



DIGITAL ACCESS TO  
SCHOLARSHIP AT HARVARD  
DASH.HARVARD.EDU

HARVARD  
LIBRARY



# Species Distribution Modeling of Telescopefishes (Actinopterygii: Giganturidae)

## Citation

Richarte, Darlene Renee. 2022. Species Distribution Modeling of Telescopefishes (Actinopterygii: Giganturidae). Master's thesis, Harvard University Division of Continuing Education.

## Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37373333>

## Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

## Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#)

Species Distribution Modeling of Telescopefishes (Actinopterygii: Giganturidae)

Darlene Renee Richarte

A Thesis in the Field of Biology  
for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

November 2022



## Abstract

Speciation is the cornerstone of all biodiversity on Earth. It is the foundation of evolutionary biology that new species are derived from ancestral species. The world's oceans cover over 70% of the Earth's surface and account for more than ten times the area than that of terrestrial habitats. However, little is known of the conditions promoting speciation in this realm. The absence of visible barriers in this zone make the mechanisms shaping oceanic diversity a particularly fascinating topic of research. For my research project, I investigated the ecological divergence of two closely related midwater fishes, *Gigantura chuni* and *Gigantura indica*. I built two species distribution models (SDMs)-one for each of the sister-species pair of telescopefishes. SDMs are conceptually based on ecological niche theory and these principles are used to produce distributional models. Using occurrence records together with environmental data, I used the machine-learning program Maxent to calibrate models optimized for predictive performance. I tested the validity of the models via cross-validation, area under the curve (AUC) of the receiver operating characteristic (ROC), and omission test rate values. The models performed well and successfully identified the three top contributing environmental variables relative to habitat suitability for each species. I also examined how reliably each model predicted its sister-species' occurrence by projecting presence data on to the other species' model and vice-versa. The *G. chuni* model predicted the occurrence of *G. indica* 73.40% of the time, while the *G. indica* model predicted the presence of *G. chuni* at a rate

of 69.32%. These results illustrate the sister-species pair of *G. chuni* and *G. indica* are using similar environmental variables. Ecological niche similarity relative to environmental requirements is a line of evidence for niche conservatism. Without visible barriers to promote physical and reproductive isolation, we might hypothesize new conditions by which populations diverge from ancestral lines.

## Frontispiece



MCZ:Ich:60585

*Gigantura indica* Brauer, 1901

collected in Papua New Guinea

by Museum of Comparative Zoology, Harvard University

(Licensed under <http://creativecommons.org/licenses/by-nc-sa/3.0/>)

## Acknowledgments

I would like to thank Dr. Scott Edwards of Harvard University, Department of Organismic and Evolutionary Biology and Dr. Shannon DeVaney of Los Angeles Pierce College, Department of Life Sciences for their willingness to serve as co-chairs of my committee and guide me through the thesis process. I am sure that taking on a mentee in the midst of a pandemic was not an easy decision, but I am grateful that they both agreed. Thanks to Dr. Edwards for taking a chance that provided me with this challenging, yet rewarding research opportunity. I am hopeful that we can catch up when the life settles down a bit. Thanks to Dr. DeVaney who has been an inspiration to me both personally and professionally. I am lucky to have you as a colleague and friend for the past 12 years.

A sincere thank you to Dr. James Morris who has been an endless source of encouragement throughout the thesis process. Your guidance and support has been invaluable during a process that could be particularly stressful at times.

A special thank you to Dr. Raymond Wells and Professor Pat Farris of Los Angeles Pierce College who so generously offered their time. I am both honored and excited to have Professor Farris add her artistic works to my thesis.

I am especially grateful to the very helpful scientific community working in the field of marine species modeling. I am appreciative of Dr. Hannah Owens and Dr. Tom Jenkins, who were so generous with their time and guidance in experimental design and R code. I am grateful for Phillip Kaye and his patience as he helped me through the frustrating times working with R code.

## Table of Contents

Frontispiece.....	iii
Acknowledgments.....	vi
List of Tables .....	viii
List of Figures.....	ix
Chapter I. Introduction: Context of Study .....	1
The Deep Ocean.....	2
The Ecological Niche.....	3
Niche Conservatism.....	7
Species Distribution Modeling .....	9
Chapter II. Materials and Methods .....	15
Occurrence Records .....	15
Environmental Variables .....	16
Species Distribution Modeling .....	16
Chapter III. Results.....	18
Chapter IV. Discussion .....	20
Appendix 1. Tables.....	24
Appendix 2. Figures.....	26
References.....	35



## List of Tables

Table 1. Environmental Variables Considered in This Study.....	24
Table 2. Pearson Correlation Table .....	25

## List of Figures

Figure 1. BAM Diagram .....	26
Figure 2. Representative Species Distribution Model .....	27
Figure 3. Telescopefish, <i>Gigantura indica</i> .....	28
Figure 4. Occurrence data plotted in QGIS for each species .....	29
Figure 5 – Projected maps for habitat suitability .....	30
Figure 7. Response Curves- <i>Gigantura chuni</i> .....	32
Figure 8. Response Curves- <i>Gigantura indica</i> .....	33

## Chapter I.

### Introduction: Context of Study

Speciation is the origination of new species that occurs with the splitting of a single line into two or more lineages (Zimmer & Emlen, 2016). Organismic evolution occurs multi-dimensionally in space and on timescales potentially spanning over millions of years. Thus, a major consideration is discerning whether proximate, or current, observed conditions accurately reflect the ultimate factors contributing to the speciation event (Losos & Glor, 2003; Peterson, 2011a). The characteristics of evolutionary change make deducing the mechanisms by which species evolve especially challenging. Consequently, it is necessary to explore modes of speciation in an indirect manner.

To this end, the ecological niche has emerged as an operative tool for investigating the mechanisms of speciation with the caveat that the condition currently observed has been constant since the time of divergence (Losos & Glor, 2003). Although distributions are transient and change over time, the principles of ecological niche theory can inform us on the distributional ecology of a species. Moreover, the ecological niche of an organism relative to abiotic factors regulates the biogeographic occurrence of a population (Wiens & Graham, 2005). The occurrence records, or where a species occurs in the landscape, together with environmental variables can be used to model the ecological niche of a species. Once a climatic niche envelope is constructed for a species, it can be projected back to geographic space and provide an index of habitat suitability for a species (Elith et al., 2006; Pearson, Richard, 2008; Peterson, 2011a; Peterson & Soberón, 2012a). The conservative nature of the ecological niche further allows us to

determine the distributional ecology of a species as well as explore the mode by which populations diverged from their ancestral lines.

## The Deep Ocean

The deep ocean is the largest ecosystem on Earth, composing more than ten times the habitat of terrestrial systems. The deep sea lies just below the epipelagic (the first 200 meters from the sea surface or limit of the depth of light penetration) and is subdivided into the mesopelagic and bathypelagic regions. The mesopelagic extends from 200 to approximately 1000 meters below the surface to a 10 °C isotherm, and is a dimly lit zone with insufficient light to support photosynthesis. Below the mesopelagic zone is the bathypelagic zone that ranges from 1000 to 6000 meters or the 4 °C isotherm of the water column. This zone receives no sunlight and thus photosynthesis does not occur at this depth (Bruun, 1956; Herring, 2002). Despite the fundamental differences in productivity, temperature and depth, the layers of the ocean remain biologically and physically connected. Diel vertical migrations (DVM) of marine plankton and phytoplankton facilitate biological interactions. Likewise, currents, countercurrents, upwelling and downwelling events also maintain physical contact between the vertical dimensions of the deep sea.

The ways in which species originate in the open ocean are not well-documented in the published literature. Although allopatry, or spatial isolation, has been reported to be the most common form of lineage splitting (Barracough & Vogler, 2000), the absence of geographic barriers in the open ocean to facilitate this mode of speciation warrants further inquiry. Miya and Nishida (1997) examined several museum specimens of *Cyclothone*, a genus of bristlemouth fishes found globally in the ocean's midwaters.

Using neighbor-joining and maximum parsimony phylogenetic analysis, they discovered genetic differences among sister populations that corresponded to the locality where they were collected. These results demonstrated a correlation between genetic lineage and geographic location. More importantly, they provide a line of evidence that lineage sorting in the absence of geomorphological barriers could be the early stages of speciation in the deep sea ecosystem.

### The Ecological Niche

Measuring the distribution and density of organisms across the landscape, along with the causal factors, are key objectives to the study of ecology. The term “niche” is often used in literature to describe the association organisms have with their environment. However, the definition of an ecological niche of a species has transformed since its first introduction by Joseph Grinnell over a century ago. Grinnell, a pioneer in the field, performed some of the earliest research in distributional ecology. His study of the California Thrasher (*Toxostoma redivivum*) principally focused on the effect abiotic factors have on geographic range. Although Grinnell noted geographic features such as deserts and mountains that constrained the distribution of the thrasher, he asserted that physical elements like temperature and humidity were the primary drivers controlling the “associational niche” (Grinnell, 1917). Charles Elton also examined the distribution of species in their respective environments. He studied an array of animals from marine phytoplankton, to insects and rabbits. He illustrated how biotic associations, specifically trophic food webs governed species’ distribution in a community setting (Elton, 1927). Both early researchers used the niche concept to describe species-environment

relationships, albeit emphasizing different elements responsible for observed distributions.

The innovative work of G. Evelyn Hutchinson provided the contextual framework and link needed connect the physical and biotic worlds. Hutchinson proposed that organisms exist in the biotic realm as well as the physical environment simultaneously. Thus, the geographic location of where a species occurs is represented as a function of the abiotic conditions at that location. Hutchinson called these physical features, “scenopoetic variables”, as he reasoned them to be non-interactive and slow changing in nature. These physical dimensions comprise an abstract construct of environmental space collectively known as the  $n$ -hypervolume. (Hutchinson, 1957). In the marine environment, such scenopoetic variables include water temperature, salinity and pH. Colwell and Rangel (2009) have described the ability to use the geographic location of a species and represent it as a function of the physical conditions of that locale simultaneously as “Hutchinson’s Duality”. This dual relationship between geographic and environmental space allows the use of geo-referenced occurrence data with associated coordinates to be represented abstractly in environmental space or  $E$  space (Guisan & Zimmermann, 2000; Pearson, Richard, 2008; Peterson & Soberón, 2012a). Fundamental to understanding the ecological niche is being mindful of the elements that characterize each space and their inherent limitations. Species occur and exist in geographic space and we can use this information to create and analyze the ecological niche model in environmental space (Peterson & Soberón, 2012a).

The patterns of occurrence and the respective distribution of species across the landscape occur in geographic space,  $G$ . Soberón and Peterson (2005) explain this

distribution as a function of the biological and physical limitations of a species relative to areas accessible to the species. The Biotic-Abiotic-Movement or BAM diagram (Figure 1) is a heuristic representation of the ecological constraints that drive the observed distribution of species. The B in the diagram references the biotic relationships encountered by species such as prey-predator and competitive interactions. The A circle indicates abiotic factors such as temperature, humidity and precipitation that are influential in determining where species occur in the environment. M refers to movement and those areas accessible for dispersal by a species. Regions of accessibility are not often addressed in studies modeling species distributions; however, it is important to incorporate species mobility to build more informative and robust models. Moreover, incorporating species' mobility is especially relevant when projecting models to future timeframes or studying their invasive potential (Miller & Holloway, 2015; Peterson & Soberón, 2012b; Soberón & Nakamura, 2009; Warren et al., 2021). The intersection of A and B only can be interpreted as the area that are both biotically and abiotically suitable for a species, but inaccessible for dispersion by the species due to the location being too distant or impeded by geographic barriers. This region where only the biotic and abiotic requirements of a species are met is known as the invadable space of a species ( $G_i$ ). Although species can potentially exist where biotic and abiotic conditions are suitable, they are unable to persist in the area on a long-term basis (Soberón & Peterson, 2005).

The 3-way intersection of BAM describes the area where all three elements have been satisfied and the habitat is suitable for species to maintain stable populations. This segment is called the occupied distributional niche and denoted by  $G_o$  (Peterson & Soberón, 2012a). The BAM schematic provides the necessary framework for

understanding the elements limiting the distribution of species in geographic space. It is important to consider all elements of the BAM configuration when calibrating species distribution models. This entails integrating biological information available about a species and selecting a study area that represents areas accessible to the species (Peterson & Soberón, 2012b).

Transitioning to the other side of Hutchinson’s duality is environmental space. In this abstract realm, we construct the ecological niche of a species. According to Hutchinson, the fundamental niche consists of n-dimensions in environmental space. These dimensions are the abiotically suitable environmental variables where a species can survive and reproduce (Hutchinson, 1957). Thus, the fundamental niche of a species is the subset of environmental variables currently regulating the positive growth rate of the population in environmental space (Jackson & Overpeck, 2000).

The fundamental niche model considers the A circle of the BAM diagram wherein various physical environmental layers reflect the physiological tolerance of a species, but does not account for biological interactions (Kearney & Porter, 2009). The following set theory equation represents the fundamental niche and subsets thereof:

$$N_F \supseteq N^*_F = N_F \cap n(M) \supseteq N_R = N^*_F \cap n(B)$$

$N_F$  denotes the fundamental ecological niche of a species.  $N^*_F$  is the existing fundamental niche that is reduced to include only those environments in  $N_F$  that species’ are able to access. The realized niche ( $N_R$ ) is equal or less than the existing fundamental niche and includes the biotically suitable areas for the species ((Peterson et al., 2022; Peterson & Soberón, 2012b; Soberón, 2007). Although many researchers have analyzed the ecological niche of a species under the assumption that the realized niche is smaller



than the fundamental niche (Crotty & Bertness, 2015), the work of Soberón and Arroyo-Peña (2017) quantified the fundamental niches of over 100 species of reptiles and amphibians using a single niche variable and occurrence data. Their results supported the long-held hypothesis that the fundamental niche is indeed larger than the realized niche. It is important to bear in mind that SDMs use abiotic environmental variables and thereby are a kin to a subset of the Hutchinsonian fundamental niche. The fundamental niche and inclusive groups establish important relationships that should be integrated into experimental design when modeling distributions and performing analyses.

### Niche Conservatism

The characteristics of the fundamental niche of a lineage are conserved over evolutionary time in a process known as niche conservatism. The tendency of species to maintain their ecological characteristics are explained by the observation that species are more likely to be found within their fundamental niche. This results in physiological tolerances limiting the geographic range of a species (Wiens et al., 2010; Wiens & Graham, 2005). Although some researchers assert that the ecological niche of a species is more labile and readily subject to differentiation than reported (Losos & Glor, 2003), most published literature support the conserved nature of ecological niches (Wiens, 2004). The research of A.T. Peterson et al. (1999) examined the role of niche conservatism in sister species pairs of mammals, butterflies and birds of southern Mexico. They used environmental dimensions together with algorithmic software to model the fundamental niche of each species pair. By projecting the model onto geographic maps, they were able to predict the potential distribution and range of each species pair. More importantly, the models successfully predicted the geographic

distribution of its sister-species counterpart. While niche conservatism is widespread across taxa, this study suggests that the conservative nature is most evident at the species level and over moderate evolutionary timescales. The high level of niche conservatism in these timeframes is indicative of speciation occurring at geographical level rather than an ecological one (Peterson, 2011b). Further, the lack of differentiation in the fundamental niche can be used to make inferences about the mode of speciation among closely related species.

The research of K. McNyset (2008) employed ecological niche modeling to investigate niche conservatism in several sister species of freshwater fishes such as darters, minnow and basses of the southeastern United States. Using abiotic dimensions together with point-occurrence data extracted from collection databases and published literature, ecological niche models were built for several sister-species pairs of fishes. Many of the sister-species pair models demonstrated reciprocal predictivity, with the exception of some species of black basses. The ability of a model to correctly project the distribution of its sister counterpart occurring in different locales, indicates that they share the same ecological niche in relation to environmental factors and, therefore their ecological niches have been highly conserved over time. Climatic tolerance in niche conservatism has been reported to promote vicariance (Kozak & Wiens, 2006) and the niche conservatism observed in this study can be attributed to allopatry rather than ecological adaptation.

The geographic range of a species is limited with respect to climatic tolerances due to niche conservatism (Wiens & Graham, 2005) and these properties can be explored via ecological niche modeling. Thus, analyzing and modeling the distribution of species

can provide further insight into how they change over time.

### Species Distribution Modeling

Species distribution models (SDMs) are an important tool for estimating the distributional range of suitable habitat for a species of interest. SDMs take the geographic locations of where species are found and extract information about the abiotic conditions at those locations. The combination of environmental data and occurrence records is integrated using algorithmic software that produces a correlative model index of habitat suitability ((Elith & Franklin, 2013; Elith & Leathwick, 2009; Peterson, 2001). Figure 2 illustrates the process of building an ecological niche model from environmental variables and occurrence records and projecting the model on to new geographic regions. For marine SDMs, oceanic environmental variables include measurements such as sea surface temperature, salinity and dissolved oxygen (DO) content. Similarly, occurrence data for many marine species can be obtained from databases like the Global Biodiversity Information Facility (GBIF), FishBase or World Registers of Marine Species (WoRMS). SDMs can be useful in estimating a species' niche (Warren, D. and Seifert, S., 2011) and once a working model is achieved, can have a variety of applications. To wit, SDMs have been used to make predictions under future climate change conditions, to assess the potential spread of invasive species and to assist researchers in species conservation efforts (Elith et al., 2011; Guisan & Thuiller, 2005; Pearson, Richard, 2008). Further, integrating GIS-based information with occurrence and environmental variables has given investigators new tools for approaching related research questions (Kozak et al., 2008). Since most of the ocean has not been sampled, SDMs are an especially useful tool for modeling organisms of the deep sea (Bentlage et al., 2013).

S. DeVaney (2016) used SDMs to examine the distribution of deep sea eels, *Cyema atrum* and *Eurypharynx pelecanoides*. Using occurrence and environmental data, she effectively built models that characterized suitable habitat based on logistic output - a transformation of the relative occurrence rate that predicts the probability of a species' presence (Phillips and Dudik, 2008). By projecting the models in new regions of the ocean, she identified areas where future sampling efforts could be concentrated to identify new populations of eels.

Other applications for modeling species distributions is to provide insight about the invasive potential of an introduced species. Species modeling has been used to estimate the environmental niche of invasive carp based on its native distributional range. These constructs were successful in identifying suitable conditions in which the invader is most likely to establish and spread (DeVaney et al., 2009). Equipped with this information, management agencies can recognize the potential threat of an introduced species and take steps to mitigate their impact on the new environment.

SDMs have been instrumental tools for modeling rarely-seen fish species where occurrence data is limited and possibly unreliable. Phillips et al. (2017) sought to identify the distribution of cryptic ocean sunfishes, *Mola mola* and *Mola ramsayi* for biodiversity conservation purposes. The International Union for Conservation (IUCN) Red List designates these fishes as “vulnerable” and “not assessed”, respectively. The marine SDM identified these species as highly mobile fishes with a global range that fluctuated in response to seasonality. Clustered occurrences were centered in coastal regions and were highly dependent on sea surface temperatures being within a range of 16 to 23 °C. These data offer baseline information on the distribution of ocean sunfishes and suggest

how their distribution might shift under projected increases in sea surface temperatures worldwide.

The Maxent software program is well-suited for modeling the ecological niche of a variety of species using presence-only data (Elith, et al 2006). Maxent uses the principles of maximum entropy to find the most uniform distribution across the study region given specific constraints (Phillips et al., 2006, 2017). The machine-learning algorithm uses an iterative process to fit a background or training region against known presences. The background consists of unknown presences and does not represent a definitive absence if no data is available at that point. This method of modeling is especially useful for oceanic studies, as much of the ocean remains unsampled.

The purpose of my research project is to effectively model the distribution of the sister species pair, *Gigantura indica* and *Gigantura chuni* (Order Aulopiformes) and test whether the models accurately represent their ecological niches. Telescopefishes are morphologically unique and characterized by multiple skeletal reductions and toothed elements on the upper jaw (Konstantinidis & Johnson, 2016). Another distinctive feature of giganturids are their anteriorly-directed tubular eyes specialized for life in the twilight of the mesopelagic. The eyes of *Gigantura* evolved independently of other fishes inhabiting this zone. A unique behavioral characteristic is that telescopefish orient themselves vertically in the water column possibly to as a predatory feeding strategy (Marshall, 1971). This species group was selected for its distinctive physical appearance in adult form. As such, they are unlikely to be misidentified by collectors or in museum records. This measure was taken not only to remove identification ambiguity within our samples, but also to provide a level a confidence in our occurrence data.

The published literature on the ecology *G. chuni* and *G. indica* is scarce and little is known about their feeding strategies or food sources. Richards et al. (2019) recently examined the trophic ecology of eight meso- and bathypelagic teleosts, including *G. chuni* and *G. indica*. Using isotope analysis of particulate organic matter (POM) and nitrogen, they found that most of the organic carbon consumed by the deep sea fishes in the study comes from the surface layers of the ocean. Another new finding by this study discovered that *G. chuni* and *G. indica* were top predators in the trophic level of the deep ocean food web studied. Further, an analysis of the stomach contents of telescopefishes included zooplankton-consuming fishes, decapod shrimp and squid. Integrating biological information like dietary requirements into correlative SDMs provides a better understanding of niches dynamics as well as further insight into the observed distribution of species (Porteiro & Sutton, 2007; Pulliam, 2000; Richards et al., 2019). Figure 3 is a representative depiction of *G. indica*. Both species occur circumglobally (IUCN 2013a; IUCN, 2013b) in the deep pelagic zone and do not undergo diel vertical migrations (Richards, et al., 2019).

To ascertain the accuracy of the models built, I will with employ several validation techniques, including cross-validation, assess the area under the curve (AUC) of the receiver operator characteristic (ROC) and analyze Maxent outputs such as average test omission rates. I expect the models will predict a significant number of known presence points that fall within the distributional range established by the model. If the model accurately predicts a significant segment of the species' occurrence data or known presences, I can be confident that the model is predictively accurate and a good representation of the species' ecological niche. Further, I will examine whether the

models of sister-species pair, *G. chuni* and *G. indica* exhibit evidence of niche conservatism. I will accomplish this by projecting the occurrence data for *G. chuni* onto the *G. indica* model and vice versa. The ability of the models to predict the occurrence of their counterpart sister species would demonstrate a degree of niche conservatism.

Building species distribution models from species' occurrence records has several limitations (Barve et al., 2011) especially as it relates to sampling biases and autocorrelation (Legendre, 1993). One assumption often made by modeling algorithms using presence-only data is that the study area has been randomly sampled (Merow et al., 2013). However, the ocean covers a massive expanse and only a small fraction of the environment has been sampled thus far. Biases exist in occurrence records due to uneven sampling or a high degree of sampling in areas readily accessible to researchers and other workers surveying high interest areas or species (Reese et al., 2005). The impact biases have on Maxent models differs depending on intensity, the specific type of bias and on the species under study. Overall, biased data hinders our ability to determine the true relationship between occurrence and environmental data (Fourcade et al., 2014; Warren et al., 2021). In an effort to reduce sampling biases, I will use Target Groups Sampling (TGS) in my study. TGS uses the presence records of a proxy species that is likely to be sampled at the same time as our focal species. Under the assumption that our target species would have been recorded by the survey had it been present, we use this information to inform the background of our model where previous sampling efforts have occurred (Merow et al., 2013; Phillips et al., 2006). Barber et al (2020) constructed several distribution models of hoverflies (Syrphidae) using presence-only occurrence data. Researchers divided the models into subgroups, one group that implemented TGS

to reduce sampling bias and another group of models without TGS. The results of this study found that TGS was an effective tool for mitigating the effects of sampling bias in SDMs. I will use the bristlemouth fish, *Cyclothone* as a proxy for the occurrence of telescopefishes. *Cyclothone* is composed of 13 mesopelagic species (Herring, 2002) and has been reported to be the most abundant vertebrate species in the world. *Cyclothone* occurs in all seas including the Arctic and has a worldwide distribution (Nelson, 2006). Given the ubiquitous nature of *Cyclothone* together with the information that the fishes are caught in the same sampling gear as *G. Chuni* and *G. indica*, it follows that *Cyclothone* it is a good candidate for approximating sampling sites for TGS. I received occurrence records for *Cyclothone* from a 2015 GBIF download with 20,869 occurrence points (GBIF.org 2015c).

Lastly, my research involves analysis of information and records maintained in several public databases. The algorithmic software, Maxent, is also available on an open-source basis. No portion of the protocol involves the use of human subjects.



## Chapter II.

### Materials and Methods

#### Occurrence Records

To model the distribution of *Gigantura indica* and *Gigantura chuni*, I downloaded occurrence data from the Global Diversity Information Facility (GBIF) for each species. I obtained 454 representative presence records for *Gigantura indica* (GBIF.org 2021a) and 191 records for *Gigantura chuni* (GBIF.org 2021b). This raw data received from network databases was cleaned to reduce inherent bias and attain high quality data for input into our model (Hijmans et al., 1999). Thus, I visually inspected the data and removed records without coordinate information, duplicated coordinates and those points wherein the data were not recorded to at least two decimal places in either the latitude or longitude. Subsequently, I examined my occurrence data for both species in QGIS 3.22.5 (QGIS Development Team 2021). Figure 4 represents the geographic occurrence for both species of telescopesfishes. Using QGIS for a visual representation for both species, I removed data points occurring on land. Further, I identified tightly clustered points, defined as those points occurring within two or fewer grid cells of one another, and randomly omitted one or more of the records. This resulted in a single data point at a location. After data cleaning, there remained 185 coordinate-reference occurrence records for *Gigantura indica* and 88 for *Gigantura chuni*.

## Environmental Variables

I uploaded a set of 18 marine environmental variables from the Bio-Oracle database (Assis et al., 2018; Sbrocco & Barber, 2013a) via the ENMTools (v.1.0.6; Warren et al., 2021) package available in R. The layers included biochemical and physical elements influential in the distribution of marine organisms. The environmental data sets have a spatial resolution of 5 arcminutes on a global scale. Table 1 lists the environmental variables considered in this study. A Pearson correlation test was performed to identify correlated layers that would interfere with the predictive performance of the model (Cobos et al., 2019). Table 2 lists the environmental variables and their corresponding correlation to one another. For layer pairs with correlation values  $>0.65$ , I ran five trail analyses and selected the variable with the greatest contribution across trials. A set of five environmental variables was used for each species in the final Maxent models. Table 1 identifies the environmental layers in the final models.

## Species Distribution Modeling

For my final Maxent models, I input the occurrence data for each species along with the corresponding set of variables into the R package ENMevaluate (v.2.0.3, Kass, et al., 2021). The package uses a derivative of the Akaike Information Criterion, a standard test in machine learning algorithms for model comparison. The software analyzes the predictive performance of a model by determining how well the training data fits the model. It is a comparative method that identifies the best model and settings to be used with the datasets provided. I used the linear, quadratic, product, threshold and hinge features. The regularization multiplier that determines the degree of complexity for response curves was set at 1. In addition, I used cross-validation methods to test the

models. Cross-validation takes subsets of my occurrence data and uses that information to train the model. Another subset of the data that was not used in model training is subsequently utilized to test the model without replacement. I selected the jackknife option in Maxent to determine the percent each variable contributed to the resulting model. The jackknife method is an iterative process that adds or subtracts regularized gain in the training algorithm. While jackknifing and bootstrapping are both methods of resampling, jackknifing is more precise than bootstrapping at estimating variance in a data set (Shao & Tu, 1996). I ran four replicate runs for each species and opted for cumulative output (cloglog). Cloglog the newest default output in Maxent and is a transform of the logistic output format. According to the latest version of Phillips et al. (2006), the cloglog output is easy to conceptualize because it provides an estimation of the probability of occurrence relative to suitable habitat with values ranging between 0 and 1 (Phillips et al., 2017; Phillips, S.J., 2017). As will all outputs, interpretation is subject to quality of input data and experimental design.

## Chapter III.

### Results

The results of my study illustrate that *Gigantura chuni* and *Gigantura indica* have global distributions concentrated throughout the tropics with distinct latitudinal limits beyond the Tropic of Cancer and subtropical regions. Figure 5 shows the ecological niche for each species projected back onto geographic space. The cumulative output ranges from 0 to 1. The range indicates how suitable the environmental variables are for each species, zero predicting low suitability and one predicting high suitability. The blue color with a range of (0.00-0.10) show low values of cumulative output, followed by green (0.10-0.25), yellow (0.25-0.40), orange (0.40-0.70), and red (0.70-1.0). The black dots on each map represent the occurrence points for each respective species.

The area under the curve (AUC) of the receiver operating characteristic (ROC) for *G. chuni* model is .885 and the AUC for the *G. indica* model .80. Although AUC is a standard threshold-independent metric used in species distribution modeling to evaluate the predictive performance of resulting models, it should be considered with caution. It has been reported (Anderson & Raza, 2010; Lobo et al., 2008) that AUC values may be inflated for study areas covering large extents.

I set my model parameters to have Maxent use the jackknife statistical process to determine the top contributing environmental variables for habitat suitability. Figure 6 lists the variable contribution for each species of telescopefishes. For *G. chuni*, the top three contributing variables are sea surface chlorophyll (23.6%), bottom sea temperature (22.7%) and bottom nitrate concentration (19.8%). Likewise, for *G. indica*, the top three

contributing variables are sea surface chlorophyll (43.5%), sea surface silicate (21.5%) and water column mean pH concentration (14.8%).

In addition, Figure 7 provides the response curves for the top three contributing variables for *G. chuni*. Figure 7a. shows that as sea surface chlorophyll increases, the cumulative output or habitat suitability only increases within the narrow range of 0.0 - 0.4 mg/m<sup>3</sup> chlorophyll concentration. Figure 8 shows the response curves for *G. indica*. Suitable habitat for *G. indica* is evident in a similar range of 0.0 - 0.55 mg/m<sup>3</sup> chlorophyll concentration (Figure 8a).

After four replicate runs for each species, the average test omission rate for *G. chuni* was recorded at 13.80% and 15.73% for *G. indica* (p-value < <0.05). These values were computed by subtracting the 10<sup>th</sup> percentile training presence from the minimum training presence.

Figure 9 demonstrates the reciprocal predictivity of the models for each species. Figure 9A is a projected map of the *G. chuni* model in geographic space with an overlay of the *G. indica* occurrence data. The model has an omission percentage of 26.60%. Similarly, Figure 9B is a projected map of the *G. indica* model with the *G. chuni* presence data superimposed on the landscape. The presence of *G. chuni* is predicted by the *G. indica* model 69.32% of the time with an omission percentage of 30.68%.

## Chapter IV.

### Discussion

The objective of my thesis was to examine the ecological correlates of speciation of pelagic fishes in the open ocean, in the absence of physical barriers. My specific aims included accurately modeling the distribution of the sister-species pair of telescopefishes, *Gigantura chuni* and *Gigantura indica*. I hypothesized that I could build the ecological niche model for both fishes using occurrence records together with environmental data. Ecological niche models are formulated under the assumption that the currently observed distribution is indicative of a species' ecological requirements. Projecting the niche models back to geographic space yields a distributional range of suitable habitat for *G. chuni* and *G. indica*.

A second aim of my thesis was to test the predictive performance of each niche model produced. Model evaluation metrics include AUC, environmental variable response curves, cross-validation, and omission test rates values. The projected models provide an index of habitat suitability similar to conditions where the specimens have been collected. Environmental tolerances limit the geographic range of a species and can provide insight about niche conservatism and speciation processes.

The species distribution models constructed for *G. chuni* and *G. indica* show good predictive performance. While AUC alone is not reliable metric for large study areas, it does provide support for good model performance overall. Further, with an average test omission rate of 13.80%, the model for *G. chuni* makes successful predictions 86.2% of

the time. The model for *G. indica* had an average test omission rate of 15.73% and thus makes successful predictions 84.27% of the time. Calibrating the models using the cross-validation method wherein a different set of occurrence data is used to train and test the models provides further confidence that the models are robust and able to accurately make predictions when presented with new data.

The projected map for *G. chuni* (Figure 5.A) reveals that the species has a circumglobal distribution with high cumulative outputs in the South Atlantic and Indian oceans. The model predicts that the most suitable habitat relative to the environmental variables is in the Southern Atlantic Ocean. For *G. chuni*, the least suitable habitat is beyond 32° N and 42° S south of the equator. The top three environmental variables influencing the distribution of *G. chuni* are sea surface chlorophyll at the low end of the data range, bottom water temperature and the concentration of benthic nitrate in the lower range of the response curve (Figures 6 and 7). Figure 5.B shows that *G. indica* also has global distribution with high cumulative outputs in all world oceans, especially the South Atlantic, Indian and Pacific oceans. The top contributing environmental variables for *G. indica* are sea surface chlorophyll and sea surface silicate both at the low end of the data range as well as pH levels of the ocean water column (Figures 6 and 8). The least suitable habitat for *G. indica* lies beyond 35° N and 36° S of the equator. Based on these SDMs, the suitable habitat for this species-pair of telescopefishes lies within well-defined latitudinal limits.

An interesting result of this study reveals that of the 18 variables considered for modeling, the top contributing variable for both taxa is sea surface chlorophyll at similar concentrations (See Figure 6). This result supports the findings of Richards et al (2019).

This study revealed that several meso- and bathypelagic species were dependent on nutrient concentrations derived from the surface of the ocean. Dependency on sea surface chlorophyll could be a consequence of those nutrients being a food source of *Gigantura*'s prey items. This observation demonstrates that although *G. chuni* and *G. indica* are organisms that reside deep within the ocean, vertical connectivity between ocean layers, especially as it relates to transport of nutrients from the epipelagic zone to layers below, strongly influences the geographic distribution of telescopefishes.

According to the techniques used to validate the models, the distribution models for *G. chuni* and *G. indica* are robust and have good predictive value. To investigate the mode of speciation between this sister-species taxon, I employed reciprocal predictivity methods to determine how well the models of one species of telescopefish predicts the occurrence of its sister complement. To accomplish this, I projected the presence data of *G. chuni* on to *G. indica* model and vice-versa. The *G. chuni* model predicted the occurrence of *G. indica* 73.40% of the time, while the *G. indica* model predicted the presence of *G. chuni* at a rate of 69.32%. These results indicate that the species pair of *G. chuni* and *G. indica* are using similar environmental variables. Ecological niche similarity relative to environmental requirements is a line of evidence for niche conservatism. While *G. chuni* and *G. indica* demonstrate a significant degree of ecological reciprocity between their ecological niches, the occurrence overlay of one sister-species pair does not predict all presence points of its sister pair. Assessing how well ecological niches have been conserved over time operates under the assumption that the currently observed distribution of species has remained constant since the time of speciation. However, since we do not have specific information regarding the time of lineage divergence, we can



hypothesize that the differences in the ecological niches of *G. chuni* and *G. indica* could be a result of ecological adaptation after the speciation event.

The principles underpinning ecological niche conservatism are most pronounced at the species level and over shorter timescales (Peterson, 2011b). Further, in the open ocean, without visible barriers to promote physical and reproductive isolation, we might surmise possible new conditions by which populations diverge from ancestral lines. One such mechanism involves lineage sorting and potential population isolation by invisible barriers. Invisible barriers like a thermocline (Castro & Huber, 2019), halocline (Polyakov et al., 2018) or other nutrient gradients might drive the segregation of oceanic midwater species. If this is the case, the mode of speciation for many deep sea organisms is likely vicariance. Future studies should explore the effects of unseen barriers in the open ocean and their effect on speciation of midwater organisms.

## Appendix 1.

### Tables

Table 1. Environmental Variables Considered in This Study

<b>Environmental Variable</b>	<b>Units</b>	<b>Environmental Variable Used in Final Model</b>
Surface chlorophyll	mg/m <sup>3</sup>	GI, GC
Surface current velocity	m/s	GI, GC
Surface dissolved oxygen	mol/m <sup>3</sup>	
Surface nitrate	mol/m <sup>3</sup>	
Surface phosphate	mol/m <sup>3</sup>	
Surface primary productivity	g*m <sup>-3</sup> day <sup>-1</sup>	
Surface salinity	PSS	
Surface silicate	mol/m <sup>3</sup>	GI, GC
Surface sea temperature	°C	
Surface pH	-	GI
Calcite (mean)	mol/m <sup>3</sup>	
Benthic dissolved oxygen	mol/m <sup>3</sup>	
Benthic nitrate	mol/m <sup>3</sup>	GI, GC
Benthic phosphate	mol/m <sup>3</sup>	
Benthic salinity	PSS	
Benthic silicate	mol/m <sup>3</sup>	
Benthic temperature	°C	GC
Bathymetry	-	

*Note.* *Gigantura Chuni* denoted as GC and *Gigantura indica* by GI.

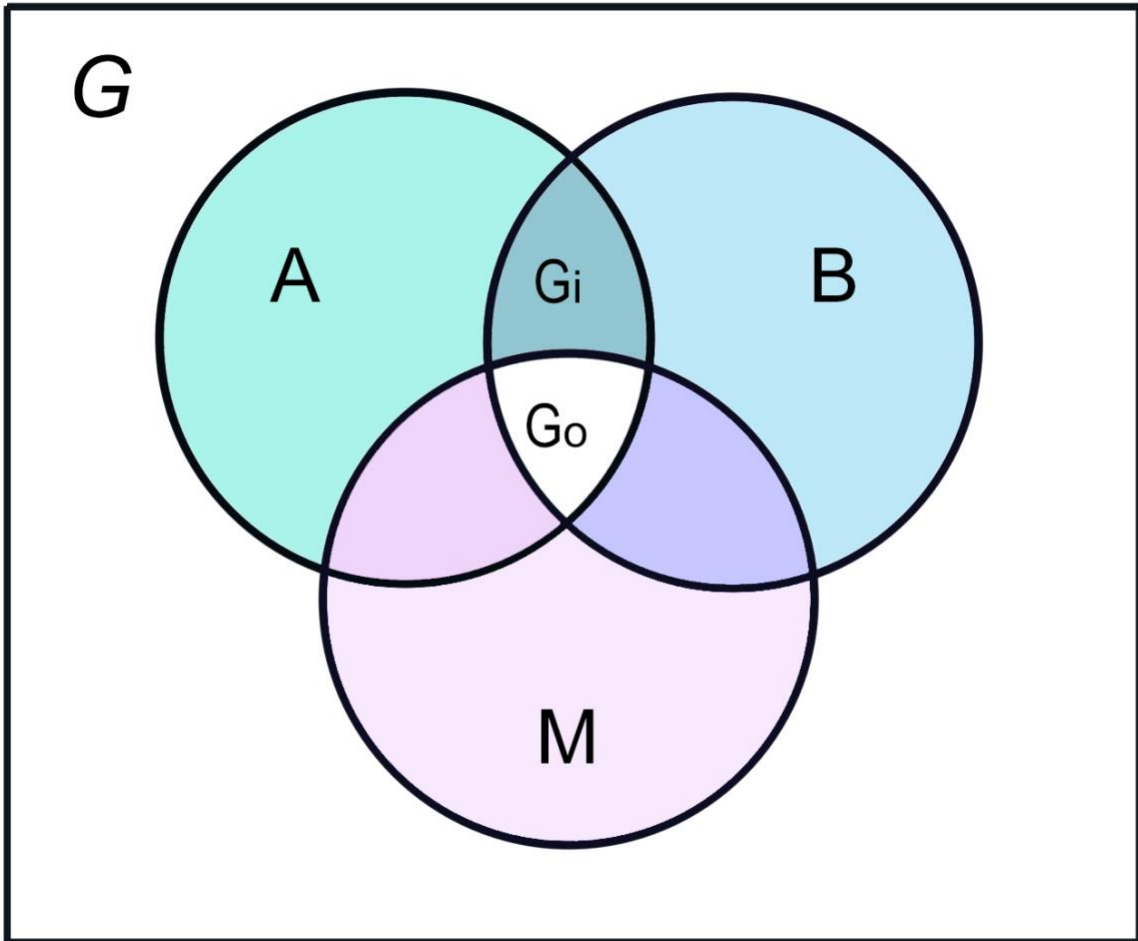
Source of environmental layers: Bio-ORACLE (Assis et al., 2018; Sbrocco & Barber, 2013b; Tyberghein et al., 2012).



Appendix 2.

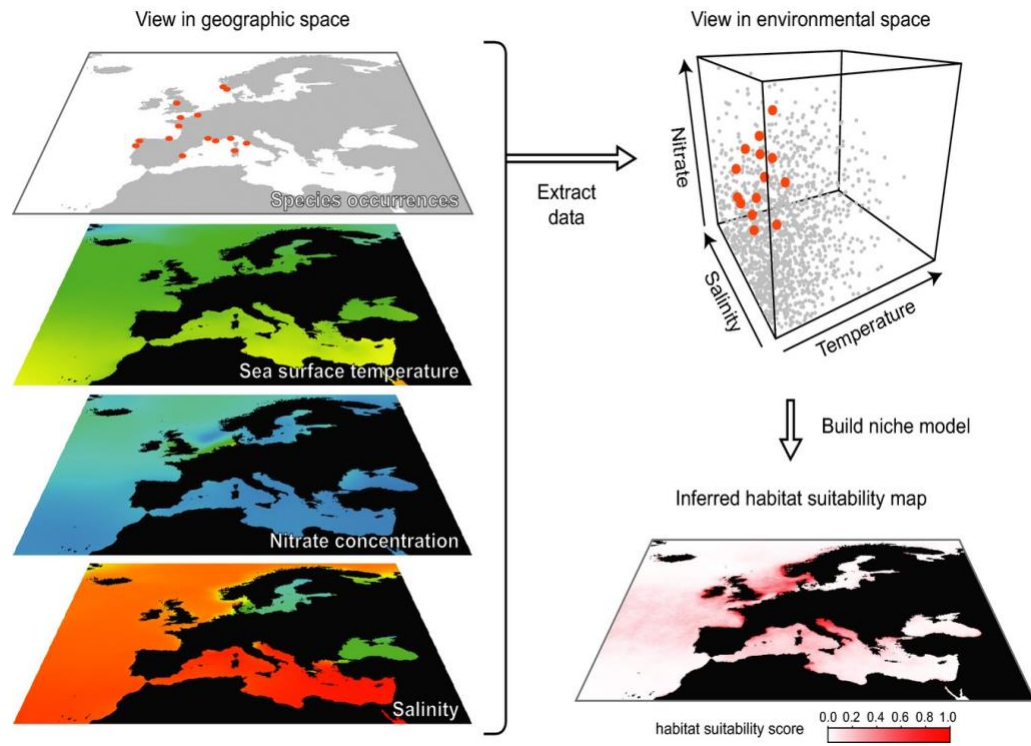
Figures

Figure 1. BAM Diagram



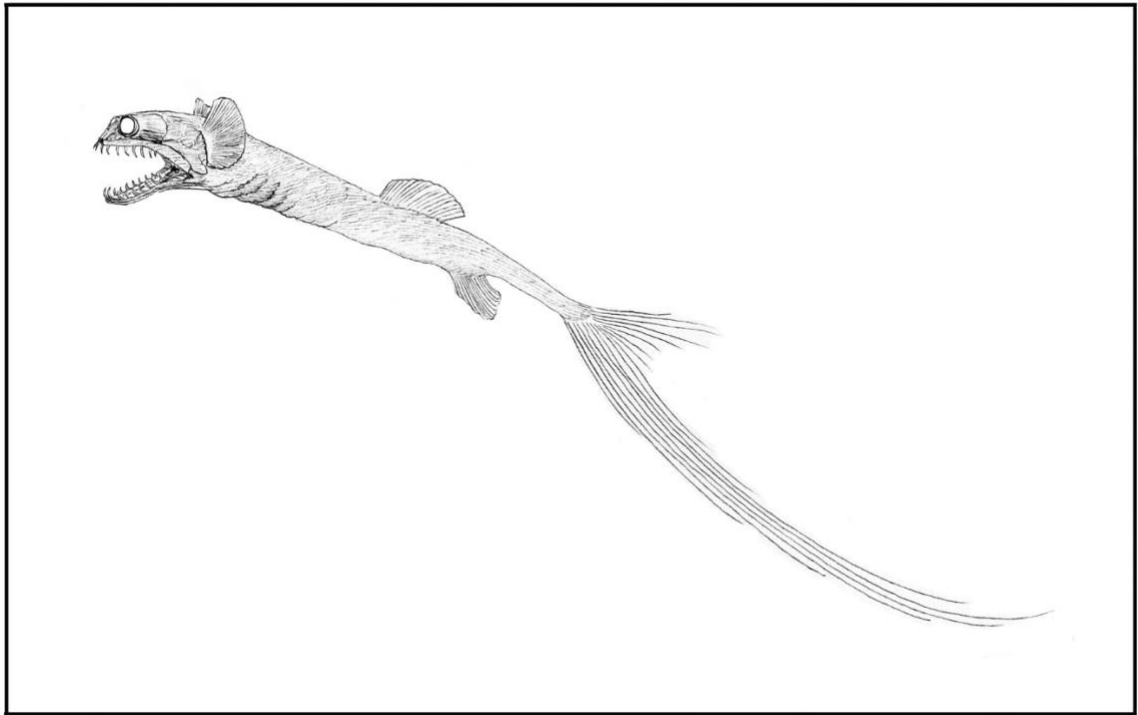
*Diagram of factors influencing species distribution in geographic space.  
Diagram courtesy of Professor Pat Farris of Los Angeles Pierce College, Department of  
Life Sciences (Retired).*

Figure 2. Representative Species Distribution Model



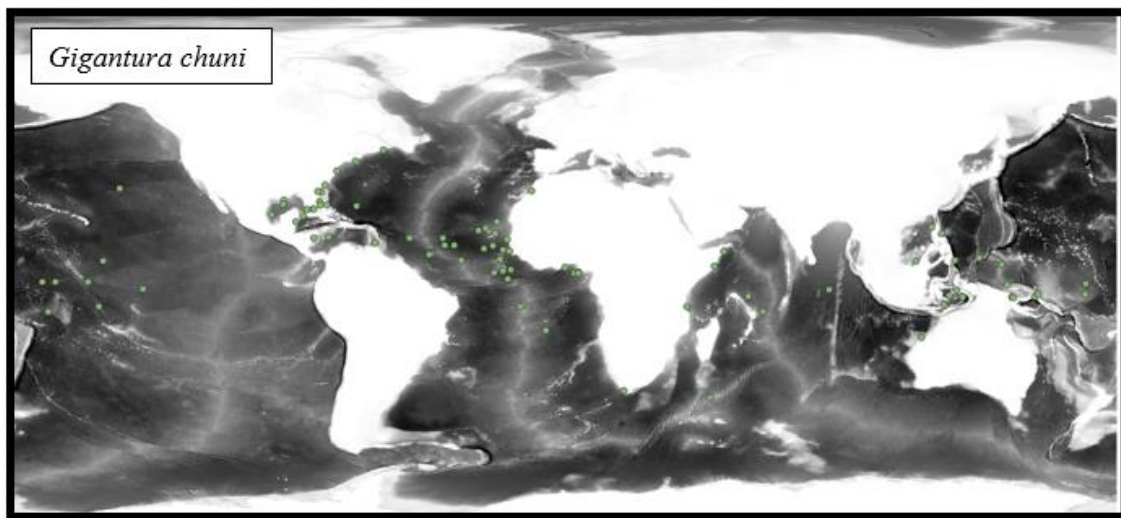
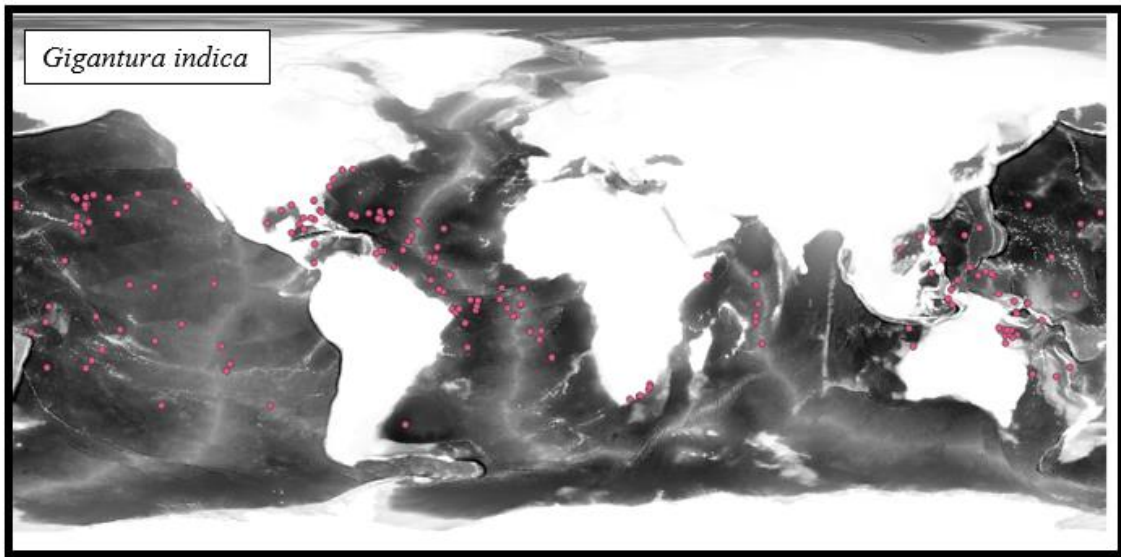
*Diagram adapted from Marcelino and Verbruggen, 2015.*

Figure 3. Telescopefish, *Gigantura indica*



*Drawing courtesy of Professor Pat Farris of Los Angeles Pierce College.*

Figure 4. Occurrence data plotted in QGIS for each species

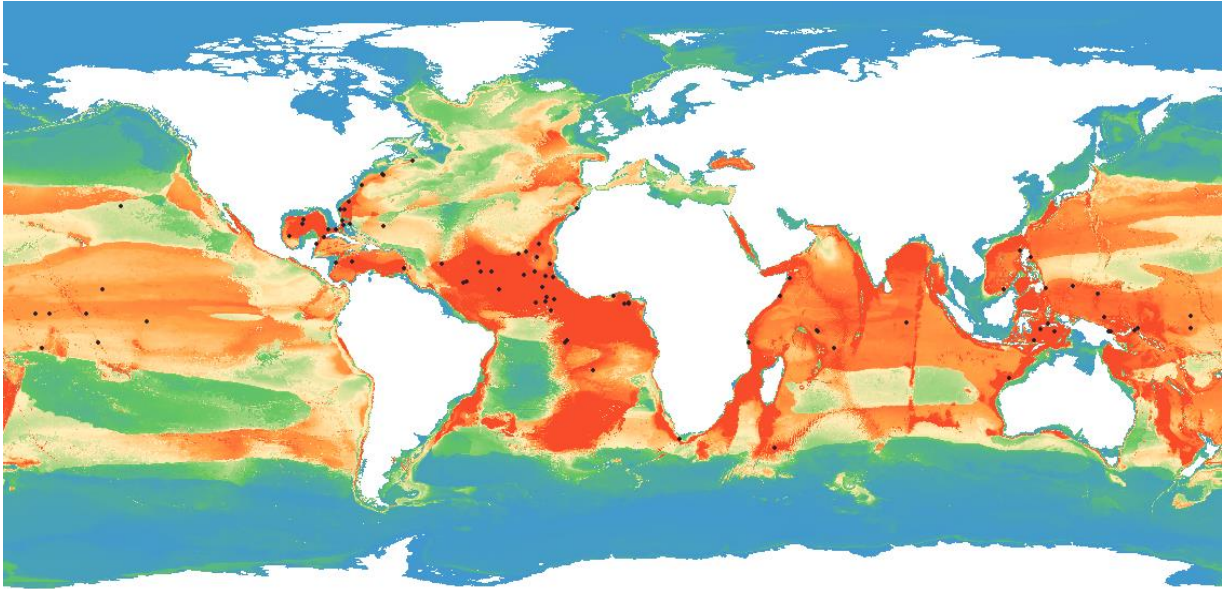


*Distribution of Gigantura indica in QGIS (top). Distribution of Gigantura chuni (bottom).*

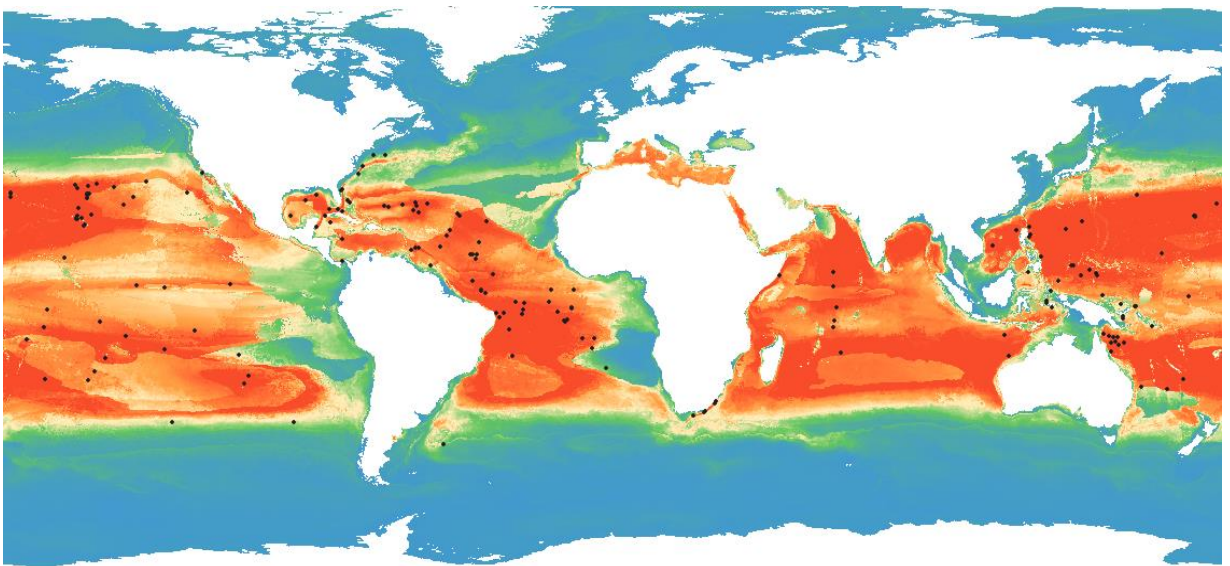


Figure 5 – Projected maps for habitat suitability

A.



B.



*Projected maps for Gigantura chuni (A) and Gigantura indica (B). The cumulative output ranges between 0 and 1 and represents probability of presence relative to suitable habitat.*

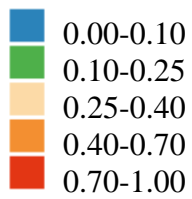




Figure 6. Environmental variable contribution.

a. *G. chuni*

Variable	Percent contribution	Permutation importance
Sea surface chlorophyll	23.6	10.9
Bottom temperature	22.7	46.9
Bottom nitrate	19.8	20.4
Sea surface silicate	19.2	16.2
Sea surface current velocity	14.7	5.6

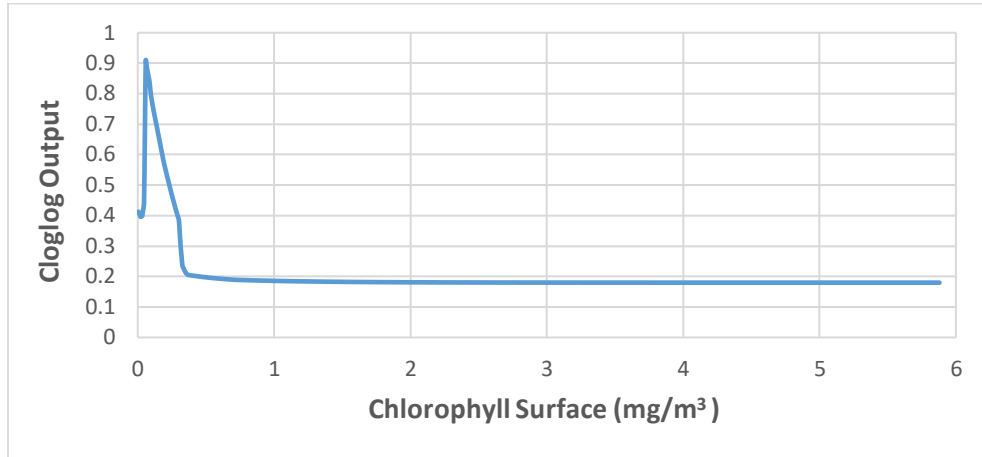
b. *G. indica*

Variable	Percent contribution	Permutation importance
Sea surface chlorophyll	43.5	45.1
Sea Surface silicate	21.5	18.3
Water column pH	14.8	12.1
Bottom nitrate	12.1	8.8
Sea Surface current velocity	8.1	15.8

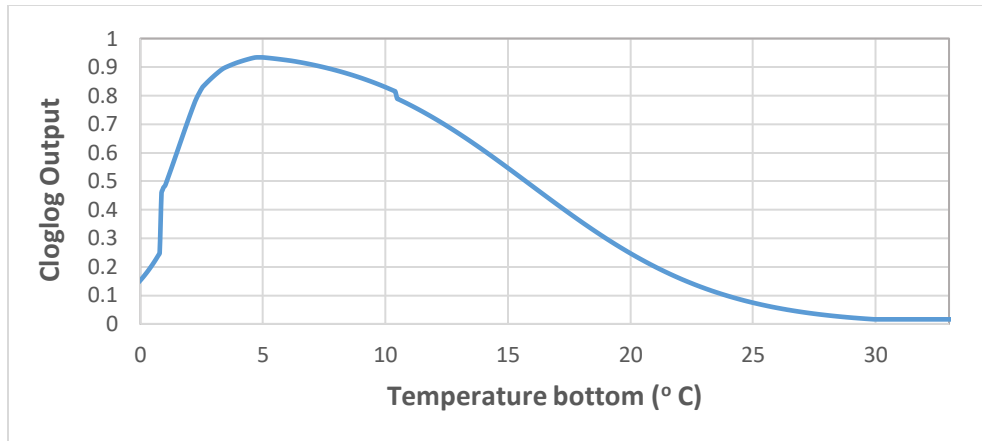
Variable contribution for each species. *G. Chuni* (a); *G. indica* (b).

Figure 7. Response Curves-*Gigantura chuni*

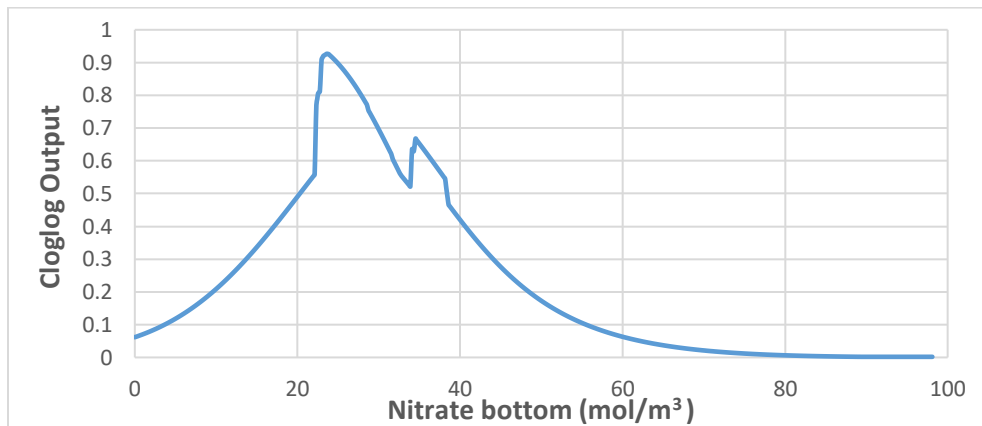
a.



b.

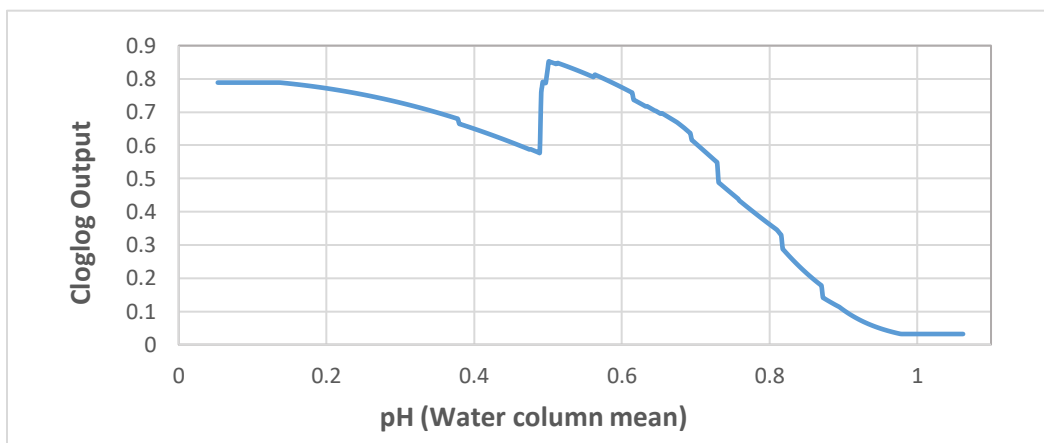
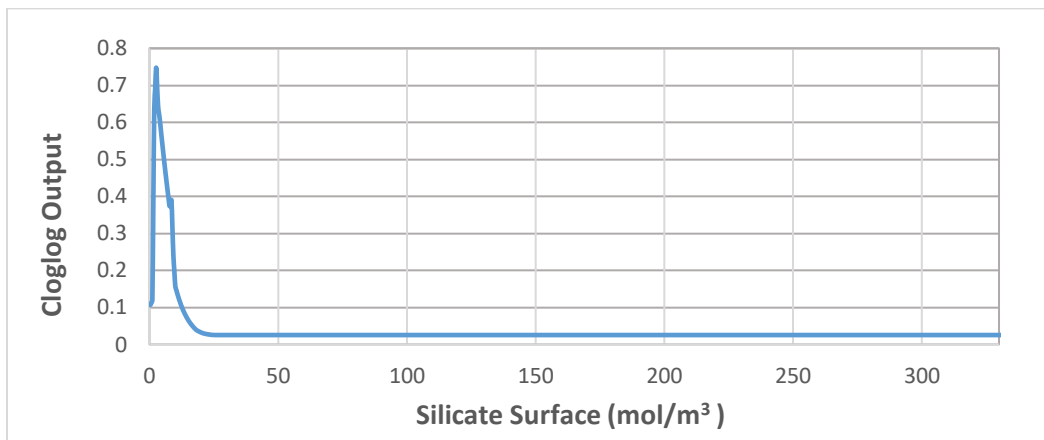
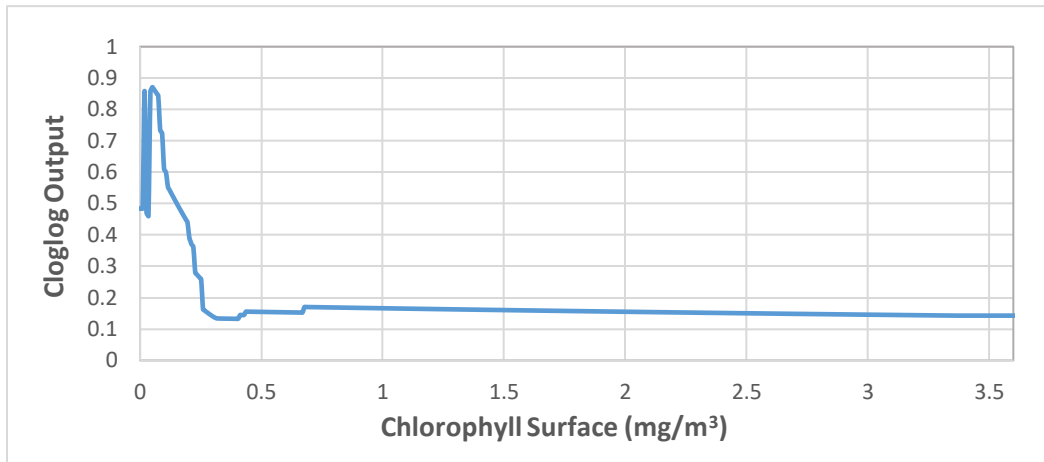


c.



Response curve for top three contributing variables for *Gigantura chuni*.

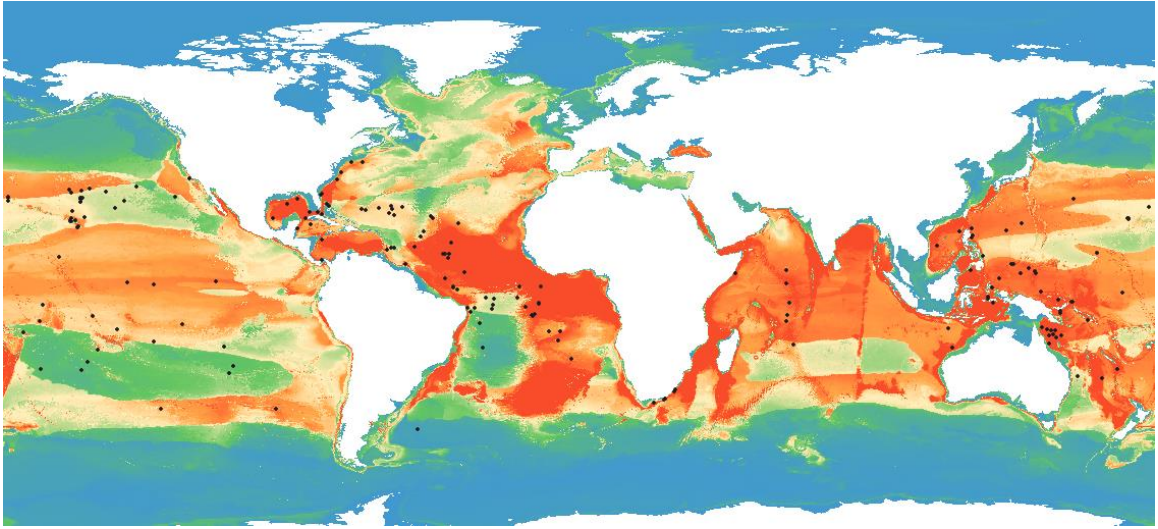
Figure 8. Response Curves-*Gigantura indica*



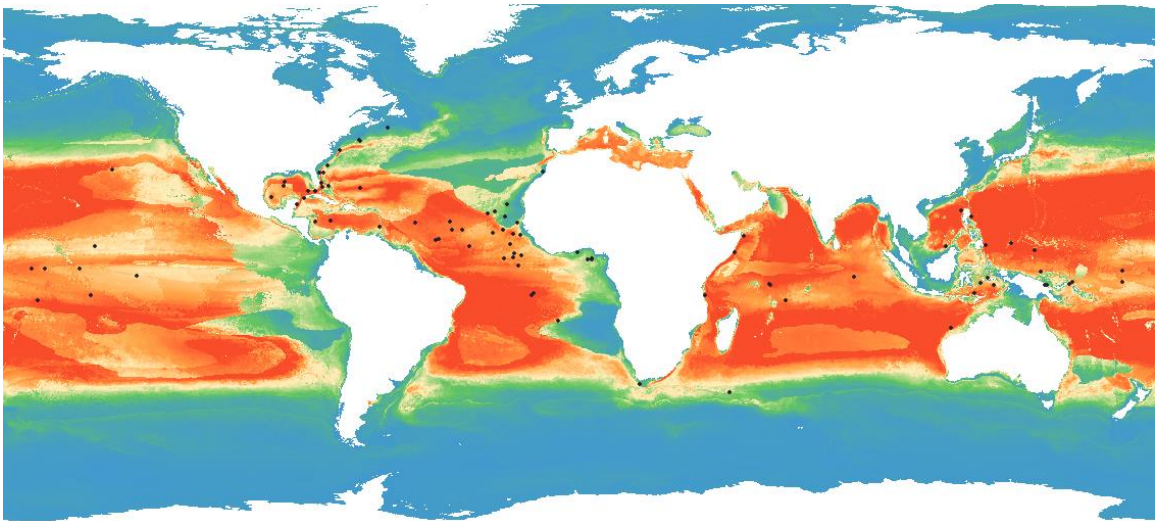
Response curve for top three contributing variables for *Gigantura indica*.

Figure 9. Reciprocal predictivity of *Gigantura chuni* and *Gigantura indica* models.

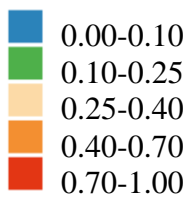
A.



B.



*Reciprocal geographic predictions of sister-species pair. Projected map for Gigantura chuni model with overlay of Gigantura indica occurrence data (A.) Projected map of Gigantura indica model with overlay of Gigantura chuni occurrence data (B).*



## References

- Anderson, R. P., & Raza, A. (2010). The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: Preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela: Effect of study region on models of distributions. *Journal of Biogeography*, 37(7), 1378–1393. <https://doi.org/10.1111/j.1365-2699.2010.02290.x>
- Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., De Clerck, O., & Tittensor, D. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, 27(3), 277–284. <https://doi.org/10.1111/geb.12693>
- Barber, R. A., Ball, S. G., Morris, R. K. A., & Gilbert, F. (2022). Target-group backgrounds prove effective at correcting sampling bias in Maxent models. *Diversity and Distributions*, 28(1), 128–141. <https://doi.org/10.1111/ddi.13442>
- Barracough, T. G., & Vogler, A. P. (2000). Detecting the Geographical Pattern of Speciation from Species-Level Phylogenies. *The American Naturalist*, 155(4), 419–434. <https://doi.org/10.1086/303332>
- Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., Soberon, J., & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222(11), 1810–1819. <https://doi.org/10.1016/j.ecolmodel.2011.02.011>
- Bentlage, B., Peterson, A. T., Barve, N., & Cartwright, P. (2013). Plumbing the depths: Extending ecological niche modelling and species distribution modelling in three dimensions: 3D ENM and SDM. *Global Ecology and Biogeography*, 22(8), 952–961. <https://doi.org/10.1111/geb.12049>
- Bioclim: The first species distribution modelling package, its early applications and relevance to most current MaxEnt studies.* (n.d.). <https://doi.org/10.1111/ddi.12144>
- Brunn, A. Fr. (1956). The Abyssal Fauna: Its Ecology, Distribution and Origin. *Nature*, 177(4520), 1105–1108. <https://doi.org/10.1038/1771105a0>
- Castro, P., & Huber, M. E. (2019). *Marine biology* (Eleventh edition). McGraw-Hill Education.
- Cobos, M. E., Peterson, A. T., Osorio-Olvera, L., & Jiménez-García, D. (2019). An exhaustive analysis of heuristic methods for variable selection in ecological niche modeling and species distribution modeling. *Ecological Informatics*, 53, 100983. <https://doi.org/10.1016/j.ecoinf.2019.100983>

- Colwell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences*, 106(supplement\_2), 19651–19658. <https://doi.org/10.1073/pnas.0901650106>
- Crotty, S. M., & Bertness, M. D. (2015). Positive interactions expand habitat use and the realized niches of sympatric species. *Ecology*, 96(10), 2575–2582. <https://doi.org/10.1890/15-0240.1>
- DeVaney, S. C. (2016). Species Distribution Modeling of Deep Pelagic Eels. *Integrative and Comparative Biology*, 56(4), 524–530. <https://doi.org/10.1093/icb/icw032>
- DeVaney, S. C., McNyset, K. M., Williams, J. B., Peterson, A. T., & Wiley, E. O. (2009). A Tale of Four “Carp”: Invasion Potential and Ecological Niche Modeling. *PLoS ONE*, 4(5), e5451. <https://doi.org/10.1371/journal.pone.0005451>
- Elith, J., & Franklin, J. (2013). Species Distribution Modeling. In *Encyclopedia of Biodiversity* (pp. 692–705). Elsevier. <https://doi.org/10.1016/B978-0-12-384719-5.00318-X>
- Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
- Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. *Diversity and Distributions*, 17(1), 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Elton, C. (1927). Animal Ecology. *Animal Ecology*.
- Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. *PLoS ONE*, 9(5), e97122. <https://doi.org/10.1371/journal.pone.0097122>
- GBIF.org. (2022a). Retrieved April 12, 2022, from <https://www.gbif.org/occurrence/download/0057202-210914110416597>
- GBIF.org. (2022b). Retrieved April 12, 2022, from <https://www.gbif.org/occurrence/download/0057194-210914110416597>
- GBIF.org. (2015c). Retrieved December 19, 2015, from <http://doi.org/10.15468/dl.y8umvf>

- Gigantura indica*, Telescopefish. (n.d.). Retrieved April 16, 2021, from <https://www.fishbase.se/summary/Gigantura-indica>
- Grinnell, J. (1917). The Niche-Relationships of the California Thrasher. *The Auk*, 34(4), 427–433. <https://doi.org/10.2307/4072271>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2–3), 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- Herring, P. J. (2002). *The biology of the deep ocean*. Oxford University Press.
- Hijmans, R. J., Schreuder, M., De la Cruz, J., & Guarino, L. (1999). Using GIS to check coordinate of germplasm accessions. *Genetic Resources and Crop Evolution*, 46(3), 291–296. <https://doi.org/10.1023/A:1008628005016>
- Hutchinson, G. E. (1957). Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22(0), 415–427. <https://doi.org/10.1101/SQB.1957.022.01.039>
- IUCN. (2013a). *Gigantura chuni*: Nunoo, F., Russell, B., Poss, S. & Bannermann, P.: *The IUCN Red List of Threatened Species 2015: e.T13462737A15603175* [Data set]. International Union for Conservation of Nature. <https://doi.org/10.2305/IUCN.UK.2015-4.RLTS.T13462737A15603175.en>
- IUCN. (2013b). *Gigantura indica*: Bannermann, P., Nunoo, F., Poss, S. & Russell, B.: *The IUCN Red List of Threatened Species 2015: e.T190180A15603010* [Data set]. International Union for Conservation of Nature. <https://doi.org/10.2305/IUCN.UK.2015-4.RLTS.T190180A15603010.en>
- Kass, J., Muscarella, R., Galante, P., Bohl, C., Buitrago-Pinilla, G., Boria, R., Soley-Guardia, M., Anderson, R. (2021). *ENMEvaluate* (2.0.3) [R Package]. <https://jamiemkass.github.io/ENMeval/articles/ENMeval-2.0.0-vignette.html>.
- Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12(4), 334–350. <https://doi.org/10.1111/j.1461-0248.2008.01277.x>
- Konstantinidis, P., & Johnson, G. D. (2016). Osteology of the telescopefishes of the genus *Gigantura* (Brauer, 1901), Teleostei: Aulopiformes. *Zoological Journal of the Linnean Society*. <https://doi.org/10.1111/zoj.12469>
- Kozak, K. H., Graham, C. H., & Wiens, J. J. (2008). Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution*, 23(3), 141–148. <https://doi.org/10.1016/j.tree.2008.02.001>



- Kozak, K. H., & Wiens, John J. (2006). Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution*, 60(12), 2604–2621. <https://doi.org/10.1111/j.0014-3820.2006.tb01893.x>
- Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6), 1659–1673. <https://doi.org/10.2307/1939924>
- Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17(2), 145–151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>
- Losos, J. B., & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology & Evolution*, 18(5), 220–227. [https://doi.org/10.1016/S0169-5347\(03\)00037-5](https://doi.org/10.1016/S0169-5347(03)00037-5)
- Marcelino, V. R., & Verbruggen, H. (2015). Ecological niche models of invasive seaweeds. *Journal of Phycology*, 51(4), 606–620. <https://doi.org/10.1111/jpy.12322>
- Marshall, N. B. (1971). *Explorations in the Life of Fishes*. <https://doi.org/10.4159/harvard.9780674865129>
- MyNyset, K. M. (2009). Ecological niche conservatism in North American freshwater fishes. *Biological Journal of the Linnean Society*, 96(2), 282–295. <https://doi.org/10.1111/j.1095-8312.2008.01121.x>
- Merow, C., Smith, M. J., & Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36(10), 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Miller, J. A., & Holloway, P. (2015). Incorporating movement in species distribution models. *Progress in Physical Geography: Earth and Environment*, 39(6), 837–849. <https://doi.org/10.1177/0309133315580890>
- Miya, M., & Nishida, M. (1997). Speciation in the open ocean. *Nature*, 389(6653), 803–804. <https://doi.org/10.1038/39774>
- Nelson, J. S. (2006). *Fishes of the world* (4th ed). John Wiley.
- Pearson, Richard. (2008). Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. *American Museum of Natural History*. <http://ncep.amnh.org>
- Peterson, A. T. (2001). Predicting Species' Geographic Distributions Based on Ecological Niche Modeling. *The Condor*, 103(3), 599–605. <https://doi.org/10.1093/condor/103.3.599>
- Peterson, A. T. (Ed.). (2011a). *Ecological niches and geographic distributions*. Princeton University Press.



- Peterson, A. T. (2011b). Ecological niche conservatism: A time-structured review of evidence: Ecological niche conservatism. *Journal of Biogeography*, 38(5), 817–827.  
<https://doi.org/10.1111/j.1365-2699.2010.02456.x>
- Peterson, A. T., Aiello-Lammens, M., Amatulli, G., Anderson, R., Cobos, M., Diniz-Filho, J. A., Escobar, L., Feng, X., Franklin, J., Gadelha, L., Georges, D., Guéguen, M., Gueta, T., Ingenloff, K., Jarvie, S., Jiménez, L., Karger, D., Kass, J., Kearney, M., Zurell, D. (2022). ENM2020: A Free Online Course and Set of Resources on Modeling Species' Niches and Distributions. *Biodiversity Informatics*, 17. <https://doi.org/10.17161/bi.v17i.15016>
- Peterson, A. T., & Soberón, J. (2012a). Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right. *Natureza & Conservação*, 10(2), 102–107.  
<https://doi.org/10.4322/natcon.2012.019>
- Peterson, A. T., & Soberón, J. (2012b). Integrating fundamental concepts of ecology, biogeography, and sampling into effective ecological niche modeling and species distribution modeling. *Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology*, 146(4), 789–796.  
<https://doi.org/10.1080/11263504.2012.740083>
- Peterson, A. T., Soberón, J., & Sánchez-Cordero, V. (1999). Conservatism of Ecological Niches in Evolutionary Time. *Science*, 285(5431), 1265–1267.  
<https://doi.org/10.1126/science.285.5431.1265>
- Phillips, S. J. 2017. A Brief Tutorial on Maxent. Available from url:  
[http://biodiversityinformatics.amnh.org/open\\_source/maxent/](http://biodiversityinformatics.amnh.org/open_source/maxent/). Accessed on 2022-6-30.
- Phillips, N. D., Reid, N., Thys, T., Harrod, C., Payne, N. L., Morgan, C. A., White, H. J., Porter, S., & Houghton, J. D. R. (2017). Applying species distribution modelling to a data poor, pelagic fish complex: The ocean sunfishes. *Journal of Biogeography*, 44(10), 2176–2187.  
<https://doi.org/10.1111/jbi.13033>
- Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: An open-source release of Maxent. *Ecography*, 40(7), 887–893.  
<https://doi.org/10.1111/ecog.03049>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3–4), 231–259.  
<https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, 31(2), 161–175.  
<https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Polyakov, I. V., Pnyushkov, A. V., & Carmack, E. C. (2018). Stability of the arctic halocline: A new indicator of arctic climate change. *Environmental Research Letters*, 13(12), 125008.  
<https://doi.org/10.1088/1748-9326/aaec1e>

- QGIS Development Team. 2022. QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>".
- Porteiro, F. M., & Sutton, T. (2007). Midwater Fish Assemblages and Seamounts. In T. J. Pitcher, T. Morato, P. J. B. Hart, M. R. Clark, N. Haggan, & R. S. Santos (Eds.), *Seamounts: Ecology, Fisheries & Conservation* (pp. 101–116). Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470691953.ch6>
- Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters*, 3(4), 349–361. <https://doi.org/10.1046/j.1461-0248.2000.00143.x>
- Reese, G.C., Wilson, K. R., Hoeting, J. A., & Flather, C. H. (2005). Factors affecting species distribution predictions: A simulation modeling experiment. *Ecological Applications*, 15(2), 554–564.
- Richards, T. M., Gipson, E. E., Cook, A., Sutton, T. T., & Wells, R. J. D. (2019). Trophic ecology of meso- and bathypelagic predatory fishes in the Gulf of Mexico. *ICES Journal of Marine Science*, 76(3), 662–672. <https://doi.org/10.1093/icesjms/fsy074>
- Sbrocco, E. J., & Barber, P. H. (2013). MARSPEC: Ocean climate layers for marine spatial ecology: *Ecological Archives* E094-086. *Ecology*, 94(4), 979–979. <https://doi.org/10.1890/12-1358.1>
- Shao, J., & Tu, D. (1996). *The Jackknife and bootstrap* (2., corr. print). Springer.
- Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10(12), 1115–1123. <https://doi.org/10.1111/j.1461-0248.2007.01107.x>
- Soberón, J., & Arroyo-Peña, B. (2017). Are fundamental niches larger than the realized? Testing a 50-year-old prediction by Hutchinson. *PLOS ONE*, 12(4), e0175138. <https://doi.org/10.1371/journal.pone.0175138>
- Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, 106(supplement\_2), 19644–19650. <https://doi.org/10.1073/pnas.0901637106>
- Soberón, J., & Peterson, A. T. (2005). Interpretation of Models of Fundamental Ecological Niches and Species' Distributional Areas. *Biodiversity Informatics*, 2(0). <https://doi.org/10.17161/bi.v2i0.4>
- Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & De Clerck, O. (2012). Bio-ORACLE: A global environmental dataset for marine species distribution modelling: Bio-ORACLE marine environmental data rasters. *Global Ecology and Biogeography*, 21(2), 272–281. <https://doi.org/10.1111/j.1466-8238.2011.00656.x>

- Warren, D. L., Dornburg, A., Zapfe, K., & Iglesias, T. L. (2021). The effects of climate change on Australia's only endemic Pokémon: Measuring bias in species distribution models. *Methods in Ecology and Evolution*, 12(6), 985–995. <https://doi.org/10.1111/2041-210X.13591>
- Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21(2), 335–342. <https://doi.org/10.1890/10-1171.1>
- Warren, D. Matzke, N., Cardillo, M., Baumgartner, J., Beaumont, L., Huron, N., Simoes, M., Iglesias, T., Dinnage, R. (2021). *ENMTools* (1.0.6) [R Package]. (<<https://orcid.org/0000-0002-8747-2451>>)
- Wiens, J. J. (2004). Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, 58(1), 193–197. <https://doi.org/10.1111/j.0014-3820.2004.tb01586.x>
- Wiens, J. J., Ackerly, D. D., Allen, A. P., Anacker, B. L., Buckley, L. B., Cornell, H. V., Damschen, E. I., Jonathan Davies, T., Grytnes, J.-A., Harrison, S. P., Hawkins, B. A., Holt, R. D., McCain, C. M., & Stephens, P. R. (2010). Niche conservatism as an emerging principle in ecology and conservation biology: Niche conservatism, ecology, and conservation. *Ecology Letters*, 13(10), 1310–1324. <https://doi.org/10.1111/j.1461-0248.2010.01515.x>
- Wiens, J. J., & Graham, C. H. (2005). Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. *Annual Review of Ecology, Evolution, and Systematics*, 36(1), 519–539. <https://doi.org/10.1146/annurev.ecolsys.36.102803.095431>
- Zimmer, C. (2019). *Evolution: Making sense of life* (3rd edition). W. H. Freeman and Co. (2016).