

MODELING SPECIES DISTRIBUTIONS: APPLICATIONS AND METHODS FOR
MARINE BIOGEOGRAPHY AND CONSERVATION

By

Hannah L. Owens

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the
Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the
degree of Doctor of Philosophy.

Chairperson A. Townsend Peterson

Edward O. Wiley

Mark Mort

Wm. Leo Smith

Stephen Egbert

Date Defended: 6th April, 2015

The Dissertation Committee for Hannah L. Owens
certifies that this is the approved version of the following dissertation:

MODELING SPECIES DISTRIBUTIONS: APPLICATIONS AND METHODS FOR
MARINE BIOGEOGRAPHY AND CONSERVATION

Chairperson A. Townsend Peterson

Date approved: 6 April, 2015

Abstract

I employed multidisciplinary approaches for understanding distributions of marine fishes in the present, past, and future, and for considering more broadly the historical role of primary research in policy decisions. In chapter 1, I generated ecological niche models (ENMs) for *Latimeria chalumnae* using two different modeling algorithms; these models anticipated occurrences of *L. chalumnae*'s sister species, *L. menadoensis*, but sample sizes were not large enough to assess the statistical significance of these species' niche similarity. Furthermore, the range of coelacanths may extend beyond their presently known distributions; future exploration may reveal additional populations of these elusive fishes. In chapter 2, I developed a holistic biogeographic history of codfishes in the subfamily Gadinae. I found both ecological niche and geographic distributions of gadine fishes to be largely conservative, but two clades, tomcods and crown cods, included both Pacific and Atlantic species. Divergence in both clades was estimated to have begun in the Pliocene; environmental tolerance reconstructions support temperate origins of both clades, and cyclical Arctic ice formation may have driven divergence. In chapter 3, I addressed the role of general circulation model (GCM) bias as a significant source of uncertainty in estimates of species' potential distributional responses to climate change. ENMs of 15 gadine species calibrated using an observation-based dataset and a dataset derived from the CCSM4 GCM showed areas of disagreement concordant with known GCM biases. Consciousness of bias in GCM data will allow researchers and policy makers to identify areas of particular concern for biodiversity more realistically. Finally, in chapter 4, I explored the evolution of the relationship between policy makers and researchers through the lens of the U.S. Fisheries Service. Building cooperative capacity between these two groups led to a more systematic understanding of the oceans, and thus to success in reducing numbers of overfished stocks.

Acknowledgements

A great many people and institutions have contributed significantly as I have pursued my doctoral degree at the University of Kansas. Just as the evolution of lineages may be shaped by drift just as easily as by global cataclysm, my graduate school experience would not be the same without even the smallest bit of assistance I have accrued over the last six years. Whether it was a quick chat over a pre-seminar cookie or a late-night trouble-shooting session or a quick response to a crisis halfway across the country, no act of collegiality has gone un-appreciated. I have been incredibly fortunate to meet so many intelligent, passionate, and engaging people, and I look forward to a lifetime of continued friendship and collaboration.

First and foremost, I thank my advisors, Ed Wiley and Town Peterson, for their inspiration, guidance, patience, and loose reigns to explore the complicated landscape that is biology. I still remember my first meeting with Ed, when I knew I wanted to do something with climate change and niche modeling, and he suggested I work on cods. I had no idea then what a wild adventure it would be. As a true polymath, Ed has demonstrated that it is possible to have a rich and varied life outside academia, and still succeed in inspiring generations of ichthyologists and systematists. Town's talent for seeing patterns in seemingly disparate data streams and distilling complex ideas into simple, compelling narratives is a model to which I aspire. He has set a high bar with for commitment to open-source science, research capacity building in developing nations, and student support. I also appreciate Town's willingness to support my brief trip to do fieldwork in the Philippines, collecting birds and parasites. Sure, it wasn't part of my dissertation, but as I learned from Town, biological collections are the backbone of what we do, and sometimes it is important to set aside proximate goals for ultimate ones.

Many thanks are also due to the other current and former members of my committee—Sharon Billings, Mark Mort, Steve Egbert, and Leo Smith—for guiding me to be a more well-rounded, successful academic. Thanks to Sharon for pushing me to better understand the geophysical processes at work in ocean systems, and also for hiking with me in Greenland; we never found ptarmigan, but I did find a much better appreciation of the tundra. I thank Mark for reminding me that public outreach and education are vital parts of life at a university, and for his help in building network contacts. Thanks to Steve for knowing when to ask a hard question, and when to ask an easy one that sounds hard. Leo has offered great advice and feedback for professional development, as well as on matters of fish systematics.

My dissertation activities, including laboratory work, a museum visit, several fellowships, an internship, and travel to scientific meetings, were generously funded through grants and awards from several institutions and agencies: the Nation Science Foundation Interdisciplinary Graduate Education and Research Traineeship Climate Change, Humans, and Nature in the Global Environment program at KU; an NSF Doctoral Dissertation Improvement Grant (Award DEB-0068967); the National Center for Atmospheric Research visiting researcher funds; the University of Kansas Ecology and Evolutionary Biology Summer Research Fellowships and Student Travel Awards; the University of Kansas Biodiversity Institute's Panorama Fund; and the International Biogeography Society student travel award. Further ongoing collaborative work at NCAR is funded through an NSF Collaborative Research grant (OCE-1243013).

Thank you to all the individuals who have shared their data through open-source databases, particularly those maintained by the Global Biodiversity Information Facility, Ocean Biogeographic Information System, and the International Nucleotide Sequence Database

Collaboration. I am also much obliged to James Maclaine and Oliver Crimmen, curators of fish at the British Museum of Natural History, for permitting me to sift through their uncurated material to collect additional occurrence data for my cod work, and to Vladimir Loskot at the Zoological Institute at the Russian Academy of Sciences for a copy of the ZIN catalog of Gadiformes. Thank you to Marcus Anders Krag of the Natural History Museum of Denmark for lending tissues; Aimee Stewart in the KUBI Informatics Division for sharing environmental data layers; and T.G. Bornman and colleagues at the African Coelacanth Ecosystem Programme at the South African Institute for Aquatic Biodiversity for sharing coelacanth submersible sighting coordinates.

I am also indebted to many members of the Oceanography Section in the Climate and Global Dynamics Division at the NCAR Earth Systems Laboratory. First and foremost, I thank Joanie Kleypas for sponsoring my work as a visiting research intern at NCAR. Joanie remained engaged through the long process of assembling the manuscript that turned into chapter three of this dissertation. Her interest and verve were inspiring, and I enjoyed our conversations about modeling, climate change, conservation, and being a woman in science just as much as our conversations about where to go birding near Boulder. Susan Barnes, Peter Buss, Enrique Curchitser, and Justin Small gave me abundant, patient assistance as I learned my way around climate data; Gokhan Danabasoglu and Steve Yeager have since included me on additional projects; Mary Pronk made everything run smoothly from an administrative standpoint; and countless other postdocs, ocean modelers, and visiting scientists made me feel welcome during my work at NCAR. Johannes Feddema at KU shared his contacts at NCAR that made this internship possible.

While at KU, I have also had the privilege of participating the Ecological Niche Modeling working group, a catalyst for many stimulating discussions and projects that helped me to push the theoretical envelope of ENM methods. I have been especially fortunate to work with Jorge Soberón, who has been very generous with his time and thoughts when I had questions or was attempting to sort out theoretical aspects of my research; I always enjoy talking with him, whether about research, historical maps, or action movies. I learned the nuts and bolts of niche modeling from Andrés Lira, as well as where to find midnight coffee and donuts in Mexico City. Lindsay Campbell has been a fabulous person to share a cube farm with, and I am endlessly appreciative of her perspective on ideas and problems. Erin Saupe and Cori Myers are two of the hardest workers I know, yet they are still always a pleasure to be around. Narayani and Vijay Barve have provided invaluable computational and programming support, as well as delicious Indian food. Thanks also to Lynnette Dornak, Kate Ingenloff, Joe Manthey, Carlos Yañez-Arenas, Tashitso Anzama, Chris Hensz, Sean Maher, Alberto Jiménez, Mona Papeş, and near-countless visiting scholars from Mexico, Brazil, and beyond for helping shape such a diverse and dynamic group.

The C-CHANGE IGERT program at KU provided me with incredible opportunities to learn and grow as an interdisciplinary climate change scientist. Especial thanks are due to Project Director Joane Nagel, who constantly challenged me to reach beyond my comfort zone while remaining true to myself, no matter how challenging things got; Joane's dedication was the glue that held the program together. Natalie Parker did an incredible job coordinating activities across departments, institutions, and countries. Instructors Joy Ward, Kees van der Veen, Kelly Kindscher, Jay Johnson, Chris Brown, and Greg Cushman and trainees Rebecca Crosthwait, Adam Sundberg, Laci Gerhart-Barley, Ferdouz Cochran, Marie-Odile Fortier, Michelle Mary,

Brian Rumsey, Lindsay Witthaus, and Alexis Reed made a particular impact, but the nature of interdisciplinary work is such that if I were to list everyone associated with the program's activities during my time as a fellow to whom I was grateful, this acknowledgements section would be longer than my dissertation.

I also tremendously benefitted from association with current and past members of the KU Ichthyology Division, where the line between paleontology and neotology blurs. Andy Bentley has been an earnest and dedicated collections manager, providing valuable links to contacts and collections at other institutions. I am also grateful to labmates Matt Davis, Sarah Gibson, Kathryn Mickle, Francine Abe, Matt Girard, and Mike Doosey for everything they have taught me about extant and extinct fishes, as well as their friendship and support, and to my undergraduate research assistant Chelsea Burgess for her dedication to niche modeling fishes. I was also fortunate to be secondarily adopted by the KU Ornithologists, including Mike Andersen, Carl Oliveros, Robin Jones, Luke Campillo, Luke Klicka, Jacob Cooper, Árpí Nyári, Richard Williams, Rob Moyle, and Mark Robbins, who always made me feel welcome among their ranks both in research and over beers, despite my fishy proclivities.

The wider community of faculty and students at KU provided an amazing environment in which to pursue a dissertation, providing help, advice, support, and companionship. Whether as supervisors, instructors, mentors, collaborators, and/or friends, in no particular order, faculty and staff members John Kelly, Linda Trueb, Mark Holder, Rafe Brown, Andrew Short, Paulyn Cartwright, Kirsten Jensen, Maria Orive, Ben Sikes, Kris Krishtalka, Rich Glor, Jen Humphrey, and Kitty Steffens have played a particularly formative role in the last six years of my life. Chris Haufler is so dedicated to supporting EEB graduate students that he called my dad to assure him they were doing everything they could to help when I was detained at the Canadian border on a

mistaken arrest warrant. Jaime Keeler, Lori Schlenker, and Aagje Ashe work incredibly hard to make sure paperwork and countless administrative details are all in order. My extended grad school family, which has made the last six years smoother and more enjoyable, whether through Kraft Klub, dinner parties, movie nights, coffee dates, beer brewing sessions, playing soccer for the Mosasaur Balls, catsitting, airport rides, or adventures both near and far include Kendra and Julius Mojica, Joanna Cielocha, Patrick Monnahan, Jay Palikij, Katie Sparks, Boryana Koseva, Wes Gapp, Curtis Congreve, Cam and Jessi Siler, Charles Linkem, Gabrielle Bassin, Andrea Crowther, Mari Pesek, Marjie Hensz, Az Klymiuk, Brian Finley, Michelle and Daniel Casey, Una Farrell, Alana Alexander, Sally Chang, Michael Pope, Jeet Sukumaran, Raul Diaz, Jeff Cole, Jack Colicchio, Paula Chateau, Caroline Cooper, Jamie Oaks, Stephen Baca, Chan Kinn Onn, Jesse Grismer, Crystal Myer, Ali Fuiten, Anthony Barley, Hayley Kilroy, Taro Eldredge, Steve Roels, Sarah Bodbyl-Roels, Scott Travers, Karen Olson, Laura Breitreutz, Jacob Carter, Kathy Roccaforte, Ashley Brenna, Andrew Mongue, Chris Hamm, Devin Bloom, and Kerry Cobb.

My family has been very supportive and understanding through my time in grad school, even when I disappear to far-flung places and return babbling about obscure animals and esoteric branches of math. My parents, Ken and Emily, nurtured my curiosity and love of the outdoors and museums from a very early age. My sister, Frannie, is an endless font of straight talk and book recommendations. I never would have imagined that summers spent visiting my freshwater Owens grandparents and marine Karraker grandparents would have led here.

Last, I would never be where I am today without the love and support of my husband and best friend, Pete Hosner. His passions for biology and discovery have encouraged me to reach farther in my own research; through our many conversations, he always provides a unique and valuable point of view. We have built a warm and happy home full of critters and eccentric

hobbies, and I have enjoyed every minute of it. Pete also had the terribly good sense to pester me into adopting a cat. Meatball has been a great addition to our little family. I am looking forward to the next step on our journey together. You guys are pretty great.

Table of Contents

Abstract	iii
Acknowledgements	iv
Table of Contents	xi
List of Figures	xii
List of Tables	xiii
List of Appendices	xiv
Introduction	xv
Chapter 1. Predicting suitable environments and potential occurrences for coelacanths (<i>Latimeria</i> spp.)	1
<i>Abstract</i>	2
<i>Introduction</i>	3
<i>Methods</i>	4
<i>Results</i>	9
<i>Discussion</i>	12
<i>Conclusion</i>	14
Chapter 2. Evolution of codfishes (Teleostei: Gadinae) in geographical and ecological space: evidence that physiological limits drove diversification of subarctic fishes	15
<i>Abstract</i>	16
<i>Introduction</i>	18
<i>Methods</i>	21
<i>Results</i>	28
<i>Discussion</i>	31
Chapter 3. Predicting future distributions of gadine fishes using global climate model data: exploring spatial and environmental variable bias	39
<i>Abstract</i>	40
<i>Introduction</i>	41
<i>Methods</i>	43
<i>Results</i>	49
<i>Discussion</i>	52
Chapter 4. Closing the gap between researchers and policymakers: lessons from the history of fisheries management in the United States	56
<i>Abstract</i>	57
<i>Introduction</i>	58
<i>Birth of the U.S. Commission of Fish and Fisheries</i>	59
<i>Modern U.S. National Fisheries Science and Management</i>	63
<i>Conclusion</i>	66
References	67
Appendix I	78
Appendix II	81
Appendix III	87
Appendix IV	90
Appendix V	92

List of Figures

Figure 1.1. Map of areas identified as suitable for the species in model projections for <i>L. chalumnae</i> projected across the Indian Ocean Basin, with a detail map of Sulawesi in Indonesia.	8
Figure 1.2. Maps of areas identified as suitable for the species in model projections for <i>L. chalumnae</i> projected worldwide.	10
Figure 1.3. MESS map for <i>L. chalumnae</i> .	11
Figure 1.4. Exploration of model rule parameters in environmental space for <i>L. chalumnae</i> .	12
Figure 2.1. Results of phylogenetic inference of relationships among codfishes in the subfamily Gadinae (Beast consensus tree).	20
Figure 2.2. Maxent ecological niche modelling results for each species in the subfamily Gadinae.	24
Figure 2.3 Biogeographical range reconstructions using Bayesian inference tree topology under the DIVA-like model with unconstrained dispersal.	30
Figure 2.4. Time-calibrated Bayesian inference phylogeny of the subfamily Gadinae.	33
Figure 3.1. Modeled suitable habitat extent for GCM- and observation-calibrated ecological niche models for <i>Arctogadus glacialis</i> , <i>Boreogadus saida</i> , <i>Gadus morhua</i> , and <i>Melanogrammus aeglefinus</i> through time.	47
Figure 3.2. Modeled suitable habitat extent for GCM and observation-calibrated ecological niche models of <i>Gadus chalcogrammus</i> , <i>Gadus macrocephalus</i> , <i>Gadus ogac</i> , and <i>Micromesistius poutassou</i> through time.	48
Figure 3.3. Map of regional prediction disagreement at each time slice.	50
Figure III.A. Gadine tolerance to maximum sea ice concentration.	88
Figure III.B Biogeographical range reconstructions under alternate models of dispersal, extinction and cladogenesis with unconstrained dispersal.	89
Figure IV.A. Modeled suitable habitat extent for GCM and observation-calibrated ecological niche models of <i>Eleginus gracilis</i> , <i>Merlangius merlangus</i> , <i>Microgadus proximus</i> , <i>Microgadus tomcod</i> , <i>Pollachius pollachius</i> , <i>Pollachius virens</i> , and <i>Trisopterus minutus</i> through time.	91

List of Tables

Table 1.1. Occurrence point statistics.	5
Table 2.1. Results of tests for phylogenetic signal, temperate and Arctic habitat characterization, and ancestral character state reconstructions for the crown cod and tomcod clades in the subfamily Gadinae.	32
Table 3.1. Estimated area (km ²) and mean latitude of potential distributions of gadine species included in this study.	44
Table 3.2. Statistical comparison of models.	49
Table II.A Genetic resources used in analysis.	82
Table II.B. Modeltest results and evolutionary models used in analysis.	84
Table II.C. Sources of raw occurrence data with numbers of records.	85
Table II.D. Range coding for biogeographical range analysis.	86

List of Appendices

Appendix I. Detailed phylogenetic methods.	78
Appendix II. Supplementary tables.	81
Appendix III. Supplementary figures.	87
Appendix IV. Supplementary figures.	90
Appendix V. R code for comparison statistics.	92

Introduction

Species' geographic distributions are the result of both short term and long-term ecological, geological, climatological, and evolutionary processes. Through modeling species' distributions and calibrating and re-calibrating these models to empirical data, it is possible to develop a better understanding of broad-scale macroecological and evolutionary drivers of geographic biodiversity patterns. This approach includes inference of species' present potential distributions, as well as assessment of effects of past and future climate change on species' distributions, and, consequently, their evolutionary histories. It is only by unraveling these complex patterns and processes that we may begin to understand how ongoing climate change will reshape biodiversity. The purpose of this thesis is to incorporate and integrate multiple sources of data, including occurrence data in the form of georeferenced museum collections, field observations, and fisheries records; molecular phylogenetics; and observed and modeled climate data at multiple time scales to develop synthetic understanding of present, past, and future fish distributions.

For the past 20 years, the hypothesis that species' abiotic ecological niche tolerances are evolutionarily conservative has been a driving theme in the growing literature on distributional ecology. Ecological niche models of one species are often able to predict occurrences of closely related species (Peterson et al. 1999), implying that ecological niches of species are evolutionarily conserved among lineages and through evolutionary time. Many subsequent studies have tested this hypothesis, with mixed positive and negative results (Peterson 2011). In the first chapter of my dissertation, I explored the idea of niche conservatism in the critically endangered and enigmatic West Indian Ocean and Indonesian coelacanths, and identified potentially suitable and accessible habitat for these species. This work also sought to understand

environmental variables that associate with potential distributions of coelacanths and potentially to facilitate discovery of additional populations in the future. Although sample sizes of locality information for both species were far too small to assess the degree to which ecological niche in the genus *Latimeria* is conserved statistically, qualitative assessment suggested that I could not reject the hypothesis of niche conservatism in this study.

Recently, ecological niche conservatism has also become a key component of total-evidence macroecological analysis as a potential mechanism underlying patterns of change in species' distributions through history. In the second chapter of my dissertation, I investigated the evolutionary role of ecological niche tolerance limits in shaping species' biogeographic histories using comparative phylogenetic methods. Synthetic approaches like these are quite novel and are in only the early stages of widespread use. However, they have significant potential to unite pattern- and process-based macroecology under multiple lines of evidence; my study was the first application of these unified methods in a marine system. My results showed that ecological niches of gadine codfishes are largely conservative; reconstructed ancestral environmental tolerances supported a temperate origin of the clade, with a single specialization of the Arctic and Polar cod clade to Arctic habitats. This study was the also the first of its kind to incorporate comparative phylogenetic methods to test for evolutionary signal in ecological niche traits; thermal tolerance measures were more phylogenetically significant than other environmental tolerances. This result suggested evolution of thermal tolerance played a key role in the diversification of gadine fishes, possibly in conjunction with Pleistocene climate change. This study is also the first to suggest that cyclical glaciation has not only played a role in structuring populations of gadine codfishes, but also may have been a driver of speciation through cyclical changes in the availability of suitable cod habitat in the Arctic.

In the course of my dissertation work, I also strove to provide improvements to methods and interpretations of ecological modeling (e.g. Manthey et al. 2015, Owens et al. 2013, Saupe et al. 2012) with the goal of better characterizing the assumptions behind these methods and uncertainty in niche inferences. In the third chapter of my dissertation, I investigated the biases inherent in using general circulation model (GCM) data as a source for environmental variables applied in ecological niche modeling. This study began with work I did as an intern at the National Center of Atmospheric Research in Boulder, Colorado, where I processed raw GCM outputs into climate layers that were usable by ENM algorithms. Although many studies have used GCM data to project potential suitable habitat for species in both the past and future, my study compared results of ENMs calibrated using present-day observational data with those calibrated using GCM data. The results indicated strong spatial biases in model projections that coincided with known GCM accuracy biases, which is a source of error that has not been considered in previous studies that have sought to anticipate the effects of climate change on species' future distributions.

Predicting potential future distributions of species is particularly vital for developing local, national, and international policies that will shape approaches to conservation and natural resource management strategies for adaptation and mitigation to ongoing climate change. In the final chapter of my dissertation, I reviewed the history of the relationship between fishery research and resource management policy in the United States as an illustration of the consequences of relationships between researchers as producers and policy makers as consumers. I concluded that, to shape appropriate responses to climate change in the future, a need exists for strong, collaborative relationships between scientists and policy makers.

These four chapters are interdisciplinary, incorporating methods and data from comparative phylogenetics, geographic ecology, climate modeling, and historical analysis. However, they represent necessary parts in the construction of a holistic understanding of the drivers and contexts of species' distributions. Species' distributions are not only a reflection of the optimal combinations of abiotic and biotic factors, but also their evolutionary history. By comprehending better the interplay of these factors, it is my sincere hope that we will better appreciate how humans will shape the future of biodiversity, and appropriately measure our actions to serve as better stewards of our planet.

Chapter 1*

Predicting suitable environments and potential occurrences for coelacanths (*Latimeria* spp.)

*Owens, HL, AC Bentley, AT Peterson. 2012. Predicting suitable environments and potential occurrences for coelacanths (*Latimeria* spp.). *Biodiversity and Conservation* 21: 577-587. doi: 10.1007/s10531-011-0202-1.

Abstract

Extant coelacanths (*Latimeria chalumnae*) were first discovered in the western Indian Ocean in 1938; in 1998, a second species of coelacanth, *Latimeria menadoensis*, was discovered off the north coast of Sulawesi, Indonesia, expanding the known distribution of the genus across the Indian Ocean Basin. This study uses ecological niche modeling techniques to estimate dimensions of realized niches of coelacanths and generate hypotheses for additional sites where they might be found. Coelacanth occurrence information was integrated with environmental and oceanographic data using the Genetic Algorithm for Rule-set Production (GARP) and a maximum entropy algorithm (Maxent). Resulting models were visualized as maps of relative suitability of sites for coelacanths throughout the Indian Ocean, as well as scatterplots of ecological variables. Our findings suggest that the range of coelacanths could extend beyond their presently known distribution and suggests alternative mechanisms for currently observed distributions. Further investigation into these hypotheses could aid in forming a more complete picture of the distributions and populations of members of genus *Latimeria*, which in turn could aid in developing conservation strategies, particularly in the case of *L. menadoensis*.

Introduction

The order Coelacanthiformes, notable as an apparent link between lungfishes and tetrapods, was originally known only from fossils that were more than 80 million years old (Holder et al. 1999). In 1938, the first known specimen of an extant species of coelacanth, *Latimeria chalumnae*, was discovered off the east coast of Africa (Smith 1939). *Latimeria chalumnae* is now known to inhabit a range encompassing the east coast of Africa from Kenya to South Africa, and extending east to Madagascar and the Comoros Islands. In 1997, a second species of coelacanth, *L. menadoensis*, was discovered off the northeast coast of Sulawesi, Indonesia (Erdmann et al. 1998). *Latimeria menadoensis* cannot be differentiated conclusively from its African sister species on the basis of morphology, but the species diverge substantially enough in their genetics that they are recognized as unique lineages (Holder et al. 1999). The IUCN currently lists *L. chalumnae* as critically endangered and *L. menadoensis* as vulnerable (IUCN 2011). Further investigations into the evolutionary relationships, biogeography, life history and appropriate conservation status of this genus are hampered by their rarity in their natural environment and their inaccessibility: coelacanths typically live at depths of 100–300 m in underwater caves on steep, rocky cliffs, emerging only at night to feed (Fricke and Hissmann 2000).

This study seeks to contribute to the understanding of distributions of *Latimeria* by generating hypotheses for additional sites where the environment might be suitable for coelacanths using ecological niche modeling (ENM). The distribution of a species is limited by the interactions between biotic and abiotic factors, as well as dispersal capability—the realized niche of a species (Soberón 2007). ENMs ideally arrive at an estimation of the realized niche of a species after being trained in a geographic area limited to habitats that are accessible to the may

nonetheless be implicitly represented in the model because they strongly correlate with abiotic factors, or disappear because such fine scale interactions disappear in large-scale analysis (Soberón and Nakamura 2009). Projections of such models into other geographic areas are primarily an expression of abiotic niche—the combinations of environmental factors that, based on the model's estimations, are most similar to areas where the species is known to occur.

ENM is a technique that has been implemented successfully for prediction and subsequent field verification of additional localities of known endangered species (Siqueira et al. 2009) and to focus searches for new species (Raxworthy et al. 2004). Such studies often are subject to very low sample sizes, which pose methodological challenges but are still useful, especially if researchers adopt a conservative interpretation of model results as areas similar to those from which a species is known (Pearson et al. 2007). While ENM applications to marine ecosystem studies are not new (e.g. Wiley et al. 2003), this methodology has yet to be applied explicitly to the problem of locating suitable habitat for reclusive marine species. In the present case, to the extent that coelacanth niche characteristics are conservative in their evolution (c.f. Peterson 2011), such models may help in focusing future searches for new populations—or even additional species—of coelacanths.

Methods

Occurrence locality records for *L. chalumnae* were downloaded from the Ocean Biogeographic Information System (OBIS) database via the Global Biodiversity Information Facility (GBIF) biodiversity information portal (<http://www.gbif.org>); data were quality controlled by removing duplicate records, records sharing cells at the resolution of our data layers, and data points which did not fall within the area covered by these layers (e.g. terrestrial

Table 1.1. Occurrence point statistics. Occurrence points localities are followed by the source of the locality: submersible sighting—Sub.; GBIF records—GBIF; or scientific literature—Lit.

Species	Latitude	Longitude	Source	Jackknife Success		Percent Predicted Area		Full Model Suitability	
				GARP	Maxent	GARP	Maxent	GARP	Maxent
<i>Latimeria chalumnae</i>	-27.53	32.72	Sub.	Y	Y	1.42%	0.73%	10	0.53
<i>Latimeria chalumnae</i>	-27.50	32.72	Sub.	Y	Y	1.38%	0.62%	10	0.95
<i>Latimeria chalumnae</i>	-17.32	38.63	GBIF	N	Y	0.71%	4.57%	10	0.49
<i>Latimeria chalumnae</i>	-11.82	43.02	GBIF	N	N	0.34%	0.82%	10	0.24
<i>Latimeria chalumnae</i>	-5.30	39.13	GBIF	Y	Y	1.01%	4.53%	10	0.77
<i>Latimeria chalumnae</i>	-5.26	39.14	GBIF	Y	Y	1.09%	4.53%	10	0.78
<i>Latimeria chalumnae</i>	-5.15	39.18	GBIF	Y	Y	1.06%	2.82%	10	0.78
<i>Latimeria chalumnae</i>	-3.23	40.23	GBIF	N	N	0.99%	0.63%	10	0.77
<i>Latimeria menadoensis</i>	1.62	124.72	Lit.	N/A	N/A	N/A	N/A	3	0.63
<i>Latimeria menadoensis</i>	1.63	124.63	Lit.	N/A	N/A	N/A	N/A	3	0.64

records). This information was supplemented with data from submersible sightings (South African Institute for Aquatic Biodiversity/African Coelacanth Ecosystem Programme/JAGO-Team), which were also reduced to unique localities. Two *L. menadoensis* locality records were taken from Erdmann (1999) and Erdmann et al. (1999). All localities used are listed in Table 1.1.

To limit over-fitting ENMs (Pearson et al. 2007) the number of environmental variables was restricted to 13. Data on world ocean bathymetry were drawn from Amante and Eakins (2009); slope and aspect were calculated from bathymetry in ArcGIS 9.3 (ESRI, Redlands, CA) to incorporate documented preferences of these fish for steep slopes (Fricke and Hissmann 2000). Worldwide sediment thickness estimates, used as a proxy for substrate type, were supplied by the National Geophysical Data Center (Divins 2009). Owing to scarcity of detailed knowledge of definitive ecological preferences of the species, we used datasets with previously demonstrated predictive power for a number of marine fish species (Wiley et al. 2003) summarizing benthic temperature, salinity, dissolved oxygen, percent oxygen saturation, apparent oxygen utilization, phosphate, silicate, nitrate, and chlorophyll which were derived from NOAA’s World Oceanic Atlas 1998 (NOAA 1999). Preliminary ENM runs using parameters as described below were run,

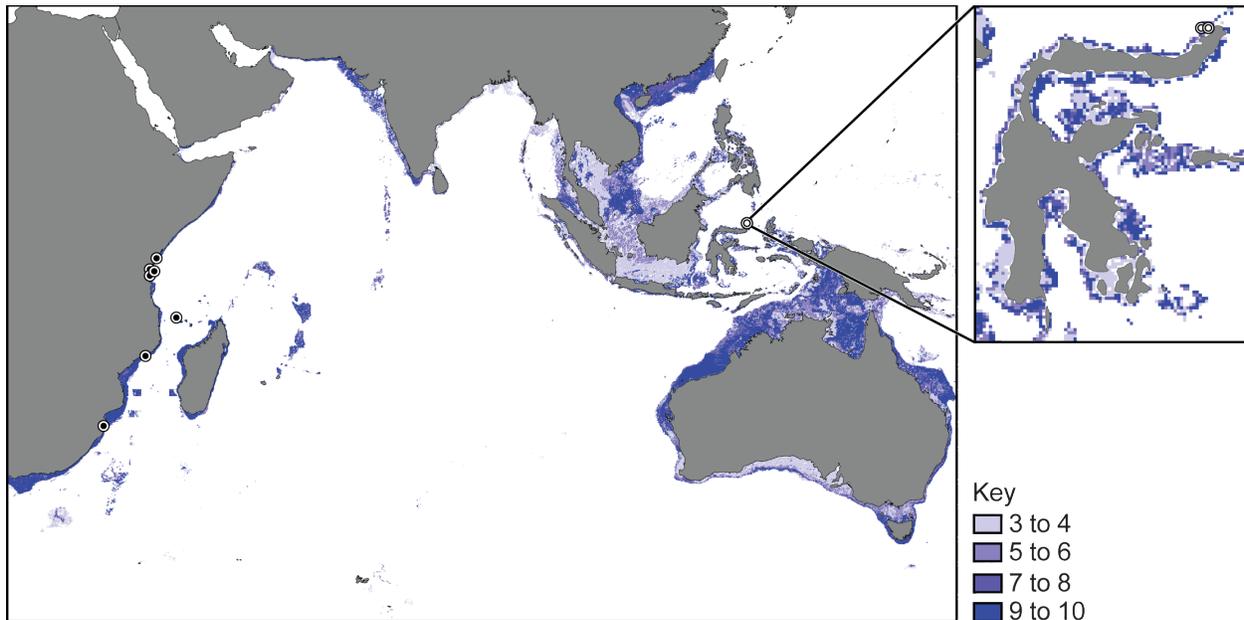
jackknifing environmental variables to investigate the amount of noise introduced by each variable. Suitability scores of each jackknifed model were qualitatively compared to the known range of *L. chalumnae* to assess the degree to which individual variables influenced the model's ability to predict the range of the species.

Latimeria chalumnae occurrence data were integrated with environmental data via two Common ENM algorithms: a maximum entropy algorithm (Maxent; Phillips et al. 2006) and a genetic algorithm (GARP; Stockwell and Peters 1999). Models were trained using a region encompassing the western Indian Ocean from the approximate tip of the Indostanic Peninsula in the northeast to the Cape of Good Hope in the southwest. The GARP algorithm develops a model by choosing a rule iteratively to describe the occurrence-environment relationship, testing the rule's accuracy based on an independent random subset of occurrence points, and consequently evolving, accepting, or rejecting that rule. Desktop GARP (ver. 1.1.6; www.nhm.ku.edu/desktopgarp, Stockwell and Peters 1999) was used to develop these models, deriving 1,000 replicate models with 1,250 pseudoabsence points, a 0.01 convergence limit and a maximum of 1,000 iterations. Best subsets of model replicates were selected using 50% of the occurrence points for intrinsic model testing, with an omission error tolerance of 0%, producing 20 models for that omission tolerance and a commission error tolerance of 50%, resulting in a sample of 10 models (Anderson et al. 2003). Maxent estimates the suitability of each grid cell by generating a probability distribution of maximum entropy from environmental variable layers on that map subject to the constraints of observed presences. Maxent (ver. 3.2.19; www.cs.princeton.edu/~schapire/maxent, Phillips et al. 2006) models were developed using 10,000 background points, a maximum of 1,000 iterations, a convergence threshold of 0.00001, and a random 50% of the data points set aside for intrinsic testing. Maxent generates an

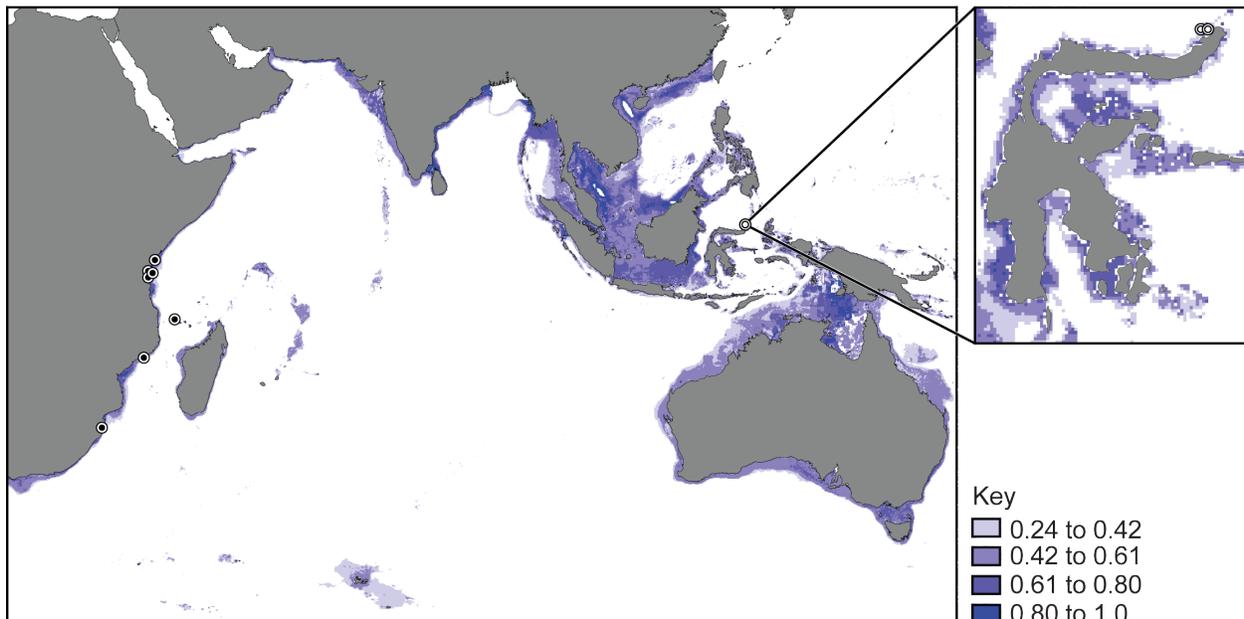
additional layer for “clamping” the model (i.e., extending the terminal values of suitability beyond the limits of environmental variables represented in the calibration region), incorporating combinations of environmental variables that do not exist in the training region in predictions that tends to lead to over-prediction; no clamping was tolerated in generating Maxent ecological suitability maps—cells with nonzero clamping scores were removed from the final projection. Maxent also calculates a multivariate environmental suitability surface (MESS) map indicating areas where environmental variables occur outside the range of values in the training region; ENM suitability projections in these regions are unreliable (Elith et al. 2010). ENMs were not developed for *L. menadoensis* owing to paucity of locality information available for this species.

As a consequence of the relatively small number of available locality records for *L. chalumnae*, typical independent model validation approaches involving partitioning the data into training and testing subsets were inappropriate; instead, we used a jackknife approach to validate ENM that is specifically designed for situations of small sample size (Pearson et al. 2007). In this method, independent GARP and Maxent models were generated iteratively, excluding one locality in each turn. The lowest suitability score of a presence point, or lowest presence threshold (LPT), for each model was then used to determine areas of predicted presence. The proportion of the training area predicted as present and the failure or success of the model to predict jackknifed points were then used to calculate the probability of the observed degree of coincidence between independent test data and predicted areas of suitability for *L. chalumnae*, as described by Pearson et al. (2007).

Figure 1.1. Map of areas identified as suitable for the species in model projections for *L. chalumnae* projected across the Indian Ocean Basin, with a detail map of Sulawesi in Indonesia. *Latimeria chalumnae* localities are indicated by a filled dot (●) and *L. menadoensis* localities are indicated by a hollow dot (○). Suitability scores are represented by shades of blue, with darker shades indicating greater suitability. A rectangle of missing data exists in the East China Sea extending northeast from Taiwan up through the Ryuku Islands.



GARP Suitable Habitat



Maxent Suitable Habitat

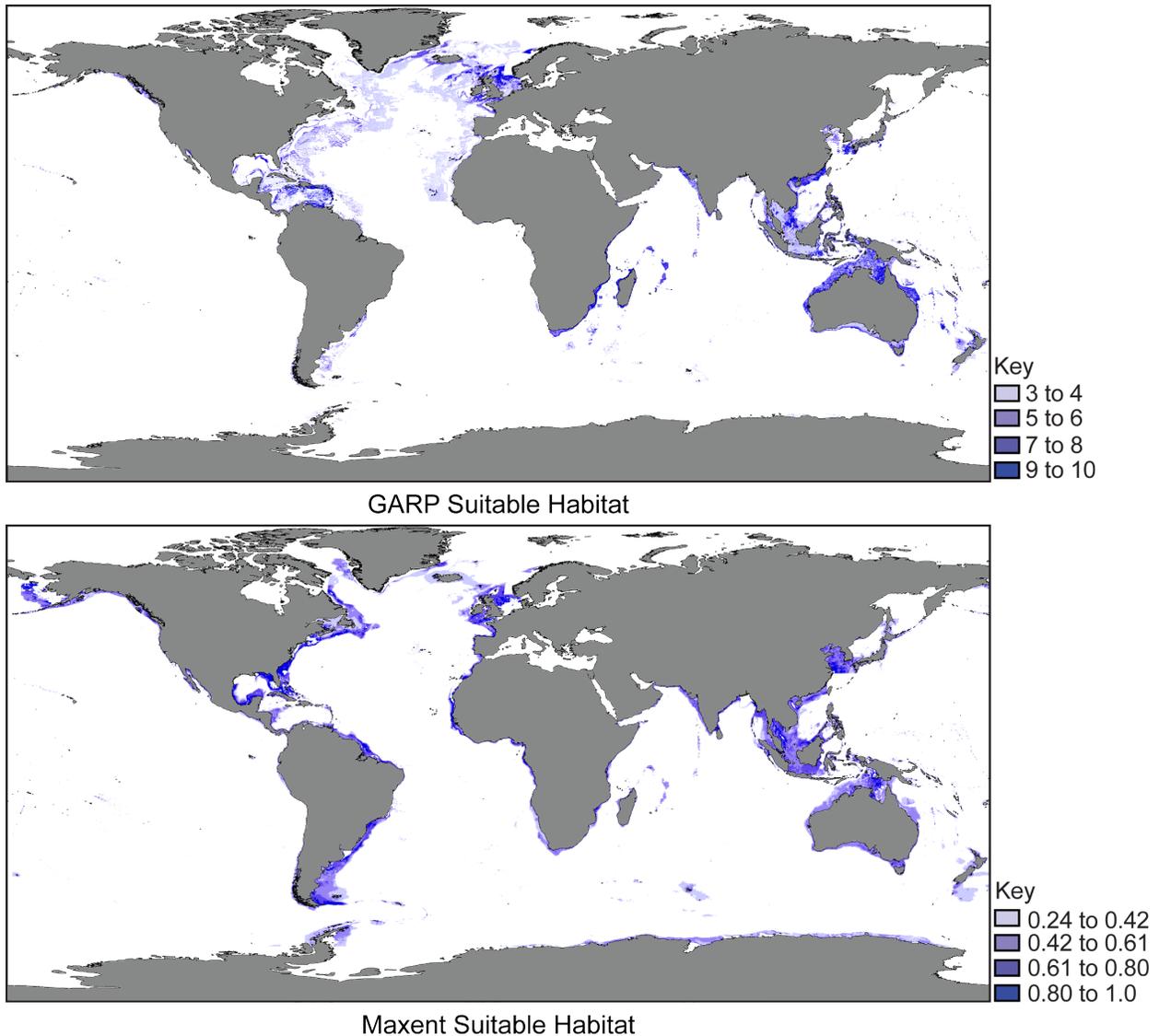
To provide a basis for comparison between our ENMs and previously collected ecological information, a coarse-resolution exploration of model rule parameters in environmental space for *L. chalumnae* was visualized by taking a random sample of 5,000 points from the training region. At each point, the abiotic variable values and the Maxent and GARP suitability scores were extracted, and scatterplot visualizations of the niche of these fish developed. Two scatterplots were generated for each model using environmental variables measured by Fricke and Hissmann (2000) describing the ecology of *L. chalumnae* in Jesser Canyon off the coast of South Africa: ocean depth versus salinity and temperature versus dissolved oxygen concentration. Each point represented a combination of variables that exists in the environment and was classified as unsuitable, suitable, or intermediate. For GARP models, points in which none of the 10 best models predicted potential for coelacanth occurrence were categorized as unsuitable, points in which all of the best models predicted potential for coelacanth occurrence were categorized as suitable, and all other points were categorized as representing intermediate suitability. For the Maxent model, suitability thresholds were chosen to yield the same percentage of each classification as the GARP model—for example, if 95% of the points were unsuitable according to GARP suitability scores, the points with the lowest 95% of Maxent suitability scores were also characterized as unsuitable.

Results

Qualitative comparison of preliminary ENM runs in which environmental variables were jackknifed with the known range of *L. chalumnae* indicates that none of the variables incorporated introduced a disproportionate amount of noise into model results. Predictions of the potential distribution of *L. chalumnae* in the western Indian Ocean as measured by the Pearson

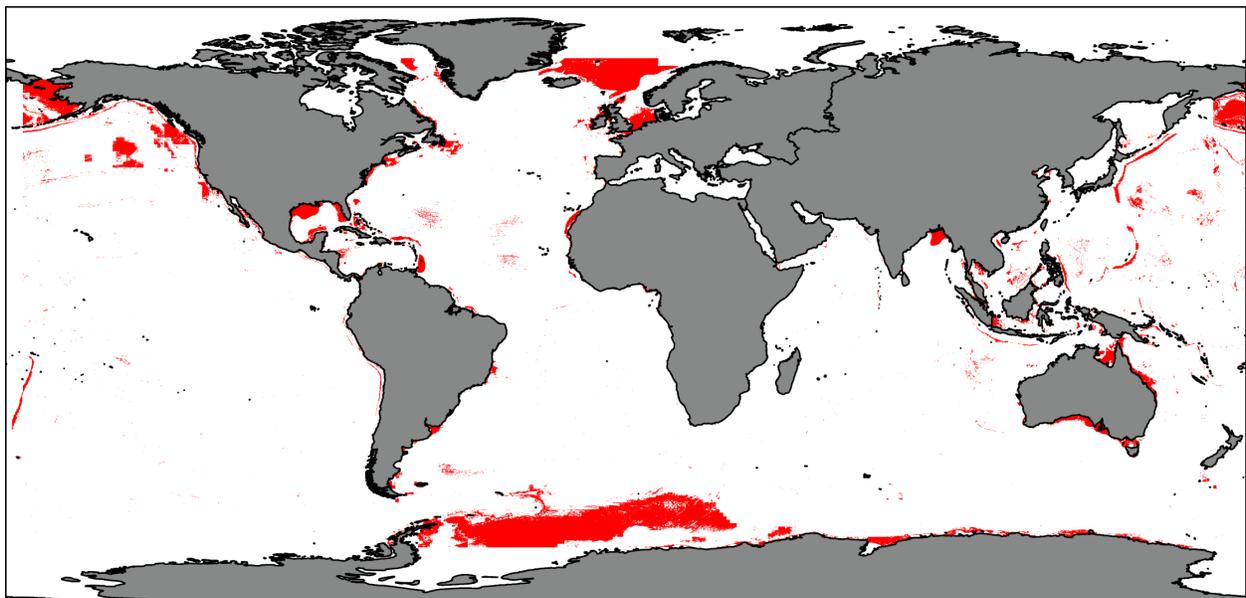
jackknife-based test procedure were significantly better than random expectations ($P > 0.01$) for both GARP and Maxent (Table 1.1). All 10 best subset GARP models trained using *L. chalumnae* occurrence points predicted habitat suitability for all *L. chalumnae* occurrences, and *L. menadoensis* occurrences were predicted by 3 of the 10 models. Maxent-estimated suitability at occurrence points for *L. chalumnae* ranged from 0.24 to 0.78, while suitability for *L. menadoensis* ranged from 0.63 to 0.64.

Figure 1.2. Maps of areas identified as suitable for the species in model projections for *L. chalumnae* projected worldwide.



When all *L. chalumnae* occurrence points were pooled to generate models identifying areas of suitable habitat across the Indian Ocean and western Pacific Ocean, these models identified potentially suitable sites scattered over the known range of the species where it has not as-yet been recorded (Figure 1.1). These areas include most of the east coast of sub-Saharan Africa, as well as along the Mascarene Plateau, and the coasts of India, Indonesia, the Philippines, and northern Australia. Worldwide projections of suitable habitat (Fig.1.2) also indicate areas of suitability far from known coelacanth localities, including off the coasts of Argentina and the Lesser Antilles. Environmental differences between the training region and the worldwide projections are expressed in the form of a Multivariate Environmental Similarity Surface (MESS) map (Fig. 1.3).

Figure 1.3. MESS map for *L. chalumnae*. Cells shown in red indicate areas where at least one environmental variable value occurs outside the range of values in the training region.



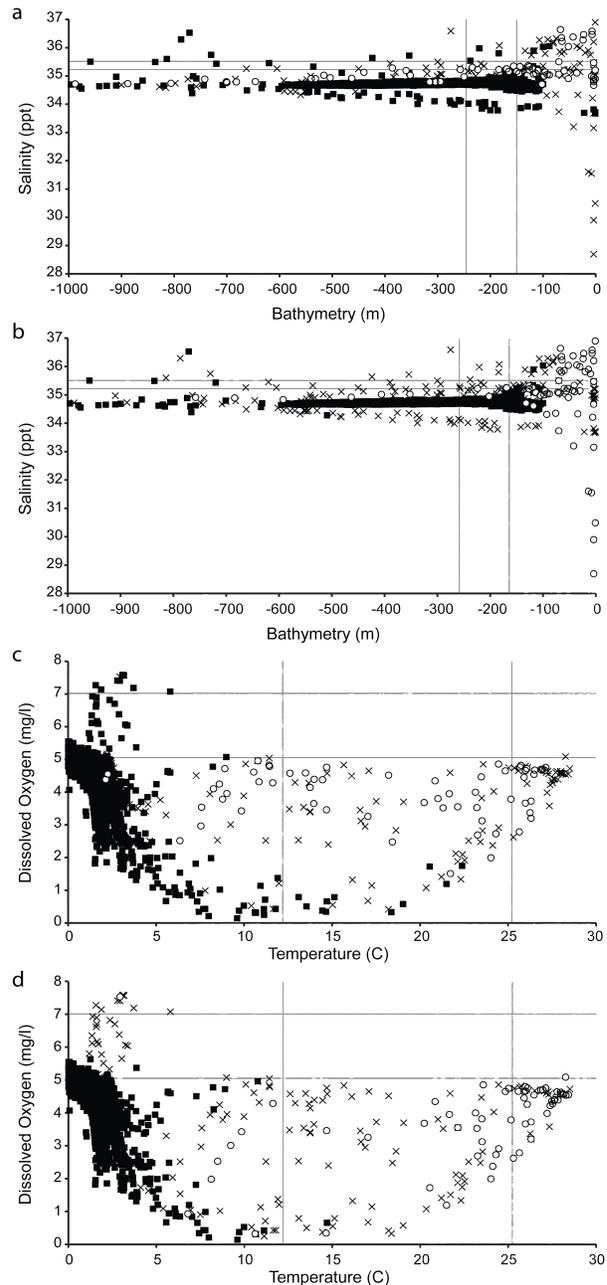
Ecological suitability maps were similar for both the GARP and Maxent models; however, some differences are notable in the suitability ranges of bathymetry, temperature, dissolved oxygen concentration, and salinity between the two models (Fig. 1.4). Perhaps most notable is the disagreement between Maxent and GARP as to whether low-temperature high-

oxygen environments were unsuitable or merely unlikely habitat for *L. chalumnae*. Combinations of field measurements of these variables reported in Fricke and Hissmann's (2000) study of coelacanth ecology were not well represented in the sample (six points from bathymetry versus salinity plots, none from temperature versus dissolved oxygen).

Discussion

Species in general occur at sites that satisfy three sets of considerations (Soberón & Peterson 2005; Pulliam 2000). First, abiotic conditions must be suitable—these physical characteristics of environments are the focus of the analyses in this paper. Second, the biotic realm must be appropriate (i.e. the correct suite of positive interactor species present, and negative interactor species absent)—in this paper, because detailed information on biotic interactions is lacking, we implicitly assume that biotic dimensions will have abiotic correlates. Finally, a site must be accessible for dispersal to

Figure 1.4. Exploration of model rule parameters in environmental space for *L. chalumnae*. X's: environmental combinations at intermediate levels of predicted suitability; black squares: unsuitable variable combinations; white circles: highly suitable variable combinations. Gray lines: observed ecological variable ranges experienced by *L. chalumnae* in Jesser Canyon, South Africa (Fricke and Hissmann 2000). a, b Bathymetry (m) versus salinity (ppt); a GARP, b Maxent. c, d Temperature versus dissolved oxygen concentration; c GARP, d Maxent.



and colonization by the species: sites that are readily accessible will likely be inhabited by populations of the same species, while less accessible sites will either be uninhabited or perhaps inhabited by related species.

Owing to the small sample size of occurrence points used to generate ecological niche models, it would be unreasonable to expect these models to describe the complete realized niche of *L. chalumnae*; however, as they do describe dimensions of ecological space in which the species is known to occur, they are still of some utility. The models generated herein predict areas of suitable habitat well beyond the known localities of the two coelacanth species. Among these areas are several previously postulated as harboring coelacanths (although sightings remain unconfirmed), including locations off the northern coast of Madagascar and the islands of Mwali and Maore in the Comoros (Stobbs 2002). Taking into account projection uncertainty as expressed by the MESS map in Figure 1.3, additional areas in the western Indian Ocean that show promise as potential coelacanth localities include parts of the Seychelles and the Mascarene archipelago, as well as the Malay Archipelago. Further investigation of these localities, informed by regional geology (i.e. the presence of caves) may provide insight into biotic and accessibility factors that influence the range of the coelacanths. Additional information gleaned by these investigations could contribute to a more complete picture of how best to conserve the rare *Latimeria* species.

There has been a great deal of speculation in the literature as to the nature of the disjunct distribution of the genus *Latimeria* in the Indian Ocean. Springer (1999) hypothesized that the genus had been continuously distributed off the shores of Africa and Eurasia, but that the collision of India with Eurasia had led to a vicariance event when the major rivers of India began depositing large amounts of silt in the Indian Ocean, rendering those areas of habitat unsuitable.

Our findings lend support to Springer’s hypothesis—suitable coelacanth habitat extends almost continuously along the coasts of the northern rim of the Indian Ocean, broken up by large areas of unsuitable habitat at the mouths of the Ganges and Indus Rivers.

When one compares the performance of GARP and Maxent models in predicting both the training species, *L. chalumnae*, and the second species, *L. menadoensis*, it becomes apparent that these algorithms do not behave entirely similarly. All ten GARP models predicted training points to be within suitable habitat, whereas only three predicted suitable habitat for *L. menadoensis*; in contrast, Maxent gave a wide range of suitability scores at training points (from 0.24 to 0.95), with the *L. menadoensis* points falling squarely into the suitability range (at 0.63–0.64). Maxent was able to predict one more jackknife point successfully than GARP, which echoes a pattern from previous studies (Pearson et al. 2007). Unfortunately, the occurrence sample size for *L. menadoensis* is too small to test niche conservatism conclusively in the group, or the differing abilities of the algorithms to predict sister species.

Conclusion

Coelacanths are rare and reclusive fish about which little is known, so no definitive idea of the full extent of the range exists for either *L. chalumnae* or *L. menadoensis*. Ecological niche model predictions of suitable areas based on occurrence data for *L. chalumnae* through the oceans of the world, combined with rigorous efforts to ground-truth the models, may prove useful in searches for new populations of coelacanths.

Chapter 2*

Evolution of codfishes (Teleostei: Gadinae) in geographical and ecological space:
evidence that physiological limits drove diversification of subarctic fishes

*Owens, HL. 2015. Evolution of codfishes (Teleostei: Gadinae) in geographical and ecological space: evidence that physiological limits drove diversification of subarctic fishes. *Journal of Biogeography*. DOI: 10.1111/jbi.12483.

Abstract

The aim of this study was to develop a holistic biogeographical history of codfishes in the subfamily Gadinae based on historical relationships, ecological niche, and evolution of physiological tolerances. Two alternative diversification scenarios were tested in two co-distributed, Northern Hemisphere clades: (1) clade ancestors were temperate, and environmental niche has been conserved over evolutionary time, implying that speciation was driven by vicariance associated with ice sheet formation; and (2) clade ancestors were Arctic, and species convergently adapted to temperate environmental conditions, implying that speciation was driven by repeated adaptation to temperate environments. Fifty-five new sequences of four genes from 23 tissue samples were combined with 10 GenBank sequences to generate a time-calibrated phylogenetic hypothesis. Combining the phylogeny with information on species' ecological niche tolerances inferred from correlational models, I reconstructed ancestral environmental tolerances of each of the focal clades. These results were combined with Bayesian area-based biogeographical analysis and regional palaeoclimatic history to develop a holistic biogeographical history of Gadinae. Of 18 environmental variables describing species' tolerances to salinity, temperature, sea ice concentration, and mixed layer depth, only mean, maximum and minimum sea bottom temperature, and mean and minimum sea surface temperature showed phylogenetic signal across Gadinae. Both ecological niche and geographical distributions of gadine fishes are largely conservative, but two clades contain both Pacific and Atlantic species. Focal clade divergence time estimates suggest a Pliocene origin for both, with further Pleistocene divergence. Reconstructed ancestral environmental tolerances of crown cods and tomcods support a temperate origin of both groups. The timing of diversification of these two clades and the intolerance of temperate species to sea ice suggest that cyclical Arctic ice

formation drove divergence. Future sea ice reduction may have dramatic consequences for distributions and persistence of commercially important species, when currently allopatric temperate species come into secondary contact.

Introduction

In biogeography, much is made of the ability of geophysical barriers to isolate populations, leading to allopatric speciation (Coyne & Orr, 2004). However, effects of geophysical barriers may be closely tied to environmental gradients coincident with such barriers, to the extent that species' environmental tolerance limits may be important drivers for the diversification of lineages (e.g. elevational gradients, Graham et al. 2004; bathymetric gradients, Quattrini et al. 2013). Fishes of the subfamily Gadinae present an ideal group within which to investigate interconnected roles of geophysical barriers and physiological tolerance in allopatric diversification. Gadines are broadly distributed across Northern Hemisphere oceans, mostly in sub-Arctic to temperate waters; several Pacific species are separated from their putative closest relatives in the Atlantic by the Arctic Ocean (Coulson et al. 2006). However, *Boreogadus* and *Arctogadus* are euryhaline and eurythermal (Cohen et al. 1990). Both genera are associated with Arctic sea ice, and *Boreogadus* can also be found along Arctic coasts and in brackish lagoons and estuaries (Cohen et al. 1990). The ability of the Arctic Ocean to serve as both a barrier and an endemic habitat to gadine fishes makes its role in gadine biogeographical history uncertain – does it serve primarily as a geophysical barrier to temperate species, or as an environmental barrier that may have been more permeable in a warmer past?

Several gadine fishes, including Atlantic and Pacific cod and Alaska pollock (*Gadus morhua*, *Gadus macrocephalus*, and *Gadus chalcogrammus*), are commercially important: in 2011, 5.7 million tonnes of gadine fish were harvested from marine fisheries (FAO 2013). Additionally, gadine fishes are a key trophic link in northern ocean ecosystems, both as consumers and as prey (Link & Garrison 2002; Gradinger & Bluhm 2004). As a result of their ecological and commercial importance, many aspects of gadine biology have seen detailed study.

Resolving gadine phylogenetic relationships and inferring their biogeographical history has been an area of particular ongoing effort (e.g. Carr et al. 1999; Pogson & Mesa 2004; Coulson et al. 2006; Roa-Varón & Ortí 2009). While numerous population-level studies of gadine species' biogeography have been conducted (e.g. Bigg et al. 2008; Canino et al. 2010), species-level phylogenetic studies have differed in taxon sampling and gene locus selection, leading to uncertainty over placement of Arctic species and their biogeographical history. The group is thought to have originated in the North Atlantic (Svetovidov, 1948; Carr et al. 1999), with lineages dispersing across the Arctic to the Pacific twice (Pogson & Mesa, 2004; Coulson et al. 2006) or four times (Carr et al. 1999).

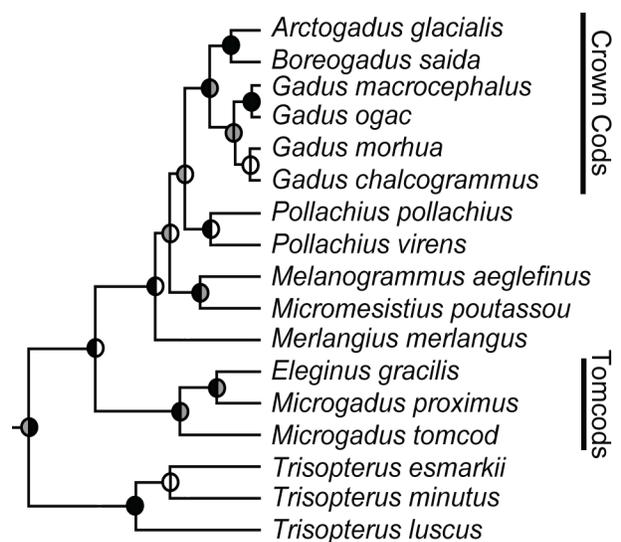
These biogeographical inferences were based on phylogenetic patterns in the context of changing ocean currents, connectivity, temperature and salinity. However, no study has used statistical methods to reconstruct the biogeographical history of Gadinae or explicitly considered how evolution of tolerances with respect to coarse-grained environmental characteristics (Grinnellian ecological niches; Grinnell, 1917) have contributed to moulding current and past distributional patterns. Given the dramatic changes in both Arctic and temperate oceans of the Northern Hemisphere, beginning with global cooling in the Miocene, the physiological limits of species were probably important factors in the dispersal and vicariance of codfishes.

A promising approach to investigating the effects of complex climatic factors on biogeographical history of organisms is through uniting historical, pattern-based analysis methods with ecological, process-based methods. One particularly fruitful line of enquiry has been to assess species' tolerances to abiotic environmental factors via correlational ecological niche models and, using comparative phylogenetic methods, to reconstruct the evolutionary history of environmental tolerances of a clade. These analyses have provided insights into how

changes in ecology and climate may have driven distributional patterns and lineage divergence (Yesson & Culham 2006a; Evans et al. 2009). The extension of these methods using range-based biogeographical reconstructions adds an explicit geographical dimension to such analyses, especially because the changeable nature of Earth's climate means that suitable past environments may have occupied geographical distributions differing from those in which they are currently found. While such techniques have been used to investigate the role of niche evolution in the biogeographical history of terrestrial plants (e.g. Yesson & Culham 2006a,b; Evans et al. 2009; Smith & Donoghue 2010), frogs (Graham et al. 2004), salamanders (Vieites et al. 2009), birds (Nyári & Reddy 2013) and other groups, they have yet to be implemented in a marine fish system.

In this study, I tested two alternative scenarios thought to have produced current amphiboreal distribution patterns in two key gadine clades, the crown cods ((*Arctogadus* + *Boreogadus*) *Gadus*) and tomcods ((*Microgadus proximus* + *Eleginus gracilis*) *M. tomcod*) (Fig. 2.1): (1) the ancestor of each key clade was temperate; (2) the ancestor of each key clade was Arctic. If the ancestor of a key clade was temperate, dispersal across the Arctic Ocean may have been possible only episodically. In this case, progenitors of modern species may have tracked temperate

Figure 2.1. Results of phylogenetic inference of relationships among codfishes in the subfamily Gadinae (Beast consensus tree). Circles at nodes indicate support: left half indicates Bayesian inference posterior probabilities, right half indicates maximum likelihood bootstrap support. Black shading denotes posterior probability of 1.0/100% bootstrap support, grey shading probability > 0.7/ > 70% support, white shading probability < 0.7/ < 70% support. Outgroups omitted for simplicity, branch lengths not to scale.



climatic conditions back and forth from the Atlantic to the Pacific, with glacial periods increasing sea ice extent in Arctic waters and isolating populations. While this model of diversification through periodic isolation and reconnection (Haffer 1969), commonly referred to as the glacial speciation pump, has been invoked in terrestrial (Weir & Schluter 2004), freshwater (April et al. 2013; Houston et al. 2014), and marine systems (Bigg et al. 2008; Maggs et al. 2008; Canino et al. 2010), it has yet to be applied in the context of diversification above the species level in gadine fishes. Alternatively, if a clade's ancestor was Arctic, a rapid Arctic radiation followed by dispersal southwards to current distributional areas may have been more likely. In this situation, increased sea ice could have served as a bridge between the Atlantic and Pacific, instead of a barrier. To explore these competing hypotheses, I reconstructed niche tolerances of tomcod and crown cod ancestors, and compared reconstructions to modern temperate and Arctic conditions. These results were compared with biogeographical range reconstructions benefitting from recent advances in biogeographical model testing (Matzke 2013a) to develop a fully integrated pattern and process analysis of the biogeographical history of gadine fishes.

Methods

Phylogeny of Gadinae. To estimate the phylogeny of the Gadinae for this study, I used sequences from 17 of the 22 currently recognized species in the clade (Nelson 2006). *Lota lota*, *Molva molva* and *Gaidropsarus ensis* were employed as outgroups based on their phylogenetic proximity to Gadinae (Roa-Varón & Ortí 2009). Because *Theragra finnmarchica* and *Arctogadus borisovi* are considered junior synonyms of *Gadus chalcogrammus* and *Arctogadus glacialis*, respectively (Jordan et al. 2003; Byrkjedal et al. 2008), they were excluded from

sampling. *Eleginus navaga*, *Gadiculus argenteus* and *Micromesistius australis* were not included because of lack of available tissue or sequence data in GenBank.

DNA was extracted from muscle tissue stored at $-20\text{ }^{\circ}\text{C}$ with a non-commercial guanidine thiocyanate method (Esselstyn et al. 2008). Two mitochondrial loci, *12S* (Li & Ortí 2007) and *CYTB* (Coulson et al. 2006), and two nuclear loci, *ZICI* (Li et al. 2007) and *RAG1* (Li & Ortí 2007), were amplified using polymerase chain reactions. PCR profiles to amplify *CYTB*, *12S* and *ZICI* followed those used in corresponding primer citations (Coulson et al. 2006; Li & Ortí 2007; Li et al. 2007); a custom profile was used to amplify both nested primer sets of *RAG1*. Further details are provided in Appendix I.

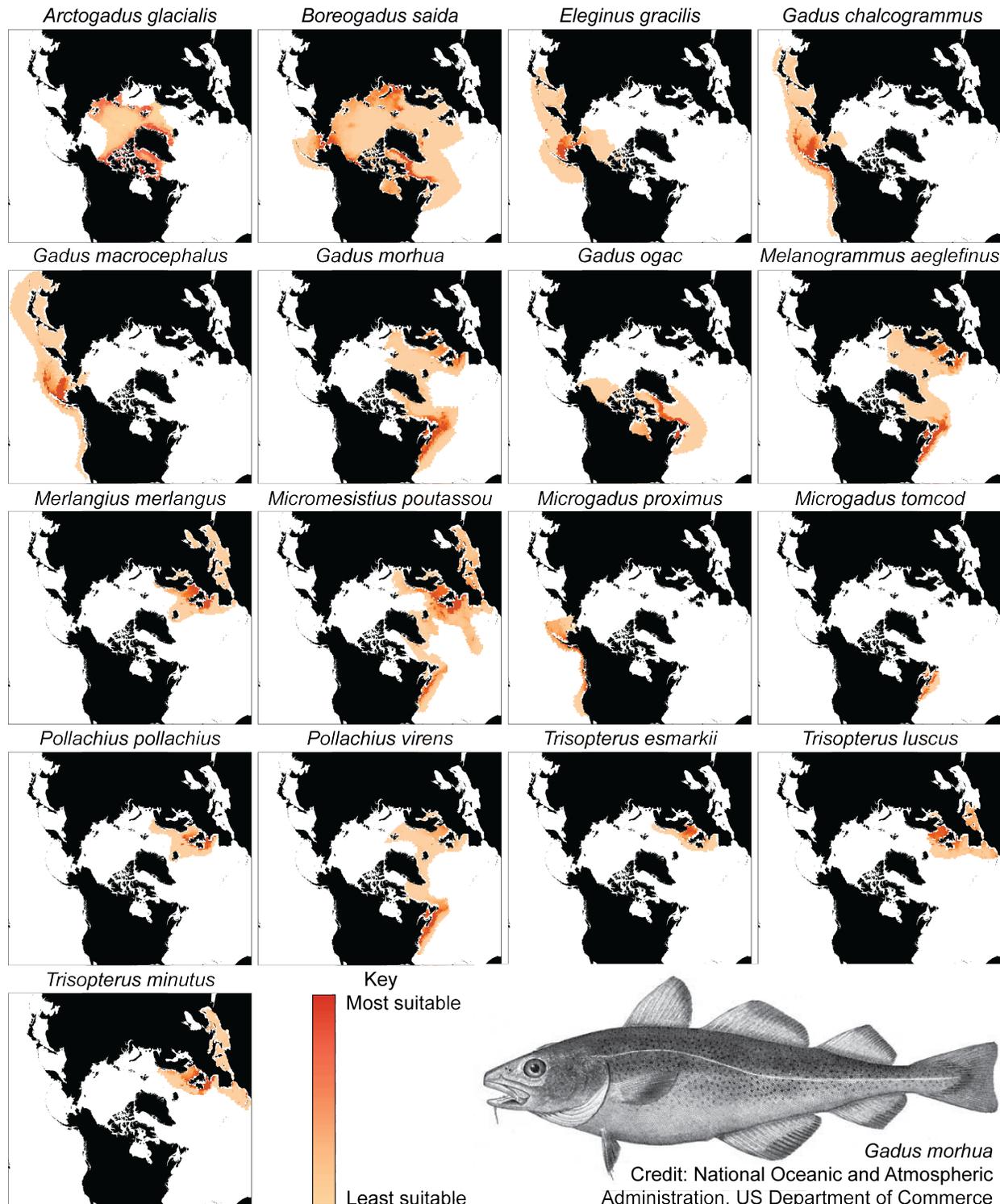
The resulting sequences were assembled using consensus by plurality, combined with data from GenBank (Appendix II.A), and aligned and inspected visually using Sequencher (Gene Codes, Ann Arbor, MI, USA). New sequences derived for this study are available on GenBank under accession numbers KP644319-KP644396 (Appendix II.A). Seven data partitions were analysed using the following models: HKY+I+G for full codons of *12S*, HKY+G for *CYTB* codon positions 1+2 and 3, and HKY for *RAG1* and *ZICI* codon positions 1+2 and 3 (Appendix II.B). Phylogenetic analyses were conducted using Bayesian inference (BI) and maximum likelihood (ML), as implemented in BEAST 1.7.5 (Drummond et al. 2012) and GARLI 2.0 (Zwickl 2006), respectively. The BI of ultrametric species tree topologies was estimated using unlinked gene trees and an uncorrelated lognormal relaxed molecular clock model for all loci except the slowly evolving *RAG1*, for which a strict clock model was applied. The BI phylogeny was scaled using a mutation rate of 1.93% per million years, as measured in *Gadus morhua* (Árnason 2004), to provide rough estimates of clade divergence times: further details are provided in Appendix I and Appendix II.B. The ML search settings included 50 attachments per taxon, generation

threshold for *topoterm* 5000, and score threshold for termination 0.001. Phylogenetic trees are available on TreeBASE under study number 15969

(<http://purl.org/phylo/treebase/phyloids/study/TB2:S15969>).

Ecological niche modelling. Occurrence data were first downloaded from the Global Biodiversity Information Facility (GBIF; <http://www.gbif.org/>) and the Ocean Biogeographic Information System (OBIS; <http://www.obis.org/>). When downloaded occurrence points were compared with documented broad-scale distributions of each species, sampling in the Eastern Hemisphere Arctic Ocean was found to be particularly sparse; consequently, I sought out additional data from the Zoological Institute at the Russian Academy of Sciences (ZIN; Balushkin & Prirodina 2008) and the Natural History Museum in London (NHMUK; Appendix II.C) to fill perceived sampling gaps. Further data quality control steps included the removal of duplicate records and records judged to be inaccurate based on known distributions of species; occurrence data were then downsampled to match the spatial resolution of environmental data using ENMTools v1.3 (Warren et al. 2010). Calibration regions (Fig. 2.2) were designed for each species to reflect the distribution of known occurrences, as well as dispersal capabilities of species, to approximate species' capability to sample suitable and unsuitable environments (Barve et al. 2011).

Figure 2.2. Maxent ecological niche modelling results for each species in the subfamily Gadinae. The extent of the orange shaded area indicates the region used to train each model; darker orange indicates relatively higher habitat suitability within the training region. Image *Gadus morhua*.



Environmental data were derived from a variety of sources. Climatic and topographic data were downloaded from NOAA's National Geophysical Data Center at 1° resolution: bathymetric data were derived from the ETOPO-1 Global Relief Model (<http://www.ngdc.noaa.gov/mgg/global/>; Amante & Eakins 2009); salinity and temperature data were derived from the 2009 World Ocean Atlas (http://www.nodc.noaa.gov/OC5/WOA09/pr_woa09.html; Antonov et al. 2010; Locarnini et al. 2010). Layers summarizing maximum, minimum and mean surface and bottom salinity and temperature were derived from climatology depth layers. Maximum, minimum and mean mixed layer depth were derived from salinity and temperature data. Maximum, minimum and mean sea ice concentration data were downloaded from NOAA's National Snow and Ice Data Center (<http://nsidc.org/data/G02172>). After preliminary modelling runs, during which variable contributions were examined, I determined that mean sea ice concentration could be omitted from every species' model (maximum and minimum sea ice concentration were retained).

Ecological niche models (ENMs) were calibrated for each species using the maximum entropy algorithm Maxent v3.3.3k (Phillips et al. 2006). Ten bootstrap replicate runs with a maximum of 10,000 iterations were conducted using a random seed with 50% of occurrence points. To avoid extrapolation in model features, no clamping, extrapolation, or threshold or hinge features were permitted (see Owens et al. 2013). Models were projected across the Northern Hemisphere to allow direct comparisons of niche tolerances among species.

Ancestral niche reconstruction and hypothesis testing. I used the BioGeoBEARS likelihood framework (Matzke 2013b) to perform biogeographical range analyses using dispersal–extinction–cladogenesis (DEC), DIVA-like [a likelihood interpretation of the dispersal–

vicariance analysis (DIVA) model; Ronquist & Sanmartín 2011], and BAYAREA-like (a simplified likelihood interpretation of the BAYAREA model; Landis et al. 2013) using the phylogenetic tree topology recovered using Bayesian inference. The three models all have free dispersal and extinction parameters, but differ in their treatment of cladogenetic events in which ancestral and daughter ranges overlap (Matzke 2013a). BAYAREA assumes ranges to be conserved at cladogenesis (i.e. the daughter lineages of a widespread lineage will have an identical range when they diverge), whereas DEC assumes one daughter lineage range will always have a range limited to a single biogeographical area (i.e. a scenario in which a widespread lineage diverged into two widespread daughters is disallowed). DIVA does not allow daughter ranges to overlap with each other unless they are identical to a limited-range ancestor.

To assess the most likely model, I compared the results of DEC, DIVA-like and BAYAREA-like by comparing their log-likelihoods and results of biogeographical range state reconstructions. I used the BI ultrametric tree and limited maximum number of allowable areas a species could occupy to two, with four possible geographical areas: north-eastern Atlantic, north-western Atlantic, Arctic and Pacific (Appendix II.D). Additionally, I investigated the effects of using a model with unconstrained dispersal versus a model requiring dispersal through the Arctic for taxa to shift from the Atlantic to the Pacific or vice versa.

Ancestral character state reconstruction began with generating predicted niche occupancy profiles with respect to each environmental variable used to develop ENMs. Median Northern Hemisphere projections of niche suitability were integrated with Northern Hemisphere environmental variable data and divided into 10 equal-width bins using PHYLOCLIM (Heibl 2011) in R (R Core Team, <http://www.r-project.org/>) to create a predicted niche occupancy (PNO) profile (e.g. Appendix III.A). The mean for each environmental variable weighted by ENM

suitability was then used to test for phylogenetic signal using Blomberg's K (Blomberg et al. 2003) as implemented in phytools (Revell 2012) in R; variables with non-significant phylogenetic signal (i.e. $P > 0.05$) were eliminated from further analysis. For each remaining environmental variable, 100 weighted samples of suitable variable space were drawn from the PNO profile of each tip species. For each draw, ancestral character states were estimated using maximum likelihood under a Brownian motion model of evolution on a sample of 1600 ultrametric trees from the posterior distribution of the BEAST tree search. Mean results for each draw were calculated to derive final character reconstructions for each environmental variable; this procedure was performed using phyloclim. Script and data for this analysis can be found in the DRYAD repository <http://doi.org/10.5061/dryad.352th>).

Finally, I tested whether ancestral crown cod and tomcod environmental tolerances were consistent with present-day Arctic or temperate cod niches. Occurrence records for all gadine species were lumped into a single null dataset and used to draw values from the environmental data layer set. Points were categorized as 'Arctic' or 'temperate' according to Longhurst biogeographical provinces (VLIZ 2009). The Wilcoxon rank sum test as implemented in R was used to test whether distributions of environmental values in the two biogeographical provinces were significantly different, and to calculate nonparametric confidence intervals for each environmental variable in the Arctic and temperate datasets. Reconstructed mean tolerances of ancestors of the crown cods and the tomcods were then compared to 95% confidence intervals characterizing Arctic and temperate habitats to test between the two hypotheses. Script and data for this analysis can be found in the DRYAD repository (<http://doi.org/10.5061/dryad.352th>).

Results

Sequences of five genes yielded a data matrix of 3488 aligned bases (*CYTB* 1161, *12S* 695, *RAG1* 819, *ZIC1* 813). Of these bases, 891 were variable (*CYTB* 478, *12S* 129, *RAG1* 186, *ZIC1* 98) and 534 were parsimony informative (*CYTB* 349, *12S* 64, *RAG1* 88, *ZIC1* 33). BI recovered a tree topology most congruent with that of Coulson et al. (2006) (Fig. 2.1). The tree topology inferred by ML was congruent with that of BI, but with lower resolution (Fig. 2.1).

The modelled geographical extent of suitable habitat for each species was congruent with known ranges (Fig. 2.2; Cohen et al. 1990). *Trisopterus* species co-occur and are restricted to the north-eastern Atlantic. *Microgadus tomcod* is limited to the western Atlantic south of Newfoundland, whereas *M. proximus* is found in the eastern Pacific south of the Bering Sea, and is replaced allopatrically by its sister species, *Eleginus gracilis*, in the Bering and Chukchi seas. *Merlangius merlangus* is restricted to the north-eastern Atlantic, while sisters *Melanogrammus aeglefinus* and *Micromesistius poutassou* co-occur across the eastern and western Atlantic. *Pollachius* species co-occur in the eastern Atlantic; the distribution of *P. virens* extends to the western Atlantic, whereas *P. pollachius* does not. *Arctogadus glacialis* is fully Arctic, with the northernmost distribution of any gadine, and is replaced allopatrically by its sister, *Boreogadus saida*, to the south. Finally, two sister lineages in *Gadus* are distributed across the northern Pacific (*G. chalcogrammus* and *G. macrocephalus*) and northern Atlantic (*G. morhua* and *G. ogac*).

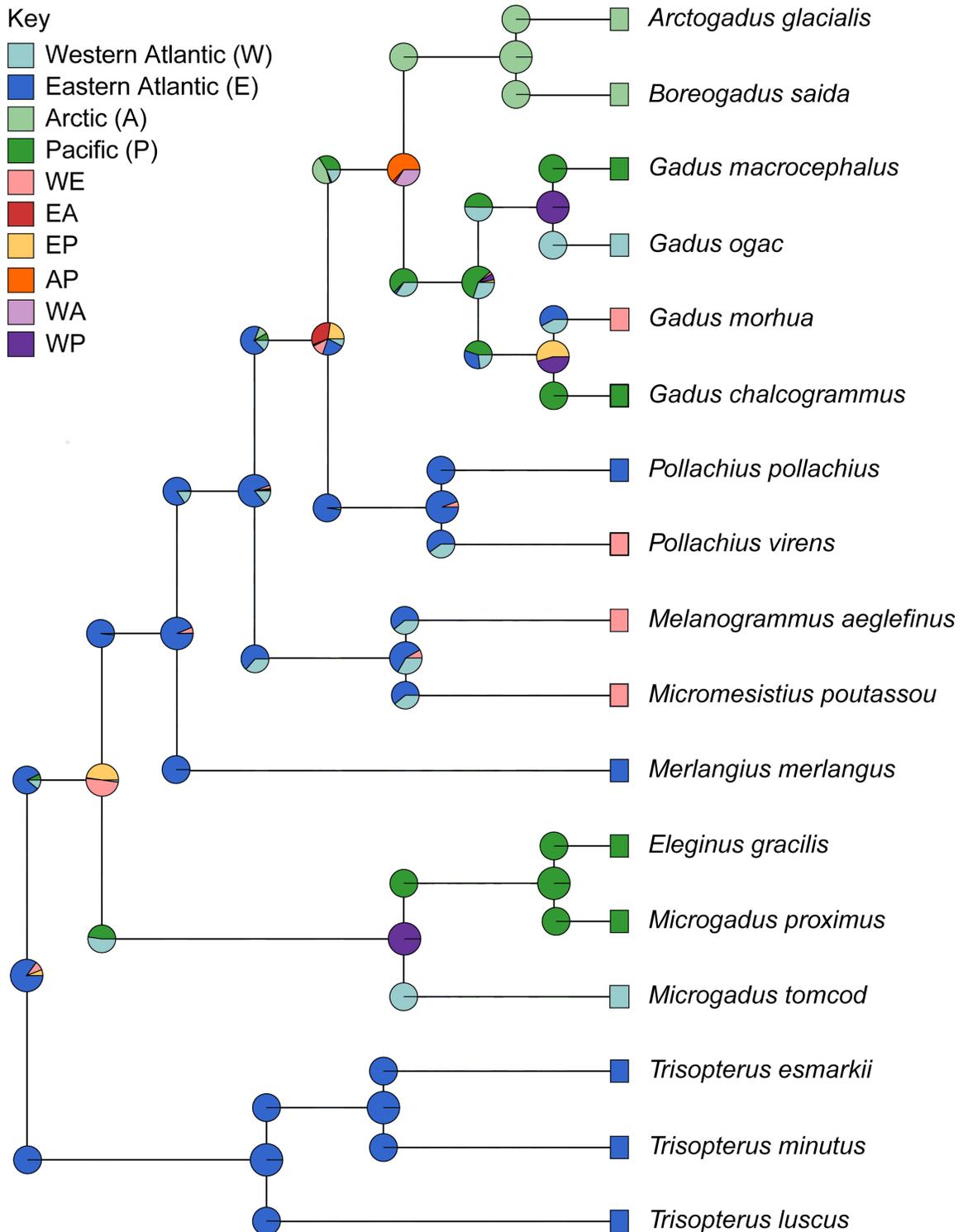
The unconstrained-dispersal DIVA-like model had the best fit to the data, with a log-likelihood of -35.5 (dispersal, $d = 2.11$; extinction, $e = 0.0$), followed by unconstrained-dispersal DEC, with a log-likelihood of -38.0 ($d = 1.87$, $e = 0.0$), and unconstrained BAYAREA-like, with a log-likelihood of -45.5 ($d = 1.35$, $e = 5.0$). Constrained dispersal models resulted in much

lower log-likelihoods: -48.3, 54.6 and -55.6 for DIVA-like, DEC and BAYAREA-like, respectively. Comparing ancestral state reconstructions under the various models further supports use of the unconstrained DIVA-like model for biogeographical range evolution in gadine fishes, as this model led to the least ambiguous reconstructions (Fig. 2.3, Appendix III.B).

Under this model, it is ambiguous whether the tomcod ancestor is more likely to have originated in the western Atlantic or Pacific before expanding in range to inhabit both of these areas and experiencing a vicariance event that split the Pacific ancestor of *Eleginus gracilis* and *Microgadus proximus* from *Microgadus tomcod*. The crown cod ancestor was most probably Arctic, expanding to encompass the Pacific at the time of cladogenesis; the ancestor of *Boreogadus* and *Arctogadus* subsequently retreated into the Arctic, while the ancestor of *Gadus* probably retracted into the Pacific. Results further suggest that *G. morhua* and *G. ogac* dispersed to the Atlantic independently from the Pacific, but both sister pairs of Atlantic/Pacific *Gadus* probably had ancestors that dispersed across the Pacific and Western Atlantic, possibly bridging via the Arctic.

Of the 18 environmental variables used to construct ecological niche models, 5 had significant phylogenetic signal ($P < 0.05$): mean, maximum and minimum bottom temperature, and mean and minimum surface temperature (Table 2.1). For each variable with phylogenetic signal, Arctic and temperate gadine locality characteristics were significantly different ($P < 0.05$). Reconstructed mean tolerances of crown cod and tomcod were not consistent with the characteristics of Arctic habitat: i.e. temperature variable mean values were higher than, and outside the confidence interval of, Arctic conditions. Reconstructed mean tolerances for both tomcod and crown cod ancestors were closer to confidence intervals for modern temperate conditions than for Arctic conditions in all cases. Mean ancestral crown cod tolerance was within

Figure 2.3. Biogeographical range reconstructions using Bayesian inference tree topology under the DIVA-like model with unconstrained dispersal. First four colours in key represent single regions; last six colours represent combinations of two of first four regions. Boxes at tree tips represent species' current distributional areas; pie charts at nodes represent relative probabilities of biogeographical range at cladogenesis; pie charts on branches represent relative probabilities of each daughter lineage's biogeographical range. Branch lengths not to scale.



95% confidence intervals of mean, maximum and minimum temperate surface temperatures; however, no mean ancestral tomcod tolerance fell within confidence intervals for gadine-specific modern temperate environments (Table 2.1).

Discussion

The phylogeny of codfishes presented herein is the most taxonomically complete hypothesis presented to date. As had been suggested by Carr et al. (1999) and Roa-Varón & Ortí (2009), this analysis contributes additional support for monophyly of the tomcods (Fig. 2.1), adding further evidence for merging *Eleginus* into *Microgadus*. Additionally, the results support previous evidence that *Gadus chalcogrammus* is sister to *G. morhua* and thence to the *G. macrocephalus*/*G. ogac* clade; however, as in previous studies, the polytomy among these three lineages was not resolved (Carr et al. 1999; Teletchea et al. 2006; Roa-Varón & Ortí 2009). Deeper genomic sampling may be required for full resolution.

When geographical distributions of gadine fishes estimated from ENMs (Fig. 2.2) are examined in the context of the phylogeny (Fig. 2.1), they appear largely conserved, with nine species currently co-occurring in the north-eastern Atlantic. Range-based reconstructions (Fig. 2.3) estimate the north-eastern Atlantic as the most likely range of the most recent common ancestor of Gadinae; a North Atlantic origin of Gadinae has long been hypothesized (Svetovidov 1948; Carr et al. 1999; Pogson & Mesa 2004; Coulson et al. 2006). The role of glacial refugia in the north-east Atlantic is often suggested as a driver of genetic diversity in the region (Provan 2013), but estimated divergence times for gadines presently distributed in the north-east Atlantic may pre-date glaciation cycles (O'Regan et al. 2011). Alternatively, the closing of the Strait of Gibraltar and subsequent Messinian Salinity Crisis (MSC; between 6.0 and 5.3 Ma), during

Table 2.1. Results of tests for phylogenetic signal, temperate and Arctic habitat characterization, and ancestral character state reconstructions for the crown cod and tomcod clades in the subfamily Gadinae. N/A indicates variable was not phylogenetically significant, and reconstruction was not done.

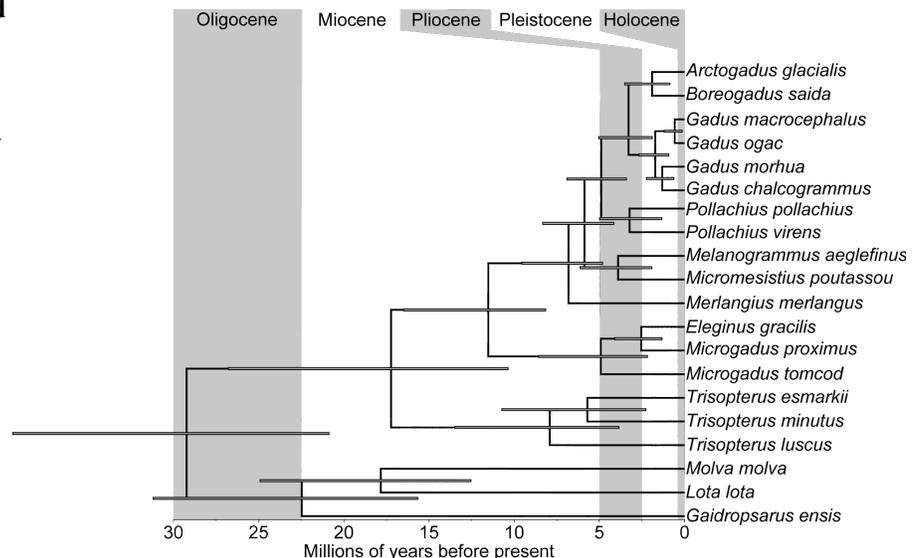
Environmental Variable	Blomberg's <i>K</i>	Arctic/temperate similarity <i>P</i>	95% Confidence intervals		Mean tolerance	
			Temperate	Arctic	MEM	ABG
Bathymetry	7.800×10^{-2}	1.619×10^{-2}	-1.935×10^2 to -1.550×10^2	-2.345×10^2 to -1.955×10^2	N/A	N/A
Maximum sea ice concentration	4.430×10^{-1}	3.887×10^{-113}	9.750×10^1 to 9.750×10^1	1.000×10^2 to 1.000×10^1	N/A	N/A
Minimum sea ice concentration	4.310×10^{-1}	1.763×10^{-5}	0.000×10^0 to 0.000×10^0	2.250×10^1 to 9.000×10^1	N/A	N/A
Average mixed layer depth	3.930×10^{-1}	1.187×10^{-24}	3.807×10^1 to 4.010×10^1	2.822×10^1 to 3.050×10^1	N/A	N/A
Maximum mixed layer depth	5.450×10^{-1}	3.630×10^{-32}	7.824×10^1 to 8.310×10^1	5.092×10^1 to 5.620×10^1	N/A	N/A
Minimum mixed layer depth	8.670×10^{-1}	1.269×10^{-60}	1.074×10^1 to 1.117×10^1	6.938×10^0 to 7.332×10^0	N/A	N/A
Average bottom salinity	3.080×10^{-1}	8.510×10^{-6}	3.409×10^1 to 3.427×10^1	3.451×10^1 to 3.469×10^1	N/A	N/A
Maximum bottom salinity	2.800×10^{-1}	7.275×10^{-11}	3.425×10^1 to 3.442×10^1	3.475×10^1 to 3.484×10^1	N/A	N/A
Minimum bottom salinity	3.340×10^{-1}	1.365×10^{-3}	3.395×10^1 to 3.412×10^1	3.429×10^1 to 3.454×10^1	N/A	N/A
Average surface salinity	3.120×10^{-1}	1.172×10^{-1}	3.2590×10^1 to 3.280×10^1	3.212×10^1 to 3.267×10^1	N/A	N/A
Maximum surface salinity	3.090×10^{-1}	6.804×10^{-5}	3.299×10^1 to 3.321×10^1	3.335×10^1 to 3.366×10^1	N/A	N/A
Minimum surface salinity	3.380×10^{-1}	1.207×10^{-12}	3.216×10^1 to 3.237×10^1	3.065×10^1 to 3.149×10^1	N/A	N/A
Average bottom temperature	5.000×10^{-3}	2.213×10^{-59}	4.607×10^0 to 5.125×10^0	1.366×10^0 to 1.816×10^0	4.303×10^0	3.931×10^0
Maximum bottom temperature	7.000×10^{-3}	5.658×10^{-53}	5.617×10^0 to 6.281×10^0	2.185×10^0 to 2.878×10^0	4.877×10^0	4.393×10^0
Minimum bottom temperature	5.000×10^{-3}	6.222×10^{-51}	3.661×10^0 to 4.083×10^0	0.712×10^0 to 1.204×10^0	3.565×10^0	3.155×10^0
Average surface temperature	3.600×10^{-2}	1.024×10^{-135}	7.540×10^0 to 7.967×10^0	1.226×10^0 to 1.679×10^0	8.224×10^0	7.545×10^0
Maximum surface temperature	7.700×10^{-2}	9.157×10^{-160}	1.318×10^1 to 1.378×10^1	4.994×10^0 to 5.717×10^0	N/A	N/A
Minimum surface temperature	5.000×10^{-3}	1.422×10^{-83}	3.108×10^0 to 3.608×10^0	-1.052×10^0 to -0.748×10^0	4.031×10^0	3.373×10^0

which the Mediterranean Sea experienced intense desiccation, may have driven gadine vicariance. Indeed, the MSC has already been suggested as a driver of speciation within the genus *Trisopterus* (Gonzalez et al. 2012).

In the late Miocene or early Pliocene, tomcods and crown cods began to diversify (Fig. 2.4) and disperse into their modern-day amphiboreal distributions. These distributions, considered in the context of the estimated diversification timing of the groups, suggest that dispersal or vicariance events played a role in shaping present distributions of these species, either following a model of Arctic adaptive radiation and dispersal or cycles of vicariance and dispersal driven by periodic glaciation. To date, glacial-cycle-driven vicariance has been the dominant model of speciation for amphiboreal fish taxa (Mecklenburg et al. 2011), while adaptive radiation is more commonly invoked to explain diversity of fish taxa of the Southern Ocean (Clarke & Crame 2010).

Results of biogeographical range analyses provide some clarification of the possible model by which present-day gadine diversity and distribution patterns of tomcods, crown cods and Gadinae as a whole emerged, with the DIVA-like model being most likely, suggesting that

Figure 2.4. Time-calibrated Bayesian inference phylogeny of the subfamily Gadinae. An uncorrelated lognormal relaxed molecular clock model was used for all loci except *RAG1*, for which a strict clock model was applied. The phylogeny was then scaled using a mutation rate of 1.93% per million years as measured in *Gadus morhua* cytochrome *b* gene (Árnason 2004). Bars represent posterior distribution of divergence-time estimates.



vicariance events were potentially more significant in the emergence of current-day biogeographical patterns than widespread dispersal and cladogenesis within overlapping ranges. However, a great deal of uncertainty remains in estimates of ancestral geographical range shifts in Gadinae. It is possible that dispersal may have happened so quickly on a relative geological time scale that vicariance merely appears more likely than dispersal and subsequent extinction. Indeed, diversification of the genus *Gadus* may have begun only in the last 2.5 million years (Fig. 2.4). By considering processes that underlie patterns of change in biogeographical range, such problems may be ameliorated (Crisp et al. 2011).

Clarification of the biogeographical history of Gadinae, particularly that of the tomcods and crown cods, may be possible when the hypothesized evolution of gadine physiological tolerance limits is taken into account. While bathymetry, mixed layer depth, salinity and temperature were all informative in generating ENMs, only temperature variables showed phylogenetic signal. This pattern may indicate that, for Gadinae, temperature adaptations are under higher selective pressure than adaptations to other environmental characteristics of their coarse-grained ecological niches. Consequently, these results suggest that physiological limitations to survival in extreme temperatures played the most significant role in governing the distributions of gadine fishes. A recent analysis comparing distributions of marine fishes with their thermal tolerance limits provided evidence that in marine systems, latitudinal distributions of fishes tend to conform to their thermal tolerance limits (Sunday et al. 2012). It is also important to note that in the present study, evidence suggests temperate species are largely intolerant to any but the lowest concentrations of sea ice, regardless of their phylogenetic position (Appendix III.A); this may account for lack of phylogenetic signal for sea ice tolerance

in the analysis. Conversely, lack of phylogenetic signal in environmental tolerance variables may simply be the result of inadequate characterization due to noise in the data.

Ancestral tomcod and crown cod niche tolerances were not statistically similar to those of modern Arctic cods. Rather, ancestral tolerances of both clades appeared much closer to those of modern temperate cods. This result is congruent with the consistently temperate biogeographical range reconstruction for the tomcod ancestor, which was probably Western Atlantic, and dispersed to include the Pacific at the time of cladogenesis. However, the biogeographical range of the crown cod ancestor was reconstructed as Arctic, expanding to the Pacific at the time of cladogenesis. This is unlikely, because the crown cods are estimated to have diverged in the Pliocene (Fig. 2.4), when proxy estimates of global temperature indicate that the climate was cooler than the present day (O'Regan et al. 2011). When viewed in the context of divergence time estimates and the palaeoceanographic history of the region, these patterns of ancestral niche tolerance suggest that the first emergence of Arctic sea ice may have contributed to diversification in these clades. According to the molecular clock estimate, tomcods began to diversify between 2.5 and 7.5 Ma, when Arctic ice sheets were first forming (O'Regan et al. 2011). The crown cods, too, first began to diversify near this time, with further diversification by the genus *Gadus* during the Pleistocene, when fluctuating levels of sea ice may have led to episodic ice-free passages between the Atlantic and Pacific (O'Regan et al. 2011). This may suggest that cyclical glaciation could have served as a speciation pump; the first time such a model has been invoked to elucidate diversification of tomcods and crown cods above the species level.

It should be noted that caution must be exercised when interpreting phylogenetic reconstructions of coarse-grained ecological niches. It is by now well established that correlative

niche models (of which Maxent is one) serve to characterize the realized niche of a species; that is, the intersection of biotic and abiotic factors which a species can access, and not the fundamental niche, which is the full suite of factors in which a species is capable of persisting (Soberón & Peterson 2005), unless considerable effort is made to generalize models. In cases where the suitable range of a given environmental factor for a species is close to the limits of values available for background sampling in the model, it is unlikely that the full fundamental niche of the species has been characterized, and any further analysis in which full characterization of the fundamental niche is to be assumed must be approached with caution (Zurell et al. 2012; Owens et al. 2013).

For example, in the present study, the extremes of the surface temperature fundamental niche of the eurythermal taxa *Arctogadus* and *Boreogadus*, have probably not been fully characterized. The lowest surface temperature at which these species can be found is the minimum temperature at which water remains liquid, and the environmental layers used to train ecological niche models in this study lack the resolution necessary to capture this lower limit. This result may bias ancestral state reconstructions on which estimates of the fundamental niche of *Arctogadus* / *Boreogadus* were based. Indeed, in view of the peripheral nature of their niches within their accessible areas, crown cod ancestral coarse-grained niche tolerance minimum and mean surface temperature may in reality be lower than that reconstructed in this study (Table 2.1). However, the effects of such a bias in this study are minimal, as only two environmental factors out of the 306 estimated (18 environmental factors for each of 17 species) may not have fully characterized fundamental tolerances due to the broad areas across which niches were characterized.

The finding that temperature probably drives gadine distributional patterns, with sea ice as a possible barrier to contact between Atlantic and Pacific species, suggests potentially significant conservation and economic consequences from future climate change. Rapid reduction of summer sea ice extent (Overland & Wang 2013) will reduce available nursery habitat for the sea-ice associated *Arctogadus glacialis* and *Boreogadus saida*. These species are essential prey for marine mammals and seabirds. In the Lancaster Sound region alone, approximately 148,000 tonnes of *Boreogadus saida* are consumed annually by ringed seals (70,000 tonnes), narwals (23,000 tonnes), belugas (23,000 tonnes), thick-billed murre (12,180 tonnes), and others (Welch et al. 1992). It is possible that faced with a shortage of gadine prey, predators will either suffer reductions in their numbers or shift to other food sources, which could in turn have additional consequences. While it is beyond the scope of this study to project the future distributions of potential gadine habitat, it is also possible that the geographical ranges of many commercially important gadine species will shift to track suitable temperatures. Indeed, the ratio of Arctic to subarctic fish and invertebrate species recovered in bottom trawl surveys from 1982 to 2006 has declined throughout the Bering Sea (Mueter & Litzow 2008). Furthermore, the opening of ice-free Arctic passages may allow Pacific and Atlantic gadines once more to disperse freely between the Pacific and Atlantic through the Arctic.

This study provides a unique combination of ancestral ecological niche tolerance reconstruction with cutting edge methods of ancestral range reconstruction, and thereby opens several key insights into gadine evolution and biogeography. However, the specific timing of diversification events and contemporaneous climate and geological changes remains unresolved. Recent improvements in fossil calibration methods (Heath et al. 2014) may provide a powerful tool by which to understand better how well the timing of glaciation and diversification events

coincide. Furthermore, using estimates from fossil calibrations, reconstructed ecological niche tolerances at key points in the evolution of Gadinae could be projected onto palaeoclimatological datasets (Peterson & Nyári 2007) to explore past periods of dispersal and isolation that may serve to inform the hypothesis that a glacial cycle speciation pump led to speciation across oceans. Understanding the evolutionary history of the Gadinae may provide valuable insight into how climate change will affect the future of this economically and ecologically important clade of fishes.

Chapter 3*

Predicting future distributions of gadine fishes using global climate model data:

exploring spatial and environmental variable bias

*Owens, HL, JA Kleypas. Predicting future distributions of gadine fishes using global climate model data: exploring spatial and environmental variable bias.

Abstract

The aim of this study was to anticipate effects of climate change on future distributions of suitable conditions for gadine codfishes and analyze challenges inherent in the use of GCM climate model data for making such predictions. Using present-day locality data, we generated ecological niche models (ENMs) of 15 species of fishes in the subfamily Gadinae based on two environmental datasets: an observation-based dataset and a dataset derived from the CCSM4 general circulation model. Both sets of niche models were projected into modeled future climates at 2054 and 2100 based on the RCP 8.5 climate scenario, the most extreme of the RCP scenarios for future greenhouse gas emissions. We generated maps of areas of disagreement between models based on observational and modeled climate data. Suitable habitat for *Arctogadus glacialis* is predicted to disappear by 2100; other gadine fishes are anticipated to experience various degrees of expansion and northward shifts in suitable habitat by 2100. We detected a strong spatial bias disagreement between observation- and GCM-based ENMs concordant with known GCM biases. Our study explicitly connects spatial bias in ENM results with that of GCM-derived environmental data inputs. Consciousness of bias in GCM data is key to using ENMs to forecast potential future distributions of suitable conditions under which species may persist. Future reductions in GCM bias, in tandem with careful evaluation of ENM projections, will allow policy makers and researchers to more confidently identify areas of particular concern for biodiversity.

Introduction

Distributions of many marine species have shifted over the last few decades, often associated with warming temperatures (Perry et al. 2005; Dulvy et al. 2008; Leu et al. 2011; Cheung et al. 2013; Vergés et al. 2014), but also in response to changes in fishing pressure (Garrison & Link 2000), prey availability (Weijerman et al. 2005), solar constant (Weijerman et al. 2005), salinity (Kimura et al. 2001; Weijerman et al. 2005) and sea ice cover (Leu et al. 2011), among others. As marine ectotherms tend to occupy the full range of their physiological tolerance to temperature, they may be particularly susceptible to changing climatic conditions (Sunday et al. 2012). This can manifest as changes in species' distributions as they track suitable climatic conditions (Parmesan & Yohe 2003). Changes in species' distributions have important management and conservation implications for marine fishes, not only because of the geographic shifts themselves, but also because of the effects of climate on population dynamics, including alterations in spawning stock composition and biomass (Brander 2005; Ottersen et al. 2006), and shortened duration of planktonic larval life stages, which may reduce dispersal ability and population connectivity (O'Connor et al. 2007; Kristiansen et al. 2014).

Cods, pollocks and haddocks of the subfamily Gadinae are among species for which climate-change-induced distributional changes have already been documented (Cormon et al. 2014; Engelhard et al. 2014). These species typically inhabit temperate, continental waters in the high latitudes of the Atlantic and Pacific oceans, but Arctic species *Arctogadus glacialis* and *Boreogadus saida* associate strongly with sea ice, and are among the most northerly-distributed marine species in the world (Cohen et al. 1990). Many gadine species are economically important wild fishery species: in 2012, 6.1 million tons of gadine fish were harvested (FAO 2014). Additionally, gadine fishes are key consumers and prey in northern ocean ecosystems

(Link & Garrison 2002, Gradinger & Bluhm 2004). Of the 22 recognized species in the clade (Nelson 2006), seven have been assessed for conservation status by the International Union for the Conservation of Nature; two, *Gadus morhua* and *Melanogrammus aeglefinus*, are currently ranked as “Vulnerable” (IUCN 2014). Further pressure may be exerted on these and other gadine species as high latitudes are expected to undergo the most pronounced climate-change-induced alterations over the coming decades (Cheung et al. 2009). Being able to predict how gadine distributions may change is vital to building adaptive strategies for management and protection of these species.

Ecological niche modeling (ENM) is increasingly used to understand factors shaping present geographic ranges of species, and to anticipate the effects of climate change on species’ distributions. ENM estimates the scenopoetic, or Grinnellian, ecological niche of a species (Soberón 2007) based on environmental characteristics associated with known occurrences (Peterson et al. 2011). Such models then may be transferred onto different landscapes or onto estimated environmental conditions in the past or future to infer the distribution of suitable habitat for the species of interest. In late January 2015, a search on Google Scholar (<http://www.scholar.google.com>) yielded 195 studies in 2014 alone in which researchers used ENM to predict future distributions of species’ suitable habitat using climate model data. Predictions of climate change effects on the distribution of species’ suitable habitat are of particular interest for conservation planning: by projecting species’ potential distributions into future climate scenarios, researchers and policy makers may be able to prioritize critical conservation areas by identifying areas where suitable habitats are likely to persist or where new areas of conservation importance may emerge (Alagador et al. 2014).

While ENM predictions using climate model data remain a powerful estimator of climate change-induced distributional change, their use for making conservation and management decisions remains controversial (Sinclair et al. 2010). Such models may erroneously estimate a species' ecological niche owing to factors such as inappropriately defined study areas (Barve et al. 2011) and limitations set by competitors (e.g. Troia & Gido 2014) or symbionts (e.g. Mueller et al. 2011). Such models are only as good as the environmental data that goes into them (Hall 2014). Data from different sources, at different scales, and summarizing different aspects of the environment may perform in dramatically different ways when identifying species' potential distributions (Peterson & Nakazawa 2007). However, the error deriving from biases in environmental variable data remains an under-appreciated source of uncertainty in such models.

This study explores a new marine dataset derived from the Community Climate System Model version 4.0 (CCSM4) developed by the National Center for Atmospheric Research (NCAR). We use these data to predict future distributions of suitable habitat for 15 species of gadine fishes. Whereas previous researchers have focused on fine-scale mechanistic modeling of the effects of climate change on gadine fishes (e.g. Kristiansen et al. 2014; Mueter et al. 2011), this study focuses on broad-scale changes in distributions of suitable scenopoetic niche space for gadine fishes. In particular, we consider the implications of uncertainty inherent in climate model data when interpreting distributional change in conservationally important species, comparing our results with known biases in the CCSM4 climate model.

Methods

We developed ENMs for 15 recognized species in the subfamily Gadinae (Nelson 2006) (Table 3.1). We eliminated *Arctogadus borisovi* and *Theragra finnmarchica* from the dataset, as

Table 3.1. Estimated area (km²) and mean latitude of potential distributions of gadine species included in this study.

Species	Observational environmental data ENM				GCM environmental data ENM			
	Area 2054	Area 2100	Mean Latitude 2054	Mean Latitude 2100	Area 2054	Area 2100	Mean Latitude 2054	Mean Latitude 2100
<i>Artogadus glacialis</i>	0	0	N/A	N/A	449,385	0	76.19	N/A
<i>Boreogadus saida</i>	14,182,652	35,128,678	63.50	68.06	11,927,679	13,037,150	66.50	68.50
<i>Eleginus gracilis</i>	20,260,569	16,711,478	65.76	66.91	1,488,961	2,272,406	64.00	67.00
<i>Gadus chalcogrammus</i>	37,077,118	32,431,453	58.00	60.50	5,887,450	7,965,118	60.53	62.00
<i>Gadus macrocephalus</i>	36,130,504	38,991,828	58.00	60.00	6,431,001	7,134,933	59.32	62.00
<i>Gadus morhua</i>	66,394,880	85,426,166	53.89	54.34	24,294,598	32,138,459	48.50	51.50
<i>Gadus ogac</i>	0	0	N/A	N/A	9,376,952	12,556,218	52.88	55.35
<i>Melanogrammus aeglefinus</i>	43,854,802	59,003,488	54.00	56.30	11,886,873	13,242,078	46.02	50.96
<i>Merlangius merlangus</i>	14,975,851	22,844,476	51.45	55.29	3,195,305	3,274,866	49.83	52.04
<i>Microgadus proximus</i>	13,800,782	16,561,752	54.50	58.00	1,597,442	923,112	63.02	65.46
<i>Microgadus tomcod</i>	6,005,362	15,217,461	47.89	57.06	2,527,088	2,351,868	50.95	55.44
<i>Micromesistius poutassou</i>	197,505,618	178,564,169	46.00	48.50	20,740,797	25,844,706	47.95	49.00
<i>Pollachius pollachius</i>	5,090,823	5,571,058	52.65	55.39	1,130,978	484,043	55.61	58.15
<i>Pollachius virens</i>	92,297,640	104,904,189	50.02	53.45	12,879,809	11,173,488	45.50	48.00
<i>Trisopterus minutus</i>	28,750,815	29,160,434	49.34	53.93	4,068,023	4,271,353	50.29	51.78

they are considered junior synonyms of *Arctogadus glacialis* and *Gadus chalcogrammus*, respectively (Byrkjedal et al. 2008; Jordan et al. 2003). *Eleginus navaga*, *Gadiculus argenteus*, *Micromesistius australis*, *Trisopterus esmarkii* and *Trisopterus luscus* were not included owing to insufficient data. Occurrence data were adapted from a dataset produced for a previous study, where details on point acquisition and quality control are provided (Owens 2015); points were downloaded from the Global Biodiversity Information Facility (GBIF; <http://www.gbif.org/>) and the Ocean Biogeographic Information System (OBIS; <http://www.obis.org/>), augmented with data from the Zoological Institute at the Russian Academy of Sciences (ZIN; Balushkin & Prirodina 2008) and the Natural History Museum in London (NHMUK) for regions not represented in GBIF and OBIS. We designed calibration regions to reflect distributions of known species occurrences and dispersal capabilities, to approximate each species' capability to sample suitable and unsuitable environments (Barve et al. 2011).

The present-day observational climatic and bathymetric dataset (henceforward observational dataset) was derived from a variety of sources at a spatial resolution of 1°.

Bathymetric data were acquired from the ETOPO-1 Global Relief Model at NOAA's National Geophysical Data Center (<http://www.ngdc.noaa.gov/mgg/global/>; Amante & Eakins 2009). Average decadal climatologies of salinity and temperature from 1955 to 2006 were downloaded from the 2009 World Ocean Atlas (http://www.nodc.noaa.gov/OC5/WOA09/pr_woa09.html; Antonov et al. 2010; Locarnini et al. 2010). Layers summarizing maximum, minimum, and mean surface and bottom salinity and temperature were derived from the climatology data; maximum, minimum, and mean mixed layer depth were calculated based on salinity and temperature data. Maximum, minimum, and mean sea ice concentrations were derived from a 1972 to 2007 climatology from NOAA's National Snow and Ice Data Center (<http://nsidc.org/data/G02172>).

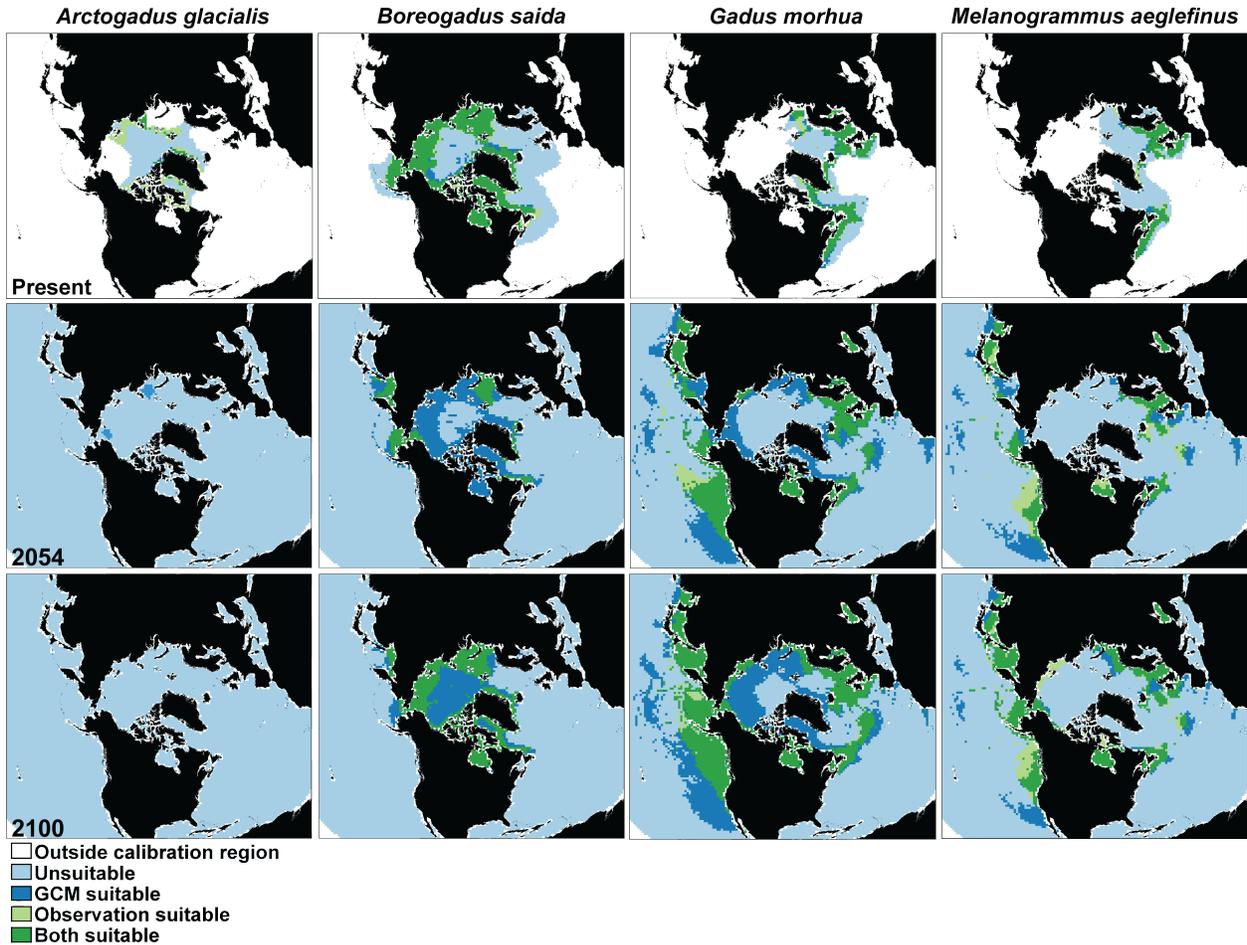
The general circulation model dataset (henceforward GCM dataset) was derived from CCSM4, which couples atmosphere, land, ocean, and sea ice dynamics into a single general circulation model (Gent et al. 2011). The ocean-model component is the CCSM2.1.1 Parallel Ocean Program (POP2), which has a nominal 1° horizontal resolution (constant at 1.125° in longitude and varying from 0.27° at the equator to about 0.64° in latitude) and 60 vertical levels. The sea ice model shares the same horizontal grid as POP. Monthly climatologies from 1955 to 2005 for salinity, potential temperature, and sea ice concentration were derived from model outputs. As with observed climatologies, layers summarizing maximum, minimum, and mean surface and bottom salinity and temperature; maximum, minimum, and mean mixed layer depth; and maximum, minimum, and mean sea ice concentrations were derived from these data. Mixed layer depth was calculated in the same manner as for the observational dataset, and potential temperature was converted to *in situ* temperature to render it comparable to observed temperature. The resulting layers were then re-gridded from the POP2 grid to a 1° geographical coordinate grid to match the observational data and render it compatible with ENM software.

Climate projections were also developed based on the RCP 8.5 future climate scenario, which represents the highest greenhouse gas emissions of the RCP family of future climate scenarios (Riahi et al. 2011); climatologies were derived from means of five ensemble members for both the early and late periods of the 21st Century projection: years 2005-2054 and 2054-2100. Observed bathymetry as derived for the observational dataset was also included in each of these datasets to render ENMs comparable.

Two sets of ENMs were developed using the correlative niche modeling algorithm Maxent v3.3.3k (Phillips et al. 2006) for each species, one calibrated using the observational dataset, and one calibrated using the GCM dataset. Each ENM was developed using 10 bootstrapped replicates, with a maximum of 10,000 iterations each. For each replicate, 50% of occurrence points were chosen at random for intrinsic model testing; 25% of occurrence points were chosen at random for *Arctogadus glacialis* and *Microgadus proximus*, as these species had more limited occurrence data. To avoid overfit ENMs with unrealistic environmental response curves, no threshold or hinge features were permitted. To reduce overfitting further, we reduced the number of variables included in final environmental datasets by assessing autocorrelation between variables and variable contributions in preliminary model runs; the final datasets included only bathymetry, maximum and minimum sea ice concentration and mean surface temperature and salinity.

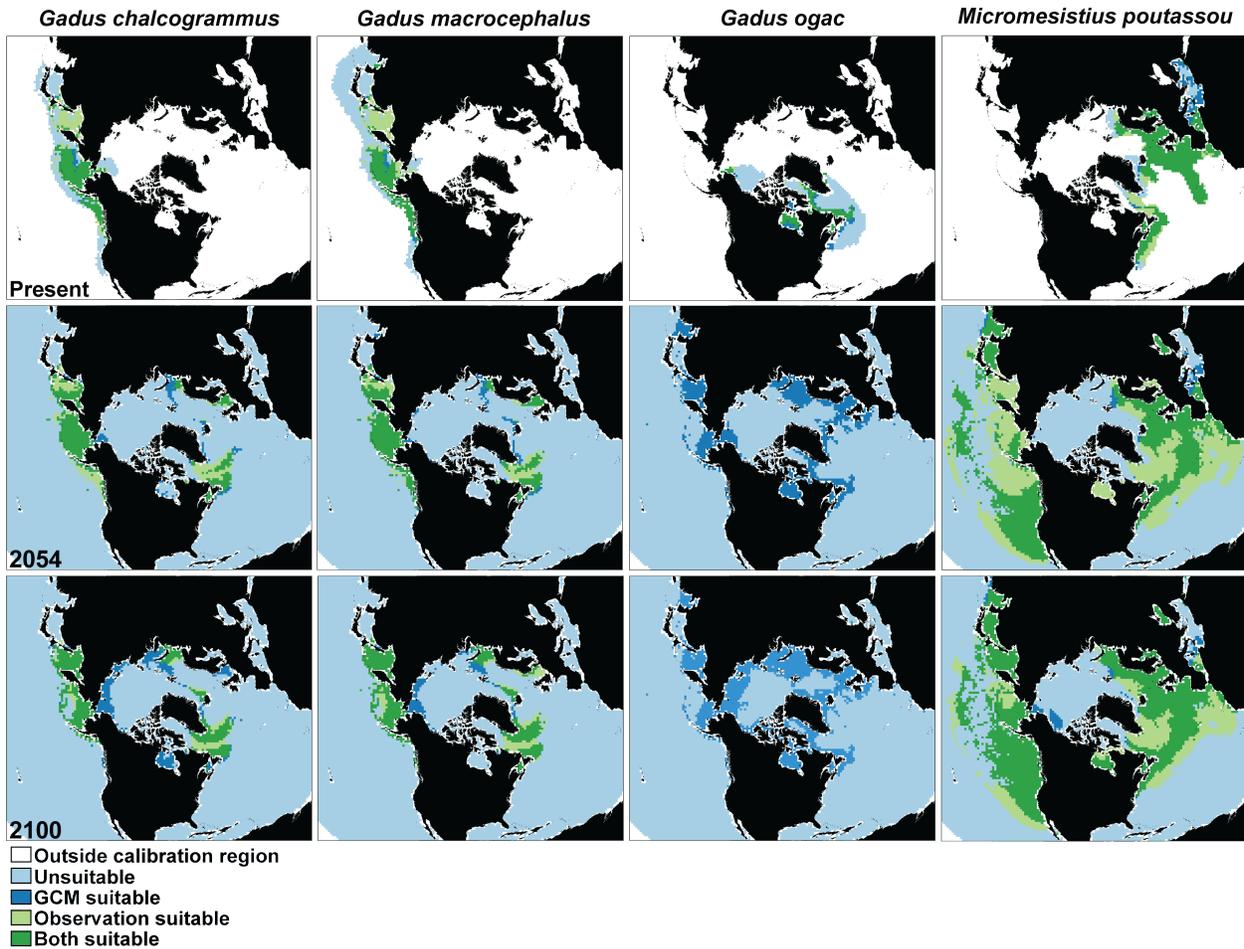
The observational and GCM-based ENMs for each species were then used to project the species' potential distributions in the present (2005), in 2054, and 2100. To avoid unpredictable model transfer behavior (Owens et al. 2013), we did not permit extrapolation or clamping. The suitable habitat for each species' modeled distributions was defined using a 95% occurrence threshold using R (R Core Team 2012; Figure 3.1-2; Appendix IV.A; Appendix V). In this

Figure 3.1. Modeled suitable habitat extent for GCM- and observation-calibrated ecological niche models for *Arctogadus glacialis*, *Boreogadus saida*, *Gadus morhua*, and *Melanogrammus aeglefinus* through time. Colored area in the present denotes the calibration region; white areas are outside calibration regions. Future time slices are projected across the Northern Hemisphere to allow estimation of potential range expansions.



method of thresholding, the highest 95% of suitability scores at occurrence points used to calibrate the model are classified as “suitable”, and values falling below the 95% threshold are classified as “unsuitable”. While more complex methods of thresholding have demonstrated better skill in classifying suitable and unsuitable environments (Jiménez-Valverde & Lobo 2007; Liu et al. 2005), our dataset lacks true absence data, and therefore, a simple fixed threshold technique was preferable (Bean et al 2012). The 95% threshold value was chosen based on comparison between ENM-classified “suitable” habitats and maps of species’ known

Figure 3.2. Modeled suitable habitat extent for GCM and observation-calibrated ecological niche models of *Gadus chalcogrammus*, *Gadus macrocephalus*, *Gadus ogac*, and *Micromesistius poutassou* through time. Colored area in the present denotes the calibration region; white areas are outside calibration regions. Future time slices are projected across the Northern Hemisphere to allow estimation of potential range expansions.



distributions (Cohen et al. 1990). In the present, models of suitable environmental space were limited to the calibration area extent to avoid unpredictable extrapolation; future time slice comparisons were made across the entirety of the northern hemisphere.

For each species' ENM projections, future suitable habitat area and mean latitude (Table 3.1) were calculated according to the 95% calibration presence threshold maps (Figure 3.1-2; Appendix IV.A; Appendix V). Cohen's Kappa was calculated using R (Appendix V) for each species at each time slice to quantify agreement between ENM projections derived from

Table 3.2. Statistical comparison of models. Cohen’s Kappa measures agreement between GCM- and observation-calibrated niche models at each time slice. Variable contributions denote relative contributions (in percent of total) of environmental variables in niche model calibration. Max Sea Ice: maximum sea ice concentration; Min Sea Ice: minimum sea ice concentration; SSS: sea surface salinity; SST: sea surface temperature.

Species	Kappa		Observational variable contributions							GCM variable contributions			
	Present	2055	2100	Depth	Max Sea Ice	Min Sea Ice	Mean SSS	Mean SST	Depth	Max Sea Ice	Min Sea Ice	Mean SSS	Mean SST
<i>Arctogadus glacialis</i>	0.29	0.00	0.00	27.4	4.3	51.7	13.2	3.4	57.0	18.0	12.1	11.5	1.3
<i>Boreogadus saida</i>	0.83	0.25	0.42	13.8	35.4	39.0	6.8	4.9	17.8	30.4	33.3	15.9	2.6
<i>Eleginus gracilis</i>	0.75	0.35	0.63	66.2	25.5	0.1	0.8	7.4	69.6	7.4	12.1	2.7	7.2
<i>Gadus chalcogrammus</i>	0.63	0.69	0.60	32.3	27.4	0.0	15.4	24.9	42.6	28.4	4.8	9.4	14.8
<i>Gadus macrocephalus</i>	0.58	0.75	0.71	37.6	20.8	0.0	2.0	39.6	48.6	23.1	0.0	1.1	27.2
<i>Gadus morhua</i>	0.72	0.55	0.48	47.1	9.3	0.0	28.0	15.6	29.8	7.7	9.0	31.2	22.3
<i>Gadus ogac</i>	0.72	0.00	0.00	40.1	32.9	4.8	6.5	15.7	54.8	11.5	18.6	7.1	7.9
<i>Melanogrammus aeglefinus</i>	0.84	0.59	0.73	19.6	17.4	0.0	37.7	25.4	35.1	22.1	0.8	19.9	22.1
<i>Merlangius merlangus</i>	0.82	0.80	0.67	74.1	8.9	0.0	7.8	9.2	42.4	7.5	0.0	49.3	0.8
<i>Microgadus proximus</i>	0.54	0.31	0.75	73.7	6.9	0.0	18.5	0.9	30.5	12.1	0.1	32.8	24.5
<i>Microgadus tomcod</i>	0.76	0.41	0.07	94.3	1.2	0.0	0.2	4.3	51.4	17.0	0.0	16.8	14.7
<i>Micromesistius poutassou</i>	0.49	0.58	0.55	41.5	26.3	0.0	10.9	21.2	30.7	23.7	0.6	25.9	19.1
<i>Pollachius pollachius</i>	0.74	0.75	0.46	70.0	11.5	0.0	10.5	8.1	38.0	4.7	0.0	49.1	8.2
<i>Pollachius virens</i>	0.72	0.45	0.38	20.2	10.6	0.0	55.4	13.8	34.1	6.4	0.5	26.3	32.7
<i>Trisopterus minutus</i>	0.81	0.72	0.72	45.8	1.0	0.0	36.7	2.6	46.5	0.9	0.0	49.4	3.1

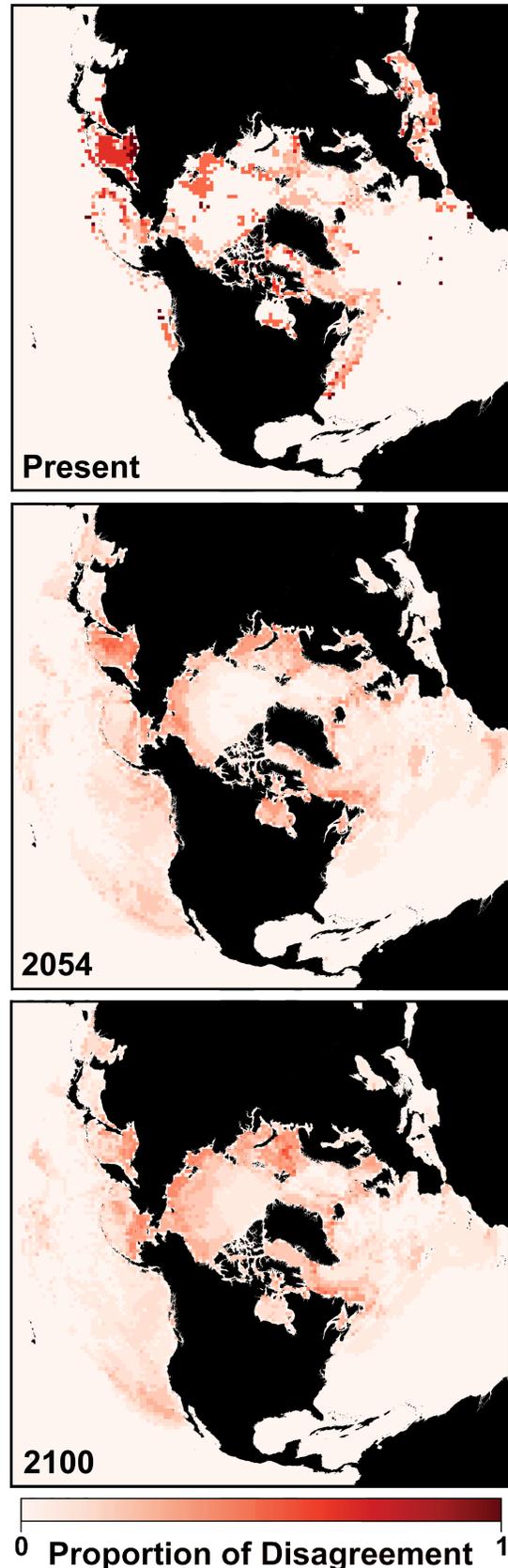
observational and GCM datasets (Table 3.2). We compared Kappa scores with the relative ENM contribution of each environmental variable to determine if certain environmental variables were more likely to cause incongruent distributional predictions than others. Further, to test whether variables in observational and GCM datasets differed significantly, we performed a paired Wilcoxon signed rank test as implemented in R (Appendix V). We also mapped ENM disagreement for all species at each time slice to explore spatially-based areas of disagreement (false positives and false negatives) between models calibrated using observational and GCM datasets that may contribute to uncertainty in predictions of species’ future distributions (Figure 3.3; Appendix V).

Results

Thresholded ENMs calibrated using either observational data or GCM data produced present-day distributional predictions that were generally congruent with known distributions (Fig. 3.1-2, Appendix IV.A). As present species’ distributions were limited to the extent of the

Figure 3.3. Map of regional prediction disagreement at each time slice. Rasters of disagreement between ENMs calibrated using observational and GCM datasets were computed for all 15 species at each time point; from those, proportional disagreement was calculated for each time point. These maps provide guidance in the interpretation of ENM projections by showing where the lowest biases are located, and thus, where results may be interpreted with the greatest confidence.

model calibration region for each species, they were not compared directly to cross-hemisphere predictions at future time points. Estimated future distributions of suitable habitat tended to be more extensive for models calibrated using observational data. In 21 of 30 cases, extents of potential distributional areas increased between 2054 and 2100: models calibrated using the observational dataset predicted an area increase for 11 species, whereas those calibrated using the GCM dataset predicted an increase for 10 species (Table 3.1). In all but one projection, mean latitude of distributional areas was predicted to increase between 2054 and 2100. For the remaining species, *Arctogadus glacialis*, neither of the niche model projections identified any suitable areas at either future time point (Fig. 3.1). No clear pattern of difference in mean latitude predictions for the two sets of models could be



detected (Table 3.1).

There were no consistent trends in Cohen's Kappa across species through time (Table 3.2). Depth was the most significant contributor in 19 of 30 ENMs, followed by mean sea surface salinity (7 models), minimum sea ice concentration (3 models), and mean sea surface temperature (1 model). Minimum sea ice concentration was the least important variable in 24 models, followed by mean sea surface temperature (4 models), mean sea surface salinity (1 model) and maximum sea ice concentration (1 model). The relative contributions of environmental variables were not consistent between GCM- and observation-calibrated models. Observational- and GCM-calibrated models agreed on the most significant variable for 6 of 15 species (depth, 5 species; minimum sea ice concentration, 1 species); the least significant variable was the same for 13 of 15 species (minimum sea ice concentration, 11 species; mean sea surface temperature, 1 species). We also found no significant relationship between model disagreement and variable contribution (Table 3.2). Results of the Wilcoxon signed rank test indicate that only mean surface temperature was not significantly different ($P > 0.05$) between the present observational dataset and the GCM dataset.

Furthermore, the model sets were neither spatially congruent in the present nor in future projections (Figures 3.1-2, Appendix IV.A). Observational- and GCM-calibrated ENMs of species that occur solely in the eastern Atlantic (*Merlangius merlangus*, *Pollachius pollachius*, and *Trisopterus minutus*) had the highest average agreement in the present, and were relatively high in agreement through future projections; models of species that occur in the Arctic (*Arctogadus glacialis* and *Boreogadus saida*) had the lowest agreement in the present, and were relatively low through future projections. Particularly notable areas of disagreement are located

in the Sea of Okhotsk, the Laptev Sea, the western Bering Sea, the Barents Sea and in the Gulf Stream off the eastern coast of North America (Fig. 3.3).

Discussion

Our results suggest that spatial incongruity between environmental data sources is a more significant and consistent source of variation in ENM results than environmental-variable-based incongruity. Relative variable importance was not predictive of congruence (Cohen's Kappa) between the sets of ENMs. Particularly striking was our observation that the ENM contributions of depth, which were identical in each dataset, and surface temperature, which was not significantly different between datasets, did not correlate with Cohen's Kappa. These patterns may be the result of Maxent's response not only to mean values of the environmental data, but also to their variance (Merow et al. 2013). That is, it is possible that, even though the two datasets do not have significantly similar means, their variances across the landscape were generally similar, leading to similar patterns in predicted species distributions.

Exploration of the CCSM4 GCM model from which we derived our GCM dataset sheds some light on disagreements between observational and GCM-calibrated ENMs. First, GCM forecast runs typically include an ensemble of runs (identical experiments with slightly different starting conditions) to take into account natural climate variability. We have used five ensemble members here, recognizing that the choice of ensemble member can result in differences that reflect such variability. Second, strong disagreements in our present-day regional disagreement map (Fig. 3.3) likely reflect regional biases in CCSM4. One region is the Gulf Stream, where CCSM4 has an elevated temperature and salinity bias (Danabasoglu et al. 2012). Another is in the Sea of Okhotsk between Japan and Russia's Kamchatka Peninsula, where CCSM4 has a cold

temperature bias in the northwest (Danabasoglu et al. 2012) that corresponds directly with the region of greatest disagreement between our present-time ENM outputs (Fig. 3.3). Finally, in both the Barents Sea and the Sea of Okhotsk, CCSM4 model biases of maximum sea ice concentration are too high (Jahn et al. 2012).

Consciousness of spatial and variable-based challenges in developing climate models is key to interpreting projections of future distributions of suitable environments for species of interest. The incongruent predictions of suitable present-day habitat for most of the Arctic gadine species call for caution when applying GCM climate projections in forecasting suitable habitat. This is particularly true with respect to finer-scale regional predictions, such as the regions identified above. However, in cases such as that of *Arctogadus glacialis*, for which both sets of ENM projections predict the disappearance of suitable habitat for the species by 2100 (Fig. 3.1), the signal of the predicted trend is probably robust. For such Arctic species, a realistic approach to studying the effects of climate change must begin with sufficient baseline data on current distributions and ecosystem services provided by these species, so that mitigation efforts can be focused on systems most likely to be impacted by climate-change-induced species loss (Post et al. 2009). Both ENMs calibrated using observational and GCM datasets suggest the possibility that extirpated *Arctogadus glacialis* will be replaced in the Arctic by its sister species, *Boreogadus saida* (Fig. 3.1). Unlike *A. glacialis*, *B. saida* is not dependent on sea ice for reproduction (Cohen et al. 1990), such that reduction of Arctic sea ice (Overland & Wang 2013; Vavrus et al. 2012) is less a challenge and more an opportunity. Although our ENMs do not explicitly account for species competition, in habitats where the two species are currently sympatric and partition food resources spatially (Christiansen et al. 2012), *B. saida* may enjoy relaxation in competitive pressure, permitting expansion and exploitation of new food resources.

Both sets of models for IUCN Vulnerable gadine species *Gadus morhua* and *Melanogrammus aeglefinus*, both Atlantic species, indicate northward shifts in mean latitude and expansion in suitable habitat between 2054 and 2100 (Fig. 3.1). The models also indicate suitable habitat in the Pacific for both species, but without clear connectivity in suitable habitat between the Pacific and Atlantic potential ranges for either species (Fig. 3.1). It is unlikely that *G. morhua* and *M. aeglefinus* will face competition from other gadine species as their distributions shift. The most common and widespread Pacific species, *G. chalcogrammus* and *G. macrocephalus*, are predicted to experience a reduction in suitable habitat within their native range, but with new habitats opening up in the Atlantic (Fig. 3.2). However, their ability to expand into Atlantic waters would require dispersal across an even broader swath of unsuitable ocean than their Atlantic counterparts may experience (Fig. 3.2). Suitable habitat for *G. ogac*, an Atlantic species sympatric with *G. morhua* and *M. aeglefinus* with an already limited range, has an uncertain future—the GCM-calibrated ENM predicts a northward expansion, whereas the observation-calibrated ENM predicts that suitable habitat for *G. ogac* will disappear by 2100 (Fig. 3.2). *Micromesistius poutassou*, a species with larval stages that share a similar feeding ecology with the two IUCN Vulnerable species, *G. morhua* and *M. aeglefinus* (Turner 1984), may present the biggest threat; *M. poutassou* is predicted by both ENMs to expand in suitable habitat within its native range, with improved connectivity between populations compared to the present day (Fig. 3.2). Additional problems for *G. morhua* and *M. aeglefinus* (and indeed, for any of the species) may arise if prey species do not experience parallel distributional shifts; previously, researchers employing mechanistic models have predicted a decline in larval *G. morhua* recruitment as cold-temperature-dependent food resources decrease as the oceans warm (Kristiansen et al. 2014).

This study demonstrates both promise and limitations in using GCM climate model data to anticipate effects of future climate change on gadine species' distributions. Particularly, we offer unique methodological insight as to the sources of spatial bias in GCM-calibrated ENMs by comparing biases in environmental inputs with those of ENM results. Considerable uncertainty remains in ENMs projected to future conditions, especially in areas where model bias has been difficult to account for or correct. Still, as the CCSM climate model continues to improve, attempts to understand changes in gadine distributions and predict future distributions of these species will likely also benefit. In the meantime, while remaining conscious of the biases implicit in GCM-based ENM predictions using the methods employed herein, policy makers and researchers still may be able to identify areas of particular concern for gadine species, and better prioritize conservation efforts as the Earth's climate continues to change.

Chapter 4*

Closing the gap between researchers and policymakers:
lessons from the history of fisheries management in the United States

*Owens, HL. 2014. Closing the gap between researchers and policymakers: lessons from the history of fisheries management in the United States. *Society and Natural Resources* 27: 1339-1345. doi: 10.1080/08941920.2014.933930.

Abstract

Increases in sea surface temperature have led to distributional changes in many commercially exploited fish species. These changes have already led to conflict over mackerel fisheries, arising from demand for fair resource apportionment and desire to manage the fishery sustainably. In order to develop adaptable management strategies for complex ocean fishery systems, policymakers and researchers must move beyond a reactive producer–consumer relationship to develop proactive, supportive collaborations. The history of U.S. national fisheries management is presented as an example of this transition. Building cooperative capacity over the last two centuries has led to a more systematic understanding of the oceans, and has led to success in reducing the number of overfished stocks. Similar development of cooperation between policymakers and researchers on an international scale may be the surest way to develop management strategies adaptable enough to withstand challenges posed by future climate change.

Introduction

In the summer of 2010, conflict erupted over northeastern Atlantic Mackerel, a commercially valuable migratory fish species. In August of that year, Iceland and the Faroe Islands raised their mackerel catch quotas from 2,000 to 130,000 tons and from 25,000 to 85,000 tons, respectively (Cendrowicz 2010). The two governments claimed their original quotas were based on obsolete data, from a time when the mackerel fishery extended only marginally into their territory; the new quotas reflected increasing numbers of mackerel migrating into their waters as a result of climate change (Cendrowicz 2010). Scotland, where mackerel is the most economically important fishery, immediately called for action from the European Union (EU), including a blockade of the offending nations at EU ports (BBC News 2010). When the EU did not comply with this request, fishers in the Scottish port of Peterhead blocked offloading of mackerel by a Faroese fishing boat (Cendrowicz 2010). As of mid-2012, a compromise had not yet been reached, despite ongoing negotiations (Hannesson 2013).

Capture fisheries are the last major source of wild food for humans, accounting for 88.6 million tons of food fish in 2010 (FAO 2012). Wild-caught fish are products of an environment that humans cannot control directly, and are subject to changes in predation, food availability, seasonal and yearly temperatures and long-term climate change. The recent northeastern Atlantic mackerel fishery conflict, colorfully dubbed the “Mackerel Wars,” may provide a preview of future conflicts if fisheries resource distributions continue to shift with climate change and fisheries management strategies are not altered in response. A recent study by Cheung and colleagues (2013) indicated that species catch is strongly correlated with sea surface temperature, with mean preferred temperature of exploited species increasing globally by 0.19°C per decade since 1970. The stakeholder demand for rapid policy response to such perceived shifts in the

distribution of fisheries resources is increasing. However, the data necessary for such decisions are largely lacking; in the case of northeastern Atlantic Mackerel, long-term data on migrations in disputed areas and perceived changes therein to date do not yet exist (Hannesson 2013).

The demand for decision making with incomplete data is a central issue in natural resource management. Policymakers demand the most complete data available from researchers, but must also move beyond demanding specific, single-issue facts in reaction to specific issues to supporting new avenues of scientific inquiry, including the exploration of complex ocean systems. This essay examines the history of marine fisheries management in the United States from the late 19th century, through the mid-20th century to the present, in order to illustrate the benefits of this transition. The factors and processes that produce healthy, productive ocean ecosystems and fisheries are complicated. Understanding these systems in a future of increasing uncertainty will require cooperation among policymakers and researchers, and must remain a priority in order to successfully implement dynamic management strategies that will have the potential to adapt to emerging challenges posed by climate change.

Birth of the U.S. Commission of Fish and Fisheries

The United States government became interested in management of the nation's fisheries in the late 19th century in response to noticeable declines in the northeastern fishery, the country's primary source of commercial fish (Hobart 1995). Although technological innovations in fish harvesting such as trap-nets were suggested as a cause, positive evidence was lacking (Baird 1873). Still, several states attempted to outlaw such fishing methods, but the legislation foundered, and it was determined that further scientific research into the problem would be necessary (Baird 1873). As a result, in 1871, Congress created the U.S. Commission of Fish and

Fisheries, originally consisting of one man, Spencer Fullerton Baird (Hobart 1995). Baird's duty was to investigate reports of fisheries declines in the northeastern United States and offer insight on possible causes and solutions to this problem (Hobart 1995).

Two years later, Baird submitted a report to Congress on his findings (Baird 1873), identifying several possible causes for fisheries diminution. Bluefish overpredation was one possible culprit—according to one particularly graphic account retold by Baird (1873), “Sometimes among a school of herring or menhaden thousands of blue-fish will be seen, biting off the tail of one and then another, destroying ten times as many fish as they really need for food, and leaving in their track the surface of the water covered with the blood and fragments of the mangled fish.” Climate variability may have contributed to this phenomenon—bluefish are a warm-water species that underwent drastic changes in size and abundance in the Northeast from year to year, possibly as a result of variable weather conditions (Baird 1873). Counterintuitive to the previous point, several fishermen suggested in Baird's interviews that cold snaps had proven detrimental to populations of several food fish species (Baird 1873).

Despite reports of the vagaries of climate variability and voracious predators, the timing of fisheries decline suggested that while natural threats to fisheries existed, the heart of the problem was overfishing. Decreases in fisheries were first observed in the 1850s and were gradual until 1865, when yields began decreasing more sharply, coincident with introduction and increasing use of trap- and pound-nets (Baird 1873). Baird concluded that overpredation by bluefish and overfishing, in combination, were the main causes of fisheries decline (Baird 1873). Since it was undesirable to destroy bluefish to save species they preyed upon (bluefish were a valuable food fish in their own right), he concluded that the easiest and most expedient way to

preserve northeastern commercial fish populations would be to ban, or at least control, use of fish traps (Baird 1873).

His recommendation was extremely controversial, owing to vast improvements in catch per unit effort and resulting economic benefits from using fish traps (Baird 1873). J. Talbot Pitman, a Providence lawyer in favor of banning the traps, attested, “If, as now alleged, [fish are] diminishing gradually from other causes, and will ultimately disappear, because there is a tradition that they had once before disappeared, about one hundred years ago, and without any known cause . . . let us not hasten the evil day, by reducing their numbers every year while they do remain, through means of these traps” (Baird 1873). Nathaniel Atwood, fisherman and entrepreneur (Del Deo 2010), testified on behalf of fish trappers, making the point that fisheries were important resources, and regulation would be detrimental to future economic growth in the region (Baird 1873). Furthermore, Atwood presented an anecdotal history of natural declines and recoveries in many fish populations of the region that had little, if anything, to do with human exploitation, adding, “it is my candid opinion that man cannot destroy a race of fishes” (Baird 1873).

This discourse highlights the struggle of policymakers to balance preservation of natural resources with economic growth, confounded by uncertainty as to population structure and the role of fluctuations in climate in perceived population declines. From Atwood’s perspective, the recent declines in northeastern fisheries were just a small snapshot of a larger cycle of variability, and the burden of proof of harm lay with would-be regulators of fish traps. Pitman, on the other hand, took a precautionary stance, arguing that previous climate variations did not excuse the actions of trap fishermen, and that the burden of proof of benign effect lay with would-be users of fish traps. In the end, Congress found Atwood’s arguments more compelling, and used

historical variability of fish populations and lack of solid evidence of the harm of fish traps as a justification for continuing the unregulated trapping of food fishes (Hobart 1995). This debate highlights the reluctance of policymakers at the time to fully acknowledge the complicated interplay of factors influencing fisheries' health.

The creation of the U.S. Fish Commission was not only in response to declines in northeastern fishery stocks, but also a response to a need for information vital to international fishing treaty negotiations. This process began with the signing of the Washington Treaty of 1871 (Baird 1879), requiring appointment of three fisheries agents: one by the United States, one by Great Britain, and one agreed upon by both parties (Hobart 1995). These three agents were to convene in Halifax at their earliest convenience to determine equitable coastal fishing rights for the United States and the Dominion of Canada (Baird 1879). Baird (1879) attended the Halifax Fisheries Commission and reported that testimony of each country's witnesses was, "as might be imagined, very opposite as to the value of the privileges conceded by the two contracting parties." He expressed frustration that the United States was extremely deficient in the necessary statistics to state its case compared to Canada, which had a history of making detailed yearly reports on all aspects of its fisheries (Baird 1879). Britain's superior advance knowledge of Canadian fisheries conferred political power and economic advantage over valuable oceanic resources. The convention found that the United States owed Great Britain more than \$5,000,000 in the day's currency to make amends for unequal fisheries harvesting (Baird 1879).

U.S. Fisheries Commission research programs were expanded rapidly as a result of this defeat, but interest in proactive information gathering on wild fish stocks following the Canadian model gradually waned. It was not until the 1970s that a renewed interest in fisheries

conservation and international resource management efforts brought the need for a shift from reactive, decision-maker-driven scientific research to proactive primary research (Hobart 1995).

Modern U.S. National Fisheries Science and Management

The late 1960s and early 1970s saw an increase not only in environmental concerns (Hobart 1995), but also in researching and developing U.S. marine resources for food and resource extraction and protecting U.S. claims on marine territories (Commission on Marine Science, Engineering and Resources [CMSER] 1969). This period marked a transition from passive, reactionary research on ocean resources to active exploration of the ocean. In 1966, the U.S. Commission on Marine Science, Engineering, and Resources was formed to examine the frontiers of these interests, and to recommend plans for issues ranging from how to utilize marine resources more effectively to how to protect marine environments (CMSER 1969). The efforts of this committee were summed up in *Our Nation and the Sea: A Plan for National Action*, known popularly as the Stratton Report, which summarized key issues relating to marine resources and made recommendations on how to address them (CMSER 1969).

Under recommendation from the Stratton Report, the Bureau of Commercial Fish and Fisheries was reorganized and renamed the National Marine Fisheries Service (NMFS) (Schoning 1973). The stated purpose of the organization was to develop and protect the nation's fisheries resources:

The global oceans, which constitute nearly three-fourths of the surface of our planet, are today the least-understood, the least-developed, and the least-protected parts of our earth. Food from the oceans will increasingly be a key element in the world's fight against hunger. The mineral resources of the ocean beds and of the

oceans themselves are being increasingly tapped to meet the growing world demand. We must understand the nature of these resources, and assure their development without either contaminating the marine environment or upsetting its balance. (Schoning 1973, 3–4)

With this call to action, the NMFS was charged not only with developing a proactive, exploratory research plan, but also with understanding the ocean as a system instead of a collection of relatively independent parts.

In 1976, Congress passed the Fishery Conservation and Management Act of 1976, commonly known as the Magnuson–Stevens Act (MSA), which established a legal framework for fisheries management in U.S. waters (Hobart 1995). The act instituted enforcement of international fishery agreements and established a conservation zone extending 200 miles off U.S. shores, within which the United States reserved the right to manage foreign fishing, stating, “International fishery agreements have not been effective in preventing or terminating the overfishing of these valuable fishery resources” (USC 16 §§ 1801–1884 1976). The act broke U.S. marine fisheries holdings into eight regions, requiring each region to develop management plans for each of its fisheries, and requiring each plan to be based on the best research available (USC 16 §§ 1801–1884 1976). However, the MSA did not provide any guidance on how fishery research was to be conducted or evaluated, or allow for support of useful extra-agency research activities, such as those conducted by academic and nonprofit institutions.

Despite the best intentions of the MSA, U.S. fisheries continued to struggle with impacts of overfishing. Between 1983 and 1993, cod catches declined by 55%, Yellowtail Flounder by 83%, and Haddock by 94% (Hennessey and Healey 2000). The year 1996 marked an important transition, not just for the MSA and U.S. national fisheries, but also for how the federal

government of the United States responded to ecological change: Congress enacted the Sustainable Fisheries Act (SFA) to amend the MSA (Hogarth 2003). The SFA added requirements that overfished stocks be replenished in less than 10 years and that ecosystem-based management practices be considered (U.S. Public Law 104-297 1996). Also, a section codifying fisheries management and research practices was added, including requirements that fisheries observers be trained under the National Sea Grant College program, and that grants be awarded to the University of Hawaii to, among other things, conduct scientific research in the Pacific Insular Area (U.S. Public Law 104-297 1996). These changes marked a shift toward increased cooperation with academic researchers, not only to train government researchers, but also to supply primary data to the NMFS. Some positive progress toward fishery recovery has been made as a result of changes made by the SFA: As of 2012, 56 of the 75 original stocks designated “overfished” are no longer listed as such (NMFS 1997; NMFS 2012).

While the SFA did not completely solve overfishing problems in the United States, it heralded a new era of commitment to improving scientific understanding of fisheries, including effects of environment and overfishing, as well as effects of fisheries on communities. In 2000, President Clinton signed the Oceans Act of 2000, which created the U.S. Commission on Ocean Policy, a think tank tasked with providing the President and Congress with recommendations on how to develop “a coordinated and comprehensive national ocean policy” (Clinton 2000). This action carried through to the George W. Bush administration, and led to reauthorization of the MSA with even more authority and motivation to protect valuable fisheries resources, as well as further plans to include academic institutions in collaborative fishery research (U.S. Public Law 109-479 2007).

U.S. ocean policy continues to evolve, and progress is still being made. In July 2010, President Obama issued an executive order creating a national ocean policy for the United States and the National Ocean Council, a cooperative body charged with the task of translating that policy into action (Obama 2010). In the spring of 2013, the National Ocean Council published its recommendations, which included plans for federal and nongovernmental participation in ocean research, maintaining and expanding infrastructure for data management and research, and managing an interface between science and decision making (National Ocean Council 2013). Most importantly, the policy implementation plan recognizes the need for thorough information gathering, as well as collaborative input from scientists and stakeholders, in order to develop sound fisheries management policies (National Ocean Council 2013).

Conclusion

In a time when climate change is drastically altering the distribution patterns of the world's fishes, it is vital that proactive, cooperative fisheries research remain a priority for all nations. Policymakers must shift their approach to scientific research from consuming raw data on single issues to promoting proactive, cooperative exploration of complex systems. With policy based on solid, holistic scientific understanding of ocean systems, it will be possible to develop dynamic management strategies with potential to adapt to emerging challenges posed by climate change. Only then will truly sustainable international fisheries management be possible, and future conflicts over resources be avoided.

References

- Alagador D, Cerdeira JO, Araujo MB. 2014. Shifting protected areas: scheduling spatial priorities under climate change. *Journal of Applied Ecology* 51:703–713.
- Amante C., Eakins B.W. 2009. *ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis*. NOAA Technical Memorandum NESDIS NGDC–24. National Geophysical Data Center, NOAA. doi:10.7289/V5C8276M.
- Anderson RP, Lew D, Peterson AT. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162:211–232.
- Antonov JI, Seidov D, Boyer TP, Locarnini RA, Mishonov AV, Garcia HE, Baranova OK, Zweng MM, Johnson DR. 2010. *World Ocean Atlas 2009, Vol. 2: Salinity (ed. by S. Levitus)*. NOAA, Washington, DC.
- April J, Hanner RH, Dion-Côté AM, Bernatchez L. 2013. Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. *Molecular Ecology* 22:409–422.
- Árnason E. 2004. Mitochondrial cytochrome *b* DNA variation in the high-fecundity Atlantic cod: trans-Atlantic clines and shallow gene genealogy. *Genetics* 4:1871–1885.
- Baird SF. 1873. *Report on the condition of the sea fisheries of the south coast of New England in 1871 and 1872*. U.S. Government Printing Office, Washington, DC.
- Baird, S. F. 1879. *Report of the commissioner for 1877*. U.S. Government Printing Office, Washington, DC.
- Balushkin, A. Prirodina, V. 2008. Catalog of specimens in the collection of the Zoological Institute, Russian Academy of Sciences. Osteichthyes, Gadiformes. *Explorations of the Fauna of the Seas* 62:1–172. Zoological Institute, Russian Academy of Sciences, St. Petersburg, Russia.
- Barve N, Barve V, Jiménez-Valverde A, Maher SP, Peterson AT, Soberón J, Villalobos F. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* 222:1810–1819.
- BBC News. 2010. Iceland and Faroes face 'mackerel wars' sanctions. BBC News. <http://www.bbc.co.uk/news/uk-scotland-north-east-orkney-shetland-11987778>. Accessed 14 December 2010.
- Bean WT, Stafford R, Brashares JS. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. *Ecography* 35:250–258.

- Bigg GR, Cunningham CW, Ottersen G, Pogson GH, Wadley MR, Williamson P. 2008. Ice-age survival of Atlantic Cod: agreement between palaeoecology models and genetics. *Proceedings of the Royal Society B: Biological Sciences* 275:163–173.
- Blomberg SP, Garland T, Ives AR. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717–745.
- Brander KM. 2005. Cod recruitment is strongly affected by climate when stock biomass is low. *ICES Journal of Marine Science* 62:339–343.
- Byrkjedal I, Rees DJ, Christiansen JS, Fevolden SE. 2008. The taxonomic status of *Theragra finnmarchica* Koefoed, 1956 (Teleostei: Gadidae): perspectives from morphological and molecular data. *Journal of Fish Biology* 73:1183–1200.
- Canino MF, Spies IB, Cunningham KM, Hauser L, Grant WS. 2010. Multiple ice-age refugia in Pacific Cod, *Gadus macrocephalus*. *Molecular Ecology* 19:4339–4351.
- Carr SM, Kivlichan DS, Pepin P, Crutcher DC. 1999. Molecular systematics of gadid fishes: implications for the biogeographic origins of Pacific species. *Canadian Journal of Zoology* 77: 19–26.
- Cendrowicz, L. 2010. The Mackerel Wars: Europe's fish tiff with Iceland. Time Inc. <http://www.time.com/time/world/article/0,8599,2014161,00.html>. Accessed 14 December 2010.
- Cheung WWL, Lam VWY, Sarmiento JL, Kearney K, Watson R, Pauly D. 2009. Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries* 10:235–251.
- Cheung WWL, Watson R, Pauly D. 2013. Signature of ocean warming in global fisheries catch. *Nature* 497:365–368.
- Christiansen JS, Hop H, Nilssen EM, Joensen J. 2012. Trophic ecology of sympatric Arctic gadoids, *Arctogadus glacialis* (Peters, 1872) and *Boreogadus saida* (Lepechin, 1774), in NE Greenland. *Polar Biology* 35:1247–1257.
- Clarke A, Crame JA. 2010. Evolutionary dynamics at high latitudes: speciation and extinction in polar marine faunas. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:3655–3666.
- Clinton, W. J. 2000. Statement on signing the Oceans Act of 2000. The American Presidency Project. <http://www.presidency.ucsb.edu/ws/index.php?pid=1470>. Accessed 15 December 2010.
- Cohen DM, Inada T, Iwamoto T, Scialabba N. 1990. Gadiform fishes of the world (Order Gadiformes): an annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to data. *FAO Species Catalogue* 10:1–442. Food and Agriculture Organization of the United States, Rome.

- Commission on Marine Science, Engineering, and Resources. 1969. Our nation and the sea. A plan for national action. Report of the Commission on Marine Science, Engineering and Resources. U.S. Government Printing Office, Washington, DC.
- Cormon X, Loots C, Vaz S, Vermard Y, Marchal P. 2014. Spatial interactions between Saithe (*Pollachius virens*) and Hake (*Merluccius merluccius*) in the North Sea. *ICES Journal of Marine Science* 71:1342–1355.
- Coulson MW, Marshall HD, Pepin P, Carr SM. 2006. Mitochondrial genomics of gadine fishes: implications for taxonomy and biogeographic origins from whole-genome data sets. *Genome* 49:1115–1130.
- Coyne JA, Orr HA. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.
- Crisp MD, Trewick SA, Cook LG. 2011. Hypothesis testing in biogeography. *Trends in Ecology and Evolution* 26:66–72.
- Danabasoglu G, Bates SC, Briegleb BP, Jayne SR, Jochum M, Large WG, Peacock S, Yeager SG. 2012. The CCSM4 ocean component. *Journal of Climate* 25:1361–1389.
- Del Deo, J. C. 2010. Atwood collection: Overview. Provincetown history preservation project. Town of Provincetown, MA. <http://www.provincetownhistoryproject.com/browse/view?browseCollection=5>. Accessed 20 November 2010.
- Divins DL. 2009. NGDC total sediment thickness of the world’s oceans and marginal seas. Available at: <http://www.ngdc.noaa.gov/mgg/sedthick/sedthick.html>. Accessed 15 December 2009.
- Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29:1969–1973.
- Dulvy NK, Rogers SI, Jennings S, Stelzenmuller V, Dye SR, Skjoldal HR. 2008. Climate change and deepening of the North Sea fish assemblage: a biotic indicator of warming seas. *Journal of Applied Ecology* 45:1029–1039.
- Elith J, Kearney M, Phillips S. 2010. The art of modeling range-shifting species. *Methods in Ecology and Evolution* 1:330–342.
- Engelhard GH, Righton DA, Pinnegar JK. 2014. Climate change and fishing: a century of shifting distribution in North Sea cod. *Global Change Biology* 20:2473–2483.
- Erdmann M. 1999. An account of the first living coelacanth known to scientists from Indonesian waters. *Environmental Biology of Fishes* 54:439–443.
- Erdmann M, Caldwell R, Moosa MK. 1998. Indonesian ‘king of the sea’ discovered. *Nature* 395:335.
- Erdmann M, Caldwell R, Jewett S, Tjakrawidjaja A. 1999. The second recorded living

- coelacanth from north Sulawesi. *Environmental Biology of Fishes* 54:445–451.
- Esselstyn JA, Garcia HJ, Saulog MG, Heaney LR. 2008. A new species of *Desmalopex* (Pteropodidae) from the Philippines, with a phylogenetic analysis of the Pteropodini. *Journal of Mammalogy* 89:815–825.
- Evans MEK, Smith SA, Flynn RS, Donoghue MJ. 2009. Climate, niche evolution, and diversification of the "bird-cage" evening primroses (*Oenothera*, sections *Anogra* and *Kleinia*). *The American Naturalist* 173:225–240.
- FAO. 2012. *Yearbook 2010: Fishery and Aquaculture Statistics*. Food and Agriculture Organization, Rome, Italy.
- FAO. 2013. *FishStatJ - software for fishery statistical time series*. FAO Fisheries and Aquaculture Department. Available at: <http://www.fao.org/fishery/statistics/software/fishstatj/en>. Accessed 30 January 2014.
- FAO. 2014. *FishStatJ - software for fishery statistical time series*. FAO Fisheries and Aquaculture Department. Available at: <http://www.fao.org/fishery/statistics/software/fishstatj/en>. Accessed 18 December 2014.
- Fricke H, Hissmann K. 2000. Feeding ecology and evolutionary survival of the living coelacanth *Latimeria chalumnae*. *Marine Biology* 136:379–386.
- Garrison LP, Link JS. 2000. Fishing effects on spatial distribution and trophic guild structure of the fish community in the Georges Bank region. *ICES Journal of Marine Science* 57:723–730.
- Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, Lawrence DM, Neale RB, Rasch PJ, Vertenstein M, Worley PH, Yang ZL. 2011. The Community Climate System Model version 4. *Journal of Climate* 24:4973–4991.
- Gonzalez EG, Cunha RL, Sevilla RG, Ghanavi HR, Krey G, Bautista JM. 2012. Evolutionary history of the genus *Trisopterus*. *Molecular Phylogenetics and Evolution* 62:1013–1018.
- Gradinger RR, Bluhm B.A. 2004. In-situ observations on the distribution and behavior of amphipods and Arctic cod (*Boreogadus saida*) under the sea ice of the High Arctic Canada Basin. *Polar Biology* 27:595–603.
- Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58:1781–1793.
- Grinnell J. 1917. The niche-relationships of the California Thrasher. *The Auk* 34:427–433.
- Haffer J. 1969. Speciation in Amazon forest birds. *Science* 165:131–137.
- Hall A. 2014. Projecting regional change. *Science* 346:1461–1462.

- Hannesson, R. 2013. Sharing the Northeast Atlantic mackerel. *ICES Journal of Marine Science* 70:259–269.
- Hennessey, T., and M. Healey. 2000. Ludwig's ratchet and the collapse of New England groundfish stocks. *Coastal Management* 28:187–213.
- Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences of the United States of America* 111:2957–2966.
- Heibl C. 2011. phyloclim: integrating phylogenetics and climatic niche modelling. Available at: <http://cran.r-project.org/web/packages/phyloclim/index.html>. Accessed 20 September 2010.
- Hobart WL. 1995. *Baird's legacy: The history and accomplishments of NOAA's National Marine Fisheries Service, 1871–1996*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Seattle, WA.
- Hogarth, W. T., ed. 2003. *Implementing the Sustainable Fisheries Act: Achievements from 1996 to the present*. U.S. National Marine Fisheries Service. http://www.nmfs.noaa.gov/sfa/SFA-Report-FINAL7_1.pdf. Accessed 20 November 2010.
- Holder MT, Erdmann MV, Wilcox TP, Caldwell RL, Hillis DM. 1999. Two living species of coelacanths? *Proceedings of the National Academy of Sciences United States of America* 96:12616–12620.
- Houston DD, Shiozawa DK, Smith BT, Riddle BR. 2014. Investigating the effects of Pleistocene events on genetic divergence within *Richardsonius balteatus*, a widely distributed western North American minnow. *BMC Evolutionary Biology* 14:111.
- IUCN. 2011. *The IUCN Red List of Threatened Species*. International Union for Conservation of Nature. Available at: www.iucnredlist.org. Accessed 14 September 2011.
- IUCN. 2014. *The IUCN Red List of Threatened Species*. International Union for Conservation of Nature. Available at: www.iucnredlist.org. Accessed 20 January 2015.
- Jahn A, Sterling K, Holland MM, Kay JE, Maslanik JA, Bitz CM, Bailey DA, Stroeve J, Hunke EC, Lipscomb WH. 2012. Late-twentieth-century simulation of Arctic sea ice and ocean properties in the CCSM4. *Journal of Climate* 25:1431–1452.
- Jiménez-Valverde A, Lobo JM. 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica* 31:361–369.
- Jordan AD, Moller PR, Nielsen JG. 2003. Revision of the Arctic cod genus *Arctogadus*. *Journal of Fish Biology* 62:1339–1352.
- Kimura S, Inoue T, Sugimoto T. 2001. Fluctuation in the distribution of low-salinity water in the North Equatorial Current and its effect on the larval transport of the Japanese eel. *Fisheries Oceanography* 10:51–60.

- Kristiansen T, Stock C, Drinkwater KF, Curchitser EN. 2014. Mechanistic insights into the effects of climate change on larval cod. *Global Change Biology* 20:1559–1584.
- Leu E, Søreide JE, Hessen DO, Falk-Petersen S, Berge J. 2011. Consequences of changing sea-ice cover for primary and secondary producers in the European Arctic shelf seas: timing, quantity, and quality. *Progress in Oceanography* 90:18–32.
- Li CH, Ortí G. 2007. Molecular phylogeny of clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* 44:386–398.
- Li CH, Ortí G, Zhang G, Lu GQ. 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evolutionary Biology* 7:1–11.
- Link JS, Garrison LP. 2002. Trophic ecology of Atlantic cod *Gadus morhua* on the northeast US continental shelf. *Marine Ecology Progress Series* 227:109–123.
- Liu C, Berry PM, Dawson TP, Pearson RG. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28:385–393.
- Locarnini RA, Mishonov AV, Antonov JI, Boyer TP, Garcia HE, Baranova OK, Zweng MM, Johnson DR. 2010. *World Ocean Atlas 2009*, Vol. 1: *Temperature* (ed. by S. Levitus). NOAA Atlas NESDIS 68. US Government Printing Office, Washington, DC.
- Maggs CA, Castilho R, Foltz D, Henzler C, Jolly MT, Kelly J, Olsen J, Perez KE, Stam W, Väinölä R, Viard F, Wares J. 2008. Evaluating signatures from glacial refugia for North Atlantic marine taxa. *Ecology* 89:S108–S122.
- Manthey J, Campbell LP, Saupe EE, Soberón J, Hensz CM, Myers CE, Owens HL, Ingenloff K, Peterson AT, Barve N, Lira-Noriega A, Barve V. 2015. A test of niche centrality as a determinant of population trends and conservation status in threatened and endangered North American birds. *Endangered Species Research* 26:201–208.
- Matzke NJ. 2013a. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* 5:242–248.
- Matzke NJ. 2013b. *BioGeoBEARS: biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts*. PhD Thesis, University of California, Berkeley, CA.
- Mecklenburg CW, Møller PR, Steinke D. 2011. Biodiversity of arctic marine fishes: taxonomy and zoogeography. *Marine Biodiversity* 41:109–140.
- Merow C, Smith MJ, Silander JA. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36:1058–1069.
- Mueller UG, Mikheyev AS, Hong E, Sen R, Warren DL, Solomon SE, Ishak HD, Cooper M, Miller JL, Shaffer KA. 2011. Evolution of cold-tolerant fungal symbionts permits winter

- fungiculture by leafcutter ants at the northern frontier of a tropical ant–fungus symbiosis. *Proceedings of the National Academy of Sciences* 108:4053–4056.
- Mueter FJ, Bond NA, Ianelli JN, Hollowed AB. 2011. Expected declines in recruitment of Walleye Pollock (*Theragra chalcogramma*) in the eastern Bering Sea under future climate change. *ICES Journal of Marine Science* 68:1284–1296.
- Mueter, F.J. Litzow, M.A. 2008. Sea ice alters the biogeography of the Bering Sea continental shelf. *Ecological Applications* 18:309–320.
- National Ocean Council. 2013. National ocean policy implementation plan. http://www.whitehouse.gov/sites/default/files/national_ocean_policy_implementation_plan.pdf. Accessed 14 September 2013.
- National Marine Fisheries Service. 1997. Status of fisheries of the United States. <http://www.nmfs.noaa.gov/sfa/statusoffisheries/Archives/StatusofFisheriesReportCongress1997.htm>. Accessed 22 November 2013.
- National Marine Fisheries Service. 2012. Status of stocks 2012. http://www.nmfs.noaa.gov/sfa/statusoffisheries/2012/2012_SOS_RTC.pdf. Accessed 22 November 2013.
- Nelson JS. 2006. *Fishes of the World*. John Wiley and Sons, Inc, Hoboken, NJ.
- NOAA. 1999. *World Ocean Atlas 1998*. National Oceanographic Data Center, Silver Spring, MD. (3CD-ROM set).
- Nyári ÁS, Reddy S. 2013. Comparative phylogenetic analysis and evolution of ecological niches in the Scimitar Babblers (Aves: Timaliidae: *Pomatorhinus*). *PLoS ONE* 8:e55629.
- Obama B. 2010. Executive order 13547 of July 19, 2010: Stewardship of the ocean, our coasts, and the Great Lakes. Fed. Reg. 75:43023.
- O'Connor MI, Bruno JF, Gaines SD, Halpern BS, Lester SE, Kinlan BP, Weiss JM. 2007. Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences* 104:1266–1271.
- O'Regan M, Williams CJ, Frey KE, Jakobsson M. 2011. A synthesis of the long-term paleoclimatic evolution of the Arctic. *Oceanography* 3:66–80.
- Ottersen G, Hjermann DØ, Stenseth NC. 2006. Changes in spawning stock structure strengthen the link between climate and recruitment in a heavily fished cod (*Gadus morhua*) stock. *Fisheries Oceanography* 15:230–243.
- Overland JE, Wang M. 2013. When will the summer Arctic be nearly sea ice free? *Geophysical Research Letters* 40:2097–2101.

- Owens HL. 2015. Evolution of codfishes (Teleostei: Gadinae) in geographical and ecological space: evidence that physiological limits drove diversification of subarctic fishes. *Journal of Biogeography* doi:10.1111/jbi.12483.
- Owens HL, Campbell LP, Dornak LL, Saupe EE, Barve N, Soberón J, Ingenloff K, Lira-Noriega AS, Hensz CM, Myers CE, Peterson AT. 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling* 263:10–18.
- Parmesan C, Yohe G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421:37–42.
- Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT. 2007. Predicting species distributions from a small number of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34:102–117.
- Perry AL, Low PJ, Ellis JR, Reynolds JD. 2005. Climate change and distribution shifts in marine fishes. *Science* 308:1912–1915.
- Peterson AT. 2011. Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography* 38:817–827.
- Peterson A, Nakazawa Y. 2007. Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography* 17:135–144.
- Peterson AT, Nyári ÁS. 2007. Ecological niche conservatism and Pleistocene refugia in the Thrush-like Mourner, *Schiffornis* sp., in the Neotropics. *Evolution* 62:173–183.
- Peterson A, Soberón J, Pearson RG, Anderson RP, Martínez-Meyer E, Nakamura M, Araújo MB. 2011. *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton, NJ.
- Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231–259.
- Pogson GH, Mesa KA. 2004. Positive Darwinian selection at the pantophysin (Pan I) locus in marine gadid fishes. *Molecular Biology and Evolution* 21:65–75.
- Posada, D. Crandall, K.A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Post E, Forchhammer MC, Bret-Harte MS, Callaghan TV, Christensen TR, Elberling B, Fox AD, Glig O, Hik DS, Høye TT, Ims RA, Jeppesen E, Klein DR, Madsen J, McGuire AD, Rysgaard S, Schindler DE, Stirling I, Tamstorf MP, Tyler NJC, van der Wal R, Welker J, Wookey PA, Schmidt NM. 2009. Ecological dynamics across the Arctic associated with recent climate change. *Science* 325:1355–1358.

- Provan J. 2013. The effects of past, present and future climate change on range-wide genetic diversity in northern North Atlantic marine species. *Frontiers of Biogeography* 5:60–66.
- Pulliam R. 2000. On the relationship between niche and distribution. *Ecological Letters* 3:349–361.
- Quattrini AM, Georgian SE, Byrnes L, Stevens A, Falco R, Cordes EE. 2013. Niche divergence by deep-sea octocorals in the genus *Callogorgia* across the continental slope of the Gulf of Mexico. *Molecular Ecology* 22:4123–4140.
- R Core Team. 2012. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raxworthy CJ, Martinez-Meyer E, Horning N, Nussbaum RA, Schneider GE, Ortega-Huerta MA, Peterson AT. 2004. Predicting distributions of known and unknown reptile species in Madagascar. *Nature* 426:837–841.
- Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3:217–223.
- Riahi K, Rao S, Krey V, Cho C, Chirkov V, Fischer G, Kindermann G, Nakicenovic N, Rafaj P. 2011. RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change* 109:33–57.
- Roa-Varón A, Ortí G. 2009. Phylogenetic relationships among families of Gadiformes (Teleostei, Paracanthopterygii) based on nuclear and mitochondrial data. *Molecular Phylogenetics and Evolution* 52:688–704.
- Ronquist F, Sanmartín I. 2011. Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics* 42:441–464.
- Saupe E, Barve V, Myers C, Soberón J, Barve N, Hensz C, Peterson A, Owens H, Lira-Noriega A. 2012. Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling* 237:11–22.
- Schoning RW. 1973. *Report of the National Marine Fisheries Service for the calendar years 1970 and 1971*. U.S. Government Printing Office, Washington, DC.
- Sinclair SJ, White MD, Newell GR. 2010. How useful are species distribution models for managing biodiversity under future climates? *Ecology and Society* 15:8.
- Siqueira MF, Durigan G, de Marco Júnior P, Peterson AT. 2009. Something for nothing: using landscape similarity and ecological niche modeling to find rare species. *Journal for Nature Conservation* 17:25–32.
- Smith JLB. 1939. A living fish of the Mesozoic type. *Nature* 143:455–456.

- Smith SA, Donoghue MJ. 2010. Combining historical biogeography with niche modeling in the *caprifolium* clade of *Lonicera* (Caprifoliaceae, Dipsacales). *Systematic Biology* 59:322–341.
- Soberón J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecological Letters* 10:1–9.
- Soberón J, Nakamura M. 2009. Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences United States of America* 106:19644–19650.
- Soberón J, Peterson AT. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics* 2:1–10.
- Springer VG. 1999. Are the Indonesian and western Indian Ocean coelacanths conspecific: a prediction. *Environmental Biology of Fishes* 54:453–456.
- Stobbs R. 2002. The coelacanth and the Comores: challenging the myth. *Science in Africa* 16: <http://www.scienceinafrica.co.za/2002/april/home.htm>. Accessed 1 Feb 2011.
- Stockwell D, Peters D. 1999. The GARP modeling systems: problems and solutions to automated spatial prediction. *International Journal of Geographic Information Science* 13:143–158.
- Sunday JM, Bates AE, Dulvy NK. 2012. Thermal tolerance and the global redistribution of animals. *Nature Climate Change* 2:686–690.
- Svetovidov A. 1948. Gadiformes. *Fauna of the USSR* 9:1–222.
- Teletchea F, Laudet V, Hänni C. 2006. Phylogeny of the Gadidae (sensu Svetovidov, 1948) based on their morphology and two mitochondrial genes. *Molecular Phylogenetics and Evolution* 38:189–199.
- Troia MJ, Gido KB. 2014. Towards a mechanistic understanding of fish species niche divergence along a river continuum. *Ecosphere* 5:41.
- Turner JT. 1984. The feeding ecology of some zooplankters that are important prey items of larval fish. *NOAA Technical Report, National Marine Fisheries Service*, 7: 1–28.
- USC 16 §§ 1801–1884. 1976. Fishery conservation and management act: Magnuson–Stevens act.
- U.S. Public Law 104-297. 1996. Sustainable fisheries act.
- U.S. Public Law 109-479. 2007. Magnuson–Stevens reauthorization act of 2006.
- Vavrus SJ, Holland MM, Jahn A, Bailey DA, Blazey BA. 2012. Twenty-first-century Arctic climate change in CCSM4. *Journal of Climate* 25:2696–2710.
- Vergés A, Steinberg PD, Hay ME, Poore AG, Campbell AH, Ballesteros E, Heck KL, Booth DJ, Coleman MA, Feary DA. 2014. The tropicalization of temperate marine ecosystems:

- climate-mediated changes in herbivory and community phase shifts. *Proceedings of the Royal Society B: Biological Sciences* 281:20140846.
- Vieites DR, Nieto-Roman S., Wake DB. 2009. Reconstruction of the climate envelopes of salamanders and their evolution through time. *Proceedings of the National Academy of Sciences of the United States of America* 106:19715–19722.
- VLIZ. 2009. *Longhurst biogeographical provinces*. Flanders Marine Institute. Available at: <http://www.marineregions.org/>. Accessed 4 February 2014.
- Warren DL, Glor RE, Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* 33:607–611.
- Weijerman M, Lindeboom H, Zuur AF. 2005. Regime shifts in marine ecosystems of the North Sea and Wadden Sea. *Marine Ecology Progress Series* 298:21–39.
- Weir JT, Schluter D. 2004. Ice sheets promote speciation in boreal birds. *Proceedings of the Royal Society of London B: Biological Sciences* 271:1881–1887.
- Welch HE, Bergmann MA, Siferd TD, Martin KA, Curtis MF, Crawford RE, Conover RJ, Haakon H. 1992. Energy flow through the marine ecosystem of the Lancaster Sound region, Arctic Canada, *Arctic* 45:343–357.
- Wiley EO, McNyset KM, Peterson AT, Robins CR, Stewart AM. 2003. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography* 16:120–127.
- Yesson C, Culham A. 2006a. Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. *Systematic Biology* 55:785–802.
- Yesson C, Culham A. 2006b. A phyloclimatic study of *Cyclamen*. *BMC Evolutionary Biology* 6:72.
- Zurell D, Elith J, Schröder B. 2012. Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. *Diversity and Distributions* 18:628–634.
- Zwickl DJ. 2006. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. PhD Thesis, University of Texas, Austin, TX.

Appendix I

Detailed phylogenetic methods from (Chapter 2; Owens 2015)

Gene sequencing

PCR reactions for *I2S*, *CYTB* and *ZIC1* consisted of 1 μ L extract, 1.5 μ L TAQ buffer, 0.125 μ L TAQ, 0.75 μ L dNTP and MgCl₂, 0.25 μ L forward and reverse primers, and dH₂O, to a final volume of 13 μ L. For *RAG1*, reactions consisted of 5 μ L extract, 1.25 μ L forward and reverse primers, one PuReTaq Ready-To-Go™ PCR Bead (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA), and 17.5 μ L dH₂O, for a final volume of 25 μ L. *RAG1* was amplified using a custom PCR profile, beginning with an initial cycle of 180 s at 94 °C, followed by 10 cycles of 45 s at 94 °C, 45 s at 52 °C, and 75 s at 72 °C, followed by 30 cycles of 45 s at 94 °C, 45 s at 47 °C, and 75 s at 72 °C, with a final extension of 7 min at 72 °C. PCR bands were visualized on 1% low melting-point agar gels, and PCR products were purified with QIAquick PCR purification kits (Qiagen, Valencia, CA, USA). DNA sequencing was done using an ABI BigDye Terminator ver. 3.1 cycle sequencing kit (Perkin Elmer, Waltham, MA, USA) for dye-terminator chemistry following manufacturer's instructions. Cycle-sequencing reactions were precipitated in 75% solution of ethanol following ABI protocol. Dried cycle-sequencing reactions were resuspended and electrophoresed on an AB 3730 Genetic Analyzer (Applied Biosystems, Grand Island, NY, USA). Both strands were sequenced to verify accuracy of sequences.

Phylogenetic inference

The Akaike information criterion (AIC) as implemented in JMODELTEST 2.1 (Posada & Crandall, 1998) was used to determine appropriate data partitioning and model selection for phylogenetic inference. Nine data partitions were tested, and those with the simplest model of evolution were used to conduct several preliminary BEAST analyses (Appendix S1: Table S3). Analyses failed to converge after 5 million generations due to CG and GT transversions in *I2S* and AT

transversions in *RAG1* position 3 as indicated by TRACER 1.5 (<http://beast.bio.ed.ac.uk/tracer>).

As such, simpler models were used in final BEAST analysis. The standard deviation of the *RAG1* uncorrelated lognormal clock model in preliminary BEAST analysis included 0, so a strict clock model was applied in the final analyses of that gene.

For final BEAST analyses, two independent runs of 10^6 generations were conducted, and the resulting tree and log files for each run were combined using LOGCOMBINER 1.6.1 (<http://beast.bio.ed.ac.uk/logcombiner>). Convergence of model parameter values was assessed by examination of convergence and likelihood stationarity in TRACER 1.5 from combined posterior samples to ensure adequate mixing of the Markov chain Monte Carlo (MCMC). A cutoff of 20% was used for the burn-in. The posterior probability density of the combined tree and log files was summarized as a maximum clade credibility tree using TREEANNOTATOR 1.6.1 (<http://beast.bio.ed.ac.uk/TreeAnnotator>).

Appendix II

Supplementary tables from (Chapter 2; Owens 2015)

Appendix II.A Genetic resources used in analysis. Bold voucher numbers indicate tissues sequenced in this study.

Species	Voucher number	Sample number	12S	CYTB	ZIC1	RAG1
<i>Arctogadus glacialis</i>	ZMUC P375407		KP644355	KP644321	KP644377	
	ZMUC P375424		KP644354			
<i>Boreogadus saida</i>	ZMUC P375316		KP644366	KP644323	KP644375	
	ZMUC P375149		KP644367	KP644322		
<i>Eleginus gracilis</i>	No voucher			AB078150		
	UW 44916					FJ215236
<i>Gadus chalcogrammus</i>	KU 28401	KU 2067	KP644356	KP644332	KP644379	KP644396
	KU 28196	KU 3212	KP644357	KP644331		
	UW 47697					FJ215294
<i>Gadus macrocephalus</i>	KU 28470	KU 3234	KP644362	KP644335		
	KU 28473	KU 3250	KP644363	KP644338		
	UW 47711					FJ215241
<i>Gadus morhua</i>	MCZ 155780	KU 2937	KP644360	KP644333	KP644382	KP644390
	No voucher	KU 3776	KP644361	KP644334	KP644384	
<i>Gadus ogac</i>	ZMUC P374606		KP644358	KP644336	KP644374	
	ZMUS P375315		KP644359	KP644337	KP644386	
<i>Gaidropsarus ensis</i>	ZMUC P375409		KP644342		KP644388	KP644391
	No voucher			DQ174049		
<i>Lota lota</i>	ZMUC P375475		KP644347	KP644340	KP644373	KP644392
<i>Melanogrammus aeglefinus</i>	MCZ 155777	KU 2939	KP644364	KP644328	KP644371	
	No voucher	KU 3773	KP644365	KP644327	KP644383	
	No voucher					FJ215262
<i>Merlangius merlangus</i>	No voucher	KU 3774	KP644353	KP644320	KP644372	KP644393
<i>Microgadus proximus</i>	KU 23722	KU 9491	KP644370		KP644387	
	KU 23722	KU 824	KP644369			
	No voucher			DQ365944		
	UW 74300					FJ215274
<i>Microgadus tomcod</i>	KU 34104	KU 5884	KP644368	KP644319	KP644380	
<i>Micromesistius poutassou</i>	No voucher	KU 3777	KP644348	KP644326	KP644385	
<i>Molva molva</i>	ZMUC P373715					KP644394
	ZMUC P374175		KP644346	KP644341	KP644389	
<i>Pollachius pollachius</i>	ZMUC P375896		KP644352	KP644324	KP644378	

	ZMUC P375793		KP644351	KP644325	
<i>Pollachius virens</i>	KU 28071	KU 359			KP644395
	KU 28071	KU 9105	KP644350	KP644330	KP644376
	No voucher	KU 3779	KP644349	KP644329	
<i>Trisopterus esmarkii</i>	ZMUC P375906		KP644344		
	MNHN 2005.1674			EU492145	
<i>Trisopterus luscus</i>	ZMUC P375806		KP644343		
	MNHN 2005.1677			EU492067	
<i>Trisopterus minutus</i>	ZMUC P373712		KP644345	KP644339	KP644381

Appendix II.B. Modeltest results and evolutionary models used in analysis.

Gene	Partition	AIC _c	BIC	Model used	Clock	Notes
<i>12S</i>	123pos	GTR+I+G	GTR+I+G	HKY+I+G	UCLN	poor convergence under GTR (CG and GT transversions)
<i>CYTB</i>	1pos	GTR+I+G	K80+G			
<i>CYTB</i>	2pos	GTR+I	F81			
<i>CYTB</i>	12pos	HKY+I+G	HKY+I+G	HKY+G	UCLN	
<i>CYTB</i>	3pos	GTR+I+G	HKY+G	HKY+G	UCLN	
<i>ZIC1</i>	12pos	GTR	JC	HKY	UCLN	
<i>ZIC1</i>	3pos	HKY	HKY	HKY	UCLN	
<i>RAG1</i>	12pos	GTR+G	HKY+I	HKY	Strict	
<i>RAG1</i>	3pos	GTR+I	GTR	HKY	Strict	poor convergence under GTR (AT transversion), UCLN StDev includes 0, justifying use of strict clock

AIC_c, Akaike information criterion corrected for small sample size; BIC, Bayesian information criterion, UCLN, uncorrelated lognormal.

Appendix II.C. Sources of raw occurrence data with numbers of records. Natural History Museum, London (NHMUK), Zoological Institute at the Russian Academy of Sciences (ZIN), Global Biodiversity Information Facility (GBIF), and Ocean Biogeographic Information System (OBIS).

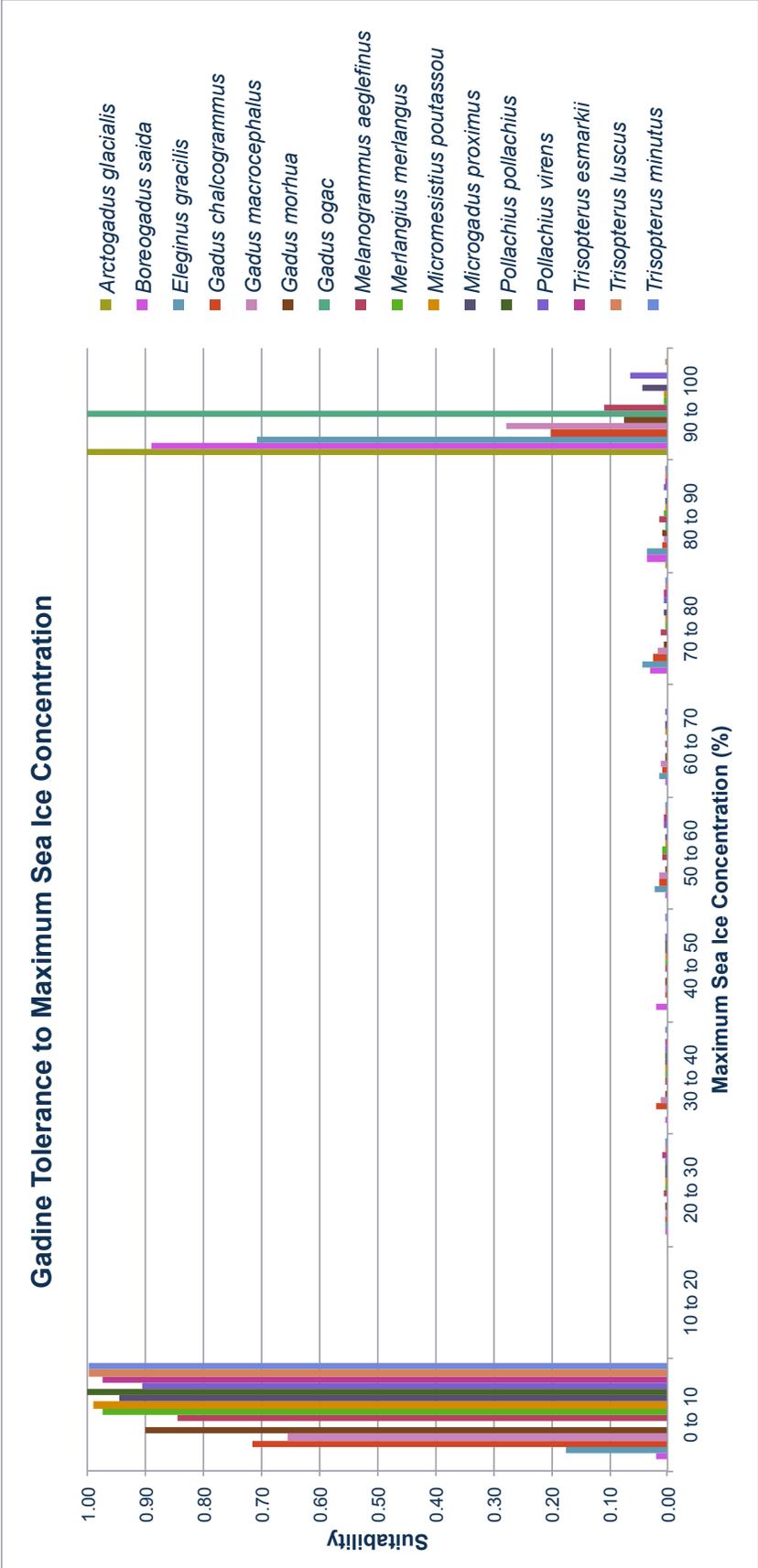
Species	Source			
	GBIF	OBIS	NHMUK	ZIN
<i>Arctogadus glacialis</i>	147	297		18
<i>Boreogadus saida</i>	6510	18301	12	170
<i>Eleginus gracilis</i>	489	493		9
<i>Gadus chalcogrammus</i>	30589	16686	4	20
<i>Gadus macrocephalus</i>	15292	19128	4	4
<i>Gadus morhua</i>	76230	743857	4	36
<i>Gadus ogac</i>	3245	615	1	5
<i>Melanogrammus aeglefinus</i>	25934	486259	15	9
<i>Merlangius merlangius</i>	12438	687911	20	12
<i>Microgadus proximus</i>	172	449		
<i>Microgadus tomcod</i>	369	1980		1
<i>Micromesistius poutassou</i>	654	45968	11	
<i>Pollachius pollachius</i>	9377	5708	1	
<i>Pollachius virens</i>	21814	79790		4
<i>Trisopterus esmarkii</i>	823	136361	5	2
<i>Trisopterus luscus</i>	4996	21570	14	
<i>Trisopterus minutus</i>	3560	98699	22	

Appendix II.D. Range coding for biogeographical range analysis.

	Western Atlantic (W)	Eastern Atlantic (E)	Arctic (A)	Pacific (P)
<i>Arctogadus glacialis</i>	0	0	1	0
<i>Boreogadus saida</i>	0	0	1	0
<i>Gadus macrocephalus</i>	0	0	0	1
<i>Gadus ogac</i>	1	0	0	0
<i>Gadus morhua</i>	1	1	0	0
<i>Theragra chalcogramma</i>	0	0	0	1
<i>Pollachius pollachius</i>	0	1	0	0
<i>Pollachius virens</i>	1	1	0	0
<i>Melanogrammus aeglefinus</i>	1	1	0	0
<i>Micromesistius poutassou</i>	1	1	0	0
<i>Merlangius merlangus</i>	0	1	0	0
<i>Eleginus gracilis</i>	0	0	0	1
<i>Microgadus proximus</i>	0	0	0	1
<i>Microgadus tomcod</i>	1	0	0	0
<i>Trisopterus esmarkii</i>	0	1	0	0
<i>Trisopterus minutus</i>	0	1	0	0
<i>Trisopterus luscus</i>	0	1	0	0

Appendix III

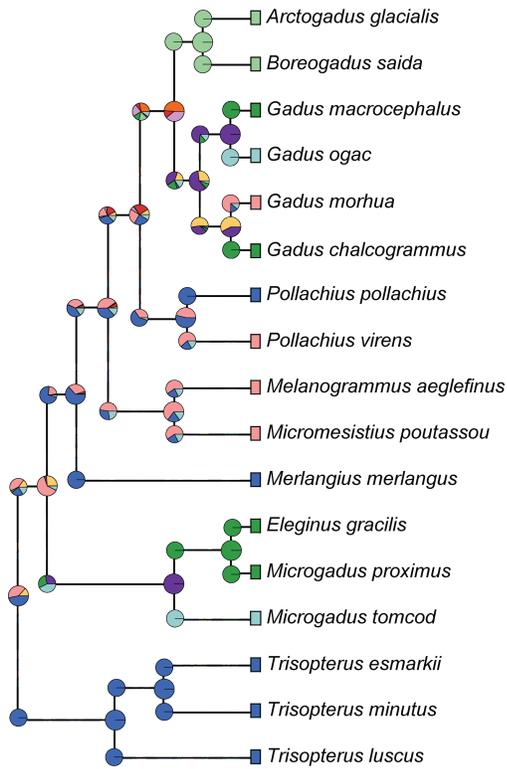
Supplementary figures from (Chapter 2; Owens 2015)



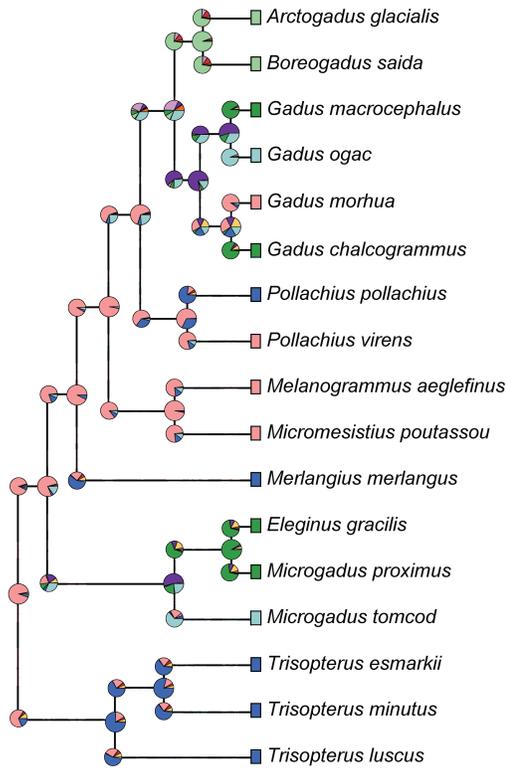
Appendix III.A. Gadine tolerance to maximum sea ice concentration. Suitability of binned maximum sea ice concentration was estimated from ecological niche models of each species combined with maximum sea ice concentration using the predicted niche occupancy function in the PHYLOCLIM package (Heibl 2011) in R (R Core Team, <http://www.r-project.org/>).

Appendix III.B Biogeographical range reconstructions under alternate models of dispersal, extinction and cladogenesis with unconstrained dispersal.

(a). DEC



(b). BAYAREA-like



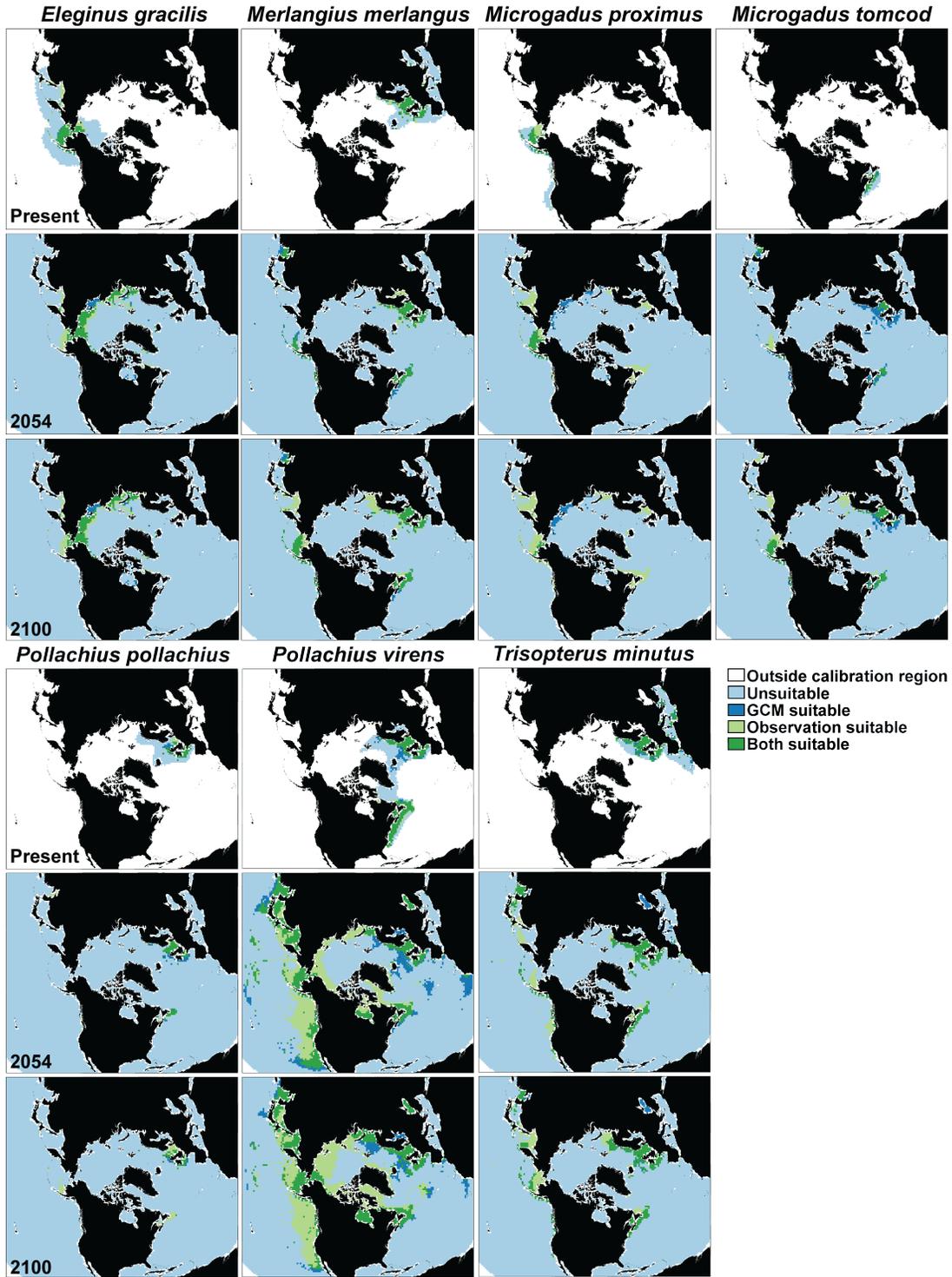
Key

- Western Atlantic (W)
- Eastern Atlantic (E)
- Arctic (A)
- Pacific (P)
- WE
- EA
- EP
- AP
- WA
- WP

Appendix IV

Supplementary figures from (Chapter 3)

Appendix IV.A. Modeled suitable habitat extent for GCM and observation-calibrated ecological niche models of *Eleginus gracilis*, *Merlangius merlangus*, *Microgadus proximus*, *Microgadus tomcod*, *Pollachius pollachius*, *Pollachius virens*, and *Trisopterus minutus* through time. Colored area in the present denotes the calibration region; white areas are outside calibration regions. Future time slices are projected across the Northern Hemisphere to allow estimation of potential range expansions.



Appendix V

R code for comparison statistics from (Chapter 3)

```

library(gtools);
library(raster);
library(sfsmisc);
library(RColorBrewer);

#####
# Threshold raw niche models to 95% training presence #
#####

#Get observation medians
setwd("~/Desktop/NCAR_Paper/CCSM4/PresentObservation/Medians/")
obsList <- list.files(pattern="*.asc")
rastObsList <- vector("list", length(obsList));
count <- 1;
while (count <= length(obsList)){
  rastObsList[[count]] <- raster(obsList[[count]]);
  count <- count + 1;
}

#Get model medians
setwd("~/Desktop/NCAR_Paper/CCSM4/PresentModel/Medians/")
modList <- list.files(pattern="*.asc")
rastModList <- vector("list", length(modList));
count <- 1;
while (count <= length(modList)){
  rastModList[[count]] <- raster(modList[[count]]);
  count <- count + 1;
}
rm(modList, obsList)

#Get 95% thresholds
##Get point data
setwd("~/Desktop/NCAR_Paper/CCSM4/Localities/");
fileList <- list.files(pattern="*.csv");
pointDataList <- vector("list", length(fileList));
count <- 1;
while (count <= length(fileList)){
  pointDataList[[count]] <- read.csv(fileList[[count]]);
  count <- count + 1;
}

modPointData <- pointDataList
obsPointData <- pointDataList
rm(pointDataList)

#Get 95% suitability values
threshObsList <- vector("list", length(obsPointData))
threshObsValues <- vector("list", length(obsPointData))
count <- 1
while (count <= length(obsPointData)){

```

```

  threshObsList[[count]] <- cbind(obsPointData[[count]][,2:3], extract(rastObsList[[count]],
obsPointData[[count]][,2:3]))
  colnames(threshObsList[[count]]) <- c("Longitude", "Latitude", "SuitabilityValue")
  threshObsList[[count]] <- threshObsList[[count]][order(threshObsList[[count]]$SuitabilityValue,
na.last=NA, decreasing = T),]
  threshObsList[[count]] <-
threshObsList[[count]][1:as.integer(length(threshObsList[[count]]$Longitude)*.95),]
  threshObsValues[[count]] <- min(threshObsList[[count]]$SuitabilityValue)
  count <- count + 1
}

threshModList <- vector("list", length(modPointData))
threshModValues <- vector("list", length(modPointData))
count <- 1
while (count <= length(modPointData)){
  threshModList[[count]] <- cbind(modPointData[[count]][,2:3], extract(rastModList[[count]],
modPointData[[count]][,2:3]))
  colnames(threshModList[[count]]) <- c("Longitude", "Latitude", "SuitabilityValue")
  threshModList[[count]] <- threshModList[[count]][order(threshModList[[count]]$SuitabilityValue,
na.last=NA, decreasing = T),]
  threshModList[[count]] <-
threshModList[[count]][1:as.integer(length(threshModList[[count]]$Longitude)*.95),]
  threshModValues[[count]] <- min(threshModList[[count]]$SuitabilityValue)
  count <- count + 1
}

#Reclassify suitability surfaces
reclassifiedObsRasters <- vector("list", length(obsPointData))
count <- 1
while (count <= length(obsPointData)){
  reclass <- c(-Inf, threshObsValues[[count]]*.999, 0, threshObsValues[[count]]*.999, Inf, 1)
  reclassMatrix <- matrix(reclass, ncol=3, byrow=TRUE)
  reclassifiedObsRasters[[count]] <- reclassify(rastObsList[[count]], reclassMatrix, NAflag = -9999)
  count <- count + 1
}
rm(reclass, reclassMatrix, rastObsList, threshObsList, obsPointData, threshObsValues)

#Reclassify suitability surfaces
reclassifiedModRasters <- vector("list", length(modPointData))
count <- 1
while (count <= length(modPointData)){
  reclass <- c(-Inf, threshModValues[[count]]*.999, 0, threshModValues[[count]]*.999, Inf, 1)
  reclassMatrix <- matrix(reclass, ncol=3, byrow=TRUE)
  reclassifiedModRasters[[count]] <- reclassify(rastModList[[count]], reclassMatrix, NAflag = -9999)
  count <- count + 1
}
rm(reclass, reclassMatrix, rastModList, threshModList, modPointData, threshModValues)

#Trim both to M
setwd("~/Desktop/NCAR_Paper/CCSM4/CodMs/ASCIIs/");
mList <- list.files(pattern="*.asc");

```

```

maskedModRasters <- vector("list", length(reclassifiedModRasters))
maskedObsRasters <- vector("list", length(reclassifiedObsRasters))
count <- 1
while(count <= length(mList)){
  setwd("~/Desktop/NCAR_Paper/CCSM4/CodMs/ASCIIs/");
  #Get mask layers
  maskRast <- raster(x = mList[[count]])
  #Get total extent
  uext <- extent(c(-180, 180, 0, 90))
  xext <- extent(maskRast)
  uext <- union(uext, xext)
  # Global Area empty rasterLayer
  rt <- raster(uext)
  res(rt) <- 1
  rt[] <- NA
  # Merge each rasterLayer to Global Extent area
  maskRast <- merge(maskRast, rt, tolerance = 1e+6)
  setwd("~/Desktop/NCAR_Paper/CCSM4/95_Thresholds/")
  maskedModRasters[[count]] <- mask(x = reclassifiedModRasters[[count]], maskRast)
  writeRaster(maskedModRasters[[count]], filename = paste(unlist(strsplit(fileList[count], split =
"[.]csv")), "_95Reclass_Mod", sep = ""), format = "ascii", overwrite = T)
  maskedObsRasters[[count]] <- mask(x = reclassifiedObsRasters[[count]], maskRast)
  writeRaster(maskedObsRasters[[count]], filename = paste(unlist(strsplit(fileList[count], split = "[.]csv")),
"_95Reclass_Obs", sep = ""), format = "ascii", overwrite = T)
  count <- count + 1
}
rm(maskRast, uext, xext, rt, count)

#####
# Make prediction error map #
#####

#Reclass masked rasters to show map of error
errorReclass <- vector("list", length(maskedModRasters))
count <- 1
while (count <= length(errorReclass)){
  errorReclass[[count]] <- maskedObsRasters[[count]] + maskedModRasters[[count]]
  reclass <- c(-Inf, .999, 0, .999, 1.5, 1, 1.6, Inf, 0)
  reclassMatrix <- matrix(reclass, ncol=3, byrow=TRUE)
  errorReclass[[count]] <- reclassify(errorReclass[[count]], reclassMatrix, NAflag = -9999)
  errorReclass[[count]][is.na(errorReclass[[count]])] <- 0
  count <- count + 1
}

sumOfErrorReclass <- errorReclass[[1]] + errorReclass[[2]] + errorReclass[[3]] + errorReclass[[4]] +
errorReclass[[5]] + errorReclass[[6]] + errorReclass[[7]] + errorReclass[[8]] + errorReclass[[9]] +
errorReclass[[10]] + errorReclass[[11]] + errorReclass[[12]] + errorReclass[[13]] + errorReclass[[14]] +
errorReclass[[15]]

#Reclass masked rasters to show map of Ms to make maps of correct and incorrect predictions
MReclass <- vector("list", length(maskedModRasters))

```

```

count <- 1
while (count <= length(MReclass)){
  MReclass[[count]] <- maskedObsRasters[[count]] + maskedModRasters[[count]]
  reclass <- c(-Inf, .999, 1, .999, 1.5, 1, 1.6, Inf, 1)
  reclassMatrix <- matrix(reclass, ncol=3, byrow=TRUE)
  MReclass[[count]] <- reclassify(MReclass[[count]], reclassMatrix, NAflag = -9999)
  MReclass[[count]][is.na(MReclass[[count]])] <- 0
  count <- count + 1
}

sumOfMReclass <- MReclass[[1]] + MReclass[[2]] + MReclass[[3]] + MReclass[[4]] + MReclass[[5]] +
MReclass[[6]] + MReclass[[7]] + MReclass[[8]] + MReclass[[9]] + MReclass[[10]] + MReclass[[11]] +
MReclass[[12]] + MReclass[[13]] + MReclass[[14]] + MReclass[[15]]

#Plot the % correct and % incorrect
plot(sumOfErrorReclass/sumOfMReclass, col = rev(brewer.pal(11, "RdYlBu")), main = "Percent of
Erroneous Predicitons")

#####
# Calculate Cohen's Kappa for species predictions #
#####

#Get cross tabulations for each species
crossTabList <- vector("list", length(maskedModRasters))
count <- 1
while (count <= length(crossTabList)){
  crossTabList[[count]] <- crosstab(maskedObsRasters[[count]], maskedModRasters[[count]])
  crossTabList[[count]] <- crossTabList[[count]][complete.cases(crossTabList[[count]]),]
  colnames(crossTabList[[count]]) <- c("Obs", "Mod", "Frequency")
  count <- count + 1
}

#Get measurements from cross tabulations
sumCells <- vector("list", length(crossTabList))
obsAgreement <- vector("list", length(crossTabList))
predPosObs <- vector("list", length(crossTabList))
predPosMod <- vector("list", length(crossTabList))
probAgreement <- vector("list", length(crossTabList))
sensitivity <- vector("list", length(crossTabList))
specificity <- vector("list", length(crossTabList))
kappaList <- vector("list", length(crossTabList))
count <- 1
while ( count <= length(crossTabList)){
  sumCells[[count]] <- sum(crossTabList[[count]][,3])
  obsAgreement[[count]] <- (crossTabList[[count]][1,3] + crossTabList[[count]][4,3])/sumCells[[count]]
  predPosObs[[count]] <- (crossTabList[[count]][2,3] + crossTabList[[count]][4,3])/sumCells[[count]]
  predPosMod[[count]] <- (crossTabList[[count]][3,3] + crossTabList[[count]][4,3])/sumCells[[count]]
  probAgreement[[count]] <- (predPosObs[[count]]*predPosMod[[count]])+((1-predPosObs[[count]])*(1-
predPosMod[[count]]))
  sensitivity[[count]] <-
crossTabList[[count]][4,3]/(crossTabList[[count]][3,3]+crossTabList[[count]][4,3])
}

```

```

specificity[[count]] <- crossTabList[[count]][1,3] / (crossTabList[[count]][1,3] +
crossTabList[[count]][2,3])
kappaList[[count]] <- (obsAgreement[[count]]-probAgreement[[count]])/(1-probAgreement[[count]])
count <- count + 1
}

statTable <- cbind(unlist(strsplit(mList, split = "[.]asc")), sensitivity, specificity, kappaList)
write.csv(statTable, "~/Desktop/NCAR_Paper/CCSM4/StatTablePresent.csv")
writeRaster(sumOfCorrectReclass/sumOfMReclass,
"~/Desktop/NCAR_Paper/CCSM4/sumOfCorrectPredictionsPresent.asc", format = "ascii", overwrite =
T)
writeRaster(sumOfErrorReclass/sumOfMReclass,
"~/Desktop/NCAR_Paper/CCSM4/sumOfErroneousPredictionsPresent.asc", format = "ascii", overwrite
= T)

#####
# Calculate change statistics for each species #
#####

#Get present
setwd(dir = "~/Desktop/NCAR_Paper/CCSM4/95_Thresholds/")
fileNames <- list.files(pattern="*.asc")
presObs <- vector("list", length(fileNames)/2);
presMod <- vector("list", length(fileNames)/2);
count <- 1;
while (count <= length(fileNames)){
  if (odd(count)){
    countAdj <- count %/% 2 + 1
    plot(raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84"), main =
paste(fileNames[[count]], count, "Mod"))
    presMod[[countAdj]] <- raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84");
  } else {
    countAdj <- count / 2
    plot(raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84"), main =
paste(fileNames[[count]], count, "Obs"))
    presObs[[countAdj]] <- raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84");
    count <- count + 1;
  }
}

#Get 2055
setwd(dir = "~/Desktop/NCAR_Paper/CCSM4/rcp8_5Projection/2055/95_Thresholds/")
fileNames <- list.files(pattern="*.asc")
Obs2055 <- vector("list", length(fileNames)/2);
Mod2055 <- vector("list", length(fileNames)/2);
count <- 1;
while (count <= length(fileNames)){
  if (odd(count)){
    countAdj <- count %/% 2 + 1
    print(paste(fileNames[[count]], count, "Mod"))
    Mod2055[[countAdj]] <- raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84");
  } else {

```

```

countAdj <- count / 2
Obs2055[[countAdj]] <- raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84");
count <- count + 1;
}

```

```

#Get 2100
setwd(dir = "~/Desktop/NCAR_Paper/CCSM4/rcp8_5Projection/2100/95_Thresholds/")
fileNames <- list.files(pattern="*.asc")
Obs2100 <- vector("list", length(fileNames)/2);
Mod2100 <- vector("list", length(fileNames)/2);
count <- 1;
while (count <= length(fileNames)){
  if (odd(count)){
    countAdj <- count %% 2 + 1
    print(paste(fileNames[[count]], count, "Mod"))
    Mod2100[[countAdj]] <- raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84");
  } else {
    countAdj <- count / 2
    Obs2100[[countAdj]] <- raster(fileNames[[count]], crs = "+proj=longlat +datum=WGS84");
    count <- count + 1;
  }
}

```

```

#Calculate area of each extent
setwd("~/Desktop/NCAR_Paper/CCSM4/Localities/");
fileNames <- list.files(pattern="*.csv");

```

```

modPresArea <- vector("list", length(fileNames));
obsPresArea <- vector("list", length(fileNames));
mod2055Area <- vector("list", length(fileNames));
obs2055Area <- vector("list", length(fileNames));
mod2100Area <- vector("list", length(fileNames));
obs2100Area <- vector("list", length(fileNames));
count <- 1
while (count <= length(fileNames)){
  obsPresArea[[count]] <- cellStats(area(presObs[[count]]) * presObs[[count]], sum);
  modPresArea[[count]] <- cellStats(area(presMod[[count]]) * presMod[[count]], sum);
  mod2055Area[[count]] <- cellStats(area(Mod2055[[count]]) * Mod2055[[count]], sum);
  obs2055Area[[count]] <- cellStats(area(Obs2055[[count]]) * Obs2055[[count]], sum);
  mod2100Area[[count]] <- cellStats(area(Mod2100[[count]]) * Mod2100[[count]], sum);
  obs2100Area[[count]] <- cellStats(area(Obs2100[[count]]) * Obs2100[[count]], sum);
  count <- count + 1;
}

```

```

#Calculate mean latitudinal change
modPresMeanLat <- vector("list", length(fileNames));
obsPresMeanLat <- vector("list", length(fileNames));
mod2055MeanLat <- vector("list", length(fileNames));
obs2055MeanLat <- vector("list", length(fileNames));
mod2100MeanLat <- vector("list", length(fileNames));
obs2100MeanLat <- vector("list", length(fileNames));
count <- 1

```

```

while (count <= length(fileNames)){
  obsPresMeanLat[[count]] <- mean(unique(rasterToPoints(presObs[[count]],
fun=function(x){x==1}),2));
  modPresMeanLat[[count]] <- mean(unique(rasterToPoints(presMod[[count]],
fun=function(x){x==1}),2));
  obs2055MeanLat[[count]] <- mean(unique(rasterToPoints(Obs2055[[count]],
fun=function(x){x==5}),2));
  mod2055MeanLat[[count]] <- mean(unique(rasterToPoints(Mod2055[[count]],
fun=function(x){x==1}),2));
  obs2100MeanLat[[count]] <- mean(unique(rasterToPoints(Obs2100[[count]],
fun=function(x){x==5}),2));
  mod2100MeanLat[[count]] <- mean(unique(rasterToPoints(Mod2100[[count]],
fun=function(x){x==1}),2));
  count <- count + 1;
}

changeTable <- cbind(strsplit(fileNames,split = ".csv"), obsPresArea, modPresArea, obs2055Area,
mod2055Area, obs2100Area, mod2100Area, obsPresMeanLat, modPresMeanLat, obs2055MeanLat,
mod2055MeanLat, obs2100MeanLat, mod2100MeanLat)
rm(obsPresArea, modPresArea, obs2055Area, mod2055Area, obs2100Area, mod2100Area,
obsPresMeanLat, modPresMeanLat, obs2055MeanLat, mod2055MeanLat, obs2100MeanLat,
mod2100MeanLat)
changeTable <- matrix(unlist(changeTable), ncol = 13, byrow = F)
colnames(changeTable) <- c("Species", "Obs Area Present", "Mod Area Present", "Obs Area 2055",
"Mod Area 2055", "Obs Area 2100", "Mod Area 2100", "Obs Lat Present", "Mod Lat Present", "Obs Lat
2055", "Mod Lat 2055", "Obs Lat 2100", "Mod Lat 2100");

write.csv(changeTable, "~/Desktop/NCAR_Paper/CCSM4/HistogramTable.csv", row.names = F);
changeTable <- read.csv("~/Desktop/NCAR_Paper/CCSM4/HistogramTable.csv");

setwd(dir = "~/Desktop/NCAR_Paper/CCSM4/PredictedDistributions/")
reclassifiedObsPres <- vector("list", length(fileNames));
count <- 1
while (count <= length(fileNames)){
  reclass <- c(-Inf, 0, 0, 0.1, Inf, 5)
  reclassMatrix <- matrix(reclass, ncol=3, byrow=TRUE)
  reclassifiedObsPres[[count]] <- reclassify(presObs[[count]], reclassMatrix, NAflag = -9999)
  writeRaster((reclassifiedObsPres[[count]] + presMod[[count]]), filename =
paste(unlist(strsplit(fileNames[count], split = "[.]csv")), "_Present", sep = ""), format = "ascii", overwrite
= T);
  writeRaster((Obs2055[[count]] + Mod2055[[count]]), filename = paste(unlist(strsplit(fileNames[count],
split = "[.]csv")), "_2055", sep = ""), format = "ascii", overwrite = T);
  writeRaster((Obs2100[[count]] + Mod2100[[count]]), filename = paste(unlist(strsplit(fileNames[count],
split = "[.]csv")), "_2100", sep = ""), format = "ascii", overwrite = T);
  count <- count +1
}

#####
# Wilcoxon Rank Sum test #
#####

```

```

#Get observation data
setwd("~/Desktop/NCAR_Paper/CCSM4/PublicationFolder/VariableComparison/Observation/");
names <- list.files(pattern="*.asc");
observations <- vector("list", length(names));
count <- 1;
while (count <= length(names)){
  observations[[count]] <- raster(names[[count]]);
  count <- count + 1;
}

#Get model data
setwd("~/Desktop/NCAR_Paper/CCSM4/PublicationFolder/VariableComparison/Model/");
names <- list.files(pattern="*.asc");
model <- vector("list", length(names));
count <- 1;
while (count <= length(names)){
  model[[count]] <- raster(names[[count]]);
  count <- count + 1;
}

#Make a sampling grid
samplingGrid <- rasterToPoints(x = model[[1]]);

#Get data
dataMatrix <- matrix(data=NA, nrow = nrow(samplingGrid), ncol= 12);
dataMatrix[,1:2] <- samplingGrid[,1:2]

obsExtractList <- vector("list", length(names));
modelExtractList <- vector("list", length(names));
count <- 1
while (count <= length(names)){
  obsExtractList[[count]] <- extract(observations[[count]], samplingGrid[,1:2])
  modelExtractList[[count]] <- extract(model[[count]], samplingGrid[,1:2])
  count <- count + 1;
}

dataMatrix[,3:12] <- cbind(obsExtractList[[1]], modelExtractList[[1]], obsExtractList[[2]],
modelExtractList[[2]], obsExtractList[[3]], modelExtractList[[3]], obsExtractList[[4]],
modelExtractList[[4]], obsExtractList[[5]], modelExtractList[[5]])

colnames(dataMatrix) <- c(colnames(samplingGrid)[1:2], "etopoObs", "etopoMod", "isMaxObs",
"isMaxMod", "isMinObs", "isMinMod", "ssMeanObs", "ssMeanMod", "tsMeanObs", "tsMeanMod")

#Do test of variable importance
etopo <- wilcox.test(dataMatrix[,3], dataMatrix[,4], paired = T, na.action = "omit")#sanity
isMax <- wilcox.test(dataMatrix[,5], dataMatrix[,6], paired = T, na.action = "omit")
isMin <- wilcox.test(dataMatrix[,7], dataMatrix[,8], paired = T, na.action = "omit")
ssMean <- wilcox.test(dataMatrix[,9], dataMatrix[,10], paired = T, na.action = "omit")
tsMean <- wilcox.test(dataMatrix[,11], dataMatrix[,12], paired = T, na.action = "omit")

```

```
sumTable <- cbind(c("etopo", "isMax", "isMin", "ssMean", "tsMean"), c(etopo$p.value, isMax$p.value,  
isMin$p.value, ssMean$p.value, tsMean$p.value))
```