THE USE OF SPECIES DISTRIBUTION MODELS TO INFORM AMPHIBIAN CONSERVATION IN WESTERN CANADA

KAEGAN J. FINN
Bachelor of Science, University of Alberta, 2021

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

MASTER OF SCIENCE

in

BIOLOGICAL SCIENCES

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

THE USE OF SPECIES DISTRIBUTION MODELS TO INFORM AMPHIBIAN CONSERVATION IN WESTERN CANADA

KAEGAN J. FINN

Date of Defence: August 16, 2023

Dr. J.A. Lee-Yaw Thesis Supervisor University of Ottawa Ottawa, Ontario	Assistant Professor	Ph.D.
Dr. J. McCune Thesis Co-Supervisor	Assistant Professor	Ph.D.
Dr. M. Bogard Thesis Examination Committee Member	Assistant Professor	Ph.D.
Dr. C. Goater Thesis Examination Committee Member	Professor	Ph.D.

ACKNOWLEDGEMENTS

This research was conducted on Treaty 7 territory: the traditional and ancestral territory of the Niitsítapi (including Kainai, Piikani, and Siksika Nations), Tsuut'ina, Ktunaxa, Stoney Nakota Sioux, and Métis First Nations. I also acknowledge all Indigenous lands from which amphibian records used in this thesis are reported and recognize these Nations as the traditional stewards of the Land.

I would like to thank Dr. Julie Lee-Yaw for her guidance, support, and trust. Since day one, Julie has been quick to provide praise at the highs, and kind words of encouragement through the lows. It has been a sincere pleasure and honour being one of her first students, and I look forward to watching her Lab grow and change in the years to come. Within the Lee-Yaw Lab I would like to thank Danial Hunter for the good times, rhymes, and insightful chats, Dylan Brassard for keeping me sane through long office days, and Dr. Arianna Kuhn and Shay Marks for help in the field. Thanks especially to Jayna Bergman for her major contributions to this thesis, for being a great friend, mentee, mentor, and host, and most of all for tolerating my pestering.

Many helping hands made the Summer 2022 field season the most fun I've ever had. First, thank you to my field assistant and friend, Dustin Snider. I have no doubt that things would not have gone nearly as smooth or be half as enjoyable had Dustin not been the keen, thorough, resilient, and positive person he is. Dustin, you're a champion By Right! Thanks also to Brian Eaton, Toby Clark, Maurice Needham, Lorna Selinger, Wonnita Andrus, Mike Taje, and Simon Wilkins for assistance with logistics, safety, and access. Lastly, thanks to Wendy Ryan for making our final day in the field the best day of the summer.

I am grateful for the support of my supervisory committee; Dr. Cameron Goater, Dr. Matthew Bogard, and especially Dr. Jenny McCune and members of her lab for taking me in amidst the lab move. Thank you especially to Amy Weidenfeld for being the best office buddy around, and Kirsty McFadyen for your support through the final push.

I would also like to thank my colleagues at Parks Canada for their support of this research. Thank you to Kim Pearson whose work inspired much of this thesis, to Helena Mahoney, Bradley Yee, and the Ecological Integrity Monitoring crew for introducing me to long-toed salamanders, and to Stephanie Crowshoe for the suggestion to look at how climate change may impact my model predictions. Thanks also to Dr. Uldis Silins, Chris Williams, Erin Cherlet, and Jeremy Fitzpatrick of the Southern Rockies Watershed Project for introducing me to southwestern Alberta, and for hosting us at the bunkhouse.

I could not have arrived at this point without the support of family and friends. Thank you first to Mom, Dad, Chelsea, and Penny for your love, and especially Mom for the many (12!) visits to Lethbridge. Your support means the world. Thank you, Aidan Sheppard, for being the brother I never had, for your constant encouragement, and your help with getting this thesis dialed. Thank you, Eamon Turner, for being my spiritual companion and Flow Bro, and for countless life-long memories made outdoors. Thanks also to Kevin Rossi for the good times outside, and a critical piece of salamander intel. For many other fun times and inspiring conversations, thank you Ferf McCurry, Nat Hontar, Eric Jenkins, Jaxon Reiter, Justin Dubiel, Aaron Veale, and Larry Powell.

Last but certainly not least, I would like to thank the Critters, Mountains, Waters, and Winds of southwestern Alberta for being great teachers and friends along this journey. Thanks to you, this place will always feel like home.

ABSTRACT

Species distribution models (SDMs) have become a widely used tool for understanding species distributions and informing conservation actions. However, user decisions made during model development can impact model predictions and performance. In this thesis, I address the sensitivity of model predictions and performance to several modeling decisions. I first reviewed modeling decisions used in studies that validated SDMs with independent data and found that the impacts of some decisions (*i.e.,* choice of geographic study extent) have yet to be fully tested using independent data. I then used independent data to test the impacts of study extent on SDMs for six amphibian species at the edge of their respective ranges in western Canada. I found that model predictions were highly sensitive to the study extent used, and that many models failed to accurately predict independent occurrence regardless of the study extent used. Following from this result, I explored the joint impacts of choice of study extent, the type of environmental data used, and whether sampling bias was accounted for in SDMs developed to inform conservation translocations of long-toed salamanders. I found that these decisions impacted the prioritization of potential release sites, and that models developed using random background points and local study extents tended to perform best. In this study I also demonstrated an approach for incorporating future climatic predictions into the selection of potential release sites and found that trade-offs exist between developing accurate models and those that can make future predictions without extrapolation beyond the conditions with which models were generated. Overall, my thesis contributes several validated SDMs for use in amphibian conservation in western Canada while simultaneously adding to our understanding of how key modeling decisions can impact SDM predictions, performance, and downstream conservation decisions.

PREFACE

Chapter 2 was developed in collaboration with BScH student, Jayna Bergman. Jayna and I developed the code together and co-wrote a first draft of this chapter. I completed several rounds of revision with my supervisor and the final version of this chapter presented here represents a substantial update from what was submitted as part of J. Bergman's BScH. Given our close collaboration, J. Bergman and I will be co-first authors on this manuscript when it is submitted. Data and code developed as part of that chapter was also used for my third chapter and Jayna will be a second author on this study when it is submitted. However, I wrote and edited this chapter alone (with input from my supervisor).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SDM	Species Distribution Model
AUC	Area Under the receiver operator Curve
TSS	True Skill Statistic
WLNP	Waterton Lakes National Park
EINP	Elk Island National Park
CCWPP	Castle Provincial Park and Castle Wildland Provincial Park
SSP	Shared Socioeconomic Pathway
MESS	Multivariate Environmental Similarity Surface

CHAPTER 1: INTRODUCTION

1.1. Background

Understanding species' distributions is a fundamental goal of ecology (Guisan and Zimmermann, 2000). Broadly speaking, species occur at the intersection of suitable abiotic and biotic conditions that are accessible to them (Peterson, 2011). However, uncovering specific variables that govern species occurrence and the configuration of suitable habitat at a given time, place, or scale, remains a central pursuit of modern ecological research (Gaston, 2003; Sexton *et. al.,* 2009). Identifying the drivers of species distributions is also of interest for conservation, as we strive to predict how species niches will shift in response to climate change (Parmesan *et. al.,* 2005), forecast advances of invasive species (Peterson and Vieglais, 2001), and prioritize habitats for protection (Stryszowska *et. al.,* 2016; Muscatello *et. al.,* 2021).

Species distribution models (SDMs) are a widely used tool for understanding species distributions (Elith *et. al.,* 2006; Phillips *et. al.,* 2006; Randin *et. al.,* 2006). Two types of data are required to develop an SDM: 1) occurrence records of a species of interest (either presence or presence-absence) and 2) environmental data, typically collected by extracting information about environmental conditions underlying the locality records using geographic information systems (Phillips *et. al.,* 2006). Occurrence records are generally sourced from online data repositories (*e.g.,* the Global Biodiversity Information System; GBIF) which include museum records and observations from researchers and citizen-scientists (*e.g.,* iNaturalist; Beck *et. al.,* 2014). Environmental GIS data are typically sourced from free climate databases (*e.g.,* WorldClim; Fick and Hijmans, 2017), though non-climate data (*e.g.,* vegetative cover or land use, often sourced from the Environmental Systems Research Institute (ESRI) online catalogue or in North America, the United States Geological Survey (USGS) are also frequently used (Phillips *et. al.,* 2006). An algorithm is then used to distinguish between environmental conditions where a species is present and conditions at background points (presence-only models) or places where the species is known to be absent (presence-absence models; Elith *et. al.,* 2006; Phillips *et. al.,* 2006). Models can then be projected across geographic space to map suitable habitat or predict probability of occurrence (Phillips *et. al.,* 2006).

Although SDMs have the potential to be useful for both fundamental ecological inquiry (Peterson, 2011) and conservation planning (Guisan *et. al.,* 2013), models should be developed with care, as they are sensitive to user-decisions made during model development (Thuiller *et. al.,* 2004; VanDerWal *et. al.,* 2009; Anderson and Raza, 2010; Benito *et. al.,* 2013; Merow *et. al.,* 2013, Norberg *et. al.,* 2019). For example, Benito *et. al.,* (2013) and Norberg *et. al.,* (2019) demonstrated how otherwise identical models developed using different algorithms resulted in very different model outputs. Likewise, Thuiller *et. al.,* (2004) and VanDerWal *et. al.,* (2009) showed that the geographic area from which the locality data used to calibrate a model are drawn (*i.e.,* the study extent) can impact model outputs. Although several 'best-practices' for making such decisions have been proposed (Elith *et. al.,* 2011; Merow *et. al.,* 2013; Araújo *et. al.,* 2019), ideally, the impacts of different modeling decisions are explored through rigorous sensitivity testing (Araújo *et. al.,* 2019; Sofaer *et. al.,* 2019).

In addition to the need for careful decision-making during model development, care is also needed at the model evaluation stage. Most studies test SDM performance using withheld input data in a process known as cross-validation, also referred to as ‘internal validation’ (Merow *et. al.,* 2013). However, such methods often overestimate model accuracy because withheld locality records share the same biases as the localities used to train the model (Elith *et. al.,* 2006; Lobo *et. al.,* 2008; McCune, 2016; Lee-Yaw *et. al.,* 2022). For example, Matutini *et. al.,* (2021) found that although sixteen out of eighteen models (89%) developed for nine species of amphibians in France passed internal tests of performance based on withheld locality data, half of these models failed to accurately predict presence-absence data from independent field surveys. Likewise, Elith and Burgman *et. al.,* (2002) developed 64 SDMs for rare plants in Australia, and found that although most models were reasonably accurate, independent estimates of model accuracy were lower than internal estimates 80% of the time. In the most comprehensive review of SDM performance to date, Lee-Yaw *et. al.,* (2022) found that SDMs accurately predicted independent occurrence only half of the time and did poorly at predicting other metrics of population performance. Thus, further testing of SDMs using independent data is required to fully assess the utility of these models.

Although previous studies have separately considered the impacts of different modeling decisions on SDMs (Thuiller *et. al.,* 2004; VanDerWal *et. al.,* 2009; Benito *et. al.,* 2013; Guisan *et. al.,* 2013; Norberg *et. al.,* 2019) and the performance of SDMs when tested with independent data (Elith and Burgman, 2002; Lobo *et. al.,* 2008; Matutini *et. al.,* 2021; Lee-Yaw *et. al.,* 2022), it remains unclear how different modeling decisions impact the accuracy of models when evaluated with independent data. In the next section, I review studies that have validated SDMs using independent presence-absence data. In doing so, I not only explore the impacts of different modeling decisions on the ability of

SDMs to predict independent occurrence, but also identify gaps in our understanding of these impacts, as well as taxonomic and geographic biases in the literature.

1.2. Literature Review

1.2.2. Studies reviewed and data extracted.

To assess the current state of understanding on how model decisions impact SDM performance, I compiled studies that validated SDMs using independent data. I started with the 101 studies included in Lee-Yaw *et. al.,* (2022) that specifically tested the ability of SDMs to predict independent presence-absence data. To retrieve relevant studies published more recently, I conducted a Web of Science search on March 26th, 2023, using the same search string used by Lee-Yaw *et. al.,* (2022) for papers testing occurrence, but with the year of publication restricted to 2021 onwards (Table A1.1). This search returned an additional 156 studies. After applying the screening criteria used by Lee-Yaw *et. al.,* (2022; Table A1.1), 15 new studies were retained and added to those included in Lee-Yaw *et. al.,* (2022). I further screened the resulting 116 studies to include only those that: 1) developed SDMs for terrestrial species and 2) reported independent assessments of model performance based on the area under the receiver operator curve ("independent-AUC") in text or in tables. Although other metrics of SDM performance are often reported (reviewed by Shabani *et. al.,* 2018), AUC is the most frequently reported evaluation metric in the literature (Tessarolo *et. al.,* 2021) and has been used by other studies aiming to summarize model performance across very different systems (*e.g.,* Newbold *et. al.,* 2010; Kharouba *et. al.,* 2013). AUC ranges from 0 to 1, where AUC = 0.5 represents predictive ability no better than random chance, and AUC > 0.7 represents the threshold of 'acceptable' predictive ability (Swets, 1988).

My screening procedure resulted in 39 studies with which to examine the impacts of model decisions on independent AUC. From each of these 39 studies, I recorded independent-AUC for each model presented, the species and country or countries for which each model was generated, and the modelling decisions used to generate each model. The 39 studies included allowed for examination of the impacts of six different modelling decisions on model accuracy: choice of modelling algorithm, type of sampling bias, choice of environmental variables used, choice of study extent, model parameters (regularization), and environmental variable resolution. In total, the 39 studies reviewed developed and independently validated 1,197 SDMs for 292 species across 35 countries (Table 1.1).

Table 1.1. Summary of the 39 studies included in this review that developed and validated SDMs with independent data on occurrence.

Taxonomic Group	Number of Species	Number of models	Country / Countries	Model Decision(s) Tested *	Reference
Plant	3	3	Iran	None	Ardestani et al. (2015)
Plant	1	4	Iran	Algorithm	Safaei et al. (2018)
Bird	14	14	Canada	Algorithm	Börger and Nudds (2014)
Bird	1	2	Spain	Bias	Brotons et al. (2012)
Fungi	1	1	Chile	Algorithm	Bacigalupe et al. (2019)
Invertebrate	1	4	Egypt, Lebanon, Jordan, Syria, Israel, Turkey, Saudi Arabia, and Iraq	Variables, Algorithm	Conley et al. (2014)
Invertebrate	2	2	South Africa, Eswatini, Mozambique	None	DeBeer et al. (2021)
Amphibian	9	162	France	Bias	Matutini et al. (2021)
Mammal	1	5	South Africa	Bias	Pédarros et al. (2020)
Mammal	1	16	U.S.A.	Bias	Tye et al. (2017)
Plant	2	8	Greenland	Algorithm	Chardon et al. (2022)
Mammal	9	12	England	Resolution, Bias	Bellamy et al. (2013)
Plant	8	32	Australia	Algorithm	Elith and Burgman, (2002)
Bird	4	24	Finland	Variables, Resolution	Heikkinen et al. (2007)
Plant	1	78	Poland	Study Extent, Bias, Algorithm	Konowalik and Nosol (2021)
Mammal	1	3	Portugal	Algorithm	Rebelo and Jones (2010)
Amphibian	1	16	U.S.A.	Study Extent, Bias, Parameterization	Searcy and Shaffer (2014)
Amphibian	1	4	South Africa	None	Becker et al. (2021)
Plant	4	8	U.S.A.	Bias, Algorithm	Buechling and Tobalske (2011)
Mammal	1	1	Australia	None	Burns et al. (2020)
Plant	70	276	Italy	Parameterization	Fois et al. (2018)
Plant	4	36	U.S.A.	Variables, Resolution, Study Extent	Gogol-Prokurat (2011)
Plant	8	24	Switzerland	Variables, Algorithm	Le Lay et al. (2010)
Plant	8	7	Canada	None	McCune (2016)
Plant	24	192	Canada	Variables, Parameterization	McCune et al. (2020)
Amphibian	1	1	U.S.A.	None	Peterman et al. (2012)
Amphibian	1	36	U.S.A.	Study Extent, Bias, Algorithm, Parameterization	Radomski et al. (2022)
Bird	99	198	Spain	Bias	Brotons et al. (2007)
Mammal	1	2	Italy	Algorithm	Cianfrani et al. (2010)
Plant	1	1	Norway	None	Edvardsen et al. (2011)

Mammal	1	2	Italy	Variables, Algorithm	Galluzzi et al. (2017)
Bird	1	9	Georgia, Russia, Turkey, Azerbaijan, Armenia, and Iran	Variables, Algorithm	Gavashelishvili and Javakhishvili (2010)
Mammal	1	2	Brazil	Bias	Gine and Faria (2018)
Plant	1	1	Spain	None	Jiménez-Alfaro et al. (2012)
Mammal	1	1	U.S.A.	None	Kirk and Zielinski (2009)
Bird	1	2	Canada	Study Extent, Parameterization	Lathrop et al. (2018)
Mammal	1	1	Germany	None	Magg et al. (2016)
Plant	1	2	France	Variables	Marage et al. (2008)
Mammal	1	2	U.S.A.	Variables	Sacks et al. (2017)

*Sensitivity of models to different decisions tested by each study were as follows:
None = no model decisions were tested
Algorithm = Choice of SDM algorithm
Bias = Whether sampling bias was accounted for, and to the type of bias correction used
Resolution = Grain size of environmental raster layers
Study extent = Geographic area from which presences and absences or background points are sourced
Variables = The number and/or type of environmental variables
Parameterization = Regularization parameter for Maxent models

1.2.3. Summary of literature and results

Assuming the studies retrieved are representative of the literature (see Lee-Yaw *et. al.,* 2022), my search revealed several gaps in terms of which taxonomic groups and geographic regions are represented among independently validated SDMs. Of the 292 species represented, almost half (46%) were plants. In contrast, fungi, insects, and amphibians were poorly represented (0.3%, 1%, and 6% of species each respectively; Fig. 1.1). Although studies occurred across a substantial portion of the globe (Fig. 1.1), there was a bias toward high-latitude regions, with no studies occurring between 20 and -20 degrees of latitude. Even within well-represented, high-latitude areas, substantial gaps in geographic coverage were apparent. For example, the vast majority of Asia and much of North America, including western Canada, were not represented among studies included in this review. Due to variation in species traits (*i.e.,* body size, dispersal ability) and the importance of non-habitat factors for shaping species distributions in some regions (*e.g.,* biotic interactions at warm range limits; Paquette and Hargreaves, 2021), SDMs may be more accurate for some taxonomic groups or in certain ecoregions (Araújo and Pearson, 2005; Seoane *et. al.,* 2005; but see Kharouba *et. al.,* 2013; Morán-Ordóñez *et. al.,* 2017; McCune *et. al.,* 2020). Thus, our understanding of the utility of these models may benefit from expanded assessment of their performance in underrepresented taxonomic groups and geographic areas.

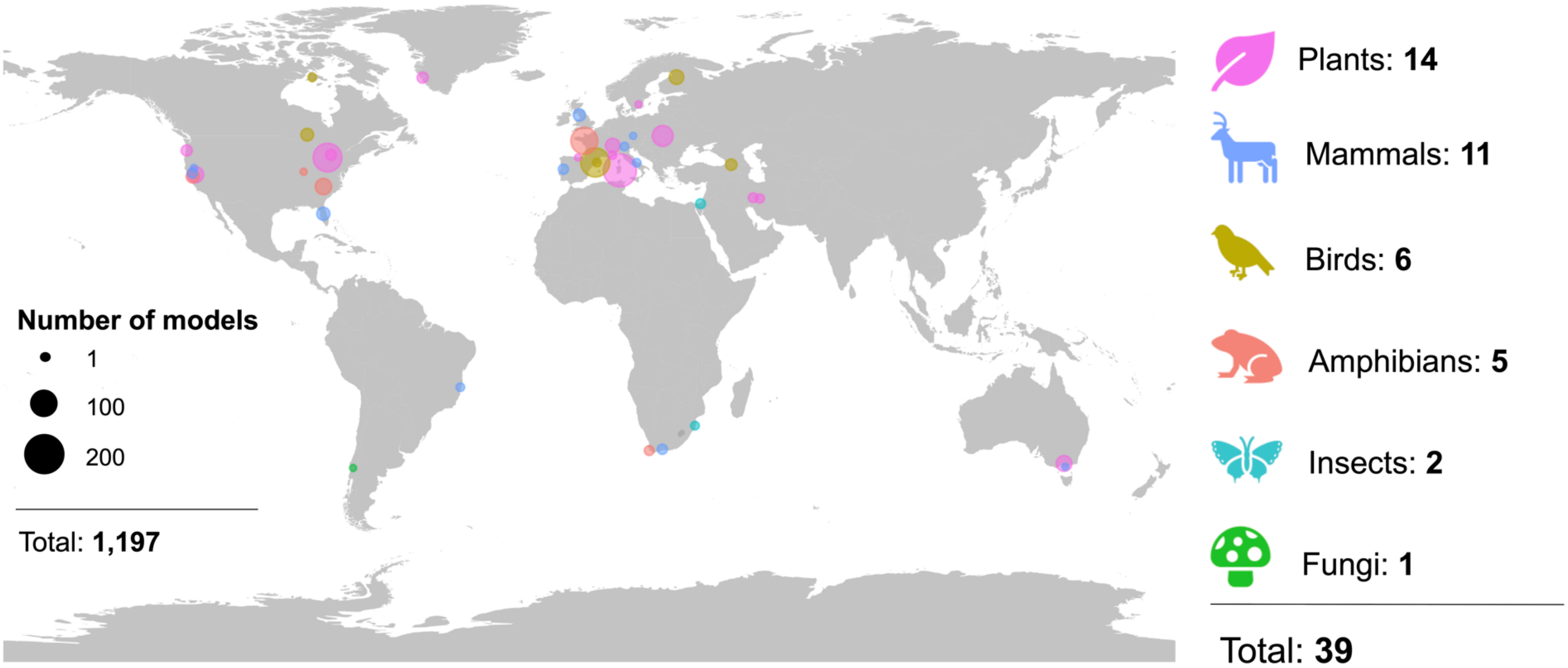


Figure 1.1. Geographic and taxonomic summary of studies that have developed and tested species distribution models with independent data for terrestrial species and have reported independent-AUC.

On average, the 1,197 models included in this review had strong ability to predict independent occurrence (mean independent-AUC = 0.85). Across the subset of 140 models from 19 studies that also validated SDMs using withheld locality records, internal estimates of predictive performance were higher on average than independent estimates (Fig. 1.2). This pattern is consistent with observations from previous studies (*e.g.,* Newbold *et. al.,* 2010; Fois *et. al.,* 2018; McCune *et. al.,* 2020), although in contrast to other studies, most models in the present review (79%) were above the threshold for considering a model suitable regardless of whether validation was based on withheld data or independent data (top right quadrant of Fig. 1.2). This may reflect a publication bias in the literature assessing SDMs with independent data as well-performing models may be more likely to be published (Lee-Yaw *et. al.,* 2022). Nevertheless, 15% of the models passed internal validation but not independent validation (bottom right quadrant of Fig. 1.2) and most models were below the 1:1 line, highlighting the tendency for internal-AUC to overestimate model accuracy (Elith and Burgman *et. al.,* 2002; Elith *et. al.,* 2006) and emphasizing the value of testing models with independent data.

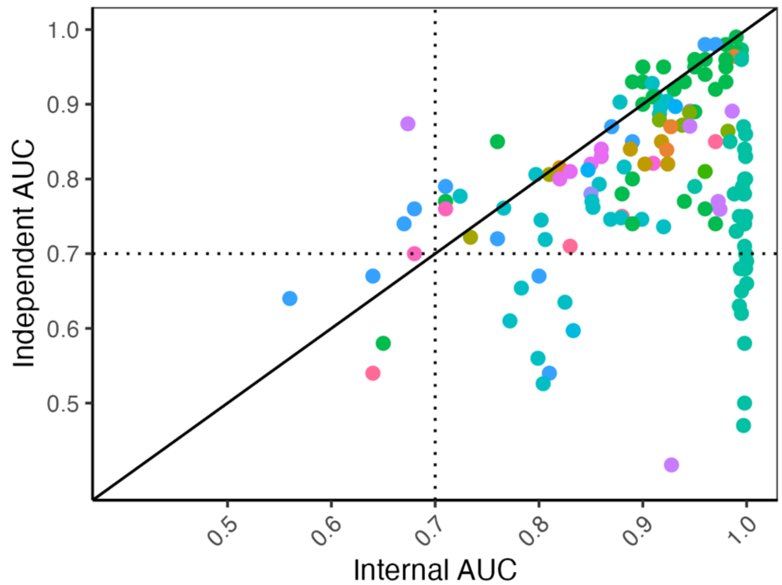


Figure 1.2. Internal versus independent area under the receiver operator curve (AUC) scores for the 140 models for which both values were reported. Each point represents a single model. Colours represent different studies (n = 19). Dashed horizontal and vertical lines represent the threshold for an 'acceptable' model (AUC > 0.7, Swets, 1988). The solid line represents the 1:1 line.

Other reviews have examined the impacts of modeling algorithm, the environmental variables included, sampling bias correction, and the choice of study extent on SDM performance based on internal-AUC (*e.g.,* Thuiller *et. al.,* 2004; Elith *et. al.,* 2011; Fourcade *et. al.,* 2014; Shabani *et. al.,* 2018; Muscatello *et. al.,* 2021). These same modeling decisions can be examined with respect to impacts on independent-AUC based on the studies reviewed here. Although statistical analyses are needed to formally assess these relationships, visual inspection of the data provides an initial look at general patterns. With respect to modeling algorithm, Maxent has previously been found to outperform other SDM algorithms (Aguirre-Gutierrez *et. al.,* 2013; Bradie and Leung, 2017; Elith *et. al.,* 2010). Among the current set of models, use of Maxent resulted in models that had lower independent-AUC on average than other modeling approaches including random forest and Generalized Linear Models and Generalized Linear Mixed Models (Fig. 1.3a). However, the amount of variation among models calibrated using Maxent was high and some of the models with the highest individual AUC scores were developed using Maxent (Fig. 1.3a). Consistent with the literature (Merow *et. al.,* 2013; Araújo *et. al.,* 2019; Seaborn *et. al.,* 2021), models based strictly on climate variables tended to produce lower AUC values on average than models that included or relied on non-climate variables (Fig. 1.3b). On average, the use of input thinning and bias files to account for sampling bias did not result in models that performed better than models that did not account for sampling bias (*i.e.,* Random background models; Fig. 1.3c), although the amount of variation was high. However, in line with the findings of previous assessments of the impacts of bias correction on model performance (*e.g.,* Hertzog *et. al.,* 2014; Barber *et. al.,* 2022), use of target-group background (localities of species related to the focal species that are expected to share the same sampling bias; Phillips *et. al.,* 2009) consistently produced well-performing models (*i.e.,* independent-AUC > 0.7) and these models had the highest mean independent-AUC (Fig. 1.3c). Finally, differences in model performance based on study extent were small, though models developed using large, range-wide study extents or regional extents had the highest average independent-AUC (Fig. 1.3d). Thus, although the above trends suggest that the impacts of some modeling decisions appear similar when evaluated by either internal or independent-AUC, further study may be required to fully understand how these decisions impact independent measures of model accuracy.

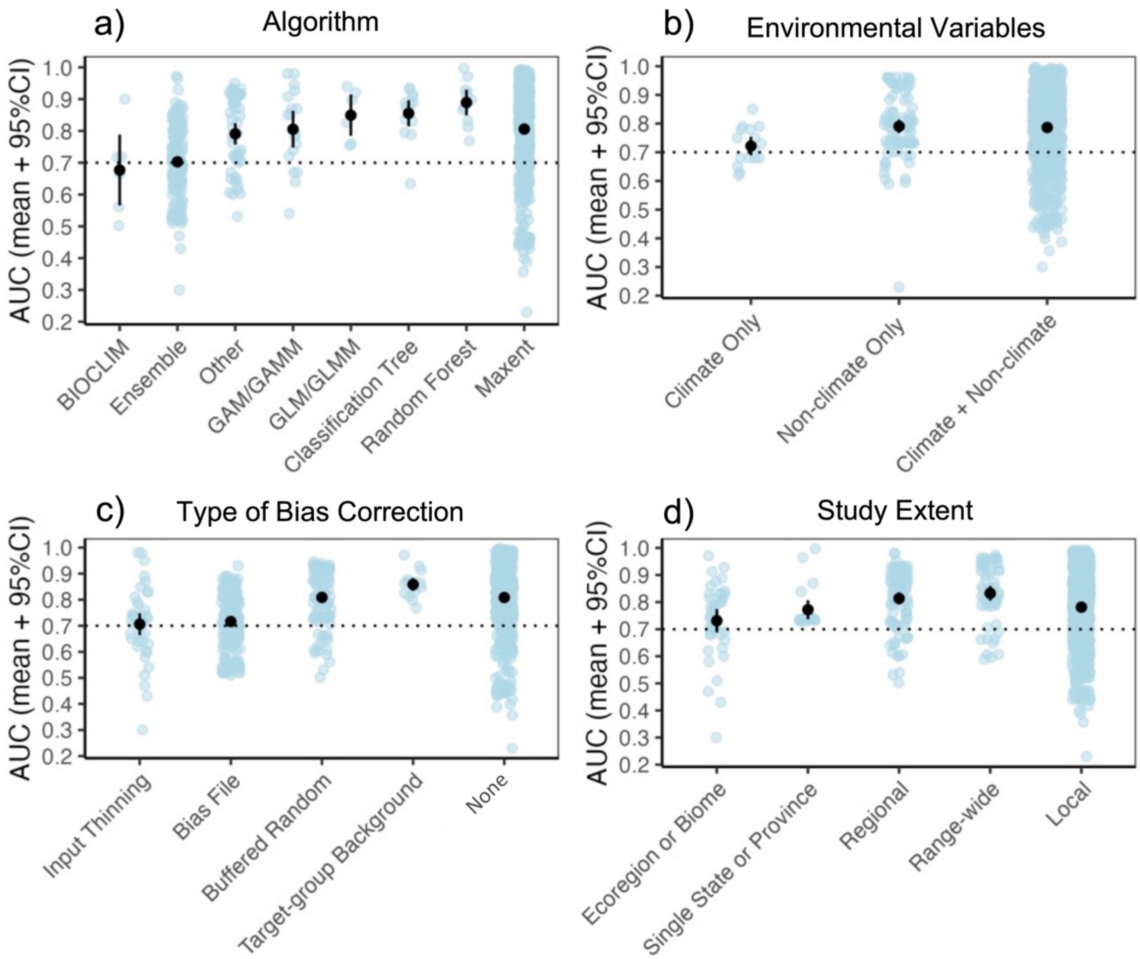


Figure 1.3. Impacts of model decisions on model performance derived from independent presence-absence survey data. Black dots represent means, and error bars represent 95% confidence intervals. Blue points represent values for the individual models from 39 studies included in this review. Definitions of the specific choices within each of these categories can be found in Table A1.2.

Although the above summary provides some guidance as to which modeling decisions tend to produce the most accurate models, testing the sensitivity of model performance to these decisions has been encouraged as a strategy to produce highly accurate SDMs (Araújo *et. al.,* 2019; Sofaer *et. al.,* 2019). The most frequently tested modeling decision among the studies reviewed here was choice of algorithm (14 studies, 36%), followed by type of bias correction (11 studies, 28%) and the type of environmental variables included in the models (9 studies, 23%; Fig. 1.4). In contrast, choice of study extent was only tested in five studies (13%), despite previous work suggesting that this decision can have a substantial impact on model performance (Thuiller *et al.,* 2004; VanDerWal *et al.,* 2009; Barbet-Massin *et al..,* 2010; Sánchez-Fernández *et al.,* 2011; Raes, 2012; Vale *et al.,* 2014). Three studies also tested the impacts of the resolution of the environmental variables used in the models, and five studies using Maxent additionally tested sensitivity to the regularization parameter (Fig. A1.1). However most notably, roughly a quarter of studies (26%) did not conduct sensitivity tests for any modeling decision (Fig. 1.4). Thus, despite repeated recommendations to perform sensitivity tests (*e.g.,* Araújo *et. al.,* 2019; Sofaer *et. al.,* 2019) and the potential for alternative modeling choices to impact model accuracy (Fig. 1.3), testing the impacts of model decisions on model performance using independent data has yet to become standard practice. Furthermore, of the studies that did conduct sensitivity tests, over half (55%) tested the impacts of just a single decision (Table 1.1). Recent studies have observed interacting impacts of model decisions on model performance (*e.g.,* Walker, 2018; Connor *et. al.,* 2019) highlighting the need for more studies that test multiple decisions simultaneously. For most of the models reviewed presently, average independent-AUC was above 0.7 regardless of the specific choices made during model calibration. However, the decisions that produced the most accurate SDMs on average did not always produce the best model within studies. Thus, there are unlikely to be singular modelling decisions that apply well to all studies, making it critical to conduct sensitivity testing when using SDMs.

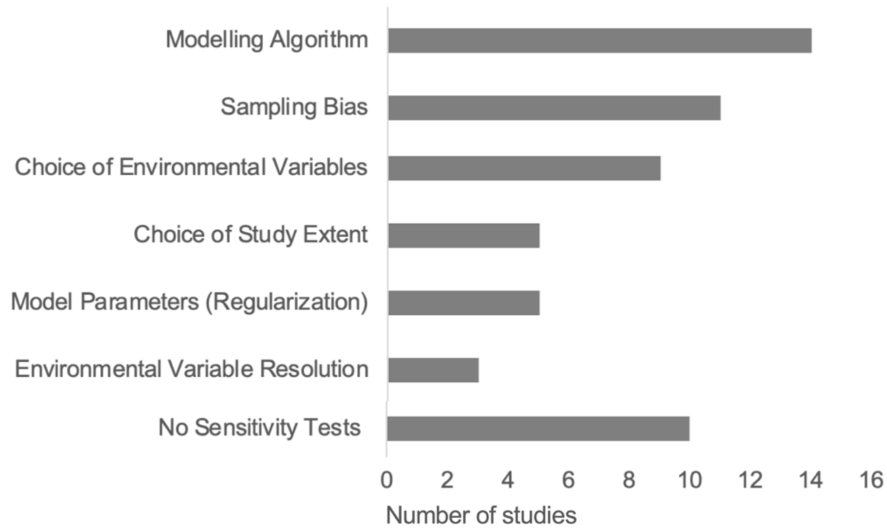


Figure 1.4. Frequency with which each modeling decision was tested across 39 studies validating SDMs with independent presence-absence data. Note that studies can be counted more than once, as some studies tested the impacts of more than one of these decisions.

1.3. Thesis Goals and Objectives

SDMs are one of the most widely used tools for understanding species distributions in ecology and conservation biology (Araújo *et. al.,* 2019). However, the above review reveals several gaps in understanding around factors impacting the accuracy of these models. Specifically, the sensitivity of models to certain modeling decisions (*e.g.,* the choice of study extent) requires further testing with independent data, particularly in underrepresented taxa and geographic areas. There is also a need for more studies that test multiple modeling decisions simultaneously to understand how they interact and impact downstream conservation decisions. Toward addressing these goals, the rest of my thesis is structured as follows:

- Chapter 2: Recognizing a need for a greater understanding of how choice of study extent impacts SDMs (see above review), I explore how this decision impacts model predictions and performance for six amphibian species in protected areas at the edge of their respective ranges in western Canada.

- Chapter 3: Towards expanding assessments of the joint impacts of multiple modeling decisions on SDMs, I simultaneously examine the impacts of choice of study extent, whether sampling bias is accounted for, and the type of environmental variables used on model predictions and performance for long-toed salamanders in southwestern Alberta. In this chapter, I additionally explore how these decisions impact the downstream prioritization of sites for a planned conservation translocation of this species.
- Chapter 4: Finally, I summarize my findings in light of the broader SDM literature, with particular focus on the use of SDMs at range edges.

My thesis focuses specifically on amphibians in western Canada. Amphibians are globally the most threatened vertebrate group, with ~40% of species experiencing some level of decline (Stuart *et. al.,* 2004; IUCN/SSC, 2013; Alroy, 2015), and 54% of species in Canada listed as either Endangered, Threatened or Special Concern (SARA public registry, June 2023). SDMs have been used for this group to understand drivers of their distributions (Seaborn *et. al.,* 2021), predict species invasions (Forti *et. al.,* 2017), and map species richness (Milanovich *et. al.,* 2012). However, few studies have rigorously tested the impacts of modeling decisions on SDMs developed specifically to inform amphibian conservation. My thesis fills this gap while simultaneously adding to our understanding of how key modeling decisions impact SDM predictions and performance, and downstream conservation decisions.

CHAPTER 2: THE IMPACTS OF STUDY EXTENT ON SPECIES DISTRIBUTION MODELS FOR SIX AMPHIBIANS AT THE EDGE OF THEIR RANGES IN WESTERN CANADA

2.1. Abstract

Species distribution models (SDMs) are a widely used tool in ecology and conservation biology. However, several decisions made during model development can impact model predictions and performance, including the choice of study extent. One context in which the choice of study extent may be particularly important is when using SDMs to predict occurrence in peripheral parts of species' ranges as environmental conditions at the edge of the range may differ from those experienced by the species across the range. In this study, I compared SDMs developed using alternative study extents for six amphibian species at the edge of their ranges in Alberta, Canada. I found that choice of study extent had substantial, yet unpredictable impacts on model predictions and the variables considered to be most important for predicting presence for each species. I further evaluated model performance using independent data and found that only three of the twelve models for which independent data were available were sufficiently accurate as to be useful. These results suggest that it is important to consider the impacts of study extent on model predictions and performance, and that in addition to independent testing of models, the consideration of alternative methods for identifying suitable habitat may be required at range edges when these models fail.

2.2. Introduction

Understanding species distributions is a fundamental goal in ecology (Guisan and Zimmermann, 2000). Species distribution models (SDMs) are widely used to both describe species distributions and to identify environmental variables that influence the distribution of populations (Elith *et. al.,* 2006; Phillips *et. al.,* 2006; Randin *et. al.,* 2006). Developing an SDM involves using information about environmental conditions at locations where a species is present and using an algorithm that distinguishes between those conditions, and either background conditions (presence-only models) or conditions where the species is absent (presence-absence models; Elith *et. al.,* 2006; Phillips *et. al.,* 2006). The resulting model can then be projected across geographic space to identify areas that are highly suitable for the focal species, or where the focal species is likely to occur (Merow *et. al.,* 2013). These models have been used to find new populations of rare species (Groff *et. al.,* 2014; McCune, 2016), choose populations for protection (Muscatello *et. al.,* 2021; Stryszowska *et. al.,* 2016), and select release sites for translocations (Guisan *et. al.,* 2013; Malone *et. al.,* 2018; Maes *et. al.,* 2019).

Several input decisions made during model development can impact model predictions and performance (Elith *et. al.,* 2011; Merow *et. al.,* 2013; Araújo *et. al.,* 2019). One such decision is the choice of study extent (Thuiller *et. al.,* 2004; Phillips *et. al.,* 2009; VanDerWal *et. al.,* 2009; Anderson and Raza, 2010; Elith *et. al.,* 2011; Merow *et. al.,* 2013). This decision defines the geographic area from which the locality data (*i.e.,* presence-absence or presence-background) used to calibrate a model are drawn and thus influences the range of environmental conditions that are considered (Thuiller *et. al.,* 2004; VanDerWal *et. al.,* 2009; Merow *et. al.,* 2013). Several studies have demonstrated that the total area predicted to be suitable for a species can be sensitive to this decision (Thuiller *et. al.,* 2004; VanDerWal *et. al.,* 2009; Barbet-Massin *et. al.,* 2010; Sánchez-Fernández *et. al.,* 2011; Raes, 2012; Vale *et. al.,* 2014; Connor *et. al.,* 2019). Yet, testing the sensitivity of models to this decision is not common practice. More importantly, few studies have assessed how this decision impacts the ability of models to predict independent presence-absence data (*i.e.,* model accuracy; but see Gogol-Prokurat, 2011; Trumbo *et. al.,* 2011; Searcy and Shaffer, 2014; Konowalik and Nosol, 2021).

Depending on context, certain study extents may offer advantages over others. For example, using the full geographic range of a focal species as the study extent is expected to capture the environmental conditions that broadly define the species' niche and the range of conditions that the species can occupy (Thuiller *et. al.,* 2004; Elith *et. al.,* 2011). On the other hand, more limited extents may allow models to hone in on conditions that influence local distributions (Elith *et. al.,* 2011) and may result in more accurate models when the area of interest harbours conditions that are not represented elsewhere in the range (Vale *et. al.,* 2014) and/or when populations are locally adapted (Halbritter *et. al.,* 2015; Hällfors *et. al.,* 2016). Studies that develop these "regional" SDMs often rely on political boundaries (*e.g.,* state, provincial, or county borders) due to jurisdictional or regional conservation objectives (Pineda and Lobo, 2009; Raes, 2012). However, biologically informed regional study extents may be more appropriate in some cases (Peterson, 2011; Araújo and Peterson, 2012). For example, ecoregions capture distinct climatic and biological communities and thus may be an appropriate way to define study extent when the goal is to model locally adapted populations or to home in on localized environmental conditions (Raes, 2012; Smith *et. al.,* 2019). Thus, models developed using biologically informed study extents may improve model accuracy in some cases (Hällfors *et. al.,* 2016; Smith *et. al.,* 2019; Chardon *et. al.,* 2020; Jinga *et. al.,* 2021).

One context in which the choice of study extent may be particularly important is when using SDMs to predict occurrence at the periphery of species' ranges. On the one hand, peripheral populations may occupy marginal conditions within the range and may be locally adapted to these conditions (Kawecki, 2008; Hargreaves *et. al.,* 2014; Vale *et. al.,* 2014; Bontrager *et. al.,* 2021). Thus, restricting the study extent to the political jurisdiction or ecoregion(s) found at the edge of the range might result in more accurate models (Searcy and Shaffer, 2014; Hällfors *et. al.,* 2016). On the other hand, range limits do not always correspond with marginal or extreme conditions (Oldfather *et. al.,* 2019; Bontrager *et. al.,* 2021). Constraints on dispersal (Svenning *et. al.,* 2008), or an intensification of biotic interactions (Paquette and Hargreaves, 2021) may exclude species from some suitable locations at the edge of their range. Thus, locality records at the range edge may fail to adequately represent the full set of conditions that can be tolerated by species in these areas (Barbet-Massin *et. al.,* 2010; Sánchez-Fernández *et. al.,* 2011; Raes, 2012). In such cases, use of a broader or range-wide extent may result in models that better characterise suitable habitat at the edge of the range (Barbet-Massin *et. al.,* 2010; Raes, 2012; Thuiller *et. al.,* 2004).

In high-latitude countries, many at-risk species are at the edge of their range (Hunter and Hutchison, 1994). Understanding the impact of modeling decisions on the predictions and performance of SDMs at species' range limits thus has implications for conservation directives that are based on SDMs in these jurisdictions. Focusing on model predictions for two federally protected areas, I explore the impacts of study extent on

SDMs for six of the ten amphibian species that occur in Alberta, Canada. Amphibians are globally the most threatened vertebrate group, with ~40% of species experiencing declines (Stuart *et. al.,* 2004, Pimm *et. al.,* 2014; Alroy, 2015), and in Alberta, there are general gaps in our understanding of the distributions and habitat requirements of these species. In this study, I specifically ask: 1) What impact does choice of study extent have on model predictions for amphibians at the edge of their range? 2) Can models accurately predict independent presence-absence data at the edge of the range? And if so, 3) do some study extents outperform others in this regard?

2.3. Methods

2.3.1. Focal species and area of interest

This study focuses on six pond-breeding amphibians that have range limits in Alberta, Canada. The long-toed salamander (*Ambystoma macrodactylum*), Columbia spotted frog (*Rana luteiventris*), and western toad (*Anaxyrus boreas*) are at the eastern edge of their range in Alberta, whereas the western tiger salamander (*Ambystoma mavortium*), boreal chorus frog (*Pseudacris maculata*), and Canadian toad (*Anaxyrus hemiophrys*) are at the western edge of their range in the province (Fig. 2.1). Although most of these species are listed as Sensitive in Alberta, all are geographically widespread in North America. Thus, there are sufficient locality records for each of these species to build models using increasingly restricted study extents without reducing the number of locality records to numbers that can be problematic for models (Wisz *et. al.,* 2008).

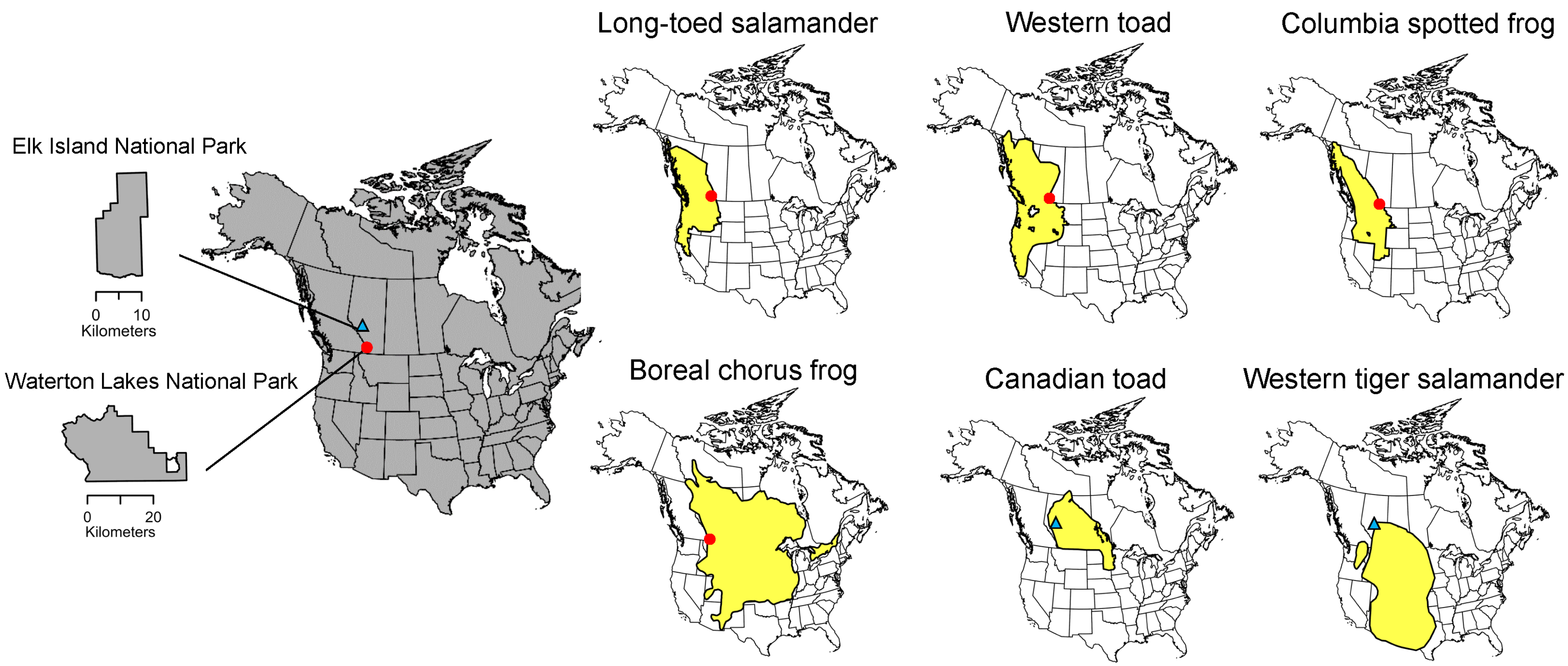


Figure 2.1. Locations of the two focal protected areas relative to the geographic distributions of the six amphibian species included in this study. Geographic range of each species are based on their International Union for the Conservation of Nature (IUCN) range polygons. Available from: iucnredlist.org.

To evaluate the impact of study extent on model predictions and accuracy for these species, I focused on two protected areas in Alberta, each harbouring peripheral populations of a subset of the focal species: Waterton Lakes National Park (WLNP), located along the transition between the Rocky Mountains and prairies in the southwestern corner of Alberta, and Elk Island National Park (EINP), located in the prairies and aspen parklands of central Alberta (Fig. 2.1). WLNP is engaged in several amphibian conservation efforts including reintroductions, the installation and maintenance of amphibian road crossing infrastructure, and a long-term monitoring program. Data from the latter monitoring program provides an opportunity to test the accuracy of models based on publicly available locality records (see below) with an independent presence-absence dataset for four of the six species in this study (the long-toed salamander, western toad, Columbia spotted frog, and boreal chorus frog). Independent presence-absence data were not available for assessing model accuracy for the remaining two species in EINP (western tiger salamanders and Canadian toads). However, these species allowed me to test the impacts of study extent on model predictions for a second protected area, expanding the context of my results.

2.3.2. Alternative study extents

For each of the six focal species, I tested the impacts of three different study extents on SDM predictions and performance in the respective focal protected area: a) a "range-wide" extent that encompassed the full geographic range of each species based on their IUCN range polygons (IUCN: http://www.iucnredlist); b) a "political" extent corresponding to the Alberta Environment and Parks Land Division Region (Government of Alberta; https://open.alberta.ca/opendata) to which either WLNP or EINP belong; and

c) an "ecoregion" extent based on the WWF Terrestrial Ecoregions of the World (Olson *et. al.,* 2001; https://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world) to which either WLNP or EINP belong (examples of the different extents are shown for long-toed salamanders and Canadian toads in Fig. 2.2). I was interested in the conditions that differentiate presence-absence within the range (rather than conditions that set the overall range limits of each species). Thus, all study extents were clipped to the species' full range in ArcGIS Pro to avoid overfit models that result from the inclusion of background points from far outside of the species' range.

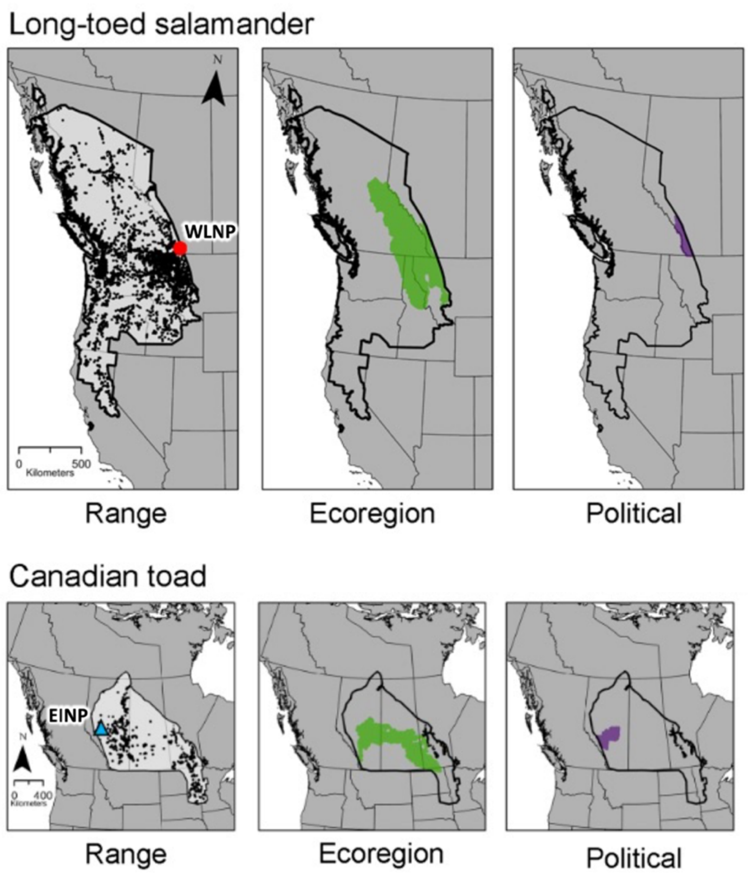


Figure 2.2. The range-wide (grey shading in left-panel and black polygon in all panels), ecoregion (green), and political (purple) study extents used to model the distributions of two amphibians at the edge of their ranges in southwestern Alberta. The left-most panels show the focal protected areas, Waterton Lakes National Park (WLNP) and Elk Island National Park (EINP), as a red dot and blue triangle respectively. Small black points in the left-most panels indicate the available locality data. See Fig. A2.1 for the different study extents used for all species included in this study.

2.3.3. Model development

SDMs were developed using Maxent (version 3.2.3; Philips *et. al.,* 2006) in R: *dismo* (R Core Team Version 4.2.1; Hijmans *et. al.,* 2021). Maxent has been widely used to develop presence-only SDMs and has been shown to perform as well or better than other algorithms for predicting occurrence (Elith *et. al.,* 2010; Aguirre-Gutierrez *et. al.,* 2013; Bradie and Leung, 2017). Input locality records were collected from the Global Biodiversity Information Facility (GBIF: https://www.gbif.org/ on October 21, 2021), in addition to collection of records from other researchers and government databases from across Canada and the United States. Records were filtered to exclude those collected before 1990, and with coordinate uncertainty > 1 km to match the spatiotemporal resolution of environmental data (see below). Duplicate records (*e.g.,* multiple observations of a species at the same site) were removed from the dataset and only one locality per cell of the environmental layers was retained for each species. After filtering, the total number of input locality records used to build the models ranged from 34 to 14,034 depending on species and study extent (Table 2.1). Locality data for amphibians are behind government paywalls for Wyoming and Nevada and I excluded these states from model calibration for the range-wide extents of western toads and Columbia spotted frogs due to lack of data. All locality filtering steps were conducted in R, using the packages *sf* (Pebesma, 2018), *tidyverse* (Wickham *et. al.,* 2019), and *dismo* (Hijmans *et. al.,* 2021).

SDMs were developed using sixteen bioclimatic variables. These variables were selected *a priori* for their biological relevance to pond-breeding amphibians and have been shown to be important predictors in previously published SDMs for one or more of the focal species (Table A2.1). Bioclimatic variables were downloaded from ClimateNA (Wang *et. al.,* 2016: https://adaptwest.databasin.org/pages/adaptwest-climatena/) at a spatial resolution of 30 arc seconds (1km) and represent 30-year climate normals between 1990 and 2020. Although factors other than broad-scale, average, climatic conditions are expected to impact amphibian occurrence (Thuiller *et. al.,* 2004), other variables (*e.g.,* vegetative cover) are often correlated with general climatic conditions (Araújo and Peterson, 2012) and models based on broad-scale, climate normals have been shown to accurately predict ectotherm distributions, including amphibian distributions (Franklin *et. al.,* 2009; Trumbo *et. al.,* 2011; Groff *et. al.,* 2014). Furthermore, low-resolution climate layers are among the most widely accessible GIS layers and thus commonly used in the SDM literature (Bradie and Leung, 2017; Fourcade *et. al.,* 2018). Testing the impacts of study extent on models developed using these variables is thus relevant to the broader literature.

As recommended by Merow *et. al.,* (2013), I tuned both the feature class and regularization multiplier values for each model using cross-validation prior to producing the final models. Six different values of the regularization parameter (0.25, 0.5, 1, 1.5, 2, and 4), and nine different feature class combinations (L, LQ, H, LQH, LQP, LQT, LQHP, LQPT, and LQHPT; where L = linear, Q = quadratic, H = hinge, P = product and T = threshold) were tested using R: *ENMeval* (Kass *et. al.,* 2021). The optimal feature combination and regularization parameter for each study extent (Table 2.1) was then selected for use in the final models using the Akaike information criterion correction (AICc; Muscarella *et. al.,* 2014).

Table 2.1. Details of the species distribution models used to test the impact of study extent on model predictions and performance for six amphibian species at the edge of their range.

Protected area of interest	Species	Study Extent	No. Input localities	No. Random background points	Features*	Regularization
Waterton Lakes National Park	Long-toed salamander	Range	5083	10,000	LQH	0.25
		Ecoregion	2020	10,000	LQPT	0.25
		Political	121	5000	LQP	0.25
	Western toad	Range	7275	10,000	H	0.25
		Ecoregion	1475	10,000	LQHP	0.25
		Political	128	5000	L	0.25
	Boreal chorus frog	Range	14,034	10,000	LQHPT	0.25
		Ecoregion	86	10,000	L	0.25
		Political	1185	5000	LQHP	0.25
	Columbia spotted frog	Range	6487	10,000	LQHPT	1.5
		Ecoregion	2677	10,000	LQT	1
		Political	139	5000	LQ	0.25
Elk Island National Park	Canadian toad	Range	947	10,000	LQH	0.25
		Ecoregion	339	10,000	LQT	1
		Political	34	5000	LQ	2
	Western tiger salamander	Range	2465	10,000	LQHP	0.25
		Ecoregion	240	10,000	LQHP	0.5
		Political	66	5000	LQ	0.25

*L = linear, Q = quadratic, H = hinge, P = product and T = threshold.

2.3.4. Model predictions

I projected the models for each species across the focal region of interest (WLNP or EINP depending on species) and visually compared the continuous prediction surfaces based on Maxent's logistic output. The logistic output is independent of spatial scale and is the most appropriate tool for comparing models developed using different study extents (Merow *et. al.,* 2013). When using Maxent's logistic output, a 'prevalence' value (τ) must be set (Merow *et. al.,* 2013). In the absence of additional independent data with which to estimate prevalence (Elith *et. al.,* 2011), I used the Maxent default setting of $\tau = 0.5$. I acknowledge that this may not be the most appropriate value for the species studied here and there may be differences in the actual prevalence of species in the focal study areas. However, inaccurately setting prevalence is expected to create problems only when predictions are compared across different regions or species in absolute terms (Merow and Silander, 2014; Guillera-Arroita *et. al.,* 2015), which is not how models are used in this study.

I converted continuous model predictions into binary surfaces of suitable habitat to facilitate comparison of the models for each species. For each model, I applied a threshold for considering a cell as suitable that achieved a 10% omission rate of the input presences. This approach to setting a threshold is common when prevalence is unknown and has been shown to outperform other thresholds at the edge of species' ranges (Vale *et. al.,* 2014). I used formal map comparisons procedures from the Map Comparison Kit version 3.2.3 (Visser and de Nijs, 2006) to compare the binary model predictions for each species. Specifically, I calculated pairwise Fuzzy Kappa between surfaces using the Map Comparison Kit V.3.2.3 (Visser and de Nijs, 2006; Hagen-Zanker, 2009). Fuzzy Kappa ranges from 1 to -1, with positive values indicating total similarity, values of zero indicating total dissimilarity, and negative values indicating both dissimilarity and differences in spatial autocorrelation between map surfaces (Hagen-Zanker, 2009). Finally, I overlaid the binary prediction surfaces for each species to visually identify areas of agreement across models generated using different study extents.

2.3.5. Model performance

All models were first evaluated using the area under the receiver operating characteristic curve (AUC) based on five-fold cross-validation (Guisan and Zimmermann, 2000; "internal-AUC"). I was also able to test model accuracy using independent presence-absence data for the four species in WLNP based on surveys conducted between 1997 and 2017 at 40 sites in WLNP. Each site was in a different 1 km cell. Species were considered present at a site if at least one observation was recorded at that site during these surveys. I used this dataset to calculate independent-AUC for each of the models for these species. Threshold-dependent metrics are commonly reported in the literature, and for comparison purposes, I also included estimates of sensitivity, specificity, and the true skill statistic (TSS = sensitivity + specificity -1; Allouche *et al.,* 2006) calculated using the independent dataset and the binary surfaces described above. All threshold-dependent and -independent evaluation metrics were calculated in R: *PresenceAbsence* (Freeman and Moisen, 2008). I note that limited sample sizes and differences in the monitoring protocols used by the park across years precluded evaluation of the models using other methods (*e.g.,* formal occupancy models: Devarajan *et. al.,* 2020 or testing calibration performance: Vaughan and Ormerod, 2005).

2.4. Results

2.4.1. Impacts of study extent on model predictions

Within species, choice of study extent had substantial impacts on model predictions. Visual comparison of prediction surfaces for each species showed high variability in both the range of suitability scores as well as the spatial distribution of suitable habitat (Fig. 2.3; Fig. 2.4). Except for the boreal chorus frog, models based on range-wide study extents produced more uniform predictions for the target conservation areas, whereas models based on smaller extents produced more variable surfaces. The percent agreement among the summed binary prediction surfaces varied across species, ranging from 10% (long-toed salamander) to 76% (western tiger salamander; Table A2.2; Table A2.3). Estimates of Fuzzy Kappa based on the binary surfaces further illustrate the variability in the amount and distribution of habitat predicted to be suitable for each species. For example, whether individual cells were suitable or not for western toads in WLNP was heavily dependent on which model was used (negative Fuzzy Kappa values; Fig. 2.4). However, the degree to which model extent impacted the binary surfaces varied by species, for instance, there was a high degree of overlap in predicted suitable habitat across models for western tiger salamanders in EINP (fuzzy kappa between 0.580 and 1.00; Fig. 2.4)

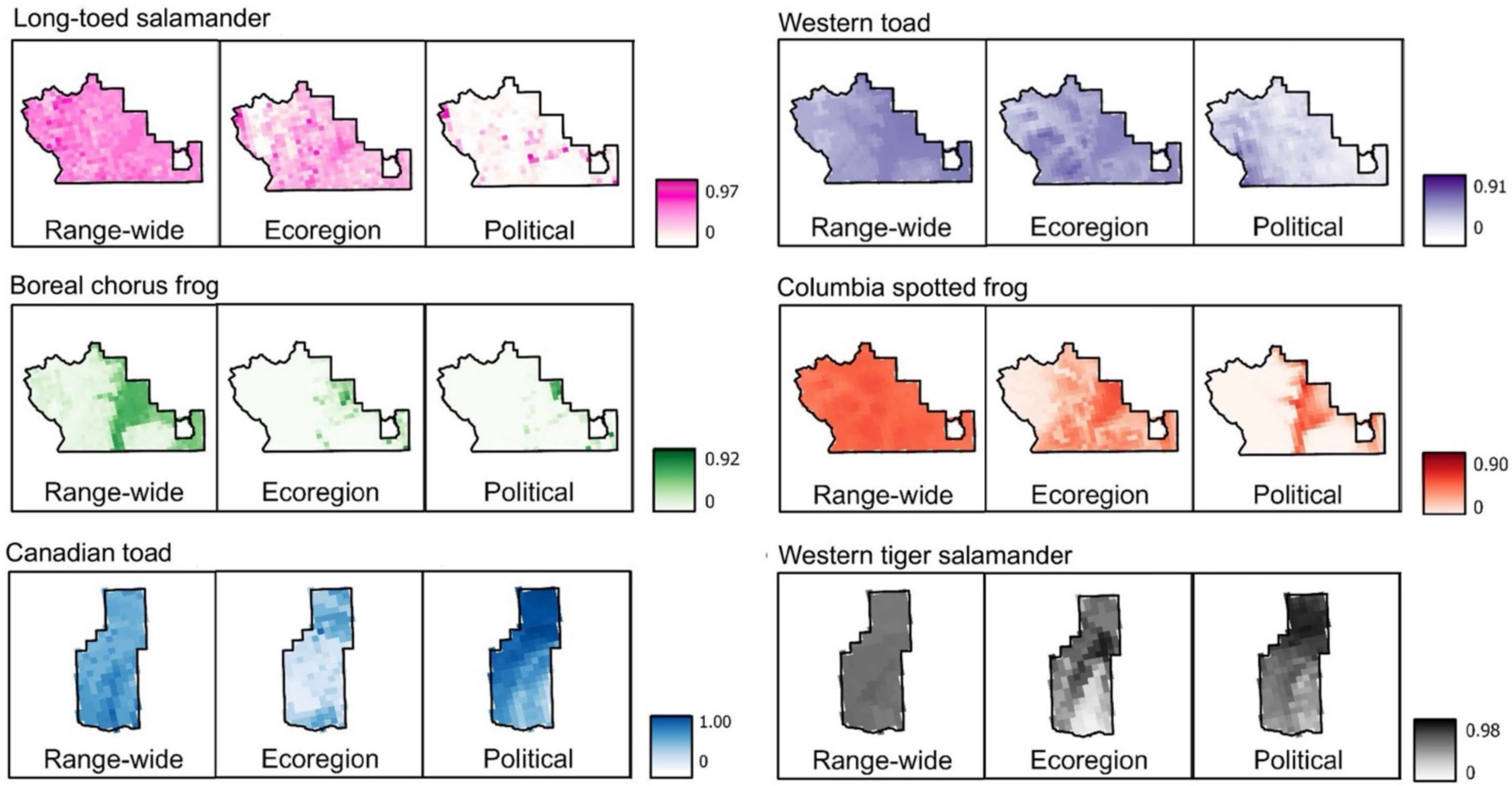


Figure 2.3. Habitat suitability for six amphibian species based on species distribution models developed using different study extents. Model predictions are shown for Waterton Lakes National Park for the long-toed salamander, western toad, boreal chorus frog, and Columbia spotted frog and Elk Island National Park for the Canadian toad and western tiger salamander. Darker colours in all panels indicate higher predicted suitability. Models were developed in Maxent using sixteen bioclimatic variables

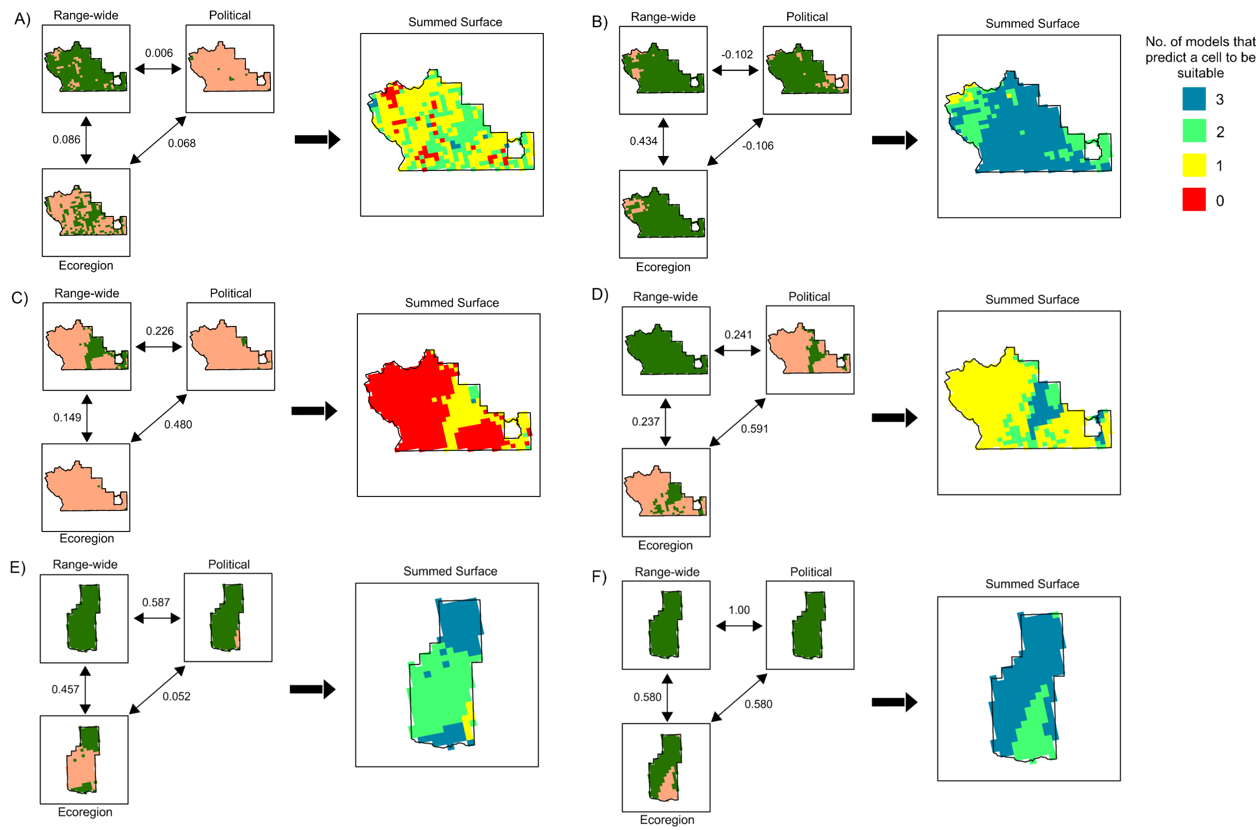


Figure 2.4. Binary predictions based on species distribution models developed using different study extents for six amphibian species in Alberta, Canada (smaller maps in each panel). Binary surfaces are shown for Waterton Lakes National Park for (A) long-toed salamanders, (B) western toads, (C) boreal chorus frogs, and (D) Columbia spotted frog, and for Elk Island National Park for (E) Canadian toads and (F) western tiger salamanders. Models were developed in Maxent using sixteen bioclimatic variables. The 10% omission rate of input localities for each model was used as the threshold for calling cells suitable (green) or unsuitable (pink). Pairwise Fuzzy Kappa values are shown above the arrows between surfaces. Larger maps in each panel show model agreement (number of models predicting a cell as suitable).

Notably, model settings also impacted the variables considered to be important by Maxent. Considering permutation importance in particular, all species had one or more variables that came out in the top three most important predictors for at least two of the models (Table 2.2). There was likewise overlap across the models for each species in the variables with the lowest permutation importance (Table 2.2). Nevertheless, there were substantial differences in the permutation importance of at least one variable depending on study extent for all species (Table 2.2). For example, mean temperature of the coldest month (MCMT) was among the top three variables for the political and ecoregion models for Canadian toads but had low permutation importance for the range-wide model. Likewise, number of frost-free days (NFFD) was among the top variables for the political model for Western toads but was within the bottom four variables for the ecoregion and range-wide models for this species. Thus, alongside impacts on the amount and distribution of predicted suitable habitat, choice of study extent influenced the variables contributing most to predicting the probability of presence of each species.

Table 2.2 Permutation importance (%) of sixteen bioclimate variables for Maxent models generated for six amphibian species using different study extents.

Species	Study Extent	MAT	MWMT	MCMT	Tave_at	Tave_wt	Tave_sm	Tave_sp	SHM	NFFD	MSP	MAP	PPT_sm	PPT_wt	PPT_sp	TD	RH
Long-toed salamander	Range	4.5	25.8	2.6	8.1	7.2	9.8	8.1	3.4	2.7	3	8.8	6.6	3	3.7	2.3	0.4
	Ecoregion	1.3	15.5	3.7	9.9	9.3	7.5	4.7	8.1	4	7.1	0.7	18.6	1.4	2.3	1.2	4.7
	Political	0.6	7.8	1.8	22.6	2.9	8.9	9.3	12.3	2.8	11.9	0.4	12.1	4	0.9	0.2	1.5
Western toad	Range	0	15.6	4.8	15.7	3.9	0.9	14.4	2.8	1.1	3.7	7.6	18.3	3.1	1	4.2	3
	Ecoregion	0	7.8	0.9	4	10.7	1.2	12.3	10.4	1.1	1.3	1.7	20.9	1.1	10.1	4.2	12.4
	Political	0	6.8	0	10.2	10	18	10.8	0.7	14.4	1.7	3.9	13.8	0	0.4	0.9	8.5
Boral chorus frog	Range	5.4	3.7	3.1	4.1	7.5	4.2	17.3	1.7	10.5	6.2	15.6	4.8	2	5.5	6	2.5
	Ecoregion	0.4	0.4	3.2	28.8	14.8	1.6	4.7	22.6	1.3	0.3	1.8	9.5	6.6	2	1	1
	Political	3.8	7.8	4.5	2	9.4	4.2	4	8.5	13.2	5.8	4.1	6.1	6.6	7.7	5	7.4
Columbia spotted frog	Range	0	21	1.3	17.5	0	18.3	10.4	5.2	12.1	0	0	2.8	2.3	7.8	1.2	0.1
	Ecoregion	0	10.5	5.1	0	0	0.3	4.6	8.2	1.6	0.3	4.3	13	0.4	14.8	2.5	34.3
	Political	0	0.1	0.2	28.3	0	11.1	19.6	2.7	9.2	0	1.7	18	1.3	2.3	1.5	3.8
Canadian toad	Range	1	8.7	2.2	5.4	12.1	4.9	17.7	1.6	2.5	2.9	12.9	7.6	1.1	6.9	11.7	0.6
	Ecoregion	2.7	1.4	12.3	6.9	9.7	13.1	13.6	2.4	3	5.5	3	5.5	1.9	6.5	10.4	2.1
	Political	3.9	0.9	18.3	0.9	7.9	6.8	4.2	1	10.7	5	6.9	8.7	8.8	10.3	1.1	4.5
Tiger salamander	Range	2.7	1.9	4.5	20.9	7.8	5.3	9.3	1.6	12.1	3	11.6	4.3	3.2	7.9	2.6	1.3
	Ecoregion	1.3	2.2	5.5	12.8	5	4.7	1.1	8.9	4.2	11.8	15.3	2.8	3.1	8.8	12.2	0.4
	Political	3.8	11.1	6.7	3.3	1.6	0.6	0	7.8	17.9	9.1	14.1	13	0.6	4.3	2.8	3.3

MAT = mean annual temperature, MWMT = mean temperature of the warmest month, MCMT = mean coldest month temperature, Tave_at = autumn (September-November) mean temperature, Tave_wt = winter (December – February) mean temperature, Tave_sm = summer (June-August) mean temperature, Tave_sp = spring (March-May) mean temperature, SHM = summer heat moisture index, NFFD = number of frost-free days MSP = mean summer (June-August) precipitation, MAP = mean annual precipitation, PPT_sm = summer (June-August) mean precipitation, PPT_wt = winter (December – February) mean precipitation, PPT_sp = spring (March-May) mean precipitation, TD = continentality, RH = relative humidity,

2.4.2. Impacts of study extent on model performance

Cross-validated AUC ranged from 0.675 to 0.933 and was greater than 0.7 for all models except for the Canadian toad model built with a political study extent (Table 2.3). In contrast to cross-validated results, few models passed independent validation (Table 2.3). Specifically, independent-AUC was below the cut-off for accepting a model (< 0.7, Swets, 1988) for nine out of the 12 models, and none of the models for long-toed salamanders or Columbia spotted frogs passed independent validation. Of the three models that passed independent validation, two were developed using political study extents (Table 2.3). The political extent model for the Columbia spotted frog was also close to the cut-off for considering a model useful (AUC = 0.67, Table 2.3). In no case did using the range-wide study extent produce a model passing independent validation at the edge of the range, whereas the ecoregion model for boreal chorus frogs was acceptable (AUC = 0.72). Results based on threshold-dependent metrics showed similar variability within species (Table 2.3). However, models tended to have either a high sensitivity (true positive rate) or a high specificity (true negative rate), resulting in overall low TSS values.

Table 2.3. Results from cross-validated and independent evaluation of Maxent models developed from different study extents for six amphibian species in protected areas at the edge of their range. Bold text indicates models which had 'fair' discriminatory ability; as defined by independent-AUC above 0.70 (Swets, 1988). Sensitivity, Specificity, and TSS are based on binary surfaces created using the 10% omission rate of input presences, and 40 independently surveyed sites.

Protected area of interest	**Species**	**Study Extent**	**Cross-validated AUC**	**Independent AUC**	**Sensitivity**	**Specificity**	**TSS**
Waterton Lakes National Park	Long-toed salamander	Range	0.820	0.529	0.969	0	-0.031
		Ecoregion	0.846	0.537	0.656	0.625	0.281
		Political	0.867	0.561	0.063	1	0.063
	Western toad	Range	0.738	0.566	1	0	0
		Ecoregion	0.805	0.640	0.971	0	-0.029
		Political	0.829	**0.766**	0.657	0	-0.343
	Boreal chorus frog	Range	0.933	0.662	0.767	0.400	0.167
		Ecoregion	0.890	**0.720**	0	1	0
		Political	0.789	**0.710**	0.167	1	0.167
	Columbia spotted frog	Range	0.839	0.567	1	0	0
		Ecoregion	0.857	0.581	0.500	0.357	-0.143
		Political	0.828	0.666	0.500	0.786	0.286
Elk Island National Park	Canadian toad	Range	0.882	-	-	-	-
		Ecoregion	0.900	-	-	-	-
		Political	0.675	-	-	-	-
	Western tiger salamander	Range	0.801	-	-	-	-
		Ecoregion	0.848	-	-	-	-
		Political	0.729	-	-	-	-

2.5. Discussion

I explored how the choice of study extent impacted the predictions and performance of species distribution models at the edge of species' ranges for six amphibians in two National Parks in Canada. I found that SDM predictions differed substantially depending on the study extent used to develop the models. Furthermore, models tested with independent presence-absence data varied in their accuracy, with most models performing poorly in this regard. My results emphasize the need to think carefully about the choice of study extent when developing SDMs and reveal some of the impacts this decision can have on model outputs. My results also highlight the challenges of developing accurate models in peripheral parts of species' ranges, where species may not be in equilibrium with climate.

2.5.1. Impacts of choice of study extent on model predictions

I found substantial differences in SDM predictions in protected areas at the edge of the range depending on the study extent used for all six amphibian species included in this study. This was especially true when comparing continuous predictions, though there were often substantial differences in binary predictions of suitable versus unsuitable habitat depending on the species. Similar effects of study extent on the amount and distribution of suitable habitat predicted from SDMs have been observed in other species from different taxonomic groups in a variety of contexts (VanDerWal *et. al.,* 2009; Anderson and Raza, 2010; Anderson and Gonzalez, 2011; Searcy and Shaffer, 2014; Walker 2018; Connor *et. al.,* 2019; Schmidt *et. al.,* 2020). In terms of studies working at the edge of species' ranges specifically, Trumbo *et. al.,* (2011) found that predictions were substantially different between models developed using state borders versus range-wide study extents for four amphibian species at their range-edge in Missouri, USA. Likewise, Vale *et. al.,* (2014) found differences in both the amount and distribution of suitable habitat predicted to occur in peripheral areas of the ranges of two mammals and one amphibian in west Africa when regional versus continent-wide study extents were used to develop models. Collectively, these results suggest that choice of study extent can have a strong impact on SDM predictions and can thus have consequences for using SDMs to identify critical habitat (Guisan and Thuiller, 2005), select sites for protection or restoration (Westwood *et. al.,* 2020), or predict extirpation risk (Barbet-Massin *et. al.,* 2010).

In line with the findings of previous studies (Thuiller *et. al.,* 2004, Trivedi *et. al.,* 2008; VanDerWal *et. al.,* 2009; Trumbo *et. al.,* 2012; Searcy and Shaffer, 2014, Walker, 2018), I observed differences in the relative importance of environmental variables across different study extents, offering a clear explanation for the observed differences in model predictions (Elith *et. al.,* 2011, Merow *et. al.,* 2013). Specifically, to develop the response curve for a given variable, Maxent compares the ratio of the probability density of the variable at presence locations to the probability density of the variable at background locations (see Fig. 1 in Merow *et. al.,* 2013) and thus models capitalize on variables that maximize the contrast between conditions at presence versus background conditions. There are two ways for the response curves of a given variable to change as the study extent becomes more restrictive: a narrowing of environmental conditions represented at presence locations relative to the background sample and/or truncation of the environmental conditions represented at background locations. Both may result simply from considering a smaller geographic extent. However, the former could also reflect local adaptation (Trumbo *et. al.,* 2011; Searcy and Shaffer, 2014) or the restriction of populations to a subset of the niche (possibly marginal conditions) at the edge of the range because of novel or increasingly intense biotic interactions (*e.g.,* MacArthur, 1984, Guisan *et. al.,* 2006, Halbritter *et. al.,* 2015, Hargreaves and Eckert, 2019). Whether the species included in this study are locally adapted or otherwise limited in the habitats they occupy at the edge of their range awaits further investigation. The differences in the variables that were most important at range-wide versus political extents based on the SDMs presented here may help guide such studies. For example, for long-toed salamanders, average autumn temperature was the most important variable at the political study extent but was unimportant at larger extents (Table 2.2). Thus, it may be worth following up with field and/or field studies testing whether the species is locally adapted with respect to this variable at their range edge.

Of note were differences between the two protected areas in terms of the amount of agreement among models for the species considered. Models for the two species in EINP showed more agreement in their predictions than models for the species in WLNP. The reasons for these differences are not entirely clear. However, apart from the ecoregion model for Canadian toads, predictions for the two species in EINP tended to be uniformly high. It is possible that EINP is simply a highly suitable area for the two species considered, and that, regardless of the variables emphasized by the different models, the models for each species capture this suitability. Alternatively, differences in topographic complexity (*i.e.,* there is a 1,635 m elevational difference between the highest and lowest point of WLNP, and only a 47 m difference in EINP) or other features of the environment may shape the relative evenness and variability of predictions observed across the two focal regions. Regardless of the explanation, my results indicate that choice of study extent may impact model predictions in some areas more than others–a possibility that would be interesting to explore across a wider array of species and protected areas.

2.5.2. Implications for developing SDMs at species' range limits.

Ultimately, the ability of models to predict independent presence-absence data is the best criterion for deciding whether a model is biologically meaningful for mapping occurrence or making management decisions. Regardless of the study extent used, most of the models developed in the present study performed poorly at the edge of the range as assessed by independent-AUC. Of the twelve models for which independent data were available, only three passed independent validation and even these had AUC values that were on the low-end for considering a model useful (*i.e.,* AUC > 0.7 but < 0.8; Swets, 1988). Species may be faced with unique challenges at range limits including novel abiotic environments (Lee-Yaw *et. al.,* 2016), an intensification of biotic interactions (Angert *et. al.,* 2013; Paquette and Hargreaves, 2021), constraints on dispersal (Graham *et. al.,* 2010, Kissling *et. al.,* 2012), and Allee effects (Keitt *et. al.,* 2001; Merker and Chandler *et. al.,* 2021). These factors may limit the extent to which climate governs site-level presence-absence at the edge of the range and thus the ability of SDMs based on climate variables to predict occurrence at range limits. Whether there is a general breakdown of the predictive abilities of SDMs towards range limits in these and other species awaits studies comparing model accuracy across different positions in the range.

At the same time, my results and those of others who have specifically tested the accuracy of SDMs at range limits (Luoto *et. al.,* 2005; Trumbo *et. al.,* 2011; McCune, 2016; McCune *et. al.,* 2020), suggest that it is possible to develop reasonably accurate models for use at the edge of the range for some species. However, in each of these cases, appropriate models were identified through careful exploration of different model inputs and settings, emphasizing the need for testing of different modeling decisions, including choice of study extent.

Although my results underscore the value of testing the impacts of study extent on model predictions and performance prior to using models, it is notable that of the three models that passed independent validation, two were developed using the smallest study extent (the political extent). Trumbo *et. al.,* (2011) also found models developed using smaller extents (state boundaries) outperformed models using range-wide extents for four amphibian species at the edge of their range in Missouri. Thus, when the goal is to predict occurrence at the edge of the range, local study extents may be more appropriate than larger study extents, even when such extents are not biologically informed. However, as the poor-performing models for long-toed salamanders and Columbia spotted frogs indicate, using a local extent does not guarantee that models will be accurate. Thus, rather than focusing on universal recommendations for choosing a study extent, I emphasize the need to both test the impacts of study extent on model predictions and to evaluate models with independent data.

2.5.3. Limitations and future directions

In addition to study extent, other modeling decisions, including different Maxent settings (Elith and Burgman, 2002; Muscarella *et. al.,* 2014) or the choice of modeling algorithm (Konowalik and Nosol, 2021) can have strong impacts on model predictions and performance (Anderson and Gonzalez, 2011; Barbet-Massin *et. al.,* 2012; Walker 2018; Hallgren *et. al.,* 2019; Barber *et. al.,* 2022). Likewise, the set of environmental variables considered, and the spatial resolution of variables can influence model predictions and performance (Bellamy *et. al.,* 2013; Heikkenin *et. al.,* 2017; Connor *et. al.,* 2019), especially at the edge of the range where it may be important to specifically include variables that account for the effects of biotic interactions and dispersal limitation (Wisz *et. al.,* 2013; Pollock *et. al.,* 2014; Gamliel *et. al.,* 2020; König *et. al.,* 2021). As my focus was specifically on the impacts of study extent, I did not explore the impacts of these other modeling decisions. However, I acknowledge the potential for these decisions to have impacted my findings. For example, Walker (2018) and Connor *et. al.,* (2019) found that the effect of the resolution of environmental variables on model performance interacted with the size of study extent, and that smaller study extents performed especially well when high resolution environmental layers were also used. Thus, although my results clearly demonstrate the potential for study extent to influence both model predictions and performance, other modeling decisions have the potential to interact with these impacts and more studies that test the impacts of multiple modeling decisions simultaneously are needed.

I also note that my ability to fully evaluate model performance was limited by the availability of comprehensive amphibian survey data in Alberta. I lacked independent survey data for the two species in EINP and the survey data for WLNP came from a limited number (40) of sites. Although, the number of survey sites in the latter case is in line with survey sizes used by other studies to evaluate SDMs (*e.g.,* Sacks *et. al.,* 2017; Safaei *et. al.,* 2018; Radomski *et. al.,* 2022), model accuracy is ideally evaluated using larger surveys (n ≈ 100) that are specifically designed to span the full range of predicted suitability values from SDMs (Vaughan and Omerod, 2005; McCune, 2016, McCune *et. al.,* 2020; Rosner-Katz *et. al.,* 2020). In addition to fully assessing the discrimination performance of models, such datasets also provide an opportunity to evaluate model calibration performance–an important test for downstream applications that require ranking site suitability (Gogol-Prokurat, 2011; Merow *et. al.,* 2013; McCune, 2016; Lee-Yaw *et. al.,* 2022). Thus, my estimates of overall model performance may have improved (or worsened) had models been assessed using a larger, targeted dataset. Ironically, SDMs are often touted as a readily accessible tool to understand species distributions that bypass the need for expensive and time-intensive surveys (*e.g.,* Guisan and Thuiller, 2005). However, the need for large, well-stratified independent datasets to validate SDMs before their use in downstream applications (*e.g.,* Gogol-Prokurat, 2011; McCune, 2016; Araújo *et. al.,* 2019; Rosner-Katz *et. al.,* 2020; Lee-Yaw *et. al.,* 2022) emphasizes the continued need for robust field surveys. Advances in passive species detection technologies including auditory recording units, camera traps, and eDNA are expected to make such surveys increasingly feasible and thus may lead to increased uptake in independent validation of SDMs (*e.g.,* Hagens *et. al.,* 2018; Domahidi *et. al.,* 2019; Burns *et. al.,* 2020; Riaz *et. al.,* 2020).

2.5.4. Conclusions

Species distribution models are a widely used tool for answering a variety of questions about the distribution of species. In Canada and other high-latitude countries, using SDMs to address both fundamental and applied questions often requires developing models for populations at the edge of the range. However, my results suggest that care is needed in this context as model decisions, including choice of study extent, can impact model predictions and performance. Current advice in the literature is to choose a study extent based on the conditions one wishes to contrast with presences (Elith *et. al.,* 2011; Saupe *et. al.,* 2012; Merow *et. al.,* 2013), ideally with knowledge of novel environments in the area of interest (Elith *et. al.,* 2011). However, in line with "best practices" for other model settings and decisions (Araújo *et. al.,* 2019), an alternative approach is to explore the impacts of study extent on model predictions and any downstream conservation or management decisions. Indeed, as discussed above, such sensitivity analyses may be useful for exploring variables that are broadly versus locally important for predicting the presence of species and may thus allow investigators to understand the drivers of species distributions more fully. However, as my results indicate, even when thorough sensitivity analyses are conducted, we may still fail to produce accurate models for some species. Thus, I caution against blind acceptance of SDMs in peripheral parts of species' ranges and emphasize the need for independent testing of models and consideration of alternative methods for identifying suitable habitat when these models fail.

CHAPTER 3: USING SPECIES DISTRIBUTION MODELS TO INFORM CONSERVATION TRANSLOCATIONS OF LONG-TOED SALAMANDERS

3.1. Abstract

Conservation translocations are an important tool for combating species declines and population losses. However, before organisms can be translocated, suitable release sites must be identified. Species distribution models (SDMs) can facilitate this step. Yet, these models can be highly sensitive to several modeling decisions. In this study, I developed SDMs to inform long-toed salamander (*Ambystoma macrodactylum*) translocations in southwestern Alberta and used independent data to test the sensitivity of model predictions and performance to three key modeling decisions: 1) type of environmental variables used, 2) choice of study extent, and 3) sample bias correction. I found that model predictions and performance varied substantially depending on the different decisions made during model generation. Models developed using local study extents were more accurate than those based on range-wide extents and whether sampling bias was accounted for impacted model performance more than the type of environmental variables included in the models. I also found that these modeling decisions impacted the prioritization of potential release sites, and I demonstrated an approach for incorporating future climatic predictions into release site selection. This study demonstrates the utility of SDMs in conservation translocation planning and underscores the need for practitioners to think carefully about the impacts of input decisions on model predictions and performance when developing SDMs for this purpose.

3.2. Introduction

Climate change, exotic species, and habitat loss have led to declines in biodiversity, with extinctions currently occurring 1000 times above background rates (Butchart *et. al.,* 2010; Pimm *et. al.,* 2014). Species recovery strategies increasingly call for conservation translocations (Swan *et. al.,* 2018; Berger-Tal *et. al.,* 2020), defined as "the intentional movement and release of a living organism where the primary objective is a conservation benefit" (IUCN/SSC, 2013). These efforts include reintroductions, relocations, and assisted migration, and are likely to be increasingly necessary to ensure the persistence of many species (Berger-Tal *et. al.,* 2020), including in protected areas (Parks *et. al.,* 2022). Thus, there is a growing need for the development and evaluation of tools that can aid in the planning and execution of conservation translocations.

The first step in planning a conservation translocation is the identification of suitable release sites (IUCN/SSC, 2013; McCoy *et. al,* 2014; Soorae, 2018). Species distribution models (SDMs) can be used to predict and map suitable habitat (Peterson, 2011) and are a freely available tool that can help in this regard (Hunter-Ayad *et. al.,* 2020; Eyre *et. al.,* 2022). SDMs have been used in conservation translocations to identify suitable release sites for a number of species across the globe, including bighorn sheep (*Ovis canadensis sierrae*) in the Sierra Nevada Mountains (Johnson *et. al,* 2007), fish species in Great Smoky Mountains National Park (Malone *et. al.,* 2018;), and the Chequered Skipper butterfly (*Carterocephalus palaemon*) in the U.K. (Maes *et. al,* 2019). Furthermore, the use of SDMs to prioritize release sites may lead to improved translocation success. For instance, Bellis *et. al,* (2020) developed SDMs for previously translocated amphibians, reptiles, and terrestrial insects to evaluate whether SDM-derived suitability could predict translocation outcomes. They found that not only was SDM-based habitat suitability positively associated with the probability of translocation success, but that relative to other correlates (*e.g.,* life-stage and number of individuals released), the predicted suitability of release sites explained 48% of the variation in translocation outcome (Bellis *et. al,* 2020). Thus, SDMs represent a promising tool for evaluating and prioritizing release sites for conservation translocations.

At the same time, SDMs must be developed with care as these models can be sensitive to decisions made during model development. For example, previous studies have observed that SDM predictions and performance can be sensitive to the environmental variables included in the models (Merow *et. al,* 2013; Araújo *et. al,* 2019). Likewise, whether sampling bias in the input locality data is accounted for can have a substantial impact on model predictions (Phillips *et. al.,* 2009; Merow *et. al.,* 2013; Searcy and Shaffer, 2014; Araújo *et. al.,* 2019, Barber *et. al.,* 2022). Finally, the choice of study extent, which defines the geographic area from which presence and either background or pseudo-absence data are considered and thus the range of environmental conditions considered in the model can impact model predictions (Chapter 2; VanDerWal *et. al.,* 2009; Anderson and Raza, 2010, Raes, 2012) and performance (Chapter 2; Searcy and Shaffer, 2014; Schmidt *et. al.,* 2020; Connor *et. al.,* 2019). Although several 'best-practices' for making these and other modeling decisions have been proposed (Elith *et. al.,* 2011; Merow *et. al.,* 2013; Araújo *et. al.,* 2019), ideally, the impacts of different modeling decisions are explored through rigorous sensitivity tests (Araújo *et. al.,* 2019; Sofaer *et. al.,* 2019). This step is of particular importance when SDMs are used to inform conservation translocations (Gogol-Prokurat, 2011; Westwood *et. al.,* 2020) as poor choice of release sites based on inaccurate SDMs may lead to translocation failure and wasted resources (Fischer and Lindenmayer, 2000; Berger-Tal *et. al.,* 2020).

Amphibians are globally the most threatened vertebrate group, with ~40% of species experiencing some level of decline (Stuart *et. al.,* 2004; IUCN/SSC, 2013; Alroy, 2015). Conservation translocations have been effective for slowing declines and restoring populations of some amphibian species (Denton *et. al.,* 1997; Sarrazin and Legendre; 2000; Wilson *et. al.,* 2008) and are expected to be used with increased frequency in the coming years (Swan *et. al.,* 2018; Novak *et. al.,* 2021; Scheele *et. al.,* 2021). However, conservation translocations of amphibians fail an estimated 50% of the time (Griffiths and Pavajeau, 2008; Harding *et. al.,* 2016), often because of poor habitat quality at release sites (Germano and Bishop, 2009). Thus, there is a need to improve release site selection for amphibian translocations and SDMs may aid in this process (Bellis *et. al.,* 2020).

Long-toed salamanders (*Ambystoma macrodactylum*) are a small pond-breeding amphibian found throughout western North America. Though common across much of their range, long-toed salamanders are listed as a *Species of Special Concern* in Alberta, Canada (Graham and Powell, 1999). One of the primary threats to long-toed salamanders is the stocking of fish in historically fishless high-elevation waterbodies for recreational purposes (Funk and Dunlap, 1999; Pearson and Goater, 2008). This has resulted in the extirpation of salamanders from many high-elevation parts of their range, including in Alberta (Pearson and Goater, 2008; Kenison *et. al.,* 2016). Given that salamanders are often the top predator in these environments, their loss from these ecosystems is likely to have had substantial impacts on these freshwater communities (Davic and Welsh, 2004; Pilliod *et. al.,* 2010; Kenison *et. al.,* 2016). More generally, increased awareness around the impacts of fish stocking on aquatic systems has motivated efforts to restore high-elevation waterbodies to their natural state, particularly within protected areas in Alberta (Parker *et. al.,* 2001; Banting *et. al.,* 2021). As part of these efforts, Parks Canada is considering the translocation of long-toed salamanders to sites that were historically fish-free in southwestern Alberta and where stocked fish have been or could be removed (Parks Canada, pers. comm.). However, limited resources and the controversial nature of such activities in the province (*e.g.,* Parker *et. al.,* 2001) make it imperative that any translocation approaches are carefully planned to maximize chances of success.

This study aimed to develop SDMs for prioritizing potential translocation release sites for long-toed salamanders in southwestern Alberta. My goals were to 1) test the impacts of key modeling decisions on SDM predictions for long-toed salamanders in southwestern Alberta; 2) evaluate the ability of SDMs to predict independent presence-absence data (*i.e.,* model performance/accuracy); and 3) use well-performing models to rank potential release sites in protected areas in the province under both current and future climatic conditions. In addition to informing this specific translocation program, my study demonstrates a rigorous approach to model development and validation that can serve as a guide for the use of SDMs in conservation planning more generally.

3.3. Methods

3.3.1. General approach to developing SDMs

Presence-only SDMs rely on input locality records to capture environmental conditions that support a species of interest (Peterson, 2011). I compiled long-toed salamander locality data from several sources including the Global Biodiversity

Information Facility (GBIF: https://www.gbif.org/, DOI10.15468/dl.k8yj8x), other researchers (see Lee-Yaw and Irwin, 2015), and conservation databases for all states and provinces in which the species occurs (Alberta, B.C., Washington, Oregon, California, Idaho, and Montana). Records were filtered to match the spatiotemporal resolution of the environmental data (see below) by excluding records collected before 1990 and with coordinate uncertainty > 1 km. Duplicate observations from the same site were removed and the dataset was thinned such that only one locality record per raster cell of the environmental layers was retained. Depending on the study extent used (see below), the total number of long-toed salamander input records used to develop models ranged from 136 to 5329 (Table A3.1).

SDMs were developed using Maxent (version 3.2.3; Philips *et. al.,* 2006). Maxent has been widely used for developing SDMs (Chapter 1), and often outperforms other presence-only modeling algorithms (Phillips and Dudik, 2008; Aguirre-Gutierrez *et. al.,* 2013; Elith *et. al.,* 2010; Bradie and Leung, 2017; but see Chapter 1). I developed a series of models to explore the impacts of three key decisions (type of environmental variables used, choice of study extent, and sample bias correction; see below) on SDM predictions and performance. Rather than holding other aspects of model parameterization constant across models and thus possibly restricting model performance (*e.g.,* Anderson and Gonzales, 2011; Barber *et. al.,* 2022), I tuned both the regularization parameter (possible values: 0.25, 0.5, 1, 1.5, 2, and 4) and set of features included (possible feature combinations: L, LQ, H, LQH, LQP, LQT, LQHP, LQPT, and LQHPT; where L = linear, Q = quadratic, H = hinge, P = product and T = threshold) to maximize model performance within the confines of the input decisions being considered. Tuning was conducted using R (R Core Team Version 4.2.1.) in the *ENMeval* package (Kass *et. al.,* 2021) with the optimal regularization parameter and set of features selected using the Akaike information criterion correction (AICc, Muscarella *et. al.,* 2014; Table A3.1). All other data processing and modeling procedures were conducted using R: *sf* (Pebesma, 2018) and *dismo* (Hijmans *et al.,* 2021).

3.3.2. Sensitivity tests

3.3.2.1. Type of environmental variables included.

I compared predictions of SDMs based solely on climatic variables (climate-only models) to those that included additional non-climatic variables (climate+ models). Thirteen biologically relevant climatic variables were initially considered for inclusion in my models based on the published literature and the natural history of the salamanders (Table A3.2.). The climate normals for each variable for the 1990-2020 period were downloaded from Climate NA (Wang *et. al.,* 2016; https://adaptwest.databasin.org/pages/adaptwest-climatena/) at a spatial resolution of 30 arc seconds (~1 km) and were reprojected to North America Albers Equal Area Conic. Non-climatic variables considered in the climate+ models included net primary productivity (downloaded at a resolution of 500 x 500 m from MOD17A3HGF/ Terra Net Primary Production; Running and Zhao, 2019), land cover (downloaded at a resolution of 30 x 30 m from 2015 Land Cover of North America at 30 Meters; CEC, 2020), and topographic wetness (derived from a Digital Elevation Model DEM; Available from: https://www.usgs.gov/centers/eros). Non-climatic layers were resampled or derived and reprojected to match the resolution and projection of the climatic data using *Spatial*

Analyst in ArcGIS Pro v.3.0.1. To reduce collinearity among predictors, I used pairwise Pearson's correlation tests to identify and remove highly correlated predictors among the continuous variables (*i.e., r* > 0.70, or < -0.70; Table A3.3). For pairs of variables that were highly correlated, the variable correlated with the fewest other variables was retained. A final set of nine variables were retained for modeling, comprised of six climatic variables (MAP = mean annual precipitation, NFFD = number of frost-free days, PPT_sm = summer precipitation (June-Aug), RH = relative humidity, SHM = summer heat moisture index, and TD = difference between mean warmest and coldest month temperature), and three non-climatic variables (landcover, NPP = net primary productivity, and wetness). Climate-only models were based on the six climatic variables, whereas climate+ models were based on all nine variables (Table A3.2).

3.3.2.2. Sample bias correction.

Presence-only SDMs compare environmental conditions underlying the presences of species to those generally available across the study extent to identify suitable conditions for the focal species (Merow *et. al.,* 2013). To characterize the latter, the default Maxent approach is to use conditions at randomly sampled background points across the study extent (Phillips *et. al.,* 2009). However, because input localities are often collected in a biased fashion (increased sampling effort near roads, trails, towns; Phillips *et. al.,* 2009; Barber *et. al.,* 2022) the use of random background points can result in models that are overfit to the input localities. To address this issue, several bias correction strategies have been proposed (reviewed by Barber *et. al.,* 2022), of which the use of target-group background points (*i.e.,* locality records from other species expected to share the same survey biases as the focal species; Phillips *et. al.,* 2009) has been shown to outperform others (Barber *et. al.,* 2022; Chapter 1).

I developed models using both random background and target-group background. Random background models were based on the Maxent default of 10,000 random background points except where the study extent included less than 10,000 grid cells, in which case 5,000 random points were used (Table A3.1). Target-group background points included locations were any pond-breeding amphibian other than long-toed salamanders had been observed based on records in the Global Biodiversity Information Facility (GBIF: https://www.gbif.org/ on October 21, 2021; Table A3.4). Target-group background points were filtered in the same way as the long-toed salamander locality records (see above) resulting in 215 to 13,868 target-group background points depending on study extent (Table A3.1).

3.3.2.3. Choice of study extent

I tested the impacts of four different study extents on model predictions and performance: a) a "range-wide" extent defined by the full geographic range of the long-toed salamander (IUCN: http://www.iucnredlist.org), b) an "ecoregion" extent based on the boundaries of the North Central Rockies Forests ecoregion (Olson *et. al.,* 2001 https://www.worldwildlife.org/publications/terrestrial-ecoregions-of-the-world), c) a "genetic subspecies" extent based on the genetic boundaries described by Lee-Yaw and Irwin, (2015) for the eastern long-toed salamander (*A. m. krausei*), and d) a "political" extent defined by the Alberta Environment and Parks Bow-Crow land region south of the Bow River (Available from: https://open.alberta.ca/opendata). These study extents differed not only in size, but in the degree to which they are biologically informed (*e.g.,* range boundaries versus political boundaries) and in whether they might be expected to account for potential impacts of local adaptation within the range (*e.g.,* ecoregion and genetic extents versus range-wide extent). Because I was only interested in the conditions that shape presence-absence within the range (as opposed to conditions that set the overall range limits of the species), all study extents were clipped to the species' geographic range. Final polygons representing the different study extents were produced using the *Spatial Analyst* toolbox in ArcGIS Pro v.3.0.1.

3.3.3. Model predictions

Altogether, I developed a total of 16 SDMs, representing all combinations of the two environmental variable sets (climate-only and climate+), two types of background (random and target-group), and four study extents (range, ecoregion, genetic, and political). To explore the impact of these model decisions on model predictions, I projected each model across the smallest region common to all study extents (*i.e.,* the political study extent) and visually compared the continuous prediction surfaces based on Maxent's logistic output. The logistic output is independent of scale and is the most appropriate output type for comparing models developed using different study extents (Merow *et. al.,* 2013). Using the logistic output requires specifying a scaling parameter, prevalence (Phillips and Dudik, 2008). This parameter is ideally based on independent data (Royle *et. al.,* 2012; Merow *et. al.,* 2013) and in this study, I set prevalence to 0.58 based on an independent estimate of long-toed salamander prevalence across 13 years of monitoring and 41 sites in Waterton Lakes National Park located at the southern end of the area of interest (Hunter, 2022).

I additionally converted all continuous model predictions into binary surfaces of suitable habitat for model comparison. For each model, I used the 10% omission rate of the training presences as the threshold for considering a cell to be suitable as this approach has worked well in other studies focused on range-edge populations (Vale *et. al.*, 2014; McCune 2016). To visually assess model agreement, I created a stacked surface from the binary predictions of all models. To formally assess spatial agreement across binary model predictions, I calculated pairwise average similarity and Fuzzy Kappa between surfaces using the Map Comparison Kit V.3.2.3 (Visser and de Nijs, 2006; Hagen-Zanker, 2009). Average similarity represents the global proportion of cells with the same value (0 or 1) in two given surfaces, expressed as a percentage (Visser and de Nijs, 2006). Fuzzy Kappa ranges from 1 to -1, with values of 1 indicating total similarity, values of zero indicating total dissimilarity, and negative values indicating both dissimilarity and differences in spatial autocorrelation between map surfaces (Hagen-Zanker, 2009).

3.3.4. Model performance

Before models can be used in conservation decision making, model accuracy should be assessed using independent data (Lee-Yaw *et. al.,* 2022; Johnson *et. al.,* 2023). Thus, in addition to internal five-fold cross-validation (Phillips *et. al.,* 2009), I conducted independent presence-absence surveys throughout southwestern Alberta for purposes of evaluating model accuracy. A set of 240 lentic water bodies that did not coincide with existing records of long-toed salamanders or fish (based on provincial records; Alberta Fisheries and Wildlife Management Information System; 2022) were considered for this survey and represented almost all putatively fish-free waterbodies in this region. These sites were identified using Alberta Biodiversity Monitoring Institute's wetland inventory, the Alberta's Provincial Base Waterbody Polygon Dataset (Available from: open.alberta.ca/opendata/), topographic maps, and Google Earth Pro v.7.3.6.9345. From this set of 240, 83 sites were selected for presence-absence surveys that could be feasibly surveyed in a single season while maximizing geographic coverage of the study area. Although it was not possible to design a single random survey stratified according to the predictions of all models (*e.g.,* Rosner-Katz *et. al.,* 2020), the selected sites captured the range of values of each of the environmental variables included in the models and are thus representative of a range environmental conditions in the study area (Fig. A3.1).

Field surveys took place between May and August 2022. Standard visual encounter and dip net surveys were conducted using established protocols (See Appendix II for survey details). Briefly, two surveyors circumnavigated each waterbody in opposite directions, visually scanning the water surface, pond bottom, and submerged debris for the presence of long-toed salamanders (mainly observed as eggs and larvae at this time of year) for a maximum of two people-hours. Surveyors spent an additional two people-hours dip-netting the entire perimeter of the pond at those sites where salamanders were not detected visually. Although single-visit surveys precluded calculation of detectability and accounting for potential false absences (*e.g.,* Royle and Link, 2006), waterbodies in this part of the species' range typically have high water clarity, and previous estimates of the probability of detection based on a single visit at other sites in the Rocky Mountains are relatively high for this species (P= 0.74; Pilliod *et. al.,* 2010). Data from nine sites were excluded from the final dataset due to flowing water (salamanders are generally found in standing waterbodies) or the presence of fish. Thus, data from 74 survey sites (40 presences and 34 absences) were used to evaluate SDMs.

Discriminatory ability (the capacity to predict presences and absences) of the 16 SDMs was tested using the threshold-independent metric AUC (area under the receiver operator curve, Phillips *et. al.,* 2009), where models with AUC > 0.7 were considered to have acceptable discriminatory ability (Swets, 1988). As a secondary assessment of discriminatory ability, I used the threshold-dependent metric TSS (true skill statistic, Allouche *et. al.,* 2006; using thresholds for binary surfaces described above). Threshold-independent and dependent metrics were calculated using the *PresenceAbsence* package in R (Freeman and Moisen, 2008). In addition to these measures of model discrimination, I also tested whether average predicted suitability was significantly higher at observed presences than absences using a Kruskal-Wallis test (Kruskal and Wallis, 1952). Finally, as my main goal was to use SDM-derived suitability to rank potential translocation sites, I used binomial logistic regression to test the calibration performance of each model (*i.e.,* whether predicted suitability scaled with the probability of presence; Gogol-Prokurat 2011; Lee-Yaw *et. al.,* 2022). Models were considered to have adequate calibration performance if the relationship between predicted habitat suitability and the probability of presence was positive, linear, and statistically significant ($p < 0.05$; after using Bonferroni correction to account for multiple testing, Gogol-Prokurat, 2011). Only models with both adequate discrimination and calibration performance were used to rank potential translocation release sites.

3.3.5. Ranking of potential translocation release sites

To evaluate potential release sites for long-toed salamanders in protected areas of southwestern Alberta, I first identified all (n = 178) lentic water bodies within Waterton Lakes National Park (WLNP), and Castle Provincial, and Castle Wildland Provincial Parks (CCWPP) using Alberta Biodiversity Monitoring Institute's wetland inventory, the Alberta's Provincial Base Waterbody Polygon Dataset (Available from: open.alberta.ca/opendata/), topographic maps, and Google Earth Pro v.7.3.6.9345. Because long-toed salamander populations have been disproportionately lost from high-elevation sites (Pearson and Goater, 2008), I only considered waterbodies above 1800 m in elevation as potential release sites. I also excluded sites known to support existing populations of long-toed salamanders, resulting in a total of 36 potential release sites (12 in WLNP and 24 in CCWPP, Table 3.3), inclusive of six sites already under consideration for the release of individuals (Parks Canada pers. comm.).

Predicted suitability of these 36 potential release sites was extracted from models with high calibration performance (see above). Sites were first ranked from 1 to 36 based on the continuous predicted suitability scores produced by Maxent's logistic output for each model. I then calculated the average of the ranking for each site across models and used these values to assign a final overall rank for each site. I also estimated the future suitability of potential release sites based on moderate and extreme global CO2 emissions scenarios (SSP: Shared Socioeconomic Pathway 2-4.5 and, SSP4-8.5 respectively; Ebi *et. al.,* 2014). I used the Atmosphere-Ocean General Circulation Models (AOGCMs) assembled by the Coupled Model Intercomparison Project (CMIP6) for a 20-year period ~50 years in the future (2060-2080; available from Climate NA; Wang *et. al.,* 2016;

Mahony *et. al.,* 2022) to generate future surfaces based on these scenarios for each of the climate variables included in my models. Because future predictions for landcover, topographic wetness, and net primary productivity are not readily available, I relied on climate-only models (that passed independent model evaluation above) to generate future suitability surfaces for the focal region. I used multivariate environmental similarity surfaces (MESS; Elith *et. al.,* 2010) to identify cells within the focal area of interest where the future climatic conditions are predicted to be outside of the range of conditions that were used to calibrate each model (*i.e.,* "non-analog" conditions) and where model extrapolation would be needed to make predictions about the suitability of sites. For each model, I then determined whether potential translocation release sites were predicted to have suitable, unsuitable, or non-analog climatic conditions in the future. The same threshold values used to generate present-day binary surfaces were used to classify sites as either suitable or unsuitable for each model.

Three climate-only models passed independent model evaluation (see results) and thus could be used to forecast suitable habitat using projections of future climatic conditions in this area. The best-performing of these three models (political model; Model 2; see results) was not considered further as the MESS revealed that all potential translocation release sites were predicted to have non-analog conditions under both global change scenarios (Fig. A3.3). Sites were ultimately considered suitable (or unsuitable) for a given emissions scenario (SSP2-4.5 or SSP4-8.5) if both remaining SDMs predicted the site to be suitable (or unsuitable) into the future. Otherwise, the future suitability of a site was considered "unclear" if one model predicted the site to be suitable and the other classed it as unsuitable or non-analog.

3.4. Results

3.4.1. Model predictions

Most models predicted suitable habitat for long-toed salamanders along the western border of Alberta, with multiple models showing high concentrations of suitable habitat at low elevations in WLNP and CCWPP and to the immediate north of these protected areas (Fig. 3.1). Nevertheless, the choice of input variables, background type, and study extent all had substantial impacts on model predictions, as well as on variable importance (Table A3.5). Visual inspection of the continuous prediction surfaces revealed high variability in the predicted distribution of suitable habitat across models (Fig. 3.1). Models developed with the political study extent tended to predict a narrow band of suitable habitat, concentrated along the Alberta-BC border, whereas predictions based on larger study extents were more even across the prediction area (Fig. 3.1). The percentage of area predicted to be suitable for long-toed salamanders ranged from 29% to 94% across models (Fig. 3.1) using the 10% omission rate of input presences as the threshold for defining cells as suitable/unsuitable. Formal map comparisons of the binary prediction surfaces yielded average similarity scores ranging from 0.367 to 0.991, and fuzzy kappa values between -0.384 and 0.915 (Table 3.1), again indicating moderate to large differences in the amount and distribution of suitable habitat across models.

The choice of background data (*i.e.,* whether sampling bias was accounted for) and study extent impacted model predictions more than the set of environmental variables used. For example, for pairs of models that used the same background type (random points or target-group) and study extent but different environmental variables, the amount of area predicted to be suitable differed by an average of only 5%, whereas for pairs of models that used different background types or study extents but the same environmental variables, the amount of area predicted to be suitable differed by an average of 25% (Fig. 3.1.). Likewise, formal map comparisons revealed greater similarity between pairs of models developed using different environmental variables (average similarity = 0.424, Fuzzy kappa = 0.776) than pairs of models using different background types (average similarity = 0.045, Fuzzy Kappa = 0.594) or study extents (average similarity = 0.340, Fuzzy Kappa = 0.741; Table 3.1). With respect to the choice of study extent, the observed variability in model predictions was largely driven by differences between models based on the political extent and those based on other, larger study extents. For example, the amount of area predicted to be suitable by the four models using random background data and climate-only environmental variables was similar for the range, ecoregion, and genetic models (55%, 60%, and 63% respectively), but much lower for the political model (29%; similar trends are observed down the other columns of maps in Fig. 3.1). Likewise, formal map comparisons revealed that when the environmental variable set and type of background were held constant across models, there were greater similarities between range, ecoregion, and genetic models (average similarity = 0.461, Fuzzy kappa = 0.843) than between political models and those of the three larger study extents (average similarity = 0.218, Fuzzy Kappa = 0.640; Table 3.1)

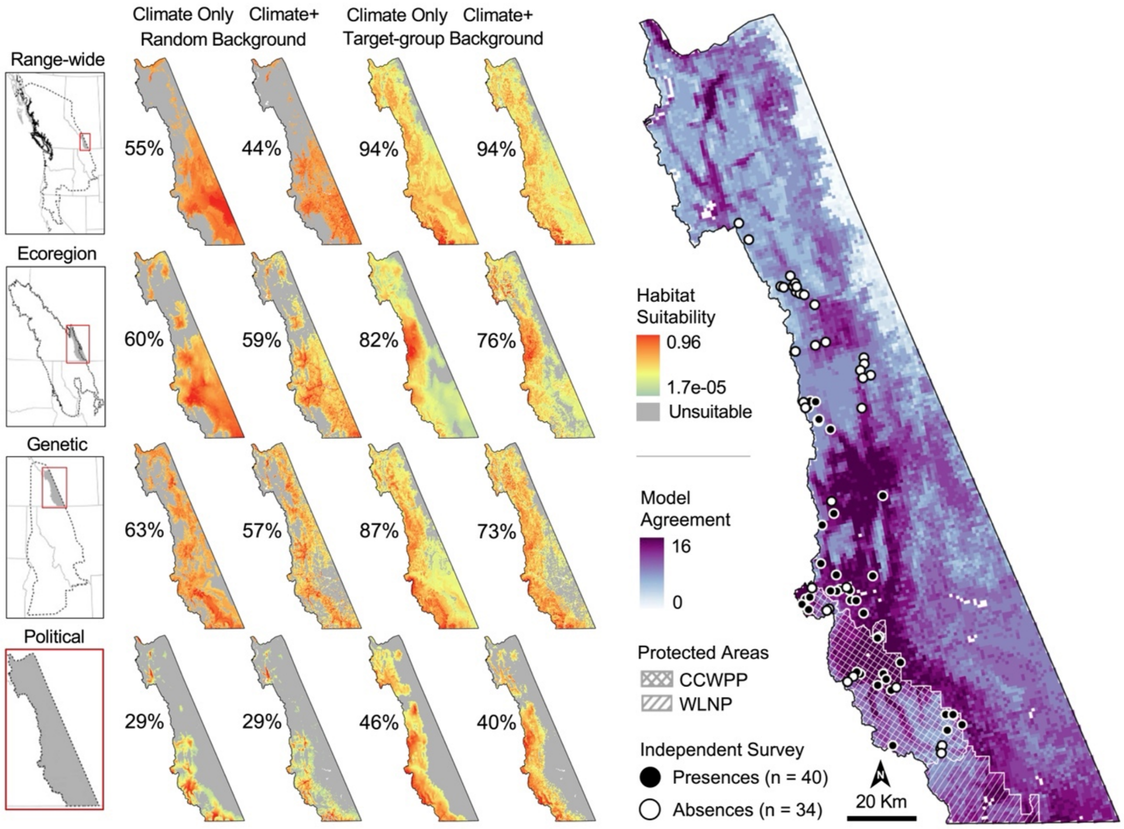


Figure 3.1. Predicted suitability of southwestern Alberta for the long-toed salamander based on Maxent models developed using different sets of environmental variables, types of background data, and choices of study extent. Continuous predictions of long-toed salamander habitat suitability using Maxent's logistic output are shown on the left with darker and warmer colours (red) representing highly suitable habitat. Grey shading represents unsuitable habitat defined by the predicted suitability value below which 10% of input presences occur. Percentages indicate the area of the study region predicted to be suitable (not greyed out). The stacked agreement surface of binary suitable/unsuitable versions of the sixteen suitability maps is shown on the right. Dark purple represents cells that all models predicted to be suitable, white represents cells that were never predicted to be suitable. Points show presences (filled circles) and absences (open circles) from an independent field survey conducted in Summer 2022 and used to validate models. Waterton Lakes National Park (WLNP) and Castle Provincial, and Castle Wildland Provincial Parks (CCWPP), protected areas where translocations are planned or possible, are indicated with hashed and cross-hatched shading respectively. The political extent is truncated to the east by the edge of the long-toed salamander's range as defined by the International Union for the Conservation of Nature (IUCN) range polygon.

Table 3.1. Pairwise fuzzy kappa (top right) and average similarity (bottom left) between binary maps of predicted suitability for long-toed salamanders in southwestern Alberta based on Maxent models developed using different types of background data, environmental variable sets, and study extents. The 10% omission rate of input presences was used as a cut-off to define suitable habitat for each model. Pairwise comparisons were conducted using the Map Comparison Kit V.3.2.3 (Visser and de Nijs, 2006; Hagen-Zanker, 2009). Increasingly warm colours indicate more positive numbers (more similarity between binary prediction surfaces) increasingly cool colours indicate more negative numbers (more differences between binary prediction surfaces).

			Random Background								Target-Group Background							
			Climate Only				Climate +				Climate Only				Climate +			
			Range	Ecoregion	Genetic	Political	Range	Ecoregion	Genetic	Political	Range	Ecoregion	Genetic	Political	Range	Ecoregion	Genetic	Political
Random Background	Climate Only	Range		0.782	0.424	0.13	0.808	0.757	0.232	0.136	0.12	-0.095	0.012	-0.384	0.111	-0.079	-0.359	-0.225
		Ecoregion	0.907		0.439	0.608	0.713	0.915	0.234	0.201	0.127	0.013	0.032	-0.216	0.136	0.044	-0.311	-0.068
		Genetic	0.782	0.794		0.642	0.318	0.456	0.583	0.207	0.065	0.033	0.108	-0.132	0.076	0.111	0.072	0.005
		Political	0.596	0.187	0.204		0.153	0.17	0.222	0.864	0.048	0.121	0.089	0.479	0.055	0.131	0.19	0.571
	Climate +	Range	0.918	0.877	0.744	0.665		0.728	0.199	0.214	0.094	0.038	0.053	-0.335	0.096	-0.001	-0.026	-0.2
		Ecoregion	0.902	0.966	0.821	0.621	0.892		0.298	0.668	0.137	0.056	0.077	-0.224	0.148	0.096	-0.299	-0.067
		Genetic	0.715	0.72	0.884	0.677	0.72	0.776		0.738	-0.031	-0.062	-0.02	-0.046	-0.025	0.049	0.215	0.114
		Political	0.618	0.634	0.679	0.955	0.713	0.214	0.278		0.04	0.113	0.095	0.438	0.049	0.128	0.184	0.56
Target-Group Background	Climate Only	Range	0.621	0.661	0.673	0.367	0.525	0.667	0.593	0.373		0.43	0.62	0.094	0.914	0.283	0.233	0.075
		Ecoregion	0.536	0.61	0.659	0.478	0.536	0.637	0.596	0.49	0.884		0.601	0.657	0.461	0.561	0.379	0.261
		Genetic	0.609	0.647	0.722	0.459	0.562	0.678	0.656	0.488	0.941	0.908		0.61	0.654	0.516	0.438	0.169
		Political	0.369	0.422	0.511	0.774	0.423	0.459	0.564	0.766	0.531	0.305	0.194		0.1	0.285	0.426	0.833
	Climate +	Range	0.626	0.671	0.687	0.382	0.538	0.681	0.611	0.395	0.991	0.891	0.947	0.539		0.319	0.254	0.092
		Ecoregion	0.591	0.658	0.743	0.547	0.588	0.704	0.722	0.582	0.827	0.876	0.886	0.674	0.84		0.442	0.684
		Genetic	0.489	0.532	0.736	0.592	0.493	0.58	0.778	0.625	0.798	0.817	0.861	0.743	0.811	0.859		0.735
		Political	0.457	0.518	0.593	0.834	0.525	0.548	0.663	0.841	0.474	0.614	0.576	0.927	0.494	0.304	0.399	

3.4.2. Model performance

In addition to evaluating the impacts of model inputs on model predictions, I asked whether models differed in terms of their accuracy, and thus ultimate utility for informing the selection of release sites. Internal-AUC based on cross-validation ranged from 0.65 to 0.89 and all but two models had internal-AUC > 0.7 (Fig. 3.2a). AUC based on independent presence-absence surveys (independent-AUC) was consistently lower than internal-AUC (Fig. 3.2a), ranging from 0.51 to 0.83 and was above 0.7 for only seven models (Table 3.2). Values for the threshold-dependent evaluation metric, TSS, ranged from -0.05 to 0.51 and models with high AUC scores also tended to have high discriminatory performance based on TSS (Table 3.2). Likewise, for the seven models with independent-AUC > 0.7, Kruskal-Wallis tests revealed that the mean predicted suitability of independent-presences was significantly higher than that of independent-absences (Table 3.2). Binomial logistic regression also revealed significant relationships between predicted suitability and probability of presence for these seven models (Table 3.2), and visual inspection of model fits indicated that these relationships were positive and roughly linear (Fig. A3.2). These seven models explained a relatively high proportion of the variance in probability of occurrence (percent null deviance explained ranged from 12 to 25 %; Table 3.2).

There were clear patterns with respect to which modeling inputs produced accurate models (Fig. 3.2). First, all but one of the seven models with high discriminatory and calibration performance were developed using random background points (no sampling bias correction; Table 3.2). Second, although it was possible to develop a well-performing model using the range-wide, ecoregion, or political study extents depending on other model inputs, models based on the political extent performed the best on average (Fig, 3.2d). In contrast, none of the models developed using the genetic extent passed independent evaluation (Table 3.2). The two models with the highest AUC and 'best' overall calibration performance (*i.e.,* based on percent null deviance explained and AIC; Table 3.2) were both developed using random group background data and the political study extent. Consistent with the results for model predictions, model performance was least impacted by which environmental variable set was used (Table 3.2; Fig. 3.2), with both climate-only and climate+ models represented roughly equally in the group of seven models that passed both discrimination and calibration performance (Table 3.2).

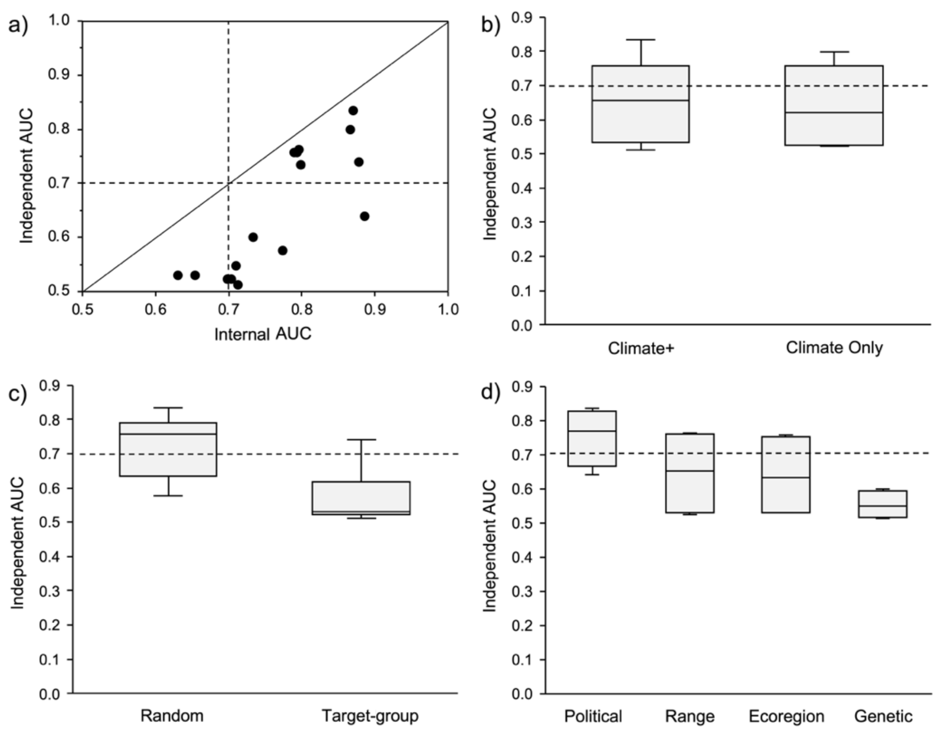


Figure 3.2. Summary of independent-AUC (area under the receiver operator curve) scores for sixteen long-toed salamander species distribution models developed in Maxent to inform conservation translocations in southwestern Alberta. Independent-AUC is based on presence-absence data from 74 sites and plotted in relation to a) internal-AUC (based on five-fold cross-validation), where each point represents a single model b) the set of environmental variables used, c) type of background points used, and d) the choice of study extent. Dashed lines at AUC = 0.7 represent the threshold for an 'acceptable' model (Swets, 1988).

Table 3.2. Model evaluation based on independent presence-absence data for Maxent models developed for long-toed salamanders in southwestern Alberta. Models are presented and numbered in descending order based on independent-AUC. Models above the solid line passed at least one evaluation test. Models above the dashed line are the two 'best' models based on observed breaks in discrimination (AUC > 0.8), and calibration performance (ΔAIC < 2).

	Model Decisions			**Discrimination Performance**					**Calibration Performance**				
						Kruskal-Wallis test			*Binomial Logistic Regression*				
	Environmental Variables	Background Point Type	Study Extent	AUC*	TSS†	Mean predicted suitability (±SD) of presences (n=40) and absences (n=36)	Raw *p*	Bonferroni Adjusted *p* ‡	Percent Null Deviance Explained (%)	Raw *pr(>|z|)*	Bonferroni Adjusted *pr(>|z|)* ‡	AIC	ΔAIC
1	Climate +	Random	Political	0.83	0.51	0.38 (±0.20), 0.14 (±0.17)	0.000001	**0.00001**	25	0.00004	**0.0003**	80.8	0.0
2	Climate Only	Random	Political	0.80	0.50	0.40 (±0.26), 0.13 (±0.15)	0.00001	**0.0002**	23	0.0001	**0.001**	82.6	1.8
3	Climate +	Random	Range	0.76	0.29	0.40 (±0.08), 0.32 (±0.09)	0.0001	**0.002**	14	0.001	**0.01**	91.4	10.6
4	Climate Only	Random	Ecoregion	0.76	0.40	0.38 (±0.11), 0.28 (±0.11)	0.0001	**0.002**	14	0.001	**0.007**	91.4	10.6
5	Climate Only	Random	Range	0.76	0.42	0.38 (±0.07), 0.31 (±0.09)	0.0002	**0.002**	14	0.002	**0.01**	92.0	11.2
6	Climate +	Target-group	Political	0.74	0.12	0.57 (±0.11), 0.47 (±0.10)	0.0004	**0.007**	15	0.006	**0.04**	90.8	10.0
7	Climate +	Random	Ecoregion	0.73	0.30	0.40 (±0.12), 0.29 (±0.14)	0.0006	**0.009**	12	0.002	**0.01**	94.3	13.5
8	Climate Only	Target-group	Political	0.64	0.29	0.55 (±0.12), 0.46 (±0.15)	0.008	0.12	8	0.0005	0.008	97.4	16.6
9	Climate Only	Random	Genetic	0.60	0.00	0.40 (±0.09), 0.36 (±0.10)	0.14	1.00	3	0.08	1.00	102.7	21.9
10	Climate +	Random	Genetic	0.58	0.14	0.55 (±0.12), 0.46 (±0.15)	0.26	1.00	2	0.12	1.00	103.6	22.8
11	Climate +	Target-group	Range	0.55	0.00	0.43 (±0.11), 0.38 (±0.16)	0.48	1.00	2	0.92	1.00	104.5	23.7
12	Climate Only	Target-group	Ecoregion	0.53	0.00	0.54 (±0.05), 0.54 (±0.04)	0.35	1.00	0	0.65	1.00	105.8	25.0
13	Climate +	Target-group	Ecoregion	0.53	0.15	0.53 (±0.06), 0.52 (±0.08)	0.66	1.00	0	0.58	1.00	105.9	25.1
14	Climate Only	Target-group	Genetic	0.52	-0.02	0.52 (±0.05), 0.53 (±0.05)	0.63	1.00	0	0.46	1.00	106.0	25.2
15	Climate Only	Target-group	Range	0.52	0.00	0.52 (±0.03), 0.52 (±0.04)	0.94	1.00	0	0.22	1.00	106.1	25.3
16	Climate +	Target-group	Genetic	0.51	-0.05	0.52 (±0.06), 0.53 (±0.06)	0.86	1.00	1	0.81	1.00	105.5	24.7

* Independent-AUC (area under the receiver operator curve) scores based on presence-absence data from 74 sites

† True Skill Statistic, threshold set by the 10% omission rate, which is the suitability value below which 10% of input presences occur

‡ Bold values remain significant after correcting for multiple testing

3.4.3. Ranking of potential translocation release sites

Results from independent model evaluation suggest that seven of the models presented here can be used not only to accurately predict the occurrence of long-toed salamanders (*i.e.,* based on discrimination performance), but also to rank potential release sites for translocations based on their predicted suitability (*i.e.,* based on calibration performance). I used these seven models to rank the 36 sites of interest in CCWPP and WLNP (Fig. 3.3; Table 3.3). Based on average site rank, the top ten release sites for potential translocations included six sites in CCWPP and four sites in WLNP (Fig. 3.3; Table 3.3). Two climate-only models (the ecoregion; Model 4 and range-wide models; Model 5) were used to assess the future suitability of sites. Five of the top ten release sites identified using current conditions (West Castle Pond, Lost Lake, Crypt Lake, Boundary Pond, and Alderson Lake) were predicted to be suitable under the moderate future emissions scenario (SSP2-4.5; Fig. 3.3, Table 3.3). However, only three of these sites (Crypt Lake, Boundary Pond, and Alderson Lake) were expected to remain suitable under the extreme future emissions scenario (SSP4-8.5; Fig. 3.3, Table 3.3). The future suitability of many potential release sites was unclear, as models either disagreed on the suitability of the site, or model extrapolation into novel conditions (*i.e.,* non-analog conditions) was required to make predictions based on one or both models (Fig. 3.3; Table 3.3).

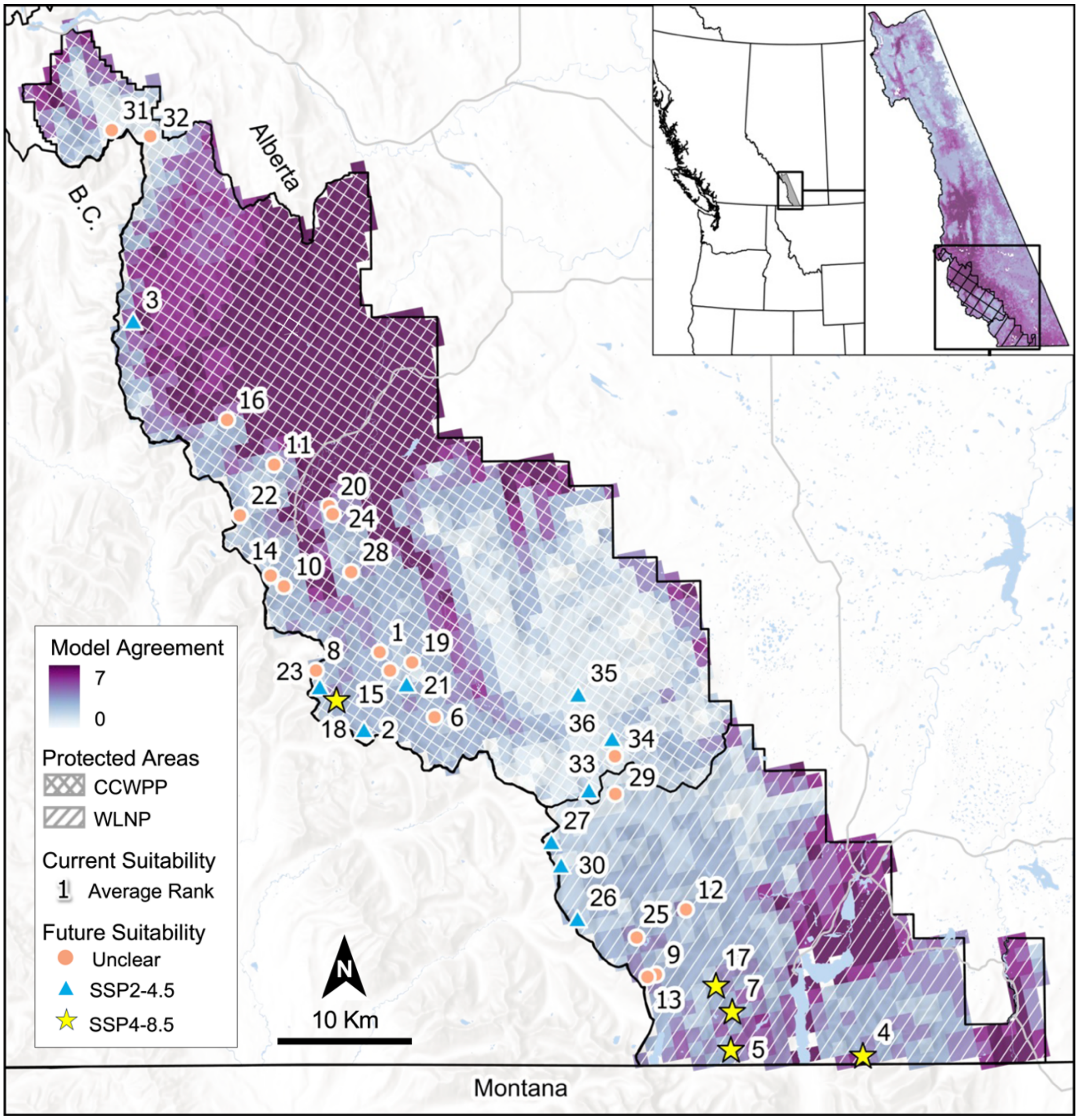


Figure 3.3. Potential release sites for long-toed salamander conservation translocations in Waterton Lakes National Park (WLNP) and Castle and Castle Wildlands Provincial Parks (CCWPP) in southwestern Alberta. Numbered labels correspond to the average rank of sites across seven models that passed tests of model discrimination and calibration under current climatic conditions. Dark purple represents cells that all seven models predicted to be suitable, white represents cells that were not predicted to be suitable by any model. Symbols represent the suitability of the site under future conditions, where circles represent sites for which future suitability was unclear (both models for the same emissions scenario disagreed on the site's suitability), triangles represent sites predicted to be suitable by both moderate emissions scenario (SSP2-4.5) models, and stars represent sites predicted to be suitable by both extreme emissions scenario (SSP4-8.5) models.

Table 3.3. The suitability of 36 potential release sites for long-toed salamander conservation translocations in in Waterton Lakes National Park (WLNP) and Castle and Castle Wildlands Provincial Park (CCWPP) in southwestern Alberta based on Maxent models with adequate discrimination and calibration performance. Models are numbered by AUC (area under the receiver operator curve) score, where Model 1 had the highest AUC score. Continuous suitability scores for each site under current conditions is shown for each model and were used to calculate average site rank across the seven models. These averages were then arranged 1-36 to determine average rank position Binary predictions as to whether sites are predicted to be suitable (1= yes, 0 = no) under moderate (SSP2-4.5) and extreme (SSP4-8.5) future emissions scenarios are also shown, where N-A stands non-analog and represents sites for which models were required to extrapolate into non-analog climatic space as determined by MESS maps (Elith *et. al.,* 2010). The top ten ranked sites are indicated by the solid horizontal line.

		Current Suitability Rank (predicted suitability) *								**Future Suitability †**			
										SSP2-4.5		SSP4-8.5	
Site Name	Protected Area	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Average Rank Position (±SD)	Model 4	Model 5	Model 4	Model 5
Grizzly Lake	CCWPP	4 (0.28)	3 (0.7)	3 (0.31)	6 (0.66)	15 (0.37)	3 (0.78)	7 (0.32)	1 (±4)	0	1	0	1
West Castle Pond	CCWPP	3 (0.23)	2 (0.72)	2 (0.26)	16 (0.73)	6 (0.37)	2 (0.80)	13 (0.37)	2 (±6)	**1**	**1**	0	1
Lost Lake	CCWPP	8 (0.38)	9 (0.47)	7 (0.39)	1 (0.48)	18 (0.33)	1 (0.83)	3 (0.32)	3 (±6)	**1**	**1**	0	1
Crypt Lake	WLNP	15 (0.26)	5 (0.33)	1 (0.35)	13 (0.57)	3 (0.47)	19 (0.57)	6 (0.39)	4 (±7)	**1**	**1**	**1**	**1**
Boundary Pond	WLNP	1 (0.27)	13 (0.74)	12 (0.30)	7 (0.38)	17 (0.31)	6 (0.71)	8 (0.32)	5 (±5)	**1**	**1**	**1**	**1**
South Scarpe Lake	CCWPP	7 (0.27)	8 (0.55)	9 (0.26)	10 (0.49)	12 (0.32)	8 (0.69)	12 (0.34)	6 (±2)	0	1	N-A	1
Alderson Lake	WLNP	6 (0.35)	17 (0.63)	8 (0.60)	2 (0.36)	20 (0.33)	13 (0.63)	1 (0.31)	7 (±7)	**1**	**1**	**1**	**1**
Rainy Ridge 2	CCWPP	23 (0.19)	6 (0.23)	6 (0.21)	21 (0.55)	5 (0.35)	15 (0.63)	20 (0.37)	8 (±8)	0	1	0	1
Lower Rowe Lake	WLNP	2 (0.27)	18 (0.72)	26 (0.28)	9 (0.33)	28 (0.25)	5 (0.72)	9 (0.20)	9 (±10)	1	0	**1**	**1**
Paradise Lake	CCWPP	18 (0.20)	15 (0.30)	10 (0.22)	19 (0.38)	7 (0.32)	18 (0.58)	17 (0.36)	10 (±5)	0	1	N-A	1
Syncline Pond	CCWPP	13 (0.23)	12 (0.33)	25 (0.17)	17 (0.39)	8 (0.25)	9 (0.69)	24 (0.36)	11 (±7)	0	1	N-A	1
Ruby Lake	WLNP	9 (0.32)	1 (0.47)	29 (0.42)	4 (0.80)	29 (0.20)	36 (0.40)	2 (0.20)	12 (±15)	1	0	1	0
Upper Rowe Lake	WLNP	5 (0.28)	20 (0.70)	33 (0.37)	5 (0.33)	33 (0.15)	12 (0.64)	5 (0.15)	13 (±13)	1	0	**1**	**1**
Haig Lake	CCWPP	26 (0.17)	14 (0.11)	5 (0.21)	26 (0.38)	4 (0.35)	23 (0.54)	18 (0.38)	14 (±9)	0	1	0	1
Ruby Lake 2	CCWPP	21 (0.19)	16 (0.25)	14 (0.19)	20 (0.37)	9 (0.30)	17 (0.61)	22 (0.35)	15 (±5)	0	1	0	1

McCarty Pond	CCWPP	20 (0.27)	27 (0.27)	18 (0.28)	12 (0.25)	19 (0.28)	16 (0.62)	10 (0.31)	16 (±6)	0	1	0	1
Buchanan Pond	WLNP	31 (0.33)	4 (0.08)	23 (0.28)	3 (0.61)	26 (0.25)	26 (0.50)	11 (0.25)	17 (±11)	**1**	**1**	**1**	**1**
Jake Smith Lake	CCWPP	10 (0.17)	19 (0.39)	16 (0.15)	27 (0.33)	23 (0.30)	4 (0.73)	27 (0.27)	18 (±9)	**1**	**1**	**1**	**1**
Lys lake	CCWPP	12 (0.15)	26 (0.36)	4 (0.16)	31 (0.25)	11 (0.36)	20 (0.55)	26 (0.35)	19 (±10)	0	1	N-A	1
Barnaby Lake	CCWPP	24 (0.27)	7 (0.12)	21 (0.10)	11 (0.52)	1 (0.26)	35 (0.44)	31 (0.42)	20 (±13)	0	1	N-A	1
North Scarpe Lake	CCWPP	17 (0.17)	22 (0.32)	13 (0.21)	28 (0.30)	16 (0.31)	22 (0.54)	19 (0.32)	21 (±5)	**1**	**1**	0	1
Saint Eloi Tarn	CCWPP	27 (0.19)	24 (0.11)	11 (0.10)	22 (0.28)	13 (0.31)	10 (0.65)	33 (0.34)	22 (±9)	0	1	0	1
Rainy Ridge 3	CCWPP	25 (0.17)	10 (0.11)	19 (0.09)	25 (0.43)	14 (0.28)	14 (0.63)	34 (0.34)	23 (±8)	**1**	**1**	0	1
Southfork Lakes	CCWPP	28 (0.21)	11 (0.10)	15 (0.07)	18 (0.39)	2 (0.30)	32 (0.45)	36 (0.41)	24 (±12)	0	1	N-A	1
Lineham Lakes	WLNP	14 (0.27)	23 (0.33)	36 (0.38)	8 (0.29)	36 (0.08)	28 (0.48)	4 (0.08)	25 (±13)	1	0	1	0
Lone Lake	WLNP	11 (0.18)	29 (0.37)	28 (0.17)	24 (0.23)	27 (0.23)	21 (0.54)	25 (0.24)	26 (±6)	**1**	**1**	N-A	1
Twin Lakes	WLNP	22 (0.16)	25 (0.24)	17 (0.19)	29 (0.25)	25 (0.29)	27 (0.49)	21 (0.25)	27 (±4)	**1**	**1**	N-A	1
Turner's Pond	CCWPP	30 (0.18)	21 (0.08)	27 (0.17)	23 (0.33)	10 (0.24)	34 (0.44)	23 (0.35)	28 (±8)	0	1	N-A	1
Goat Lake	WLNP	19 (0.15)	28 (0.27)	35 (0.23)	33 (0.23)	35 (0.13)	7 (0.70)	16 (0.11)	29 (±11)	1	0	N-A	1
Blue Grouse Basin	WLNP	29 (0.16)	31 (0.09)	20 (0.25)	30 (0.14)	24 (0.27)	29 (0.47)	14 (0.26)	30 (±6)	**1**	**1**	N-A	1
Andy Good Pond	CCWPP	36 (0.26)	30 (0.00)	22 (0.12)	14 (0.17)	22 (0.25)	24 (0.53)	30 (0.29)	31 (±7)	0	1	0	1
Coulthard Tarn	CCWPP	35 (0.23)	35 (0.00)	24 (0.14)	15 (0.08)	21 (0.25)	25 (0.52)	28 (0.29)	32 (±7)	0	1	0	1
Sheep Lake	CCWPP	16 (0.10)	34 (0.33)	30 (0.10)	35 (0.10)	30 (0.19)	11 (0.65)	32 (0.15)	33 (±9)	**1**	**1**	N-A	1
Yarrow Tarn	CCWPP	33 (0.15)	32 (0.03)	34 (0.24)	32 (0.13)	34 (0.15)	30 (0.45)	15 (0.11)	34 (±7)	1	0	N-A	1
Blue (Bovin) Lake	CCWPP	32 (0.10)	36 (0.06)	31 (0.13)	36 (0.07)	32 (0.17)	33 (0.44)	29 (0.15)	35 (±3)	**1**	**1**	N-A	1
Spionkop Tarn	CCWPP	34 (0.12)	33 (0.01)	32 (0.09)	34 (0.12)	31 (0.16)	31 (0.45)	35 (0.15)	36 (±2)	**1**	**1**	N-A	1

*Rankings based on each model individually, habitat suitability value in parenthesis. See Table 3.2 for detailed description of each model

† Only projections for Models 4 & 5 (climate-only + random-background models using Ecoregion & Range study extents respectively) are shown, because these were the models for which future climate projections could be mapped, and without extrapolation at every site (e.g., Model 2; Fig. A3.3)

3.5. Discussion

This study explored the use of SDMs to inform release site selection for potential long-toed salamander translocations in southwestern Alberta, Canada. After extensive sensitivity testing and field validation, I successfully developed seven models with sufficient accuracy and calibration performance to rank 36 possible release sites based on their predicted habitat suitability. I was also able to forecast predictions under two future climate scenarios using two of my models. Joint consideration of current and future predicted suitability revealed three sites that are currently suitable and predicted to remain suitable under both moderate and extreme future emissions scenarios. This study has thus resulted in an actionable list of initial release sites at which to restore populations of a *Species of Special Concern* in Alberta. The extensive surveys included in this study have also improved our understanding of the distribution of long-toed salamanders in this region, expanding the number of known populations that can serve as sources for the translocation. In addition to informing this conservation effort, the rigorous model development and validation approach taken here highlights several key considerations for the use of SDMs in conservation translocation planning more generally.

3.5.2. Lessons learned for the use of SDMs in conservation translocations.

3.5.2.1. Modelling decisions substantially impact model predictions.

My results clearly demonstrate the impact of different modeling decisions on SDM predictions. Although each of the three modeling decisions that I explored influenced model predictions to some degree, the impacts of study extent were particularly striking. This finding is in line with results from other studies (Searcy and

Shaffer, 2014, Vale *et. al.,* 2014) and the results presented in Chapter 2. As discussed in Chapter 2, these impacts are likely caused by changing environmental variable importance as the shape of size of study extent is manipulated. In line with the findings of previous research (*e.g.,* Barber *et. al.,* 2022), I also found that whether sampling bias was accounted for also had substantial impacts on model predictions. However, unlike previous findings (*e.g.,* Hertzog *et. al.,* 2014; Searcy and Shaffer, 2014), I found that models developed with target-group background points performed poorly. Finally, with respect to the set of environmental variables used, I found that the inclusion of non-climatic variables in addition to climatic variables did not impact model predictions as much as the other two decisions I explored. This finding agrees with those of Bradie and Leung, (2017) and Seaborn *et. al.,* (2021), who found that the inclusion of non-climatic variables did little to impact predictions of suitable habitat for several amphibian species. However, the inclusion of non-climate variables may influence predictions of suitable habitat more strongly for other taxa (*e.g.,* endotherms), as the distributions of ectotherms including amphibians tend to be more closely regulated by climate than other taxa (Moreno-Rueda and Pizarro, 2007; Aragón *et. al.,* 2010). Nevertheless, my findings illustrate how different input decisions can have substantial impacts on SDM predictions for a focal species. These effects should be explored and documented when using these models to inform site selection for translocations and other conservation activities.

3.5.2.2. Some models are more accurate than others.

The performance of SDMs is not often evaluated using independent data; yet recent estimates suggest that SDMs are inaccurate between 17 and 47% of the time depending on their application (Lee-Yaw *et. al.,* 2022). My results further demonstrate the need to validate models with independent data. Specifically, less than half of the models that I developed were able to accurately predict independent presence-absence data, despite their ability to predict withheld data. In addition to discrimination performance, I independently evaluated the calibration performance of models, a step that is rarely taken (Lee-Yaw *et. al.,* 2022). In doing so, I found that the same seven models with adequate discrimination performance also had high calibration performance. Assessing model calibration is particularly important in the context of prioritizing release sites for conservation translocations, as only predictions from models with adequate calibration performance can be used to rank sites on a continuous scale of suitability (Gogol-Prokurat, 2011). As such, my recommendations are in line with previous and repeated calls (*e.g.,* Elith and Burgman, 2002; Lee-Yaw *et. al.,* 2022) to validate model predictions with independent data prior to their use in downstream conservation applications.

3.5.2.3. SDMs developed using local study extents may be most appropriate.

Although it was possible to develop an accurate model based on the ecoregion or range-wide study extents, four of the models that passed independent evaluation, including the two "best" models were developed using a political extent. This finding contradicts the results presented in Chapter 1, whereby models based on range-wide extents had the highest mean independent-AUC. However, other studies have reported similar findings (VanDerWal *et. al.,* 2009; Trumbo *et. al.,* 2011; Searcy and Shaffer, 2014). For example, both Searcy and Shaffer, (2014) and Trumbo *et. al.,* (2011) found that local study extents outperformed larger study extents for different amphibian species. Local extents may do particularly well when an area of interest harbours conditions that are not represented elsewhere in the range (Vale *et. al.,* 2014) and/or in cases where populations are locally adapted (Trumbo *et. al.,* 2011; Searcy and Shaffer, 2014; Hällfors *et. al.,* 2016; Connor *et. al.,* 2019). Furthermore, previous studies have observed that models developed using small study extents perform well at predicting occurrence in small focal areas, possibly because these extents capture environmental variability that governs occurrence locally (Fois *et al.,* 2018; Chauvier *et. al.,* 2022). Thus, local extents may be most appropriate when the modeling goal is to obtain highly accurate predictions of suitable habitat for a relatively small region (*e.g.,* within a protected area), as is often the case for conservation translocations (Langridge *et. al.,* 2020).

3.5.2.4. "Best practices" are not universally best.

Several studies have put forth "best practices" for the development of SDMs, including for the type of environmental variables used (*i.e.,* using both climate-only and non-climate environmental variables; Austin and Van Niel, 2011; Araújo *et. al.,* 2019) and strategies to account for sampling bias (*e.g.,* target-group background points; Hertzog *et. al.,* 2014; Barber *et. al.,* 2022). However, following so-called "best practices" does not always result in the best models and in some cases, can result in inaccurate models. For example, in contrast with previous studies (*e.g.,* Hertzog *et. al.,* 2014; Searcy and Shaffer, 2014), I found that models accounting for sampling bias (*i.e.,* target-group background models) performed worse than their random-background counterparts on average. This result may reflect the relatively common occurrence of long-toed salamanders compared to other pond breeding amphibians (*i.e.,* environmental conditions where long-toed salamanders and other amphibians occur are similar, making it hard for models to distinguish between presences and target-group background points) and/or the low number of target-group background points captured by some study extents (*e.g.,* Gaul *et. al.,* 2020). Likewise, several models I developed using climate-only data performed well, and in some cases outperformed their climate + non-climate counterparts (*e.g.,* random-background ecoregion models). Although I did not test all recommended 'best practices', these examples illustrate that suggested approaches to generating SDMs do not universally produce the best performing models. Instead, it is important to test the impacts of key modeling decisions on the accuracy of SDMs and any downstream conservation decisions.

3.5.2.5. Consideration of future site suitability may influence site selection.

My results demonstrate the need to consider both the present-day suitability of potential release sites, and how the suitability of these sites may shift under a changing climate. Specifically, I found that only half of the top ten ranked sites based on current conditions were expected to be suitable 50 years from now based on a moderate future emissions scenario (SSP2-4.5), and only three of these sites were predicted to be suitable under an extreme emissions scenario (SSP4-8.5). In a similar exercise, Bellis *et. al.,* (2021) examined the suitability of 66 release sites from previously conducted translocations of amphibians, reptiles, and terrestrial insects under present-day and future climatic conditions. They found that although present-day release site suitability was positively correlated with translocation success, the predicted suitability of many release sites was predicted to decrease under both optimistic and pessimistic warming scenarios (Bellis *et. al.,* 2021). Thus, consideration of present-day and future release site suitability may be required to ensure not only successful establishment of translocated populations, but long-term persistence.

3.5.2.6. Some accurate models may not be useful for predicting future suitability.

Using MESS surfaces, I found that one of the best-performing models under current conditions (Model 2) could not be projected into the future without extrapolation into non-analog environmental space. In this study, it was possible to avoid extrapolation by focusing on other high-performing models developed with larger study extents when forecasting the suitability of release sites into the future. However, as discussed above, use of larger study extents may result in less accurate models in some cases. Likewise, although models developed using a combination of climate and non-climate variables often performed well, these models could not be forecasted into the future due to unknown future landcover and lack of available layers on primary productivity or wetness. Thus, there may be trade-offs between models that can accurately predict present-day conditions and those that can be forecasted into the future.

3.5.4. Limitations and future directions

My results highlight the potential value of carefully developed SDMs for selecting release sites for conservation translocations. However, several limitations of this study warrant discussion. High-resolution layers were not available across all study extents, so I made use of environmental variables at a coarse resolution (~1 km). Although coarse-grain variables are commonly used to develop SDMs (Manzoor *et. al.,* 2018), and this resolution may be biologically appropriate for long-toed salamanders as they do not typically disperse more than 1 km from breeding sites (Pearson, 2003), several studies have shown that the resolution of input variables can impact model predictions and/or performance (Guisan *et. al.,* 2006; Connor *et. al.,* 2020; Chauvier *et. al.,* 2022). For example, Chauvier *et. al.,* (2022) compared SDMs for 180 plant species across the European Alps at resolutions ranging from 100 m to 40 km. They found that models developed using finer resolution predictors tended to produce better performing models, a result largely consistent across other studies that explored the impacts of variable resolution on performance (Guisan *et. al.,* 2006; Connor *et. al.,* 2020; Chauvier *et. al.,* 2022; but see Farashi and Alizadeh-Noughani, 2018). The continued development of freely available, fine-scale environmental variables across political and jurisdictional borders (*e.g.,* NASA's Earth Science Data Systems (ESDS) program) may allow more trans-boundary models to be developed at fine resolutions in the future, which may in turn lead to more accurate models for use in conservation planning.

Importantly, I note that SDMs developed here did not incorporate biotic interactions that may influence population dynamics and/or be important to consider when selecting appropriate release sites. For example, aquatic invertebrates (*e.g., Dytiscidae;* personal observation), birds (*e.g., Perisoreus canadensis;* Murray *et. al.,* 2005), and other amphibian species (*e.g., Ambystoma tigrinum;* Blaustein *et. al.,* 2001) have been observed predating upon long-toed salamander larvae, and other co-distributed amphibian species also compete with this species (Blaustein *et. al.,* 2001). Despite the potential importance of biotic interactions for shaping species distributions (Wisz *et. al.,* 2013; Paquette and

Hargreaves, 2021), limited data availability often prevents such interactions from being incorporated into SDMs (Dormann *et. al.,* 2018). Even when these data are available, how best to integrate biotic interactions into SDMs remains unclear (König *et. al.,* 2021). Recent modeling advancements such as joint and multi-species SDMs (*e.g.,* Chauvier *et. al.,* 2022; Poggiato *et. al.,* 2022) are emerging as tools to accomplish this goal, but additional research is required to determine if this approach consistently improves SDM performance in all contexts, including for informing conservation translocations.

Finally, I caution that although several of the models I developed could accurately predict independent presence/absence, conditions dictating occurrence may be different from conditions required to establish new populations (Baer and Maron, 2020; Greiser *et. al.,* 2020). Indeed, SDMs often fall short in their ability to predict measures of population performance beyond occurrence (*i.e.,* abundance, growth rates, and genetic diversity; Lee-Yaw *et. al.,* 2022) which may ultimately determine both the establishment and long-term persistence of translocated populations (Fischer and Lindenmayer, 2000). Post-release monitoring of translocated long-toed salamander populations at sites that span the range of predicted suitability rankings will ultimately test whether the models I present here can predict translocation success.

3.5.4. Conclusions

Conservation translocations are expected to become increasingly necessary in the coming decades to restore extirpated populations and slow species declines (Swan *et. al.,* 2018; Novak *et. al.,* 2021; Scheele *et. al.,* 2021). Thus, conservation tools such as SDMs that aid in the effective implementation of these programs are needed. The models presented here represent some of the most rigorously tested and validated SDMs specifically designed for informing conservation translocations to date. While I here focused on the ability of SDMs to aid in the selection and prioritization of suitable release sites, SDMs may aid in other aspects of translocation planning as well. For example, it may be important to consider current or future levels of connectivity between release sites and other occupied sites or suitable habitat to either facilitate the spread of introduced individuals into more habitat (if desired by the management objectives, *e.g.,* assisted migration; Johnson *et. al.,* 2007; Parks *et. al.,* 2022) or to reduce the chances of released individuals migrating away from release sites into unsuitable habitat (Germano and Bishop, 2009). Thus, in addition to assessing the suitability of a particular site, SDMs can also be used to assess the suitability of areas surrounding the release site at distances relevant to the dispersal ability of the focal species (*e.g.,* Westwood *et. al.,* 2020). Likewise, SDMs can be projected beyond the region(s) for which they were developed, which can aid in matching environmental conditions between source and release sites (Malone *et. al.,* 2018; Maes *et. al.,* 2019). Thus, SDMs represent a promising tool for conservation translocation planning, and this study can serve as a template for the rigorous development and validation of such models for this purpose.

CHAPTER 4: CONCLUSION

4.1. Summary of main findings

SDMs have become widely used to understand species distributions and to inform conservation management (Elith *et. al.,* 2006; Phillips *et. al.,* 2006; Guisan *et. al.,* 2013). However, SDMs are sensitive to several user-decisions made during model development (Araújo *et. al.,* 2019; Sofaer *et. al.,* 2019). Much of what we know about the sensitivity of SDMs to these decisions is derived from studies that use internal tests of model performance, despite known issues with this evaluation approach (Elith *et. al.,* 2006; Lobo *et. al.,* 2008; Lee-Yaw *et. al.,* 2022). The lack of independent tests of the accuracy of SDMs (but see Chapter 1) may be particularly problematic in some contexts, such as at the edge of species' ranges, or when these models are used to inform downstream conservation decisions (Gogol-Prokurat, 2011; McCune, 2016).

In this thesis, I used independent data to examine the impacts of key modeling decisions on SDM predictions, performance, and downstream conservation decisions. Across all chapters, I found that independent estimates of model accuracy were consistently lower than internal estimates, highlighting the tendency for internal validation to overestimate model accuracy. In Chapter 2, I showed that model predictions and performance were strongly impacted by the choice of study extent for six amphibian species in protected areas at the edge of their respective ranges in western Canada. Although for some species, local extents performed best (a finding in line with those of previous range-edge SDM studies; Trumbo *et. al.,* 2011; Searcy and Shaffer, 2014; Vale *et. al.,* 2014), this trend was not consistent across all models and many models failed to accurately predict independent occurrence regardless of the study extent used. These findings demonstrate challenges presented when developing SDMs, particularly at species' range limits, and underscore the need to test the impacts of study extent on models generated for use in these in these contexts.

In Chapter 3, I simultaneously examined the impacts of choice of study extent, whether sampling bias is accounted for, and the type of environmental variables used on model predictions and performance for long-toed salamander SDMs in southwestern Alberta. For this study I conducted a large and targeted presence-absence field survey, designed explicitly to evaluate the impacts of these decisions on model performance. I found that some modeling decisions (*i.e.,* choice of study extent and type of background data) impacted model predictions and performance more than others (*i.e.,* the set of environmental variables used), with models developed using random-background points and local study extents tending to produce the most accurate models for this species in the focal area of interest. Some of these findings were in line with those of previous studies (*i.e.,* local extents performed well, Trumbo *et. al.,* 2011; Searcy and Shaffer, 2014, and non-climate data contributed little to model performance; Bradie and Leung, 2017; Seaborn *et. al.,* 2021), but others were surprising. For example, based on previous studies (Hertzog *et. al.,* 2014, Searcy and Shaffer, 2014, and Barber *et. al.,* 2022), I expected target-group background models to outperform random background models. However, I found that random background models tended to perform best. These results emphasize that there is no one-size fits all approach to developing SDMs, and that sensitivity tests are often necessary to develop the most appropriate model for a given context.

Additionally, in Chapter 3 I demonstrated an approach to using the results from my sensitivity tests and model validation to rank sites for conservation purposes—in this case, potential release sites for long-toed salamander translocations. I showed that the predicted suitability of release sites varied depending on the model selected, and that many highly ranked potential translocation sites were not predicted to be suitable under future emissions scenarios. Finally, I also found that in some cases, decisions that produced accurate models (*i.e.,* use of local study extents and inclusion of non-climatic environmental variables) precluded the use of the models for forecasting future site suitability. These results can serve to guide the use of SDMs to inform conservation translocations in the future.

4.2. The use of SDMs at range edges

In high-latitude countries such as Canada, many species occur at the edge of their range (Hunter and Hutchinson, 1994). As niches shift poleward in response to climate change, more species ranges are expected to intersect the borders of high-latitude countries (Parmesan *et. al.,* 2005; Titley *et. al.,* 2021). Thus, SDMs are expected to be increasingly used to better understand the distributions and conservation of peripheral populations. At the same time, peripheral parts of species' ranges may present unique challenges for the development and use of SDMs, as species' distributions may not be in equilibrium with climate, populations may be locally adapted, and/or unique environmental conditions may be present (Kawecki, 2008; Hargreaves *et. al.,* 2014; Bontrager *et. al.,* 2021). Therefore, understanding how SDMs (and decisions made during their development) perform at the edge of species' ranges has implications for use of these models for conservation in Canada.

Results from this thesis demonstrate how 'best practices' for model development (*e.g.,* the use of bias-corrected background data and a combination of climate and non-climate environmental variables) don't always produce the most accurate models at the range edge. At the same time, I found that models developed using local study extents tended to perform well for species considered presently (see also Trumbo *et. al.,* 2011; Searcy and Shafer, 2014; Vale *et. al.,* 2014). However, species are often rare at their range limits (Hunter and Hutchison, 1994; McCune, 2016) and there may be too few locality records at species' range limits to develop meaningful models with very local extents in some cases (*e.g.,* Wisz *et. al.,* 2008, but see Hernandez *et. al.,* 2006; Pearson *et. al.,* 2007, McCune*,* 2016). Critically, I found that in some cases, range-edge SDMs failed to accurately predict independent occurrences regardless of the modeling decisions used (Chapter 2). It may be possible to improve the accuracy of range-edge models by incorporating biotic interactions (*e.g.,* joint, and multi-species SDMs; Poggiato *et. al.,* 2022) or dispersal (Allouche *et. al.,* 2008; Miller *et. al.,* 2015). However, further studies specifically comparing the accuracy of SDMs at range edges to those developed at the core of species ranges are needed to fully assess whether these methods are necessary, or whether for some species, the utility of these tools simply breaks down at the range edge.

4.4. Statement of impact

SDMs are an important tool in ecology and conservation biology. My results, from two studies examining the impact of different modeling decisions on SDM predictions, performance, and downstream conservation decisions, demonstrate the importance of testing the impacts of different modeling decisions with independent data. Overall, my thesis suggests that the use of sensitivity tests and independent validation should become standard practice to maximize the value of SDMs, especially when these tools are used to inform conservation actions.

References

Aguirre-Gutiérrez, J., Carvalheiro, L. G., Polce, C., van Loon, E. E., Raes, N. *et. al.,* (2013). Fit-for-Purpose: Species Distribution Model Performance Depends on Evaluation Criteria – Dutch Hoverflies as a Case Study. *PLoS ONE*, *8*(5), e63708.

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa, and the true skill statistic (TSS): Assessing the accuracy of distribution models. *Journal of Applied Ecology*, *43*(6), 1223–1232.

Alroy, J. (2015). Current extinction rates of reptiles and amphibians. *Proceedings of the National Academy of Sciences*, *112*(42), 13003–13008.

Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, *222*(15), 2796–2811.

Anderson, R. P., & Raza, A. (2010). The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: Preliminary tests with montane rodents (genus Nephelomys) in Venezuela: Effect of study region on models of distributions. *Journal of Biogeography*, *37*(7), 1378–1393.

Angert, A. L., LaDeau, S. L., & Ostfeld, R. S. (2013). Climate change and species interactions: Ways forward: Climate change and species interactions. *Annals of the New York Academy of Sciences*, *1297*(1), 1–7.

Aragón, P., Lobo, J. M., Olalla-Tárraga, M. Á., & Rodríguez, M. Á. (2010). The contribution of contemporary climate to ectothermic and endothermic vertebrate distributions in a glacial refuge: Influences of climate on animal groups. *Global Ecology and Biogeography*, *19*(1), 40–49.

Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F. *et. al.,* (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, *5*(1), eaat4858.

Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, *93*(7), 1527–1539.

Araújo, M. B., & Pearson, R. G. (2005). Equilibrium of species' distributions with climate. *Ecography*, 28(5), 693–695.

Ardestani, E. G., Tarkesh, M., Bassiri, M., & Vahabi, M. R. (2015). Potential habitat modeling for reintroduction of three native plant species in central Iran. *Journal of Arid Land*, *7*(3), 381–390.

Austin, M. P., & Van Niel, K. P. (2011). Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, *38*(1), 1–8.

Bacigalupe, L. D., Vásquez, I. A., Estay, S. A., Valenzuela-Sánchez, A., Alvarado-Rybak, M. *et. al.,* (2019). The amphibian-killing fungus in a biodiversity hotspot: Identifying and validating high-risk areas and refugia. *Ecosphere*, *10*(5).

Baer, K. C., & Maron, J. L. (2020). Ecological niche models display nonlinear relationships with abundance and demographic performance across the latitudinal distribution of *Astragalus utahensis* (Fabaceae). *Ecology and Evolution*, *10*(15), 8251–8264.

Banting, A. L. K., Taylor, M. K., Vinebrooke, R. D., Carli, C. M., & Poesch, M. S. (2021). Assisted colonization of a regionally native predator impacts benthic invertebrates in fishless mountain lakes. *Conservation Science and Practice*, *3*(2).

Barber, R. A., Ball, S. G., Morris, R. K. A., & Gilbert, F. (2022). Target-group backgrounds prove effective at correcting sampling bias in Maxent models. *Diversity and Distributions*, *28*(1), 128–141.

Barbet-Massin, M., Thuiller, W., & Jiguet, F. (2010). How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography*, *33*(5), 878–886.

Beck, J., Böller, M., Erhardt, A., & Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, *19*, 10–15.

Becker, F. S., Slingsby, J. A., Measey, J., Tolley, K. A., & Altwegg, R. (2022). Finding rare species and estimating the probability that all occupied sites have been found. *Ecological Applications*, *32*(2).

Bellamy, C., Scott, C., & Altringham, J. (2013). Multiscale, presence-only habitat suitability models: Fine-resolution maps for eight bat species. *Journal of Applied Ecology*, *50*(4), 892–901.

Bellis, J., Bourke, D., Maschinski, J., Heineman, K., & Dalrymple, S. (2020). Climate suitability as a predictor of conservation translocation failure. *Conservation Biology*, *34*(6), 1473–1481.

Bellis, J., Longden, M., Styles, J., & Dalrymple, S. (2021). Using macroecological species distribution models to estimate changes in the suitability of sites for threatened species reintroduction. *Ecological Solutions and Evidence*, *2*(1).

Benito, B. M., Cayuela, L., & Albuquerque, F. S. (2013). The impact of modelling choices in the predictive performance of richness maps derived from species-distribution models: Guidelines to build better diversity models. *Methods in Ecology and Evolution*, *4*(4), 327–335.

Berger-Tal, O., Blumstein, D. T., & Swaisgood, R. R. (2020). Conservation translocations: A review of common difficulties and promising directions. *Animal Conservation*, *23*(2), 121–131.

Blaustein, A. R., Hatch, A., Belden, L. K., & Wildy, E. L. (2001). Influence of abiotic and biotic factors on amphibians in ephemeral ponds with special reference to long-toed salamanders (*Ambystoma macrodactylum*). *Israel Journal of Zoology*, *47*(4), 333–346.

Bontrager, M., Usui, T., Lee-Yaw, J. A., Anstett, D. N., Hargreaves, A. L., Muir, C. D., & Angert, A. L. (2021). Adaptation across geographic ranges is consistent with strong selection in marginal climates and legacies of range expansion. *Evolution, 75*(6), 1316–1333.

Booth, T. H., Nix, H. A., Busby, J. R., & Hutchinson, M. F. (2014). BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Diversity and Distributions, 20*(1), 1-9.

Börger, L., & Nudds, T. D. (2014). Fire, humans, and climate: Modeling distribution dynamics of boreal forest waterbirds. *Ecological Applications*, *24*(1), 121–141.

Bradie, J., & Leung, B. (2017). A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. *Journal of Biogeography*, *44*(6), 1344–1361.

Breheny, P., & Burchett, W. (2017). Visualization of Regression Models Using visreg. *The R Journal*, *9*(2), 56.

Brotons, L., De Cáceres, M., Fall, A., & Fortin, M.-J. (2012). Modeling bird species distribution change in fire prone Mediterranean landscapes: Incorporating species dispersal and landscape dynamics. *Ecography*, *35*(5), 458–467.

Brotons, L., Herrando, S., & Pla, M. (2007). Updating bird species distribution at large spatial scales: Applications of habitat modelling to data from long-term monitoring programs. *Diversity and Distributions*, *13*(3), 276–288.

Buechling, A., & Tobalske, C. (2011). Predictive Habitat Modeling of Rare Plant Species in Pacific Northwest Forests. *Western Journal of Applied Forestry*, *26*(2), 71–81.

Burns, P. A., Clemann, N., & White, M. (2020). Testing the utility of species distribution modelling using Random Forests for a species in decline: Utility of SDMs for a declining species. *Austral Ecology*, *45*(6), 706-716.

Butchart, S., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., *et. al.,* (2010). Global biodiversity: Indicators of recent declines. *Science*, *328*(5982), 1164–1168.

Chardon, N. I., Nabe-Nielsen, J., Assmann, J. J., Dyrholm Jacobsen, I. B., Guéguen, M., Normand, S., & Wipf, S. (2022). High resolution species distribution and abundance models cannot predict separate shrub datasets in adjacent Arctic fjords. *Diversity and Distributions*, *28*(5), 956–975.

Chardon, N. I., Pironon, S., Peterson, M. L., & Doak, D. F. (2020). Incorporating intraspecific variation into species distribution models improves distribution predictions but cannot predict species traits for a wide-spread plant species. *Ecography*, *43*(1), 60–74.

Chauvier, Y., Descombes, P., Guéguen, M., Boulangeat, L., Thuiller, W., & Zimmermann, N. E. (2022). Resolution in species distribution models shapes spatial patterns of plant multifaceted diversity. *Ecography*, *2022*(10), e05973.

Cianfrani, C., Le Lay, G., Hirzel, A. H., & Loy, A. (2010). Do habitat suitability models reliably predict the recovery areas of threatened species? *Journal of Applied Ecology*, *47*(2), 421–430.

Commission for Environmental Cooperation (CEC). (2020). 2015 Land cover of North America at 30 meters. North American Land Change Monitoring System. Canada Centre for Remote Sensing (CCRS), U.S. Geological Survey (USGS), Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Comisión Nacional Forestal (CONAFOR), Instituto Nacional de Estadística y Geografía (INEGI). Ed. 2.0, Raster digital data [30-m].

Conley, A. K., Fuller, D. O., Haddad, N., Hassan, A. N., Gad, A. M., & Beier, J. C. (2014). Modeling the distribution of the West Nile and Rift Valley Fever vector *Culex pipiens* in arid and semi-arid regions of the Middle East and North Africa. *Parasites & Vectors*, *7*(1), 289.

Connor, T., Viña, A., Winkler, J. A., Hull, V., Tang, Y., Shortridge, A. *et. al.,* (2019). Interactive spatial scale effects on species distribution modeling: The case of the giant panda. *Scientific Reports*, *9*(1), 14563.

Davic, R. D., & Welsh, H. H. (2004). On the ecological roles of salamanders. *Annual Review of Ecology, Evolution, and Systematics*, *35*(1), 405–434.

De Beer, C. J., Dicko, A. H., Ntshangase, J., Moyaba, P., Taioe, M. O., *et. al.,* (2021). A distribution model for *Glossina brevipalpis* and *Glossina* austeni in southern Mozambique, Eswatini and South Africa for enhanced area-wide integrated pest management approaches. *PLOS Neglected Tropical Diseases*, *15*(11), e0009989.

Denton, J. S., Hitchings, S. P., Beebee, T. J. C., & Gent, A. (1997). A recovery program for the natterjack toad (*Bufo calamita*) in Britain. *Conservation Biology*, *11*(6), 1329–1338.

Devarajan, K., Morelli, T. L., & Tenan, S. (2020). Multi-species occupancy models: Review, roadmap, and recommendations. *Ecography*, *43*(11), 1612–1624.

Domahidi, Z., Shonfield, J., Nielsen, S. E., Spence, J. R., & Bayne, E. M. (2019). Spatial distribution of the Boreal Owl and Northern Saw-whet Owl in the Boreal region of Alberta, Canada. *Avian Conservation and Ecology*, *14*(2), art14.

Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., *et. al.,* (2018). Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. *Global Ecology and Biogeography*, *27*(9), 1004–1016.

Ebi, K. L., Kram, T., Van Vuuren, D. P., O'Neill, B. C., & Kriegler, E. (2014). A new toolkit for developing scenarios for climate change research and policy analysis. *Environment: Science and Policy for Sustainable Development*, *56*(2), 6–16.

Edvardsen, A., Bakkestuen, V., & Halvorsen, R. (2011). A fine-grained spatial prediction model for the red-listed vascular plant *Scorzonera humilis*. *Nordic Journal of Botany*, *29*(4), 495–504.

Elith J., and Burgman M.A. (2002). Predictions and their validation: rare plants in the Central Highlands, Victoria, Australia. In: Predicting Species Occurrences: Issues of Accuracy and Scale, Editors: Scott J.M., Heglund P.J., Morrison M.L., Raphael M.G., Wall W.A., and Samson F.B. pp.303-314. Island Press, Covelo, CA.

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S. *et. al.,* (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129–151.

Elith, J., Kearney, M. R., & Phillips, S. J. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*, 330–342.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. *Diversity and Distributions*, *17*(1), 43–57.

Eyre, A. C., Briscoe, N. J., Harley, D. K. P., Lumsden, L. F., McComb, L. B., & Lentini, P. E. (2022). Using species distribution models and decision tools to direct surveys and identify potential translocation sites for a critically endangered species. *Diversity and Distributions*, *28*(4), 700–711.

Farashi, A., & Alizadeh-Noughani, M. (2018). Effects of models and spatial resolutions on the species distribution model performance. *Modeling Earth Systems and Environment*, *4*(1), 263–268.

Ficetola, G. F., Rondinini, C., Bonardi, A., Baisero, D., & Padoa-Schioppa, E. (2015). Habitat availability for amphibians and extinction threat: A global analysis. *Diversity and Distributions, 21*(3), 302-311.

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

Fischer, J., & Lindenmayer, D. B. (2000). An assessment of the published results of animal relocations. *Biological Conservation*, *96*(1), 1–11.

Fois, M., Cuena-Lombraña, A., Fenu, G., & Bacchetta, G. (2018). Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues, and future directions. *Ecological Modelling*, *385*, 124–132.

Forti, L. R., Becker, C. G., Tacioli, L., Pereira, V. R., Santos, A. C. F. A. *et. al.,* (2017). Perspectives on invasive amphibians in Brazil. *PLOS ONE*, *12*(9), e0184703.

Fourcade, Y., Besnard, A. G., & Secondi, J. (2018). Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Biogeography*, *27*(2), 245–256.

Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. *PLoS ONE*, *9*(5), e97122.

Franklin, J., Wejnert, K. E., Hathaway, S. A., Rochester, C. J., & Fisher, R. N. (2009). Effect of species rarity on the accuracy of species distribution models for reptiles and amphibians in southern California. *Diversity and Distributions*, *15*(1), 167–177.

Freeman, E. A., & Moisen, G. (2008). PresenceAbsence: An *R* package for presence absence analysis. *Journal of Statistical Software*, *23*(11).

Funk, W. C., & Dunlap, W. W. (1999). Colonization of high-elevation lakes by long-toed salamanders (*Ambystoma macrodactylum*) after the extinction of introduced trout populations. *Canadian Journal of* Zoology, *77*, 1759-1767.

Galluzzi, M., Armanini, M., Ferrari, G., Zibordi, F., Nocentini, S., & Mustoni, A. (2017). Habitat Suitability Models, for ecological study of the alpine marmot in the central Italian Alps. *Ecological Informatics*, *37*, 10–17.

Gamliel, I., Buba, Y., Guy-Haim, T., Garval, T., Willette, D., Rilov, G., & Belmaker, J. (2020). Incorporating physiology into species distribution models moderates the projected impact of warming on selected Mediterranean marine species. *Ecography*, *43*(7), 1090–1106.

Gaston, K. J. (2003). *The Structure and Dynamics of Geographic Ranges*. Oxford University Press.

Gaul, W., Sadykova, D., White, H. J., Leon-Sanchez, L., Caplat, P., Emmerson, M. C., & Yearsley, J. M. (2020). Data quantity is more important than its spatial bias for predictive species distribution modelling. *PeerJ*, *8*, e10411.

Gavashelishvili, A., & Javakhishvili, Z. (2010). Combining radio-telemetry and random observations to model the habitat of Near Threatened Caucasian grouse *Tetrao mlokosiewiczi*. *Oryx*, *44*(4), 491–500.

Germano, J. M., & Bishop, P. J. (2009). Suitability of amphibians and reptiles for translocation. *Conservation Biology*, *23*(1), 7–15.

Giné, G. A. F., & Faria, D. (2018). Combining species distribution modeling and field surveys to reappraise the geographic distribution and conservation status of the threatened thin-spined porcupine (*Chaetomys subspinosus*). *PLOS ONE*, *13*(11), e0207914.

Gogol-Prokurat, M. (2011). Predicting habitat suitability for rare plants at local spatial scales using a species distribution model. *Ecological Applications*, *21*(1), 33–47.

Goldberg, C. S., & Waits, L. P. (2010). Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular ecology, 19*(17), 3650-3663.

Graham, C. H., Silva, N., & Velásquez-Tibatá, J. (2010). Evaluating the potential causes of range limits of birds of the Colombian Andes. *Journal of Biogeography*, *37*(10), 1863-1875.

Graham, K. L., & Powell, G. L. (1999). Status of the long-toed salamander *(Ambystoma macrodactylum)* in Alberta. Alberta Environmental Protection, Fisheries and

Wildlife Management Division, and Alberta Conservation Association. Wildlife Status Report No. 22, Edmonton, Alberta. 19 pp.

Greiser, C., Hylander, K., Meineri, E., Luoto, M., & Ehrlén, J. (2020). Climate limitation at the cold edge: Contrasting perspectives from species distribution modelling and a transplant experiment. *Ecography*, *43*(5), 637–647.

Griffiths, R. A., & Pavajeau, L. (2008). Captive breeding, reintroduction, and the conservation of amphibians: Amphibian captive breeding. *Conservation Biology*, *22*(4), 852–861.

Groff, L. A., Marks, S. B., & Hayes, M. P. (2014). Using ecological niche models to direct rare amphibian surveys: A case study using the Oregon spotted frog (*Rana pretiosa*). *Herpetological Conservation and Biology*, 15.

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., *et. al.,* (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, *24*, 276–292.

Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J. Mc. C., Aspinall, R., & Hastie, T. (2006). Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology*, *43*(3), 386–392.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., *et. al.,* (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424–1435.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186.

Hagens, S. V., Rendall, A. R., & Whisson, D. A. (2018). Passive acoustic surveys for predicting species' distributions: Optimising detection probability. *PLOS ONE*, *13*(7), e0199396.

Hagen-Zanker, A. (2009). An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation. *International Journal of Geographical Information Science*, *23*(1), 61–73.

Halbritter, A. H., Billeter, R., Edwards, P. J., & Alexander, J. M. (2015). Local adaptation at range edges: Comparing elevation and latitudinal gradients. *Journal of Evolutionary Biology*, *28*(10), 1849–1860.

Hällfors, M. H., Liao, J., Dzurisin, J., Grundel, R., Hyvärinen, M., *et. al.,* (2016). Addressing potential local adaptation in species distribution models: Implications for conservation under climate change. *Ecological Applications*, *26*(4), 1154–1169.

Hallgren, W., Santana, F., Low-Choy, S., Zhao, Y., & Mackey, B. (2019). Species distribution models can be highly sensitive to algorithm configuration. *Ecological Modelling*, *408*, 108719.

Hao, T., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2020). Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography*, *43*(4), 549-558.

Harding, G., Griffiths, R. A., & Pavajeau, L. (2016). Developments in amphibian captive breeding and reintroduction programs: Amphibian Breeding and Reintroduction. *Conservation Biology*, *30*(2), 340–349.

Hargreaves, A. L., & Eckert, C. G. (2019). Local adaptation primes cold-edge populations for range expansion but not warming-induced range shifts. *Ecology Letters*, *22*(1), 78–88.

Hargreaves, A. L., Samis, K. E., & Eckert, C. G. (2014). Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *The American Naturalist*, *183*(2), 157–173.

Heikkinen, R. K., Luoto, M., Virkkala, R., Pearson, R. G., & Körber, J.-H. (2007). Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, *16*(6), 754–763.

Hernandez, P. A., Graham, C. H., Master, L. L., & Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography, 29*(5), 773-785.

Hertzog, L. R., Besnard, A., & Jay-Robert, P. (2014). Field validation shows bias-corrected pseudo-absence selection is the best method for predictive species-distribution modelling. *Diversity and Distributions*, *20*(12), 1403–1413.

Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J., & Hijmans, M. R. J. (2017). Package 'dismo'. *Circles, 9*(1), 1-68.

Hunter, D. (2022). Effects of a severe wildfire on amphibian populations in Waterton Lakes National Park. University of Lethbridge, Lethbrirdge, Canada.

Hunter, M. L., & Hutchinson, A. (1994). The virtues and shortcomings of parochialism: conserving species that are locally rare, but globally common. *Conservation Biology*, *8*(4), 1163–1165.

Hunter-Ayad, J., Ohlemüller, R., Recio, M. R., & Seddon, P. J. (2020). Reintroduction modelling: A guide to choosing and combining models for species reintroductions. *Journal of Applied Ecology*, *57*(7), 1233–1243.

IUCN/SSC. (2013). Guidelines for reintroductions and other conservation translocations. Version 1.0. IUCN Species Survival Commission, Gland, Switzerland.

Jiménez-Alfaro, B., Draper, D., & Nogués-Bravo, D. (2012). Modeling the potential area of occupancy at fine resolution may reduce uncertainty in species range estimates. *Biological Conservation*, *147*(1), 190–196.

Jinga, P., Liao, Z., & Nobis, M. P. (2021). Species distribution modeling that overlooks intraspecific variation is inadequate for proper conservation of marula (*Sclerocarya birrea, Anacardiaceae*). *Global Ecology and Conservation*, *32*, e01908.

Johnson, H., Bleich, V. C., & Stephenson, T. R. (2007). Modelling Sierra Nevada Bighorn Sheep habitat: Applying resource selection functions to species recovery. California Department of Fish and Game, Bishop, CA.

Johnson, S., Molano-Flores, B., & Zaya, D. (2023). Field validation as a tool for mitigating uncertainty in species distribution modeling for conservation planning. *Conservation Science and Practice*, e12978.

Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla-Buitrago, G. E., *et. al.,* (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution*, *12*(9), 1602–1608.

Kawecki, T. J. (2008). Adaptation to marginal habitats. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 321–342.

Keinath, D. A., Andersen, M. D., & Beauvais, G. P. (2010). Range and modeled distribution of Wyoming's species of greatest conservation need. Report prepared by the Wyoming Natural Diversity Database, Laramie Wyoming for the Wyoming Game and Fish Department, Cheyenne, Wyoming and the US Geological Survey, Fort Collins, Colorado.

Keitt, T. H., Lewis, M. A., & Holt, R. D. (2001). Allee effects, invasion pinning, and species' borders. *The American Naturalist*, *157*(2), 203–216.

Kenison, E. K., Litt, A. R., Pilliod, D. S., & McMahon, T. E. (2016). Role of habitat complexity in predator–prey dynamics between an introduced fish and larval long-toed salamanders (*Ambystoma macrodactylum*). *Canadian Journal of Zoology*, *94*(4), 243–249.

Kharouba, H. M., McCune, J. L., Thuiller, W., & Huntley, B. (2013). Do ecological differences between taxonomic groups influence the relationship between species' distributions and climate? A global meta-analysis using species distribution models. *Ecography*, *36*(6), 657–664.

Kirk, T. A., & Zielinski, W. J. (2009). Developing and testing a landscape habitat suitability model for the American marten (*Martes americana*) in the Cascades mountains of California. *Landscape Ecology*, *24*(6), 759–773.

Kissling, W. D., Dormann, C. F., Groeneveld, J., Hickler, T., Kühn, I., *et. al.,* (2012). Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents: Modelling multispecies interactions. *Journal of Biogeography*, *39*(12), 2163–2178.

König, C., Wüest, R. O., Graham, C. H., Karger, D. N., Sattler, T., *et. al.,* (2021). Scale dependency of joint species distribution models challenges interpretation of biotic interactions. *Journal of Biogeography*, *48*(7), 1541–1551.

Konowalik, K., & Nosol, A. (2021). Evaluation metrics and validation of presence-only species distribution models based on distributional maps with varying coverage. *Scientific Reports*, *11*(1), 1482.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *260*(7), 583-621.

Langridge, J., Sordello, R., & Reyjol, Y. (2020). Outcomes of wildlife translocations in protected areas: What is the type and extent of existing evidence? A systematic map protocol. *Environmental Evidence*, *9*(1), 16.

Lathrop, R. G., Niles, L., Smith, P., Peck, M., Dey, A., Sacatelli, R., & Bognar, J. (2018). Mapping and modeling the breeding habitat of the Western Atlantic Red Knot (*Calidris canutus rufa*) at local and regional scales. *The Condor*, *120*(3), 650–665.

Le Lay, G., Engler, R., Franc, E., & Guisan, A. (2010). Prospective sampling based on model ensembles improves the detection of rare species. *Ecography*, *33*(6), 1015–1027.

Lee-Yaw, J. A., & Irwin, D. E. (2015). The importance (or lack thereof) of niche divergence to the maintenance of a northern species complex: The case of the long-toed salamander (*Ambystoma macrodactylum*). *Journal of Evolutionary Biology*, *28*(4), 917–930.

Lee-Yaw, J. A., Kharouba, H. M., Bontrager, M., Mahony, C., Csergő, A. M., *et. al.,* (2016). A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecology Letters*, *19*(6), 710–722.

Lee-Yaw, J. A., McCune, J. L., Pironon, S., & Sheth, S. N. (2022). Species distribution models rarely predict the biology of real populations. *Ecography*, *2022*(6), e05877.

Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, *17*(2), 145–151.

Luoto, M., Pöyry, J., Heikkinen, R. K., & Saarinen, K. (2005). Uncertainty of bioclimate envelope models based on the geographical distribution of species. *Global Ecology and Biogeography*, *14*(6), 575–584.

MacArthur, R. H. (1984). Geographical ecology: patterns in the distribution of species. *Princeton University Press*.

Maes, D., Ellis, S., Goffart, P., Cruickshanks, K. L., van Swaay, C. A. M., *et. al.,* (2019). The potential of species distribution modelling for reintroduction projects: The case study of the Chequered Skipper in England. *Journal of Insect Conservation*, *23*(2), 419–431.

Magg, N., Müller, J., Heibl, C., Hackländer, K., Wölfl, S., *et. al.,* (2016). Habitat availability is not limiting the distribution of the Bohemian–Bavarian lynx (*Lynx lynx*) population. *Oryx*, *50*(4), 742–752.

Mahony, C. R., Wang, T., Hamann, A., & Cannon, A. J. (2022). A global climate model ensemble for downscaled monthly climate normals over North America. *International Journal of Climatology*, *42*(11), 5871–5891.

Malone, E. W., Perkin, J. S., Leckie, B. M., Kulp, M. A., Hurt, C. R., & Walker, D. M. (2018). Which species, how many, and from where: Integrating habitat suitability, population genomics, and abundance estimates into species reintroduction planning. *Global Change Biology*, *24*(8), 3729–3748.

Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size – finer may not always be better. *Scientific Reports*, *8*(1), 7168.

Marage, D., Garraud, L., & Rameau, J. (2008). The influence of management history on spatial prediction of *Eryngium spinalba*, an endangered endemic species. *Applied Vegetation Science*, *11*(1), 139–148.

Matutini, F., Baudry, J., Pain, G., Sineau, M., & Pithon, J. (2021). How citizen science could improve species distribution models and their independent assessment. *Ecology and Evolution*, *11*(7), 3028–3039.

McCoy, E. D., Osman, N., Hauch, B., Emerick, A., & Mushinsky, H. R. (2014). Increasing the chance of successful translocation of a threatened lizard: Improving chances of success in reptile translocations. *Animal Conservation*, *17*, 56–64.

McCune, J. L. (2016). Species distribution models predict rare species occurrences despite significant effects of landscape context. *Journal of Applied Ecology*, *53*(6), 1871–1879.

McCune, J. L., Rosner-Katz, H., Bennett, J. R., Schuster, R., & Kharouba, H. M. (2020). Do traits of plant species predict the efficacy of species distribution models for finding new occurrences? *Ecology and Evolution*, *10*(11), 5001–5014.

Merker, S. A., & Chandler, R. B. (2021). An experimental test of the Allee effect range limitation hypothesis. *Journal of Animal Ecology*, *90*(3), 585–593.

Merow, C., & Silander, J. A. (2014). A comparison of Maxlike and Maxent for modelling species distributions. *Methods in Ecology and Evolution*, *5*(3), 215–225.

Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, *36*(10), 1058–1069.

Milanovich, J. R., Peterman, W. E., Barrett, K., & Hopton, M. E. (2012). Do species distribution models predict species richness in urban and natural green spaces? A case study using amphibians. *Landscape and Urban Planning*, *107*(4), 409–418.

Morán-Ordóñez, A., Lahoz-Monfort, J. J., Elith, J., & Wintle, B. A. (2017). Evaluating 318 continental-scale species distribution models over a 60-year prediction horizon: What factors influence the reliability of predictions? *Global Ecology and Biogeography*, *26*(3), 371–384.

Moreno-Rueda, G., & Pizarro, M. (2023). Snake species richness and shrubland correlate with the short-toed eagle (*Circaetus gallicus*) distribution in southeastern Spain. *Annales Zoologici Fennici, 44*(4), 314-320.

Murray, M. P., Pearl, C. A., & Bury, R. B. (2005). Apparent Predation by Gray Jays (*Perisoreus canadensis*) on Long-toed Salamanders (*Ambystoma macrodactylum*), in the Oregon Cascade Range. *The Canadian Field-Naturalist*, *119*(2), 291.

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MAXENT ecological niche models. *Methods in Ecology and Evolution*, *5*(11), 1198–1205.

Muscatello, A., Elith, J., & Kujala, H. (2021). How decisions about fitting species distribution models affect conservation outcomes. *Conservation Biology*, *35*(4), 1309–1320.

Newbold, T., Reader, T., El-Gabbas, A., Berg, W., Shohdi, W. M., *et. al.,* (2010). Testing the accuracy of species distribution models using species records from a new field survey. *Oikos, 119*(8), 1326-1334.

Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., *et. al.,* (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs*, *89*(3).

Novak, B. J., Phelan, R., & Weber, M. (2021). U.S. conservation translocations: Over a century of intended consequences. *Conservation Science and Practice*, *3*(4).

Oldfather, M. F., Kling, M. M., Sheth, S. N., Emery, N. C., & Ackerly, D. D. (2019). Range edges in heterogeneous landscapes: Integrating geographic scale and climate complexity into range dynamics. *Global Change Biology*, *26*(3), 1055–1067.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., *et. al.,* (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51*(11), 933.

Paquette, A., & Hargreaves, A. L. (2021). Biotic interactions are more often important at species' warm versus cool range edges. *Ecology Letters*, *24*(11), 2427-2438.

Parker, B. R., Schindler, D. W., Donald, D. B., & Anderson, R. S. (2001). The effects of stocking and removal of a nonnative salmonid on the plankton of an alpine lake. *Ecosystems*, *4*(4), 334–345.

Parks, S. A., Holsinger, L. M., Littlefield, C. E., Dobrowski, S. Z., Zeller, K. A., *et. al.,* (2022). Efficacy of the global protected area network is threatened by disappearing climates and potential transboundary range shifts. Environmental Research Letters, 17(5), 054016.

Parmesan, C., Gaines, S., Gonzalez, L., Kaufman, D. M., Kingsolver, J., *et. al.,* (2005). Empirical perspectives on species borders: From traditional biogeography to global change. *Oikos*, *108*(1), 58–75.

Pearson, K. J., & Goater, C. P. (2008). Distribution of long-toed salamanders and introduced trout in high- and low-elevation wetlands in southwestern Alberta, Canada. *Ecoscience*, *15*(4), 453–459.

Pearson, K.J. (2003). Distribution and habitat associations of the long-toed salamander (*Ambystoma macrodactylum*) in the Oldman River Drainage. Alberta Conservation

Association, Alberta Sustainable Resource Development, Fish & Wildlife Division, Wildlife Conservation and Biodiversity Section. Edmonton, Alberta.

Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography, *34*, 102–117.

Pebesma, E. (2018). Simple features for R: standardized support for spatial vector data. *The R Journal*, *10*(1), 439.

Pédarros, É., Coetzee, T., Fritz, H., & Guerbois, C. (2020). Rallying citizen knowledge to assess wildlife occurrence and habitat suitability in anthropogenic landscapes. *Biological Conservation*, *242*, 108407.

Peterman, W. E., Crawford, J. A., & Kuhns, A. R. (2013). Using species distribution and occupancy modeling to guide survey efforts and assess species status. *Journal for Nature Conservation*, *21*(2), 114–121.

Peterson, A. T. (2011). Ecological niches and geographic distributions. Princeton University Press.

Peterson, A. T., & Vieglais, D. A. (2001). Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *BioScience*, *51*(5), 363.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190*,(3-4), 231-259.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161–175.

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

Pilliod, D. S., Hossack, B. R., Bahls, P. F., Bull, E. L., Corn, P. S., *et. al.m,* (2010). Non-native salmonids affect amphibian occupancy at multiple spatial scales: Multi-scale fish effects on amphibians. *Diversity and Distributions*, *16*(6), 959–974.

Pilliod, D. S., Arkle, R. S., Robertson, J. M., Murphy, M. A., & Funk, W. C. (2015). Effects of changing climate on aquatic habitat and connectivity for remnant populations of a wide-ranging frog species in an arid landscape. *Ecology and Evolution, 5(*18), 3979-3994.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., *et. al.,* (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187), 1246752.

Pineda, E., & Lobo, J. M. (2009). Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *Journal of Animal Ecology*, *78*, 182–190.

Poggiato, G., Andréoletti, J., Shirley, L., & Thuiller, W. (2022). Integrating food webs in species distribution models improves ecological niche estimation and predictions. *Authorea*.

Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., *et. al.,* (2014). Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, *5*(5), 397–406.

Qian, H., Wang, X., Wang, S., & Li, Y. (2007). Environmental determinants of amphibian and reptile species richness in China. *Ecography, 30*(4), 471–482.

Radomski, T., Beamer, D., Babineau, A., Wilson, C., Pechmann, J., & Kozak, K. H. (2022). Finding what you don't know: Testing SDM methods for poorly known species. *Diversity and Distributions*, *28*, 1769-1780.

Raes, N. (2012). Partial versus full species distribution models. *Natureza & Conservação*, *10*(2), 127–138.

Randin, C. F., Dirnböck, T., Dullinger, S., Zimmermann, N. E., Zappa, M., & Guisan, A. (2006). Are niche-based species distribution models transferable in space? *Journal of Biogeography*, *33*(10), 1689–1703.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rebelo, H., & Jones, G. (2010). Ground validation of presence-only modelling with rare species: A case study on barbastelles *Barbastella barbastellus* (Chiroptera: Vespertilionidae). *Journal of Applied Ecology*, *47*(2), 410–420.

Riaz, M., Kuemmerlen, M., Wittwer, C., Cocchiararo, B., Khaliq, I., *et. al.,* (2020). Combining environmental DNA and species distribution modeling to evaluate reintroduction success of a freshwater fish. *Ecological Applications*, *30*(2).

Rosner-Katz, H., McCune, J. L., & Bennett, J. R. (2020). Using stacked SDMs with accuracy and rarity weighting to optimize surveys for rare plant species. *Biodiversity and Conservation*, *29*(11–12), 3209–3225.

Roy, K., Hunt, G., Jablonski, D., Krug, A.Z. & Valentine, J.W. (2009) A macroevolutionary perspective on species range limits. *Proceedings of the Royal Society 276*, 1485– 1493.

Royle, J. A., Chandler, R. B., Yackulic, C., & Nichols, J. D. (2012). Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution*, *3*(3), 545–554.

Royle, J. A., & Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, *87*(4), 835–841.

Running, S., & Zhao, M. (2019). - MOD17A3HGF MODIS/Terra net primary production gap-filled yearly L4 global 500 m SIN grid V006, NASA EOSDIS Land Processes DAAC.

Sacks, B. N., Statham, M. J., & Wittmer, H. U. (2017). A Preliminary Range-Wide Distribution Model for the Sacramento Valley Red Fox. *Journal of Fish and Wildlife Management*, *8*(1), 28–38.

Safaei, M., Tarkesh, M., Bashari, H., & Bassiri, M. (2018). Modeling potential habitat of *Astragalus verus Olivier* for conservation decisions: A comparison of three correlative models. *Flora*, *242*, 61–69.

Sánchez-Fernández, D., Lobo, J. M., & Hernández-Manrique, O. L. (2011). Species distribution models that do not incorporate global data misrepresent potential distributions: A case study using Iberian diving beetles. *Diversity and Distributions*, *17*(1), 163–171.

Sarrazin, F., & Legendre, S. (2000). Demographic approach to releasing adults versus young in reintroductions. *Conservation Biology*, *14*(2), 488–500.

Saupe, E. E., Barve, V., Myers, C. E., Soberón, J., Barve, N., *et. al.,* (2012). Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, *237–238*, 11–22.

Scheele, B. C., Hollanders, M., Hoffmann, E. P., Newell, D. A., Lindenmayer, D. B., *et. al.,* (2021). Conservation translocations for amphibian species threatened by chytrid fungus: A review, conceptual framework, and recommendations. *Conservation Science and Practice*, *3*(11), e524.

Schmidt, H., Radinger, J., Teschlade, D., & Stoll, S. (2020). The role of spatial units in modelling freshwater fish distributions: Comparing a subcatchment and river network approach using MaxEnt. *Ecological Modelling*, *418*, 108937.

Seaborn, T., Goldberg, C. S., & Crespi, E. J. (2021). Drivers of distributions and niches of North American cold-adapted amphibians: Evaluating both climate and land use. *Ecological Applications*, *31*(2), e2236.

Searcy, C. A., & Shaffer, H. B. (2014). Field validation supports novel niche modeling strategies in a cryptic endangered amphibian. *Ecography*, *37*(10), 983–992.

Seoane, J., Carrascal, L. M., Alonso, C. L., & Palomino, D. (2005). Species-specific traits associated to prediction errors in bird habitat suitability modelling. *Ecological Modelling*, *185*(2–4), 299–308.

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 415–436.

Shabani, F., Kumar, L., & Ahmadi, M. (2018). Assessing accuracy methods of species distribution models: AUC, specificity, sensitivity and the True Skill Statistic. *Global Journal of Human Social Science, 18*(1), 7-18.

Smith, A. B., Beever, E. A., Kessler, A. E., Johnston, A. N., Ray, C., *et. al.,* (2019). Alternatives to genetic affinity as a context for within-species response to climate. *Nature Climate Change*, *9*(10), 787–794.

Sofaer, H. R., Jarnevich, C. S., Pearse, I. S., Smyth, R. L., Auer, S., *et. al.,* (2019). Development and delivery of species distribution models to inform decision-making. *BioScience*, *69*(7), 544–557.

Soorae, P.S., (2018). Global reintroduction perspectives: 2018. Case studies from around the globe (6th ed.). IUCN, International Union for Conservation of Nature.

Stryszowska, K. M., Johnson, G., Mendoza, L. R., & Langen, T. A. (2016). Species distribution modeling of the threatened Blanding's turtle's (*Emydoidea blandingii*) range edge as a tool for conservation planning. *Journal of Herpetology*, *50*(3), 366–373.

Stuart, S. N., Chanson, J. S., Cox, N. A., Young, B. E., Rodrigues, A. S. L., *et. al.,* (2004). Status and trends of amphibian declines and extinctions worldwide. *Science*, *306*(5702), 1783–1786.

Svenning, J. C., Normand, S., & Skov, F. (2008). Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, *31*(3), 316–326.

Swan, K. D., Lloyd, N. A., & Moehrenschlager, A. (2018). Projecting further increases in conservation translocations: A Canadian case study. *Biological Conservation*, *228*, 175–182.

Swets, J. A. (1988). Measuring the Accuracy of Diagnostic Systems. *Science*, *240*(4857), 1285–1293.

Tessarolo, G., Lobo, J. M., Rangel, T. F., & Hortal, J. (2021). High uncertainty in the effects of data characteristics on the performance of species distribution models. *Ecological Indicators*, *121*, 107147.

Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, *27*(2), 165–172.

Titley, M. A., Butchart, S. H., Jones, V. R., Whittingham, M. J., & Willis, S. G. (2021). Global inequities and political borders challenge nature conservation under climate change. *Proceedings of the National Academy of Sciences, 118*(7), e2011204118.

Trivedi, M. R., Berry, P. M., Morecroft, M. D., & Dawson, T. P. (2008). Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Global Change Biology*, *14*(5), 1089–1103.

Trumbo, D. R., Burgett, A. A., & Knouft, J. H. (2011). Testing climate-based species distribution models with recent field surveys of pond-breeding amphibians in eastern Missouri. *Canadian Journal of Zoology*, *89*(11), 1074–1083.

Trumbo, D. R., Burgett, A. A., Hopkins, R. L., Biro, E. G., Chase, J. M., & Knouft, J. H. (2012). Integrating local breeding pond, landcover, and climate factors in predicting amphibian distributions. *Landscape Ecology*, *27*(8), 1183–1196.

Tye, C. A., McCleery, R. A., Fletcher, R. J., Greene, D. U., & Butryn, R. S. (2017). Evaluating citizen vs. Professional data for modelling distributions of a rare squirrel. *Journal of Applied Ecology*, *54*(2), 628–637.

Vale, C. G., Tarroso, P., & Brito, J. C. (2014). Predicting species distribution at range margins: Testing the effects of study area extent, resolution and threshold selection in the Sahara-Sahel transition zone. *Diversity and Distributions*, *20*(1), 20–33.

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.

Vaughan, I. P., & Ormerod, S. J. (2005). The continuing challenges of testing species distribution models: Testing distribution models. *Journal of Applied Ecology*, *42*(4), 720–730.

Visser, H., & De Nijs, T. (2006). The Map Comparison Kit. *Environmental Modelling & Software*, *21*(3), 346–358.

Walker, C. M. (2018). Spatial and Temporal Components of Environmental Predictors in Species Distribution Modeling: A Case Study Using Bell's Vireo *(Vireo bellii)*. Oklahoma State University, Stillwater, United States.

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), e0156720.

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

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., *et. al.,* (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Wilson, G. A., Fulton, T. L., Kendell, K., Scrimgeour, G., Paszkowski, C. A., & Coltman, D. W. (2008). Genetic diversity and structure in Canadian northern leopard frog (*Rana pipiens*) populations: Implications for reintroduction programs. *Canadian Journal of Zoology, 86*, (12), 863-874.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., & NCEAS Predicting Species Distributions Working Group. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), 763–773.

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., *et. al.,* (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, *88*(1), 15–30.

Appendix I: Supplemental Tables and Figures

CHAPTER 1: INTRODUCTION

Table A1.1. Details of Web of Science search conducted to retrieve papers that independently validated species distribution models.

Web of Science search string*	Number of papers retrieved	Screening criteria from Lee-Yaw *et. al.,* (2022)	Number of papers remaining after Lee-Yaw *et. al.,* (2022) criteria	New screening criteria	Number of studies retained
(((TS=("distribution model*" OR "niche model*" OR "habitat suitability model*" OR "habitat model*" OR "bioclimatic envelope*") AND TS=("independent data" OR "independent presence-absence" OR "independent survey*" OR "ground validation" OR "field survey*" OR "independent evaluation" OR "independent field data" OR "independent species data" OR "independent grid cells" OR "field validat*" OR "field-validated" OR "independent sampling" OR "field sampling" OR "evaluation data" OR "independent record*" OR "model-based sampling" OR "field occurrence data") AND LA=(English)) AND DT=(Article))), **Timespan: 2021-01-01 to 2023-03-26**	156	Studies needed to include more than one independent observation from both suitable and unsuitable sites. Excluded studies in which the independent data were based on fossils, anecdotal sightings (*e.g.,* surveys of local people), or expert opinion without actual observations. Independent observations had to come from the same region and time for which the model was developed (studies involving spatial or temporal transfer were excluded). Studies modeling ensembles of species rather than individual species (e.g., modeling vegetation types) were excluded.	15 *from new search* + 101 *from Lee-Yaw et. al., (2022)* = **116** *total*	Excluded studies that developed models for marine/aquatic species (terrestrial organisms only) Excluded studies that did not report independent AUC scores for individual models.	39

*Identical to Lee-Yaw *et. al.,* (2022) string except for timespan.

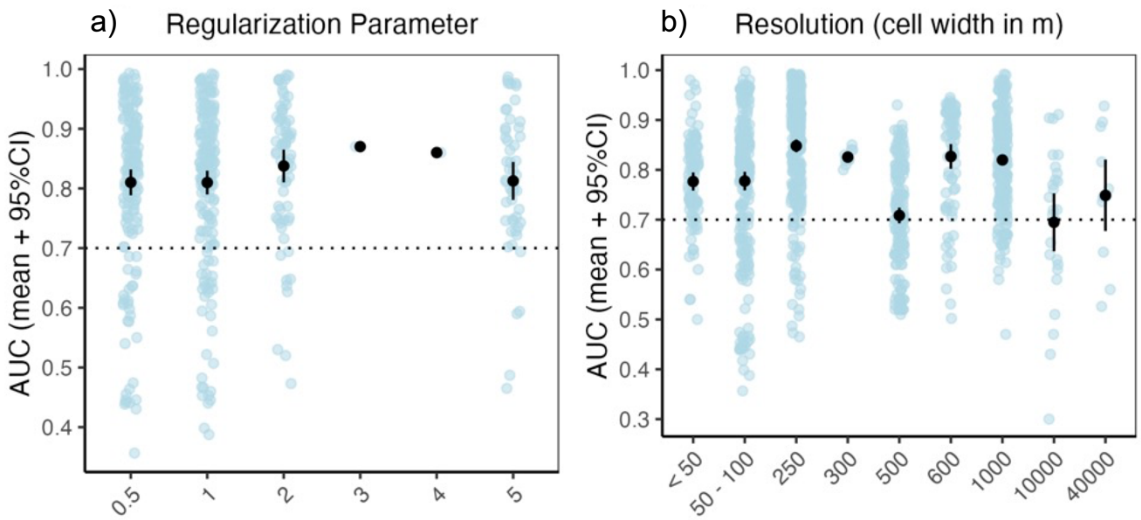


Figure A1.1. Impacts of model decisions not explored in detail in-text on model performance (AUC: area under the receiver operator curve) identified across 39 studies that independently validated SDMs. Black dots represent means, and error bars represent 95% confidence intervals. Blue points represent the independent-AUC scores of individual models from 39 studies included in this review. Definitions of these decisions can be found in Table A1.2.

Table A1.2. Glossary of modelling decision terms used in Chapter 1.

1. Algorithm	
BIOCLIM:	Early species distribution modeling package that relates the bioclimatic envelope (range of tolerances) of a species to climatic interpolations (Booth *et. al.,* 2014).
Ensemble:	An approach where predictions from two or more model algorithms are combined to create a final prediction (Hao *et. al.,* 2020).
Other:	This category captured other seldom-used modeling approaches such as ANUCLIM, Mahalanobis Distance models, Artificial Neural Networks and Bayesian additive regression trees
GAM/GAMMM	A type of generalized linear model in which the response variable depends linearly on unknown smooth functions of predictor variables.
GLM/GLMM	A generalized version of a linear model. Generalization is achieved by using link functions, which allow the magnitude of the variance of each measurement to be a function of its predicted value.
Classification Tree:	Supervised machine learning method that partitions data based on conditional statements to identify environmental conditions a species is most likely to occupy, and results in categorical suitability maps.
Random Forests:	Ensemble machine learning method that uses either classification trees or regression trees algorithms to produce either categorical or numerical species distribution maps.
Maxent:	JavaScript algorithm based on maximum-entropy that aims to predict the probability of suitable habitat occurring in each cell of a raster based on input presences, background points, and environmental variables (Phillips *et. al.,* 2006)

2. Environmental Variables	
Climate Only:	GIS climate variables such as temperature, precipitation, and relative humidity were the only variables provided to the modeling algorithm.
Non-climate Only:	GIS variables other than climate such as land cover, vegetation community, and topography were the only variables provided to the modeling algorithm.
Climate + non-climate:	Models were based on climatic variables as well as other GIS variables.

3. Type of Bias Correction	

Input Thinning:	Input localities of the focal species were thinned to reduce sampling biases. Thinning could either be uniform (*i.e.*, no points within 5 km of one another) or based on a density surface (see below).
Bias File:	Sampling bias was first mapped through production of a sampling density surface (kernel density or similar), using input localities, or GIS information on the distribution of roads, towns, human footprint, distance from these features, or species localities themselves. The selection of background points is then weighted based on this density surface, such that more points are taken in regions of high sampling effort.
Buffered Random:	Random points selection is restricted to buffered areas around input localities, or highly sampled areas (*i.e.,* roads, towns) using GIS software to manipulate the number of randomly developed background points that would occur in that area to correct for sampling bias.
Target-group Background:	Background points are records from other species expected to share the same survey biases as background points (Phillips *et al.,* 2009).
Random:	No bias correction was performed, and background points were randomly developed across the study extent.

4. Study Extent	
Ecoregion or Biome:	Study extent was delineated based on the boundaries of an ecoregion, biome, vegetation district, or other similar biological unit.
Single State or Province:	Study extent was delineated based on the boundaries of a single state or province, or other sub-country political boundary.
Regional:	Study extent was delineated based on the boundaries of multiple states, provinces, or countries.
Range-wide:	Study extent was delineated based on the total known global range of the species.
Local:	User-defined areas within a single state or province, often reflecting the intended application of the study

5. Regularization Parameter	A complexity penalty in the Maxent algorithm. Larger values reduce the number of feature classes available to maxent with which to fit environmental variable response curves and result in increasingly simplistic models (Merow *et. al.,* 2013). The Maxent default is 1.

5. Resolution	Size of the cells in the GIS raster data used to develop models

CHAPTER 2: THE IMPACTS OF STUDY EXTENT ON SPECIES DISTRIBUTION MODELS FOR SIX AMPHIBIANS AT THE EDGE OF THEIR RANGES IN WESTERN CANADA

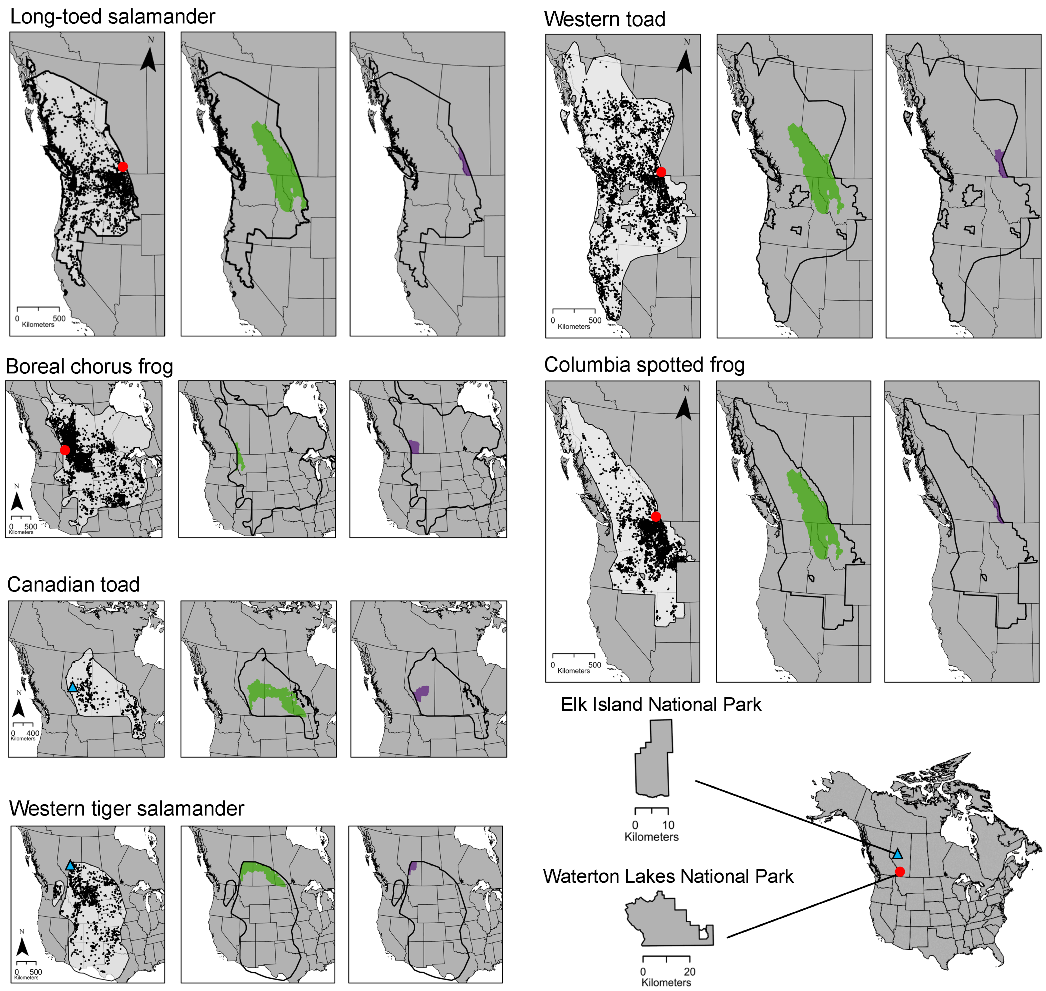


Figure A2.1. Alternative study extents for the six amphibians included in this study. The three study extents considered include: Range-wide (outlined in black), Ecoregion (green), and Political (purple).

Table A2.1. Sixteen climatic variables used for all species distribution modes in Chapter 2, their acronym, and justification for inclusion in this study based on biological relevance and previous inclusion in amphibian SDMs. The species in this study to which the findings of studies in the 'literature justification' column apply are listed in the 'species with known relevance' column. LTSA = long-toed salamander, WETO = western toad, BCFR = boreal chorus frog, CSFR = Columbia spotted frog, CATO = Canadian toad, TISA = tiger salamander.

Variable Definition	**Climate NA Abbreviation**	**Biological Justification**	**Literature Justification**	**Species with known relevance**
Mean annual temperature (°C)	MAT	Determines timing of life stage events (breeding and overwintering: Trumbo *et al*., 2011)	Bradie and Leung, 2017; Seaborn *et al*., 2021	BCFR, CATO, CSFR, LTSA, WETO, TISA
Mean temperature of the warmest month (°C)	MWMT	Eggs and larval stages are exposed to summer heat (Stuart *et al*., 2004)	Roy, 2009	LTSA, TISA, CATO, BCFR
Mean temperature of the coldest month (°C)	MCMT	Affects overwinter survival (Stuart *et al*., 2004)	Keinath *et al*., 2010; Roy, 2009	BCFR, CATO
Autumn (Sept – Nov) mean temperature (°C)	Tave_at	Determines survival of juveniles leaving breeding ponds (Blaustein *et al*., 2001)	Pilliod *et al*., 2015	CSFR
Winter (Dec – Feb) mean temperature (°C)	Tave_wt	Affects overwinter survival and start date of breeding season (Blaustein *et al*., 2001)	Lee-Yaw and Irwin, 2015; Pilliod *et al*., 2015	LTSA, CSFR
Summer (June – Aug) mean temperature (°C)	Tave_sm	Eggs and larval stages are directly exposed to summer heat (Stuart *et al*., 2004)	Lee-Yaw and Irwin, 2015; Pilliod et al., 2015; Keinath *et al*., 2010	LTSA, CSFR, WETO
Spring (March – May) mean temperature (°C)	Tave_sp	Determines start of breeding season (Blaustein *et al*., 2001)	Pilliod *et al*., 2015	CSFR

Summer heat moisture index (MWMT/(MSP/1000))	SHM	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Lee-Yaw and Irwin, 2015	BCFR, CATO, CSFR, LTSA, WETO, TISA
Number of frost-free days	NFFD	Determines start of breeding season (Blaustein *et al*., 2001)	Lee-Yaw and Irwin, 2015; Keinath *et al*., 2010	LTSA, TISA
Mean summer (May – Sep) precipitation (mm)	MSP	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Seaborn *et al*., 2021; Pilliod *et al*., 2015	BCFR, CATO, CSFR, LTSA, WETO, TISA
Mean annual precipitation (mm)	MAP	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Bradie and Leung, 2017; Seaborn *et al*., 2021	BCFR, CATO, CSFR, LTSA, WETO, TISA
Summer (June – Aug) precipitation (mm)	PPT_sm	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Lee-Yaw and Irwin, 2015; Keinath *et al*., 2010	BCFR, CATO, CSFR, LTSA, WETO, TISA
Winter (Dec – Feb) precipitation (mm)	PPT_wt	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Pilliod *et al*., 2015; Keinath *et al*., 2010	CSFR, TISA, WETO
Spring (March – May) precipitation (mm)	PPT_sp	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Pilliod *et al*., 2015; Keinath *et al*., 2010	BCFR, CSFR, WETO
Temperature Difference (difference between MCMT and MWMT, a measure of continentality: °C)	TD	Determines timing of life stage events (breeding and overwintering: Trumbo *et al*., 2011)	Seaborn *et al*., 2021; Keinath *et al*., 2010	BCFR, CATO, CSFR, LTSA, WETO, TISA
Relative Humidity	RH	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Keinath *et al*. 2010	BCFR, CATO, CSFR, LTSA, WETO, TISA

Table A2.2. Counts and percentages of cells predicted to be suitable across three models of different study extents (range, ecoregion and political) for species in Waterton Lakes National Park (number of cells = 479). Agreement across three binary surfaces calculated as the number of cells always predicted to be suitable plus the number of cells never predicted to be suitable divided by the total number cells.

No. of models that predict a cell to be suitable	Long-toed salamander		Western toad		Boreal chorus frog		Columbia spotted frog	
	No. of cells	Percentage	No. of cells	Percentage	No. of cells	Percentage	No. of cells	Percentage
0	41	8.56	0	0	355	74.11	0	0
1	262	54.70	12	2.51	115	24.01	353	73.70
2	167	34.86	118	24.63	8	1.67	74	15.45
3	9	1.88	349	72.86	1	0.21	52	10.86
Agreement	50	10.44	349	72.86	356	74.32	52	10.86

Table A2.3. Counts and percentages of cells predicted to be suitable across three models of different study extents (range, ecoregion and political) for species in Elk Island National Park (number of cells = 185). Agreement across three binary surfaces calculated as the number of cells always predicted to be suitable plus the number of cells never predicted to be suitable divided by the total number cells.

No. of models that predict a cell to be suitable	Canadian toad		Tiger salamander	
	No. of cells	Percentage	No. of cells	Percentage
0	0	0	0	0
1	7	3.78	0	0
2	109	58.92	44	23.78
3	69	37.30	141	76.22
Agreement	69	37.30	141	76.22

Table A3.1. Model settings for 16 species distribution models developed for long-toed salamanders in southwestern Alberta.

Modeling Decisions			*Number of Points*		*Tuning*	
Environmental Variables*	Background Points	Study Extent	Presences	Background	Features‡	Regularization
Climate Only	Random	Range	5329	10000	LQHPT	0.5
		Ecoregion	2096	10000	LQHP	0.5
		Genetic	2462	10000	LQHPT	0.5
		Political	136	5000	LQHPT	1.5
	Target-Group	Range	5329	13868	LQHPT	0.5
		Ecoregion	2096	2733	LQHP	0.5
		Genetic	2462	4590	LQHPT	0.5
		Political	136	215	LQ	0.5
Climate + non-climate	Random	Range	5329	10000	LQHPT	0.5
		Ecoregion	2096	10000	LQHP	0.5
		Genetic	2462	10000	LQHPT	0.5
		Political	136	5000	LQ	0.5
	Target-Group	Range	5329	13868	LQHPT	0.5
		Ecoregion	2096	2733	LQHPT	0.5
		Genetic	2462	4590	LQHPT	0.5
		Political	136	215	L	0.5

* See main text and Table 3.2 for details of variables included in the models

‡ L = linear, Q = quadratic, H = hinge, P = product and T = threshold

Table A3.2. Nineteen climatic and non-climatic variables considered and used to generate species distribution models for the long-toed salamander.

Variable Definition	Abbreviation	Biological Justification	Retained for Model(s)*
Mean annual temperature (°C)	MAT	Determines timing of life stage events (breeding and overwintering: Trumbo *et al*., 2011)	NA
Mean temperature of the warmest month (°C)	MWMT	Eggs and larval stages are exposed to summer heat (Stuart *et al*., 2004)	NA
Mean temperature of the coldest month (°C)	MCMT	Effects overwinter survival (Stuart *et al*., 2004)	NA
Autumn (Sept – Nov) mean temperature (°C)	Tave_at	Determines survival of juveniles leaving breeding ponds (Blaustein *et al*., 2001)	NA
Winter (Dec – Feb) mean temperature (°C)	Tave_wt	Effects overwinter survival and start date of breeding season (Blaustein *et al*., 2001)	NA
Summer (June – Aug) mean temperature (°C)	Tave_sm	Eggs and larval stages are directly exposed to summer heat (Stuart *et al*., 2004)	NA
Spring (March – May) mean temperature (°C)	Tave_sp	Determines start of breeding season (Blaustein *et al*., 2001)	NA
Summer heat moisture index (MWMT/(MSP/1000))	SHM	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Climate Only, Climate + non-climate
Number of frost-free days	NFFD	Determines start of breeding season (Blaustein *et al*., 2001)	Climate Only, Climate + non-climate
Mean summer (May – Sep) precipitation (mm)	MSP	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	NA

Mean annual precipitation (mm)	MAP	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Climate Only, Climate + non-climate
Summer (June – Aug) precipitation (mm)	PPT_sm	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Climate Only, Climate + non-climate
Winter (Dec – Feb) precipitation (mm)	PPT_wt	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	NA
Spring (March – May) precipitation (mm)	PPT_sp	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	NA
Temperature Difference (difference between MCMT and MWMT, a measure of continentality: °C)	TD	Determines timing of life stage events (breeding and overwintering: Trumbo *et al*., 2011)	Climate Only, Climate + non-climate
Relative Humidity	RH	Determines availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015)	Climate Only, Climate + non-climate
Net Primary Productivity	NPP	Net Primary Productivity is positively correlated with amphibian species richness (Qian et al. 2009)	Climate + non-climate
Land Cover	Landcover	Accounts for habitats with high landscape resistance (urban areas & agricultural lands; Goldberg & Waits 2010)	Climate + non-climate
Topographic Wetness	Wetness	Predicts availability and hydroperiod of breeding ponds (Ficetola *et al*., 2015	Climate + non-climate

* Variables retained had pairwise correlation coefficients < 0.7, or > -0.7 and based on Pearson correlation tests (see Table A3.3) were retained. See main text for additional details.

Table A3.3. Pairwise Pearson correlation coefficients *(r)* for the 19 environmental variables considered for use in species distribution models for long-toed salamanders. Bold values indicate values > 0.7, or < -0.7 and pairs of variables for which decisions as to which variable to keep had to be made. Variable abbreviations as per Table A3.2,

	MAP	MAT	MCMT	MSP	MWMT	NFFD	NPP	PPT_sm	PPT_sp	PPT_wt	RH	SHM	Tave_at	Tave_sm	Tave_sp	Tave_wt	TD	Wetness
Landcover	0.231	-0.229	-0.123	0.294	-0.273	-0.206	-0.060	0.282	0.154	0.171	0.184	-0.116	-0.223	-0.280	-0.263	-0.132	-0.167	-0.098
MAP		-0.029	0.237	**0.805**	-0.289	0.212	0.352	0.675	**0.960**	**0.956**	0.666	-0.403	-0.007	-0.275	-0.093	0.213	-0.695	-0.323
MAT			**0.921**	-0.392	**0.927**	**0.925**	0.090	-0.540	0.109	0.150	-0.256	**0.701**	**0.996**	**0.936**	**0.983**	**0.935**	-0.147	0.264
MCMT				-0.196	**0.724**	**0.929**	0.184	-0.375	0.370	0.412	0.006	0.523	**0.930**	**0.732**	**0.858**	**0.998**	-0.510	0.097
MSP					-0.544	-0.176	0.223	**0.969**	0.675	0.611	0.602	-0.557	-0.386	-0.526	-0.407	-0.218	-0.401	-0.303
MWMT						**0.755**	-0.114	-0.650	-0.146	-0.123	-0.498	**0.778**	**0.915**	**0.997**	**0.932**	**0.747**	0.224	0.349
NFFD							0.325	-0.328	0.325	0.367	0.078	0.495	**0.924**	**0.779**	**0.911**	**0.935**	-0.371	0.141
NPP								0.205	0.290	0.358	0.480	-0.150	0.084	-0.068	0.132	0.182	-0.402	-0.051
PPT_sm									0.526	0.448	0.571	-0.616	-0.539	-0.631	-0.534	-0.394	-0.281	-0.294
PPT_sp										**0.962**	0.588	-0.334	0.137	-0.141	0.025	0.346	**-0.704**	-0.315
PPT_wt											0.589	-0.266	0.179	-0.112	0.067	0.389	**-0.735**	-0.290
RH												-0.497	-0.257	-0.475	-0.278	-0.019	-0.629	-0.408
SHM													0.689	**0.775**	**0.709**	0.546	0.231	0.365
Tave_at														**0.922**	**0.969**	**0.943**	-0.174	0.245
Tave_sm															**0.949**	**0.756**	0.208	0.357
Tave_sp																**0.877**	-0.050	0.326
Tave_wt																	-0.479	0.118
TD																		0.298

Table A3.4. Total number of localities of other pond-breeding amphibians (target-group) obtained from GBIF (https://www.gbif.org/; Search date: October 21, 2021) from across the range of the long-toed salamander. Species distribution models based on target-group background used up to the total number of localities depending on study extent (see Table A3.1 for numbers of localities used in the different models).

Scientific Name	Number of Localities (after filtering)
Anaxyrus boreas	3951
Rana luteiventris	3555
Pseudacris regilla	1962
Pseudacris sierra	880
Lithobates catesbeianus	782
Ambystoma gracile	513
Pseudacris maculata	505
Rana cascadae	356
Lithobates sylvaticus	295
Dicamptodon copei	200
Ambystoma mavortium	198
Rana aurora	197
Spea intermontana	134
Rhyacotriton cascadae	91
Rana pretiosa	78
Lithobates clamitans	61
Lithobates pipiens	55
Anaxyrus woodhousii	49
Ambystoma tigrinum	4
Anaxyrus hemiophrys	2
Total:	13,868

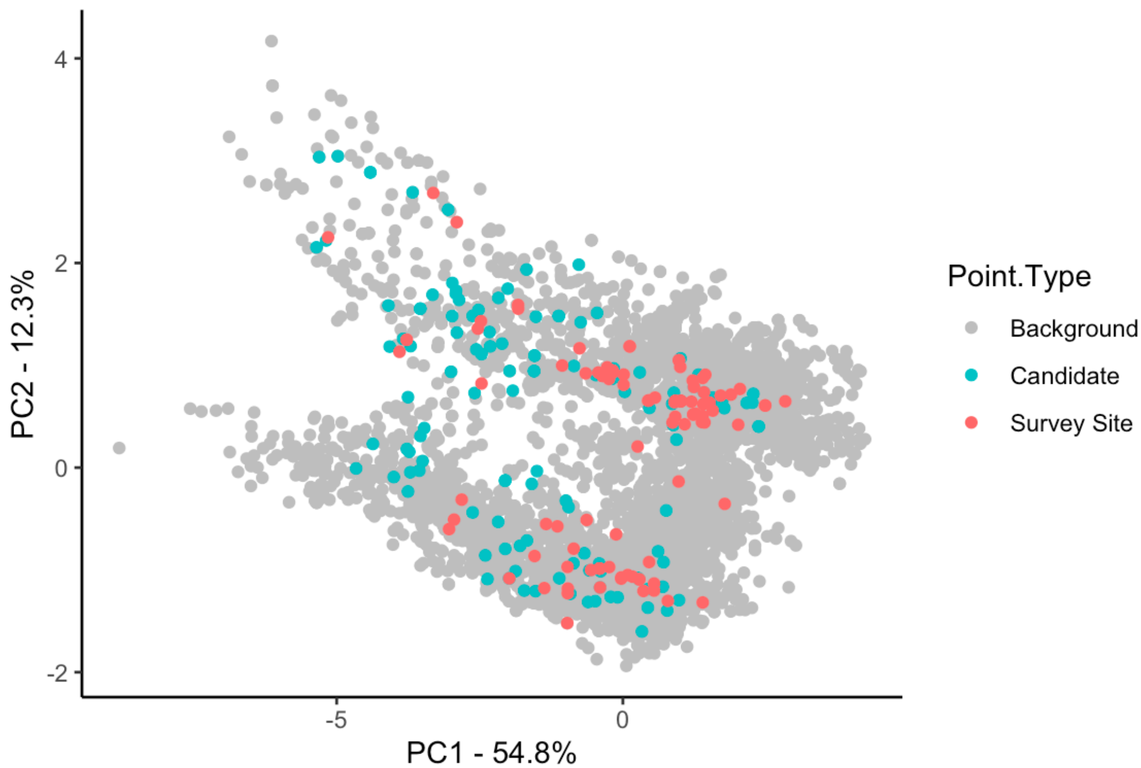


Figure A3.1. Principal component analysis (PCA) of the environmental space represented by random background points and candidate and final long-toed salamander survey sites in southwestern Alberta. Variables included in PCA were the same six used to develop climate-only models (MAP = mean annual precipitation, NFFD = number of frost-free days, PPT_sm = summer precipitation (June-Aug), RH = relative humidity, SHM = summer heat moisture index, and TD = difference between mean warmest and coldest month temperature).

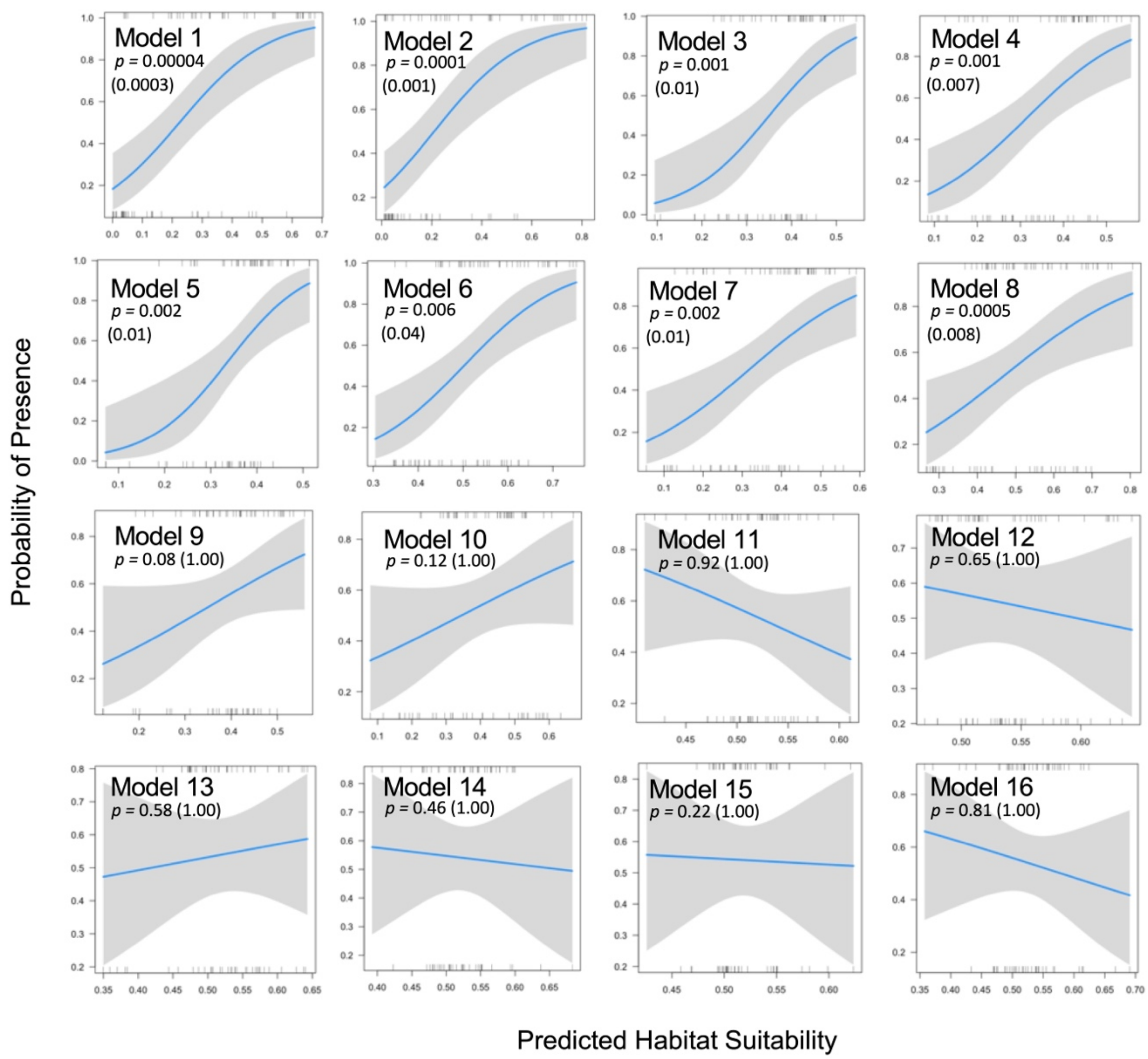
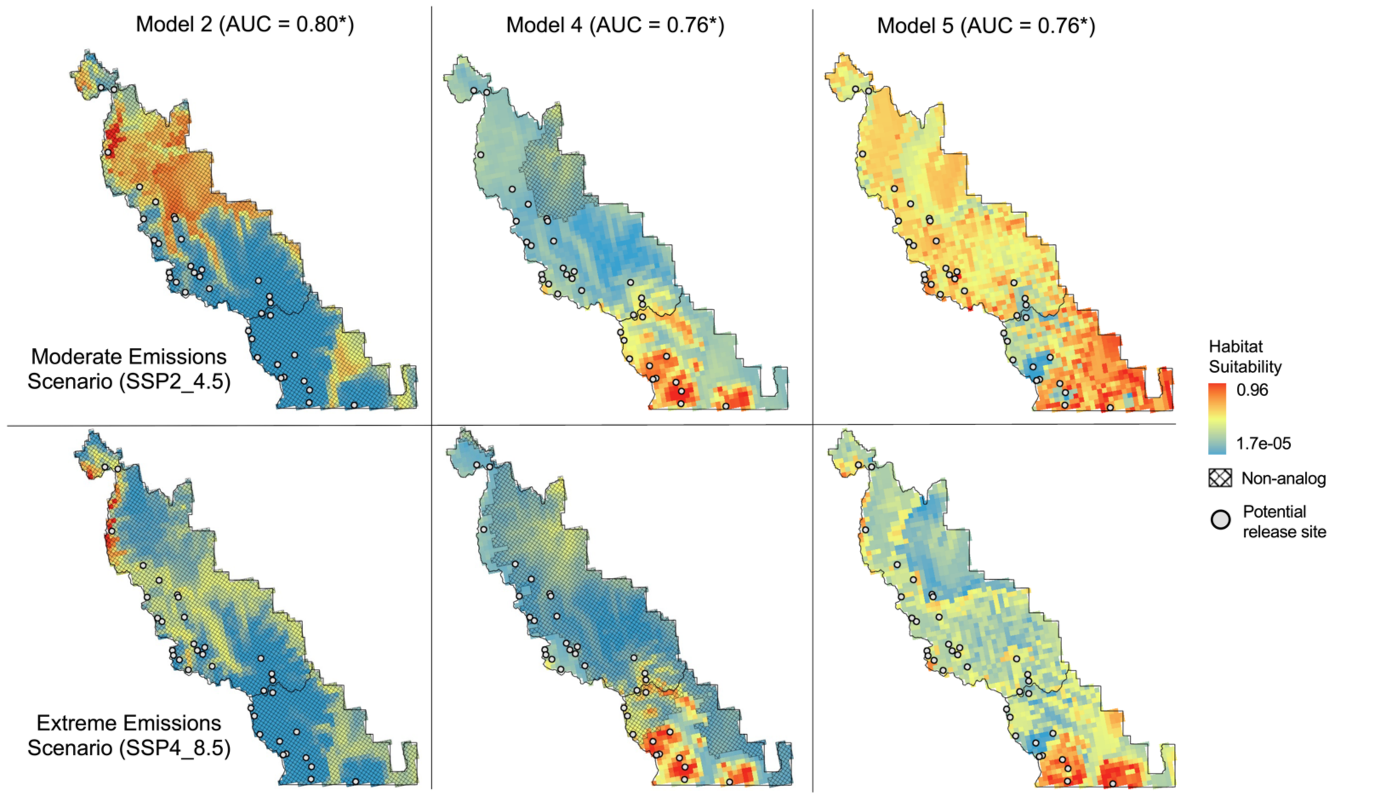


Figure A3.2. Visualizing results from binomial logistic regressions modeling the probability of presence of long-toed salamanders in southwestern Alberta versus predicted habitat suitability from sixteen different Maxent models generated using different environmental variable sets, types of background data, and study extents. Models are numbered based on their independent AUC-score, where Model 1 had the highest and Model 16 had the lowest value (see Table 3.2). Blue lines represent the fitted models, grey shading represent confidence bands of the partial residuals. P-values derived from likelihood ratio tests. Raw *p*-values are listed first, followed by Bonferroni adjusted *pr(>|z|)* values in parentheses. Tick marks at top of each plot represent presences, marks at bottom represent absences. Plots created in R: *visreg* (Breheny and Burchett, 2017).



* Independent-AUC (area under the receiver operator curve) scores based on presence-absence data from 74 sites and models for current time. See Table 3.2 for details.

Figure A3.3. Maps of future long-toed salamander habitat suitability in Waterton Lakes National Park and Castle and Castle Wildland Provincial Parks in Alberta, Canada based on climate-only Maxent models with adequate discrimination and calibration performance to rank potential release sites. Suitable habitat is based on Maxent's logistic output, with darker and warmer colours (red) representing highly suitable habitat. Predictions are shown for Model 2: Climate-only + Random background + Political study extent, Model 4: Climate-only + Random background + Ecoregion study extent, and Model 5: Climate-only + Random background + Range study extent. Black crosshatching represents places where model extrapolation is required (non-analog conditions) based on multivariate environmental similarity surfaces (MESS; Elith *et. al.,* 2010). Grey points represent potential translocation release sites.

Table A3.5. Permutation importance (%) of nine environmental variables across sixteen models of long-toed salamander habitat in southwestern Alberta developed using different environmental variable sets, types of background data, and study extents. Green and red cells represent the variables that were most and least important in each model respectively. See Table A3.2 for variable definitions.

Model	MAP	NFFD	PPT_sm	RH	SHM	TD	Landcover	NPP	Wetness
ClimNonClim+RB+Range	22.38	7.187	13.868	3.441	35.225	8.373	1.775	5.293	2.453
ClimNonClim+RB+ecotone	10.896	8.287	40.345	12.373	16.032	7.35	1.766	0.89	2.057
ClimNonClim+RB+genetic	19.0778	4.933	19.332	5.964	23.201	10.123	3.228	8.25	5.888
ClimNonClim+RB+political	28.298	24.166	4.288	0	31.006	7.078	3.625	1.236	0.301
ClimNonClim+TGB+Range	13.588	8.672	27.56	8.053	12.312	20.489	1.21	5.216	2.896
ClimNonClim+TGB+ecotone	8.479	6.412	24.112	5.02	32.902	14.462	0.062	4.394	4.152
ClimNonClim+TGB+genetic	15.835	11.536	12.019	12.087	21.199	15.638	4.103	5.161	2.418
ClimNonClim+TGB+political	27.809	22.336	29.172	8.637	7.083	0.007	4.451	0.393	0.108
ClimOnly+RB+Range	18.858	9.404	18.559	4.11	42.719	6.347	NA	NA	NA
ClimOnly+RB+ecotone	11.378	9.247	39.376	17.231	15.597	7.168	NA	NA	NA
ClimOnly+RB+genetic	14.402	9.107	29.508	8.866	23.773	14.342	NA	NA	NA
ClimOnly+RB+political	27.474	18.197	9.837	1.778	32.273	10.438	NA	NA	NA
ClimOnly+TGB+Range	9.84	12.833	26.571	10.451	14.766	25.536	NA	NA	NA
ClimOnly+TGB+ecotone	11.482	6.833	26.288	5.085	34.414	15.895	NA	NA	NA
ClimOnly+TGB+genetic	15.686	11.126	16.794	15.109	26.537	14.746	NA	NA	NA
ClimOnly+TGB+political	28.032	6.516	30.557	0	34.438	0.454	NA	NA	NA

Appendix II: Field Survey Protocol

Equipment

- Waders/ Rubber Boots
- Dip nets
- Field datasheet binders
- Field datasheets
- Mechanical pencils
- Viewing containers (Tupperware)
- Decontamination supplies
- Bear Spray
- First Aid Kit
- inReach
- Cellphone
- Field camera
- Measuring tape
- Personal food, medications etc.

Conditions

Surveys will be conducted between 9:00 and 17:00 MDT. Extend if lighting is good and air temperature is warm (>12˚C). Identify **conditions** on datasheet using the descriptions below:

- **GOOD** search conditions = no rain, little to no wind, warm air temps (>15˚C), good lighting;
- **FAIR** search conditions = intermittent rain, low-med wind, moderately warm air temps (12-15˚C), fair lighting;
- **POOR** search conditions = consistent rain, med-high winds, cool air temps (10-12˚C) poor lighting;
- **DON'T SURVEY** during hard rain, if snow covers the surface or edges of ponds, or if air temperature is <10˚C.

Pre-Survey

1. Upon arriving at the edge of a survey pond, mark access point on inReach/ GPS/ MapsCanada App, and record **latitude/longitude in decimal degrees** and **elevation in metres** on datasheet.

2. Fill in the **Date**, **Crew**, and **Site** fields on datasheet. Use the MM/DD/YY date format, initials for Crew, and verbatim site name spelling from **site list**. If site was encountered incidentally, create as meaningful of a name as possible (*e.g. 'Castle Divide Approach'*).

3. Using the field camera take site photos in the following order:
   a. DS: Photo of datasheet (portrait orientation)
   b. POND: Photo of pond as seen from access point. Bottom half of frame should be pond top half should be terrestrial environment
   c. L: Shoreline of pond directly to the left of access point
   d. R: Shoreline of pond directly to the right of access point
   e. B: Terrestrial environment directly behind access point (180˚ from POND photo)

   Check the 'Site photos' box on datasheet when complete.

4. If eDNA sample is to be taken, do this now. See '*ABMI wetland eDNA Field Protocol*'
   Check the '**eDNA'** box on datasheet when complete.

Surveys (4 people hr max, 2 hr visual encounter + 2 hr dip netting)

1. Record **start time** (24 hr, MDT).
2. To commence the visual survey surveyors should begin heading in opposite directions from the access point. Surveyors will each conduct an independent search of the shoreline, meeting back at the access point at the end of the 1 hr (max) search period.
3. Each surveyor should try to search as much of the pond's perimeter as possible, including the shallow water zone (<1.5 m deep), the shoreline (the physical line between land and water), and the shore zone (land within 3 m of the shoreline). Surveyors should walk and/or wade slowly through these three zones in a zigzag course parallel to the shoreline. Stop every 5 m to scan ahead, both along the shore and shoreline. Explore overhanging shoreline edge or debris on moist soil; replace in original position. In the shallow water zone, scan water surface, pond bottom, and submerged debris. Stop every 5 m to search a sample of emergent vegetation for eggs by closely examining stems and gently lifting leaves.
4. Later in the season when no eggs are observed or expected to be disturbed, surveyors may engage in a 2 hr (max) 'blind' dip-netting to search for larvae **after the 1 hr (2 people hr) visual search has concluded**. Larvae can be held temporarily in Tupperware viewing containers for identification and counting. Handling larvae directly should be avoided whenever possible. Record time started dip netting in comments if applicable.

    a. Use dipnets to sweep all microhabitats of pond up to 1.5m in depth.

    b. Use quick 'strikes', bouncing the reinforced flat edge of dip net on pond bottom during sweep (*surveyors will be trained on this strategy*).

    c. Continue at 1–2 m intervals (depending on pond size and habitat complexity) until the entire pond has been sampled, or for a maximum of 1 hr.
5. Record all amphibians seen or heard during visual encounter and dip netting surveys. Record **species**, **life stage**, and **number** of individuals. Record other species observed in the **OTHER** row (amphibians, birds, etc.) Remember: the goal is to complete two independent surveys during the same site visit. To avoid surveyor bias, surveyors should not share any sightings with each other while the survey is being conducted.
6. For larger ponds that cannot be circumnavigated within the time limit, describe the surveyed area in the comments (*e.g., 'SW shoreline'*), and record a track of surveyed area, or mark start and end points using the inReach/GPS/MapsCanada App.
7. Record **end time** (24 hr, MDT).

Post-survey

1. Complete the habitat assessment by circling one of the options for each categorical variable under the **Pond** and **Terrestrial** headings. Use the **Detail** line if none of the available options apply, or if additional notes are required.

    a. For **Disturbance**, provide a short description of type (*see example below*).

    b. For **Dist. forest cover (m)**, use measuring tape to measure shortest distance from water's edge to nearest live tree trunk. If not possible, visually estimate and use 'est.' to indicate this. If nearest cover is > 100 m from edge, write '> 100 m'.

2. Provide details on **Weather** in as much detail as possible. At minimum, note sky conditions (overcast, sun and cloud etc.), wind conditions, estimated temp (˚C) and recent or current precipitation. Be sure to note if weather changed throughout the survey.

3. Check the '**Tissue'** box if tissue samples were collected. Detailed tissue collection notes will be provided in a second notebook.

4. Check the '**Fish observed'** box if applicable. Provide details in **Comments**.

5. Record additional **Comments** (Interesting notes about the site, incidental observations, deviations from this protocol etc.).

6. Ensure all gear is collected before leaving site. You are done!

Example Datasheet

Date: 5/09/2022 Crew: KF, DS

Site: Example Pond

Lat (°N)	Lon (°W)
4 9 . 1 2 2 5 6 7	1 1 4 . 1 4 6 1 5 6

Elev: 1,966 m Start: 1 2 : 4 3 End: 1 3 : 3 6 MDT

Site photos: ✓ eDNA: ✓ Tissue: ✓ Fish observed: ☐

Species	Present	Larvae (#)	Egg Masses (#)	Adults (#)
LTSA	✓		12	
WETO				
CSFR				
OTHER	Possible BCFR heard			

Pond: Details

Water depth: (<1, 1-2, > 2 m)

Water clarity: (low, med, high)

Substrate: (boulder, cobble, pebble, silt..)

Cover type: (grass, willow, CWD, rock…)

% Shore cover: (<25%, 25-50%, 50-75%, >75%)

Permanence: (ephemeral, permanent, semi-permanent)

Terrestrial:

Disturbance (low, med, high) type: Paved road 20 m away

Surrounding hab.: (Forest, grass, shrub…)

Dist. forest cover (m): 120

Canopy type: (Conifer, aspen, mixedwood…)

Conditions (Poor, fair, good, or excellent) & Weather:

Overcast, light breeze ~15˚C, no recent precip.

Comments:

Egg Masses observed along SW shore of pond attached to grass & sedges. Avg 20 eggs per mass. No obvious development