Modelling current and projected niche shifts of the blacklegged tick, *Ixodes scapularis,* in eastern Canada employing community science data and global climate change drivers

Jacob Westcott

A Thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

Master of Science in Applied Geomatics

School of Science and the Environment, Grenfell Campus, Memorial University of Newfoundland

May, 2024

Corner Brook, Newfoundland and Labrador

Abstract

Climate change rapidly drives species range dynamics, prompting many terrestrial organisms to shift northward to higher latitudes and forcing new species-environment and species-species interactions. The tick vector *Ixodes scapularis,* commonly known as the blacklegged tick, has historically been endemic to the United States but is establishing a persistent population in Canada, potentially exposing people to a novel zoonotic pathogen, *Borrelia burgdorferi*, the causative agent of Lyme disease. The collection of tick records (within Canada and the United States) between 2017 and 2022 via citizen-/community-science programs and the usage of high-resolution 1km climate data enabled me to produce robust, ensemble ecological niche models. I carried out 4,704 model iterations across two datasets, 12 algorithms, and 10 climate profiles using 40 environmental variables. I extrapolated select models over three time periods, 2011-2040, 2041-2070, and 2071-2100, across two projected climate scenarios, SSP5-8.5 and SSP3-7.0, incorporating 2,094 future predictions of *I. scapularis* distribution. My ensembles (AUC: 0.9565 ± 0.0065; TSS: 0.8435 ± 0.0155; Kappa: 0.819 ± 0.014) identified temperature, precipitation, biomass production (NPP), length of the growing season, climate moisture index, and the number of yearly degree days as the variables that best explained the distribution of *I. scapularis*. Further changes to these climate conditions will result in continued *I. scapularis* range expansion, with estimates ranging from ~205% (409,475 km^2 to 1,247,689 km^2) up to ~248% (447,532 km^2 to 1,556,760 km^2) before the end of the century. These distributional niche changes coincide with a northern latitude limit reaching as far as ~48°N by 2040, ~50°N by 2070, and ~52°N by 2100. These findings highlight the invasive potential of *I. scapularis*, with implications for public health and changing ecosystem dynamics.

General Summary

Climate change is causing various species to migrate northward. Among them is the blacklegged tick (*Ixodes scapularis*), notorious for transmitting Lyme disease. The continuous intrusion of this tick into Canada poses a significant threat. Using environmental data and documented blacklegged tick sightings; I have developed predictive models to identify the most suitable habitats for the blacklegged tick in Canada now and in the foreseeable future. My research highlights that the temperature, precipitation, plant density, moisture levels, and prolonged warmth of a location are key factors for determining where the blacklegged tick can live. As climate change alters these factors, the suitable habitat for the blacklegged tick in Canada is projected to expand more than threefold by 2100. Additionally, their range is expected to extend an additional 4°N.

Acknowledgements

First and foremost, I would like to extend my heartfelt thanks to my supervisors, Dr. Joe Bowden and Dr. Karen Doody. Your guidance, expertise, mentorship, and support have helped shape this research and my academic (and professional) growth.

I would also like to acknowledge the combined financial support provided by Memorial University of Newfoundland and Labrador and Natural Resources Canada, which enabled me to conduct this research and make meaningful contributions to the fields of biogeographical modelling and zoonotic epidemiology.

Furthermore, I am grateful to Jade Savage and the eTick network, as well as iNaturalist and those who have contributed to that community science database for providing and collecting invaluable *I. scapularis* records. The working data collected, thanks to both experts and public participants, have been instrumental in the analysis and findings of my research. Collaboration with such organisations (and the public) is essential in advancing our understanding of species range dynamics.

I would like to thank Natural Resources Canada and Doug Piercey for providing me with a dedicated workstation and being readily available to consult if I had any questions or concerns.

Lastly, I am profoundly grateful for my parents' unwavering support through my academic journey and path to self-discovery. Their love, guidance, assistance, and encouragement have empowered me to break through self-imposed limits.

Table of Contents

List of Figures

List of Tables

List Appendices

List of Abbreviations

Abbrev	Full Name
AHM	Annual Heat Moisture Index (MAT + 10) / (MAP / 1000)
AOGCM	Atmosphere-Ocean General Circulation Models
AR6	Sixth Assessment Report
AUC	Area under the Reciever Operating Characteristic (ROC) Curve
Bio1	Mean Annual Temperature (°C)
Bio10	Mean Daily Air Temperature of the Warmest Quarter (°C)
Bio12	Annual Precipitation (kg m^{-2})
Bio13	Precipitation Amount of the Wettest Month (kg m^{-2})
Bio14	Precipitation Amount of the Driest Month (kg m^{-2})
Bio15	Precipitation Seasonal (kg m^{-2})
Bio16	Mean Monthly Precipitation of the Wettest Quarter (kg m^{-2})
Bio17	Mean Monthly Precipitation of the Driest Quarter (kg m^{-2})
Bio18	Precipitation of the Warmest Quarter (kg m^{-2})
Bio2	Mean Diurnal Air Temperature Range (°C)
Bio3	Isothermality (°C)
Bio4	Temperature Seasonality (°C/100)
Bio5	Maximum Daily Mean Air Temperature of the Warmest Month (°C)
Bio7	Annual Range of Air Temperature (°C)
Bio8	Mean Daily Air Temperature of the Wettest Quarter (°C)

BRT	Boosted Regression Trees
CART	Classification and Regression Trees
CMD	Hargreaves Climate Moisture Index (mm)
CMI	Hogg's Climate Moisture Index (mm)
CMIP6	Coupled Model Intercomparison Project Phase 6
cp	Complexity Parameters
DD	Degree Days
DD_0	Degree Days Below 0°C
DD1040	Degree Days Between 10°C and 40°C
DD18	Degree Days Above 18°C
DD5	Degree Days Above 5°C
ELEV	Elevation (m)
ENFA	Ecological Niche Factor Analysis
ENM	Ecological Niche Models
EXT	Mean Extreme Maximum Temperature Over 30-Years (°C)
FCF	Frost-Cover-Frequency
FDA	Flexible Discrimination Analysis
GAM	Generalised Additive Model
GARP	Genetic Algorithm for Rule Set Production
GCS	Geographic Coordinate System

GDDLGD0	Last (Growing Degree) Day above 0°C (Julian days)
GDDLGD10	Last (Growing Degree) Day above 10°C (Julian days)
GDDLGD5	Last (Growing Degree) Day above 5°C (Julian days)
GFLD-ESM4	Geophysical Fluid Dynamic Laboratory: Earth System Model 4
GLM	Generalised Linear Model
GLMPoly	Generalised Linear Model with Polynomial Regression
GSL	Growing Season Length
GSP	Growing Season Precipitation (kg m^{-2})
GST	Growing Season Temperatures (°C)
IPCC	Intergovernmental Panel on Climate Change
Kappa	Cohen's kappa
MAHAL	Mahalanobis Model
MAP	Mean Annual Precipitation (mm)
MAT	Mean Annual Temperature (°C)
Maxent	Maximum Entropy Modelling
MCMT	Mean Temperature of the Coldest Month (°C)
MDA	Multiple Discrimination Analysis
MLP	MultiLayer Perceptron
MSP	Mean Summer Precipitation (mm)
MWMT	Mean Temperature of the Warmest Month (°C)

NL	Newfoundland and Labrador
NPP	Net Primary Productivity (g C m^{-2} yr^{-1})
PEI	Prince Edward Island
PAS	Precipitation as Snow (mm)
PPT_at	Autumn Precipitation (mm)
PPT_sm	Summer Precip (mm)
PPT_sp	Spring Precipitation (mm)
PPT_wt	Winter Precipitation (mm)
RBF	Radial Basis Function
RF	Random Forests
RH	Mean Annual Relative Humidity (%)
RPART	Recursive Partitioning and Regression Trees
SDM	Species Distribution Models
SHM	Summer Heat Moisture (MWMT / (MSP / 1000)
SSP	Shared Socioeconomic Pathways
SVM	Support Vector Machine
SWE	Snow-Water Equivalent (kg m^{-2})
TD	Temperature Difference between Warmest and Coldest Month (i.e., Measure of Continentality) (°C)
tHRF	Tick Histamine Release Factor
TSS	True Skill Statistic
VIF	Variance Inflation Factor
WGS 1984	World Geodetic System of 1984

Chapter 1: Literature Review

Introduction

Since the pre-industrial period, there has been a 1.1°C increase in the temperature of the Earth's biosphere. Without intervention, an additional 2°C to 5°C of warming is expected within the 21st century (Harvey et al., 2022). This warming will produce compounding effects, impacting various environmental, ecological, socio-economic, and socio-political components (Abbass et al., 2022). Many ecosystems (e.g., terrestrial, coastal, freshwater, etc.) are experiencing significant and potentially irreversible changes (IPCC, 2023); many local plant and animal species are threatened with extinction, spearheaded by climate changes (Parmesan et al., 2022). Close to half of the species assessed within the IPCC Sixth Assessment Report (AR6) have moved to higher elevations and farther toward the poles (2023). Many species cannot adapt quickly enough in response to the rapid onset of climate change (IPCC, 2023). These rapid changes are driving changes in species' biology with implications for their physiology, which ultimately dictates phenology (timing of life events, e.g., spring emergence), body size, fitness, and changing species distributions – universal species responses to climate change (McCauley & Mabry, 2011). Species range shifts, for example, could indicate that a species is declining, as is often the case for rare species with a specific suite of ecological niche requirements. Similarly, species with broad thermal limits or ecological requirements (e.g., many pest species) are capable of rapidly expanding their range (Lancaster, 2016). Range expansion could affect novel species interactions, with implications for ecology, economics and human health (Alexander et al., 2015; Gilman et al., 2010; Wikel, 2018).

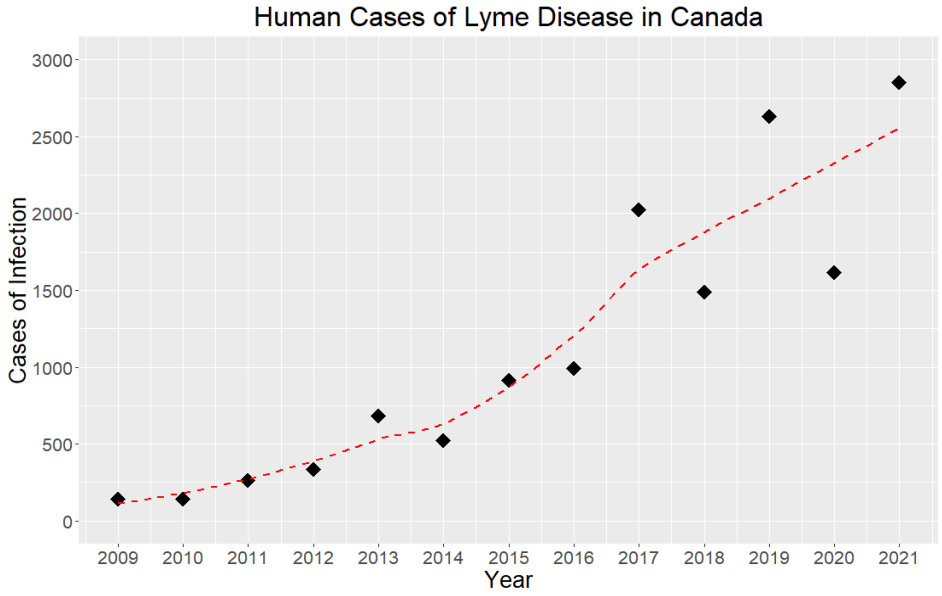
Due to their physiology being mediated by the external environment, terrestrial arthropods (e.g., spiders, insects, mites) respond readily to environmental change (Harrison et al., 2012). As such, terrestrial arthropods exhibit significant rapid changes in response to climate change, including shifts in phenology, body size and range (e.g., Bowden et al., 2015; Levi et al., 2015; Porretta et al., 2013). Humans are especially concerned with arthropods that could significantly impact the ecology of the landscape, economies and human health as a result of changes to species distributions. Arthropods that can transmit zoonotic pathogens, such as ticks, rapidly respond to warmer climates by increasing locally in abundance via more favourable climates (Ogden et al., 2021), expanding their range northward (Porretta et al., 2013), and advancing their phenology (Levi et al., 2015).

Ixodes scapularis, commonly known as the blacklegged tick (or deer tick), is of growing concern in Canada from a macroecological and epidemiological perspective. While capable of spreading seven human pathogens (Eisen & Eisen, 2018), *I. scapularis* is the primary vector for the spirochete bacterium *Borrelia burgdorferi*, the etiological microbe responsible for borreliosis *(*i.e., Lyme disease; Belli et al., 2017). Lyme borreliosis is the most common vector-borne infectious disease in North America (Levi, 2015), responsible for over 300,000 cases annually in the United States (Talbot et al., 2019). Transmitted via tick bites, early symptoms of Lyme disease may consist of a characteristic skin rash (i.e., erythema migrans), fever, chills, musculoskeletal aches, fatigue, headaches, and swollen lymph nodes (Centers for Disease Control and Prevention [CDC], 2022). If left untreated, the infection can spread, exacerbating previous symptoms or introducing new acute ones, namely swelling of the central nervous system, heart or eye complications, cognitive impairment, facial paralysis, nerve damage, arthritis, and, in rare cases, death (CDC, 2022).

In North America*, I. scapularis* has historically been endemic to the eastern United States (Its close relative, *Ixodes pacificus,* or the western blacklegged tick, is primarily distributed within the Pacific U.S.; Eisen et al., 2016a). However, climate change is pushing the geographical limits of *I. scapularis* northward into Canada. Areas once considered unsuitable are becoming viable habitats for establishing sustained populations (Slatculescu et al., 2020). *I. scapularis* sightings have been reported across various Canadian provinces. Southern Ontario, Quebec, and Nova Scotia stand out as the most heavily infested and expanding regions, while populations also inhabit New Brunswick and Manitoba (Chilton et al., 2020; eTick, 2023; iNaturalist, 2023). In addition, *I. scapularis* has been sighted sporadically in insular Newfoundland and Labrador (NL), Prince Edward Island (PEI), and lower Saskatchewan (eTick, 2023; iNaturalist, 2023).

Over the past 70 years, species have moved approximately 17 (±3) km per year towards higher latitudes due to climate change (IPCC, 2021a).. This shift has enabled many tick host-species ranges to extend northward, leading to spatial diffusion from present tick localities (Morrill, 1988). The leading edge of *I. scapularis'* range is expanding nearly three times as fast as the average species, at ~48 (±2) km/year (Clow et al., 2017; Talbot et al., 2019). Terrestrial animals, such as deer and smaller mammals, are capable of transporting ticks shorter distances, while birds are the leading cause of long-range expansion (Halsey et al., 2018; Keirans et al., 1996). The rising population of both *I. scapularis* and infected hosts/carriers in Canada has corresponded with an increased number of human Lyme disease cases in the country, with 144 recorded exposures in 2009 to 2,850 in 2021 (Figure 1, Public Health Agency of Canada [PHAC], 2022).

Figure 1. Number of reported human cases of Lyme disease in Canada from 2009 to 2021, adapted from "Lyme disease: Surveillance" (PHAC, 2022). The dashed red line depicts the upward trend in recorded human cases.



The Biology and Ecology of I. scapularis

I. scapularis have a life cycle of about two years, sometimes prolonged to three. They are hematophagous arthropods (Nuss et al., 2021) with three life stages after eclosion: larva, nymph, and adult. Each transition requires a blood meal before developing into the next phase. Female adult *I. scapularis* ticks are oviparous insects, laying up to ~3000 eggs near the end of their lifespan (Halsey & Miller, 2018). However, this value is highly variable and contingent on ambient temperatures (Gaff et al., 2020). While transovarial transmission (i.e., the transmission

of *B. burgdorferi* bacteria from an adult to an egg) is possible, there is less than a 1% chance of it occurring (Rosenthal & Coburn, 2008).

I. scapularis is a host-generalist identified on over 125 North American vertebrate species (Halsey et al., 2018; Keirans et al., 1996). *I. scapularis* larvae or nymphs primarily feed on small mammals, which several are pathogen-reservoir hosts (e.g., mice, squirrels, chipmunks, and shrews; Dumas et al., 2022) Adult *I. scapularis* primarily feed on larger mammals (Halsey & Miller, 2018) such as the white-tailed deer, *Odocoileus virginianus* (Huang et al., 2019). Feeding success rates range from ~3% to 49%, depending on host species and grooming behaviour (Halsey et al., 2018). The probability of *B. burgdorferi* transmission via a single infected nymphal tick to a vulnerable host species is high (Gaff et al., 2020), particularly among northern *I. scapularis* ticks that have engorged for a minimum of 3 days (Goddard et al., 2015). The transmission of *B. burgdorgeri* can occur within a shorter period, with *Ixodes* ticks capable of transmitting the pathogen in as little as 16 hours and consistently within 24 hours (Cook, 2014). Once a tick becomes infected with the bacterium, transmission is transstadial, where the pathogen persists within each subsequent life stage (Rosenthal & Coburn, 2008). Ticks aim to feed to satiety; failing to do so means they lack the proper nutrients and energy to moult into nymphs or adults (Halsey & Miller, 2018). *I. scapularis* can spend up to ~98% of its life span off host vertebrates (Mathisson et al., 2021). They can survive for months without feeding because of low basal metabolic processes (Mathisson et al., 2021) and within refugia that protect them against harmful climatic conditions (Burtis et al., 2019). They have an acute ability to detect if a host is within 1m² and can discern temperature (Halsey & Miller, 2018). The phenology of *Ixodes* ticks is variable geographically and differs across species (Levi et al., 2015). For *I. scapularis*, peak nymph activity is within late spring and early summer, with larval activity

peaking in the late summer. (Burtis et al., 2023; Levi et al., 2015; Ogden et al., 2018). Adult *I. scapularis* has a bimodal distribution in activity, peaking in spring and fall (Hamer et al., 2012). Under optimal conditions, *I. scapularis* can complete a lifecycle in less than a year. In some extreme cases in southeastern Canada (i.e., one of the coldest regions of *I. scapularis* known range), the life cycle can take 3-4 years to complete (Eisen et al., 2016b).

The close linkages between climate and phenology suggest that climate change could induce earlier nymph and larva activity in *I. scapularis* and, consequently, favour increased transmission of Lyme borreliosis (Levi et al., 2015; Ogden et al., 2008a) *I. scapularis* exhibits asynchronous phenology, unlike *Ixodes pacificus* and European genera of *Ixodes*. If variations among geographic, climatic, and other natural processes generate a longer lag (greater asynchronous) in activity, nymph *I. scapularis* ticks may vector *B. burgdorferi* within host populations, preceding the arrival of larvae ticks, who can then acquire the pathogen from the infected hosts (Levi et al., 2015). This could facilitate increased transmissions of Lyme disease. If the particular strain of *B. burgdorferi* found in the Canadian *I. scapularis* ticks resembles the variant of *I. scapularis* found in the northeastern United States, known for their prolonged period of infectivity in vertebrate hosts (Levi et al., 2015).

The transmission dynamics of *I. scapularis* are facilitated by their host-seeking behaviour. *I. scapularis* employ an ambush strategy (Mathisson et al., 2021) called questing. A questing tick ascends natural objects (e.g., vegetation, rocks), outstretchs its front legs, and waits to clasp onto a host to feed (Leal et al., 2020). Travelling (and questing) comes at a risk to a tick's life as they may expend a potentially detrimental amount of energy at the potential cost of receiving no blood meal (Leal et al., 2020), or they may need to modify questing behaviour to avoid desiccation (Couper et al., 2021). When not questing, ticks must find protection under

suitable biotic and abiotic conditions or the risk of death is significantly increased. *I. scapularis* ticks are positively associated with deciduous forests with closed canopies, while they have an aversion to grasslands, coniferous forests, and open areas (Mathisson et al., 2021). Within refugia, they have an affinity for soils with good water drainage (Burtis et al., 2019) and seek vegetation with dense undergrowth (e.g., *Berberis thunbergii*, *Ilex verticillata*; *Lonicera* and *Osmunda* species; Mathisson et al., 2021) that protects against variable weather and creates humid microclimates to prevent desiccation (Mathisson et al., 2021). Snow and leaf litter that covers the soil surface act as insulators against seasonal changes (e.g., dormant overwintering), particularly against ambient temperature and relative humidity (for which developmental and extinction thresholds exist; Ogden, 2004). Many generalist arthropod groups (e.g., spiders, beetles, ants) prey upon *I. scapularis*, though these interactions are likely haphazard (Burtis & Pflueger., 2017).

While climate change remains a significant factor in the distribution of ticks, globalisation, particularly human travel, can impact their movement. Anthropogenic alterations to land use, such as habitat fragmentation, agricultural expansion, and deforestation (Harvey et al., 2022), can disrupt natural ecosystems, leading to changes in microclimates and resource availability (Léger et al., 2013). Additionally, increased movement of humans, domestic livestock, and pets can lead to rising interactions between ticks and host species in novel territories (Léger et al., 2013). These disruptions can result in shifts in ecological dynamics, influencing the abundance of reservoir host species or tick density (Léger et al., 2013). However, the broader expansion of *I. scapularis* distribution in Canada is less likely due to the local spread of ticks from terrestrial hosts but instead from northward migration from avian species (Ogden et al., 2008b). Each spring in Canada, ~2% of the estimated 3 billion birds that migrate northward

into central and eastern Canada transport and scatter *I. scapularis* ticks; this is between 50 - 175 million *I. scapularis* ticks (Ogden et al., 2008b). The long distances that migratory birds undergo provide enough time (2-4 days) for larvae and nymph ticks to become engorged, acquire *B. burgdorferi*, and subsequently transmit the bacteria at the next blood meal (Ogden et al., 2008c). In bird-borne ticks, infection prevalence is relatively low, up to 10%, consequently delaying the establishment of transmission cycles in new populations (Ogden et al., 2008b; Ogden et al., 2013). In eastern Canada, the tick population likely precedes bacterial invasion, typically occurring three to five years later (Ogden et al., 2013).

The effects temperature has on tick behaviour and survival have been well documented (e.g., Brunner et al., 2012; Burtis et al., 2016; Ogden et al., 2004). Fieler et al. (2021) showed that *I. scapularis* has the lowest survival at both low and high temperatures compared to other *Ixodes* species (*e.g., Amblyomma americanum, Dermacentor variabilis, Rhipicephalus sanguineus, Amblyomma maculatum, Dermacentor variabilis*), with continual temperatures between 4°C and under 32°C needed for reproduction and moulting success (Fieler et al., 2021). If *I. scapularis* are kept at temperatures upwards of 30°C, tick premoulting duration and moulting success rates are adversely impacted (Ogden et al., 2004). *I. scapularis* are killed outright when exposed to temperatures below -10°C to -12°C, whereas even lower subzero temperatures coupled with high humidity may induce lethal inoculative freezing (Burks et al., 1996; Nabbout et al., 2023). The colder temperatures historically limit the spread of *I. scapularis,* restricting them to only the lowest northern latitudes of Canada (e.g., Northwestern Ontario; Lindsay et al., 1995), where overwintering was and is possible, provided individual tick thermal limits are above the aforementioned threshold temperatures (Nabbout et al., 2023). Clark (1995) found that the temperatures in which *I. scapularis* can move are strongly correlated with

tick size. The temperature range for coordinated host-seeking behaviour in *I. scapularis* is above 13.9°C ± 3.0°C for nymphs, 9.2°C ± 4.1°C for males, and 11.2°C ± 3.4°C for females. (Clark, 1995). Clark (1995) further observed that the activity threshold, the temperature at which all activity ceases, is 9.8°C ± 3.4°C in nymph ticks, 8.5°C ± 3.0°C in males, and 6.2°C ± 3.6°C in females. To resume questing, the temperature must range between 4°C and 21°C on consecutive days (Clark, 1995), with immature ticks capable of questing at higher temperatures (Mount et al., 1997). Duffy and Campbell (1994) also observed a questing activity threshold of ~4°C; however, Schultz et al. (2001) found *I. scapularis* questing at temperatures as low as -0.6°C, suggesting that underlying factors (e.g., soil temperature) may explain the differences in temperatures (Shultz et al., 2001). Additional ways *I. scapularis* can persist in the cold is by employing inherent physiological and behavioural qualities (Nabbout et al., 2023), such as producing antifreeze molecules (Neelakanta et al., 2010), seeking leaf litter for insulation and protection, or autophagy (Nabbout et al., 2023). *B. burgdorferi*-infected *I. scapularis* may experience behavioural changes (e.g., increased phototaxis and lure to vertical surfaces; Benelli, 2020) and physiological adaptations (e.g., increased activation of tick histamine release factor [tHRF], lower mobility; Benelli, 2020; Nabbout et al., 2023). Female *I. scapularis* infected with *B. burgdorferi* may have increased resilience to survive overwintering in comparison to uninfected female ticks (Nabbout et al., 2023), potentially aiding in tick population growth (e.g., Elias et al., 2022). When the temperature is constantly below 4°C, female ticks do not lay eggs (Lindsay et al., 1995). The number of eggs a female produces strongly correlates with temperature (Gaff et al., 2020). If temperatures are 8°C and maintain high humidity (tested at > 95% humidity; Lindsay et al., 1995), female *I. scapularis* may lay eggs, but they will not metamorphose into larvae at temperatures lower than 10°C, potentially at 12°C (Lindsay et al., 1995). Lindsay et al.

(1995) used these values to develop threshold development temperatures, maintaining that 132±37 degree days (DD) above 6°C is required for oviposition in fed adult *I. scapularis* that overwinter, and 636±6 DD above 11°C for eggs to develop into larva. Rand et al. (2004) used the same threshold values for pre-oviposition and pre-eclosion, finding that degree days are apportioned around 243 and 429 DD, respectively. The discrepancies in values across each of the subsequent studies express some of the challenges in recording thermal dynamics and the developmental response of *I. scapularis* at the northern extremes of their range, whereas when latitude increased, the length of each stage progressively extended (Lindsay et al., 1995; Rand et al., 2004). Regardless, the likelihood of many *Ixodida* species completing ontogenesis is very low, at < 1% (Burtis et al., 2019). Diapause (delayed development) in immature *I. scapularis* ticks changes in response to the photoperiod (Ogden et al., 2004), with seasonal activity shown to be between 10-15 hours of daylight (Mount et al., 1997). These values are not immutable, and *I. scapularis* may have biological adaptations in different geographic spaces. They have shown capabilities to respond and adapt to environmental conditions, notably temperature (Mount et al., 1997).

Ecological Niche Modelling

Although not explicitly coined, the niche concept has been an ecological term since the 19th century. It was present in Darwinian ideas, where ecosystems are competitive and structured on survival within the interconnected assemblage of habitats, resources, and predators (Pocheville, 2015). The term 'niche' or 'ecological niche' is generally credited to being first contextualised (in a research capacity) by Joseph Grinnell in 1924, who defined it as the ecological position of a species and all the factors (e.g., biotic and abiotic) that condition its

existence (Pocheville, 2015; Vandermeer, 1972). By exploring communities in different regions, Grinnell found evidence of similar niches within geographically distant places (Pocheville, 2015) through evolutionary equivalency, where species that play the same ecological role but in independent locations occur (Pianka, 2011; Pocheville, 2015). Elton, another researcher exploring evolutionary equivalencies, furthered the niche concept theory by integrating the importance of interactions between species within a community (Sales et al., 2021). In 1957, a revolutionary change in ecology occurred (Schoener, 1989) when George Hutchinson postulated the niche around species' attributes rather than the environment's (Pocheville, 2015; Sales et al., 2021). This change was a nuanced shift from previous theories. Researchers such as Grinnell and Eltons saw the niche as independent and occupied *by* a species; Hutchinson saw the niche as a property *of* the species (Pocheville, 2015; Sales et al., 2021).

Hutchinson's work conceptualises and quantifies the ecological niche as an n-dimensional volume within a hyperspace (Colwell & Rangel, 2009; Hutchinson, 1957). The multidimensional volume exhibits the range of environmental conditions (each "n" is an environmental axis in the suite of environmental factors that make up a species niche, e.g., temperature, humidity), while the hyperspace separates the physical entity of place from the environment (Colwell & Rangel, 2009). If place is defined by geographic coordinates, and n-environmental attributes delineate the niche space of a species, it enables the potential to project a species' actual or potential geographical distribution and niche, present or future (Colwell & Rangel, 2009).

Also attributed to Hutchinson is the conception of the fundamental and realised niche (Hutchinson, 1957). The refined niche theory suggests that a species' fundamental niche is delineated by its physiological range of tolerance to environmental variables without biotic interactions (Soberón & Arroyo-Peña, 2017; Zurell, 2017) and permits the species to exist

indefinitely (Pocheville, 2015; Hutchinson, 1957) with population growth (Zurell, 2017). The realised niche is confined within the boundaries of the fundamental niche and is the niche *actually* occupied by the species (Pocheville, 2015). Hutchinson's realised niche was constrained only by species interactions (Colwell & Rangel, 2009) but has since extended to include the addition of dispersal, demographic processes, and interspecific species interactions, which may shrink the realised niche while source-sink dynamics, facilitation, and time-delayed extinctions may expand the realised niche (Zurell et al., 2017).

The finer concepts of the niche are enigmatic (Mathiopoulos, 2022), with ambiguity on how to measure, predict or analyse niches (Dormann et al., 2012; Sales et al., 2021). Ecological niche models (ENM), species distribution models (SDM), bioclimatic envelope models, and habitat suitability models are terms used synonymously (sometimes controversially; McInerny & Etienne, 2013; Peterson & Soberón, 2012; Zurell et al., 2020) to describe species' occupancy, habitat, niche, or invasiveness in ecological or biogeographical environments (Melo-Merino et al., 2020). These models are often employed to answer questions on geographical ranges, climate change impacts, or conservation strategies (Melo-Merino et al., 2020).

Three types of niche models stand out enough to delineate unique types: correlative models, mechanistic models, and process-based or hybrid models (Neineri et al., 2015; Peterson et al., 2015; Tourinho & Vale, 2023). Mechanistic and correlative models are conceptually different. Correlative models use implicit processes and focus on prediction (Tourinho & Vale, 2023). These models aim to depict the patterns, rather than the mechanisms, of the association between species occurrences and environmental data (Dormann et al., 2012). On the niche spectrum, correlative models usually estimate between the realised niche and the existing niche; a subset of the fundamental niche represents real-world geography (not just the suite of

environmental conditions, e.g., *I. scapularis* in Ontario, Quebec, Eastern US) paramount to the species (Peterson et al., 2015). In contrast, mechanistic models explicitly represent processes (Tourinho & Vale, 2023). They focus on understanding a species' environmental requirements and using biophysical models to determine the fundamental niche of a species. Mechanistic models often require detailed, accurate, and complex parameters such as microclimatic conditions, species physiology, and species behavioural characteristics to estimate a species niche (Peterson et al., 2015). Process-based models combine correlative and mechanistic approaches by integrating implicit and explicit causal aspects into models (Tournho & Vale, 2023). They estimate species distribution by additionally incorporating dispersal capabilities and biotic interactions into the modelling process (Melo-Merino et al., 2020; Peterson et al., 2015). Process-based models do not prioritise estimating species niches (Peterson et al., 2015). Instead, they use niche estimates, potentially produced by integrating correlative and mechanistic models, and incorporate dispersal characteristics to simulate the actions that produce an *occupied distribution* of a species: the area in which a species may be found and potentially persist (Peterson et al., 2015).

I used correlative models to estimate the *I. scapularis* niche. Correlative models use empirical, telemetry-based species occurrence data and environmental variables to infer a species niche. The models produced are only as good as the input data used to build them (Collins et al., 2017). Parameters used within these models have no *a priori* ecological context (Dormann et al., 2012) or, at present, little ecological hypotheses (Gobeyn et al., 2019). Instead, the characteristics of the environments where the target species is recorded can identify the conditions in which the same target species can survive (Peterson et al., 2015). Machine learning and statistical

algorithms are employed to uncover spatial patterns within the data and to estimate species' responses to environmental predictors (Gobeyn et al., 2019).

Higher-resolution and precise environmental layers are becoming increasingly produced and accessible. Most available environmental layers are derived from statistical interpolation from weather stations, satellite data, or statistical downscaling of climate layers (Morales-Barbero & Vega-Álvarez, 2019). Downscaling and satellite environmental data are more accurate than statistical interpolation because weather stations which track environmental conditions are unevenly distributed. However, inputs from satellite data may still deviate from observed environmental values, particularly some regional temperature, humidity, and precipitation estimates (Morales-Barbero & Vega-Álvarez, 2019). Discrepancies, uncertainties, or biases inherent in most climatic input data highlight the significance of selecting a suitable database and caution against assuming that a single database is best suited for all regions and scales (Morales-Barbero & Vega-Álvarez, 2019).

Species occurrence data is also becoming increasingly available with improved sharing initiatives, better ways to generate data, and broader conservation and scientific communities aiding in expanding biodiversity databases (Aubry et al., 2017; Graham et al., 2008). If we first consider Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970), then it logically follows that a greater abundance of presence points enhances the precision of assessing spatial autocorrelation between variables. Otherwise, there is no environmental gradient; all environmental variables are considered equal to assessing the species distribution, which within natural contexts is exceedingly unlikely (Sillero & Barbosa, 2020). Collected species records often show a spatial clustering of points in areas that are more accessible to larger human populations (e.g., near

urban zones, cities, and roads), which leads to a spatial bias. This bias explains ecological patterns in a way that favours geographic space and strongly autocorrelated environmental variables (Franklin, 2010), which can lead to overfitting when assessing model accuracy, coefficients, residuals (e.g., more type 1 errors), and statistical inferences (Franklin, 2010; Sillero & Barbosa, 2020). The conventional approach to alleviating bias is to thin the dataset, either by a nearest neighbour distance equation or through random sampling (Aiello-Lammens et al., 2015). However, thinning the data could be problematic. Eliminating a substantial portion of data that captures small-scale environmental conditions or endogenous spatial autocorrelation data may make models less accurate and realistic (Franklin, 2010). Determining the ideal quantity of records to exclude relies on empirical trial and error, which is time-consuming and computationally intensive (Aiello-Lammens et al., 2015).

Correlative-based models can utilise species records in three ways: presence-only, presence-background or presence-absence models, with each method using inherently different mathematical approaches. As the name suggests, present-only models, sometimes classified as profile methods (Hijman & Elith, 2023), only require precise species sightings, limiting the environmental conditions to those within a bounding envelope of their known distribution. Various distance rules and polygonal envelopes (e.g., rectilinear shapes; Barbet-Massin et al., 2012) help define the environmental hyperspace of the species; any unknown sites are compared to a percentile distribution of environmental values within that hyperspace and given a predictive score (Hijman & Elith., 2023; Piñeiro et al., 2007; Senay et al., 2013). Some examples of presence-only models are BIOCLIM, Mahalanobis models (MAHAL), and DOMAIN (Hijman & Elith, 2023; Senay et al., 2013). Presence-background models take the same observation data and generate "background" points, new records that indiscriminately sample environmental and

geographic spaces within the designated extent (Liu et al., 2019). The models look at the environmental conditions of known records relative to those in alternative spaces and infer a final prediction (Senay et al., 2013). The widely used maximum entropy approach (MAXENT) is an example of this type, along with ecological niche factor analysis (ENFA; Senay et al., 2013) and genetic algorithms for rule set production (GARP; Barbet-Massin et al., 2012). Presence-absence models, sometimes called group discrimination methods, apply a different methodology, availing of absence data. Absence points, often called true or real absences, are positions where the environmental conditions are unsuitable for the species to survive (Liu et al., 2019). Statistically, it is better to model using negative classes, which are absence points, in conjunction with positive classes, which are the occurrence points, than just positive classes (Manevitz & Yousef., 2001; Senay et al., 2013). The main caveat is that true absences are rare; they often occur from active surveillance by entomologists or health authorities, which can be expensive, time-consuming and labour-intensive to cover a limited area (Soucy et al., 2018). Additionally, the absence of the species can stem from factors beyond environmental unsuitability, such as unique ecosystem dynamics (e.g., barriers, predation, competition, temporal anomalies) or instances where it remained undetected during survey efforts (Senay et al., 2013).

The complexities associated with acquiring authentic absence data have prompted the adoption of pseudo-absences, artificially generated absence points serving as the negative or opposing class in datasets (Jiménez & Soberón, 2020). These pseudo-absences are strategically positioned within environmental or spatial dimensions to maximise the detection of environmental or spatial disparities, namely, within dissimilar environmental contexts that capture local variations rather than in incongruous locations where the species would obviously

be absent (e.g., ticks in the ocean; Senay et al., 2013). Only a few studies have explored the placement and abundance of pseudo-absences in-depth (Barbet-Massin et al., 2013; Liu et al., 2019) despite their influence on algorithms (Senay et al., 2013). Within the literature, pseudo-absence selection methods are often varied, with contradicting claims and results per method (Senay et al., 2013). If the input data is reliable, abundant, and all-encompassing of a species' bionomic conditions, presence-only or present-background methods can be helpful, but if there is bias or uncertainty with the input data, pseudo-absence methods may be better at reducing overfitting and accurately approximate the output when provided new information (e.g., a changing climate; Senay et al., 2013). In general, presence-absence algorithms are more robust and perform the best of the correlative approaches, along with the presence-background method MAXENT (Liu et al., 2019).

Given that variations in mathematical algorithms lead to increased sensitivity in processing derivatives, resulting in disparity amongst model results (Araújo & New, 2007), an ensemble framework was introduced to help alleviate model-based uncertainty in correlative models. Based on a succession of individual models, the ensemble (consensus) approach is founded on the theory that individual predictors are replicas of potential states of the real distribution (Marmion et al., 2009). However, each state can vary considerably per modelling technique - particularly in biogeographical or species distribution studies (Marmion et al., 2009). When combined within a single estimator, the variance is reduced, producing an ensemble that indicates trends or points of agreement within the models (Zhu et al., 2021), feasibly improving accuracy and providing a more concise picture. Some standard techniques to transition individual models into ensembles are to take the mean or median of single models or the weighted average or principal component analysis median, subdividing single models based on fixed criteria

(Marmion et al., 2009). In addition, models have shown the potential to lose accuracy when extrapolated into environmental conditions outside those used to train the models (e.g., changing of novel climate conditions common in most future projections). In these situations, simple models show greater transferability (lower interpolative), while complex models show greater explanatory value (lower transferability, Zhu et al., 2021). However, both transferability and explanatory power can not be simultaneously maximised (Bokulich, 2013), so using a combination of models depends on the context itself. There is no *de facto* model algorithm (Qiao et al., 2015); a consensus model helps create modelled scenarios closest to the truth by incorporating all circumstances (Araújo & New, 2007).

Chapter 2: Ecological Niche Modelling of the Blacklegged Tick (*Ixodes scapularis*, Say) in Eastern Canada with the Aid of Community Science Data

Introduction

Atmospheric greenhouse gases, extreme weather events, and sea-land temperatures are experiencing concentrations, frequencies or accelerated rates not seen for millions of years, decades, or millennia, respectively. Without intervention, global warming in this century will reach or surpass 1.5°C to 2°C above pre-industrial levels. (IPCC, 2021b). This warming, attributable to human-induced greenhouse gas emissions (IPCC, 2021b), will lead to a shift in climatic zones, with implications for the distribution of species (i.e., species range, seasonal phenomena, abundance, IPCC, 2019).

Many terrestrial species have shifted poleward in response to changes in climate envelopes (climates in which given species can live). The shifts observed in species distributions have been rapid, exhibiting an average displacement of 17 (±3) km toward higher latitudes and 11 (±2) m towards higher elevations per decade throughout the preceding 70 years (IPCC, 2021a). The changes in species distribution have significant implications for changes in species-environment and species-species interactions. One novel species-species interaction that is of concern to people is between humans and vectors (e.g., ticks) of zoonotic pathogens such as *B. burgdorferi* (Lyme disease).

Pest species, such as ticks, are responding rapidly to climate change with changes in their distribution, abundance, phenology, microbiota, development rate, species phenotype, host preference, and genotype (Backus et al., 2021; Nuttall, 2021; Ogden et al., 2021). *Dermacentor variabilis* Say, 1821; *Amblyomma americanum* Linnaeus, 1758; and *Ixodes ricinus* are all examples of tick species experiencing range expansion (Nuttall, 2021). *Haemaphysalis longicornis* Neumann, 1901, and *Hyalomma marginatum marginatum,* Koch, 1844, are species that have been detected in novel environments outside of their traditional ranges, such as within new countries (Nuttall, 2021). *I. scapularis* is of particular concern in northeastern North America as it is the primary vector of Lyme disease (Bouchard et al., 2019). Over the past three decades, there has been a notable northward expansion of *I. scapularis* into central and eastern Canada, particularly in Ontario, where decades prior, there was minimal evidence of an established population (Robinson et al., 2022). *I. scapularis* is capable of both short and long-range dispersal (via migratory birds, Ogden et al., 2008b). Climate change and host species movements are considered key contributors to subsequent range expansion (Robinson et al., 2022), with studies showing that temperature, rainfall, elevation (Leighton et al., 2012), canopy

cover, and proportion of forested landcover (Slatculescu et al., 2020) are important environmental drivers of *I. scapularis* occurrence and abundance.

Ecological niche models effectively develop predictions of current and future species distributions (Zhu et al., 2021), thereby aiding in the formulation of conservation planning efforts (Esser et al., 2019). Leveraging environmental and georeferenced empirical occurrence data, these models are often employed to forecast a species shifting niche under different climate scenarios (Zhu et al., 2021). With increasingly accurate and precise environmental data and the advancement of novel statistical methods (i.e., the introduction of machine learning techniques), consensus (ensemble) modelling approaches are routinely used (Zhu et al., 2021) to combine the results from individual modelling algorithms to produce a more robust, explicit result, including multiple modelling parameters (Melo-Merino et al., 2020).

Studying the expansion of *I. scapularis* into eastern Canada is invaluable to understanding other broader ecological shifts influenced by climate change. As *I. scapularis* is highly sensitive to environmental conditions, movement into novel territories is an indicator of shifting climatic patterns and can provide insight into changing tick-borne disease dynamics. Whereas previous studies have typically focused on evaluating the expansion of *I. scapularis* range in Canada at a provincial scale, this study advances the field by employing a more extensive spatial scope. It stands out for its thorough utilisation of occurrence records and environmental variables, as well as the comprehensive use of algorithms and model outputs, all within the broader context of eastern Canada. I use ecological niche models to determine the key drivers that explain the current distribution of *I. scapularis* and how changes to those drivers (i.e., climate change) will influence the species' range and ability to establish across eastern Canada this century. In addition, I aimed to determine which environmental drivers used within

the models best explain the current climatic distribution of *I. scapularis* and how they respond to the environmental drivers, their optimum conditions relative to covariate values, and their tolerance to changing conditions. I test whether the leading edge of *I. scapularis'* range will push further into Canada, colonising new, previously uninhabitable landscapes that could act as dispersal-foci (Rehm et al., 2015). From a contemporary context, I test if established *in situ I. scapularis* populations in southeastern Canada will become more prevalent as their range expands where, typically, optimal conditions for a species are found near the center of their range (Rehm et al., 2015).

Methods

Data

I combined a comprehensive dataset of georeferenced verified tick occurrence data from 2017-2022 within Manitoba, eastern Canada (Ontario, Quebec, Newfoundland and Labrador, Prince Edward Island, Nova Scotia and New Brunswick), and the United States (North and South Dakota, Nebraska, Kansas, Oklahoma, Minnesota, Iowa, Missouri, Arkansas, Louisiana, Wisconsin, Illinois, Tennessee, Mississippi, Michigan, Indiana, Kentucky, Alabama, Ohio, Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New Jersey, Delaware, Maryland, Washington D.C, West Virginia, North and South Carolina, Georgia, and Florida) from iNaturalist and eTick (eTick, 2021; iNaturalist, 2022).

Table 1. The identified whereabouts and quantity of *I. scapularis* documented from 2017 to 2022.

Country	Province/State	Count of *I. scapularis* records
CAN	Ontario	4766
CAN	Quebec / Québec	2013
CAN	Nova Scotia / Nouvelle-Écosse	1044
USA	New York	472
USA	Pennsylvania	376
USA	Massachusetts	375
CAN	New Brunswick / Nouveau-Brunswick	328
USA	Vermont	325
USA	Ohio	156
USA	Wisconsin	121
USA	New Jersey	119
USA	Virginia	113
USA	Minnesota	113
USA	Maryland	111
USA	Connecticut	102
USA	Michigan	96
USA	New Hampshire	85
USA	Texas	85
USA	Illinois	78
CAN	Prince Edward Island / Île-du-Prince-Édouard	72
USA	Florida	66
USA	Maine	53
USA	West Virginia	52
USA	Tennessee	36
USA	North Carolina	36
USA	Indiana	35

USA	Georgia	35
USA	Kentucky	29
USA	Alabama	27
CAN	Newfoundland and Labrador / Terre-Neuve-et-Labrador	26
USA	Oklahoma	21
USA	South Carolina	19
USA	Missouri	17
USA	Mississippi	15
USA	Rhode Island	15
USA	Delaware	14
USA	Iowa	14
USA	Arkansas	14
CAN	Manitoba	11
USA	Louisiana	7
USA	District of Columbia	6
USA	Kansas	3
USA	North Dakota	1
USA	South Dakota	1
USA	Nebraska	1
	Total	11504

Records were not collected equally or uniformly across the five years, with ~75% collected in 2020 and 2021; the 2022 records were only collected during a six-month period, from January to June (for a complete breakdown, see Appendix H). Most acquisitions were acquired during spring and early summer (April, May, and June) and in October and November during the fall season. January contained the lowest number of identifications, with only a select few recorded within Canada (lower Ontario and Quebec).

I removed records with coordinate systems treated as private or obscured from within the iNaturalist dataset; if requested by the uploader, iNaturalist will conceal *I. scapularis* sightings randomly within a 0.2° x 0.2° (~22 x 22 km) proximity for privacy reasons (Loarie, 2021). This large envelope diminishes the precision of the location data and prevents the records from matching with the environmental data (~1 km resolution) at the observed tick location. I filtered additional sightings if data entries had erroneous positional accuracy of upward of ~1 km or showed apparent errors or inconsistencies. *ArcGIS Pro 3.0* (ESRI, 2020) and *QGIS 3.28.4* (QGIS Development Team, 2022) were used to overlay merged point data (eTick and iNaturalist records) over a North American shapefile: "Canada - United States Boundary," provided by Natural Resources Canada [NRCan] and the U.S. Geological Survey ([USGC], 2010).

Bioclimatic data was sourced from two locations: *AdaptWest* (AdaptWest Project, 2022; Mahony et al., 2022; Wang et al., 2022; Wang et al., 2016) and *Chelsa-Bioclim+* (Brun et al., 2022; Karger et al., 2023; Karger et al., 2021a; Karger et al., 2021b), in GeoTiff format. Both datasets have a spatial resolution of ~1 km, derived by downscaling climate projections from the Coupled Model Intercomparison Project Phase 6 (CMIP6) database and standardised by values reported in the 6th IPCC Assessment Report (AdaptWest Project, 2022). The Chelsa-Bioclim+ raster dataset used was modelled on the Geophysical Fluid Dynamic Laboratory: Earth System Model 4 (GFDL-ESM4), with historical data derived from 1981-2010. Future climatologies are from 2011-2040, 2041-2070, and 2071-2100 (Brun et al., 2022) using emission scenarios of SSP5-8.5 and SSP3-7.0. The AdaptWest rasters were interpolated by *ClimateNA v7.3*, with future values averaged over 13 CMIP6 Atmosphere-Ocean General Circulation Models (CMIP6-AOGCM) through 30-year periods. Historical data exist from 1991-2020, followed by predictive climate normals for the future timeframes: 2011-2040, 2041-2070, and 2071-2100

using SSP5-8.5 and SSP3-7.0 values (AdaptWest Project, 2022). Additional data included the "Global Man-made Impervious Surface Dataset From Landsat, v1 (2010)" mosaiced by Brown de Colstounet et al. (2017) with a Kappa coefficient of ~0.92 (evaluated by a cross-validated random forest; Zhang et al., 2020). The impervious surface data for Canada and the United States is available at a resolution of 30m.

Procedure

There is a lack of congruences within different climate datasets (Morales-Barbero & Vega-Álvarez, 2019). For this reason, I utilised two datasets and two SSP scenarios within the study to produce more reliable predictions. The Chelsa-Bioclim+ dataset, which I considered the primary dataset, outlines the results presented in this thesis. I based the essential findings and outcomes on the information and data provided by this dataset. The other dataset, AdaptWest (herein mentioned as AdaptWest, the complementary dataset, or the secondary dataset), provided a secondary function to my research. The complementary dataset, which provided more up-to-date historical data (and an expanded array of future models), was used to cross-reference the results of the first dataset, certifying that environmental variables and severity of range expansion were similar, enhancing the robustness and credibility of the models while incorporating additional environmental factors not covered within the primary set (e.g., degree days, seasonal precipitations).

The environmental data was used first to explore *I. scapularis'* contemporary distribution and later forecasted to identify their potential geographic expansion (over 30-year increments, i.e., 2011-2040, 2041-2070, and 2071-2100). The forecasting data used SSP3-7.0 and SSP5-8.5 values to account for a spectrum of future emission scenarios; SSP5-8.5 is a high-emission

scenario, while SSP3-7.0 provides more conservative, middle-high estimates (Meinshausen et al., 2020).

I projected rasters and shapefiles into *Lambert Azimuthal Equal Area* spatially referenced projection system from the standard Geographic Coordinate System (GCS), World Geodetic System 1984 (WGS 1984). The "Canada - United States Boundary" shapefile (NRCan & USGC, 2010) was clipped within *ArcGIS Pro 3.0* (ESRI, 2020) to a spatial extent based explicitly on an understanding of *I. scapularis* fundamental niche (the environmental niche or E-space; Melo-Merino et al., 2020). The extent allowed me to integrate a considerable number of pseudo-absence points situated within optimal ranges of the known distribution (VanDerWal et al., 2009).

The AdaptWest and Chelsa-Bioclim+ rasters were masked within *ArcGIS Pro 3.0* (ESRI, 2020) to match the study's extent. Utilising the raster calculator tool, I rescaled the values. Specific environmental parameters underwent internal scale adjustments and offsets to convert them into integers, simplifying the download process (Karger et al., 2022). Some versions of GDAL or ArcGIS do not automatically rectify these offsets and scale changes, necessitating manual adjustments to conform to conventionally known values (adhering to CHELSA V2.1 technical documentation; Karger et al., 2022).

The "Global Man-made Impervious Surface Dataset From Landsat, v1 (2010)" was imported into *ArcGIS Pro 3.0* (ESRI, 2020) and converted to a polygon layer to overlay against the *I. scapularis* records to thin the distribution points.

Statistical Analysis

My framework consisted of running complementary R-scripts under modified conditions, targeting culpable factors of model bias, overfitting, and uncertainty - increasing model strength and realism. Each time the script was processed (herein called "sequences"), the ensuing sequences used different environmental variables to capture further the broad spectrum of biotic and abiotic factors that mediate *I. scapularis* distribution. Reproducing models leads to noticeable patterns in the resulting raster images, response curves, and permutation importance, culminating in establishing a final predictability paradigm (i.e., ensemble).

I integrated statistical modelling using *R Programming Language* (R Core Team, 2020) within the *RStudio* environment (RStudio Team, 2020). The following packages (and their correspondent dependencies) automated tasks: *sdm* (Naimi & Araujo, 2016)*, rgdal (*Bivand et al., 2022)*, raster (*Hijmans, 2022)*, plyr* (Wickham, 2011)*, usdm* (Naimi et al., 2023)*, caret* (Kuhn, 2022)*, biomod2* (Thuiller et al., 2022)*, ENMeval* (Kass et al., 2021)*, gridExtra* (Auguie, 2017)*, dplyr* (Wickham et al., 2022)*, CENFA* (Rinnan, 2021)*,* and *corrplot* (Wei & Simko, 2021)*.* The reproducible code is available in Supplementary Materials.

The georeferenced records (turned point data) and environmental variables were transferred and staged within the RStudio program. Environmental variables were tested for multicollinearity problems in each sequence, with pairwise exclusion of highly correlated variables using variance inflation factor (VIF) and VIFstep. Given the ecological insights highlighted in the literature concerning the significance of certain variables in influencing *I. scapularis* distribution (e.g., temperature as a causative factor), I used a correlation threshold of < 0.8 and a VIFstep threshold of < 10 between each paired variable in each sequence (Dormann et al., 2012; Naimi et al., 2014, Naimi, 2023). However, I prioritised maintaining a a correlation

coefficient of < 0.7 whenever feasible, aligning with recommendations found in some literature (Naimi et al., 2014). The remaining variables (through VIF and trained analysis) were overlaid with the point data. Raster cell values corresponding to the established environmental variables were extracted and recorded, while any records with null attributes were excluded.

I used twelve algorithms in the production of the ensemble of models: Generalised Linear Models (*GLM*), Generalised Additive Models (*GAM*), Generalised Linear Models with Polynomial Regression (GLMPoly), Multivariate Adaptive Regression Splines (*MARS*), Multiple Discriminant Analysis (*MDA*), Flexible Discrimination Analysis (*FDA*), Recursive Partitioning (*RPART*), Classification and Regression Trees (*CART*), Boosted Regression Trees; otherwise known as Generalised Boosted Regression Models (*BRT/GBM*), Random Forests (*RF*), Support Vector Machines (*SVM*), Radial Basis Function Network (*RBF*), and Multilayer Perceptron (*MLP*). Applicable algorithms were selected and tuned using the *caret* package to find best-fit, modifiable, manually-set hyperparameters that control the learning process of a model and can increase model performance (Kuhn, 2022; Yang & Shami, 2020). Using a 5-fold cross-validation procedure run eight times with a tuning length of 50 (i.e. 50 combinations of each configurable tuning parameter were tested), the values of the hyperparameters, which can be continuous, integers, or categorical (Hutter et al., 2015), were selected based on the highest receiver operating characteristic curve (ROC) value. GAM models were built using splines, with 'select' and 'method' as tuning hyperparameters. MDA models use subclasses, while FDA and MARS (using bagging and generalised, cross-validating pruning) have hyperparameters of the number of 'degrees' and 'nprunes.' RPART is CART using recursive partitions (instead of the tree command; Morgan, 2014), which uses complexity parameters ('cp') to calibrate the algorithm. BRT becomes stochastic gradient boosting models, optimised with 'n.trees,' 'interaction.depth',

‘shrinkage,’ and ‘n.minobsinnode.’ RF became a regularised model with: ‘mtry’, ‘coefRef’, and ‘coefImp’ hyperparameters. SVMs are built with polynomial kernels and augmented with ‘degree’, ‘scale’ and ‘C’ values. Lastly, RBF uses a ‘negativeThreshold’ value while MLP uses multiple layers: ‘layer1’, ‘layer2’, and ‘layer3’.

AdaptWest & Chelsa-Bioclim+ Sequences 1-3

The position and quantity of pseudo-absences have been extensively studied (and scrutinised, Chefaoui & Lobo, 2008; Dubos et al., 2022; Hazen et al., 2021; Liu et al., 2019; Senay et al., 2013; VanDerWal et al., 2009; Wisz & Guisan, 2009). I based the conceptual generation of pseudo-absences for each algorithm on the working of Barbet-Massin et al. (2012), which provides insight into the optimal quantity and distribution of pseudo-absence points, considering factors such as spatial or climatic biases within the data (Barbet-Massin et al., 2012). Subsequent studies, such as the work conducted by Liu et al. (2019), have affirmed the results presented by Barbet-Massin et al. (2012). I iterated the execution of an R-script three times, each time selecting identical environmental variables but varying the presence records and the ratio of pseudo-absences to reflect these changes. The configuration that yielded the most robust models and indicated a comprehensive understanding of *I. scapularis'* environmental dynamics, particularly evident in how closely the models' response curves aligned with *I. scapularis'* preferred environmental conditions, was selected for subsequent sequence and model constructions.

The initial sequence utilised all available records, potentially resulting in multiple records within each 1 km aggregate grid of environmental values. In the second scenario, records were thinned without altering the spatial distribution, focusing solely on reducing abundance. This

nuanced adjustment led to a reduction of records by more than half. The third change prioritised keeping records outside the scope of human activities, such as within secluded regions with less human traffic, manufacturing a distribution influenced less by accessibility and thinned to one record per grid cell. Sequences one to three yielded approximately 12,000; 5,500; and 3,500 I. scapularis records, respectively

The following variables were used from the Chelsa-Bioclim+ dataset for sequences 1-3: (a) mean diurnal air temperature range (Bio2), (b) mean daily air temperature of the wettest quarter (Bio8), (c) precipitation amount of the wettest month (Bio13), (d) precipitation seasonality (Bio15), (e) net primary productivity (NPP; based on Lieth's "Modelling the primary productivity of the earth", 1975).

The following variables were used from the AdaptWest dataset: (a) Elevation (ELEV), (b) Mean annual precipitation (MAP), (c) Precipitation as snow (PAS), (d) Mean annual relative humidity (RH), (e) Temperature difference between the mean temperature of the coldest month (MCMT) and the mean temperature of the warmest month (MWMT), (f) Annual heat moisture index (AHM).

For AdaptWest and Chelsa-Bioclim+ sequence one (i.e., all records, biased distribution), GLM, GAM, GLMPoly, and MARS used 10,000 pseudo-absence points, randomly distributed across the study area. MDA, FDA, RPART, and CART used 4,000 pseudo-absence points, using a 2°far distribution from presence records (i.e., a $\geq$ 222 km buffer from known occurrence points). BRT, RF, RBF, and MLP algorithms used 11,500 pseudo-absence points, similar to the count of *I. scapularis* records, with an SRE distribution (i.e., a minimum bounding geometric polygon is positioned around the presences, dispersing pseudo-absences outside that envelope). In the second sequence, where records were reduced to one point per grid cell (~5,500 points),

the same set of algorithms and pseudo-absence distributions were applied. However, the allocation of pseudo-absence points varied: GLM, GAM, GLMPoly, and MARS models employed 10,000 pseudo-absences, while MDA, FDA, RPART, and CART models utilised 1,750, and BRT, RF, RBF, and MLP models used 5,500. Sequence three used the impervious surface shapefile with interval buffers to mollify human sampling bias from heavily frequented areas. Each interval extended by 10 meters progressively, reaching a distance of up to 50 meters from any artificially created surfaces or structures.

Area	Remaining	Removed
default overlay	7,951 observations	3,772 observations
10–meter buffer	7556 observations	4,157 observations
20–meter buffer	7,266 observations	4,457 observations
30–meter buffer	7,067 observations	4,659 observations
40–meter buffer	6,846 observations	4,877 observations
50–meter buffer	6,703 observations	5,013 observations

Table 2. Filtered out georeferenced points situated within or in proximity to an impervious surface shapefile.

I used a 50-meter buffer; enough points remained to accurately assess *I. scapularis* environmental conditions (within their niche) while also addressing potential errors or temporal gaps in the spatial alignment of relevant layers (e.g., point vectors, polygon vectors, environmental variables). The remaining points were lowered to one point per grid cell, resulting in approximately 3,500 observations remaining. Pseudo-absences were then allocated as follows:

10,000 for GLM, GAM, GLMPoly, and MARS; 1,100 for MDA, FDA, RPART, and CART; and 3,400 for BRT, RF, RBF, and MLP algorithms.

The results of sequences 1-3 determined the subsequent number of presences used in sequences 4-8 (i.e., ~3,500 presence records), with sequence three superseding one and two in the ensembles. The resulting rasters, response curves, and variable significance were comprehensively similar; however, sequences one and two showed additional signs of overfitting. This manifested in specific algorithms showing inflated AUC, sensitivity, and specificity values, along with difficulties in isolating finer-scale environmental conditions where one variable exerted an unrealistic disproportionate influence. Given the identical environmental variables shared among sequences 1-3, I retained only one sequence. Sequence three demonstrated superior performance, excluding sequences one and two from further consideration, thereby employing only sequence three in the final ensemble. Additionally, to better optimise the amalgamation of models used, I augmented the combinations of algorithms in subsequent sequences to refine the performance of underperforming model types. This alteration affected the count and placement of pseudo-absences in separate algorithms, aiming to improve the performance of underperforming model types.

AdaptWest & Chelsa-Bioclim+ Sequences 4-7

AdaptWest & Chelsa-Bioclim+ Sequences 4

Sequence four for AdaptWest used (a) degree days above 18°C (DD18), (b) degree days below 0°C (DD_0), (c) mean annual relative humidity (RH), (d) Hargreaves climate moisture

deficit (CMD; Hargreaves & Allen, 2003), (e) mean annual precipitation (MAP). Sequence four for the Chelsa-Bioclim+ models used: (a) isothermality (Bio3), (b) annual precipitation amount (Bio12), (c) frost cover frequency (FCF), (d) growing season length (GSL; per "TREELIM"; Paulsen & Körner, 2014) and, (e) growing season temperature (GST; in keeping with "TREELIM"; Paulsen & Körner, 2014). The amended pseudo-absence selections for both datasets continue to group GLM, GAM and GLMPoly models as previously (with 10,000 random records). However, MARS gets separated with considerably fewer pseudo-absence records (1,100; 2° away). MDA and FDA models used 1,100 absences with an SRE distribution – removing the RPART and CART models, which are integrated with the same count and spatial diffusion as BRT, RF, SVM, RBF, and MLP models (3,425; SRE distributed pseudo-absences).

AdaptWest & Chelsa-Bioclim+ Sequences 5

Sequence five used the same pseudo-point parameters, with six Chelsa-Bioclim+ covariates: (a) temperature seasonality (Bio4), (b) mean daily air temperature of the warmest month (Bio5), (c) accumulated precipitation of the growing season (GSP; adhering to "TREELIM" values; Paulsen & Körner, 2014), (d) and the last growing degree days above 0°C (GDDLGD0), (e) above 5°C (GDDLGD5), (f) and above 10°C (GDDLGD10). AdaptWest sequence five used: (a) Hogg's climate moisture index (CMI; NRCan, 2020), (b) mean precipitation amount during the summer (MSP; precipitation amounts between May to September), (c) mean spring precipitation (PPT_sp), (d) and the number of degree-days above 5°C (DD5).

AdaptWest & Chelsa-Bioclim+ Sequences 6

The distribution parameters of the MARS algorithm were changed to "environmentally random." Chelsa-Bioclim+ sequence six used: (a) annual range of air temperature (Bio7; difference between max temperature of warmest month and minimum temperature of the coldest month), (b) mean daily air temperature of the warmest quarter (Bio10), (c) precipitation amount of the driest month (Bio14), mean monthly precipitation of the warmest quarter (Bio18), and snow water equivalent (SWE; melted amount of snow turned to liquid water) as variables. AdaptWest sequence six used (a) mean annual temperature (MAT), (b) summer heat moisture (SHM), (c) summer precipitation (PPT_sm; precipitation amounts between June and August), (d) precipitation during the winter (PPT_wt).

AdaptWest & Chelsa-Bioclim+ Sequences 7

Chelsa-Bioclim+ sequence seven used three environmental variables: (a) Mean annual temperature (Bio1), (b) mean monthly precipitation amount of the wettest quarter (Bio16), (c) mean monthly precipitation amount of the driest quarter (Bio17). The AdaptWest dataset used: (a) degree-days above 10°C and below 40°C (DD1040), (b) extreme maximum temperature over 30 years (EXT), (c) spring precipitation (PPT_sp), (d) summer precipitation (PPT_sm), (e) autumn precipitation (PPT_at).

Evaluation Metrics

Algorithms were trained with an 80:20 data-partition, training-test split of the data (replicated four times - constrained due to limitations in computer memory). A training dataset is used to 'teach' a model; the algorithms 'learn' patterns and relationships within the data. The test dataset is used to evaluate model performance post-training. Resampling methodologies were adopted to evaluate the models and assess their efficacy. Resampling techniques involve iteratively selecting samples to constitute the test set. Three resampling techniques were employed: subsampling, bootstrapping, and 5-fold cross-validation methods (Naimi & Araujo, 2016). Each algorithm generated twenty-eight models. Resampled models were trichotomised into four subsampled, four bootstrapped, and twenty cross-validated methods. The evaluation of each model performance, derived from the dependent test data, considered three metrics: Area Under the ROC (receiving-operating characteristics) Curve (AUC), True Skill Statistic (TSS), and Cohen's kappa (Kappa).

Liu et al. (2013) suggested using pseudo-absences as true-absence data to create confusion matrices (Jiménez & Soberón, 2020), which provides a means to evaluate models by deriving the sensitivity (correctly predicted presence locations) and specificity (correctly predicted absence locations) of a classification and verifies those against the false-positive rate (i.e., 1- specificity, Shabani et al., 2018).

Cohen's kappa (Kappa) measures the degree of concordance between the observed and predicted and determines if this agreement is higher than what would occur from chance alone (Liu et al., 2009). Kappa classifies predictions as either presence or absence but also measures

model calibration by evaluating the ratio of correct presences summed over all observations (Steen et al., 2021). A Kappa value can range from -1.0 to 1.0, with the suggested standard interpretation of "values ≤ 0 as indicating no agreement and 0.01-0.20 as none to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement" (McHugh et al., 2012, p. 4). While kappa is a standardised measurement across studies, the consequential weight of the agreement depends on the discipline in question (e.g., clinical trials in healthcare: a substantial agreement would still lead to approximately 20% to 40% of evaluations being wrong; an unpropitious result for quality control, such as in the assessment of a new medicinal drug (McHugh et al., 2012). A suggested agreement quartile for species-climate envelope modelling is that Kappa values ≤ 0.40 indicate poor agreement, values between 0.40-0.75 indicate good agreement, and values ≥ 0.75 indicate excellent agreement (Araújo et al., 2005)."

The AUC of a model represents the likelihood of the model ranking a randomly selected presence site higher than a randomly chosen absence site (Liu et al., 2011). A score of 0.5 suggests the model performs no better than random chance (Elith et al., 2006; Raes & ter Steege., 2007). A higher AUC value signifies that algorithms learned well and performed better than a random classification or chance occurrence (Jiménez & Soberón., 2020; O'Neill et al., 2023); a score of 1.0 signifies a perfect fit of the model (Elith et al., 2006; Raes & ter Steege., 2007). When predicting species niche using presence-absence data, an ordinal value ≤ 0.7 could indicate a poor model, a value between 0.7-0.9 could indicate a good model, and ≥ 0.9 could point to an excellent model (Ringwaldt et al., 2023; Zhang et al., 2015).

TSS demonstrates a robust correlation with AUC metric values and is regarded as a more realistic statistic for assessing the accuracy of biodiversity models (Shabani et al., 2018). TSS

and AUC are both prevalence-free methods; however, they can be distinguished because TSS is a threshold-dependent statistic. In niche modelling, it is sometimes advantageous to use a binary distribution classification over a continuous result (e.g., for hotspot analysis or presence-absence maps; Shabani et al., 2018). TSS can be a better metric for evaluating models developed on the same dataset (Liu et al., 2015). TSS uses a threshold to dichotomise a confusion matrix's true presences and false absences (i.e., true presence values fall above the threshold; false absence values fall below the threshold; Liu et al., 2015). I chose a threshold, MaxSSS, that maximises the sensitivity and specificity (Max[se+sp]) of the present/absence data. MaxSSS optimises objectivity, discriminability, and equality; helps evaluate TSS values that may have inflated AUC or Kappa metrics; and is considered one of, if not the best, threshold methods for ENMs using present/absent data (Liu et al., 2015; Liu et al., 2013; Liu et al., 2005)

AUC and TSS are discrimination metrics that are ideal for predicting the probability of species presence/absence or the relative suitability of a location. Kappa is considered a calibration metric effective at dealing with uncertainty or prevalence (Steen et al., 2021). Models with metrics below the threshold values of AUC < 0.80 (Li et al., 2023), TSS < 0.70 (e.g., Thuiller et al., 2019), and Kappa < 0.6 (e.g., as suggested in McHugh's work, 2012) were removed.

Models within each sequence, which utilise distinct environmental variables, underwent separate evaluation from other sequences. Models failing to meet the threshold criteria (i.e., AUC < 0.80, TSS < 0.70, and Kappa < 0.6) were eliminated. The top-performing model from each algorithm was singled out; in cases where multiple top models demonstrated comparable performance metrics, they were amalgamated to reduce noise, averaging covariate influence, response curves, and niche predictions. Each algorithm within each sequence now yields a single

averaged model, used to produce raster layers of each algorithm for the following periods: contemporary, 2011-2040, 2041-2070, and 2071-2100, under RCP3-7.0 and RCP5-8.5 scenarios. Raster layers of each algorithm in each period of time underwent another averaging to generate ensembles, iterated five times for each sequence. These ensembles were subsequently merged with the other sequences of the same period, culminating in unified ensembles for both present and future timeframes. I repeated this process for the AdaptWest dataset (i.e., the secondary dataset).

Results

Evaluation (Contemporary and Future Tick Response to Climate)

Each ensemble is made up of raster cells. Each ~1 km grid square represents ~1 km of the area on the ground. These grid cells have a predictive value between zero and one, denoting niche suitability for *I. scapularis.* As varied predictions represent potential states of actual distributions, individual grid cells may exhibit good predictions while inaccurately predicting others (Marmion et al., 2009). Combining ensembles distinguishes a prevailing viewpoint or agreement among diverse model outputs. Each sequence resulted in a present-day raster, three SSP3-7.0 scenarios, and three SSP5-8.5 scenarios (2011-2040, 2041-2070, and 2071-2100). From evaluating the Chelsa-Bioclim+ data set, the following algorithms were used in each ensemble:

Table 3. Comparison of model statistics (with standard error indicated by ±) among the selected algorithms for each ensemble.

Algorithm	Count	AUC	TSS	Kappa
GAM	17	0.924 ± 0.007	0.727 ± 0.014	0.6755 ± 0.0235
GLMPoly	13	0.9175 ± 0.0095	0.720 + 0.18	0.657 ± 0.038
MARS	6	0.930 ± 0.014	0.7445 ± 0.0145	0.7005 ± 0.0375
MDA	10	0.970 ± 0.012	0.9025 ± 0.0185	0.850 ± 0.023
FDA	17	0.9295 ± 0.0525	0.8115 ± 0.0885	0.754 ± 0.099
RPART	9	0.954 ± 0.029	0.8285 ± 0.1015	0.791 ± 0.139
CART	17	0.974 ± 0.012	0.898 ± 0.026	0.898 ± 0.026
BRT	23	0.992 ± 0.002	0.9205 ± 0.0105	0.9205 ± 0.0105
RF	22	0.993 ± 0.003	0.9365 ± 0.0265	0.9365 ± 0.0265
SVM	13	0.9875 ± 0.0045	0.918 ± 0.018	0.918 ± 0.018
RBF	20	0.960 ± 0.017	0.8485 ± 0.0285	0.8485 ± 0.0286
MLP	16	0.985 ± 0.005	0.9135 ± 0.0225	0.9135 ± 0.0225

I conducted 4,704 model iterations across two datasets, employing up to 12 algorithms, investigated 10 climate profiles and considered 40 environmental variables. Under the two projected climate scenarios (SSP5-8.5 and SSP3-7.0), I encompassed a total of 2,094 future predictions of *I. scapularis* distribution. After evaluating all the models, the Chelsa-Bioclim+ ensembles (i.e., the primary dataset) consisted of 183 models employing 11 algorithms, while the AdaptWest (i.e., the complementary dataset) ensembles were produced using 166 models employing 12 algorithms. The final ensembled results demonstrated an average performance with AUC: 0.963, TSS: 0.859, and Kappa: 0.833. Complementary ensembles exhibited an averaged performance with AUC: 0.950, TSS: 0.828, and Kappa: 0.805. Models employing

GLMs exhibited high AUC values but were characterised by underperforming TSS and Kappa values, thus being excluded from the final Chelsa-Bioclim ensemble, with only two selected within the AdaptWest ensembles. A GLM introduced with polynomial regression exhibited significantly better performance, although slightly inferior to GAMs that were identified as the most robust in their amalgam.

MDA and FDA models demonstrated consistent performance within their grouping, with MARS models showing overfitting tendencies when used with the same number of pseudo-absences as GLMs, GAMs, and GLMPolys. However, MARS models exhibited improved performance and reduced overfitting using fewer pseudo-absences. Nonetheless, the efficacy of this claim was hindered by challenges in appropriately setting MARS pruning hyperparameters (i.e., nprunes and degrees), resulting in overgeneralisation and difficulties in interpreting variable responses and fluctuations.

BRTs and RFs consistently emerged as the top-performing algorithms across sequences, followed closely by SVM models. Radial Basis Function RBF models were noted for their inconsistency, failing periodically in production or displaying significant disparities in performance outcomes. MLP models were reliable, exhibiting high evaluation metrics, albeit slightly less robust than SVM, RF, and BRT algorithms. RPART and CART algorithms demonstrated enhanced performance when an equal number of pseudo-absence and presence points were utilised, compared to scenarios with a limited number of pseudo-absence points, with RPART showing marginal superiority. These findings are reflected in Table 3 and Appendices F, H, and I.

Range Expansion and Land Suitability Changes

To analyse the expansion rate specifically within eastern Canada, area calculations were performed by reclassifying values into predefined intervals following a breakdown similar to the one used by Yang et al., 2023 but with an additional classification for the highest suitability. These classifications are as follows: unsuitable niche (0-0.1), low-suitable niche (0.1- 0.3), medium-suitable niche (0.3-0.6), highly-suitable niche (0.6-0.8), optimal suitable niche (0.8-1).

Table 4. Changes in total land suitability per each timestamp. Total area was computed by converting rasters into polygons with the above classification values and summing their respective area values.

Time Period	Suitability	Total Area (km^2)	Relative Percentage (%)
Current (Historic)	Unsuitable	2,432,522 km^2	64.81 %
Current (Historic)	Low Suitability	613,459 km^2	16.34 %
Current (Historic)	Medium Suitability	259,804 km^2	6.92 %
Current (Historic)	High Suitability	126,529 km^2	3.37%
Current (Historic)	Optimal Suitability	321,003 km^2	8.56%
2011-2040 (SSP5-8.5)	Unsuitable	1,793,798 km^2	47.79%
2011-2040 (SSP5-8.5)	Low Suitability	705,938 km^2	18.81%
2011-2040 (SSP5-8.5)	Medium Suitability	486,881 km^2	12.97%
2011-2040 (SSP5-8.5)	High Suitability	286,599 km^2	7.64%
2011-2040 (SSP5-8.5)	Optimal Suitability	480,101 km^2	12.80%
2041-2070 (SSP5-8.5)	Unsuitable	1,321,509 km^2	35.20%
2041-2070 (SSP5-8.5)	Low Suitability	731,502 km^2	19.49%

2041-2070 (SSP5-8.5)	Medium Suitability	533,652 km^2	14.22%
2041-2070 (SSP5-8.5)	High Suitability	502,100 km^2	13.38%
2041-2070 (SSP5-8.5)	Optimal Suitability	664,554 km^2	17.71%
2071-2100 (SSP5-8.5)	Unsuitable	716,294 km^2	19.08%
2071-2100 (SSP5-8.5)	Low Suitability	851,324 km^2	22.68%
2071-2100 (SSP5-8.5)	Medium Suitability	628,939 km^2	16.76%
2071-2100 (SSP5-8.5)	High Suitability	864,089 km^2	23.02%
2071-2100 (SSP5-8.5)	Optimal Suitability	692,671 km^2	18.45%

Current I. scapularis Distribution

In my contemporary ensemble (Figure 2), the peak suitability reaches approximately 94%, with the most favourable conditions for *I. scapularis* observed in southern and southeastern Canada, notably in lower Ontario and Quebec (covering the majority of major urban centers). Suitability is sporadic in Nova Scotia, near Summerside and Kensington in Prince Edward Island (P.E.I.), and in regions of New Brunswick between Fredericton, Moncton, and Saint John. The suitability for I. scapularis diminishes beyond approximately 47°N, with Newfoundland and Labrador exhibiting low to medium suitability except in the Burin Peninsula. Much of Manitoba shows insufficient suitability, with most cell values below 30%, except for a southeastern region near Winnipeg, which is slightly more suited for *I. scapularis*. Of the combined ~3,753,317 km^2 of land measured within Eastern Canada (and Manitoba), only ~450,000 km^2, or ~12% of our extent, is categorised as having high or optimal suitability for *I. scapularis,* with the majority of land low to unsuitable (~75%, Table 4).

The historical and contemporary ensemble, utilising the AdaptWest environmental dataset (Figure A1), averages future environmental conditions from 13 climate normals. It showcases a nuanced reduction in predictive accuracy compared to the Chelsa-Bioclim+ ensembles, notably evident in lower Ontario and Quebec, where estimates remain largely unchanged. Newfoundland and Labrador exhibit similar trends, with the province devoid of high suitability. Nova Scotia and New Brunswick also demonstrate comparative estimations, with subtle variations in values, particularly in the lower and upper-east regions of Nova Scotia. In Prince Edward Island, AdaptWest ensembles present the island as optimal for *I. scapularis* habitation, displaying raster values exceeding 90% throughout.

The Chelsea-Bioclim+ ensemble predicts ~321,000 km^2 (8.56%) of land characterised by optimal coverage for *I. scapularis,* with an additional ~126,500 km^2 exhibiting high niche suitability (Table 4). The AdaptWest dataset presents ~317,000 km^2 as optimal and 92,500 km^2 as high (Tables G2 and G3).

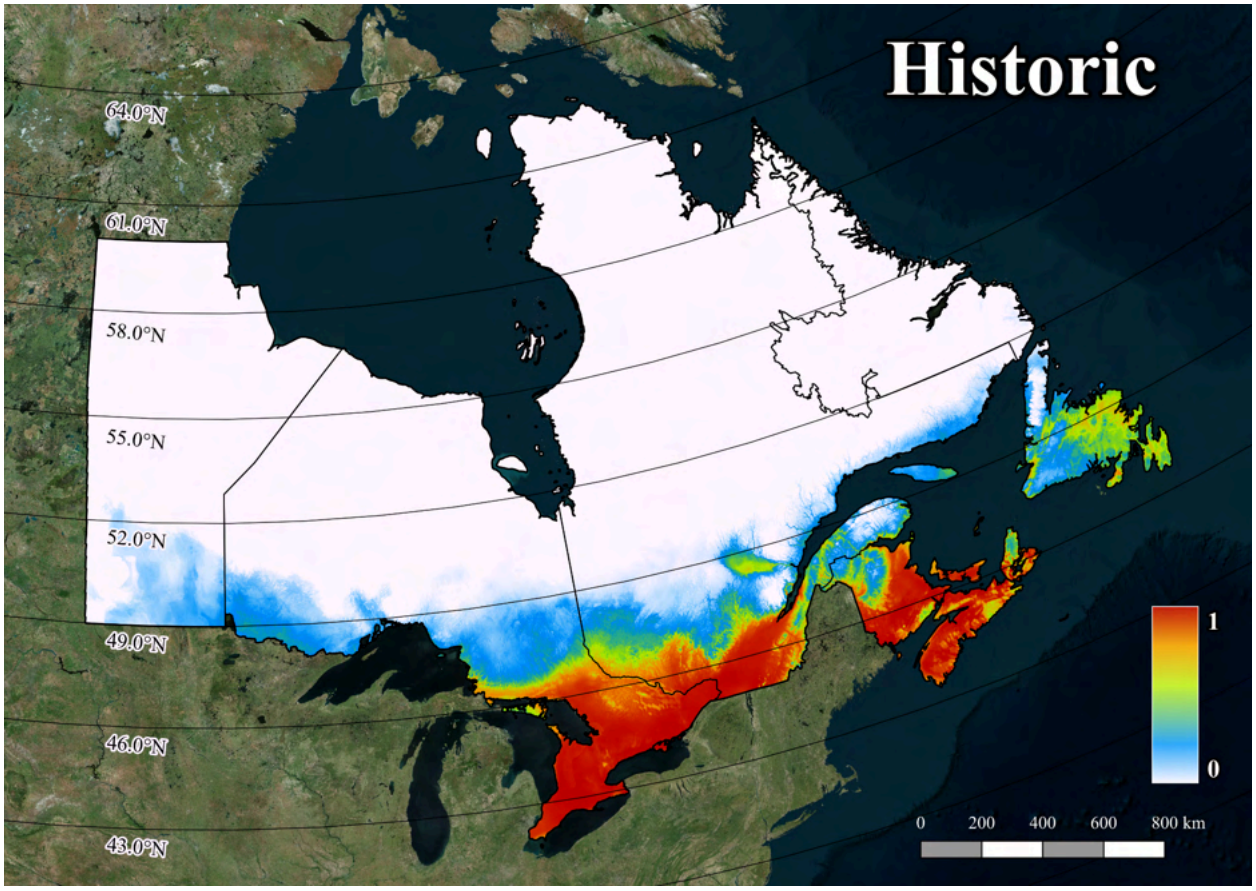


Figure 2. Ensemble produced using historic environmental data. Each ~1 km cell received a predictive value between 0-1. Zero indicates no suitability or likelihood of *I. scapularis*; one indicates the highest suitability or likelihood of occurrence. Additional classification breaks down *I. scapularis* suitability as follows: unsuitable niche (white: 0-0.1), low-suitable niche (blue: 0.1-0.3), medium suitable niche (green: 0.3-0.6), high-suitable niche (orange: 0.6-0.8), optimally suitable niche (red: 0.8-1.0).

Variable Permutation and Key Response

The main text illustrates the eight primary determinant factors within Chelsa-Bioclim+ ensembles. These determinants are identified through a permutation of climate variables. Variable significance is listed in Table 5, and the response curves of eight key variables are

illustrated in Figure 3. These factors are the primary drivers behind both the ensemble results discussed within the text and the mapped *I. scapularis* distribution showcased in the results section. The additional response curves for the Chelsa-Bioclim+ dataset are provided in Appendix D, while the response curves for the AdaptWests dataset are included in Appendix E. For details regarding the individual variable importance in the AdaptWest ensembles, refer to Appendix A, Table A1.

Variables of particular importance are net primary productivity (NPP), contributing ~10.29% to the overall ensemble results, mean annual temperature (Bio1), responsible for ~10.14%, and the growing season length (GSL), for ~8.44%. Additionally, the mean monthly precipitation of the driest quarter (Bio17) explains ~7.95% of the ensemble variation, while the mean daily air temperature of the warmest quarter (Bio10) accounts for ~7.72%. The environmental variable representing precipitation during the driest month (Bio14) elucidates 6.46% of the variance within the ensembles. Additionally, annual precipitation (Bio12) accounts for 5.61% of the total, and temperature seasonality (Bio4) explains 5.46%.

Net primary productivity (NPP) response exhibits a distinct pattern characterised by an asymmetrical, unimodal distribution. The peak of this distribution occurs at approximately 1250 $gCm^{-2}yr^{-1}$, with a confidence level of ~75%. Notably, to the left of the inflection point, the response demonstrates a steep slope, starting from 0% probability at approximately 2500 $gCm^{-2}yr^{-1}$, indicating a heightened sensitivity to changes in NPP compared to the right side of the curve, where the decline is more gradual (refer to Figure 3). The modelled influence of mean annual temperature on the distribution of *I. scapularis* suggests an optimal thermal range between 5°C and 15°C, coinciding with a high probability of ~80%. Beyond this range, the probability diminishes, reaching near 0% suitability at temperatures below -10°C and above

25°C (Figure 3). Regions with a short growing season length (GSL) indicate a low probability of *I. scapularis* suitability. As the GSL increases, so does the probability, which exhibits a sharp increase in responsiveness for a GSL of 100 to 200 days. The optimal GSL is 250 days, beyond which the responsiveness tapers off (Figure 3). The mean monthly precipitation of the driest quarter (Bio17) shows a notable increase in probability from 0 to 150mm month^{-1} before plateauing between 200 and 350mm month^{-1} at ~90% probability, followed by a gradual decline in probability with further increases in precipitation. The mean air temperature of the warmest quarter (Bio10) exhibited similarities with the response pattern observed in Bio5 (i.e., the mean air temperature of the warmest month). However, Bio10 displayed more pronounced probabilities and a heightened sensitivity to temperature variations. For instance, Bio10 demonstrated a low probability below 10°C, followed by a steep gradient from 10°C to 20°C, resulting in a notable increase in probability percentage from ~15% to ~90%. The peak probability occurred between 20°C and 23°C, followed by a substantial decrease again (Figure 3). The precipitation observed during the driest month (Bio14), which accounts for 6.46% of the ensemble, displays a bell-shaped curve with a nearly symmetrical distribution. This pattern reveals optimal precipitation levels ranging between 50 mm and 150 mm at ~90% to ~95% probability. as illustrated in Figure 3. Conversely, the annual precipitation (Bio12), featuring a variable importance of 5.61% according to Table 5, exhibits a steep response gradient as it ascends from 0 to 1000mm. Notably, suitability dramatically increases from ~0% to 80%. Bio12's response curve gradually peaks and levels at ~90% between 1250mm and 2500mm of precipitation (Figure 3). Temperature seasonality (Bio4) explains 5.46% of the ensemble. Bio4 expounds on the degree of variation or fluctuations across different monthly temperatures. The

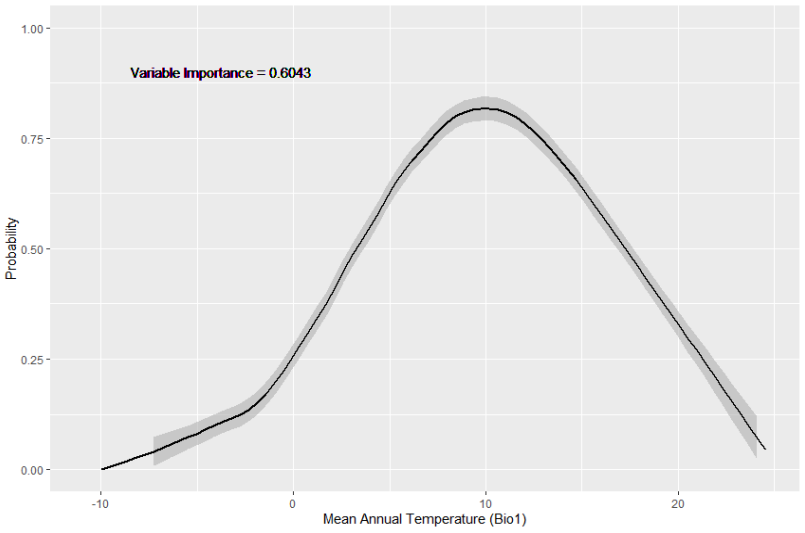
response of Bio4 is characterised by a unimodal curve, demonstrating a narrow range between its peak and trough probabilities, ~20% to ~40%, respectively (Figure 3).

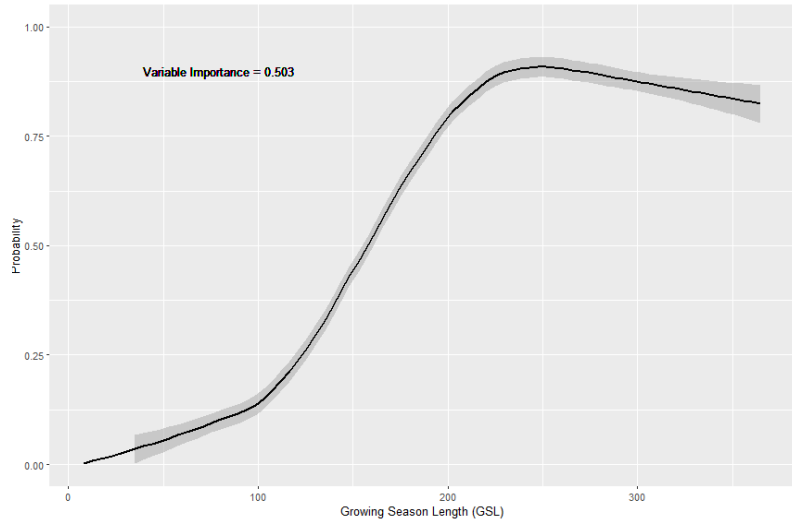
Table 5. Environmental variable permutations in Chelsa-Bioclim+ ensembles

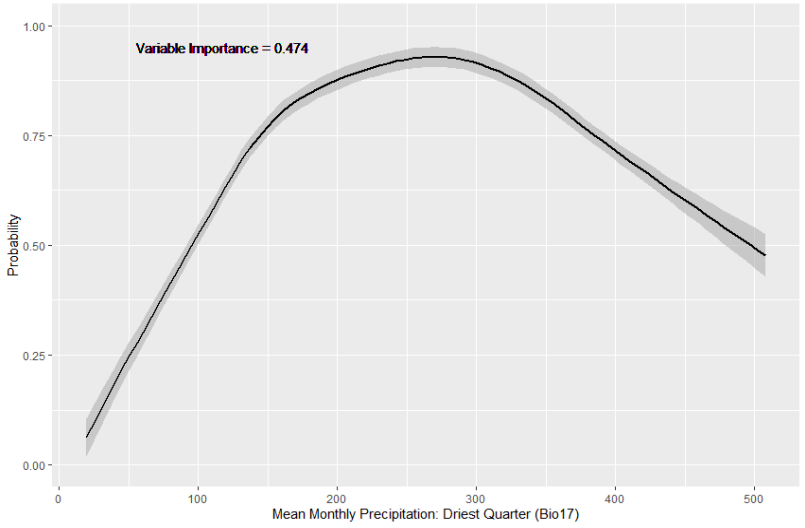
	Chelsa-Bioclim+ Variables		
Variable	Abbrev	Relative Sequence Variable Importance	Ensembled Variable Importance
Net Primary Productivity	NPP	61.33%	10.29%
Mean Annual Temperature	Bio1	60.43%	10.14%
Growing Season Length	GSL	50.30%	8.44%
Mean Monthly Precipitation - Driest Quarter	Bio17	47.40%	7.95%
Mean Daily Air Temperature - Warmest Quarter	Bio10	46%	7.72%
Precipitation - Driest Month	Bio14	38.52%	6.46%
Annual Precipitation	Bio12	33.46%	5.61%
Temperature Seasonality	Bio4	32.52%	5.46%
Last (Growing Degree) Day above 10°C	GDDLGD10	30.87%	5.18%
Maximum Daily Mean Air Temperature - Warmest Month	Bio5	28.08%	4.71%

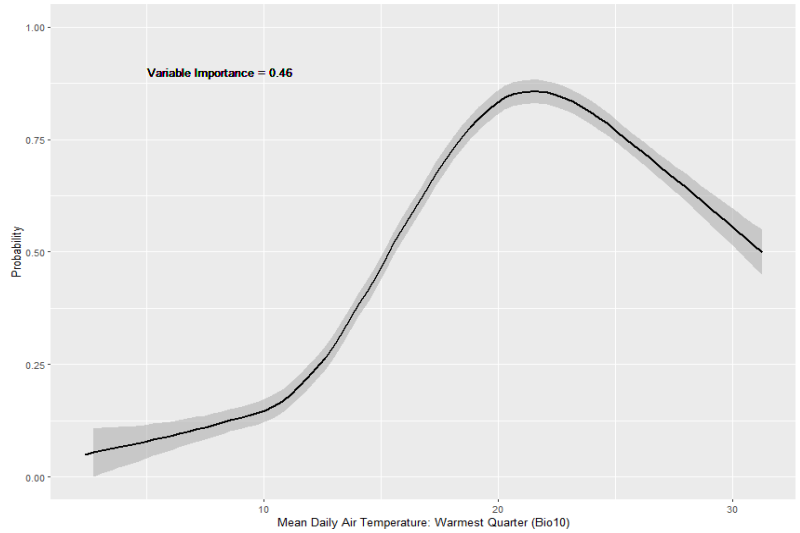
Growing Season Precipitation	GSP	25.36%	4.25%
Last (Growing Degree) Day above 0°C	GDDLGD0	20.85%	3.50%
Last (Growing Degree) Day above 5°C	GDDLGD5	18.80%	3.15%
Precipitation Seasonality	Bio15	18.55%	3.11%
Mean Diurnal Air Temperature Range	Bio2	13.71%	2.30%
Isothermality	Bio3	13.58%	2.28%
Snow-Water Equivalent	SWE	10.77%	1.81%
Frost-Cover-Frequency	FCF	12.46%	2.09%
Growing Season Temperatures	GST	10.82%	1.82%
Annual Air Temperature Range	Bio7	9.56%	1.60%
Mean Monthly Precipitation - Wettest Quarter	Bio16	5.53%	0.93%
Mean Daily Air Temperature - Wettest Quarter	Bio8	4.96%	0.83%
Precipitation - Wettest Month	Bio13	2.16%	0.36%
	Total:	596.02	100.00%

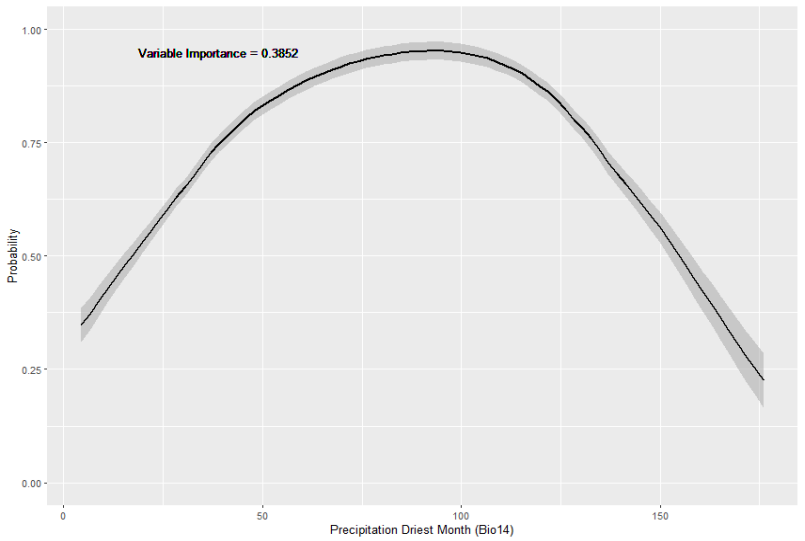

Variable Importance = 0.6133
Probability
1.00
0.75
0.50
0.25
0.00
500
1000
1500
2000
Net Primary Productivity (NPP)


Variable Importance = 0.6043
Probability
1.00
0.75
0.50
0.25
0.00
-10
0
10
20
Mean Annual Temperature (Bio1)


Variable Importance = 0.503
Probability
1.00
0.75
0.50
0.25
0.00
0
100
200
300
Growing Season Length (GSL)


Variable Importance = 0.474
Probability
1.00
0.75
0.50
0.25
0.00
0
100
200
300
400
500
Mean Monthly Precipitation: Driest Quarter (Bio17)


Variable Importance = 0.46
Probability
1.00
0.75
0.50
0.25
0.00
10
20
30
Mean Daily Air Temperature: Warmest Quarter (Bio10)


Variable Importance = 0.3852
Probability
1.00
0.75
0.50
0.25
0.00
0
50
100
150
Precipitation Driest Month (Bio14)

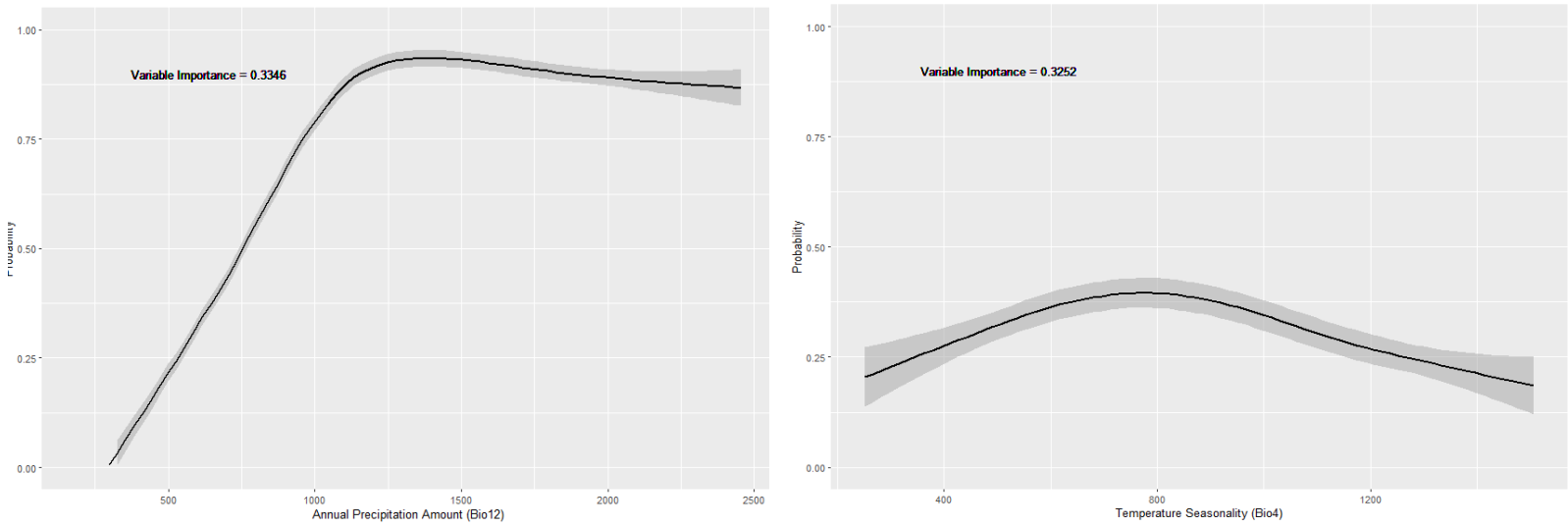


Figure 3. The response curves corresponding to the eight most influential variables in the Chelsa-Bioclim+ ensembles

Future I. scapularis Distribution

By 2040, the niche areas (i.e., above 80% *I. scapularis* suitability) will experience notable expansion from contemporary locales, shifting northeasterly, growing by ~160,000 km^2 and reaching an upper northern threshold of ~48°N (Figure 4). The only notable shift from 90% suitability zones is from near Lake Erie (e.g., around Hamilton, London, and Toronto), where the compatibility decreases by 10%. All low, medium, and high suitability areas have expanding zones, with low suitability areas reaching ~52°N, medium zones reaching ~50°N, and high reaching ~49°N. Manitoba still reaches low to medium environmental compatibility, though the clustering of medium-valued raster cells does increase. Newfoundland will experience a significant and rapid expansion in niche suitability for *I. scapularis.* This expansion is marked by

continued growth in the Burin area, along with the emergence of niche zones in various regions: the Avalon Peninsula (e.g., Conception Bay South, Holyrood, and Butter Pot Provincial Park), the northern coast of central Newfoundland (e.g., near Terra Nova National Park, Lewisport, and Notre Dame Provincial Park) and Western Newfoundland (e.g., Corner Brook and Deer Lake). The only area that reaches 90% suitability in Newfoundland is close to the west coast near Stephenville and Stephenville Crossing. Prince Edward Island is characterised by consistently high suitability throughout, with only the eastern point in medium suitability zones. *I. scapularis* optimal range in Ontario, which historically reaches as far north as Sudbury, expands an additional ~100 km north in central and northeastern Ontario (as far as Lady Evelyn Smoothwater Provincial Park and Temiskaming Shores); western Ontario expands across Lake Huron and near the entire length, and area, of Lake Superior Provincial Park. The optimal limits for *I. scapularis* within western and central Quebec reach the top of La Vérendrye Wildlife Reserve. Heading eastward from Saint Georges along the St. Lawrence River to the Quebec/Maine border and onward to New Brunswick, the niche suitability for the blacklegged tick now largely exceeds 80%. The further northern bank of the St. Lawrence River also provides optimal conditions for *I. scapularis* (from Quebec City to Baie-Comeau).

By 2070, the northern range of *I. scapularis* niche extends to ~49°N (high suitability reaches ~51°N, medium suitability reaches ~52°N, low suitability reaches ~53°N, Figure 5). Discernable changes of improved suitability occur on the northern bank of the St. Lawrence River, situated near a rare cold spot proximate to Parc national de la Jacques-Carter, Quebec. The total area of optimal suitability (over 80% suitability) is increasing from ~480,000 km^2 (12.8% of total cover) to ~665,000 km^2 (or ~17.70%), whereas the percentage of raster cells above ninety marginally decreases (Table 5). Most previous composites of near-valued cells are expanding in

size; central Newfoundland hot spots continue to extend, with a few additional locations reaching 90% suitability (e.g., near Terra Nova National Park, Port Blandford).

Near the turn of the century, the northern threshold and the optimal niche space percentage will increase to ~50°N and ~18.50%, respectively. High suitability reaches upwards of ~52°N, medium suitability and low suitability reach ~53°N and ~55°N (Figure 6). While the percentage of land predicted to be over 80% has increased, the percentage of land over 90% has decreased significantly. Reduced conditions to medium values occur in lower Ontario, lower Nova Scotia, and parts of Lower Quebec. Labrador now has land that reaches high suitability, with one location reaching optimal suitability, near Happy Valley-Goose Bay.

Between 2011 and 2040, the supplementary ensembles (Adaptwest, Appendix A), in relation to the principal ensembles (Chelsa-Bioclim+, Figures 2, 4, 5, and 6), reached higher northern latitudes (~52°N) at lower suitabilities; comparable at medium suitability (i.e., reaching ~49-50°N), but extend less in high and optimal *I. scapularis* suitability zones (~48°N, Figure A2). Much of the ninety-percentile rasters have reduced certainty but are still within the margins of optimal suitability (e.g., Prince Edward Island is still compromised wholly of raster cells within the optimal suitability categorisation). Most optimal suitability conditions are similar to those in the aforementioned rasters (e.g., situated within lower Ontario and Quebec, eastern New Brunswick and central Nova Scotia). Only one large conglomeration of raster cells consists of values over 90%, near Windsor, Nova Scotia, covering most of the division of Hants County. The largest Newfoundland subsection with optimal conditions for *I. scapularis* is near Botwood and Bishop's Falls, following the Bay d'Espoir highway (route 360) for ~40 km. High suitability is found by Central and near the western hubs of Corner Brook, Pasadena, and Stephenville.

By 2070 (Figure A3), notable hotspots shift north once more. Low suitability reaches ~44-45°N, medium and high reach ~51°N and ~50°N. The optimal niche reaches ~48-49°N; All Nova Scotia and New Brunswick are considered high or optimal suitability, and over half the island of Newfoundland. The highly suitable sites in Newfoundland by 2040 continue to hold the highest suitability, escalating in values to optimal *I. scapularis* environmental conditions. Near Happy Valley-Goose Bay in Labrador, medium suitability is reached, the first within the region (a notable feature predicted within the primary dataset). The one location reaching 90% predictability is to the left and right of Quebec City. Near this location, the cold spot identified within the primary dataset (i.e., near Parc National de la Jacques-Carter, Quebec) is additionally within this collection of models.

By the end of the century (Figure A4), nearly all of Canada contains at least low suitability (i.e., between 10-30% niche suitability, reaching 58N-59°N). Medium suitability reaches ~55-56°N (52N in Ontario), and high suitability (higher in Quebec and Labrador than in Ontario) reaches ~53-54°N. The optimal suitability locations have dropped significantly, with Nova Scotia largely (and P.E.I completely) consisting of high suitability cover. The northwest portion of New Brunswick now consists of the optimal conditions for *I. scapularis*, joined at the Quebec-New Brunswick regional border. Quebec contains the highest concentration of optimal conditions, still to the east of Quebec City but also creeping along the northern coast of the St. Lawrence River and sporadically up the Gulf of St. Lawrence until reaching the Quebec-Labrador coastal border. Lake Superior Provincial Park in Ontario is the one provincial location with quintessential *I. scapularis* habitat conditions. Newfoundland exhibits a slight northwesterly shift, with the optimal location moving from Pasadena, Corner Brook, and Stephenville to East Arm Bay and the coastline of Gros Morne National Park. The western

interior and up the Baie Verte Peninsula also contain optimal conditions, while the western and eastern coast contain only medium suitability conditions. In Labrador, the one optimal location (coincidently identified as significant across late-century ensembles) is near Happy Valley-Goose Bay, in which high suitability conditions expand, reaching the Quebec border and the Gulf of St. Lawrence coastline.

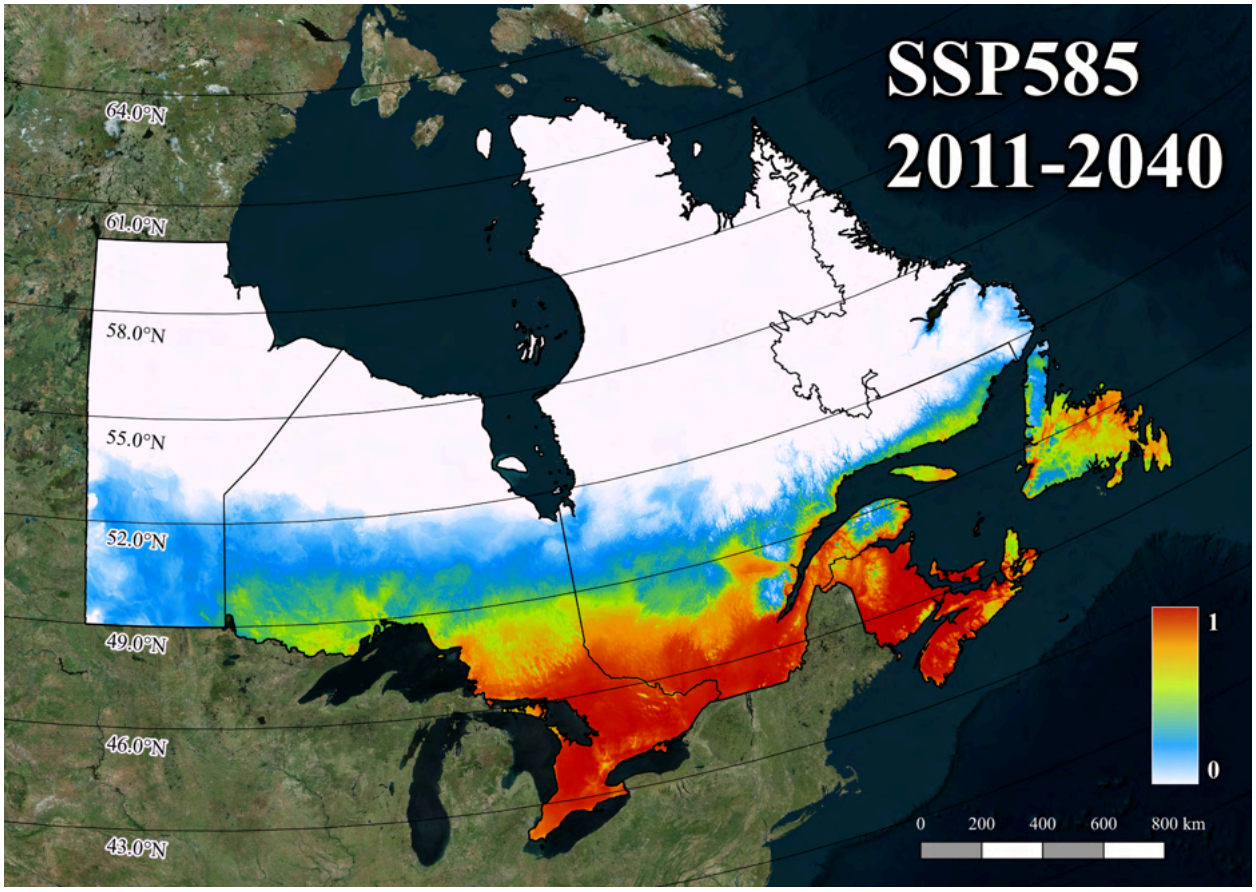


Figure 4. Ensemble produced for the time-period 2011-2040 using SSP5-8.5 environmental data. Each ~1 km cell received a predictive value between 0-1. Zero indicates no suitability or likelihood of *I. scapularis*; one indicates the highest suitability or likelihood of occurrence. Additional classification breaks down *I. scapularis* suitability as follows: unsuitable niche (white: 0-0.1),

low-suitable niche (blue: 0.1-0.3), medium suitable niche (green: 0.3-0.6), high-suitable niche (orange: 0.6-0.8), optimally suitable niche (red: 0.8-1.0).

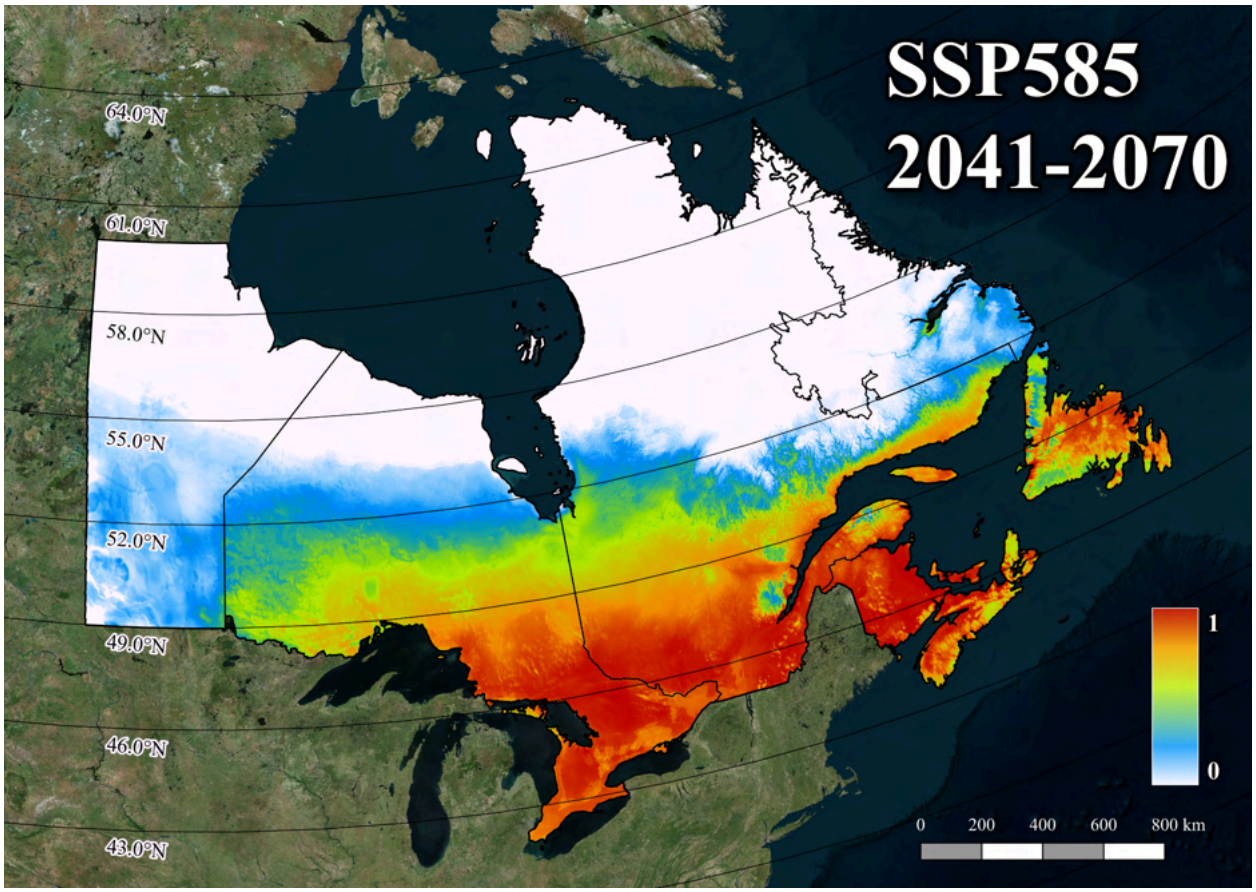


Figure 5. Ensemble produced for the time-period 2041-2070 using SSP5-8.5 environmental data. Each ~1 km cell received a predictive value between 0 - 1. Zero indicates no suitability or likelihood of *I. scapularis*; one indicates the highest suitability or likelihood of occurrence. Additional classification breaks down *I. scapularis* suitability as follows: unsuitable niche (white: 0 - 0.1), low-suitable niche (blue: 0.1 - 0.3), medium suitable niche (green: 0.3 - 0.6), high-suitable niche (orange: 0.6 - 0.8), optimally suitable niche (red: 0.8 - 1.0).

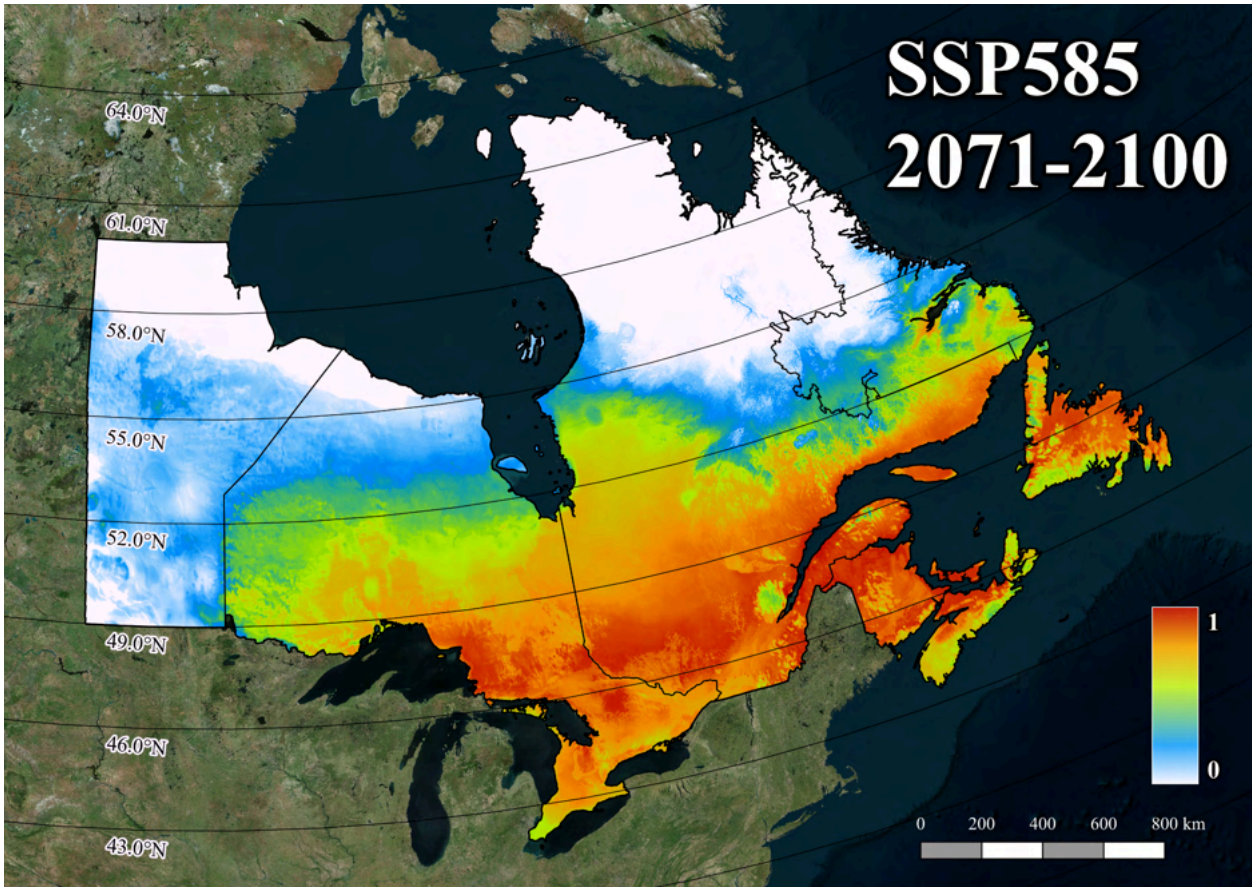


Figure 6. Ensemble produced for the time-period 2071-2100 using SSP5-8.5 environmental data. Each ~1 km cell received a predictive value between 0-1. Zero indicates no suitability or likelihood of *I. scapularis*; one indicates the highest suitability or likelihood of occurrence. Additional classification breaks down *I. scapularis* suitability as follows: unsuitable niche (white: 0-0.1), low-suitable niche (blue: 0.1-0.3), medium suitable niche (green: 0.3-0.6), high-suitable niche (orange: 0.6-0.8), optimally suitable niche (red: 0.8-1.0).

Future I. scapularis Distribution Under SSP3-7.0 Scenarios

The future predictive ensembled using SSP3-7.0 scenario environmental data (Appendix B) depicts a similar projection of *I. scapularis* range expansion from 2011-2040 to that of the SSP5-8.5 scenario (Figure B1 compared to Figure 4). The northern ranges match, aside from latitudinal growth in low suitability in Labrador; minor discrepancies exist in the expansion of optimal suitability habitats (e.g., hot spots are slightly larger in Ontario, Quebec, Nova Scotia, P. E. I, and New Brunswick using SSP5-8.5 values - slightly smaller in Newfoundland). By 2070 (Figure B2), SSP3-7.0 future values provide a more conservative estimate of range expansion, with the leading edge of the suitability envelope not extending as far north with reduced suitability, particularly in the growth of niche zones (e.g., central Newfoundland, lower and central Ontario, lower Nova Scotia, and central Quebec). Within this division of time, there were more apparent changes than within the previous epoch (i.e., 2011-2040, Figure B1). Once again, by 2100 (Figure B3), the northern range has extended, with lower estimates in the SSP3-7.0 scenario compared to the SSP5-8.5 projections (Figure B3 set against Figure 6). However, the lower concentrations scenario seems to have *I. scapularis* favouring the lower, southern edges of Ontario, Quebec, New Brunswick and Nova Scotia, which, in addition, these provinces (and P.E.I) also have a larger concentration of 90% predictive cells then the SSP5-8.5 ensemble.

The separation between the SSP5-8.5 and SSP3-7.0 AdaptWest datasets (Appendix A and Appendix C) is, again, inconsequential up to 2040, with differences in the northern limits of low and medium suitabilities - SSP5-8.5 reaching farther. From 2041-2070, the ensembles continue the trend of slight separation with SSP5-8.5 advancing further north, but the locations of optimal *I. scapularis* niche differ slightly. Both datasets have ~372,000 km^2 (or 9.90% of the extent; Tables G2 and G3) as optimal. However, SSP5-8.5 has a larger concentration in Newfoundland

and New Brunswick and slightly less within the other provinces. The most significant differences exist within the AdaptWest 2071-2100 rasters, where a stark contrast exists between optimal zones (SSP5-8.5s 7.14% to SSP3-7.0s 2.98%; Tables G2 and G3); the hot spot distributions are similar, with the exception of a few additional localities within the SS5-8.5 scenario (e.g., Lake Superior National Park, Algonquin National Park, Happy-Valley Goose Bay). However, the corresponding locations have an expanded magnitude, containing P.E.I and northeastern Nova Scotia as well).

Discussion

Climate change is occurring at an alarming rate; as such, the blacklegged tick (*I. scapularis*) is responding rapidly. I have shown here that temperature, precipitation, net primary production (i.e., the quantity of biomass/carbon production), length of the growing season, climate moisture index (i.e., moisture balance), and the number of degree days in a year explain well the current distribution of ticks in eastern North America and that climate change will continue to strongly influence the distribution of *I. scapularis* within the 21st century. As the prevalent environmental conditions change over time, so will the likelihood of *I. scapularis* occurring, with a 583% increase in highly suitable land and a 116% increase in optimum suitable land at higher latitudes by the end of the century (Figure 6 and Table 4). The rising suitability coincides with the projected expansion of *I. scapularis'* optimal niche, which reaches northern latitudes of ~48°N by 2040, ~49°N by 2070, and ~52°N by 2100 (see Figures 2, 4, 5 and 6). The forecasted ensembles show good discrimination between categorised zones, which is corroborated by the expressed range expansion, range direction, hotspots, and many of the same

environmental factors within the complementary dataset (AdaptWest; Appendix A, Appendix C, Appendix E and Appendix G), notwithstanding, some more conservative estimates in suitability.

Environmental Influences on Distribution

Temperature and precipitation conditions, either alone or acting in tandem, considerably influence the prevalence of *I. scapularis* (Schulze et al., 2009). My models reveal that 1,250mm to 2,000mm of annual precipitation is ideal, with less precipitation during the year progressively suboptimal for survivability (Figure 3, Appendix E). Within each morphological phase, *I. scapularis* must find a balance between wet and dry precipitation conditions, particularly during the winter (ideal amounts of 200 to 400mm, Appendix E) and spring (ideally 200 to 300mm, Appendix E). These seasons have a stronger association with tick presence than the fall and summer. This is potentially due to mild, wet winters increasing overwintering survivability (Hayes et al., 2015), while heavier precipitation in late spring (and early summer) can coincide with increased nymphal questing and longevity, potentially halting larval diapause and aiding in the emergence of new ticks (McCabe & Bunnell, 2004). The models suggest snow accumulation could also be a factor during the colder months, acting, along with leaf litter, as an insulator to the cold (Burtis & Pflueger., 2017; Templer et al., 2012), with greater survival rates under the accretion of leaf litter and snow cover (Linske et al., 2019). If climate change generates a reduced or delayed winter snowpack, there could also be increased mortality amongst arthropod species (Templer et al., 2012). Annual snowfall amounts of 250mm are ideal for *I. scapularis* niche (Appendix E), with even minor increments of snowfall preferable to none. Additionally, increased rainfall can induce higher relative humidity, stabilise fluctuating temperatures, and increase moisture content in the duff layer of forests (McCabe & Bunnell, 2004). The

contribution of rainfall is further supported by the increased significance of monthly precipitation during the driest month and driest quarter (relative to wetter attributes) where precipitation is scarce, making precipitation events critical (Figure 3, Table 5, and Appendix D). However, ever-increasing rainfall leads to more frequent and extreme rainfall events, which could hinder *I. scapularis* survivability and expansion (Ogden et al., 2006a).

Temperature is often the most important variable determining the distribution of *I. scapularis* (Clark, 1995), with most temperature-based variables (e.g., annual mean temperature, degree-days above a threshold temperature, length of the growing season, temperature seasonality, last day of the year above a threshold temperature) strongly correlated with *I. scapularis* (Table 5; Appendix D and Appendix E). In my models (both datasets), the suitable annual mean temperature range is between ~6.0 to ~12.0 °C ± 1.0°C (Figure 3 and Appendix E), coinciding well with a documented uncoordinated activity threshold of 9.2 ± 4.1 and 11.2 ±3.4 for females and males and an activity threshold of 6.2 ± 3.6 for females, 8.5 ± 3.0 for males, respectively (Clarke, 1995). A temperature at or below -10°C exhibits a 0% probability of tick occurrence(Figure 3 and Appendix E). This finding is supported by empirical evidence indicating that *I. scapularis* ticks can only survive at temperatures of -10.0°C for brief durations (i.e., a few hours; Lindsay et al., 1998), while temperatures ranging from -11.0°C to -16.0°C are lethal (Brunner et al., 2012).

The models additionally suggest that locations characterised by a higher frequency of degree days below freezing (Appendix E) are rendered unfavourable for *I. scapularis*, thereby affirming the predictive capabilities of the models. Moreover, the duration of optimal temperatures plays a crucial role, as warmer climates are associated with an accelerated life cycle and reproductive rate for the tick population (Eisen et al., 2016b). The ensembles also indicate

that an extended growing season, particularly upward of 200 days, leads to peak suitability (Figure 3). *I. scapularis* can complete its life cycle with optimal conditions in less than 230 days (Eisen et al., 2016b). As the duration of the growing season is anticipated to become longer throughout the 21st century (NRCan, 2022), *I. scapularis* may experience accelerated activity and development. This phenomenon is particularly notable in relation to degree days, wherein growing degree days (defined as degree days above 5 °C, Appendix E) exhibit a greater relative significance (and a progressed response to changing conditions) compared to other degree day variables measured (including degree days above 10 °C and less than 40 °C, as well as degree days above 18 °C; Appendix E). One possibility for this is because the growing season, and likewise growing degree days, occur during the spring and summer, a period that coincides with resumed activity (i.e., from May to June) and development and engagement (i.e., from May to September) in *I. scapularis* (Lindsay et al., 1998).

Net primary production (NPP) is a measure of annual plant biomass production (Green & Byrne, 2004). Associated vegetation types with lower NPP are deserts, tundra, and some shrublands, with the highest NPP values within tropical forests (Melillo et al., 1993). Our models suggest peak suitability at ~ 1,250g C m^{-2} yr^{-1} to ~1,400g C m^{-2} yr^{-1}, which is around the middle range of values (Table 5). A greater NPP exhibits increased suitability and is less responsive to changes in conditions relative to lower NPP values. Contemporarily, temperate forests and tall grasslands could be considered the intermediate zone in terms of NPP (and the most suitable for *I. scapularis*), with temperate and boreal forests expected to experience increased NPP values in response to climate change (Melillo et al., 1993). Hogg's climate moisture index (CMI) reports the difference between precipitation and potential evapotranspiration (i.e., water vapour loss from vegetated land cover; NRCan, 2020) and is independent of forest type and vegetation

indices (NRCan, 2020; Wang et al., 2014). A CMI value above zero indicates wet and moist conditions (e.g., can nourish a closed-canopy forest; NRCan, 2020); a negative CMI value suggests drier conditions. Most of the eastern Boreal ecozone has a CMI value exceeding 20 over the past half-century or longer (Wang et al., 2014). My analysis suggests *I. scapularis* favour locations with positive CMI values, particularly above 50 CMI (Appendix E). Anything below 50 CMI created a threshold reaction to increased sensitivity to change, and any negative CMI values could conceivably impact *I. scapularis's* ability to persist or survive due to water stress (Berger et al., 2014; Rodgers et al., 2007). Over the next century, locations in eastern Canada that experience high precipitation levels will likely continue to have a high CMI, while other locations in eastern Canada may become marginally wetter, furthering a potential increase in suitable land cover (Wang et al., 2014).

SSP5-8.5 and SSP3-7.0 Climate Scenarios

SSPs tell of differences in the future state of affairs with the changes in climate mitigation/adaptation challenges, climate change (O'Neill et al., 2016), or socioeconomic trends (Riahi et al., 2017). They aid in our understanding of potential futures and provide a reference point for researchers to integrate results across studies (Riahi et al., 2017). The numbers following the SSP initialism (e.g., the 2-4.5 in SSP2-4.5) stand for stratospheric-adjusted radiative forcing, ranging from 2.6 to 8.5 W m^{-2} (Tebaldi et al., 2021). There are four (sometimes five, Meinshausen et al., 2020) high-priority scenarios: SSP1-2.6, SSP2-4.5, SSP4-6.0, and SSP5-8.5 (O'Neill et al., 2016; Tebaldi et al., 2021). SSP1-2.6 is considered the sustainable path, with interests in inclusion, equality, economic growth, less consumption, and low difficulties in adaptation and mitigation (Riahi et al., 2017). SSP2-4.5 simulations are middle-of-the-road,

where trends (e.g., social, economic, political) do not deviate far from historical patterns. SSP3-7.0 is a medium-high scenario that incorporates "regional rivalry" dynamics (Meinshausen et al., 2020). Economic and population growth is slow, material consumption is high, resources and opportunities are unequal, there is a low environmental priority, and adaptation and mitigation potential are considered highly challenging (Riahi et al., 2017). SSP5-8.5 is within the upper echelon of predictive forcing levels, relying on high fossil fuel usage. Markets remain competitive with increased globalisation; the push for socioeconomic and economic growth leads to a resource-intensive and energy-intensive society. Mitigation is a challenge, but adaptation practices are sometimes successful (e.g., accomplished use of managing local ecological systems; Riahi et al., 2017). SSP5-8.5 scenarios correlate with a wide range of CO_2 concentrations that are higher than the previous RCP8.5 values (Meinshausen et al., 2020).

A significant amount of literature contentiously refers to RCP8.5's climate predictions as the "business as usual" scenario, which suggests that it represents likely future outcomes if climate mitigation strategies are not employed (Hausfather & Peters, 2020). I prioritised the high-emission SSP5-8.5 scenario because, from a risk management standpoint, overpredicting *I. scapularis'* range is preferable to underpredicting risk (Soucy et al., 2018). The confidence level associated with future scenarios remains subject to debate as they integrate fundamental assumptions (Nazarenko et al., 2022). For instance, projections assume greenhouse gas emissions will continue along the same accelerating, anthropogenically-driven growth trajectory observed in recent decades. Additionally, there are indications that many nations are unlikely to fulfill their commitments outlined in political agreements, such as their pledged contributions to the Paris Agreement (Liu & Raftery, 2021). Moreover, there is uncertainty in the forecasting of economic growth (Christensen et al., 2018), feedback loops (Kemp et al., 2022), compound

hazards/events (Mathews et al., 2019), and tipping cascades (Klose et al., 2021). This sets a dangerous precedent to underestimate the magnitude of climate change and the consequential risk *I. scapularis* brings. While contentious, some reports suggest that projected global warming in CMIP6 is lower than the observed, placing the observed values closer to the upper SSP5-8.5 emission scenario (Carvalho et al., 2022). Additionally, SSPs do not capture uncertainty in long-run productivity growth rates. This is problematic because the high-emission scenarios (e.g., SSP5-8.5 and RCP8.5) often negate this factor, rejecting a subset of climate forcing values that indicate (through a series of economic projections) that emission concentration will surpass a radiative forcing of 8.5 W/m^2 used in upper-tier baseline emission scenarios (Christensen et al., 2018). The probability of emission values reaching greater severity than those used in RCP8.5 is greater than 35% (Christensen et al., 2018), warranting consideration for prioritising extreme outcomes (i.e., SSP5-8.5).

Nevertheless, I simulated SSP3-7.0 scenarios to provide an alternative, high-priority experiment founded on different baseline assumptions. Still considered a high-emission scenario (IPCC, 2021b), my results show scarce differences between the two primary SSP scenarios up to 2040 (Figure 4 and Figure B1). From thereon to 2100, there are no new focal points, but instead, dynamic expansions from current foci (both northward and omnidirectional), with greater expansion apparent with SSP5-8.5 ensembles over SSP3-7.0 (Figures 5 and 6; Figures B2 and B3.

Implications

Climate variables, which follow a more uniform pattern over larger extents, are critical for range expansion and colonising new areas (Slatculescu et al., 2020). In contrast, ecological

variables, which are localised and may experience high variability in conditions, are more impactful in ensuring that a population can establish and maintain itself (Slatculescu et al., 2020). While I have shown the extent to which suitable *I. scapularis* habitat will change through climate, so too will the rate of species invasions (Leighton et al., 2012). The species populations will likely expand from the current Canadian foci (i.e., southeastern Canada), with a high density of suitable hosts and rodent reservoir species of zoonotic, tick-borne pathogen (Ogden et al., 2006b). My models suggest that the northern spread into Canada will veer easterly instead of towards the Canadian prairies, likely because they are excessively dry and arid, a condition that will persist with climate change, even if temperature conditions become more favourable for *I. scapularis* (Ogden et al., 2006b). The mass migration of birds each spring into Canada aids this expansion, passing through north-central and northeastern states where there is a high prevalence of *I. scapularis* and Lyme borreliosis (Ogden et al., 2006b). The long-range dispersal capabilities of birds are an efficient way for ticks to expand their range without dealing with restrictive dispersal constraints (e.g., ecological connectivity, land cover) and could explain any suitable habitats that may appear stochastic. The southern (or trailing) edge of the range will also consolidate in lower Ontario, Quebec, New Brunswick and Nova Scotia; although, after the 21st century, there are suggestions that the ideal conditions for *I. scapularis* will advance again.

In southeastern Manitoba, established populations of *I. scapularis* have been documented (e.g., on the Lake of the Woods; Krakowetz et al., 2011). Despite their presence, these habitats were not shown as suitable within my contemporary models, and the projections indicate their suitability in shifting away from Manitoba in all foreseeable scenarios. The persistence of *I. scapularis* in this region could be attributed to the presence of woodland habitats, an aspect not considered in the model due to the omission of land cover variables (Ogden et al., 2008c). Furthermore, the intricate nature of microclimates, pivotal for tick survival, could explain why ticks could persist in Manitoba despite low climate suitability (Ogden et al., 2008c). Conversely, areas flagged as ideal habitats for *I. scapularis* might be limited due to barriers or obstacles hindering their colonization. Although my models indicate that significant parts of southwestern Ontario provide optimal conditions for the tick's persistence, certain pockets within this region, particularly those in or near major urban hubs such as Toronto, Ottawa, and Quebec, suffer from extensive urbanization or agricultural activities. Despite favourable climatic conditions, these locales may not offer suitable habitats due to human development.

However, in general, accurate modelling of *I. scapularis* range and subsequent expansion makes it feasible to infer patterns within the landscape, enabling the implementation of proactive preventive measures. This allows authorities, such as public health, to efficiently allocate resources and raise awareness of affected regions to safeguard the population. Additionally, advanced forewarning before *I. scapularis* achieves permanency in a region provides time for authorities to enact wildlife management strategies for targeted intervention strategies (e.g., through acaricides or habitat modification) to mitigate the impacts on sensitive species and ecosystems. Decision-makers must prioritize explicit articulation of the public costs, the intended scale of intervention implementation, and the documented impacts on ecosystems when

presenting interventions to target populations. Such clarity is crucial as these aspects can influence the acceptability of the intervention (Aenishaenslin et al., 2015).

Limitations

Multi-model-averaged climate projections, such as those used within the AdaptWest dataset, are an approach that has emerged in prominence in recent years to combat ambiguities in climate change predictions (Shen et al., 2023). The deviation in values within inter-models can sometimes capture missing or over/under-estimated variability within climatic processes, increasing model precision (Shen et al., 2023). I consider this an implicit reason for producing the complementary dataset (AdaptWest); however, it could also explain why there are fewer optimal cells compared to the Chelsa-Bioclim+ dataset because the multi-model AdaptWest dataset represents the range of equilibrium and transient climate sensitivities (Mahony et al., 2022), which, when considering the models are trained on empirical historical data, could provide conflicting patterns and make identifying of these patterns in forecasted conditions nuanced.

Using high-performing quantitative benchmarks to evaluate a model's predictive ability does not ensure model certainty, particularly when selecting and relying on an appropriate framework to accurately represent your data to answer research questions or substantiate any underlying assumptions. Thoughtful considerations for the count and placement of absences and the integrity of input data are essential to reduce bias or noise generated within the models. Randomly generated pseudoabsences may lead to inaccurate predictions. Therefore, giving thoughtful consideration to how pseudoabsences are chosen is essential, with a myriad of literature available (and growing) to help with the process (Chefaoui & Lobo, 2008; Dubos et al.,

2022; Hazen et al., 2021; Liu et al., 2019; Senay et al., 2013; VanDerWal et al., 2009; Wisz & Guisan, 2009).

In addition to the variables I used, land cover, soil conditions, microhabitats, host species density, and leaf litter quantity are known factors influencing tick activity, adaptability, and survivability (Clow et al., 2017). My study used a cell resolution of ~1 km, which is too coarse to pick up the intricacies of small-scale observable changes (e.g., microclimates, slopes). Additionally, I refrained from using categorical variables and those that do not exhibit an accurate shift or alteration in conditions over time (e.g., land cover). Steady-state variables may represent an underlying continuum or are interdependent on other unexamined attributes, which introduce complexities that are not efficiently modelled or assessed. For example, *I. scapularis* could show favouritism for a particular soil type when, in reality, there was a propensity for soil depth, which was not addressed. Another possibility could be that land cover pivotally exerts an influence on *I. scapularis* distribution, or is it the underlying mechanisms of land cover that drive the observed significance, such as confounding variables of temperature or precipitation.

A caveat of using environmental data is that to control for some of the limitations of the data, ecological relationships (i.e., species-species and species-environmental interactions) have to be considered in isolation and in equilibrium, which is rarely the case. In particular, the dynamic nature of *I. scapularis* to travel on host species and traverse restrictive boundaries means the utmost consideration has to go into vetting the input data, such as ensuring that the *I. scapularis* records are founded on discernible ecological judgement, as well the values of the environmental data must accurately cohere to a given place and time. Nevertheless, the outcome reflects the input; ensuring adequate quality and quantity of input data leads to a more robust result; over time, input data will improve, and so too will the results.

Additionally, biogeographical models are prone to uncertainty when extrapolated to novel environmental conditions (in this instance, using SSP3-7.0 and SSP5-8.5 scenarios). SSP models attempt to simulate how conditions will change over time with increased greenhouse gas concentrations and newly incorporating socio-economic indicators. Our models, and therefore the distribution of *I. scapularis*, are trained on historical data within a calibration range, opposite to forecasted scenarios that fundamentally contain some level of ambiguity, potentially using climate values outside the range of values in which I trained the models on (i.e., non-analogue conditions).

Reusing a variable might slightly bias the results; using the same variable (albeit in different sequences) impacts the variable's weight in the final ensemble, thereby influencing the mapped distribution. The reprocessing of certain variables (apparent only in the secondary dataset) was for two reasons: firstly, relative humidity (as a percentage; Appendix A) is used in two sequences (sequences three and four). Studies show that high humidity is a factor paramount to the survival of *I. scapularis* (e.g., Berger et al., 2014; Brunner et al., 2012; Ginsberg et al., 2020; Stafford, 1994). The results of sequence three showed R.H. explaining only ~6.8%, a low estimate compared to its role in nature. Because of its significance, I included R.H. to see if the strength of other factors (i.e., the ecological indicators of annual snowfall, heat moisture, and precipitation) was heavily correlated with *I. scapularis* distribution and overshadowed the impact and response of humidity. Additionally, strong confounding variables (e.g., temperature) limited the viable variables to process a sequence; I deemed it necessary to recycle variables capable of acting as a counterbalance to assess critical variables.

Contemporary models can identify areas where *I. scapularis* might be present but undetected (Fouquet et al., 2010). To verify the validity of the models, the foremost available information on local knowledge and species behaviour should be used (Frans et al., 2022). Model data is split into training and test sets. The training set is used to tune the models, and the test set is used to verify model accuracy and its usefulness if used on unseen data (e.g., climate projection data). During the tuning of the training set, there is a risk that they 'learn' and 'memorise' patterns specific within the data subsets (Quinn et al., 2020). Instead of learning general concepts from the data, models may lose efficiency when generalising to novel information or data, resulting in overfitting problems and inferior forecasting (Quinn et al., 2020). Ideally, the test set is collected separately from the training set to avoid these issues. When a separate test set is unavailable, as in this study, then the test set should be evaluated using resampling methods (e.g., cross-validation; Quinn et al., 2020). Resources permitting, evaluating the test set using an external, unseen collection of *I. scapularis* records, using true-absences collected in the field, not artificially collected ones, would be beneficial. Moreover, it is important to determine ground-truth areas as “optimal” or those that have a “high” likelihood of *I. scapularis* occurrence (even at small scales). This can be achieved with the collection of empirical data to assess whether the modern niche ensembles produce an actual representation of the state of *I. scapularis* as predicted and if there is potential to use the maps for further analyses (e.g., risk analysis).

Conclusion

The small size of terrestrial arthropods means they are highly susceptible to some degree of phenotypic plasticity under warming temperatures and shifting moisture regimes; this particularly appears to be the case with those that are considered pest species. This frequently results in changes to physiology, behaviour, phenology, distribution, abundance, species assemblages, and species interactions (Harvey et al., 2022). Climate change is subsequently shifting many species northward; this includes species of public health concern like *I. scapularis*, with hundreds of documented records within previously uninhabited areas of Canada over the last five years. My investigation into the ecological niche and range expansion of *I. scapularis* within eastern Canada has provided valuable insights into the complex interplay between environmental factors and the distribution of tick species. Through the culmination of my research, which is a comprehensive assessment of key ecological variables and advanced geospatial modelling techniques, it is evident that records alone do not definitively confirm the establishment of or help us predict the future distribution of *I. scapularis* populations in Canada. Therefore, we should consider the increased interplay between tick surveillance and information gained by geospatial models to further our collective ability to validate models and improve our knowledge of *I. scapularis* range dynamics and host-species interactions.

With habitats altered over time, I have shown that *I. scapularis* will extend significantly into novel environments, particularly to the north, within the 21st century, underlying the importance of ongoing monitoring and adaptation to changing environmental conditions. The foreseeable trajectory of *I. scapularis,* with the potential risk of Lyme disease this entails, serves as a poignant reminder of the intricate and multifaceted relationship between biological distributions

and environmental parameters, emphasising the need for future research into risk prevention measures and better understanding of species-environment interactions.

Cited Information

References

Abbass, K., Qasim, M.Z., Song, H., Murshed, H., Mahmood., & Younis I. (2022). A review of the global climate change impacts, adaptation, and sustainable mitigation measures. *Environ Sci Pollut Res*, *29*, 42539–42559. https://doi.org/10.1007/s11356-022-19718-6

Aenishaenslin, C., Michel, P., Ravel, A., Gern, L., Waaub, J.-P., Milord, F., & Bélanger, D. (2015). Acceptability of tick control interventions to prevent lyme disease in Switzerland and Canada: A mixed-method study. *BMC Public Health*, *16*(1). https://doi.org/10.1186/s12889-015-2629-x

AdaptWest Project. (2022). Gridded current and projected climate data for North America at 1km resolution, generated using the ClimateNA v7.30 software (Wang et al., 2022). [Data file]. Adaptwest.data basin.org

Aiello-Lammens, M. A., Boria, A. R., Radosavljevic, A., Vilela, B., Anderson, P. R. (2015) spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38(5). 10.1111/ecog.01132

Alexander, J. M., Diez, J. M., & Levine, J. M. (2015). Novel competitors shape species' responses to climate change. *Nature*, *525*(7570), 515-518. https://www.nature.com/articles/nature14952

Araújo, M. B., Pearson, R. G., Thuiller, W., & Erhard, M. (2005). Validation of species–climate impact models under climate change. *Global Change Biology*, *11*(9), 1504-1513. https://doi.org/10.1111/j.1365-2486.2005.01000.x

Araújo, M., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, *22*(1), 42–47. https://doi.org/10.1016/j.tree.2006.09.010

Auguie B (2017). GridExtra: miscellaneous functions for "grid" graphics. *R.package version 2.3*. https://CRAN.R-project.org/package=gridExtra

Aubry, K. B., Raley, C. M., & McKelvey, K. S. (2017). The importance of data quality for generating reliable distribution models for rare, elusive, and cryptic species. *PLOS ONE*, *12*(6), e0179152. https://doi.org/10.1371/journal.pone.0179152

Backus, L. H., López Pérez, A. M., & Foley, J. E. (2021). Effect of temperature on host preference in two lineages of the brown dog tick, *Rhipicephalus sanguineus*. *The American journal of tropical medicine and hygiene*, *104*(6), 2305–2311. https://doi.org/10.4269/ajtmh.20-1376

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

Belli, A., Sarr, A., Rais, O., & Voordouw, M. (2017). Ticks infected via co-feeding transmission can transmit Lyme borreliosis to vertebrate hosts. *Sci Rep*. https://doi.org/10.1038/s41598-017-05231-1

Benelli, G. (2020). Pathogens manipulating tick behavior—through a glass, darkly. *Pathogens*, *9*(8). https://doi.org/10.3390/pathogens9080664

Berger, K. A., Ginsberg, H. S., Gonzalez, L., & Mather, T. N. (2014) Relative humidity and activity patterns of *Ixodes scapularis* (Acari: Ixodidae), *Journal of Medical Entomology*, *51*(4), 769–776. https://doi.org/10.1603/ME13186

Bivand, R., Keitt, T., & Rowlingson, B. (2022). Rgdal: Bindings for the 'geospatial' data abstraction library. *R.package version 1.5–3.2.* https://CRAN.R-project.org/package=rgdal

Bokulich, A. (2013). Explanatory models versus predictive models: Reduced complexity modeling in geomorphology. *The European Philosophy of Science Association Proceedings*, *2*, 115-128. 10.1007/978-3-319-01306-0_10

Bouchard, C., Dibernardo, A., Koffi, J., Wood, H., Leighton, P., & Lindsay, L. (2019). Increased risk of tick-borne diseases with climate and environmental changes. *Canada Communicable Disease Report*, *45*(4), 83–89. https://doi.org/10.14745/ccdr.v45i04a02

Bowden, J. J., Eskildsen, A., Hansen, R. R., Olsen, K., Kurle, C, M., & Høye, T. T. (2015). High-Arctic butterflies become smaller with rising temperatures. *Biol. Lett, 11.* http://doi.org/10.1098/rsbl.2015.0574

Brown de Colstoun, E. C., Huang, C., Wang, P.,Tilton, J. C., Tan, B., Phillips, J., Niemczura, S., Ling, P.-Y., & Wolfe, R. E. (2017). Global man-made impervious surface (GMIS) dataset from Landsat. *NASA Socioeconomic Data and Applications Center (SEDAC).* https://doi.org/10.7927/H4P55KKF

Brun, P., Zimmermann, N.E., Hari, C., Pellissier, L., & Karger, D.N. (2022). Global climate-related predictors at kilometre resolution for the past and future. *Earth System Science Data*. https://doi.org/10.5194/essd-2022-212

Brunner, J. L., Killilea, M., & Ostfeld, R. S. (2012). Overwintering survival of nymphal *Ixodes scapularis* (Acari: Ixodidae) under natural conditions. *Journal of Medical Entomology*, *49*(5), 981-987. https://doi.org/10.1603/ME12060

Burks, C. S., Stewarts, R. L., Needham, G. R., & Lee, R. E. (1996). The role of direct chilling injury and inoculative freezing in cold tolerance of *Amblyomma americanum*, *Dermacentor variabilis* and *Ixodes scapularis*. *Physiological Entomology*, *21*(1), 44–50. https://doi.org/10.1111/j.1365-3032.1996.tb00833.x

Burtis, J. C., Bjork, J., Johnson, T. L., Schiffman, E., Neitzel, D., & Eisen, R. J. (2023). Seasonal activity patterns of host-seeking *Ixodes scapularis* (Acari: Ixodidae) in Minnesota, 2015–2017. *Journal of Medical Entomology*, *60*(4), 769–777. https://doi.org/10.1093/jme/tjad048

Burtis, J. C, Yavitt, J. B, Fahey, T. J, & Ostfeld, R.S. (2019). Ticks as soil-dwelling arthropods: An intersection between disease and soil ecology, *Journal of Medical Entomology*, *56*(6), 1555–1564. https://doi.org/10.1093/jme/tjz116

Burtis, J. C., & Pflueger, C. (2017). Interactions between soil‐dwelling arthropod predators and *Ixodes scapularis* under laboratory and field conditions. *Ecosphere*, *8*(8). https://doi.org/10.1002/ecs2.1914

Burtis, J. C., Sullivan, P., Levi, T., Oggenfuss, K., Fahey, T. J., & Ostfeld, R. S. (2016). The impact of temperature and precipitation on blacklegged tick activity and Lyme disease

incidence in endemic and emerging regions. *Parasites & Vectors*, *9*(1), 606. https://doi.org/10.1186/s13071-016-1894-6

Carvalho, D., Rafael, S., Monteiro, A., Rodrigues, V., Lopes, M., & Rocha, A. (2022). How well have CMIP3, CMIP5 and CMIP6 future climate projections portrayed the recently observed warming. *Scientific Reports*, *12*. https://doi.org/10.1038/s41598-022-16264-6

Centers for Disease Control and Prevention. (2022). Lyme Disease. https://www.cdc.gov/lyme/index.html

Chefaoui, R. M., & Lobo, J. M. (2008). Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*, *210*(4), 478–486. https://doi.org/10.1016/j.ecolmodel.2007.08.010

Chilton, N. B., Curry, P. S., Lindsay, L. R., Rochon, K., Lysyk, T. J., & Dergousoff, S. J. (2020). Passive and Active Surveillance for Ixodes scapularis (Acari: Ixodidae) in Saskatchewan, Canada. *Journal of Medical Entomology*, *57*(1), 156-163. https://doi.org/10.1093/jme/tjz155

Christensen, P., Gillingham, K., & Nordhaus, W. (2018). Uncertainty in forecasts of long-run economic growth. *Proceedings of the National Academy of Sciences*, *115*(21), 5409–5414. https://doi.org/10.1073/pnas.1713628115

Clark, D. D. (1995). Lower temperature limits for activity of several Ixodid ticks (Acari: Ixodidae): Effects of body size and rate of temperature change. *Journal of Medical Entomology*, *32*(4), 449-452. https://doi.org/10.1093/jmedent/32.4.449

Clow, K. M., Ogden, N. H., Lindsay, L. R., Michel, P., Pearl, D. L., & Jardine, C. M. (2017). The influence of abiotic and biotic factors on the invasion of *Ixodes scapularis* in Ontario,

Canada. *Ticks and Tick-Borne Diseases*, *8*(4), 554–563 https://doi.org/10.1016/j.ttbdis.2017.03.003

Collins, S. D., Abbott, J. C., & McIntyre, N. E. (2017). Quantifying the degree of bias from using county-scale data in species distribution modeling: Can increasing sample size or using county-averaged environmental data reduce distributional overprediction? *Ecology and Evolution*, *7*(15), 6012-6022. https://doi.org/10.1002/ece3.3115

Colwell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences*, *106*(supplement_2), 19651-19658. https://doi.org/10.1073/pnas.0901650106

Connolly, R, S., Keith, A, S., Colwell, K, R., & Rahbek, C. (2017). Process, mechanism, and modeling in macroecology. *Trends in Ecology & Evolution, 32*(11), 835-844. https://doi.org/10.1016/j.tree.2017.08.011

Cook, M. (2014). Lyme Borreliosis: A review of data on transmission time after tick attachment. *International Journal of General Medicine*, 1–8. https://doi.org/10.2147/ijgm.s73791

Couper, L. I., MacDonald, A. J., & Mordecai, E. A. (2021). Impact of prior and projected climate change on US Lyme disease incidence. *Global Change Biology*, *27*(4), 738-754. https://doi.org/10.1111/gcb.15435

Cuddington, K., Fortin, J., Gerber, L. R., Hastings, A., Liebhold, A., & Ray, C. (2013). Process-based models are required to manage ecological systems in a changing world. *Ecosphere*, *4*(2), 1-12. https://doi.org/10.1890/ES12-00178.1

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2012).

Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dubos, N., Préau, C., Lenormand, M., Papuga, G., Monsarrat, S., Denelle, P., Louarn, M. L., Heremans, S., May, R., Roche, P., & Luque, S. (2022). Assessing the effect of sample bias correction in species distribution models. *Ecological Indicators*, *145*. https://doi.org/10.1016/j.ecolind.2022.109487

Duffy, D. C., & Campbell, S. R. (1994). Ambient air temperature as a predictor of activity of adult *Ixodes scapularis* (Acari: Ixodidae). *Journal of Medical Entomology*, *31*(1), 178-180. https://doi.org/10.1093/jmedent/31.1.178

Dumas, A., Bouchard, C., Dibernardo, A., Drapeau, P., Lindsay, L. R., Ogden, N. H., & Leighton, P. A. (2022). Transmission patterns of tick-borne pathogens among birds and rodents in a forested park in southeastern Canada. *PLOS One*, *17*(4), 1–25. https://doi.org/10.1371/journal.pone.0266527

Eisen, R. J., & Eisen, L. (2018). The Blacklegged tick, *Ixodes scapularis*: An increasing public health concern. *Trends in Parasitology*, *34*(4), 295–309. https://doi.org/10.1016/j.pt.2017.12.006

Eisen, R. J., Eisen, L., & Beard, C. B. (2016a). County-Scale Distribution of Ixodes scapularis and Ixodes pacificus (Acari: Ixodidae) in the Continental United States. *Journal of Medical Entomology*, *53*(2), 349. https://doi.org/10.1093/jme/tjv237

Eisen, R. J., Eisen, L., Ogden, N. H., & Beard, C. B. (2016b). Linkages of weather and climate with *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae), enzootic transmission of

Borrelia burgdorferi, and Lyme disease in North America, *Journal of Medical Entomology, 53*(2), 250–261. https://doi.org/10.1093/jme/tjv199

Elias, S. P., Witham, J. W., Schneider, E. F., Rand, P. W., Hunter, M. L., Lubelczyk, C., & Smith, R. P. (2022). Emergence of *Ixodes scapularis* (Acari: Ixodidae) in a small mammal population in a coastal oak-pine forest, Maine, USA. *Journal of Medical Entomology*, *59*(2), 725-740. https://doi.org/10.1093/jme/tjab209

Elith, J., Graham, C., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., M. Overton, J. M., Peterson, A. T., Phillips, S. J., Richardson, K., . . . Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129-151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Engler, R., Guisan, A., & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, *41*(2), 263–274. https://doi.org/10.1111/j.0021-8901.2004.00881.x

ESRI (2020). *ArcGIS Pro* (Version 3.0). https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

Esser, L. F., Neves, D. M., & Jarenkow, J. A. (2019). Habitat‐specific impacts of climate change in the Mata Atlântica biodiversity hotspot. *Diversity and Distributions*, *25*(12), 1846–1856. https://doi.org/10.1111/ddi.12984

eTick. (2021). Public platform for image-based identification and population monitoring in Canada [Data file]. http://www.eTick.ca

eTick. (2023). Public platform for image-based identification and population monitoring in Canada. http://www.eTick.ca

Fieler, A. M., Rosendale, A. J., Farrow, D. W., Dunlevy, M. D., Davies, B., Oyen, K., Xiao, Y., & Benoit, J. B. (2021). Larval thermal characteristics of multiple Ixodid ticks. *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology*, *257*. https://doi.org/10.1016/j.cbpa.2021.110939

Frans, V. F., Augé, A. A., Fyfe, J., Zhang, Y., McNally, N., Edelhoff, H., Balkenhol, N., & Engler, J. O. (2022). Integrated SDM database: Enhancing the relevance and utility of species distribution models in conservation management. *Methods in Ecology and Evolution*, *13*(1), 243–261. https://doi.org/10.1111/2041-210x.13736

Franklin, J. (2010). Mapping species distributions: Spatial inference and prediction. *Cambridge University Press*, 1–320. https://doi.org/10.1017/CBO9780511810602

Fouquet, A., Ficetola, G. F., Haigh, A., & Gemmell, N. (2010). Using ecological niche modelling to infer past, present and future environmental suitability for *Leiopelma Hochstetteri*, an endangered New Zealand native frog. *Biological Conservation*, *143*(6), 1375–1384. https://doi.org/10.1016/j.biocon.2010.03.012

Gaff, H., Eisen, R. J., Eisen, L., Nadolny, R., Bjork, J., & Andrew J Monaghan, A, J. (2020) LYMESIM 2.0: An updated simulation of blacklegged tick (Acari: Ixodidae) population dynamics and enzootic transmission of *Borrelia burgdorferi* (Spirochaetales: Spirochaetaceae). *Journal of Medical Entomology*, *57*(3), 715–727. https://doi.org/10.1093/jme/tjz252

Gilman, S. E., Urban, M. C., Tewksbury, J., Gilchrist, G. W., & Holt, R. D. (2010). A framework for community interactions under climate change. *Trends in ecology & evolution*, *25*(6), 325-331. https://www.cell.com/trends/ecology-evolution/fulltext/S0169-5347(10)00061-3?large_figure=true

Ginsberg, H. S., Rulison, E. L., Miller, J. L., Pang, G., Arsnoe, I. M., Hickling, G. J., Ogden, N. H., LeBrun, R. A., & Tsao, J. I. (2020). Local abundance of *Ixodes scapularis* in forests: Effects of environmental moisture, vegetation characteristics, and host abundance. *Ticks and Tick-Borne Diseases*, *11*(1), 101271. https://doi.org/10.1016/j.ttbdis.2019.101271

Gobeyn, S., Mouton, A. M., Cord, A. F., Kaim, A., Volk, M., & Goethals, P. L. (2019). Evolutionary algorithms for species distribution modelling: A review in the context of machine learning. *Ecological Modelling*, *392*, 179-195. https://doi.org/10.1016/j.ecolmodel.2018.11.013

Goddard, J., Embers, M., Hojgaard, A., & Piesman, J. (2015). Comparison of tick feeding success and vector competence for Borrelia burgdorferi among immature *Ixodes scapularis* (Ixodida: Ixodidae) of both southern and northern clades. *Journal of Medical Entomology*, *52*(1), 81–85. https://doi.org/10.1093/jme/tju005

Graham, C. H., Elith, J., Hijmans, R. J., Guisan, A., Peterson, A. T., & Loiselle, B. A. (2008). The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, *45*(1), 239-247. https://doi.org/10.1111/j.1365-2664.2007.01408.x

Green, C., & Byrne, K. A. (2004). Biomass: Impact on carbon cycle and greenhouse gas emissions. *Encyclopedia of Energy*, 223–236. https://app.knovel.com/hotlink/pdf/id:kt004KV472/encyclopedia-energy-volumes/biomass-carbon-cycling

Halsey, S. J., Allan, B. F., & Miller, J. R. (2018). The role of *Ixodes scapularis*, *Borrelia burgdorferi* and wildlife hosts in Lyme disease prevalence: A quantitative review. *Ticks and Tick-Borne Diseases*, *9*(5), 1103-1114. https://doi.org/10.1016/j.ttbdis.2018.04.006

Halsey, S. J., & Miller, J. R. (2018). A spatial agent-based model of the disease vector *Ixodes scapularis* to explore host-tick associations. *Ecological Modelling*, *387*, 96–106. https://doi.org/10.1016/j.ecolmodel.2018.09.005

Hamer, S. A., Hickling, G. J., Sidge, J. L., Walker, E. D., & Tsao, J. I. (2012). Synchronous phenology of juvenile *Ixodes scapularis*, vertebrate host relationships, and associated patterns of *borrelia burgdorferi* ribotypes in the Midwestern United States. *Ticks and Tick-Borne Diseases*, *3*(2), 65–74. https://doi.org/10.1016/j.ttbdis.2011.11.004

Hargreaves, G.H., & Allen, R.G. (2003). History and evaluation of Hargreaves Evapotranspiration Equation. *Journal of Irrigation and Drainage Engineering*, *129* (1), 53–63. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:1(53)

Harrison, J., Woods, H., & Stephen, R. (2012). Ecological and environmental physiology of insects. 10.1093/acprof:oso/9780199225941.001.0001

Harvey, J. A., Tougeron, K., Gols, R., Heinen, R., Abarca, M., Abram, P. K., Basset, Y., Berg, M., Boggs, C., Brodeur, J., Cardoso, P., de Boer, J. G., De Snoo, G. R., Deacon, C., Dell, J. E., Desneux, N., Dillon, M. E., Duffy, G. A., Dyer, L. A., … Chown, S. L. (2022).

Scientists' warning on climate change and insects. *Ecological Monographs*, *93*(1). https://doi.org/10.1002/ecm.1553

Hausfather, Z., & Peters, G. P. (2020). Emissions – the 'business as usual' story is misleading. *Nature*, *577*(7792), 618–620. https://doi.org/10.1038/d41586-020-00177-3

Hayes, L. E., Scott, J. A., & Stafford, K. C. (2015). Influences of weather on *Ixodes scapularis* nymphal densities at long-term study sites in Connecticut. *Ticks and Tick-Borne Diseases*, *6*(3), 258–266. https://doi.org/10.1016/j.ttbdis.2015.01.006

Hazen, E. L., Abrahms, B., Brodie, S., Carroll, G., Welch, H., & Bograd, S. J. (2021). Where did they not go? Considerations for generating pseudo-absences for telemetry-based habitat models. *Movement Ecology*, *9*(1). https://doi.org/10.1186/s40462-021-00240-2

Hijman, R. J., & Elith, J. (2023). Species distribution models. *Spatial Data Science with R and "Terra"*. https://rspatial.org/raster/sdm/raster_SDM.pdf

Hijmans, R. (2022). Raster: Geographic data analysis and modeling. *R.package version 3.6-3*. https://CRAN.R-project.org/package=raster.

Huang, C., Kay, S. C., Davis, S., Tufts, D. M., Gaffett, K., Tefft, B., & Diuk-Wasser, M. A. (2019). High burdens of Ixodes scapularis larval ticks on white-tailed deer may limit Lyme disease risk in a low biodiversity setting. Ticks and Tick-Borne Diseases, 10(2), 258-268. https://doi.org/10.1016/j.ttbdis.2018.10.013

Hutter, F., Lücke, J., & Schmidt-Thieme, L. (2015). Beyond manual tuning of hyperparameters. *KI - Künstliche Intelligenz*, *29*(4), 329–337. https://doi.org/10.1007/s13218-015-0381-0

iNaturalist. (2022). Eastern Black-legged tick (*Ixodes scapularis*) [Data file]. https://www.inaturalist.org/taxa/60598-Ixodes-scapularis

iNaturalist. (2023). https://www.inaturalist.org/

IPCC. (2023) *Climate Change 2023: Synthesis report*. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team, Lee, H., & Romero, J. (Eds.). Cambridge University Press. 35–115. 10.59327/IPCC/AR6-9789291691647

IPCC. (2021a). *Climate Change 2021: The physical science basis.* Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., & Zhou, B. (eds.). Cambridge University Press. 10.1017/9781009157896.

IPCC. (2021b) Summary for Policymakers. In: *Climate Change 2021: The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., & Zhou, B. (eds.). Cambridge University Press, 3−32, doi:10.1017/9781009157896.001.

IPCC. (2019). *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.* P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, & J. Malley, (Eds.), 1–874.

https://www.ipcc.ch/site/assets/uploads/2019/11/SRCCL-Full-Report-Compiled-191128.pdf

Jiménez, L., & Soberón, J. (2020). Leaving the area under the receiving operating characteristic curve behind: An evaluation method for species distribution modelling applications based on presence‐only data. *Methods in Ecology and Evolution*, *11*(12), 1571–1586. https://doi.org/10.1111/2041-210x.13479

Karger, D. N., Brun, P., & Zimmermann, N. E. (2022). (tech.). *BIOCLIM+* A novel set of global climate-related predictors at kilometre-resolution: Technical specification, 1–24. https://www.envidat.ch/dataset/21d662b7-9c59-41da-aa82-7d8a879b8db7/resource/aa676d1f-5f7f-479b-af20-26066b0537d1/download/chelsa_file_specification_bioclim_plus.pdf

Karger D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E,, Linder, H.P., Kessler, M. (2023). Climatologies at high resolution for the Earth's land surface areas. *Dryad Digital Repository* [Data file]. http://dx.doi.org/doi:10.5061/dryad.kd1d4

Karger, D.N., Lange, S., Hari, C., Reyer, C. P. O., Zimmermann, N.E. (2021a). CHELSA-W5E5 v1.0: W5E5 v1.0 downscaled with CHELSA v2.0. *ISIMIP Repository*. https://doi.org/10.48364/ISIMIP.836809

Karger, D.N., Wilson, A.M., Mahony, C., Zimmermann, N.E., Jetz, W. (2021b). Global daily 1km land surface precipitation based on cloud cover-informed downscaling. *Scientific Data.* doi.org/10.1038/s41597-021-01084-6

Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla-Buitrago, G. E., Boria, R. A., Soley-Guardia, M., & Anderson, R. P. (2021). ENMeval 2.0: Redesigned for

customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.13628

Keirans, J.E., Hutcheson, J.H., Durden, L. A, & Klompen, J. S. H. (1996). *Ixodes (Ixodes) scapularis* (Acari: Ixodidae): Redescription of all active Stages, distribution, hosts, geographical variation, and medical and veterinary importance, *Journal of Medical Entomology*, *33*(3), 297–318. https://doi.org/10.1093/jmedent/33.3.297

Kemp, L., Xu, C., Depledge, J., Ebi, K. L., Gibbins, G., Kohler, T. A., Rockström, J., Scheffer, M., Schellnhuber, H. J., Steffen, W., & Lenton, T. M. (2022). Climate endgame: Exploring catastrophic climate change scenarios. *Proceedings of the National Academy of Sciences*, *119*(34). https://doi.org/10.1073/pnas.2108146119

Klose, A. K., Wunderling, N., Winkelmann, R., & Donges, J. F. (2021). What do we mean, 'tipping cascade'? *Environmental Research Letters*, *16*(12). https://doi.org/10.1088/1748-9326/ac3955

Krakowetz, C. N., Lindsay, L. R., & Chilton, N. B. (2011). Genetic diversity in *Ixodes scapularis* (Acari: Ixodidae) from six established populations in Canada. *Ticks and Tick-Borne Diseases*, *2*(3), 143–150. https://doi.org/10.1016/j.ttbdis.2011.05.003

Kuhn, M. (2022) Caret: Classification and regression training. *R.package version 6.0-93*. https://CRAN.R-project.org/package=caret

Lancaster, L. (2016). Widespread range expansions shape latitudinal variation in insect thermal limits. *Nature Clim Change, 6*, 618–621. https://doi.org/10.1038/nclimate2945

Leal, B., Zamora, E., Fuentes, A., Thomas, D. B., & Dearth, R. K. (2020). Questing by tick larvae (Acari: Ixodidae): A review of the influences that affect off-host survival. *Annals*

of the Entomological Society of America, *113*(6), 425–438. https://doi.org/10.1093/aesa/saaa013

Léger, E., Vourc'h, G., Vial, L., Chevillon, C., & McCoy, K. D. (2013). Changing distributions of ticks: Causes and consequences. *Exp Appl Acarol*, *59*(1–2), 219–244. https://doi.org/10.1007/s10493-012-9615-0

Leighton, P. A., Koffi, J. K., Pelcat, Y., Lindsay, L. R., & Ogden, N. H. (2012). Predicting the speed of tick invasion: An empirical model of range expansion for the Lyme disease vector *Ixodes scapularis* in Canada. *Journal of Applied Ecology*, *49*(2), 457–464. https://doi.org/10.1111/j.1365-2664.2012.02112.x

Levi, T., Keesing, F., Oggenfuss, K., & Ostfeld, R. S. (2015). Accelerated phenology of Blacklegged ticks under climate warming. *Phil. Trans. R. Soc. B*, *370*. https://doi.org/10.1098/rstb.2013.0556

Lieth, H. (1975). Modeling the primary productivity of the world. Lieth, H., Whittaker, R.H. (Eds.) *Primary Productivity of the Biosphere* (14, pp. 237–263). Springer https://doi.org/10.1007/978-3-642-80913-2_12

Lindsay, L. R., Barker, I. K., Surgeoner, G. A., Mcewen, S. A., Gillespie, T. J., & Addison, E. M. (1998). Survival and development of the different life stages of *Ixodes scapularis* (Acari: Ixodidae) held within four habitats on Long Point, Ontario, Canada. *Journal of Medical Entomology*, *35*(3), 189-199. https://doi.org/10.1093/jmedent/35.3.189

Lindsay, L. R., Barker, I. K., Surgeoner, G. A., McEwen, S. A., Gillespie, T. J., & Robinson, J. T. (1995). Survival and development of *Ixodes scapularis* (Acari: Ixodidae) under various climatic conditions in Ontario, Canada. *Journal of Medical Entomology*, *32*(2), 143–152. https://doi.org/10.1093/jmedent/32.2.143

Li, C., Zhu, H., Luo, H., Zhou, S., Kong, J., Qi, L., & Rao, C. (2023). Spread prediction and classification of Asian Giant Hornets based on GM-logistic and CSRF models. *Mathematics*, *11*(6), 1332. https://doi.org/10.3390/math11061332

Linske, M. A., Stafford, K. C., Williams, S. C., Lubelczyk, C. B., Welch, M., & Henderson, E. F. (2019). Impacts of deciduous leaf litter and snow presence on nymphal *Ixodes scapularis* (Acari: Ixodidae) overwintering survival in coastal New England, USA. *Insects*, *10*(8), 227. https://doi.org/10.3390/insects10080227

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, *28*(3), 385–393. https://doi.org/10.1111/j.0906-7590.2005.03957.x

Liu, C., Newell, G., & White, M. (2019). The effect of sample size on the accuracy of species distribution models: Considering both presences and pseudo-absences or background sites. *Ecography*, *42*(3), 535–548. https://doi.org/10.1111/ecog.03188

Liu, C., Newell, G., & White, M. (2015). On the selection of thresholds for predicting species occurrence with presence‐only data. *Ecology and Evolution*, *6*(1), 337–348. https://doi.org/10.1002/ece3.1878

Liu, C., White, M., & Newell, G. (2009). Measuring the accuracy of species distribution models: a review. *Proceedings 18th World IMACs/MODSIM Congress, 4241*. https://www.researchgate.net/profile/Canran-Liu-2/publication/288962465_Measuring_the_accuracy_of_species_distribution_models_a_review/links/5bf3456492851c6b27cadf31/Measuring-the-accuracy-of-species-distribution-models-a-review.pdf

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence‐only data. *Journal of Biogeography*, *40*(4), 778–789. https://doi.org/10.1111/jbi.12058

Liu, C., White, M., & Newell, G. (2011). Measuring and comparing the accuracy of species distribution models with presence–absence data. *Ecography*, *34*(2), 232-243. https://doi.org/10.1111/j.1600-0587.2010.06354.x

Liu, P, R., & Raftery, A, E. (2021) Country-based rate of emissions reductions should increase by 80% beyond nationally determined contributions to meet the 2 °C target. *Commun, 2, 1-29.* https://doi.org/10.1038/s43247-021-00097-8

Loarie, S. (Ed.). (2021). *Geoprivacy*. iNaturalist. https://www.inaturalist.org/pages/geoprivacy

Mahony, C, R., Wang, T., Hamann, A., & Cannon, A, J. (2022). A CMIP6 ensemble for downscaled monthly climate normals over North America. *International Journal of Climatology 42, (11)*, 5871-5891. https://doi.org/10.1002/joc.7566

Manevitz, L. M., & Yousef, M. (2001). One-class SVMs for document classification. Journal of machine Learning research, 2, 139-154. https://www.jmlr.org/papers/volume2/manevitz01a/manevitz01a.pdf?ref=https://codemonkey.link

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K., & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, *15*(1), 59–69. https://doi.org/10.1111/j.1472-4642.2008.00491.x

Mathisson, D. C., Kross, S. M., Palmer, M. I., & Diuk-Wasser, M. A. (2021). Effect of vegetation on the abundance of tick vectors in the northeastern United States: A Review of the

literature. *Journal of Medical Entomology*, *58*(6), 2030–2037. https://doi.org/10.1093/jme/tjab098

Matthews, T., Wilby, R.L., & Murphy, C. (2019). An emerging tropical cyclone–deadly heat compound hazard. *Nat. Clim. Chang, 9*, 602–606. https://doi.org/10.1038/s41558-019-0525-6

McCabe, G. J., & Bunnell, J. E. (2004). Precipitation and the occurrence of Lyme disease in the northeastern United States. *Vector-Borne and Zoonotic Diseases*, *4*(2), 143–148. https://doi.org/10.1089/1530366041210765

McCauley, S. J., & Mabry, K. E. (2011). Climate change, body size, and phenotype dependent dispersal. *Trends in Ecology and Evolution*, *26*(11). https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1352&context=bio_fac

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276-282. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/

McInerny, G. J., & Etienne, R. S. (2013). Niche" or "distribution" modelling? A response to Warren. *Trends Ecol. Evol*, *28*, 191-192.

Meinshausen, M., Nicholls, Z. R., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., … Wang, R. H. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, *13*(8), 3571–3605. https://doi.org/10.5194/gmd-13-3571-2020

Melillo, J. M., McGuire, A. D., Kicklighter, D. W., Moore, B., Vorosmarty, C. J., & Schloss, A. L. (1993). Global climate change and terrestrial net primary production. *Nature*, *363* (6426), 234–240. https://doi.org/10.1038/363234a0

Melo-Merino, S. M., Reyes-Bonilla, H., & Lira-Noriega, A. (2020). Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. *Ecological Modelling*, 1–12. https://doi.org/10.1016/j.ecolmodel.2019.108837

Morales‐Barbero, J., & Vega‐Álvarez, J. (2019). Input matters matter: Bioclimatic consistency to map more reliable species distribution models. *Methods in Ecology and Evolution*, *10*(2), 212–224. https://doi.org/10.1111/2041-210x.13124

Morgan, J. (2014). Classification and regression tree analysis.1–16. https://www.bu.edu/sph/files/2014/05/MorganCART.pdf

Morrill, R., Gaile, G.L., & Thrall, G.I. (1988). *Spatial Diffusion*. Thrall, G.I. (Eds.). WVU Research Repository. https://researchrepository.wvu.edu/rri-web-book/11/

Mount, G. A., Haile, D. G., & Daniels, E. (1997). Simulation of Blacklegged tick (Acari: Ixodidae) population dynamics and transmission of *Borrelia burgdorferi*. *Journal of Medical Entomology*, *34*(4), 461–484. https://doi.org/10.1093/jmedent/34.4.461

Natural Resources Canada. (2022). *Growing Season*. https://natural-resources.canada.ca/climate-change/climate-change-impacts-forests/forest-change-indicators/growing-season/18470

Natural Resources Canada. (2020). *Drought: Climate Moisture Index (CMI)* [Hogg, E.H., Price, D.T., McKenney, D.W.]. Forest Change Data Catalogue. https://cfs.nrcan.gc.ca/fc-data-catalogue/read/1

Nabbout, A. E., Ferguson, L. V., Miyashita, A., & Adamo, S. A. (2023). Female ticks (*Ixodes scapularis*) infected with *Borrelia burgdorferi* have increased overwintering survival, with implications for tick population growth. *Insect Science*. https://doi.org/10.1111/1744-7917.13205

Naimi B., & Araujo M.B. (2016). Sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, *39*, 368-375. doi:10.1111/ecog.01881

Naimi B., Hamm N, A., Groen T, A., Skidmore A, K., & Toxopeus A, G. (2014). Where is positional uncertainty a problem for species distribution modelling. *Ecography*, *37*(2), 191-203. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Naimi, B. (2023). Usdm: Uncertainty analysis for species distribution modelling. *R. package version 2.1-6.* https://cran.r-project.org/package=usdm

Natural Resources Canada, & U.S Geological Survey. (2010). Canada - United States Boundary [Data file]. https://open.canada.ca/en

Nazarenko, L. S., Tausnev, N., Russell, G. L., Rind, D., Miller, R. L., Schmidt, G. A., Bauer, S. E., Kelley, M., Ruedy, R., Ackerman, A. S., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T. L., Cook, B. I., Cruz, C. A., … Yao, M. (2022). Future climate change under SSP emission scenarios with GISS‐E2.1. *Journal of Advances in Modeling Earth Systems*, *14*(7). https://doi.org/10.1029/2021ms002871

Neelakanta, G., Sultana, H., Fish, D., Anderson, J. F., & Fikrig, E. (2010). *Anaplasma phagocytophilum* induces *Ixodes scapularis* ticks to express an antifreeze glycoprotein gene that enhances their survival in the cold. *Journal of Clinical Investigation*, *120*(9), 3179–3190. https://doi.org/10.1172/jci42868

Nuss, A., Sharma, A., & Gulia-Nuss, M. (2021). Genetic manipulation of ticks: A paradigm shift in tick and tick-borne diseases research. *Frontiers in Cellular and Infection Microbiology*, *11*. https://doi.org/10.3389/fcimb.2021.678037

Nuttall, P. A. (2021). Climate change impacts on ticks and tick-borne infections. *Biologia*, *77*(6), 1503–1512. https://doi.org/10.1007/s11756-021-00927-2

Ogden, N. H., Beard, C. B., Ginsberg, H. S., & Tsao J. I. (2021) Possible effects of climate change on Ixodid ticks and the pathogens they transmit: Predictions and observations, *Journal of Medical Entomology*, *58*(4), 1536–1545. https://doi.org/10.1093/jme/tjaa220

Ogden, N.H., Bigras-Poulin, M., Hanincová, K., Maarouf, A., O'Callaghan, C. J., & Kurtenbach, K. (2008a). Projected effects of climate change on tick phenology and fitness of pathogens transmitted by the North American tick *Ixodes scapularis*. *Journal of Theoretical Biology*, *254*(3), 621–632. https://doi.org/10.1016/j.jtbi.2008.06.020

Ogden, N. H., Lindsay, L. R., Beauchamp, G., Charron, D., Maarouf, A., O'Callaghan, C. J., Waltner-Toews, D., & Barker, I. K. (2004). Investigation of relationships between temperature and developmental rates of tick *Ixodes scapularis* (Acari: Ixodidae) in the laboratory and field. *Journal of Medical Entomology*, *41*(4), 622–633. https://doi.org/10.1603/0022-2585-41.4.622

Ogden, N. H., Lindsay, R. L., Hanincová, K., Barker, I. K., Bigras-Poulin, M., Charron, D. F., Heagy, A., Francis, C. M., O'Callaghan, C. J., Schwartz, I., & Thompson, R. A. (2008b). Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocytophilum* in Canada. *Applied and Environmental Microbiology*, *74*(12). https://doi.org/10.1128/AEM.01982-07

Ogden, N. H., Lindsay, L. R., & Leighton, P. A. (2013). Predicting the rate of invasion of the agent of Lyme disease Borrelia burgdorferi. *Journal of Applied Ecology*, *50*(2), 510-518. https://doi.org/10.1111/1365-2664.12050

Ogden, N. H., Maarouf, A., Barker, I. K., Bigras-Poulin, M., Lindsay, L. R., Morshed, M. G., O'Callaghan, C. J., Ramay, F., Waltner-Toews, D., & Charron, D. F. (2006a). Climate change and the potential for range expansion of the Lyme disease vector *Ixodes scapularis* in Canada. *International Journal for Parasitology*, *36*(1), 63–70. https://doi.org/10.1016/j.ijpara.2005.08.016

Ogden, N. H., Pang, G., Ginsberg, H. S., Hickling, G. J., Burke, R. L., Beati, L., & Tsao, J. I. (2018). Evidence for geographic variation in life-cycle processes affecting phenology of the Lyme disease vector *Ixodes scapularis* (Acari: Ixodidae) in the United States. *Journal of Medical Entomology, 55*(6), 1386–1401. https://doi.org/10.1093/jme/tjy104

Ogden, N. H, St-Onge, L., Barker, I. K., Brazeau, S., Bigras-Poulin, M., Charron, D. F., Francis, C. M., Heagy, A., Lindsay, R., Maarouf, A., Michel, P., Milord, F., O'Callaghan, C. J., Trudel, L., & Thompson, A. (2008c). Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. *International Journal of Health Geographics*, *7*(1). https://doi.org/10.1186/1476-072x-7-24

Ogden, N. H., Trudel, L., Artsob, H., Barker, I. K., Beauchamp, G., Charron, D. F., Drebot, M. A., Galloway, T. D., Thompson, R. A., & Lindsay, L. R. (2006b). *Ixodes scapularis* ticks collected by passive surveillance in Canada: Analysis of geographic distribution and infection with Lyme borreliosis agent *Borrelia burgdorfer*i. *Journal of Medical Entomology*, *43*(3), 600-609. https://doi.org/10.1093/jmedent/43.3.600

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

O'Neill, D., Häkkinen, H., Neumann, J., Shaffrey, L., Cheffings, C., Norris, K., & Pettorelli, N. (2023). Investigating the potential of social media and citizen science data to track changes in species' distributions. *Ecology and Evolution*, *13*(5), e10063. https://doi.org/10.1002/ece3.10063

Parmesan, C., Morecroft, M.D., Trisurat, Y., R. Adrian, Anshari, G.Z., Arneth, A., Gao, Q., Gonzalez, P., Harris, R., Price, J., Stevens, N., & Talukdar, G.H. (2022) Terrestrial and freshwater ecosystems and their services. In: C*limate change 2022: impacts, adaptation, and vulnerability. contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change.* Pörtner, H.-O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (Eds.). Cambridge University. 197–377. doi:10.1017/9781009325844.004.

Paulsen, J., & Körner, C. (2014). A climate-based model to predict potential treeline position around the globe. *Alpine Botany*, *124*(1), 1–12. https://doi.org/10.1007/s00035-014-0124-0

Peterson, T, A., Papes, M., & Soberón, J. (2015). Mechanistic and correlative models of ecological niches. *European Journal of Ecology*, *1*(2), 10.1515/eje-2015-0014

Peterson, A. T., & Soberón, J. (2012). Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservação*, *10*(2), 102-107.

Pianka, E. (2011). Convergent evolution and ecological equivalence. In Hutchins, M., Geist, V., & Pianka, E.,(Eds.) *Grzimek's Animal Life Encyclopedia, Evolution* (pp. 235–242). Gale Publishing Group.

Piñeiro, R., Aguilar, J. F., Munt, D. D., & Feliner, G. N. (2007). Ecology matters: Atlantic–Mediterranean disjunction in the sand-dune shrub *Armeria pungens* (Plumbaginaceae). *Molecular Ecology*, *16*(10), 2155-2171. https://doi.org/10.1111/j.1365-294X.2007.03280.x

Pocheville, A. (2015). The Ecological Niche: History and Recent Controversies. In Heams, T., Huneman, P., Lecointre, G., Silberstein, M. (eds) *Handbook of Evolutionary Thinking in the Sciences*. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9014-7_26

Porretta, D., Mastrantonio, V., Amendolia, S., Gaiarsa, S., Epis, S., Genchi, C., ... & Urbanelli, S. (2013). Effects of global changes on the climatic niche of the tick *Ixodes ricinus* inferred by species distribution modelling. *Parasites & Vectors*, *6*, 1-8. https://doi.org/10.1186/1756-3305-6-271

Public Health Agency of Canada. (2022). Lyme disease: Surveillance. https://www.canada.ca/en/public-health/services/diseases/lyme-disease/surveillance-lyme-disease.html

QGIS Development Team (2022). Geographic Information System (Version 3.28.4). *Open Source Geospatial Foundation Project*. http://qgis.org.

Qiao, H., Soberón, J., & Peterson, A. T. (2015). No silver bullets in correlative ecological niche modelling: Insights from testing among many potential algorithms for niche estimation.

Methods in Ecology and Evolution, 6(10), 1126-1136. https://doi.org/10.1111/2041-210X.12397

Quinn, T. P., Le, V., & Cardilini, A. P. (2020). Test set verification is an essential step in model building. *Methods in Ecology and Evolution*, *12*(1), 127–129. https://doi.org/10.1111/2041-210x.13495

Raes, N., & ter Steege, H. 2007. A null-model for significance testing of presence-only species distribution models. *Ecography, 30,* 727–736. https://doi.org/10.1111/j.2007.0906-7590.05041.x

Rand, P. W., Holman, M. S., Lubelczyk, C., Lacombe, E. H., DeGaetano, A. T., & Smith, R. P., Jr. (2004). Thermal accumulation and the early development of *Ixodes scapularis*. *J Vector Ecol*, 29(1), 164–176. https://pubmed.ncbi.nlm.nih.gov/15266754/

R Core Team (2020). R: A language and environment for statistical computing. https://www.R-project.org/.

Rehm, E. M., Olivas, P., Stroud, J., & Feeley, K. J. (2015). Losing your edge: Climate change and the conservation value of range-edge populations. *Ecology and Evolution*, *5*(19), 4315-4326. https://doi.org/10.1002/ece3.1645

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., … Tavoni, M. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Ringwaldt, E. M., Brook, B. W., Buettel, J. C., Cunningham, C. X., Fuller, C., Gardiner, R., Hamer, R., Jones, M., Martin, A. M., & Carver, S. (2023). Host, environment, and anthropogenic factors drive landscape dynamics of an environmentally transmitted pathogen: Sarcoptic mange in the bare-nosed wombat. *Journal of Animal Ecology*, *92*(9), 1786-1801. https://doi.org/10.1111/1365-2656.13960

Rinnan, D.S. (2021). CENFA: Climate and Ecological Niche Factor Analysis. *R. package ver 1.1.1*. https://CRAN.R-project.org/package=CENFA

Robinson, E. L., Jardine, C. M., Koffi, J. K., Russell, C., Lindsay, L. R., Dibernardo, A., & Clow, K. M. (2022). Range Expansion of *Ixodes scapularis* and *Borrelia burgdorferi* in Ontario, Canada, from 2017 to 2019. *Vector-Borne and Zoonotic Diseases*, *22*(7), 361–369. https://doi.org/10.1089/vbz.2022.0015

Rodgers, S. E., Zolnik, C. P., & Mather, T. N. (2007). Duration of Exposure to Suboptimal Atmospheric Moisture Affects Nymphal Blacklegged Tick Survival, *Journal of Medical Entomology*, *44*(2), 372–375. https://doi.org/10.1093/jmedent/44.2.372

Rosenthal, M., & Coburn, J. (2008). Lyme Disease. In H. Heggenhougen (Ed.), International Encyclopedia of Public Health (pp. 168–176). Academic Press. Retrieved from https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/ixodes-scapularis.

RStudio Team (2020). RStudio: Integrated Development for R. http://www.rstudio.com/

Sales, L. P., Hayward, M. W., & Loyola, R. (2021). What do you mean by "niche"? Modern ecological theories are not coherent on rhetoric about the niche concept. Acta Oecologica, 110, 103701. https://doi.org/10.1016/j.actao.2020.103701

Schoener, T. W. (1987). The Ecological Niche. In Cherrett, J. M., (Eds.) *Ecological concepts: the contribution of ecology to an understanding of the natural world* (pp. 79-113). Blackwell Scientific Publications.

Schulze, T. L., Jordan, R. A., Schulze, C. J., & Hung, R. W. (2009). Precipitation and temperature as predictors of the local abundance of *Ixodes scapularis* (acari: Ixodidae) nymphs. *Journal of Medical Entomology*, *46*(5), 1025–1029. https://doi.org/10.1603/033.046.0508

Schulze, T. L., Jordan, R. A., & Hung, R. W. (2001). Effects of selected meteorological factors on diurnal questing of *Ixodes scapularis* and *Amblyomma americanum* (Acari: Ixodidae). *Journal of Medical Entomology*, *38*(2), 318-324. https://doi.org/10.1603/0022-2585-38.2.318

Senay, S. D., Worner, S. P., & Ikeda, T. (2013). Novel three-step pseudo-absence selection technique for improved species distribution modelling. *PLOS One*, 8(8). https://doi.org/10.1371/journal.pone.0071218

Shen, Z., Zhou, W., Li, J., & Chan, J. C. (2023). A frequent ice-free Arctic is likely to occur before the mid-21st century. *Npj Climate and Atmospheric Science*, *6*(1), 1-7. https://doi.org/10.1038/s41612-023-00431-1

Sillero, N., & Barbosa, A. M. (2020). Common mistakes in ecological niche models. *International Journal of Geographical Information Science*, *35*(2), 213–226. https://doi.org/10.1080/13658816.2020.1798968

Slatculescu, A. M., Clow, K. M., McKay, R., Talbot, B., Logan, J. J., Thickstun, C. R., Jardine, C. M., Ogden, N. H., Knudby, A. J., & Kulkarni, M. A. (2020). Species distribution models for the eastern blacklegged tick, *Ixodes scapularis*, and the Lyme disease

pathogen, *Borrelia burgdorferi*, in Ontario, Canada. *PLOS One. 15*(9). https://doi.org/10.1371/journal.pone.0238126

Soberón, J., & Arroyo-Peña, B. (2017). Are fundamental niches larger than the realized? Testing a 50-year-old prediction by Hutchinson. *PLOS ONE*, *12*(4), e0175138. https://doi.org/10.1371/journal.pone.0175138

Soucy, J.-P. R., Slatculescu, A. M., Nyiraneza, C., Ogden, N. H., Leighton, P. A., Kerr, J. T., & Kulkarni, M. A. (2018). High-resolution ecological niche modeling of *Ixodes scapularis* ticks based on passive surveillance data at the northern frontier of Lyme disease emergence in North America. *Vector-Borne and Zoonotic Diseases*, *18*(5), 235–242. https://doi.org/10.1089/vbz.2017.2234

Stafford, K. C. (1994). Survival of immature *Ixodes scapularis* (Acari: Ixodidae) at different relative humidities. *Journal of Medical Entomology*, *31*(2), 310-314. https://doi.org/10.1093/jmedent/31.2.310

Steen, V. A., Tingley, M. W., C. Paton, P. W., & Elphick, C. S. (2021). Spatial thinning and class balancing: Key choices lead to variation in the performance of species distribution models with citizen science data. *Methods in Ecology and Evolution*, *12*(2), 216-226. https://doi.org/10.1111/2041-210X.13525

Talbot, B., Slatculescu, A., Thickstun, C.R., Koffi, J. K., Leighton, P. A., Mckay, R., & Kulkarni, M. A. (2019). Landscape determinants of density of blacklegged ticks, vectors of Lyme disease, at the northern edge of their distribution in Canada. *Sci Rep, 9*. https://doi.org/10.1038/s41598-019-50858-x

Tebaldi, C.,Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P. Knutti, R., Lowe, J., O'Neill, B., Sanderson, B., van Vuuren, D., Riahi, K., Meinshausen, M., Nicholls, Z.,

Tokarska, K. B., Hurtt, G., Kriegler, E., Lamarque, J.-F., Meehl, G., Moss, . . . Ziehn, T. (2021). Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. *Earth System Dynamics, 12*(1). 253-293. 10.5194/esd-12-253-2021

Templer, P. H., Schiller, A. F., Fuller, N. W., Socci, A. M., Campbell, J. L., Drake, J. E., & Kunz, T. H. (2012). Impact of a reduced winter snowpack on litter arthropod abundance and diversity in a northern hardwood forest ecosystem. *Biology and Fertility of Soils*, *48*(4). 413–424. https://doi.org/10.1007/s00374-011-0636-3

Thuiller, W., Georges, D., Gueguen, M., Engler., R, Breiner., F., Lafourcade, B., & Patin, R. (2022). Biomod2: Ensemble platform for species distribution modeling. *R.package version 4.1-2*. https://CRAN.R-project.org/package=biomod2.

Thuiller, W., Guéguen, M., Renaud, J., Karger, D. N., & Zimmermann, N. E. (2019). Uncertainty in ensembles of global biodiversity scenarios. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09519-w

Tobler, W. R. (1970) A computer movie simulating urban growth in the Detroit region, *Economic Geography, 46,* 234-240, DOI: 10.2307/143141

Tourinho, L., & Vale, M. (2022). Choosing among correlative, mechanistic, and hybrid models of species' niche and distribution. *Integrative Zoology*, *18*(1), 93–109. https://doi.org/10.1111/1749-4877.12618

VanDerWal, J., Shoo, L. P., Graham, C., & Williams, S. E. (2009). Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, *220*(4), 589–594. https://doi.org/10.1016/j.ecolmodel.2008.11.010

Wang, Y., Hogg, E. H., Price, D. T., Edwards, J., & Williamson, T. (2014). Past and projected future changes in moisture conditions in the Canadian boreal forest. *The Forestry Chronicle, 90*(05), 678–691. https://doi.org/10.5558/tfc2014-134

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2016). Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLOS One*, *11*(6). https://doi.org/10.1371/journal.pone.0156720

Wang, T., Hamann, A., Spittlehouse, D., & Carroll, C. (2022). ClimateNA (Version 7.30). https://climatena.ca/

Wei, T., & Simko, V. (2021). Corrplot: Visualization of a correlation matrix. *R. package version 0.92*. https://github.com/taiyun/corrplot

Wickham, H., François, R., Henry, L., & Müller, K. (2022). Dplyr: A grammar of data manipulation. *R.package version 1.0.10.* https://CRAN.R-project.org/package=dplyr.

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*(1), 1–29. https://doi.org/10.18637/jss.v040.i01

Wikel, S. K. (2018). Ticks and tick-borne infections: complex ecology, agents, and host interactions. *Veterinary Sciences*, *5*(2), 60. https://doi.org/10.3390/vetsci5020060

Wisz, M. S., & Guisan, A. (2009). Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, *9*(1). https://doi.org/10.1186/1472-6785-9-8

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061

Zhang, X., Liu, L., Wu, C., Chen, X., Gao, Y., Xie, S., & Zhang, B. (2020). Development of a global 30 m impervious surface map using multisource and multitemporal remote sensing datasets with the Google Earth Engine Platform. *Earth System Science Data*, *12*(3), 1625–1648. https://doi.org/10.5194/essd-12-1625-2020

Zhang, L., Liu, S., Sun, P., Wang, T., Wang, G., Zhang, X., & Wang, L. (2015). Consensus Forecasting of Species Distributions: The Effects of Niche Model Performance and Niche Properties. *PLOS ONE*, *10*(3), e0120056. https://doi.org/10.1371/journal.pone.0120056

Zhu, G., Fan, J., & Peterson, A. T. (2021). Cautions in weighting individual ecological niche models in ensemble forecasting. *Ecological Modelling*, *448*. https://doi.org/10.1016/j.ecolmodel.2021.109502

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

Appendix A

Complementary (AdaptWest) Ensembles (SSP5-8.5) and Variable Importance

SSP5-8.5 ensembles produced using environmental data from complementary dataset, AdaptWest. Each ~1 km cell received a predictive value between 0-1. Zero indicates no suitability or likelihood of *I. scapularis*; one indicates the highest suitability or likelihood of occurrence. Additional classification breaks down *I. scapularis* suitability as follows: unsuitable niche (white: 0-0.1), low-suitable niche (blue: 0.1-0.3), medium suitable niche (green: 0.3-0.6), high-suitable niche (orange: 0.6-0.8), optimally suitable niche (red: 0.8-1.0)

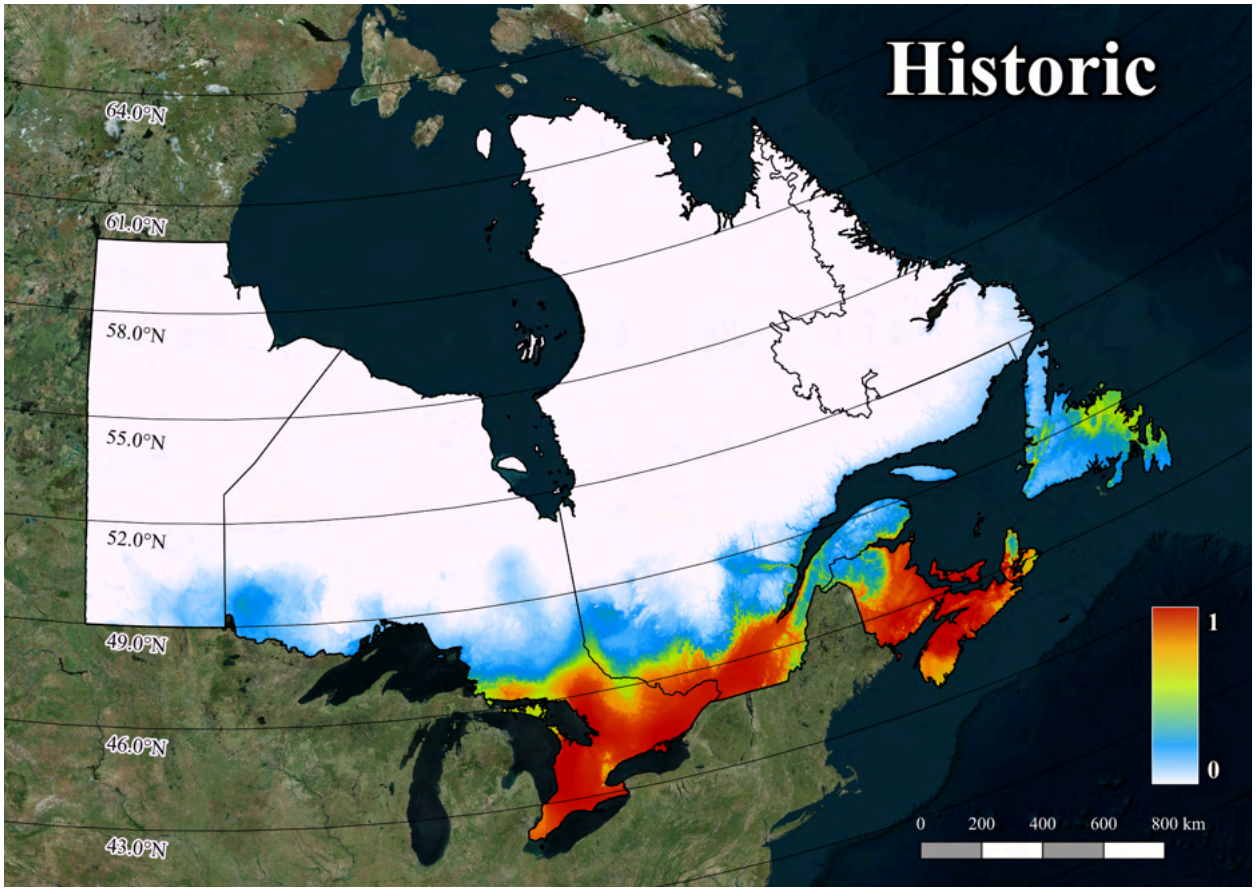


Figure A1: Contemporary ensemble of the AdaptWest dataset

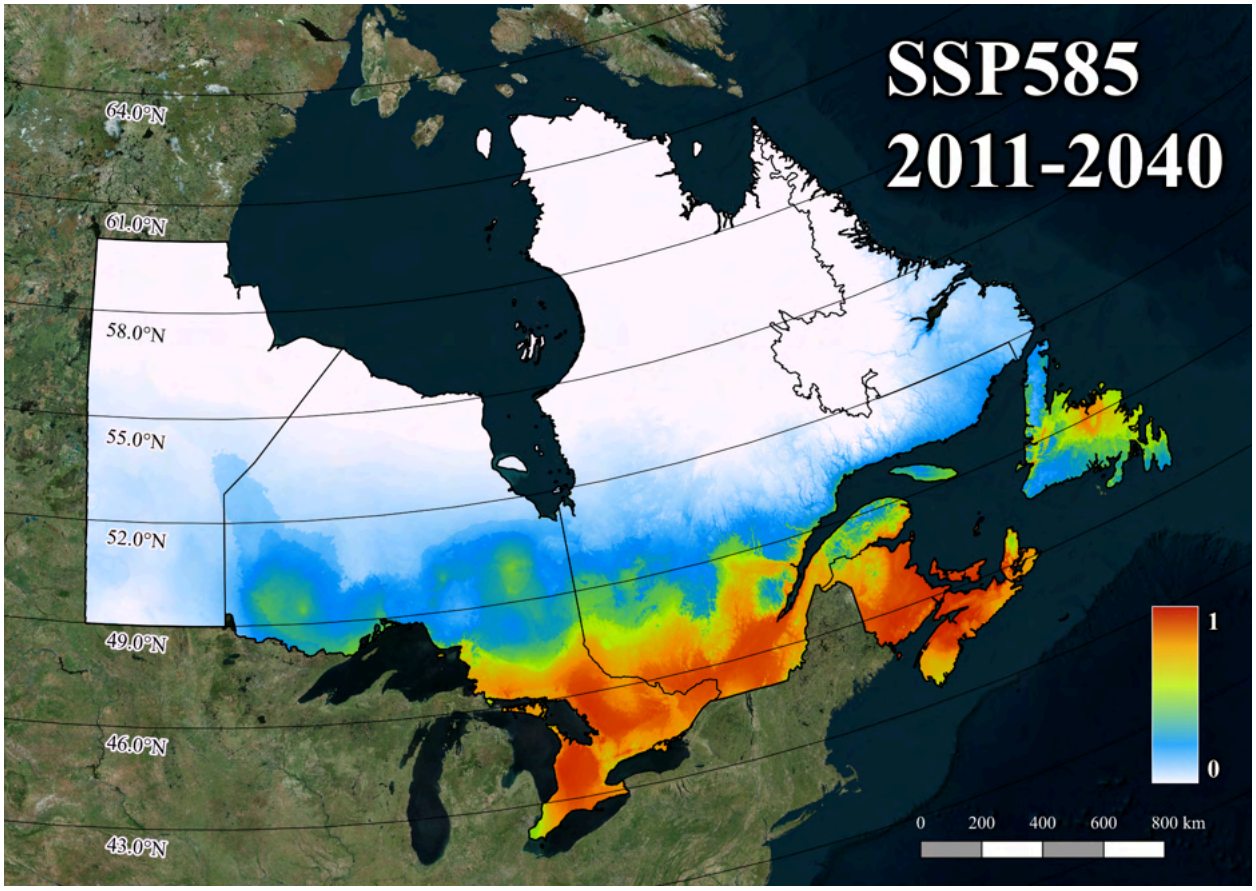


Figure A2: Ensemble produced for 2011-2040 using Adaptwest dataset and SSP5-8.5 environmental data.

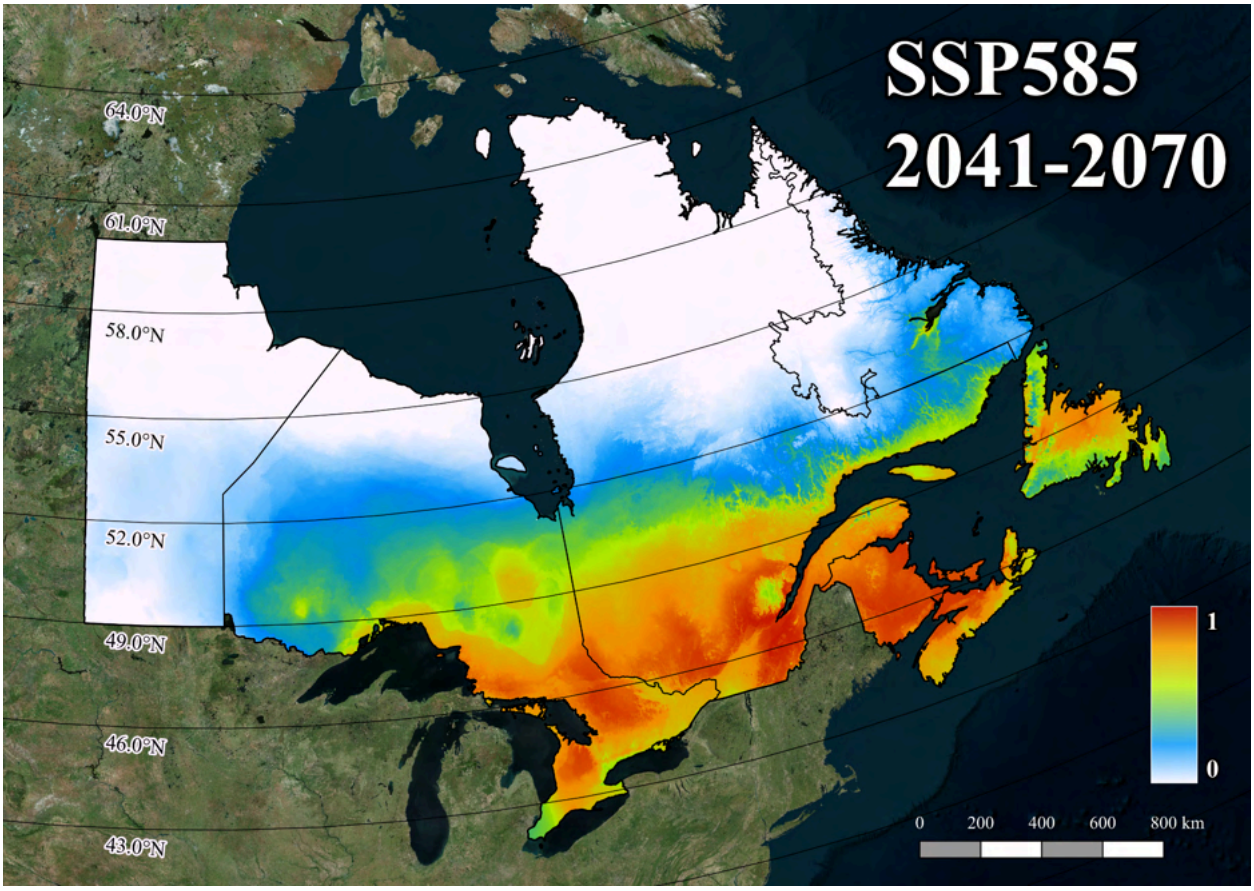


Figure A3: Ensemble produced for 2041-2070 using AdaptWest dataset and SSP5-8.5 environmental data.

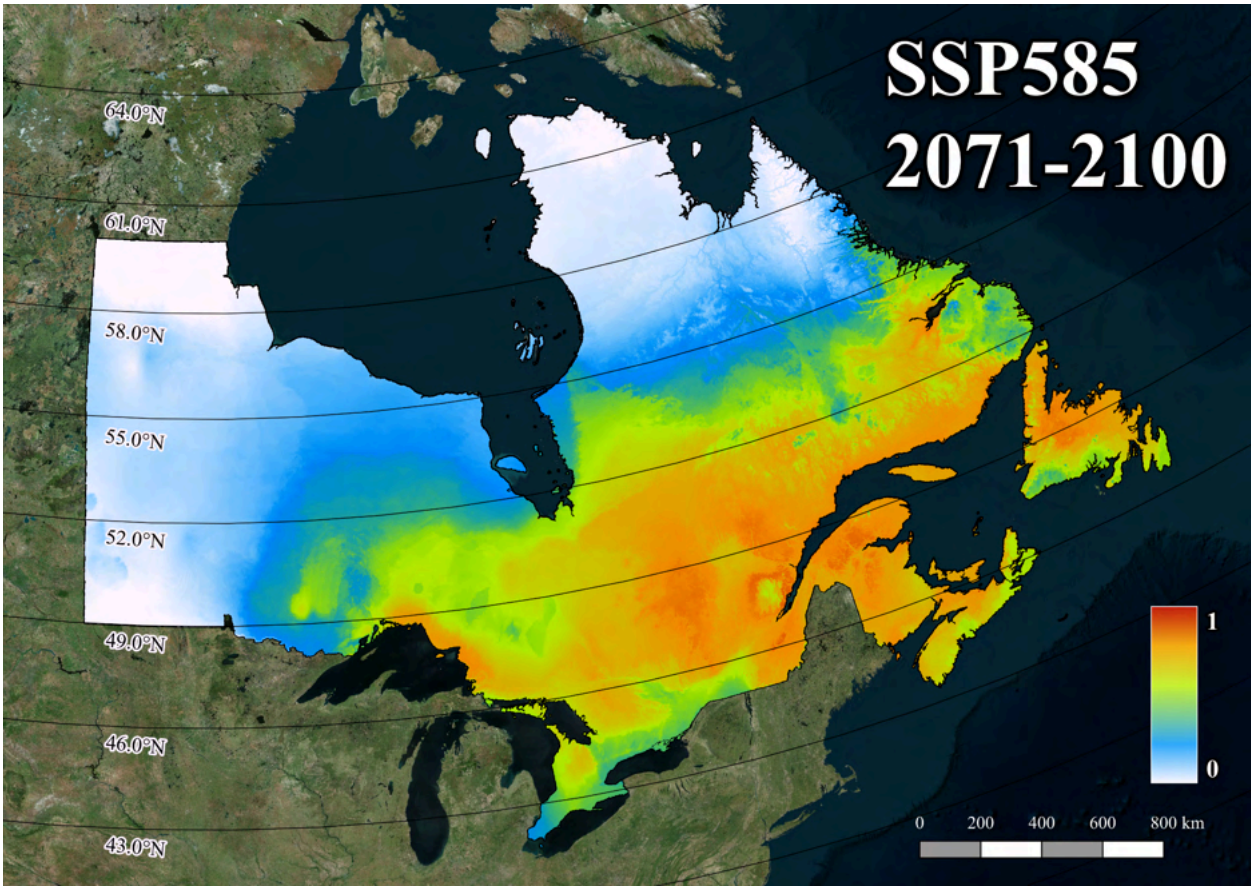


Figure A4: Ensemble produced for 2071-2100 using AdaptWest dataset and SSP5-8.5 environmental data.

Table A1: Environmental variable permutations in the complementary (AdaptWest) dataset ensembles.

	AdaptWest Variables		
Variable	Abbrev	Relative Sequence Variable Importance	Ensembled Variable Importance
Degree Days Above 5°C	DD5	54.84%	10.00%
Hogg's Climate Moisture Index	CMI	52.84%	9.63%
Mean Annual Temperature	MAT	52.09%	9.50%
Mean Annual Precipitation	MAP	49.93%	9.10%
Winter Precipitation	PPT_wt	42.70%	7.79%
Degree Days Above 18°C	DD18	39.99%	7.29%
Degree Days Between 10°C and 40°C	DD1040	36.27%	6.61%
Precipitation as Snow	PAS	34.26%	6.25%
Degree Days Below °0C	DD_0	34.17%	6.23%
Mean Extreme Maximum Temperature Over Past 30-Years	EXT	28.15%	5.13%
Annual Heat Moisture	AHM	26.22%	4.78%
Spring Precipitation	PPT_sp	24.69%	4.50%
Autumn Precipitation	PPT_at	16.63%	3.03%
Summer Heat Moisture	SHM	16.39%	2.99%
Temperature Difference between Warmest and Coldest Month (i.e., Measure of Continentality)	TD	10.45%	1.91%
Summer Precip	PPT_sm	9.04%	1.65%
Hargreaves Climate Moisture Index	CMD	8.47%	1.54%
Annual Relative Humidity	RH	7.72%	1.41%
Elevation	DEM	3.57%	0.64%
	Total:	548.42%	100%

Appendix B

Primary Ensembles using SSP3-7.0

SSP3-7.0 (i.e., medium-high scenario) ensembles produced using environmental data from primary dataset, Chelsa-Bioclim+. Each ~1 km cell received a predictive value between 0-1. Zero indicates no suitability or likelihood of *I. scapularis*; one indicates the highest suitability or likelihood of occurrence. Additional classification breaks down *I. scapularis* suitability as follows: unsuitable niche (white: 0-0.1), low-suitable niche (blue: 0.1-0.3), medium suitable niche (green: 0.3-0.6), high-suitable niche (orange: 0.6-0.8), optimally suitable niche (red: 0.8-1.0)

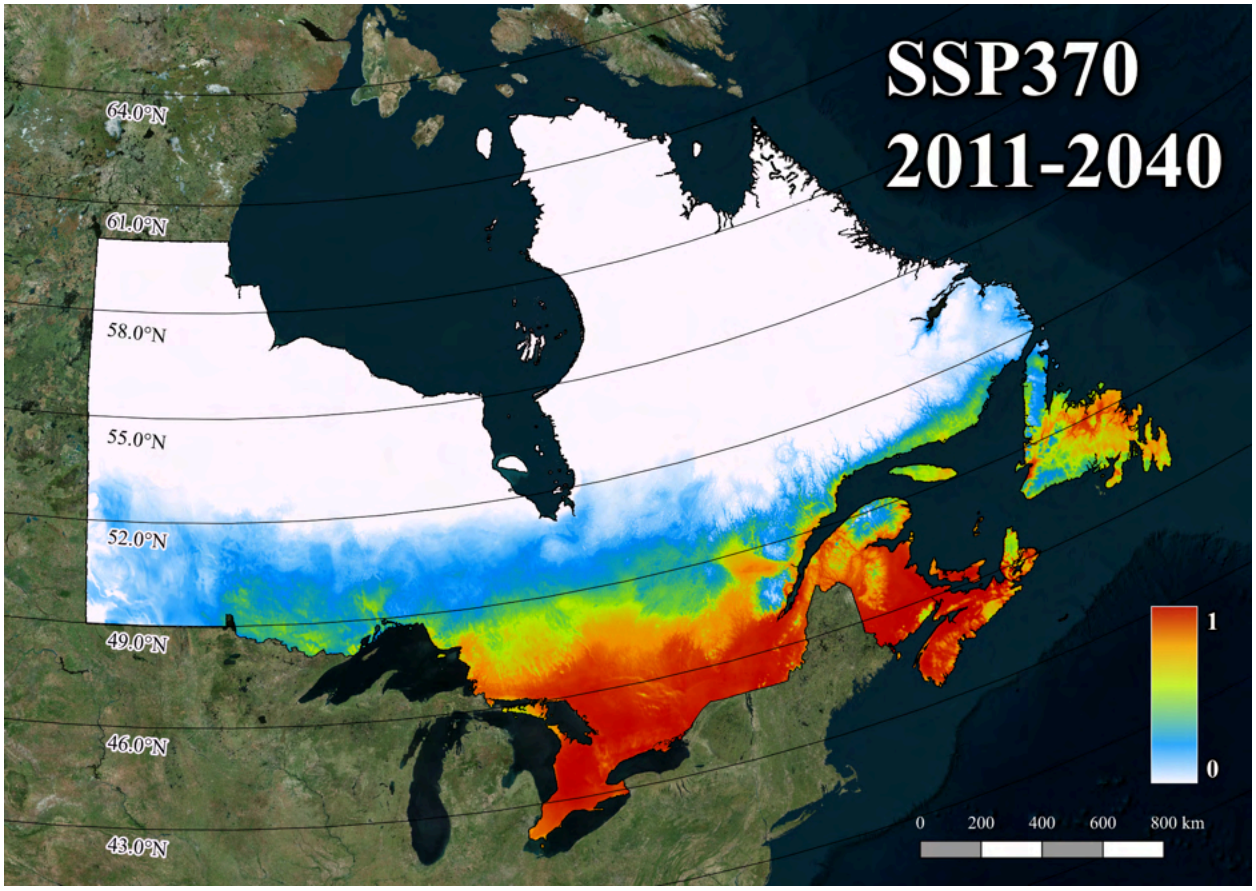


Figure B1: Ensemble produced for 2011-2040 using primary dataset and SSP3-7.0 environmental data.

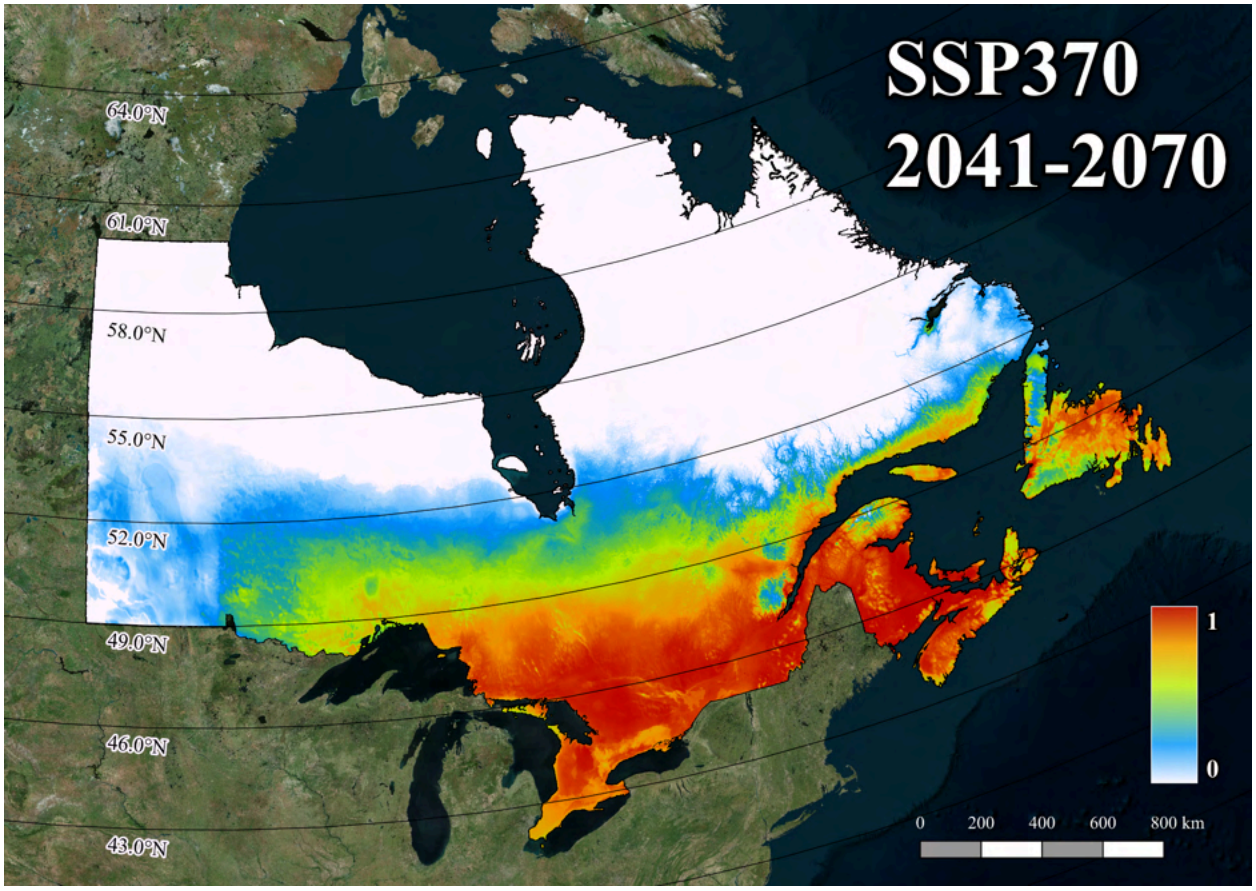


Figure B2: Ensemble produced for 2041-2070 using primary dataset and SSP3-7.0 environmental data.

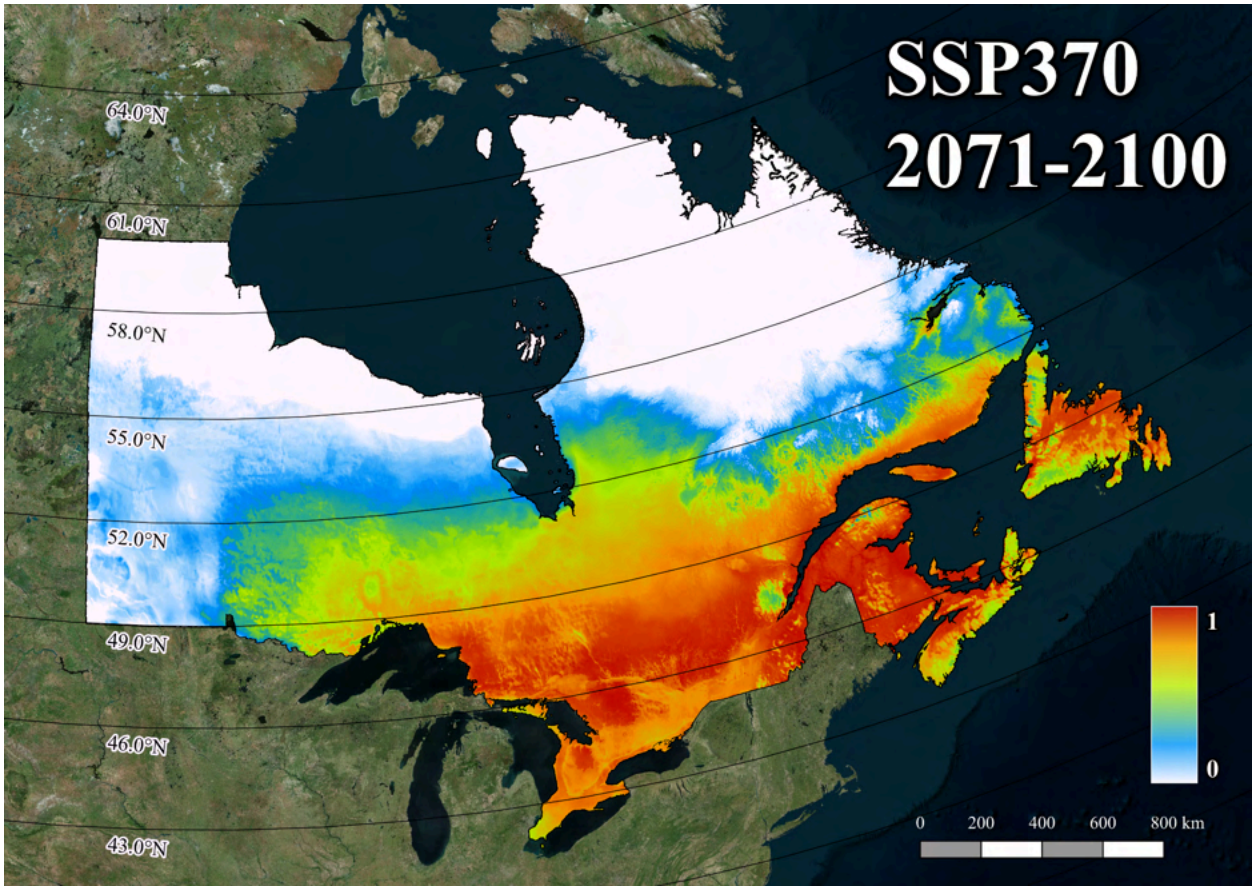


Figure B3: Ensemble produced for 2041-2070 using primary dataset and SSP3-7.0 environmental data.

Appendix C

SSP3.7-0 (i.e., medium-high scenario) ensembles produced using environmental data from complementary dataset, AdaptWest. Each ~1 km cell received a predictive value between 0-1. Zero indicates no suitability or likelihood of *I. scapularis*; one indicates the highest suitability or likelihood of occurrence. Additional classification breaks down *I. scapularis* suitability as follows: unsuitable niche (white: 0-0.1), low-suitable niche (blue: 0.1-0.3), medium suitable niche (green: 0.3-0.6), high-suitable niche (orange: 0.6-0.8), optimally suitable niche (red: 0.8-1.0)

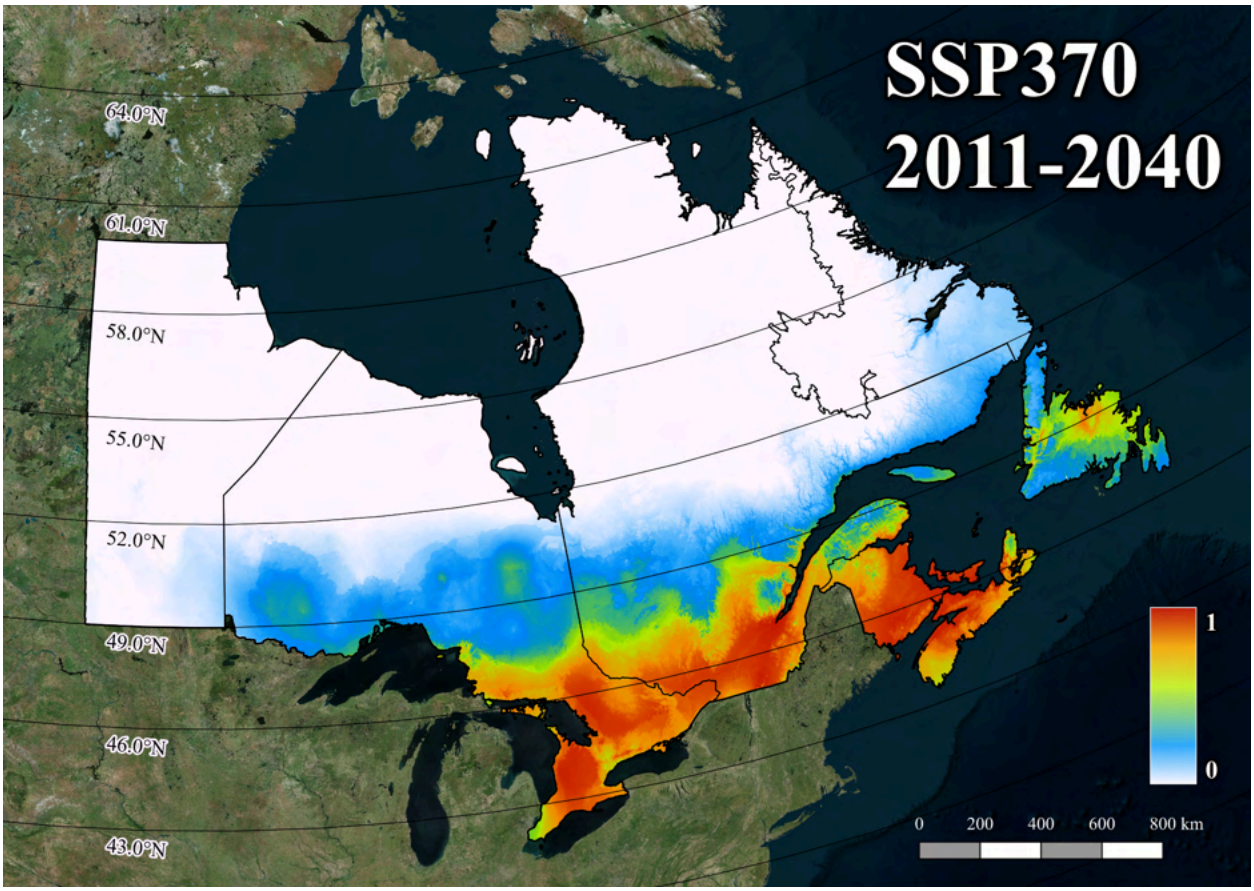


Figure C1: Ensemble produced for 2011-2040 using AdaptWest dataset and SSP3-7.0 environmental data.

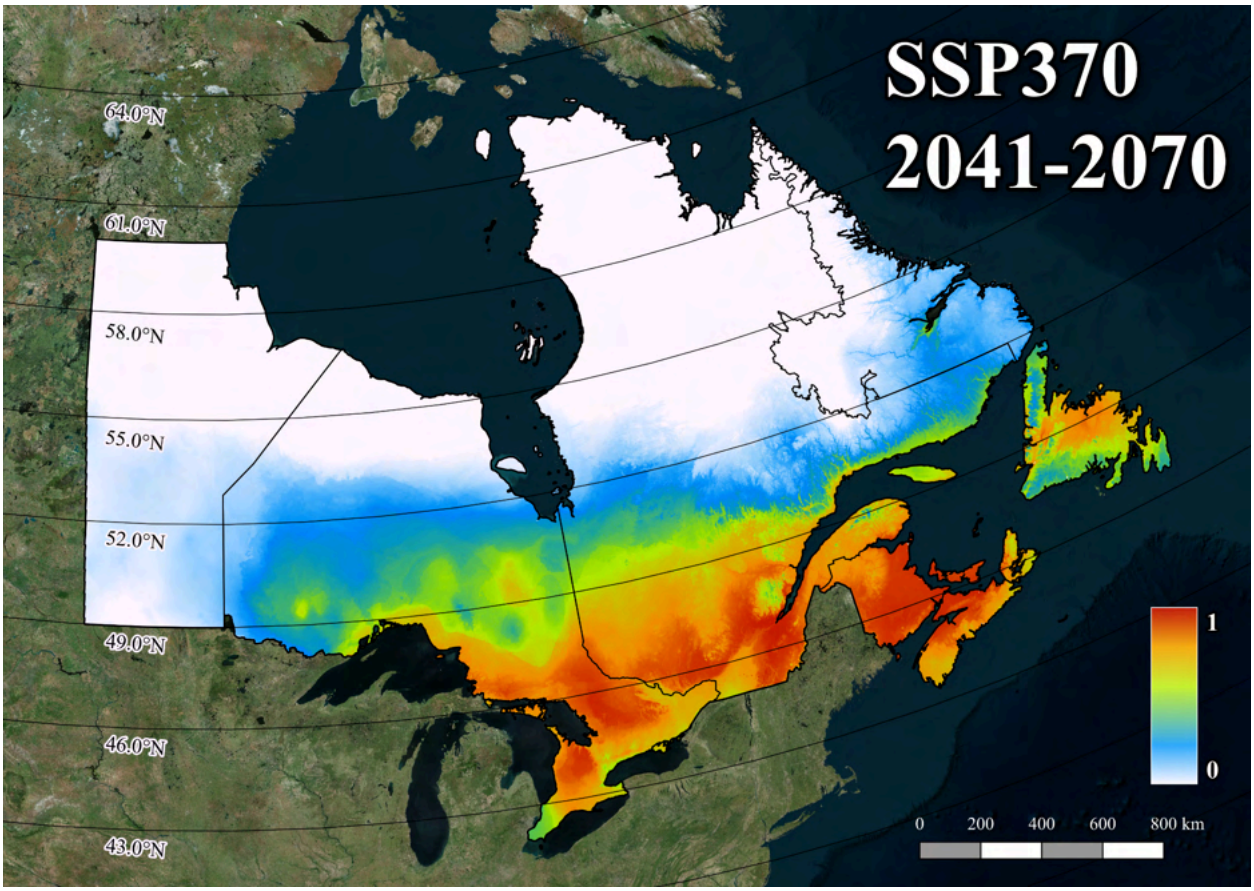


Figure C2: Ensemble produced for 2041-2070 using AdaptWest dataset and SSP3-7.0 environmental data.

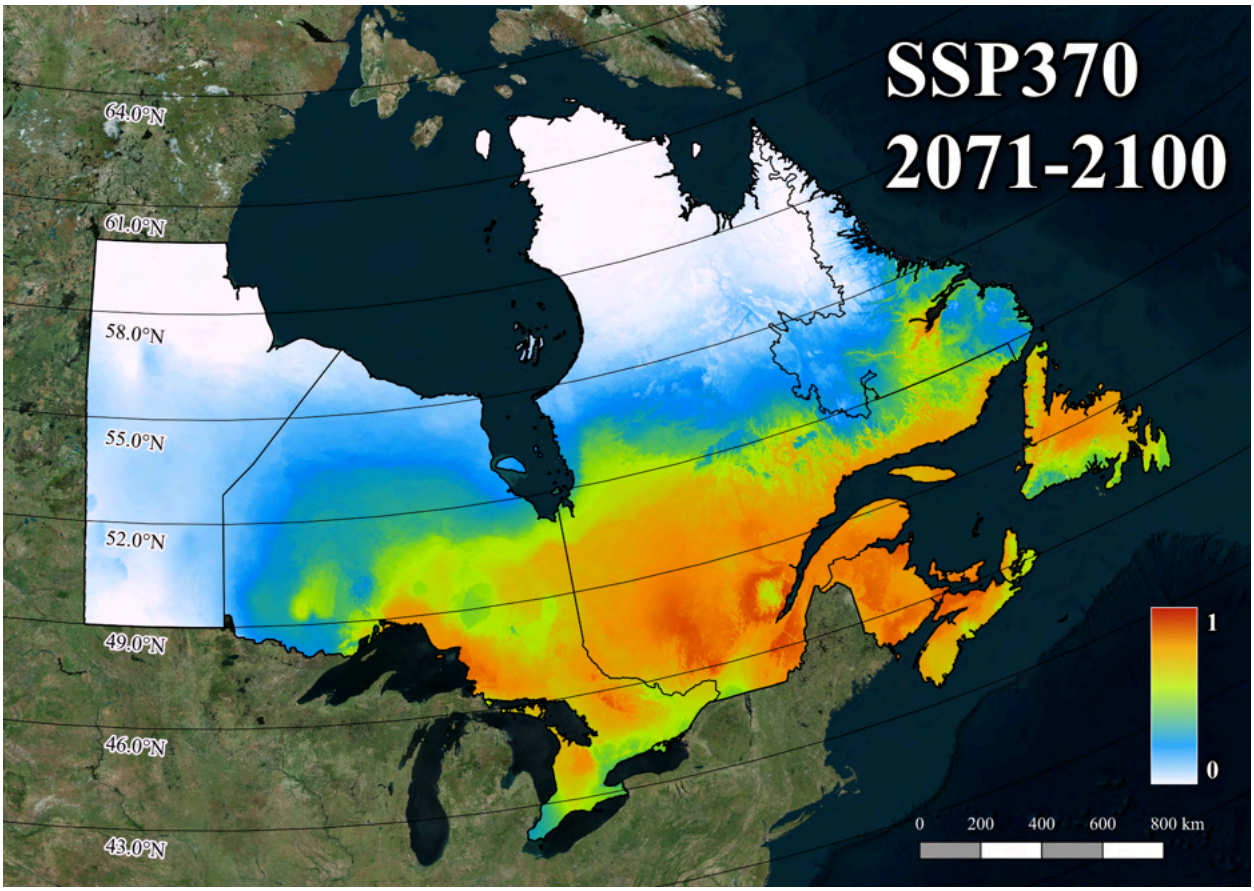
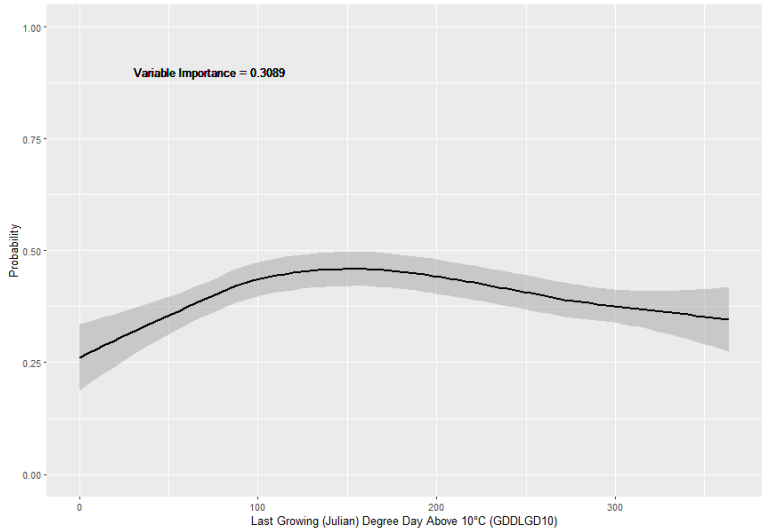


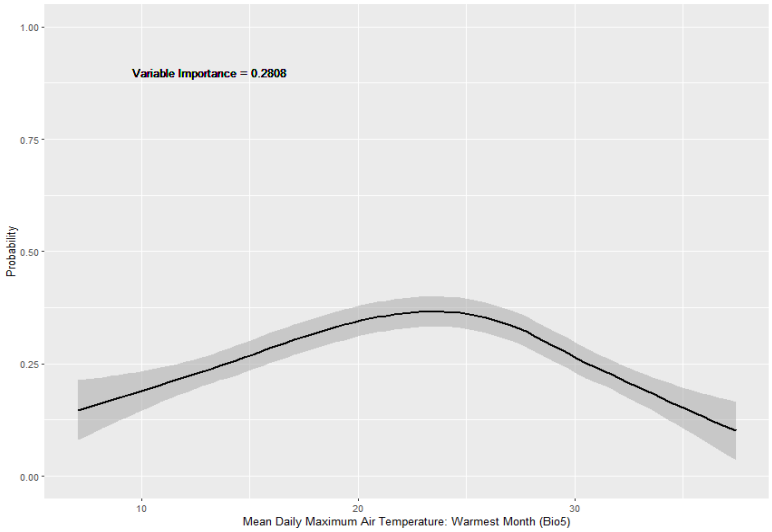
Figure C3: Ensemble produced for 2071-2100 using AdaptWest dataset and SSP3-7.0 environmental data.

Appendix D

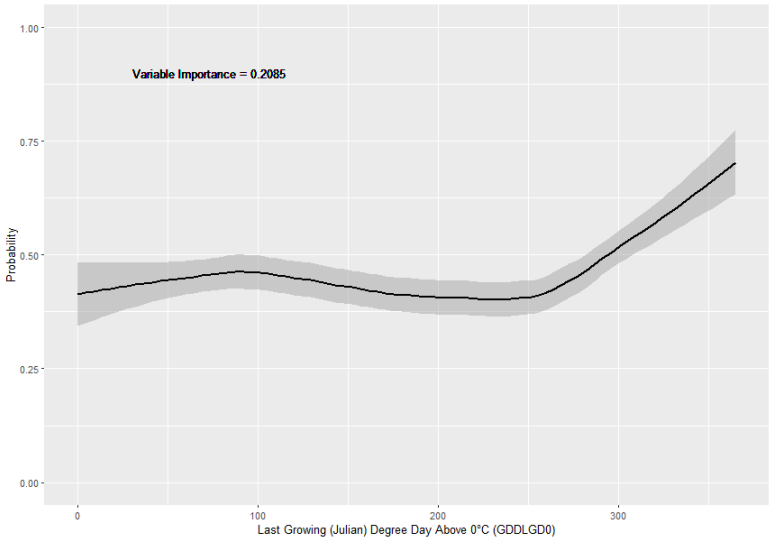
Remaining Primary Models Response Curves (and relative variable importance)

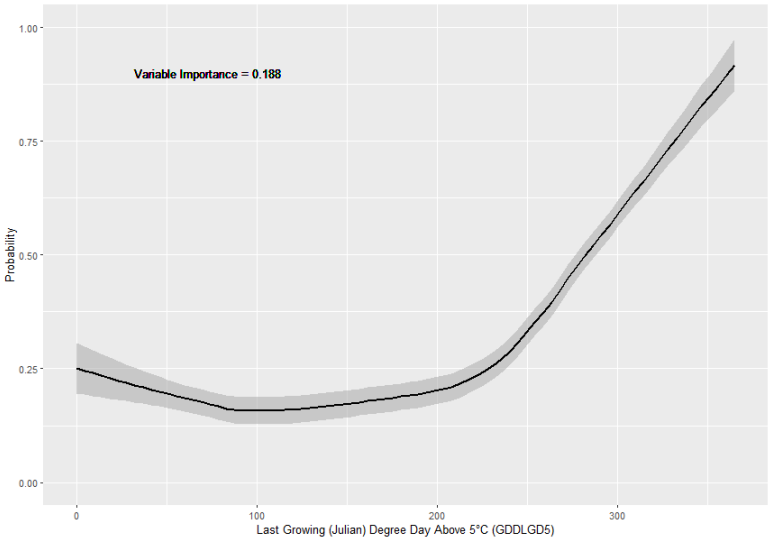
The remaining graphs that were not shown in the main text predict or forecast the response of *I. scapularis* to changes in stimulus conditions within each sequence of the primary dataset in descending order of importance.

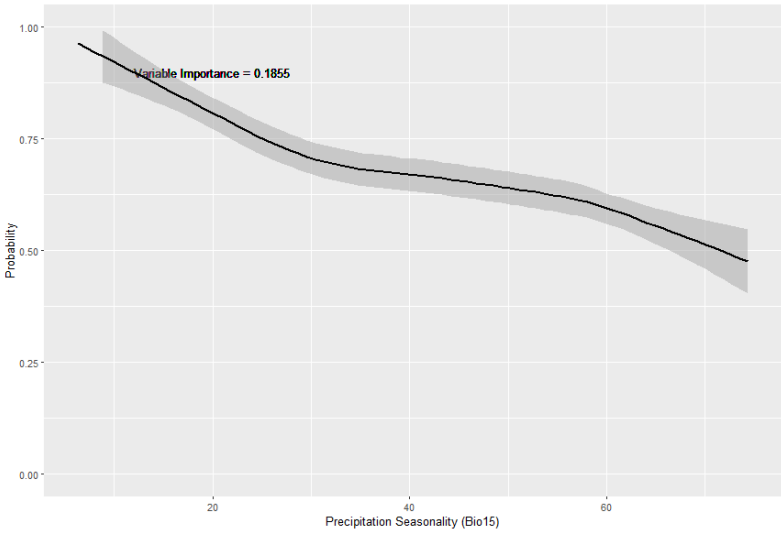

Variable Importance = 0.3089
Probability
1.00
0.75
0.50
0.25
0.00
0
100
200
300
Last Growing (Julian) Degree Day Above 10°C (GDDLGD10)

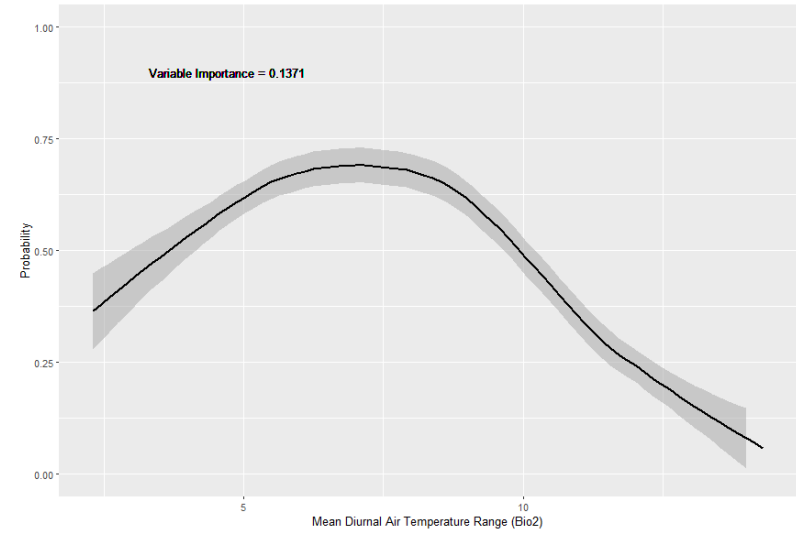

Variable Importance = 0.2808
Probability
1.00
0.75
0.50
0.25
0.00
10
20
30
Mean Daily Maximum Air Temperature: Warmest Month (Bio5)

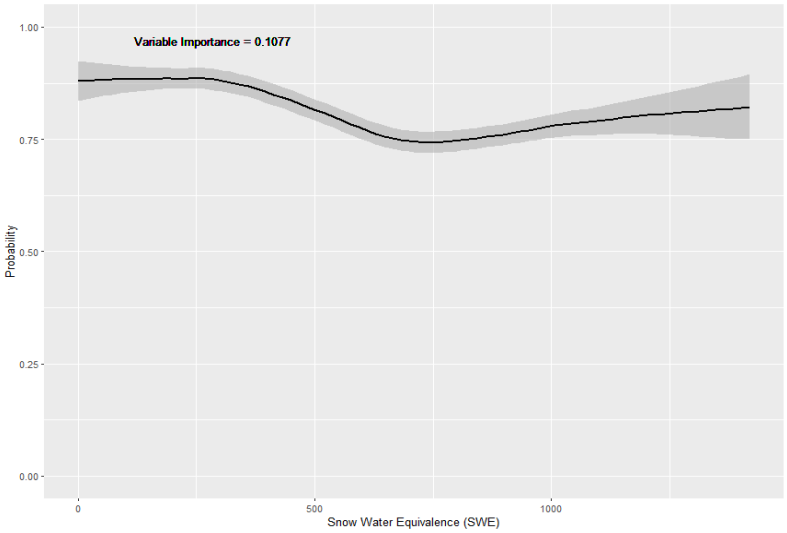

Variable Importance = 0.2536
Probability
1.00
0.75
0.50
0.25
0.00
0
500
1000
1500
2000
Growing Season Precipitation (GSP)

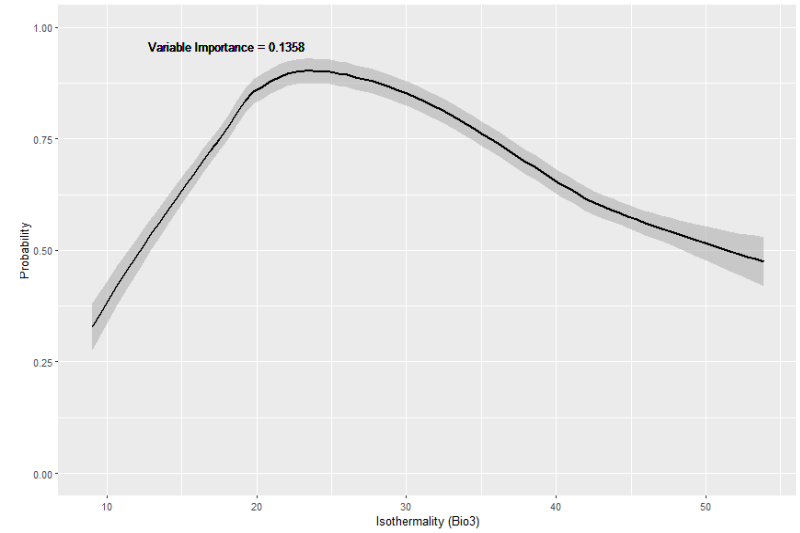

Variable Importance = 0.2085
Probability
1.00
0.75
0.50
0.25
0.00
0
100
200
300
Last Growing (Julian) Degree Day Above 0°C (GDDLGD0)

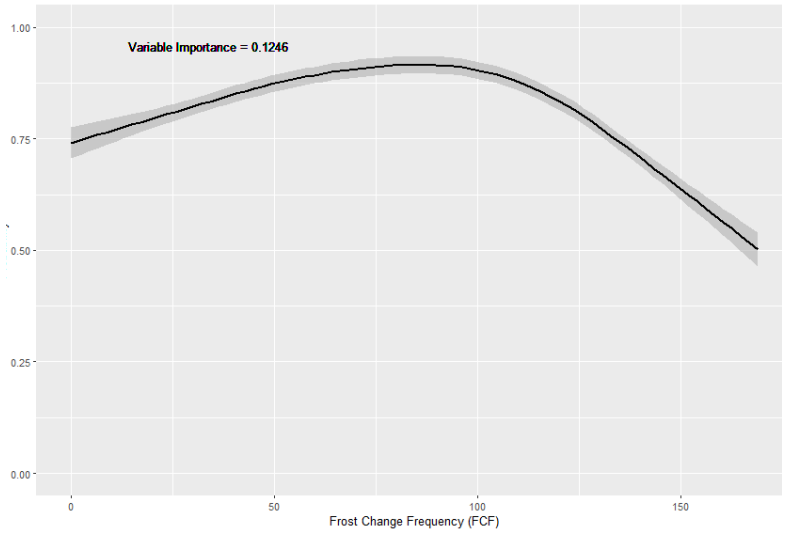

Variable Importance = 0.188
Probability
1.00
0.75
0.50
0.25
0.00
0
100
200
300
Last Growing (Julian) Degree Day Above 5°C (GDDLGD5)

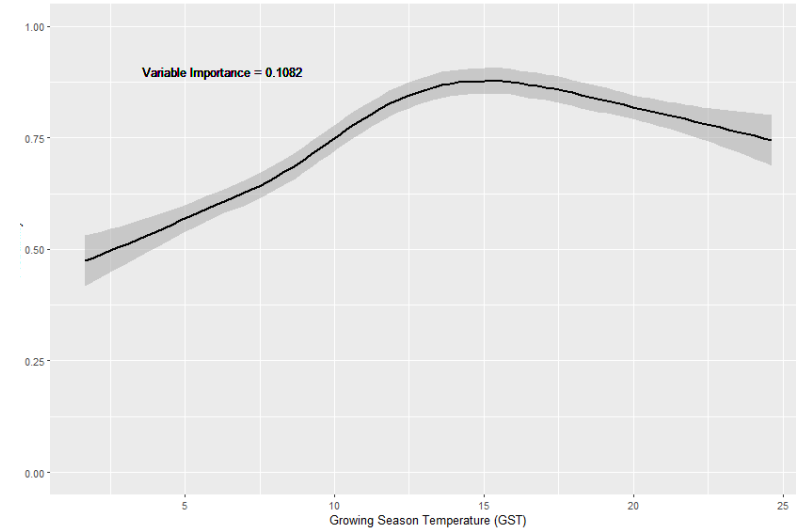

Variable Importance = 0.1855
Probability
1.00
0.75
0.50
0.25
0.00
20
40
60
Precipitation Seasonality (Bio15)

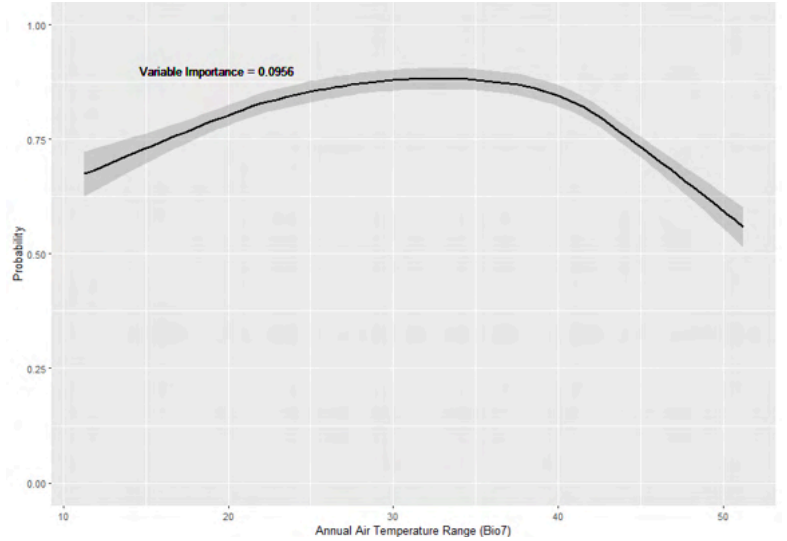

Variable Importance = 0.1371
Probability
1.00
0.75
0.50
0.25
0.00
5
10
Mean Diurnal Air Temperature Range (Bio2)

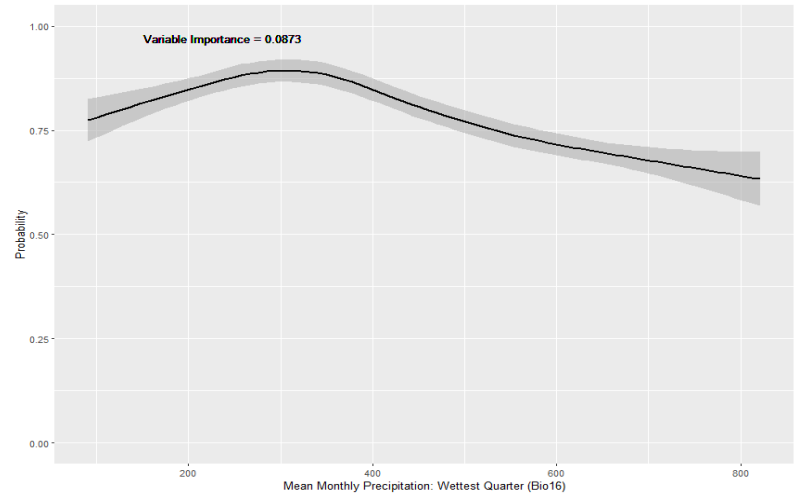

Variable Importance = 0.1077
Probability
1.00
0.75
0.50
0.25
0.00
0
500
1000
Snow Water Equivalence (SWE)

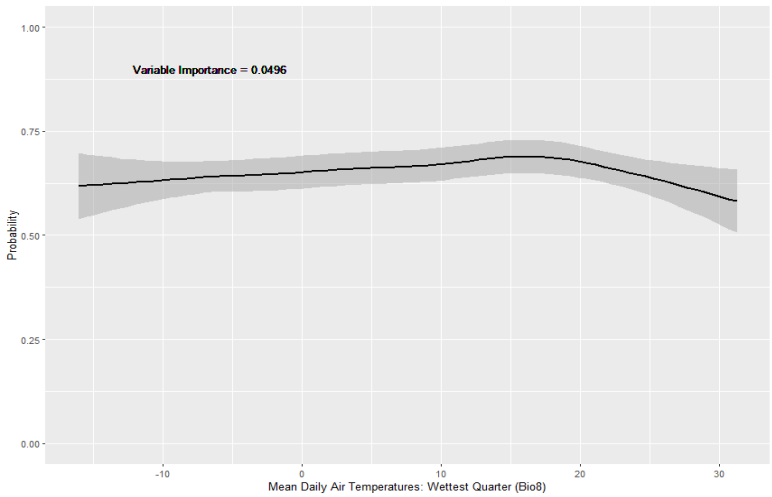

Variable Importance = 0.1358
Probability
1.00
0.75
0.50
0.25
0.00
10
20
30
40
50
Isothermality (Bio3)


Variable Importance = 0.1246
1.00
0.75
0.50
0.25
0.00
0
50
100
150
Frost Change Frequency (FCF)


Variable Importance = 0.1082
1.00
0.75
0.50
0.25
0.00
5
10
15
20
25
Growing Season Temperature (GST)


Variable Importance = 0.0956
Probability
1.00
0.75
0.50
0.25
0.00
10
20
30
40
50
Annual Air Temperature Range (Bio7)

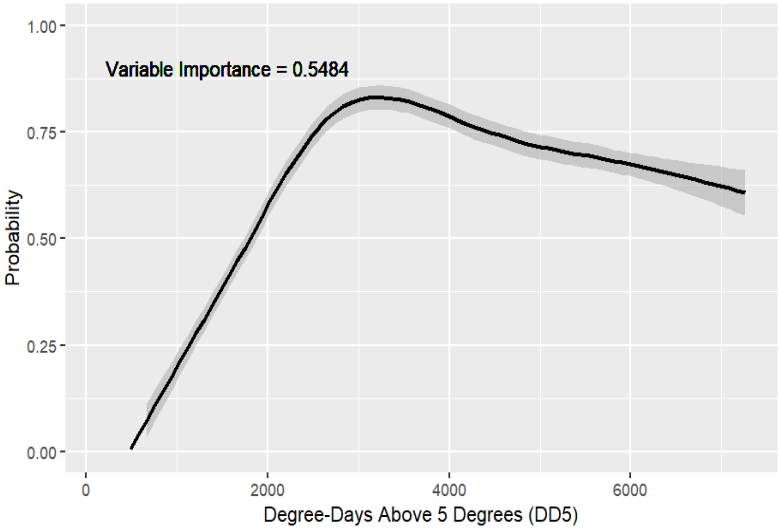

Variable Importance = 0.0873
1.00
0.75
0.50
0.25
0.00
Probability
200
400
600
800
Mean Monthly Precipitation: Wettest Quarter (Bio16)

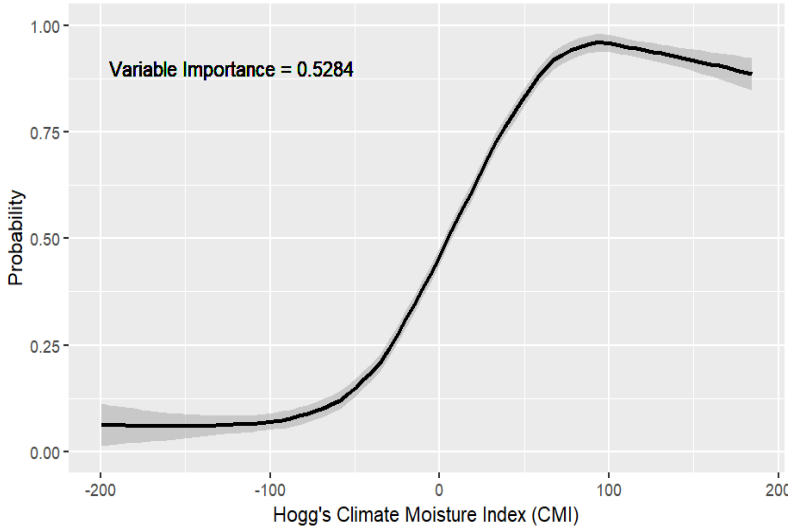

Variable Importance = 0.0496
1.00
0.75
0.50
0.25
0.00
Probability
-10
0
10
20
30
Mean Daily Air Temperatures: Wettest Quarter (Bio8)

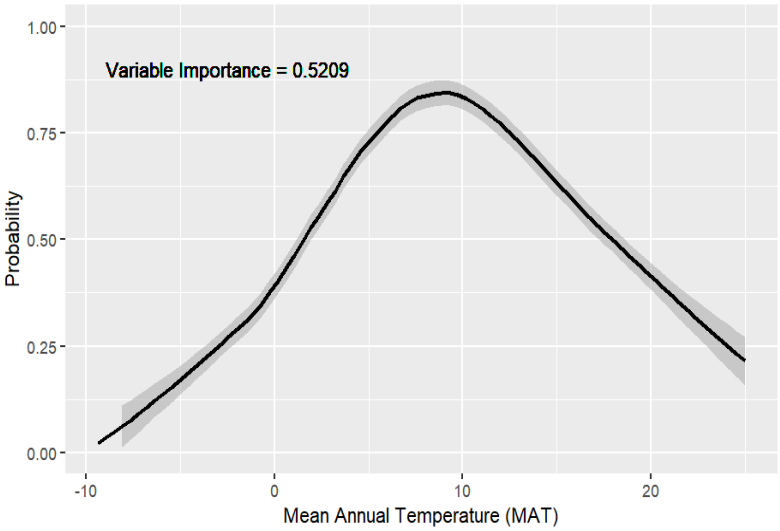
Appendix E

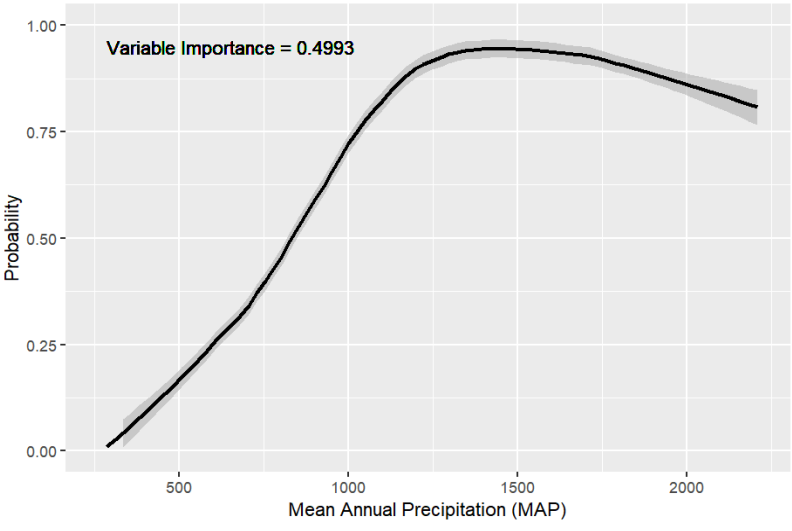
Complementary (AdaptWest) Models Response Curves (and relative variable importance)

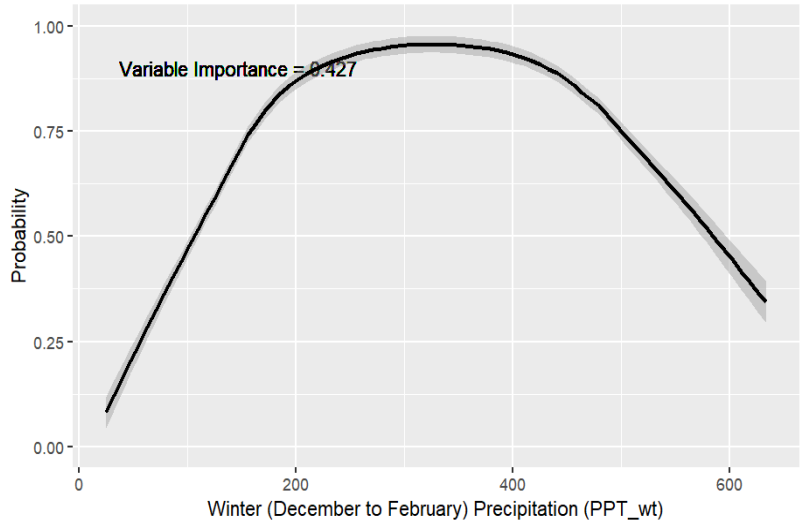
The graphs predict or forecast the response of *I. scapularis* to changes in stimulus conditions within each sequence of the secondary dataset in descending order of importance.

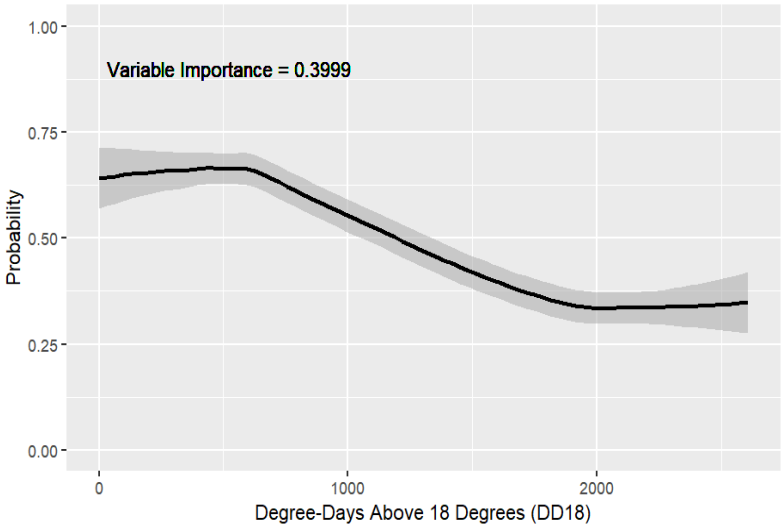


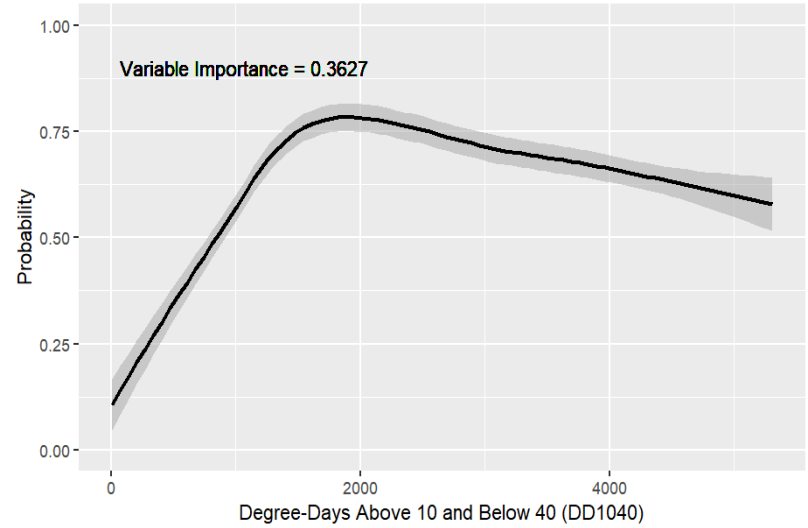


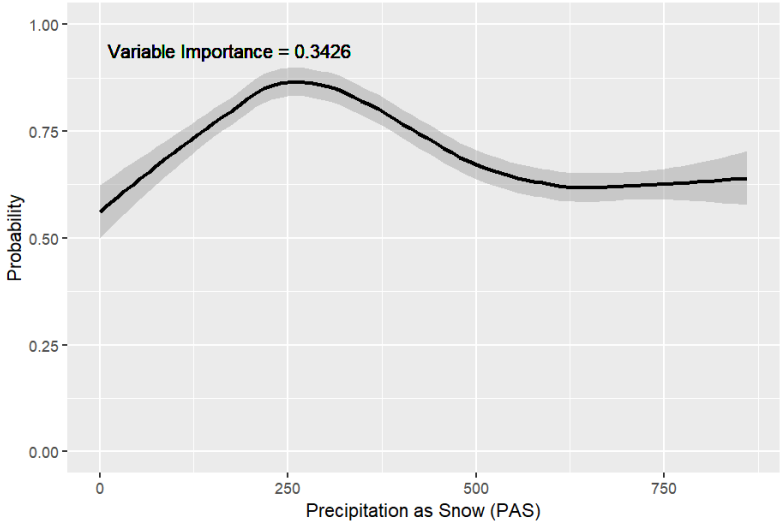


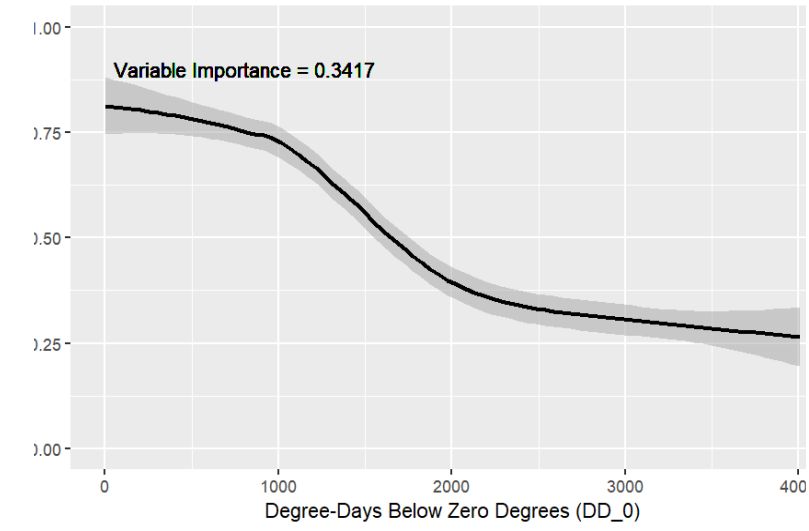


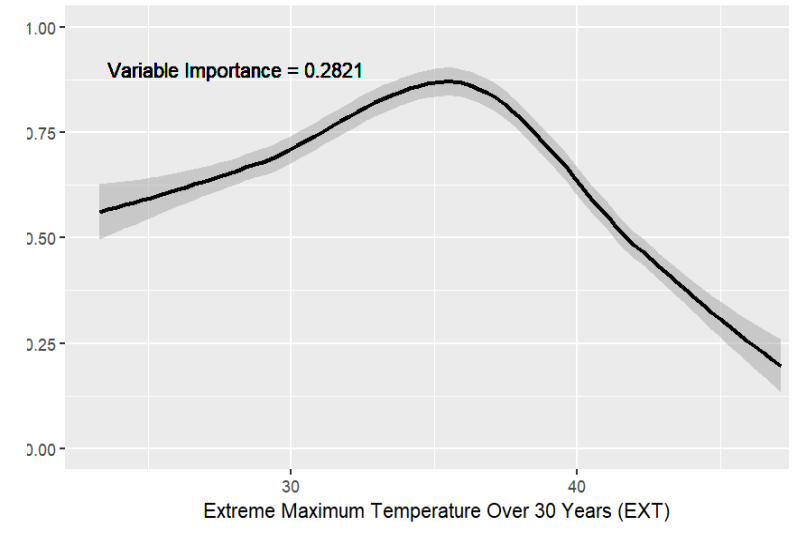

Variable Importance = 0.427
Probability
1.00
0.75
0.50
0.25
0.00
0
200
400
600
Winter (December to February) Precipitation (PPT_wt)

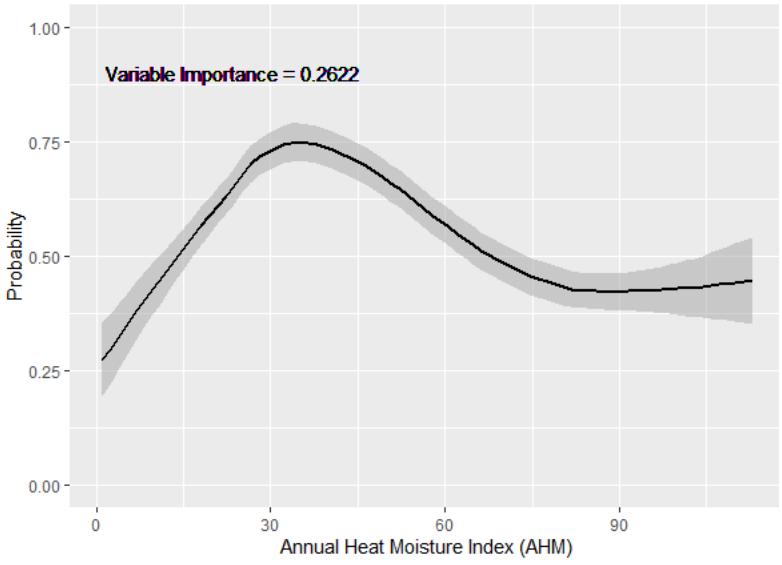

Variable Importance = 0.3999
Probability
1.00
0.75
0.50
0.25
0.00
0
1000
2000
Degree-Days Above 18 Degrees (DD18)

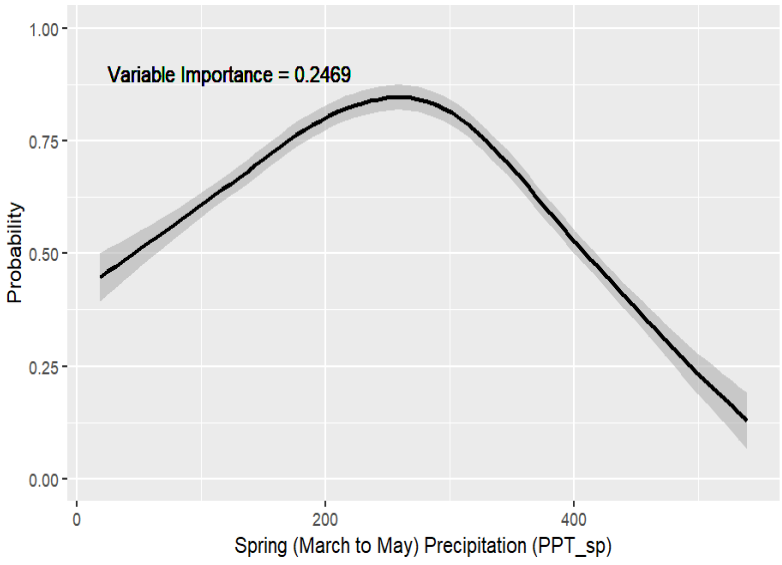

Variable Importance = 0.3627
Probability
1.00
0.75
0.50
0.25
0.00
0
2000
4000
Degree-Days Above 10 and Below 40 (DD1040)

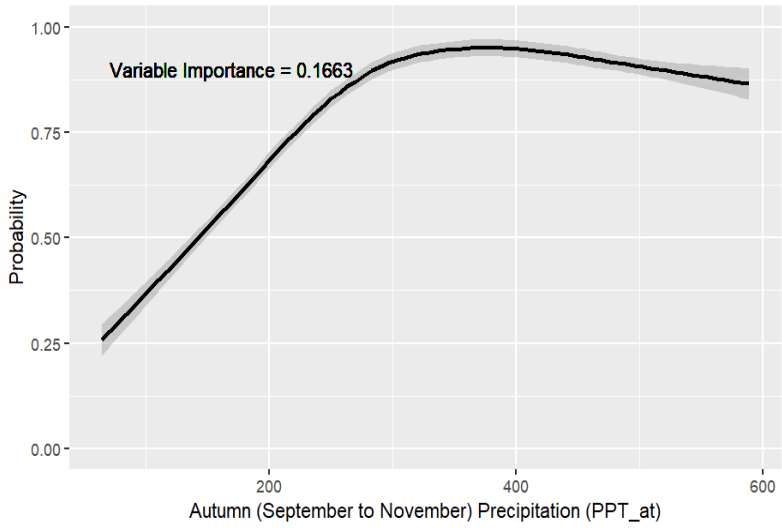

Variable Importance = 0.3426
Probability
1.00
0.75
0.50
0.25
0.00
0
250
500
750
Precipitation as Snow (PAS)

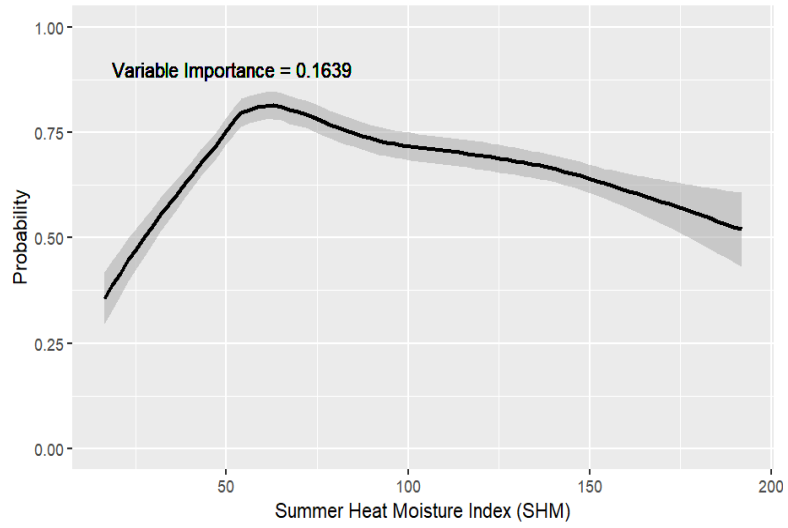

Variable Importance = 0.3417
1.00
0.75
0.50
0.25
0.00
0
1000
2000
3000
400
Degree-Days Below Zero Degrees (DD_0)

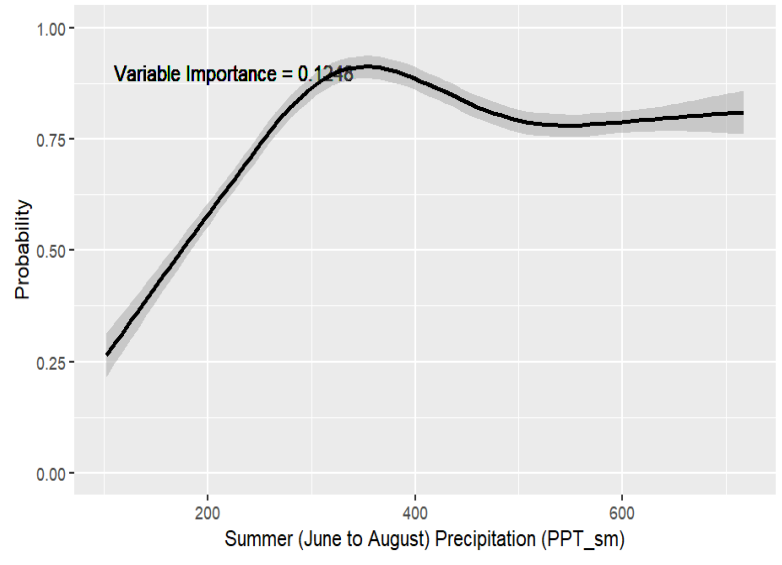

Variable Importance = 0.2821
1.00
0.75
0.50
0.25
0.00
30
40
Extreme Maximum Temperature Over 30 Years (EXT)

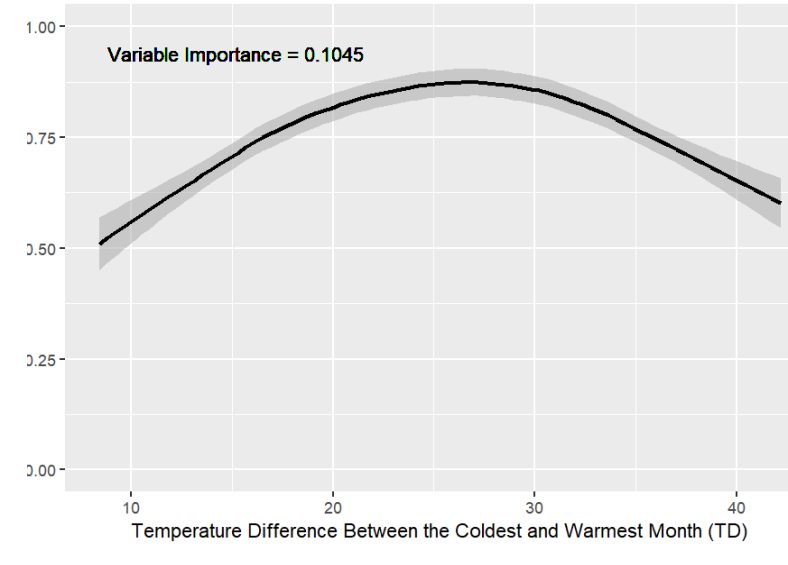

Variable Importance = 0.2622
Probability
1.00
0.75
0.50
0.25
0.00
0
30
60
90
Annual Heat Moisture Index (AHM)

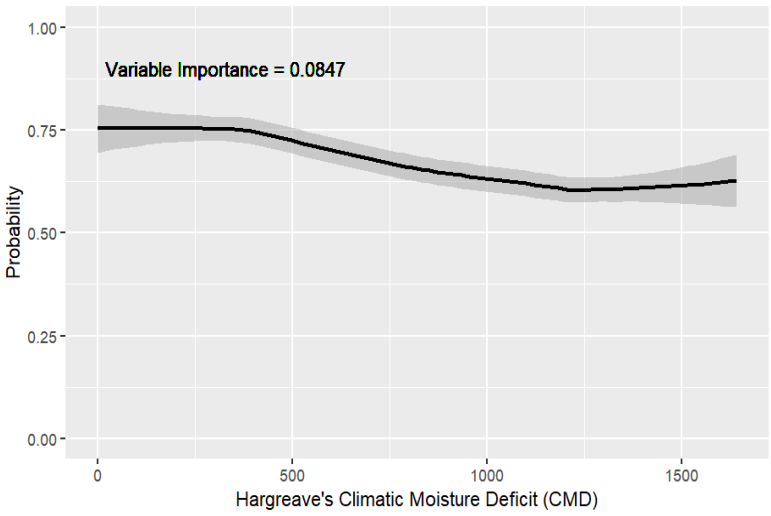

Variable Importance = 0.2469
Probability
1.00
0.75
0.50
0.25
0.00
0
200
400
Spring (March to May) Precipitation (PPT_sp)

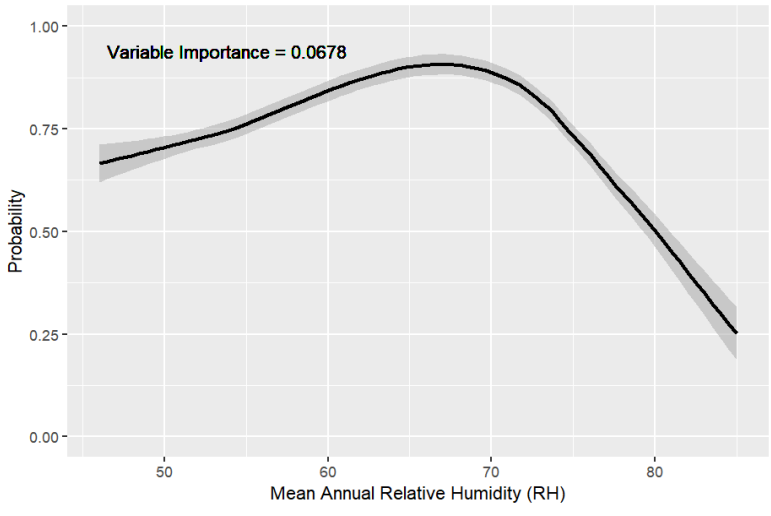

Variable Importance = 0.1663
Probability
1.00
0.75
0.50
0.25
0.00
200
400
600
Autumn (September to November) Precipitation (PPT_at)

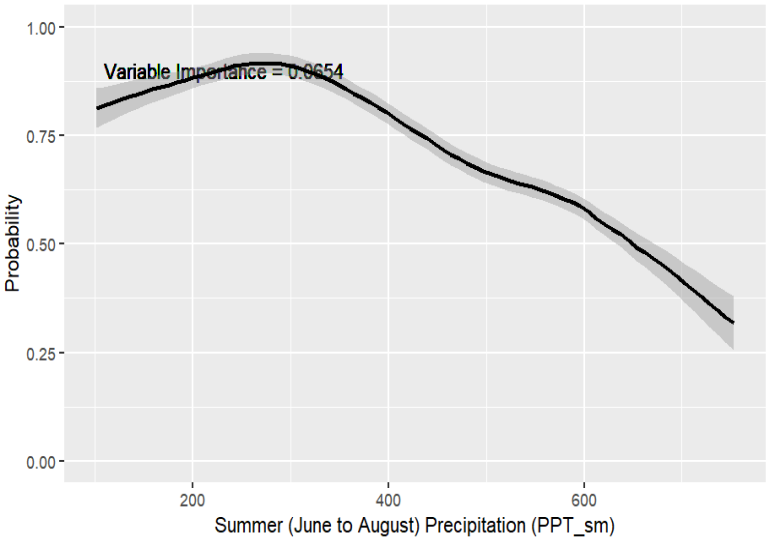

Variable Importance = 0.1639
Probability
1.00
0.75
0.50
0.25
0.00
50
100
150
200
Summer Heat Moisture Index (SHM)

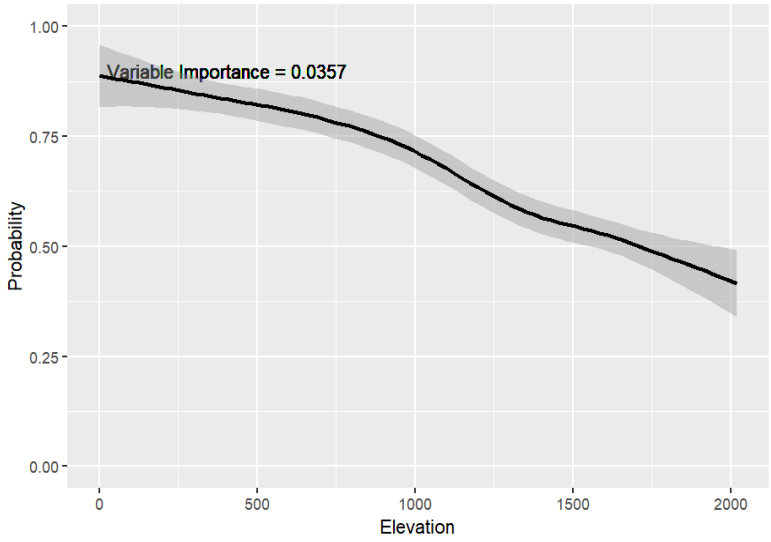

Variable Importance = 0.1248
Probability
1.00
0.75
0.50
0.25
0.00
200
400
600
Summer (June to August) Precipitation (PPT_sm)


Variable Importance = 0.1045
1.00
0.75
0.50
0.25
0.00
10
20
30
40
Temperature Difference Between the Coldest and Warmest Month (TD)


Variable Importance = 0.0847
Probability
1.00
0.75
0.50
0.25
0.00
0
500
1000
1500
Hargreave's Climatic Moisture Deficit (CMD)


Variable Importance = 0.0678
Probability
1.00
0.75
0.50
0.25
0.00
50
60
70
80
Mean Annual Relative Humidity (RH)


Variable Importance = 0.0654
Probability
1.00
0.75
0.50
0.25
0.00
200
400
600
Summer (June to August) Precipitation (PPT_sm)


Variable Importance = 0.0357
Probability
1.00
0.75
0.50
0.25
0.00
0
500
1000
1500
2000
Elevation

Appendix F

Algorithms selected for use within complementary (AdaptWest) ensembles with major statistics

Table F1. Comparison of model statistics (± standard error) between the remaining algorithms chosen for each ensemble in the complementary dataset.

Algorithm	Count	AUC	TSS	Kappa
GLM	2	0.879 ± 0.002	0.7010	0.607 ± 0.004
GAM	26	0.9225 ±0.0105	0.714 ± 0.034	0.652 ± 0.052
GLMPoly	11	0.914±0.008	0.702±0.021	0.637 ±0.035
MARS	3	0.9325 ± 0.0015	0.763 ± 0.012	0.709 ± 0.032
MDA	8	0.9605 ±0.01905	0.845 ± 0.045	0.7995 ± 0.0475
FDA	9	0.8985 ± 0.0785	0.763 ± 0.140	0.7295 ± 0.1025
RPART	11	0.9805 ± 0.0035	0.928 ± 0.056	0.937 ± 0.037
CART	16	0.960 ± 0.012	0.8845 ± 0.0125	0.9295 ± 0.0575
BRT	13	0.9875 ± 0.0045	0.918 ± 0.017	0.9185 ± 0.0175
RF	25	0.9925 ± 0.0035	0.9255 ± 0.0225	0.9255 ± 0.025
SVM	6	0.984 ± 0.004	0.909 ± 0.019	0.909 ± 0.019
RBF	17	0.9555 ± 0.0155	0.8475 ± 0.0255	0.847 ± 0.025
MLP	19	0.985 ± 0.005	0.9095 ± 0.0185	0.9095 ± 0.0185

Appendix G

Changes in total land suitability per each timestamp per each additional SSP scenario (i.e., Chelsa-Bioclim+ dataset using SSP 3-7.0; AdaptWest dataset using SSP3-7.0 and SSP5-8.5)

Table G1. Changes in total land suitability in the (Chelsea-Bioclim+ dataset using SSP3-7.0.

Time Period	Suitability	Total Area (km^2)	Relative Percentage (%)
Current (Historic)	Unsuitable	2,432,522 km^2	64.81%
Current (Historic)	Low Suitability	613,459 km^2	16.33%
Current (Historic)	Medium Suitability	259,804 km^2	6.92%
Current (Historic)	High Suitability	126,529 km^2	3.37%
Current (Historic)	Optimal Suitability	321,003 km^2	8.56%
2011-2040 (SSP3-7.0)	Unsuitable	1,788,838 km^2	47.66%
2011-2040 (SSP3-7.0)	Low Suitability	777,284 km^2	20.71%
2011-2040 (SSP3-7.0)	Medium Suitability	457,502 km^2	12.19%
2011-2040 (SSP3-7.0)	High Suitability	272,212 km^2	7.25%
2011-2040 (SSP3-7.0)	Optimal Suitability	457,481 km^2	12.18%
2041-2070 (SSP3-7.0)	Unsuitable	1,507,336 km^2	40.16%
2041-2070 (SSP3-7.0)	Low Suitability	695,700 km^2	18.54%
2041-2070 (SSP3-7.0)	Medium Suitability	526,056 km^2	14.02%
2041-2070 (SSP3-7.0)	High Suitability	424,689 km^2	11.32%
2041-2070 (SSP3-7.0)	Optimal Suitability	599,536 km^2	15.97%
2071-2100 (SSP3-7.0)	Unsuitable	1,018,118 km^2	27.13%
2071-2100 (SSP3-7.0)	Low Suitability	768,076 km^2	20.46%

2071-2100 (SSP3-7.0)	Medium Suitability	622,371 km^2	16.58%
2071-2100 (SSP3-7.0)	High Suitability	632,165 km^2	16.84%
2071-2100 (SSP3-7.0)	Optimal Suitability	712,587 km^2	18.99%

Table G2. Changes in total land suitability in the secondary (AdaptWest) dataset using SSP5-8.5.

Time Period	Suitability	Total Area (km^2)	Relative Percentage (%)
Current (Historic)	Unsuitable	2,433,392 km^2	64.71%
Current (Historic)	Low Suitability	728,614 km^2	19.38%
Current (Historic)	Medium Suitability	189,052 km^2	5.03%
Current (Historic)	High Suitability	92,534 km^2	2.46%
Current (Historic)	Optimal Suitability	316,938 km^2	8.43%
2011-2040 (SSP5-8.5)	Unsuitable	1,197,774 km^2	31.85%
2011-2040 (SSP5-8.5)	Low Suitability	1,449,894 km^2	38.56%
2011-2040 (SSP5-8.5)	Medium Suitability	494,122 km^2	13.14%
2011-2040 (SSP5-8.5)	High Suitability	309,595 km^2	8.23%
2011-2040 (SSP5-8.5)	Optimal Suitability	308,901 km^2	8.21%
2041-2070 (SSP5-8.5)	Unsuitable	1,001,291 km^2	26.63%
2041-2070 (SSP5-8.5)	Low Suitability	1,077,897 km^2	28.67%
2041-2070 (SSP5-8.5)	Medium Suitability	693,718 km^2	18.45%
2041-2070 (SSP5-8.5)	High Suitability	610,194 km^2	16.23%
2041-2070 (SSP5-8.5)	Optimal Suitability	377,188 km^2	10.03%
2071-2100 (SSP5-8.5)	Unsuitable	117,971 km^2	3.14%

2071-2100 (SSP5-8.5)	Low Suitability	1,329,188 km^2	35.35%
2071-2100 (SSP5-8.5)	Medium Suitability	944,997 km^2	25.13%
2071-2100 (SSP5-8.5)	High Suitability	1,256,237 km^2	33.41%
2071-2100 (SSP5-8.5)	Optimal Suitability	111,895 km^2	2.98%

Table G3. Changes in total land suitability in the secondary (AdaptWest) dataset using SSP3-7.0.

Time Period	Suitability	Total Area (km^2)	Relative Percentage (%)
Current (Historic)	Unsuitable	2,433,392 km^2	64.71%
Current (Historic)	Low Suitability	728,614 km^2	19.38%
Current (Historic)	Medium Suitability	189,052 km^2	5.03%
Current (Historic)	High Suitability	92,534 km^2	2.46%
Current (Historic)	Optimal Suitability	316,938 km^2	8.43%
2011-2040 (SSP5-8.5)	Unsuitable	1,851,959 km^2	49.25%
2011-2040 (SSP5-8.5)	Low Suitability	917,943 km^2	24.41%
2011-2040 (SSP5-8.5)	Medium Suitability	401,259 km^2	10.67%
2011-2040 (SSP5-8.5)	High Suitability	275,642 km^2	7.33%
2011-2040 (SSP5-8.5)	Optimal Suitability	313,485 km^2	8.34%
2041-2070 (SSP5-8.5)	Unsuitable	1,208,590 km^2	32.14%
2041-2070 (SSP5-8.5)	Low Suitability	1,024,294 km^2	27.24%
2041-2070 (SSP5-8.5)	Medium Suitability	635,489 km^2	16.90%
2041-2070 (SSP5-8.5)	High Suitability	519,582 km^2	13.82%
2041-2070 (SSP5-8.5)	Optimal Suitability	372,233 km^2	9.90%

2071-2100 (SSP5-8.5)	Unsuitable	325,239 km^2	8.65%
2071-2100 (SSP5-8.5)	Low Suitability	1,335,207 km^2	35.51%
2071-2100 (SSP5-8.5)	Medium Suitability	852,153 km^2	22.66%
2071-2100 (SSP5-8.5)	High Suitability	979,081 km^2	26.04%
2071-2100 (SSP5-8.5)	Optimal Suitability	268,608 km^2	7.14%

Appendix H

Original count of *I. scapularis* records in each province or state over the five-year period

Year	Country	Province/State	Ixodes Scapularis
2017	USA	Delaware	0
2017	USA	Kentucky	0
2017	USA	Virginia	0
2017	USA	Pennsylvania	8
2017	USA	South Carolina	1
2017	USA	South Dakota	0
2017	USA	Ohio	5
2017	USA	New Jersey	6
2017	USA	Alabama	1
2017	USA	North Carolina	0
2017	USA	Missouri	1
2017	USA	North Dakota	0
2017	CAN	Prince Edward Island / Île-du-Prince-Édouard	0

2017	USA	Georgia	1
2017	USA	Kansas	0
2017	CAN	New Brunswick / Nouveau-Brunswick	0
2017	USA	Tennessee	2
2017	USA	Rhode Island	0
2017	CAN	Quebec / Québec	158
2017	USA	Maine	1
2017	USA	Minnesota	2
2017	USA	Connecticut	5
2017	USA	New Hampshire	1
2017	USA	Florida	4
2017	CAN	Ontario	18
2017	USA	Texas	0
2017	USA	Iowa	0
2017	USA	Wisconsin	1
2017	USA	Maryland	3
2017	USA	District of Columbia	0
2017	USA	Vermont	27
2017	USA	Nebraska	0
2017	CAN	Newfoundland and Labrador / Terre-Neuve-et-Labrador	0
2017	USA	New York	30
2017	CAN	Manitoba	1
2017	USA	Illinois	1

2017	USA	Mississippi	0
2017	USA	Oklahoma	1
2017	USA	Louisiana	0
2017	USA	Massachusetts	14
2017	USA	West Virginia	1
2017	USA	Indiana	2
2017	CAN	Nova Scotia / Nouvelle-Écosse	1
2017	USA	Michigan	1
2017	USA	Arkansas	2
2018	USA	Delaware	1
2018	USA	Kentucky	0
2018	USA	Virginia	5
2018	USA	Pennsylvania	28
2018	USA	South Carolina	0
2018	USA	South Dakota	0
2018	USA	Ohio	10
2018	USA	New Jersey	3
2018	USA	Alabama	4
2018	USA	North Carolina	0
2018	USA	Missouri	2
2018	USA	North Dakota	0
2018	CAN	Prince Edward Island / Île-du-Prince-Édouard	0
2018	USA	Georgia	1
2018	USA	Kansas	0

2018	CAN	New Brunswick / Nouveau-Brunswick	0
2018	USA	Tennessee	4
2018	USA	Rhode Island	1
2018	CAN	Quebec / Québec	126
2018	USA	Maine	1
2018	USA	Minnesota	3
2018	USA	Connecticut	19
2018	USA	New Hampshire	1
2018	USA	Florida	6
2018	CAN	Ontario	31
2018	USA	Texas	7
2018	USA	Iowa	0
2018	USA	Wisconsin	3
2018	USA	Maryland	1
2018	USA	District of Columbia	1
2018	USA	Vermont	24
2018	USA	Nebraska	0
2018	CAN	Newfoundland and Labrador / Terre-Neuve-et-Labrador	0
2018	USA	New York	30
2018	CAN	Manitoba	0
2018	USA	Illinois	5
2018	USA	Mississippi	1
2018	USA	Oklahoma	0

2018	USA	Louisiana	2
2018	USA	Massachusetts	22
2018	USA	West Virginia	0
2018	USA	Indiana	3
2018	CAN	Nova Scotia / Nouvelle-Écosse	8
2018	USA	Michigan	5
2018	USA	Arkansas	2
2019	USA	Delaware	0
2019	USA	Kentucky	2
2019	USA	Virginia	9
2019	USA	Pennsylvania	60
2019	USA	South Carolina	1
2019	USA	South Dakota	0
2019	USA	Ohio	14
2019	USA	New Jersey	16
2019	USA	Alabama	10
2019	USA	North Carolina	5
2019	USA	Missouri	5
2019	USA	North Dakota	0
2019	CAN	Prince Edward Island / Île-du-Prince-Édouard	0
2019	USA	Georgia	7
2019	USA	Kansas	1
2019	CAN	New Brunswick /	45

		Nouveau-Brunswick	
2019	USA	Tennessee	11
2019	USA	Rhode Island	2
2019	CAN	Quebec / Québec	552
2019	USA	Maine	10
2019	USA	Minnesota	17
2019	USA	Connecticut	11
2019	USA	New Hampshire	6
2019	USA	Florida	7
2019	CAN	Ontario	448
2019	USA	Texas	9
2019	USA	Iowa	0
2019	USA	Wisconsin	12
2019	USA	Maryland	14
2019	USA	District of Columbia	0
2019	USA	Vermont	101
2019	USA	Nebraska	0
2019	CAN	Newfoundland and Labrador / Terre-Neuve-et-Labrador	0
2019	USA	New York	69
2019	CAN	Manitoba	3
2019	USA	Illinois	14
2019	USA	Mississippi	2
2019	USA	Oklahoma	4
2019	USA	Louisiana	0

2019	USA	Massachusetts	62
2019	USA	West Virginia	4
2019	USA	Indiana	5
2019	CAN	Nova Scotia / Nouvelle-Écosse	14
2019	USA	Michigan	13
2019	USA	Arkansas	0
2020	USA	Delaware	3
2020	USA	Kentucky	9
2020	USA	Virginia	42
2020	USA	Pennsylvania	111
2020	USA	South Carolina	6
2020	USA	South Dakota	0
2020	USA	Ohio	47
2020	USA	New Jersey	28
2020	USA	Alabama	2
2020	USA	North Carolina	11
2020	USA	Missouri	4
2020	USA	North Dakota	1
2020	CAN	Prince Edward Island / Île-du-Prince-Édouard	3
2020	USA	Georgia	7
2020	USA	Kansas	1
2020	CAN	New Brunswick / Nouveau-Brunswick	132

2020	USA	Tennessee	7
2020	USA	Rhode Island	3
2020	CAN	Quebec / Québec	440
2020	USA	Maine	13
2020	USA	Minnesota	25
2020	USA	Connecticut	33
2020	USA	New Hampshire	29
2020	USA	Florida	9
2020	CAN	Ontario	1563
2020	USA	Texas	26
2020	USA	Iowa	5
2020	USA	Wisconsin	30
2020	USA	Maryland	28
2020	USA	District of Columbia	4
2020	USA	Vermont	84
2020	USA	Nebraska	1
2020	CAN	Newfoundland and Labrador / Terre-Neuve-et-Labrador	11
2020	USA	New York	114
2020	CAN	Manitoba	3
2020	USA	Illinois	20
2020	USA	Mississippi	5
2020	USA	Oklahoma	4
2020	USA	Louisiana	5
2020	USA	Massachusetts	101
2020	USA	West Virginia	5

2020	USA	Indiana	7
2020	CAN	Nova Scotia / Nouvelle-Écosse	370
2020	USA	Michigan	21
2020	USA	Arkansas	2
2021	USA	Delaware	1
2021	USA	Kentucky	14
2021	USA	Virginia	39
2021	USA	Pennsylvania	110
2021	USA	South Carolina	8
2021	USA	South Dakota	1
2021	USA	Ohio	51
2021	USA	New Jersey	54
2021	USA	Alabama	6
2021	USA	North Carolina	12
2021	USA	Missouri	3
2021	USA	North Dakota	0
2021	CAN	Prince Edward Island / Île-du-Prince-Édouard	69
2021	USA	Georgia	10
2021	USA	Kansas	0
2021	CAN	New Brunswick / Nouveau-Brunswick	146
2021	USA	Tennessee	8
2021	USA	Rhode Island	4

2021	CAN	Quebec / Québec	724
2021	USA	Maine	14
2021	USA	Minnesota	32
2021	USA	Connecticut	20
2021	USA	New Hampshire	29
2021	USA	Florida	22
2021	CAN	Ontario	2593
2021	USA	Texas	31
2021	USA	Iowa	8
2021	USA	Wisconsin	44
2021	USA	Maryland	42
2021	USA	District of Columbia	1
2021	USA	Vermont	51
2021	USA	Nebraska	0
2021	CAN	Newfoundland and Labrador / Terre-Neuve-et-Labrador	15
2021	USA	New York	160
2021	CAN	Manitoba	2
2021	USA	Illinois	18
2021	USA	Mississippi	7
2021	USA	Oklahoma	7
2021	USA	Louisiana	0
2021	USA	Massachusetts	96
2021	USA	West Virginia	33
2021	USA	Indiana	13

2021	CAN	Nova Scotia / Nouvelle-Écosse	608
2021	USA	Michigan	38
2021	USA	Arkansas	6
2022	USA	Delaware	9
2022	USA	Kentucky	4
2022	USA	Virginia	18
2022	USA	Pennsylvania	59
2022	USA	South Carolina	3
2022	USA	South Dakota	0
2022	USA	Ohio	29
2022	USA	New Jersey	12
2022	USA	Alabama	4
2022	USA	North Carolina	8
2022	USA	Missouri	2
2022	USA	North Dakota	0
2022	CAN	Prince Edward Island / Île-du-Prince-Édouard	0
2022	USA	Georgia	9
2022	USA	Kansas	1
2022	CAN	New Brunswick / Nouveau-Brunswick	5
2022	USA	Tennessee	4
2022	USA	Rhode Island	5

2022	CAN	Quebec / Québec	13
2022	USA	Maine	14
2022	USA	Minnesota	34
2022	USA	Connecticut	14
2022	USA	New Hampshire	19
2022	USA	Florida	18
2022	CAN	Ontario	113
2022	USA	Texas	12
2022	USA	Iowa	1
2022	USA	Wisconsin	31
2022	USA	Maryland	23
2022	USA	District of Columbia	0
2022	USA	Vermont	38
2022	USA	Nebraska	0
2022	CAN	Newfoundland and Labrador / Terre-Neuve-et-Labrador	0
2022	USA	New York	69
2022	CAN	Manitoba	2
2022	USA	Illinois	20
2022	USA	Mississippi	0
2022	USA	Oklahoma	5
2022	USA	Louisiana	0
2022	USA	Massachusetts	80
2022	USA	West Virginia	9
2022	USA	Indiana	5

2022	CAN	Nova Scotia / Nouvelle-Écosse	43
2022	USA	Michigan	18
2022	USA	Arkansas	2
		Total	11504

Appendix I

Primary (Chelsa-Bioclim+) datasets selected models and metrics

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
3	GAM	29	0.926	0.72	0.593	0.255	0.301	0.713	0.658	0.851	0.862	0.388	0.666	0.678	0.944	0.859	0.321
3	GAM	36	0.931	0.728	0.573	0.254	0.269	0.729	0.655	0.885	0.845	0.403	0.67	0.66	0.956	0.855	0.341
3	GAM	40	0.93	0.731	0.577	0.255	0.281	0.732	0.666	0.878	0.855	0.408	0.678	0.674	0.953	0.86	0.332
3	GAM	51	0.929	0.728	0.575	0.255	0.263	0.723	0.657	0.872	0.852	0.397	0.669	0.668	0.951	0.857	0.333
3	GAM	55	0.927	0.722	0.586	0.255	0.309	0.717	0.668	0.847	0.871	0.396	0.675	0.691	0.943	0.864	0.312
3	GLMPOLY	68	0.927	0.719	0.592	0.255	0.252	0.732	0.652	0.895	0.837	0.405	0.669	0.652	0.959	0.851	0.35
3	GLMPOLY	69	0.925	0.711	0.595	0.255	0.292	0.727	0.666	0.869	0.859	0.404	0.676	0.678	0.95	0.861	0.327
3	FDA	36	0.931	0.728	0.573	0.254	0.269	0.729	0.655	0.885	0.845	0.403	0.67	0.66	0.956	0.855	0.341
3	FDA	40	0.93	0.731	0.577	0.255	0.281	0.732	0.666	0.878	0.855	0.408	0.678	0.674	0.953	0.86	0.332
3	FDA	51	0.929	0.728	0.575	0.255	0.263	0.723	0.657	0.872	0.852	0.397	0.669	0.668	0.951	0.857	0.333
3	RPART	68	0.927	0.719	0.592	0.255	0.252	0.732	0.652	0.895	0.837	0.405	0.669	0.652	0.959	0.851	0.35
3	RPART	69	0.925	0.711	0.595	0.255	0.292	0.727	0.666	0.869	0.859	0.404	0.676	0.678	0.95	0.861	0.327
3	BRT	2	0.99	0.925	0.250	0.5	0.468	0.913	0.912	0.936	0.977	0.747	0.913	0.976	0.938	0.956	0.48
3	BRT	6	0.991	0.929	0.236	0.5	0.584	0.912	0.912	0.944	0.968	0.741	0.912	0.967	0.945	0.956	0.488
3	RF	30	0.991	0.929	0.237	0.5	0.493	0.918	0.918	0.945	0.974	0.757	0.919	0.973	0.946	0.959	0.486
3	RF	33	0.99	0.923	0.260	0.502	0.422	0.911	0.911	0.938	0.974	0.743	0.912	0.973	0.939	0.956	0.484
3	RF	34	0.993	0.937	0.207	0.5	0.520	0.926	0.926	0.956	0.97	0.773	0.926	0.969	0.957	0.963	0.493
3	RF	35	0.991	0.935	0.236	0.501	0.505	0.929	0.929	0.955	0.974	0.78	0.929	0.974	0.955	0.964	0.491
3	RF	36	0.991	0.933	0.225	0.512	0.522	0.922	0.922	0.945	0.978	0.767	0.922	0.978	0.944	0.961	0.495
3	RF	43	0.989	0.923	0.255	0.5	0.550	0.914	0.914	0.924	0.99	0.762	0.916	0.989	0.929	0.957	0.468
3	MLP	114	0.98	0.908	0.412	0.5	0.608	0.895	0.895	0.924	0.971	0.71	0.896	0.969	0.927	0.947	0.477
3	MLP	115	0.982	0.902	0.360	0.5	0.551	0.891	0.891	0.933	0.958	0.696	0.891	0.957	0.934	0.945	0.488

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
3	MLP	127	0.981	0.914	0.349	0.5	0.375	0.896	0.896	0.936	0.961	0.708	0.897	0.96	0.937	0.948	0.488
3	MLP	134	0.983	0.911	0.340	0.5	0.565	0.898	0.898	0.921	0.977	0.719	0.899	0.975	0.925	0.949	0.472
4	GAM	30	0.917	0.695	0.626	0.255	0.256	0.723	0.652	0.879	0.845	0.396	0.666	0.659	0.953	0.853	0.34
4	GAM	35	0.919	0.699	0.620	0.262	0.241	0.729	0.655	0.891	0.838	0.403	0.67	0.661	0.956	0.852	0.353
4	GAM	56	0.927	0.715	0.592	0.255	0.241	0.735	0.658	0.893	0.842	0.41	0.673	0.658	0.958	0.855	0.346
4	GLMPOLY	69	0.907	0.67	0.655	0.255	0.280	0.703	0.619	0.885	0.819	0.37	0.638	0.625	0.954	0.835	0.361
4	GLMPOLY	70	0.911	0.685	0.646	0.255	0.411	0.714	0.669	0.839	0.875	0.393	0.674	0.696	0.941	0.866	0.308
4	GLMPOLY	76	0.916	0.69	0.637	0.255	0.354	0.702	0.641	0.853	0.85	0.373	0.651	0.66	0.944	0.85	0.33
4	GLMPOLY	84	0.908	0.676	0.655	0.255	0.366	0.7	0.645	0.842	0.858	0.372	0.653	0.669	0.941	0.854	0.321
4	MDA	16	0.982	0.888	0.359	0.757	0.808	0.902	0.857	0.939	0.964	0.687	0.861	0.988	0.835	0.945	0.719
4	MDA	25	0.977	0.858	0.419	0.757	0.907	0.921	0.859	0.93	0.991	0.722	0.867	0.997	0.819	0.945	0.707
4	MDA	26	0.958	0.857	0.557	0.757	0.790	0.887	0.85	0.942	0.945	0.662	0.853	0.982	0.838	0.942	0.726
4	MDA	27	0.978	0.875	0.380	0.757	0.815	0.896	0.851	0.937	0.959	0.675	0.855	0.986	0.83	0.942	0.72
4	FDA	44	0.978	0.837	0.394	0.757	0.948	0.889	0.853	0.943	0.945	0.666	0.856	0.982	0.842	0.944	0.727
4	FDA	53	0.982	0.851	0.344	0.757	0.974	0.9	0.842	0.927	0.973	0.676	0.849	0.991	0.81	0.938	0.709
4	RPART	8	0.983	0.922	0.314	0.495	0.500	0.908	0.908	0.939	0.969	0.734	0.909	0.967	0.942	0.954	0.481
4	CART	36	0.962	0.895	0.391	0.495	0.823	0.888	0.888	0.917	0.971	0.697	0.89	0.969	0.923	0.944	0.468
4	CART	46	0.97	0.9	0.362	0.5	0.633	0.889	0.889	0.928	0.961	0.693	0.889	0.959	0.931	0.944	0.484
4	BRT	63	0.994	0.943	0.195	0.501	0.434	0.932	0.932	0.971	0.961	0.786	0.932	0.962	0.97	0.966	0.506
4	BRT	64	0.993	0.942	0.200	0.495	0.585	0.929	0.93	0.944	0.985	0.788	0.93	0.984	0.947	0.965	0.475
4	BRT	71	0.991	0.933	0.224	0.5	0.515	0.921	0.921	0.947	0.974	0.763	0.921	0.973	0.949	0.961	0.487
4	BRT	74	0.992	0.932	0.218	0.5	0.709	0.921	0.921	0.937	0.984	0.77	0.922	0.983	0.94	0.961	0.477
4	RF	89	0.993	0.942	0.200	0.497	0.436	0.937	0.937	0.957	0.98	0.801	0.937	0.979	0.959	0.969	0.486
4	RF	90	0.994	0.944	0.187	0.493	0.463	0.938	0.938	0.964	0.974	0.801	0.938	0.973	0.965	0.969	0.489
4	RF	91	0.995	0.95	0.180	0.501	0.573	0.946	0.946	0.971	0.975	0.821	0.946	0.975	0.971	0.973	0.5
4	RF	92	0.995	0.952	0.173	0.495	0.523	0.944	0.944	0.957	0.987	0.82	0.944	0.986	0.959	0.972	0.48
4	RBF	158	0.97	0.774	0.618	0.5	0.643	0.877	0.877	0.924	0.953	0.669	0.878	0.952	0.926	0.939	0.486
4	RBF	164	0.968	0.81	0.561	0.501	0.596	0.864	0.864	0.909	0.955	0.646	0.865	0.953	0.913	0.932	0.478
4	MLP	171	0.988	0.92	0.281	0.5	0.342	0.912	0.912	0.952	0.961	0.741	0.912	0.96	0.952	0.956	0.496
4	MLP	175	0.987	0.922	0.283	0.501	0.341	0.913	0.913	0.966	0.947	0.742	0.913	0.948	0.965	0.956	0.511
4	MLP	176	0.982	0.914	0.374	0.495	0.319	0.911	0.911	0.949	0.962	0.739	0.911	0.961	0.951	0.956	0.489
4	MLP	183	0.986	0.919	0.304	0.5	0.675	0.912	0.912	0.942	0.971	0.744	0.913	0.97	0.943	0.956	0.486
5	GAM	47	0.927	0.718	0.588	0.255	0.356	0.731	0.681	0.857	0.875	0.413	0.688	0.7	0.947	0.87	0.312
5	GAM	56	0.927	0.722	0.587	0.255	0.324	0.738	0.682	0.868	0.87	0.419	0.69	0.695	0.951	0.869	0.319
5	GLMPoly	65	0.911	0.681	0.650	0.255	0.360	0.709	0.653	0.85	0.86	0.383	0.662	0.674	0.943	0.857	0.321
5	GLMPoly	75	0.908	0.669	0.660	0.255	0.310	0.709	0.627	0.885	0.825	0.377	0.645	0.633	0.954	0.84	0.356
5	MDA	8	0.974	0.856	1.166	0.76	0.725	0.888	0.84	0.933	0.955	0.658	0.845	0.985	0.818	0.938	0.72

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
5	MDA	27	0.979	0.856	0.566	0.757	0.862	0.888	0.827	0.92	0.968	0.652	0.834	0.989	0.795	0.931	0.704
5	FDA	31	0.974	0.806	0.437	0.757	0.957	0.881	0.81	0.908	0.973	0.637	0.82	0.99	0.773	0.924	0.694
5	FDA	36	0.974	0.818	0.447	0.76	0.941	0.894	0.836	0.926	0.968	0.664	0.842	0.989	0.805	0.936	0.711
5	FDA	48	0.975	0.819	0.421	0.757	0.947	0.884	0.815	0.911	0.973	0.643	0.824	0.99	0.778	0.926	0.696
5	FDA	53	0.981	0.834	0.382	0.757	0.943	0.892	0.839	0.928	0.964	0.663	0.845	0.988	0.812	0.937	0.711
5	RPART	21	0.973	0.897	0.385	0.5	0.857	0.902	0.902	0.909	0.993	0.745	0.905	0.992	0.916	0.951	0.458
5	CART	29	0.954	0.88	0.438	0.5	0.904	0.872	0.872	0.905	0.966	0.665	0.873	0.964	0.911	0.936	0.469
5	CART	39	0.949	0.876	0.451	0.5	0.779	0.872	0.872	0.912	0.959	0.661	0.872	0.957	0.916	0.936	0.477
5	CART	49	0.959	0.883	0.424	0.5	0.762	0.87	0.87	0.923	0.947	0.655	0.87	0.946	0.925	0.935	0.488
5	BRT	61	0.993	0.942	0.193	0.509	0.442	0.931	0.931	0.96	0.972	0.785	0.931	0.972	0.959	0.966	0.502
5	BRT	62	0.992	0.936	0.218	0.48	0.476	0.926	0.926	0.961	0.965	0.772	0.926	0.962	0.964	0.963	0.48
5	BRT	63	0.992	0.934	0.216	0.499	0.534	0.923	0.923	0.947	0.976	0.768	0.923	0.975	0.949	0.961	0.485
5	BRT	64	0.993	0.938	0.209	0.505	0.521	0.933	0.933	0.954	0.979	0.791	0.933	0.978	0.954	0.966	0.492
5	RF	89	0.991	0.928	0.233	0.509	0.450	0.91	0.91	0.942	0.969	0.738	0.91	0.969	0.941	0.955	0.494
5	RF	92	0.99	0.926	0.242	0.505	0.530	0.912	0.911	0.938	0.974	0.744	0.912	0.973	0.939	0.956	0.486
5	SVM	115	0.983	0.912	0.294	0.5	0.411	0.901	0.901	0.924	0.977	0.724	0.902	0.975	0.928	0.95	0.474
5	SVM	117	0.987	0.917	0.268	0.509	0.443	0.9	0.9	0.934	0.966	0.717	0.9	0.966	0.934	0.95	0.492
5	SVM	119	0.985	0.912	0.295	0.499	0.398	0.9	0.9	0.933	0.967	0.717	0.901	0.966	0.935	0.95	0.483
5	SVM	122	0.984	0.917	0.278	0.5	0.456	0.905	0.905	0.923	0.982	0.737	0.907	0.981	0.927	0.953	0.47
5	SVM	123	0.986	0.917	0.275	0.5	0.458	0.904	0.904	0.93	0.974	0.728	0.905	0.973	0.933	0.952	0.478
5	SVM	129	0.983	0.916	0.283	0.5	0.399	0.908	0.908	0.933	0.975	0.737	0.909	0.974	0.936	0.954	0.479
5	SVM	134	0.985	0.918	0.276	0.5	0.463	0.908	0.908	0.939	0.969	0.734	0.908	0.968	0.941	0.954	0.485
5	RBF	145	0.949	0.759	0.696	0.509	0.590	0.828	0.827	0.893	0.935	0.579	0.828	0.934	0.894	0.913	0.486
5	RBF	147	0.947	0.775	0.694	0.499	0.593	0.827	0.827	0.884	0.943	0.581	0.828	0.939	0.891	0.913	0.47
5	RBF	150	0.953	0.78	0.650	0.5	0.587	0.825	0.825	0.888	0.937	0.575	0.826	0.934	0.893	0.912	0.475
5	RBF	151	0.944	0.754	0.696	0.5	0.596	0.822	0.822	0.896	0.926	0.568	0.822	0.923	0.899	0.911	0.485
5	RBF	157	0.956	0.769	0.669	0.5	0.509	0.828	0.828	0.926	0.902	0.577	0.828	0.904	0.924	0.914	0.512
5	RBF	162	0.943	0.757	0.727	0.5	0.565	0.82	0.82	0.907	0.914	0.564	0.82	0.913	0.907	0.91	0.496
5	MLP	173	0.983	0.911	0.308	0.509	0.452	0.9	0.9	0.932	0.969	0.718	0.9	0.968	0.932	0.95	0.489
5	MLP	178	0.986	0.92	0.269	0.5	0.611	0.908	0.908	0.918	0.99	0.751	0.91	0.989	0.924	0.954	0.464
5	MLP	189	0.985	0.917	0.294	0.5	0.452	0.909	0.909	0.946	0.964	0.735	0.91	0.963	0.947	0.955	0.491
5	MLP	190	0.984	0.915	0.302	0.5	0.598	0.902	0.902	0.921	0.981	0.731	0.904	0.98	0.926	0.951	0.47
5	MARS	5	0.916	0.75	0.541	0.767	0.723	0.915	0.817	0.732	0.708	0.431	0.709	0.943	0.745	0.892	0.744
5	MARS	10	0.936	0.786	0.484	0.757	0.664	0.926	0.827	0.753	0.738	0.465	0.739	0.943	0.781	0.902	0.743
5	MARS	17	0.941	0.778	0.496	0.757	0.755	0.911	0.85	0.761	0.729	0.464	0.732	0.95	0.754	0.896	0.726
6	GAM	29	0.928	0.727	0.586	0.255	0.289	0.734	0.684	0.86	0.875	0.417	0.69	0.701	0.948	0.871	0.313
6	GAM	54	0.925	0.718	0.590	0.255	0.356	0.727	0.689	0.839	0.888	0.414	0.693	0.719	0.942	0.876	0.297

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
6	GLMPOLY	57	0.923	0.718	0.628	0.255	0.394	0.735	0.695	0.847	0.888	0.422	0.699	0.721	0.944	0.877	0.299
6	GLMPOLY	82	0.916	0.703	0.644	0.255	0.403	0.723	0.681	0.841	0.882	0.406	0.686	0.709	0.942	0.871	0.302
6	MDA	5	0.985	0.893	0.532	0.761	0.684	0.905	0.871	0.95	0.955	0.7	0.873	0.985	0.858	0.951	0.734
6	MDA	11	0.982	0.876	0.455	0.757	0.743	0.908	0.873	0.949	0.959	0.705	0.875	0.986	0.858	0.951	0.728
6	FDA	30	0.883	0.714	0.665	0.757	0.823	0.743	0.712	0.907	0.836	0.44	0.714	0.945	0.742	0.89	0.726
6	FDA	31	0.878	0.701	0.698	0.757	0.720	0.749	0.714	0.904	0.845	0.445	0.717	0.948	0.738	0.89	0.722
6	FDA	53	0.878	0.716	0.683	0.757	0.773	0.755	0.719	0.905	0.85	0.453	0.723	0.949	0.742	0.892	0.721
6	FDA	55	0.877	0.7	0.695	0.757	0.788	0.752	0.707	0.893	0.859	0.443	0.712	0.952	0.721	0.885	0.71
6	RPART	13	0.98	0.923	0.330	0.5	0.571	0.915	0.915	0.947	0.968	0.749	0.916	0.967		0.958	0.49
6	RPART	18	0.98	0.924	0.327	0.5	0.571	0.918	0.918	0.942	0.977	0.759	0.919	0.976		0.959	0.482
6	RPART	27	0.978	0.924	0.340	0.5	0.600	0.917	0.917	0.95	0.966	0.751	0.917	0.966		0.958	0.492
6	CART	30	0.975	0.916	0.320	0.5	0.577	0.899	0.899	0.936	0.964	0.714	0.9	0.962		0.95	0.486
6	CART	31	0.979	0.909	0.326	0.5	0.563	0.895	0.895	0.939	0.956	0.704	0.895	0.955		0.947	0.491
6	CART	45	0.972	0.908	0.344	0.5	0.512	0.895	0.895	0.936	0.959	0.704	0.895	0.958		0.947	0.488
6	CART	55	0.974	0.913	0.326	0.5	0.923	0.895	0.895	0.918	0.977	0.713	0.896	0.975		0.947	0.47
6	BRT	59	0.993	0.938	0.227	0.5	0.410	0.923	0.923	0.955	0.968	0.764	0.923	0.967		0.961	0.493
6	BRT	60	0.991	0.935	0.235	0.5	0.452	0.926	0.926	0.955	0.971	0.772	0.926	0.97		0.963	0.492
6	BRT	71	0.991	0.932	0.246	0.5	0.382	0.927	0.927	0.958	0.969	0.775	0.927	0.969		0.964	0.494
6	BRT	72	0.992	0.936	0.229	0.5	0.668	0.924	0.924	0.933	0.991	0.784	0.926	0.991		0.962	0.471
6	BRT	73	0.992	0.935	0.233	0.5	0.420	0.923	0.923	0.953	0.969	0.765	0.923	0.969		0.961	0.492
6	BRT	74	0.991	0.936	0.235	0.5	0.440	0.924	0.924	0.947	0.977	0.771	0.924	0.976		0.962	0.485
6	BRT	83	0.991	0.932	0.242	0.5	0.538	0.921	0.921	0.944	0.977	0.765	0.922	0.976		0.961	0.484
6	RF	89	0.995	0.95	0.176	0.527	0.485	0.944	0.944	0.968	0.977	0.816	0.944	0.979		0.972	0.521
6	RF	90	0.994	0.953	0.172	0.532	0.547	0.953	0.951	0.965	0.988	0.841	0.952	0.989		0.976	0.519
6	RF	91	0.995	0.951	0.181	0.504	0.480	0.947	0.947	0.969	0.978	0.823	0.947	0.978		0.973	0.499
6	RF	92	0.995	0.956	0.160	0.499	0.521	0.952	0.952	0.963	0.99	0.843	0.953	0.989		0.976	0.486
6	SVM	115	0.991	0.941	0.219	0.5	0.317	0.936	0.936	0.962	0.974	0.796	0.936	0.973		0.968	0.494
6	SVM	116	0.99	0.938	0.221	0.5	0.586	0.931	0.931	0.947	0.984	0.791	0.932	0.983		0.966	0.482
6	SVM	130	0.991	0.939	0.215	0.5	0.523	0.93	0.93	0.942	0.988	0.792	0.931	0.988		0.965	0.477
6	RBF	155	0.974	0.837	0.489	0.5	0.554	0.87	0.87	0.915	0.955	0.657	0.871	0.953		0.935	0.48
6	RBF	156	0.975	0.844	0.475	0.5	0.574	0.874	0.874	0.92	0.955	0.665	0.875	0.953		0.937	0.482
6	RBF	157	0.977	0.845	0.481	0.5	0.577	0.866	0.866	0.905	0.961	0.652	0.867	0.958		0.933	0.472
6	RBF	158	0.973	0.834	0.495	0.5	0.531	0.867	0.867	0.918	0.949	0.65	0.868	0.947		0.934	0.485
6	MLP	171	0.989	0.937	0.251	0.5	0.650	0.934	0.934	0.953	0.981	0.795	0.935	0.98		0.967	0.486
6	MARS	13	0.944	0.754	0.513	0.757	0.792	0.753	0.664	0.844	0.909	0.43	0.682	0.967	0.651	0.86	0.661
6	MARS	22	0.935	0.759	0.519	0.757	0.679	0.733	0.68	0.879	0.855	0.415	0.687	0.95	0.694	0.873	0.701
6	MARS	27	0.924	0.759	0.548	0.757	0.767	0.74	0.663	0.854	0.886	0.416	0.677	0.959	0.661	0.862	0.674

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
7	GAM	32	0.925	0.717	0.594	0.255	0.353	0.721	0.679	0.84	0.881	0.403	0.684	0.708	0.941	0.87	0.303
7	GAM	42	0.931	0.727	0.576	0.255	0.343	0.729	0.677	0.857	0.872	0.409	0.684	0.696	0.947	0.868	0.315
7	GAM	50	0.927	0.726	0.587	0.255	0.308	0.741	0.687	0.869	0.873	0.424	0.695	0.7	0.951	0.872	0.317
7	GAM	51	0.93	0.734	0.575	0.255	0.390	0.734	0.699	0.841	0.893	0.424	0.702	0.729	0.942	0.88	0.294
7	GAM	56	0.931	0.735	0.572	0.255	0.311	0.732	0.673	0.869	0.864	0.411	0.682	0.685	0.95	0.865	0.323
7	GLMPOLY	70	0.927	0.713	0.602	0.255	0.453	0.72	0.68	0.837	0.883	0.403	0.684	0.71	0.94	0.871	0.301
7	GLMPOLY	79	0.924	0.719	0.604	0.255	0.429	0.732	0.687	0.851	0.881	0.416	0.693	0.71	0.945	0.873	0.306
7	GLMPOLY	84	0.926	0.719	0.602	0.255	0.448	0.725	0.689	0.836	0.889	0.412	0.693	0.721	0.941	0.876	0.296
7	MDA	8	0.976	0.865	0.535	0.756	0.696	0.888	0.835	0.927	0.961	0.654	0.84	0.986	0.809	0.935	0.711
7	MDA	26	0.968	0.86	0.494	0.757	0.829	0.884	0.835	0.93	0.955	0.65	0.84	0.985	0.814	0.936	0.715
7	FDA	31	0.982	0.819	0.419	0.757	0.986	0.89	0.82	0.913	0.977	0.655	0.83	0.992	0.782	0.928	0.696
7	FDA	32	0.982	0.835	0.370	0.757	0.970	0.898	0.84	0.926	0.973	0.674	0.847	0.991	0.808	0.937	0.708
7	FDA	39	0.98	0.815	0.429	0.757	0.968	0.88	0.828	0.926	0.955	0.641	0.833	0.984	0.805	0.933	0.712
7	FDA	53	0.978	0.796	0.493	0.757	0.982	0.895	0.835	0.923	0.973	0.667	0.842	0.991	0.801	0.935	0.705
7	RPART	10	0.979	0.927	0.306	0.5	0.643	0.93	0.93	0.942	0.988	0.792	0.931	0.988	0.944	0.965	0.477
7	RPART	11	0.981	0.928	0.297	0.5	0.714	0.93	0.93	0.942	0.988	0.792	0.931	0.988	0.944	0.965	0.477
7	CART	30	0.973	0.926	0.285	0.5	0.570	0.92	0.92	0.933	0.987	0.77	0.921	0.986	0.936	0.96	0.473
7	CART	32	0.973	0.916	0.319	0.5	0.966	0.91	0.91	0.929	0.981	0.745	0.911	0.98	0.932	0.955	0.474
7	CART	34	0.979	0.93	0.277	0.505	0.934	0.924	0.924	0.941	0.983	0.774	0.924	0.983	0.942	0.962	0.483
7	CART	38	0.986	0.927	0.256	0.5	0.602	0.915	0.915	0.929	0.987	0.761	0.917	0.986	0.932	0.958	0.471
7	CART	39	0.981	0.927	0.259	0.5	0.529	0.921	0.921	0.931	0.99	0.776	0.923	0.989	0.935	0.961	0.471
7	CART	40	0.978	0.928	0.273	0.5	0.583	0.921	0.921	0.933	0.988	0.774	0.923	0.988	0.936	0.961	0.473
7	CART	43	0.977	0.921	0.300	0.5	0.584	0.914	0.914	0.931	0.982	0.754	0.915	0.982	0.935	0.957	0.475
7	CART	50	0.98	0.924	0.277	0.5	0.961	0.915	0.915	0.923	0.993	0.769	0.918	0.992	0.928	0.958	0.465
7	BRT	62	0.993	0.942	0.209	0.505	0.712	0.929	0.928	0.935	0.994	0.796	0.93	0.993	0.938	0.964	0.475
7	BRT	63	0.992	0.937	0.219	0.489	0.436	0.927	0.927	0.956	0.971	0.775	0.927	0.969	0.959	0.964	0.482
7	BRT	67	0.994	0.946	0.197	0.5	0.606	0.939	0.939	0.946	0.993	0.816	0.94	0.992	0.948	0.969	0.477
7	BRT	68	0.99	0.936	0.221	0.5	0.506	0.929	0.929	0.94	0.988	0.789	0.93	0.988	0.943	0.964	0.476
7	BRT	71	0.99	0.935	0.228	0.5	0.775	0.926	0.926	0.93	0.996	0.794	0.928	0.995	0.934	0.963	0.468
7	BRT	78	0.991	0.934	0.229	0.5	0.711	0.926	0.926	0.933	0.993	0.789	0.927	0.992	0.937	0.963	0.47
7	RF	90	0.996	0.961	0.147	0.505	0.518	0.963	0.963	0.972	0.991	0.871	0.963	0.991	0.972	0.981	0.495
7	RF	91	0.997	0.959	0.151	0.489	0.477	0.954	0.955	0.968	0.987	0.846	0.955	0.986	0.97	0.977	0.479
7	RF	92	0.995	0.949	0.180	0.499	0.490	0.944	0.944	0.96	0.984	0.818	0.944	0.983	0.961	0.972	0.488
7	RF	95	0.994	0.949	0.175	0.5	0.632	0.949	0.949	0.949	1	0.853	0.95	1	0.951	0.974	0.475
7	RF	99	0.991	0.94	0.206	0.5	0.756	0.93	0.93	0.931	0.999	0.809	0.932	0.998	0.936	0.965	0.467
7	RF	105	0.993	0.942	0.199	0.5	0.396	0.934	0.934	0.958	0.977	0.793	0.934	0.976	0.958	0.967	0.491
7	SVM	118	0.991	0.936	0.233	0.505	0.298	0.93	0.93	0.952	0.978	0.785	0.93	0.978	0.953	0.965	0.491

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
7	SVM	123	0.99	0.937	0.225	0.5	0.503	0.93	0.93	0.949	0.981	0.786	0.93	0.98	0.95	0.965	0.484
7	SVM	133	0.992	0.938	0.229	0.5	0.438	0.931	0.931	0.945	0.987	0.794	0.932	0.986	0.947	0.966	0.479
7	RBF	145	0.96	0.776	0.628	0.511	0.678	0.84	0.838	0.883	0.957	0.606	0.841	0.955	0.887	0.919	0.472
7	RBF	146	0.97	0.811	0.570	0.505	0.669	0.853	0.853	0.9	0.953	0.627	0.854	0.951	0.904	0.926	0.478
7	RBF	150	0.962	0.799	0.573	0.5	0.594	0.825	0.825	0.897	0.928	0.573	0.825	0.926	0.9	0.912	0.484
7	RBF	155	0.966	0.806	0.593	0.5	0.685	0.829	0.829	0.862	0.968	0.599	0.834	0.964	0.875	0.915	0.447
7	RBF	158	0.954	0.776	0.653	0.5	0.681	0.822	0.822	0.88	0.942	0.573	0.823	0.938	0.887	0.911	0.469
7	RBF	160	0.959	0.771	0.655	0.5	0.667	0.838	0.838	0.893	0.945	0.599	0.839	0.942	0.899	0.919	0.474
7	RBF	164	0.963	0.789	0.616	0.5	0.683	0.832	0.832	0.874	0.958	0.596	0.835	0.954	0.884	0.916	0.458
7	RBF	165	0.969	0.808	0.578	0.5	0.626	0.854	0.854	0.911	0.943	0.625	0.854	0.941	0.914	0.927	0.484
7	MLP	179	0.989	0.943	0.237	0.5	0.725	0.936	0.936	0.945	0.991	0.808	0.937	0.991	0.947	0.968	0.477
7	MLP	183	0.988	0.934	0.250	0.5	0.669	0.931	0.931	0.939	0.993	0.801	0.933	0.992	0.942	0.966	0.473
7	MLP	189	0.99	0.94	0.238	0.5	0.595	0.934	0.934	0.936	0.999	0.817	0.936	0.998	0.94	0.967	0.469

Appendix J

Complementary (AdaptWest) datasets selected models and metrics

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
3	BRT	2	0.988	0.916	0.286	0.5	0.436	0.901	0.901	0.943	0.959	0.718	0.902	0.958	0.944	0.951	0.492
3	BRT	12	0.983	0.908	0.312	0.5	0.525	0.904	0.904	0.934	0.971	0.728	0.905	0.969	0.936	0.952	0.482
3	GAM	31	0.912	0.688	0.661	0.254	0.347	0.68	0.626	0.828	0.853	0.35	0.635	0.656	0.936	0.846	0.32
3	GAM	34	0.916	0.691	0.625	0.253	0.273	0.688	0.606	0.87	0.818	0.352	0.625	0.618	0.949	0.831	0.357
3	GAM	36	0.914	0.691	0.627	0.251	0.306	0.692	0.628	0.847	0.846	0.36	0.639	0.648	0.943	0.846	0.328
3	GAM	39	0.918	0.697	0.616	0.254	0.223	0.696	0.606	0.887	0.809	0.36	0.628	0.612	0.955	0.829	0.368
3	GAM	40	0.916	0.691	0.623	0.254	0.272	0.688	0.623	0.847	0.841	0.355	0.634	0.644	0.942	0.842	0.334
3	GAM	42	0.92	0.7	0.612	0.254	0.230	0.688	0.603	0.876	0.812	0.352	0.623	0.613	0.951	0.828	0.363
3	GAM	43	0.919	0.7	0.634	0.254	0.263	0.696	0.615	0.875	0.822	0.362	0.633	0.625	0.951	0.835	0.355
3	GAM	45	0.916	0.691	0.627	0.254	0.260	0.687	0.605	0.871	0.817	0.351	0.624	0.617	0.949	0.83	0.358
3	GAM	53	0.916	0.697	0.646	0.253	0.265	0.681	0.607	0.854	0.827	0.345	0.622	0.626	0.943	0.834	0.346
3	GAM	56	0.912	0.688	0.634	0.253	0.237	0.689	0.6	0.881	0.808	0.352	0.622	0.609	0.952	0.826	0.367
3	GLM POLY	66	0.911	0.681	0.641	0.254	0.263	0.681	0.61	0.853	0.829	0.346	0.624	0.628	0.943	0.835	0.344
3	GLM POLY	67	0.913	0.684	0.634	0.254	0.243	0.683	0.602	0.868	0.816	0.346	0.62	0.615	0.948	0.829	0.358

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
3	GLM POLY	71	0.917	0.695	0.625	0.254	0.294	0.692	0.626	0.85	0.842	0.359	0.638	0.646	0.943	0.844	0.334
3	MLP	120	0.984	0.911	0.335	0.512	0.552	0.9	0.899	0.925	0.975	0.72	0.9	0.975	0.925	0.949	0.486
3	MLP	124	0.98	0.903	0.346	0.5	0.535	0.891	0.891	0.925	0.966	0.7	0.892	0.965	0.928	0.946	0.479
3	MLP	131	0.981	0.909	0.409	0.5	0.606	0.894	0.894	0.924	0.971	0.708	0.895	0.969	0.927	0.947	0.476
3	RBF	88	0.964	0.818	0.541	0.5	0.580	0.863	0.863	0.928	0.935	0.64	0.863	0.935	0.928	0.932	0.496
3	RBF	105	0.963	0.817	0.600	0.5	0.640	0.85	0.85	0.903	0.947	0.619	0.851	0.945	0.907	0.925	0.478
3	RBF	111	0.97	0.819	0.545	0.5	0.652	0.857	0.857	0.892	0.965	0.64	0.859	0.962	0.9	0.929	0.464
3	RF	30	0.989	0.925	0.251	0.5	0.622	0.913	0.913	0.924	0.99	0.761	0.915	0.989	0.928	0.957	0.467
3	RF	33	0.992	0.939	0.212	0.504	0.518	0.926	0.926	0.951	0.975	0.774	0.926	0.975	0.951	0.963	0.491
3	RF	34	0.993	0.937	0.213	0.49	0.420	0.926	0.926	0.961	0.965	0.771	0.926	0.963	0.963	0.963	0.489
3	RF	35	0.993	0.939	0.209	0.501	0.446	0.928	0.928	0.959	0.969	0.776	0.928	0.969	0.959	0.964	0.496
3	RF	36	0.995	0.947	0.187	0.512	0.509	0.943	0.942	0.966	0.977	0.813	0.943	0.978	0.964	0.971	0.505
3	RF	47	0.989	0.921	0.254	0.5	0.664	0.903	0.903	0.918	0.985	0.736	0.905	0.984	0.923	0.951	0.466
4	GLM	15	0.881	0.597	0.745	0.253	0.294	0.701	0.603	0.901	0.8	0.366	0.629	0.604	0.96	0.825	0.377
4	GLM	28	0.877	0.591	0.754	0.253	0.314	0.701	0.611	0.889	0.812	0.366	0.633	0.616	0.956	0.831	0.366
4	GAM	42	0.933	0.733	0.575	0.253	0.243	0.748	0.663	0.91	0.838	0.425	0.681	0.655	0.965	0.856	0.351
4	GAM	56	0.931	0.73	0.572	0.253	0.253	0.746	0.662	0.907	0.839	0.423	0.68	0.656	0.964	0.856	0.35
4	GLM POLY	62	0.916	0.687	0.630	0.25	0.364	0.702	0.638	0.851	0.851	0.372	0.649	0.655	0.945	0.851	0.325
4	GLM POLY	70	0.918	0.69	0.629	0.253	0.329	0.7	0.625	0.867	0.833	0.367	0.64	0.637	0.949	0.842	0.344
4	GLM POLY	84	0.921	0.694	0.625	0.253	0.344	0.702	0.628	0.869	0.834	0.37	0.643	0.639	0.949	0.843	0.344
4	MDA	16	0.98	0.839	0.550	0.755	0.900	0.885	0.838	0.931	0.955	0.652	0.842	0.984	0.817	0.936	0.713
4	MDA	19	0.974	0.84	0.543	0.755	0.914	0.874	0.834	0.934	0.941	0.636	0.837	0.98	0.821	0.935	0.719
4	MDA	25	0.969	0.874	0.887	0.755	0.679	0.865	0.847	0.951	0.914	0.637	0.847	0.971	0.859	0.942	0.739
4	FDA	47	0.962	0.847	0.399	0.755	0.866	0.858	0.833	0.944	0.914	0.618	0.835	0.971	0.841	0.936	0.734
4	FDA	48	0.951	0.814	0.467	0.755	0.889	0.874	0.838	0.938	0.936	0.639	0.841	0.978	0.831	0.938	0.724
4	RPART	3	0.979	0.911	0.346	0.497	0.714	0.9	0.9	0.92	0.98	0.726	0.902	0.979	0.925	0.95	0.467
4	RPART	8	0.984	0.921	0.304	0.486	0.625	0.905	0.906	0.933	0.972	0.73	0.906	0.969	0.939	0.953	0.468
4	RPART	21	0.98	0.912	0.340	0.497	0.538	0.903	0.903	0.942	0.961	0.721	0.903	0.959	0.944	0.952	0.488
4	CART	29	0.962	0.888	0.410	0.497	0.973	0.876	0.877	0.907	0.97	0.676	0.878	0.967	0.913	0.938	0.466
4	CART	32	0.957	0.886	0.404	0.497	0.527	0.875	0.876	0.902	0.973	0.676	0.878	0.971	0.909	0.938	0.462
4	CART	43	0.959	0.886	0.406	0.497	0.618	0.876	0.877	0.914	0.962	0.671	0.878	0.96	0.919	0.938	0.474
4	CART	54	0.956	0.884	0.421	0.497	0.971	0.875	0.875	0.9	0.975	0.677	0.878	0.973	0.908	0.938	0.46
4	BRT	64	0.992	0.937	0.220	0.486	0.384	0.924	0.924	0.96	0.964	0.767	0.924	0.962	0.962	0.962	0.485

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
4	RF	89	0.993	0.94	0.203	0.499	0.472	0.93	0.93	0.961	0.97	0.782	0.93	0.969	0.961	0.965	0.495
4	RF	90	0.995	0.946	0.185	0.509	0.412	0.941	0.94	0.96	0.981	0.81	0.941	0.981	0.96	0.97	0.498
4	RF	91	0.991	0.934	0.225	0.505	0.497	0.923	0.923	0.948	0.975	0.767	0.923	0.974	0.949	0.961	0.492
4	RF	92	0.996	0.953	0.164	0.486	0.450	0.948	0.948	0.973	0.976	0.826	0.948	0.974	0.974	0.974	0.485
4	MLP	172	0.986	0.919	0.285	0.497	0.428	0.908	0.908	0.932	0.977	0.739	0.909	0.975	0.935	0.954	0.475
4	MLP	176	0.985	0.922	0.296	0.486	0.619	0.912	0.913	0.933	0.98	0.749	0.914	0.977	0.939	0.957	0.464
4	MLP	182	0.985	0.917	0.298	0.497	0.533	0.909	0.909	0.922	0.987	0.749	0.911	0.986	0.927	0.954	0.465
4	MLP	189	0.985	0.916	0.311	0.497	0.493	0.907	0.907	0.941	0.966	0.732	0.908	0.965	0.943	0.954	0.485
4	MLP	193	0.98	0.905	0.411	0.497	0.790	0.904	0.905	0.922	0.982	0.736	0.906	0.981	0.927	0.952	0.467
5	GLM	6	0.791	0.421	0.943	0.259	0.264	0.539	0.452	0.811	0.728	0.206	0.48	0.51	0.917	0.749	0.412
5	GLM	15	0.787	0.407	0.939	0.253	0.268	0.551	0.453	0.831	0.721	0.216	0.486	0.502	0.926	0.748	0.419
5	GAM	41	0.926	0.724	0.607	0.253	0.334	0.73	0.675	0.862	0.869	0.409	0.683	0.69	0.949	0.867	0.317
5	GAM	42	0.932	0.736	0.571	0.253	0.313	0.739	0.686	0.865	0.875	0.422	0.693	0.7	0.95	0.872	0.313
5	GAM	43	0.932	0.733	0.566	0.253	0.312	0.735	0.67	0.876	0.859	0.412	0.681	0.678	0.953	0.863	0.328
5	GAM	44	0.929	0.728	0.575	0.253	0.301	0.74	0.677	0.878	0.863	0.42	0.687	0.684	0.954	0.866	0.325
5	GAM	56	0.933	0.739	0.569	0.253	0.281	0.745	0.68	0.884	0.862	0.425	0.691	0.684	0.956	0.867	0.327
5	GLM POLY	70	0.925	0.721	0.598	0.253	0.307	0.738	0.676	0.875	0.864	0.418	0.686	0.685	0.953	0.866	0.324
5	GLM POLY	84	0.926	0.726	0.592	0.253	0.380	0.758	0.715	0.866	0.892	0.451	0.719	0.731	0.951	0.885	0.3
5	MDA	11	0.965	0.834	0.623	0.756	0.923	0.896	0.847	0.932	0.964	0.672	0.852	0.988	0.822	0.94	0.713
5	MDA	13	0.981	0.85	0.394	0.755	0.955	0.899	0.833	0.918	0.982	0.675	0.842	0.994	0.794	0.933	0.697
5	MDA	19	0.98	0.86	0.653	0.755	0.789	0.902	0.833	0.916	0.986	0.682	0.843	0.995	0.792	0.933	0.695
5	MDA	20	0.977	0.854	0.465	0.755	0.844	0.896	0.846	0.932	0.964	0.672	0.851	0.988	0.822	0.94	0.713
5	MDA	22	0.974	0.862	0.659	0.755	0.853	0.898	0.83	0.916	0.982	0.672	0.839	0.994	0.791	0.932	0.696
5	FDA	44	0.836	0.653	0.775	0.755	0.743	0.708	0.702	0.922	0.786	0.413	0.702	0.93	0.765	0.889	0.749
5	RPART	7	0.984	0.926	0.279	0.496	0.571	0.911	0.911	0.944	0.968	0.74	0.912	0.966	0.946	0.956	0.484
5	RPART	8	0.981	0.927	0.301	0.501	0.625	0.915	0.915	0.939	0.976	0.752	0.916	0.975	0.941	0.958	0.482
5	RPART	21	0.976	0.917	0.339	0.498	0.643	0.91	0.911	0.929	0.981	0.746	0.912	0.98	0.933	0.955	0.472
5	CART	36	0.971	0.907	0.327	0.501	0.992	0.897	0.897	0.902	0.995	0.741	0.901	0.995	0.91	0.949	0.454
5	CART	50	0.967	0.903	0.355	0.498	0.994	0.894	0.894	0.904	0.99	0.726	0.898	0.989	0.913	0.947	0.455
5	BRT	62	0.99	0.932	0.233	0.493	0.522	0.92	0.92	0.944	0.976	0.762	0.921	0.975	0.947	0.96	0.477
5	BRT	64	0.992	0.937	0.214	0.501	0.473	0.924	0.924	0.95	0.975	0.77	0.925	0.974	0.951	0.962	0.488
5	BRT	77	0.988	0.93	0.242	0.498	0.693	0.922	0.922	0.931	0.991	0.78	0.924	0.991	0.935	0.961	0.468
5	RF	89	0.991	0.938	0.221	0.498	0.560	0.935	0.935	0.95	0.985	0.8	0.936	0.984	0.952	0.968	0.48
5	RF	90	0.994	0.949	0.190	0.493	0.454	0.944	0.944	0.966	0.979	0.818	0.944	0.978	0.967	0.972	0.487

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
5	RF	91	0.994	0.946	0.187	0.496	0.558	0.937	0.937	0.956	0.981	0.801	0.937	0.98	0.958	0.969	0.484
5	RF	92	0.995	0.954	0.160	0.501	0.560	0.951	0.951	0.961	0.99	0.841	0.952	0.99	0.962	0.976	0.486
5	SVM	114	0.98	0.924	0.274	0.498	0.721	0.922	0.922	0.932	0.99	0.778	0.924	0.989	0.936	0.961	0.47
5	SVM	133	0.988	0.935	0.246	0.498	0.576	0.928	0.928	0.943	0.985	0.785	0.929	0.985	0.945	0.964	0.477
5	RBF	147	0.971	0.825	0.536	0.496	0.569	0.873	0.873	0.929	0.944	0.66	0.873	0.942	0.931	0.937	0.489
5	RBF	148	0.966	0.821	0.576	0.501	0.605	0.86	0.86	0.89	0.97	0.648	0.863	0.967	0.898	0.93	0.461
5	RBF	160	0.968	0.821	0.575	0.498	0.606	0.868	0.868	0.909	0.959	0.655	0.869	0.957	0.914	0.934	0.473
5	MLP	170	0.987	0.923	0.258	0.498	0.579	0.912	0.912	0.932	0.98	0.748	0.913	0.978	0.936	0.956	0.475
5	MLP	189	0.99	0.93	0.236	0.498	0.385	0.919	0.919	0.954	0.965	0.757	0.919	0.964	0.955	0.96	0.493
5	MLP	190	0.986	0.92	0.281	0.498	0.745	0.913	0.913	0.931	0.982	0.753	0.915	0.981	0.935	0.957	0.472
5	MARS	1	0.931	0.782	0.494	0.756	0.742	0.775	0.741	0.912	0.864	0.482	0.744	0.954	0.76	0.9	0.722
5	MARS	17	0.932	0.75	0.530	0.755	0.819	0.751	0.677	0.86	0.891	0.43	0.69	0.961	0.674	0.868	0.676
5	MARS	28	0.934	0.759	0.524	0.755	0.776	0.758	0.712	0.894	0.864	0.45	0.717	0.953	0.725	0.887	0.709
6	GAM	29	0.925	0.717	0.596	0.254	0.390	0.718	0.683	0.829	0.889	0.404	0.687	0.718	0.939	0.874	0.293
6	GAM	42	0.932	0.734	0.569	0.253	0.431	0.721	0.704	0.81	0.911	0.42	0.705	0.755	0.934	0.885	0.272
6	GAM	55	0.928	0.726	0.583	0.253	0.343	0.728	0.683	0.847	0.881	0.411	0.689	0.707	0.944	0.872	0.303
6	GLM POLY	57	0.917	0.699	0.622	0.254	0.430	0.711	0.666	0.837	0.875	0.39	0.672	0.694	0.94	0.865	0.306
6	GLM POLY	77	0.919	0.699	0.617	0.253	0.422	0.714	0.662	0.847	0.868	0.391	0.67	0.685	0.943	0.862	0.314
6	GLM POLY	83	0.922	0.71	0.606	0.253	0.346	0.723	0.664	0.862	0.862	0.399	0.673	0.679	0.948	0.862	0.322
6	MDA	12	0.973	0.863	0.493	0.756	0.780	0.865	0.842	0.947	0.918	0.633	0.843	0.973	0.849	0.94	0.736
6	MDA	14	0.981	0.887	0.431	0.755	0.596	0.895	0.881	0.963	0.932	0.698	0.882	0.978	0.891	0.956	0.744
6	MDA	15	0.975	0.866	0.554	0.755	0.855	0.877	0.849	0.946	0.932	0.65	0.85	0.977	0.847	0.942	0.731
6	MDA	21	0.981	0.877	0.453	0.755	0.738	0.898	0.872	0.953	0.945	0.692	0.873	0.982	0.867	0.951	0.733
6	FDA	41	0.977	0.816	0.416	0.755	0.980	0.893	0.827	0.916	0.977	0.663	0.836	0.992	0.79	0.931	0.697
6	FDA	42	0.969	0.809	0.454	0.755	0.984	0.877	0.8	0.9	0.977	0.63	0.812	0.992	0.76	0.919	0.685
6	RPART	18	0.983	0.919	0.325	0.498	0.636	0.905	0.905	0.94	0.965	0.726	0.905	0.964	0.942	0.952	0.485
6	RPART	25	0.976	0.908	0.388	0.498	0.647	0.905	0.905	0.931	0.974	0.73	0.905	0.972	0.934	0.952	0.477
6	RPART	28	0.981	0.914	0.331	0.498	0.625	0.902	0.902	0.937	0.965	0.72	0.902	0.964	0.939	0.951	0.484
6	CART	37	0.972	0.891	0.374	0.498	0.459	0.874	0.874	0.938	0.936	0.661	0.874	0.935	0.939	0.937	0.5
6	CART	46	0.964	0.887	0.399	0.498	0.786	0.88	0.88	0.922	0.958	0.675	0.88	0.956	0.925	0.94	0.48
6	CART	53	0.961	0.883	0.411	0.498	0.468	0.872	0.872	0.923	0.949	0.659	0.873	0.947	0.926	0.936	0.485
6	CART	55	0.966	0.886	0.399	0.498	0.808	0.877	0.877	0.912	0.965	0.673	0.878	0.963	0.917	0.938	0.471

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
6	CART	56	0.963	0.887	0.399	0.498	0.449	0.878	0.878	0.934	0.945	0.669	0.878	0.943	0.935	0.939	0.493
6	BRT	61	0.993	0.937	0.211	0.48	0.545	0.922	0.922	0.947	0.975	0.766	0.923	0.972	0.952	0.961	0.467
6	BRT	62	0.992	0.931	0.230	0.502	0.419	0.916	0.916	0.959	0.957	0.749	0.916	0.958	0.959	0.958	0.502
6	BRT	64	0.99	0.928	0.244	0.506	0.530	0.916	0.916	0.932	0.985	0.761	0.917	0.984	0.934	0.958	0.479
6	BRT	74	0.992	0.93	0.230	0.498	0.551	0.918	0.918	0.943	0.975	0.757	0.918	0.974	0.945	0.959	0.482
6	BRT	84	0.991	0.929	0.235	0.498	0.565	0.915	0.915	0.935	0.98	0.754	0.916	0.978	0.938	0.957	0.476
6	RF	89	0.995	0.95	0.177	0.48	0.497	0.945	0.945	0.962	0.983	0.821	0.946	0.981	0.966	0.973	0.471
6	RF	90	0.994	0.943	0.193	0.502	0.476	0.933	0.933	0.965	0.968	0.788	0.933	0.968	0.965	0.966	0.5
6	RF	91	0.994	0.94	0.196	0.473	0.496	0.924	0.924	0.955	0.969	0.768	0.924	0.965	0.96	0.962	0.468
6	RF	92	0.992	0.94	0.209	0.506	0.482	0.934	0.934	0.953	0.982	0.795	0.934	0.982	0.953	0.967	0.491
6	RF	102	0.993	0.937	0.213	0.498	0.482	0.928	0.928	0.959	0.969	0.777	0.928	0.969	0.96	0.964	0.493
6	SVM	130	0.985	0.914	0.301	0.498	0.499	0.902	0.902	0.937	0.965	0.72	0.902	0.964	0.939	0.951	0.484
6	SVM	132	0.98	0.902	0.336	0.498	0.595	0.891	0.891	0.918	0.974	0.705	0.893	0.972	0.923	0.946	0.47
6	SVM	140	0.986	0.907	0.315	0.498	0.266	0.89	0.89	0.944	0.946	0.693	0.89	0.945	0.945	0.945	0.497
6	RBF	143	0.946	0.764	0.697	0.498	0.668	0.83	0.83	0.894	0.936	0.583	0.831	0.933	0.899	0.915	0.478
6	RBF	147	0.958	0.787	0.638	0.473	0.599	0.828	0.829	0.903	0.925	0.578	0.829	0.915	0.914	0.915	0.467
6	RBF	153	0.953	0.756	0.678	0.498	0.599	0.827	0.827	0.891	0.936	0.578	0.828	0.932	0.897	0.913	0.476
6	RBF	154	0.948	0.78	0.632	0.498	0.640	0.844	0.845	0.912	0.933	0.607	0.845	0.931	0.914	0.922	0.488
6	RBF	156	0.94	0.758	0.716	0.498	0.617	0.827	0.827	0.907	0.92	0.575	0.827	0.918	0.909	0.913	0.492
6	RBF	160	0.949	0.763	0.679	0.498	0.645	0.822	0.823	0.89	0.933	0.57	0.823	0.929	0.895	0.911	0.477
6	RBF	161	0.958	0.775	0.648	0.498	0.650	0.849	0.849	0.9	0.949	0.618	0.85	0.946	0.905	0.924	0.474
6	RBF	164	0.957	0.774	0.680	0.498	0.639	0.828	0.828	0.888	0.94	0.582	0.829	0.936	0.894	0.914	0.472
6	RBF	165	0.94	0.734	0.753	0.498	0.657	0.836	0.836	0.897	0.939	0.593	0.836	0.935	0.902	0.918	0.477
6	MLP	174	0.986	0.913	0.305	0.502	0.426	0.897	0.897	0.941	0.956	0.709	0.897	0.956	0.941	0.949	0.494
6	MLP	178	0.985	0.909	0.297	0.498	0.473	0.897	0.897	0.921	0.977	0.718	0.899	0.975	0.925	0.949	0.47
6	MLP	182	0.982	0.908	0.335	0.498	0.473	0.897	0.897	0.937	0.961	0.71	0.898	0.959	0.939	0.949	0.486
6	MLP	186	0.99	0.925	0.250	0.498	0.380	0.909	0.909	0.95	0.959	0.733	0.909	0.958	0.951	0.955	0.493
6	MLP	189	0.981	0.909	0.332	0.498	0.546	0.897	0.897	0.931	0.966	0.712	0.898	0.965	0.934	0.949	0.48
6	MLP	193	0.982	0.903	0.336	0.498	0.691	0.896	0.896	0.922	0.974	0.713	0.897	0.972	0.926	0.948	0.472
6	MLP	195	0.984	0.914	0.310	0.498	0.475	0.903	0.903	0.932	0.971	0.725	0.904	0.969	0.935	0.952	0.479
7	GAM	33	0.924	0.709	0.605	0.259	0.324	0.719	0.658	0.866	0.854	0.394	0.669	0.674	0.948	0.857	0.333
7	GAM	41	0.921	0.701	0.610	0.254	0.344	0.708	0.657	0.842	0.866	0.384	0.664	0.681	0.942	0.86	0.314
7	GAM	42	0.923	0.709	0.604	0.254	0.305	0.733	0.668	0.876	0.857	0.41	0.679	0.675	0.953	0.862	0.329
7	GAM	46	0.921	0.696	0.607	0.254	0.336	0.714	0.654	0.857	0.857	0.388	0.664	0.671	0.946	0.857	0.324
7	GAM	49	0.923	0.703	0.608	0.254	0.365	0.714	0.657	0.853	0.861	0.388	0.666	0.676	0.945	0.859	0.32
7	GAM	56	0.93	0.722	0.576	0.254	0.295	0.728	0.655	0.882	0.846	0.402	0.669	0.66	0.955	0.855	0.339
7	GLM POLY	70	0.911	0.681	0.655	0.254	0.421	0.716	0.672	0.839	0.877	0.397	0.677	0.699	0.941	0.867	0.305

Sequence	Algorithm	modelID	AUC	COR	Deviance	Prevalence	threshold	TSS	Kappa	sensitivity	specificity	NMI	phi	ppv	npv	ccr	prevalence
7	GLM POLY	77	0.906	0.669	0.671	0.254	0.425	0.706	0.657	0.838	0.868	0.382	0.663	0.683	0.94	0.86	0.311
7	MDA	2	0.957	0.799	0.509	0.756	0.883	0.812	0.767	0.912	0.9	0.529	0.771	0.966	0.767	0.909	0.713
7	MDA	8	0.954	0.786	0.536	0.754	0.902	0.808	0.752	0.898	0.91	0.518	0.759	0.968	0.744	0.901	0.699
7	MDA	10	0.941	0.777	0.681	0.756	0.830	0.814	0.781	0.924	0.891	0.54	0.783	0.963	0.79	0.916	0.724
7	MDA	19	0.954	0.771	0.659	0.755	0.903	0.841	0.775	0.9	0.941	0.566	0.784	0.979	0.753	0.91	0.694
7	MDA	21	0.951	0.803	0.542	0.755	0.827	0.805	0.774	0.923	0.882	0.528	0.777	0.96	0.789	0.913	0.726
7	RPART	6	0.981	0.918	0.331	0.501	0.500	0.902	0.902	0.956	0.947	0.719	0.902	0.947	0.955	0.951	0.505
7	RPART	11	0.977	0.913	0.340	0.498	0.667	0.91	0.911	0.937	0.974	0.741	0.911	0.973	0.939	0.955	0.48
7	CART	34	0.948	0.867	0.470	0.501	0.972	0.861	0.86	0.893	0.967	0.647	0.863	0.965	0.9	0.93	0.464
7	CART	35	0.958	0.879	0.409	0.504	0.994	0.865	0.864	0.873	0.993	0.683	0.871	0.992	0.885	0.932	0.443
7	CART	37	0.956	0.879	0.428	0.498	0.974	0.862	0.862	0.881	0.981	0.662	0.867	0.979	0.892	0.931	0.448
7	CART	39	0.965	0.897	0.378	0.498	0.974	0.877	0.877	0.89	0.987	0.693	0.881	0.985	0.9	0.938	0.45
7	CART	42	0.951	0.873	0.446	0.498	0.977	0.86	0.861	0.881	0.98	0.658	0.865	0.977	0.892	0.93	0.449
7	BRT	61	0.993	0.937	0.211	0.48	0.545	0.922	0.922	0.947	0.975	0.766	0.923	0.972	0.952	0.961	0.467
7	BRT	62	0.992	0.931	0.230	0.502	0.419	0.916	0.916	0.959	0.957	0.749	0.916	0.958	0.959	0.958	0.502
7	BRT	64	0.99	0.928	0.244	0.506	0.530	0.916	0.916	0.932	0.985	0.761	0.917	0.984	0.934	0.958	0.479
7	BRT	74	0.992	0.93	0.230	0.498	0.551	0.918	0.918	0.943	0.975	0.757	0.918	0.974	0.945	0.959	0.482
7	BRT	84	0.991	0.929	0.235	0.498	0.565	0.915	0.915	0.935	0.98	0.754	0.916	0.978	0.938	0.957	0.476
7	RF	89	0.993	0.939	0.210	0.494	0.463	0.936	0.936	0.961	0.975	0.796	0.936	0.974	0.962	0.968	0.487
7	RF	90	0.995	0.95	0.173	0.501	0.465	0.941	0.941	0.971	0.97	0.809	0.941	0.97	0.971	0.971	0.502
7	RF	91	0.992	0.936	0.215	0.504	0.454	0.928	0.928	0.95	0.978	0.78	0.928	0.977	0.951	0.964	0.49
7	RF	92	0.994	0.945	0.192	0.493	0.498	0.943	0.943	0.966	0.977	0.813	0.943	0.976	0.967	0.971	0.488
7	RF	93	0.989	0.925	0.240	0.498	0.559	0.925	0.925	0.938	0.987	0.781	0.926	0.986	0.942	0.963	0.474
7	RF	95	0.99	0.932	0.224	0.498	0.582	0.925	0.925	0.938	0.987	0.781	0.926	0.986	0.942	0.963	0.474
7	SVM	123	0.988	0.93	0.259	0.498	0.700	0.921	0.921	0.934	0.987	0.772	0.922	0.986	0.938	0.96	0.472
7	RBF	151	0.961	0.79	0.639	0.498	0.605	0.852	0.852	0.918	0.934	0.62	0.852	0.933	0.92	0.926	0.49
7	RBF	162	0.971	0.822	0.567	0.498	0.640	0.85	0.85	0.887	0.964	0.628	0.853	0.96	0.896	0.925	0.46

Supplementary Material

Example script:

This is the standard workflow within RStudio for sequence seven of the primary dataset

```
install.package('plyr')
install.package('usdm')
install.package('caret')
install.package('biomod2')
install.package('gridExtra')
install.package('dplyr')
install.package('CENFA')
install.package('ENMeval')
install.package('sdm')
installAll()

library(sdm)
library(rgdal)
library(raster)
library(plyr)
library(usdm)
library(caret)
library(biomod2)
library(ENMeval)
library(gridExtra)
library(dplyr)
library(CENFA)
library(corrplot)
```

```
 Ixodes_Scapularis_Shapefile <-
readOGR("C:/Users/jwestcot/AGIS9000/Ixodes_Scapularis/Shapefile","Ixodes_Scapularis")

Raster_Stack <- stack(List_Rasters)

vifcor_Rasters <- vifcor(Raster_Stack, th = 0.7)

Rasters_Present <- exclude(Raster_Stack,vifcor_Rasters)

PearsonR <- layerStats(Raster_Stack, 'pearson', na.rm=T)

Corr_Matrix = PearsonR$'pearson correlation coefficient'

sdmData_GLM_GAM_GLMPoly <- sdmData(formula = I_Scap~., train=
Ixodes_Scapularis_Shapefile, predictors=Rasters_Present, bg = list(n=10000,
method='gRandom'))

IScap_Shapefile_Turned_DF <- as.data.frame(Ixodes_Scapularis_Shapefile)

Background_Grid_MDA_FDA <- BIOMOD_FormatingData(resp.var = rep(1,
nrow(IScap_Shapefile_Turned_DF)),
expl.var = Rasters_Present,
resp.xy = IScap_Shapefile_Turned_DF[,c("coords.x1","coords.x2")], # colnames of coords in DF
resp.name = "Ixodes_Scapularis", # species name in DF
PA.strategy = "sre",
PA.nb.rep = 1,
PA.nb.absences = 1100,
na.rm = FALSE)

Background_MDA_FDA <- cbind(sp = c(rep(1,nrow(IScap_Shapefile_Turned_DF)),rep(0,
1100)), Background_Grid_MDA_FDA@coord)

names(Background_MDA_FDA)[1] <- names(IScap_Shapefile_Turned_DF)[1]

coordinates(Background_MDA_FDA) <- 2:3

sdmData_MDA_FDA <- sdmData(formula = I_Scap~., train= Background_MDA_FDA)

Background_Grid_MARS <- BIOMOD_FormatingData(resp.var = rep(1,
nrow(IScap_Shapefile_Turned_DF)),
expl.var = Rasters_Present,
resp.xy = IScap_Shapefile_Turned_DF[,c("coords.x1","coords.x2")],
resp.name = "Ixodes_Scapularis",
PA.strategy = "disk",
```

```
PA.dist.min = 220000,
PA.nb.rep = 1,
PA.nb.absences = 1100,
na.rm = FALSE)

Background_MARS <- cbind(sp = c(rep(1,nrow(IScap_Shapefile_Turned_DF)),rep(0, 1100)),
Background_Grid_MARS@coord)

names(Background_MARS)[1] <- names(IScap_Shapefile_Turned_DF)[1]

coordinates(Background_MARS) <- 2:3

sdmData_MARS <- sdmData(formula = I_Scap~., train= Background_MARS)

Background_Grid_RPART_CART_BRT_RF_SVM_RBF_MLP <-
BIOMOD_FormatingData(resp.var = rep(1, nrow(IScap_Shapefile_Turned_DF)),
expl.var = Rasters_Present,
resp.xy = IScap_Shapefile_Turned_DF[,c("coords.x1","coords.x2")],
resp.name = "Ixodes_Scapularis",
PA.strategy = "sre",
PA.nb.rep = 1,
PA.nb.absences = 3425,
na.rm = FALSE)

Background_RPART_CART_BRT_RF_SVM_RBF_MLP <- cbind(sp =
c(rep(1,nrow(IScap_Shapefile_Turned_DF)),rep(0, 3425)),
Background_Grid_RPART_CART_BRT_RF_SVM_RBF_MLP@coord)

names(Background_RPART_CART_BRT_RF_SVM_RBF_MLP)[1] <-
names(IScap_Shapefile_Turned_DF)[1]

coordinates(Background_RPART_CART_BRT_RF_SVM_RBF_MLP) <- 2:3

sdmData_RPART_CART_BRT_RF_SVM_RBF_MLP <- sdmData(formula = I_Scap~., train=
Background_RPART_CART_BRT_RF_SVM_RBF_MLP)

DF_GLM_GAM_GLMPoly <- as.data.frame(sdmData_GLM_GAM_GLMPoly)

DF_MDA_FDA <- as.data.frame(sdmData_MDA_FDA)

DF_MARS <- as.data.frame(sdmData_MARS)

DF_RPART_CART_BRT_RF_SVM_RBF_MLP <-
as.data.frame(sdmDaa_RPART_CART_BRT_RF_SVM_RBF_MLP)
```

```
DF_GLM_GAM_GLMPoly <- subset(DF_GLM_GAM_GLMPoly, select = -rID)
DF_GLM_GAM_GLMPoly <- subset(DF_GLM_GAM_GLMPoly, select = -coords.x1)
DF_GLM_GAM_GLMPoly <- subset(DF_GLM_GAM_GLMPoly, select = -coords.x2)

DF_MDA_FDA <- subset(DF_MDA_FDA, select = -rID)
DF_MDA_FDA <- subset(DF_MDA_FDA, select = -coords.x1)
DF_MDA_FDA <- subset(DF_MDA_FDA, select = -coords.x2)

DF_MARS <- subset(DF_MARS, select = -rID)
DF_MARS <- subset(DF_MARS, select = -coords.x1)
DF_MARS <- subset(DF_MARS, select = -coords.x2)

DF_RPART_CART_BRT_RF_SVM_RBF_MLP <-
subset(DF_RPART_CART_BRT_RF_SVM_RBF_MLP, select = -rID)
DF_RPART_CART_BRT_RF_SVM_RBF_MLP <-
subset(DF_RPART_CART_BRT_RF_SVM_RBF_MLP, select = -coords.x1)
DF_RPART_CART_BRT_RF_SVM_RBF_MLP <-
subset(DF_RPART_CART_BRT_RF_SVM_RBF_MLP, select = -coords.x2)

DF_GLM_GAM_GLMPoly$Ixodes_Scapularis <-
as.factor(DF_GLM_GAM_GLMPoly$Ixodes_Scapularis)
DF_MDA_FDA$Ixodes_Scapularis <- as.factor(DF_MDA_FDA$Ixodes_Scapularis)
DF_MARS$Ixodes_Scapularis <- as.factor(DF_MARS$Ixodes_Scapularis)
DF_RPART_CART_BRT_RF_SVM_RBF_MLP$Ixodes_Scapularis <-
as.factor(DF_RPART_CART_BRT_RF_SVM_RBF_MLP$Ixodes_Scapularis)

levels(DF_GLM_GAM_GLMPoly$Ixodes_Scapularis) <- c("A", "P") # A = Absence, P =
Presence
levels(DF_MDA_FDA$Ixodes_Scapularis) <- c("A", "P")
levels(DF_MARS$Ixodes_Scapularis) <- c("A", "P")
levels(DF_RPART_CART_BRT_RF_SVM_RBF_MLP$Ixodes_Scapularis) <- c("A", "P")

fitControl <- trainControl(method = "repeatedcv",
number = 5,
repeats = 8,
classProbs = TRUE,
summaryFunction = twoClassSummary,
search = "random")

GAM_Fit <- caret::train(Ixodes_Scapularis~., data = DF_GLM_GAM_GLMPoly,
method = "gam",
metric = "ROC",
tuneLength = 50,
```

```
trControl = fitControl)
GLMPoly_Fit <- caret::train(Ixodes_Scapularis~., data = DF_GLM_GAM_GLMPoly,
method = "gam",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)

MDA_Fit <- caret::train(Ixodes_Scapularis~., data = DF_MDA_FDA,
method = "mda",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)
FDA_Fit <- caret::train(Ixodes_Scapularis~., data = DF_MDA_FDA,
method = "fda",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)

MARS_Fit <- caret::train(Ixodes_Scapularis~., data = DF_MARS,,
method = "bagEarth",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)

RPART_Fit <- caret::train(Ixodes_Scapularis~., data =
DF_RPART_CART_BRT_RF_SVM_RBF_MLP,
method = "rpart",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)
BRT_Fit <- caret::train(Ixodes_Scapularis~., data =
DF_RPART_CART_BRT_RF_SVM_RBF_MLP,
method = "gbm",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)
RF_Fit <- caret::train(Ixodes_Scapularis~., data =
DF_RPART_CART_BRT_RF_SVM_RBF_MLP,
method = "RRF",
```

```
metric = "ROC",
tuneLength = 50,
trControl = fitControl)
SVM_Fit <- caret::train(Ixodes_Scapularis~., data =
DF_RPART_CART_BRT_RF_SVM_RBF_MLP,
method = "svmPoly",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)
RBF_Fit <- caret::train(Ixodes_Scapularis~., data =
DF_RPART_CART_BRT_RF_SVM_RBF_MLP,
method = "rbfDDA",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)
MLP_Fit <- caret::train(Ixodes_Scapularis~., data =
DF_RPART_CART_BRT_RF_SVM_RBF_MLP,
method = "mlpML",
metric = "ROC",
tuneLength = 50,
trControl = fitControl)

SDM_GLM_GAM_GLMPoly <-
sdm(Ixodes_Scapularis~.,data=sdmData_GLM_GAM_GLMPoly,methods=c('glm','gam','glmpol
y'), replication = c('sub','boot','cv'), cv.folds = 5,n=4,test.percent=20, modelSettings =
list(gam=list(select = FALSE, method = 'GCV.Cp')))

SDM_MARS <- sdm(Ixodes_Scapularis~.,data=sdmData_MARS,methods=c('mars'), replication
= c('sub','boot','cv'), cv.folds = 5,n=4,test.percent=20, modelSettings = list(mars=list(nprune = 2,
degree = 2)))

SDM_MDA_FDA <-
sdm(Ixodes_Scapularis~.,data=sdmData_MDA_FDA,methods=c('mda','fda'), replication =
c('sub','boot','cv'), cv.folds = 5,n=4,test.percent=20, modelSettings = list(mda=list(subclasses =
47),fda = list(degree = 2, nprune = 13)))

SDM_RPART_CART_BRT_RF_SVM_RBF_MLP <-
sdm(Ixodes_Scapularis~.,data=sdmData_RPART_CART_BRT_RF_SVM_RBF_MLP,methods=
c('rpart','cart','brt','svm','rbf','mlp'), replication = c('sub','boot','cv'), cv.folds =
```

```
5,n=4,test.percent=20, modelSettings = list(rpart=list(cp = 9.735202e-05), brt = list(n.trees =
1311, interaction.depth = 9, shrinkage = 0.01702334, n.minobsinnode = 20),rf = list(mtry = 5,
coefReg = 0.7684281, coefImp = 0.2433694),svm = list(degree = 3, scale = 0.3088387, C =
112.9314),rbf = list(negativeThreshold = 0.08416721), mlp = list(layer1 = 19, layer2 = 14, layer3
= 4)))

gui(SDM_GLM_GAM_GLMPoly)  # user interface to see results of models

gui(SDM_MARS)

gui(SDM_MDA_FDA)

gui(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP)

eval_glm_gam_glmpoly <- lapply(c("SDM_GLM_GAM_GLMPoly"),
function(x) {
getEvaluation(get(x),wtest='test.dep',stat=c('AUC','COR','Deviance','obs.prevalence',
'threshold','TSS','Kappa','sensitivity','specificity','NMI','phi','ppv','npv','ccr','prevalence'),opt=2) } )
n <- 28 # number of models in each algorithm
algo_glm_gam_glmpoly <- c('GLM','GAM','GLMPoly') ######whatever one you want
eval_algo_glm_gam_glmpoly <- data.frame(algo = c(rep(algo_glm_gam_glmpoly,
each=n)),bind_rows(eval_glm_gam_glmpoly))
good_models_glm_gam_glmpoly <-
eval_algo_glm_gam_glmpoly[eval_algo_glm_gam_glmpoly["AUC"] >= 0.9 &
eval_algo_glm_gam_glmpoly["TSS"] >= 0.88 & eval_algo_glm_gam_glmpoly["Kappa"] >=
0.82,]
good_models_glm_gam_glmpoly

eval_mars <- lapply(c("SDM_MARS"),
function(x) {
getEvaluation(get(x),wtest='test.dep',stat=c('AUC','COR','Deviance','obs.prevalence',
'threshold','TSS','Kappa','sensitivity','specificity','NMI','phi','ppv','npv','ccr','prevalence'),opt=2) } )
n <- 28 # number of models in each algorithm
algo_mars <- c('MARS')
eval_algo_mars <- data.frame(algo = c(rep(algo_mars, each=n)),bind_rows(eval_mars))
good_models_mars <- eval_algo_mars[eval_algo_mars["AUC"] >= 0.9 &
eval_algo_mars["TSS"] >= 0.72 & eval_algo_mars["Kappa"] >= 0.67,]
good_models_mars
```

```
eval_mda_fda <- lapply(c("SDM_MDA_FDA"),
function(x) {
getEvaluation(get(x),wtest='test.dep',stat=c('AUC','COR','Deviance','obs.prevalence',
'threshold','TSS','Kappa','sensitivity','specificity','NMI','phi','ppv','npv','ccr','prevalence'),opt=2) } )
n <- 28 # number of models in each algorithm
algo_mda_fda <- c('MDA','FDA') ######whatever one you want
eval_algo_mda_fda <- data.frame(algo = c(rep(algo_mda_fda,
each=n)),bind_rows(eval_mda_fda))
good_models_mda_fda <- eval_algo_mda_fda[eval_algo_mda_fda["AUC"] >= 0.9 &
eval_algo_mda_fda["TSS"] >= 0.88 & eval_algo_mda_fda["Kappa"] >= 0.82,]
good_models_mda_fda

eval_rpart_cart_brt_rf_svm_rbf_mlp <-
lapply(c("SDM_RPART_CART_BRT_RF_SVM_RBF_MLP),
function(x) {
getEvaluation(get(x),wtest='test.dep',stat=c('AUC','COR','Deviance','obs.prevalence',
'threshold','TSS','Kappa','sensitivity','specificity','NMI','phi','ppv','npv','ccr','prevalence'),opt=2) } )
n <- 28 # number of models in each algorithm
algo_rpart_cart_brt_rf_svm_rbf_mlp <- c('RPART','CART','BRT','RF','SVM','RBF','MLP')
######whatever one you want
eval_algo_rpart_cart_brt_rf_svm_rbf_mlp <- data.frame(algo = c(rep(algo_,
each=n)),bind_rows(eval_mda_fda))
good_models_rpart_cart_brt_rf_svm_rbf_mlp <-
eval_algo_rpart_cart_brt_rf_svm_rbf_mlp[eval_algo_rpart_cart_brt_rf_svm_rbf_mlp["AUC"]
>= 0.9 & eval_algo_rpart_cart_brt_rf_svm_rbf_mlp["TSS"] >= 0.88 &
eval_algo_rpart_cart_brt_rf_svm_rbf_mlp["Kappa"] >= 0.82,]
good_models_rpart_cart_brt_rf_svm_rbf_mlp

ID_GLM <- good_models_glm_gam_glmpoly[good_models_glm_gam_glmpoly$algo
=="GLM",]
ID_GAM <- good_models_glm_gam_glmpoly[good_models_glm_gam_glmpoly$algo
=="GAM",]
ID_GLMPoly <- good_models_glm_gam_glmpoly[good_models_glm_gam_glmpoly$algo
=="GLMPoly",]

ID_MARS <- good_models_mars[good_models_mars$algo =="MARS",]

ID_MDA <- good_models_mda_fda[good_models_mda_fda$algo =="MDA",]
ID_FDA <- good_models_mda_fda[good_models_mda_fda$algo =="FDA",]
```

```
ID_RPART <-
good_models_rpart_cart_brt_rf_svm_rbf_mlp[good_models_rpart_cart_brt_rf_svm_rbf_mlp$alg
o =="RPART",]
ID_CART <-
good_models_rpart_cart_brt_rf_svm_rbf_mlp[good_models_rpart_cart_brt_rf_svm_rbf_mlp$alg
o =="CART",]
ID_BRT <-
good_models_rpart_cart_brt_rf_svm_rbf_mlp[good_models_rpart_cart_brt_rf_svm_rbf_mlp$alg
o =="BRT",]
ID_RF <-
good_models_rpart_cart_brt_rf_svm_rbf_mlp[good_models_rpart_cart_brt_rf_svm_rbf_mlp$alg
o =="RF",]
ID_SVM <-
good_models_rpart_cart_brt_rf_svm_rbf_mlp[good_models_rpart_cart_brt_rf_svm_rbf_mlp$alg
o =="SVM",]
ID_RBF <-
good_models_rpart_cart_brt_rf_svm_rbf_mlp[good_models_rpart_cart_brt_rf_svm_rbf_mlp$alg
o =="RBF",]
ID_MLP <-
good_models_rpart_cart_brt_rf_svm_rbf_mlp[good_models_rpart_cart_brt_rf_svm_rbf_mlp$alg
o =="MLP",]

List_Rasters_SSP585_2011_2040 <- list.files(path =
"D:/Westcott/Raster_Data/Sequence_7/SSP585/2011_2040/SQ/PJ/Msk", pattern = 'tif$',
full.names = T)
List_Rasters_SSP585_2041_2070 <- list.files(path =
"D:/Westcott/Raster_Data/Sequence_7/SSP585/2041-2070/Sq/PJ/Msk", pattern = 'tif$',
full.names = T)
List_Rasters_SSP585_2071_2100 <- list.files(path =
"D:/Westcott/Raster_Data/Sequence_7/SSP585/2071-2100/Sq\PJ\Msk", pattern = 'tif$',
full.names = T)

Rasters_SSP585_2011_2040 <- stack(List_Rasters_SSP585_2011_2040)
Rasters_SSP585_2041_2070 <- stack(List_Rasters_SSP585_2041_2070)
Rasters_SSP585_2071_2100 <- stack(List_Rasters_SSP585_2071_2100)

List_Rasters_SSP370_2011_2040 <- list.files(path =
"D:/Westcott/Raster_Data/Sequence_7/2011_2040/SQ/PJ/Msk", pattern = 'tif$', full.names = T)
```

```
List_Rasters_SSP370_2041_2070 <- list.files(path =
"D:/Westcott/Raster_Data/Sequence_7/2041-2070/Sq/PJ/Msk", pattern = 'tif$', full.names = T)
List_Rasters_SSP370_2071_2100 <- list.files(path =
"D:/Westcott/Raster_Data/Sequence_7/2071-2100/Sq\PJ\Msk", pattern = 'tif$', full.names = T)

Rasters_SSP370_2011_2040 <- stack(List_Rasters_SSP370_2011_2040)
Rasters_SSP370_2041_2070 <- stack(List_Rasters_SSP370_2041_2070)
Rasters_SSP370_2071_2100 <- stack(List_Rasters_SSP370_2071_2100)

Ensemble_GAM <- ensemble(SDM_GLM_GAM_GLMPoly, Rasters_Present, filename =
"IS_CB_S7_GAM.tif", setting=list(method = 'unweighted', id = ID_GAM$modelID))
Ensemble_GAM_SSP585_2011_2040 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP585_2011_2040, filename = "IS_CB_S7_GAM_SSP585_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_GAM$modelID))
Ensemble_GAM_SSP585_2041_2070 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP585_2041_2070, filename = "IS_CB_S7_GAM_SSP585_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_GAM$modelID))
Ensemble_GAM_SSP585_2071_2100 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP585_2071_2100, filename = "IS_CB_S7_GAM_SSP585_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_GAM$modelID))
Ensemble_GAM_SSP370_2011_2040 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP370_2011_2040, filename = "IS_CB_S7_GAM_SSP370_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_GAM$modelID))
Ensemble_GAM_SSP370_2041_2070 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP370_2041_2070, filename = "IS_CB_S7_GAM_SSP370_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_GAM$modelID))
Ensemble_GAM_SSP370_2071_2100 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP370_2071_2100, filename = "IS_CB_S7_GAM_SSP370_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_GAM$modelID))

Ensemble_GLMPoly <- ensemble(SDM_GLM_GAM_GLMPoly, Rasters_Present, filename =
"IS_CB_S7_GLMPoly.tif", setting=list(method = 'unweighted', id = ID_GLMPoly$modelID))
Ensemble_GLMPoly_SSP585_2011_2040 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP585_2011_2040, filename = "IS_CB_S7_GLMPoly_SSP585_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_GLMPoly$modelID))
Ensemble_GLMPoly_SSP585_2041_2070 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP585_2041_2070, filename = "IS_CB_S7_GLMPoly_SSP585_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_GLMPoly$modelID))
```

```
Ensemble_GLMPoly_SSP585_2071_2100 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP585_2071_2100, filename = "IS_CB_S7_GLMPoly_SSP585_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_GLMPoly$modelID))
Ensemble_GLMPoly_SSP370_2011_2040 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP370_2011_2040, filename = "IS_CB_S7_GLMPoly_SSP370_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_GLMPoly$modelID))
Ensemble_GLMPoly_SSP370_2041_2070 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP370_2041_2070, filename = "IS_CB_S7_GLMPoly_SSP370_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_GLMPoly$modelID))
Ensemble_GLMPoly_SSP370_2071_2100 <- ensemble(SDM_GLM_GAM_GLMPoly,
Rasters_SSP370_2071_2100, filename = "IS_CB_S7_GLMPoly_SSP370_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_GLMPoly$modelID))

Ensemble_MDA <- ensemble(SDM_MDA_FDA, Rasters_Present, filename =
"IS_CB_S7_MDA.tif", setting=list(method = 'unweighted', id = ID_MDA$modelID))
Ensemble_MDA_SSP585_2011_2040 <- ensemble(SDM_MDA_FDA,
Rasters_SSP585_2011_2040, filename = "IS_CB_S7_MDA_SSP585_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_MDA$modelID))
Ensemble_MDA_SSP585_2041_2070 <- ensemble(SDM_MDA_FDA,
Rasters_SSP585_2041_2070, filename = "IS_CB_S7_MDA_SSP585_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_MDA$modelID))
Ensemble_MDA_SSP585_2071_2100 <- ensemble(SDM_MDA_FDA,
Rasters_SSP585_2071_2100, filename = "IS_CB_S7_MDA_SSP585_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_MDA$modelID))
Ensemble_MDA_SSP370_2011_2040 <- ensemble(SDM_MDA_FDA,
Rasters_SSP370_2011_2040, filename = "IS_CB_S7_MDA_SSP370_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_MDA$modelID))
Ensemble_MDA_SSP370_2041_2070 <- ensemble(SDM_MDA_FDA,
Rasters_SSP370_2041_2070, filename = "IS_CB_S7_MDA_SSP370_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_MDA$modelID))
Ensemble_MDA_SSP370_2071_2100 <- ensemble(SDM_MDA_FDA,
Rasters_SSP370_2071_2100, filename = "IS_CB_S7_MDA_SSP370_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_MDA$modelID))

Ensemble_FDA <- ensemble(SDM_SDM_MDA_FDA, Rasters_Present, filename =
"IS_CB_S7_FDA.tif", setting=list(method = 'unweighted', id = ID_FDA$modelID))
Ensemble_FDA_SSP585_2011_2040 <- ensemble(SDM_MDA_FDA,
Rasters_SSP585_2011_2040, filename = "IS_CB_S7_FDA_SSP585_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_FDA$modelID))
```

```
Ensemble_FDA_SSP585_2041_2070 <- ensemble(SDM_MDA_FDA,
Rasters_SSP585_2041_2070, filename = "IS_CB_S7_FDA_SSP585_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_FDA$modelID))
Ensemble_FDA_SSP585_2071_2100 <- ensemble(SDM_MDA_FDA,
Rasters_SSP585_2071_2100, filename = "IS_CB_S7_FDA_SSP585_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_FDA$modelID))
Ensemble_FDA_SSP370_2011_2040 <- ensemble(SDM_MDA_FDA,
Rasters_SSP370_2011_2040, filename = "IS_CB_S7_FDA_SSP370_2011_2040.tif",
setting=list(method = 'unweighted', id = ID_FDA$modelID))
Ensemble_FDA_SSP370_2041_2070 <- ensemble(SDM_MDA_FDA,
Rasters_SSP370_2041_2070, filename = "IS_CB_S7_FDA_SSP370_2041_2070.tif",
setting=list(method = 'unweighted', id = ID_FDA$modelID))
Ensemble_FDA_SSP370_2071_2100 <- ensemble(SDM_MDA_FDA,
Rasters_SSP370_2071_2100, filename = "IS_CB_S7_FDA_SSP370_2071_2100.tif",
setting=list(method = 'unweighted', id = ID_FDA$modelID))

Ensemble_RPART <- ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP,
Rasters_Present, filename = "IS_CB_S7_RPART.tif", setting=list(method = 'unweighted', id =
ID_RPART$modelID))
Ensemble_RPART_SSP585_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2011_2040,
filename = "IS_CB_S7_RPART_SSP585_2011_2040.tif", setting=list(method = 'unweighted', id
= ID_RPART$modelID))
Ensemble_RPART_SSP585_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2041_2070,
filename = "IS_CB_S7_RPART_SSP585_2041_2070.tif", setting=list(method = 'unweighted', id
= ID_RPART$modelID))
Ensemble_RPART_SSP585_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2071_2100,
filename = "IS_CB_S7_RPART_SSP585_2071_2100.tif", setting=list(method = 'unweighted', id
= ID_RPART$modelID))
Ensemble_RPART_SSP370_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2011_2040,
filename = "IS_CB_S7_RPART_SSP370_2011_2040.tif", setting=list(method = 'unweighted', id
= ID_RPART$modelID))
Ensemble_RPART_SSP370_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2041_2070,
filename = "IS_CB_S7_RPART_SSP370_2041_2070.tif", setting=list(method = 'unweighted', id
= ID_RPART$modelID))
```

```
Ensemble_RPART_SSP370_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2071_2100,
filename = "IS_CB_S7_RPART_SSP370_2071_2100.tif", setting=list(method = 'unweighted', id
= ID_RPART$modelID))

Ensemble_CART <- ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP,
Rasters_Present, filename = "IS_CB_S7_RPART.tif", setting=list(method = 'unweighted', id =
ID_CART$modelID))
Ensemble_CART_SSP585_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2011_2040,
filename = "IS_CB_S7_CART_SSP585_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_CART$modelID))
Ensemble_CART_SSP585_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2041_2070,
filename = "IS_CB_S7_CART_SSP585_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_CART$modelID))
Ensemble_CART_SSP585_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2071_2100,
filename = "IS_CB_S7_CART_SSP585_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_CART$modelID))
Ensemble_CART_SSP370_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2011_2040,
filename = "IS_CB_S7_CART_SSP370_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_CART$modelID))
Ensemble_CART_SSP370_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2041_2070,
filename = "IS_CB_S7_CART_SSP370_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_CART$modelID))
Ensemble_CART_SSP370_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2071_2100,
filename = "IS_CB_S7_CART_SSP370_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_CART$modelID))

Ensemble_BRT <- ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP,
Rasters_Present, filename = "IS_CB_S7_BRT.tif", setting=list(method = 'unweighted', id =
ID_BRT$modelID))
Ensemble_BRT_SSP585_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2011_2040,
```

```
filename = "IS_CB_S7_BRT_SSP585_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_BRT$modelID))
Ensemble_BRT_SSP585_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2041_2070,
filename = "IS_CB_S7_BRT_SSP585_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_BRT$modelID))
Ensemble_BRT_SSP585_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2071_2100,
filename = "IS_CB_S7_BRT_SSP585_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_BRT$modelID))
Ensemble_BRT_SSP370_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2011_2040,
filename = "IS_CB_S7_BRT_SSP370_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_BRT$modelID))
Ensemble_BRT_SSP370_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2041_2070,
filename = "IS_CB_S7_BRT_SSP370_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_BRT$modelID))
Ensemble_BRT_SSP370_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2071_2100,
filename = "IS_CB_S7_BRT_SSP370_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_BRT$modelID))

Ensemble_RF <- ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_Present,
filename = "IS_CB_S7_RF.tif", setting=list(method = 'unweighted', id = ID_RF$modelID))
Ensemble_RF_SSP585_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2011_2040,
filename = "IS_CB_S7_RF_SSP585_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_RF$modelID))
Ensemble_RF_SSP585_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2041_2070,
filename = "IS_CB_S7_RF_SSP585_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_RF$modelID))
Ensemble_RF_SSP585_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2071_2100,
filename = "IS_CB_S7_RF_SSP585_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_RF$modelID))
Ensemble_RF_SSP370_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2011_2040,
```

```
filename = "IS_CB_S7_RF_SSP370_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_RF$modelID))
Ensemble_RF_SSP370_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2041_2070,
filename = "IS_CB_S7_RF_SSP370_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_RF$modelID))
Ensemble_RF_SSP370_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2071_2100,
filename = "IS_CB_S7_RF_SSP370_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_RF$modelID))

Ensemble_SVM <- ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP,
Rasters_Present, filename = "IS_CB_S7_SVM.tif", setting=list(method = 'unweighted', id =
ID_SVM$modelID))
Ensemble_SVM_SSP585_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2011_2040,
filename = "IS_CB_S7_SVM_SSP585_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_SVM$modelID))
Ensemble_SVM_SSP585_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2041_2070,
filename = "IS_CB_S7_SVM_SSP585_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_SVM$modelID))
Ensemble_SVM_SSP585_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2071_2100,
filename = "IS_CB_S7_SVM_SSP585_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_SVM$modelID))
Ensemble_SVM_SSP370_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2011_2040,
filename = "IS_CB_S7_SVM_SSP370_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_SVM$modelID))
Ensemble_SVM_SSP370_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2041_2070,
filename = "IS_CB_S7_SVM_SSP370_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_SVM$modelID))
Ensemble_SVM_SSP370_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2071_2100,
filename = "IS_CB_S7_SVM_SSP370_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_SVM$modelID))
```

```
Ensemble_RBF <- ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP,
Rasters_Present, filename = "IS_CB_S7_RBF.tif", setting=list(method = 'unweighted', id =
ID_RBF$modelID))
Ensemble_RBF_SSP585_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2011_2040,
filename = "IS_CB_S7_RBF_SSP585_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_RBF$modelID))
Ensemble_RBF_SSP585_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2041_2070,
filename = "IS_CB_S7_SVM_SSP585_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_RBF$modelID))
Ensemble_RBF_SSP585_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2071_2100,
filename = "IS_CB_S7_SVM_SSP585_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_RBF$modelID))
Ensemble_RBF_SSP370_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2011_2040,
filename = "IS_CB_S7_RBF_SSP370_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_RBF$modelID))
Ensemble_RBF_SSP370_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2041_2070,
filename = "IS_CB_S7_RBF_SSP370_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_RBF$modelID))
Ensemble_RBF_SSP370_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2071_2100,
filename = "IS_CB_S7_RBF_SSP370_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_RBF$modelID))

Ensemble_MLP <- ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP,
Rasters_Present, filename = "IS_CB_S7_MLP.tif", setting=list(method = 'unweighted', id =
ID_MLP$modelID))
Ensemble_MLP_SSP585_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2011_2040,
filename = "IS_CB_S7_MLP_SSP585_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_MLP$modelID))
Ensemble_MLP_SSP585_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2041_2070,
filename = "IS_CB_S7_MLP_SSP585_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_MLP$modelID))
```

```
Ensemble_MLP_SSP585_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP585_2071_2100,
filename = "IS_CB_S7_MLP_SSP585_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_MLP$modelID))
Ensemble_MLP_SSP370_2011_2040 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2011_2040,
filename = "IS_CB_S7_MLP_SSP370_2011_2040.tif", setting=list(method = 'unweighted', id =
ID_MLP$modelID))
Ensemble_MLP_SSP370_2041_2070 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2041_2070,
filename = "IS_CB_S7_MLP_SSP370_2041_2070.tif", setting=list(method = 'unweighted', id =
ID_MLP$modelID))
Ensemble_MLP_SSP370_2071_2100 <-
ensemble(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, Rasters_SSP370_2071_2100,
filename = "IS_CB_S7_MLP_SSP370_2071_2100.tif", setting=list(method = 'unweighted', id =
ID_MLP$modelID))

write.csv(ID_GAM,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapular
is/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/
Tables/IS_CB_S7_GAM_ID.csv")
write.csv(ID_GLMPoly,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scap
ularis/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Sev
en/Tables/IS_CB_S7_GLMPoly_ID.csv")
write.csv(ID_MDA,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapulari
s/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/T
ables/IS_CB_S7_MDA_ID.csv")
write.csv(ID_FDA,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapulari
s/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/T
ables/IS_CB_S7_FDA_ID.csv")
write.csv(ID_RPART,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapul
aris/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seve
n/Tables/IS_CB_S7_RPART_ID.csv")
write.csv(ID_CART,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapular
is/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/
Tables/IS_CB_S7_CART_ID.csv")
write.csv(ID_BRT,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapularis
/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/Ta
bles/IS_CB_S7_BRT_ID.csv")
```

```
write.csv(ID_RF,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapularis/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/Tables/IS_CB_S7_RF_ID.csv")
write.csv(ID_SVM,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapularis/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/Tables/IS_CB_S7_SVM_ID.csv")
write.csv(ID_RBF,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapularis/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/Tables/IS_CB_S7_RBF_ID.csv")
write.csv(ID_MLP,"C:/Users/jwestcot/Permission_Check/Black_Legged_Tick/Ixodes_Scapularis/R/Ixodes_Scapularis/Ixodes_Scapularis_Scenario_Seven/Ixodes_Scapularis_Scenario_Seven/Tables/IS_CB_S7_MLP_ID.csv")

RC_GAM <- rcurve(SDM_GLM_GAM_GLMPoly, id = c(32,42,50,51,56), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_GLMPoly <- rcurve(SDM_GLM_GAM_GLMPoly, id = c(70,79,84), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_MDA <- rcurve(SDM_MDA_FDA, id = c(70,79,84), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_MDA <- rcurve(SDM_MDA_FDA, id = c(8,26), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_FDA <- rcurve(SDM_MDA_FDA, id = c(31,32,39,53), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_RPART <- rcurve(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(10,11), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_CART <- rcurve(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(30,32,34,38,39,40,43,50), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_BRT <- rcurve(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(62,63,67,68,71,78), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_RF <- rcurve(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(90,91,92,95,99,105), gg=T, mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
```

```
RC_SVM <- rcurve(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(118,123,133), gg=T,
mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1,
vjust=0.5),text=element_text(size=10))
RC_RBF <- rcurve(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id =
c(145,146,150,155,158,160,164,165), gg=T, mean=T, confidence=T, main = 'Response Curve')+
theme(axis.text.x=element_text(angle=90, hjust=1, vjust=0.5),text=element_text(size=10))
RC_MLP <- rcurve(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(179,183,189), gg=T,
mean=T, confidence=T, main = 'Response Curve')+ theme(axis.text.x=element_text(angle=90, hjust=1,
vjust=0.5),text=element_text(size=10))

RC_BIND <-
rbind(RC_GAM$data,RC_GLMPoly$data,RC_MDA$data,RC_FDA$data,RC_RPART$data,RC_CART$
data,RC_BRT$data,RC_RF$data,RC_SVM$data,RC_RBF$data,RC_MLP$data)

RC_V_BIO1 <- RC_BIND[RC_BIND$variable=="Band_1.1",]
RC_V_BIO16 <- RC_BIND[RC_BIND$variable=="Band_1.2",]
RC_V_BIO17 <- RC_BIND[RC_BIND$variable=="Band_1.3",]


VARIMP_GAM <- getVarImp(SDM_GLM_GAM_GLMPoly, id = c(32,42,50,51,56), wtest = 'test.dep')
VARIMP_GLMPoly <- getVarImp(SDM_GLM_GAM_GLMPoly, id = c(70,79,84), wtest = 'test.dep')
VARIMP_MDA <- getVarImp(SDM_MDA_FDA, id = c(8,26), wtest = 'test.dep')
VARIMP_FDA <- getVarImp(SDM_MDA_FDA, id = c(31,32,39,53), wtest = 'test.dep')
VARIMP_RPART <- getVarImp(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(10,11), wtest
= 'test.dep')
VARIMP_CART <- getVarImp(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id =
c(30,32,34,38,39,40,43,50), wtest = 'test.dep')
VARIMP_BRT <- getVarImp(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id =
c(62,63,67,68,71,78), wtest = 'test.dep')
VARIMP_RF <- getVarImp(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id =
c(90,91,92,95,99,105), wtest = 'test.dep')
VARIMP_SVM <- getVarImp(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(118,123,133),
wtest = 'test.dep')
VARIMP_RBF <- getVarImp(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id =
c(145,146,150,155,158,160,164,165), wtest = 'test.dep')
VARIMP_MLP <- getVarImp(SDM_RPART_CART_BRT_RF_SVM_RBF_MLP, id = c(179,183,189),
wtest = 'test.dep')

Mean_VI_Bio1 <-
(VARIMP_GAM@varImportanceMean$corTest$corTest[1]+VARIMP_GLMPoly@varImportanceMean$
corTest$corTest[1]+VARIMP_MDA@varImportanceMean$corTest$corTest[1]+VARIMP_FDA@varImp
ortanceMean$corTest$corTest[1]+VARIMP_RPART@varImportanceMean$corTest$corTest[1]+VARIMP
```

```
_CART@varImportanceMean$corTest$corTest[1]+VARIMP_BRT@varImportanceMean$corTest$corTes
t[1]+VARIMP_RF@varImportanceMean$corTest$corTest[1]+VARIMP_SVM@varImportanceMean$cor
Test$corTest[1]+VARIMP_RBF@varImportanceMean$corTest$corTest[1]+VARIMP_MLP@varImporta
nceMean$corTest$corTest[1])/11
Mean_VI_Bio16 <-
(VARIMP_GAM@varImportanceMean$corTest$corTest[2]+VARIMP_GLMPoly@varImportanceMean$
corTest$corTest[2]+VARIMP_MDA@varImportanceMean$corTest$corTest[2]+VARIMP_FDA@varImp
ortanceMean$corTest$corTest[2]+VARIMP_RPART@varImportanceMean$corTest$corTest[2]+VARIMP
_CART@varImportanceMean$corTest$corTest[2]+VARIMP_BRT@varImportanceMean$corTest$corTes
t[2]+VARIMP_RF@varImportanceMean$corTest$corTest[2]+VARIMP_SVM@varImportanceMean$cor
Test$corTest[2]+VARIMP_RBF@varImportanceMean$corTest$corTest[2]+VARIMP_MLP@varImporta
nceMean$corTest$corTest[2])/11
Mean_VI_Bio17 <-
(VARIMP_GAM@varImportanceMean$corTest$corTest[3]+VARIMP_GLMPoly@varImportanceMean$
corTest$corTest[3]+VARIMP_MDA@varImportanceMean$corTest$corTest[3]+VARIMP_FDA@varImp
ortanceMean$corTest$corTest[3]+VARIMP_RPART@varImportanceMean$corTest$corTest[3]+VARIMP
_CART@varImportanceMean$corTest$corTest[3]+VARIMP_BRT@varImportanceMean$corTest$corTes
t[3]+VARIMP_RF@varImportanceMean$corTest$corTest[3]+VARIMP_SVM@varImportanceMean$cor
Test$corTest[3]+VARIMP_RBF@varImportanceMean$corTest$corTest[3]+VARIMP_MLP@varImporta
nceMean$corTest$corTest[3])/11

PLOT_BIO1 <- RC_V_BIO1 %>% ggplot(aes(Value, Response)) + ylim(0,1) +
geom_text(x = ((max(RC_V_BIO1$Value)-min(RC_V_BIO1$Value))*0.2 + min(RC_V_BIO1$Value)),
y = 0.9,
label = paste0("Variable Importance = ",round(Mean_VI_Bio1, 4)), parse = F) +
labs(x = "Mean Annual Precipitation (Bio1)", y = "Probability") +
geom_smooth(color="Black", span = 0.50, method = "loess", method.args = list(degree=1))
PLOT_BIO16 <- RC_V_BIO16 %>% ggplot(aes(Value, Response)) + ylim(0,1) +
geom_text(x = ((max(RC_V_BIO16$Value)-min(RC_V_BIO16$Value))*0.2 +
min(RC_V_BIO16$Value)), y = 0.9,
label = paste0("Variable Importance = ",round(Mean_VI_Bio16, 4)), parse = F) +
labs(x = "Mean Monthly Precipitation: Wettest Quarter (Bio16)", y = "Probability") +
geom_smooth(color="Black", span = 0.50, method = "loess", method.args = list(degree=1))
PLOT_BIO17 <- RC_V_BIO17 %>% ggplot(aes(Value, Response)) + ylim(0,1) +
geom_text(x = ((max(RC_V_BIO17$Value)-min(RC_V_BIO17$Value))*0.2 +
min(RC_V_BIO17$Value)), y = 0.9,
label = paste0("Variable Importance = ",round(Mean_VI_Bio17, 4)), parse = F) +
labs(x = "Mean Monthly Precipitation: Warmest Quarter (Bio17)", y = "Probability") +
geom_smooth(color="Black", span = 0.50, method = "loess", method.args = list(degree=1))
```