Telling the North American beaver tale: modelling *Castor canadensis* distribution in Mi'kma'ki (Nova Scotia, Canada)

by

Geneva Bahen

Submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Environmental Science and Environment, Sustainability, and Society

at

Dalhousie University

Halifax, Nova Scotia

April 2023

Supervisor:

Dr. Alana Westwood

School for Resource and Environmental Studies

Dalhousie University

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The American beaver (*Castor canadensis*) is a keystone species of both significant ecological and biocultural importance in Mi'kma'ki (Nova Scotia). In North America, several environmental covariates are known to influence *C. canadensis* habitat selection, including distance to watercourse, stream gradient, and distance to preferred hardwoods; however, specific distances and species vary greatly throughout their continental range, prompting the need for local study. In Nova Scotia, occurrence data have never been systematically collected as *C. canadensis* is a common and unthreatened species, resulting in geographic knowledge gaps. This thesis remedied this knowledge gap by using Maximum Entropy modelling (Maxent) to create a species distribution model using environmental conditions present in areas of known occurrences to predict areas of *C. canadensis* distribution across the province. Input layers consisted of predominantly citizen-science occurrence data, and environmental covariates which characterized geomorphology and forest composition. Correlation analysis and reverse stepwise elimination were used to generate two models: an ecological model, and a human footprint model, where the latter investigated the influence of anthropogenic disturbance. Each model was an average of 10 replicates, using 500 iterations, 10,000 background pseudo-absence points, and a jackknife test to measure variable importance. The ecological model produced a high averaged area under the receiver operating characteristic curve (AUC) for the replicated runs (0.80 +/- 0.02), where the strongest contributors to distribution based on the permutation importance were 'Watercourse' (26.6%), 'Elevation' (15.1%), 'Red Oak' (11.8%), 'Aspen' (11.4%), and 'Gray Birch' (9.9%). Response curves indicated proximity to watercourses, low elevation, proximity to aspen, and distance from red oak and gray birch were important habitat associations. The human footprint model showed a positive relationship between occurrence points and human footprint, likely due to citizen-science collected occurrence data. Areas of known historic usage overlap at a landscape scale with distribution. The habitat association findings are consistent with previous studies suggesting watercourses, elevation and specific hardwood tree species are the main drivers of distribution, highlighting important areas of the predicted distribution of *C. canadensis* in the Wabanaki-Acadian Forest. Modelling this distribution is a practical conservation and management tool, which can contribute to future efforts to map biocultural connectivity in Mi'kma'ki (Nova Scotia), inform land use management, and provide insight into future species' sampling needs. ***Keywords**: American beaver, Castor canadensis, SDM, Maxent, Nova Scotia*

LIST OF ABBREVIATIONS

AC CDC	Atlantic Canada Conservation Data Center
AUC	Area Under the Receiver Operating Characteristic Curve
C. canadensis	*Castor canadensis*
CHFI	Canadian Human Footprint Index
DEM	Digital Elevation Model
ESRI	Environmental Systems Research Institute
FVC	Forest Vegetation Composition
GM	Geomorphologic
HF	Human Footprint
IBPES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
IPCA	Indigenous Protected and Conserved Area
IPCC	Intergovernmental Panel On Climate Change
SDM	Species Distribution Model
Maxent	Maximum Entropy
NAD83 UTM	North American Datum of 1983 Universal Transverse Mercator
NCCSC	Nature Conservancy Canada Stream Classification
NSDNR	Nova Scotia Department of Natural Resources
NSHN	Nova Scotia Hydrographic Network
NWWG	Northwestern Working Group
RFP	Request for Proposals
UINR	Unama'ki Institute of Natural Resources
UNEP	United Nations Environment Programme

ACKNOWLEDGEMENTS

I acknowledge that this thesis was conducted in Mi’kma’ki, the ancestral and unceded territory of the Mi’kmaq people, and I am grateful to live and learn on these lands. We are all treaty people. I would like to thank my supervisor, Dr. Alana Westwood, for the support, guidance, and direction that she provided throughout this project. I would like to extend a special acknowledgement to Riley Scanlan Lane for providing mentorship and friendship. I would like to acknowledge Dr. Tarah Wright for her guidance and positivity as the Honour’s course coordinator. Thank you to my friends and family for endless support.

CHAPTER 1 - INTRODUCTION

1.1 Motivation

Biodiversity and its ecological benefits underpin planetary health and human well-being (IPBES, 2019; UNEP, 2021). Unfortunately, global biodiversity is declining at unprecedented rates, largely attributed to compounding anthropogenic pressures (Venter et al., 2016; IPBES, 2019; UNEP, 2021; Hirsh-Pearson, 2022). Natural wetlands are one of the most biodiverse ecosystems and have lost up to 87% of their extent since the 18th century due to anthropogenic pressures, limiting their ecological benefits (Davidson, 2014; IPBES, 2019; Dertein et al., 2020). These rapid declines far exceed our capacity to monitor the associated spatial and temporal changes in species populations (Davidson, 2014; IPBES, 2019; Dertein et al., 2020). In a Western conservation framework, threatened species tend to receive higher priority than non-threatened species for data collection due to the strategic allocation of limited resources (Boakes et al., 2010; Baker et al., 2019). However, deficient data concerning important but nonthreatened species reduces our ability to track current population dynamics and reflect the intricacies of current ecosystems for future study (Currie, 1991; Boakes et al., 2010). Species distribution models (SDM) are one approach used to remedy our limited data collection capacities, which apply an understanding of species niche requirements and available environmental data to predict species distribution at large spatial scales (Guisan & Zimmerman, 2000; Werkowska et al., 2017; Lee-Yaw et al., 2022).

One such species that is unthreatened and under-documented is the American beaver (*Castor canadensis*). This species actively works to mediate the biodiversity crisis and restore wetland habitat (Müller-Schwartze, 2011). *C. canadensis* build dams in watercourses out of organic materials to create predator-resistant residences, which can increase water levels, carbon and nutrient sinks, primary production, biodiversity, and numerous other ecosystem services (Naiman et al. 1988; Hyvönen & Nummi, 2008; Larsen et al., 2021). The works of *C. canadensis* promote climate resilience, as the subsequent flooding from their dams reconnects and restores floodplains (Jordan & Fairfax, 2022). *C. canadensis* are a keystone species for these reasons and one of few 'ecosystem engineers', as they physically alter wetlands to enrich habitat for themselves and other species (Menge et al., 1995; Müller-Schwartze, 2011). Though unthreatened, as a pillar of environmental integrity, Indigenous identity, and history in North

America, further efforts are required to understand the main drivers of *C. canadensis* distribution. To safeguard these important ecosystem services, this thesis developed SDMs to identify the landscape-scale drivers of *C. canadensis* distribution in the Atlantic Canadian province of Nova Scotia, a previously unstudied region.

1.2 Background

Poor data concerning a keystone species, though currently common and unthreatened, may pose ecological consequences in an uncertain future (Boakes et al. 2010; Baker et al., 2019). Fossil records demonstrate that *C. canadensis* has existed as the only extant beaver on the continent for at least 12,000 years, physically shaping the ecological landscape (Müller-Schwartze, 2011). As a semiaquatic rodent, *C. canadensis* are known to thrive in suitable wetland habitat where rich vegetation grows near the riparian zone, the strip of terrestrial land adjacent to watercourses (Müller-Schwartze, 2011). As habitat composition varies greatly throughout North America, subpopulations of *C. canadensis* have grown to rely on diverse food sources, watercourse characteristics and environmental conditions (Müller-Schwartze, 2011). Summarized in Touhiri et al. (2018) and Müller-Schwartze (2011), previous studies of regional drivers of *C. canadensis* distribution in North America have used many SDM techniques with varying results by region. Past studies found strong drivers of distribution to include geomorphological characteristics such as watercourse proximity, watercourse gradient, elevation, and forest composition characteristics, such as woody and deciduous vegetation (Touhiri et al., 2018; Dittbrenner et al., 2018). However, these characteristics vary greatly throughout North America and no studies have investigated *C. canadensis*-habitat interactions in Atlantic Canada.

C. canadensis have been intertwined with Mi'kmaq communities for millennia, playing a key role both ecologically and culturally (Fauset, 1925; Michelson, 1925; Parsons, 1925). The Mi'kmaq are the Indigenous peoples of the traditional and unceded territory of Mi'kma'ki (the Atlantic Canadian provinces). Traditional Mi'kmaq ways of life relied wholly on the surrounding natural environment, consequently developing strong connections to and reverence for the features, processes and ecological communities that sustained them (Historica Canada, 2022). *C. canadensis* are identified in many stories of Mi'kmaq oral history such as that of Kluscap petting the megafauna beaver until it became an appropriate non-threatening size (UINR, 2007), among many others (Fauset, 1925; Michelson, 1925; Parsons, 1925). Both oral and archeological histories reveal seasonal patterns of resource harvesting and migration of the Mi'kmaq, reliant on oceanic resources in the spring and summer, and inland resources in the fall and winter, which includes mammals such as *C. canadensis* for nutrition, clothing, and tools (UINR, 2007; Historica Canada, 2022).

C. canadensis have been and continue to be valuable to Mi'kmaq communities, and governing bodies such as the Unama'ki Institute of Natural Resources (UINR) (UINR, 2007). A recent Requests for Proposals (RFP) from the UINR expressed interest in modelling biocultural connectivity between Kluskap Indigenous Protected and Conserved Area (IPCA) and the Bras d'Or Biosphere Reserve in northern Nova Scotia (UINR, 2022). This RFP emphasized the importance of biocultural diversity, as Mi'kmaw principles understand the "web of life" to include the interconnectedness of both ecological and cultural features (UINR, 2022). *C. canadensis* are an ecological keystone species of significant cultural importance; however, in Nova Scotia the quantitative spatial data concerning this species is of poor quality and coverage, posing difficulty for these proposed biocultural connectivity modelling efforts (UINR, 2022).

Aside from their relationship with the Mi'kmaq, *C. canadensis* and humans have a long-shared and turbulent history of coexistence. As a furbearer, *C. canadensis* pelts were a significant driver of the fur trade leading to the European exploration and colonization of Turtle Island (North America) (Carlos et al., 2011). The fur trade drove the once omnipresent *C. canadensis* from an estimated 60-400 million individuals in North America prior to European settlement towards near-extirpation (Seton, 1929), from which the continental population has recovered to upwards of 9 million since the 1980s (Naiman et al., 1988). While populations are considered recovered from a population stability and growth perspective (Naiman et al., 1988), they are still a fraction of what they were once estimated to be before near-extirpation (Seton, 1929). Today, one factor which is understood to limit population growth is land use change, and the impact of the expanding human footprint on wetland ecosystems (Müller-Schwartze, 2011). Many wetlands throughout North America have been destroyed or fragmented by roads, agriculture, harvest, pollution and other human activities (NWWG, 1988; Dahl, 1990; IPCC, 2001; IPBES, 2019). This human footprint often leads to human-beaver conflict as *C. canadensis* recolonize their historic wetland habitat, now rife with infrastructure (Naiman et al., 1988; Baker

& Hill, 2003; Müller-Schwartze, 2011). The extent of the human footprint influence on *C. canadensis* populations throughout North America remains unknown.

In Nova Scotia, resources for conservation and management are often deployed on species listed under laws to protect species at risk, to prevent extinction and promote recovery. However, it is becoming increasingly valuable to understand the requirements of species before they become listed as endangered or at risk, to be proactive in understanding, protecting, and sustaining healthy populations of species with intrinsic and extrinsic value (Boakes et al., 2010, Baker et al., 2019; UINR, 2022). *C. canadensis* are a keystone species who restore wetland habitat and increase local biodiversity (Wright et al., 2002; Larsen et al., 2021), and therefore understanding and sustaining their populations is linked to many other wetland dependent species that benefit from their presence. In Nova Scotia, occurrence or distribution data have never been systematically compiled as *C. canadensis* is unthreatened, and thus their specific habitat preferences in the province remain unknown. While *C. canadensis* is no longer at risk of extirpation, it remains a keystone species of cultural importance to the Mi'kmaq and is a symbol of the sovereignty of Canada (*National Symbol of Canada Act,* 1985). The increasing pressures of climate change and the human footprint on wetland ecosystems deem it critical to be proactive and augment the poor-quality data in this unstudied region by using statistical extrapolation to better predict distribution and understand local habitat associations (Boakes et al., 2010, Baker et al., 2019). This can provide insight into how best to support a keystone species whose activities foster biodiversity, safeguarding wetland habitat throughout a turbulent relationship with the human footprint.

1.3 Introduction to study

The primary goal of this study was to augment the poor-quality occurrence data in Nova Scotia by developing an SDM for *C. canadensis* to identify the major landscape-scale drivers of their distribution in the province. A secondary objective was to characterize local *C. canadensis* habitat associations, as these associations are known to vary by region and study design (Müller-Schwartze, 2011; Touhiri et al., 2018; Baker et al., 2019), and Nova Scotia is characterized by the unique Wabanaki-Acadian Forest and distinctive topographic features, which have not been previously studied. A third objective was to investigate how anthropogenic pressures influence

C. canadensis distribution in the province. To remedy these geographic knowledge gaps, this study created a single species SDM which used environmental variables present in areas of known occurrences to extrapolate predicted distribution across the landscape (Guisan & Zimmerman, 2000; Elith et al., 2010). The research questions and associated hypotheses are:

1. Which landscape-scale covariates are the main drivers of American beaver (C. canadensis) predicted distribution in Mi'kma'ki (Nova Scotia)?

> H_0: Each landscape-scale covariate included in the model has no influence over explaining the variability in predicted distribution.
>
> H_a: Each landscape-scale covariate included in the model has an influence over explaining the variability in predicted distribution.

2. Are there any patterns or clusters in predicted distribution?

> H_0: The predicted probability of occurrence is evenly distributed across the landscape.
>
> H_a: The predicted probability of occurrence is unevenly distributed across the landscape.

3. How does human footprint influence distribution?

> H_0: The human footprint has no effect on predicted probability of occurrence.
>
> H_{a1}: The human footprint has a positive relationship with predicted probability of occurrence.
>
> H_{a2}: The human footprint has a negative relationship with predicted probability of occurrence.

To answer these research questions, a single-species SDM was constructed using Maximum Entropy (Maxent) modelling software (Version 3.4.4) as it is a robust presence-only SDM method (Phillips et al., n.d.). The SDM combined georeferenced occurrence data with multiple raster layers extracted from environmental datasets, composing an environmental covariate group data layer, which characterized *C. canadensis* niche requirements. Environmental covariates were informed based on a literature review of *C. canadensis* ecology, previous modelling efforts, and relevance to Nova Scotia. The SDM output was analyzed to understand the landscape-scale drivers and pattern of *C. canadensis* distribution in Nova Scotia.

To understand the relationship between human footprint and predicted distribution, the final model was run with a Human Footprint Index spatial data layer representing cumulative anthropogenic pressures (Hirsh-Pearson, 2022). While ground-truthing SDMs to validate their predictions are generally regarded as good practice, a second external dataset did not exist in Nova Scotia, as data concerning this species is limited (Araújo et al., 2019; Lee-Yaw et al., 2022). Instead, the model was compared to a spatial data layer representing historic beaver flowage ponds, representing wetland habitat alterations created by beaver dams (NSDNR, 2021). This layer was interpreted from aerial imagery and can act as a proxy of their historic distribution (NSDNR, 2021). This study increases our understanding of the specific drivers of *C. canadensis* distribution in Nova Scotia, can locate areas of high and low predicted occurrence and can inform further data collection efforts.

CHAPTER 2 - LITERATURE REVIEW

2.1 Literature search strategy

To effectively produce an SDM for a species, it is essential that there is a sufficient understanding of species ecology, past modelling efforts, species distribution modelling and associated methods. This review contextualized the knowledge gaps related to insufficient data for *C. canadensis* and illuminated Maxent modelling as a best practice to remedy this gap. This review searched the databases: SCOPUS, Biological Abstracts, and ScienceDirect for published material as of October 2022. If relevant modelling studies were referenced in those sources identified, they were also investigated. Key search terms used in all databases included:

- "Castor canadensis" OR "American beaver" AND "ecology" OR "habitat"
- "Castor canadensis" OR "Castor fiber" AND "species distribution model*"
- "MaxEnt" OR "Maximum Entropy*"
- "Species distribution model*"

2.2 Species ecology

There are two living species of beaver today, including the American beaver (*Castor canadensis*) and the Eurasian beaver (*Castor fiber*) (Müller-Schwartze, 2011). *C. canadensis* has existed in North America since the Pleistocene (2.588 million – 12,000 years) and has been the only extant beaver on the continent since the Holocene (12,000 – present) (Müller-Schwartze, 2011). *C. canadensis* occupies most of the continent, apart from arctic tundra regions, southwestern deserts, and peninsular Florida (Baker & Hill, 2003). *C. canadensis* are ecosystem engineers as they modify wetland habitats by building predator-resistant dams and lodges in watercourses (Müller-Schwartze, 2011). *C. canadensis* adults range from 40-50lbs and are adapted with distinct characteristics that aid in their functioning, such as a waterproof fur pelt, large incisors adapted to gnaw and fell trees, hand-like front feet, webbed back feet, and a flat, scaly tail to balance on land, steer when swimming, and use as a noise-diversion defense technique (Müller-Schwartze, 2011). *C. canadensis* live in colonies with sizes averaging 4 individuals in northern regions to 8 in southern regions, though this is dependent on habitat quality, as some populations in notably poor-quality habitat are known to have fewer kits than those in higher quality habitat (Müller-Schwartze, 2011).

C. canadensis are a semiaquatic mammal and depend on an interface of both suitable aquatic and terrestrial habitats, called the riparian zone (Wang et al., 2019). Riparian areas in the *C. canadensis* species range vary greatly across the continent, in the types of watercourses, topography, and vegetation present, though *C. canadensis* will consistently thrive among diverse riparian vegetation of many compositions (Müller-Schwarze, 2011). *C. canadensis* dam structures are built out of riparian materials such as felled trees, sticks, stones, mud, and grass (Wright et al., 2002; Müller-Schwartze, 2011).

While *C. canadensis* are capable of modifying stream flow, water level, and hydrology once a dam has been established, certain hydrologic characteristics of watercourses need to be present to establish a colony (Touihri et al., 2018). These watercourses must be sufficient in supplying year-round swimming, and safety from predators within their dams (Müller-Schwarze, 2011), but not be "too deep" to inhibit construction (Slough & Sadleir, 1977). Watercourses must not have extreme water level fluctuations, which flood or damage dam structures (Slough & Sadleir, 1977; Allen, 1983; Müller-Schwarze, 2011). Similarly, watercourses classified as 1st to 4th order by Strahler (1957) are preferred, as increasing flow order can damage dams (Touihri et al., 2018). Lastly, streams and rivers are preferred, rather than lakes, as these provide an opportunity for increased aquatic vegetation and are conducive to protecting dam structures from wave action (Allen, 1983).

C. canadensis are a central place forager, operating out of their lodges to gather food and materials, as travelling far distances from their lodges to forage expends energy and exposes them to predation risk (Müller-Schwartze, 2011). *C. canadensis* will travel to forage in both the aquatic habitat and terrestrial areas surrounding their dams up to 100 m from the shoreline according to Donkor & Fryxell (2000) and up to 200 m from the water according to Allen (1983) which depends on their specific dam site, surrounding river valley, and food sources (Allen, 1983). However, Müller-Schwarze (2011) summarized the findings of many studies, suggesting that determinant food sources should grow within 30 m from occupied watercourses. Discrepancies likely exist due to site characteristics and forest variability across their continental range. While *C. canadensis* are known to travel between 100 - 200 m from a watercourse for food, dams are likely to be located within 30 m of determinant food sources (Müller-Schwarze, 2011).

As a sentinel species of the wetland environment, *C. canadensis* modify the landscape to suit their niche requirements and subsequently provide essential functions and services to other biodiversity. While *C. canadensis* dams and lodges are created to act as a residence, these dams alter the ecosystem to create deep 'flowage ponds', which reduce water velocity and increase the water level, which subsequently increases food supply from the growth of riparian and wetland vegetation (Larsen et al., 2021). There are a variety of ecosystem services that are impacted by this construction (Naiman et al., 1988). Larsen et al. (2021) summarizes a few of these impacts, including the ability of these dams to alter hydrology, increase nutrient and water resident times, sequester carbon, increase nutrient sinks, provide lotic and lentic habitat transitions, increase primary production, promote aquatic habitat, increase biodiversity, and increase habitat complexity (Larsen et al., 2021). Jordan & Fairfax (2022) note how local flooding restores floodplains, which promote heterogenous water temperatures and extreme climactic event resilience, such as that from fire, drought, and flooding (Jordan & Fairfax, 2022).

C. canadensis rely on unique riparian vegetation and tree species for food, though the primary sources of food differ depending on seasonal availability and region (Baker & Hill, 2003; Müller-Schwarze, 2011). A commonality among all *C. canadensis* individuals is that their primary consumed foods are deciduous trees, aquatic plants, and shrubs, additionally felling trees and seedlings to create their dams (Baker & Hill, 2003). Seasonal variation in diet includes a tendency towards herbaceous plants in their active growing seasons, and woody bark (the cambium layer) in the winter months (Baker & Hill, 2003). It is not only species present, but the composition of vegetation groundcover which has been identified as an indicator of *C. canadensis* damming sites (Dieter & McCabe, 1989; Müller-Schwarze, 2011). *C. canadensis* are best described as a "picky generalist" species as they can exist in a wide range of environmental conditions throughout their expansive range; however, there are consistent preferences throughout this range (Müller-Schwartze, 2011). The presence of wetland and riparian habitat suitable for dam-building, and the presence of forest composition and vegetation for consumption are the primary drivers of habitat selection (Müller-Schwartze, 2011; Touhiri et al., 2018). However, the specific niche requirements of *C. canadensis* are disputed, as the range of suitable conditions differ by region.

2.3 Previous modelling efforts

In the search of Biological Abstracts and SCOPUS, there were no SDMs for *C. canadensis*. The search of ScienceDirect yielded one study of habitat preferences and forage selection of *C. canadensis,* Mahoney & Stella (2020). As few studies were found in this direct search, those referenced specifically as past modelling efforts in Müller-Schwartze (2011), a *C. canadensis* ecology textbook, were investigated. These ten studies were organized by year, and their locations, model method, and categories of covariates included were compiled (Table 2.1).

Table 2.1 Studies identified in literature review to have used deterministic covariates to model *C. canadensis* distribution in North America. Covariate Category code: GM = geomorphologic characteristics; FVC = vegetation composition; HF = human footprint; OTHER = another variable not included in GM, FVC, or HF.

Study	**Location**	**Model Type**	**Covariate Category**			
			GM	*FVC*	*HF*	*OTHER*
Allen (1983)	Colorado	Habitat Suitability Index Model	X	X	X	X
Howard & Larson (1985)	Massachusetts	Stream Habitat Classification	X	X	X	X
Beier & Barrett (1987)	Nevada & California	Stepwise Logic Regression	X	X		X
McComb et al. (1990)	Oregon	Habitat Suitability Model	X	X		
Cotton (1990)	Quebec	Linear Regression Model	X	X	X	X
Barnes & Mallik (1997)	Ontario	Dam Abundance Model	X	X		
Suzuki & McComb (1998)	Oregon	Habitat Classification Model	X			
Curtis & Jensen (2004)	New York	Stepwise Logistic Regression	X	X	X	X
Mumma et al. (2018)	British Columbia	Multinomial Logistic Regression Model	X		X	X
Mahoney & Stella (2020)	New York	Mixed-effect Logistic Regression Model		X		

The individual modelling efforts had significant variations in environmental predictor covariates, model type used, and region of study (Table 2.1). Most studies were conducted in the United States, whereas only three were conducted in Canada (Table 2.1). While the study design depends on the goals of the study, different designs, model types, input covariates, and regional habitat variability have led to inconsistencies regarding the drivers of *C. canadensis* distribution. Nine of the ten studies reviewed included geomorphological characteristics, eight included forest vegetation composition characteristics, and five included some type of human footprint variable (Table 2.1). Six included another covariate exclusive of these categories, such as wolf populations in Mumma et al. (2018), or stream substrates in Curtis & Jensen (2004). There was extensive variation in how each covariate was represented, which influences the conclusions which can be drawn when comparing study results.

In addition to the models reviewed in Table 2.1, a review paper by Touhiri et al. (2018) summarized ten modelling studies, some of which were excluded from this literature review, which identified important factors influencing *C. canadensis* distribution in past efforts. Among these studies, environmental covariates included home range size, forage distance, watercourse type, stream order, stream gradient, stream size and depth, watershed size, valley width, substrate type, riparian slope, deciduous species, tree stem diameter, fire and anthropogenic features (Touhiri et al., 2018). A second review paper by Dittbrenner et al. (2018) summarized ten modelling studies with three commonalities with those from Touhiri et al. (2018). The common covariates regarding geomorphology were valley width, stream length, stream gradient, stream depth, steam width, bank slope, stream substrates, stream order, and basin size (Dittbrenner et al. 2018). The forest characteristics studied included vegetation composition, vegetation density, canopy cover, canopy height, stem diameter, habitat and vegetation area, and a shoreline development ratio (Dittbrenner et al., 2018). We can gather from these major findings that the primary covariates used to model *C. canadensis* distribution in previous efforts have used a combination of geomorphological and forest vegetation composition conditions.

This section will explore the findings of these past modelling efforts for the *C. canadensis* to highlight the importance of small-scale studies based on a robust understanding of species ecology to understand local factors that influence population distribution.

One of the earliest attempts to 'model' *C. canadensis* habitat was conducted by Atwater (1940) in Montana to determine *C. canadensis* distribution using environmental and topographic variables. These initial qualitative studies provided the groundwork for Allen (1983) to produce a quantitative habitat suitability index model to investigate habitat associations in British Columbia. Allen (1983) found positive trends with *C. canadensis* occurrence and shrub characteristics, tree diameter at breast height, water lily, and woody vegetation dominated by aspen, willow, cottonwood or alder. Allen (1983) found a negative trend between occurrence and stream gradients higher than 6%. Studies alike since then have progressed from qualitative studies and have quantitatively investigated relationships between *C. canadensis* and geomorphologic, forest, and other environmental characteristics (e.g., Howard & Larson, 1985; Beier & Barrett, 1987; Suzuki & McComb, 1998).

Watercourse characteristics, such as hydrology, stream width, bank slope and stream gradient have been recognized to be influential in *C. canadensis* distribution (e.g., Suzuki & McComb, 1998; Curtis & Jensen, 2004; Touhiri et al., 2018; Dittbrenner et al., 2018). Stream width is a geomorphologic characteristic of watercourses which has been found to be a determinant of *C. canadensis* dam locations; however, specific widths differ by location of population sampled. A summary of findings across the continent found that *C. canadensis* prefer to build dams in streams over 1.4 m wide, with a preference for 8 m in width (Müller-Schwarze, 2011). However, Suzuki & McComb (1998) found that *C. canadensis* colonies in their small-scale studies tended to prefer watercourse widths of 4–6 m or less (Suzuki & McComb, 1998). Müller-Schwarze (2011) compiled an average of studies across the continent, whereas Suzuki & McComb (1998) studied specific subpopulations.

As described in Baldwin (2013), many previous studies which have investigated bank slope as a determinant of *C. canadensis* site selection found that increasing steepness of the riverbank was negatively correlated with occurrence, although there is no distinct number highlighting the preferable or intolerable slope. While baseline geomorphologic factors influence the first establishment of a dam in a specific watercourse, specific ranges governing local populations are dependent on available habitat.

Stream gradient has been commonly investigated and found to be a significant driver of *C. canadensis* occurrences in most models (e.g. Allen, 1983; McComb et al., 1990). *C. canadensis* generally tend to prefer low stream gradients (Howard & Larson, 1985; Beier & Barret, 1987). According to a study conducted in Drift Creek Basin, Lincoln County, Oregon by Suzuki & McComb (1998), a 3% gradient is optimal, and no dams were recorded on watercourses with a 10% gradient or higher. However, Retzer et al. (1956) reported that 68% of studied *C. canadensis* colonies in Colorado were situated on streams with a gradient of less than 6%, and that gradients up to 14% were habitable. Increasing stream gradient tends to have a negative influence on *C. canadensis* occurrence, though specific grade ranges and impacts depend on the location studied and stream gradients available.

Watercourse valley width has been investigated in previous models (e.g. McComb et al., 1990; Cotton, 1990; Suzuki & McComb, 1998). Generally, a wider valley provides more suitable habitat for *C. canadensis* (McComb et al., 1990; Cotton, 1990; Suzuki & McComb, 1998; Touhiri et al., 2018). It is suggested that this preference is due to narrow valleys providing poor support for the vegetation that are the primary food sources of *C. canadensis,* such as poplar, willow, alder, and birch (Northcott, 1964). While wider watercourse valleys are generally accepted to be preferable, there is discrepancy in the specific widths which are most suitable. *C. canadensis* tend to prefer to build in watercourses where the valley width is greater than 10m (Suzuki & McComb, 1998). However, McComb et al. (1990) found that the valleys occupied by *C. canadensis* had a mean width of 13.5m, which contrasts the latter Suzuki & McComb (1998) findings which found a mean width of 22.7m. While specific widths are not consistent among populations studied, there is a positive trend of *C. canadensis* sites and wider valleys, likely associated with habitable conditions for tree species of interest (Northcott, 1964).

The preferred food sources of *C. canadensis* differ by regional availability, from willow stands, to brush, to southern beech, to pine in rare cases depending on the forest composition (Müller-Schwarze, 2011). In a study by Aleksiuk (1970) on *C. canadensis* diets in the Mackenzie Delta, NWT, it was found that preferred food sources during growing seasons were the foliage of willow, whereas in the winter months, the bark of willow, poplar, and alder were the primary source of protein and calories (Aleksiuk, 1970; Baker & Hill, 2003). Contrary to Aleksiuk's findings, when aspen and poplar are present and abundant, these species are more desirable than willow (Jenkins, 1981; Baker & Hill, 2003; Müller-Schwarze, 2011). A commonality of these trees is that they are hardwoods, with easy bark to peel off and can be gnawed with ease (Müller-Schwarze, 2011). In Ontario, a study was conducted which provided food sources to *C. canadensis*, which concluded that the primary dietary preference was aspen, water lily, raspberry, alder, and red maple, of the species provided (Doucet & Fryxell, 1993). The least preferred sources of food are conifers, with resiny bark, such as pine and spruce (Müller-Schwarze, 2011). Primary food sources for *C. canadensis* in Novak (1987) found that preferred vegetation included a variety of riparian plant species, such as water lily, grasses, cattails, rushes, and other herbaceous species; however, deciduous woody plants are the primary limiting food source that determines suitable habitat (Grinnell et al., 1937; Baker and Hill, 2003). Generally, it is accepted that hardwood tree species and woody vegetation provide a better source of food than softwood species.

Due to the variability in diet among populations, perhaps hydrologic and geomorphologic factors are better determinants of *C. canadensis* occurrence rather than variables which are associated with diet (Jenkins, 1981). In a following study by Beier & Barrett (1987), it was found that stream depth, stream gradient and stream width were the most important factors related to habitat use, which is consistent with many of the findings from Allen (1983). In both Jakes et al. (2007) and Lapointe St-Pierre et al., (2017), it was noted that the specific characteristics of riparian vegetation in a study area are difficult to represent without field data and are consequently difficult to use in predictive models (Franklin, 1995). Based on previous studies of the preferred *C. canadensis* diet, woody deciduous species can define *C. canadensis* habitat depending on regional species and local availability but are less effective at determining habitat than geomorphologic factors.

Lastly, several studies investigated covariates related to human influence on *C. canadensis* distribution (Table 2.1). One study found a positive relationship between anthropogenic features, such as culverts, and *C. canadensis* occurrence, where *C. canadensis* were likely to construct a dam blocking small culverts due to the opportunistic watercourse constriction (Curtis & Jensen, 2004). Contrary to this finding, Allen (1983) found a negative relationship between a shoreline development ratio and *C. canadensis* occurrence. As the impacts of the human footprint come in many forms, their studied impacts on distribution differ by region and study design.

2.4 Species distribution modelling

In the field of ecology, SDMs are a practical tool which can be used to understand species distributions across large spatial scales which can help understand the impact of climate change on species distributions, direct conservation resources, and understand species habitat associations (Guisan & Zimmerman, 2000; Elith et al., 2010; Yackulic et al., 2013). SDMs apply ecological niche theory, as they predict distribution based on the relationship between known occurrences and environmental conditions at those locations (Austin, 2007; Westwood, 2016; Lee-Yaw et al., 2022). An underlying assumption of SDMs is that the input occurrence data represents species in suitable habitat conditions (Westwood, 2016). SDMs can therefore predict distribution across a landscape, based on the composition of suitable habitat found at known occurrence points (Guisan & Zimmerman, 2000). Many SDM methods use different input data types, algorithms, and software; therefore, method selection differs based on available inputs, desired outputs, and intended applications (Westwood, 2016).

Maxent is an open-source SDM method which is advantageous when compared to other techniques as it only requires species presence data, rather than systematically collected presence-absence data to generate a model with good predictive performance (Phillips et al., n.d.; Phillips et al., 2006; BCCVL, 2021). Maxent has proven success with modelling presence-only species records as the software uses randomly generated pseudo-absence background points to sample environmental conditions in 'unsuitable' habitat (Elith et al., 2010). Each occurrence point is assumed to represent the ecological composition of the realized niche of a species, whereas those at background pseudo-absence points ideally represent that which is less suitable (Phillips et al., 2006). Maxent requires an understanding of the species ecology to select determinant environmental covariates, which are ideally sourced from multiple lines of evidence (Araújo et al., 2019), as it operates under the assumption that determinant covariates are not omitted from the model (Guisan & Zimmermann, 2000; Yackulic et al., 2013). Maxent applies the theory of maximum entropy to predict probability of species occurrence across a geographic landscape (Phillips et al., 2006). The program predicts probability of occurrence with the most uniform distribution (the maximum entropy), while satisfying known constraints which are based on the difference of environmental conditions at the presence and background points (Phillips et al., 2006). While SDM methods which incorporate known absences (presence-absence data) generally show higher performance accuracy than Maxent's presence-only methods, these methods were not applicable due to the presence-only data available in the province (Phillips et al., 2006; Westwood, 2016). Maxent's user-friendly interface deems it an effective modelling tool for this application due to the presence-only data available, the ease of use, and the programs demonstrated robustness (Philips et al., 2004; Phillips et al., 2006; Elith et al., 2010; Yackulic et al., 2013).

There have been many studies which investigated the species distribution, or habitat suitability of *C. canadensis* across their continental range (Table 2.1). Most notably, this review highlighted how specific geomorphologic and forest composition factors influencing one local population may not be determinant of the distribution of another population, due to *C. canadensis* being a generalist species and the variation in habitat throughout North America (Müller-Schwarze 2011). Therefore, as no study has been conducted in any province within Atlantic Canada, it is currently not understood which factors in Nova Scotia drive *C. canadensis* distribution. As Nova Scotia is home to the unique Wakanaki-Acadian Forest, there is a need to study *C. canadensis* distribution within this region to understand their specific local habitat associations. Aside from their national significance and ecological integrity, augmenting this poor data can aid future modelling and conservation efforts within the province, and help to reflect the complexities of current ecosystems for future studies (Boakes et al., 2010; UINR, 2022). This study will be the first to use the Maxent to understand which covariates influence *C. canadensis* distribution in Nova Scotia.

CHAPTER 3 - METHODS

3.1 Study area

The study area I used to investigate *C. canadensis*-environment interactions is Nova Scotia, an eastern coastal province in Canada (Figure 3.1). Nova Scotia is situated in Mi'kma'ki, the unceded traditional territory of the Mi'kmaq people. The peninsular province is joined to North America by a thin stretch of land called the Chignecto Isthmus (Figure 3.1). Climactically, Nova Scotia is in the Atlantic Maritime Ecozone, meaning the province's proximity to the north Atlantic Ocean yields moderate average winter and summer temperatures, with mean annual temperatures between 5°C and 7°C (NSMNH, n.d.; Ecological Stratification Working Group, 1995). The coastal nature of the province similarly influences precipitation, producing mean annual precipitation levels of 900 mm inland and 1500 mm in coastal areas (NSMNH, n.d.). Nova Scotia's total landcover includes over 5.5 million hectares of both land and freshwater (NSDNR, 2008). In 2016, a landcover inventory of the province found the dominant landcover was forests (75.8%), followed by natural non-forested (7.8%), agriculture (4.9%), inland water (4.2%), wetlands (2.9%) and other combined anthropogenic cover composed the remainder (NSDNR, 2016). Nova Scotia's dominant forest type is the Wabanaki-Acadian Forest, a temperate forest ecosystem with a unique mix of coniferous and deciduous tree species, which contrasts the boreal forest, dominating most other Canadian provinces (Ecological Stratification Working Group, 1995; NSDNR, 2017).

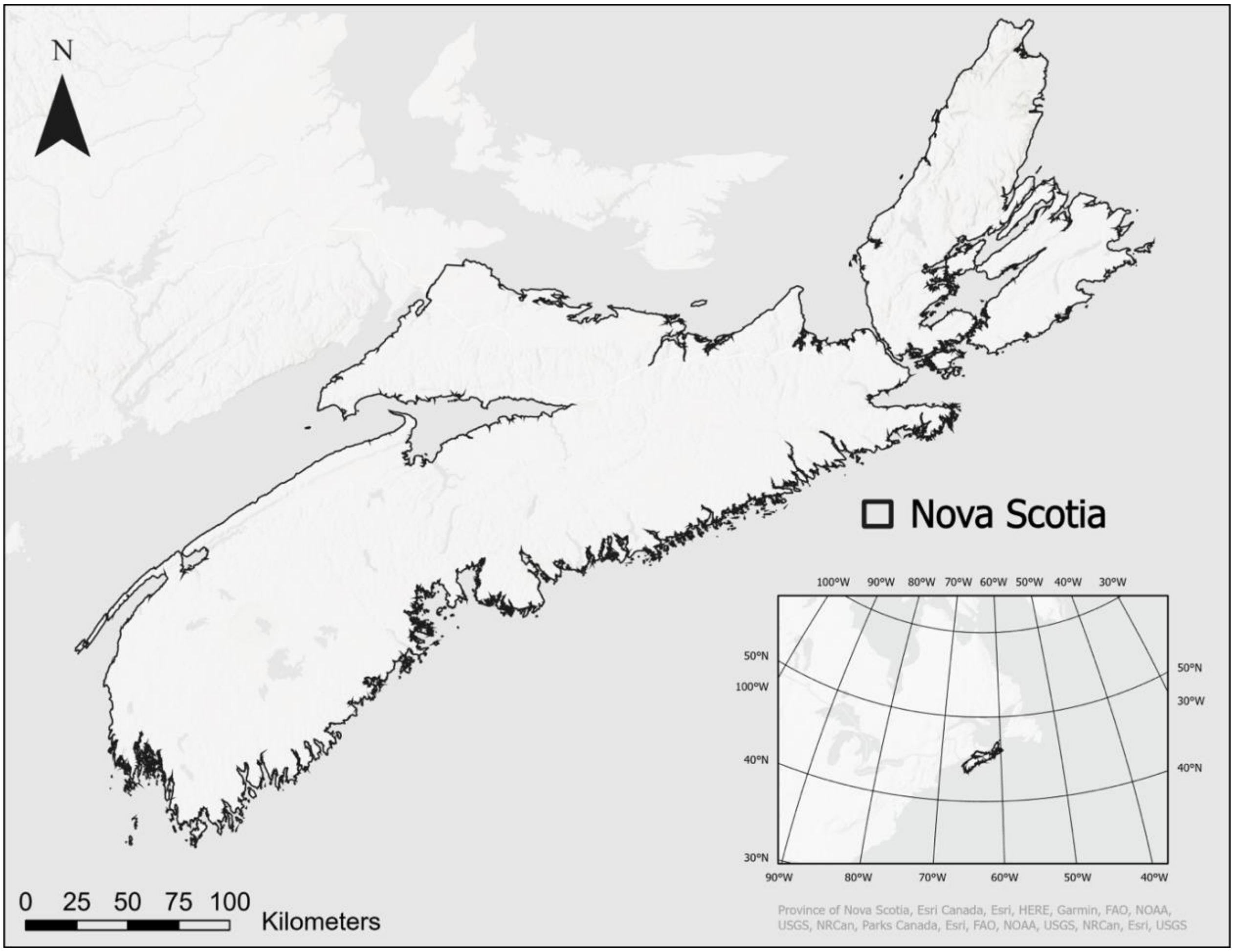

Figure 3.1 Study area for the Maxent modelling of the American beaver (*Castor canadensis*) in Nova Scotia, Canada. An inset map highlights the location of the province as it connects to Eastern North America.

3.2 Occurrence data

C. canadensis is a common and unthreatened species in Nova Scotia, and to our knowledge has not been a subject of deliberate, systematic searches for the purpose of recording occurrences, abundance, or occupancy. The occurrence data used in this study was a combination of 375 research grade occurrence points from iNaturalist (iNaturalist, 2022), and 39 occurrence points from Atlantic Canada Conservation Data Center (AC CDC) (AC CDC, 2022), which yielded a combined 414 occurrence points. iNaturalist relies on citizen-scientists to capture species occurrence records, and the AC CDC dataset included haphazard logged *C. canadensis* occurrences when conducting field work for other species (iNaturalist, 2022; AC CDC, 2022). All research grade iNaturalist points in Nova Scotia as of November 1, 2022, were used, logged between 2008-2021 (iNaturalist, 2022). The AC CDC dataset had a total of 131 occurrence records in Nova Scotia as of October 2022; however, many of these records were not georeferenced (AC CDC, 2022). There were 39 datapoints of the original 131 with sufficient information to be manually georeferenced by searching for and assigning the provided location a coordinate, resulting in a final dataset of 414 aggregated occurrences (Figure 3.2).

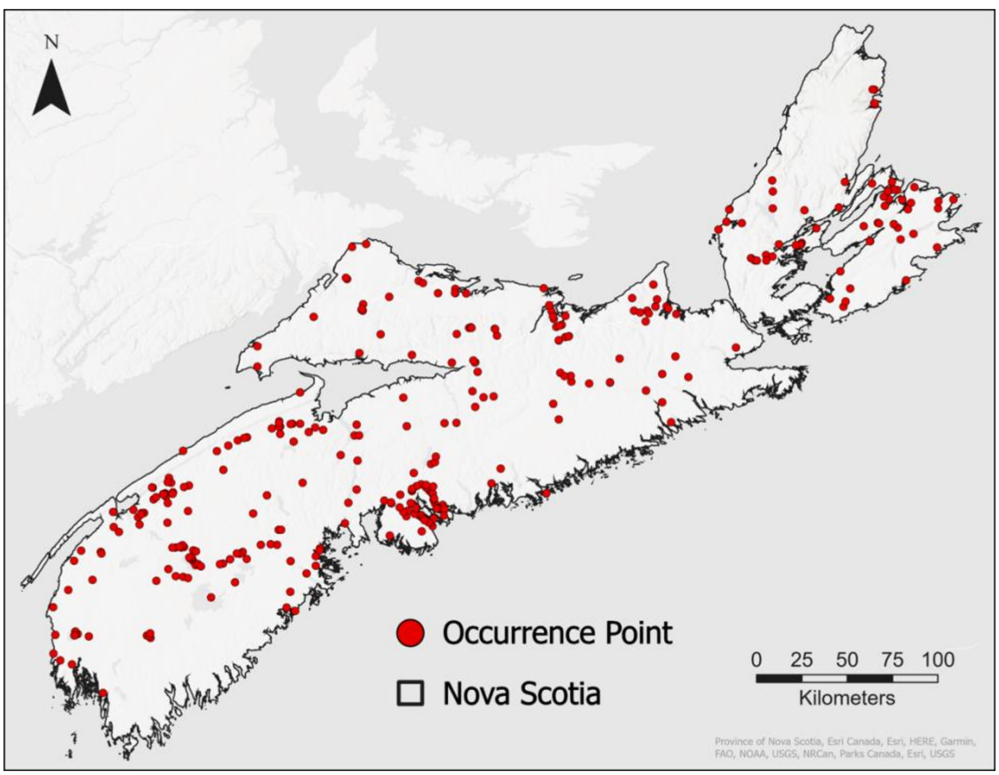

Figure 3.2 The location of the aggregated occurrence data points for the Maxent modelling of the American beaver (*Castor canadensis*) in Nova Scotia, Canada.

3.3 Environmental covariate group data layer

The literature review of *C. canadensis* ecology and previous SDM efforts for modelling this species informed the selection process of environmental covariates which were best able characterize niche habitat from available geospatial data. Of the ecologically-relevant covariates identified and those of known importance in previous modelling efforts, the covariates were included if they were both relevant to the Nova Scotian landscape and if there was an open-source spatial dataset available covering the full extent of the study area to reliably represent the habitat feature of interest. The spatial datasets used to generate these spatial layers are summarized in Table 3.1.

Table 3.1 Spatial datasets used to create the environmental covariate group data layer and subsequent Maxent SDM for the American beaver (*Castor canadensis*) in Nova Scotia, Canada.

Spatial Dataset	Brief Description	Year	Resolution	Citation
Nature Conservancy Canada Stream Classification Layer v 2.0	Freshwater classification of stream characteristics.	2019	Vector (Line)	Millar et al. (2019)
Nova Scotia Enhanced Digital Elevation Model v 2.0	Hydrologically corrected 20m resolution digital elevation model.	2006	20m	Nova Scotia Department of Natural Resources (2006)
Nova Scotia Forest Inventory	Aerial photography interpretations of forest characteristics. Interpreted and digitized from 1:10,000 air photos.	1992-2021	Vector (Polygon)	NSDNR (2021)
Nova Scotia Hydrographic Network	Watercourse characteristics, type, and locations.	2022	Vector (Line)	Service Nova Scotia and Internal Services (2020)
Statistics Canada Nova Scotia Census Files	Provincial boundary shapefiles.		Vector (Polygon)	Statistics Canada (2021)
The Canadian Human Footprint Index	Quantification of human pressures on terrestrial ecosystems.	2022	300m	Hirsh-Pearson et al. (2022)

While the literature review of previous modelling efforts identified an extensive list of covariates which have been used to model *C. canadensis* distribution, many were not included in this study. Many geomorphological covariates such as stream order and stream depth were not able to be included due to a lack of available data in the province, whereas others such as bank slope were not included due to the 300 m^2 spatial resolution of this study as this would not effectively represent the species-environment interaction at a generalized scale. The Nova Scotia Forest Inventory (NSFI) is the best classified forest data available in the study area, though the temporal range of data included in the dataset spans nearly 30 years, and the forest characteristics included are limited (NSDNR, 2021). Forest stands dominant with a specific tree species, and specific forest types (e.g., hardwood, mixedwood) were considered to be more effective at representing the foraging habits of *C. canadensis* than covariates such as canopy height, as canopy height was perceived to change more over time than dominant tree species in a forest stand. Additionally, many covariates such as vegetation density and stem diameter are not included in the NSFI and were therefore not included. Jenkins (1981) identified that perhaps vegetation composition is not as strong of a determinant of distribution when compared to watercourse characteristics; however, foraging habits and the availability of preferred food sources are intrinsic to species niche requirements, and therefore these species-environment relationships were included, as there were relevant interactions which could be captured with the data available.

There were 25 candidate covariates generated for the SDM (Appendix A). There were two distinct categories of covariates created. Firstly, there were forest-specific covariates, including specific deciduous species of known nutritional value, and forest composition types. Secondly, there were geomorphological covariates, including watercourse characteristics and elevation. The extraction of each covariate layer was done in ArcGIS Pro Version 2.9.5 (ESRI Inc., 2021). The categorical data layers from the Nature Conservancy Canada Stream Classification Layer v 2.0 (NCCSC) warranted the use of the 'Euclidean Allocation' tool as an extraction method, which provides a value for each raster cell of the closest category (Appendix A) (ESRI Inc., 2021). Continuous data layers from the Nova Scotia Hydrographic Network (NSHN) and the NSFI required extraction using the 'Euclidean Distance' tool, which provides a value for each raster cell of the distance to the attribute of interest (Appendix A) (ESRI Inc., 2021). Continuous data layers from the Canadian Human Footprint Index (CHFI), and the Nova

Scotia Enhanced Digital Elevation Model (DEM) required resampling but no extraction (Appendix A) (ESRI Inc., 2021).

To ensure that each extracted covariate which composed the environmental covariate group data layer was congruent, the covariate layers were constructed using a 300m spatial resolution, as this was the cell size of the coarsest input layer, and reduced computational processing requirements (Hirsh-Pearson, 2022). Each data input was converted to the uniform projection: NAD83 UTM Zone 20N (Seely, 2011). In ArcGIS Pro, each input layer was projected to the uniform projection, extracted, masked, and snapped to a 300m raster delineating Nova Scotia (ESRI Inc., 2021). Specific details regarding how each covariate was extracted can be found in Appendix A.

Several of the 25 candidate environmental covariates generated had correlations, where cell values in one covariate layer had a relationship to the equivalent cell in another covariate layer (ArcGIS, n.d.). The correlation between the environmental covariates increases the similarity of the inputs into the model, which limits the assessment of each covariate's predictive capacity and contribution (Phillips et al., 2017). A correlation matrix was executed in ArcGIS Pro using the 'Band Collection Statistics' tool (Appendix B) which produced a table of Pearsons R correlation coefficient values between each raster layer (ArcGIS, n.d.; ESRI Inc., 2021). The output of the correlation matrix provided values from –1.0 to +1.0, where layers which had absolute values greater than 0.7 were considered strongly correlated, and values between 0.4 and 0.7 were considered moderately correlated (e.g., Baker, 2022). The less ecologically important covariate, as determined based on known life history for *C. canadensis*, was removed among strongly correlated layers. The implications of moderate correlations were considered, and the less ecologically relevant covariate was removed, if needed (e.g., Baker, 2022).

Environmental covariates with strong correlations (greater than |0.7|) included Balsam Poplar and Willow (r = 0.99), Gradient (Complex) and Gradient (Simple) (r = 0.97), Black Cherry and Balsam Poplar (r = 0.91), Black Cherry and Willow (r = 0.89), and Black Cherry and Gray Birch (r = 0.72) (Appendix B). Balsam Poplar, Willow, and Black Cherry were correlated due to sparse coverage as these layers were built with 31 or fewer polygons where these species had 60% dominance or higher (NSDNR, 2021). These layers were all removed due to their perceived lack of influence on site selection on a landscape scale and their high correlations.

Gradient (Complex) and Gradient (Simple) had a strong correlation, as they are different classifications of the same data (Millar et al., 2019) (Appendix B). Gradient (Simple) was removed to retain more information about how gradient influences site selection.

Environmental covariates with moderate correlations included Tidal, Alkalinity, Temp, Size, and Gradient (Complex), which were generated from the same source dataset (Table 3.1), and each had moderate correlations among each other (Appendix B). Tidal, Alkalinity, and Temp were removed due to having the highest correlations and being less ecologically relevant than Size (Complex), and Gradient (Complex) (Müller-Schwarze, 2011; Touhiri et al., 2018). Ash and Gray Birch were moderately correlated (r = 0.67), where Ash was removed due to having separate moderate correlations between Red Maple and Red Oak (Appendix B). Mixedwood and Hardwood were moderately correlated (r = 0.62), where Mixedwood was removed due to being less ecologically relevant than hardwood species (Müller-Schwarze, 2011, Touhiri et al., 2018). Sugar Maple and Yellow Birch were moderately correlated (r = 0.56), where Sugar Maple was removed due to being less ecologically relevant than birch species (Müller-Schwarze, 2011; Touhiri et al., 2018) (Appendix B). White Birch and Yellow Birch, Brush and Alder, and Elevation and Human Footprint all had correlation values between 0.4 and 0.45 but were regarded as ecologically significant enough for the purposes of this study to remain in the environmental covariate group data layer (Appendix B).

Overall, an evaluation of the correlation matrix resulted in the removal of Balsam Poplar, Willow, Black Cherry, and Gradient (Simple), Tidal, Alkalinity, Temp, Ash, Mixedwood, and Sugar Maple. There were 25 candidate covariates developed for this modelling effort, and after a process of correlation analysis, there were 14 remaining (Table 3.2).

Table 3.2 Environmental candidate covariates used as initial inputs to model species distribution for the American beaver (*Castor canadensis*) in Nova Scotia, Canada.

Data Source	**Covariate**	**Rationale**	**Attribute of Interest**	**Extraction Method**	**Layer Type**
NCC Stream Classification Layer v 2.0	Gradient (Complex)	Beavers tend to build dams in a specific range of gradients.	Grad_Comp	Euclidean Allocation	Categorical (7 classes)
	Size (Complex)	Larger streams support increased vegetation.	Size_Comp	Euclidean Allocation	Categorical (6 classes)
NS Enhanced Digital Elevation Model	Elevation	Influence on dam site suitability and vegetation.	Elevation	Resample (Cubic)	Continuous
NS Forest Inventory	Alder	Distance to food/materials.	FORNON38 FORNON39	Euclidean Distance	Continuous
	Aspen	Distance to food/materials. including Large Tooth and Trembling Aspen.	TA	Euclidean Distance	Continuous
	Brush	Distance to food/materials.	FORNON33	Euclidean Distance	Continuous
	Gray Birch	Distance to food/materials.	GB	Euclidean Distance	Continuous
	Hardwood	Distance to hardwood stand.	FORNON 8	Euclidean Distance	Continuous
	Red Maple	Distance to food/materials.	RM	Euclidean Distance	Continuous
	Red Oak	Distance to food/materials.	RO	Euclidean Distance	Continuous
	Softwood	Distance to softwood stand.	FORNON 2	Euclidean Distance	Continuous
	White Birch	Distance to food/materials.	YB	Euclidean Distance	Continuous
	Yellow Birch	Distance to food/materials.	WB	Euclidean Distance	Continuous
NS Hydrographic Network	Watercourse	Distance to nearest watercourse habitat.	N/A	Euclidean Distance	Continuous
The Canadian Human Footprint	Human Footprint	Influence of anthropogenic disturbance from human footprint.	Cumulative Threat Layer	Resample (Bilinear)	Continuous

3.4 Maximum entropy modelling procedure and reverse stepwise elimination

Maxent software (Version 3.4.4) inputs include presence-only species occurrence data in the form of georeferenced points, and multiple environmental covariate group data layers (Phillips et al., n.d.). The latter requires the creation of congruent raster layers, which represent species-environment interactions which are proven to influence species distribution, ideally from multiple lines of evidence (Araújo et al., 2019). The statistical relationship between the occurrence data and environmental covariate group data layer predicts probability of occurrence across the landscape, creates species-covariate response curves, and evaluates the importance of each covariate in the model output.

Maxent provides the user with a variety of default and advanced settings to model species distribution (Phillips et al., n.d.). For this study, the default Maxent output format 'cloglog', or 'complimentary log-log regression' was used, which predicts probability of species occurrence across the landscape as an index that ranges from 0.0 to 1.0. The 'cloglog' output format automatically selects the type of equation (e.g. linear, quadratic, hinge) which best mathematically represents the species-covariate relationships (Phillips et al., 2017). The output includes a measure of predictive strength for the model called the area under the receiver operating characteristic curve (AUC) (Phillips et al., 2017). A model AUC value of 1.0 indicates the model has perfect predictive capacity, whereas an AUC of 0.5 indicates that the model is no better than random (Phillips et al., 2017).

To generate each model, each model run was executed using 90% of the occurrence data to train the model and the remaining 10% was reserved to test the model. For each model run, 10 replicate models were produced which each reserved a different 10% of the occurrence data for testing. Each model run analysis used a cross validated arithmetic average of 10 replicate models. Additionally, a jackknife test was conducted, which measures the importance, relative contribution, and influence of each covariate on the AUC (Phillips et al., 2017). To eliminate a subset of the sample bias in the occurrence point data layer, the "Remove duplicate presence records" option was checked in the Maxent settings. This option eliminated species occurrence data within the same $300m^2$ raster cell, which reduced the impact of the sampling bias which stemmed from citizen-science collected occurrence data (Phillips et al., 2017).

The default regularization multiplier of 1.0 was used, with 500 maximum iterations, and 10,000 randomly generated background pseudo-absence points. While adjusting the regularization multiplier and number of iterations from the default parameters are known to reduce model overfitting, impact the accuracy, and subsequent interpretation of results (Merow et al. 2013), the meaningful manipulation of these parameters required external processing and the use of supplementary packages beyond the scope of this study (e.g., Baker, 2022). Additionally, Merow et al. (2013) highlight how the manipulation of the locations and number of randomly generated background sample points can reduce sample bias and impact the accuracy of a model; however, the specific best practices are disputed, and the creation of a bias file to direct background sample points required external processing using packages which were considered out of the scope of this study (Merow et al., 2013).

After inputting the candidate covariates (Table 3.2) into Maxent, a process of reverse stepwise elimination was used to distill the covariate group data layer to create a model with only the most deterministic covariates, iteratively removing those which contributed the least to the predictive capacity of the model (e.g., Bale et al., 2020; Baker, 2022). Each cross validated model run generates AUC statistics for the test and training data as a result of the jackknife test. The reverse stepwise elimination process evaluated each model's output, and iteratively removed any variables which were worse than random predictors of distribution. This measure was evaluated using the jackknife test results of the AUC on test data. For each covariate, the jackknife test produces a model output built with only the one covariate, and a model built without that covariate, and produces AUC on test data values for each model (Phillips et al., 2017). The covariates with the least capacity to predict distribution were iteratively removed if they had an AUC on test data equal to or less than 0.5, meaning they were worse than random predictors.

The first model run used the whole environmental covariate group data layer and had an average test AUC for the replicate runs of 0.81 +/- 0.02. Hardwood had a moderate permutation importance of 4.3; however, an analysis of the AUC on test data jackknife results revealed that the Hardwood covariate when evaluated individually had an AUC < 0.5, meaning it performed no better than random as a predictor for the model. Hardwood was removed from the environmental covariate group data layer for the second model run, to reduce overfitting of the model (Phillips et al., 2017).

The second model run used the environmental covariate group data layer without Hardwood and had a lower average test AUC for the replicated runs than the first model run (0.80 +/-0.24). White Birch had a low percent contribution and permutation importance; similarly, an analysis of the jackknife test results revealed that White Birch on its own had an AUC < 0.5, meaning it preformed no better than random as a predictor of occurrence. In this model run, Size (Complex) also had a low percent contribution and permutation importance, though an analysis of the jackknife test results for the AUC on test data revealed it had a 0.62 AUC value. Due to Size (Complex) having an AUC > 0.5, this variable was considered ecologically valuable to retain, and only White Birch was removed from the environmental covariate group data layer for third model run.

The third model used the environmental covariate group data layer without Hardwood and White Birch. This model had an average test AUC for the replicated runs of 0.80 +/- 0.02 (Appendix C). As each variable had a jackknife AUC test value greater than 0.5, this final suite of variables composed the last iteration of the model (Table 3.3).

Table 3.3 Final environmental covariates used as inputs to model species distribution for the American beaver (*Castor canadensis*) in Nova Scotia, Canada.

Data Source	**Covariate**	**Rationale**	**Attribute of Interest**	**Extraction Method**	**Layer Type**
NCC Stream Classification Layer v 2.0	Gradient (Complex)	Beavers tend to build dams in a specific range of gradients.	Grad_Comp	Euclidean Allocation	Categorical (7 classes)
	Size (Complex)	Larger streams support increased vegetation.	Size_Comp	Euclidean Allocation	Categorical (6 classes)
NS Enhanced Digital Elevation Model	Elevation	Influence on dam site suitability and vegetation.	Elevation	Resample (Cubic)	Continuous
NS Forest Inventory	Alder	Distance to food/materials.	FORNON38 FORNON39	Euclidean Distance	Continuous
	Aspen	Distance to food/materials. including Large Tooth and Trembling Aspen.	TA	Euclidean Distance	Continuous
	Brush	Distance to food/materials.	FORNON33	Euclidean Distance	Continuous
	Gray Birch	Distance to food/materials.	GB	Euclidean Distance	Continuous
	Red Maple	Distance to food/materials.	RM	Euclidean Distance	Continuous
	Red Oak	Distance to food/materials.	RO	Euclidean Distance	Continuous
	Softwood	Distance to softwood stand.	FORNON 2	Euclidean Distance	Continuous
	Yellow Birch	Distance to food/materials.	YB	Euclidean Distance	Continuous
NS Hydrographic Network	Watercourse	Distance to nearest watercourse habitat.	N/A	Euclidean Distance	Continuous
The Canadian Human Footprint	Human Footprint	Influence of anthropogenic disturbance from human footprint.	Cumulative Threat Layer	Resample (Bilinear)	Continuous

*Human Footprint covariate not used in the ecological model and only used in the human footprint model.

3.5 Investigation and validation

The final suite of covariates in Table 3.3 composed the ecological model output. A second model was generated which added the CHFI layer to create a human footprint model. The human footprint layer represents factors such as built environments, crops and pastureland, night light pollution, population density, resource extraction, forestry, and roads, with their associated 'threat' value summed as an index score (Hirsh-Pearson et al., 2022). This model quantified the influence of the human footprint on predicted distribution. The human footprint model was generated, and both the output statistics and visual model were compared to the ecological model to assess how the distribution changed.

External datasets are often used as a validation measure for SDMs, which could be alternative presence-only data, or newly created data to test whether areas of high predicted probability of occurrence had higher species counts than areas of low predicted probability (e.g., West et al., 2016; Allen & McMullin, 2019; Smith et al., 2021; Baker, 2022). For *C. canadensis* in Nova Scotia, all available occurrence data points were combined to one occurrence data layer. However, a unique advantage to modelling an ecosystem engineer is that *C. canadensis* modify the landscape with their damming activities and create flooded areas known as beaver flowage ponds, which in some cases can be seen in aerial imagery (Müller-Schwartze, 2011; NSDNR, 2021). To use an external dataset to validate the ecological model, a polygon layer representing historic beaver flowage ponds was used from the NSFI (NSDNR, 2021). This dataset has a "non-forest classification code" for 'beaver flowage' (NSDNR, 2021). This classification is separate from the 'wetland' classification and includes "any area that is or has been occupied by beaver ... this designation refers only to the water flowage area or grassy areas created by the beaver dam." and is listed as FORNON code '71' (NSDNR, 2021).

Beaver flowage polygons were extracted from the NSFI dataset, and four buffers of varying distances were created around the polygons, including 100m, 200m, 500m, and 1000m (NSDNR, 2021). The original polygons, as well as the four buffer layers were sampled using the 'Zonal Statistics as Table' tool in ArcGIS Pro (ESRI Inc., 2021). These zonal statistics calculated the mean and maximum probability of occurrence for each ecological model cell within a flowage polygon and associated buffer. Ideally, if the ecological model output and the flowage polygons were complementary, the mean or maximum probability within each polygon would have a probability of occurrence greater than the average probability of occurrence across the study area. Sampling this dataset revealed whether the model built with recent occurrence data predicted occurrence in areas of known historic *C. canadensis* use.

CHAPTER 4 – RESULTS

I created two predicted probability of occurrence models for *C. canadensis* using the Maxent algorithm, the aggregated presence-only occurrence points, and the environmental covariate group data layer for the specified model. The ecological model revealed the significant species-environment interactions driving distribution in Nova Scotia, whereas the human footprint model revealed the influence of the human footprint on the predicted distribution. A historic layer of beaver flowage was used to sample the ecological model to understand the relationship between predicted and historic distribution.

4.1 Habitat associations

4.1.1 Ecological model

The ecological model had a relatively high averaged AUC for the replicated runs (0.80 +/- 0.02) (Appendix C). Generally, an AUC value between 0.7 and 0.8 is considered acceptable, an AUC between 0.8 and 0.9 is considered excellent, and an AUC between 0.9 and 1.0 is considered 'outstanding' (Hosmer, 1989 in Mandrekar, 2010). The ecological model AUC can be considered acceptable-excellent by these standards and is closer to a perfect prediction (1.0) than a null model (0.5). A visual representation of the averaged AUC for the ecological model can be found in Appendix C. There were 12 environmental covariates used in the ecological model which each had varying contributions to the predicted distribution (Table 4.1). The variables that collectively had the most permutation importance, explaining over half the variance in the model, were 'Watercourse' (26.6%), 'Elevation' (15.1%), 'Red Oak' (11.8%), and 'Aspen' (11.4%) (Table 4.1).

Table 4.1 Percent contribution and permutation importance for twelve environmental covariates averaged across 10 cross validated Maxent replicate models, used to predict American beaver (*Castor canadensis*) probability of occurrence in Nova Scotia, Canada.

Environmental Covariate	Average Percent Contribution (%)	Average Permutation Importance (%)
Watercourse	48.3	26.6
Elevation	9.2	15.1
Red Oak	4.3	11.8
Aspen	9.3	11.4
Gray Birch	6.5	9.9
Yellow Birch	4.8	6.8
Brush	6.8	5.0
Gradient (Complex)	3.0	4.4
Alder	1.2	4.3
Softwood	5.0	3.3
Red Maple	1.1	1.0
Size (Complex)	0.7	0.3

Based on the covariate response curves in Figures 4.1, 4.2 and 4.3, inferences about the strength and type of species-covariate interactions can be deduced. *C. canadensis* had the highest probability of occurrence adjacent to a watercourse (Figure 4.1). This relationship was found to be the strongest predictor of occurrence (Table 4.1). The watercourse gradient classes which were found to be most favorable were lakes, low gradients, and moderate-high gradients (Figure 4.2). Additionally, the categorical watercourse size classes which had the highest predicted occurrence were small and medium tributary rivers (Figure 4.2). Watercourse gradient had a higher permutation importance than watercourse size (Table 4.1). *C. canadensis* were predicted to be most likely to occur in areas of low elevation, as probability of occurrence dropped below 0.5 at elevations of 200 meters or more above sea level (Figure 4.1). These geomorphological covariates had a combined permutation importance of 46.4%, explaining nearly half of the variance in the ecological model (Table 4.1).

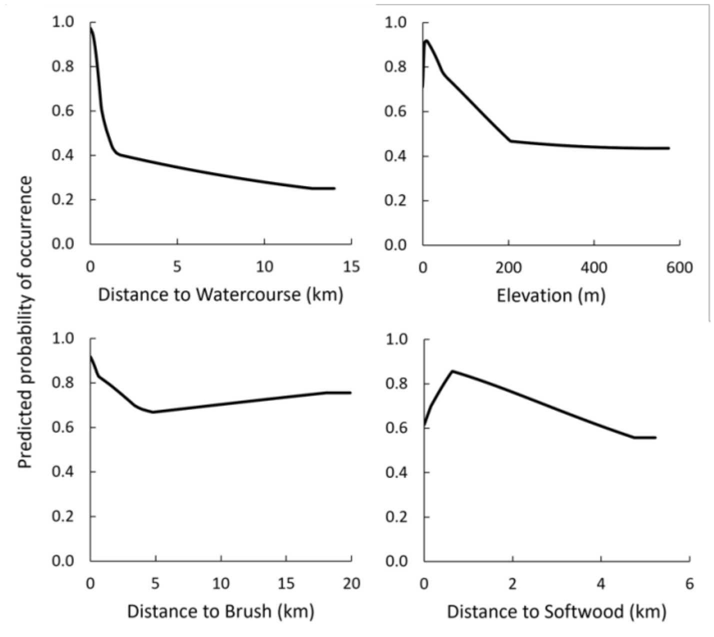

Figure 4.1 Model response curves for the geomorphologic and forest stand type species-covariate relationships for the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada. Response curves derived from an arithmetic average of 10 replicate cross validated Maxent models.

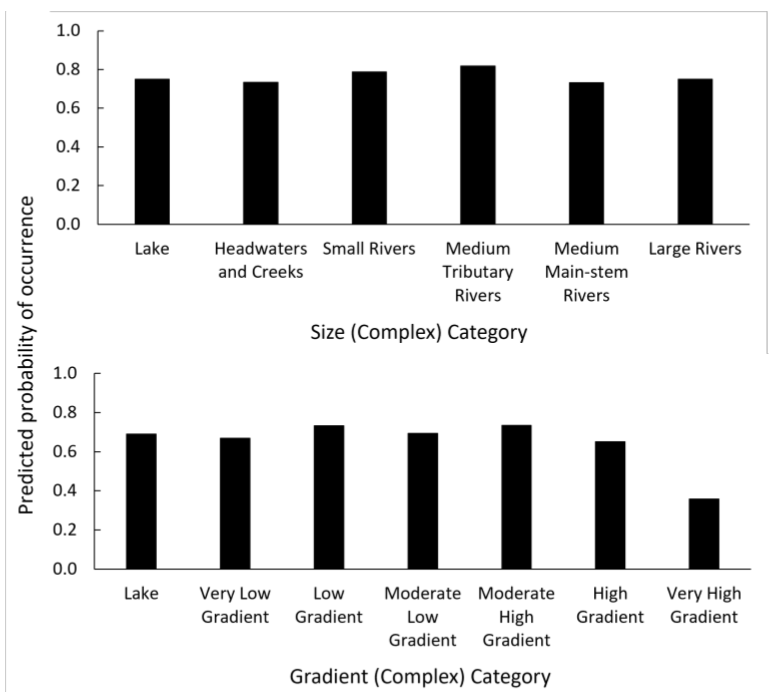

Figure 4.2 Model response curves for the categorical watercourse characteristic species-covariate relationships for the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada. Response curves derived from an arithmetic average of 10 replicate cross validated Maxent models.

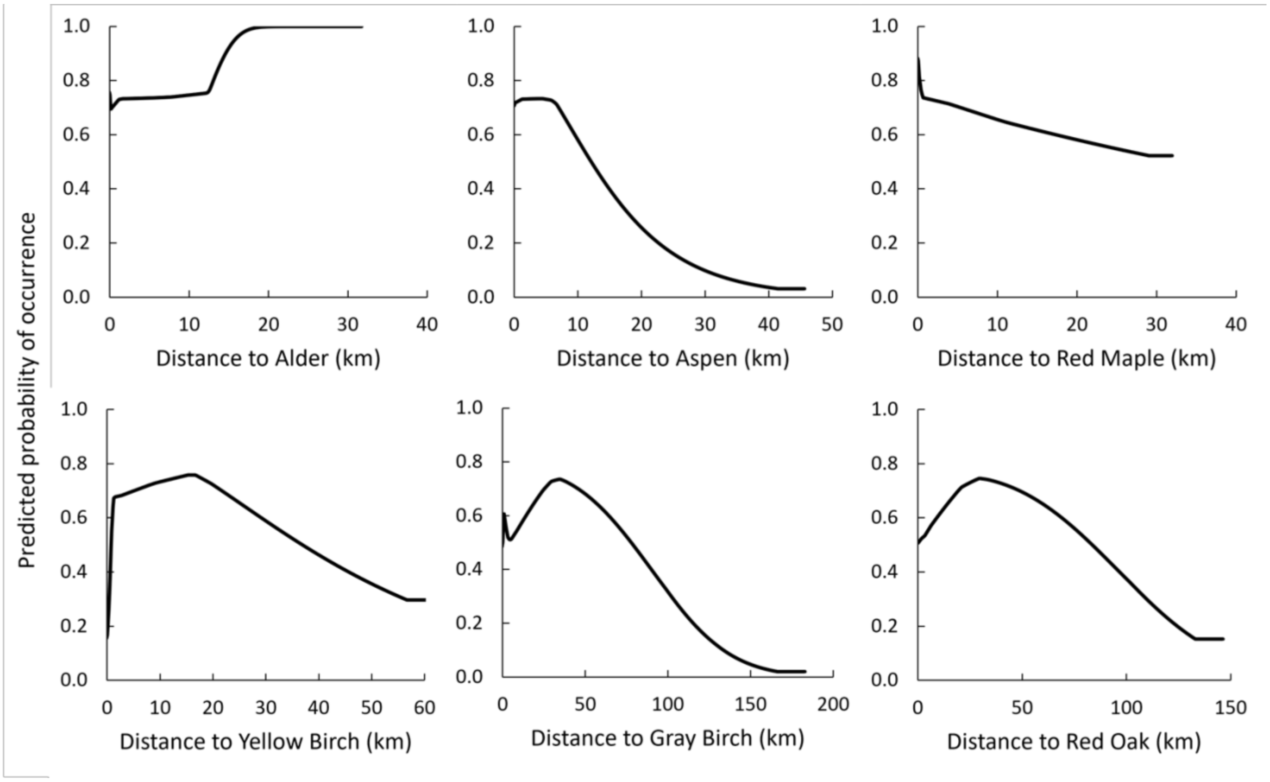

Figure 4.3 Model response curves for the tree stand-specific species-covariate relationships for the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada. Response curves derived from an arithmetic average of 10 replicate cross validated Maxent models. Each panel represents the variation in predicted probability of occurrence as distance increases from tree stands dominant at 60% or higher with the listed species.

Forest types and specific species combined accounted for 53.6% of the permutation importance of the ecological model, of which brush and softwood stands explained 8.3% (Table 4.1). *C. canadensis* were predicted to occur between 0 and 1 km of brush forest stands, as predicted occurrence decreased as distance increased between 1 and 5 km (Figure 4.1). They were most likely to occur approximately 800m from softwood stands, as predicted occurrence decreased for closer and further proximity, indicating that there is likely a preferable forest type which exists most often between known *C. canadensis* presences and softwood forest stands (Figure 4.2). Brush stands had a higher permutation importance than softwood stands and contributed more to predicted distribution (Table 4.1).

Specific tree species had a combined permutation importance of 45.3%, which explained nearly half of the variation in the distribution (Table 4.1). Predicted occurrence was approximately 0.75 adjacent to Alder stands, which remained relatively constant until distance increased to 15 km, where the predicted occurrence increased indefinitely (Figure 4.3). For Aspen, predicted occurrence was highest between 0 and 7 km, after which it dropped to nearly 0.0 at 40 km (Figure 4.3). Red Maple exhibited a similar response, where predicted occurrence was very high directly adjacent to these stands, after which it decreased to approximately 0.5 as distance increased (Figure 4.3). The probability of occurrence peaked between 1 and 20 km away from Yellow Birch, and approximately 40 km away from Gray Birch (Figure 4.3). For Red Oak stands, probability of occurrence is highest at 35 km away from these stands and is equal to 0.51 adjacent to these stands (Figure 4.3). Of the specific tree species included, Alder, Aspen, and Red Maple had the highest predicted occurrence directly adjacent to the stands.

The jackknife test of the AUC on test data shows a comparison between model performance using only a specific covariate, and model performance using the environmental covariate group data layer without that specific covariate (Table 4.2). The removal of the Watercourse covariate from the model would result in the greatest decrease in model AUC and it similarly had the highest AUC when considered on its own (Table 4.2). Among the other strong predictors were Aspen, Elevation, Gray Birch and Red Oak as the removal of these covariates in the model drop the AUC value by 0.01, indicating that these covariates increase the capacity of the model to predict test data (Table 4.2). This shows that these five covariates are strong predictors in the model, and that their contributions to the model are not explained in the other covariate layers. The weakest predictors of distribution when considered on their own were Red Maple, Alder, Aspen, Red Oak, and Yellow Birch, with AUC values less than 0.6 when running single covariate models (Table 4.2).

Table 4.2 The jackknife test results of the area under the receiver operating characteristic curve (AUC) on test data for the ecological model of predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada. Responses derived from an arithmetic average of 10 replicate cross validated Maxent species distribution models.

Environmental Covariate	AUC With Only Covariate	AUC Without Covariate
Alder	0.57	0.80
Aspen	0.58	0.79
Brush	0.63	0.81
Elevation	0.66	0.79
Gradient (Complex)	0.61	0.81
Gray Birch	0.60	0.79
Red Maple	0.55	0.81
Red Oak	0.58	0.79
Size (Complex)	0.62	0.81
Softwood	0.62	0.80
Watercourse	0.74	0.77
Yellow Birch	0.59	0.80
All Covariates		0.81

4.1.2 Human footprint model

The human footprint model for *C. canadensis* in Nova Scotia had a high averaged AUC for the replicated runs (0.83 +/- 0.044) (Appendix C). This AUC is between 0.8 and 0.9 and can be considered excellent (Hosmer, 1989 in Mandrekar, 2010). There were 13 environmental covariates used in the model that had varying contributions to the predicted probability of occurrence (Table 4.3). The addition of the human footprint layer resulted in a decrease in the permutation importance of 'Watercourse' from 26.6% in the ecological model to 22.2% in the human footprint model (Table 4.1 and Table 4.3). 'Watercourse' had the highest permutation importance (22.2%), though the human footprint layer had the second highest (18.8%), which superseded the rest of the environmental covariates (Table 4.3). The removal of Aspen, Elevation, Gray Birch, Red Oak, Watercourse, and the Human Footprint Index resulted in a drop of the AUC on test data (Table 4.4).

Table 4.3 Human footprint model percent contribution and permutation importance for 13 covariates averaged across 10-cross validated Maxent replicate models used to predict American beaver (*Castor canadensis*) probability of occurrence in Nova Scotia, Canada.

Environmental Covariate	**Average Percent Contribution (%)**	**Average Permutation Importance (%)**
Watercourse	43.2	22.2
Human Footprint Index	31.4	18.8
Gray Birch	3.2	11.7
Elevation	1.8	9.1
Aspen	3.9	9.0
Red Oak	1.8	6.9
Alder	1.6	6.7
Yellow Birch	2.4	4.6
Softwood	4.7	3.3
Gradient (Complex)	2.5	3.0
Brush	0.9	2.1
Red Maple	0.4	1.8
Size (Complex)	2.2	0.8

Table 4.4 Human footprint model jackknife test results of the area under the receiver operating characteristic curve (AUC) on test data for the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada. Responses derived from an arithmetic average of 10 replicate cross validated Maxent species distribution models.

Environmental Covariate	**AUC With Only Covariate**	**AUC Without Covariate**
Alder	0.56	0.83
Aspen	0.58	0.82
Brush	0.63	0.83
Elevation	0.66	0.82
Gradient (Complex)	0.62	0.83
Gray Birch	0.59	0.82
Red Maple	0.55	0.83
Red Oak	0.58	0.82
Size (Complex)	0.63	0.83
Softwood	0.63	0.83
Watercourse	0.74	0.80
Yellow Birch	0.59	0.83
Human Footprint Index	0.72	0.82
With all covariates		0.83

Based on the ecological model covariate response curves, each covariate response curve generated in the human footprint model exhibited a similar species-environment relationship to those identified previously (Appendix C). Alternatively, the human footprint index presented a 0.28 probability of occurrence at value of 0, meaning that areas with no human footprint had a 0.28 predicted occurrence (Figure 4.4). The probability of occurrence peaked at an index value of approximately 30, with a near-1.0 occurrence (Figure 4.4). Beyond an index value of 30, the probability of occurrence dropped to 0.82 (Figure 4.4). This model response curve shows areas where the human footprint index value is equal to 30 is where the highest predicted occurrence for *C. canadensis* is, and areas with a human footprint index value of zero are areas of low predicted occurrence (Figure 4.4).

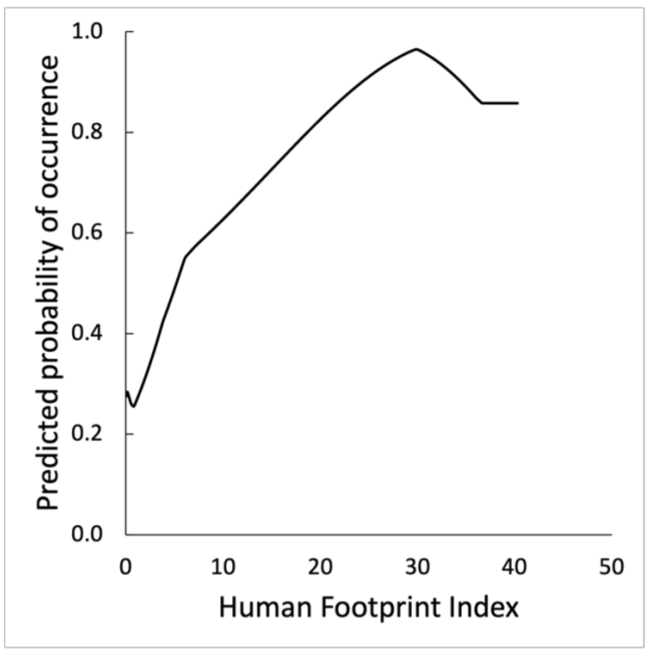

Figure 4.4 Model response curve for the human footprint index for the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada. Response curves derived from an arithmetic average of 10 replicate cross validated Maxent models.

4.2 Predicted distribution

Both models used the Maxent default output format of ‘cloglog’ or ‘complimentary log-log regression’, which produced a probability of occurrence between 0 and 1 for the study area (Phillips et al., 2017). The spatial distribution of the probability of occurrence for the ecological model and the human footprint model are shown in Figure 4.5.

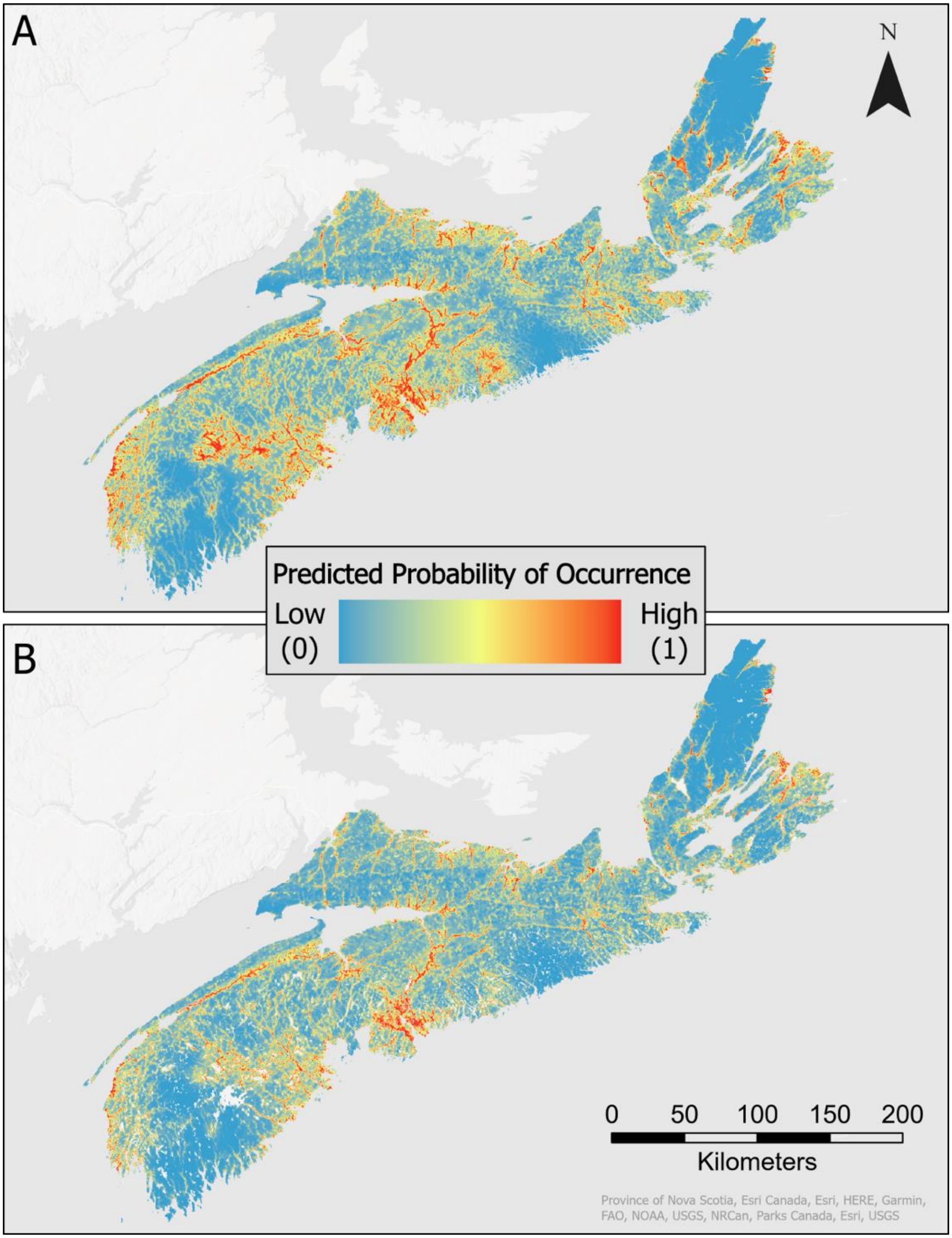

Figure 4.5 Comparison of the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada where panel A) ecological model; and B) human footprint model, derived by running the ecological model with a resampled cumulative anthropogenic threat layer from the Canadian Human Footprint Index. Both models derived from an arithmetic average of 10 replicate cross validated Maxent models.

The ecological model reveals patterns in the predicted distribution of occurrence throughout Nova Scotia based on geomorphologic and forest characteristics, shown in Panel A (Figure 4.5). In the ecological model, the areas with the highest predicted probability of occurrence appear to follow watercourses in a branch-like pattern (Figure 4.5). Large high probability areas appear to surround but not include the Halifax peninsula, extending west towards Dartmouth, Cole Harbour, and Lawrencetown, and north towards Fall River and Beaver Bank (Figure 4.5). A large high probability region follows the Shubenacadie River and nearby watercourses towards Cobequid Bay in Minas Basin (Figure 4.5). Kejimkujik National Park is a large high probability area which extends southeast towards Greenfield (Figure 4.5). On Cape Breton Island, areas of high probability include Tangier Grand Lake Wilderness Area, Lake Ainslie, Glace Bay, and Sydney (Figure 4.5). Tatamagouche, northwest of Truro, and the St. Croix River which feed into Minas Basin were also high probability areas for *C. canadensis* occurrence (Figure 4.5). Similarly, the areas surrounding the Annapolis River were shown as areas with high predicted occurrence (Figure 4.5).

In the ecological model output, there are large areas where the model predicted very low or no occurrence (Figure 4.5). These low or no occurrence areas included the area east of Kemptville and the Tobeatic Wilderness Area to the southwestern tip of Nova Scotia, the wilderness areas between Shelbourne and Sherbrooke, and the Cape Breton Highlands (Figure 4.5).

In panel B, the human footprint model reveals how the predicted distribution changes when the Human Footprint covariate was added (Figure 4.5). In this model, the densest area with the highest predicted probability of occurrence appears to surround but not include the Halifax peninsula (Figure 4.5). A large high probability region follows the Shubenacadie River towards Truro but appears to include fewer tributaries than the ecological model and fewer areas with very high predicted occurrence (Figure 4.5). Kejimkujik National Park appears to have moderate predicted occurrence, which extends southeast towards Greenfield (Figure 4.5). On Cape Breton Island, areas of high probability include Glace Bay, Ingonish and Sydney (Figure 4.5). The areas surrounding the Annapolis River were also shown as areas with high predicted occurrence (Figure 4.5).

The large areas identified as low predicted occurrence in the human footprint model are similar to those identified in the ecological model, which the high predicted occurrence areas are dissimilar (Figure 4.5). Both model outputs show regions of predicted occurrence >0.5 and areas with low or no predicted occurrence in similar regions, seen in the overlap of the location of large blue regions (Figure 4.5). However, the extent of high probability of occurrence areas are reduced greatly in the human footprint model, seen in the reduction of coverage of areas shown in red (Figure 4.5).

4.3 Model validation

For *C. canadensis* in Nova Scotia, all available occurrence data were combined to one occurrence data layer; however, the NSFI 'beaver flowage' data layer was used to externally validate the ecological model. The location of the *C. canadensis* flowage polygons as well as the occurrence data layer are displayed in Figure 4.6.

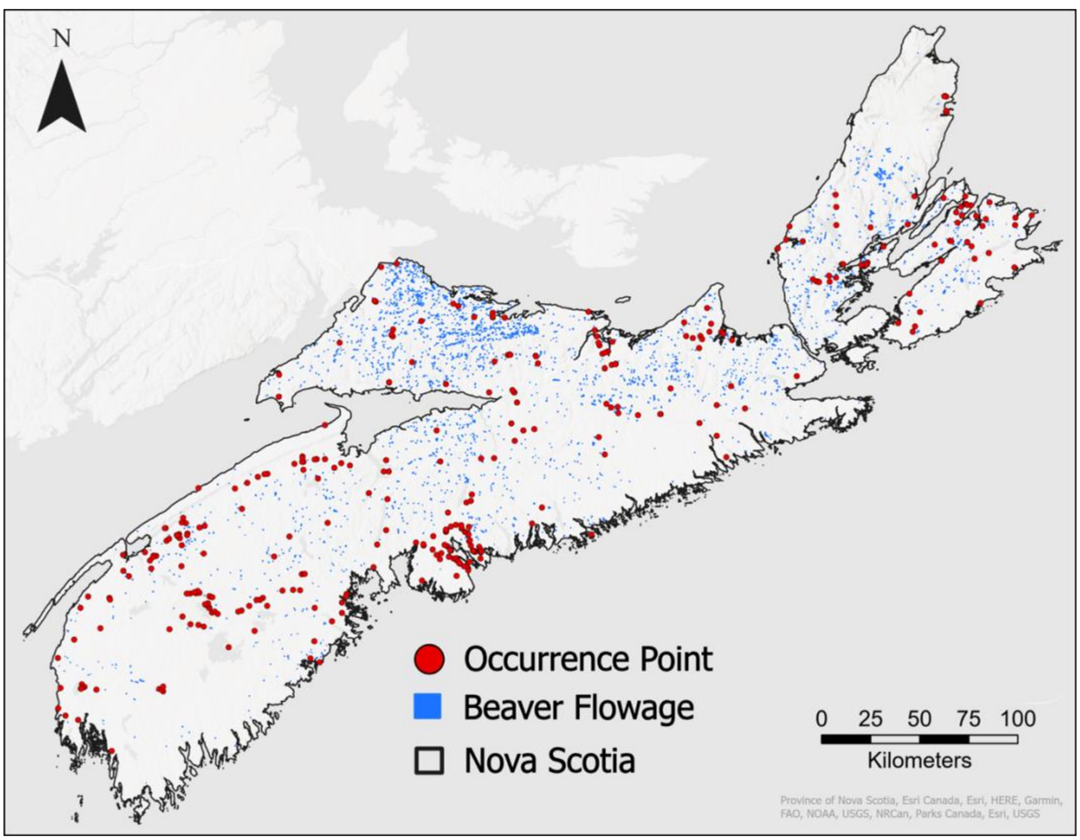

Figure 4.6 Locations American beaver (*Castor canadensis*) occurrence points used to build the Maxent SDM in Nova Scotia, Canada and the location of historic beaver flowage.

For this validation analysis, only the ecological model was sampled. The beaver flowage polygons are not evenly distributed in the province and have a high visual concentration near the northwestern portion of Nova Scotia (Figure 4.6). The occurrence points seem to appear in areas in reasonable proximity to beaver flowage polygons, however, *C. canadensis* flowage polygons appear in many areas where there are no occurrence points (Figure 4.6). This comparison shows that the areas with no occurrence points (Figure 4.6), the areas of low predicted probability in both models (Figure 4.5), and the areas with no or few *C. canadensis* flowage polygons represent similar landscape-scale regions (Figure 4.6). These three regions in the southern tip, northern tip, and the wilderness areas between Shelbourne and Sherbrooke have low representation of beaver use or predicted occurrence, compared to the rest of Nova Scotia (Figure 4.6).

Of the 414 occurrence points used in the SDM, only 6 were found within a beaver flowage polygon (1.45%) (Table 4.5). Similarly, when a 100 m buffer was applied to the flowage polygons, 14 of the 414 points used were found within this area (3.38%) (Table 4.5). A 200 m buffer included 19 points (4.59%), a 500 m buffer included 32 points (7.73%), and a 1000 m buffer included 58 points (14.01%) (Table 4.5). An analysis of the mean nearest distance of the occurrence points to the beaver flowage polygons had a value of 3.7 km +/- 2.6, indicating that occurrence points are most often found over 3 km away from a flowage polygon.

Table 4.5 Five buffer distances around beaver flowage polygons where the percentage of the occurrence points used to build the Maxent ecological model were found within the polygons.

Polygon Buffer Type	Occurrence Points Within Flowage Polygons	Percentage (%)
No Buffer	6	1.45%
100 m Buffer	14	3.38%
200 m Buffer	19	4.59%
500 m Buffer	32	7.73%
1000 m Buffer	58	14.01%

Sampling the original flowage polygons, as well as those generated by the four buffer distances using ‘Zonal Statistics’ in ArcGIS Pro revealed the mean and maximum predicted probability within these polygons (Figure 4.9 & Figure 4.10). The mean values average between 0.20 and 0.25 for all the polygons sampled, and the spread of the values decreases as size of the polygons increases (Figure 4.7). Alternatively, the mean distribution of the maximum values within each polygon increases from no buffer to a buffer distance of 1000 m (Figure 4.8). The buffer distance with the highest mean maximum value of predicted occurrence was 1000 m, with a range including +/- 1SD between 0.38 and 0.86, indicating that the model often predicted suitable habitat >0.5 in a 1 km radius surrounding areas of historic beaver use (Figure 4.8). This value is greater than the mean probability of occurrence within all of Nova Scotia (Figure 4.9).

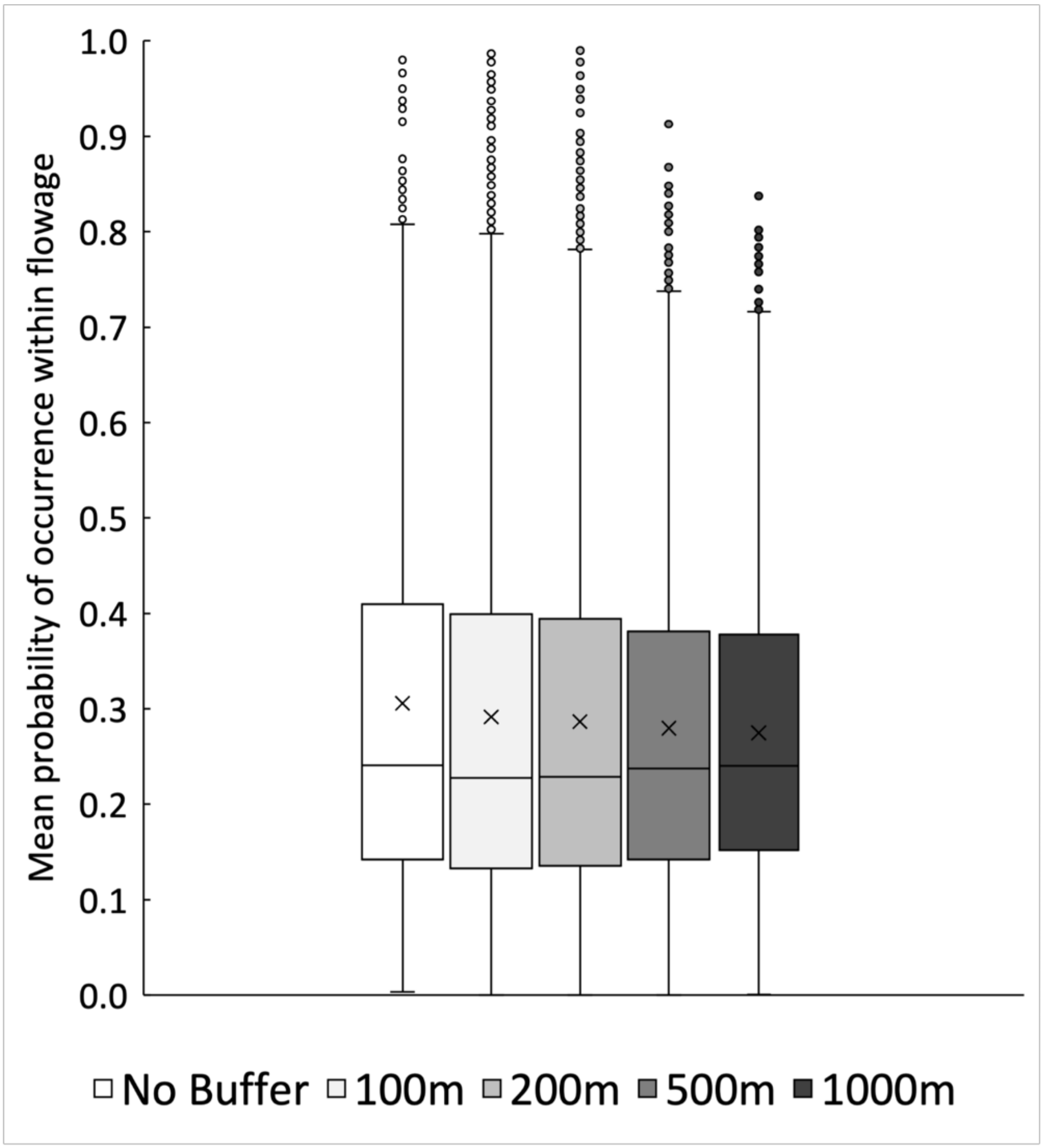

Figure 4.7 Mean predicted probability of occurrence value (0 – low, 1 – high) for the ecological model raster cells in areas associated with beaver flowage, using no buffer, 100m buffer, 200m buffer, 500m buffer, and 1000m buffers. The mean predicted occurrence in Nova Scotia is 0.27 +/- 0.24.

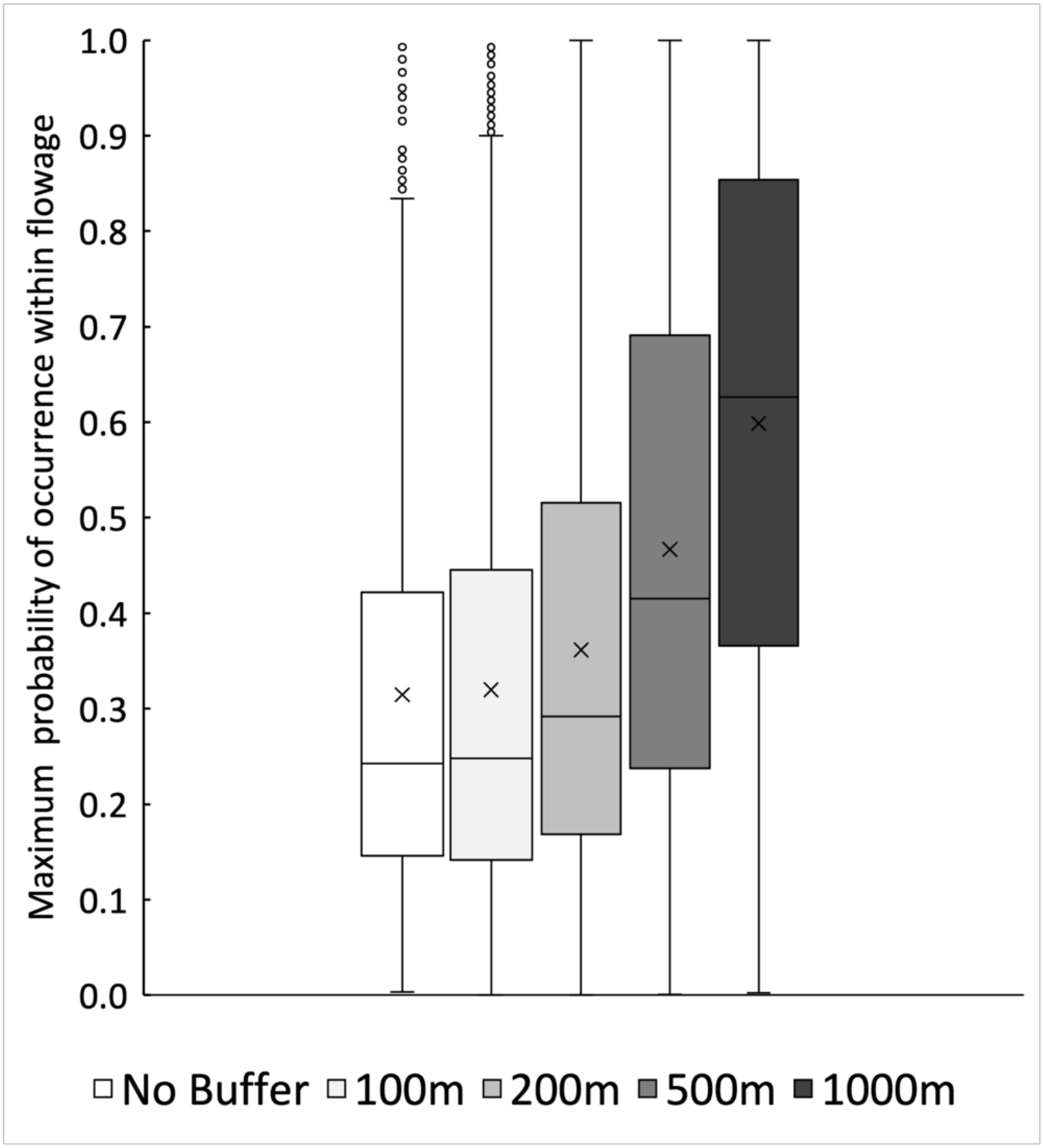

Figure 4.8 Mean distribution of the maximum predicted probability of occurrence value (0 – low, 1 – high) for the ecological model raster cells within areas associated with beaver flowage, using no buffer, 100m buffer, 200m buffer, 500m buffer, and 1000m buffers. The mean predicted occurrence in Nova Scotia is 0.27 +/- 0.24.

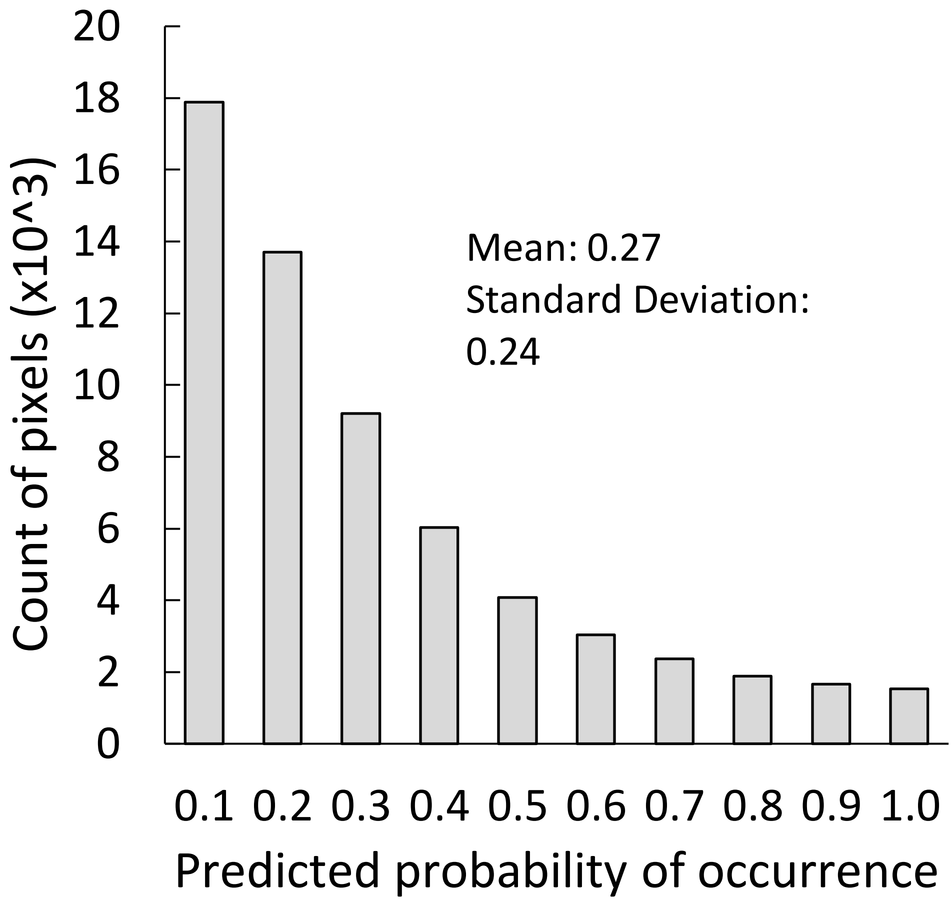

Figure 4.9 Histogram of the distribution of predicted probability of occurrence for *C. canadensis* in the ecological model in Nova Scotia (mean = 0.27 +/- 0.24).

To understand how the ecological model predicted occurrence in relation to areas of historic use, each raster cell within Nova Scotia was given a value representing the distance to the nearest flowage polygon using the 'Euclidean Distance' tool in ArcGIS Pro. A map of the distance to flowage across Nova Scotia is provided in panel A, compared to the ecological model raster in panel B (Figure 4.10). Visualization of the distance to flowage produced an output visually similar to the ecological model, where the furthest distances from areas of known beaver flowage (dark blue) align with the lowest predicted occurrence areas (light blue). This relationship suggests that on a landscape scale, the Cape Breton Highlands, the Fundy Coast, the Rossignol and Clyde River regions, and the area near Sheet Harbour are both furthest from areas historically occupied by beaver and have the lowest probability of occurrence based on the ecological model (Figure 4.10).

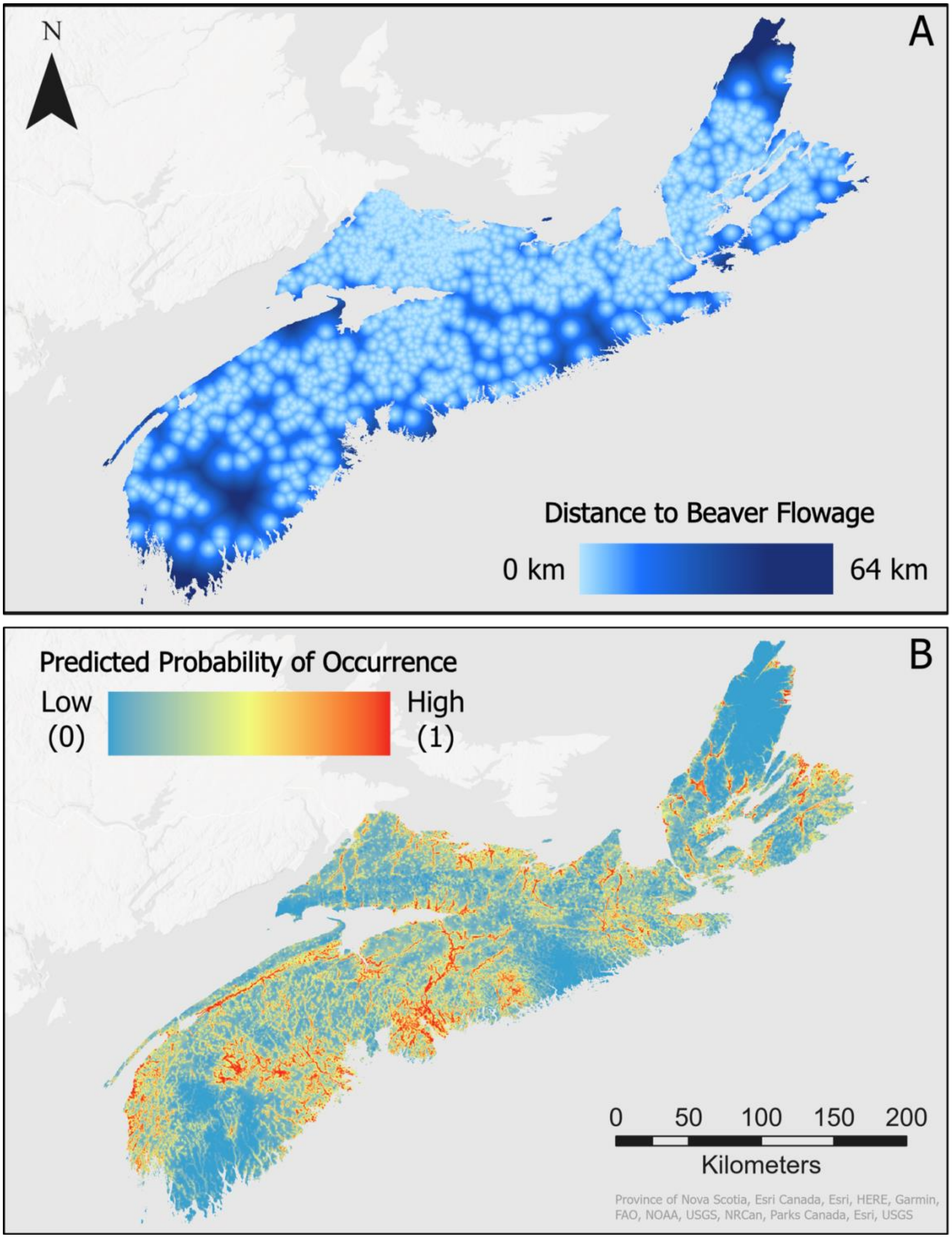

Figure 4.10 Comparative map depicting the distance to beaver flowage in Nova Scotia, Canada (panel A) and the ecological model of predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada (panel B). Panel A was derived using Euclidean Distance to flowage polygons. Panel B is the output of an averaged 10 replicate cross validated Maxent model.

CHAPTER 5 - DISCUSSION

This study created the first SDM for *C. canadensis* in Nova Scotia using covariates of known importance across their continental range. This study also investigated the relationship between predicted distribution and both human footprint data and areas of historic *C. canadensis* occupancy. The objectives of the study were to augment poor quality occurrence data in order to identify the landscape-scale drivers of distribution, identify areas of high or low predicted occurrence, and investigate the relationship of the human footprint on distribution. Succinctly, all objectives of the study were met.

5.1 Landscape-scale drivers

The ecological model showed that proximity to watercourse contributed the most to the predicted distribution (Table 4.1 and Table 4.2). This was expected as *C. canadensis* are wetland dam builders and are reliant on suitable watercourse habitat (Müller-Schwartze, 2011). The watercourse characteristics with the highest predicted probability of occurrence were those with low to moderate-high gradients, with the least preferable being very high gradients (Figure 4.2). These findings parallel those of past studies conducted outside of Nova Scotia, which identified a preference for low stream gradients, and an aversion to relatively higher gradients (e.g. Howard & Larson, 1985; Beier & Barret, 1987; McComb et al, 1990). Due to the spatial resolution of the input data and subsequent analysis, this study does not infer specific site suitability but rather identifies areas where watercourses are likely to provide suitable gradients. Elevation contributed the second most to the predicted distribution, where *C. canadensis* were more likely to occur at lower elevations, as water flows towards these low elevation areas due to gravity (Table 4.1 and Figure 4.2).

The influence of forest composition on the habitat selection of *C. canadensis* is an intrinsically difficult covariate type to model. *C. canadensis* are a generalist species with varied diets (Jenkins, 1981; Beier & Barrett, 1987), and characteristics of riparian vegetation are difficult to represent without field data (Franklin, 1995; Jakes et al. 2007; Lapointe St-Pierre et al., 2017). There are limitations associated with aerial interpretation; however, the NSFI used to represent the forest composition covariates is the best available data in Nova Scotia representing these pre-classified environmental conditions (Franklin, 1995; NSDNR, 2021). The forest-related habitat associations are derived from the NSFI, and specific tree species covariates only represent stands dominant at 60% or higher (Appendix A). *C. canadensis* may be equally likely to exist in areas where the tree species are present at lower percentages, as *C. canadensis* do not consume a whole stand worth of trees, but rather selectively forage a few (Johnston & Naiman, 1990). However, selecting a threshold of 60% dominance or higher ensured the model was mostly reflecting the tree species of interest.

Of the forest stand types, *C. canadensis* were predicted to occur adjacent to brush coverage, and 800m away from softwood stands (Figure 4.1). The brush forest class represents open areas with at least 25% woody plant cover (NSDNR, 2021). This is consistent with previous findings which emphasize the importance of deciduous woody plants and shrubs as a limiting factor of site suitability (Grinnell et al., 1937; Allen, 1983; Baker, 2003). The species response to softwood stands indicated that it is likely that a preferable forest stand type exists between softwood and *C. canadensis* occurrence (Figure 4.1). The NSFI classifies forest stand types as softwood, mixedwood, or hardwood (NSDNR, 2021). The negative response to softwood most likely indicates that the preferable stand type is therefore mixedwood or hardwood (NSDNR, 2021). The 'Hardwood' covariate was removed due to low predictive capacity overall, which may indicate that it is not the general category of hardwood forests, but individual hardwood species within the forest that contribute to predicted distribution (Table 4.1 and Table 4.2).

C. canadensis had the highest probability of occurrence directly adjacent to Red Maple, Alder, and Aspen stands, of the specific species included in the model (Figure 4.3). This finding is consistent with Doucet & Fryxell (1993), who determined the primary tree species of interest to *C. canadensis* in Ontario were aspen, alder, and red maple, of those provided. However, Red Maple had a very low permutation importance (Table 4.1). This may be because red maple is dominant throughout all of Nova Scotia and can occur in both softwood and hardwood dominant forest types (Webb & Marshall, 1999). Aspen was the species with the highest permutation importance which had a predicted probability of occurrence over 0.5 adjacent to the dominant stands (Table 4.1 and Figure 4.3). This finding parallels Jenkins (1981) and Müller-Schwarze (2011), who conclude that aspen is a preferable species when present and abundant. The predicted probability of occurrence is approximately 0.5 or less adjacent to stands of yellow birch, gray birch, and red oak (Figure 4.3). However, these tree species had relatively high permutation importance, which likely indicates that their absence is preferable, or the predicted composition of suitable habitat is generally further from habitat including these species (Table 4.1).

Answering my primary research question, the ecological model of predicted *C. canadensis* occurrence in Nova Scotia derived the main landscape-scale drivers of their geographic distribution. We can reject the null hypothesis and accept the alternative hypothesis that each landscape-scale covariate included in the model has an influence over explaining the variability in predicted distribution.

5.2 Predicted distribution

The ecological model generated a unique predicted distribution across the Nova Scotian landscape. We can reject the null hypothesis of the second research question and accept the alternative hypothesis that predicted probability of occurrence of *C. canadensis* is unevenly distributed. The ecological model and the human footprint model, as well as the distance to beaver flowage revealed four large areas which were identified as low predicted occurrence with few beaver flowage ponds. These areas included the Fundy Coast, the Cape Breton Highlands, areas near Sheet Harbour, and the Southwest Nova Scotia uplands (Rossignol and Clyde River). These are landscape-scale regions, corresponding approximately to the terrestrial Ecoregions and Ecodistricts of Nova Scotia (Figure 5.1).

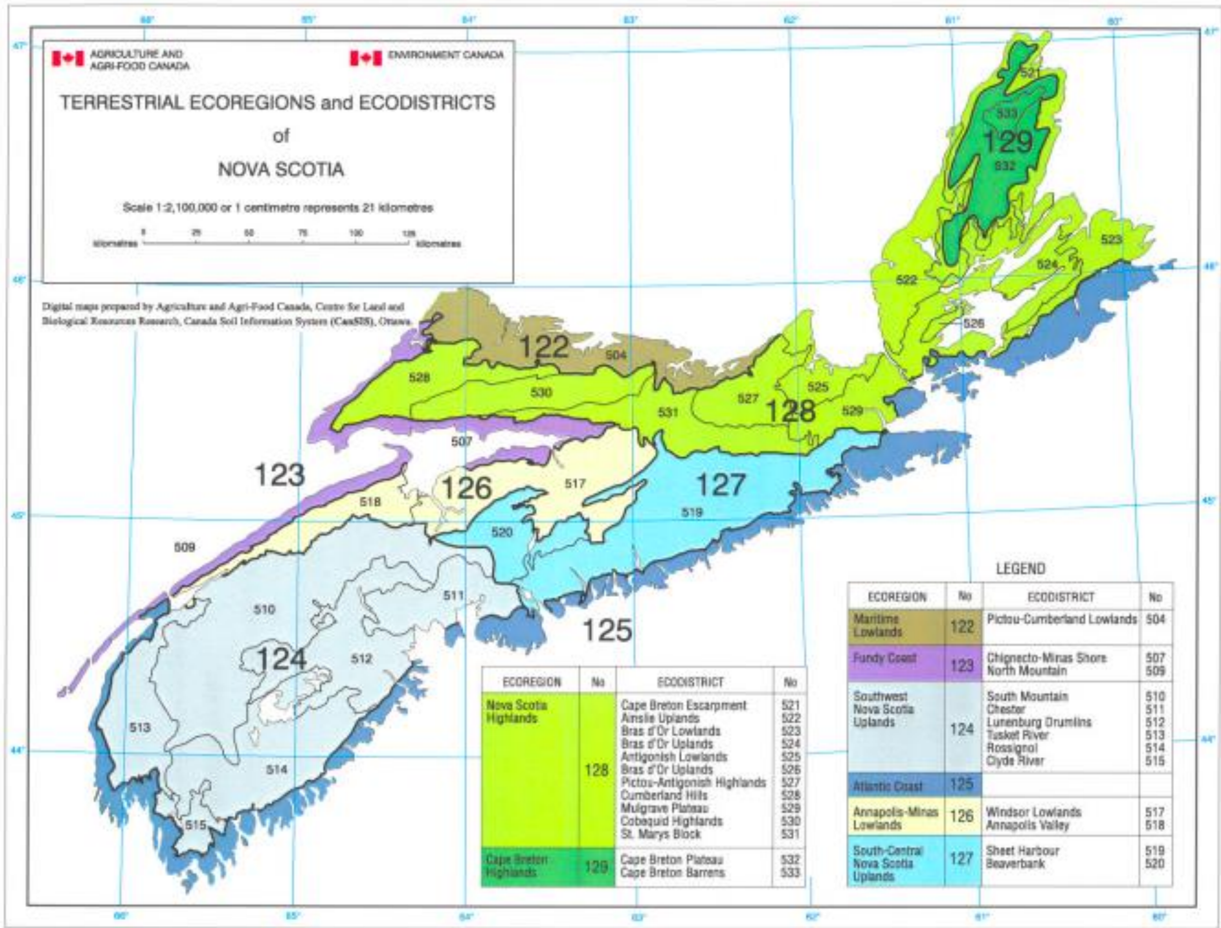

Figure 5.1 Terrestrial Ecoregions and Ecodistricts of Nova Scotia (adapted from Webb & Marshall, 1999).

In the Southwest Nova Scotia uplands Ecoregion, the Clyde River Ecodistrict (515) is characterized by softwood forests and barrens, and the Rossignol Ecodistrict (514) primarily used for forestry, has been extensively burned, and is dominated by coniferous species in the center and hardwoods toward the east (Webb & Marshall, 1999) (Figure 5.1). These combined factors may contribute to the low predicted distribution and the minimal historic beaver flowage in these areas, as we understand *C. canadensis* to prefer hardwood species (Müller-Schwartze, 2011). Similarly, the Cape Breton Highlands Ecoregion is characterized by high elevations and softwood forests, though we understand low elevations to be important to predicted distribution, which may explain the low levels of predicted occurrence in this region (Table 4.1 and Figure 4.1). We can deduct possible similar explanations for the Sheet Harbour Ecodistrict (519), where the low elevation areas support softwoods, and only the high elevation areas and hill-tops support hardwood species (Webb & Marshall, 1999). Based on the findings of the ecological model, the type of habitat available in the Sheet Harbour Ecodistrict would not be highly conducive to high predicted *C. canadensis* occurrence, as this model predicted higher occurrence in areas with low elevation, near watercourses, where hardwood species such as aspen are present.

5.3 Influence of human footprint

To answer the final research question, a human footprint model was generated to understand if human footprint influenced predicted distribution. We can reject the null hypothesis and accept the first alternative hypothesis that the human footprint had a positive relationship with predicted occurrence. The human footprint AUC was higher than the ecological model AUC, which means the added layer increased the ability of the model to predict omitted test data (Appendix C). The human footprint model predicted *C. canadensis* occurrence highest in places with a relatively high cumulative threat index value, and predicted occurrence was lowest in areas with minimal human footprint (Figure 4.5 and Figure 5.1). The human footprint layer includes factors such as built environments, crops and pastureland, night light pollution, population density, resource extraction, forestry, and roads (Hirsh-Pearson et al., 2022).

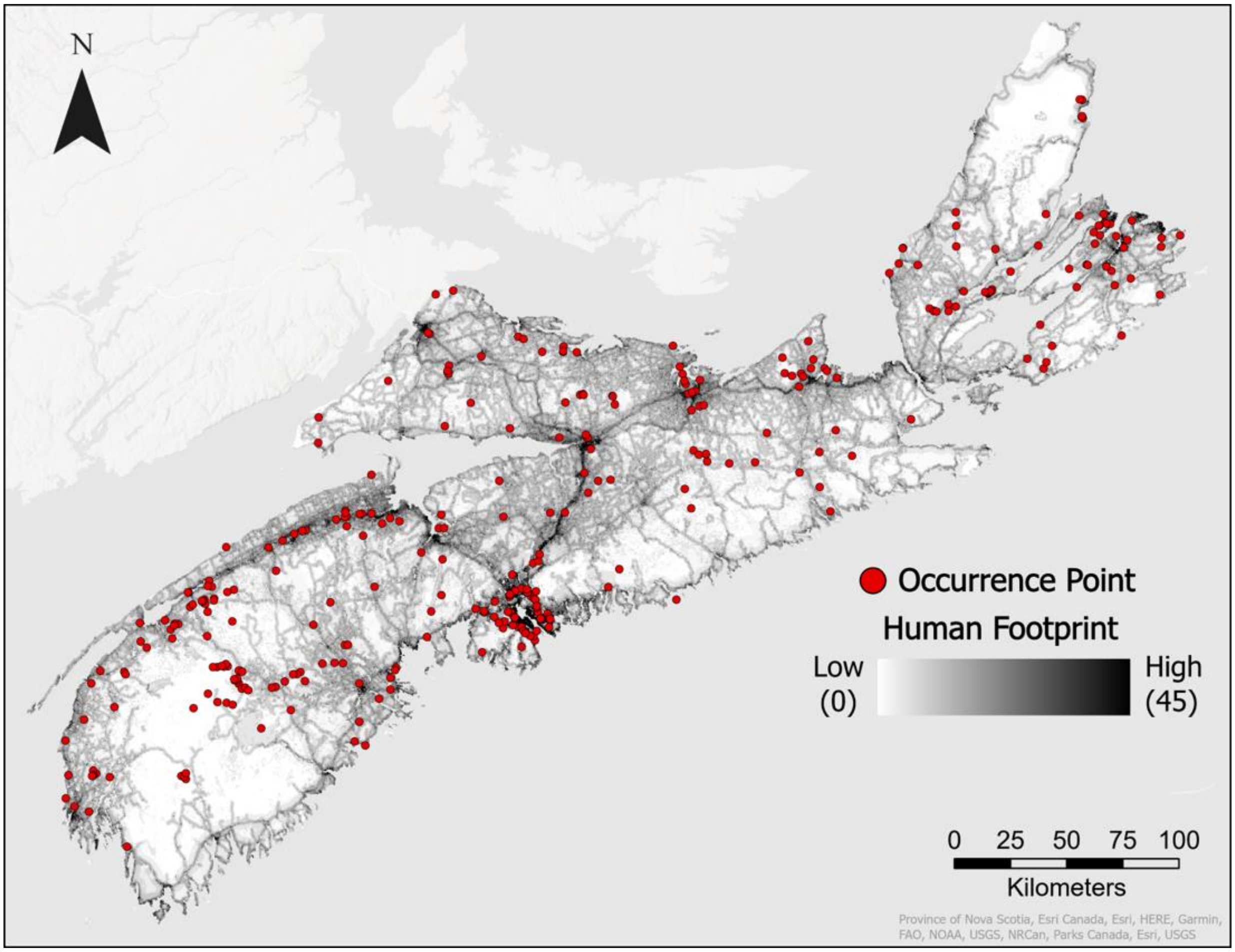

Figure 5.2 Comparison of the American beaver (*Castor canadensis*) occurrence data layer with the human footprint cumulative threat index in Nova Scotia, Canada.

The occurrence layer included primarily data from iNaturalist, a citizen-science generated observation platform. Citizen science data is known to have spatial and temporal biases due to differences in observer habits, and the frequent interactions between users and species in areas of high human population density, such as urban areas, along roads and highways, and within public parks (Boakes et al., 2010; Dickinson et al., 2010). Additionally, *C. canadensis* occupy wetland habitats which are often difficult to access, leading to fewer beaver-human interactions in undeveloped areas, and therefore these areas receive less sampling effort. A visual analysis of the occurrence points and the human footprint cumulative threat index value indicates a similar spatial distribution between beaver occurrence and high human footprint areas (Figure 5.1). It is possible that this relationship can be attributed to spatial bias in the occurrence data layer, that was largely generated through citizen science, resulting in greater beaver observations near roads and high population density. Consequently, this means the model is less representative of inaccessible areas, such as undeveloped wetlands. Interestingly, these areas with low levels of human footprint (Figure 5.2), few occurrence points (Figure 5.2), and low predicted occurrence (Figure 4.5) were relatively far from historic beaver flowage (Figure 4.10).

One fundamental assumption of Maxent is that the sampling is either random or representative throughout the study area (Phillips et al., 2009; Yackulic et al., 2013; Kramer-Schadt et al., 2013). However, in Yackulic et al. (2013) it was found that Maxent studies frequently violate this assumption when using real presence-only data. Regarding the iNaturalist data in this study, spatial heterogeneity of accessible areas near roads and urban areas resulted in the oversampling of these areas and under sampling of areas far removed from accessible centers (Dickinson et al., 2010). The comparison of the occurrence data layer and the human footprint shows this in practice (Figure 5.1). Understanding how this impacts the interpretation of the predicted distribution is critical to effectively convey how this modelling effort can be applied.

Tobler's First Law of Geography states that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). According to this law, as the occurrence data was predominantly collected in highly urban areas and along road networks, this led to the composition of habitat in these areas to be over-represented as inputs to the model, as there is a greater number of occurrence points collecting similar environmental conditions. Consequently, this means the model predicted occurrence biased towards these environmental conditions. Sampling bias is known to impact the accuracy of predictive models, and correcting for this bias has been shown to improve model accuracy (Kadmon et al., 2004; Bystriakova et al. 2012). The occurrence data in this study is an accidental, non-random sample, as not every *C. canadensis* individual had an equal probability of being sampled and the occurrence data is therefore over representative of a subset of *C. canadensis* individuals likely to encounter humans. As a result, the ecological model also overrepresents these covariate conditions and may not represent *C. canadensis* in their idealized niche habitat or across all habitable landscapes in Nova Scotia, but rather suitable habitat that is also conducive to beaver-human interactions. The human footprint model highlights this bias further to show areas of likely beaver-human interactions, as it effectively adds an input representing places of human occupation.

5.4 Limitations

The results of this study are the relative predicted probability of occurrence of *C. canadensis* across the Nova Scotian landscape, relative to the limitations of the study including the non-systematic sampling of occurrence points, the high detectability of *C. canadensis* in urban areas, and the constraints of the model inputs (Yackulic et al. 2013; Westwood, 2016). As noted in Westwood (2016), accurately reporting model limitations is important so we can understand the capacity of the model and be sure to not overstate its application.

In addition to the sampling bias associated with the nature of iNaturalist, there is user error associated with citizen scientists not logging the exact location of a *C. canadensis* individual in ideal riparian habitat, but rather logging their own location during the interaction, as pictures can be taken from far distances. Likewise, the tools used to collect the iNaturalist data are often not fieldwork GPS devices, which can introduce error logging the appropriate location. As the occurrence data was predominantly from a citizen-science database, spatial biases of the occurrence data were expected to impact the distribution of the model. As previously mentioned, this non-systematically collected occurrence data may not reflect the conditions of source habitat. While there are spatial filtering capabilities to account for this, these were not within the scope of this study.

Maxent used randomly generated background sample points to assess the values of the environmental covariate group data layer at areas which represent pseudo-absences (Phillips et al., 2017). As the occurrence data was not systematically collected, these non-directed 10,000 background points likely capture covariate values which may reflect highly suitable habitat which were not sampled due to their distance from accessible areas.

Limitations exist for using the *C. canadensis* flowage layer as an external dataset. While the beaver flowage polygons represent historic *C. canadensis* landscape modification, the polygons are derived from the NSFI, which was used to construct the forest-composition covariates. This means the model cells containing these flowage polygons were not associated with the covariate group data layer inputs, which may have contributed to the low predicted occurrence within the polygons when sampled. Additionally, the NSFI is interpreted from aerial imagery, which lends itself to significant human error associated with the visual classification of wetland areas (Gallant, 2015).

CHAPTER 6 - CONCLUSION

As a keystone species of significant ecological and biocultural importance, this study was an important first effort to understand how *C. canadensis* is distributed within the Wabanaki-Acadian Forest. A primary goal was to identify the local landscape-scale drivers of this distribution, as the province has unique habitat contrasting those previously studied. This study revealed the most significant drivers of *C. canadensis* distribution in the province, such as proximity to watercourses and aspen, low elevation, and distance from red oak and yellow birch. Additionally, this model reveals how predicted distribution compares to historic beaver flowage in the province and provides valuable insights into interactions between humans and *C. canadensis* today.

This modelling effort was not equally representative of the diversity of the Nova Scotian landscape, overrepresenting *C. canadensis*-environment interactions in urban areas and underrepresenting these interactions in natural, inaccessible areas due to the spatial bias of the occurrence data. For future studies to better understand habitat associations and distribution of *C. canadensis* throughout the province using existing data, I propose that these efforts account for these biases adequately (Yackulic et al., 2013). The least resource demanding options for this include manipulation of the occurrence data through spatial filtering, or the manipulation of the background sample points using a bias mask (Phillips et al., 2009; Yackulic et al., 2013; Kramer-Schadt et al., 2013). Spatially filtering the occurrence points to remove many points within a small area would limit the extent of the influence of these over-represented areas on the model predictions. I propose a future effort use a group of the Hirsh-Pearson et al. (2022) human footprint layers including 'built environments', 'human population density', and 'roads' to generate a spatial bias file, which can be used to direct background sample points to oversampled areas, which may more effectively represent 'sink habitat' conditions (Elith et al., 2010). In addition to this direction of background sample points, using external packages to adjust the Maxent model iterations and regularization multipliers could impact the accuracy of the model (Merow et al., 2013). Ground truthing this model would aid in its interpretation and how we can derive more conclusive measures of the strength of its predictions (e.g. West et al., 2016; Allen & McMullin, 2019; Westwood et al., 2019; Smith et al., 2021; Baker, 2022). Lastly, executing a true systematic or random sampling effort would generate the most reliable results for an ecologically reflective model across the Nova Scotian landscape.

While this study does not conclusively represent all *C. canadensis* distribution in Nova Scotia, it predicts occurrence based on likely human-beaver interactions. The relationship between humans and *C. canadensis* remains important, as a keystone species and a national icon (*National Symbol of Canada Act,* 1985). This model contributes to a better understanding of our relationship with a keystone species of great ecological integrity, Indigenous identity, and Canadian history, highlighting places where there are likely interactions between humans and *C. canadensis*. "All models are wrong, but some are useful" is a famous quote by George E. P Box (Box, 1980). This model, while fundamentally skewed towards areas of likely human-beaver interactions, remains a valuable new dataset and is the first of its kind in Atlantic Canada.

Currently, known efforts are underway to model biocultural connectivity within Nova Scotia (UINR, 2022). This model can be used as an input for models alike, for a species of significant ecological and cultural importance, to highlight those specific interactions between people and *C. canadensis*. This model can be a practical modelling tool for the AC CDC's Atlantic Canada Species at Risk Habitat Modelling Community of Practice, which aims to foster knowledge sharing in Atlantic Canada and support species at risk (AC CDC, 2023).

As the first SDM for *C. canadensis* in Nova Scotia, this model lays significant groundwork. The model and future renditions can contribute to informing practices such forest management, in the case of protecting wildlife habitat in timber harvests (Stoffyn-Egli & Duinker, 2013), or for urban land use planning. As climate change and anthropogenic pressures continue to strain wetland habitats (Salimi et al., 2021), beavers have emerged as a wetland restoration technique (DeVries et al., 2012). This keystone species is known to enhance valuable ecosystem services, such as increasing riparian biodiversity (Wright et al., 2002), sediment deposition (Pollock et al., 2007), and floodplain connectivity (Jordan and Fairfax, 2022). We can and should factor the protection of these ecosystem services into future development, and this model may act as a proxy for where *C. canadensis* are likely to safeguard these wetland services in development prone areas (Thompson et al., 2021; Larsen et al., 2021; Jordan & Fairfax, 2022).

REFERENCES

AC CDC (2022). Atlantic Canada Conservation Data Center Castor canadensis occurrence dataset [Dataset] Sackville, New Brunswick: Atlantic Canada Conservation Data Center Office.

AC CDC (2023). *Atlantic Canada Species At Risk Habitat Modelling Community of Practice.* Atlantic Canada Conservation Data Center. Retrieved from: https://atlantichabitatmodels.github.io/communityofpractice/

Aleksiuk, M. (1970). The Seasonal Food Regime of Arctic Beavers. *Ecology*, *51*(2), 264–270. https://doi.org/10.2307/1933662

Allen, A. (1983). *Habitat suitability index models: beaver.* Western Energy and Land Use Team Division of Biological Service Research and Development Fish and Wildlife Service U.S. *Semspub. FWS/OBS-82/10.30*, 20 pp. https://semspub.epa.gov/work/01/463451.pdf

Allen, J., & McMullin, R. (2019). Modeling algorithm influence on the success of predicting new populations of rare species: ground-truthing models for the Pale-Belly Frost Lichen (Physconia subpallida) in Ontario. *Biodiversity and Conservation*, *28*(7), 1853–1862. https://doi.org/10.1007/s10531-019-01766-z

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), 11 pp. https://doi.org/10.1126/sciadv.aat4858

ArcGIS (n.d.). *ArcGIS: How Band Collection Statistics Work*. ESRI ArcGIS Pro. Retrieved March 31, 2023, from https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/how-band-collection-statistics-works.htm

Atwater, M. (1940). South Fork (Montana) Beaver Survey: 1939. *The Journal of Wildlife Management*, *4*(1), 100 pp. https://doi.org/10.2307/3796274

Austin, M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. In *Ecological Modelling*, *200*(1–2), 1–19. https://doi.org/10.1016/j.ecolmodel.2006.07.005

Baker, B. (2003). Beaver (*Castor canadensis*) in heavily browsed environments. *U.S. Geological Survey, Fort Collins Science Center, Fort Collins, CO 80526-8118, USA*, *46*(2), 173–181. https://www.semanticscholar.org/paper/Beaver-(Castor-canadensis)-in-Heavily-Browsed-Baker/d956106870f59fe0210a1ac32fa29f3c01515434

Baker, B., & Hill, E. (2003). *Wild Mammals of North America: Biology, management, and conservation: Beaver (Castor canadensis)*. Johns Hopkins University Press. 288-310. https://pubs.er.usgs.gov/publication/87287

Baker, D., Garnett, S., O'Connor, J., Ehmke, G., Clarke, R., Woinarski, J., & McGeoch, M. (2019). Conserving the abundance of nonthreatened species. *Conservation Biology*, *33*(2), 319–328. https://doi.org/10.1111/cobi.13197

Baker, T. (2022). *Developing Species Distribution Models for Wood Turtle (Glyptemys insculpta) in Atlantic Canada*. [Master's thesis, Dalhousie University]. Dalspace. https://dalspace.library.dal.ca/bitstream/handle/10222/81940/ThomasBaker2022.pdf?sequence=3

Baldwin, J. (2013). Problematizing Beaver Habitat Identification Models for Reintroduction Application in the Western United States. *Association of Pacific Coast Geographers Yearbook, 75*, 104-119. https://www.beaverinstitute.org/wp-content/uploads/2017/08/Baldwin-2013-Beaver-Habitat-Identification-Models-for-Reintroduction-APCG-Yearbook-2013-2.pdf

Bale, S., Beazley, K., Westwood, A., & Bush, P. (2020). The benefits of using topographic features to predict climate-resilient habitat for migratory forest landbirds: An example for the Rusty Blackbird, Olive-sided Flycatcher, and Canada Warbler. *The Condor*, *122*(1), 1-19. https://doi.org/10.1093/condor/duz057

Barnes, D., & Mallik, A. (1997). Habitat factors influencing beaver dam establishment in a northern Ontario watershed. *Journal of Wildlife Management, 61*(4), 1371-1377. https://doi.org/10.2307/3802140

BCCVL (2021). *MAXENT Algorithm Information (SDMs)*. Biodiversity and Climate Change Virtual Laboratory. https://support.bccvl.org.au/support/solutions/articles/6000083216-

maxent#:~:text=Maxent%2C%20which%20stands%20for%20maximum,environmental%20variables%20of%20known%20locations

Beier, P., & Barrett, R. (1987). Beaver Habitat Use and Impact in Truckee River Basin, California. *The Journal of Wildlife Management*, *51*(4), 794 pp. https://doi.org/10.2307/3801743

Boakes, E., McGowan, P., Fuller, R, Chang-qing, D., Clark, N., O'Connor, K., & Mace, G. (2010). Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data. *PLoS Biology, 8*(6), 1-11. https://doi.org/10.1371/journal.pbio.1000385

Box, G. (1980). Sampling and Bayes' Inference in Scientific Modelling and Robustness. *Journal of the Royal Statistical Society. Series A (General)*, *143*(4), 1-383. https://doi.org/10.2307/2982063

Bystriakova, N., Peregrym, M., Erkens, R., Bezsmertna, O., & Schneider, H. (2012). Sampling bias in geographic and environmental space and its effect on the predictive power of species distribution models. *Systematics and Biodiversity*, *10*(3), 305–315. https://doi.org/10.1080/14772000.2012.705357

Carlos, A., & Lewis, F. (2011). Commerce by a Frozen Sea: Native Americans and the European Fur Trade. *University of Pennsylvania Press*, 1–14. http://econ.queensu.ca/pub/faculty/lewis/econ331/Carlos&Lewis-%20Commerce%20by%20a%20Frozen%20Sea-1.pdf

Cotton, F., & Oderwala, R. (1990). *Potential Beaver Colony Density In Parts of Quebec*. [Master's thesis, Virginia Tech]. Vtechworks. https://vtechworks.lib.vt.edu/handle/10919/42093

Currie, D. (1991). Energy and Large-Scale Patterns of Animal- and Plant-Species Richness. *The American Naturalist*, *137*(1), 27–49. https://doi.org/10.1086/285144

Curtis, P., & Jensen, P. (2004). Habitat features affecting beaver occupancy along roadsides in New York State. *Journal of Wildlife Management*, *68*(2), 278-287. https://ezproxy.library.dal.ca/login?url=https://www.proquest.com/scholarly-journals/habitat-features-affecting-beaver-occupancy-along/docview/234274225/se-2

Dahl, T. (1990). *Wetlands: Losses in the United States 1780's to 1980's*. U.S. Department of the Interior, Fish and Wildlife Service, Washington, DC, USA. 13 pp. https://www.fws.gov/wetlands/documents/wetlands-losses-in-the-united-states-1780s-to-1980s.pdf

Davidson, N. (2014). How much wetland has the world lost? Long-term and recent trends in global wetland area. *Marine and Freshwater Research*, *65*(10), 934 pp. https://doi.org/10.1071/MF14173

Dertien, J., Self, S., Ross, B., Barrett, K., & Baldwin, R. (2020). The relationship between biodiversity and wetland cover varies across regions of the conterminous United States. *PLOS ONE*, *15*(5), 1-18. https://doi.org/10.1371/journal.pone.0232052

DeVries, P., Fetherston, K., Vitale, A., & Madsen, S. (2012) Emulating Riverine Landscape Controls of Beaver in Stream Restoration, Fisheries, *37*(6), 246-255. https://doi.org/10.1080/03632415.2012.687263

Dickinson, J., Zuckerberg, B., & Bonter, D. (2010). Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics*, *41*, 149–172. http://www.jstor.org/stable/27896218

Dieter, C., & McCabe, T. (1989). Factors Influencing Beaver Lodge-site Selection on a Prairie River. *American Midland Naturalist*, *122*(2), 408 pp.. https://doi.org/10.2307/2425928

Dittbrenner, B., Pollock, M., Schilling, J., Olden, J., Lawler, J., & Torgersen, C. (2018). Modeling intrinsic potential for beaver (*Castor canadensis*) habitat to inform restoration and climate change adaptation. *PLOS ONE*, *13*(2), 15 pp. https://doi.org/10.1371/journal.pone.0192538

Donkor, N., & Fryxell, J. (2000). Lowland boreal forests characterization in Algonquin Provincial Park relative to beaver (Castor canadensis) foraging and edaphic factors. *Plant Ecology*, *148*(1), 1–12. https://doi.org/10.1023/A:1009860512339

Doucet, C., & Fryxell, J. (1993). The Effect of Nutritional Quality on Forage Preference by Beavers. *Oikos, 67*(2), 201–208. https://doi.org/10.2307/3545464

Ecological Stratification Working Group (1995). *A National Ecological Framework for Canada. Agriculture and Agri-Food Canada*. Research Branch, Centre for Land and Biological Resources Research and Environment Canada, State of the Environment Directorate, Ecozone Analysis Branch, Ottawa/Hull. ISBN: 0-662-24107-X https://sis.agr.gc.ca/cansis/publications/manuals/1996/A42-65-1996-national-ecological-framework.pdf

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*(4), 330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

ESRI Inc. (2021). ArcGIS Pro (Version 2.9.5). Esri Inc. Retrieved from: https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

Fauset, A. H. (n.d.). Folklore from the Half-Breeds in Nova Scotia*. The Journal of American Folklore, 38(148).*

Franklin, J. (1995). Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography, 19*(4), 74–499. https://journals.sagepub.com/doi/epdf/10.1177/030913339501900403

Gallant, A. (2015). The Challenges of Remote Monitoring of Wetlands. *Remote Sensing*, *7*(8), 10938–10950. https://doi.org/10.3390/rs70810938

Grinnel, J., Dixon, S., & Linsdale, J. (1937). *Fur-bearing Mammals of California: Their Natural History, Systematic Status, and Relations to Man*. University of California Press, 2, 777 pp. https://catalogue.nla.gov.au/Record/952105/Copyright

Guisan, A., & Zimmermann, N. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hirsh-Pearson, K., Johnson, C., Schuster, R., Wheate, R., & Venter, O. (2022). Canada's human footprint reveals large intact areas juxtaposed against areas under immense anthropogenic pressure. *FACETS, 7,* 398–419. https://doi.org/10.1139/facets-2021-0063

Historica Canada. (2022, September 7). *Mi'kmaq*. The Canadian Encyclopedia. https://www.thecanadianencyclopedia.ca/en/article/micmac-mikmaq

Hosmer, D. & Lemeshow, S. (1989) *Applied logistic regression*. Wiley. https://books.google.ca/books?hl=en&lr=&id=64JYAwAAQBAJ&oi=fnd&pg=PR13&ots=DtjK4W6mhN&sig=bDTNkLNPzRHkXK13CR-fjD8JYYo&redir_esc=y#v=onepage&q&f=false

Howard, R., & Larson, J. (1985). A Stream Habitat Classification System for Beaver. *The Journal of Wildlife Management*, *49*(1), 19. https://doi.org/10.2307/3801833

Hyvönen, T., & Nummi, P. (2008). Habitat dynamics of beaver Castor canadensis at two spatial scales. *Wildlife Biology*, *14*(3), 302–308. https://doi.org/10.2981/0909-6396(2008)14[302:HDOBCC]2.0.CO;2

iNaturalist (2022). Castor canadensis occurrence download open research grade Nova Scotia [Dataset]. https://www.inaturalist.org/observations/export?flow_task_id=308268

IPBES (2019). Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. *IPBES secretariat, Bonn, Germany,* 1148 pages. https://doi.org/10.5281/zenodo.3831673

IPCC (2001). Working Group II: Impacts, Adaptation and Vulnerability; Chapter 15. North America. https://archive.ipcc.ch/ipccreports/tar/wg2/index.php?idp=1

Jakes, A., Snodgrass, J., & Burger, J. (2007). *Castor canadensis* (Beaver) impoundment associated with geomorphology of southeastern streams. *Southeastern Naturalist*, *6*(2), 271–282. https://doi.org/10.1656/1528-7092(2007)6[271:CCBIAW]2.0.CO;2

Jenkins, S. (1981). Problems, progress, and prospects in studies of food selection by beavers. *Worldwide Furbearer Conference Proceedings Appalachian Environmental Lab,* 1, 559–579.

Johnston, C., & Naiman, R. (1990). Browse selection by beaver: effects on riparian forest composition. *Canadian Journal of Forest Research, 20*(7), 1036-1043. https://cdnsciencepub.com/doi/pdf/10.1139/x90-138

Jordan, C., & Fairfax, E. (2022). Beaver: The North American freshwater climate action plan. *WIREs Water*, *9*(4). https://doi.org/10.1002/wat2.1592

Kadmon, R., Farber, O., & Danin, A. (2004). Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, *14*(2), 401–413. https://doi.org/10.1890/02-5364

Kramer-Schadt, S., Niedballa, J., Pilgrim, J., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A., Augeri, D., Cheyne, S., Hearn, A., Ross, J., Macdonald, D., Mathai, J., Eaton, J., Marshall, A., Semiadi, G., Rustam, R., Wilting, A. (2013). The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, *19*(11), 1366–1379. https://doi.org/10.1111/ddi.12096

Lapointe St-Pierre, M., Labbé, J., Darveau, M., Imbeau, L., & Mazerolle, M. (2017). Factors Affecting Abundance of Beaver Dams in Forested Landscapes. *Wetlands*, *37*(5), 941–949. https://doi.org/10.1007/s13157-017-0929-x

Larsen, A., Larsen, J., & Lane, S. (2021). Dam builders and their works: Beaver influences on the structure and function of river corridor hydrology, geomorphology, biogeochemistry and ecosystems. *Earth-Science Reviews*, *218*, 46 pp. https://doi.org/10.1016/j.earscirev.2021.103623

Lee-Yaw, J., McCune, J., Pironon, S., & Sheth, S. (2022). Species distribution models rarely predict the biology of real populations. *Ecography, 2022*(6), 16 pp. https://doi.org/10.1111/ecog.05877

Mahoney, M., & Stella, J. (2020). Stem size selectivity is stronger than species preferences for beaver, a central place forager. *Forest Ecology and Management*, *475*, 12 pp. https://doi.org/10.1016/j.foreco.2020.118331

Mandrekar, J. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology*, *5*(9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d

McComb, W., Sedell, J., Buchholz, T., Mccomb, W., Sedell, J., & Buchholz, T. (1990). Dam-site selection by beavers in an eastern Oregon basin Recommended Citation. *Great Basin Naturalist 50*(3), 273-281. https://scholarsarchive.byu.edu/gbn/vol50/iss3/9

Menge, B. (1995). Indirect effects in marine rocky intertidal interaction webs: patterns and importance. *Ecological Monographs*, *65*(1) 21–74. https://doi.org/10.2307/2937158

Merow, C., Smith, M., & Silander, J. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, *36*(10), 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Michelson, T. (1925). Micmac Tales. *The Journal of American Folklore, 38*(147).

Millar, W., Olivero-Sheldon, A., Nussey, P., & Noseworthy, J. (2019). A Stream Classification for the Northern Appalachian–Acadian Region of Canada, Version 2.0. [Dataset]. Fredericton, New Brunswick: Nature Conservancy of Canada, Atlantic Regional Office. Retrieved from https://2c1forest.databasin.org/datasets/3fa5eb769b99496fad0c05c838c8823d/

Müller-Schwartze. (2011). *The Beaver: Its Life and Impact* (2nd ed.). Cornell University Press. https://www.jstor.org/stable/10.7591/j.ctt7zdm0

Mumma, M., Gillingham, M., Johnson, C., & Parker, K. (2018). Where beavers (*Castor canadensis*) build: testing the influence of habitat quality, predation risk, and anthropogenic disturbance on colony occurrence. *Canadian Journal of Zoology*, *96*(8), 897–904. https://doi.org/10.1139/cjz-2017-0327

Naiman, R., Johnston, C., & Pastor, J. (1988). Alteration of North American streams by beaver. *BioScience*, *38*, 750–765. http://www.jstor.org/stable/1310784

National Symbol of Canada Act, Pub. L. No. c. N-17 (1985).

NHMNS (n.d.). Natural History of Nova Scotia: The dynamics of Nova Scotia's climate. *Nova Scotia Museum Publications*, *1*, 94–103.

Northcott, T. (1964). *An investigation of the factors affecting carrying capacity of selected areas in Newfoundland for the beaver, Castor Canadensis caecator bangs.* [Master's thesis, Memorial University of Newfoundland]. MUN research library. https://research.library.mun.ca/7260/

Novak, M. (1987). *Wild furbearer management and conservation in North America*. Ontario Trappers Association & Ontario; Ontario Ministry of Natural Resources.

Nova Scotia Department of Natural Resources (2006) DP ME 55, Version 2, 2006, Enhanced Digital Elevation Model, Nova Scotia, Canada. Nova Scotia Department of Natural Resources (NSDNR). [Dataset]. https://novascotia.ca/natr/meb/download/dp055.asp

NSDNR (2008). State of the Forest: Report 1995-2005, Nova Scotia Forests in Transition. Nova Scotia Department of Natural Resources (NSDNR). Nova Scotia Natural Resources. https://novascotia.ca/natr/forestry/reports/State-Of-Forest-Report-April-2008.pdf

NSDNR (2016). *State of the Forest*. (Report No. 2). Nova Scotia Department of Natural Resources (NSDNR). Province of Nova Scotia. ISBN: 978-1-55457-737-8

NSDNR (2017). Ecological Land Classification for Nova Scotia. Nova Scotia Department of Natural Resources (NSDNR). https://novascotia.ca/natr/forestry/ecological/pdf/Ecological-Land-Classification-guide.pdf

Nova Scotia Department of Natural Resources and Renewables (2021). Nova Scotia Interpreted Forest Inventory - Current Forest Data (Web Version - 2021) [Dataset]. https://novascotia.ca/natr/forestry/gis/dl_forestry.asp

NWWG (1988): Wetlands of Canada. Ecological Land Classification Series #24, National Wetlands Working Group, Environment Canada, Ottawa, ON, Canada, 452 pp.

Parsons, E. (1925). Micmac Folklore. *The Journal of American Folklore, 38*(147).

Phillips, S., Dudik, M, & Schapire, R. (n.d.). [Internet] Maxent software for modelling species niche and distributions (Version 3.4.4). https://biodiversityinformatics.amnh.org/open_source/maxent/

Phillips, S., Dudík, M., & Schapire, R. (2004). A maximum entropy approach to species distribution modeling. *Twenty-First International Conference on Machine Learning - ICML '04*, 83. https://doi.org/10.1145/1015330.1015412

Phillips, S., Anderson, R., & Schapire, R. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3–4), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S., Dudík, M., Elith, J., Graham, C., 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

Phillips, S., Anderson, R., Dudík, M., Schapire, R., & Blair, M. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, *40*(7), 887–893. https://doi.org/10.1111/ecog.03049

Pollock, M., Beechie, T., Jordan, C. (2007). Geomorphic changes upstream of beaver dams in Bridge Creek, an incised stream channel in the interior Columbia River basin, eastern Oregon. *Earth Surface Processes and Landforms, 32*(8), 1174 – 1185

Retzer, J. (1956). Alpine soils of the rocky mountains. *Journal of Soil Science*, *7*(1), 22–32. https://doi.org/10.1111/j.1365-2389.1956.tb00858.x

Salimi, S., Almuktar, S., & Scholz, M. (2021). Impact of climate change on wetland ecosystems: A critical review of experimental wetlands. *Journal of Environmental Management*, *286*, 15 pp. https://doi.org/10.1016/j.jenvman.2021.112160

Seely, B. (2011). Nova Scotia Coordinate Referencing System (NSCRS): Future policy strategy. Service Nova Scotia and Municipal Relations, Geographic Information Services. Amherst, Nova Scotia. 49 pp.

Seton, J. (1929). Lives of Game Animals: Rodents, Etc. *Doubleday, Doran*, *4*(2).

Service Nova Scotia and Internal Services (2020). Nova Scotia Hydrographic Network [Dataset] https://data.novascotia.ca/Environment-and-Energy/Nova-Scotia-Hydrographic-Network/dk27-q8k2

Slough, B., & Sadleir, R. (1977). A land capability classification system for beaver (*Castor canadensis* Kuhl). *Canadian Journal of Zoology*, *55*(8), 1324–1335. https://doi.org/10.1139/z77-172

Smith, J., Kelly, N., & Renner, I. (2021). Validation of presence-only models for conservation planning and the application to whales in a multiple-use marine park. *Ecological Applications*, *31*(1). https://doi.org/10.1002/eap.2214

Statistics Canada (2016). 2016 Census - Boundary files [Datset]. Government of Canada. https://www12.statcan.gc.ca/census-recensement/alternative_alternatif.cfm?l=eng&dispext=zip&teng=lpr_000b16a_e.zip&k=%20%20%20%2027960&loc=http://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/files-fichiers/2016/lpr_000b16a_e.zip

Stoffyn-Egli, P., & Duinker, P. (2013). An Ecological Approach to Riparian-Buffer Definition, and Implications for Timber Harvests in Nova Scotia, Canada. *Journal of Sustainable Development*, 6, 111. https://doi.org/10.5539/jsd.v6n12p111

Strahler, A. (1957). Quantitative analysis of watershed geomorphology. *Transactions, American Geophysical Union*, *38*(6), 913. https://doi.org/10.1029/TR038i006p00913

Suzuki, N., & McComb, W. (1998). Habitat classification models for beaver (Castor canadensis) in the streams of the central Oregon Coast Range. *Northwest Science*, *72*(2), 102–110. https://hdl.handle.net/2376/1217

Thompson, S., Vehkaoja, M., Pellikka, J., & Nummi, P. (2021). Ecosystem services provided by beavers *Castor* spp. *Mammal Review*, *51*(1), 25–39. https://doi.org/10.1111/mam.12220

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234-240. https://web.archive.org/web/20190308014451/http://pdfs.semanticscholar.org/eaa5/eefed d4fa34b7de7448c0c8e0822e9fdf956.pdf

Touihri, M., Labbé, J., Imbeau, L., & Darveau, M. (2018). North American Beaver (*Castor canadensis* Kuhl) key habitat characteristics: review of the relative effects of geomorphology, food availability and anthropogenic infrastructure. *Écoscience*, *25*(1), 9–23. https://doi.org/10.1080/11956860.2017.1395314

UINR (2007, April 16). *Kluscap's science*. Unama'ki Institute of Natural Resources. https://www.uinr.ca/kluscaps-science/

UINR (2020). *About*. Unama'ki Institute of Natural Resources. https://www.uinr.ca/about/

UNEP (2021). First draft of the post-2020 global biodiversity framework. https://www.cbd.int/doc/c/abb5/591f/2e46096d3f0330b08ce87a45/wg2020-03-03-en.pdf

UINR (2022). *Unama'ki Institute of Natural Resources Request for Proposal*. UINR. https://www.uinr.ca/wp-content/uploads/2021/12/UINR-RFP-Connectivity-06-Dec-2021-280-1.pdf

Venter, O., Sanderson, E., Magrach, A., Allan, J., Beher, J., Jones, K., Possingham, H., Laurance, W., Wood, P., Fekete, B., Levy, M., & Watson, J. (2016). Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications, 7*(1), 11 pp. https://doi.org/10.1038/ncomms12558

Wang, G., McClintic, L., & Taylor, J. (2019). Habitat selection by American beaver at multiple spatial scales. *Animal Biotelemetry*, *7*(1), 10 pp. https://doi.org/10.1186/s40317-019-0172-8

Webb, K., Marshall, I (1999). *Ecoregions and ecodistricts of Nova Scotia*. Agriculture and Agri-Food Canada & Environment Canada. 47 pp. https://sis.agr.gc.ca/cansis/publications/surveys/ns/nsee/nsee_report.pdf

Werkowska, W., Márquez, A., Real, R., & Acevedo, P. (2017). A practical overview of transferability in species distribution modeling. *Environmental Reviews*, *25*(1), 127–133. https://doi.org/10.1139/er-2016-0045

West, A., Kumar, S., Brown, C., Stohlgren, T., & Bromberg, J. (2016). Field validation of an invasive species Maxent model. *Ecological Informatics*, *36*, 126–134. https://doi.org/10.1016/j.ecoinf.2016.11.001

Westwood, A. (2016). *Conservation of three forest landbird species at risk: characterizing and modelling habitat at multiple scales to guide management planning.* [Doctoral thesis, Dalhousie University]. International Rusty Blackbird Working Group. https://rustyblackbird.org/wp-content/uploads/Westwood-Alana-PhD-BIOL-March-2016.pdf

Westwood, R., Westwood, A. , Hooshmandi, M., Pearson, K., LaFrance, K., & Murray, C. (2020). A field-validated species distribution model to support management of the critically endangered Poweshiek skipperling (*Oarisma poweshiek*) butterfly in Canada. *Conservation Science and Practice, 2*(3). https://doi.org/10.1111/csp2.163

Wright, J., Jones, C., & Flecker, A. (2002). An ecosystem engineer, the beaver, increases species richness at the landscape scale. *Oecologia*, *132*(1), 96–101. https://doi.org/10.1007/s00442-002-0929-1

Yackulic, C., Chandler, R., Zipkin, E., Royle, J., Nichols, J., Campbell Grant, E., & Veran, S. (2013). Presence-only modelling using MAXENT: when can we trust the inferences? *Methods in Ecology and Evolution, 4*(3), 236–243. https://doi.org/10.1111/2041-210x.12004

APPENDICES

Appendix A – Covariate Extraction Methods

Table A.1 Original candidate environmental covariates generated to model species distribution for the American beaver (*Castor canadensis*) in Nova Scotia, Canada.

Data Source	Covariate	Rationale	Attribute of Interest	Extraction Method	Layer Type
NCC Stream Classification Layer v 2.0	Tidal	Tidal influence as proxy for saltwater toxicity.	Tidal	Euclidean Allocation	Categorical (Yes/No)
	Temperature	Watercourse temperature as indicator of stream depth.	Temp	Euclidean Allocation	Categorical (3 classes)
	Alkalinity	Alkalinity can impact site choice.	Alk	Euclidean Allocation	Categorical (3 classes)
	Size (Complex)	Larger streams support increased vegetation.	Size_Comp	Euclidean Allocation	Categorical (6 classes)
	Gradient (Simple)	Beavers tend to build dams within a specific range of gradients.	Grad_Simp	Euclidean Allocation	Categorical (4 classes)
	Gradient (Complex)	Beavers tend to build dams in a specific range of gradients.	Grad_Comp	Euclidean Allocation	Categorical (7 classes)
NS Hydrographic Network	Watercourse	Distance to nearest watercourse habitat.	N/A	Euclidean Distance	Continuous
NS Forest Inventory	Aspen	Distance to food/materials including Large Tooth and Trembling	TA	Euclidean Distance	Continuous
	Balsam Poplar	Distance to food/materials	BP	Euclidean Distance	Continuous
	Yellow Birch	Distance to food/materials	WB	Euclidean Distance	Continuous
	Gray Birch	Distance to food/materials	GB	Euclidean Distance	Continuous
	White Birch	Distance to food/materials	YB	Euclidean Distance	Continuous
	Black Cherry	Distance to food/materials	BC	Euclidean Distance	Continuous
	Ash	Distance to food/materials, including Black and White Ash.	AS	Euclidean Distance	Continuous
	Sugar Maple	Distance to food/materials	SM	Euclidean Distance	Continuous
	Red Maple	Distance to food/materials	RM	Euclidean Distance	Continuous
	Red Oak	Distance to food/materials	RO	Euclidean Distance	Continuous
	Willow	Distance to food/materials	WI	Euclidean Distance	Continuous

Data Source	Covariate	Rationale	Attribute of Interest	Extraction Method	Layer Type
	Alder	Distance to food/materials	FORNON 38, FORNON 39	Euclidean Distance	Continuous
	Brush	Distance to food/materials	FORNON_33	Euclidean Distance	Continuous
	Softwood	Distance to softwood stand.	FORNON 2	Euclidean Distance	Continuous
	Mixedwood	Distance to mixedwood stand.	FORNON 5	Euclidean Distance	Continuous
	Hardwood	Distance to hardwood stand.	FORNON 8	Euclidean Distance	Continuous
NS Enhanced Digital Elevation Model	Elevation	Influence on suitability of dam site suitability and vegetation.	Elevation	Resample (Cubic)	Continuous
The Canadian Human Footprint	Human Footprint	Influence of cumulative threats from human footprint.	Cumulative Threat Layer	Resample (Bilinear)	Continuous

The method of extraction for each covariate constructed for the Maxent model is listed below organized by source dataset, which can be referenced in Table A.1.

Nova Scotia Forest Inventory:

Distance to Alder: NS Forest Inventory polygons were selected with 'FORNON' codes of 38 and 39, representing "Alders less than 75% cover" and "75% or greater cover" respectively. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the nearest stand of alder.

Distance to Aspen: NS Forest Inventory polygons were selected with 'SPECIES' codes of 'TA06, TA07, TA08, TA09, and TA10'. This denotes stands where Large Tooth and Trembling Aspen are present at 60% or higher. After the selection was exported as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the nearest stand of aspen.

Distance to Balsam Poplar: NS Forest Inventory polygons were selected with 'SPECIES' codes of 'BP06, BP07, BP08, BP09, and BP10' with BP denoting "Balsam Poplar" and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the ‘Euclidean Distance’ tool was used to create a 300m^2 raster showing each cell’s distance to the nearest stand of Balsam Poplar.

Distance to Yellow Birch: NS Forest Inventory polygons were selected with ‘SPECIES’ codes of ‘YB06, YB07, YB08, YB09, and YB10’ with YB denoting “Yellow Birch” and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the ‘Euclidean Distance’ tool was used to create a 300m^2 raster showing each cell’s distance to the nearest stand of Yellow Birch.

Distance to Gray Birch: NS Forest Inventory polygons were selected with ‘SPECIES’ codes of ‘GB06, GB07, GB08, GB09, and GB10’ with GB denoting “Gray Birch” and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the ‘Euclidean Distance’ tool was used to create a 300m^2 raster showing each cell’s distance to the nearest stand of Gray Birch.

Distance to White Birch: NS Forest Inventory polygons were selected ‘SPECIES’ codes of ‘WB06, WB07, WB08, WB09, and WB10’ with WB denoting “White Birch” and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the ‘Euclidean Distance’ tool was used to create a 300m^2 raster showing each cell’s distance to the nearest stand of White Birch.

Distance to Black Cherry: NS Forest Inventory polygons were selected with ‘SPECIES’ codes of ‘BC06, BC07, BC08, BC09, and BC10’ with BC denoting “Black Cherry” and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the ‘Euclidean Distance’ tool was used to create a 300m^2 raster showing each cell’s distance to the nearest stand of Black Cherry.

Distance to White and Black Ash: NS Forest Inventory polygons were selected with ‘SPECIES’ codes of ‘AS06, AS07, AS08, AS09, and AS10’ with AS denoting “White Ash” and “Black Ash”, and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the ‘Euclidean Distance’ tool was used to create a 300m^2 raster showing each cell’s distance to the nearest stand of White and Black Ash.

Distance to Sugar Maple: NS Forest Inventory polygons were selected with ‘SPECIES’ codes of ‘SM06, SM07, SM08, SM09, and SM10’ with SM denoting “Sugar Maple”, and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the

'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the nearest stand of Sugar Maple.

Distance to Red Maple: NS Forest Inventory polygons were selected with 'SPECIES' codes of 'RM06, RM07, RM08, RM09, and RM10' with RM denoting "Red Maple", and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the nearest stand of Red Maple.

Distance to Red Oak: NS Forest Inventory polygons were selected 'SPECIES' codes of 'RO06, RO07, RO08, RO09, and RO10' with RO denoting "Red Oak", and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the nearest stand of Red Oak.

Distance to Willow: NS Forest Inventory polygons were selected with 'SPECIES' codes of 'WI06, WI07, WI08, WI09, and WI10' with WI denoting "Willow", and 06 denoting 60% composition. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the nearest stand of Willow.

Distance to Brush: NS Forest Inventory polygons were selected with 'FORNON' code 83, meaning it has less than 25% merchantable tree cover and greater than 25% woody plant cover. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the nearest brush stands.

Distance to Softwood: NS Forest Inventory polygons were selected with 'COVER_TYPE' code of 2, meaning it is 75% softwood species by basal area. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the softwood stands.

Distance to Mixedwood: NS Forest Inventory polygons were selected with 'COVER_TYPE' code of 5, meaning it is 26% - 74% softwood species by basal area. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a $300m^2$ raster showing each cell's distance to the mixedwood stands.

Distance to Hardwood: NS Forest Inventory polygons were selected with 'COVER_TYPE' code of 8, meaning it is less than 25% softwood species by basal area. After exporting the selection as a feature class, the stands were buffered 10 m and the 'Euclidean Distance' tool was used to create a 300m^2 raster showing each cell's distance to the hardwood stands.

Nova Scotia Stream Classification Network:

Tidal Influence: Using the 'Buffer' tool, a 200m buffer around the line features was created. Using the 'Reclassify' tool, the nominal categorical text data stored in the 'Tidal' field was reclassified into numeric categorical data using the 'Calculate Field' tool in Field View. The Python "reclass()" script was used to reclassify the data as: "Lake : No Class" = 1, "No" = 2, "Yes" = 3. The 'Euclidean Allocation' tool was used to create a 300m^2 raster showing each cell's closest Tidal category.

Temperature: Using the 'Buffer' tool, a 200m buffer around the NSSCN line features was created. Using the 'Reclassify' tool, the nominal categorical text data stored in the 'Temp' field was reclassified into numeric categorical data using the 'Calculate Field' tool in Field View. The Python "reclass()" script was used to reclassify the data as: "Lake : No Class" = 1, "Cold" = 2, "Cool" = 3, "Warm" = 4. The 'Euclidean Allocation' tool was used to create a 300m^2 raster showing each cell's closest Temp category.

Alkalinity: Using the 'Buffer' tool, a 200m buffer around the NSSCN line features was created. Using the 'Reclassify' tool, the nominal categorical text data stored in the 'Alk' field was reclassified into numeric categorical data using the 'Calculate Field' tool in Field View. The Python "reclass()" script was used to reclassify the data as: "Lake : No Class" = 1, "Low Alkalinity" = 2, "Moderate Alkalinity" = 3, "High Alkalinity" = 4. The 'Euclidean Allocation' tool was used to create a 300m^2 raster showing each cell's closest alkalinity category.

Size (Complex): Using the 'Buffer' tool, a 200m buffer around the NSSCN line features was created. Using the 'Reclassify' tool, the nominal categorical text data stored in the 'Size_Comp' field was reclassified into numeric categorical data using the 'Calculate Field' tool in Field View. The Python "reclass()" script was used to reclassify the data as: "Lake : No Class" = 1, "Headwaters and Creeks" = 2, "Small Rivers" = 3, "Medium Tributary Rivers" = 4, "Medium Main-stem Rivers = 5", "Large Rivers" = 6. The 'Euclidean Allocation' tool was used to create a 300m^2 raster showing each cell's closest watercourse size category.

Gradient (Complex): Using the ‘Buffer’ tool, a 200m buffer around the NSSCN line features was created. Using the ‘Reclassify’ tool, the nominal categorical text data stored in the ‘Grad_Comp’ field was reclassified into numeric categorical data using the ‘Calculate Field’ tool in Field View. The Python “reclass()” script was used to reclassify the data as: “Lake : No Class” = 1, “Very Low Gradient” = 2, “Low Gradient” = 3, “Moderate Low Gradient” = 4, “Moderate-High Gradient” = 5, “High Gradient” = 6, “Very High Gradient” = 7. The ‘Euclidean Allocation’ tool was used to create a 300m^2 raster showing each cell’s closest watercourse complex gradient category.

Gradient (Simple): Using the ‘Buffer’ tool, a 200m buffer around the NSSCN line features was created. Using the ‘Reclassify’ tool, the nominal categorical text data stored in the ‘Grad_Simp’ field was reclassified into numeric categorical data using the ‘Calculate Field’ tool in Field View. The Python “reclass()” script was used to reclassify the data as: “Lake : No Class” = 1, “Low Gradient” = 2, “Moderate Gradient” = 3, “High Gradient” = 4. The ‘Euclidean Allocation’ tool was used to create a 300m^2 raster showing each cell’s closest watercourse simple gradient category.

Nova Scotia Hydrographic Network:

Distance to Watercourse: The ‘Select by Attributes’ tool was used to select features representing freshwater lakes and rivers. The watercourses were buffered 10 m, and the ‘Euclidean Distance’ tool was used to create a 300m^2 raster showing each cell’s distance to a river or lake.

Nova Scotia Digital Elevation Model:

Elevation: The “Resample” tool was used to convert the layer from 20m^2 to 300m^2 cell size using ‘Cubic’ resampling technique.

Canadian Human Footprint:

Human Footprint: The “Resample” tool was used to convert the layer from 20m^2 to 300m^2 cell size using ‘Bilinear’ resampling technique.

Appendix B – Covariate Correlation Matrix

Table B.1 Correlation matrix depicting raster layer correlations between the environmental covariate group data layer, where covariate is shown in black (corresponding name in below), correlation values > 0.7 are shown in red, and values between 0.7 > 0.4 are shown in yellow.

Covariate	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	1.00	-0.07	0.23	-0.06	-0.36	-0.20	-0.17	-0.15	-0.23	-0.09	-0.04	0.02	-0.26	-0.14	-0.14	-0.06	0.06	0.06	0.08	0.10	0.19	0.23	-0.34	-0.41	-0.04
2	-0.07	1.00	-0.09	0.62	0.22	0.34	0.18	0.14	0.25	0.31	0.11	0.39	0.38	0.29	0.14	0.29	-0.11	-0.10	-0.08	0.03	0.00	-0.06	0.06	-0.01	-0.07
3	0.23	-0.09	1.00	0.03	-0.08	-0.06	-0.02	-0.02	-0.06	0.05	0.00	-0.02	-0.14	-0.08	-0.03	0.00	0.00	0.00	0.00	-0.03	0.02	0.11	-0.09	-0.08	0.02
4	-0.06	0.62	0.03	1.00	0.09	0.23	0.06	0.01	0.20	0.37	0.17	0.23	0.33	0.19	0.00	0.13	-0.06	-0.06	-0.07	-0.03	0.01	0.00	0.02	0.10	-0.04
5	-0.36	0.22	-0.08	0.09	1.00	0.24	0.34	0.38	0.21	-0.06	-0.22	0.20	0.27	0.36	0.39	0.38	-0.27	-0.26	-0.09	0.02	-0.22	-0.26	0.41	0.01	-0.10
6	-0.20	0.34	-0.06	0.23	0.24	1.00	0.38	0.25	0.67	0.46	0.46	0.40	0.34	0.18	0.22	0.37	0.04	0.02	-0.02	-0.04	0.02	-0.04	0.09	0.23	0.05
7	-0.17	0.18	-0.02	0.06	0.34	0.38	1.00	0.91	0.72	0.11	-0.19	0.20	-0.01	0.17	0.89	0.51	-0.16	-0.15	-0.02	0.18	-0.06	-0.14	0.15	-0.01	-0.07
8	-0.15	0.14	-0.02	0.01	0.38	0.25	0.91	1.00	0.50	-0.04	-0.29	0.19	-0.03	0.24	1.00	0.57	-0.24	-0.22	-0.01	0.24	-0.09	-0.18	0.16	-0.10	-0.11
9	-0.23	0.25	-0.06	0.20	0.21	0.67	0.72	0.50	1.00	0.36	0.25	0.24	0.30	0.11	0.47	0.32	-0.03	-0.03	-0.02	0.00	0.00	-0.08	0.06	0.13	-0.02
10	-0.09	0.31	0.05	0.37	-0.06	0.46	0.11	-0.04	0.36	1.00	0.46	-0.02	0.18	-0.06	-0.08	-0.02	0.21	0.19	-0.01	-0.11	0.07	0.11	0.03	0.43	0.16
11	-0.04	0.11	0.00	0.17	-0.22	0.46	-0.19	-0.29	0.25	0.46	1.00	-0.02	0.14	-0.24	-0.33	-0.15	0.23	0.21	0.01	-0.21	0.11	0.18	-0.14	0.27	0.12
12	0.02	0.39	-0.02	0.23	0.20	0.40	0.20	0.19	0.24	-0.02	-0.02	1.00	0.17	0.31	0.20	0.56	-0.12	-0.11	-0.01	0.16	0.09	0.02	0.03	-0.33	-0.08
13	-0.26	0.38	-0.14	0.33	0.27	0.34	-0.01	-0.03	0.30	0.18	0.14	0.17	1.00	0.23	-0.04	0.01	-0.10	-0.11	-0.10	-0.10	-0.08	-0.19	0.18	0.09	-0.06
14	-0.14	0.29	-0.08	0.19	0.36	0.18	0.17	0.24	0.11	-0.06	-0.24	0.31	0.23	1.00	0.26	0.44	-0.20	-0.18	-0.02	0.11	-0.05	-0.13	0.19	-0.14	-0.09
15	-0.14	0.14	-0.03	0.00	0.39	0.22	0.89	1.00	0.47	-0.08	-0.33	0.20	-0.04	0.26	1.00	0.58	-0.27	-0.24	-0.01	0.25	-0.10	-0.19	0.17	-0.14	-0.12
16	-0.06	0.29	0.00	0.13	0.38	0.37	0.51	0.57	0.32	-0.02	-0.15	0.56	0.01	0.44	0.58	1.00	-0.18	-0.16	0.00	0.20	0.01	0.00	0.11	-0.21	-0.08
17	0.06	-0.11	0.00	-0.06	-0.27	0.04	-0.16	-0.24	-0.03	0.21	0.23	-0.12	-0.10	-0.20	-0.27	-0.18	1.00	0.97	0.33	0.40	0.65	0.56	-0.11	0.17	0.39
18	0.06	-0.10	0.00	-0.06	-0.26	0.02	-0.15	-0.22	-0.03	0.19	0.21	-0.11	-0.11	-0.18	-0.24	-0.16	0.97	1.00	0.38	0.46	0.68	0.59	-0.11	0.13	0.37
19	0.08	-0.08	0.00	-0.07	-0.09	-0.02	-0.02	-0.01	-0.02	-0.01	0.01	-0.01	-0.10	-0.02	-0.01	0.00	0.33	0.38	1.00	0.62	0.57	0.51	-0.08	-0.14	0.02
20	0.10	0.03	-0.03	-0.03	0.02	-0.04	0.18	0.24	0.00	-0.11	-0.21	0.16	-0.10	0.11	0.25	0.20	0.40	0.46	0.62	1.00	0.68	0.54	-0.04	-0.27	0.02
21	0.19	0.00	0.02	0.01	-0.22	0.02	-0.06	-0.09	0.00	0.07	0.11	0.09	-0.08	-0.05	-0.10	0.01	0.65	0.68	0.57	0.68	1.00	0.69	-0.15	-0.18	0.20
22	0.23	-0.06	0.11	0.00	-0.26	-0.04	-0.14	-0.18	-0.08	0.11	0.18	0.02	-0.19	-0.13	-0.19	0.00	0.56	0.59	0.51	0.54	0.69	1.00	-0.18	-0.19	0.15
23	-0.34	0.06	-0.09	0.02	0.41	0.09	0.15	0.16	0.06	0.03	-0.14	0.03	0.18	0.19	0.17	0.11	-0.11	-0.11	-0.08	-0.04	-0.15	-0.18	1.00	0.17	0.01
24	-0.41	-0.01	-0.08	0.10	0.01	0.23	-0.01	-0.10	0.13	0.43	0.27	-0.33	0.09	-0.14	-0.14	-0.21	0.17	0.13	-0.14	-0.27	-0.18	-0.19	0.17	1.00	0.17
25	-0.04	-0.07	0.02	-0.04	-0.10	0.05	-0.07	-0.11	-0.02	0.16	0.12	-0.08	-0.06	-0.09	-0.12	-0.08	0.39	0.37	0.02	0.02	0.20	0.15	0.01	0.17	1.00

Covariate number associated with Table B.1:

1	Cumulative Treat	6	Ash	11	Red Oak	16	Yellow Birch	21	Tidal
2	Hardwood	7	Black Cherry	12	Sugar Maple	17	Grad Comp	22	Alkalinity
3	Softwood	8	Balsam Poplar	13	Aspen	18	Grad Simp	23	Brush
4	Mixedwood	9	Gray Birch	14	White Birch	19	Size	24	Elevation
5	Alder	10	Red Maple	15	Willow	20	Temp	25	Watercourse

Appendix C – Ecological Model and Human Footprint Model Results

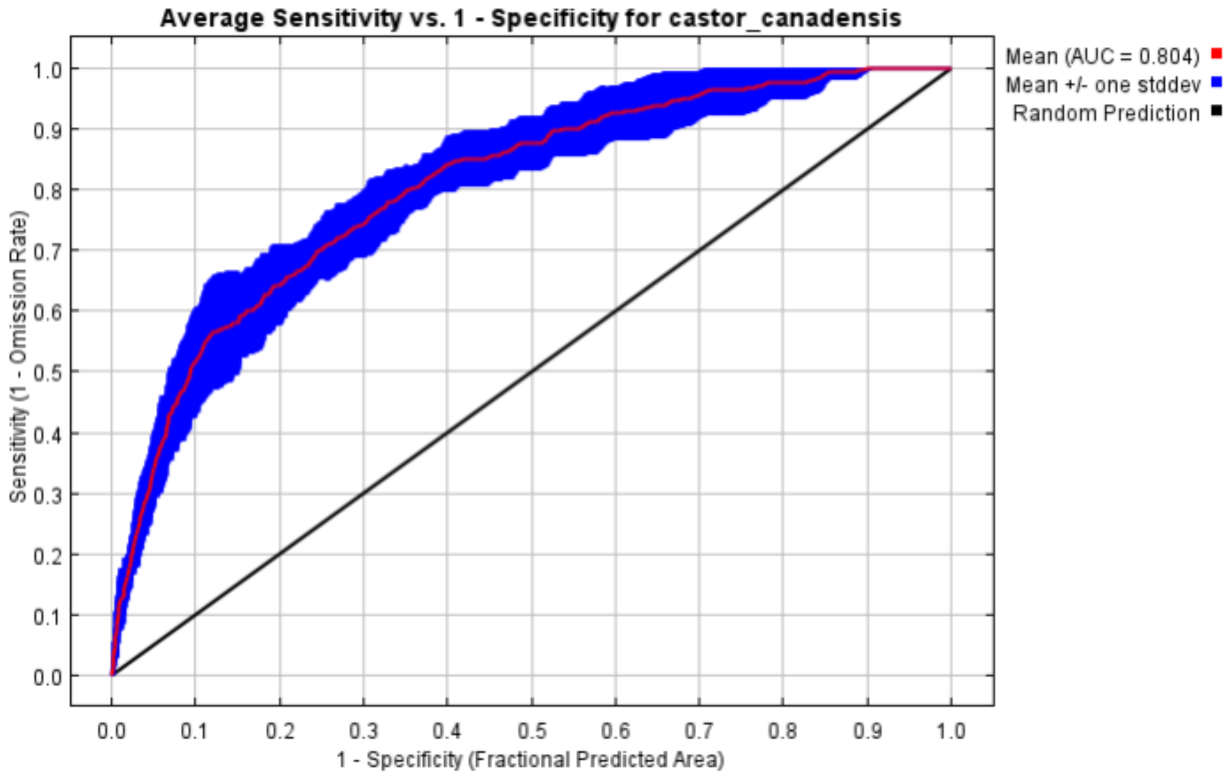

Figure C.1 Area under the receiver operating characteristic curve (AUC) for the ecological Maxent model of predicted probability of occurrence for American beaver (*Castor canadensis*) in Nova Scotia, Canada. Test AUC derived from an arithmetic average of 10 replicate cross validated models.

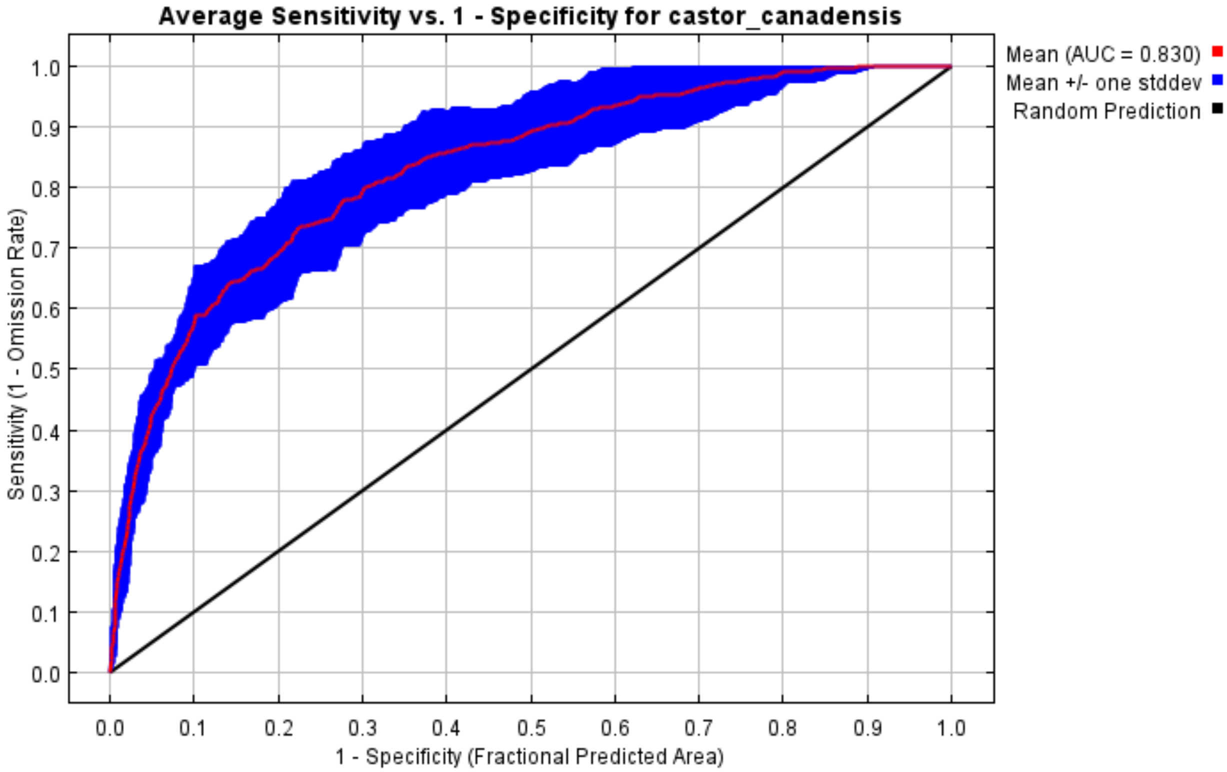

Figure C.2 Area under the receiver operating characteristic curve (AUC) for the human footprint Maxent model of predicted probability of occurrence for American beaver (*Castor canadensis*) in Nova Scotia, Canada. Test AUC derived from an arithmetic average of 10 replicate cross validated models.

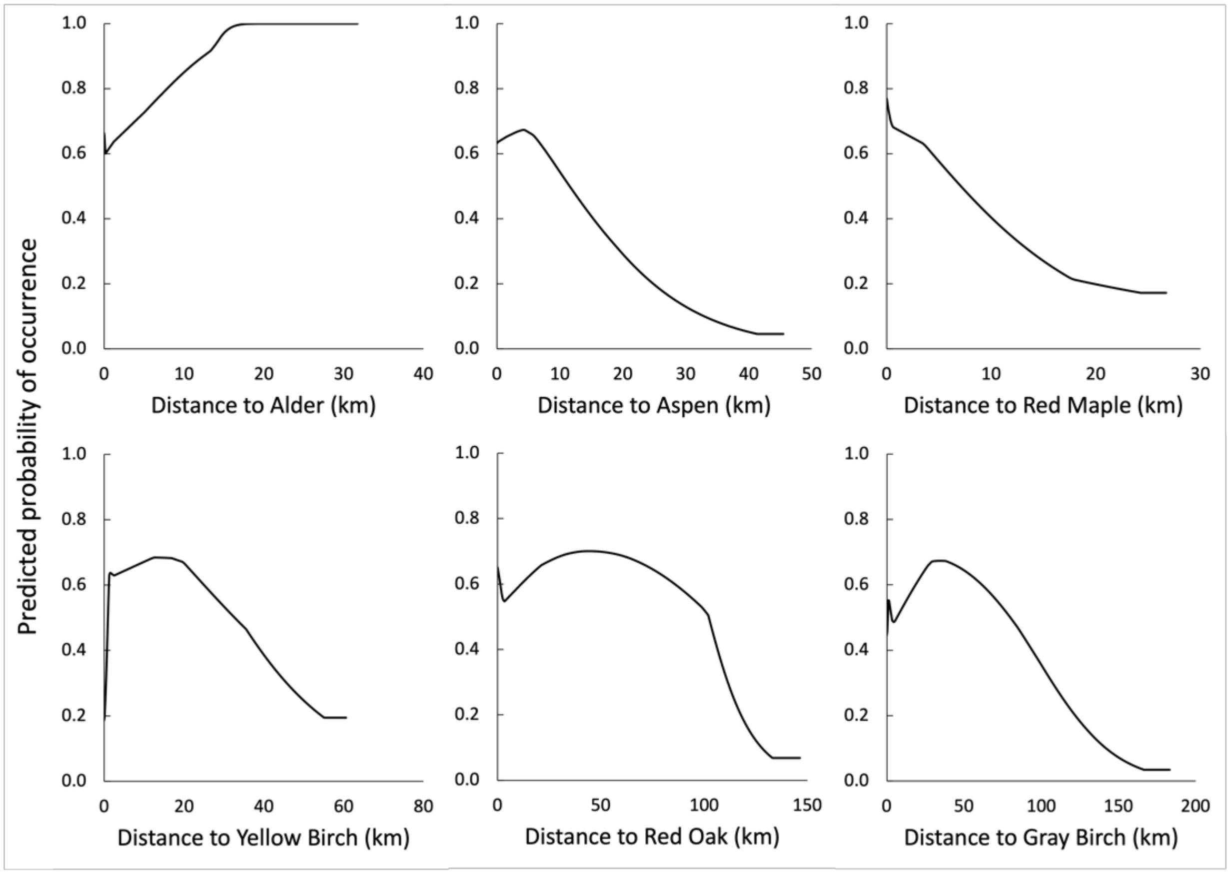

Figure C.3 Human footprint model response curves for the tree stand-specific species-covariate relationships for the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada, created using the cumulative threat layer in addition to the environmental covariate group data layer. Response curves derived from an arithmetic average of 10 replicate cross validated Maxent models. Each panel represents the variation in predicted probability of occurrence as distance increases from tree stands where the listed species is dominant with 60% composition or higher.

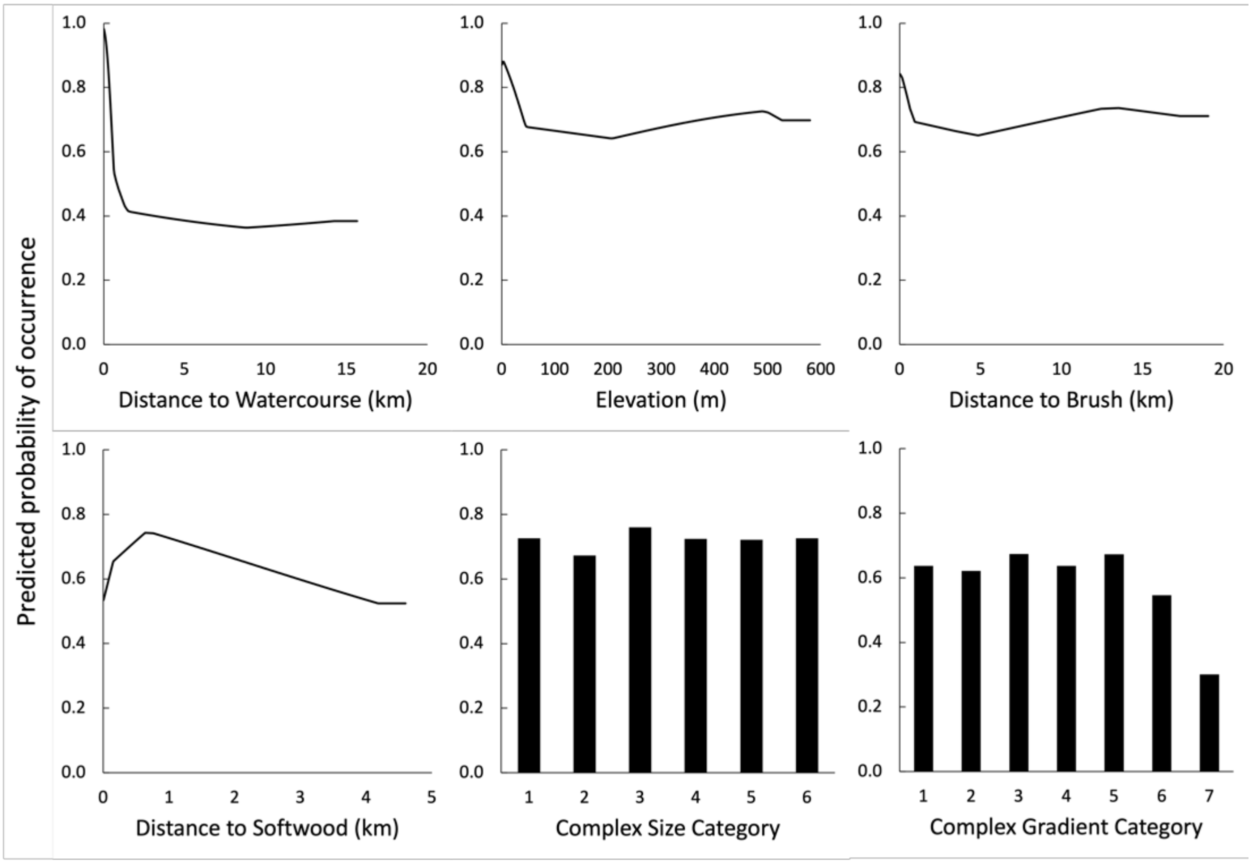

Figure C.4 Human footprint model response curves for the watercourse, elevation, and forest type specific species-covariate relationships for the predicted probability of occurrence of American beaver (*Castor canadensis*) in Nova Scotia, Canada created using the cumulative threat layer in addition to the environmental covariate group data layer. Response curves derived from an arithmetic average of 10 replicate cross validated Maxent species distribution models. Each panel represents the variation in predicted probability of occurrence as covariate values vary, where Size (Complex) Categories are Lake = 1, Headwaters and Creeks = 2, Small Rivers = 3, Medium Tributary Rivers = 4, Medium Main-stem Rivers = 5, Large Rivers = 6.; and Gradient (Complex) Categories are Lake = 1, Very Low Gradient = 2, Low Gradient = 3, Moderate-Low Gradient = 4, Moderate-High Gradient = 5, High Gradient = 6, Very High Gradient = 7.