDNA model had a TSS of 0.39 and an AUC of 0.75, whereas the traditional model had a TSS of -0.05 and an AUC of 0.48. Specificity and sensitivity were both higher in the eDNA-based model (Table 2).

Models built with traditionally sampled data from 68 randomly selected sites consistently performed somewhere in between the traditional and eDNA models with the exception of *Salvelinus*, in which case this dataset produced the best model for the species (Table 2). Similarly, models built with traditional and eDNA data performed worse than models built with just traditional data, but better than models built with just eDNA data (Table 2).

Table 2: Model performance for models built with traditionally sampled data and those with eDNA data, as well as a combination of both data types for five fish species in Alaska's North Slope. Best performing models (based on TSS) are in boldface.

	Species	# of Presences	Sensitivity	Specificity	AUC	TSS
Traditional (n = 130)	**Burbot**	**9**	**0.89**	**0.88**	**0.85**	**0.77**
	Arctic Grayling	85	0.61	0.33	0.48	-0.05
	Least Cisco	**10**	**1.00**	**0.93**	**0.97**	**0.93**
	Chum Salmon	**5**	**0.40**	**0.88**	**0.56**	**0.28**
	Salvelinus	12	0.50	0.81	0.75	0.31
Traditional (n = 68)	Burbot	9	0.75	0.87	0.87	0.62
	Arctic Grayling	45	0.71	0.57	0.60	0.28
	Least Cisco	10	1.00	0.90	0.95	0.90
	Chum Salmon	5	0.40	0.79	0.66	0.19
	Salvelinus	**12**	**0.58**	**0.88**	**0.85**	**0.46**
eDNA (n = 68)	Burbot	23	0.39	0.62	0.56	0.01
	Arctic Grayling	**57**	**0.84**	**0.55**	**0.75**	**0.39**
	Least Cisco	22	0.59	0.63	0.62	0.22
	Chum Salmon	7	0.43	0.77	0.55	0.20
	Salvelinus	9	0.33	0.75	0.61	0.08
eDNA + Traditional (198)	Burbot	32	0.64	0.75	0.70	0.39
	Arctic Grayling	142	0.72	0.44	0.54	0.17
	Least Cisco	32	0.79	0.78	0.79	0.57
	Chum Salmon	12	0.41	0.82	0.55	0.24
	Salvelinus	24	0.41	0.78	0.68	0.19

3.2 Performance of final species distribution models all species

Models for Burbot, Least Cisco, Bering Cisco, Alaska Blackfish, Round Whitefish, Ninespine Stickleback, Humpback Whitefish, and Broad Whitefish built with all traditional data (n = 130) were robust with AUCs from 0.85 to 0.97 and TSSs from 0.53 to

0.92 (Table 3). Models for Arctic Grayling, Slimy Sculpin and *Salvelinus* had AUCs of 0.75, 0.71, and 0.85 respectively, but did not perform very well because they failed to correctly predict presences, resulting in low TSSs and sensitivity (Table 3). The model for Chum Salmon performed poorly with an AUC of 0.48 and a TSS of 0.28 (Table 3).

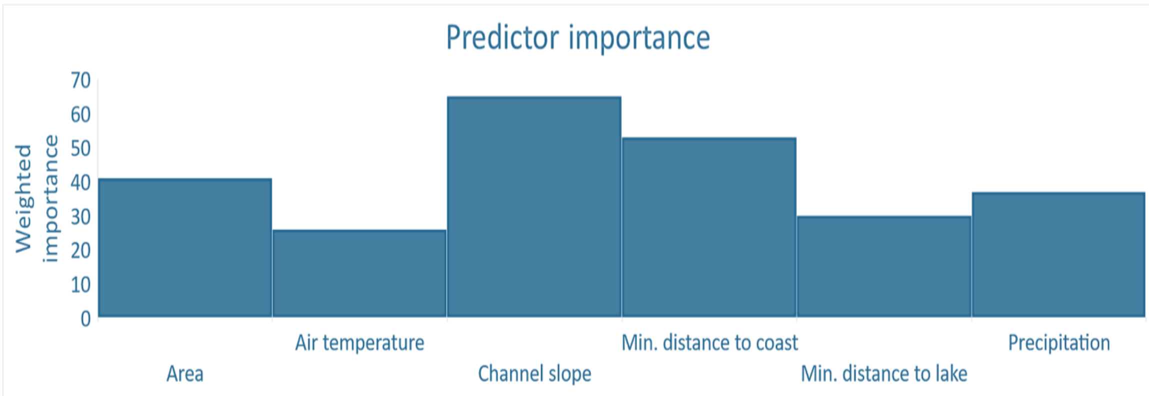


Figure 5: Weighted importance of the six core predictors used in final models.

Table 3: Model performance results for Random Forest models for twelve Alaskan fish species. Performance measures include: sensitivity, specificity, Area Under the Curve (AUC), and True Skill Statistic (TSS). Predictors are abbreviated as per Table 1. Sample size for all species was 130. Models built with eDNA data are denoted with a *, and models built with trimmed traditional data **.

Species	**# Present**	**Sensitivity**	**Specificity**	**AUC**	**TSS**	**Predictors in order of importance**
Burbot	9	0.89	0.88	0.85	0.77	Slope, CoastMin, Area, LakeMin, ObsTemp, ObsPrecip
Arctic Grayling*	57	0.84	0.55	0.75	0.39	Area, ObsTemp, LakeMin, Slope, ObsPrecip, CoastMin
Least Cisco	10	1.00	0.93	0.97	0.93	Slope, CoastMin, ObsTemp, Area, ObsPrecip, LakeMin
Chum Salmon	5	0.40	0.88	0.56	0.28	ObsPrecip, Area, Slope, LakeMin, CoastMin, ObsTemp

Table 3 (cont.)

*Salvelinus***	12	0.58	0.88	0.85	0.46	Slope, ObsPrecip, Area, CoastMin, LakeMin, ObsTemp
Bering Cisco	7	1.00	0.88	0.96	0.88	CoastMin, Slope, ObsTemp, LakeMin, Area, ObsPrecip
Slimy Sculpin	31	0.48	0.79	0.71	0.27	Slope, CoastMin, ObsPrecip, Area, LakeMin, ObsTemp
Alaska Blackfish	8	0.75	0.82	0.87	0.57	Slope, CoastMin, LakeMin, ObsPrecip, ObsTemp, Area
Round Whitefish	10	0.70	0.89	0.86	0.59	Slope, CoastMin, ObsPrecip, Area, ObsTemp, LakeMin
Ninespine Stickleback	39	0.74	0.79	0.85	0.53	CoastMin, Slope, LakeMin, ObsPrecip, Area, ObsTemp
Humpback Whitefish	10	0.90	0.90	0.94	0.80	Slope, CoastMin, Area, ObsPrecip, ObsTemp, LakeMin
Broad Whitefish	13	1.00	0.91	0.96	0.91	Slope, CoastMin, Area, ObsPrecip, LakeMin, ObsTemp

Channel slope and minimum distance to the coast were the most important predictors of species occurrences (Figure 5). Slope was the most important predictor for 75% of species and among the top three predictors for 1 of the species (Table 3, Figure 5, Appendix A). Minimum distance to the coast was the most important predictor for two species and among the top two predictors for nine of 12 species (Table 3, Figure 5, Appendix A). Mean annual air temperature and the minimum distance to an unfrozen lake were the least important predictors (Table 3, Figure 5, Appendix A).

Species' distributions across the North Slope were generally similar, with all predictions including higher likelihoods of occurrence in the coastal lowlands. All five whitefish species models (Broad Whitefish, Humpback Whitefish, Round Whitefish,

Bering Cisco, and Least Cisco) showed these species were more likely to inhabit stream reaches near or connected to the coastline. Few of the models showed these five species occurring inland very far, except in the Meade and Colville river systems (Figure 6). Burbot were predicted to be present in similar coastal systems but also further upstream in the Meade and Okpiksak rivers (Figure 6). Alaska Blackfish and Ninespine Stickleback were predicted to occur throughout a much larger portion of the rivers and streams in the North Slope. These two species were most likely to be found in the marshes south and west of the Meade river delta (Figure 6). Slimy Sculpin were predicted to occur throughout the center of the North Slope and had the farthest inland predicted range (Figure 6).

3.3 Predictions of future species distributions

Overall, the mean changes in probabilities of occurrence across the region were negligible for each species (Table 4). Similarly, species richness was not predicted to change much as a whole throughout the North Slope (Figure 7). The sum of change estimates for each of the 3,088 stream segments was 288.9, which suggests the stream reaches across the North Slope may be able to support slightly more fish species by the end of the century. Chum Salmon (+0.03), *Salvelinus* (+0.03), Round Whitefish (+0.03), and Humpback Whitefish (+0.03) were all predicted to increase in mean probability of occurrence, whereas Burbot was predicted to decrease in probability of occurrence (-0.03) (Table 4). Changes in mean probability of occurrence for Arctic Grayling, Least Cisco, Bering Cisco, Slimy Sculpin, Alaska Blackfish, Ninespine Stickleback, and Broad Whitefish were all negligible (Table 4). Despite little change through the region on average, some stream reaches were predicted to experience sharper increases or decreases

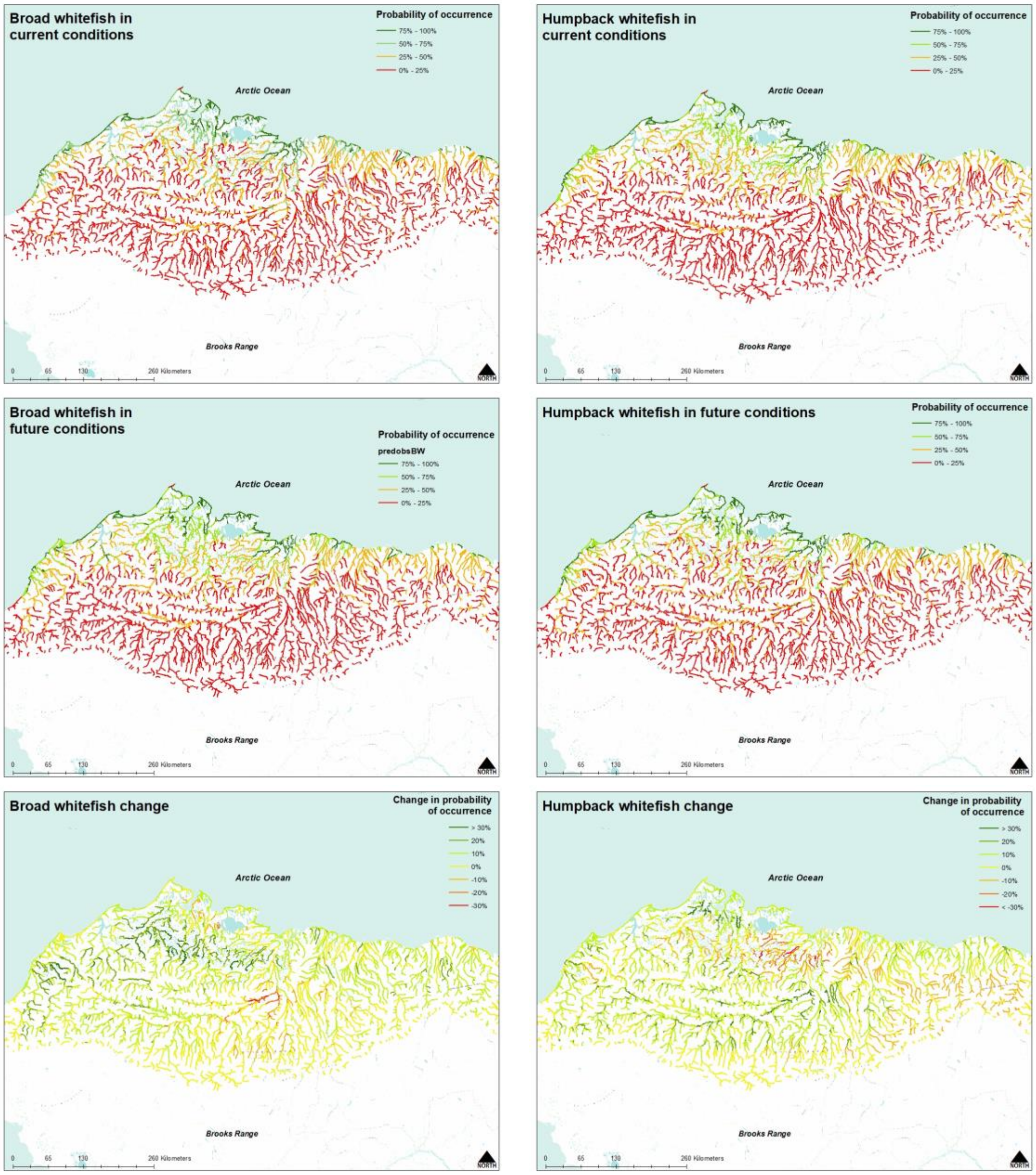


Figure 6: Current likely distribution (top), future potential distribution (middle), and the change in percent likelihood of detection (bottom) for Broad Whitefish and Humpback Whitefish.

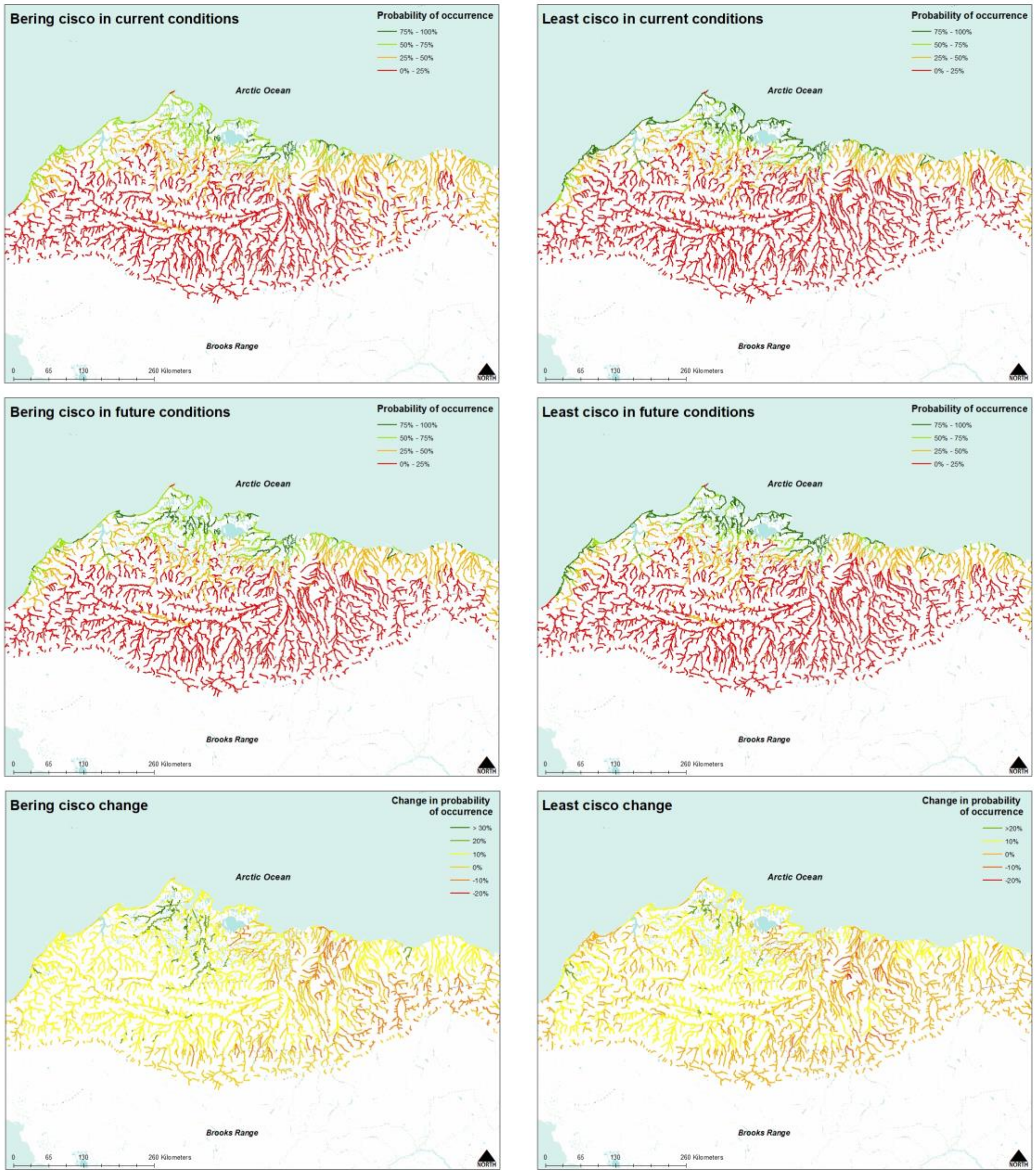


Figure 6: Current likely distribution (top), future potential distribution (middle), and the change in percent likelihood of detection (bottom) for Bering Cisco and Least Cisco.

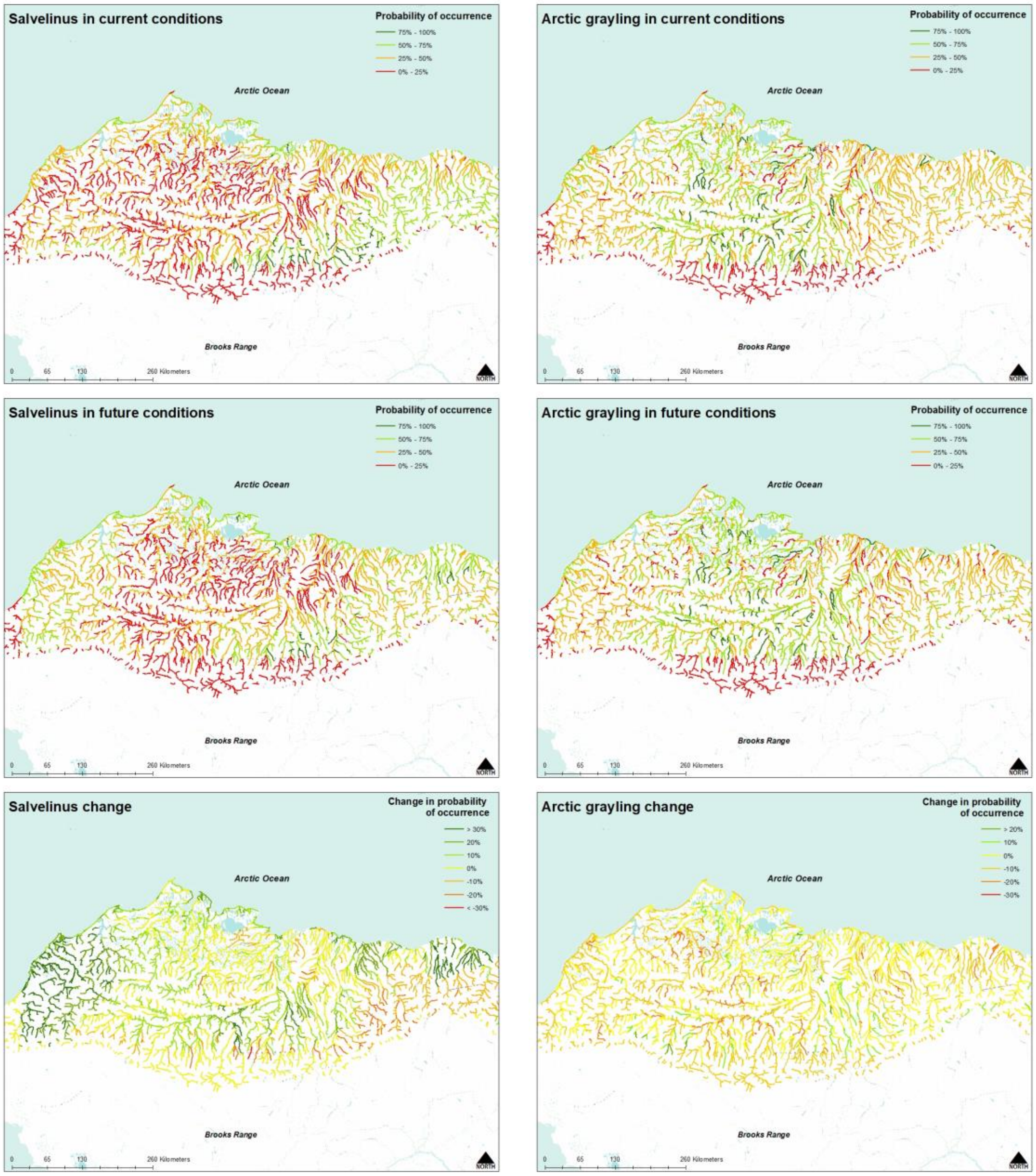


Figure 6: Current likely distribution (top), future potential distribution (middle), and the change in percent likelihood of detection (bottom) for Salvelinus and Arctic Grayling.

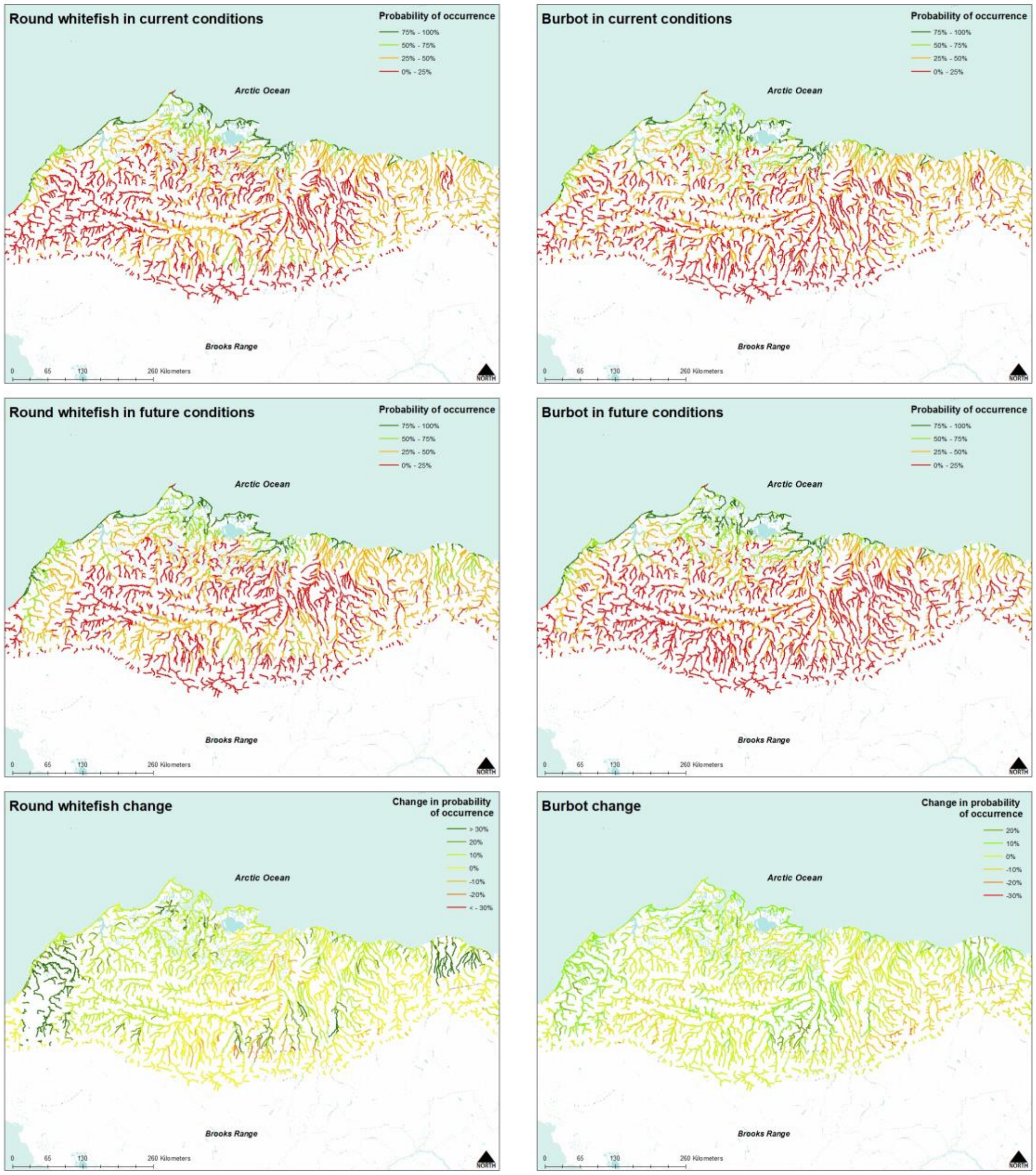


Figure 6: Current likely distribution (top), future potential distribution (middle), and the change in precent likelihood of detection (bottom) for Round Whitefish and Burbot.

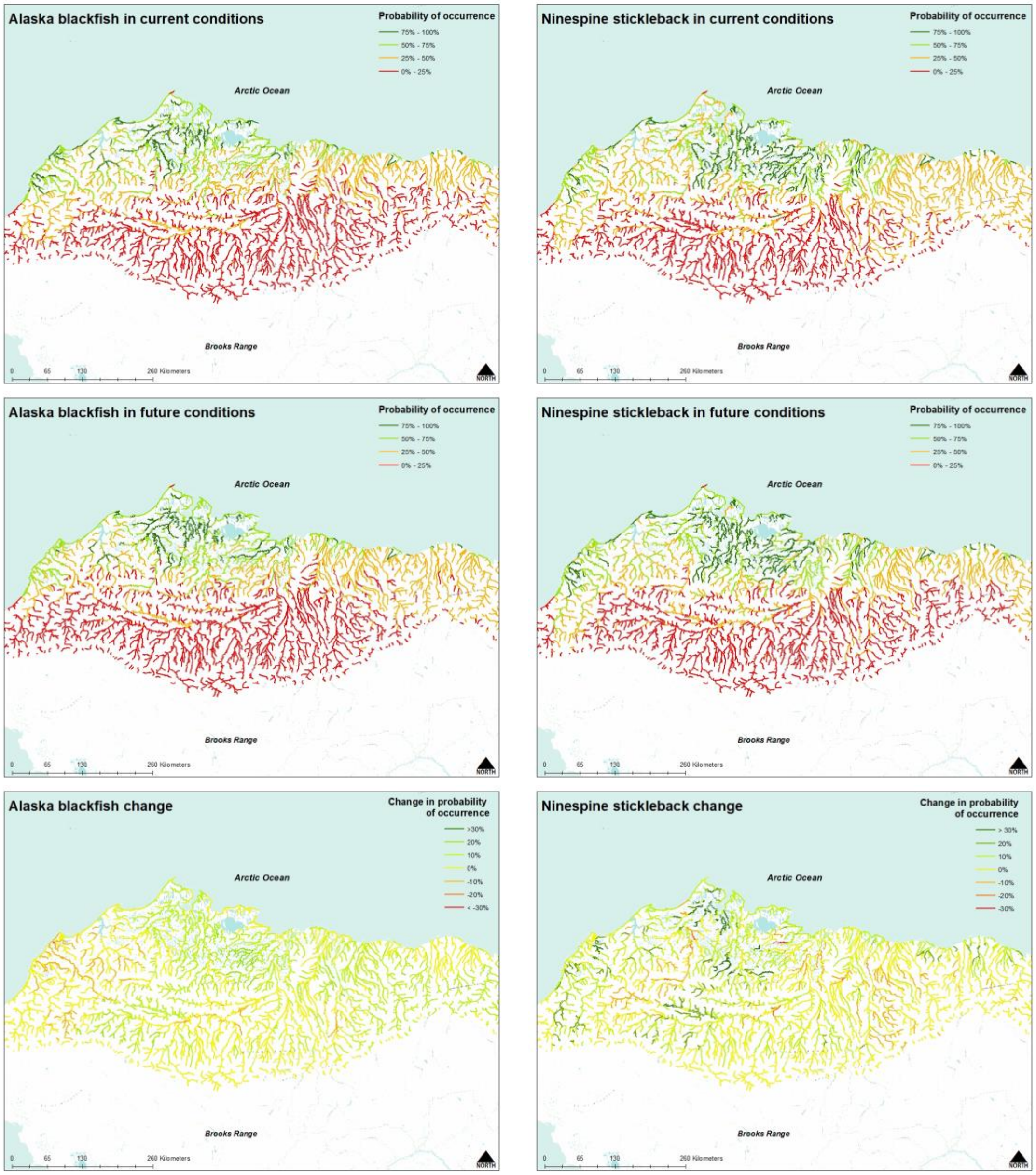


Figure 6: Current likely distribution (top), future potential distribution (middle), and the change in percent likelihood of detection (bottom) for Alaska Blackfish and Ninespine Stickleback.

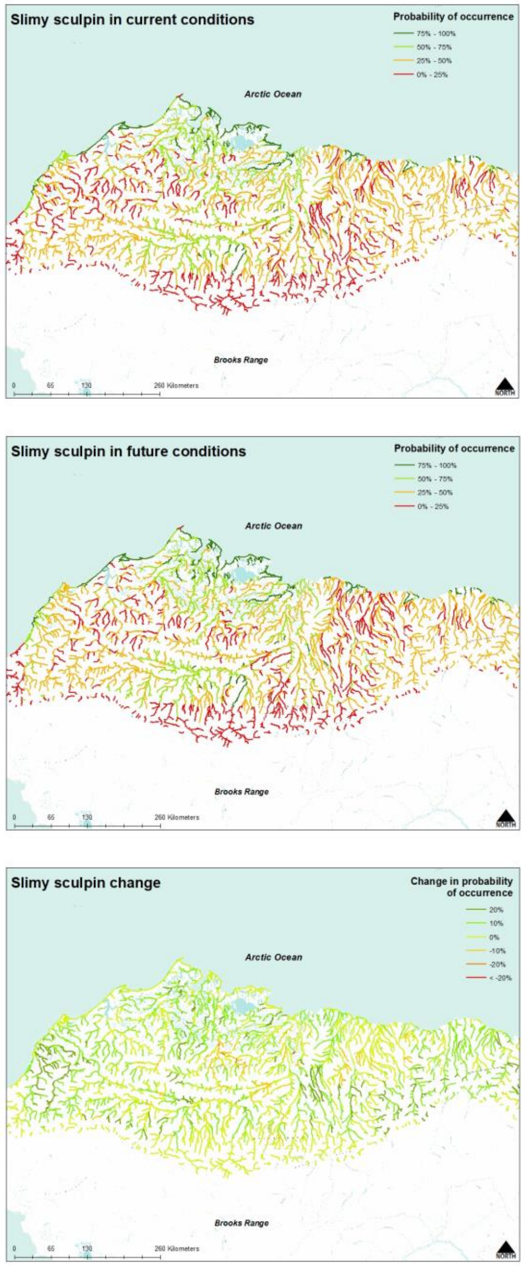


Figure 6: Current likely distribution (top), future potential distribution (middle), and the change in percent likelihood of detection (bottom) for Slimy Sculpin.

in species occurrence. For instance, *Salvelinus* probability of occurrence was predicted to increase by 30% in streams along the west coast of the North Slope (Figure 6). However, the same species was predicted to be less likely to occur by as much as 30% in the southeast (Figure 6).

Table 4: Mean probability of occurrence in an Alaskan North Slope stream for each species now and in the future, and the difference between the two.

	Mean probabilities of occurrence		
Species	Current	Future	Difference
Burbot	0.39	0.36	-0.03
Arctic Grayling	0.42	0.43	0.01
Least Cisco	0.27	0.27	0.00
Chum Salmon	0.36	0.39	0.03
Salvelinus	0.43	0.46	0.03
Bering Cisco	0.29	0.29	0.00
Slimy Sculpin	0.44	0.45	0.01
Alaska Blackfish	0.36	0.35	-0.01
Round Whitefish	0.36	0.39	0.03
Ninespine Stickleback	0.39	0.39	0.00
Broad Whitefish	0.33	0.34	0.01
Humpback Whitefish	0.32	0.35	0.03

Some species were predicted to undergo shifts in their distributions in response to climate change, whereas others were predicted to experience negligible shifts. A few stream reaches and networks stood out as places where probabilities of occurrence were likely to change more than other places. Broad Whitefish, Humpback Whitefish, Bering Cisco, Least Cisco, Arctic Grayling, Round Whitefish, and Ninespine Stickleback were all predicted to increase by +10-30 percent in the Meade and Okpiksak river networks in the center of the northern coast (Figure. 6). In the south (headwaters of the Colville river), Humpback Whitefish, Arctic Grayling, *Salvelinus*, Round Whitefish, Burbot, and Slimy Sculpin were predicted to increase by +10-30 percent (Figure 4). Broad Whitefish,

Salvelinus, Round Whitefish, and Slimy Sculpin were predicted to increase from +10-30 percent along the west coast of the North Slope (Figure 3).

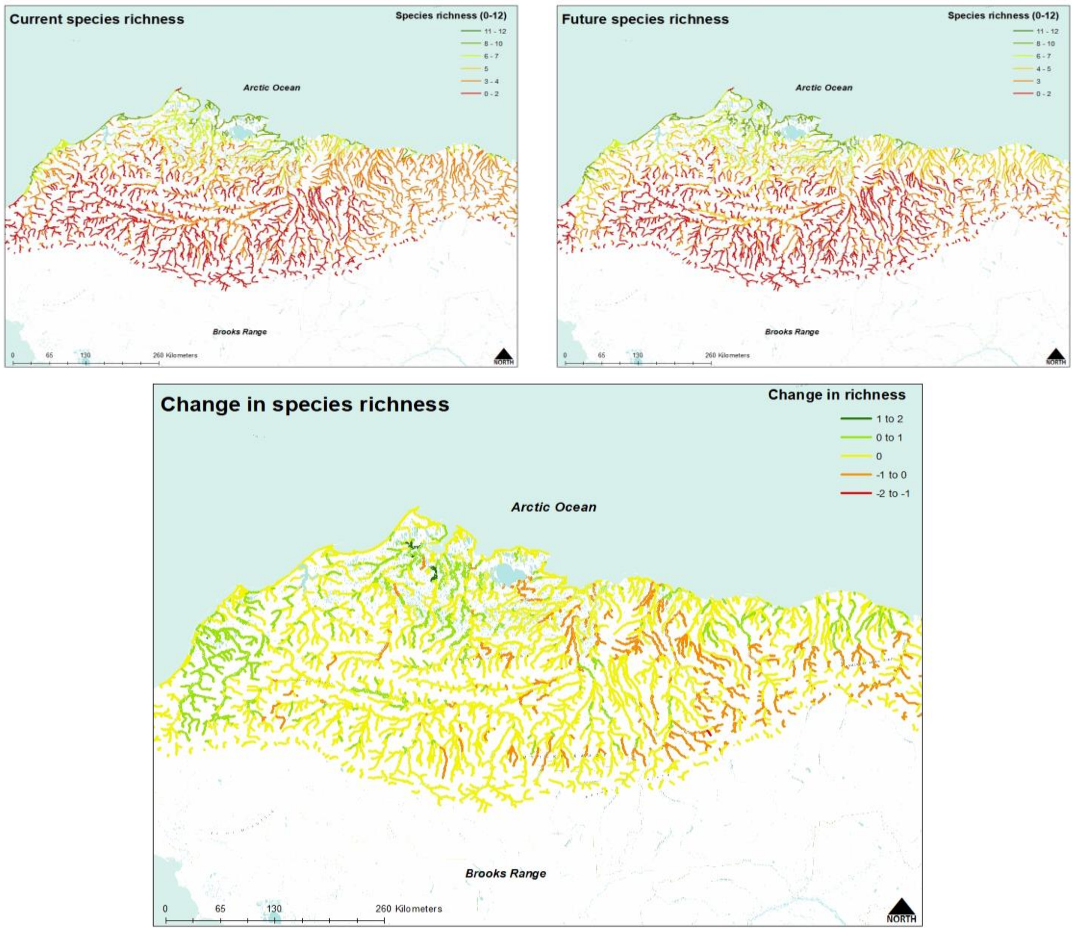


Figure 7: Modeled species richness (0 – 12) in stream segments in Alaska's North slope now (top left) and in the future (top right) and the predicted change in species richness by the end of the century (bottom).

DISCUSSION

This study was designed to build useful and robust species distribution models for twelve fish species in Alaska's North Slope. Here, I discuss the potential value of using SDMs to inform management actions in a remote region where data are difficult or expensive to collect. There were a few predictors I was surprised did or did not have a strong influence on species' distributions. I speculate as to why we saw signals from predictors that I did not expect. I also discuss the potential advantages and disadvantages of using eDNA data to produce SDMs for river-dwelling fish species. Finally, I discuss predictions of species distributions into the future and how managers might best use that knowledge to preserve species of interest.

4.1 Predicting current habitat suitability

In places such as northern Alaska, where on-the-ground data collection is difficult, SDMs are a useful tool with regard to identifying important habitat areas. In this study, I was able to produce robust models based on only remotely sensed habitat data and relatively sparse species occurrence data. Species distributions have been effectively modeled in the past with only remotely-sensed predictor variables (Gottschalk et al., 2005), but I was able to demonstrate their effectiveness in instances where species occurrences were rare. Despite a small amount of data, the Random Forest models produced accurate models (AUC > 0.8 and/or TSS > 0.5) for eight species. Accurate predictions allowed me to generate maps that show where the species are likely (or not) to occur throughout the North Slope, but especially in the NPRA where management priority is high due to increases in development. Managers can use these models and maps to

prioritize areas or stream reaches that are most important to protect and maintain populations of important fish species.

Although many remotely-sensed predictors are available in the North Slope, using predictors with potential redundancy can result in a difficult to interpret model. To make models more interpretable and useful, I built models with only six predictors whose potential relationships to fish occurrence were more interpretable. Random Forest is capable of generating robust models from large groups of predictors that may have complex interactions (Cutler et al., 2007). Unfortunately, these full models did not provide much useful information because some variables were redundant or had unclear relationships to the species. Instead of using such cumbersome and uninterpretable models, I was able to produce interpretable models based on just six predictors. These models performed similarly, or as well, as models built with all of the predictor variables and allowed more clear inferences to be made about species presence or absence in relation to habitat.

I found most species were predicted to occur in similar parts of the study region – i.e., coastal river systems. The majority of species were predicted to more likely occur in systems along or near the coastline, the only exceptions being Arctic Grayling and Slimy Sculpin. These results are consistent with the ranges of these species reported by McPhail and Lindsey, (1970). The salmonid species (whitefish, cisco, *Salvelinus*, and salmon), with the exception of Arctic Grayling, are anadromous species. It is therefore not surprising that these species were most likely predicted to occur in river deltas and coastal systems, or that their most important predictors were slope and minimum distance to the coast. Ninespine Stickleback and Alaska Blackfish were predicted to be present in coastal

regions as well, but not as often in deltas and river mouths. These two species are known to prefer coastal lowlands and inhabit brackish waters, which explains why they were not predicted to occur in the foothills to the south of the coastal plain region. Arctic Grayling and Slimy Sculpin were predicted in systems across the North Slope, although model performance for both species was weak. It is possible that there are suitable reaches that these species, sculpin especially because of their poor swimming abilities, are not able to reach. Burbot were predicted most often in coastal lowland streams as well, which is likely a function of the number of glacial lakes and ponds that occur in this region, which Burbot use for breeding and overwintering.

After modeling and mapping species distributions, I noticed that most species were predicted to occur in the coastal plain region where glacial lakes are prominent, yet minimum distance to a lake did not have a strong signal. Eight of eleven species were predicted to occur predominantly in the northern portion of the North Slope along the coast. Glacial lakes and ponds are abundant in this northern region above the Brooks Range, but minimum distance to a lake only ranked as high as the third most important predictor for only three species (Table 3). Especially for a species like Burbot, that uses lakes as spawning grounds, it was somewhat unexpected that distance to a lake was not a more important predictor of their occurrence. I suspect that, because lakes are so abundant throughout the region, there was less signal associated with connectivity to lakes as there might be in a dryer region. If more samples had been taken deeper into the Brooks Range, minimum distance to a lake may have shown more effect because lakes would be generally farther from more sample sites.

4.2 eDNA versus traditional model performance

I was able to produce robust models from sparse data, but I was not able to reproduce or improve models by using or adding eDNA data. Despite adding more presences to the data sets, using eDNA data or combining eDNA with traditional data did not produce better models. eDNA has generally been shown to be a useful tool for detecting aquatic species (e.g. Minamoto et al., 2012; Nathan et al., 2014; Rees et al., 2014; Takahara et al., 2012; Wilcox et al., 2013), but in this study eDNA data produced poorly performing models suggesting there is more to learn about eDNA before it can be used effectively in species distribution modeling. One factor that may limit the utility of eDNA is that DNA can be transported downstream to different habitats before it degrades (Jane et al., 2015). In such cases, eDNA could produce false positives that compromise model performance. In fact, Strickler et al., 2015 found in a lab experiment that most eDNA degradation occurs over three to ten days, and that eDNA persisted in colder conditions for up to fifty-eight days. In Northern Alaska, where colder water temperatures are likely to allow eDNA to persist, we need to better understand how quickly eDNA degrades in streams and rivers before it can be used effectively to model species distributions.

4.3 Future species distributions

Despite an increase in mean annual air temperature, overall changes in almost all species' distributions were negligible. With air temperatures projected to potentially increase by ten degrees C or more in parts of the North Slope, it is reasonable to expect that fish distributions would likely shift markedly. In fact, some climate change model predictions suggest streams in snowmelt-dominated regions will be highly sensitive to

climate change (Wu et al., 2012). However, more recent observations of climate change effects show that cold streams are relatively resilient to changes in air temperature (Lisi et al., 2015). The North Slope should continue to experience annual ice formation and snowfall despite rising air temperatures, and repeated snow and ice melt may mitigate changes in stream temperatures.

Changes in predictions of occurrence varied between species, with several species showing increased likelihood of occurrence in the NPRA despite similarities in predicted current distributions. Bering Cisco, Least Cisco, Broad Whitefish, Humpback Whitefish, Round Whitefish, Alaska Blackfish, and Ninespine Stickleback were all predicted to increase in probability of detection in the northern part of the study area, especially in the Meade River delta and the Okpiksak River (Figure 6). Many of these same species were predicted to decline in likelihood of detection in the eastern portion of the North Slope (Figure 6). Monthly average precipitation is projected to decline in the eastern part of the North Slope (Figure 4). Decreases in average precipitation could lead to lower flows, or a loss of connectivity, and ultimately declines in occurrence of these largely migratory species. *Salvelinus* and Round Whitefish displayed strong (> 20%) increases in probabilities of detection in the east and west ends of the North Slope (Figure 4). Arctic Grayling and Slimy Sculpin were projected to experience very little change in predicted distribution throughout the region from now to the end of the century (Figure. 6). However, models for Arctic Grayling and Slimy Sculpin performed poorly (TSSs < 0.5) making these future projections highly uncertain (Table 3).

With increasing infrastructure and a changing climate in Alaska's North Slope, it will be important for resource managers to identify critical habitat areas for fish species,

especially in the NPRA as development ramps up. All of the species modeled and mapped here were predicted to have high probabilities of detection in at least one river system within the NPRA, both now and by the end of the century. Several species were predicted to increase in likelihood of detection in the Meade River Delta and the Okpiksak river systems, which lie in the middle of the NPRA and are likely to experience significant land use changes over the next century (Figure 7). Managers should closely monitor and protect these systems, as well as the rest of the North Slope's coastal lowlands. These areas are where the majority of the species are likely to occur (Figure 7), so road building, drilling, and clearing of vegetation should be planned with fish presence in mind.

REFERENCES

Albuquerque, G., Eisemann, M., Lehmann, D.J., Holger, T., Magnor, M.A., 2009. VMV-Quality-based visualization matrices.

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Araújo, M.B., Pearson, R.G., 2005. Equilibrium of species' distributions with climate. Ecography (Cop.). 28, 693–695. https://doi.org/10.1111/j.2005.0906-7590.04253.x

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol. Modell. 157, 101–118. https://doi.org/https://doi.org/10.1016/S0304-3800(02)00205-3

Bailing, M., Li, Z., Cunzhu, L., Lixin, W., Chengzhen, J., Fuxiang, B., Chao, J., 2018. Temporal and spatial heterogeneity of drought impact on vegetation growth on the Inner Mongolian Plateau. Rangel. J. 40. https://doi.org/10.1071/RJ16097

Bonar, S.A., Hubert, W.A., Willis, D.W., 2009. Standard methods for sampling North American freshwater fishes. American Fisheries Society, Bethesda, MD.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brown, R.S., Duguay, C.R., Mueller, R.P., Moulton, L.L., Doucette, P.J., Tagestad, J.D., 2010. Use of Synthetic Aperture Radar (SAR) to Identify and Characterize Overwintering Areas of Fish in Ice-Covered Arctic Rivers: A Demonstration with Broad Whitefish and Their Habitats in the Sagavanirktok River, Alaska. Trans. Am. Fish. Soc. 139, 1711–1722. https://doi.org/10.1577/T09-176.1

Cutler, D.R., Edwards Jr., T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. RANDOM FORESTS FOR CLASSIFICATION IN ECOLOGY. Ecology 88, 2783–2792. https://doi.org/10.1890/07-0539.1

Davies, B., Walker, K., 1986. The Ecology of River Systems, XF2006265026. https://doi.org/10.1007/978-94-017-3290-1_1

Dunne, J.P., John, J.G., Shevliakova, E., Stouffer, R.J., Krasting, J.P., Malyshev, S.L., Milly, P.C.D., Sentman, L.T., Adcroft, A.J., Cooke, W., Dunne, K.A., Griffies, S.M., Hallberg, R.W., Harrison, M.J., Levy, H., Wittenberg, A.T., Phillips, P.J., Zadeh, N., 2013. GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics. J. Clim. 26, 2247–2267. https://doi.org/10.1175/JCLI-D-12-00150.1

Earnst, S.L., 2004. Status assessment and conservation plan for the yellow-billed loon (Gavia adamsii), Scientific Investigations Report. Reston, VA. https://doi.org/10.3133/sir20045258

Elith, J., Leathwick, J.R., 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27, 861–874. https://doi.org/https://doi.org/10.1016/j.patrec.2005.10.010

Ficetola, G.F., Miaud, C., Pompanon, F., Taberlet, P., 2008. Species detection using environmental DNA from water samples. Biol. Lett. 4, 423–425. https://doi.org/10.1098/rsbl.2008.0118

Gottschalk, T., Huettmann, F., Ehlers, M., 2005. Thirty years of analysing and modelling avian habitat relationships using satellite imagery data: A review. Int. J. Remote Sens. 26, 2631–2656. https://doi.org/10.1080/01431160512331338041

Grunblatt, J., Atwood, D., 2014. Mapping lakes for winter liquid water availability using SAR on the North Slope of Alaska. Int. J. Appl. Earth Obs. Geoinf. 27, 63–69. https://doi.org/https://doi.org/10.1016/j.jag.2013.05.006

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8, 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

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., Ferrier, S., Kearney, M.R., Possingham, H.P., Buckley, Y.M., 2013. Predicting species distributions for conservation decisions. Ecol. Lett. 16, 1424–1435. https://doi.org/10.1111/ele.12189

Guisan, A., Zimmermann, N., 2000. Guisan A, Zimmermann NE. Predictive habitat distribution models in ecology. Ecological Modeling. Ecol. Model. 135, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hanczar, B., Sima, C., Dougherty, E.R., Hua, J., Weinstein, J., Bittner, M., 2010. Small-sample precision of ROC-related estimates. Bioinformatics 26, 822–830. https://doi.org/10.1093/bioinformatics/btq037

Hand, D.J., 2009. Measuring classifier performance: a coherent alternative to the area under the ROC curve. Mach. Learn. 77, 103–123. https://doi.org/10.1007/s10994-009-5119-5

Harper, K.C., Harris, F., Miller, S.J., Thalhauser, J.M., Ayers, S.D., 2012. Life History Traits of Adult Broad Whitefish and Humpback Whitefish. J. Fish Wildl. Manag. 3, 56–75. https://doi.org/10.3996/022011-JFWM-011

Heinrichs, J.A., Bender, D.J., Gummer, D.L., Schumaker, N.H., 2010. Assessing critical habitat: Evaluating the relative contribution of habitats to population persistence. Biol. Conserv. 143, 2229–2237. https://doi.org/https://doi.org/10.1016/j.biocon.2010.06.009

Hill, R.A., Weber, M.H., Leibowitz, S.G., Olsen, A.R., Thornbrugh, D.J., 2016. The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. JAWRA J. Am. Water Resour. Assoc. 52, 120–128. https://doi.org/10.1111/1752-1688.12372

Hofmann, N., Fischer, P., 2002. Temperature Preferences and Critical Thermal Limits of Burbot: Implications for Habitat Selection and Ontogenetic Habitat Shift. Trans. Am. Fish. Soc. 131, 1164–1172. https://doi.org/10.1577/1548-8659(2002)131<1164:TPACTL>2.0.CO;2

Hourdin, F., Foujols, M.-A., Codron, F., Guemas, V., Dufresne, J.-L., Bony, S., Denvil, S., Guez, L., Lott, F., Ghattas, J., Braconnot, P., Marti, O., Meurdesoif, Y., Bopp, L., 2013. Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. Clim. Dyn. 40, 2167–2192. https://doi.org/10.1007/s00382-012-1411-3

Huntley, B., Green, R.E., Collingham, Y.C., Hill, J.K., Willis, S.G., Bartlein, P.J., Cramer, W., Hagemeijer, W.J.M., Thomas, C.J., 2004. The performance of models relating species geographical distributions to climate is independent of trophic level. Ecol. Lett. 7, 417–426. https://doi.org/10.1111/j.1461-0248.2004.00598.x

Jane, S.F., Wilcox, T.M., McKelvey, K.S., Young, M.K., Schwartz, M.K., Lowe, W.H., Letcher, B.H., Whiteley, A.R., 2015. Distance, flow and PCR inhibition: eDNA dynamics in two headwater streams. Mol. Ecol. Resour. 15, 216–227. https://doi.org/10.1111/1755-0998.12285

Jones, B.M., Arp, C.D., Hinkel, K.M., Beck, R.A., Schmutz, J.A., Winston, B., 2009. Arctic Lake Physical Processes and Regimes with Implications for Winter Water Availability and Management in the National Petroleum Reserve Alaska. Environ. Manage. 43, 1071–1084. https://doi.org/10.1007/s00267-008-9241-0

Kittel, T.G.F., Baker, B.B., Higgins, J. V, Haney, J.C., 2011. Climate vulnerability of ecosystems and landscapes on Alaska’s North Slope. Reg. Environ. Chang. 11, 249–264. https://doi.org/10.1007/s10113-010-0180-y

Lalkhen, A.G., McCluskey, A., 2008. Clinical tests: sensitivity and specificity. Contin. Educ. Anaesth. Crit. Care Pain 8, 221–223. https://doi.org/10.1093/bjaceaccp/mkn041

Lee, David S., Survey., N.C.B., of Natural History., N.C.S.M., 1980. Atlas of North American freshwater fishes, 1980- et seq. /. [Raleigh, N.C.] :North Carolina State Museum of Natural History,.

Liaw, A., Wiener, M., others, 2002. Classification and regression by randomForest. R news 2.

Lisi, P.J., Schindler, D.E., Cline, T.J., Scheuerell, M.D., Walsh, P.B., 2015. Watershed geomorphology and snowmelt control stream thermal sensitivity to air temperature. Geophys. Res. Lett. 42, 3380–3388. https://doi.org/10.1002/2015GL064083

Mainali, K.P., Warren, D.L., Dhileepan, K., McConnachie, A., Strathie, L., Hassan, G., Karki, D., Shrestha, B.B., Parmesan, C., 2015. Projecting future expansion of invasive species: comparing and improving methodologies for species distribution modeling. Glob. Chang. Biol. 21, 4464–4480. https://doi.org/10.1111/gcb.13038

McNyset, K.M., 2005. Use of ecological niche modelling to predict distributions of freshwater fish species in Kansas. Ecol. Freshw. Fish 14, 243–255. https://doi.org/10.1111/j.1600-0633.2005.00101.x

McPhail, J.D., Lindsey, C.C., 1970. Freshwater fishes of northwestern Canada and Alaska, Bulletin (Fisheries Research Board of Canada). Fisheries Research Board of Canada : available by mail from the Queen’s Printer.

McPhail, J.D., Paragamian, V.L., 2000. Burbot biology and life history. Trans. Am. Fish. Soc. 1.

Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M.N., Kawabata, Z., 2012. Surveillance of fish species composition using environmental DNA. Limnology 13, 193–197. https://doi.org/10.1007/s10201-011-0362-4

Nathan, L.M., Simmons, M., Wegleitner, B.J., Jerde, C.L., Mahon, A.R., 2014. Quantifying Environmental DNA Signals for Aquatic Invasive Species Across Multiple Detection Platforms. Environ. Sci. Technol. 48, 12800–12806. https://doi.org/10.1021/es5034052

Pacifici, M., Foden, W.B., Visconti, P., Watson, J.E.M., Butchart, S.H.M., Kovacs, K.M., Scheffers, B.R., Hole, D.G., Martin, T.G., Akçakaya, H.R., Corlett, R.T., Huntley, B., Bickford, D., Carr, J.A., Hoffmann, A.A., Midgley, G.F., Pearce-Kelly, P., Pearson, R.G., Williams, S.E., Willis, S.G., Young, B., Rondinini, C., 2015. Assessing species vulnerability to climate change. Nat. Clim. Chang. 5, 215.

Pearson, R., Terence, T.P., Liu, C., 2004. Modelling species distributions in Britain: A hierarchical integration of climate and land-cover data. Ecography (Cop.). 27, 285–298. https://doi.org/10.1111/j.0906-7590.2004.03740.x

Perez, C., Bonar, S., Amberg, J., Ladell, B., Rees, C., T. Stewart, W., J. Gill, C., Cantrell, C., T. Robinson, A., 2017. Comparison of American Fisheries Society (AFS) Standard Fish Sampling Techniques and Environmental DNA for Characterizing Fish Communities in a Large Reservoir. North Am. J. Fish. Manag. 37, 1010–1027. https://doi.org/10.1080/02755947.2017.1342721

Porfirio, L.L., Harris, R.M.B., Lefroy, E.C., Hugh, S., Gould, S.F., Lee, G., Bindoff, N.L., Mackey, B., 2014. Improving the use of species distribution models in conservation planning and management under climate change. PLoS One 9, e113749–e113749. https://doi.org/10.1371/journal.pone.0113749

Portt, C.B., Coker, G.A., Ming, D.L., Randall, R.G., 2006. A review of fish sampling methods commonly used in Canadian freshwater habitats. Can. Tech. Rep. Fish. Aquat. Sci.

Radinger, J., Wolter, C., 2014. Patterns and predictors of fish dispersal in rivers. Fish Fish. 15, 456–473. https://doi.org/10.1111/faf.12028

Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R.M., Gough, K.C., 2014. REVIEW: The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. J. Appl. Ecol. 51, 1450–1459. https://doi.org/10.1111/1365-2664.12306

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., Toll, D., 2004. The Global Land Data Assimilation System. Bull. Am. Meteorol. Soc. 85, 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rosenfeld, J., 2003. Assessing the Habitat Requirements of Stream Fishes: An Overview and Evaluation of Different Approaches. Trans. Am. Fish. Soc. 132, 953–968. https://doi.org/10.1577/T01-126

Strickler, K.M., Fremier, A.K., Goldberg, C.S., 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. Biol. Conserv. 183, 85–92. https://doi.org/https://doi.org/10.1016/j.biocon.2014.11.038

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., Kawabata, Z., 2012. Estimation of fish biomass using environmental DNA. PLoS One 7, e35868. https://doi.org/10.1371/journal.pone.0035868

Terrado, M., Sabater, S., Chaplin-Kramer, B., Mandle, L., Ziv, G., Acuña, V., 2016. Model development for the assessment of terrestrial and aquatic habitat quality in conservation planning. Sci. Total Environ. 540, 63–70. https://doi.org/https://doi.org/10.1016/j.scitotenv.2015.03.064

Thomsen, P.F., Kielgast, J.O.S., Iversen, L.L., Wiuf, C., Rasmussen, M., Gilbert, M.T.P., Orlando, L., Willerslev, E., 2012. Monitoring endangered freshwater biodiversity using environmental DNA. Mol. Ecol. 21, 2565–2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x

Thuiller, W., Brotons, L., Araújo, M.B., Lavorel, S., 2004. Effects of restricting environmental range of data to project current and future species distributions. Ecography (Cop.). 27, 165–172. https://doi.org/10.1111/j.0906-7590.2004.03673.x

Treglia, M.L., Fisher, R.N., Fitzgerald, L.A., 2015. Integrating multiple distribution models to guide conservation efforts of an endangered toad. PLoS One 10, 1–18. https://doi.org/10.1371/journal.pone.0131628

Vezza, P., Muñoz-Mas, R., Martinez-Capel, F., Mouton, A., 2015. Random forests to evaluate biotic interactions in fish distribution models. Environ. Model. Softw. 67, 173–183. https://doi.org/https://doi.org/10.1016/j.envsoft.2015.01.005

Watanabe, M., Suzuki, T., O'ishi, R., Komuro, Y., Watanabe, S., Emori, S., Takemura, T., Chikira, M., Ogura, T., Sekiguchi, M., Takata, K., Yamazaki, D., Yokohata, T., Nozawa, T., Hasumi, H., Tatebe, H., Kimoto, M., 2010. Improved Climate Simulation by MIROC5: Mean States, Variability, and Climate Sensitivity. J. Clim. 23, 6312–6335. https://doi.org/10.1175/2010JCLI3679.1

Wilcox, T.M., McKelvey, K.S., Young, M.K., Jane, S.F., Lowe, W.H., Whiteley, A.R., Schwartz, M.K., 2013. Robust detection of rare species using environmental DNA: the importance of primer specificity. PLoS One 8, e59520–e59520. https://doi.org/10.1371/journal.pone.0059520

Wolfe, R.J., Walker, R.J., 1987. Subsistence Economies in Alaska: Productivity, Geography, and Development Impacts. Arctic Anthropol. 24, 56–81.

Wu, H., Kimball, J.S., Elsner, M.M., Mantua, N., Adler, R.F., Stanford, J., 2012. Projected climate change impacts on the hydrology and temperature of Pacific Northwest rivers. Water Resour. Res. 48. https://doi.org/10.1029/2012WR012082

APPENDICES

APPENDIX A

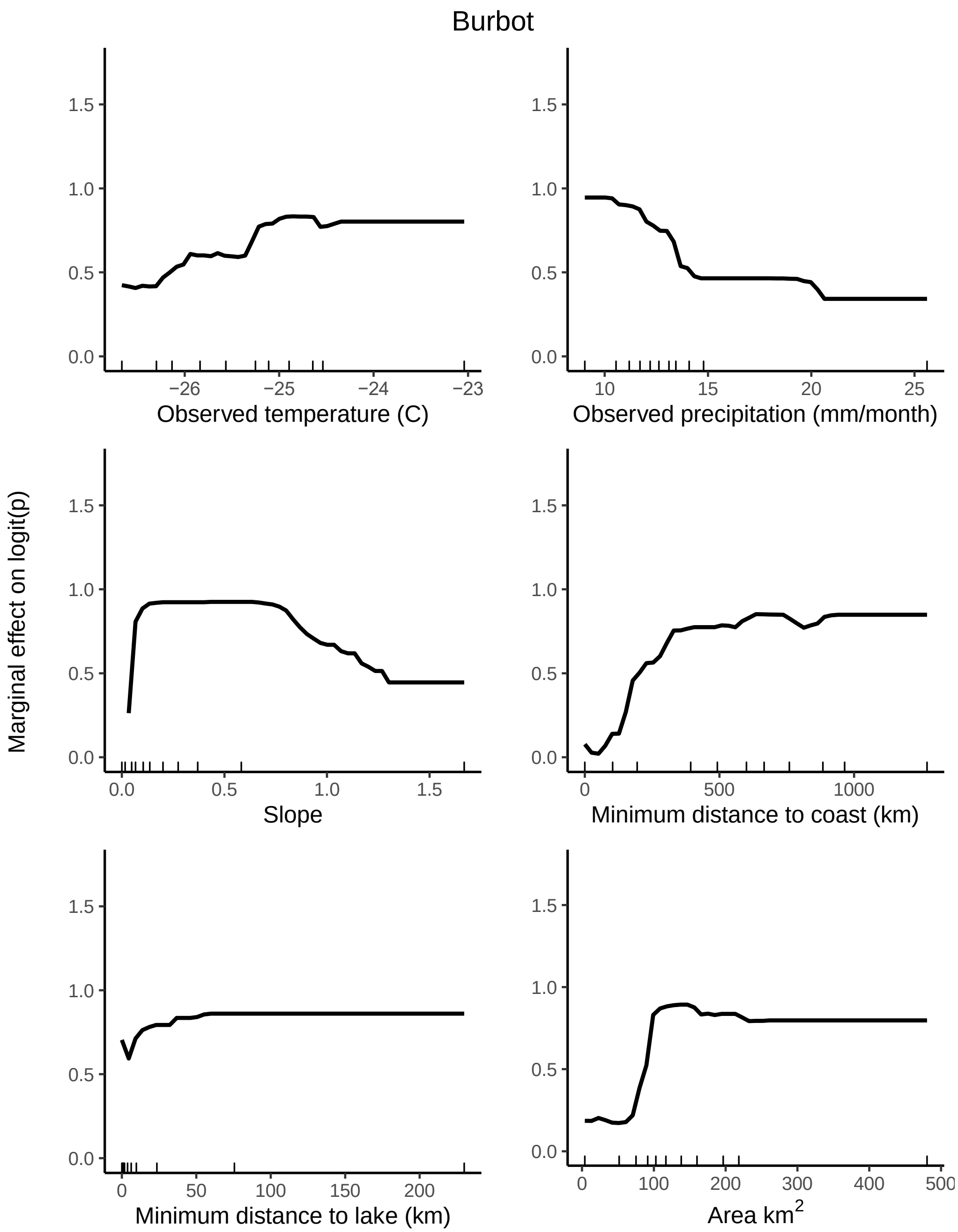


Figure A1: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Burbot. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

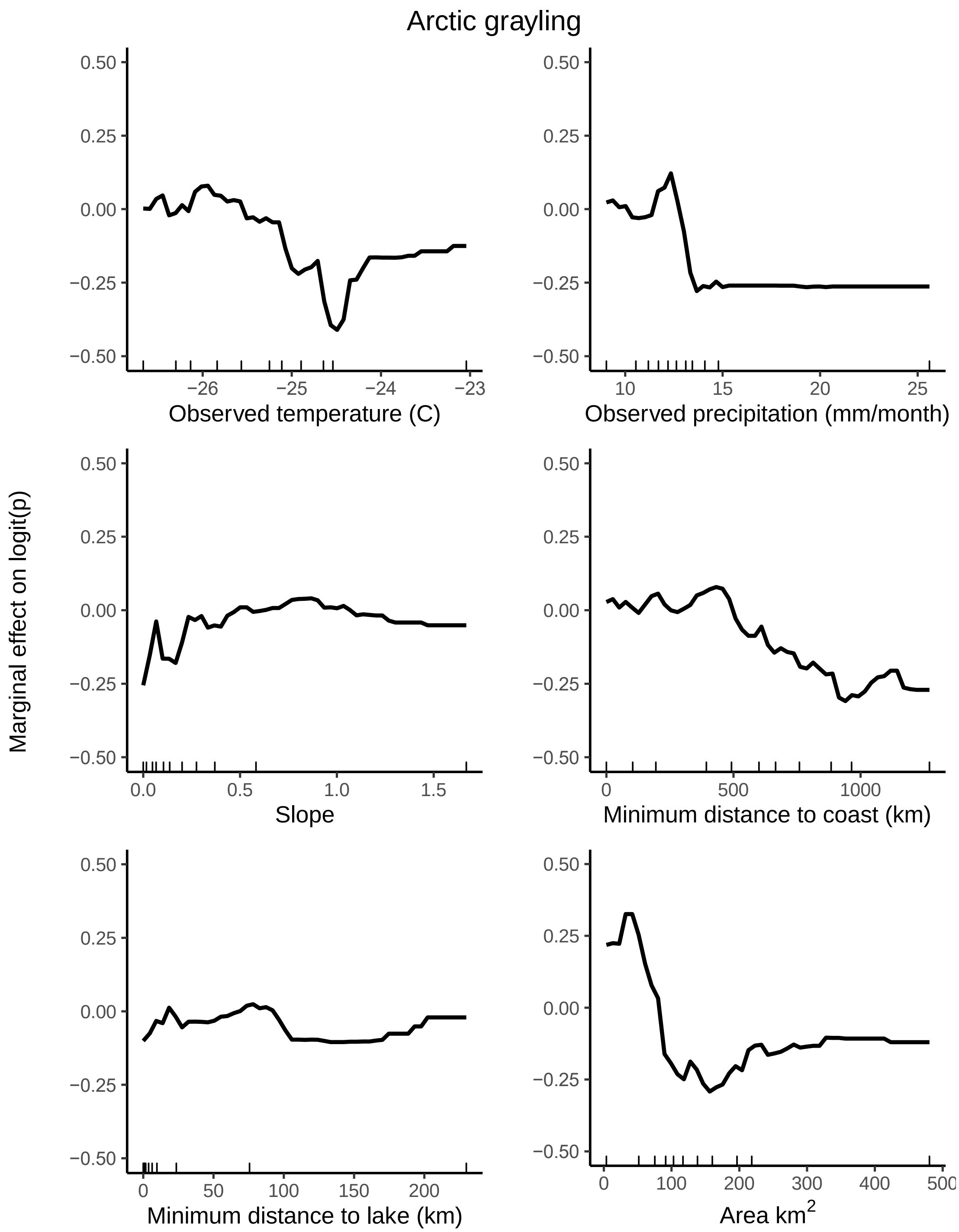


Figure A2: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Arctic Grayling. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

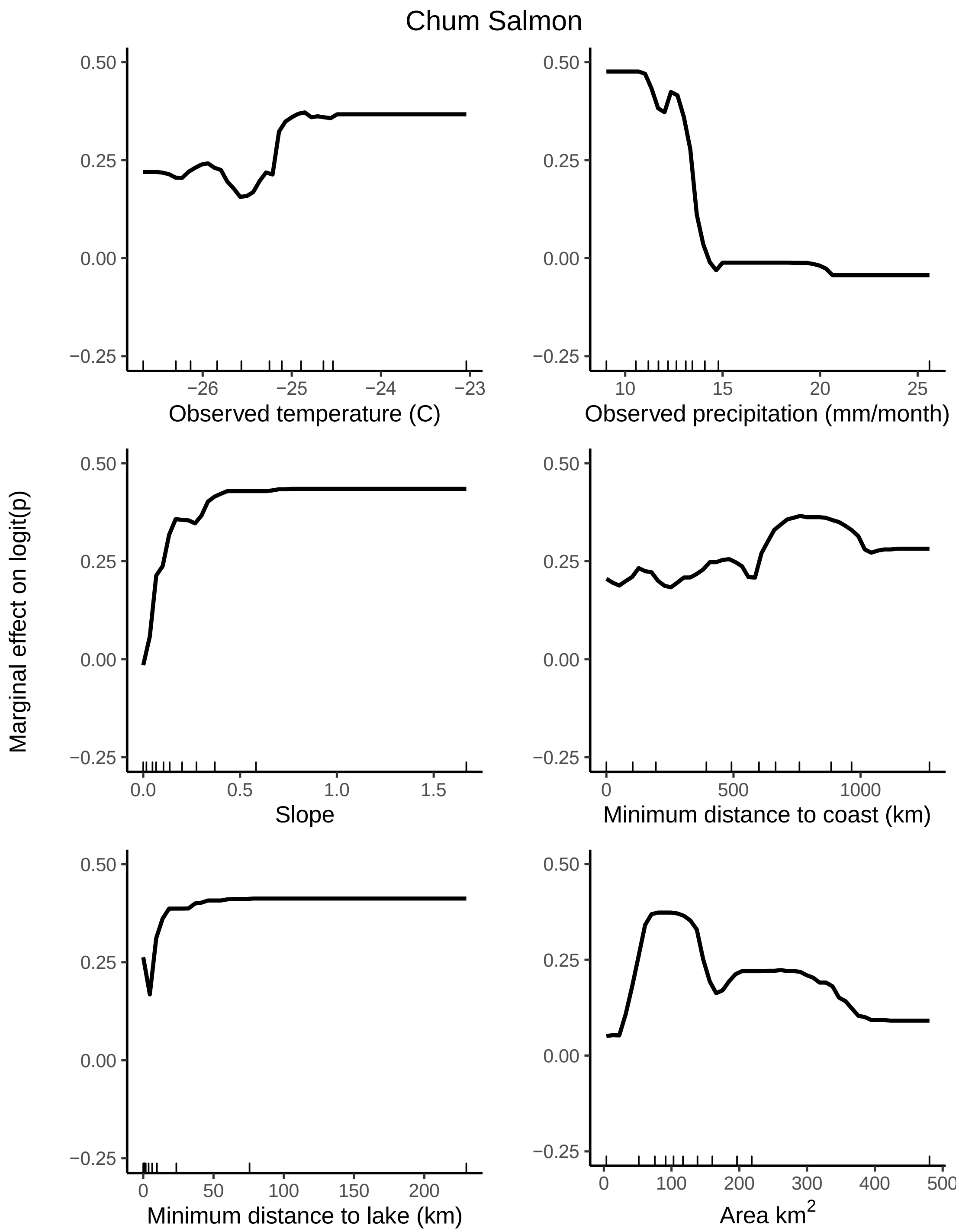


Figure A3: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Chum Salmon. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

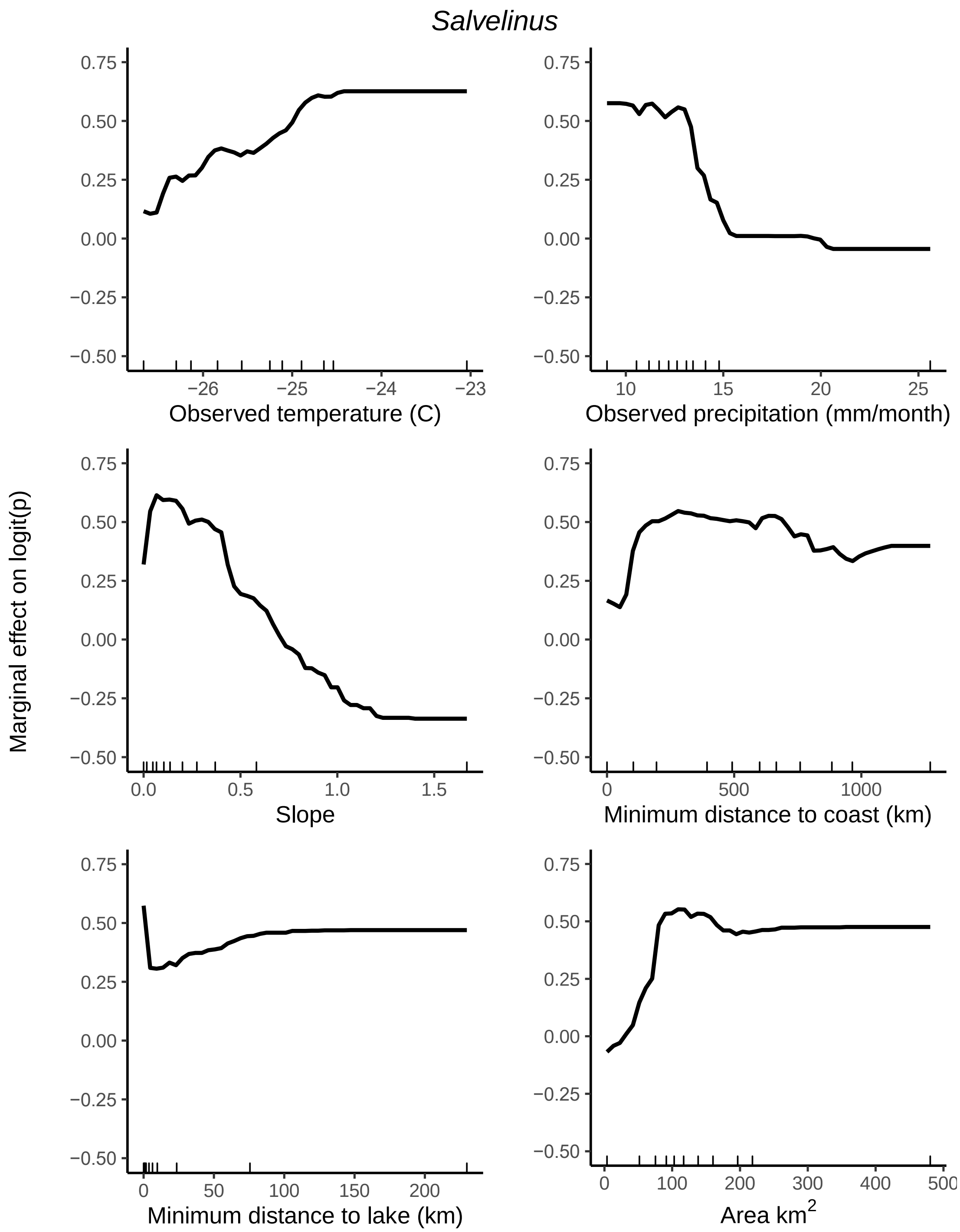


Figure A4: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of *Salvelinus*. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

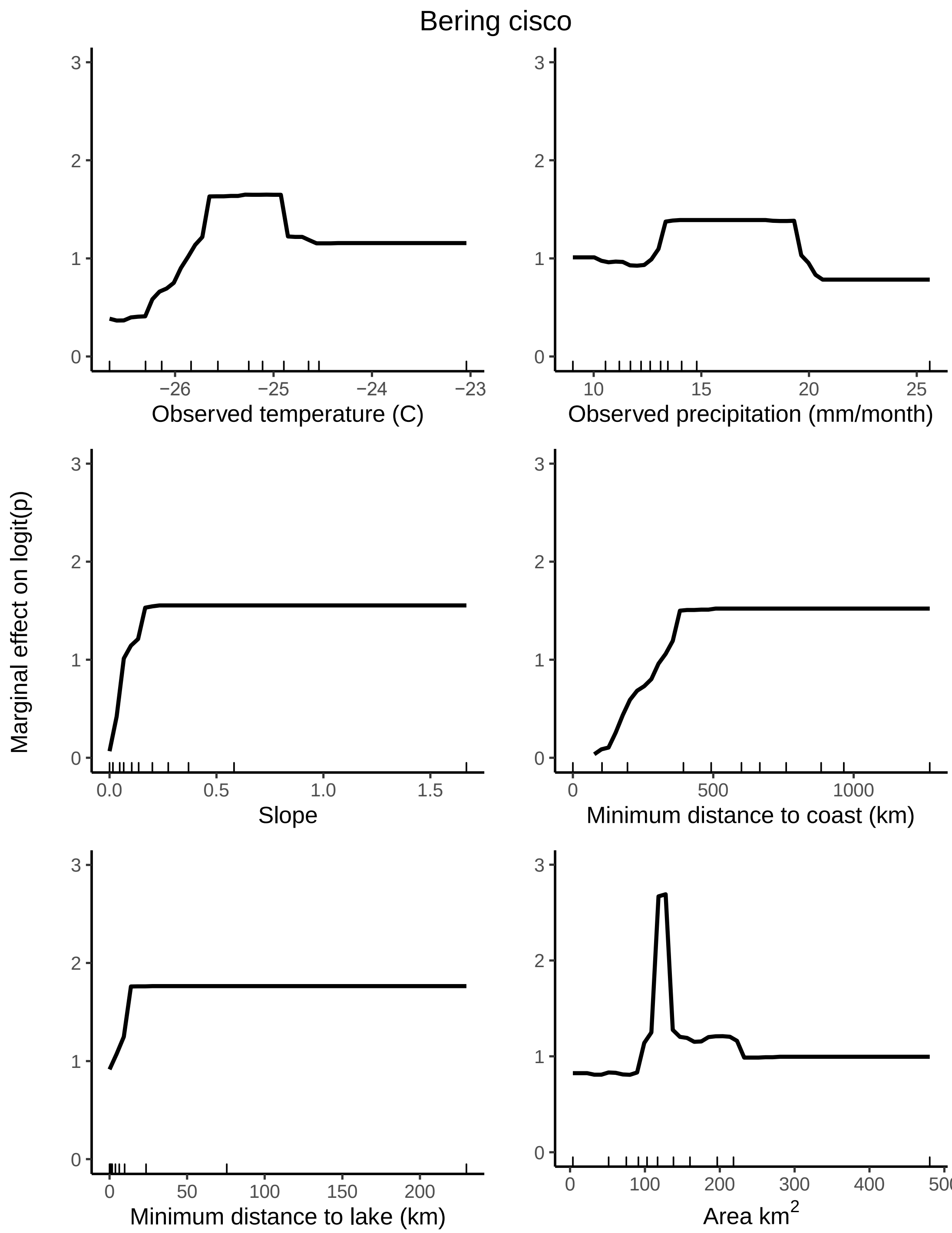


Figure A5: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Bering Cisco. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

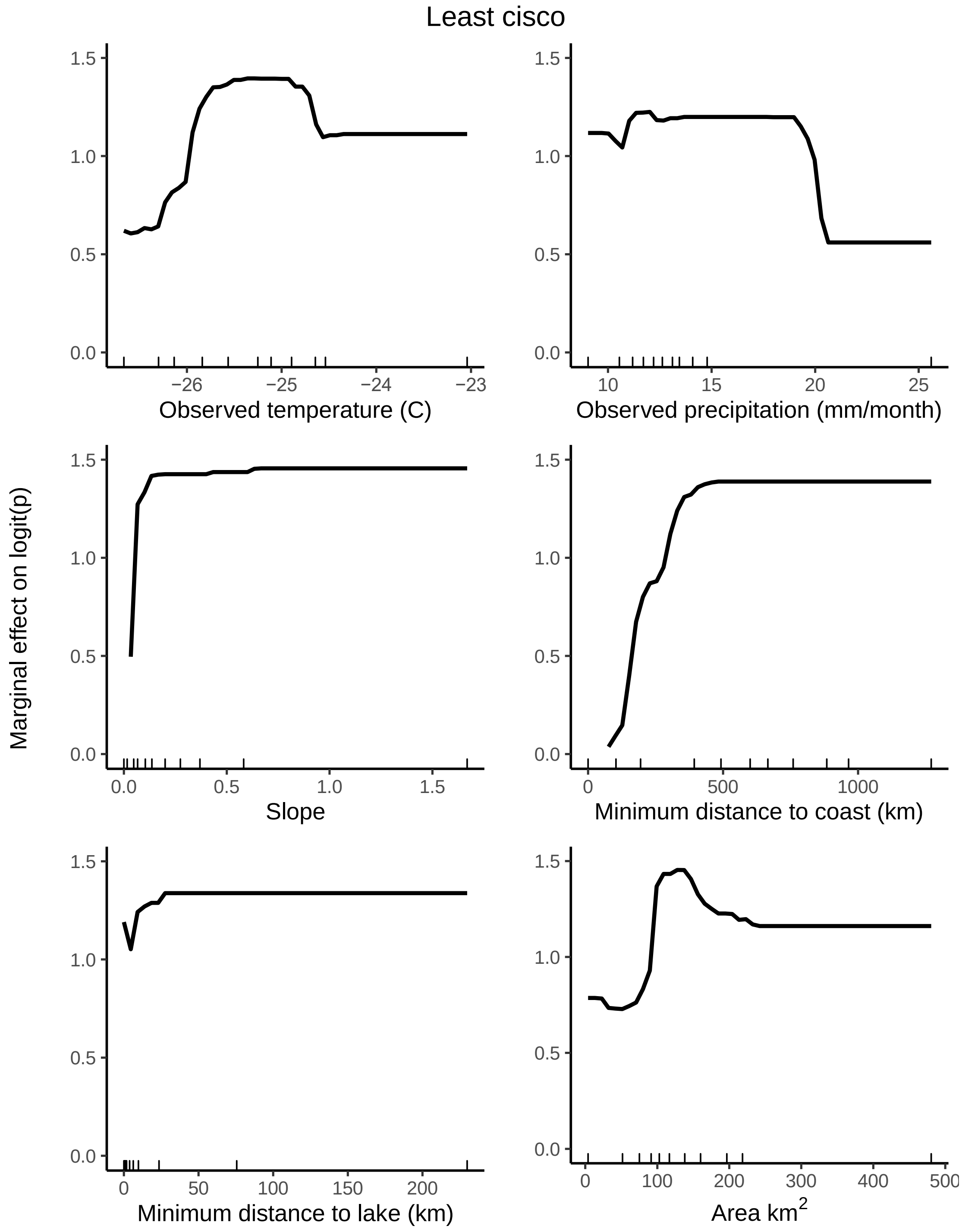


Figure A6: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Least Cisco. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

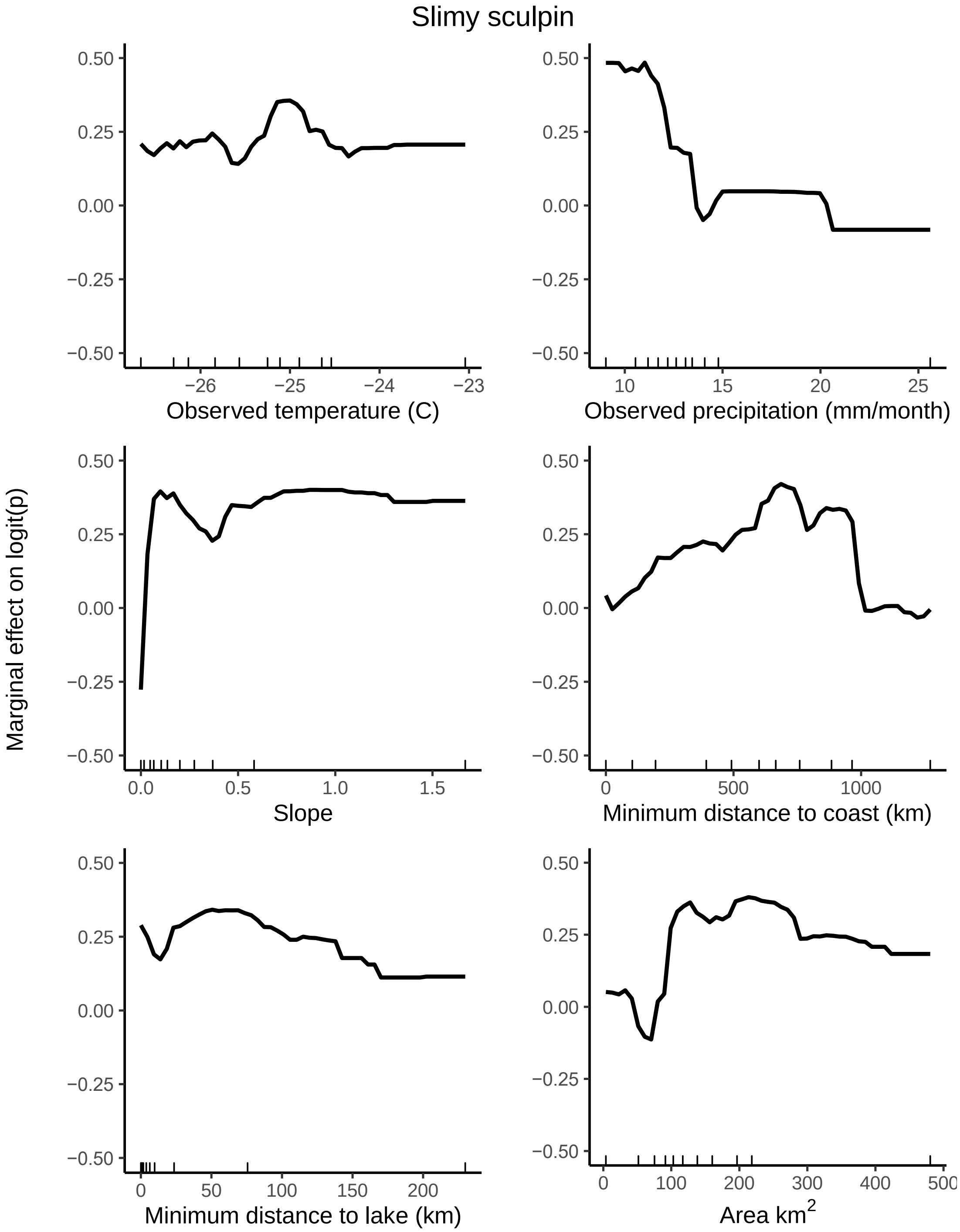


Figure A7: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Slimy Sculpin. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

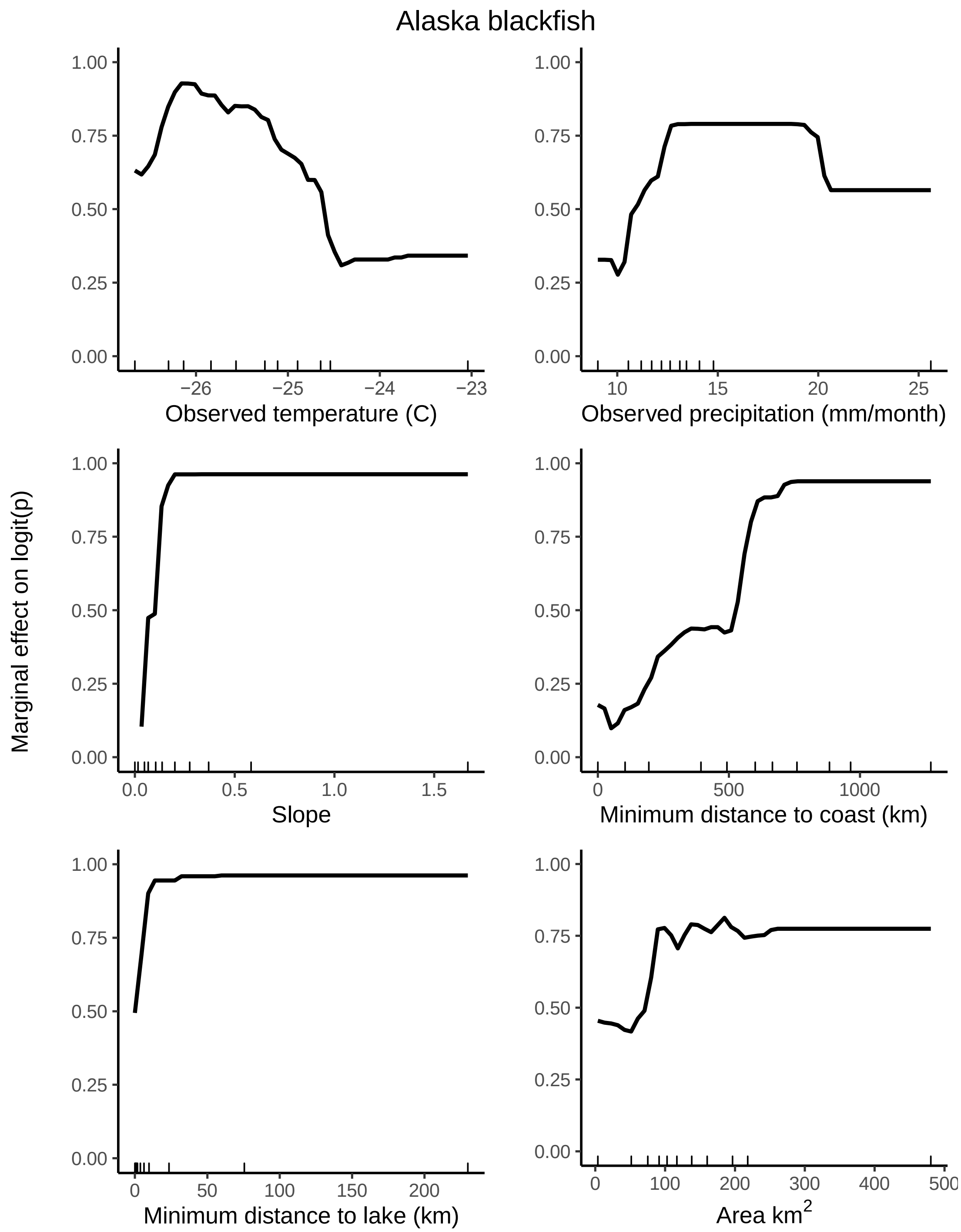


Figure A8: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Alaska Blackfish. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

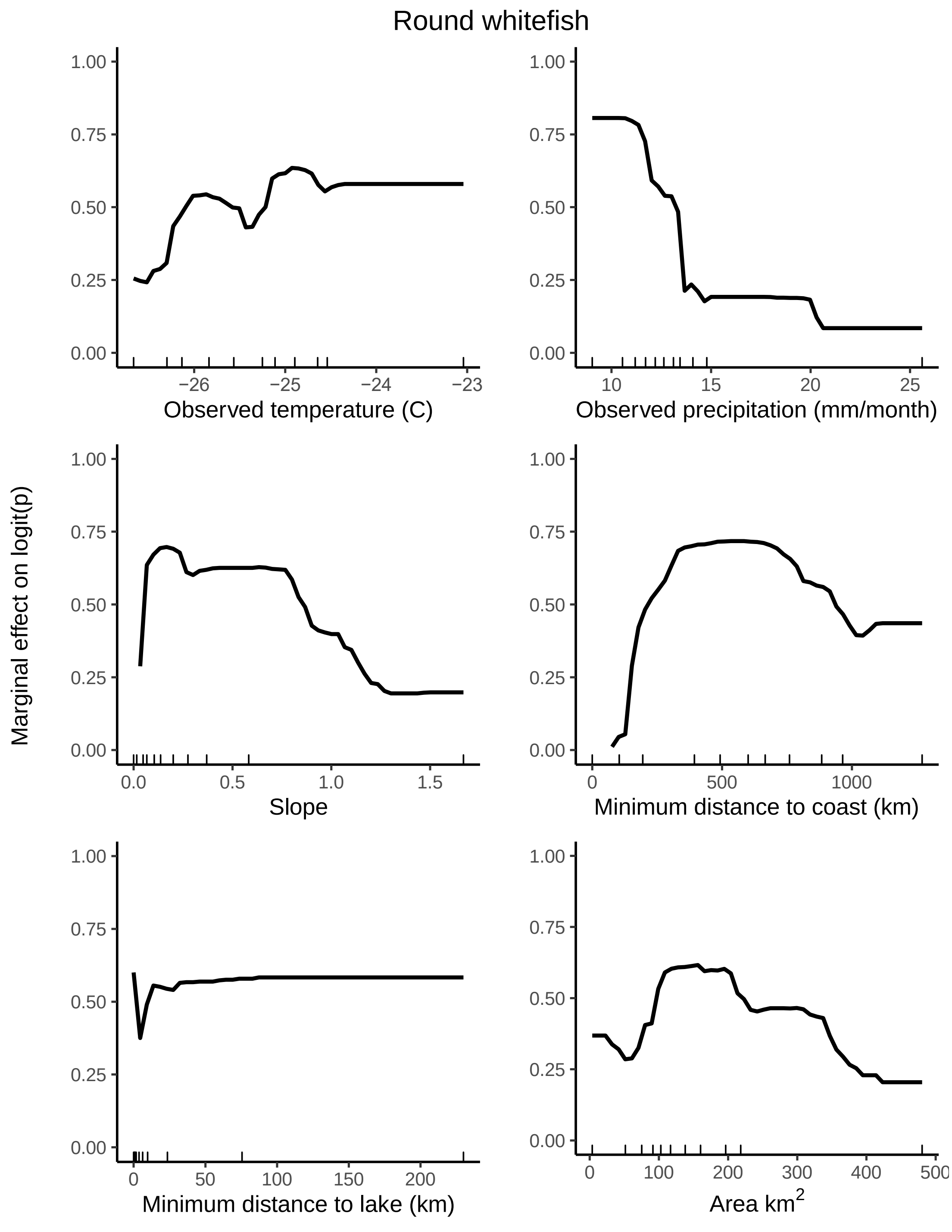


Figure A9: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Round Whitefish. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

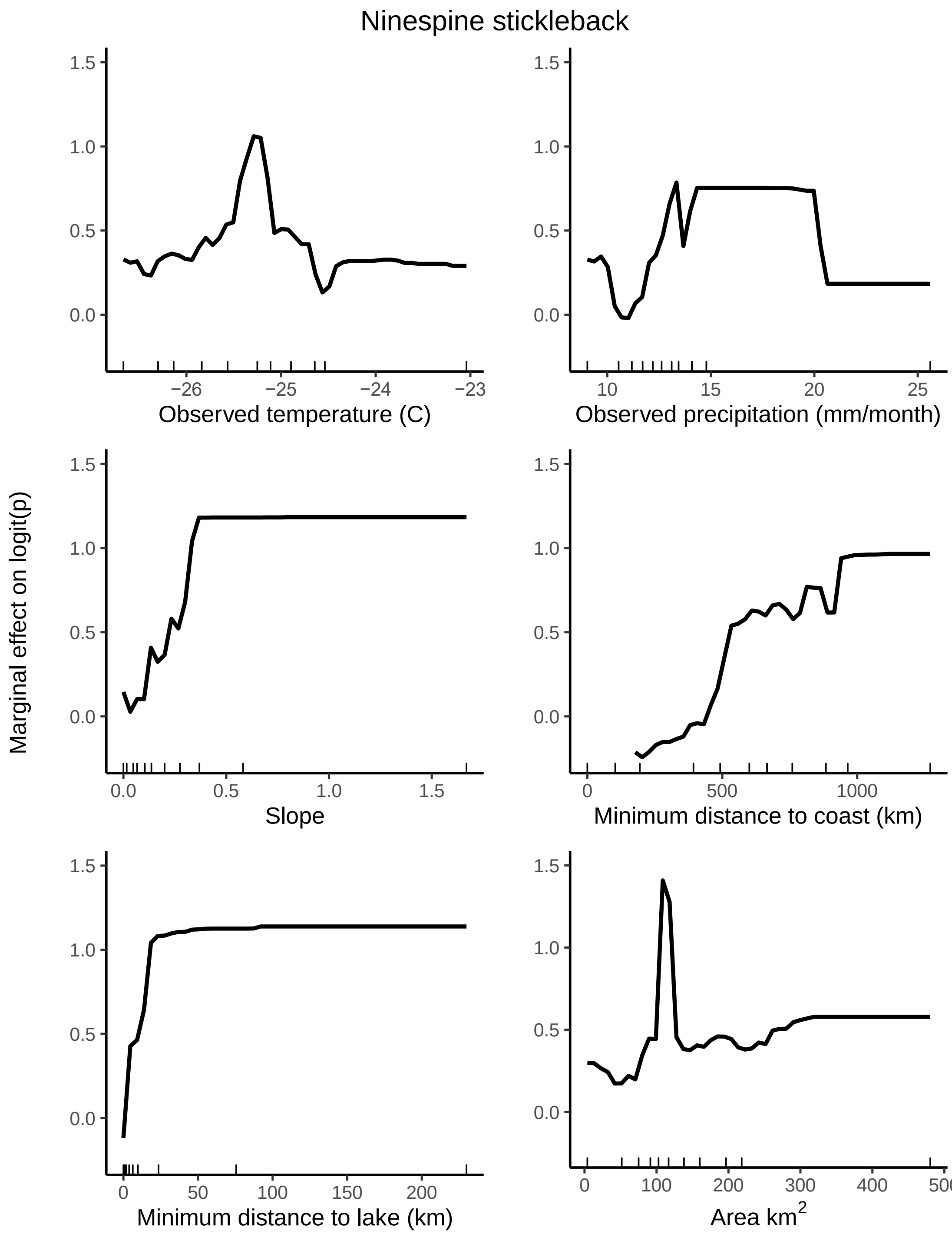


Figure A10: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Ninespine Stickleback. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

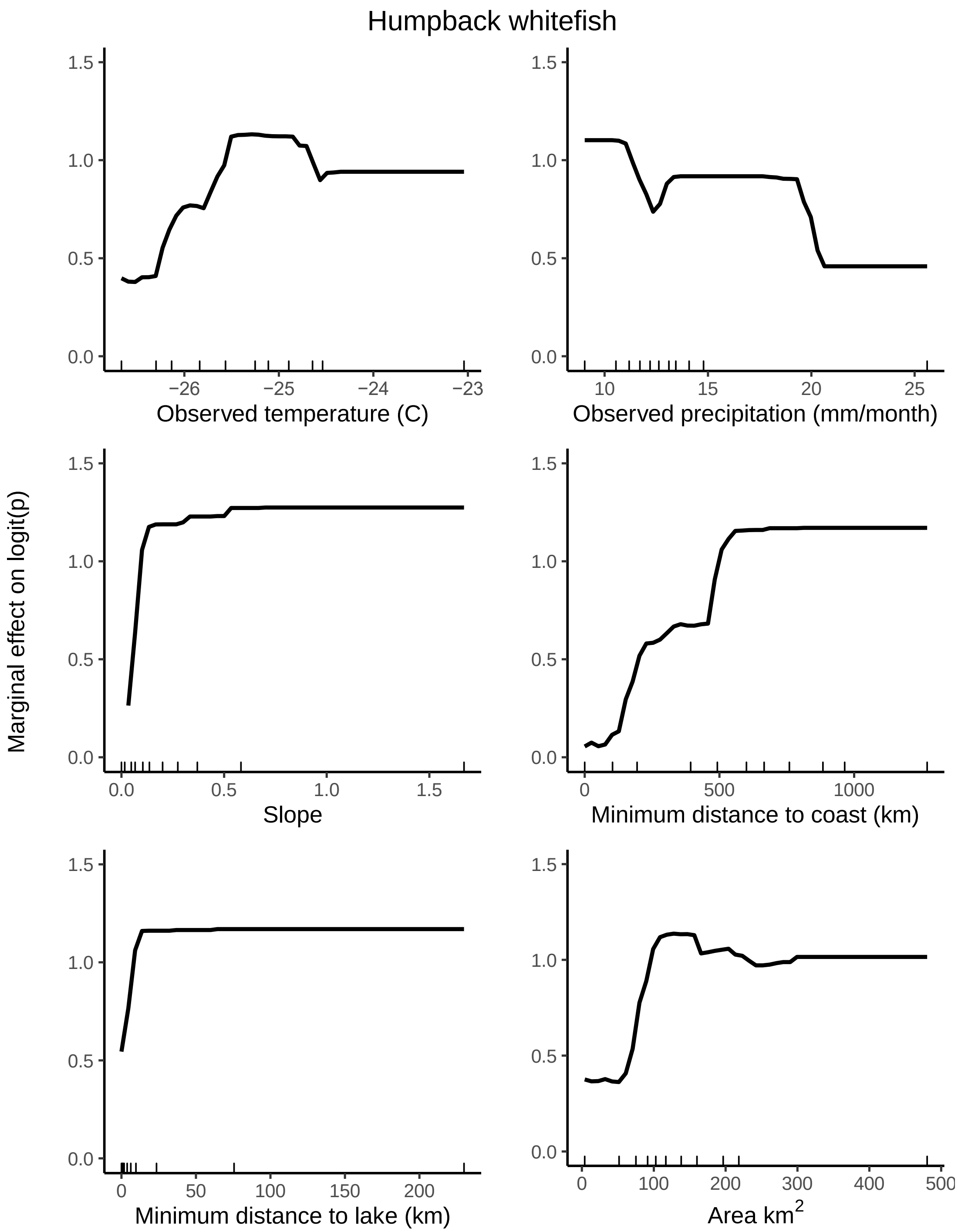


Figure A11: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Humpback Whitefish. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

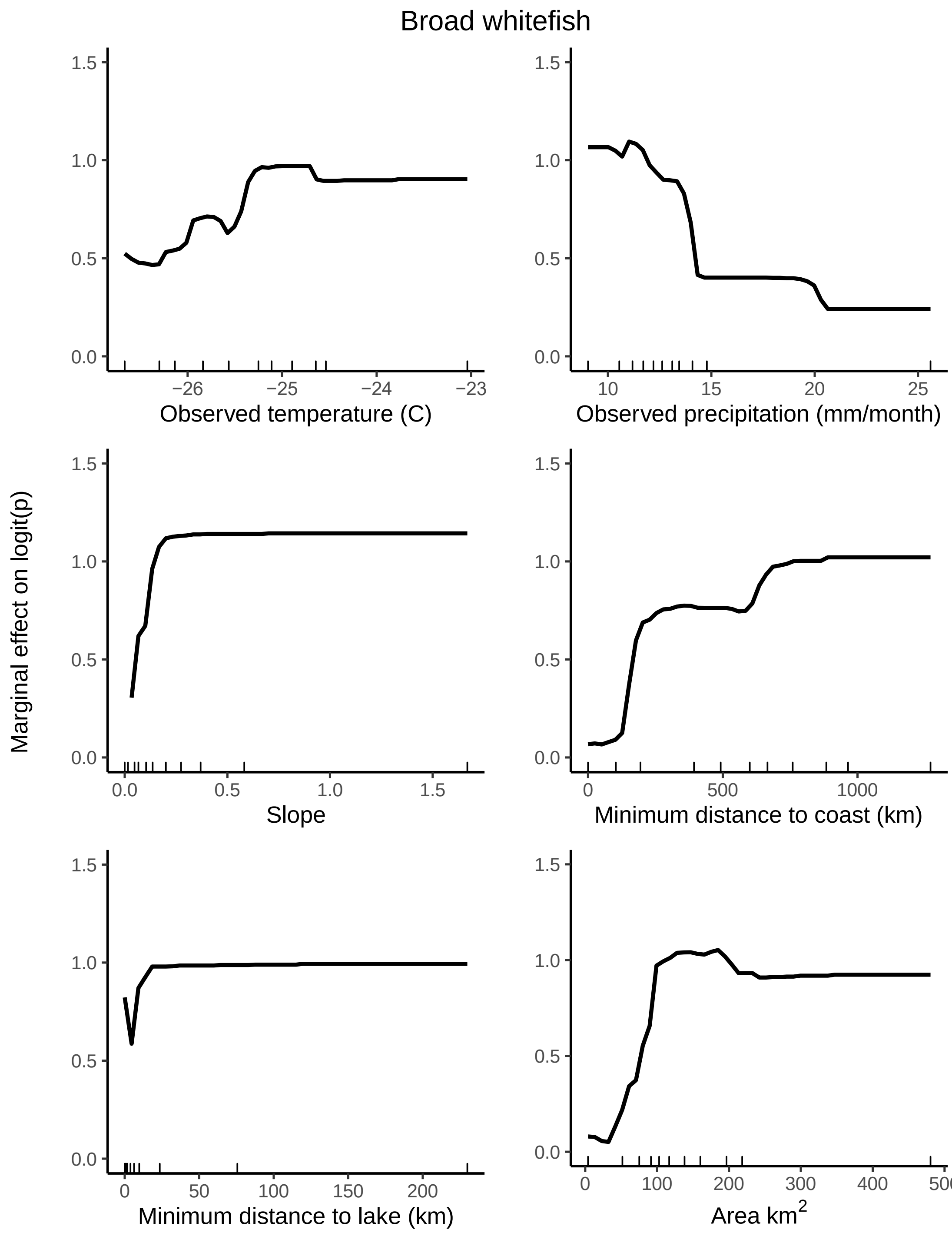


Figure A12: Partial dependence plots (PDPs) for the six predictor variables used in the species distribution model of Broad Whitefish. Observed temperature is the average observed annual air temperature from 1981-2010. Observed precipitation is the observed monthly rainfall in the same time period. Slope describes the change in elevation along the segment and the segment length. Minimum distance to coast and lake describe swimming distance to lake or ocean refugia, and Area describes the watershed area. The rug along the X-axis denotes deciles within the sample population.

APPENDIX B

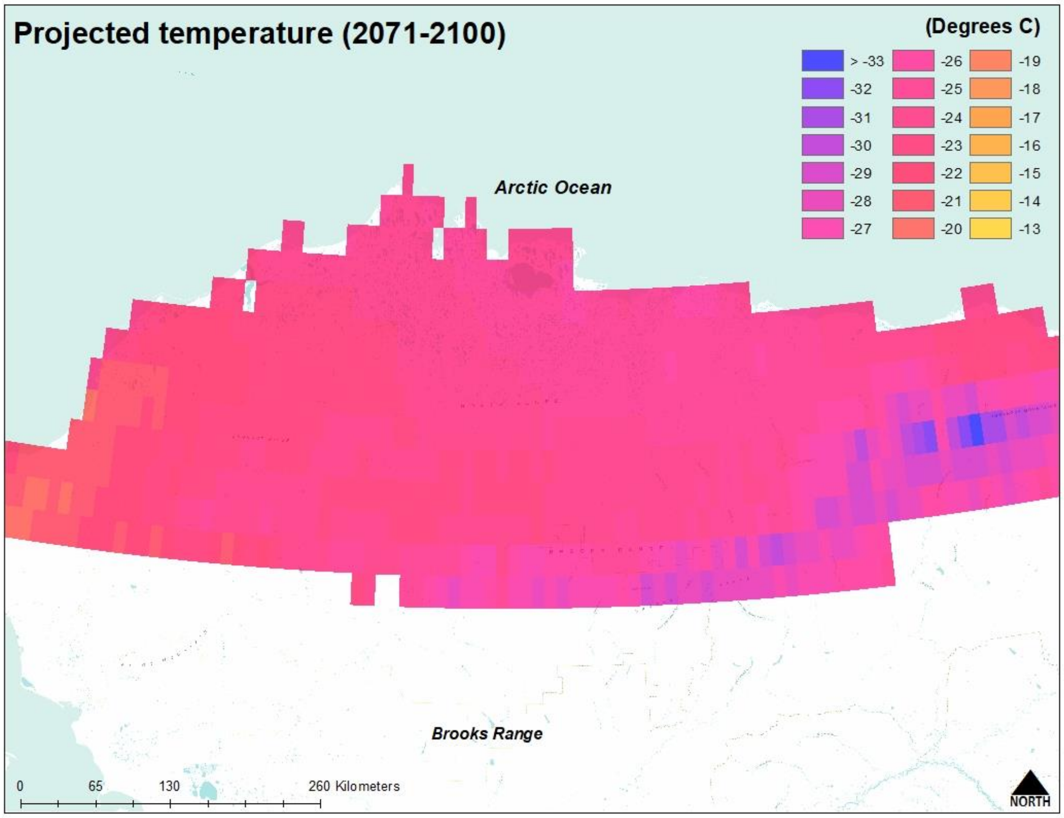


Figure B1: Projected temperature (GFDL 2 low emission scenario) in Alaska's North Slope.

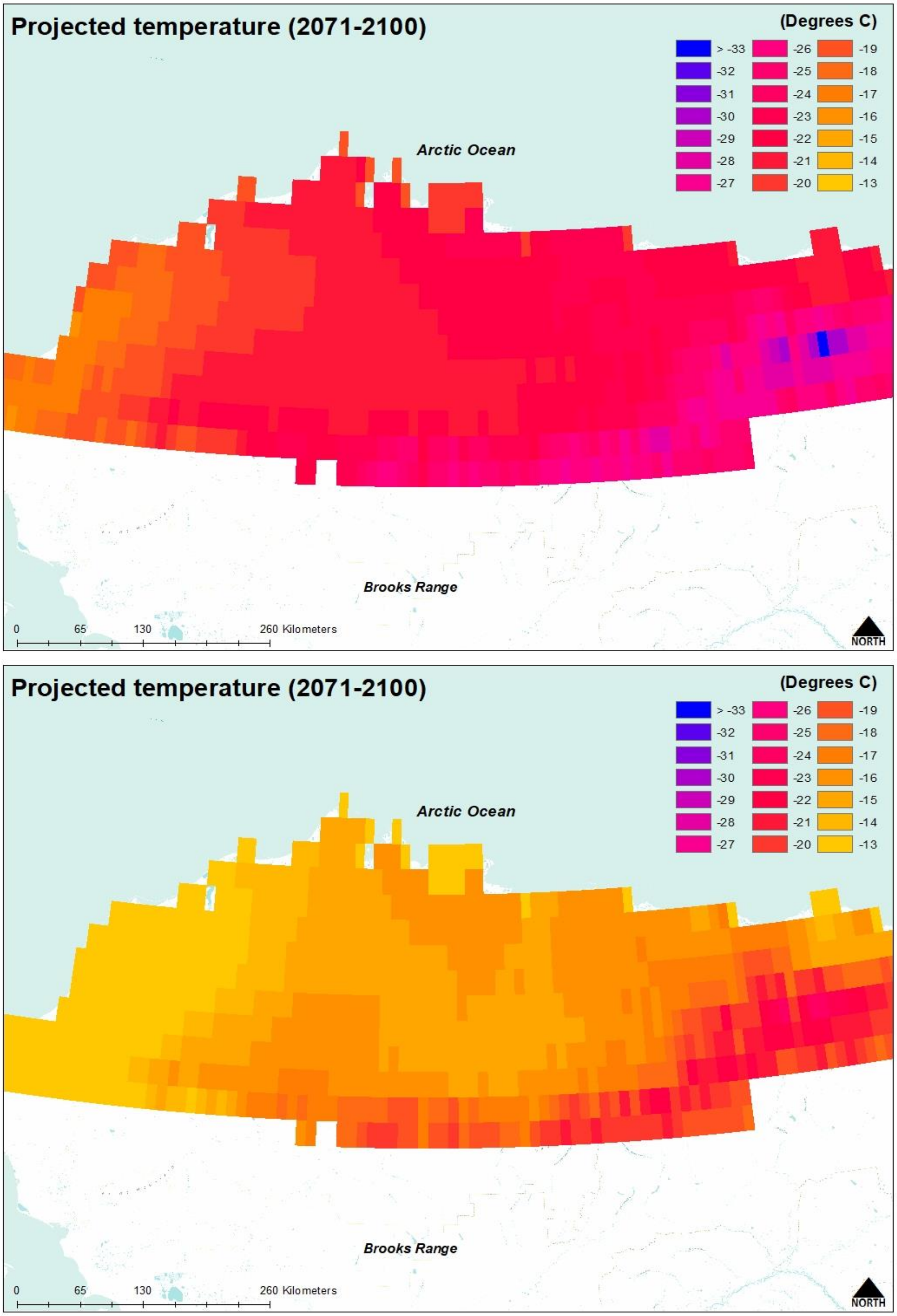


Figure B2: Projected low emission (top) and high emission (bottom) temperatures (IPSL) in Alaska's North Slope.

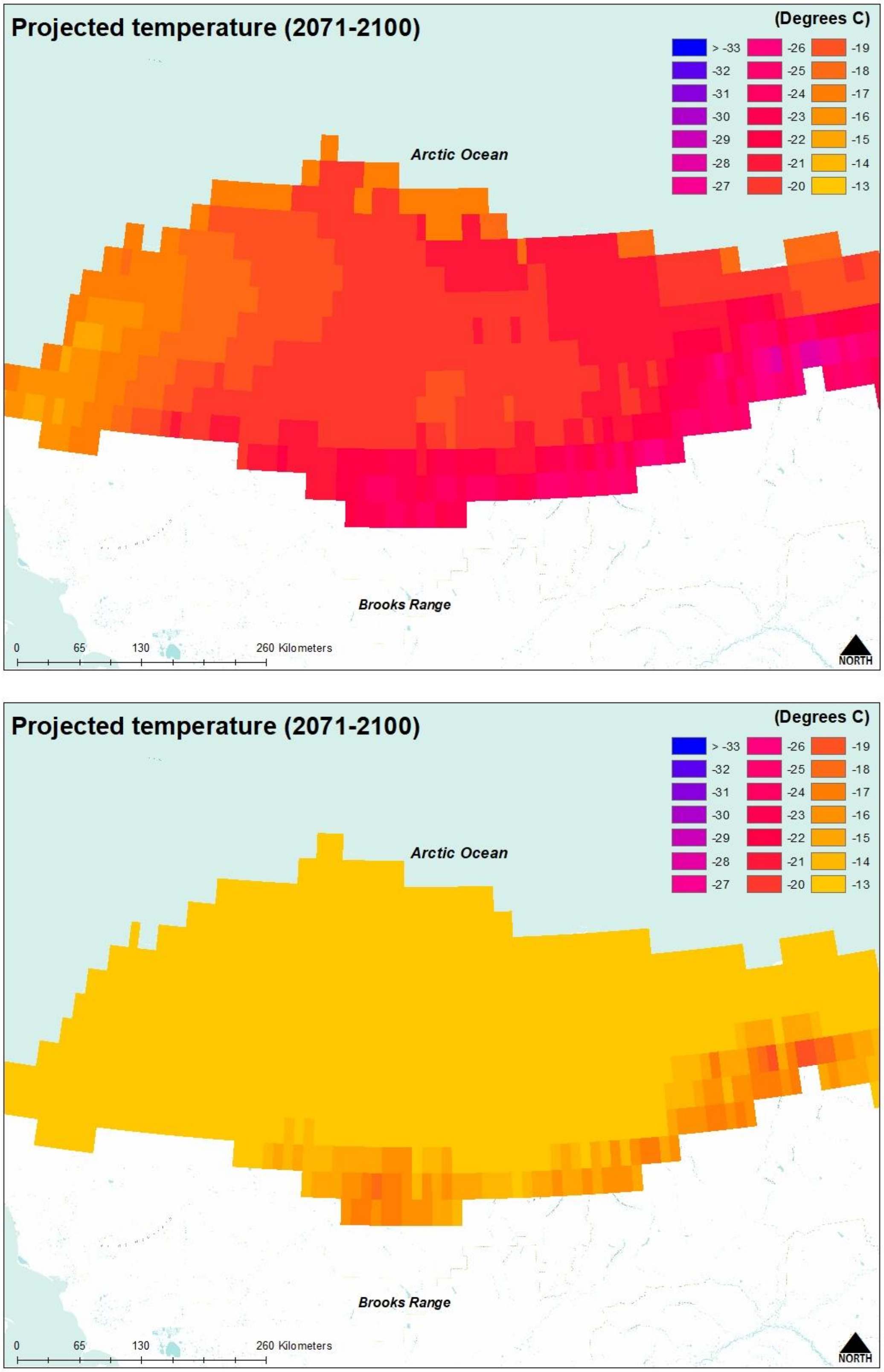


Figure B3: Projected low emission (top) and high emission (bottom) temperatures (MIROC) in Alaska's North Slope.

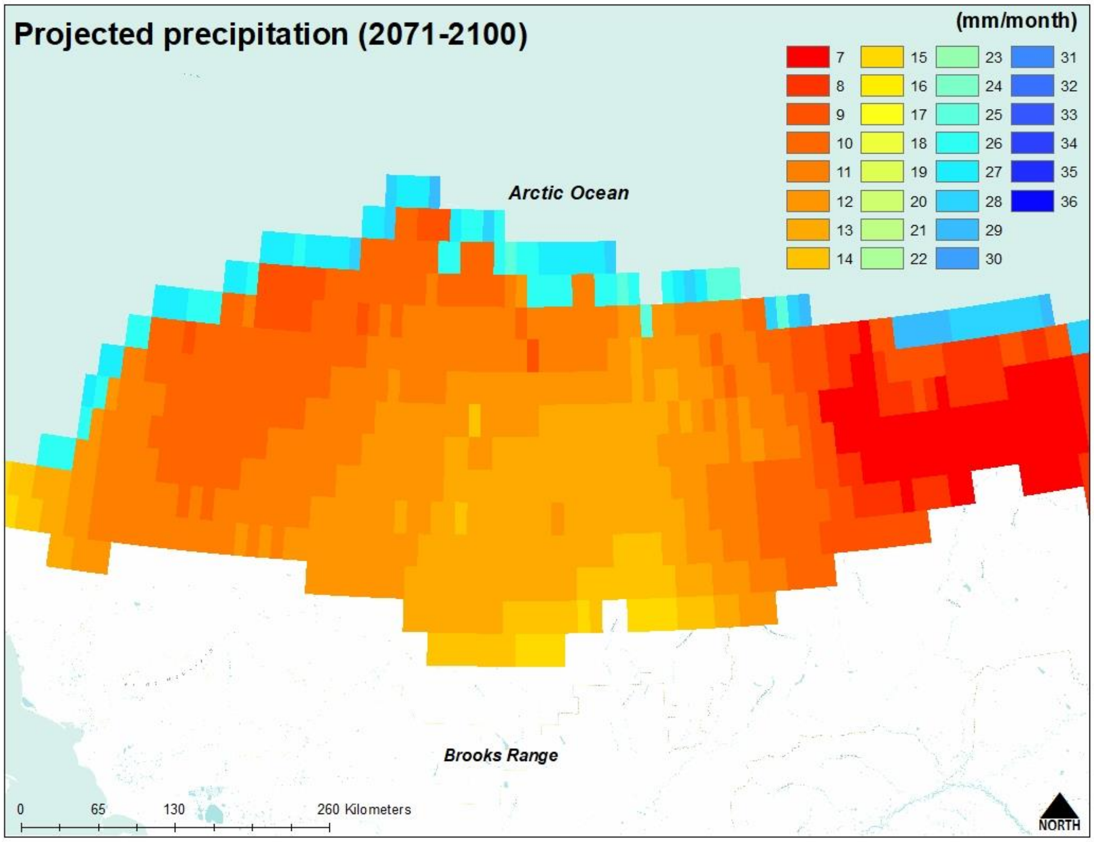
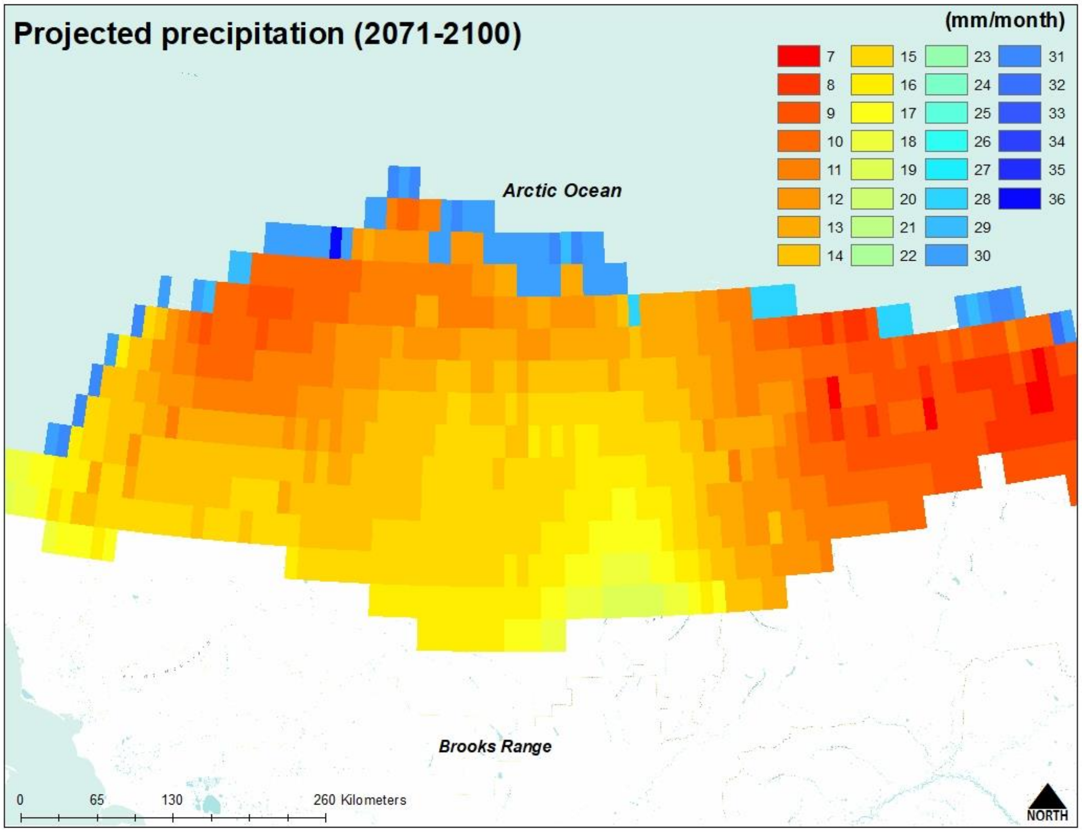


Figure B4: Projected precipitation (GFDL2 low emission scenario) in Alaska's North Slope.



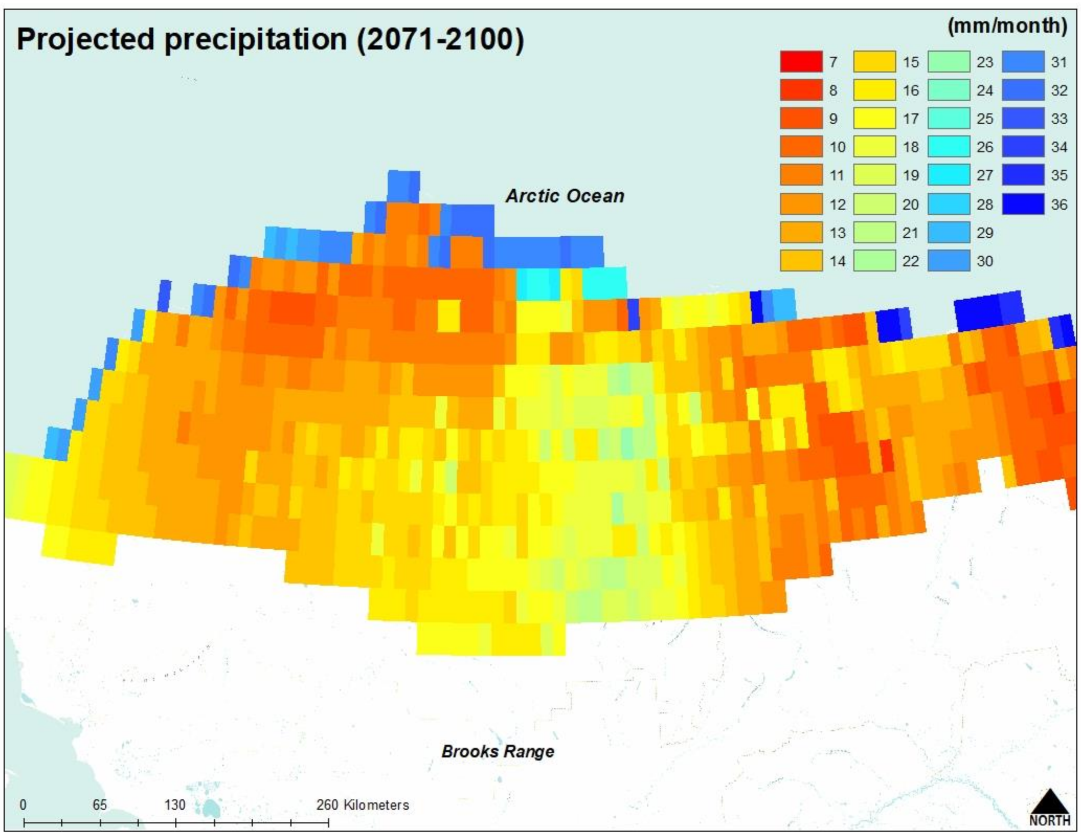


Figure B5: Projected low emission (top) and high emission (bottom) precipitation (IPSL) in Alaska’s North Slope.

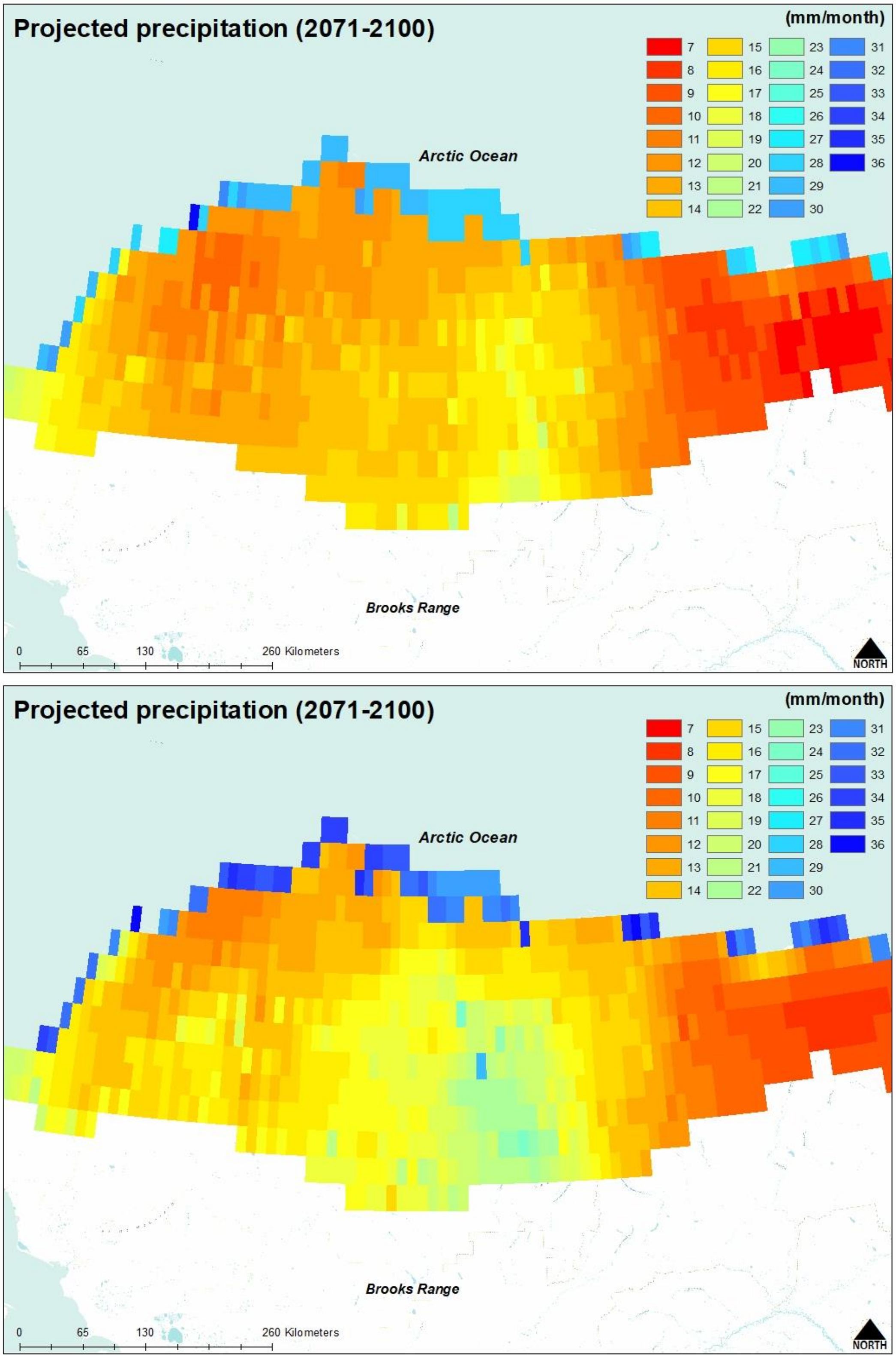


Figure B6: Projected low emission (top) and high emission (bottom) precipitation (MIROC) in Alaska's North Slope.

APPENDIX C

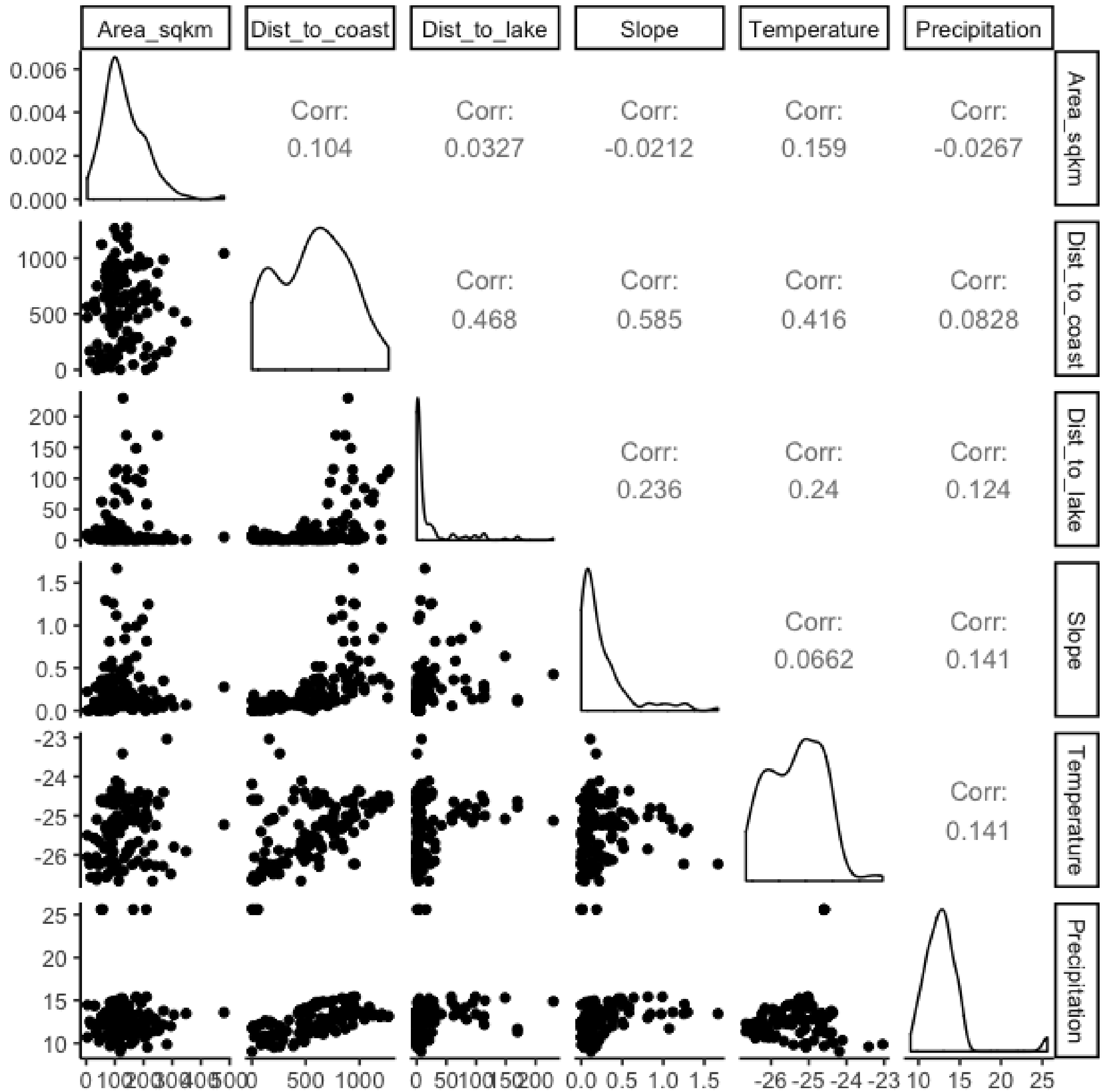


Figure C: Scatterplot matrix (SPLOM) for the six environmental predictors used in building species distribution models for fish species in Alaska's North Slope.

APPENDIX D

Table D. All (83) of the candidate remotely sensed environmental predictor variables with minimum, maximum, and the range of values.

Range	Maximum	Minimum	Predictor	Units
480	480	0	AreaSqKM	km^2
63357	65535	2178	ET_Cat	$kg/m^2/8day$
63354	65535	2181	ET_Ws	$kg/m^2/8day$
19054	22775	3721	EVI_Cat	EVI
18935	22775	3840	EVI_Ws	EVI
0	0	0	Fire_Cat	km^2
0.47	0.47	0	FireMax_Cat	km^2
0.55	0.55	0	FireSum_Cat	km^2
0	0	0	Fire_Ws	km^2
0.47	0.47	0	FireMax_Ws	km^2
0.55	0.55	0	FireSum_Ws	km^2
32730	32766	36	GPP_Cat	kg C m^2 /year
32730	32766	36	GPP_Ws	kg C m^2
11096	1097	-9999	WV_Cat	Cm
8134	1097	-7037	WV_WS	Cm
0.27	0.27	0	Alder_Cat	%
0.09	0.09	0	Arcto_Cat	%
0.79	0.79	0	Bare_Cat	%
0.68	0.68	0	Birch_Cat	%
0.97	0.97	0	Burned_Cat	%
0.67	0.67	0	Carex_Cat	%
0.63	0.63	0	Coastmarsh_Cat	%
0.01	0.01	0	Deciduous_Cat	%
0.83	0.83	0	Dwarf_Cat	%
0.66	0.66	0	Dwarfother_Cat	%
1	1	0	IceSnow_Cat	%
0.46	0.46	0	Willow_Cat	%
0.01	0.01	0	Marine_Cat	%
0.81	0.81	0	Mesicsedge_Cat	%
0.32	0.32	0	Mesiherb_Cat	%

Table D (cont.)

0.01	0.01	0	Mixleaf_Cat	%
0.27	0.27	0	Needle_Cat	%
1	1	0	Openwater_Cat	%
0.76	0.76	0	Sparsveg_Cat	%
0.85	0.85	0	Tussock_Cat	%
0.11	0.11	0	Unclassified_Cat	%
0.02	0.02	0	Wetsedgesph_Cat	%
0.79	0.79	0	Wetsedge_Cat	%
0.26	0.26	0	Woodneedle_Cat	%
0.26	0.26	0	Alder_Ws	%
0.05	0.05	0	Arcto_Ws	%
0.62	0.62	0	Bare_Ws	%
0.68	0.68	0	Birch_Ws	%
0.97	0.97	0	Burned_Ws	%
0.5	0.5	0	Carex_Ws	%
0.22	0.22	0	CoastMarsh_Ws	%
0.01	0.01	0	Deciduous_Ws	%
0.83	0.83	0	Dwarf_Ws	%
0.61	0.61	0	Dwarfother_Ws	%
1	1	0	IceSnow_Ws	%
0.26	0.26	0	Willow_Ws	%
0.01	0.01	0	Marine_Ws	%
0.81	0.81	0	Mesicsedge_Ws	%
0.14	0.14	0	Mesiherb_Ws	%
0	0	0	Mixleaf_Ws	%
0.11	0.11	0	Needle_Ws	%
1	1	0	Openwater_Ws	%
0.76	0.76	0	Sparsveg_Ws	%
0.85	0.85	0	Tussock_Ws	%
0.11	0.11	0	Unclassified_Ws	%
0.01	0.01	0	Wetsedgesph_Ws	%
0.79	0.79	0	Wetsedge_Ws	%
0.19	0.19	0	Woodneedle_Ws	%
191	428	237	PetCCC_Cat	mm
191	428	237	PetCCC_Ws	mm
197	439	242	PetMPI_Cat	mm
197	439	242	PetMPI_Ws	mm
7	-3	-10	MAGT_Cat	C
7	-3	-10	MAGT_Ws	C
0.69	0.94	0.25	ALT_Cat	cm

Table D (cont.)

0.68	0.93	0.25	ALT_WS	cm
1431679	1431679	0	CoastMin	km
1444106	1444691	585	CoastMax	km
1201175	1201175	0	CoastRange	km
1437685	1438110	425	CoastMean	km
802884	802884	0	LakeMin	km
839665	841833	2168	LakeMax	km
190063	190063	0	LakeRange	km
820836	822370	1534	LakeMean	km
125985	124284	-1701	ElevDif_m	m
9	9	0	Slope	None
11.94	-21.16	-33.1	ObsTemp	C
18.17	25.61	7.43	ObsPrecip	mm/month

APPENDIX E

Among the 83 variables were two that quantified the connectivity of sites to other potential habitats within the region: minimum distance to the coast (CoastMin) and minimum distance to a lake (LakeMin). To generate these connectivity measures, I used ARCGIS and Satellite Aperture Radar imagery (SAR imagery) to identify lakes that freeze entirely (Brown et al., 2010). I then used the flow direction and flow accumulation tools in ARCGIS to define the stream network. Next, I used a conditional statement in the raster calculator to assign a high cost for travel over land and a low cost for travel through the drainage network. Finally, I used the cost distance tool to calculate the minimum, maximum, and mean distances to from each site to unfrozen lakes and the coast.

The remaining predictors were produced using ARCGIS and available data layers for Alaska's North Slope. Watersheds were delineated using ARC for each NHD stream unit, or segment, and then predictor variables were averaged using the zonal statistics tool, to come up with values for predictors in each segment's associated watershed.