Species Distribution Models and Island Biogeography: Challenges and Prospects

Eva C Benavides Ríos

A thesis submitted to the University of Birmingham

for the degree of Doctor of Philosophy

School of Geography, Earth and Environmental Sciences
College of Life and Environmental Sciences
University of Birmingham
September 2024

ABSTRACT

The distinct characteristics of oceanic islands—restricted geographic boundaries, remoteness, lack of competitors and reduced habitat diversity—have made them global hotspots of endemism. However, these factors also mean that many island species have small populations, low genetic diversity and reduced competitive abilities. The arrival of humans to these systems exposed their fragile vulnerability, causing numerous species extinctions, many more so than on mainland environments.

In addition to their inherent vulnerability, oceanic islands are especially susceptible to environmental change due to their geographic isolation and limited physical space. Unlike many continental species, island species, especially endemics, face restricted opportunities to adapt through migration or by tracking suitable climatic conditions, given their confined distribution and isolation by surrounding aquatic barriers.

To better understand and predict species loss and range shifts in response to global change, this thesis focuses on Species Distribution Models (SDMs). It addresses the limitations of existing SDM frameworks, which are generally designed for continental species with extensive occurrence data, by tailoring the models to the unique conditions of island environments. The thesis focuses on

improving SDM applications through the use of high-resolution environmental descriptors (≤500m), given that traditional SDM approaches often rely on coarse climatic variables, which have been shown to be inadequate predictors of island species distributions in many cases. However, the use of high-resolution data raises important questions about how such predictors interact with the often-limited occurrence data available for island species.

To address these challenges, the thesis focused on the Revillagigedo Volcanic Archipelago, using it as a 'model archipelago' to provide insights relevant to a wider range of islands. Plant species were chosen as the focal taxon due to their critical ecological roles and usefulness as conservation indicators, with plant presence/absence data collected through field sampling. The thesis had two main objectives: to enhance SDM applications on islands through the use of high-resolution environmental descriptors, and to assess the robustness of fine-scale predictors, particularly in tropical island environments. The emphasis on tropical island environments was due to both the location of the model archipelago and, more broadly, the rich, yet understudied biodiversity of these regions. This research highlighted the potential of SDMs to deepen our understanding of ecosystems that, despite their global biodiversity significance, is often limited due to insufficient data.

Key findings include the identification of (1) minimum data requirements at very high resolutions (30m) for accurate presence-only species distribution modelling on islands, and (2) strategies to balance data limitations with pseudo-absence data to optimise predictability and accuracy. The research also investigates the impact of spatial inaccuracies in occurrence records, finding that while these inaccuracies are relatively less important for the SDMs of widespread species, they introduce greater uncertainty for range-restricted species; however, these uncertainties were found to be mitigated by the use of an adequate methodological framework.

The thesis further proposes methods for acquiring microclimatic descriptors that remain useful even in regions with sparse meteorological data, and provides strategies for validating these descriptors in future studies. The most informative analytical scales for these microclimatic data, for both predicting and projecting SDMs, and taking into consideration the spatial differences between training and projection areas, are identified.

The final empirical chapter of the thesis applies the lessons learned from the previous chapters in regard to the optimal implementation of SDMs in islands to analyse the impacts of future climate change on island species using a microclimatic approach. The results of this analysis confirm that endemic species' ranges are disproportionately threatened in most global biodiversity hotspots

analysed, and that the current instruments to categorise species risks (e.g., the IUCN Red List) are not able to effectively identify the vulnerability of (island) species to climate change. This research marks a significant advancement in our understanding of island species' responses to climate change at microclimatic scales and offers a crucial baseline for developing more targeted mitigation strategies in different archipelagos around the world.

Dedication

To my dear mentor and friend, the late Dr. Aurora Breceda-Solis, whose passion for plants, nature, and people, along with her unwavering support, sparked in me the passion and courage to embark on the path of science. This thesis is a testament to the power of her influence.

ACKNOWLEDGEMENTS

I feel grateful and fortunate to have been accompanied by my supervising committee. First and foremost, I thank Dr Tom Matthews for giving me the opportunity to continue learning about island ecology. I also thank him for his warmth and patience, which supported me through the challenge of writing my doctoral project in my second language. I also extend my gratitude to my secondary supervisors, Dr Laura Graham and Professor Jon Sadler, who provided kind, patient, and highly enriching guidance throughout this project.

I would like to thank the Mexican Council for Science and Technology and the University of Guadalajara for funding this doctoral project.

On a personal note, I wish to thank my partner, Kristoffer, who, despite the physical distance between us, gave me unconditional support, making us stronger.

I also want to thank my UK friends, who became my little Italian-English family during these years. You provided me with a wonderful support network and helped me find balance in this doctoral project, even when my passion for it sometimes made me lose sight of enjoying other aspects of life.

I would also like to thank my friends in Mexico, who always offered me unconditional support and helped in every way possible to ensure this project was successfully completed. Special thanks to Eliza Álvarez for tracking my stipend monthly. As I often said, "sorry to bother you—with my friendship!"

And finally, I am immensely grateful to my family, especially to my Mum and sister, who have always supported me in pursuing my dreams, even when it meant being physically apart from them.

"There must be a few experiences in the biologist’s world to compare with approaching a 'new' island by boat. The ingredients are adventure and suspense, mystery, and perhaps even a little danger. There are feelings of discoveries to be made, knowledge to be extended, curiosity to be both piqued and satisfied. Such feelings are shared not only among natural historians but any adventurous and curious traveller; the most difficult the island to reach, the keener the excitement of the visit"

-Cody, Moran, & Thompson, *Island Biogeography in the Sea of Cortés II*

List of papers

This thesis incorporates the following paper, which corresponds to one of the four empirical chapters:

- **Benavides, E[1]**., Sadler, J.[1,2], Graham[1,2,3], L., & Matthews, T. J[1,2,4]. (2024). Species distribution models and island biogeography: challenges and prospects. *Global Ecology and Conservation*, e02943.

The following co-authors contributed to the work in the following ways:

Thomas J. Matthews: Supervision, including contributing to improving the methodology and analyses. Reviewed and edited the manuscript.

Jon Sadler: Supervision, including contributing to improving results presentation and analyses. Reviewed and edited the manuscript.

Laura Graham: Supervision, including improving paper structure and analyses. Reviewed and edited the manuscript.

[1]GEES (School of Geography, Earth and Environmental Sciences), University of Birmingham, B15 2TT, UK

[2]Birmingham Institute of Forest Research, University of Birmingham, Birmingham B15 2TT, UK

[3]Biodiversity, Ecology & Conservation Group, International Institute for Applied Systems Analysis, Austria

[4]CE3C – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group / CHANGE – Global Change and Sustainability Institute and Universidade dos Açores – Faculty of Agricultural Sciences and Environment, PT-9700-042, Angra do Heroísmo, Açores, Portugal

Contents

FIGURES

TABLES

1 . Introduction

1.1 Research background

1.1.1 Global environmental change impacts on islands

Islands are at the centre of the current extinction crisis. While the world’s approximately 179,000 marine islands account for only 5% of the planet's surface, they contain 20% of all catalogued biodiversity, including disproportionately high rates of endemism (Kier et al., 2009; Medina et al., 2011; Whittaker et al., 2017). The unique and isolated nature of island ecosystems has led to the evolution of species that are found nowhere else on Earth. However, this same isolation makes these species particularly vulnerable to anthropogenic changes and disturbances (Whittaker et al., 2023).

The fragile nature of island ecosystems is highlighted by the fact that over 60% of documented terrestrial species extinctions since 1500 have been island endemics (Whittaker et al., 2017). Currently, 39% of insular species face an imminent risk of extinction (Ricketts et al., 2005), underscoring the urgent need for conservation efforts.

Three of the primary drivers of biodiversity loss on islands are habitat destruction, species introductions, and, more recently, climate change (Russell and Kueffer, 2019), with hunting also a driver of numerous island vertebrate extinctions (Whittaker et al., 2023). Species introductions in particular pose a significant threat to island biodiversity. Non-native species, whether plants, animals, or pathogens, have severely disrupted island ecosystems (Aguirre-Muñoz et al., 2011; Samuel et al., 2012; Wood et al., 2017; Cubas et al., 2018). This issue is particularly pronounced on oceanic islands—those islands that were never connected to the mainland—where many island species evolved without the competitive pressures faced by mainland species, such as large predators and herbivores (Whittaker et al., 2023). Consequently, these species often lack the behavioural and physiological adaptations needed to cope with environmental changes introduced by humans (Hanna & Cardillo, 2014; Whittaker et al., 2023). For example, introduced predators have decimated the populations of many native insular bird species, resulting in numerous extinctions (Matthews et al., 2022), and there are several examples of invasive plants outcompeting, and introduced herbivores consuming, native vegetation, altering habitats and reducing the availability of resources for native species (Walter & Levin, 2008; Gizicki et al., 2018; Parada-Díaz et al., 2022). Diseases brought by introduced species have often spread rapidly among the populations of numerous native island species that have no immunity, with the classic example being the impact of avian malaria on the Hawaiian honeycreepers (Samuel et al., 2012; Hume, 2017; Whittaker et al., 2023).

Habitat destruction on islands has often been particularly devastating due to the limited land area. Deforestation, urbanisation, and agricultural expansion have led to the loss of critical habitats for many species (Ferrer-Sánchez & Rodríguez-Estrella, 2015; Surasinghe et al., 2019; Connor et al., 2024). Unlike mainland ecosystems, where species can often migrate to new areas, island endemic species are trapped in their shrinking habitats.

Climate change adds an additional layer of threat. Rising sea levels, changing weather patterns, and increased frequency of extreme weather events can have catastrophic effects on island ecosystems (Veron et al., 2019; Macinnis et al., 2021). Many islands are low-lying and susceptible to flooding and erosion (Saintilan et al., 2023). Climate change also has a great potential to shrink or result in the complete loss of island habitats (Taylor & Kumar. 2016).

Addressing the extinction crisis on islands is a focal point for global biodiversity conservation efforts (Russell & Kueffer, 2019). However, precise ecological and geographical data, which serve as the evidence base for effective mitigation, remain sparse for the biota of most islands (Benavides et al., 2024). Tackling this issue requires innovative approaches, such as maximising the utility of available data through advanced analytical techniques and refining ecological modelling methods to better predict current species distributions and their responses to various global change scenarios.

1.1.2 Species Distribution Models: Linking Range Limits to Conservation

Understanding what governs species range limits is key to predict which species will contract, maintain, or expand their distributions and where and when these changes will occur within their geographical ranges (Bridle & Hoffman, 2022). However, island ecosystems are typically characterised by a dearth of species distributional data (Benavides et al., 2024), making it challenging to understand the ecogeographical limits of insular biota.

One tool with the potential to fill this knowledge gap are ecological niche models, also known as Species Distribution Models (SDMs). Over the past two decades, SDMs have become widely used statistical tools to predict species distributions, even for poorly known species with limited occurrence data (Radomski et al., 2022). By applying the ecological niche concept (Soberón & Nakamura 2009), SDMs describe species' environmental/climatic preferences across a geographical area by analysing the relationship between known species occurrences and a set of corresponding environmental variables (e.g., climate, topography, the presence of interacting species). SDM outputs are then used to create species distribution maps, which identify potentially suitable areas of occurrence and provide a quantitative baseline for conservation assessment, planning, and decision-making (Elith and Leathwick, 2009; Peterson and Soberón, 2012; Guisan et al., 2017; Araújo et al., 2019).

One of the primary challenges in using SDMs is navigating the vast and ever-expanding array of available methods, making it difficult for those unfamiliar with the field to select the most appropriate approach for their specific needs (Elith et al., 2006; Thuiller et al., 2009; Araújo et al., 2019). Additionally, significant differences between various methods have been documented, further complicating the selection process (Hao et al., 2019). Implementing these models is also complex, as numerous decisions must be made at each stage of the modelling process (Leroy, 2022). Although several guidelines are available (Elith et al., 2006; Guisan et al., 2017; Phillips et al., 2006; Thuiller et al., 2009), a unified SDM framework that can be generalised across all contexts is not stablished, and probably not feasible. As a result, the field remains in a state of continuous development, with approaches constantly being adapted to fit specific contexts.

In addition, the SDM field has largely focused on widespread species with abundant occurrence data, a fact that has heavily influenced SDM guidelines and "best practice" literature. However, rare species—those with small geographical ranges—present unique challenges that require specialised modelling approaches (Lomba et al., 2010; Gabor et al., 2020). These species typically have limited associated occurrence data, which restricts the number of predictor variables that can be used in model development (Breiner et al., 2015). To address this, an ensemble of small models—a method that combines weighted multiple models, each using a small subset of the relevant predictors—has been proposed as a strategy to reduce overfitting caused by using too many variables relative to the limited number of occurrence points (Lomba et al., 2010; Breiner et al., 2015). This technique offers a more effective way to model the distributions of range-restricted species, opening new avenues for SDMs.

A significant challenge in SDM application for range-restricted species is the effective use of presence-only data. Since direct absence data—which many SDM methods rely on—are generally unavailable, researchers typically create pseudo-absences. These are hypothetical locations where a species might actually occur but where no actual observations have been recorded (Lobo et al., 2010; Barbet-Massin et al., 2012; Morera-Pujol et al., 2020).

While pseudo-absences are crucial for accurate modelling, the best practices for generating them are still debated (Descombes et al., 2022) and, like other existing guidelines, are primarily tailored for species with broader ranges and larger study areas (Phillips et al., 2006; Barbet-Massin et al., 2014; Whitford et al., 2024). This focus makes them less effective for more localised contexts, such as island environments. Therefore, developing and refining pseudo-absence strategies specifically for island SDMs is an essential consideration (Benavides et al., 2024).

1.1.3 Scale-dependency in island SDMs

Despite advances in using SDMs to model range-restricted species (Breinner et al., 2015), significant methodological uncertainties remain that need to be addressed to ensure these models effectively contribute to island biogeography. A common and problematic assumption in this context is that broad climatic factors serve as informative and ecologically relevant predictors of species distributions (Pearson & Dawson, 2003). This assumption prevails largely because most SDM evaluations have traditionally been conducted at broad geographic scales (Thuiller et al., 2004; Syphard & Franklin, 2009).

However, incorporating environmental variables at finer spatial scales—such as microclimatic conditions, land cover, and terrain features—not only offers more accurate representations of the environments of many islands, where ecological processes often operate at finer scales (Whittaker et al., 2023), but also allows for assessments in islands where only a limited number of climatic grids (of a given resolution) are available. As such, several studies have demonstrated the value of finer-scale predictor variables in SDMs employed in island contexts (Bellamy et al., 2013; Heinänen et al., 2012; Lannuzel et al., 2021; Turvey et al., 2020; Segal et al., 2021).

The overreliance on broad-scale climatic variables has resulted in significant gaps in our understanding of the ecological validity of alternative environmental variables in SDM development (Austin & Van Niel, 2011; Syphard & Franklin, 2009). Additionally, there is uncertainty regarding the appropriate analytical scale needed to accurately describe species-environment relationships (Tshwene-Mauchaza & Aguirre-Gutiérrez, 2019; Tehrani et al., 2020; Ashrafzadeh et al., 2020). Addressing these issues is essential for improving the selection of predictor variables in island SDMs studies.

1.1.4 Coupling occurrence data with high-resolution predictors

The required shift to smaller analytical scales for studying island species-environment relationships is complicated by the often-limited availability of species observation records. While finer grid pixels in environmental data enable models to capture landscape variation more accurately, this increased precision typically requires a larger number of occurrence records to adequately represent a species' environmental needs. Research on small-ranged continental species, suggests that fewer occurrence records might be sufficient (van Proosdij et al., 2015). However, it remains unclear whether these lower data requirements hold true when finer predictor data and smaller spatial contexts are analysed.

Another critical challenge is that high-resolution environmental data demand occurrence records with matching spatial precision (Moudrý & Šímová, 2012; Sillero & Barbosa, 2020). However, spatial uncertainty (or coarse precision in accuracy) is a common issue in species’ occurrence data (McPherson et al., 2006). This leads to a difficult decision: whether to exclude imprecise records, which could reduce sample size and potentially diminish model accuracy (Graham et al., 2008; Fernandez et al., 2009; Naimi et al., 2011), or to include them, weighing the risk of adding noise against the potential benefit of increased sample size.

The low precision of occurrence records is particularly problematic in environments with sharp environmental gradients (Gábor et al., 2023), and for highly specialised species (Visscher et al., 2009; Gábor et al., 2020) —both common in island ecosystems. Addressing this issue is necessary, given the significant risk of including low-accuracy records in island SDMs, and the current lack of knowledge to effectively guide these decisions.

Thus, two critical questions arise for the application of SDMs on islands at fine scales in the face of data limitations: (1) how much data are needed to produce reliable high resolution SDMs across different species range sizes? And (2) what should be done with spatially uncertain occurrence information? Specifically, can including less precise data enhance predictions by increasing sample size, or does it introduce significant noise?

Given that different modelling methods vary in their sensitivity to factors such as sample size (Valavi et al., 2022), and occurrence data spatial inaccuracies (Gábor, et al., 2022, 2023) addressing these questions by identifying suitable methods could significantly improve the effectiveness of SDMs in island environments. By selecting modelling approaches that are more robust to limited data and spatial uncertainties, researchers can optimise the use of available records while reducing uncertainties in high-resolution SDM outputs.

1.1.5 Prioritising Among Priorities: Assessing threats on tropical insular hotspots

Global biodiversity hotspots are recognised on the basis of their high levels of plant endemism coupled with significant losses of vegetation cover (Myers et al., 2000; Mittermeier et al., 2013). Almost half of all global hotspots are located on islands, primarily in tropical developing countries. Despite being home to around a third of the world's critically endangered plants (Tershy et al., 2015), these regions often lack adequate funding for biodiversity research (Caujape-Castells et al., 2010; Habel et al., 2019).

While SDMs have been presented as a powerful tool for linking biogeographic knowledge with species conservation, their implementation in tropical island environments presents an additional layer of complexity. In addition to the challenges posed by sparse occurrence data, there is the issue of limited availability of environmental and climatic data at appropriate scales (Baker et al., 2017; Lannuzel et al., 2021).

The emergence of open-access satellite imagery and global datasets (e.g., NASA EOSDIS, 2024) offers promising solutions for modelling species distributions in tropical islands at fine-scale resolution. These datasets provide static variables like terrain and land cover, capturing crucial micro-environmental variation that influences species distribution patterns (Gabor et al., 2022; Heinänen et al., 2012; Lannuzel et al., 2021; Turvey et al., 2020; Segal et al., 2021). However, the use of these variables has two main limitations. First, their static nature hinders the tracking of species distribution changes over time. Second, they rely primarily on topography-related proxies, which may actually represent actual climatic variation, but could also be affected by factors like nutrient availability and soil properties (Meier et al., 2010). This reliance can obscure the specific influences on species distributions, which becomes problematic when ecological patterns beyond distributional patterns are of interest.

To address these limitations, it is essential to explore methods that directly incorporate climatic variables, enabling more precise assessments of species-climate associations in island ecosystems. Many frequently used databases (e.g., WorldClim: Fick & Hijmans, 2017), derived from interpolations of on-site meteorological data, have been useful for building SDMs (Vega et al., 2017). However, the resolution of these data (typically ≥1km) may not capture the scale at which island plants respond Chauvier et al., 2022), requiring the use of downscaling methods to better represent climatic variations in tropical island regions (Flint and Flint, 2012; Franklin et al., 2013; Khosravi et al., 2016).

Concerns exist about the accuracy of global databases in capturing local-scale climate variations in data-scarce regions like tropical environments (Soria-Auza et al., 2010; Bedia et al., 2013; Fernandez et al., 2013). These databases typically originate from global climate models, reanalysis products, or other coarse-resolution climate data sources, which may not adequately reflect local variations. This can introduce significant biases in downscaled estimates, compromising the reliability of modelling outputs. Therefore, there is a need to (i) investigate the reliability of downscaled high-resolution climatic estimates in regions with limited meteorological data and, (ii) assess the reliability and uncertainty of SDM predictions using global downscaled climate datasets in tropical environments.

The effectiveness of SDMs is typically measured by their ability to generate accurate predictions in static scenarios (spatially or temporally) and their capacity for projection into novel conditions (Araujo et al., 2019). A critical factor influencing both prediction accuracy and model transferability is the selection of predictor variable type. Existing research highlights limited understanding of how the choice of climatic or terrain variables impacts SDM outcomes (Bobrowski et al., 2017; Karger et al., 2017; Petitpierre et al., 2017). Moreover, studies examining these impacts at fine-grained scales are rare, contributing to further gaps in our knowledge (Austin & Van Niel, 2011; Stanton et al., 2012).

There is also a recognised risk that highly specific models, tailored to narrow study areas with high-resolution predictors, may suffer from overfitting (Manzoor et al., 2018). This overfitting can compromise the applicability of models when they are projected into new environmental contexts or different temporal scales.

In summary, achieving the right balance between accuracy and generalisability in high-resolution tropical island SDMs requires: 1) creating alternatives for finer-resolution predictors, 2) ensuring these predictors are statistically validated given the limited environmental and climatic data in tropical regions, and 3) identifying which types and resolutions maximise both accuracy and transferability. The end result is a model that can be reliably applied across different spatial and temporal scales, contributing to global change research and supporting conservation and management efforts in tropical island hotspots.

1.2 This study

1.2.1 Aims and scope

In light of the research gaps identified above, this study aims to (1) identify fundamental considerations related to the application of SDMs in island environments, considering occurrence data limitations and the restricted spatial context of island environments, (2) Investigate relevant data sources to address gaps in occurrence data for island species across various taxa, (3) enhance SDM implementation in island environments at fine-grained scales, (4) investigate the uncertainty associated with the use of environmental and climatic predictors in data-scarce island regions at fine analytical scales, and (5) assess how the predicted responses of insular plant species to climate change are related to factors such as species conservation status and degree of endemism.

Based on the aims outlined above, and focusing specifically on island plants, the thesis looks to address several research questions:

- Using previous research as a baseline, what are the challenges of SDM implementation in island environments, and how can we improve them? (Chapter 2)

- What sample size of occurrence data is needed to accurately predict the distributions of island species using high-resolution predictors, while accounting for varying specialisation levels? (Chapter 3)

- How does the spatial precision of occurrence data affect the accuracy of island SDM predictions at high resolutions? (Chapter 3)

- How does the choice of pseudo-absence data impact the accuracy of predicted island species distributions? (Chapter 3)
- Are downscaled predictors a reliable baseline to model island plant species distributions in tropical latitudes? (Chapter 4)
- Does the grain size of downscaled climatic estimates impact the accuracy and transferability of island SDMs to novel environments? (Chapter 4)
- Does the choice of predictor variable type (Bioclimatic vs. Terrain variables) impact the accuracy of SDMs on islands, as well as their transferability to novel environments? (Chapter 4)
- Are the predicted responses of species to future climate change related to their endemism and/or conservation status? (Chapter 5)

This thesis utilises two separate datasets to answer these questions. The first dataset comprises plant species presence/absence records from the Revillagigedo Archipelago, Mexico, obtained through my own field work, involving sampling using vegetation transects of varying length (100m-1km) depending on island size and habitat variability. The second dataset was curated from the global database GBIF (GBIF.org, 2024), covering several volcanic archipelagos within the (sub)tropics. Detailed descriptions of the datasets are provided in Section 1.3 and the relevant chapters where they are utilised.

1.2.2 Overview of chapters

The chapters are presented as standalone papers, with references included for each. Chapters 3-4 share portions of the Revillagigedo Archipelago dataset, and chapters 3-5 use similar methodologies and metrics; thus, there is some repetition of the methods and background in these chapters.

- Chapter 2 is a systematic review of the island SDM literature (224 published studies) to assess the appropriate use of SDMs in island biogeography, specifically focusing on marine islands (i.e., islands in the sea). Species were divided into different insular distribution categories (i.e., chorotypes: single island/archipelago endemics, non-endemic natives, and non-natives) in order to provide chorotype-specific SDM recommendations. The paper highlights how to navigate two fundamental considerations related to the application of SDMs in island environments. 1) Response variables, specifically the issue of small sample sizes for many island species. 2) Predictor variables, including (i) the selection of relevant environmental predictors at appropriate spatial grains, and (ii) addressing the truncation of environmental extent across the entire species range, especially for non-endemic species. This work has been published in Global Ecology and Conservation (DOI: https://doi.org/10.1016/j.gecco.2024.e02943).

- Chapter 3 utilises a portion of the Revillagigedo Archipelago dataset, along with simulated virtual species, the distributions of which resemble known species distribution patterns within the study area. This approach allows for the evaluation of methodological considerations necessary to produce accurate Presence-only Species Distribution Models (SDMs) at high resolutions, addressing the primary data challenges found in many published island SDM studies.

- Chapter 4 uses the Revillagigedo Archipelago dataset to investigate the use of downscaled global databases at a range of high resolutions (30, 100, and 500m) with species distribution models, contrasting the predicted distributions with those generated by a model of random species occupancy. Additionally, the accuracy and transferability of these models are tested to explore the potential of such datasets in studies that require model projections, such as those focused on climate change and invasive species.

- Chapter 5 utilises a dataset curated from the global GBIF database to examine the potential future effects of climate change on plant species distributions in three island biodiversity hotspots located within the tropics. The role of species endemism and conservation status on predicted responses is assessed.

1.3 Datasets and study locations

As outlined above, two primary datasets are used within this work: a plant species presence/absence records dataset from the Revillagigedo Archipelago, Mexico, obtained through field sampling, and a dataset compiled from the global database GBIF, comprising plant distribution data for several volcanic archipelagos within the (sub)tropics.

1.3.1 Plant occurrence database from Revillagigedo Archipelago, Mexico

The Revillagigedo plant dataset comprises occurrence (presence/absence) points obtained through field sampling from vegetation transects conducted between 2018 and 2022. The sampling efforts focused on the two largest islands in the archipelago, Socorro (132 km^2) and Clarion (19.8 km^2), chosen for their ecological significance as hosts to the majority of species in the archipelago (Figure 1.1).

During the first sampling period, plant records were collected in both the dry (May 2018) and wet (November 2018) seasons on Clarion Island, the least studied of the two (Benavides et al., 2019). In 2022, data were collected from Socorro Island at the beginning of the wet season (November 2022) to capture the largest possible range of both perennial and non-perennial species.

Depending on the island size and habitat variability, vegetation transects of varying lengths were implemented: 100m transects on Clarion and 500m to 1km transects on Socorro. These transects were systematically selected to capture environmental heterogeneity. To counteract potential sampling bias, I also ensured a widespread distribution of transects across the islands.

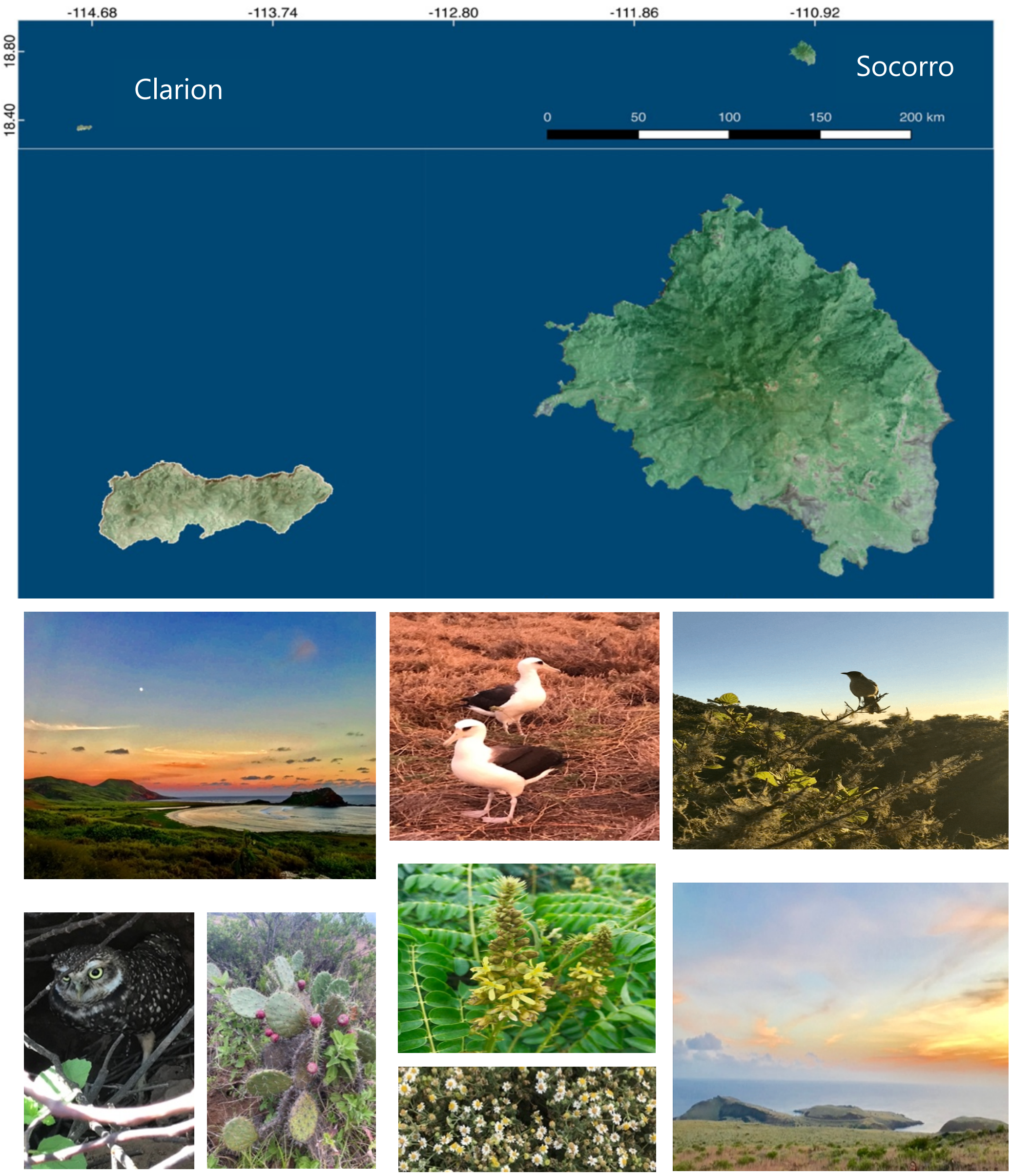


Figure 1.1 The two main islands of the Revillagigedo Archipelago, Mexico, with a selection of their biodiversity

1.3.2 GBIF island data: tropical volcanic archipelagos

The Global Biodiversity Information Facility (GBIF) is an extensive dataset resulting from collaborative efforts between data providers and taxonomists from numerous institutions, making it a valuable resource for biodiversity research and conservation (Beck et al., 2013). This database offers free access to digitised biological data from various sources, including museum collections and survey programs on a global scale. Here, we use a subset of the data extracted with the "Rgbif" R package (Chamberlain et al., 2024) within global island polygons (sourced from: Geodata R package; Hijmans et al., 2024). The informativeness of these data is discussed in more detail within Chapter 2, where we analysed which islands provide sufficient information to fit SDMs in islands across diverse taxa at the global scale. Subsequently, we selected volcanic archipelagos in (sub)tropical biodiversity hotspots (Myers et al., 2000; Mittermeier et al., 2013) that had sufficient plant occurrence data to analyse the effects of climate change at a macroecological scale in Chapter 5.

1.4 Summary and synthesis

The rapid alteration of species' geographic ranges due to human activities poses a significant threat to biodiversity, especially on islands where extinction rates are higher than in other terrestrial environments (Whittaker et al., 2017). Effective conservation strategies require a thorough understanding of the geographic and environmental limits of island biodiversity.

Given that the full distributions of many island species are unknown, species distribution models (SDMs) are essential tools for filling ecological knowledge gaps. They enable data-informed conservation and management decisions for these highly threatened components of global biodiversity. However, most SDM research has focused on species with broad ranges and ample occurrence data, typically focussing on continental species (Hickisch et al., 2019; Leroy, 2022).

Islands differ significantly from continental systems in many ways. Modelling island species distributions thus requires specific guidelines which consider the specific data limitations and characteristics of (oceanic) islands. Islands are often remote and challenging to access, with many species found in low abundance due to natural factors or human influences. Consequently, obtaining comprehensive and representative datasets of species occurrences in these environments is challenging. This is compounded by the small, bounded study areas that often exhibit considerable topographic and environmental variation. This situation raises several concerns: 1) the appropriate scale of predictors needed to capture local-scale environmental variations; 2) the quality and quantity of data required to describe species' habitat requirements considering island spatial idiosyncrasies; and 3) the reliability of predictors derived from high-resolution data products in accurately describing island environments despite the scarcity of environmental data on most islands.

The main objective of the thesis is to improve the accuracy and reliability of SDMs in island environments, while applying the insights gained from analyses of optimal island SDM implementation to better understand island plant distributions on (sub)tropical islands. Ultimately, it is hoped that the information generated from the analyses presented in this thesis will enhance conservation and management strategies for insular biodiversity going forward.

1.5 References

Aguirre-Muñoz, A., Samaniego-Herrera, A., Luna-Mendoza, L., Ortiz-Alcaraz, A., Rodríguez-Malagón, M., Méndez-Sánchez, F., ... & Latofski-Robles, M. (2011). Island restoration in Mexico: ecological outcomes after systematic eradications of invasive mammals. *Island invasives: eradication and management*, 250-258.

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

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

Baker, D. J., Hartley, A. J., Pearce-Higgins, J. W., Jones, R. G., & Willis, S. G. (2017). Neglected issues in using weather and climate information in ecology and biogeography. *Diversity and Distributions, 23*(3), 329-340.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*(2), 327-338.

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

Benavides, E., Kuethe, J. R., Ortiz-Alcaraz, A., & DE LA LUZ, J. L. (2019). *Oenothera resicum* (Onagraceae), a new species and the first record of the family from the Revillagigedo Archipelago, Mexico. *Phytotaxa*, *416*(1), 59-66.

Benavides, E., Sadler, J., Graham, L., & Matthews, T. J. (2024). Species distribution models and island biogeography: Challenges and prospects. *Global Ecology and Conservation*, e02943.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution, 6*(10), 1210-1218.

Bridle, J., & Hoffmann, A. (2022). Understanding the biology of species' ranges: when and how does evolution change the rules of ecological engagement? *Philosophical Transactions of the Royal Society B*, *377*(1848), 20210027.

Bobrowski, M., Gerlitz, L., & Schickhoff, U. (2017). Modelling the potential distribution of *Betula utilis* in the Himalaya. *Global Ecology and Conservation, 11,* 69-83.

Caujape-Castells, J., Tye, A., Crawford, D. J., Santos-Guerra, A., Sakai, A., Beaver, K., ... & Kueffer, C. (2010). Conservation of oceanic island floras: present and future global challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(2), 107-129.

Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, Ram K (2024). *rgbif: Interface to the Global Biodiversity Information Facility API*. R package version 3.8.0, https://CRAN.R-project.org/package=rgbif.

Connor, S. E., Lewis, T., van Leeuwen, J. F., Schaefer, H., Porch, N., Gomes, A. I., ... & Elias, R. B. (2024). Original plant diversity and ecosystems of a small, remote oceanic island (Corvo, Azores): Implications for biodiversity conservation. *Biological Conservation*, *291*, 110512.

Cubas, J., Martín-Esquivel, J. L., Nogales, M., Irl, S. D., Hernández-Hernández, R., López-Darias, M., ... & González-Mancebo, J. M. (2018). Contrasting effects of invasive rabbits on endemic plants driving vegetation change in a subtropical alpine insular environment. *Biological Invasions*, *20*, 793-807.

Descombes, P., Chauvier, Y., Brun, P., Righetti, D., Wüest, R. O., Karger, D. N., ... & Zimmermann, N. E. (2022). Strategies for sampling pseudo-absences for species distribution models in complex mountainous terrain. *BioRxiv*, 2022-03.

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., ... & E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), 129-151.

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.

Fernandez, M., Blum, S., Reichle, S., Guo, Q., Holzman, B., & Hamilton, H. (2009). Locality uncertainty and the differential performance of four common niche-based modeling techniques. *Biodiversity Informatics, 6*, 36-52.

Ferrer-Sánchez, Y., & Rodríguez-Estrella, R. (2015). Man-made environments relationships with island raptors: endemics do not cope with habitat changes, the case of the island of Cuba. *Biodiversity and Conservation*, *24*, 407-425.

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

Franklin, J., Davis, F. W., Ikegami, M., Syphard, A. D., Flint, L. E., Flint, A. L., & Hannah, L. (2013). Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Global Change Biology, 19*(2), 473-483.

Gábor, L., Jetz, W., Lu, M., Rocchini, D., Cord, A., Malavasi, M., ... & Moudrý, V. (2022). Positional errors in species distribution modelling are not overcome by the coarser grains of analysis. *Methods in Ecology and Evolution*, *13*(10), 2289-2302.

Gábor, L., Jetz, W., Zarzo-Arias, A., Winner, K., Yanco, S., Pinkert, S., ... & Moudrý, V. (2023). Species distribution models affected by positional uncertainty in species occurrences can still be ecologically interpretable. *Ecography*, *2023*, e06358.

Gábor, L., Moudrý, V., Lecours, V., Malavasi, M., Barták, V., Fogl, M., ... & Václavík, T. (2020). The effect of positional error on fine scale species distribution models increases for specialist species. *Ecography*, *43*(2), 256-269.

GBIF.org (2024), *GBIF Home Page*. Available from: https://www.gbif. org [September 2021].
Gizicki, Z. S., Tamez, V., Galanopoulou, A. P., Avramidis, P., & Foufopoulos, J. (2018). Long-term effects of feral goats (Capra hircus) on Mediterranean island communities: results from whole island manipulations. *Biological Invasions*, *20*, 1537-1552.

Graham, C. H., Elith, J., Hijmans, R. J., Guisan, A., Townsend Peterson, A., Loiselle, B. A., & NCEAS Predicting Species Distributions Working Group. (2008). The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology, 45*(1), 239-247.

Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat suitability and distribution models: with applications in R*. Cambridge University Press, Cambridge.

Guntenspergen, G. (2023). Widespread retreat of coastal habitat is likely at warming levels above 1.5 °C. *Nature*, *621*(7977), 112-119.

Habel, J. C., Rasche, L., Schneider, U. A., Engler, J. O., Schmid, E., Rödder, D., ... & Stork, N. E. (2019). Final countdown for biodiversity hotspots. *Conservation Letters*, *12*(6), e12668.

Hanna, E., & Cardillo, M. (2014). Island mammal extinctions are determined by interactive effects of life history, island biogeography and mesopredator suppression. *Global Ecology and Biogeography*, *23*(4), 395-404.

Hao, T., Elith, J., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2019). A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and Distributions*, *25*(5), 839-852.

Heinänen, S., Erola, J., & von Numers, M. (2012). High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. *Landscape Ecology*, *27*(4), 545-555.

Hickisch, R., Hodgetts, T., Johnson, P. J., Sillero-Zubiri, C., Tockner, K., & Macdonald, D. W. (2019). Effects of publication bias on conservation planning. *Conservation Biology, 33*(5), 1151-1163.

Hijmans RJ, Barbosa M, Ghosh A, Mandel A (2024). *geodata: Download Geographic Data*. R package version 0.6-2, https://github.com/rspatial/geodata.

Hume, J. P. (2017). *Extinct birds.* T & AD Pyser, London. .

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific data*, *4*(1), 1-20.

Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisch, P. L., Nowicki, C., ... & Barthlott, W. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences*, *106*(23), 9322-9327.

Khosravi, R., HEMAMI, M. R., Malekian, M., FLINT, A., & FLINT, L. (2016). Maxent modeling for predicting potential distribution of goitered gazelle in central Iran: the effect of extent and grain size on performance of the model. *Turkish Journal of Zoology, 40*(4), 574-585.

Lannuzel, G., Balmot, J., Dubos, N., Thibault, M., & Fogliani, B. (2021). High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia. *Biodiversity and Conservation*, *30*(4), 963-990.

Leroy, B. (2022). Choosing presence-only species distribution models. *Journal of Biogeography*, *50*(1),247-250.

Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography, 33*(1), 103-114.

Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological conservation, 143*(11), 2647-2657.

Matthews, T. J., Wayman, J. P., Cardoso, P., Sayol, F., Hume, J. P., Ulrich, W., ... & Triantis, K. A. (2022). Threatened and extinct island endemic birds of the world: Distribution, threats and functional diversity. *Journal of Biogeography*, *49*(11), 1920-1940

Macinnis-Ng, C., Mcintosh, A. R., Monks, J. M., Waipara, N., White, R. S., Boudjelas, S., Clark, C. D., Clearwater, M. J., Curran, T. J., Dickinson, K. J., Nelson, N., Perry, G. L., Richardson, S. J., Stanley, M. C., & Peltzer, D. A. (2021). Climate-change impacts exacerbate conservation threats in island systems: New Zealand as a case study. *Frontiers in Ecology and the Environment*, *19*(4), 216-224.

McPherson, J. M., Jetz, W., & Rogers, D. J. (2006). Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions—possibilities and limitations. *Ecological Modelling*, *192*(3-4), 499-522.

Medina, F. M., Bonnaud, E., Vidal, E., Tershy, B. R., Zavaleta, E. S., Josh Donlan, C., ... & Nogales, M. (2011). A global review of the impacts of invasive cats on island endangered vertebrates. *Global Change Biology*, *17*(11), 3503-3510.

Meier, E. S., Kienast, F., Pearman, P. B., Svenning, J. C., Thuiller, W., Araújo, M. B., ... & Zimmermann, N. E. (2010). Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, *33*(6), 1038-1048.

Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global biodiversity conservation: the critical role of hotspots. In *Biodiversity hotspots: distribution and protection of conservation priority areas* (pp. 3-22). Springer, Berlin Heidelberg.
Morera-Pujol, V., Mostert, P. S., Murphy, K. J., Burkitt, T., Coad, B., McMahon, B. J., ... & Ciuti, S. (2023). Bayesian species distribution models integrate presence-only and presence–absence data to predict deer distribution and relative abundance. *Ecography*, *2023*(2), e06451.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*(6772), 853-858.

Naimi, B., Skidmore, A. K., Groen, T. A., & Hamm, N. A. (2011). Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *Journal of Biogeography, 38*(8), 1497-1509.

NASA. (2024). NASA Worldview. Retrieved June 26, 2024, from https://worldview.earthdata.nasa.gov

Parada-Díaz, J., Cubas, J., del Arco Aguilar, M. J., López, Á. F., & González-Mancebo, J. M. (2022). Changing the landscape: Short-term disruption of recruitment, height growth and flowering of endemic plants on Macaronesian heathlands by invasive alien rabbits. *Biological Conservation*, *272*, 109596.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, 12(5), 361–371.

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

Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C., & Guisan, A. (2017). Selecting predictors to maximize the transferability of species distribution models: Lessons from cross-continental plant invasions. *Global Ecology and Biogeography, 26*(3), 275-287.

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

Saintilan, N., Horton, B., Törnqvist, T. E., Ashe, E. L., Khan, N. S., Schuerch, M., Perry, C., Kopp, R. E., Garner, G. G., Murray, N., Rogers, K., Albert, S., Kelleway, J., Shaw, T. A., Woodroffe, C. D., Lovelock, C. E., Goddard, M. M., Hutley, L. B., Kovalenko, K., . . .

Radomski, T., Beamer, D., Babineau, A., Wilson, C., Pechmann, J., & Kozak, K. H. (2022). Finding what you don't know: testing SDM methods for poorly known species. *Diversity and Distributions*, *28*(9), 1769-1780.

Russell, J. C., & Kueffer, C. (2019). Island biodiversity in the Anthropocene. *Annual Review of Environment and Resources*, *44*(1), 31-60.

Samuel, M. D., Hobbelen, P. H., DeCastro, F., Ahumada, J. A., LaPointe, D. A., Atkinson, C. T., ... & Duffy, D. C. (2011). The dynamics, transmission, and population impact of avian malaria in native Hawaiian birds: a modeling approach. *Ecological Applications*, *21*(8), 2960-2973.

Soberón, J., & Nakamura, M. (2009). Niches and distributional areas: concepts, methods, and assumptions*. Proceedings of the National Academy of Sciences, 106*, 19644-19650.

Soria-Auza, R.W., Kessler, M., Bach, K., Barajas-Barbosa, P.M., Lehnert, M., Herzog, S.K. & Bohner, J. (2010) Impact of the quality of climate models for modelling species occurrences in countries with poor climatic documentation: a case study from Bolivia. *Ecological Modelling, 221*, 1221– 1229.

Stanton, J. C., Pearson, R. G., Horning, N., Ersts, P., & Reşit Akçakaya, H. (2012). Combining static and dynamic variables in species distribution models under climate change. Methods in *Ecology and Evolution, 3*(2), 349-357.

Sillero, N., & Barbosa, A. M. (2021). Common mistakes in ecological niche models. *International Journal of Geographical Information Science*, *35*(2), 213-226.

Surasinghe, T., Kariyawasam, R., Sudasinghe, H., & Karunarathna, S. (2019). Challenges in biodiversity conservation in a highly modified tropical river basin in Sri Lanka. *Water*, *12*(1), 26.

Syphard, A. D., & Franklin, J. (2009). Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. *Ecography*, *32*(6), 907-918.

Tang, Y., Winkler, J. A., Viña, A., Liu, J., Zhang, Y., Zhang, X., ... & Zhao, Z. (2018). Uncertainty of future projections of species distributions in mountainous regions. *PloS one*, *13*(1), e0189496

Taylor, S., & Kumar, L. (2016). Global climate change impacts on pacific islands terrestrial biodiversity: a review. *Tropical Conservation Science*, 9(1), 203-223

Tershy, B. R., Shen, K. W., Newton, K. M., Holmes, N. D., & Croll, D. A. (2015). The importance of islands for the protection of biological and linguistic diversity. *Bioscience, 65*(6), 592-597.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369-373.

Turvey, S. T., Kennerley, R. J., Hudson, M. A., Nuñez-Miño, J. M., & Young, R. P. (2020). Assessing congruence of opportunistic records and systematic surveys for predicting Hispaniolan mammal species distributions. *Ecology and Evolution*, 10(11), 5056-5068.

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs*, *92*(1), e01486.

Vega, G., Pertierra, L. R., & Olalla-Tárraga, M. Á. (2017). MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. *Scientific Data*, *4*(1), 1-12.

Veron, S., Mouchet, M., Govaerts, R., Haevermans, T., & Pellens, R. (2019). Vulnerability to climate change of islands worldwide and its impact on the tree of life. *Scientific Reports*, *9*(1), 1-14.

Visscher, D. (2006). GPS measurement error and resource selection functions in a fragmented landscape. *Ecography*, *29*(3), 458-464.

Walter, H. S., & Levin, G. A. (2008). Feral sheep on Socorro Island: facilitators of alien plant colonization and ecosystem decay. *Diversity and Distributions*, *14*(2), 422-431.

Whitford, A. M., Shipley, B. R., & McGuire, J. L. (2024). The influence of the number and distribution of background points in presence-background species distribution models. *Ecological Modelling*, 488, 110604.

Whittaker, R. J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: taking the long view of nature's laboratories. *Science, 357*(6354), eaam8326.

Whittaker, R. J., Fernández-Palacios, J. M., & Matthews, T. J. (2023). *Island biogeography: geo-environmental dynamics, ecology, evolution, human impact, and conservation*. Oxford University Press, Oxford.

Wood, J. R., Alcover, J. A., Blackburn, T. M., Bover, P., Duncan, R. P., Hume, J. P., ... & Wilmshurst, J. M. (2017). Island extinctions: processes, patterns, and potential for ecosystem restoration. *Environmental Conservation*, *44*(4), 348-358.

2 . Species distribution models and island biogeography: challenges and prospects

Eva Benavides[a*], Jonathan Sadler[a,b], Laura Graham[a,b,c], Thomas J. Matthews[a,b,d].

[a]GEES (School of Geography, Earth and Environmental Sciences), University of Birmingham, B15 2TT, UK.

[b]Birmingham Institute of Forest Research, University of Birmingham, Birmingham, B15 2TT, UK.

[c]Biodiversity, Ecology & Conservation Group, International Institute for Applied Systems Analysis, Austria

[d]CE3C – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group / CHANGE – Global Change and Sustainability Institute and Universidade dos Açores – Faculty of Agricultural Sciences and Environment; PT-9700-042, Angra do Heroísmo, Açores, Portugal.

* Corresponding author: ecb942@student.bham.ac.uk

Abstract

Species distribution models (SDMs) are the primary tools used to model and predict changes to species' ranges, and are often used to provide a quantitative baseline for conservation measures. However, most SDM methods and frameworks have been primarily designed for use with species with relatively large amounts of occurrence data and covering broad continental ranges.

Here, we undertake a systematic review of the literature (224 published studies) to assess the appropriate use of SDMs in island biogeography, specifically focusing on marine islands. We divide species into different insular distribution categories (i.e., chorotypes: single island/archipelago endemics, non-endemic natives, and non-natives) in order to provide chorotype-specific SDM recommendations. We highlight how to navigate three fundamental considerations related to the application of SDMs in island environments. 1) Response variables, specifically the issue of small sample sizes for many island species. 2) Predictor variables, including (i) the selection of relevant environmental predictors at appropriate spatial grains, and (ii) addressing the truncation of environmental extent across the entire species range, especially for non-endemic species. 3) Model building, particularly, in the context of limited occurrence data for many island species, how to (i) approach the uncertainty related to the choice of modelling method, and (ii) avoid overfitting.

We also examine the data sources used in island SDM studies, finding that there are strong geographical biases in study location. Alongside this, we evaluate the potential of the GBIF database – a comprehensive global database of species occurrences – in island SDM research. We find that GBIF has been potentially underutilised in island SDM studies so far, and represents a useful resource for filling island distribution data gaps for several taxa going forward.

Based on the insights we obtained from our systematic review, we propose a set of SDM recommendations tailored to insular species and environments.

2.1 Introduction

As anthropogenic global environmental change progresses, suitable habitats for many species are shrinking, becoming fragmented, or being lost entirely (Kerr et al., 2007; Thurman et al., 2020), resulting in the loss of biodiversity at multiple spatial scales (Cardinale et al., 2012). Nowhere is this loss more apparent than on islands, systems that have suffered large numbers of recorded extinctions (Whittaker et al., 2017, 2023; Russell & Kueffer, 2019; Matthews & Triantis, 2021). Species persistence often depends on keeping pace with environmental change, primarily through dispersal and range-shifts (Bulgarella et al., 2014). As such, self-contained environments such as islands, which provide limited opportunities for species (especially endemics) to shift their range, are particularly vulnerable. Therefore, accurately mapping island species distributions is crucial for anticipating the impact of environmental change on insular environments (Porfirio et al., 2016; Ellis-Soto et al., 2017; Matthews & Triantis, 2021).

Ecological niche models, also termed species distribution models (hereafter SDMs), have become a widely used methodological tool for providing this information (Elith & Leathwick, 2009). SDMs use species occurrence data in combination with corresponding environmental covariates to identify potentially suitable sites for species in both native and novel ranges and across different temporal scales (Franklin, 2013; Guisan et al., 2013; Robinson et al., 2017). The flexibility of SDM approaches allows for a wide range of applications, including assessments of the ecological impacts of climate and land-use changes, and biological invasions (Parmesan & Yohe, 2003; Guisan & Thuiller, 2005; Faleiro et al., 2013): all important drivers of extinction on islands (Veitch & Clout, 2002; Ricketts et al., 2005; Russell & Kueffer, 2019; Whittaker et al., 2023).

Even though conservation efforts focused on island biodiversity may greatly benefit from the ecological information generated by SDMs, research to-date on SDMs has tended to focus on species with broad ranges and for which there is a large amount of occurrence information, corresponding typically to continental species (Hickisch et al., 2019, Leroy, 2022). Historically, there have been fewer applications of SDMs in island environments and for island species, although recently this situation has started to change (e.g., Price et al., 2012; Vergilio et al., 2016; Spiers et al., 2019; Goedecke et al., 2020; Barlow et al., 2021). The increasing implementation of SDMs in island studies provides an opportunity for evaluating how two key considerations for developing robust models for islands species have been handled: data limitations (Pearson et al., 2007; Lannuzel et al., 2021) and the restricted spatial context of island environments (Kier et al., 2009; Whittaker et al., 2023).

Data limitations in the study of island species arise i) due to the inherent small ranges of many insular species, particularly single island endemics (Fernández-

Mazuecos et al., 2014; Chiatante, 2022); ii) as a consequence of numerous anthropogenic disturbances that have severely reduced already small island populations (Helmstetter et al., 2021; Matthews et al., 2022; Nogué et al., 2021); and/or iii) simply because many islands are remote and inaccessible, hampering detailed biological surveys (i.e., a lack of extensive sample data; Ando, 2019). Regardless of the origin (incomplete sampling or naturally restricted distributions), these distribution data limitations increase model uncertainties in SDM applications, a particularly problematic issue given that many SDM modelling techniques are known to be sensitive to small sample sizes (Stockwell & Peterson, 2002; Hernadez et al., 2006; Wisz et al., 2008) and the issue of class imbalance (i.e., a disproportionate number of absences/pseudoabsences in the data; Longadge et al., 2013; Robinson et al., 2017). In addition, small sample sizes may lead to statistical overfitting, especially when used in combination with a disproportionately large number of predictors in relation to sample size (Breiner et al., 2015), compromising the transferability of model predictions in space and time, a key aspect in many SDM studies (Heinänen, et al., 2012).

Additional important considerations regarding the application of SDMs in an island context are that study areas i) are spatially distinct units surrounded by a non-permeable matrix for many organisms (i.e., water; Matthews, 2021), and ii) often contain high environmental heterogeneity across a relatively small area (Barajas-

Barbosa et al., 2020). Consequently, the ecological processes that underpin island species distributions typically operate at relatively smaller scales, for which finer scale habitat descriptors such as topoclimatic, land cover and terrain variables are needed to increase the accuracy of predictions (Meyer & Thuiller, 2006; Lanuzel et al., 2021). However, SDM studies frequently rely purely on broad-scale climatic variables. These macroclimatic variables are unlikely to be able to explain the fine-scale variation in island species distributions, particularly for small-ranged species such as insular endemics (Heinänen et al., 2012; Turvey et al., 2020; Segal et al., 2021). The increasing availability of fine-grained predictors such as remotely sensed environmental data, landcover, and topographically derived variables, particularly for certain well studied archipelagos, present a means to increase the utility of SDMs for island species (Longcore et al., 2018; Hansen et al., 2019; Lannuzel et al., 2021; Hanz et al., 2023).

Another potential issue related to studying relatively small and bounded areas is the exacerbated risk of environmental space truncation, whereby studies only include occurrence and environmental data from the insular space in the models, even though the target species may be also found outside of the focal islands (i.e., non-endemic native and non-native species), increasing the uncertainty of predictions relating to species’ responses to drivers of change (Bush et al., 2018; Rosenblad et al., 2019).

In the context of the issues discussed above, in this study we review the implementation of SDMs on islands, with a particular focus on analysing data limitations, with the aim of identifying key methodological considerations to deal with data constraints and the distinctive features of island environments/species. Given the data limitations in island SDM research, we explore whether the Global Biodiversity Information Facility database (GBIF, 2022) – a primary data source in SDM studies more broadly but less utilised in island studies – provides adequate data at the island level to address knowledge gaps and minimise geographical and taxonomical biases in island SDM studies.

2.2 Materials and Methods

2.2.1 Literature search

To obtain a general overview of SDM implementation on island systems, we conducted a comprehensive literature search for studies published over the last two decades. The bibliographic search was performed using the Scopus (https://www.scopus.com) and Google Scholar (https://scholar.google.com) databases. Papers published between 1999 (first SDM publication recorded fulfilling the selection criteria) and the cut-off date 30 September 2021 with the terms [”SDM” OR “Species Distribution Model*”] AND [“island”] in the title, keywords or throughout the paper were included. To be included in our review, a paper needed to focus exclusively on true marine islands (i.e., islands of land in the sea; Matthews, 2021). We considered both methodological and empirical studies if SDMs were employed as part of the methodological framework. Studies focused on microorganisms (e.g., parasitic/pathogenic fungi, protozoa, etc.) were excluded from the analysis as they tend to involve different types of predictor variables compared to analyses of macroorganisms. We systematically assessed studies for relevant aspects of SDM fitting, including sample size, data sources, study bias, model building, model overfitting, and environmental space truncation. Additionally, we evaluated the GBIF database – including the spatial distribution of island data in GBIF for different taxa – to assess its suitability for filling knowledge gaps in island SDM research. Geographic data were mapped using QGIS (version 3.22.10; QGIS Development Team, 2022).

2.2.2 Response variable related considerations in island SDMs

Sample size

A frequently considered minimum threshold for SDM analyses is ≥20 records per species prediction (Santini et al., 2021; van Proosdij et al., 2015; Wisz et al., 2008), although this number is also a function of the number of predictors (discussed below). Based on this threshold, we categorised the sample size of the modelled species in the sourced papers as follows: species datasets with ≤5 records were defined as 'very small', datasets with [6-20] records were defined as 'small', those with [21-100] were defined as 'moderate', and those with >100 records were classified as ‘large’. In addition, to analyse data availability per species distribution range type (‘chorotype’), we categorised the modelled island species as endemic (endemic to the specific island / archipelago being studied), non-endemic native (i.e., naturally present on an archipelago but also the mainland), or non-native (i.e., introduced to an archipelago by humans). This categorisation was undertaken based on information provided in the reviewed papers.

2.2.3 Predictor variable related considerations in island SDMs

Study grain

To evaluate if the choice of study grain was related to the (1) analysed predictor variable type, (2) study aim, or (3) taxonomic group, we recorded the study grain of the predictors in the sourced studies and assessed its association with the previously mentioned factors.

We identified 5 types of predictor variable used in SDM analyses: 1) bioclimatic variables from global datasets (Globclim); 2) bioclimatic variables from regional climate models (Regclim); 3) topography-related variables (DEM); 4) land cover variables (LC; including soils/geology and vegetation cover); and 5) biotic interaction information (Biotic). For studies with multiple variable types, all corresponding categories were assigned.

We assigned to each analysed study one of the 10 categories of study aim identified: 1) method testing = MT; 2) species geographical distribution = SGR; 3) evaluation of richness patterns = RP; 4) protected area management = PAM; 5) applied restoration efforts = RST; 6) phylogeography = PHY; and assessing the impact of anthropogenic drivers of environmental change, including, 7) climate change = CC, 8) invasive non-native species = IAS, 9) land use change = LUC; and 10) the synergistic interactions between any of the previous three = SYN.

For each study, we classified the modelled taxon as one of the following: 1) anurans, 2) arthropods, 3) birds, 4) fungi, 5) mammals, 6) non-arthropod invertebrates (i.e., snails and land worms), 7) reptiles 8), vascular plants and 9) non-vascular plants. Studies analysing more than one taxon were assigned multiple taxonomic categories accordingly.

Environmental space truncation

To identify potential cases of environmental space truncation (i.e., where an island species distribution model did not utilise mainland occurrence data, where available), we analysed the geographic extent of the occurrence points used to fit the SDM by species' chorotype (i.e., endemic, non-endemic native, or non-native). That is, we distinguished between studies where (i) the occurrence data points were obtained exclusively from the island study area, which were categorised as localised, and hence no strategy was implemented to handle environmental space truncation; and ii) additional occurrence points to the island study area were obtained, which were categorised as widespread (e.g., a non-endemic native to an island modelled with data from both the insular and non-insular ranges). This categorisation was undertaken solely using information provided in the reviewed papers (i.e., we did not independently assess species distributions to determine the geographic extent of the occurrence points). For certain studies, particularly those that analysed large numbers of species and islands, this information was not always clear, and in these cases, we made an informed judgement as to the correct category.

2.2.4 Model building considerations in island SDM studies

To assess the implications of model choice in relation to sample size, we noted, for each paper where data existed, sample size, the type of occurrence data used (presence, absences or pseudo-absences), the modelling method employed (e.g., Generalised linear models, Maxent, etc.), and the type of approach selected for modelling (single vs ensemble).

Modelling technique uncertainty

To assess the implications of model choice in relation to sample size, we noted, for each paper where data existed, sample size, the type of occurrence data used (presence, absences or pseudo-absences), the modelling method employed (e.g.,

Generalised linear models, Maxent, etc.), and the type of approach selected for modelling (single vs ensemble).

Model complexity: overfitting

We used the general rule of thumb regarding SDM overfitting: the number of presences should be 10 times larger than the number of predictors used for modelling (Franklin, 2013; Vaughan & Ormerod, 2003). For each study and model fit, we noted all studies that were potentially affected by overfitting, by calculating the ratio of the number of records per modelled species to the number of predictors. For cases which were potentially affected by overfitting (i.e., those where the ratio was < 10), we recorded whether or not any approaches had been applied to mitigate this issue. Such approaches can include varying the level of regularization, which controls model complexity (Radosavljevic &Anderson 2013), modification of beta multipliers (in the context of maxent; Elith et al., 2010), or post-modelling variable removal based on their contribution (e.g., Jacknife-partition; Girini et al., 2017).

2.2.5 Assessing the potential of GBIF for filling island SDM knowledge gaps

Data Sources and Geographic Coverage

To identify the geographical coverage of island distribution data, and thus identify potential knowledge gaps, for each study, we categorised their corresponding occurrence data sources as: (1) citizen science, (2) directly from GBIF, (3) herbaria,

(4) museum, (5) previously published literature, (6) primary field data collection (i.e., field data collected as part of a study), (7) regional occurrence databases (e.g., ATLANTIS from the Azores), (8) remote sensing (i.e., satellite and aerial imagery data), and (9) specialised datasets (e.g., focusing on a specific group of species, such as economically important species, non-natives, etc.).

To analyse the geographic coverage of studies, study regions were categorised based on the bioregionalization of coastal and shelf areas by Spalding et al. (2007) that comprises: (1) Arctic, (2) Temperate Northern Atlantic, (3) Temperate Northern Pacific, (4) Tropical Atlantic, (5) Western Indo-Pacific, (6) Central Indo-Pacific, (7) Eastern Indo-Pacific, (8) Tropical Eastern-Pacific, (9) Temperate South America, (10) Temperate Southern Africa, (11) Temperate Australasia, and (12) Southern Ocean.

Assessing the extent to which GBIF data can fill island SDM data gaps

We used the "Rgbif" R package (Chamberlain et al., 2022) to extract species occurrences from the GBIF database within global island polygons (sourced from: Geodata R package; Hijmans et al., 2023). The assessment included islands ranging in size from 0.1 to 786,000 km^2 (threshold reference: Weigelt et al., 2013; Fernandez-Palacios et al., 2021). We selected all taxa at the species level under the "accepted name" category and assigned each species to one of the nine taxonomic categories listed above. Subsequently, we quantified the number of records per species within each island polygon. We preferred the use of polygons over counting points within a grid to avoid making assumptions about the optimal spatial resolution to use.

To evaluate the appropriateness of GBIF data for fitting island SDMs, we developed two metrics. First, the "island modellability index" (M index) was calculated as the ratio between the number of species with at least 20 records (i.e., occurrence points) of a given taxon and the total number of species recorded in that taxonomic group within an island. The M index ranges from 0 (all recorded species of a taxon on an island have <20 records) to 1 (all recorded species of a taxon on an island have ≥ 20 records). This index provides a measure of data availability for a taxon as a whole, regardless of taxon richness.

The second metric is the absolute number of species (within a given taxon) with at least 20 records per island; the number of species that meet the minimum sample size threshold commonly reported in the SDM literature (Santini et al., 2021; Wisz et al., 2008).

2.3 Results

2.3.1 Literature search

Through the literature search, we sourced 224 island SDM studies that passed our search criteria. Details of the sourced studies are provided in Appendix S2.1.

2.3.2 Response variable related considerations in island SDMs

Sample size

Across the 224 reviewed studies, we obtained 3043 species-occurrence datasets, corresponding to 2787 species (endemics=1327; non-endemic natives=1480, and non-natives=236; note these numbers do not sum to 2787 as some species were included in more than one dataset). Of these, roughly half of the occurrence datasets were classified as 'very small' (<5 records: n=650/3088; 21%) or 'small' (from 6 to 20 records: n=964/3088; 31%), with relatively fewer classified as 'moderate' (21-100 records: n=887/3088; 29%), or large (>100 records: n=587/3088; 19%). Most of the 'very small' to 'small' occurrence datasets (i.e., occurrence data for a single species) corresponded to endemic species (n=934/1614; 58%), followed by non-endemic native species (n=666/1614; 41%) and non-native species (n=14/1614; 1%). Conversely, 'large' datasets rarely related to endemic species (n=126/1327; 9%) and were most common for non-native species (n=176/236; 75%).

2.3.3 Predictor variable considerations in island SDM studies

Study Grain

Predictor Type – Across the sourced studies, the grain size of the predictor variables used in SDMs varied from 1.5 m to 10 km. The majority of studies that reported grain size data (n=101/216; 47%) were developed at a 1 km grid scale (Fig. 2.1a). This preference can be attributed to the predominant use of 'Globclim' predictors, particularly the WorldClim global dataset, the finest resolution of which is 1 km (Fig. 2.1a -b), accounting for 92% of studies using 'Globclim' predictors. Finer scale resolutions (i.e., between 1.5 m and 500 m) were utilised in 30% of studies (n=65/216), with DEM variables being most commonly employed as predictors (n=40/65; 61%; Fig. 2.1a). These DEM variables were generally combined with fine-grained 'Regclim' predictors (n=34/40; Fig. 2.1a-b), rather than 'Globclim' (downscaled) variables, which were rarely used at finer resolutions (n=10/65; 15%; Fig. 2.1a-b). Landcover predictors (n=29/65; 45%) were also frequently used at resolutions below 500 m (Fig. 2.1b). Studies that used a combination of coarse climatic and fine non-climatic predictors resampled the predictors to a 1 km grain size (Fig. 2.1b).

Coarser predictor grain sizes of 2-5 km were less common, with only 36 studies (n=36/216; 17%) reporting their use. Among these studies, 'Globclim' predictors alone were predominantly utilised (n=27/36; 75%; Fig. 2.1a). The coarsest resolution of 10 km was employed in 13 studies (n=13/216; 5%), with 'Globclim' predictors commonly used (n=10/13; 78%), often in combination with non-climatic predictors (n=8/10). For further details, please refer to Appendix S2.2.

Study Aim and Grain Size – We found that most island SDM studies were conducted to assess climate change impacts (n = 41/224; 18%), followed by species geographical distributions (n = 35/224; 16%), phylogeography (n=35/224; 16%), invasive species (n=28/224; 13%) and the synergistic effects of environmental drivers (n=27/224; 12%). Nineteen papers (8%) evaluated methodological considerations of SDM applications on islands, with few studies in the remaining categories: richness patterns (n=13/224; 6%), protected area management (n=9/224; 4%), restoration (n=8/224; 3.5%), and assessing the impact of land use change (n=8/224; 3.5%) (Fig 2.1c). For those studies with study grain data available, we identified that finer predictor grain sizes (<500m) were generally used in island SDM studies focused on land use change (n=6/7; 85%), restoration (n=5/6; 83%) and method testing (n=14/17; 82%) (Fig. 2.1d). In contrast, studies that were more commonly associated with predictors measured at the 1km grain size were those focused on phylogeography (n=32/35; 91%), the synergistic effects of global environmental change drivers, (n=21/25; 84%), protected area management (n=7/9; 78%), invasive species (n=19/27; 70%), climate change impacts (n=28/41;

68%), and species geographical distributions (n=14/35; 60%) (Fig 2.1d). Further details are provided in Appendix S2.2

Taxa and Grain Size – Vascular plants were the most studied taxon in island SDM analyses (n=116/224; 52%), followed by mammals (n=31/224; 14%), and reptiles, arthropods and birds (n=26/224; each representing 12%). Anurans (n=8/224; 4%), non-vascular plants (n=7/224; 3%), non-arthropod invertebrates (n=5/224; 2%) and fungi (n=2/224; 0.9%) were all studied to a lesser degree (Fig. 2.1e). We could not identify clear patterns regarding predictor variable resolution selection by taxa (Fig. 2.1f). Further details are provided in Appendix S2.2

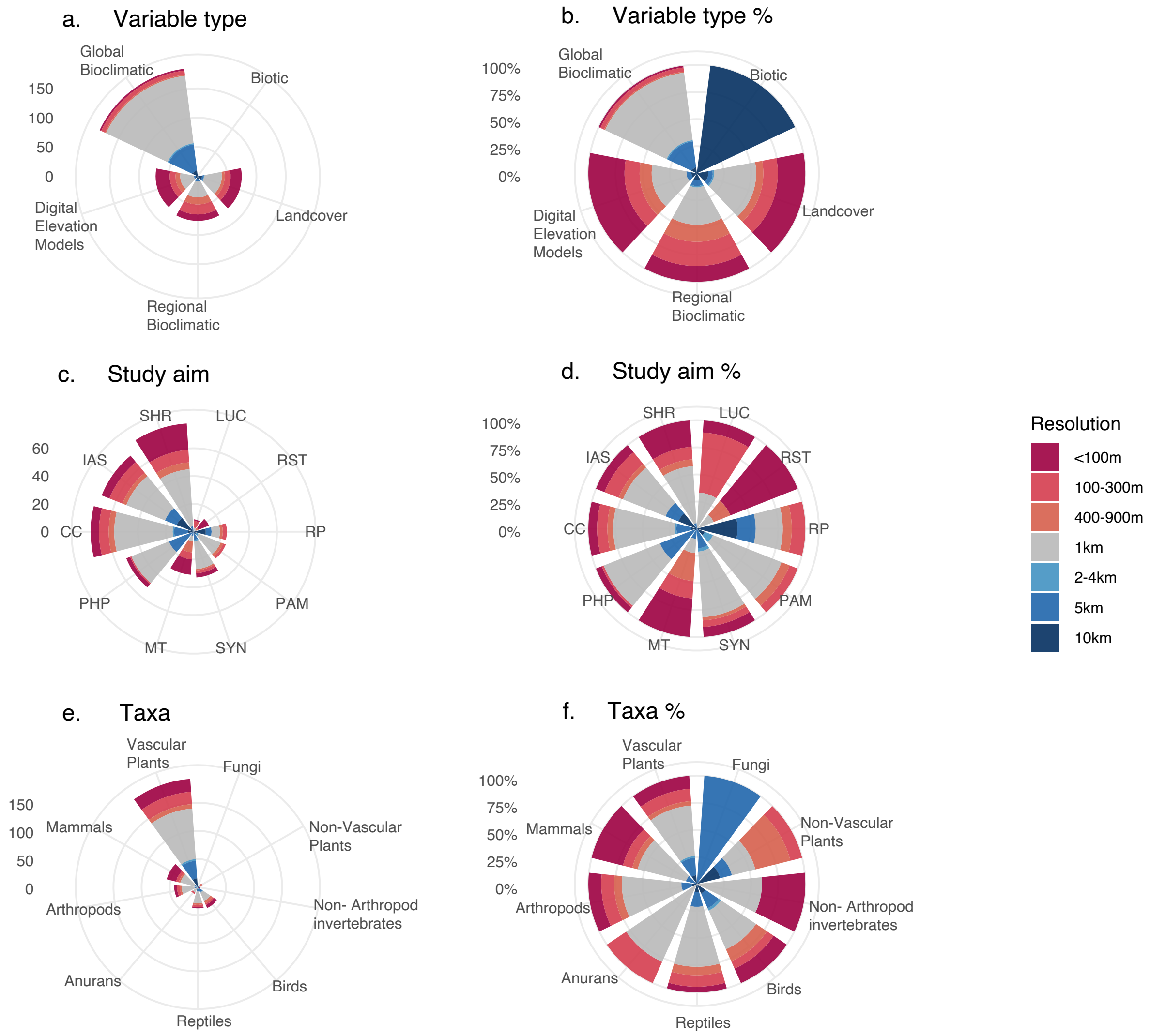


Figure 2. 1 Predictor variable resolution within published island SDM studies. The left column shows the absolute number of studies by each category analysed at each recorded resolution, while the right column shows the percentage of studies. Study aims cover climate change (CC), invasive alien species (IAS), land use change (LUC), synergy of extinction drivers (SYN), method testing (MT), protected areas management (PAM), phylogeographic (PHY), richness patterns (RP), restoration (RST), and species geographical distributions (SHR). Finer resolutions (i.e., predictor variable grain size) are represented in shades of red, while coarser resolutions are depicted in shades of blue.

Environmental space truncation

Regarding the area of a species' distribution for which occurrence points were obtained, among the models of the 1480 non-endemic native species models, 90% exclusively utilized data from the insular part of the range (i.e., localised data; n=1327/1480), with the remaining 10% using widespread occurrence and environmental data both from the insular and the non-insular parts of the distribution (i.e., here insular and non-insular refers to the focal island(s) being studied, the 'non-insular' parts of the distribution could also be on other islands that were not studied, for example, in a different archipelago). A considerable proportion of the 236 non-native species models were fitted only using data from the insular environmental space (n=101/236; 43%); one study relied on complementary physiological tolerance experiments to contrast the results of SDMs fitted with data from the insular environmental space.

While endemic species are not expected to have occurrence data from areas outside of the insular range, we found one study using additional information from the environmental space of the introduced range of natural island endemics (i.e., island endemics prior to anthropogenic introductions).

2.3.4 Model building considerations in island SDM studies

Modelling technique uncertainty

Both sample size and the nature of the input data (Presence only / Presence–absence) influenced algorithm selection in the sourced island SDM studies. Overall, a low proportion of studies included information on true species absences (n=36/224; 16%), and the use of true absence data increased with the number of occurrence points analysed (supplementary S2). The use of a single-modelling technique was the most common approach selected for modelling (n = 176/224; 79%). Of which, a marked preference for Maxent as a stand-alone method was observed (n=135/176; 77%). As datasets became smaller, Maxent was more frequently used as the sole analytical method. It was used in 83% (n=539/650) of model fits based on 'very small' datasets, but only 45% (n= 264/587) of those based on 'large' datasets. The use of multiple modelling techniques (model ensembles) was less common than single model applications, being used in 21% of papers (n=48/224). The methods most frequently used in ensembles were Generalised Linear Models (GLM; n=38/48; 79%), Random Forest (RF; n= 37/48; 77%), Boosted Regression Trees (BRT; n=31/48; 64%) and Maxent (n=30/48; 62%). It is noteworthy that all studies with 'very small' to 'moderate' sample sizes included in their ensembles at least one data hungry modelling method (i.e., those that perform better with >200 samples; van der Ploeg et al., 2014) such as RF, Artificial Neural Networks (ANN), multivariate adaptive regression splines (MARS) and Support Vector Machines (SVM). More detail on all the algorithms used in each study is provided in Appendix S2.2.

Model complexity: overfitting

We found 161 (n=161/224; 72%) papers with sufficient information for an assessment of model overfitting (number occurrence points and predictors utilised). For modelled species with low numbers of observations, we observed a general and concerning pattern: the smaller the sample sizes used for modelling species distributions, the less attention was seemingly paid to overfitting. For example, none of the four studies that used the smallest number of occurrence records for modelling (n=3) reported the use of a method for addressing overfitting, and all used a relatively large number of predictors (8– 24). 53% (n=47/89) of the studies that analysed between 4-100 occurrence records per modelled species did not report the use of any form of overfitting reduction technique, and used two to three times more predictors than the recommended ratio; only 24% (n=21/89) used the recommended ratio of predictors in relation to the sample size. The most common method for reducing overfitting was the reduction of the number of variables through Jack-knife partitioning that allowed for the identification of variables with higher contribution to the models (n = 17/89, 19%). The other two methods applied in the reviewed literature were ensembles of small models (n = 2/89; 2%) and modification of beta/regularization multipliers (in the context of Maxent) during model calibration (n = 2/89; 2%). For all the modelled species with >100 records (n=69), the number of variables was appropriate in relation to the number of observations. Further information relating to model overfitting is provided in Appendix S2.2

2.3.5 Assessing the potential of GBIF for filling island SDM knowledge gaps

Data sources and geographical coverage

By analysing the origin of data used for fitting SDMs on islands, we identified that a large proportion of the studies were developed with data newly collected from the field (n=102/224; 46%), both when used as a single source of data (n=53/102; 52%) and combined with other datasets (n=49/102; 48%). This was followed by information sourced from literature (n=68/224; 30.3%), regional databases (n=61/224; 27.2%), and the GBIF dataset (n=48/224; 21.4%). Other sources included herbaria (n=26/224; 11.6%) and museum collections (n=17/224; 7.6%), remote sensing imagery (n=16/224; 7.1%), specialised databases (n=12/224; 5.3%) and citizen science (n=8/224; 3.6%).

We identified geographical bias in published SDM studies using data from the field, literature and regional datasets (Fig. 2.2a). SDM studies using these data sources are concentrated in temperate islands regions in the Northern Atlantic, Northern

Pacific and Australasia, with fewer examples from islands with tropical affinities from the Central and Western Indo Pacific (e.g., Madagascar, New Guinea, and Sri Lanka). (Fig. 2.2a). The remaining tropical island regions were poorly represented (Fig. 2.2a). GBIF, which ranked as the fourth most utilised data source, is a more commonly used data source in certain tropical islands (e.g., New Guinea, Philippines, several islands from Eastern Indo Pacific and the tropical Atlantic; Fig.2.2b).

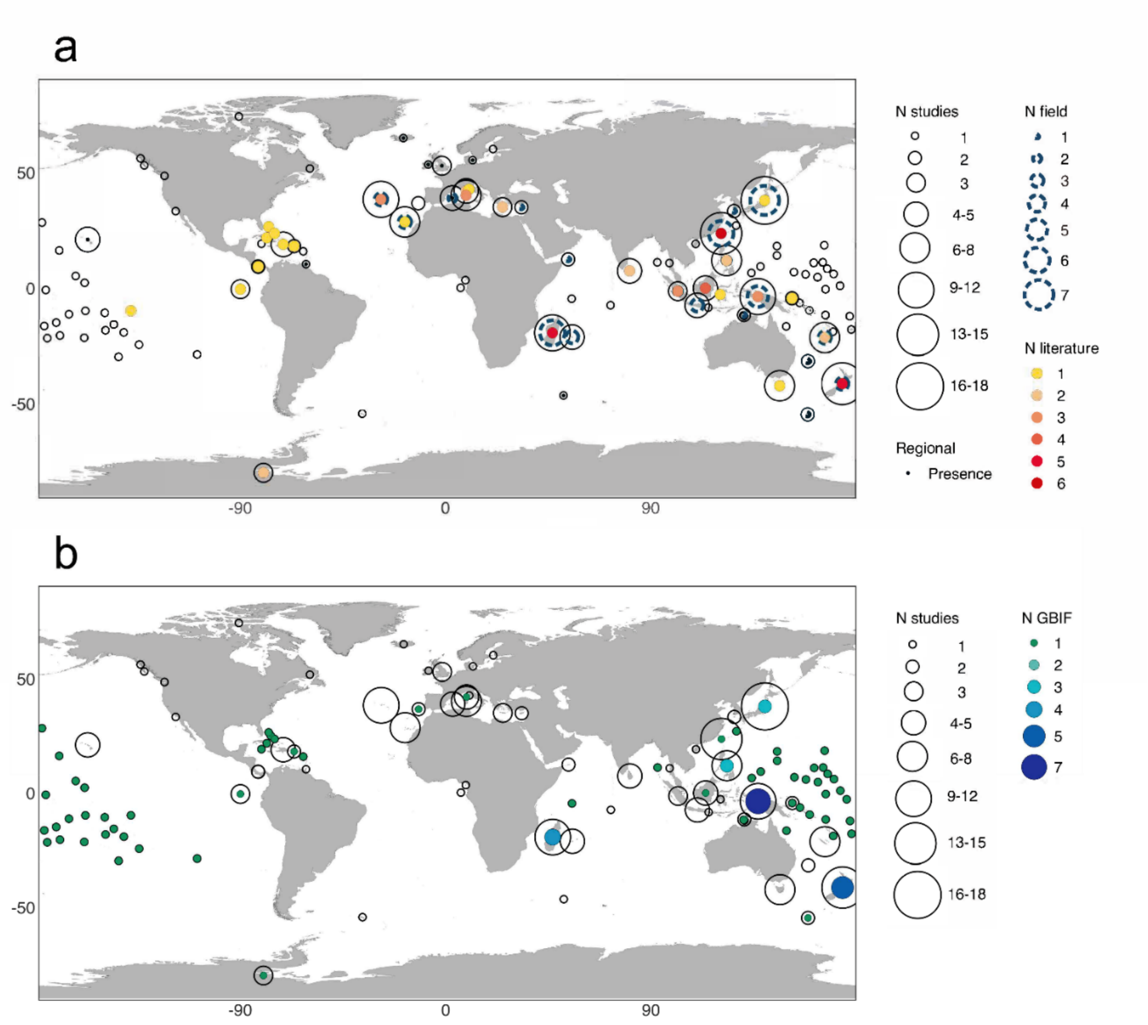


Figure 2. 2 Spatial distribution of the data sources of island SDM studies. In (a) dotted rings (N field) indicate the number of studies using primary fieldwork data, and solid rings show the total number of studies. Black dots indicate the use of regional databases with species occurrences; coloured dots indicate the number of studies using published data from the literature. In (b) coloured dots indicate the number of studies analysing GBIF data. Islands points in the eastern Pacific reporting one study with GBIF belong to a single multi-island assessment. Note that a number of studies did not identify the specific island(s) studied (simply the insular region) – in these cases we used the centroids of the study area.

Assessing the extent to which GBIF data can fill island SDM data gaps

The final analysed GBIF island occurrence database comprised 12,557,748 occurrences across multiple species (n=106,858) from 4060 islands. In this section, for ease we focus on islands falling within the top 10th percentile (M > 0.09) of the M Index (i.e., for a given taxon, the number of species with ≥20 records divided by the total number of species recorded on the island), and within the top 10th percentile (> 5 species) in relation to the number of species with at least 20 occurrence records per island. However, the full range of these metrics is displayed in Fig. 2.3. Table 1 summarises the results of our two metrics, which were calculated at the island level and for each taxon individually: the M index and the absolute number of species with ≥20 records.

Overall, our analysis revealed that birds were the best represented taxon in the GBIF island data, as indicated by both of our metrics (Table 2.1, Fig. 2.3e-f). Island anurans and mammals also showed significant availability of data per species, with numerous islands across all regions showing M index values > 0.09 (Table 1, Fig. 2.3a-b & g-h). Reptiles exhibited islands with M index values >0.09 primarily in tropical regions, while representation in temperate regions was relatively low, which is to be expected given the ecology of this taxon (Fig. 2.3o-p).

On the other hand, insular arthropods, fungi, non-arthropod invertebrates, non-vascular plants, and vascular plants were poorly represented in the GBIF dataset at a global scale (Table 2.1 Fig. 2.3). However, there were numerous islands for non-vascular and vascular plants with representative occurrence data (≥20 records per species) concentrated in understudied temperate regions (Fig. 2.3k-l & q-r).

It is worth noting that the number of species with ≥20 records per island was a particularly informative metric for birds. For vascular plants and arthropods, which ranked second and third for this metric, the high number of species with relatively large occurrence record datasets can be attributed to their overall higher species richness. A higher number of species per island, for a given taxon, increases the likelihood of there being a larger number of species with numerous occurrence records, all else being equal. For more detailed information on the results related to this metric, see Appendix S2.2.

Table 2.1 Summary of M Index Values [number of spp. with ≥20 records / total number of spp. recorded at the island level] and Total Number of spp. with ≥20 records across taxonomic groups in GBIF Island Data.

Taxonomic Group	Mean number of records per spp. in GBIF island data	Total number of islands with GBIF Data	Number of islands with M Index >0.09*	Percentage of islands with M Index >0.09*	Number of islands with ≥5 spp. with at least 20 records	Percentage of islands with ≥5 spp. with at least 20 records
Birds	33	2452	639	**26%**	677	**27.6%**
Anurans	9	929	138	**14.8%**	20	2.1%
Mammals	15	1571	249	**15.8%**	50	3.2%
Reptiles	9	549	57	10.4%	5	0.9%
Arthropods	8	1817	20	1.1%	138	**7.5%**
Non-arthropod invertebrates	4	1136	33	3%	14	1.2%
Non-vascular plants	7	1069	33	3%	43	4%
Vascular plants	5	2275	132	5.8%	266	**11.7%**
Fungi	4	1171	75	6.4%	35	0.5%

The three highest-ranked percentages for each metric are marked in **bold.** *10th percentile value of the M index

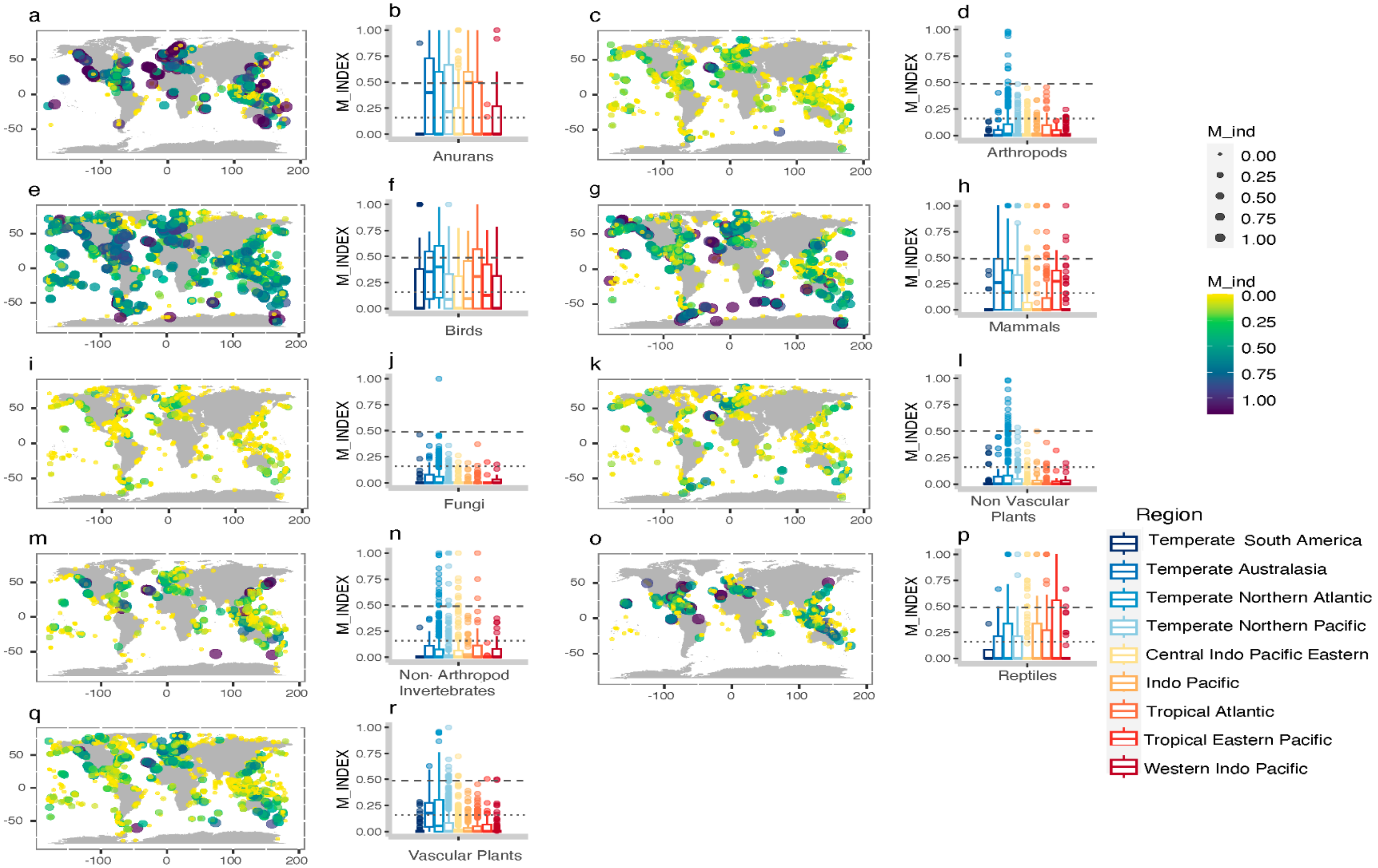


Figure 2. 3 Spatial distribution of the island modellability Index (M index, as derived from GBIF data) divided by taxa: Anurans (a-b), Arthropods (c-d), Birds (e-f), Mammals (g-h), Fungi (i-j), Non-Vascular Plants (k-l), Non-Arthropod Invertebrates (m-n), Reptiles (o-p), and Vascular Plants (q-r). Blue shades in boxplots indicate temperate environments, yellow to red shades indicate tropical environments. In the boxplots, the dashed line represents the 25th percentile of the M index, and the dotted lines represent the 10th percentile.

2.4 Discussion

Species distribution models (SDMs) can serve as a valuable methodological tool, offering insights into the distribution patterns of island species. This information is crucial for designing effective conservation strategies for this unique yet highly threatened segment of global biodiversity (Kier et al., 2009; Whittaker et al., 2017). Despite the distinct features of island environments and species, and the associated considerations for modelling the distributions of the latter, there has been a notable absence of reviews of island SDMs, particularly those focused on identifying potential sources of uncertainty and examples of best practice.

In light of this gap, and drawing from the results of our literature review, in this section we first present a set of recommendations that future island SDM studies should consider. We conclude by discussing future prospects for island SDM studies, emphasising the potential of the Global Biodiversity Information Facility (GBIF) database in advancing island SDM research.

2.4.1 The selection of relevant predictors at an appropriate analytical scale maximises the utility of limited occurrence data

It has been consistently suggested in the SDM literature that an occurrence dataset should have a minimum of 20 records per species, ideally exceeding 50, for robust model predictions (Santini et al., 2021; Wisz et al., 2008). However, our findings indicate that more than half of the analysed island SDM datasets fall below this recommended threshold (i.e., n < 20). Small sample sizes in island analyses may result from incomplete sampling due to the inherent inaccessibility of islands, or the presence of species occurring in low numbers, including endemics with naturally narrow ranges and island species whose populations have been reduced by human activities (Fernandez-Palacios et al., 2021; Helmstetter et al., 2021; Matthews et al., 2022; Nogué et al., 2021). This limitation is arguably the most significant factor hindering the proper implementation of island SDMs.

Compounding this issue, the use of coarse-grained predictors may further diminish sample sizes, especially for species with narrow and clustered distribution patterns. For example, at the extreme, if very coarse-grained predictors are used to model a species with a very narrow and clumped distribution, the species may only occur in a single grid cell. In cases such as this, the predictors will be unlikely to adequately capture the environmental variation at a relevant scale for meaningful SDMs. In addition, this results in the loss of occurrence information, data that are often already scarce.

Our review identifies a prevalent overreliance on coarse climatic variables in island SDMs, possibly influenced by the current SDM guidelines, which emphasise the importance of climate as a strong predictor of species distributions at continental scales (Pearson & Dawson, 2003; Gardner et al., 2019). However, islands typically exhibit distinct environmental conditions, such as limited land area, high isolation, and significant topographical variation over relatively small areas, especially for volcanic oceanic islands (Barajas-Barbosa et al., 2020). As a result, island species, particularly from certain taxa, often have highly specific habitat requirements (Whittaker et al., 2017, 2023). For example, evidence suggests that the interaction between island plants and their environment primarily occurs at fine scales (Bellamy et al., 2013; Fernandez-Palacios et al., 2021; Whittaker & Fernandez-Palacios., 2007). Therefore, it is crucial that researchers select predictors (and predictor variable grain sizes) that reflect these highly specific microenvironmental variations to accurately represent island species distributions. This ensures that occurrence points within the ranges of highly specialised or restricted species cover a broad a range of environmental conditions as possible. Consequently, it enhances the environmental data available for constructing SDMs and potentially optimises the utilisation of (often limited) occurrence information.

2.4.2 Island SDM analyses should avoid the default use of coarse-scale climatic predictor variables

Recent advancements in downscaling climatic variables (Khosravi et al., 2016) offer a potential solution for incorporating high-resolution environmental descriptors into island SDMs. However, these methods are not widely adopted in this field, likely due to their demand for precise remotely sensed and topographic data, along with regional climatic models, to make reliable predictions (e.g., the CIELO model for the Azores archipelago: Azevedo, 1996; Baker et al., 2017). While high-resolution remotely sensed and topographic data are increasingly accessible (e.g., ASTER DEM, Abrams et al., 2020), high-quality open-access regional climate data remain unavailable for most island systems. Consequently, island studies must choose between using either a) climatic proxies derived from topographic, land cover, or remotely sensed information (e.g., Lannuzel et al., 2022); or b) downscaled global climate data, even though the latter approach comes with higher uncertainty (Bazzicheto et al., 2021; Hanz et al., 2023). Both of these approaches hold promise for islands lacking regional climate datasets, but additional research is required to assess the associated uncertainty and relevance of downscaled variables for studying species distributions in island environments, given the novelty of this approach.

Acknowledging the challenges associated with obtaining accurate environmental datasets describing microhabitats (Fitzpatrick & Ellison, 2018), future island SDM studies should refrain from the default use of coarse-grained climatic predictors. Instead, a comprehensive search for finer-resolution variables tailored to the study area should be undertaken. Importantly, the use of high-resolution predictor variables will require corresponding occurrence data at a closely aligned resolution (Moudrý & Šímová, 2012).

The use of coarse-scale occurrence data can also influence SDMs by linking species with unsuitable environmental conditions. In island environments, two key factors can significantly worsen SDM performance due to inaccuracies resulting from occurrence points being sampled outside of a species' preferred environments (Gabor et al., 2023): 1) oceanic islands frequently exhibit diverse and highly variable terrain within small areas, resulting in sharp variations in environmental conditions over short distances (Hanz et al., 2023); and 2) many species inhabiting islands, particularly in certain taxa (e.g., plants), have very narrow niches (Rosenblad et al., 2019).

As a final consideration, once the set of potential variables is obtained, and is then paired with a sufficient sample size of species occurrence data, the final selection of predictor variables should be informed by existing knowledge or at least theoretical justification (Araujo et al., 2019). By considering factors like species ecology and the specific scientific question, researchers can make informed decisions about the most suitable predictors and their grain of analysis.

2.4.3 Selection of an appropriate extent to capture the entire environmental range of island species: a crucial consideration for model projections

Training a SDM on only a portion of a species' ecological niche can lead to problems when predicting into new climates or geographical areas (Guillera-Arroita et al., 2015; Peterson et al., 2011). By not encompassing the entire environmental range of a species, valuable information may be overlooked, and the model may make inaccurate extrapolations into environmental spaces where data are lacking. We discovered that a majority of studies of non-endemic native species, and a significant portion of assessments of non-native species, only used data from the focal-insular environmental space. However, these species, by definition, have larger distributions than the focal island(s) they inhabit. Consequently, by solely utilising data from the limited island environment, the SDMs may be unable to accurately predict potential range expansions, contractions, or shifts that could occur in new locations or under different climatic scenarios (Thuiller et al., 2004; Pang et al., 2022).

Considering our finding that the impacts of climate change and invasive species have been two of the most extensively explored topics in the island SDM literature, more focus on this issue is required in future island SDM applications that involve the spatial of temporal projection of models.

2.4.4 Mitigating model uncertainty by selecting appropriate methods for small sample sizes and presence-only data

In line with the recommendations of Araujo et al. (2019) for the optimal implementation of Species Distribution Models (SDMs), addressing uncertainty related to the choice of modelling technique is an essential consideration. While using an ensemble of diverse methods has been proposed to reduce uncertainty

(Araújo & New, 2007; Dormann et al., 2018; Araujo et al., 2019), recent evidence suggests that including inappropriate models in an ensemble can compromise performance, potentially being less effective than employing a single technique (Valavi et al., 2021). Notably, our literature review revealed that model ensembles fitted with small occurrence datasets often included at least one data-hungry method (e.g., ANN, RF, MARS, SVM) known to perform poorly with sample sizes fewer than 200 (Wisz et al., 2008; Mateo et al., 2010). Furthermore, our findings indicate that most published island SDM analyses lack empirical absence data, with researchers opting for background or pseudo-absence points (PA) as substitutes. This practice is known to further reduce the performance of data-hungry modelling approaches (Akbani et al., 2004; Robinson et al., 2018), raising concerns about the reliability of model ensembles that use such methods in island SDM analyses.

Most of the reviewed studies that employed only one modelling approach exhibited a clear preference for Maxent, a presence-only method widely recognised for its high performance with small sample sizes (van Proosdij et al., 2016). Whether implemented individually or as part of an ensemble, traditional statistical methods like Maxent, GLM, or GAM (as per Valavi et al., 2021) consistently demonstrated superior performance compared to machine learning or tree-based techniques in the context of small sample sizes (Breiner et al., 2015; Valavi et al., 2021), making them safer choices for modelling island species distributions.

Given that the effect of small sample sizes on the performance of many different modelling techniques is relatively well understood (Wisz et al., 2008; van Proosdj et al., 2015; Gaul et al., 2020), prioritising research on how this aspect interacts with pseudo-absence sampling strategies is essential for island studies. This is because true absence data are generally lacking, and modelling methods have different sensitivities to unbalanced occurrence data (a largely disproportionate ratio of pseudo-absences to presence data; Benkendorf et al., 2023). Understanding how pseudo-absence data sampling influences model predictions across diverse modelling techniques, particularly in the context of limited occurrence data, is necessary for enhancing island SDM implementation. Additionally, current pseudo-absence sampling-strategy guidelines (which focus on sampling strategies), primarily designed for large continental study areas (e.g., Descombes et al., 2022), may not directly translate to SDM implementation on islands.

2.4.5 Minimise model building uncertainty by managing complexity

Our review indicated that models fitted with a high number of predictors in relation to the (small) sample size is a prevalent issue in island SDM studies, with studies frequently using two or three times more predictors than the recommended ratio based on available occurrence points during model calibration (Vaughan &

Ormerod, 2003). This problem may result in overfitted models, which compromises their transferability into novel spatial or temporal scenarios (Manzoor et al., 2018). Alternative approaches to modelling species distributions with small datasets have been proposed. One such approach is the use of 'ensembles of small models' (Lomba et al., 2010; Breiner et al., 2015). This methodology involves fitting specific bivariate models and then averaging them as an ensemble, weighted by model performance (Breiner et al., 2015). Studies have shown that this approach significantly improves predictive power for range-restricted or data-limited species compared to standard modelling techniques used for widespread species (Breiner et al., 2015; Fitzpatrick & Ellison, 2018). Despite its potential to overcome the commonly detected issue of overfitted SDMs for island species, this strategy has rarely been implemented in island studies, with only two studies in our dataset exploring this approach.

Another possible alternative to mitigate overfitting in small occurrence datasets are joint-SDMs. This is an approach that has been argued to increase the effective sample size of species occurrence records by leveraging the power from species with more occurrences in the data, and their co-occurrence probabilities (Warton et al., 2015; Pichler & Hartig, 2021). However, implementing this method in many island environments may be less feasible due to its dependency on community-level occurrence knowledge, which is often limited, given the low number of archipelagos and islands extensively studied (Borges et al., 2018). Additionally, joint-SDMs have not consistently demonstrated increased performance relative to standard SDM techniques (Tikhonov et al., 2017; Zhang et al., 2020); therefore, caution is advised when considering this alternative.

2.4.6 GBIF data may hold potential for closing knowledge gaps in island SDM studies

By exploring the origins of data used for modelling island species distributions, we identified a geographical bias in the location of island SDM studies: a concentration of island SDM studies on temperate islands in the Pacific and Atlantic regions (e.g., Canaries, Azores and Mediterranean islands), with a smaller number of studies focused on tropical islands. This bias is likely a result of a range of historical, logistical (e.g., limited infrastructure, difficulties with undertaking field work in remote locations), and funding factors (Sutton et al., 2021). The publication of regional / archipelago-specific databases from well-studied island groups, while obviously providing useful resources, likely propagates this bias as they represent an easy-to-use data source for island SDM studies. The open access GBIF database has seemingly been used to fill the data gaps in several otherwise underrepresented regions, as most of the tropical island studies used this data source to develop SDMs (e.g., Indonesian, Caribbean and Eastern pacific islands). Thus, exploring the potential of GBIF data to fill the remaining gaps in knowledge of tropical island species distributions more broadly is a pressing concern, given that many tropical insular territories are amongst the most biodiverse and threatened regions globally, while at the same time being some of the least studied (Cayuela et al., 2009; Myers et al., 2000).

We identified high availability of data on GBIF for a range of taxa and island regions (e.g., tropical regions from the Atlantic and Pacific Ocean) that are underrepresented in the island SDM literature, suggesting that this database may have the potential to reduce knowledge gaps. In addition, numerous species of vertebrate taxa (e.g., anurans, birds [likely due to the link between GBIF and eBird], mammals, reptiles), showed good levels of representation across most island regions at the global-scale, enabling future larger-scale comparative studies to be developed. On the other hand, for plant species (both vascular and non-vascular), sufficient occurrence data (≥20 occurrences) are only concentrated in the island regions that have already been relatively well studied. For the remaining studied taxa, GBIF data does not appear to currently represent a useful source of island occurrence information.

We acknowledge that our metric (M-Index) sets a conservative threshold (≥20n), given that SDM predictions can potentially remain informative at smaller sample sizes (Pearson et al., 2007). In addition, our metric assesses the quantity of data and does not explicitly account for data quality. Global databases often come with inherent biases, including species misidentification, low spatial precision, and uneven sampling effort (Hortal et al., 2007; Beck et al., 2014). However, a growing set of statistical methods, such as ROBITT (Representativeness and Bias Identification Tool for Tabular Data; Boyd et al., 2022), are increasingly available to address these data limitations. ROBITT was specifically developed to identify and manage issues of representativeness and potential biases in datasets, particularly those compiled from diverse sources with differing aims, scopes, and methodologies. By focusing on improving data quality across several critical domains—geographic, temporal, environmental, and taxonomic—each of which can introduce unique biases that affect study outcomes, ROBITT facilitates the homogenisation and appropriate curation of GBIF data, enhancing its utility for island studies.

2.5 Conclusion

Given that the full distributions of many island species are unknown, species distribution models (SDM) represent a vitally important tool for island biologists, and are essential for enabling data-informed conservation and management decisions related to this highly threatened component of global biodiversity. However, most SDM frameworks and recommendations relate to species with relatively wide ranges, and are generally not appropriate for insular species and environments, which have several unique features (Whittaker et al., 2023). The results of the present review (i) indicate that there are a number of areas related to the application of SDMs in an island context that require careful consideration and future research, and (ii) provide the foundation for the development of an explicit island SDM framework going forward.

2.6 References

Abrams, Michael; Crippen, Robert; Fujisada, Hiroyuki. (2020). ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). *Remote Sensing*, *12(7)*, 1156.

Ando, H. (2019). Genetic and ecological conservation issues for oceanic island Birds, revealed by a combination of the latest molecular techniques and conventional fieldwork. *Ecological Research*, *34*(2), 255-264.

Araújo, M. B., and M. New. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution, 22*, 42–47.

Azevedo EB. (1996). Modelação do Clima Insular à Escala Local—Modelo CIELO Aplicado à Ilha Terceira. PhD Thesis, Universidade dos Açores, Angra do Heroı´smo, Portugal.

Baker DJ, Hartley AJ, Pearce-Higgins JW, Jones RG, Willis SG. (2017). Neglected issues in using weather and climate information in ecology and biogeography. *Diversity and Distributions*. *23*, 329–340.

Barajas-Barbosa, M.P., Weigelt, P., Borregaard, M.K., Keppel, G. & Kreft, H. (2020). Environmental heterogeneity dynamics drive plant diversity on oceanic islands. *Journal of Biogeography*, *47*, 2248-2260.

Barlow, M. M., Johnson, C.N., McDowell, M. C., M. C., Fielding, M. W., Amin, R.J., & Brewster, R. (2021). Species distribution models for conservation: Identifying translocation sites for eastern quolls under climate change. *Global Ecology and Conservation*, *29*, e01735.

Bazzichetto, M., Massol, F., Carboni, M., Lenoir, J., Lembrechts, J. J., Joly, R., & Renault, D. (2021). Once upon a time in the far south: Influence of local drivers and functional traits on plant invasion in the harsh sub-Antarctic islands. *Journal of Vegetation Science*, *32*(4), e13057.

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

Benkendorf, D. J., Schwartz, S. D., Cutler, D. R., & Hawkins, C. P. (2023). Correcting for the effects of class imbalance improves the performance of machine-learning based species distribution models. *Ecological Modelling*, *483,* 110414.

Borges, P. A., Cardoso, P., Kreft, H., Whittaker, R. J., Fattorini, S., Emerson, B. C., ... & Gabriel, R. (2018). Global Island Monitoring Scheme (GIMS): a proposal for the long-term coordinated survey and monitoring of native island forest biota. *Biodiversity and Conservation*, *27*, 2567-2586.

Boyd, R. J., Powney, G. D., Burns, F., Danet, A., Duchenne, F., Grainger, M. J., ... & Pescott, O. L. (2022). ROBITT: A tool for assessing the risk-of-bias in studies of temporal trends in ecology. *Methods in Ecology and Evolution*, *13*(7), 1497-1507.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, *6*(10), 1210-1218.

Bush, A., Catullo, R. A., Mokany, K., Thornhill, A. H., Miller, J. T., & Ferrier, S. (2018). Truncation of thermal tolerance niches among Australian plants. *Global Ecology and Biogeography*, *27*(1), 22-31.

Bulgarella, M., Trewick, S. A., Minards, N. A., Jacobson, M. J., & Morgan-Richards, M. (2014). Shifting ranges of two tree weta species (*Hemideina* spp.): competitive exclusion and changing climate. *Journal of Biogeography, 41*(3), 524-535.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P. et al. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*, 59–67.

Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, Ram K (2022). rgbif: Interface to the Global Biodiversity Information Facility API. R package version 3.7.

Chiatante, G. (2022). Spatial distribution of an assemblage of an endemic genus of birds: an example from Madagascar. *African Journal of Ecology*, *60*(1), 13-26.

Denelle, P., Weigelt, P. & Kreft, H. (2023). GIFT—An R package to access the Global Inventory of Floras and Traits. *Methods in Ecology and Evolution*, *14,* 2738-2748

Descombes, P., Chauvier, Y., Brun, P., Righetti, D., Wüest, R. O., Karger, D. N., ... & Zimmermann, N. E. (2022). Strategies for sampling pseudo-absences for species distribution models in complex mountainous terrain. *bioRxiv*, 2022-03.

Devkota, R. S., Field, R., Hoffmann, S., Walentowitz, A., Medina, F. M., Vetaas, O. R., ... & Beierkuhnlein, C. (2020). Assessing the potential replacement of laurel forest by a novel ecosystem in the steep terrain of an Oceanic Island. *Remote Sensing*, *12*(24), 4013.

Diamond, S. E., Nichols, L. M., McCoy, N., Hirsch, C., Pelini, S. L., Sanders, N. J., ... & Dunn, R. R. (2012). A physiological trait-based approach to predicting the responses of species to experimental climate warming. *Ecology*, *93*(11), 2305-2312

Dormann, C. F., Calabrese, J. M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, F., … Hartig, F. (2018). Model averaging in ecology: A review of Bayesian, information-theoretic and tactical approaches. *Ecological Monographs*, *88*, 485–504.

Ellis-Soto, D., Blake, S., Soultan, A., Guézou, A., Cabrera, F., & Lötters, S. (2017). Plant species dispersed by Galapagos tortoises surf the wave of habitat suitability under anthropogenic climate change. *PLoS One*, *12*(7), e0181333.

Elith, J. and Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 677–697.

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*(4), 330-342.

Faleiro, F. V., Machado, R. B., & Loyola, R. D. (2013). Defining spatial conservation priorities in the face of land-use and climate change. *Biological Conservation*, *158*, 248-257.

Fernández-Mazuecos, M., Jiménez-Mejías, P., Rotllan-Puig, X., & Vargas, P. (2014). Narrow endemics to Mediterranean islands: moderate genetic diversity but narrow climatic niche of the ancient, critically endangered *Naufraga* (Apiaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, *16*(4), 190-202.

Fernandes, R. F., Scherrer, D., & Guisan, A. (2018). How much should one sample to accurately predict the distribution of species assemblages? A virtual community approach. *Ecological Informatics*, *48*, 125-134.

Fernández-Palacios, J. M., Kreft, H., Irl, S. D., Norder, S., Ah-Peng, C., Borges, P. A., ... & Drake, D. R. (2021). Scientists' warning–The outstanding biodiversity of islands is in peril. *Global Ecology and Conservation*, *31*, e01847.

Fernández-Palacios, J. M., Otto, R., Borregaard, M. K., Kreft, H., Price, J. P., Steinbauer, M. J., ... & Whittaker, R. J. (2021). Evolutionary winners are ecological losers among oceanic island plants. *Journal of Biogeography*, *48*(9), 2186-2198.

Fitzpatrick, M. C., Ellison, A. M., & Adamec, L. (2018). Estimating the exposure of carnivorous plants to rapid climatic change. Carnivorous Plants: *Physiology, Ecology and Evolution*. Oxford University Press, London.

Fourcade, Y. (2016). Comparing species distributions modelled from occurrence data and from expert-based range maps. Implication for predicting range shifts with climate change. *Ecological Informatics*, *36,* 8-14.

Franklin, J. (2013). Species distribution models in conservation biogeography: developments and challenges. *Diversity and Distributions, 19*(10), 1217-1223.

Franklin, Janet; Davis, Frank W.; Ikegami, Makihiko; Syphard, Alexandra D.; Flint, Lorraine E.; Flint, Alan L.; Hannah, Lee (2013). Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Global Change Biology, 19(2),* 473–483.

Gábor, L., Jetz, W., Lu, M., Rocchini, D., Cord, A., Malavasi, M., ... & Moudrý, V. (2022). Positional errors in species distribution modelling are not overcome by the coarser grains of analysis. *Methods in Ecology and Evolution*, *13*(10), 2289-2302.

Gábor, L., Jetz, W., Zarzo-Arias, A., Winner, K., Yanco, S., Pinkert, S., ... & Moudrý, V. (2023). Species distribution models affected by positional uncertainty in species occurrences can still be ecologically interpretable. *Ecography*, *2023*(6), e06358.

Gardner, A. S., Maclean, I. M., & Gaston, K. J. (2019). Climatic predictors of species distributions neglect biophysiologically meaningful variables. *Diversity and Distributions, 25*(8), 1318-1333.

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

GBIF.org (2021), GBIF Home Page. Available from: https://www.gbif.org [March 2021].

Girini, J. M., Palacio, F. X., & Zelaya, P. V. (2017). Predictive modeling for allopatric *Strix* (Strigiformes: Strigidae) owls in South America: determinants of their distributions and ecological niche-based processes. *Journal of Field Ornithology*, *88*(1), 1-15.

Goedecke, F., mArcenò, C., Guarino, R., Jahn, R., & Bergermeier, E. (2020). Recoprocal extrapolation of species distribution models between two island–specialists perform better than generalists and geological data reduces prediction accuracy. *Ecological Indicators, 108*, 105652.

Graham, L. J., Spake, R., Gillings, S., Watts, K., & Eigenbrod, F. (2019). Incorporating fine-scale environmental heterogeneity into broad-extent models. *Methods in Ecology and Evolution*, *10*(6), 767-778.

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., ... & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, *24*(3), 276-292.

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

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., ... & Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424-1435.

Hansen, M. F., Nawangsari, V. A., van Beest, F. M., Schmidt, N. M., Fuentes, A., Traeholt, C., ... & Dabelsteen, T. (2019). Estimating densities and spatial distribution of a commensal primate species, the long-tailed macaque (*Macaca fascicularis*). *Conservation Science and Practice*, *1*(9), e88.

Hanz, D. M., Cutts, V., Barajas-Barbosa, M. P., Algar, A., Beierkuhnlein, C., Collart, F., ... & Irl, S. D. (2023). Effects of climate change on the distribution of plant species and plant functional strategies on the Canary Islands. *Diversity and Distributions 29*(9), 1157-1171.

Hao, T., Elith, J., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2019). A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. *Diversity and Distributions, 25*(5), 839-852.

Heinänen, S., Erola, J., & von Numers, M. (2012). High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. *Landscape Ecology*, *27*(4), 545-555.

Helmstetter, A. J., Cable, S., Rakotonasolo, F., Rabarijaona, R., Rakotoarinivo, M., Eiserhardt, W. L., ... & Papadopulos, A. S. (2021). The demographic history of Madagascan micro-endemics: have rare species always been rare? *Proceedings of the Royal Society B*, *288*(1959), 20210957.

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

Hickisch, R., Hodgetts, T., Johnson, P. J., Sillero-Zubiri, C., Tockner, K., & Macdonald, D. W. (2019). Effects of publication bias on conservation planning. *Conservation Biology*, *33*(5), 1151-1163.

Hsu, R. C. C., Tamis, W. L., Raes, N., de Snoo, G. R., Wolf, J. H., Oostermeijer, G., & Lin, S. H. (2012). Simulating climate change impacts on forests and associated vascular epiphytes in a subtropical island of East Asia. *Diversity and Distributions, 18*(4), 334-347.

Kafash, A., Ashrafi, S., Yousefi, M., Rastegar-Pouyani, E., Rajabizadeh, M., Ahmadzadeh, F., ... & Pellissier, L. (2020). Reptile species richness associated to ecological and historical variables in Iran. *Scientific Reports*, *10*(1), 1-11.

Kerr, J.T., Kharouba, H. M., & Currie, D. J. (2007). The macroecological contribution to global change solutions. *Science, 316*(5831), 1581-1584.

Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisch, P. L., Nowicki, C., ...& Bartglott, W. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences, 106*(23), 9322-9327.

Khosravi, R., Hemami, M. R., Malekian, M., Flint, A., & Flint, L. (2016). Maxent modeling for predicting potential distribution of goitered gazelle in central Iran: the effect of extent and grain size on performance of the model. *Turkish Journal of Zoology, 40*(4), 574-585.

Ladle, R. J., & Whittaker, R. J. (Eds.). (2011). *Conservation biogeography*. John Wiley & Sons.

Lamelas-López, L., Pardavila, X., Borges, P. A., Santos-Reis, M., Amorim, I. R., & Santos, M. J. (2020). Modelling the distribution of *Mustela nivalis* and *M. putorius* in the Azores archipelago based on native and introduced ranges*. PloS One*, *15*(8), e0237216.

Lannuzel, G., Balmot, J., Dubos, N., Thibault, M., & Fogliani, B. (2021). High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia. *Biodiversity and Conservation*, *30*(4), 963-990.

Lentini, P.E., Stirnemann, I. A., Stojanovic, D., Worthy, T. H., & Stein, J. A. (2018). Using fossil records to inform reintroduction of the kakapo as a refugee species. *Biological Conservation, 217*, 157-165.

Leroy, B. (2022). Choosing presence-only species distribution models. *Journal of Biogeography*, *50*, 247-250.

Liao, C.C., & Chen, Y. H. (2021). Improving performance of species distribution model in mountainous areas with complex topography. *Ecological Research*, *36*(4), 648-662.

Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation, 143*(11), 2647-2657.

Longadge, R., Dongre, S. S., & Malik, L. (2013). Class imbalance problem in data mining: Review. *International Journal of Computer Science and Network*, *2*, 83–87.

Longcore, T., Noujdina, N., & Dixon, P. J. (2018). Landscape modeling of the potential natural vegetation of Santa Catalina Island, California. *Western North American Naturalist*, *78*(4), 617-632.

Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size–finer may not always be better. *Scientific Reports*, *8*(1), 1-9

Mateo, R. G., Felicísimo, Á. M., & Munoz, J. (2010). Effects of the number of presences on reliability and stability of MARS species distribution models: the importance of regional niche variation and ecological heterogeneity. *Journal of Vegetation Science, 21*(5), 908-922.

Matthews, T. J. (2021). On the biogeography of habitat islands: the importance of matrix effects, noncore species, and source-sink dynamics. *The Quarterly Review of Biology*, *96*(2), 73-104.

Matthews, T. J., & Triantis, K. (2021). Island biogeography. *Current Biology*, *31*(19), R1201-R1207.

Matthews, T. J., Wayman, J. P., Cardoso, P., Sayol, F., Hume, J. P., Ulrich, W., ... & Triantis, K. A. (2022). Threatened and extinct island endemic birds of the world: Distribution, threats and functional diversity. *Journal of Biogeography*, *49*(11), 1920-1940.

Mertes, K., & Jetz, W. (2018). Disentangling scale dependencies in species environmental niches and distributions. *Ecography*, *41*(10), 1604-1615.

Meyer, C. B., & Thuiller, W. (2006). Accuracy of resource selection functions across spatial scales. *Diversity and Distributions, 12*(3), 288-297.

Moudrý, V., & Šímová, P. (2012). Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Information Science*, *26*(11), 2083-2095.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*(6772), 853-858.

Nogué, S., Santos, A. M., Birks, H. J. B., Björck, S., Castilla- Beltrán, A., Connor, S., … & Steinbauer, M. J. (2021). The human dimension of biodiversity changes on islands. *Science, 372*(6541), 488-491.

NASA/METI/AIST/Japan Space Systems, and U.S./Japan ASTER Science Team. ASTER global digital elevation model V003. (2018). Distributed by NASA EOSDIS Land Processes DAAC, https://doi.org/ 10.5067/ASTER/ASTGTM. 003.

Pang, S. E., Zeng, Y., De Alban, J. D. T., & Webb, E. L. (2022). Occurrence–habitat mismatching and niche truncation when modelling distributions affected by anthropogenic range contractions. *Diversity and Distributions*, *28*(6), 1327-1343.

Parmesan C., G. Yohe. (2003). A globally coherent fingerprint of climate change impacts across natural systems, *Nature*, *421*, 37-42

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Global Ecology and Biogeography*, *12*(5), 361–371.

Pearson, R. G., Raxworthy, C. J., Nakamura, M., & Townsend Peterson, A. (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, *34*(1), 102-117.

Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). Ecological niches and geographic distributions Princeton University Press, Princeton, NJ..

Pichler, M., & Hartig, F. (2021). A new joint species distribution model for faster and more accurate inference of species associations from big community data. *Methods in Ecology and Evolutio*n, *12*(11), 2159-2173.

Porfirio, L. L., Harris, R. M., Stojanovic, D., Webb, M. H., & Mackey, B. (2016). Projected direct and indirect effects of climate change on the swift parrot, an endangered migratory species. *Emu-Austral Ornithology*, *116*(3), 273-283.

Price, J.P., Jacobi, J.D., Gon, S.M., III, Matsuwaki, D., Mehrhoff, L., Wagner, W., Lucas, M., and Rowe, B. (2012). Mapping plant species ranges in the Hawaiian Islands—Developing a methodology and associated GIS layers: U.S. Geological Survey Open-File Report 2012–1192, 34 p., 1 appendix (species table), 1,158 maps, available at http://pubs.usgs.gov/of/2012/1192/.

QGIS Development Team. (2022). QGIS Geographic Information System. Open Source Geospatial Foundation Project. htttp://qgos.osgeo.org.

Radosavljevic, A., & Anderson, R. P. (2014). Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography*, *41*(4), 629-643.

Rapacciuolo, G., Roy, D. B., Gillings, S., Fox, R., Walker, K., & Purvis, A. (2012). Climatic associations of British species distributions show good transferability in time but low predictive accuracy for range change. *Plos One*, *7*(7), e420112.

Ricketts, T. H., Dinerstein, E., Boucher, T., Brooks, T. M., Butchart, S. H., Hoffmann, M., ... & Rodrigues, A. S. (2005). Pinpointing and preventing imminent extinctions. *Proceedings of the National Academy of Sciences*, *102*(51), 18497-18501.

Robinson, N. M., Nelson, W. A., Costello, M. J., Sutherland, J. E., & Lundquist, C. J. (2017). A systematic review of marine-based species distribution models (SDMs) with recommendations for best practice. *Frontiers in Marine Science*, *4*, 421.

Robinson, O. J., Ruiz-Gutierrez, V., & Fink, D. (2018). Correcting for bias in distribution modelling for rare species using citizen science data. *Diversity and Distributions*, *24*(4), 460-472.

Rodríguez-Castañeda G, Hof AR, Jansson R, Harding LE. (2012). Predicting the fate of biodiversity using species' distribution models: enhancing model comparability and repeatability*. Plos One, 7*(9), e44402.

Rosenblad, K. C., Perret, D. L., & Sax, D. F. (2019). Niche syndromes reveal climate-driven extinction threat to island endemic conifers. *Nature Climate Change*, *9*(8), 627-631.

Russell, J. C., & Kueffer, C. (2019). Island biodiversity in the Anthropocene. *Annual Review of Environment and Resources*, *44*, 31-60.

Santini, L., Benítez-López, A., Maiorano, L., Čengić, M., & Huijbregts, M. A. J. (2021). Assessing the reliability of species distribution projections in climate change research. *Diversity and Distributions*, *27*(6), 1035–1050.

Segal, R. D., Massaro, M., Carlile, N., & Whitsed, R. (2021). Small-scale species distribution model identifies restricted breeding habitat for an endemic island bird. *Animal Conservation*, *24*(6), 959-969.

Spiers, J. A., Oatham, M. P., Rostant, L. V., & Farrell, A. D. (2018). Applying species distribution modelling to improving conservation based decisions: a gap analysis of Trinidad and Tobago's endemic vascular plants. *Biodiversity and Conservation*, *27*(11), 2931-2949.

Stockwell, D. R., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, *148*(1), 1-13.

Stuber, E. F., & Fontaine, J. J. (2019). How characteristic is the species characteristic selection scale? *Global Ecology and Biogeography*, *28*(12), 1839-1854.

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

Thurman, L. L., Stein, B. A., Beever, E.A., Foden W., Giange, S.R., Green, N., …&Young, B.E. (2020). Persist in place or shift in space? Evaluating the adaptative capacity of species to climate change. *Frontiers in Ecology and Environment*, *18*(19), 520-528

Tikhonov, G., Abrego, N., Dunson, D., & Ovaskainen, O. (2017). Using joint species distribution models for evaluating how species-to-species associations depend on the environmental context. *Methods in Ecology and Evolution*, *8*(4), 443-452.

Turvey, S. T., Kennerley, R. J., Hudson, M. A., Nuñez-Miño, J. M., & Young, R. P. (2020). Assessing congruence of opportunistic records and systematic surveys for predicting Hispaniolan mammal species distributions. *Ecology and Evolution*, *10*(11), 5056-5068.

Vaughan, I. P., & Ormerod, S. J. (2003). Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. *Conservation Biology*, *17*(6), 1601-1611.

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2021). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs, 1*, e01486.

van Proosdij AS, Sosef MS, Wieringa JJ, Raes N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography, 39*, 542–552

Veitch, C. R., Clout, M. N. (Eds.). (2002). Turning the tide: The eradication of invasive species. Proceedings of the International Conference on Eradication of Island Invasives (No. 27). IUCN.

Vergilio, M., Fonseca, C., Calado, H., Borges, P.A., Elias, R. B., Gabriel, R., …& Cardoso, P. (2016). Assessing the efficiency of protected areas to represent biodiversity: a small island case study. *Environmental Conservation, 43*(4), 337-349.

Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. (2015). So many variables: joint modeling in community ecology. *Trends in Ecology & Evolution*, *30(*12), 766-779.

Weigelt, P., Jetz, W., & Kreft, H. (2013). Bioclimatic and physical characterization of the world's islands. *Proceedings of the National Academy of Sciences*, *110*(38), 15307-15312.

Wells, K. D. (2010). The ecology and behavior of amphibians. University of Chicago press., Chicago.

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

Whittaker, R. J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: Taking the long view of nature's laboratories. *Science*, *357*(6354), eaam8326

Whittaker, R.J., Fernández-Palacios, J.M. & Matthews, T.J. (2023). Island biogeography: geo-environmental dynamics, ecology, evolution, human impact, and conservation. Oxford University Press, Oxford.

Zhang, C., Chen, Y., Xu, B., Xue, Y., & Ren, Y. (2020). Improving prediction of rare species' distribution from community data. *Scientific Reports, 10*(1), 1-9.

. Enhancing SDM Implementation in Island

3 . Improving SDM implementation in island environments: tackling data challenges associated with fine-grained predictions

Abstract

Applying Species Distribution Models (SDM) in island environments presents unique challenges due to spatial constraints, the high habitat specialisation of many island species, and limited occurrence data. This study focuses on optimising presence-only SDMs for island ecosystems, using empirical plant species data and simulated distributions as case studies. A key emphasis is placed on the use of high-resolution environmental predictors, while addressing complications posed by scarce and often imprecise occurrence data, which can hinder the understanding of insular species distributions at appropriate analytical scales.

To optimise high-resolution SDMs in island systems, the study highlights the importance of balancing presence data with pseudo-absence/background data to enhance model accuracy and validation metrics, particularly when considering species' range size (i.e., wide vs. narrow-range species). It also assesses how these factors interact with different modelling methods. Additionally, the study explores the potential benefits of including records with positional inaccuracies to increase sample size. With appropriate methodological frameworks, such as Bayesian Additive Regression Trees (BART) and generalised Linear Models GLM, the negative impact of analysing coarse-precision occurrence data can be mitigated. These methods proved especially beneficial for narrow-ranged species, which require

larger sample sizes and are more sensitive to spatial inaccuracies. In contrast, wide-ranging species are less affected by coarse-precision occurrence data and can often be modelled with fewer, more precise data points.

In scenarios where the spatial data precision aligns with the spatial analytic resolution, findings highlight that Generalised Additive Models (GAMs) achieve the highest accuracy in an island context. Notably, these models also require smaller sample sizes, regardless of range size. Contrary to expectations, Maxent ranked second, and was found to have difficulties in modelling wide-ranging species, requiring relatively large sample sizes. By selecting the appropriate algorithm based on species prevalence and data availability, the study provides practical recommendations to improve SDM performance in island environments, enhancing model reliability to better support conservation efforts, even when working with sparse or uncertain data.

3.1 Introduction

The rapid alteration of the Earth system, driven by human activities, has led to large-scale distributional changes in biodiversity (Araujo et al., 2019). This issue is particularly acute on islands, where species are often more vulnerable due to a range of factors, including small population sizes, limited habitat availability, and unique ecological and evolutionary dynamics (Whittaker et al., 2017; Fernández-Palacios et al., 2021; Matthews et al., 2022). A comprehensive understanding of the geographic and environmental drivers of island biodiversity is essential for developing effective conservation strategies for these environments.

Islands are known centres of plant endemism, and island plants have been shown to face disproportionate risks relative to their continental relatives (Kier et al., 2009; Caujape-Castells et al., 2010). As with all ecosystems, plants are a key component of island systems, as they provide a variety of crucial ecosystem services, including water regulation, erosion control, and food provision (Borges et al., 2018). However, for many archipelagos we still lack accurate data on the distributions of plant species, information that is needed for informed conservation planning and management decisions (Cursach et al., 2020). Species Distribution Models (SDMs) have emerged as a valuable tool in this regard (Benavides Rios et al., 2024). SDMs relate species occurrence data, typically in the form of presence and absence records, with environmental predictors to generate predictions of species distributions across space and time (Guisan et al., 2017; Franklin, 2010; Araujo et al., 2019).

Effective use of SDMs is context dependent and thus the use of this methodological tool in a given study requires careful decision-making throughout the modelling process (Ferrier et al., 2022). Although several studies have proposed guidelines to facilitate SDM fitting (e.g., Guillera-Arroita et al., 2015; Guisan et al., 2017; Araujo et al., 2019), they have focused primarily on species in continental environments, limiting their applicability to island environments (Leroy, 2022; but see Benavides Rios et al., 2024).

Islands—particularly volcanic oceanic islands, which are the focus of this study—differ from continental systems in several key ways, requiring tailored guidelines for effective SDMs (Leroy, 2022; Benavides Rios et al., 2024). The limited spatial context of islands, characterised by relatively small size and greater isolation, combined with significant topographic and environmental variation over small areas (Barajas-Barbosa et al., 2020), means that we need to develop a deeper understanding of how island species respond to environmental gradients and change at fine spatial scales (Patiño et al., 2023; Benavides Rios et al., 2024). This approach contrasts with the common use of large-scale bioclimatic variables in modelling species distributions at continental scales (Pearson & Dawson, 2003; Gardner et al., 2019). Therefore, it is crucial to examine island species' distributions using high-resolution environmental data that capture microhabitat variation (Segal et al., 2021; Patiño et al., 2023), which is a key factor influencing the distribution of many island plants (Irl et al., 2020; Hanz et al., 2023; Lannuzel et al., 2022).

Another issue with using SDMs in an island plant context is that, for many archipelagos, occurrence data are often limited. This is due to small sample sizes resulting from natural factors (e.g., the presence of highly specialised insular endemics with small ranges; Chiatante, 2022), anthropogenic factors (e.g., the reduced population sizes of many island species: Nogué et al., 2021; Matthews et al., 2022), and/or the remoteness and limited accessibility of many islands (Borges et al., 2018). Small amounts of occurrence data can present significant challenges for fine-resolution SDMs. This is because finer-grained grids divide environmental variability into smaller, more specific segments, making it less likely that small samples will capture the full range of suitable environmental conditions for a species. In contrast, coarser grids encompass broader environmental information within fewer grid cells, which may allow small samples to more easily represent a wider range of the species' habitat requirements, though this comes with a higher risk of overestimation (Patiño et al., 2023). As a result, while fine-scale predictors are typically a better choice in the context of islands, their use may require larger sample sizes to ensure adequate coverage of the species' habitat requirements. However, previous studies suggest that for species with smaller ranges, such as many island endemics, fewer occurrence records may still be sufficient to accurately predict their distributions (Pearson et al., 2007; van Proosdij et al., 2015). Thus, further research is needed to understand how a species' range size influences data requirements for robust SDMs at fine spatial scales on islands.

To effectively use high-resolution environmental predictors in SDMs, the precision of the occurrence data must match the resolution of the predictors (Moudrý & Šímová, 2012). For many island species, limited and sparse occurrence data often result in large geographic uncertainties (i.e., occurrence coordinates frequently cover broad areas). Excluding such records can negatively impact predictions (Fernandez et al., 2009; Naimi et al., 2014), as reducing sample size increases model uncertainty (Gaul et al., 2020; Graham et al., 2008; Hernández et al., 2006; Soultan & Safi, 2017). Conversely, retaining these uncertain records just to increase sample size can be especially problematic for islands with steep environmental gradients—such as those on topographically diverse volcanic islands—or for highly specialised insular endemic species. In these cases, the uncertainty in the occurrence data can significantly affect SDM predictions, particularly when fine-scale predictors are used (Visscher, 2006; Gábor et al., 2020, 2023). Therefore, it is crucial to evaluate whether uncertainty in occurrence data can be tolerated and how best, given the limited data availability in island contexts, to optimise the use of occurrence data in SDMs.

A final challenge related to the application of SDMs, in both continental and island settings, is that occurrence datasets often contain only opportunistic observations and are thus limited to species presence data (Fourcade et al., 2014). However, many modelling algorithms require absence data which, if not available, are generated using pseudo-absence/background (Psa/BG) points. There are three issues with this approach in the context of islands. First, determining the number of Psa/BG points and the accompanying selection strategy lacks a clear consensus (Grimmet et al., 2020; Descombes et al., 2023). Second, existing guidelines for PsA/BG sampling are typically tailored for species with extensive ranges (e.g., areas >10,000 km^2, based on the standardised number of Psa/BG points recommended, assuming a typical 1 km grid resolution in environmental descriptors; Barbet-Massin et al., 2012; Descombes et al., 2023), making these recommendations incompatible in the context of small island regions. Finally, model validation metrics evaluated with Psa/BG data are affected by sample size, imbalance in occurrence datasets (i.e., a large ratio of Psa/BG points compared to a small sample of presence points), and species specialisation (Lobo et al., 2008; Santika, 2011). Thus, research is needed to determine the ideal ratio of Psa/BG sampling for different modelling methods, considering data availability, study area characteristics, specialisation levels (or species prevalence), and modelling technique requirements.

Based on the above, it is evident that effective SDM implementation on islands requires selecting suitable modelling strategies for use with limited occurrence and environmental data. To this end, our study aims to improve the effectiveness of SDM use in island systems by examining how the quantity and quality of occurrence data affect model predictions at fine analytical scales. We will evaluate whether certain methods improve SDM predictions with 1) varying ratios of presence to PsA/BG data, as modelling methods can have varying sensitivities to unbalanced datasets (Barbet-Massin et al., 2012; Hertzog et al., 2014; Wisz & Guisan, 2009); and 2) varying levels of occurrence data inaccuracies. We then investigate if these results differ based on i) the quantity of occurrence data available (sample size), and ii) species range size.

To address these research objectives, we will use a dataset of plant occurrence data (with detailed presence and absence data) from the volcanic Revillagigedo Archipelago (Mexico), and simulated virtual species that mimic known species distribution patterns in the study area. We seek to provide practical recommendations drawn from our findings, particularly for island endemic plant species with limited occurrence data.

3.2 Materials and Methods

3.2.1 Study area

Our study focuses on the two largest islands in the Revillagigedo archipelago, Socorro (132 km^2) and Clarión (19.8 km^2), chosen due to (i) an in-depth understanding of species distribution patterns acquired from extensive fieldwork expeditions led by one of the authors (e.g., Benavides et al., 2019), (ii) the availability of high-resolution environmental data, and (iii) the sizes of the islands, representing a relatively smaller and larger island.

3.2.2 Virtual species

To ensure robust conclusions, it was necessary to include a relatively large number of species in the analyses, including large (≥50n), unclustered, and unbiased island species distribution datasets exhibiting varied distribution patterns and levels of species specialisation (Boria et al., 2014; Fourcade et al., 2014; Araujo et al., 2019). However, challenges in extensively sampling certain parts of the selected islands in the Revillagigedo archipelago were encountered, and despite best efforts, sample sizes rarely exceeded 30 records per species. Consequently, detailed data on the distributions of plant species in the Revillagigedo Archipelago (e.g., Benavides et al., 2019) were utilised to create virtual species closely resembling natural spatial distribution patterns (see Supplementary Figure S3.1.1 in Appendix S.3.1). This approach avoids unrealistic species distributions (i.e., ensuring realistic responses to environmental gradients: Miller, 2014) and facilitates robust testing of model conceptualisation and implementation (Meynar et al., 2019).

Environmental descriptors for calibrating virtual species were derived from a 30m resolution digital elevation model (ASTER Global Digital Elevation Map v3: NASA & METI, 2018). These descriptors, including elevation, slope, aspect, and distance to the coast, served as proxies for microclimatic variation, and are considered important drivers of island plant species distributions (Irl et al., 2020; Lannuzel et al., 2021); they were generated using the 'Terra' R package (Hijmans, 2022). Climatic variables were excluded as only coarse resolution (1 km) data were available; such data are known to increase model uncertainty, particularly in tropical regions (Deblawue et al., 2016; Benavides et al., 2024).

Using the 'Virtualspecies' R package (Leroy et al., 2016), four virtual species were created for each island using the 'niche.breadth' function and applying the Principal Component Analysis (PCA) method. This function generated species with realistic responses to environmental gradients (Soultan & Safi, 2017) and allowed for control over range sizes by selecting the extent of the PCA (environmental) space occupied by the species. This was managed with the 'niche.breadth' parameter, which could be set to either 'Narrow' or 'Wide' (see S3.1.1). Each island had two virtual species with wide distributions and two with narrow distributions. Selecting four virtual species per island (eight in total) was deemed adequate due to the relatively small study areas, computational demands (e.g., the total number of core hours used in the analyses was 420 core hours per species), and an experimental design that required numerous replicates.

3.2.3 Real species data

To address the potential limitation of limited transferability inherent in conclusions derived solely from virtual species data, we conducted complementary analyses using models fitted to real species presence-absence data, essentially providing a "true species distribution" equivalent for the virtual species. Our baseline models were constructed using occurrence data collected during previous field excursions to both islands (see Appendix S3.2 Method to obtain presence/absence plant data from the Revillagigedo Archipelago, Mexico.for detailed information). From this dataset, we selected four species from each island with sufficient presence/absence data (≥10 presences/ ≥15 absences). Based on data availability and species range sizes, we chose two species with widespread distributions and two species with relatively narrow distributions. Based on field observations, widespread species were classified as inhabiting numerous habitats across the islands, while narrow-range species were exclusive to either coastal environments or a narrow elevation band. A summary of the selected species and their corresponding records is provided in Table 3.1.

Table 3.1 Analysed plant species list with range sizes and the number (N) of presence and absence records used for fitting SDMs. + Single island endemics

Island	Species	Range sizes	N Presence	N Absence
Clarion	*Euphorbia anthonyi+*	Wide	15	15
	Tribulus cistoides	Wide	14	25
	Karwinskia humboldtiana	Narrow (Elevation band)	11	25
	Teucrium townsendi+	Narrow (Coastal)	10	53
Socorro	*Croton masonii+*	Wide	17	60
	Psidium socorrense+	Wide	25	54
	Bursera epinnata	Narrow (Elevation band)	69	94
	Conocarpus erectus	Narrow (Coastal)	10	19

3.2.4 Modelling methods

Due to the small sample sizes for most analysed species (n<30), the 'ensemble of small models' (ESM) approach was selected for modelling species distributions, a method that has been shown to be suitable for narrow-ranged species with limited occurrence data (Lomba et al., 2010; Breiner et al., 2015). All potential combinations of bivariate models are fitted, and their performance is evaluated based on AUC scores. Then, a final averaged model is created, incorporating all informative small models (AUC>0.5; Lomba et al., 2010; Breiner et al., 2015).

The ESM analyses used three regression-based modelling methods that have shown satisfactory results in the ESM approach (Breiner et al., 2015, 2018): generalized linear models (GLM; binomial family), General Additive Models (GAM; Hastie & Tibshirani, 1986), and Maxent (Phillips et al., 2006). Additionally, a tree-based method, Bayesian Additive Regression Trees (BART) was included, a relatively new approach in ecological studies (Carlson, 2020). The applicability of BART within the ESM framework has not yet been tested. However, exploring this approach is valuable, as previous studies have shown it to be informative for island SDM studies, often generating satisfactory predictions (Hanz et al., 2023).

To avoid adding unnecessary complexity in the modelling process, it was chosen to maintain the default settings of the respective SDM algorithms. This decision was also based on work that suggests that parameter tuning has a lower impact on predictive capability for the selected modelling methods when sample sizes are small (<30 records: Valavi et al., 2022; Radomski et al., 2022).

3.2.5 Presence and Pseudo-Absence Ratios across Modelling Methods: Impact on Model Accuracy and Validation

Virtual species

The total number of grids at the selected resolution (30 m) for each island was quantified, and 5%, 10%, 20%, and 40% of pseudo-absence/background (Psa/BG) points were randomly sampled in cells without presence records. This method was chosen for its transparency, fewer assumptions, and proven accuracy (Stokland et al., 2011; Descombes et al., 2023). It was examined how different occurrence data availability levels affected model accuracy across varied presence to Psa/BG data ratios.

For each virtual species, 20 replicates of randomly sampled presence points at four sample sizes (5, 10, 20, and 50) were generated, totalling 80 replicates. Each replicate was paired with one of 10 random samples of Psa/BG for each of the four percentages, resulting in 800 combinations. Bivariate models were fitted for all possible combinations of 1 out of 5 bivariate predictor combinations, 1 out of 20 sample size replicates, 1 out of 10 Psa/BG point percentage replicates, and 1 out of 4 modelling methods (Fig. 3.1). In each run, 5-fold cross-validation was performed, using 20% of occurrence points for evaluation and the remaining 80% for training. Subsequently, the five bivariate combinations of small models were combined into an ESM by selecting all small models with AUC > 0.5. Thus, 64,000 small models were generated for each species, combining 80 sample size replicates, 4 Psa/BG percentages, 10 Psa/BG replicates per percentage, 5 bivariate combinations, and 4 modelling methods (Fig. 3.1).

The differences in predicted probability outputs calibrated with diverse modelling methods can make them incomparable or incompatible for combination. To address this, the probability outputs were converted into favourability measures using the "Fav" function from the Fuzzysim R package. This function calculates a prevalence-independent favourability measure for species presence, based on a model object calibrated with presence–Psa/BG data and the proportion of binarized occurrence values (i.e., model prevalence, where presence = 1 and Psa/BG = 0). The resulting output is a favourability layer ranging from 0 to 1, distinguishing between conditions that support species presence ($F > 0.5$) and those that are unfavorable ($F < 0.5$) (Acevedo & Real, 2012; Real et al., 2006), and standardising values across distribution models to enable comparisons.

To avoid the binarization of modelling outputs, the overlap between predicted and actual virtual species distributions was assessed using Schoener's D-index, a metric well-suited for evaluating SDM outputs (Rödder & Engler, 2011). This index measures how the predicted favourability of each grid cell compares with the actual favourability of the corresponding grid cell in the virtual species distribution, resulting in a final value ranging from 0 (no overlap) to 1 (complete agreement).

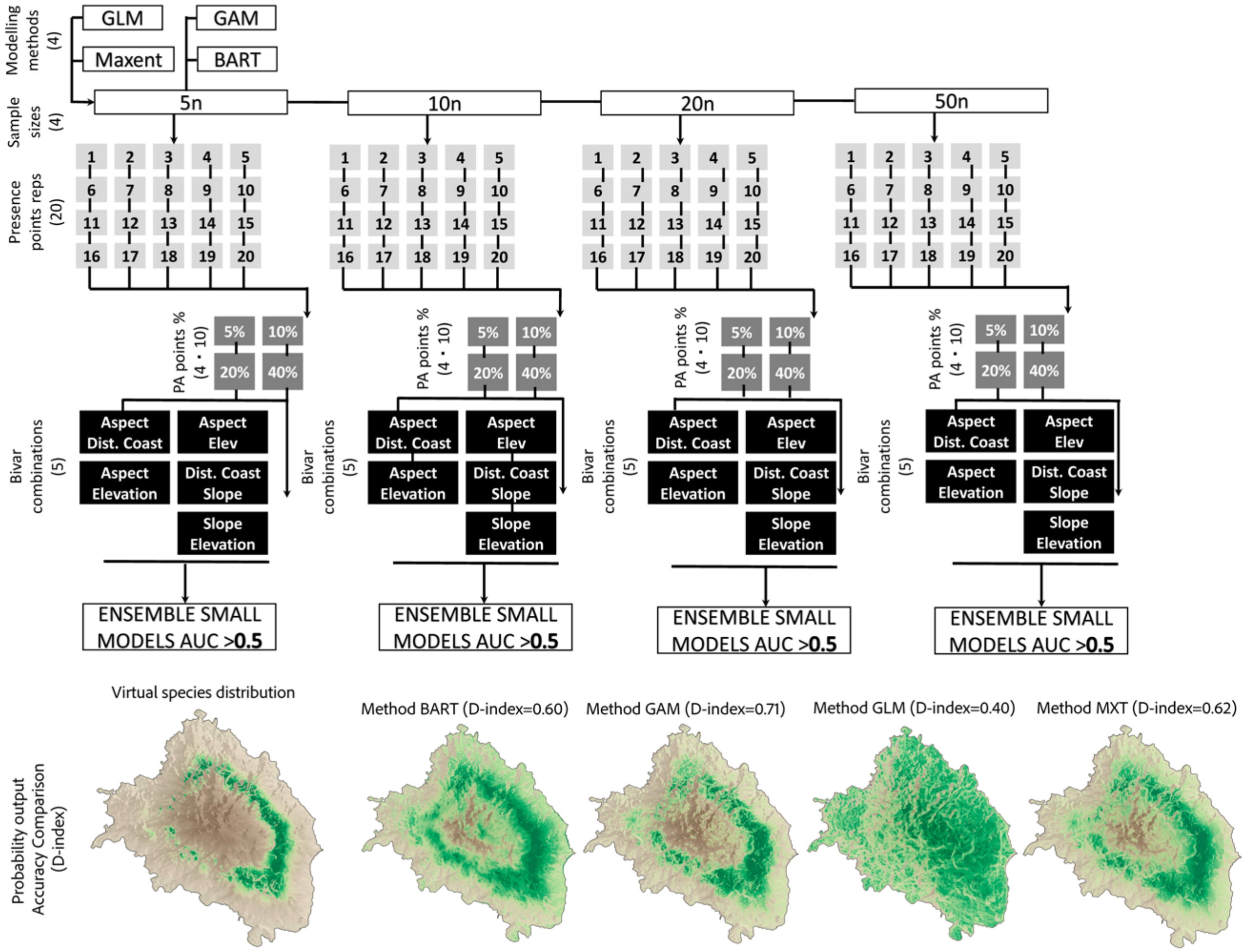


Figure 3.1 Modelling Workflow for the Analysis of Virtual Species. The final step (bottom row) displays, on the left, a probability distribution map for a narrow virtual species on Socorro Island (the largest island). The distribution is shaded in green, with increasing colour intensity representing higher occurrence probabilities. Brown shading depicts the island's topography, with hue intensity reflecting elevation. To the right of the virtual species map are examples of the four final outputs (one per method) for each species, along with their D-values, which compare the virtual "real" and predicted species distributions. The sample size for these final models was set to 10n.

To investigate the effects of different modelling conditions on model accuracy, a standard multiple linear regression was performed with model accuracy (measured by the D-index) as the response variable. The analysis included interaction terms for 'Sample Size,' 'Pseudo-absence/Background %,' 'Method,' and 'Range Size.'

Model accuracy was also evaluated using two common metrics: AUC and the Boyce index, both calculated with the 'modEVA' R package (Barbosa et al., 2013). In SDMs, AUC (Area Under the Curve) measures how well the model distinguishes between presence and absence locations, with values ranging from 0.5 (random) to 1 (perfect performance; Raes & ter Steege,2007). The Boyce index assesses the correlation between model predictions and the observed versus expected occurrence frequencies across prediction classes. It ranges from -1 to 1, where positive values indicate that presences are more frequent in areas with higher predicted suitability. Values near zero mean predictions are no better than random, while negative values suggest that presences occur more often in areas with lower predicted suitability (Boyce et al., 2002).

The purpose of assessing these validation metrics was to determine how accurately they measure actual model performance, as measured by the D-index (i.e., the overlap between real and predicted species distributions). A stratified Pearson’s correlation analysis was conducted between the D-index and the validation metrics (AUC and Boyce index). Instead of calculating a single correlation for the entire dataset, stratified correlation analysis divides the data into distinct groups based on relevant factors, and then calculates the correlation within each subgroup. The analysis was stratified according to subgroups defined by 'Method,' 'Sample Size,' and 'Pseudo-absence/Background %,' allowing for the assessment of how well these metrics reflect prediction accuracy across varying modelling conditions. This approach aimed to improve understanding of how changes in these variables influence the relationship between model accuracy and validation metrics, thus bridging the gap between predictive accuracy and validation reliability.

Real species analyses

For the real species data, a similar methodology and set of variables were used as in the virtual species regression models. Overlap (D-index) was assessed between SDMs fitted with presence-absence (PA) data, which serve as a proxy for the "real species distribution," and those using various pseudo-absence/background (Psa/BG) percentages. This evaluation aimed to understand how different modelling conditions affect model accuracy. Similar to the virtual species analysis, a standard multiple linear regression was performed, using model accuracy (measured by the D-index) as the response variable. The analysis included interaction terms for 'Pseudo-absence/Background %,' 'Method,' and 'Range Size.' However, in contrast to the virtual species linear regression, the 'Sample Size' interaction was excluded due to limited variation in sample sizes among the real species.

The impact of different modelling conditions on the reliability of validation metrics (AUC and Boyce) in real species SDMs was also assessed. Using the same approach as for virtual species, a stratified correlation analysis was conducted between the D-index and the validation metrics. This analysis was stratified by 'Pseudo-absence/Background %,' 'Range Size,' and 'Method.'

3.2.6 Interaction between sample size and range size for high-resolution predictors

As the real species occurrence datasets were fairly uniform in size, this analysis was conducted using only the virtual species data. First, the Psa/BG percentage that produced the highest SDM accuracy for each modelling method was identified, thereby removing the Psa/BG factor from further analysis (see section 3.2.5). Next, it was investigated whether the minimum required sample size varied across different methods and range sizes.

Following van Proosdij et al. (2015), the top 95% of Schoener's D values from the ESM replicates were extracted, excluding the lowest-performing 5%. Then two decision thresholds were applied to determine the minimum sample size required to achieve high overlap (Schoener's D >0.6) and very high overlap (>0.7) between real and predicted values. This approach aligns with the common practice of considering D index values >0.5 as indicative of better-than-random fits (Pirie et al., 2019; Püts et al., 2020).

Finally, the minimum sample size for each range size category (Narrow and Wide) was determined by identifying the sample size at which the lower bound of the 95th percentile of D-index values exceeded the thresholds of 0.6 and 0.7.

3.2.7 Impact of data precision on island SDM predictions at high resolutions

Virtual species

For this analysis, we resampled our virtual species presence/absence maps, obtained using the PA function from the 'Virtualspecies' R package in combination with our 30m resolution maps, to coarser resolutions: 90 metres, 500 metres, and 1 kilometre. We randomly sampled 20 replicates from each of the "small” occurrence datasets and "large” occurrence datasets from each resampled map.

As increasing resolution reduces the number of grids where a species is recorded as present, our "large" datasets contain either 50 occurrence points or, for species present in <50 grids, 80% of the resampled area where the species is present. For "small" datasets, we ensured a minimum of either 10 occurrence points or, for species present in <10 grids, 60% of the resampled area where the species is present (ensuring ≥5 occurrence points, our minimum sample size).

Unlike studies using random point displacements to simulate data uncertainty (e.g., Graham et al., 2008; Gabor et al., 2023), our method aligns with the spatial gridded organisation of occurrence points sourced from certain biodiversity databases (e.g., GBIF). In such cases, a species was recorded somewhere within the grid of a given extent, but the exact location is unknown. See Figure 3.2 for an illustration of the uncertainty simulation method.

Real species

The positioning of spatial inaccuracies for the real species presence records followed the same procedure used for the virtual species. Since the spatial accuracy of these records aligns with the study grain (with a margin of error of 25 metres from the GPS device), uncertainty was introduced by resampling the occurrence

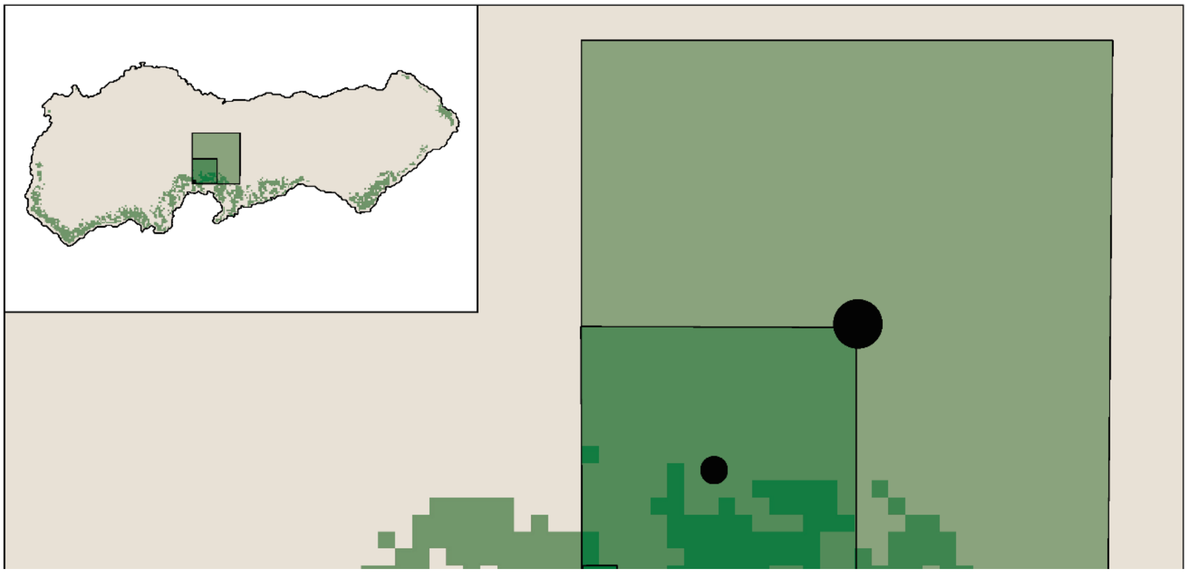


Figure 3. 2 Simulation of precision uncertainty in a gridded format at three resolutions on Clarion Island (full extent in top left corner). Brown shading represent the island's land. Species presence at 30m resolution is shown in green. Squares depict a progressively larger grid when the species presence/absence map is resampled at 90m, 500m, and 1km. The white dot represents the original presence point (at the centre of the 30m grid), while black points indicate increased uncertainty levels.

points to the centroids of raster cells at resolutions of 90 metres, 500 metres, and 1 kilometre.

3.3 Results

3.3.1 Presence and Pseudo-Absence Ratios across Modelling Methods: Impact on Model Accuracy and Validation

Model accuracy: virtual species

The hierarchical partitioning of explained variance of the predictors in our linear regressions showed that range size explained the majority of the variance in the D-index (overall model R^2=0.52; Variance explained: Range size = 95.15%, Sample size = 3.17%, Psa/BG percentage = 0.36%, method = 1.33%), therefore masking any effect of the variables of interest (Psa/BG ratio and its interactions with sample size and SDM method). As such, we partitioned the data into two datasets based on range size for the subsequent analysis. On the whole, accuracy was better for species with a wide range size (Intercept_wide = 0.85, intercept_narrow = 0.60).

For species with a narrow range, the full model containing Psa/BG ratio, sample size, SDM method and all interactions explained 13% of the variance in the D Index (R^2 = 0.13). There was a positive effect Psa/BG ratio and SDM method. Specifically, increasing the ratio of background points improves accuracy for all modelling approaches except for BART (Fig 3.3a). For GAM, this effect is not present for small sample sizes (n = 5). For GLM and Maxent, the effect is more pronounced (Fig. 3.3a, numerical results in Supplementary table 3.1.2).

For species with a wide range, the full model containing Psa/BG ratio, sample size, SDM method and all interactions explained 35% of the variance in the D Index (R^2 = 0.35). Here, increasing the number of background points had little effect on the accuracy of the models, with this being largely driven by sample size, particularly within the GAM approach (Fig. 3.3c, numerical results in Supplementary table 3.1.3).

Model accuracy: real species

The hierarchical partitioning results for the predictors from the real species regression model also showed that range size explained the majority of the variance in the D-index (overall model R^2 = 0.07; Range size = 44.35%, Psa/BG percentage= 27.97%, method = 28.26%); thus, we again partitioned the data into two datasets based on range size for the subsequent analysis. On the whole, accuracy was similar for species with both range sizes (Intercept_wide = 0.94, intercept_narrow = 0.90).

For species with a narrow range, the full model containing Psa/BG ratio, SDM method and their interaction explained 29% of the variance in the D Index (R^2 = 0.29). Increasing the percentage of background points did not improve predictive ability: most modelling approaches showed improved accuracy between 10-20% of

Psa/BG, except for GAM that showed clear improvements specifically at 10% of Psa/BG (Fig 3.3b, numerical results in Supplementary table 3.1.4).

For species with a wide range, the full model containing Psa/BG ratio, sample size, SDM method and all interactions explained 7% of the variance in the D Index (R^2 = 0.07). Here, background points at 20% Psa/BG resulted in lower accuracy, with minimal variation across the other percentages, except for BART, which showed clear improvements specifically at 20% Psa/BG (Fig. 3.3d; numerical results in Supplementary table 3.1.5).

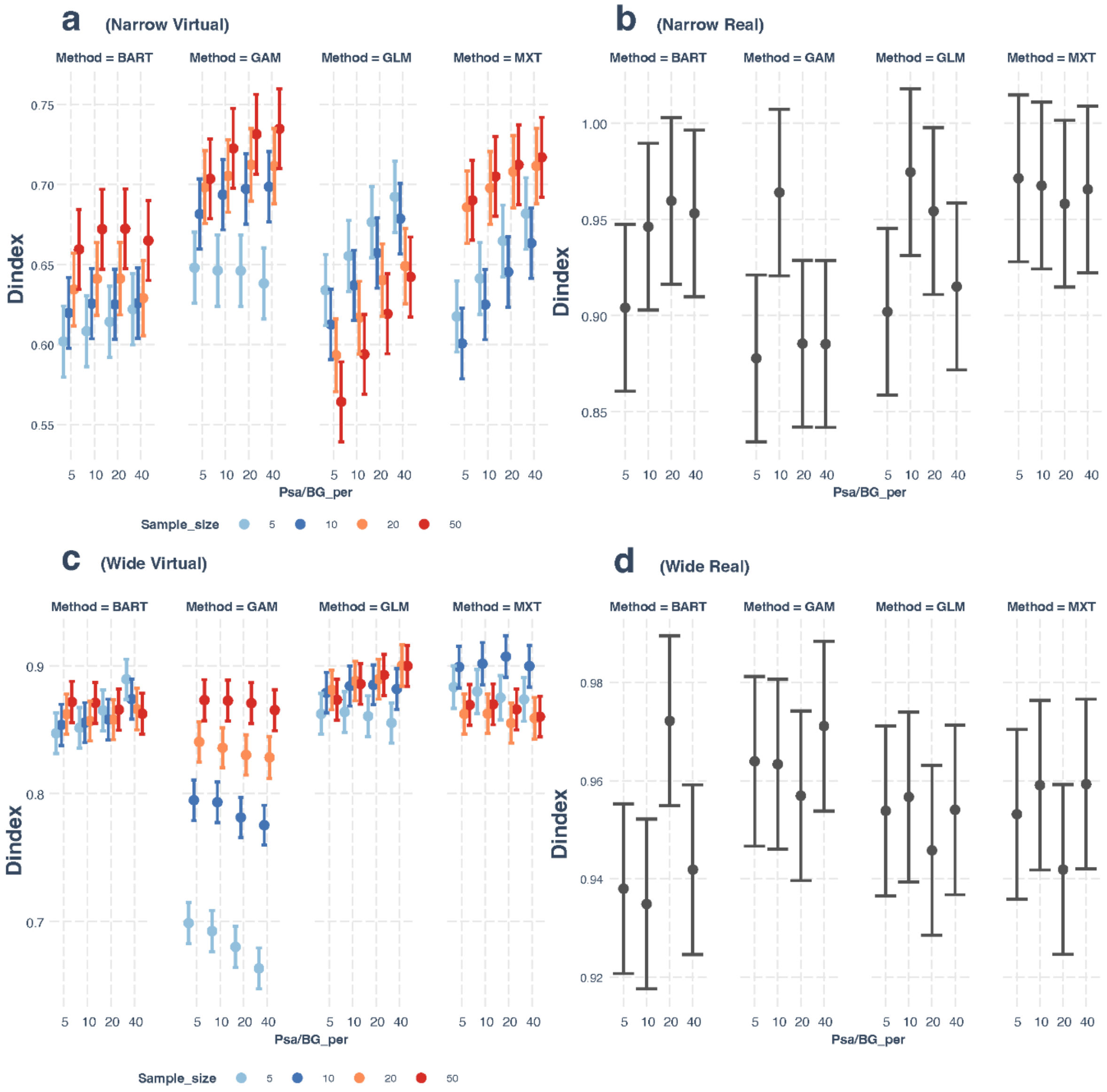


Figure 3. 3 Impact of Pseudo-Absence/Background (Psa/BG) percentage of the study area sampled on model accuracy (measured using the D-index), across different modeling methods. Figures 3a and 3c show the results of the virtual species analysis considering the ratio of sample sizes to Psa/BG percentage sampled. Figures 3b and 3d show the results of the real species analysis considering the Psa/BG percentage sampled (mean sample size of real species = 21). The circles represent regression model coefficient values, with the bars representing the standard error.

Model validation: virtual species

The correlation analysis between validation metrics (AUC and Boyce index) and model accuracy (D-index) reveals distinct patterns for wide-ranged and narrow-ranged species. Full results are shown in Figure 3.4a & 3.4b and Supplementary Table S3.1.6 &. S3.1.7. For wide-ranged species, the correlation between AUC and model accuracy (D-index) shows low variation across different Psa/BG percentages, with a consistently negative mean correlation (mean AUCcor = -0.67). The Boyce index correlation was also near zero (mean Boycecor = -0.09), though it tends to improve (i.e., shift from negative toward zero) as PsA/BG percentages increase.

For narrow-ranged species, the AUC correlation (mean AUCcor = 0.34) was generally positive and consistent across all methods, except for BART, which showed no correlation (mean BARTcor = 0.004). In BART, increasing PsA/BG percentages further reduced the correlations (mean BARTcor with PsA/BG 5% = 0.23; mean BARTcor with PsA/BG 40% = -0.28). For the other methods, the correlations changed only slightly across different Psa/BG percentages (Fig. 3.4a). The Boyce index correlations followed a similar pattern, with generally positive and consistent correlations across methods (mean Boycecor = 0.47). Excluding BART, both AUC and the Boyce index have positive correlations with the D-index, indicating better predictability for narrow-ranged species.

Model validation: real species

Full details of the described correlations are presented in Figure 3.4 b& d, and Supplementary Table S3.1.8 &. S3.1.9. The correlation analysis between validation metrics and species accuracy shows that for wide-ranged species, the correlation between the AUC from a pseudoabsence/background model and model accuracy (as measured by the D-index) remains consistently high across all methods, with only a slightly increased correlation as Psa/BG percentage increases. The mean AUC correlation across all scenarios is positive (mean AUCcor=0.86). For the Boyce index, correlations are positive (mean Boycecor=0.89), and tend to increase with higher Psa/BG percentages, with the highest correlation typically observed at 20% and 40%, particularly with GAM and BART. However, Maxent correlation values remain relatively low (mean MXTcor=0.40), regardless of Psa/BG percentage.

For narrow-ranged species, the AUC metric correlation increases with Psa/BG percentage across all methods, except at the 20% Psa/BG, which had the lowest correlation for BART, GLM, and MXT. Overall, the mean AUC correlation across all scenarios is positive (Mean AUCcor=0.86). The Boyce metric displays a similar pattern (Mean Boyceor=0.83).

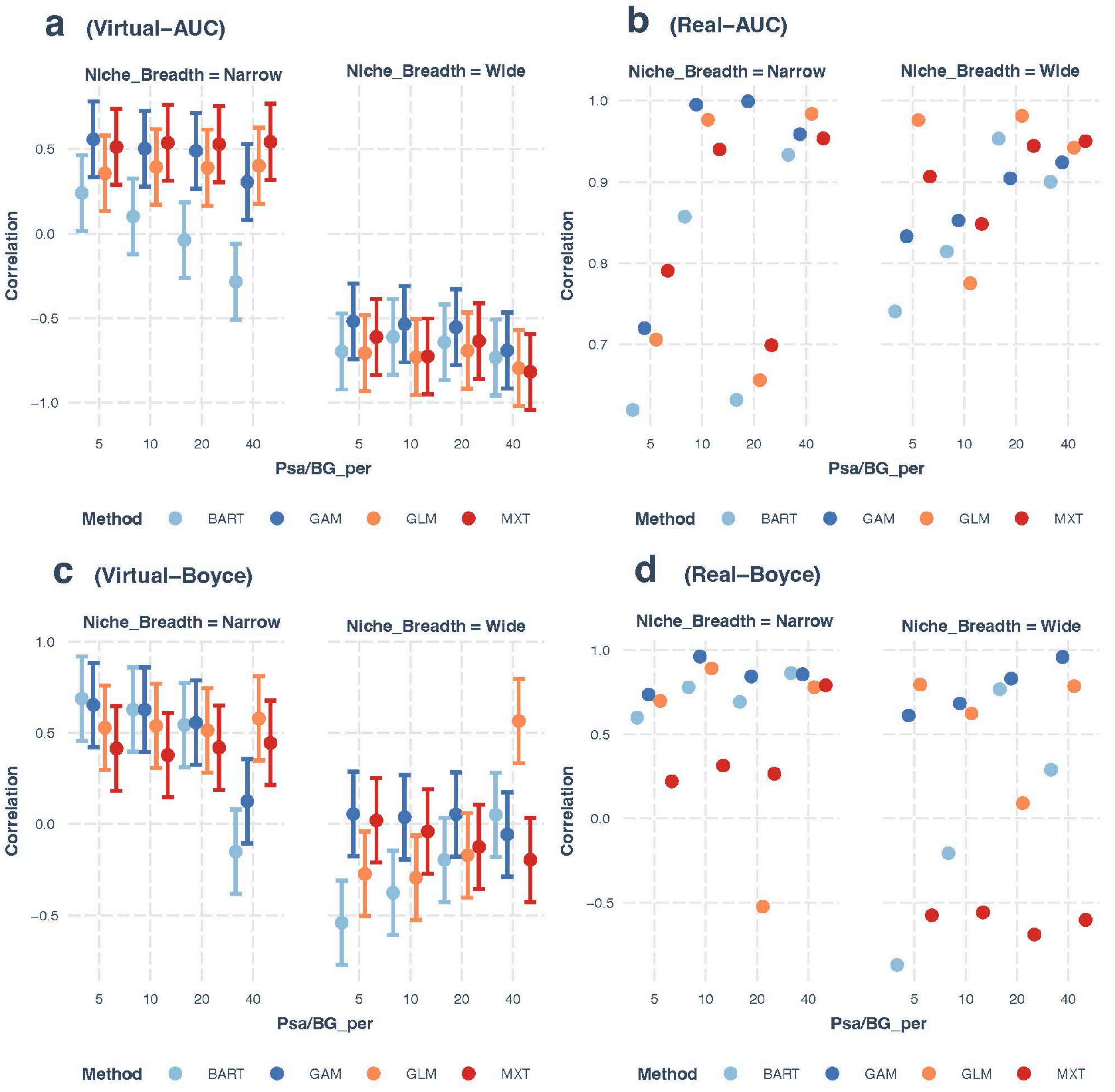


Figure 3.4 Results of the correlations between the AUC and Boyce index validation metrics and the actual accuracy, as measured by the D-index (i.e., real vs. predicted species distribution accuracy). This analysis assesses the reliability of validation metrics in reflecting actual model performance and explores whether this reliability varies across different Psa/BG percentages (x-axis) and modelling methods. Panels a and c depict the correlation between the D-index and validation metrics for virtual species; data points represent the correlation estimate, with error bars.

3.3.2 Minimum Sample Size Requirements Across Range Sizes for High-Resolution Predictors

Findings indicate that increasing the proportion of pseudo-absence/background (Psa/BG) area sampled generally improves prediction accuracy (D-index) for most species and methods (Fig 3.3). However, for models with small sample sizes fitted with GAM and GLM, this increase reduces accuracy (Fig 3.3a & c). Thus, we found 20% Psa/BG to be a good balance for exploring the minimum sample size needed for robust predictions across various range sizes and methods.

Two decision thresholds were set: one for high overlap (>0.6) between real vs. predicted species distribution and another for very high distribution overlap (>0.7). For virtual species with wide ranges, models consistently achieved D-index values exceeding 0.6 and 0.7 with a minimum sample size of 5, indicating that a significant portion of the distributional range can be predicted even with limited data. For narrow-ranged species, the required minimum sample size required to achieve a D-index > 0.6 varies by method: 20 for BART and GLM, 10 for Maxent, and 5 for GAM. Only GLM, with a sample size of 50, achieved a D-index > 0.7. For narrow-ranged island species, GAM and Maxent are thus the most suitable methods at small sample sizes, though larger datasets generally yield more robust predictions. See Figure 3.5 and Supplementary Table S3.1.10 for more detailed information.

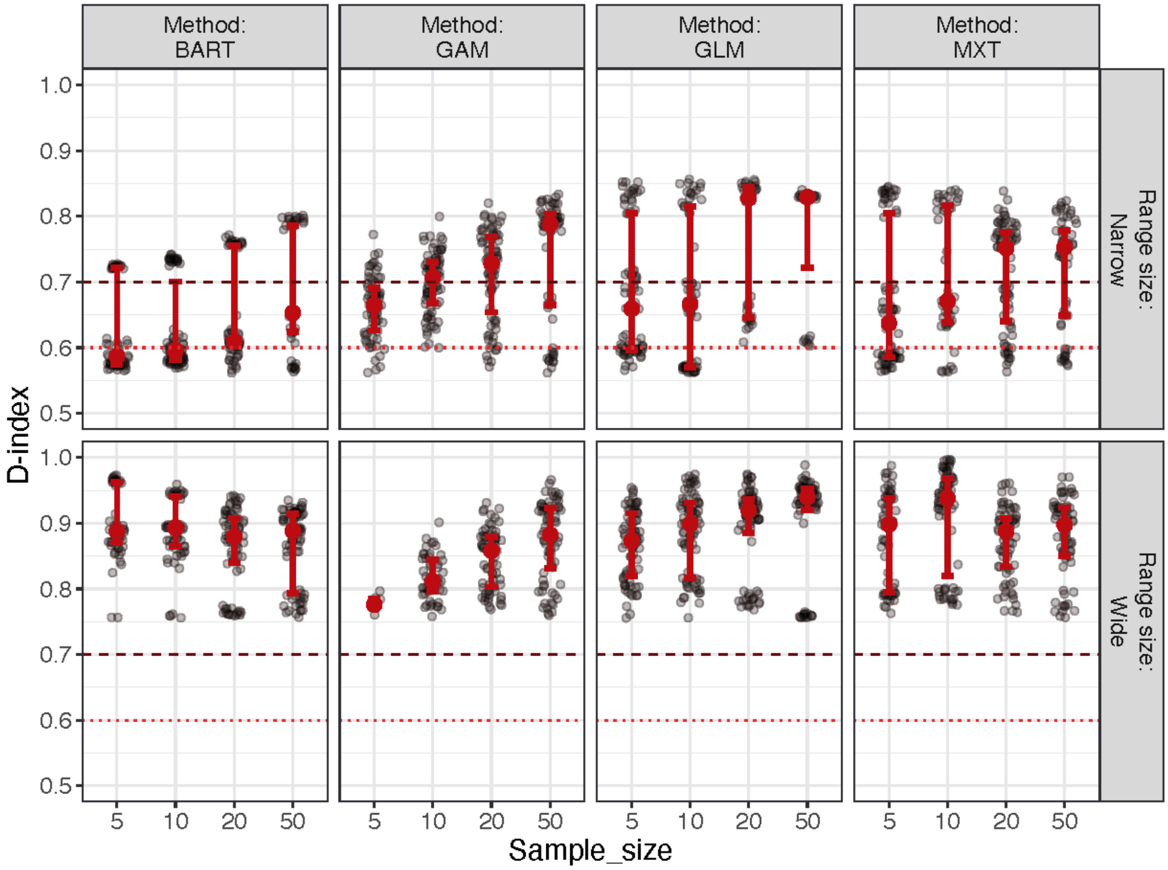


Figure 3.5 Minimum sample size needed for robust species predictions using the D index, and split by method and range size. Points represent the mean D index value, and error bars indicate the standard error. The red dotted line is the threshold for 'high niche overlap' (D-index>0.60), and the dark red dashed line for 'very high niche overlap' (D-index >0.70).

3.3.3 Impact of data precision on island SDM predictions at high resolutions

Virtual Species

As in the previous analysis, the focus was on models where 20% of the study area was sampled using the Psa/BG ratio, as this proportion provides a better balance in accuracy across all scenarios. The next step involved exploring the relationship between the predictor variables 'Method,' 'Range size,' 'Sample size,' and various levels of simulated gridded 'Uncertainty' (i.e. occurrence point displacement to the centroid of higher resolution grids: 90, 500 and 1000m), and the response variable of model accuracy, measured by the D-index. Similarly to the results presented above, hierarchical partitioning of explained variance showed that 'Range size' was the most important variable driving the predictability of the model (Model R^2 = 0.50; Range size' =96.68 %; Uncertainty 2.61%; Sample_size=0.29%; Method= 0.36%). Thus, the data was partitioned into narrow and wide-ranged species.

For species with narrow ranges, 'Method,' followed by 'Uncertainty' showed similar contribution to the variability in the D-index (Model R^2 = 0.1%; Method= 51.12%; Uncertainty 48.36%; Sample_size=0.34%). For species with wide ranges, the D-index is primarily explained by 'Method,' with ‘Uncertainty’ explaining only minimal variation (Model R^2 = 0.11; Method= 80.33%; Uncertainty 7.89%; Sample_size=11.95%). Specifically, the GAM and Maxent methods showed important declines in accuracy when sample sizes are small and the uncertainty is high (1000m) for both range sizes (figure 3.6 and Supplementary Table S3.1.11& S3.1.12 for numerical results).

Real Species analysis

In contrast to the virtual species results, 'Uncertainty' was the most significant variable explaining variation in model predictability (Model R^2 = 0.43; Uncertainty: 57.38%; Method: 21.89%; Range size: 20.75%) (see figure 3.6 and Supplementary

Table S3.1.13& S3.1.14 for numerical results). For narrow-ranged species, all methods showed large declines in accuracy with increasing uncertainty, particularly at 500-1000m. For wide-ranged species, the GAM method in particular showed declines in accuracy with increasing uncertainty, especially at 500-1000m.

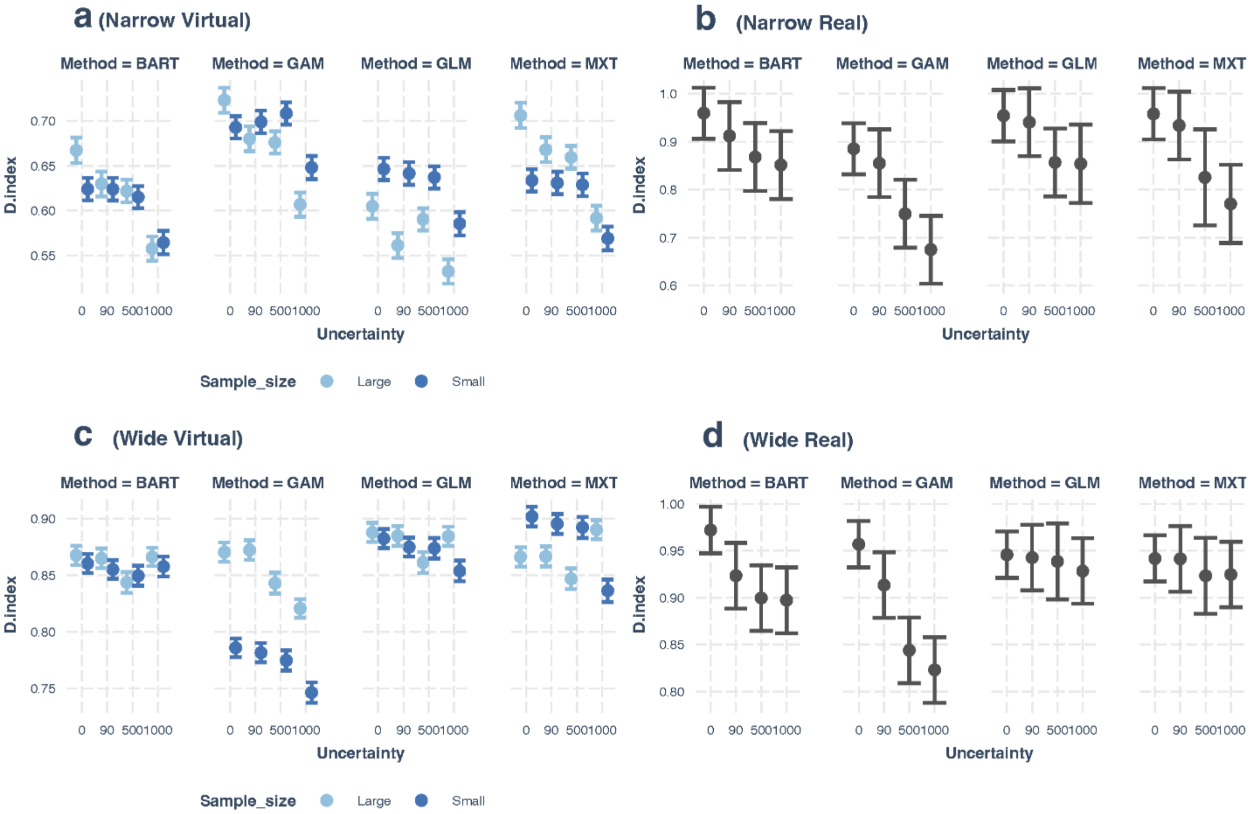


Figure 3.6 Impact of uncertainty in occurrence data on SDM accuracy (D-index). Panels 6a and 6c illustrate the effects of increasing gridded uncertainty levels (no uncertainty = 0; uncertainty = 90, 500, 1000 m) on wide and narrow ranged virtual species, considering large (50n or 80% of the species range) and small sample sizes (10n or 69% of the species range). Points represent the mean, and error bars indicate the standard error. Panels 6b and 6d present the corresponding results for real species (mean sample size = 21).

3.4 Discussion

This study explores the complexities of SDM application in island environments using plants as the focal taxon. The analyses presented highlight how challenges such as high habitat specialisation, spatial constraints, and sparse occurrence data hinder the understanding of island distribution patterns. The discussion addresses the suitability of different modelling techniques for islands, the optimal selection of pseudo-absence points, and the impact of positional uncertainty in occurrence data. Practical recommendations are provided to enhance the effectiveness of SDMs, even when working with limited and sparse data, to support conservation efforts on islands.

3.4.1 Data limitations and specialisation levels in island species: Optimising presence-only models and their validation

Several studies have investigated how species range size (i.e., prevalence in the study area) and sample size affect model accuracy (Stockwell & Peterson, 2002; Seoane et al., 2005; Hernandez et al., 2006; McPherson and Jetz, 2007; van Proosdij et al., 2016). However, the typical focus of such studies on continental areas, alongside their use of coarse-scale environmental/climatic predictors, means that their results are not necessarily generalisable to species from more restricted environments like islands.

The findings of this research partially support previous evidence (e.g., Hernandez et al., 2006; van Proosdij et al., 2016), confirming that sample size and range size influence prediction accuracy. However, a notable deviation was observed from the expected pattern that species with wide ranges require larger sample sizes and tend to be associated with less accurate predictions (see van Proosdij et al., 2016). Specifically, in the context of islands, species with wide ranges showed better predictability even with smaller sample sizes. This may be because larger study areas with more environmental variability require more data to accurately model widely distributed species, whereas smaller study areas with more constrained environmental variation, such as on many islands, allow for more accurate predictions of wide-ranged species using fewer data points. These results also provide new evidence on the challenges of modelling narrow-ranged species distributions. Narrow-ranged species often have limited data, complicating the use of SDMs (Lomba et al., 2010; Breiner et al., 2015). This task becomes even more difficult in island settings, as this study suggests that narrow-ranged species on islands may require relatively more data for accurate predictions than their continental counterparts (van Proosdij et al., 2016).

The validation metrics analysed (AUC and Boyce index) also proved highly sensitive to species' range size, often overestimating accuracy for narrow-range species and underestimating it for wide-range species. While It was found that increasing the proportion of Psa/BG often improve prediction accuracy and reduce uncertainty (Descombes et al., 2022; Valavi et al., 2022), analysis of both virtual and real species indicates that, depending on data availability and species’ range sizes, a trade-off may be required to bridge the gap between model validation and accuracy. In instances where the accuracy of spatial predictions is crucial (such as in conservation planning, where over- or under-predictions of species' ranges should be avoided), compromising some accuracy may be necessary to achieve realistic model validation by reducing the proportion of Psa/BG (see further details in Section 3.4.3).

Finally, the significant influence of range size on model performance emphasises the risk of using arbitrary thresholds to differentiate between high- and low-performance models (e.g., >0.7 for AUC and >0.4 for Boyce) in multi-species studies with varying range sizes (Raes & ter Steege,2007). To complement traditional validation techniques, testing SDM fits against null models is recommended to distinguish meaningful results from random outcomes, rather than subjectively interpreting them as a continuum of 'high' and 'low' performance (Hu & Liu, 2014; Osborne et al., 2022; see also Chapter 4).

3.4.2 The impact of low occurrence data precision on the accuracy of island SDM predictions at high-resolutions

Previous studies have shown that positional error can significantly reduce the predictive performance of species distribution models (SDMs) (Hefley et al., 2014; Tulowiecki et al., 2015; Mitchel et al., 2017; Soultan & Safi, 2017; Fernandes et al., 2019; Gábor et al., 2022). This effect is particularly strong for specialist species (Visscher, 2006; Gábor et al., 2020) and species in environmentally heterogeneous regions (Gábor et al., 2023). These factors are especially relevant for island environments because: 1) island terrain, especially on oceanic islands, can vary significantly over small areas, creating sharp environmental gradients (Hanz et al., 2023), and 2) islands floras often comprise a high proportion of narrow-ranged specialist species (Kreft et al., 2008). Based on these points, SDM performance was expected to decline with positional inaccuracies, particularly for species in topographically diverse areas and those with narrow ranges. In this study, species with broader ranges typically inhabit more diverse topographies on islands, while narrow-ranged species are often confined to less variable environments (e.g., coastal zones or specific elevation bands). The results suggest that widespread species are less affected by occupancy point positional errors, likely because these errors often fall within their suitable habitat range. In contrast, highly specialised species, even in less heterogeneous environments, are more vulnerable to positional errors when they occur in rare or restricted habitat patches, as a large uncertainty (e.g., ≥1km) can place occurrence points far outside of suitable habitats.

Our results suggest that, to mitigate the impact of positional error in SDMs in island environments, it is essential to consider species' range size. Furthermore, as we observed little difference in prediction accuracy between large and small sample sizes, data quality (i.e., uncertainty) seems to be more important than data quantity (in accordance with Engler at al., 2004). Therefore, the removal of highly uncertain records —based on the previously defined data requirements for each species' range sizes— is justified on the basis of these analyses.

3.4.3 Optimal methods for island SDMs considering data availability and uncertainty

The analyses identify effective methods for future island SDM studies. GAMs stand out as the most effective modelling approach, on analyses of both virtual and real species, showing lower sensitivity to range size, strong coherence with validation metrics (the Boyce index and AUC), and relatively low data requirements (≥10 records for narrow range species resulted in high model accuracy). However, this method is highly sensitive to occurrence records with spatial imprecision. Interestingly, Maxent, often considered the strong alternative for occurrence data and range-restricted scenarios in both island and continental studies (Pearson et al., 2007; Valavi et al., 2021), was found to have limitations in our study, with wide-ranged species consistently having low prediction accuracy. Paradoxically, it was also found that Maxent tolerates occurrence records with spatial imprecision, at least for this group of species. The BART approach generated satisfactory results overall, requiring moderately large datasets for robust predictions (≥20 occurrence records), but tolerating occurrence records with spatial imprecision to a greater degree, and providing robust SDM fits in island environments under appropriate pseudoabsence sampling conditions (Table 3.2). GLMs demonstrated high accuracy (under appropriate modelling conditions: Table 3.2) and tolerance to uncertainty, yet also required moderately large datasets (≥20 occurrence records) to provide satisfactory fits and predictions.

Table 3.2 Summary of the minimum sample sizes, background/pseudo-absence (Psa/BG) percentages, and the effect of spatial uncertainty (i.e., geographical occurrence inaccuracies) on model accuracy for island species with narrow and wide ranges across the four modelling methods: BART, GAM, GLM, and Maxent (MXT). (*) recommended percentage for multi-method or multi-species studies.

	Narrow-ranged			**Wide-ranged**		
	Minimum Sample size	**Psa/BG %**	**SDM Accuracy With Occurrence Uncertainty**	**Minimum Sample size**	**Psa/BG %**	**SDM Accuracy With Occurrence Uncertainty**
BART	20	10	Affected by large inaccuracies >1000m	10	10*/20	Low negative impact from inaccuracies

GAM	10	10	Affected by large inaccuracies >1000m	10	10	Affected by large inaccuracies >1000m
GLM	20	10	Affected by large inaccuracies >1000m	10	10	Low negative impact from inaccuracies
MXT	10	10*/40	Affected by large inaccuracies >1000m	10	10*/40	Low negative impact from inaccuracies

3.5 Conclusion

This study has illustrated the importance of spatial context and data availability in Species Distribution Modelling (SDM), offering valuable guidance for researchers aiming to enhance predictive accuracy for species in island environments. Given that island species often have narrow ranges, high specialisation, and limited occurrence data, particularly in the form of presence-only records, selecting appropriate modelling techniques and pseudoabsence sampling strategies while accounting for range size and sample size is crucial for robust island SDM implementation.

3.6 References

Acevedo, P., & Real, R. (2012). Favourability: concept, distinctive characteristics and potential usefulness. *Naturwissenschaften*, *99*, 515-522.

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

Barajas-Barbosa, M. P., Weigelt, P., Borregaard, M. K., Keppel, G., & Kreft, H. (2020). Environmental heterogeneity dynamics drive plant diversity on oceanic islands. *Journal of Biogeography*, *47*(10), 2248-2260.

Barbosa, A. M. (2015). fuzzySim: applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution*, *6*(7), 853-858.

Barbosa AM, Real R, Munoz AR, Brown JA (2013). New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions*, *19*(10)*,* 1333–1338.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*(2), 327-338.

Benavides, E., Kuethe, J. R., Ortiz-Alcaraz, A., & De La Luz, J. L. (2019). Oenothera resicum (Onagraceae), a new species and the first record of the family from the Revillagigedo Archipelago, Mexico. *Phytotaxa*, *416*(1), 59-66.

Benavides, E., Sadler, J., Graham, L., & Matthews, T. J. (2024). Species distribution models and island biogeography: Challenges and prospects. *Global Ecology and Conservation*, e02943.

Borges, P. A., Cardoso, P., Kreft, H., Whittaker, R. J., Fattorini, S., Emerson, B. C., ... & Gabriel, R. (2018). Global Island Monitoring Scheme (GIMS): a proposal for the long-term coordinated survey and monitoring of native island forest biota. *Biodiversity and Conservation*, *27*, 2567-2586.

Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, *275,* 73-77.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. (2002). Evaluating resource selection functions. *Ecological Modelling*, *157*(2-3), 281-300.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, *6*(10), 1210-1218.

Breiner, F. T., Nobis, M. P., Bergamini, A., & Guisan, A. (2018). Optimizing ensembles of small models for predicting the distribution of species with few occurrences. *Methods in Ecology and Evolution*, *9*(4), 802-808.

Caujape-Castells, J., Tye, A., Crawford, D. J., Santos-Guerra, A., Sakai, A., Beaver, K., ... & Kueffer, C. (2010). Conservation of oceanic island floras: present and future global challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(2), 107-129.

Carlson, C. J. (2020). embarcadero: Species distribution modelling with Bayesian additive regression trees in r. *Methods in Ecology and Evolution*, *11*(7), 850-858.

Chiatante, G. (2022). Spatial distribution of an assemblage of an endemic genus of birds: an example from Madagascar. *African Journal of Ecology*, *60*(1), 13-26.

Cursach, J., Far, A. J., & Ruiz, M. (2020). Geospatial analysis to assess distribution patterns and predictive models for endangered plant species to support management decisions: a case study in the Balearic Islands. *Biodiversity and Conservation*, *29*(11), 3393-3410.

Deblauwe, V., Droissart, V., Bose, R., Sonké, B., Blach-Overgaard, A., Svenning, J. C., ... & Couvreur, T. L. P. (2016). Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. *Global Ecology and Biogeography, 25*(4), 443-454.

Descombes, P., Chauvier, Y., Brun, P., Righetti, D., Wüest, R. O., Karger, D. N., ... & Zimmermann, N. E. (2022). Strategies for sampling pseudo-absences for species distribution models in complex mountainous terrain. *BioRxiv*, 2022-03.

Engler, R., Guisan, A., & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, *41*(2), 263-274.

Fernandes, R. F., Scherrer, D., & Guisan, A. (2019). Effects of simulated observation errors on the performance of species distribution models. *Diversity and Distributions*, *25*(3), 400-413.

Fernandez, M., Blum, S., Reichle, S., Guo, Q., Holzman, B., & Hamilton, H. (2009). Locality uncertainty and the differential performance of four common niche-based modeling techniques. *Biodiversity Informatics*, *6*, 36-52.

Fernández-Palacios, J. M., Kreft, H., Irl, S. D., Norder, S., Ah-Peng, C., Borges, P. A., ... & Drake, D. R. (2021). Scientists' warning–The outstanding biodiversity of islands is in peril. *Global Ecology and Conservation*, *31*, e01847.

Ferrier, S., Watson, G., Pearce, J. & Drielsma, M. (2002). Extended statistical approaches to modelling spatial patterns in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodiversity and Conservation*, *11*(12), 2275– 2307.

Fife, D. (2021). Flexplot: Graphically-Based Data Analysis. *Psychological Methods*, *27*(4), 477-496. https://doi.org/10.1037/met0000424

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

Franklin, J. (2010). Mapping species distributions: spatial inference and prediction. Cambridge University Press, Cambridge, UK.

Gábor, L., Jetz, W., Zarzo-Arias, A., Winner, K., Yanco, S., Pinkert, S., ... & Moudrý, V. (2023). Species distribution models affected by positional uncertainty in species occurrences can still be ecologically interpretable. *Ecography*, e06358.

Gardner, A. S., Maclean, I. M., & Gaston, K. J. (2019). Climatic predictors of species distributions neglect biophysiologically meaningful variables. *Diversity and Distributions*, *25*(8), 1318-1333.

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

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., ... & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, *24*(3), 276-292

Guisan, A., Graham, C. H., Elith, J., Huettmann, F., & NCEAS Species Distribution Modelling Group. (2007). Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, *13*(3), 332-340.

Graham, C. H., Elith, J., Hijmans, R. J., Guisan, A., Townsend Peterson, A., Loiselle, B. A., & NCEAS Predicting Species Distributions Working Group. (2008). The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, *45*(1), 239-247.

Grimmett, L., Whitsed, R., & Horta, A. (2020). Presence-only species distribution models are sensitive to sample prevalence: Evaluating models using spatial prediction stability and accuracy metrics. *Ecological Modelling*, *431*, 109194.

Hanz, D. M., Cutts, V., Barajas-Barbosa, M. P., Algar, A., Beierkuhnlein, C., Collart, F., ... & Irl, S. D. (2023). Effects of climate change on the distribution of plant species and plant functional strategies on the Canary Islands. *Diversity and Distributions*, *29*(9), 1157-1171.

Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science*, *1*, 297–310.

Hefley, T. J., Baasch, D. M., Tyre, A. J., & Blankenship, E. E. (2014). Correction of location errors for presence-only species distribution models. *Methods in Ecology and Evolution*, *5*(3), 207-214.

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

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

Hijmans, R. J., Bivand, R., Forner, K., Ooms, J., Pebesma, E., & Sumner, M. D. (2022). Package ‘terra’. R package. Maintainer: Vienna, Austria.

Hirzel, A. H., Helfer, V., & Metral, F. (2001). Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, *145*(2-3), 111-121.

Hu, J., & Liu, Y. (2014). Unveiling the conservation biogeography of a data-deficient endangered bird species under climate change. *PLoS One*, *9*(1), e84529.

Irl, S. D., Obermeier, A., Beierkuhnlein, C., & Steinbauer, M. J. (2020). Climate controls plant life-form patterns on a high-elevation oceanic island. *Journal of Biogeography*, *47*(10), 2261-2273.

Kier G., H.r Kreft, T. M. Lee, W. Jetz, P. L. Ibisch, C. Nowicki, J. Mutke, & W. Barthlott. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences*, *106*(23), 9322-9327.

Kreft, H., Jetz, W., Mutke, J., Kier, G., & Barthlott, W. (2008). Global diversity of island floras from a macroecological perspective. Ecology letters, 11(2), 116-127.

Lannuzel, G., Balmot, J., Dubos, N., Thibault, M., & Fogliani, B. (2021). High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia. *Biodiversity and Conservation*, *30*(4), 963-990.

Leroy, B. (2022). Choosing presence-only species distribution models. *Journal of Biogeography, 50*(1), 247-250.

Leroy, B., Meynard, C. N., Bellard, C., & Courchamp, F. (2016). virtualspecies, an R package to generate virtual species distributions. *Ecography*, *39*(6), 599-607.

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.

Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, *143*(11), 2647-2657.

Miller, J. A. (2014). Virtual species distribution models: using simulated data to evaluate aspects of model performance. *Progress in Physical Geography*, *38*(1), 117-128.

Matthews, T. J., Wayman, J. P., Cardoso, P., Sayol, F., Hume, J. P., Ulrich, W., ... & Triantis, K. A. (2022). Threatened and extinct island endemic birds of the world: Distribution, threats and functional diversity. *Journal of Biogeography*, *49*(11), 1920-1940.

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

McPherson, J., & Jetz, W. (2007). Effects of species' ecology on the accuracy of distribution models. *Ecography*, *30*(1), 135-151.

Meynard, C. N., Leroy, B., & Kaplan, D. M. (2019). Testing methods in species distribution modelling using virtual species: what have we learnt and what are we missing?. *Ecography*, *42*(12), 2021-2036.

Mitchell, P. J., Monk, J., & Laurenson, L. (2017). Sensitivity of fine-scale species distribution models to locational uncertainty in occurrence data across multiple sample sizes. *Methods in Ecology and Evolution*, *8*(1), 12-21.

Mondanaro, A., Di Febbraro, M., Castiglione, S., Melchionna, M., Serio, C., Girardi, G., ... & Raia, P. (2023). ENphylo: A new method to model the distribution of extremely rare species. *Methods in Ecology and Evolution*, *14*(3), 911-922.

Moudrý, V., & Šímová, P. (2012). Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Information Science*, *26*(11), 2083-2095.

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

NASA/METI/AIST/Japan Space Systems, and U.S./Japan ASTER Science Team. ASTER global digital elevation model V003, 2018, distributed by NASA EOSDIS Land Processes DAAC, https://doi.org/ 10.5067/ASTER/ASTGTM.

Nogué, S., Santos, A. M., Birks, H. J. B., Björck, S., Castilla- Beltrán, A., Connor, S., ... & Steinbauer, M. J. (2021). The human dimension of biodiversity changes on islands. *Science, 372*(6541), 488-491.

Osborne, O. G., Fell, H. G., Atkins, H., van Tol, J., Phillips, D., Herrera-Alsina, L., ... & Algar, A. C. (2022). Fauxcurrence: simulating multi-species occurrences for null models in species distribution modelling and biogeography. *Ecography*, *2022*(7), e05880.

Patiño, J., Collart, F., Vanderpoorten, A., Martin-Esquivel, J. L., Naranjo-Cigala, A., Mirolo, S., & Karger, D. N. (2023). Spatial resolution impacts projected plant responses to climate change on topographically complex islands. *Diversity and Distributions*, *29*(10), 1245-1262.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, *12*(5), 361-371.

Pearson, R. G., Raxworthy, C. J., Nakamura, M., & Townsend Peterson, A. (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, *34*(1), 102-117.

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

Pirie, M. D., Kandziora, M., Nürk, N. M., Le Maitre, N. C., Mugrabi de Kuppler, A., Gehrke, B., ... & Bellstedt, D. U. (2019). Leaps and bounds: geographical and ecological distance constrained the colonisation of the Afrotemperate by Erica. *BMC Evolutionary Biology*, *19*, 1-12.

Püts, M., Taylor, M., Núñez-Riboni, I., Steenbeek, J., Stäbler, M., Möllmann, C., & Kempf, A. (2020). Insights on integrating habitat preferences in process-oriented ecological models–a case study of the southern North Sea. *Ecological Modelling*, *431*, 109189.

Radomski, T., Beamer, D., Babineau, A., Wilson, C., Pechmann, J., & Kozak, K. H. (2022). Finding what you don’t know: testing SDM methods for poorly known species. *Diversity and Distributions*, *28*(9), 1769-1780.

Real, R., Barbosa, A. M., & Vargas, J. M. (2006). Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, *13*, 237-245.

Raes, N., & ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, *30*(5), 727-736.

Rödder, D., & Engler, J. O. (2011). Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. *Global Ecology and Biogeography*, *20*(6), 915-927.

Santika, T. (2011). Assessing the effect of prevalence on the predictive performance of species distribution models using simulated data. *Global Ecology and Biogeography*, *20*(1), 181-192.

Segal, R. D., Massaro, M., Carlile, N., & Whitsed, R. (2021). Small-scale species distribution model identifies restricted breeding habitat for an endemic island bird. *Animal Conservation*, *24*(6), 959-969.

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

Soultan, A., & Safi, K. (2017). The interplay of various sources of noise on reliability of species distribution models hinges on ecological specialisation. *PloS One,* 12(11), e0187906.

Stokland, J. N., Halvorsen, R., & Støa, B. (2011). Species distribution modelling—Effect of design and sample size of pseudo-absence observations. *Ecological Modelling, 222*(11), 1800-1809.

Stockwell, D. R., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, *148*(1), 1-13.

Tulowiecki, S. J., Larsen, C. P., & Wang, Y. C. (2015). Effects of positional error on modeling species distributions: a perspective using presettlement land survey records. *Plant Ecology*, *216*, 67-85.

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs*, *92*(1), e01486.

van Proosdij, A. S., Sosef, M. S., Wieringa, J. J., & Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542-552.

R. Visscher, D. (2006). GPS measurement error and resource selection functions in a fragmented landscape. *Ecography*, *29*(3), 458-464.

Warton, D. I., & Hui, F. K. (2011). The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, *92*(1), 3-10.

Weigelt, P., Jetz, W., & Kreft, H. (2013). Bioclimatic and physical characterization of the world's islands. *Proceedings of the National Academy of Sciences*, *110*(38), 15307-15312.

Whittaker, R. J., Fernández-Palacios, J. M., Matthews, T. J., Borregaard, M. K., & Triantis, K. A. (2017). Island biogeography: Taking the long view of nature's laboratories. *Science, 357*(6354).

Wisz, M. S., & Guisan, A. (2009). Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, *9*, 1-13.

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

4 . Assessing the ability of gridded microclimatic data to improve species distribution model applications in tropical island regions

Abstract

Biodiversity hotspots, areas of key focus for global conservation, are characterised by high levels of plant endemism and significant vegetation loss. Nearly half of these hotspots are mainly comprised of islands, most of which are in tropical regions, making the understanding of island plant species distributions in the tropics crucial for conservation. However, knowledge gaps persist, particularly in tropical islands, where Species Distribution Models (SDMs) can play a crucial role in predicting species' habitat requirements and potential range shifts in response to global change. Applying SDMs in tropical island environments presents challenges, mainly due to the uncertainty associated with global climatic datasets.

To overcome these challenges, this study evaluates the effectiveness of two downscaled global climatic datasets (WorldClim and CHELSA) and a non-climatic alternative—terrain-derived variables that may serve as a climatic proxy—in predicting plant species distributions in the Revillagigedo Archipelago, a tropical island biodiversity hotspot. Model accuracy and transferability were tested across various environmental baselines and spatial resolutions (30, 90, and 500 m) to examine their impact on predictive performance.

The contrasting size and topographic complexity of the two islands studied—Clarion (smaller island) and Socorro (larger island)—provided valuable insights for

improving the accuracy and transferability of SDMs across varied island landscapes. Results indicate that models utilising downscaled climatic data at finer resolutions (≤90 m) significantly enhance predictive accuracy for smaller islands, as evidenced by the results from Clarion Island. The Worldclim dataset consistently achieved both high accuracy and transferability across a wide range of contexts. Notably, model performance varied across species, with three species consistently yielding poor predictions under all tested conditions.

The analysed approaches represent promising alternatives to coarse-scale climate data for use in tropical island SDMs, and a method is presented to evaluate the informativeness of high-resolution environmental data for SDMs in these settings. The results of the study also provide practical guidance to enhance SDM accuracy and transferability, ultimately supporting more effective conservation efforts in tropical island ecosystems.

4.1 Introduction

Biodiversity hotspots, globally recognised as crucial for conservation efforts, are characterised by high levels of plant endemism and significant vegetation losses (Myers et al., 2000; Mittermeier et al., 2011). The use of plants as the basis of the biodiversity hotspots scheme underscores the essential contribution of plants to maintaining the health and resilience of ecosystems and their role as indicators for biodiversity conservation (Stork & Habel, 2014).

Notably, nearly a half of the recognised 36 hotspots are comprised primarily of islands, with a large majority (10/15= 67%) located at tropical latitudes (Myers et al., 2000; Mittermeier et al., 2011). Understanding the distribution of plant species in these tropical insular regions and the impact of environmental changes on their distributions is vital for broader conservation efforts. Despite increased global investment in biodiversity research, knowledge gaps persist, particularly in tropical island regions (Whittaker et al., 2005; Benavides et al., 2024). Species Distribution Models (SDMs) serve as valuable tools for filling these knowledge gaps, as well as predicting range shifts in response to global change (Guisan et al., 2013; Sofaer et al., 2019). In island plant conservation for example, SDMs are crucial for predicting climate change impacts, understanding habitat modification, managing invasive species, and designing effective conservation plans (Vieilledent et al., 2013; Rodríguez-Rodríguez et al., 2018; Patiño et al., 2023).

While SDMs are useful tools, their application in islands faces challenges because of limitations associated with commonly used environmental descriptors that are derived from global climatic datasets (Benavides et al., 2024). This issue is especially pronounced in tropical island environments, where the precision of global climate datasets is often questioned due to the sparse distribution of meteorological stations in tropical latitudes, which can lead to potential biases from statistical interpolation methods (Soria-Auza et al., 2010; Deblauwe et al., 2016).

Moreover, the most widely used global climatic datasets (e.g., WorldClim) are only available at resolutions that may inadequately capture the fine scale distributions of island plants (≥1km). For many such species, information on fine scale microclimatic variations is essential for understanding their distribution patterns (Lannuzel et al., 2021; Whittaker et al., 2023), and neglecting these factors can lead to misinterpretations of climatic niche requirements, potentially resulting in misleading spatial predictions (Patiño et al., 2023).

To address these challenges, researchers have explored alternatives to global climatic datasets for use with SDMs, such as using high-resolution land cover and surface variables as proxies for microclimatic variation (Heinänen et al., 2012; Turvey et al., 2020; Lannuzel et al., 2021; Segal et al., 2021; Gabor et al., 2023), and downscaling coarse-resolution global climatic datasets through statistical computation to obtain high-resolution climatic grids (Bazzichetto et al., 2021;

Patiño et al., 2023). However, research is needed to evaluate the effectiveness of available approaches for obtaining meaningful microclimatic/environmental data to improve SDM implementation in tropical island environments.

The effectiveness of different SDM approaches can be assessed by both model accuracy and transferability. The former relates to the ability of a given model to fit the data at hand (i.e., training data). Transferability in contrast relates to the ability of a model developed in one set of conditions to accurately predict species distributions in different, novel environments, and is a critical but often overlooked consideration in island SDM research. Transferability is influenced by both the choice of predictor variables (Goedecke et al., 2020; Datta et al., 2020) and the study grain (Manzoor et al., 2018). However, there is a recognised gap in our understanding regarding how the choice of environmental data source affects model transferability (Austin & Van Niel, 2011a, 2011b; Stanton et al., 2012, Datta et al., 2020).

Different predictor variable types can affect model accuracy and transferability in different ways. For example, while SDMs fitted using terrain variables (i.e., those derived from elevation models) are known to effectively model the current distribution of species, they show reduced transferability to new environments (Rosseau & Betts, 2022). On the other hand, the use of climatic variables derived from global datasets often enhances model extrapolation across space and time due to the availability of (i) data at global scales, and (ii) a range of temporal climatic scenarios (Petitpierre et al., 2017). However, their use may lead to higher uncertainty in predictions in tropical environments due to the previously mentioned issues related to the sparseness of meteorological data in these regions (Soria-Auza et al., 2010; Deblauwe et al., 2016), and the issue may be exacerbated when interpolated to finer scales. Regarding the study grain, past studies have shown that, as the analysed grid size decreases, models may face challenges such as overfitting, diminishing their transferability to new environmental contexts (Manzoor et al., 2018).

In this study, we investigate the effectiveness of two high-resolution downscaled bioclimatic global baselines for predicting species distributions in a tropical insular hotspot (Revillagigedo Archipelago) lacking regional climatic data. We use an interpolation approach that considers topographic and elevational effects on climate (Flint & Flint, 2012) and is known to be useful in characterising microclimatic conditions (Franklin et al., 2012; Khosravi et al., 2016). The selected global climatic baselines are WorldClim, widely used in SDM studies, and CHELSA (Karger et al., 2017), which has been shown to generate useful outcomes in applications focused on remote islands (Bazzichetto et al., 2021; Hanz et al., 2023). Additionally, we examine models built with terrain variables derived from a global digital elevation model, a widely implemented alternative to climatic descriptors in island SDM studies (Heinänen et al., 2012; Turvey et al., 2020; Lannuzel et al., 2021). We assess the performance of SDMs calibrated with the three aforementioned distinct environmental/climatic baselines (Downscaled-WorldClim, Downscaled-CHELSA, Terrain-derived) across a range of high-resolution study grains (30, 90, and 500 m). As outlined above, we measure effectiveness by focusing on both model accuracy and transferability.

Our study addresses three key questions:

I) How does the prediction accuracy of SDMs differ when using various high-resolution environmental data sources (Downscaled Bioclimatic and Terrain variables) to predict island plant species distributions at tropical latitudes?

II) Does the grain size of downscaled climatic estimates impact the transferability of island SDMs to new environments?

III) Does the choice of variable type (Downscaled Bioclimatic vs. Terrain variables) influence the transferability of island SDMs to novel environments?

By addressing these questions, our study aims to investigate the predictive accuracy related to the use of environmental and climatic predictors relevant to island floras, with the goal of improving our understanding of species distributions in understudied island regions. Understanding and addressing uncertainty in environmental descriptors in regions with limited data availability is crucial for advancing knowledge of species' responses to global change. Thus, the insights gained from our analyses will enhance the applicability and transferability of SDMs in island environments, a critical endeavour for strengthening conservation efforts in these unique and vulnerable ecosystems (Benavides et al., 2024). This is especially crucial in addressing major threats to islands such as climate change mitigation and invasive species management (Whittaker et al., 2023), which require both accurate and transferable SDMs.

4.2 Materials and Methods

4.2.1 Study area

The Tropical Eastern Pacific harbours numerous islands that form part of the Mesoamerica biodiversity hotspot (Myers et al., 2000). Among these, the Revillagigedo Archipelago (RA) stands out for its unique flora. Despite its relatively low species richness, attributed to its remote location (approximately 600km from the continent), one-third (39 species) of the plant species on these volcanic islands are endemic to the archipelago (Levin & Moran, 1989; Rzedowski, 1991; Benavides et al., 2019).

The study focuses on the two largest islands in the archipelago, Socorro (132 km^2) and Clarión (19.8 km^2), selected for their ecological significance as they host the majority of species in the archipelago. Despite their considerable conservation value, there has been a notable lack of comprehensive assessments regarding species distribution patterns across all taxa in the archipelago (Brattstrom, 1990; Flores-Palacios et al., 2009; Wanless et al., 2009). Addressing this knowledge gap can be effectively achieved through the optimal use of SDM methods (Benavides et al., 2024), supported by in-depth understanding of empirical species distribution patterns acquired from extensive fieldwork expeditions led by one of the authors (e.g., Benavides et al., 2019). This knowledge facilitates a robust assessment of the relevance of the tested approaches, even in the absence of a regional climate model for this area.

4.2.2 Species data

Occurrence data for both vascular and non-vascular plants were collected for model calibration and validation directly from the field. Vegetation transects of varying lengths were implemented; 10 transects of 100m on Clarion, the smallest island, and 15 transects of 500m-1km on the larger island Socorro. These transects were systematically located to capture diverse vegetation units and habitats. During the first sampling period, plant records were collected in both the dry (May 2018) and wet (November 2018) seasons on Clarion Island in quadrats of 5x5 meters that were located along the sides of the vegetation transects, to both characterise the vegetation and to record plant diversity, as this island was lacking such information

(Benavides et al., 2019). Dry season transects in Socorro Island were not required as the vegetation units have already been characterised (de la Luz et al., 1996; Flores-Palacios, 2009). Thus, data collection was conducted continuously along the transects at the beginning of the wet season (November 2022) to capture the largest possible range of both perennial and non-perennial species. To mitigate potential sampling biases, transects were widely distributed across both islands. In addition, in order to increase species occurrence sample sizes, incidental records were collected across both islands during other sampling events with different objectives within the sampling seasons. To homogenise sampling strategies across islands and to address the potential biases associated with the difficulties of sampling in hard-to-access areas, we filtered clusters of occurrence points (i.e., where plant data collection was more accessible, for example, due to the presence of roads or easily accessible patches of vegetation) using the ‘Thin’ function in the "spThin" R package (Aiello-Lammens et al., 2015). Prior to applying the filter, we conducted an analysis of the extent of occurrence for each species, calculating the minimum convex polygon around all occurrence points using the “adehabitatHR” R package (Calenge, 2011). The objective of the latter analysis was to confirm that clusters of points did not correspond to species with very small ranges (i.e., clustered distributions), considering the common occurrence of small range sizes in island species (Gaston & Fuller, 2009; Rosenblad et al., 2019).

To undertake the filtering, for both islands, criteria were established to define distribution ranges based on field observations of species' distributional ranges. Very small-ranged species were defined as those with an extent of occurrence of <5km^2, and thus, no filtering approach was applied to these species. Due to the larger area of Socorro Island (132km^2), for species with an extent of occurrence ranging from 5-30km^2, we implemented a filter requiring pairs of points to be at least 250m apart. For species with an extent of occurrence exceeding 30km^2, a more relaxed filter of 500m between pairs of points was used. Considering the smaller area of Clarion Island (19km^2), for species with an extent of occurrence >5km^2, a spatial filtering approach was applied to keep a 250m distance between pairs of points.

Using all available information per species after filtering, a dataset of native species specific to the archipelago with at least 10 occurrence records was compiled (Table 4.1). To increase the number of species analysed and those shared between islands for the transferability analysis, non-endemic natives were included, although their models were exclusively trained and projected within the archipelago.

4.2.3 Environmental data

Data from global climatic sources were downscaled, incorporating the effects of topography in the study area using the methodology of Flint & Flint (2012). This interpolation method combines a spatial gradient and inverse-distance-squared (GIDS) weighting to a global dataset through multiple linear regression.

To apply this methodology, bioclimatic variables for the study area were initially obtained from the global WorldClim and CHELSA databases at the smallest available resolution of 1 km/30 arc-sec. Subsequently, the original 1-km resolution WorldClim/CHELSA data were downscaled to three finer resolutions: 500, 90, and 30 m. In this downscaling approach, the location and elevation (obtained from a digital elevation model at the target resolution) of the new finer-resolution grid cell relative to a coarse-resolution grid cell are utilized to weight the parameters based on the following equation:

$$Z = \left[\sum_{i=1}^{N} \frac{Z_i + (X - X_i) * C_x + (Y - Y_i) * C_y + (E - E_i) * C_e}{d_i^2}\right] \Big/ \left[\sum_{i=1}^{N} \frac{1}{d_i^2}\right]$$

where Z is the estimated climatic variable at the specific location defined by easting (X) and northing (Y) coordinates and elevation (E); Z_i is the climatic variable from the 1-km grid cell i; X_i, Y_i, and E_i are easting and northing coordinates and elevation of the 1-km grid cell i, respectively; N is the number of 1-km grid cells in a specified search radius; C_x, C_y, and C_e are regression coefficients for easting, northing, and elevation, respectively; and d_i is the distance from the 500m, 90m or 30m site to the 1-km grid cell i depending on the downscaled resolution target (Flint & Flint, 2012).

The topographic information for downscaling was obtained from a digital elevation model at a 30m resolution (ASTER Global Digital Elevation Map v3: NASA & METI, 2018), which was later resampled to 90m and 500m resolutions with the ‘resample’ function from the 'Terra' R package (Hijmans, 2022). The elevation information from the digital elevation model was also utilised to create the terrain variable dataset. As well as elevation (ELEV), the terrain-derived dataset contained slope steepness (SLOPE) and slope exposure (ASPECT), which were derived from the elevation data using the 'Terra' R package. The topographic wetness index (TWI), a proxy for water availability, was calculated with the function ‘upslope.area’ from the 'dynatopmodel' R package (Smith & Lancaster, 2022). Finally, the distance to the coast (DTC) was computed as the Euclidean distance from the centre of each grid to the closest point on the coastline. All these layers were also converted to 90 m and 500 m resolutions. These variables were selected because they are known to be informative in capturing island plant species-habitat relationships, serving as useful proxies for microclimatic variation (Irl et al., 2020; Lannuzel et al., 2021).

For the two climate datasets (Downscaled-WorldClim and Downscaled-CHELSA), the multicollinearity of the 19 bioclimatic variables was analysed using the 'corSelect' function from the "fuzzySim" R package (Barbosa & Barbosa, 2020). This function calculates the pairwise Pearson’s correlations between variables, and if the correlation threshold ($r > 0.7$) is exceeded, one variable from the pair is excluded by selecting the variable with the highest variance inflation factor (VIF) in a linear regression model. A common final set of variables was selected for both islands (bio1 [Annual Mean Temperature], bio3 [Isothermality], bio5 [Max Temperature of Warmest Month], bio16[Precipitation of Wettest Quarter], bio18[Precipitation of Warmest Quarter]) using this approach, ensuring reasonable multicollinearity levels (all VIFs were < 5) (Abdelaal et al., 2019).

4.2.4 Modelling framework

SDMs were fitted using the ensemble of small models (ESM) approach (Lomba et al., 2010; Breiner et al., 2015). In this approach, all potential combinations of bivariate models are fitted, and their performance is evaluated based on their Area Under the Curve (AUC) scores. Subsequently, a final averaged model is created, which includes all informative small models (AUC >0.5).

The choice of this approach over more traditional modelling frameworks, which include all variables in model fitting, stems from the fact that all our species' occurrence datasets have ≤30 occurrence points (with the exception of one species). The implementation of ESM is therefore preferred because it reduces the risk of overfitting by adhering to a standard rule of including at least 10 records per predictor variable, effectively controlling for model complexity (Lomba et al., 2010; Breiner et al., 2015).

To reduce uncertainty arising from model selection (Araujo et al., 2019), a decision was made to fit SDMs using a single method that has demonstrated high performance for modelling island species distributions in the study area in combination with the ESM approach: GAMs (see Chapter 3). This fitting procedure resulted in one averaged model from all bivariate predictor combinations per species across three different baselines.

Although method-specific parameter tuning is sometimes recommended, the default settings of the respective SDM fitting algorithm were maintained, as evidence suggests that tuning features have a lower impact on predictive capability for the selected modelling method when sample sizes are small (<30 records: Valavi et al., 2022; Radomski et al., 2022). A presence-pseudoabsence approach was used to fit the models. In this method, pseudo-absence/background points were simulated by first quantifying the number of grids within the study area and then randomly sampling 10% of that area. This random sampling approach relies on fewer assumptions than other methods and has been shown in prior research to yield highly accurate predictions (Stokland et al., 2011; Descombes et al., 2023). The 10% sampling rate was selected because it provides a good balance between model performance and validation metrics (Chapter 3).

To validate model performance, two metrics were used: AUC and the continuous Boyce index. While an AUC >0.70 is conventionally accepted as the minimum threshold for adequate model performance (van Proosdij et al., 2015), AUC is known to be highly sensitive to sample prevalence and may struggle to accurately validate SDMs performance for widespread species (Lobo et al., 2008; Chapter 3). In the context of the study area, previous evidence suggests that AUC values >0.65 indicate adequate model performance (Chapter 3). Therefore, this slightly lower threshold was used to differentiate between informative and non-informative models.

The described modelling procedure was repeated for each of the three primary predictor datasets (Downscaled-WorldClim, Downscaled-CHELSA, Terrain-derived), and the three different resolutions: 500, 90 and 30m.

4.2.5 Statistical association of species' occurrence and the environmental/ climatic variables

The assumption that SDMs can provide useful predictions when predictor variables show statistical associations with species occurrence data was tested, following the approach outlined by Deblawe et al. (2016). For each species, a gridded sample of random (null) occurrence points was generated to replicate many properties of the observed species' geographic distribution. This sample had the same resolution, geographical extent, number of occupied cells, and spatial dependence (clustering or dispersion) as the actual observations, but the points were distributed randomly in space. Spatial dependence was summarised over a range of distances using the variance-stabilised Ripley K-function, also known as the L-function (Ripley, 1976). A detailed description of the procedure is available in Goreaud et al. (2004).

To assess the significance of each of the three environmental/climatic datasets for each species model, across the three resolutions analysed, the likelihood of obtaining an area under the receiver operating characteristic curve (AUC) equal to or greater than the AUC under the null hypothesis (that associated with the random null distribution points outlined above) was evaluated. Put simply, the P-value of a one-tailed test was calculated as the frequency at which AUC null ≥ AUC obs, drawn from a sample of 200 different null distributions in total per species (an alpha value of 0.05 was used to determine significance). It is important to note that this test only allows determination of whether a modelled distribution is 'random' (i.e., not different from that expected under the null distribution) relative to the environmental predictors included in the model.

Subsequently, a Chi-Square Test of Independence was conducted to determine if a modelled distribution being classified as 'random' or 'non-random' is related to different model properties. Specifically, the analysis focused on whether 'Model Randomness' (i.e., the result of the aforementioned null distribution comparison; levels = 'random' and 'non-random') is dependent on 'Resolution' (levels: '30', '90', and '500') and/or 'Environmental Baseline' (levels: 'WC', 'CH', and 'DEM') and the island analysed (levels: 'Socorro' and 'Clarion').

Finally, how accuracy (in terms of AUC) varied among the different model fits was assessed. For this, a (Gaussian) linear mixed effects model was fitted. The predictor variables of interest were 'Resolution,' 'Environmental Baseline,' and 'Island,' with all two-way interaction terms included. Marked species-specific differences in model performance were expected, so species identity ('Species') was incorporated as a random effect (random intercept): AUC ~ (Baseline + grain + Island)^2 + (1 | Species). The model was fitted using the ‘lme4’ R package (Bates et al., 2015).

4.2.6 Model transferability

For this analysis, species (n = 7) were selected based on their common occurrence across both islands within the archipelago. The goal was to investigate how grain size and the choice of predictor variables affect model transferability. Since all the selected species are more common on one island (resulting in more presence data points) than the other, presence/pseudoabsence models were trained using data from the island with the larger occurrence dataset. These models were subsequently projected to the island where the species are less common. The metric selected for model validation was the Area Under the Curve (AUC), calculated with the presence and absence data from the target projection island.

This validation approach enabled assessment of both the accuracy and transferability of the model in its projected areas. Evaluations were conducted across three analytical resolutions (30 m, 90 m, 500 m) and the three environmental datasets selected for this study (Terrain-derived, CHELSA, Worldclim). The linear mixed-effects model outlined above, specified as AUC ~ (Baseline + grain + Island)^2 + (1 | Species), was also applied to assess model transferability. This approach allowed analysis of whether the outcomes were consistent in terms of both accuracy and transferability.

4.3 Results

4.3.1 Species data

The final dataset comprises a total of 20 species, categorised into three chorotypes: single island endemics (n = 8), species endemic to the archipelago (n = 1), and non-endemic natives (n = 11). From the full dataset, 7 species were selected for the transferability assessment on the basis that they occurred on both islands, with 4 species being more common on Socorro Island and 3 species being more common on Clarion Island. Table 4.1 summarises the species list, the number of records obtained, and the species selected for transferability, detailing where the models were fitted and where they were projected.

Table 4.1 Species list and number of occurrence points (post filtering) per island. Island datasets selected as training areas for the transferability test (*) are highlighted in bold. This indicates that the model was trained in the island with the larger sample size and projected to the one with fewer occurrence records. (+) Non-endemic native species.

Taxon	*N. records Clarion*	*N. records Socorro*	*Transferability test*
Brickellia peninsularis Brandegee	*21*	***27***	*
Bursera epinnata+ (Rose) Engl.	-	*69*	-
Caesalpinia bonducella+ (L.) Fleming	*10*	-	-
Conocarpus erectus+ *L.*	-	*10*	-
Croton masonii I.M. Johnst.	-	*17*	-
Dodonaea viscosa+ (L.) Jacq.	*10*	***31***	*
Euphorbia anthonyi Brandegee	***15***	*6*	*
Euphorbia californica+ Benth.	*13*	-	-
Ficus cotinifolia+ Kunth	-	*15*	-
Guettarda insularis Brandegee	-	*24*	-
Ilex socorroensis Brandegee	-	*16*	-
Karwinskia humboltiana+ (Schult.) Zucc.	*11*	-	-
Nicotiana stocktonii Brandegee	*8*	*2*	*
Sideroxylon socorrense (Brandegee) T.D. Penn.	-	*19*	-
Perityle socorrosensis Rose	*10*	***13***	*
Psidium socorrense I.M. Johnst.	-	*25*	-
Teucrium townsendii Vasey & Rose	*7*	-	-
Tribulus cistoides+ L.	***14***	*10*	*
Waltheria indica+ .L.	*19*	***27***	*
Zanthoxylum fagara+ (L.) Sarg.	*13*	-	-

4.3.2 Model accuracy

Model performance was analysed using two metrics: AUC and the continuous Boyce index. A significant positive correlation between these metrics was observed (Pearson's r = 0.61, P <0.01), indicating correlated patterns of high and low-performing models across metrics. Thus, only the AUC results are discussed in the main text, with detailed information for both metrics available in Supplementary Table S.4.1.1.

The output of the model $AUC \sim (\text{Baseline} + \text{grain} + \text{Island})\text{^}2 + (1 \mid Species)$ revealed that most of the explained variation was attributable to the species-level random effect (Marginal R^2= 20%; Conditional R^2= 77%). The predictor "Resolution" had no significant effect on model performance (coefficient = 0; SE = 0; p = 0.17). Similarly, the "Island" variable did not significantly influence the AUC (coefficient = 0; SE = 0.02; p = 0.57). The environmental baselines had a significant impact on AUC. Specifically, the WorldClim (WC) baseline positively affected model performance (coefficient = 0.05; SE = 0.02; p < 0.01) relative to CHELSA, while the Terrain-derived (DEM) baseline had a negative effect (coefficient = -0.04; SE = 0.02; p = 0.02) relative to CHELSA. Overall, models using downscaled baselines produced higher AUC values across both islands, with resolutions of ≤90m proving particularly beneficial for SDM predictions on Clarion Island.

Some interactions between “Environmental Baseline,” “Resolution,” and “Island” were significant in the regression model. There was a negative interaction between the DEM baseline and Clarion Island (coefficient = -0.06; SE = 0.02; $p < 0.01$), and between the WC baseline and Clarion Island (coefficient = -0.08; SE = 0.02; $p < 0.01$). Additionally, there was a significant positive interaction between the DEM baseline and resolution (coefficient = 0.001; SE = 0.0005; $p < 0.01$), where coarser resolutions increased the AUC scores of models fitted with DEM variables. Figure 4.1 shows the AUC differences across islands, environmental baselines, and resolutions.

Regardless of the selected analytical resolution or type of environmental baseline, two species consistently had AUC values below the 0.65 threshold, which is considered the minimum acceptable for model performance in the study area. Specifically, *B. peninsularis* (mean AUC = 0.57) and *E. anthonyi* (mean AUC = 0.59) exhibited poor model performance across all tested scenarios (Fig. 4.2).

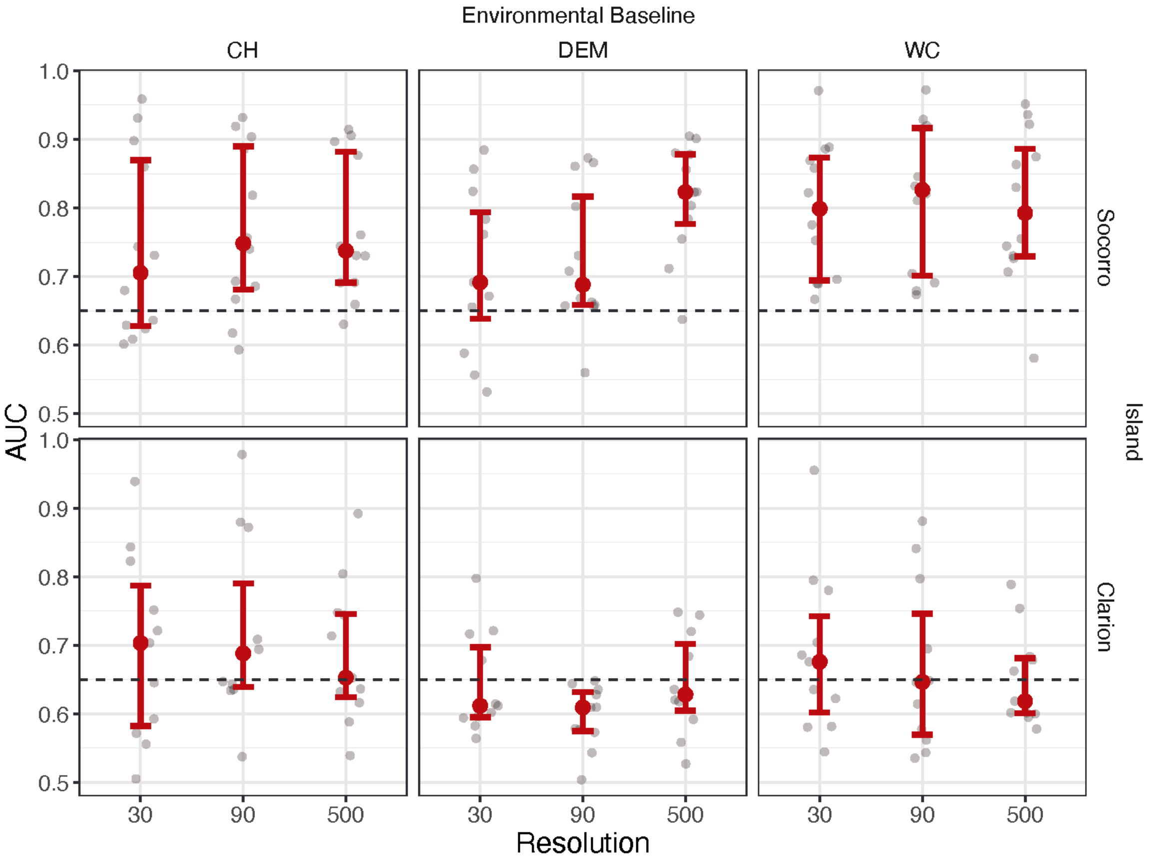


Figure 4.1 AUC results from SDMs fitted to 20 species across Environmental Baselines (WC= downscaled Worldclim; CH= downscaled CHELSA; DEM=Topographically derived), and Resolution. Red circles show mean AUC and the lines show the 95% CI. Gray points indicate the AUC values of each of the 20 species analysed. Gray dashed line marks the AUC threshold considered acceptable for model performance (0.65).

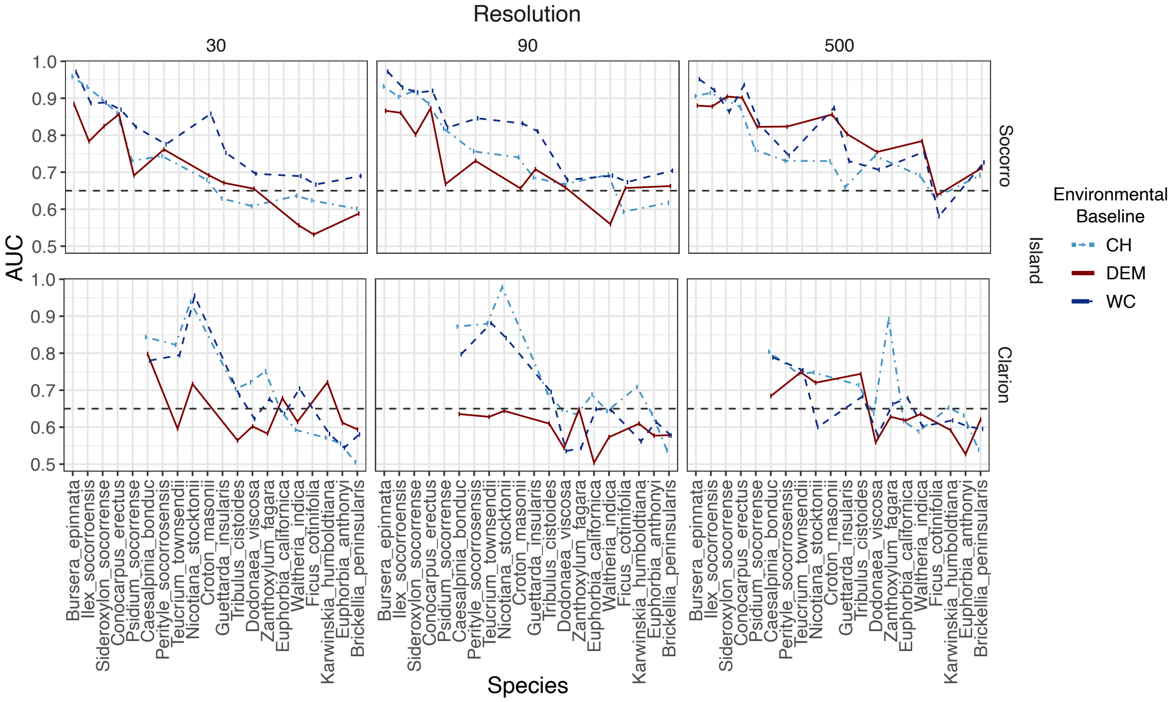


Figure 4.2 AUC values from fitted models by species and across different resolutions and environmental baselines (WC= downscaled Worldclim; CH= downscaled CHELSA; DEM=Topographically derived). Downscaled environmental (bioclimatic) baselines are represented in blue shades, while the terrain-derived baseline is shown in red. Gray dashed line marks the AUC threshold considered acceptable for model performance (0.65).

4.3.3 Statistical association between species' occurrence and the environmental/climatic variables

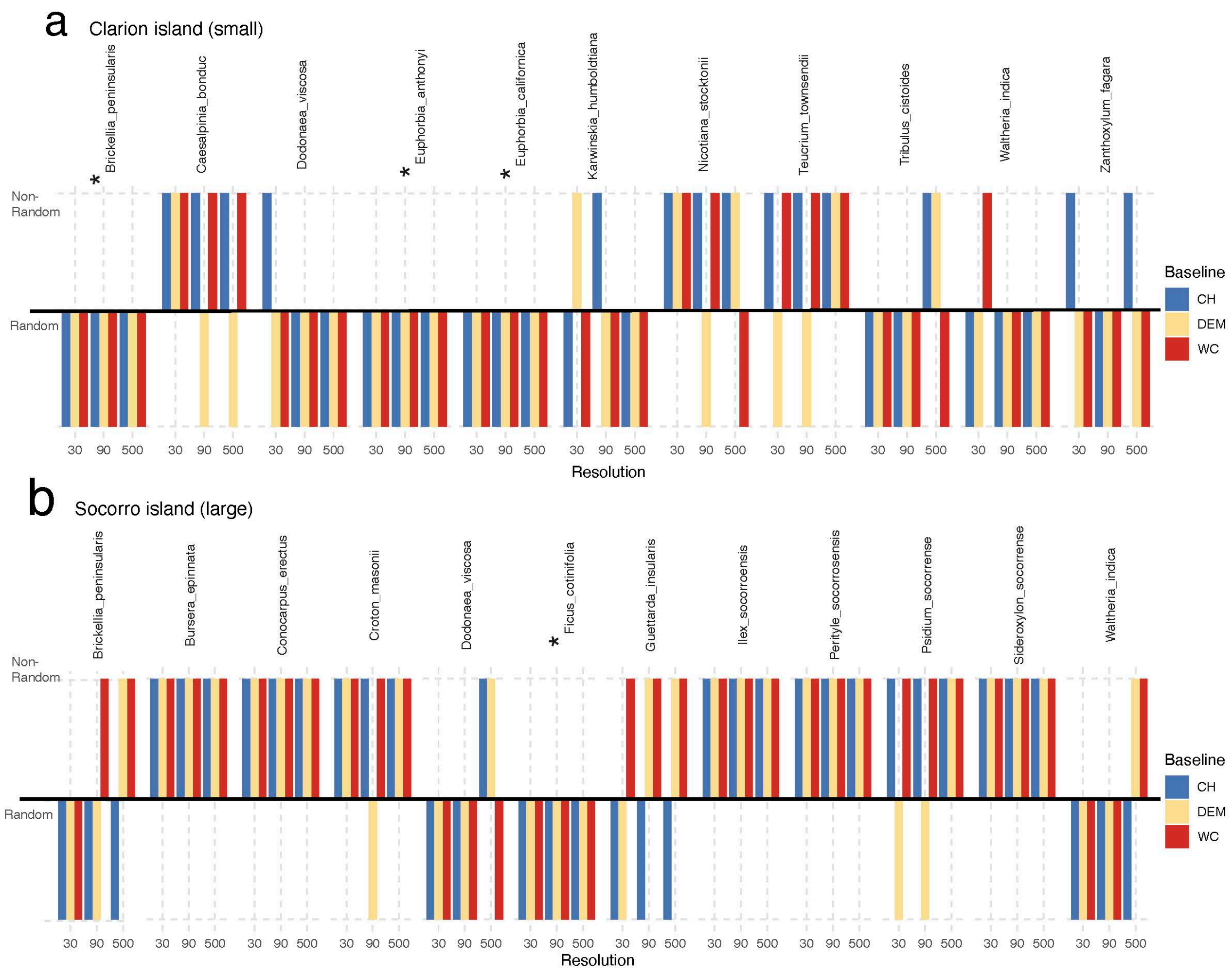
The Chi-Square test of independence revealed no significant association between 'Environmental Baseline' or 'Resolution' and 'Model Randomness' (Resolution: $\chi^2(2) = 2.98$, $p = 0.23$; Environmental Baseline: $\chi^2(2) = 1.89$, $p = 0.38$). This suggests that neither the choice of baseline nor resolution significantly impacts whether a model fit is worse than random. However, 'Model Randomness' was significantly associated with both 'Species' ($\chi^2(19) = 144.6$, $p < 0.01$) and 'Island' ($\chi^2(1) = 28.87$, $p < 0.001$), indicating species-specific relationships with the different environmental baselines and resolutions, and that there are significant differences in the proportion of random models observed between islands (discussed below).

On Clarion Island, all SDMs for three species—*B. peninsularis*, *D. viscosa*, and *E. anthonyi*—resulted in random fits, meaning that the modelled fits were no better than a randomly generated distribution. These species also exhibited low model performance (AUC < 0.65), suggesting that their distributions could not be reliably predicted by any of the environmental baselines used, regardless of resolution.

For Clarion SDMs, non-random model fits were more common at finer spatial resolutions (30m and 90m). Notably, 27% (n=3) of the species displayed non-random fits, which were evenly distributed across all resolutions and baselines. In contrast, 45% (n=5) of the species showed non-random fits in specific scenarios, particularly when using downscaled CHELSA (CH) data (Fig. 4.3). Overall, downscaled CH data contributed to non-random fits in 7 of the 8 species, followed by DEM and downscaled Worldclim (WC), which contributed to non-random fits in 5 and 4 species, respectively.

For the Socorro Island SDM fits, only one species, *F. cotinifolia*, had a random SDM fit, and this model also had poor performance (AUC < 0.65). In contrast to the Clarion Island models, where non-random SDM fits were mostly specific to a particular baseline and generally associated with smaller resolutions, on Socorro

Island, non-random SDM fits were more consistent across different resolutions and baselines (Fig. 4.3). Specifically, 58% of the species (n=7) had consistent non-random fits across most scenarios of environmental baselines and resolutions. Whereas 33% of the species (n=4) showed non-random fits in specific scenarios based on WC and DEM datasets. Figure 4.3 summarises the distribution of random and non-random SDM fits across the variables analysed.



b

Figure 4.3 Distribution of random (lower section) versus non-random (upper section) model fits across different environmental baselines and resolutions for both Clarion and Socorro islands. The X-axis represents three resolutions (30m, 90m, 500m) and three environmental baselines (CH: downscaled CHELSA, DEM: Terrain-derived, and WC: WorldClim). a - Clarion Island: a higher number of SDMs resulted in random fits, with non-random fits occurring primarily in scenarios using downscaled CHELSA data. b - Socorro Island: shows a more balanced distribution of random and non-random fits across all resolutions and baselines. (*) Species that exhibited all-random SDM fits.

4.3.4 Model transferability

Model transferability, defined as the ability of a model fitted using presence-absence data from one island to accurately predict the species' distribution on a second island, was assessed for seven species. This assessment was conducted using empirical presence-absence data from the island to which models were projected, and evaluated with the AUC (Area Under the Curve) metric. Figure 4.4 summarises the AUC-transferability results, with detailed information provided in Supplementary Table S.4.1.2.

A mixed-effect model, with AUC_transferability as the response, species as a random effect, and environmental baseline, study grain and island as fixed effects, revealed that the fixed effects explained a relatively large proportion of the variability in the AUC_transferability metric (Marginal R^2 = 62%; Conditional R^2 = 11%). Unlike model performance, whereby the species-level random effect explained most of the variation in AUC, the variation in model transferability was better explained by the environmental baseline used in the SDM fit. Specifically, the Terrain-derived (DEM) baseline had a positive impact on model performance (coefficient = 0.16; SE = 0.04; p = 0.01) relative to CHELSA, while the Downscaled WorldClim (WC) baseline also had a positive effect (coefficient = 0.09; SE = 0.04; p = 0.02) relative to CHELSA. The predictor ‘Resolution’ did not have a significant effect on model performance (coefficient = 0; SE = 0; p =0.34). Similarly, the ‘Island’ variable did not significantly influence the AUC_transferability (coefficient = -0.07; SE = 0.05; p = 0.35).

A marked difference was observed in transferability between islands. Species distribution models fitted using data from the largest island (Socorro) and projected onto the smallest island (Clarion) showed poor transferability (AUC_transferability < 0.65) for all species. For species projected into Socorro Island, the Terrain (DEM) environmental baseline generally provided the highest mean AUC_transferability across all resolutions (DEM AUC = 0.76; DEM-Res30 = 0.79; DEM-Res90 = 0.77; DEM-Res500 = 0.72). Model transferability for WorldClim (WC) also showed high average AUC_transferability scores (WC AUC = 0.71; WC-Res30 = 0.72; WC-Res90 = 0.74; WC-Res500 = 0.70). In contrast, CHELSA (CH) showed lower average AUC_transferability scores (CH AUC = 0.65; CH-Res30 = 0.60; CH-Res90 = 0.64; CH-Res500 = 0.65). Differences in model transferability across resolutions were minimal for both islands, with only slight improvements observed at the finest (30m) resolution (Figure 4.4).

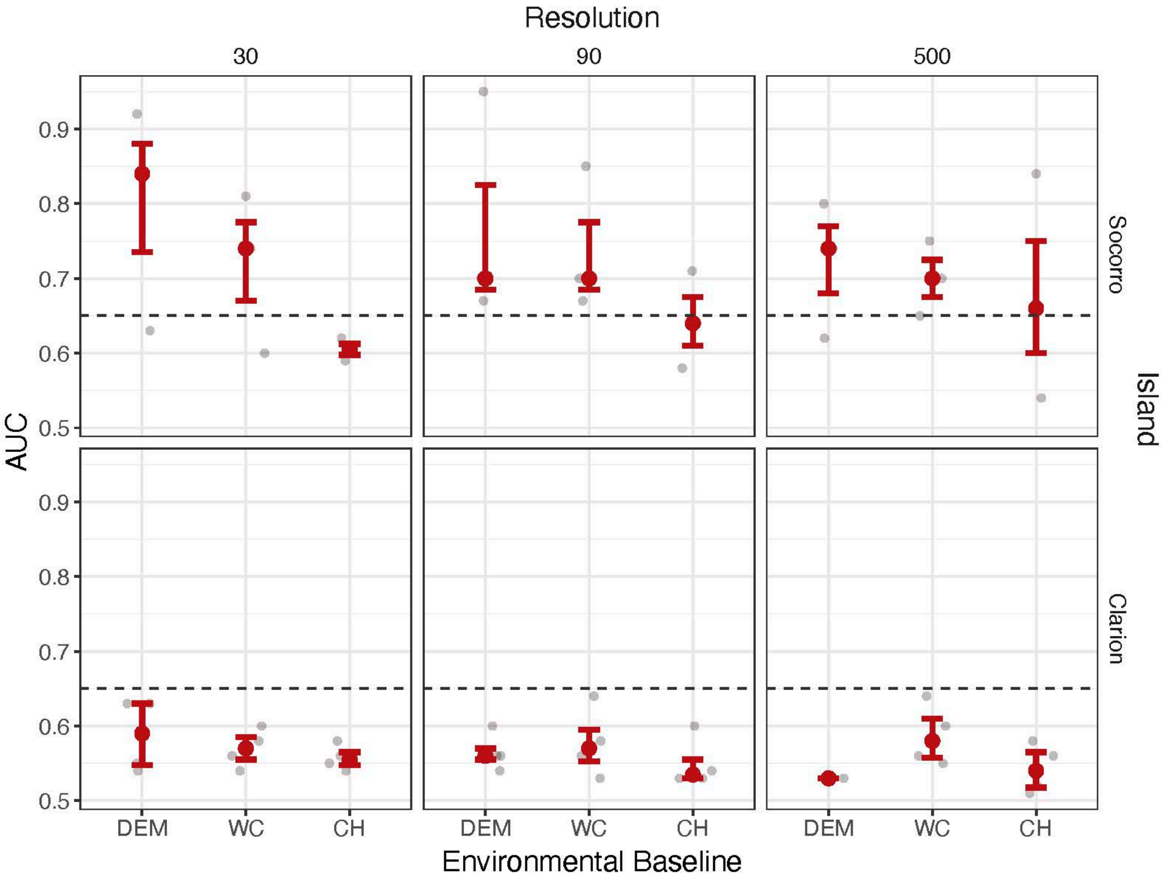


Figure 4.4 AUC-transferability results by species across different resolutions and Environmental baselines (WC= downscaled wordlclim; CH= dowscaled Chelsa; DEM=Topographically derived). Gray points are AUC_transferability values per species across diverse scenarios of resolution and baselines. Red circles are average AUC_transferability scores and the lines extend to the 95% CI. Gray dashed lines mark the AUC threshold used in the study area

4.4 Discussion

Species Distribution Models (SDMs) are essential tools for understanding the rich yet understudied biodiversity of tropical regions (Poteau et al., 2019), particularly on islands. However, their effectiveness is often hindered by uncertainties in climatic estimates (Deblauwe et al., 2016) and the inability of these estimates to capture the variability of small or topographically complex islands (Benavides et al., 2024). As a result, there is a reliance on proxies, such as topography and remotely sensed vegetation data, to fill these gaps.

This study aimed to address these environmental data limitations while focusing on a selection of ecologically relevant variables for SDMs at appropriate scales—an aspect that has been relatively understudied (Austin & Van Niel, 2011; Syphard & Franklin, 2009; Tulloch et al., 2016). Advances in data availability and computational power have made it increasingly feasible to develop detailed microclimatic descriptors by downscaling global climate data. However, generating downscaled climate data is only part of the challenge; it is also necessary to ensure that these variables, in the context of SDMs, are ecologically valid and accurately reflect species-environment relationships. To this end, plant species distributions were analysed on islands, which are often particularly sensitive to fine-scale climatic and topographic gradients (Lannuzel et al., 2021). The downscaling methods explored in this study have been shown to be effective in describing microclimatic variability, thus leading to accurate plant SDMs in many contexts (Flint & Flint, 2012; Franklin et al., 2013). However, the uncertainty associated with global climatic estimates in tropical islands (Deblauwe et al., 2016) means that the reliability of these approaches in tropical island environments remains uncertain.

It was initially expected that terrain-derived predictors would result in more accurate and informative distribution models, as local topography is known to be an important driver of microclimatic habitats relevant for plant species through factors such as slope, exposure, water availability, and temperature (Franklin et al., 2013; Chauvier et al., 2021). Contrary to expectations, the analysis revealed that, when spatial context and species properties were considered, downscaled climatic predictors provided more accurate SDMs. This suggests that such variables may capture habitat preferences not reflected by DEM-based predictors. This finding indicates a need to reconsider the reliance on terrain-derived data in high-resolution island SDM studies (e.g., Poteau et al., 2019; Lannuzel et al., 2021).

The remainder of the discussion outlines and contextualises scenarios where each type of predictor (and the corresponding analytical resolutions) is relevant for (1) predicting species distributions and (2) extrapolating models to novel conditions (i.e., model projections).

4.4.1 Model Accuracy and Species–Climate Associations Using High-Resolution Environmental/Climatic Data

The analysis of two islands with differing sizes and levels of topographic complexity provided valuable insights into how spatial context influences species associations with environmental descriptors across various analytical resolutions. Socorro, the larger island (132 km^2) with complex topography and higher elevation (1000m), presented contrasting results to Clarion, a much smaller island (19 km^2) with relatively simpler topography and lower elevation (300m).

In the case of Socorro, SDMs generally showed higher accuracy (relative to Clarion) and produced consistent species-climate associations across different environmental baselines (CH, DEM, WC) and resolutions (30m, 90m, 500m). The consistent effectiveness of microclimatic variables (WC/CH), as predictor variables in SDMs based on species data from Socorro, can be attributed to the island's larger area, which provided more data for downscaling (in the form of 1km gridded climate data), thereby reducing model uncertainty. Additionally, the island's size resulted in broader resolution predictors more effectively capturing environmental variability, leading to more stable results in terms of predictive accuracy (Austin & Van Niel, 2011).

In contrast, SDM fits based on data from the smaller island of Clarion exhibited greater uncertainty, with models showing high accuracy and generating better-than-random fits only in specific scenarios regarding resolution and variable choice. The most accurate models for Clarion were based on very fine-resolution environmental descriptors (30m), a logical outcome as finer resolution data provide a better representation of environmental variability in small study areas. Notably, the relative performance of the CH dataset was greater on Clarion Island, and was the only case where this environmental baseline outperformed WC and DEM (further explanation in the impacts of transferability section).

The models based on DEM descriptors performed relatively poorly for Clarion, likely due to its lower topographic diversity. On this smaller and less orographically diverse island, the relationship between topography and microclimate may be weaker, obscuring the role of environmental factors that influence species distributions on larger, more topographically varied islands (McInerny et al., 2011; Lembrechts et al., 2019).

4.4.2 Balancing model accuracy and transferability in regards to variable selection and analytical resolution

Evidence suggests that even if a model performs well in terms of its fit to a given training dataset, it may not necessarily exhibit high transferability (Huang et al., 2016). This is particularly important in island studies, where it is often crucial to undertake such a complementary approach to model evaluation, for two main reasons:

1) Resolution and Transferability**:** Previous research has indicated that SDMs fitted with fine-scale resolution data often exhibit lower spatial transferability (Manzoor et al., 2018). However, this study contradicted this expectation, revealing that transferability improved when models were fitted with finer resolution data. This discrepancy may be attributed to the significant size differences between the islands involved in the transferability analysis. For Socorro, the larger island, resolution was less critical for accurate modelling, whereas for Clarion, the smaller

island, a finer resolution was necessary. Therefore, when transferring models between these islands, it is crucial to match the resolution appropriately to ensure that the model performs well in both environments.

2) Type of Predictor Variables: Transferability is closely linked to the type of predictor variables used in SDMs, with studies showing differences in results depending on the utilised climatic datasets (Petitpierre et al., 2017; Manzoor et al., 2018; Goedecke et al., 2020; Datta et al., 2020). Additionally, the choice of global climatic baselines is sensitive to the amount of meteorological data available in the study area (Karger et al., 2017). Previous studies have shown that plant SDMs in regions with sparse meteorological stations performed better when calibrated with the CHELSA (1.1) dataset compared to WorldClim (Bobrowski & Schickhoff, 2017; Karger et al., 2017). Based on this, downscaled CHELSA data were initially expected to outperform WorldClim in this study. However, our findings revealed that models fitted with WorldClim generally performed better and exhibited greater transferability than those using CHELSA. This contrasting result, which is consistent with Datta et al. (2020), may be due to the aforementioned previous studies using the first version of WorldClim (version 1), whereas the present analyses were based on the latest version of WorldClim (version 2). The relative improvement of WorldClim2 could be due to enhancements in data quality, including the incorporation of remotely sensed variables like land surface temperature and cloud

cover. The latest version of CHELSA (1.2) has also made corrections to account for orographic effects, particularly concerning precipitation (Karger et al., 2017). However, these corrections in CHELSA 1.2 may compromise the transferability of climatic models when projecting to new regions if the local correlation structure among variables changes (Mesgaran et al., 2016). This issue is relevant in the present study, as Clarion Island—the smaller of the two studied islands—has less topographic variability compared to Socorro, the largest island in the archipelago. This difference in topographic variability may explain why models fitted with CHELSA data showed poor transferability in general. Nonetheless, CHELSA data were useful in modelling species distributions in some cases, even with very limited gridded data on the smaller island.

4.4.3 Study Limitations

Beyond spatial context, species identity significantly influenced model accuracy, with the species-level random effect explaining most of the variation in the mixed-effects model of SDM accuracy. Certain characteristics of the study species were associated with poor model fit and transferability. In particular, species with uninformative model fits—indicated by low AUC scores or fits no better than random—and poor transferability shared a common trait: high prevalence in the study area (i.e., widespread species). This characteristic has been identified as a challenge for SDM fitting, both in the study area (Chapter 3) and more broadly (van

Proosdij et al., 2015). Additionally, with the exception of one island endemic species (*E. anthonyii*), all the relatively poor SDM fits were associated with non-endemic native species. This is likely due to the models only capturing a truncated realised niche, which is known to negatively affect SDM fit and transferability (Suárez-Seoane et al., 2014; Benavides et al., 2024).

Although the potential issue of niche truncation for non-endemic species was recognised during the method design and species selection phases of the study, these species were included in the analyses because field observations indicated that they are dominant elements in the vegetation structure of the analysed islands. Mapping their distributions is, therefore, important for fully understanding the community assembly of island floras. There was also a practical necessity to their inclusion given that some of these species are shared between the islands, thus increasing the species pool for the transferability tests. Overall, the poor SDM fits, rather than reducing the validity of the selected predictors / baselines, confirm the previously identified challenges of SDM implementation on islands (Benavides et al., 2024), and highlight potential sources of noise that should be considered in future island SDM studies.

4.5 Conclusion

In this study, the reliability of SDMs was evaluated in a tropical insular hotspot where the availability of fine-scale climatic data is limited, and strategies to enhance their application in such challenging contexts were explored. The findings broaden the range of potential options for using high-resolution environmental baselines in SDMs for tropical island plants. It was demonstrated that selecting appropriate downscaled fine-scale predictors, guided by the spatial context of the study islands, can significantly improve the informativeness of SDMs. This finding calls into question the common reliance on topographically derived data (Benavides et al., 2024), which may not always be the most reliable or ecologically robust choice in island SDM studies.

While downscaled microclimatic predictors were found to be effective in many contexts, this result may not universally apply to other tropical islands. However, the null-hypothesis testing approach utilised here is likely to be beneficial in island SDM studies more broadly and should guide the selection of informative predictor variables in future studies. By implementing this approach in SDM studies of regions with environmental data limitations, researchers can better assess the reliability of their models. Additionally, valuable insights were gained into improving the transferability of SDMs—in an insular context—by selecting relevant variables at appropriate resolutions, which has been shown to also depend on spatial context. These findings support approaches that involve model extrapolation across different islands, such as in invasive species management and conservation translocations.

4.6 References

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, *38*(5), 541-545.

Ashrafzadeh, M. R., Khosravi, R., Adibi, M. A., Taktehrani, A., Wan, H. Y., & Cushman, S. A. (2020). A multi-scale, multi-species approach for assessing effectiveness of habitat and connectivity conservation for endangered felids. *Biological Conservation*, *245*, 108523.

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

Austin, M. P., & Van Niel, K. P. (2011b). Impact of landscape predictors on climate change modelling of species distributions: a case study with *Eucalyptus fastigata* in southern New South Wales, Australia. *Journal of Biogeography*, *38*(1), 9-19.

Barbosa A. M. (2015). "fuzzySim: applying fuzzy logic to binary similarity indices in ecology." *Methods in Ecology and Evolution*, 6(7), 853-858.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B,. Singmann, H., …& Bolker, M. B. (2015). Package 'lme4'. *Convergence*, *12*(1), 2.

Bazzichetto, M., Massol, F., Carboni, M., Lenoir, J., Lembrechts, J. J., Joly, R., & Renault, D. (2021). Once upon a time in the far south: Influence of local drivers and functional traits on plant invasion in the harsh sub-Antarctic islands. *Journal of Vegetation Science*, 32(4), e13057.

Beever, E. A., Ray, C., Wilkening, J. L., Brussard, P. F., & Mote, P. W. (2011). Contemporary climate change alters the pace and drivers of extinction. *Global Change Biology, 17*(6), 2054-2070.

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

Benavides, E., Kuethe, J. R., Ortiz-Alcaraz, A., & DE LA LUZ, J. L. (2019). *Oenothera resicum* (Onagraceae), a new species and the first record of the family from the Revillagigedo Archipelago, Mexico. *Phytotaxa*, *416*(1), 59-66.

Benavides, E., Sadler, J., Graham, L., & Matthews, T. J. (2024). Species distribution models and island biogeography: Challenges and prospects. *Global Ecology and Conservation*, e02943.

Bobrowski, M., Gerlitz, L., & Schickhoff, U. (2017). Modelling the potential distribution of *Betula utilis* in the Himalaya. *Global Ecology and Conservation*, *11*, 69-83.

Brattstrom, B. H. (1990). Biogeography of the Islas Revillagigedo, Mexico. *Journal of Biogeography*, *17*(2), 177-183.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, *6*(10), 1210-1218.

Brook, B. W., Sodhi, N. S., & Bradshaw, C. J. (2008). Synergies among extinction drivers under global change. *Trends in Ecology & Evolution*, *23*(8), 453-460.

Calenge, C. (2011). Home range estimation in R: the adehabitatHR package. *Office national de la classe et de la faune sauvage: Saint Benoist, Auffargis, France*.

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

Datta, A., Schweiger, O., & Kühn, I. (2020). Origin of climatic data can determine the transferability of species distribution models. *NeoBiota 59*, 61–76.

de la Luz, J. L. L., Breceda, A., & Benet, R. C. (1996). Las comunidades vegetales en la isla Socorro, México. *SIDA, Contributions to Botany*, 215-230.

Dobrowski, S. Z. (2011). A climatic basis for microrefugia: the influence of terrain on climate. *Global Change Biology*, *17*(2), 1022-1035.

Hanz, D. M., Cutts, V., Barajas-Barbosa, M. P., Algar, A., Beierkuhnlein, C., Collart, F., ... & Irl, S. D. (2023). Effects of climate change on the distribution of plant species and plant functional strategies on the Canary Islands. *Diversity and Distributions*, *29*(9), 1157-1171.

Heinänen, S., Erola, J., & von Numers, M. (2012). High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. *Landscape Ecology*, *27*(4), 545-555.

Huang, J., Frimpong, E. A., & Orth, D. J. (2016). Temporal transferability of stream fish distribution models: can uncalibrated SDMs predict distribution shifts over time? *Diversity and Distributions*, *22*(6), 651-662

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

Flint, L. E., & Flint, A. L. (2012). Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, *1*, 1-15.

Flores-Palacios, A., Martínez-Gómez, J. E., & Curry, R. L. (2009). La vegetación de isla Socorro, archipiélago de Revillagigedo, México. *Boletín de la Sociedad Botánica de México*, *84*, 13-23.

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

Franklin, J. 2010. *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press, Cambridge.

Franklin, J., Davis, F. W., Ikegami, M., Syphard, A. D., Flint, L. E., Flint, A. L., & Hannah, L. (2013). Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Global Change Biology*, *19*(2), 473-483.

Gábor, L., Jetz, W., Zarzo-Arias, A., Winner, K., Yanco, S., Pinkert, S., ... & Moudrý, V. (2023). Species distribution models affected by positional uncertainty in species occurrences can still be ecologically interpretable. *Ecography*, *2023*(6), e06358.

Gaston, K. J., & Fuller, R. A. (2009). The sizes of species' geographic ranges. *Journal of Applied Ecology*, *46*(1), 1-9.

Goedecke, F., Marcenò, C., Guarino, R., Jahn, R., & Bergmeier, E. (2020). Reciprocal extrapolation of species distribution models between two islands–specialists perform better than generalists and geological data reduces prediction accuracy. *Ecological Indicators*, *108*, 105652.

Goreaud, F., Loussier, B., Ngo Bieng, M.A. & Allain, R. (2004) Simulating realistic spatial structure for forest stands: a mimetic point process. *Interdisciplinary Spatial Statistics Workshop 2004,* December 2–3, 2004, Paris, France. ENGREF and University of Paris I Panthéon-Sorbonne, Paris, France.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., ... & Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424-1435.

Heinänen, S., Erola, J., & von Numers, M. (2012). High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. *Landscape Ecology*, *27*(4), 545-555

Humphreys, A. M., Govaerts, R., Ficinski, S. Z., Nic Lughadha, E., & Vorontsova, M. S. (2019).

Global dataset shows geography and life form predict modern plant extinction and rediscovery. *Nature Ecology & Evolution*, *3*(7), 1043-1047.

Khosravi, R., Hemami, M. R., Malekian, M., Flint, A., & Flint, L. (2016). Maxent modeling for predicting potential distribution of goitered gazelle in central Iran: the effect of extent and grain size on performance of the model. *Turkish Journal of Zoology*, *40*(4), 574-585.

Lannuzel, G., Balmot, J., Dubos, N., Thibault, M., & Fogliani, B. (2021). High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia. *Biodiversity and Conservation*, *30*(4), 963-990.

Lembrechts, J. J., Nijs, I., & Lenoir, J. (2019). Incorporating microclimate into species distribution models. *Ecography*, *42*(7), 1267-1279.

Levin J.A. y Moran R. 1989. The vascular flora of Isla Socorro, México. San Diego Society of Natural History, University of Michigan, 71 pp.

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

Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, *143*(11), 2647-2657.

Longcore, T., Noujdina, N., & Dixon, P. J. (2018). Landscape modeling of the potential natural vegetation of Santa Catalina Island, California. *Western North American Naturalist*, *78*(4), 617-632.

Manzoor, S. A., Griffiths, G., & Lukac, M. (2018). Species distribution model transferability and model grain size–finer may not always be better. *Scientific Reports*, *8*(1), 1-9.

McInerny, G. J., & Purves, D. W. (2011). Fine‐scale environmental variation in species distribution modelling: regression dilution, latent variables and neighbourly advice. *Methods in Ecology and Evolution*, *2*(3), 248-257.

McGeoch, M. A., Genovesi, P., Bellingham, P. J., Costello, M. J., McGrannachan, C., & Sheppard, A. (2016). Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. *Biological Invasions*, *18*, 299-314.

Mesgaran, M. B., Lewis, M. A., Ades, P. K., Donohue, K., Ohadi, S., Li, C., & Cousens, R. D. (2016). Hybridization can facilitate species invasions, even without enhancing local adaptation. *Proceedings of the National Academy of Sciences*, *113*(36), 10210-10214.

Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global biodiversity conservation: the critical role of hotspots. In *Biodiversity hotspots: distribution and protection of conservation priority areas* (pp. 3-22). Springer, Berlin Heidelberg.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853-858.

NASA JPL. *NASA Shuttle Radar Topography Mission Global 1 Arc Second*; NASA EOSDIS Land Processes DAAC: Pasadena, CA, USA, 2013.

NASA/METI/AIST/Japan Space Systems, and U.S./Japan ASTER Science Team. ASTER global digital elevation model V003, 2018, distributed by NASA EOSDIS Land Processes DAAC, https://doi.org/ 10.5067/ASTER/ASTGTM.

Oberbauer, T. A. (2002, March). Analysis of vascular plant species diversity of the Pacific Coast islands of Alta and Baja California. In *Proceedings of the Fifth California Islands Symposium. Santa Barbara Museum of Natural History, Santa Barbara, CA* (pp. 201-211).

Patiño, J., Collart, F., Vanderpoorten, A., Martin-Esquivel, J. L., Naranjo-Cigala, A., Mirolo, S., & Karger, D. N. (2023). Spatial resolution impacts projected plant responses to climate change on topographically complex islands. *Diversity and Distributions*, *29*(10), 1245-1262.

Peterson, A. T., J. Soberon, R.G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araujo. (2011). *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton.

Petitpierre B, Broennimann O, Kueffer C, Daehler C, Guisan A (2017). Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Global Ecology and Biogeography*, *26*, 275287

Pili, A. N., Tingley, R., Sy, E. Y., Diesmos, M. L. L., & Diesmos, A. C. (2020). Niche shifts and environmental non-equilibrium undermine the usefulness of ecological niche models for invasion risk assessments. *Scientific Reports*, *10*(1), 7972.

Pouteau, R., Munoz, F., & Birnbaum, P. (2019). Disentangling the processes driving tree community assembly in a tropical biodiversity hotspot (New Caledonia). *Journal of Biogeography*, *46*(4), 796-806.

Radomski, T., Beamer, D., Babineau, A., Wilson, C., Pechmann, J., & Kozak, K. H. (2022). Finding what you don't know: testing SDM methods for poorly known species. *Diversity and Distributions*, *28*(9), 1769-1780.

Ricketts, T. H., Dinerstein, E., Boucher, T., Brooks, T. M., Butchart, S. H., Hoffmann, M., ... & Wikramanayake, E. (2005). Pinpointing and preventing imminent extinctions. *Proceedings of the National Academy of Sciences*, *102*(51), 18497-18501.

Rodríguez-Rodríguez, P., de Castro, A. G. F., & Sosa, P. A. (2018). The restoration of the endangered *Sambucus palmensis* after 30 years of conservation actions in the Garajonay National Park: genetic assessment and niche modeling. *PeerJ*, *6*, e4985.

Rosenblad, K. C., Perret, D. L., & Sax, D. F. (2019). Niche syndromes reveal climate-driven extinction threat to island endemic conifers. *Nature Climate Change*, *9*(8), 627-631.

Rousseau, J. S., & Betts, M. G. (2022). Factors influencing transferability in species distribution models. *Ecography*, *2022*(7), e06060.

Ripley, B.D. (1976) The second-order analysis of stationary point processes. *Journal of Applied Probability*, *13,* 255-266.

Segal, R. D., Massaro, M., Carlile, N., & Whitsed, R. (2021). Small-scale species distribution model identifies restricted breeding habitat for an endemic island bird. *Animal Conservation*, *24*(6), 959-969.

Smith P, Metcalfe P (2023). *dynatop: An Implementation of Dynamic TOPMODEL Hydrological Model in R*. R package version 0.3.0.1010. https://github.com/waternumbers/dynatop

Soberon, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters 10*(12), 1115–23.

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

Stanton, J. C., Pearson, R. G., Horning, N., Ersts, P., & Reşit Akçakaya, H. (2012). Combining static and dynamic variables in species distribution models under climate change. *Methods in Ecology and Evolution*, *3*(2), 349-357.

Stork, N. E., & Habel, J. C. (2014). Can biodiversity hotspots protect more than tropical forest plants and vertebrates? *Journal of Biogeography*, *41*(3), 421-428.

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

Tshwene-Mauchaza, B., & Aguirre-Gutiérrez, J. (2019). Climatic drivers of plant species distributions across spatial grains in Southern Africa tropical forests. *Frontiers in Forests and Global Change*, *2*, 69.

Turvey, S. T., Kennerley, R. J., Hudson, M. A., Nuñez-Miño, J. M., & Young, R. P. (2020). Assessing congruence of opportunistic records and systematic surveys for predicting Hispaniolan mammal species distributions. *Ecology and Evolution*, *10*(11), 5056-5068.

USGS. *Global Multi-Resolution Terrain Elevation Data 2010 7.5 Arc Second*; USGS EarthExplorer: Washington, DC, USA, 2010.

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs*, *92*(1), e01486.

Vieilledent, G., Cornu, C., Sanchez, A. C., Pock-Tsy, J. M. L., & Danthu, P. (2013). Vulnerability of baobab species to climate change and effectiveness of the protected area network in Madagascar: Towards new conservation priorities. *Biological Conservation*, *166*, 11-22.

Wanless, R. M., Aguirre-Muñoz, A., Angel, A., Jacobsen, J. K., Keitt, B. S., & McCann, J. (2009). Birds of Clarion Island, Revillagigedo Archipelago, Mexico. *The Wilson Journal of Ornithology*, *121*(4), 745-751.

Weller, S. G., Sakai, A. K., Clark, M., Lorence, D. H., Flynn, T., Kishida, W., ... & Wood, K. (2018). The effects of introduced ungulates on native and alien plant species in an island ecosystem: Implications for change in a diverse mesic forest in the Hawaiian Islands. *Forest Ecology and Management, 409*, 518-526.

Whittaker, R. J., Fernández-Palacios, J. M., & Matthews, T. J. (2023). *Island biogeography: geo-environmental dynamics, ecology, evolution, human impact, and conservation*. Oxford University Press, Oxford.

Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E., & Willis, K. J. (2005). Conservation biogeography: assessment and prospect. *Diversity and Distributions*, *11*(1), 3-23.

5 . Predicting Microclimatic-driven Shifts in the Distribution of Conservation Priority Plant Species in Oceanic Island Biodiversity Hotspots

Abstract

Island endemic species are often endangered due to their small ranges and limited migration options in response to climate change and other anthropogenic pressures. Understanding insular species distributions is essential for addressing anthropogenic threats on islands, and Species Distribution Models (SDMs) have become valuable tools in this regard. In this study, SDMs were used to assess the impact of climate change on a group of indicator plant species—those that are vulnerable and/or have recognised ecosystem functions—across four archipelagos: the Canary Islands, Galápagos, Hawaiian Islands, and the Revillagigedo archipelago. By incorporating microclimatic data into the SDMs, findings show that most species are projected to experience significant reductions in their distributions, with an average range reduction of 27-37%. Spatial patterns of distribution decline varied, with some regions showing concentrated losses, while others exhibited more dispersed patterns. The analysis also revealed that species categorised as "Least Concern" or not evaluated by the IUCN were more vulnerable to climate-driven range loss than species already listed as threatened. Additionally, by analysing species chorotypes, it was confirmed that endemic species are disproportionately more at risk compared to their non-endemic native counterparts. These findings highlight gaps in current conservation assessments and provide critical insights for developing targeted strategies to protect biodiversity and ecosystem functionality in island ecosystems under climate change.

5.1 Introduction

Despite comprising only 6.7% of the Earth's surface, islands support approximately 20% of global biodiversity (Whittaker et al., 2023). However, human activities—including habitat destruction, the introduction of invasive species, overexploitation of resources, and the accelerating impacts of climate change—pose significant threats to island ecosystems, placing many species at risk of extinction (Matthews et al., 2022; Whittaker et al., 2023). Oceanic islands, in particular, face even greater vulnerability due to their geographic isolation and unique geological histories, which have resulted in high levels of species endemism (Whittaker et al., 2023). As a result, islands account for roughly 75% of recorded extinctions and support a disproportionate number of the world’s endangered species (Fernández-Palacios et al., 2021).

To address the alarming decline in island biodiversity driven by anthropogenic pressures, it is first necessary to understand where species occur on islands. Species Distribution Models (SDMs) have become widely used tools in this regard, aiding in increasing our understanding of anthropogenic impacts on biodiversity (Frans & Liu, 2024). SDMs predict species' habitat suitability by analysing the association between species occurrence and various environmental variables, allowing for the identification of suitable habitats both in the present and under future environmental change scenarios (Franklin, 2010; Guisan et al., 2013; Fourcade et al., 2014; Araujo et al., 2019).

Island endemic species are at particular risk of future climate change. The limited space available on islands provides few opportunities for species to shift their ranges to track changing climate, thereby increasing their vulnerability (Cox & Moore, 2010; Pimm, 2002; Taylor & Kumar, 2016). This threat spans various taxa and island environments. For instance, cold-adapted island endemic plants, as well as tropical insular reptiles, amphibians, and birds, are all expected to experience significant declines (up to 90-100%) in climate suitability, potentially leading to local and global extinctions even under optimistic predicted future scenarios (Dubos et al., 2022; Edwards et al., 2022; Fortini et al., 2015; Upson et al., 2016).

As primary producers and, in many cases, ecosystem engineers, plants play a crucial role in sustaining island biodiversity by providing essential services such as water regulation, erosion control, habitat, and food supply (Gravel et al., 2011; Borges et al., 2018). For instance, while certain areas on a given island may remain climatically suitable for some island birds, the loss of critical nesting habitat provided by plants could significantly reduce their chances of survival (Porfirio et al., 2016; Amin et al., 2021).

Given the importance of plants in sustaining island biodiversity, it is crucial to accurately determine their habitat suitability on islands in the face of climate change (Caujape-Castells et al., 2010; Patiño et al., 2023). This requires considering small-scale climatic variation, as oceanic islands often exhibit high environmental heterogeneity over relatively small areas, leading to diverse micro-habitats (Irl & Beierkuhnlein, 2011; Barajas-Barbosa et al., 2020; Whittaker et al., 2023; Chapter 2, Chapter 4). Consequently, the distributions of island endemic plant species are often closely related microclimatic variations, and such species often have narrow, clustered distributions (Nyakatya & McGeoch, 2008; Patiño et al., 2023) Therefore, coarse-grained environmental / climatic predictors are typically insufficient for capturing the necessary environmental variation required for effective SDMs (Benavides et al., 2024). Given the difficulties in sourcing fine-scale environmental data on islands (see Chapter 4), we still lack a comprehensive understanding of how island plant species distributions will respond to future climate change on different archipelagos. This understanding is essential if we are to effectively conserve island plant species going forward.

This study builds on the methodological foundations of the preceding three chapters to evaluate the potential impacts of future climate change on plants of high ecological and conservation value found in various (sub)tropical oceanic archipelagos, all of which are global biodiversity hotspots (Mittermeier et al., 2004).

To generate relevant information to support management decisions, this analysis aims to (1) examine how predicted distribution changes in response to future climate change vary across species chorotypes (i.e., non-endemic natives vs. archipelago endemic species) to identify whether local or global species losses will be more prevalent in the analysed areas; (2) assess how these predicted distribution changes vary across different species extinction risk categories (as recognised by global authorities such as the IUCN, and national authorities) to determine whether climate change poses an additional threat—or confirms an existing one—for species already under threat; and (3) determine whether the modelled changes in plant species distribution patterns result in significant losses or gains in specific island regions, in order to identify zones of loss or stability. By using microclimatic data in the SDMs, the aim is to improve predictions of species distributions under scenarios of future environmental change and, in doing so, inform conservation efforts in these vulnerable environments.

5.2 Materials and methods

5.2.1 Study area

This study focuses on four archipelagos of volcanic origin located in biodiversity hotspots in the tropics and sub-tropics: the Canary Islands, located in the Mediterranean Basin hotspot; the Galápagos archipelago, located in the Chocó-Darien/Western Ecuador hotspot; the Hawaiian Islands, located in the Polynesia/Micronesia hotspot; and the Revillagigedo archipelago, located in the Mesoamerica hotspot. These archipelagos were selected based on their (sub)tropical affinities, volcanic origin, recognised categorisation within the biodiversity hotspot framework, and the availability of ample and high-quality occurrence data in global biodiversity datasets (e.g., GBIF: Benavides et al., 2024).

5.2.2 Current Climatic Data

Global climatic data were downscaled following the approach outlined in Chapter 4, which incorporates the effects of topography in the study area using the methodology of Flint & Flint (2012). This interpolation method combines a spatial gradient and inverse-distance-squared (GIDS) weighting to the global dataset through multiple regression.

To apply this methodology, bioclimatic variables for the study area were initially obtained from the CHELSA database at the highest resolution available, 1-km/30 arc-sec (Karger et al., 2017). This global dataset was selected because it offers the finest resolution currently available for climate change scenarios, enabling more robust interpolation of finer grids (Flint & Flint, 2012). Additionally, it has proven effective in developing SDMs for plants in island regions with sparse and limited meteorological data (Chapter 4). Topographic information for downscaling was obtained from a digital elevation model at a 30m resolution (ASTER Global Digital Elevation Map v3: NASA & METI, 2018), which was later resampled to 500m resolution. The original 1-km resolution CHELSA data were subsequently downscaled to 500m. A resolution of 500 metres was chosen because most of the islands under analysis are relatively large, making this resolution informative (as demonstrated in Chapter 4), while striking an effective balance between maximising model informativeness and minimising computation time for both the downscaling process and the development of SDMs.

The downscaling methodology involves using gridded climatic data, initially at a coarse resolution, which are transferred to a finer-resolution grid using corresponding fine-resolution elevation data. To adjust the parameters for this finer resolution, the following equation is applied to weight the parameters

$$Z = \left[\sum_{i=1}^{N} \frac{Z_i + (X - X_i) * C_x + (Y - Y_i) * C_y + (E - E_i) * C_e}{d_i^2} \right] / \left[\sum_{i=1}^{N} \frac{1}{d_i^2} \right]$$

where Z is the estimated climatic variable at the specific location defined by easting (X) and northing (Y) coordinates and elevation (E); Z_i is the climatic variable from the 1-km grid cell i; X_i, Y_i, and E_i are easting and northing coordinates and elevation of the 1-km grid cell i, respectively; N is the number of 1-km grid cells in a specified search radius; C_x, C_y, and C_e are regression coefficients for easting, northing, and elevation, respectively; and d_i is the distance from the 500m to 1-km grid cell i (Flint & Flint, 2012).

As a final step, the multicollinearity of the 19 CHELSA downscaled bioclimatic variables was analysed using the 'corSelect' function from the "fuzzySim" R package (Barbosa & Barbosa, 2020). This function calculates the pairwise Pearson's correlation between variables, and if the correlation threshold ($r > 0.7$) is exceeded, one variable from the pair is excluded by selecting the variable with the highest variance inflation factor (VIF) in a linear regression model. A final set of variables for each group of islands was selected using this approach, ensuring reasonable multicollinearity levels (all VIFs were < 5) (Abdelaal et al., 2019). The lists of variables selected per archipelago are presented in the results section.

5.2.3 Future climatic data

To project future climate scenarios for the period 2041-2070, two climate models with different levels of climatic sensitivity were selected from the CHELSA dataset.

Climatic sensitivity refers to the degree to which Earth's temperature responds to rising greenhouse gas levels. The models chosen were UKESM1-0-LL (UK-mod) and IPSL-CM6A-LR (IP-mod). The UK-mod model is particularly important in the context of the present analyses as it detects substantial changes in the tropics (Swaminathan et al., 2022), making it a "worst-case scenario" for the study areas. This model was contrasted with the IP-mod model, which reflects less pronounced changes, allowing for a balanced depiction of potential future climates.

The period 2041-2070 was chosen because it represents a future timeframe that is far enough ahead to reflect significant changes yet not so distant as to be less relevant for current policy planning. The available alternatives, 2011-2040 and 2070-2100, were less suitable: the former includes past, present, and imminent future conditions, which offer limited new insights, while the latter extends too far into the future for practical application in present-day decision-making. The selected period strikes a balance, providing a forward-looking but still actionable timeframe for climate projection. Each model was associated with three shared socio-economic pathways (hereafter termed 'SSP'): SSP1-2.6, SSP 3–7.0 and SSP 5–8.5. These three Shared Socioeconomic Pathways (SSPs) provide varying projections of global socio-economic changes and mitigation measures for the near future, with global mean temperature increases expected to range from up to 2.4°C in SSP1-2.6 to 5.7°C in SSP5-8.5 (where CO_2 emissions reach zero by 2050 in the best-case scenario and triple by 2070 in the worst-case scenario) by 2100, depending on the mitigation measures implemented (O'Neill et al., 2016). We applied the same downscaling methodology used for present-day grids to these future climate scenarios. This equated to six future outputs per archipelago: two climatic models for each of the three concentration pathways, selecting only the non-collinear group of predictor variables resulting from the current climatic scenarios (i.e., all models for a given archipelago [i.e., current and future] are using the same predictor variables).

5.2.4 Species data

Using literature and regional databases for each archipelago, a list of priority species—those of high ecological value and/or classified under a recognised category of threat—was compiled to assess their vulnerability to climate change. To be included in the list, a species had to meet at least one of the following criteria:

1) Be categorised as threatened by either global biodiversity authorities (e.g., IUCN) or local authorities (e.g., government threat/vulnerability assessments). In regard to the IUCN Red List, threatened species were those classified as Vulnerable, Endangered or Critically Endangered (Matthews et al., 2022).

2) Be currently recognised as being under pressure from anthropogenic drivers of environmental change (e.g., habitat modification, disease, invasive species, wildfires). This category applies to species that are (i) not currently classified as

threatened but have been in the past and are still being monitored, and (ii) not yet classified as threatened but are experiencing a decline or fragmentation in their populations and/or habitats.

3) Have an identified specific important ecosystem function (e.g., providing nesting habitat, pollination resources, or edible fruits, or playing a dominant role in habitat stabilisation).

For the identified species, occurrence data were extracted from the GBIF database using the "rgbif" R package (Chamberlain et al., 2022), focusing on the selected archipelagos. Island polygons were obtained from the "Geodata" R package (Hijmans et al., 2023), covering islands ranging in size from 19 to 28,311 km^2. All taxa identified at the species level were included by searching with the 'Genus' and 'Specific epithet' functions. These functions retrieve all records associated with the genus and specific epithet categories within the specified polygon, ensuring that synonyms with corresponding occurrence data were included. The final database for each species was then curated by the "Accepted name” category within GBIF to unify species identity across synonyms. Finally, we filtered the data to retain only records with a location uncertainty of 500 meters or less.

To mitigate potential sampling bias, a widespread distribution of occurrence points was ensured by filtering clustered sampling localities using the "spThin" R package

(Aiello-Lammens et al., 2015). Prior to applying the filter, an analysis of the extent of occurrence for each species was conducted by calculating the minimum convex polygon around all occurrence points using the “adehabitatHR” R package. This was done to confirm that clustered localities did not correspond to species with very small ranges, which is common in island species. A very small range was defined as any extent of occurrence less than 10 km^2. For these species, occurrence points were not filtered. For species with ranges greater than 10 km^2, a filter of 500 m (matching the predictor variable resolution) was selected as the minimum distance to maintain between pairs of occurrence points. For more detailed information on this approach, see Chapter 3.

5.2.5 Modelling framework

Considering the availability of occurrence data and the number of variables per archipelago, models were fitted using two different approaches. Firstly, if the number of records resulted in a ratio of 10 points per variable (as recommended by Franklin, 2013: e.g., ≥50 occurrence points for 5 variables in the study area), a traditional model approach was applied, involving the inclusion of all variables to fit a single model. Conversely, if the recommended ratio of ten records per predictor variable was not met (e.g., 30 occurrence points for 5 variables in the study area), the ensemble of small models (ESM) approach was used (Lomba et al., 2010; Breiner et al., 2015). In this approach, all potential combinations of bivariate models are fitted, and their performance is evaluated based on their Area Under the Curve (AUC) scores. Subsequently, a final weighted averaged model is created, which includes all informative small models (AUC >0.5).

The implementation of ESM in these cases is preferred because it reduces the risk of overfitting and effectively controls model complexity. For both approaches, 5-fold cross-validation was performed, using 20% of occurrence points for evaluation and the remaining 80% for training. To validate model performance, both the AUC and the Boyce index were calculated using the 'modEVA' R package (Barbosa et al., 2013).

To minimise the uncertainty associated with model selection (Araújo et al., 2019), four different methodologies that have demonstrated reliable results for studying island species were applied, along with the ESM approach: GLM, GAM, Maxent, and BART (Chapter 3). Each ESM was individually fitted using each method, and the resulting models were then weighted by their AUC scores. Only models with an AUC >0.6 were included in the final ensemble. Modelling features were set to default, as evidence suggests that parameter tuning has little effect on predictive performance when sample sizes are small (Valavi et al., 2022). This is particularly relevant since the majority of our species have fewer than 30 records (Valavi et al., 2022; Radomski et al., 2022).

Since presence-only data were available for the analysed species, we simulated pseudo-absence/background (Psa/BG) data by quantifying the number of grids available within the study area and subsequently randomly sampling 10% of the study area (where presence points are not recorded). This random sampling approach relies on fewer assumptions and has been shown to yield highly accurate predictions based on prior research (Stokland et al., 2011; Descombes et al., 2023; Chapter 3). Additionally, the selected percentage has been shown to strike a good balance between real model performance and validation metrics in island SDMs (Chapter 3).

The differences in predicted probability outputs calibrated with diverse modelling methods can make them incomparable or incompatible for combination. Therefore, probability outputs were converted into favourability measures using the "Fuzzysim" R package. Favourability, which ranges from 0 to 1, distinguishes conditions supporting species presence ($F > 0.5$) from those detrimental to the species ($F < 0.5$), thereby standardising values across distribution models to enable comparisons (Acevedo & Real, 2012; Real et al., 2006).

5.2.6 Predicted climatic-driven range changes

To assess potential future distribution changes in response to climate change, maps of the current distribution were first created, focusing exclusively on suitable areas containing actual occurrence points. Only suitable patches with at least one recorded presence were retained. Even though this approach may exclude some suitable ~~areas~~ patches lacking presence data due to a lack of sampling, this trade-off was deemed acceptable to reduce overprediction and account for dispersal limitations and historical factors that may have prevented the species from occupying those areas (Anadón et al., 2015).

For future scenarios, a migration rate of 200 meters per year was incorporated, as suggested by Liao et al. (2020). To depict future climatic suitability, current suitable habitat patches were first identified, and a 10 km buffer was applied around these areas. If any future suitable patches overlapped with current patches within this buffer, they were retained as potential future habitat. This buffer accounts for the species' potential dispersal into areas projected to become climatically suitable, based on future climate models. While some unconnected areas beyond the buffer may also become suitable, it is unlikely that species will disperse to all such areas within the timeframe, given their dispersal limits (Burns, 2018).

By excluding unconnected patches, the aim was to avoid overestimating the species' realised distribution in present and future scenarios, as including such distant patches could result in unrealistic projections of range expansion or dispersal in future scenarios (Benavides et al., 2020).

Finally, differences between present and future suitability scenarios, were calculated using a Net Range Change Evaluation approach. For this, the threshold of >0.5 from the favourability (F) outputs was utilised to identify zones with suitable conditions for the species (without binarising the outputs).

The differences between present and future distributions were quantified, identifying (i) new areas with F > 0.5, defined as habitat gain, (ii) areas suitable in the present scenario (F > 0.5) but unsuitable in the future (F < 0.5), defined as habitat loss, and (iii) areas whose suitability remained within the threshold (0.5 - 1), defined as stable habitat. As a final step the relative change in the distribution area was calculated, using the following formula:

$$\left(\frac{\text{Climatically suitable area} - \text{suitable area loss} + \text{suitable area gain}}{\text{Climatically suitable area}}\right) - 1.$$

This formula calculates the net future distribution area by subtracting the original area loss from the original distribution area and adding the new area gain. The result is then divided by the original distribution area to provide a relative measure of change, with the final subtraction of 1 adjusting the result to reflect net change. A positive value indicates an overall gain in suitable area, a negative value indicates a loss, and a value of 0 indicates no net change.

A linear model was constructed for each metric to analyse the impact of three factors—Archipelago (Canaries, Hawaii, Galápagos, and Revillagigedo), future climate model (Fut_Model: UK and IP), and SSP (SSP1-2.6, SSP3-7.0, and SSP5-8.5)—on Net Range Change. The model included the main effects and all possible interactions among the three factors. An ANOVA was conducted on this model to assess the significance of each factor and their interactions. The formula used was:

$$\text{Net Range Change} \sim Archipelago \, * \, Fut\_Model \, * \, SSP$$

Finally, in line with previous studies (Beaumont et al., 2016; Rose et al., 2024), species distributional changes were categorised based on the global range of Net Range Change values (converted to percentages). Total Loss (TL: -100%), Extreme Loss (EL: >-80% and <-100%), Severe Loss (SVL: -80% to -50%), Moderate Loss (ML: -50% to 0%), and Gain (G: >0% to 19%, our highest value of habitat gain globally). These categories were analysed by archipelago, and their distribution was assessed across (a) three threat classifications: species of Least Concern (LC), species currently not assessed (NE), and species under some category of threat (THR), which included species categorised as Vulnerable, Endangered or Critically Endangered by IUCN (or its equivalent on a national risk categorisation); and (b) two species chorotypes: Endemic (EN) and Non-Endemic Natives (NEN).

5.2.7 Spatial patterns of species loss

The loss of habitat suitability for various species under future scenarios was assessed by comparing suitability maps for the present and future timeframes, focusing on identifying grid cells that exhibited a loss in suitability. Loss of suitability was defined based on the criteria used in the Net Range Change analysis. A grid cell was considered to have lost suitability if its favourability value (F) in the present was greater than 0.5 (indicating suitability) but dropped below 0.5 in the future (indicating loss of suitability).

To summarise the loss of suitability across multiple species, the 'mosaic' function from the "terra" R package was utilised. This function enabled the merging of individual species maps into a single ensemble map, indicating how many species experienced a loss of suitability in each grid cell. The final output of this process is a map that highlights the regions where habitat suitability has declined for a given number species.

5.3 Results

5.3.1 Bioclimatic datasets

The multicollinearity analysis of predictor variables revealed differences in the amount of multicollinearity between predictors, among the studied island groups. The Galápagos Islands had the highest number of non-collinear variables, with ten identified and thus selected for analysis. This was followed by the Canaries with seven non-collinear variables. Both the Revillagigedo and Hawaii island groups each had five non-collinear variables. The final set of variables used across all models comprised bio1 (mean annual air temperature), bio3 (isothermality), bio4 (temperature seasonality), bio5 (mean daily maximum air temperature of the warmest month), bio 6 (mean daily minimum air temperature of the coldest month), bio7 (annual range of air temperature), bio9 (mean daily mean air temperatures of the driest quarter), bio14 (precipitation amount of the driest month), bio15 (precipitation seasonality), bio16 (mean monthly precipitation amount of the wettest quarter), bio18 (mean monthly precipitation amount of the warmest quarter), and bio19 (mean monthly precipitation amount of the coldest quarter). Table 5.1 shows the variables selected to fit the SDMs for each archipelago.

Table 5.1 The bioclimatic predictor variables selected per archipelago based on a multicollinearity analysis.

	Bio 1	***Bio 3***	***Bio 4***	***Bio 5***	***Bio 6***	***Bio 7***	***Bio 9***	***Bio 14***	***Bio 15***	***Bio 16***	***Bio 18***	***Bio 19***
Canaries	*X*					*X*	*X*		*X*		*X*	*X*
Galápagos	*X*	*X*	*X*		*X*	*X*	*X*	*X*	*X*		*X*	*X*
Hawaii	*X*		*X*	*X*					*X*			*X*
Revillagigedo	*X*	*X*		*X*						*X*	*X*	

5.3.2 Species data

A total of 135 species across the study areas were identified as fulfilling the criteria for being considered priority species in their respective archipelagos. These species were assigned a distribution chorotype based on whether they are native or endemic to the island group. Of these, 100 had sufficient occurrence data for species distribution model (SDM) fitting (>6 records post-thinning). Specifically, in Hawaii, 32 priority species were identified, with 24 (Endemics=18; Natives=6) having sufficient information for SDM fitting. In the Galápagos, 34 priority species were identified, with 22 (Endemics=17; Natives=5) having adequate data. The Canaries had 47 priority species, of which 32 (Endemics=23; Natives=9) had sufficient information for SDM fitting. In Revillagigedo, all 22 priority species (Endemics=10; Natives=12) had sufficient data for SDM fitting. Supplementary Table S.5.1.1 presents the information used to select the species as indicators, the number of records before and after thinning, and their distribution chorotypes.

5.3.3 Species distribution models

A significant positive correlation was found between AUC and Boyce metrics (Pearson's $r = 0.5$, $P < 0.01$). Given this correlation, the AUC results are discussed from herein.

The model fitting results, as indicated by the Area Under the Curve (AUC), show that the classification performance was better than random for most species across the islands. Overall, while the models generally provided relatively good fits, some species posed classification challenges, reflected in their lower AUC values. In the Galápagos, three species—*Calandrinia galapagosa*, *Scalesia villosa*, and *Trigonopterum laricifolium*—were excluded from further analysis due to consistently low AUC values (below 0.6) in all scenarios. Similarly, in Revillagigedo, models for *Ficus cotinifolia* and *Guettarda insularis* consistently had low AUC values (<0.6), and these two species were thus excluded. In the Canary Islands, *Salix canariensis* was also removed for the same reason. Notably, the models for all species in Hawaii had AUC values above 0.6 across all scenarios.

After removing the aforementioned species from the analysis, the global mean AUC was 0.77, indicating relatively strong model performance overall. AUC values ranged from 0.6 to 0.95, with an interquartile range (IQR) showing consistency—25th percentile at 0.720, the median at 0.761, and the 75th percentile at 0.820.

When comparing individual islands, the Canary Islands showed the highest mean AUC of 0.797, with values ranging from 0.683 to 0.953 and an IQR of 0.741 to 0.850, reflecting robust model predictions. Hawaii followed with a mean AUC of 0.77, displaying a wider range of 0.60 to 0.93 and an IQR from 0.72 to 0.82. In the Galápagos, the mean AUC was 0.75, with values spanning 0.69 to 0.92 and an IQR from 0.71 to 0.75. Revillagigedo had the lowest mean AUC at 0.75, with a range of 0.62 to 0.92 and an IQR of 0.65 to 0.80.

5.3.4 Predicted climatic-driven range changes

Overall, climate change projections indicate a general pattern of negative Net Range Change across all analysed archipelagos (global results in Net Range Change are presented in Supplementary Table 5.1.2). The mean Net Range Change was slightly lower for the Galápagos and Revillagigedo Island groups compared to the others (Canaries: -37%; Hawaii: -35%; Galápagos: -27.1%; Revillagigedo: -27.5%). The results of the linear models assessing the effects of Archipelago, Climatic Model, SSP, and their interactions on Net Range Change are presented in Supplementary Table 5.1.3. The main effect of archipelago was statistically significant (F= 6.75, df=3, p<0.001), indicating that Net Range Change varied significantly across different archipelagos. The main effect of SSP was also significant (F= 3.46, df=32, p=0.03), suggesting that different trajectories of climate change mitigation influenced Net Range Change. However, neither the main effect of Future Model (F= 0.14, df=1, p=0.71) nor any of the interaction terms were statistically significant, indicating no strong evidence for interaction effects between these factors. Numerical results of the ANOVA are provided in Supplementary Table 5.14.

Additionally, the trends of impact of SSPs remained consistent across all archipelagos. There was a gradual increase in the negative Net Range Change corresponding to the impact of the SSPs for all archipelagos, except Revillagigedo where variation between SSPs was minimal. The variation across the selected climatic models (IP_mod, UK_mod) was also small, with the UK model showing slightly higher negative Net Range Change. Figure 5.1 summarises the results of Net Range Change across archipelagos, climatic models and SSP.

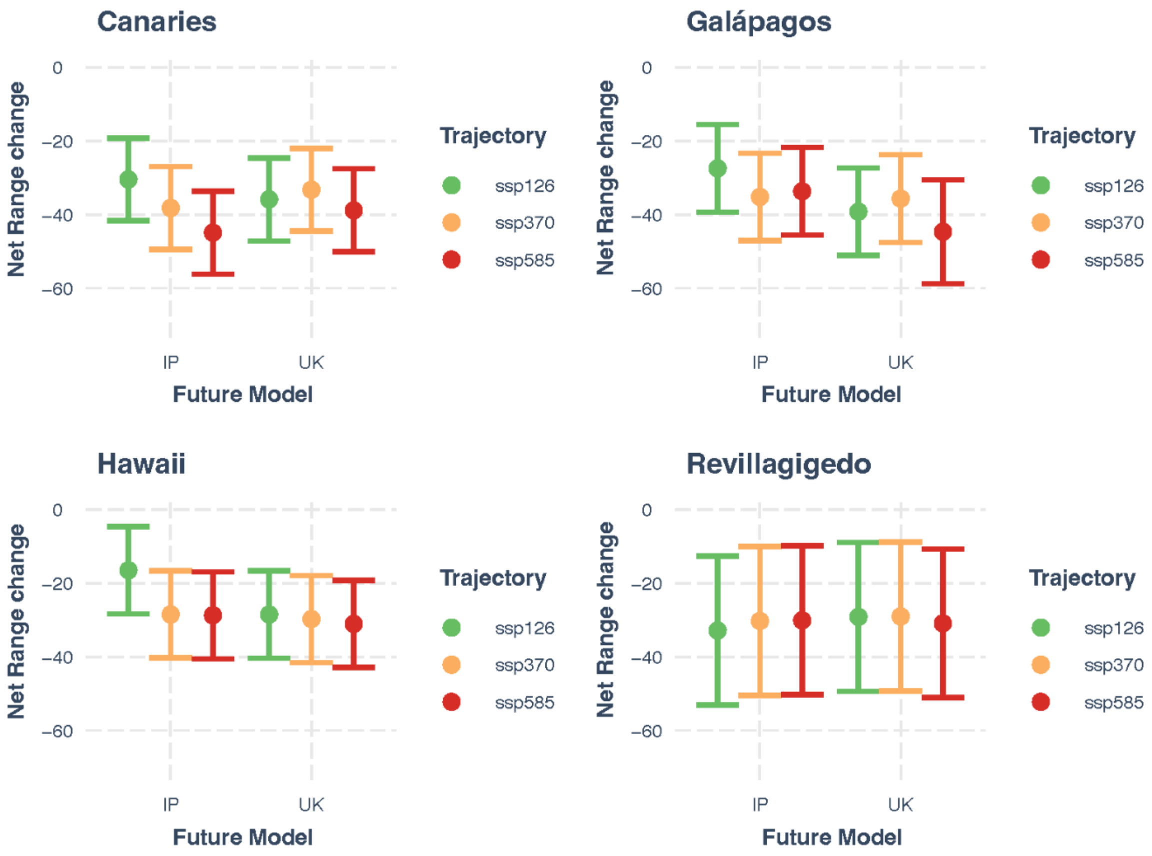


Figure 5.1 Projected Range Change Across Four Archipelagos Under Different Climate Scenarios. The figure shows the mean projected Net Range Change (converted to percentage; ± standard error) for species across four archipelagos: Canaries, Galápagos, Hawaii, and Revillagigedo. The projections are based on two

future climate models (IP and UK) and three SSP trajectories (SSP1-2.6, SSP3-7.0, and SSP5-8.5), represented by green, orange, and red points.

To analyse the frequency distribution of Net Range Change values, five categories were defined: Total Loss (TL= -100%), Extreme Loss (-100% < EL ≤ -80%), Severe Loss (-80% < SL ≤ -50%), Moderate Loss (-50% < ML ≤ 0%), and Gain (0% < G ≤ 19%, the highest value across all archipelagos). Since trends in Net Range Change vary across different archipelagos, the results are aggregated and presented at the archipelago level.

Overall, species with Moderate Loss is the predominant category across all archipelagos (Canaries: 55.7% of evaluated species in the archipelago were in this category, n=16; Galápagos: 55.5%, n=10; Hawaii: 50%, n=24; Revillagigedo: 26.3%, n=6). Species with Severe Loss are also particularly common in the Galápagos (33.4%, n=6) and Canaries (24.1%, n=7). The Canary Islands are the only archipelago with species that are classified as being threatened on the basis of climate change, specifically *Echium wildpretii*, *Heberdenia excelsa*, and *Sideroxylon canariense*. Of these, only *E. wildpretii* has predicted Extreme Loss due to climate change (Net Range Change: -84.5%), while the latter two species exhibit Moderate Loss (*H. excelsa*: -1.4%, *S. canariense*: -13.2%).

Notably, a higher number of species reporting Gains (G) are observed in Revillagigedo (35%, n=6) and Hawaii (25%, n=6), suggesting that some species in these regions may be resilient to future climate change. In contrast, Revillagigedo is the only archipelago reporting Total Losses (17% of species, n=3). Detailed information of the number of species per each category is available in Fig. 5.2 and in Supplementary table S5.1.2.

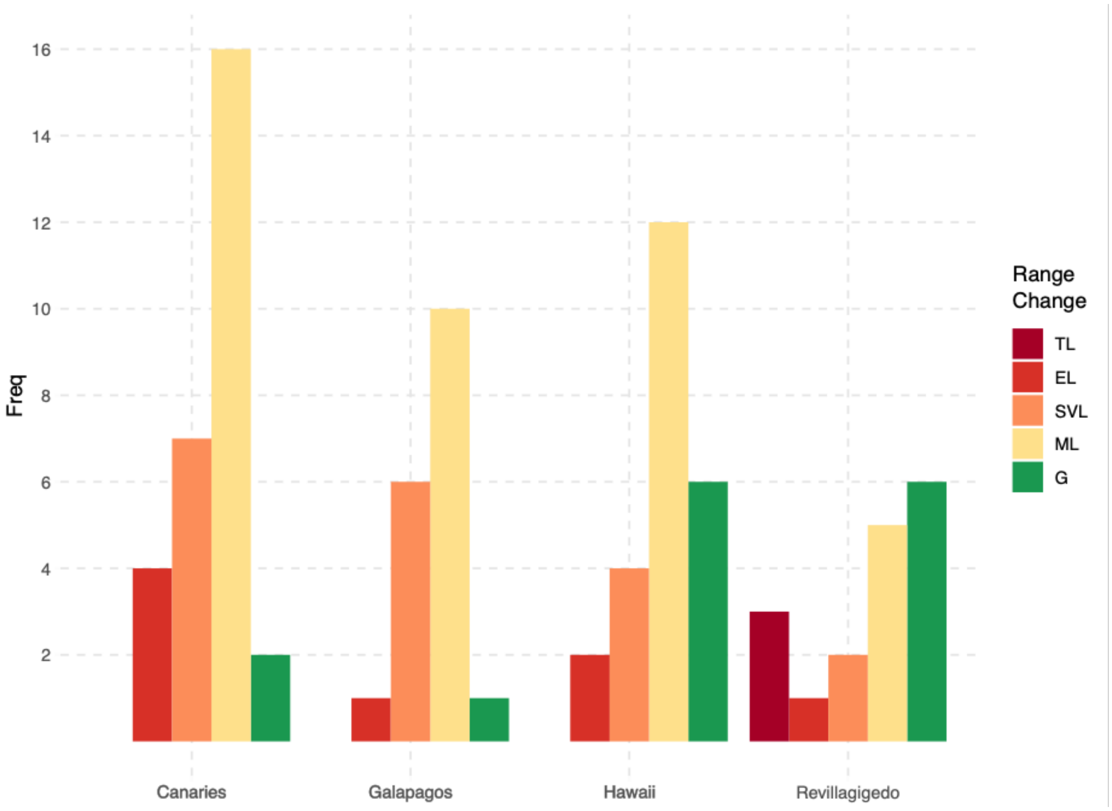


Figure 5.2 Frequency distribution of Net Range Change categories across the four archipelagos analysed. Categories are defined as Total Loss (TL: -100%), Extreme Loss (EL: >-80% and <-100%), Severe Loss (SVL: -80% to -50%), Moderate Loss (ML: -50% to 0%), and Gain (G: >0% to +19%).

The distribution of Net Range Change categories across three threat classifications was also assessed: species of Least Concern (LC), species not currently assessed (NE), and species under some category of threat (THR). Figure 5.3 illustrates the number of species in each threat group within each Net Range Change category. This section highlights species with higher proportions of Severe Loss (SVL) or worse. Overall, LC and NE species exhibit a greater proportion of SVL or worse compared to THR species.

In the LC category, a relatively high proportion of species with SVL or worse was observed, particularly in the Galápagos (33%, n=6), the Canaries (17%, n=4), and Revillagigedo (17%, n=3). For the NE category, the proportions were 23.5% (n=4) in Revillagigedo, 16% (n=4) in Hawaii, and 13% (n=4) in the Canaries. In contrast, the THR category had much lower proportions of species experiencing SVL or worse: 10.3% (n=3) in the Canaries, 5.5% (n=1) in the Galápagos, 8.3% (n=2) in Hawaii, and none in Revillagigedo.

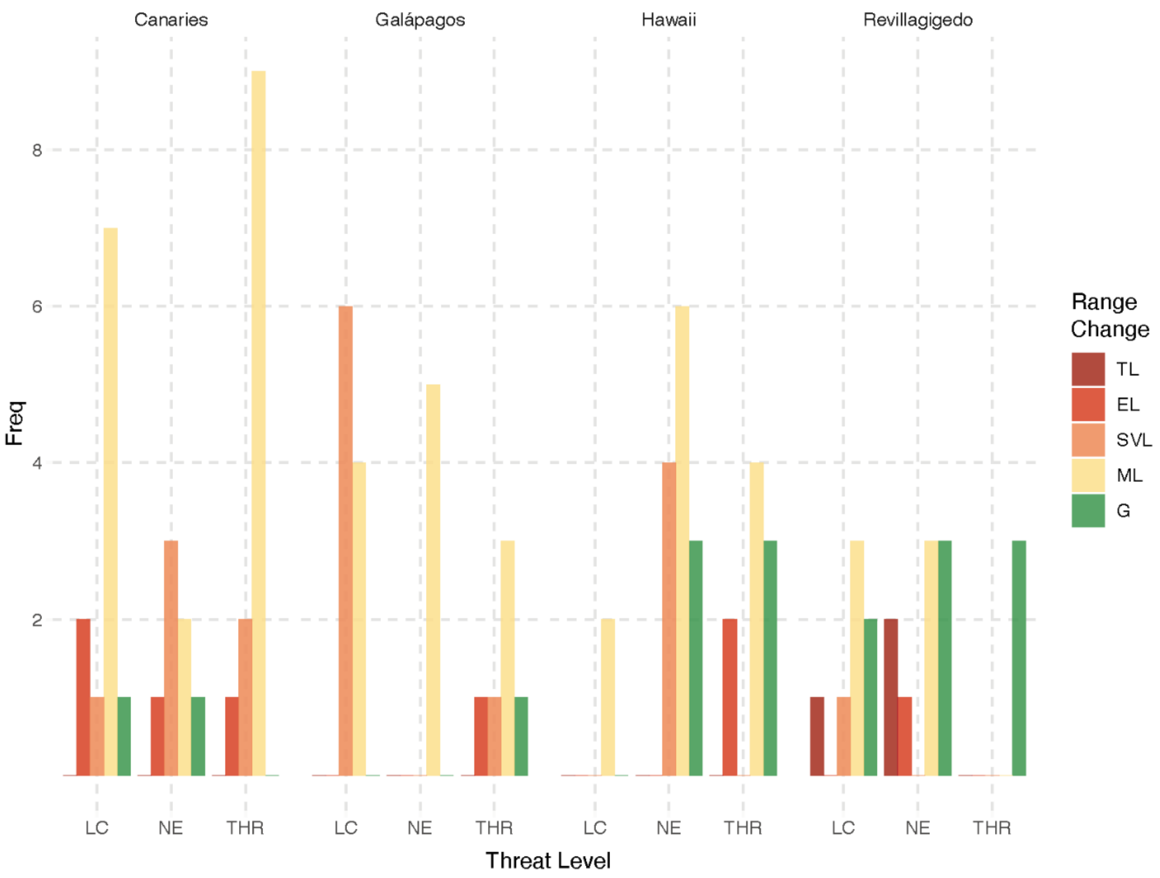


Figure 5.3 Frequency Distribution of Net Range Change Categories Across Three Risk Levels (LC: least concern, NE: Not evaluated, and THR: threatened) in the four Archipelagos analysed. Categories are defined as Total Loss (TL: -100%), Extreme Loss (EL: >-80% and <-100%), Severe Loss (SVL: -80% to -50%), Moderate Loss (ML: -50% to 0%), and Gain (G: >0% to +19%).

An analysis of Net Range Change categories was also conducted for two species chorotypes: Endemic (EN) and Non-Endemic Natives (NEN). The number of species in each range change category is illustrated in Figure 5.4. This section highlights species with higher proportions of Severe Loss (SVL) or worse. The results indicate that NEN species are generally predicted to experience higher proportions of SVL or worse compared to EN species in the Revillagigedo archipelago. However, the trend is reversed in the Canaries, Hawaii, and the Galápagos.

In the Canaries, 20.7% (n=6) of EN species are expected to experience SVL or worse, while 17.2% (n=5) of NEN species are predicted to exhibit similarly high losses. A similar pattern is observed in the Galápagos, where 27% (n=5) of EN species are expected to face SVL or worse, compared to 11% (n=2) of NEN species. In Hawaii, 21% (n=5) of EN species are predicted to experience SVL or worse, compared to only 4% (n=1) of NEN species showing significant range loss. However, Hawaii also exhibits distributional gains, particularly among EN species, with 17% (n=4) experiencing range expansion, compared to 8% (n=2) of NEN species.

Conversely, in Revillagigedo, 23% (n=4) of NEN species are projected to undergo SVL or worse, compared to 11.7% (n=2) of EN species. Notably, complete losses are reported only in Revillagigedo, all of which correspond to NEN species (n=3, 18%), suggesting these losses are local rather than global. Conversely, EN species show the highest proportion of range gains in this archipelago (29.4%, n=5).

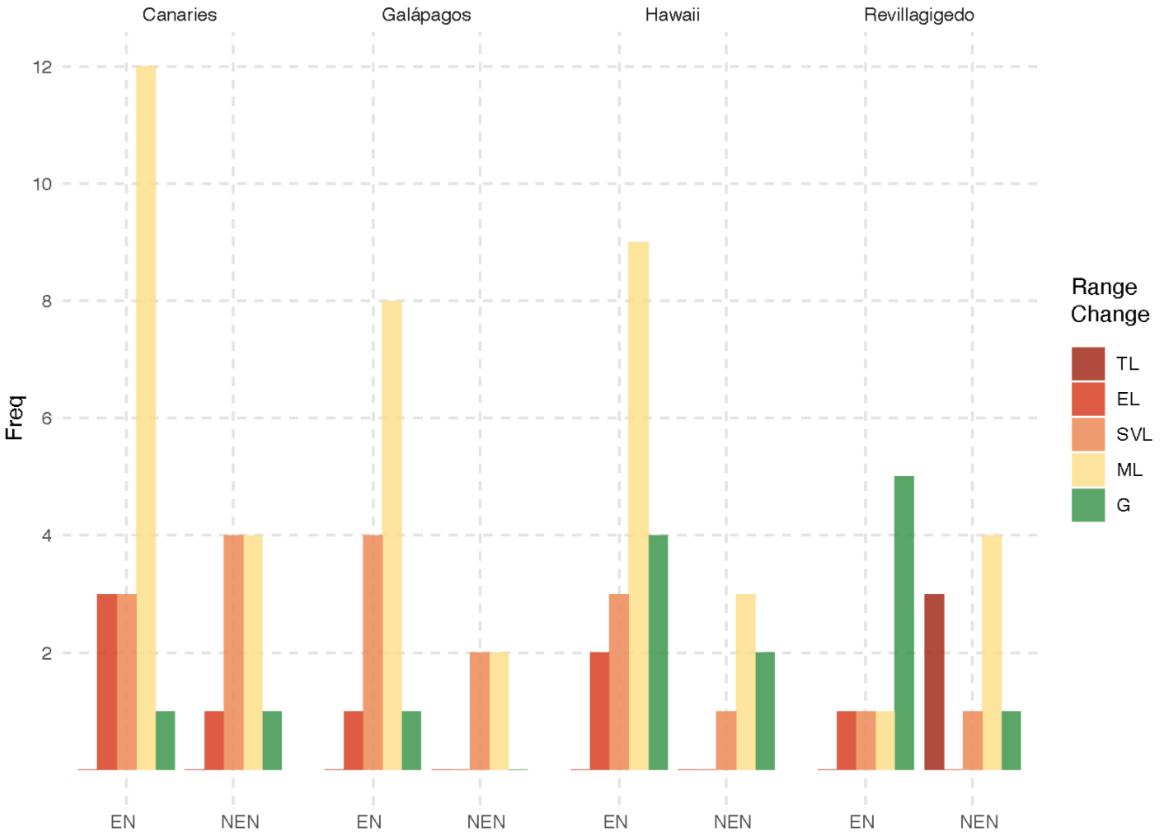


Figure 5.4 Frequency distribution of Net Range Change Categories across two distributional categories (EN: Endemic, and NEN: Non-Endemic Natives) in the four Archipelagos analysed. Categories are defined as Total Loss (TL: -100%), Extreme Loss (EL: >-80% and <-100%), Severe Loss (SVL: -80% to -50%), Moderate Loss (ML: -50% to 0%), and Gain (G: >0% to +19%).

5.3.5 Climatic-driven spatial patterns of species loss

Spatial assessment of the projected impacts of climate change on species distributions, averaged across the IP and UK models, reveals that patterns of species loss (i.e., loss of favourability on a given square) vary across archipelagos, with greater predicted loss in certain regions. Spatial patterns of species loss for the intermediate scenario of concentration pathway (SSP 370) are presented in Figure 5.5, while the detailed patterns for the other scenarios (SSP 126 & 585) are depicted in supplementary Figure S.51.5 and discussed in detail in Appendix S5.2.1.

In the Canaries we see, for the SSP370 trajectory, 24.96% of the grid squares remain unaffected, with 21.17% losing at least one species, and 12.99% losing ≥3 species (Fig 5.5). With the exception of Lanzarote and Fuerteventura (2 spp.), the rest of the islands in the archipelago show similar numbers of species losses (2-8 spp.; Fig. 5.5). The SSP126 and SSP585 scenarios show less and more severe effects, respectively, with species loss being more widespread under SSP585 (details in Appendix S5.2.1).

In the Galápagos, for the SSP370 trajectory, 35.65% of grid squares experience no species loss. In contrast, 31.88% lose at least one species, and 15.89% ≥2 species (Fig 5.5). Within the archipelago, Santa Cruz Island has the highest number of species losses (2-4 spp.), followed by Isabela and San Cristobal, both of which also show significant proportions of grids with notable species loss (1-4 spp.). The rest of the archipelago has lower (1 sp.) to no losses (Fig 5.5). As with the Canaries, the SSP126 shows lower impacts and the SSP585 more severe effects (see Appendix S5.2.1 for details).

In Hawaii, under the SSP370 trajectory, 17.51% of grid squares remain unaffected, with 27.91% losing at least one species, and 24.10% losing two or more species (Fig 5.5). Within the archipelago, Oahu has the highest number of grid squares with significant species losses (6-8 spp.), followed by Hawaii and Kauai, which also have high losses but in fewer grids. The rest of the islands are predicted to experience lower losses (1 spp.) or no losses at all (Fig 5.5). Less severe species loss is seen under SSP126, while SSP585 shows increased losses, with up to 13.52% of the area losing three or more species (see Appendix S5.2.1 for details).

Compared to the other three archipelagos, plant species on Revillagigedo appear to be more resilient to future climate change. For the SSP370 trajectory, this remains relatively stable, with 97.08% of the area unaffected. Unlike the other archipelagos analysed, species loss on Revillagigedo is not concentrated in specific zones; instead, it occurs across different regions of the islands (Fig 5.5). Species loss under SSP126 is minimal, while the SSP585 scenario shows a slight increase in species loss, affecting 19.42% of the grid squares (see Appendix S5.2.1 for details).

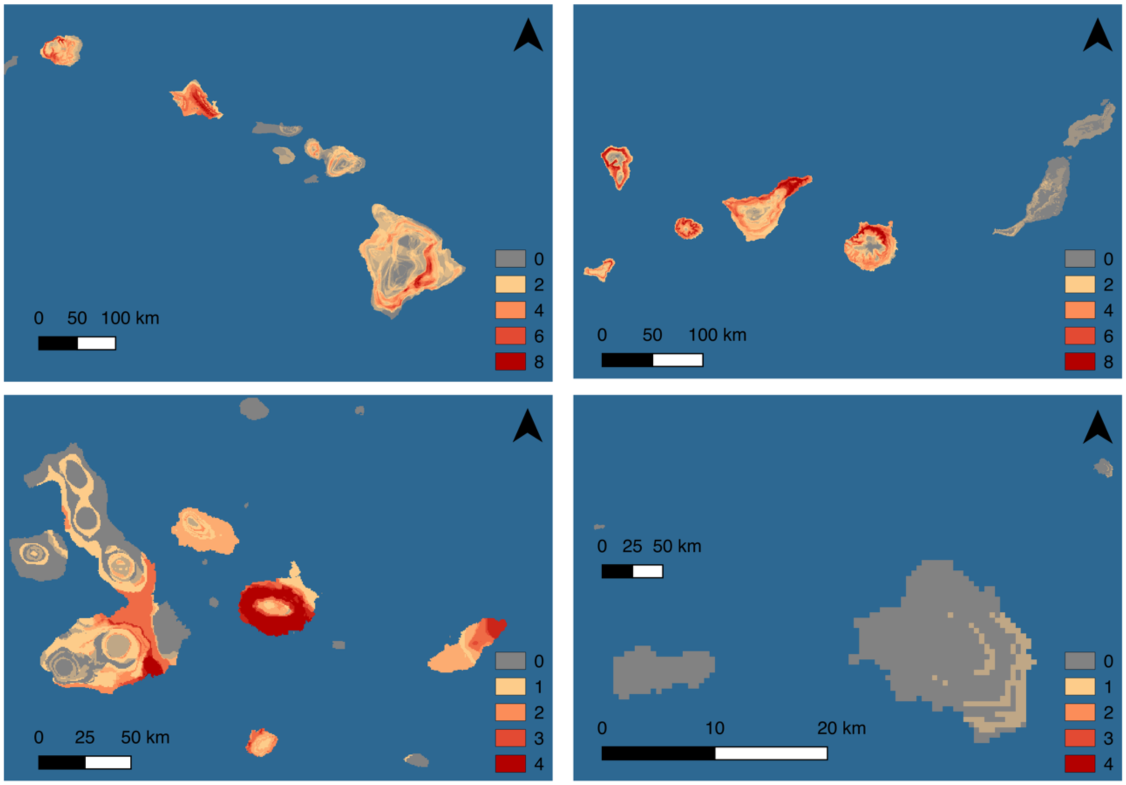


Figure 5.5 Spatial patterns of species loss by archipelago for the SSP 370 scenario (intermediate scenario) for: Hawaii (Top left), Canaries (Top Right), Galápagos (Bottom left) and Revillagigedo (Bottom Right). For Revillagigedo, the top section shows the actual distance and size of the islands, while the bottom section provides a magnified view. Increased colour intensity represents a higher number of species lost, measured as switching favourability within grid (500 m^2) squares.

5.4 Discussion

Reducing the impact of climate change on biodiversity is a key conservation goal, with efforts often focused on the most vulnerable species and ecosystems (Conradi et al., 2024). In this study, species distribution models (SDMs) were applied to insular plant species of high ecological and conservation value, using them as indicators of potential exposure to climate change in island biodiversity hotspots.

This approach provides valuable insights into how climate-driven species loss could affect distributions in these regions.

Although island ecosystems are vital to global biodiversity conservation efforts, the impacts of climate change on insular species are only beginning to be explored in depth (Veron et al., 2019; Benavides et al., 2024). This study significantly advances understanding of the vulnerability of island plant biodiversity. A key finding is that, despite variations across archipelagos, a concerning trend of range loss emerges. Species in all regions are projected to lose, on average, about one-third of their range, mirroring trends seen in other island hotspots (Pouteau & Birnbaum, 2016; Tagliari et al., 2021). Unlike most climate change studies on islands that use macroscale climate data (Benavides et al., 2024), this study is one of the first to incorporate high-resolution microclimatic variation, providing more accurate estimates of species' responses to climate change (see also Patiño et al., 2023).

Additionally, areas where multiple species experience overlapping range loss were identified, offering insights into spatial patterns that can inform more targeted and effective conservation strategies. Given the critical roles many of the studied species play within their ecosystems, these findings are especially important for maintaining both biodiversity and ecosystem functionality.

5.4.1 Impact of climate change varies across species’ conservation status

The IUCN criteria, while widely recognised for offering a comprehensive system for classifying species facing various threats, were not specifically designed to evaluate the impacts of climate change (Akçakaya et al., 2006; Keith et al., 2014). This study found that, although the IUCN framework effectively identifies species already at risk—making them ideal candidates for analysing climate change impacts—it has notable limitations in this regard. Many species, particularly those not yet evaluated or currently classified as of least concern, are expected to face severe climate-induced impacts, including the loss of more than half of their suitable ranges. This highlights the inadequacy of relying solely on the IUCN criteria to assess biodiversity risks from long-term threats like climate change.

In the analyses, only one out of three species (*Echium wildpretii* from Canary Islands) identified by the IUCN as threatened by climate change was predicted to lose the majority of its range (approximately 80%). This finding aligns with previous studies suggesting that the IUCN criteria need updating to better reflect the impacts of climate change in species assessments (Akçakaya et al., 2006; Keith et al., 2014). In accordance with Trull et al. (2018), it is emphasised that the IUCN requires a more integrated approach, incorporating specific traits that may increase species' vulnerability to climate change, as current guidelines lack these crucial considerations. Alternatively, as Akçakaya et al. (2006) suggest, existing criteria could be reassessed through the lens of climate change implications and applied to threats within this context to enhance the overall framework.

5.4.2 Impact of climate change varies across species' chorotype

Analysing the impacts of climate change on island endemic species is particularly crucial, as these species have global conservation significance. The loss of an insular endemic is permanent and represents a reduction in global biodiversity. Moreover, range restriction is a strong predictor of species extinction risk (Broennimann et al., 2006; Gaston & Fuller, 2009), making it unsurprising that insular species are especially vulnerable to climate change (Harter et al., 2015). In our study, endemic species exhibited relatively higher proportions of Severe to Extreme losses in suitable ranges (>50% suitability loss) across most of the archipelagos analysed, compared to their non-endemic counterparts, with the exception of the Revillagigedo Archipelago. This suggests a future dramatic decline in species with already small ranges, leading to heightened vulnerability from a global conservation perspective.

Conversely, while the Revillagigedo archipelago is predicted to experience greater losses among non-endemic species, and in some cases complete extirpations, it is essential to note that the species studied provide critical ecological functions in these islands. Although the extinction of these species may not represent a global loss, their disappearance could severely disrupt local ecosystems (Hautier et al., 2018), potentially destabilising ecological balance at the island level. This highlights the importance of directing conservation efforts at the local scale and developing strategies to prevent the extinction of vulnerable species. Additionally, it is crucial to note that, with the exception of the Revillagigedo Archipelago, all the archipelagos studied have human settlements. This makes it essential to consider not only the impacts of climate change but also the synergistic effects of anthropogenic threats to biodiversity (Benning et al., 2002; Ellis-Soto et al., 2017). Even for taxa not considered highly vulnerable to climate change, other factors—particularly habitat destruction caused by human activity and the introduction of invasive species—may pose immediate risks to their survival. Future island SDM research should examine the wide range of threats affecting island ecosystems, focusing on specific anthropogenic pressures and how they interact with climate change, to more effectively inform conservation strategies (Veron et al., 2019; Whittaker et al., 2023; Frans & Liu, 2024).

5.4.3 Spatial patterns of species loss

To ensure the long-term survival of biodiversity, conservation strategies must account for the spatial variability of climate change impacts (Jones et al., 2016). This approach enables the targeted prioritisation of mitigation efforts in areas where these impacts are most severe. In this study, a general trend of species loss was identified, driving the focus on pinpointing hotspots of this loss. Although the analysis did not systematically quantify the extent of these hotspots or their overlap with existing protected areas, a visual assessment indicates that most of these hotspots of species loss are roughly situated within protected natural areas in all the analysed archipelagos (Carralero, 2011; DPNG, 2014; CONANP, 2019; NARS, 2024). This suggests that while current conservation measures may help mitigate other anthropogenic threats, additional actions are likely needed to protect vulnerable species from the impacts of climate change. Future research could focus on gathering data to support the redefinition of protected area boundaries, ensuring the inclusion of climate refugia (Graham et al., 2019). Additionally, it is crucial to deepen our understanding of how the loss of these plant species may affect higher trophic levels, which could inform strategies such as relocating species to more favourable environments (Barlow et al., 2021).

5.4.4 Sources of uncertainty

Species distribution models are widely used to predict future species distributions, but they are not without limitations (Pearson & Dawson, 2003; Araújo & Luoto, 2007; Elith & Leathwick, 2009; Gotelli & Stanton-Geddes, 2015). For one, the reliability of SDM outputs is heavily influenced by the accuracy of climate projections and the choice of model, both of which are associated with considerable uncertainty.

To address and mitigate the uncertainties inherent in SDMs, several measures were implemented in the analyses tailored to the specific context of island environments. First, the impact of island topography on microclimatic variability was considered by employing a downscaling approach for both present and future bioclimatic variables. This adjustment helps account for the environmental variation that is known to drive island plant distributions (Chauvier et al., 2021; Patiño et al., 2024). Second, in depicting future scenarios, a diverse range of climatic models and projections was utilised. This approach reduces uncertainty by incorporating multiple potential climate trajectories, thereby providing a more comprehensive view of possible future conditions (Rose et al., 2024). Given these constraints of uncertainty in future projections, the need for caution in interpreting the results of future range changes is emphasised. Nonetheless, the consistency observed across various climatic models lends some robustness to the trends identified in this study.

Third, an ensemble of methods previously identified as effective for the specific ecological context of islands was employed (Chapter 3). By integrating different modelling approaches, the aim was to counteract potential biases that could arise from reliance on a single method (Rose et al., 2024). Fourth, for species with limited sample sizes, model complexity was reduced through a weighted ensemble of small models designed for small data sets (Lomba et al., 2010; Breiner et al., 2015;

Chapters 2 & 3). This strategy enhances model reliability despite the constraints of limited occurrence data. Finally, a sampling strategy of background/pseudo-absence points was selected that has previously demonstrated high prediction accuracy in island SDMs using presence-only data and specific to the selected SDM methods, further supporting the robustness of the model predictions (Chapter 3).

Through these measures, the aim was to minimise the uncertainties associated with SDMs and improve the reliability of predictions, providing more accurate insights into future species distributions in island ecosystems. However, as with all SDMs, the results should be interpreted as reflecting the species' suitable climate space within the analysed areas, rather than their full fundamental niches. This means some species' potential resilience to climate change may not be fully captured due to this limitation. Nonetheless, evidence from endemic island plants, where fundamental niche data are available (such as from populations outside their natural insular habitats), suggests that SDMs offer a close approximation of predicted impacts, even when based solely on insular occurrence data (Rosenblad et al., 2019).

5.5 Conclusions

This study enhances our understanding of the vulnerability of island plant biodiversity to climate change by incorporating microclimatic variation, resulting in more accurate predictions of species' responses compared to earlier macroscale models. The findings show a projected loss of suitable range area for plant species across various tropical island biodiversity hotspots, aligning with the findings of previous studies (Pouteau & Birnbaum, 2016; Tagliari et al., 2021). Additionally, the identification of hotspots of species' range loss provides critical insights that can inform the updating of protected area boundaries to incorporate climate change mitigation. Collectively, these results establish a valuable foundation for developing targeted conservation strategies on islands, which are crucial for maintaining biodiversity and ecosystem functionality in these fragile environments.

5.6 References

Abdelaal, M., Fois, M., Fenu, G., & Bacchetta, G. (2019). Using MaxEnt modeling to predict the potential distribution of the endemic plant *Rosa arabica* Crép. in Egypt. *Ecological Informatics*, *50*, 68-75.

Acevedo, P., & Real, R. (2012). Favourability: concept, distinctive characteristics and potential usefulness. *Naturwissenschaften*, *99*, 515-522.

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, *38*(5), 541-545.

Akçakaya, H. R., Butchart, S. H., Mace, G. M., Stuart, S. N., & Hilton-Taylor, C. R. A. I. G. (2006). Use and misuse of the IUCN Red List Criteria in projecting climate change impacts on biodiversity. *Global Change Biology*, *12*(11), 2037-2043.

Amin, R. J., Buettel, J. C., Fielding, M. W., Vaughan, P. M., & Brook, B. W. (2021). Hot, unpredictable weather interacts with land use to restrict the distribution of the Yellow-tailed Black-Cockatoo. *Emu-Austral Ornithology*, *121*(4), 323-332.

Anadón, J. D., Graciá, E., Botella, F., Giménez, A., Fahd, S., & Fritz, U. (2015). Individualistic response to past climate changes: niche differentiation promotes diverging Quaternary range dynamics in the subspecies of *Testudo graeca*. *Ecography, 38*(9), 956-966.

Arjona, Y., Nogales, M., Heleno, R., & Vargas, P. (2018). Long-distance dispersal syndromes matter: diaspore–trait effect on shaping plant distribution across the Canary Islands. *Ecography*, *41*(5), 805-814.

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

Araújo, M. B., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, *16*(6), 743-753.

Barlow, M. M., Johnson, C. N., McDowell, M. C., Fielding, M. W., Amin, R. J., & Brewster, R. (2021). Species distribution models for conservation: identifying translocation sites for eastern quolls under climate change. *Global Ecology and Conservation*, *29*, e01735.

Barajas-Barbosa, M. P., Weigelt, P., Borregaard, M. K., Keppel, G., & Kreft, H. (2020). Environmental heterogeneity dynamics drive plant diversity on oceanic islands. *Journal of Biogeography*, *47*(10), 2248-2260.

Barbosa A.M., Real, R., Munoz, A.R., Brown J. A. (2013). New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions*, 19(10), 1333–1338.

Benavides, E., Breceda, A., & Anadón, J. D. (2021). Winners and losers in the predicted impact of climate change on cacti species in Baja California. *Plant Ecology, 222*(1), 29-44.

Benavides, E., Sadler, J., Graham, L., & Matthews, T. J. (2024). Species distribution models and island biogeography: Challenges and prospects. *Global Ecology and Conservation*, e02943.

Benning, T. L., LaPointe, D., Atkinson, C. T., & Vitousek, P. M. (2002). Interactions of climate change with biological invasions and land use in the Hawaiian Islands: modeling the fate of endemic birds using a geographic information system. *Proceedings of the National Academy of Sciences, 99*(22), 14246-14249.

Beaumont, L. J., Graham, E., Duursma, D. E., Wilson, P. D., Cabrelli, A., Baumgartner, J. B., ... & VanDerWal, J. (2016). Which species distribution models are more (or less) likely to project broad-scale, climate-induced shifts in species ranges? *Ecological Modelling*, *342*, 135-146.

Borges, P. A., Cardoso, P., Kreft, H., Whittaker, R. J., Fattorini, S., Emerson, B. C., ... & Gabriel, R. (2018). Global Island Monitoring Scheme (GIMS): a proposal for the long-term coordinated survey and monitoring of native island forest biota. *Biodiversity and Conservation*, *27*, 2567-2586.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, *6*(10), 1210-1218.

Broennimann, O., Thuiller, W., Hughes, G., Midgley, G. F., Alkemade, J. R., & Guisan, A. (2006). Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Global Change Biology*, *12*(6), 1079-1093.

Burns, K. C. (2018). Time to abandon the loss of dispersal ability hypothesis in island plants: A comment on García-Verdugo, Mairal, Monroy, Sajeva and Caujapé-Castells (2017). *Journal of Biogeography*, *45*(6), 1219-1222.

Caujape-Castells, J., Tye, A., Crawford, D. J., Santos-Guerra, A., Sakai, A., Beaver, K., ... & Kueffer, C. (2010). Conservation of oceanic island floras: present and future global challenges. Perspectives in Plant Ecology, *Evolution and Systematics, 12*(2), 107-129.

Cardador, L., Díaz-Luque, J. A., Hiraldo, F., Gilardi, J. D., & Tella, J. L. (2018). The effects of spatial survey bias and habitat suitability on predicting the distribution of threatened species living in remote areas. *Bird Conservation International*, *28*(4), 581-592.

Carralero, I. (2001). La red Canaria de espacios naturales protegidos. In *Naturaleza de las islas Canarias: ecología y conservación* (pp. 399-405). Turquesa Ediciones.

Case, T. J., Case, T. J., Cody, M. L., & Ezcurra, E. (Eds.). (2002). *A new island biogeography of the Sea of Cortés*. Oxford University Press, Oxford.

Caujape-Castells, J., Tye, A., Crawford, D. J., Santos-Guerra, A., Sakai, A., Beaver, K., ... & Kueffer, C. (2010). Conservation of oceanic island floras: present and future global challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(2), 107-129.

Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, Ram K (2022). *rgbif: Interface to the Global Biodiversity Information Facility API*. R package version 3.7.

Channell, R., & Lomolino, M. V. (2000). Dynamic biogeography and conservation of endangered species. *Nature*, *403*(6765), 84-86.

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

Collen, B., Ram, M., Zamin, T., & McRae, L. (2008). The tropical biodiversity data gap: addressing disparity in global monitoring. *Tropical Conservation Science*, *1*(2), 75-88.

CONANP. (2019). Programa de Manejo: Parque Nacional Revillagigedo.

Conradi, T., Eggli, U., Kreft, H., Schweiger, A. H., Weigelt, P., & Higgins, S. I. (2024). Reassessment of the risks of climate change for terrestrial ecosystems. *Nature Ecology & Evolution, 8*, 888–900.

Cursach, J., Far, A. J., & Ruiz, M. (2020). Geospatial analysis to assess distribution patterns and predictive models for endangered plant species to support management decisions: a case study in the Balearic Islands. *Biodiversity and Conservation*, *29*(11), 3393-3410.

Da Re, D., Tordoni, E., De Pascalis, F., Negrín-Pérez, Z., Fernández-Palacios, J. M., Arévalo, J. R., ... & Bacaro, G. (2020). Invasive fountain grass (*Pennisetum setaceum* (Forssk.) Chiov.) increases its potential area of distribution in Tenerife Island under future climatic scenarios. *Plant Ecology*, *221*(10), 867-882.

Deb, C. R., Jamir, N. S., & Kikon, Z. P. (2017). Distribution prediction model of a rare orchid species (Vanda bicolor Griff.) using small sample size. *American Journal of Plant Sciences*, *8*(06), 1388.

del Arco Aguilar, M. J., González-González, R., Garzón-Machado, V., & Pizarro-Hernández, B. (2010). Actual and potential natural vegetation on the Canary Islands and its conservation status. *Biodiversity and Conservation*, *19*, 3089-3140.

De la Luz, J. L. L., BRECEDA, A., & Benet, R. C. (1996). Las comunidades vegetales en la isla Socorro, México. *SIDA, Contributions to Botany*, 215-230.

Descombes, P., Chauvier, Y., Brun, P., Righetti, D., Wüest, R. O., Karger, D. N., ... & Zimmermann, N. E. (2022). Strategies for sampling pseudo-absences for species distribution models in complex mountainous terrain. *BioRxiv*, 2022-03.

Diamond, J. M., Terborgh, J., Whitcomb, R. F., Lynch, J. F., Opler, P. A., Robbins, C. S., ... & Abele, L. G. (1976). Island biogeography and conservation: strategy and limitations. *Science*, *193*(4257), 1027-1032.

DNPG. (2014). Plan de Manejo de la Áreas Protegidas de Galápagos para el Buen Vivir. Puerto Ayora, Isla Santa Cruz- Galápagos: Dirección del Parque Nacional Galápagos, Puerto Ayora-Galápagos.

Dubos, N., Montfort, F., Grinand, C., Nourtier, M., Deso, G., Probst, J. M., ... & Crottini, A. (2022). Are narrow-ranging species doomed to extinction? Projected dramatic decline in future climate suitability of two highly threatened species. *Perspectives in Ecology and Conservation*, *20*(1), 18-28.

Edwards, W. M., Bungard, M. J., Rakotondrasoa, E. F., Razafindraibe, P., Andriantsimanarilafy, R. R., Razafimanahaka, J. H., & Griffiths, R. A. (2022). Predicted impact of climate change on the distribution of the Critically Endangered golden mantella (Mantella aurantiaca) in Madagascar. *The Herpetological Journal*, *32*(1), 5-13.

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*, 677-697.

Ellis-Soto, D., Blake, S., Soultan, A., Guézou, A., Cabrera, F., & Lötters, S. (2017). Plant species dispersed by Galapagos tortoises surf the wave of habitat suitability under anthropogenic climate change. *PLoS One*, *12*(7), e0181333.

Fernández-Palacios, J. M., Kreft, H., Irl, S. D., Norder, S., Ah-Peng, C., Borges, P. A., ... & Drake, D. R. (2021). Scientists' warning–The outstanding biodiversity of islands is in peril. *Global Ecology and Conservation*, *31*, e01847.

Ferreira, M. T., Cardoso, P., Borges, P. A., Gabriel, R., de Azevedo, E. B., Reis, F., ... & Elias, R. B. (2016). Effects of climate change on the distribution of indigenous species in oceanic islands (Azores). *Climatic Change*, *138*(3), 603-615.

Flint, L. E., & Flint, A. L. (2012). Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, *1*, 1-15.

Fortini, L. B., Vorsino, A. E., Amidon, F. A., Paxton, E. H., & Jacobi, J. D. (2015). Large-scale range collapse of Hawaiian forest birds under climate change and the need 21st century conservation options. *PloS one*, *10*(10), e0140389.

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

Franklin, J. (2010). Mapping species distributions: *Spatial inference and prediction.* Cambridge University Press, Cambridge.

Franklin, J. (2013). Species distribution models in conservation biogeography: developments and challenges. *Diversity and Distributions, 19*(10), 1217-1223.

Frans, V. F., & Liu, J. (2024). Gaps and opportunities in modelling human influence on species distributions in the Anthropocene. *Nature Ecology & Evolution, 8*(7), 1365-137

Fungjanthuek, J., Huang, M. J., Hughes, A. C., Huang, J. F., Chen, H. H., Gao, J., & Peng, Y. Q. (2022). Ecological niche overlap and prediction of the potential distribution of two sympatric *Ficus* (Moraceae) species in the Indo-Burma Region. *Forests*, *13*(9), 1420.

Gotelli, N. J., & Stanton-Geddes, J. (2015). Climate change, genetic markers and species distribution modelling. *Journal of Biogeography*, *42*(9), 1577-1585.

Galápagos Species Database, Linum cratericola", dataZone. Charles Darwin Foundation, https://datazone.darwinfoundation.org/es/checklist/?species=562. [Accessed 28 June 2024].

Gaston, K. J., & Fuller, R. A. (2009). The sizes of species' geographic ranges. *Journal of Applied Ecology*, *46*(1), 1-9.

Graham, V., Baumgartner, J. B., Beaumont, L. J., Esperón-Rodríguez, M., & Grech, A. (2019). Prioritizing the protection of climate refugia: designing a climate-ready protected area network. *Journal of Environmental Planning and Management*, *62*(14), 2588-2606.

Gravel, D., Massol, F., Canard, E., Mouillot, D., & Mouquet, N. (2011). Trophic theory of island biogeography. *Ecology Letters*, *14*(10), 1010-1016.

Guézou, A., Trueman, M., Buddenhagen, C. E., Chamorro, S., Guerrero, A. M., Pozo, P., & Atkinson, R. (2010). An extensive alien plant inventory from the inhabited areas of Galápagos. *PLoS One*, *5*(4), e10276.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., ... & Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424-1435.

Harter, D. E., Irl, S. D., Seo, B., Steinbauer, M. J., Gillespie, R., Triantis, K. A., ... & Beierkuhnlein, C. (2015). Impacts of global climate change on the floras of oceanic islands–Projections, implications and current knowledge. *Perspectives in Plant Ecology, Evolution and Systematics*, *17*(2), 160-183.

Hautier, Y., Isbell, F., Borer, E. T., Seabloom, E. W., Harpole, W. S., Lind, E. M., ... & Hector, A. (2018). Local loss and spatial homogenization of plant diversity reduce ecosystem multifunctionality. *Nature Ecology & Evolution*, *2*(1), 50-56.

Henareh Khalyani, A., Gould, W. A., Falkowski, M. J., Muscarella, R., Uriarte, M., & Yousef, F. (2019). Climate change increases potential plant species richness on Puerto Rican uplands. *Climatic Change*, *156*, 15-30.

Hijmans RJ, Barbosa M, Ghosh A, Mandel A (2024). *geodata: Download Geographic Data*. R package version 0.6-2

Irl, S. D., & Beierkuhnlein, C. (2011). Distribution of endemic plant species on an oceanic island–a geospatial analysis of La Palma (Canary Islands). *Procedia Environmental Sciences*, *7*, 170-175.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*(1), 1-20.

Keith, D. A., Mahony, M., Hines, H., Elith, J., Regan, T. J., Baumgartner, J. B., ... & Akçakaya, H. R. (2014). Detecting extinction risk from climate change by IUCN Red List criteria. *Conservation Biology*, *28*(3), 810-819.

Laurance, W. F. (2009). Beyond island biogeography theory. In *The theory of island biogeography revisited* (pp. 214-236). Princeton University Press, Princeton.

Liao, Z., Zhang, L., Nobis, M. P., Wu, X., Pan, K., Wang, K., ... & Tian, X. (2020). Climate change jointly with migration ability affect future range shifts of dominant fir species in Southwest China. *Diversity and Distributions*, *26*(3), 352-367.

Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, *143*(11), 2647-2657.

Matthews, T. J., Wayman, J. P., Cardoso, P., Sayol, F., Hume, J. P., Ulrich, W., ... & Triantis, K. A. (2022). Threatened and extinct island endemic birds of the world: Distribution, threats and functional diversity. *Journal of Biogeography, 49*(11), 1920-1940.

Mittermeier, R.A., Gil, P.R., Hoffman, M., Pilgrim, J., Brooks, T., Mittermeier, C.G., Lamoreux, J. & da Fonseca, G.A.B. (2004) *Hotspots revisited: earth's biologically richest and most endangered terrestrial ecoregions*. University of Chicago Press, Chicago, IL.

Natural Area Reserves System. (2024). https://dlnr.hawaii.gov/ecosystems/nars/ [accessed September, 2024]

Nyakatya, M. J., & McGeoch, M. A. (2008). Temperature variation across Marion Island associated with a keystone plant species (*Azorella selago* Hook.(Apiaceae)). *Polar Biology*, *31*, 139-151.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... & Sanderson, B. M. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461-3482.

Otto, R., García-del-Rey, E., Méndez, J., & Fernández-Palacios, J. M. (2012). Effects of thinning on seed rain, regeneration and understory vegetation in a *Pinus* *canariensis* plantation (Tenerife, Canary Islands). *Forest Ecology and Management*, *280*, 71-81.

Patiño, J., Collart, F., Vanderpoorten, A., Martin-Esquivel, J. L., Naranjo-Cigala, A., Mirolo, S., & Karger, D. N. (2023). Spatial resolution impacts projected plant responses to climate change on topographically complex islands. *Diversity and Distributions*, *29*(10), 1245-1262.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, *12*(5), 361-371.

Porfirio, L. L., Harris, R. M., Stojanovic, D., Webb, M. H., & Mackey, B. (2016). Projected direct and indirect effects of climate change on the Swift Parrot, an endangered migratory species. *Emu-Austral Ornithology*, *116*(3), 273-283.

Pouteau, R., & Birnbaum, P. (2016). Island biodiversity hotspots are getting hotter: vulnerability of tree species to climate change in New Caledonia. *Biological Conservation*, *201*, 111-119.

Radomski, T., Beamer, D., Babineau, A., Wilson, C., Pechmann, J., & Kozak, K. H. (2022). Finding what you don’t know: testing SDM methods for poorly known species. *Diversity and Distributions, 28*(9), 1769-1780.

Real, R., Barbosa, A. M., & Vargas, J. M. (2006). Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, *13*, 237-245.

Rose, M. B., Velazco, S. J. E., Regan, H. M., Flint, A. L., Flint, L. E., Thorne, J. H., & Franklin, J. (2024). Uncertainty in consensus predictions of plant species' vulnerability to climate change. *Diversity and Distributions*, *30*(8), e13898.

Rosenblad, K. C., Perret, D. L., & Sax, D. F. (2019). Niche syndromes reveal climate-driven extinction threat to island endemic conifers. *Nature Climate Change*, *9*(8), 627-631.

Stokland, J. N., Halvorsen, R., & Støa, B. (2011). Species distribution modelling—Effect of design and sample size of pseudo-absence observations. *Ecological Modelling*, *222*(11), 1800-1809.

Summers, D. M., Bryan, B. A., Crossman, N. D., & Meyer, W. S. (2012). Species vulnerability to climate change: impacts on spatial conservation priorities and species representation. *Global Change Biology*, *18*(7), 2335-2348.

Swaminathan, R., Parker, R. J., Jones, C. G., Allan, R. P., Quaife, T., Kelley, D. I., ... & Walton, J. (2022). The physical climate at global warming thresholds as seen in the UK Earth System Model. *Journal of Climate*, *35*(1), 29-48.

Tagliari, M. M., Danthu, P., Leong Pock Tsy, J. M., Cornu, C., Lenoir, J., Carvalho-Rocha, V., & Vieilledent, G. (2021). Not all species will migrate poleward as the climate warms: The case of the seven baobab species in Madagascar. *Global Change Biology, 27*(23), 6071-6085.

Trull, N., Böhm, M., & Carr, J. (2018). Patterns and biases of climate change threats in the IUCN Red List. *Conservation Biology*, *32*(1), 135-147.

Upson, R., Williams, J. J., Wilkinson, T. P., Clubbe, C. P., Maclean, I. M., McAdam, J. H., & Moat, J. F. (2016). Potential impacts of climate change on native plant distributions in the Falkland Islands. *PloS one*, *11*(11), e0167026.

Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs*, 92(1), e01486.

Wan, J. N., Mbari, N. J., Wang, S. W., Liu, B., Mwangi, B. N., Rasoarahona, J. R., ... & Wang, Q. F. (2021). Modeling impacts of climate change on the potential distribution of six endemic baobab species in Madagascar. *Plant Diversity*, *43*(2), 117-124.

Warren DL, Glor RE, Turelli M (2008). Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution, 62*, 2868–2883.

Whittaker, R. J., Fernández-Palacios, J. M., & Matthews, T. J. (2023). *Island biogeography: geo-environmental dynamics, ecology, evolution, human impact, and conservation*. Oxford University Press, Oxford.

6 . Synthesis and future perspectives

6.1 Research summary

This thesis aimed to deepen our understanding of Species Distribution Modelling (SDM), with a particular focus on the unique spatial and ecological complexities of island environments. Recognising the underutilised potential of SDMs in island studies, this research highlights their value as a tool for understanding biogeographical patterns on islands and linking this knowledge to effective conservation management.

Drawing on the existing literature on SDMs in an island context, this research identified the need for a tailored approach to SDM implementation for islands (Chapter 2). One of the primary findings relates to the importance of adjusting the resolution of predictor variables to effectively capture fine-scale environmental variation, which is crucial for accurate species distribution predictions in island settings. The shift from using coarse climate variables to fine-scale microclimatic factors raised important questions about how these detailed predictors interact with limited species’ occurrence data—another significant challenge identified in island occurrence datasets (Chapter 2).

The thesis explored a range of strategies aimed at improving the implementation of Species Distribution Models (SDMs) in island environments, focusing on the Revillagigedo Volcanic Archipelago in Mexico, referred to as the 'model archipelago' hereafter (Chapters 3 and 4). The goal was to generalise findings from this regional case study to other global island groups, using plant species presence/absence data collected through field sampling.

Plant species were chosen as the focal taxon due to their critical ecological roles, including the provision of essential services, such as habitat stabilisation and trophic resources for other organisms (Borges et al., 2018). The effort to improve SDM implementation in island ecosystems had two primary objectives: (1) to enhance the application of SDMs by incorporating high-resolution environmental descriptors, addressing data limitations and the challenges of small study areas (Chapter 3); and (2) after optimising SDM fitting at high resolutions in island environments, assessing the robustness of fine-scale predictor variables, particularly their wider relevance to tropical island ecosystems (Chapter 4).

Chapter 4 focused on the model archipelago's plant diversity. The chapter identified the scarcity of geographical species data, coupled with high uncertainty in environmental data due to a limited number of meteorological stations (Deblauwe et al., 2016; Benavides et al., 2024), which complicates the use of SDMs in these regions. To address this, the thesis explored methods to generate new high-resolution estimates (c.f. Flint & Flint, 2012) and assess their usefulness, thereby enhancing the availability of detailed data. These efforts confirmed the utility of such data for advancing scientific understanding and supporting more effective, data-driven conservation strategies in tropical island ecosystems.

After optimising SDM fitting for island environments and confirming the reliability of microclimatic estimates, the thesis focused on the practical application of the science to conservation (Chapter 5). The fifth chapter highlighted the potential of well-designed SDMs in island ecosystems, particularly in the context of climate change—a predicted driver of future species extinctions in these regions (Taylor & Kumar, 2016; Veron et al., 2019). The chapter pursued three main objectives: (1) to examine species distributional changes in relation to their current vulnerability status, assessing whether climate change exacerbates the risk for already threatened species; (2) to evaluate whether endemic species are more severely affected by range losses due to climate change, reflecting broader global trends; and (3) to identify hotspots of species loss to inform spatially-targeted mitigation strategies. This research represents a significant advancement in understanding how species may respond to climate change on islands at microclimatic scales, offering a critical baseline for developing more precise mitigation strategies for tropical island biodiversity hotspots.

6.2 Synthesis

Eight questions related to the implementation of SDMs focusing on the unique spatial and ecological complexities of island environments were addressed in this thesis. I return to each in turn to evaluate the findings and implications of the research:

6.2.1 Using previous research as a baseline, what are the challenges of SDM implementation in island environments, and how can we improve them?

While the potential of SDMs for informing data-driven conservation and management decisions, especially in relation to highly threatened island biodiversity has been noted previously (Fois et al., 2018; Mousikos et al., 2021), most SDM frameworks and recommendations have been developed for species with relatively wide ranges and are unsuitable for many insular species and environments, which often possess distinct characteristics (Leroy, 2022; Whittaker et al., 2023).

A review of 224 papers that implemented SDMs in island environments highlighted the absence of an appropriate SDM framework for island studies (Benavides et al., 2024). This gap has led to sub-optimal methodological decisions, particularly in the SDM fitting process. Due to the limited occurrence data often available for island species—whether due to their naturally small populations, anthropogenic-driven population declines, or difficulties in accessing remote islands—these models frequently rely on small sample sizes. This, in turn, requires the careful selection of algorithms and a reduction in model complexity, especially regarding the choice of environmental variables. Including too many variables relative to occurrence points increases the risk of overfitting, a problem that has been well documented in the island SDM literature (Dormann et al., 2012; Radomski et al., 2022; Rousseau & Betts, 2022).

Additionally, most environmental variables used in island SDMs are bioclimatic or macroclimatic, leading to a two-fold problem: 1) the number of islands that can be analysed is limited to those with sufficient macroclimatic grid data, and 2) such variables are often poor descriptors of the distributions of island species, which are often (i) dependent on microhabitat-level conditions (Patiño et al., 2023; Benavides et al., 2024) and (ii) characterised by highly specialised ecological requirements (Whittaker et al., 2023). This reliance on macroclimatic variables likely arises from the mainstreaming of continental models, where macroclimate plays a significant role in species distributions (Booth, 2022).

To improve SDM implementation in island studies, this thesis proposes careful consideration when working with small sample sizes. First, it is crucial to select appropriate modelling methods that account for limited data. Also, rather than relying on traditional models that input all variables into a single model, an ensemble of small models should be implemented to reduce the predictor-to-occurrence data ratio, helping to manage model complexity (Lomba et al., 2010; Beriner et al., 2015). Additionally, it is important to carefully select the extent of occurrence data. Non-endemic island species are often modelled using only insular data, which can truncate their ecological niche, undermining scientific questions that require an understanding beyond the realised niche (Chapter 4).

6.2.2 What sample size of occurrence (presence) data is needed to accurately predict the distributions of island species using high-resolution predictors, while accounting for varying specialisation levels?

This thesis proposes shifting to smaller analytical scales to better capture the relationships between island species and their environments. While these finer spatial resolutions provide a more detailed and accurate reflection of island environmental variation, they also present a challenge given the limited availability of species observation data. Smaller grid cells enhance model precision, but the combination of high specificity per grid and limited data points makes it more difficult to fully capture a species' environmental requirements, potentially reducing the effectiveness of this approach.

Previous research at continental scales (van Proosdij et al., 2016) has demonstrated that range-restricted species, such as those found on islands, generally require smaller sample sizes. However, it was essential to explore whether this relationship held at finer spatial resolutions. Contrary to the results of previous work on continents, the analyses in Chapter 3 revealed that range-restricted species in the model archipelago require larger sample sizes in smaller study areas with high-resolution predictors, while wide-ranging species maintained high predictability even with minimal data. These conflicting results may be because larger continental study areas, with greater environmental gradients, require more data to accurately model widely distributed species (van Proosdij et al., 2016). Conversely, in smaller insular areas with relatively flatter environmental gradients, species with broader ecological tolerances can be predicted accurately with fewer data points. This suggests that current modelling methods tend to overgeneralise, and overpredict SDMs for narrow-ranged species. As a result, more data are needed to confirm their specialised habitat requirements (Collins et al., 2017; Mendes et al., 2020).

Evaluating the performance of different SDM algorithms was also a critical focus of the thesis, given that evidence suggests that certain methods are better suited for studying range-restricted species (Breiner et al., 2015). However, modelling such species remains problematic due to their inherent rarity. This challenge, referred to as the rare-species modelling paradox (Lomba et al., 2010), highlights the difficulty of applying SDMs to species with scarce data. At the same time, it underscores the need for accurate distribution assessments, as rarity often signals a higher risk of extinction (Jeliazkov et al., 2020). While these challenges have been confirmed for insular species, this thesis improved our ability to assess and predict the distribution of range-restricted (insular) species by identifying the data requirements for more effective modelling at representative scales.

6.2.3 How does the choice of pseudo-absence data impact the accuracy of predicted island species distributions?

Most occurrence datasets, whether from islands or other environments, typically contain only presence data, requiring the creation of pseudo-absences to simulate species absences (i.e., locations where the species might have been present but were not observed; Lobo et al., 2010; Barbet-Massin et al., 2012; Morera-Pujol et al., 2020). However, there is no consensus on best practices for generating pseudo-absence data, and existing guidelines are often designed for larger study areas (Lobo et al., 2008; Barbet-Massin et al., 2012; Whitford et al., 2024), highlighting the need for strategies specific to islands (Benavides et al., 2024).

Building on previous research (Acevedo et al., 2012; Jiménez-Valverde et al., 2013; Grimmett et al., 2020; Descombes et al., 2022), the results of Chapter 3 add to the growing evidence that adjustments to pseudo-absence data—considering factors such as modelling method, spatial context, data availability, and species range—significantly affect model performance. By testing these factors using the 'model archipelago', this research offers a straightforward framework for optimising pseudo-absence data selection during model fitting. Specifically, random sampling of the study area, based on a percentage of total grids and refined according to species range sizes and data availability, led to highly accurate SDM predictions. This approach effectively addressed the problem while accommodating the needs of various modelling methods.

Despite identifying optimal pseudo-absence sampling strategies for island environments, standard metrics (e.g., AUC and the Boyce Index) still present significant limitations for evaluating model performance. In certain instances, these validation metrics did not adequately reflect the improvements achieved through optimised pseudo-absence strategies. Trade-offs exist where one might need to accept a reduction in predictive accuracy to reconcile differences between validation and performance. Moreover, as Warren et al. (2020) noted, these metrics, while effective at measuring classification accuracy (such as distinguishing between areas of high suitability for species presence from those with pseudo-absences), often fail to provide meaningful insights into habitat suitability across the landscape. This disconnect means that models may meet conventional performance standards but still struggle to predict species distributions accurately on a smaller spatial scale.

This issue persists because much of the SDM literature focuses on improving predictions through classification metrics (e.g., Guisan et al., 2007; Boria et al., 2014; Wisz et al., 2008; Kuebler et al., 2016), rather than evaluating the ability of models to capture true habitat suitability. To address this gap, species simulations were incorporated into the methodology to enhance pseudo-absence sampling in an island context, ensuring that true habitat suitability was also accurately represented (Warren et al., 2020).

The strength of the proposed guidelines relied on two key elements: (1) a thorough understanding of species distribution patterns in the model archipelago, informed by field observations, which allowed the simulation of species resembling real species distributions and ensured alignment between true model performance, validation metrics, and spatial predictions; and (2) the availability of directly collected presence-absence data, which served as a benchmark for comparing models using presence/pseudo-absence data. This approach offered valuable insights from both real and simulated species data, with the goal of improving the accuracy of presence-only models through effective pseudo-absence sampling.

Despite the advances of this thesis in improving presence-only SDMs for island contexts, the common practice of using arbitrary thresholds to differentiate between "high" and "low" performing models remains due to a lack of a solid theoretical foundation (Olden et al., 2002). Specifically, for the AUC, the most commonly used validation metric, Raes & ter Steege (2007) highlighted that for presence-only models, the maximum achievable AUC for a perfect fit never reaches 1 but is instead 1 minus half the proportion of the area occupied by the species' true distribution—a value rarely available for real species SDMs. Nonetheless, SDMs still use a baseline where an AUC of 1 indicates perfect agreement, and 0.5 suggests the model is no better than random. However, AUC primarily measures whether a model outperforms randomness rather than assessing its true predictive accuracy.

In the case study of the model archipelago, AUC values rarely exceeded 0.65 for widespread species, even when predictions closely matched actual species ranges.

This highlights three issues: (1) wide-ranging species or those with high prevalence in the study area naturally have lower validation scores, as pseudo-absence data are more likely to fall within suitable patches, making AUC prevalence-dependent (Lobo et al., 2008); (2) using AUC thresholds to classify model performance is inherently context-dependent (Lawson et al., 2014); and (3) the variation between a random AUC of 0.5 and the maximum achievable value complicates establishing a universal standard for distinguishing high- and low-performing models. This highlights the limitations of relying on arbitrary thresholds (e.g., AUC > 0.7), which often fail to capture true predictive accuracy across different species ranges.

AUC primarily focuses on classification rather than evaluating the model's ability to accurately assess habitat suitability within a given grid cell. In contrast, the Boyce Index is considered a better indicator of reliability for presence-only data (Boyce et al., 2002; Di Cola et al., 2016; Hirzel et al., 2002; Li & Guo, 2013). Thus, employing both metrics provides a more comprehensive evaluation of model performance, covering both classification accuracy and reliability. However, it is important to note that the Boyce Index is highly sensitive to small sample sizes (fewer than 30 occurrence records; Liu et al., in press), which is a common issue in island datasets and should be considered when interpreting results.

While validation metrics have limitations, they remain useful as informative guidelines. However, their methodological shortcomings should be acknowledged, and complementary approaches—such as null-model comparisons—should be integrated as standard practice within the SDM framework (Osborne et al., 2022).

6.2.4 How does the spatial precision of occurrence data affect the accuracy of island SDM predictions at high resolutions?

In island environments, where species occurrence data are often limited, using high-resolution predictors can further restrict the number of occurrences that meet specific spatial criteria. The pixel size of environmental variables should correspond to the spatial resolution of species records (Sillero & Barbosa, 2021); however, available data are often too coarse in spatial resolution (McPherson et al., 2006). While incorporating records with coarse spatial precision might seem like a practical way to increase the sample size, it risks introducing errors that could undermine the model's accuracy (Gabor et al., 2023). On the other hand, excluding such data further reduces the dataset, potentially diminishing the model’s predictive power (Moudrý et al., 2023). Addressing this trade-off required systematic testing to optimise occurrence data, ensuring that the information used enhances rather than detracts from the model's performance.

Previous evidence showed that even minor spatial inaccuracies in occurrence data can lead to significant errors in predicting species distributions in topographically heterogeneous areas (Gábor et al., 2023). Given that islands often feature highly heterogeneous environments compressed into relatively small areas, these inaccuracies are likely to be especially problematic (Barbosa et al., 2020). This issue is more frequently observed on islands of volcanic origin, which exhibit steep gradients and diverse habitats within limited space.

The model archipelago is of volcanic origin so it provided a compelling opportunity to explore the impact of low spatial accuracy records on model performance, a highly relevant issue for island SDMs that had not been systematically explored before. By comparing datasets with varying levels of spatial uncertainty (resampled to 90, 500, and 1000m resolutions), it was found that while models generally perform better with more accurate data, the effects of spatial inaccuracies can be mitigated by selecting modelling methods that are less sensitive to this uncertainty (e.g., GLM and BART). Additionally, the impact of these inaccuracies was found to vary depending on the range size of the species.

Interestingly, contrary to expectations, species in topographically diverse habitats were less affected by spatial inaccuracies, likely because their wider ranges increase the chances that imprecise occurrence points still fall within suitable areas. In contrast, species with narrow distributions, even in topographically homogeneous regions (e.g., restricted to specific elevation bands), were more vulnerable to spatial errors due to their limited habitat requirements, making spatially imprecise records more likely to fall outside suitable areas of habitat.

The findings from this and previous analyses (questions 2 & 3 above) help guide decisions on when it may be appropriate to include records with geographical uncertainty and how to mitigate any negative effects using suitable methodological approaches. For wide-ranging island species, which require fewer data points, stricter thresholds on data accuracy can be applied without significantly affecting model performance. In contrast, narrow-range island species, which need larger datasets, benefit from retaining slightly inaccurate records to increase sample size. In such cases, it is important to choose modelling methods that can tolerate inaccuracies, such as GLM and BART. These methods also perform best with moderate-sized datasets (around 20 records), making this strategy of increasing sample size particularly effective in addressing the limitations of SDMs for narrow-range island species.

Notably, this analysis used a fine resolution of 30 meters, which is rarely applied in SDMs (Benavides et al, 2024). This further supports the argument that increasing the resolution of predictors does not necessarily resolve issues caused by spatially inaccurate records (Moudrý & Šímová, 2012; Gábor et al., 2022). As shown in Chapter 4, using coarse analytical scales leads to uninformative SDMs for small island study areas (e.g., Clarion Island). In summary, selecting appropriate modelling methods and considering data needs across species ranges provides a better approach for addressing the challenge of low spatial precision in occurrence data for high-resolution island SDMs.

6.2.5 Are downscaled predictors a reliable baseline to model island plant species distributions in tropical latitudes?

Chapter 4 addressed the critical consideration of incorporating micro-environmental data into SDMs for islands, a challenge constrained by the limited availability of such predictors (Fitzpatrick & Ellison, 2018; Benavides et al., 2024). This chapter developed a methodology focused on a set of ecologically relevant variables for island SDMs, specifically tailored to scales appropriate for these environments—a consideration that has been relatively understudied (Austin & Van Niel, 2011; Syphard & Franklin, 2009; Tulloch et al., 2016).

With recent advancements in data availability and computational power, it is now possible to create detailed microclimatic descriptors by downscaling global climate data and evaluating whether these variables are ecologically significant and can accurately predict species distributions. This thesis identified a downscaling method that could be useful to describe microclimate in island environments, as this method accounts for orographic effects to derive microclimatic variables (Flint & Flint, 2012). This approach aligns well with the climatic dynamics specific to islands and is based on global bioclimatic grids, making it applicable to islands worldwide.

With recent advancements in data availability and computational power, it is now possible to create detailed microclimatic descriptors by downscaling global climate data and assess whether these variables are ecologically significant and accurately predict species distributions. This thesis identified a downscaling method that could be particularly useful for describing microclimate in island environments, as it accounts for orographic effects to derive microclimatic variables (Flint & Flint, 2012). This approach aligns well with the climatic dynamics specific to islands and is based on global bioclimatic grids (e.g., Worldclim; Fick & Hijmans, 2017; CHELSA: Karger et al., 2017), making it applicable to islands worldwide.

The most significant finding was that a null-model testing approach, introduced to evaluate the informativeness of microclimatic estimates, helped determine whether species show a statistically significant association with a diverse set of potentially informative microclimatic predictors (Deblauwe et al., 2016). This approach confirmed that downscaled microclimatic data often provided more informative spatial predictions than terrain-derived data in high-resolution island plant SDM studies (e.g., Heinänen et al., 2012; Turvey et al., 2020; Lannuzel et al., 2021; Segal et al., 2021), challenging the view that terrain-derived data are the best alternative for island studies with limited environmental data (Poteau et al., 2015). The results suggested that incorporating fine-scale microclimatic variables into SDMs is particularly important when the relationship between microclimate and topography is weaker, such as on low-elevation islands with minimal topographic variation (McInerny et al., 2011; Lembrechts et al., 2019). In such cases, relying solely on topography as a climate proxy may reduce the accuracy of model predictions.

These strategies and methodological considerations were developed to test the effectiveness, and improve the reliability, of microclimatic descriptors. While the main focus is on tropical regions, future approaches could be adapted for other island areas and tested across different taxa.

6.2.6 Does the grain size of downscaled climatic estimates impact the accuracy and transferability of island SDMs to novel environments?

While a model may show high predictive accuracy in a specific context, this does not guarantee the same level of accuracy when projected to different spatial or temporal settings (Yates et al., 2018; Rousseau & Betts, 2022). This issue arises due to factors such as local species adaptations (niche plasticity), variations in limiting factors across regions, and changes in the correlation structure of environmental variables (Dormann et al., 2012; Radomski et al., 2022). Additionally, biotic interactions add further complexity to the transferability of SDMs (Fitzpatrick et al., 2007). Despite these challenges, improving SDM transferability is crucial for two key conservation objectives in island studies: (1) assessing and mitigating ecological invasions, a major driver of species extinctions in island ecosystems, which requires spatially transferable models (Goedecke et al., 2020); and (2) improving the accuracy of predictions of species distribution responses to global environmental change, which demands temporally transferable models (Rapacciuolo et al., 2012).

Enhancing SDM transferability is a complex task requiring careful consideration of multiple factors. This thesis highlights the impact of using finer spatial scales in SDM transferability and calls for further research in this area (Petitpierre et al., 2017; Manzoor et al., 2018). One of the study's strengths is the use of independent presence-absence datasets for each island, providing a robust framework for testing transferability. Although such datasets are rare, they are vital for thoroughly assessing SDM transferability under comparable environmental conditions (Petitpierre et al., 2017). The contrasting sizes and topographical complexities of the two islands studied—one large and topographically complex, the other smaller and less topographically complex—also provided valuable insights to enhance island SDM transferability.

While the findings of this thesis align with previous research (Dormann et al., 2012; Radomski et al., 2022; Rousseau & Betts, 2022) and emphasise the critical role of scale in influencing SDM transferability, it offers two novel contributions: (1) it challenges earlier assumptions that finer analytical grains negatively impact species transferability (Manzoor et al., 2018). In fact, a high resolution of 30 metres was found to be essential for improving transferability, as the small size of one of the study islands required a much finer resolution to produce informative model projections. (2) Measuring environmental similarity between training and calibration regions is crucial for assessing uncertainty in model transferability, as changes in the correlation structure of environmental variables, caused by varying levels of topographical complexity among islands, were shown to compromise transferability (Dormann et al., 2012; Qiao et al., 2019). These two insights support the enhancement of transferability in island SDMs for future research. A key takeaway is that even at the smallest resolution tested, transferability was maintained and, in some cases, represented the only informative analytical scale (Chapter 4). This finding highlights the need for further studies using very fine resolutions, particularly for small islands. For future research on ecological invasions, the thesis highlights that poor model transferability can arise from mismatches in analytical granularity between training and transfer areas, as well as from overlooking changes in the relationships among environmental variables across these areas. This issue will be particularly important when predicting invasions from large continental areas to smaller island environments. In the context of climate change studies, the findings suggest that adjusting the scale for small islands to fine analytical resolutions can improve accuracy without compromising SDM transferability.

6.2.7 Does the choice of variable type (Bioclimatic vs. Terrain variables) impact the accuracy and transferability of island SDMs to novel environments?

In addition to the analytical scale, the choice of predictors has been identified as significantly influencing the transferability of SDMs (Datta et al., 2020; Goedecke et al., 2020). Building on existing research, this thesis examined how different climatic baselines, downscaled to fine resolutions, affect model transferability. Specifically, it compared a dataset derived from topography with downscaled versions of two widely used bioclimatic datasets: WorldClim (Fick & Hijmans, 2017) and CHELSA (Karger et al., 2017).

Testing these baselines was considered relevant due to concerns about the reliability of climatic descriptors for islands (Poteau et al., 2015). Consequently, island SDMs have often relied on topographically derived and remotely sensed environmental descriptors (e.g., Heinänen et al., 2012; Turvey et al., 2020; Lannuzel et al., 2021; Segal et al., 2021). However, as demonstrated in earlier sections of this thesis, downscaled microclimatic predictors significantly improve model accuracy, making a detailed analysis of their transferability a logical next step.

Given the methodological differences between these climatic baselines, it was hypothesised that CHELSA would outperform WorldClim. CHELSA is recognised for its relative effectiveness in areas with sparse meteorological data and high topographic complexity (Datta et al., 2020), both of which are characteristics of the study archipelago. However, despite expectations, WorldClim exhibited superior spatial transferability. This discrepancy highlighted CHELSA's sensitivity to changes in correlations between predictor variables across regions, exacerbated by the significant variation in size and topographical complexity among the islands (Mesgaran et al., 2016). For the same spatial differences, using topography-derived predictors for fitting SDMs reduced their transferability, revealing the limitations of these baselines in such contexts.

These findings suggest that the latest version of WorldClim provides a stronger foundation for transferring spatial models across regions. In contrast, CHELSA proves particularly useful when models are applied within the same region, especially in areas with limited meteorological data. The combined insights from the two transferability-related questions in this thesis offer valuable guidance for selecting environmental predictors and determining the appropriate analytical scale, whether models are applied locally or projected across regions in the context of island SDMs.

It is important to emphasise that these recommendations are guidelines for future research rather than definitive conclusions. Further investigation is needed to deepen our understanding of how island species distributions change over time and space and to refine SDM methodologies to better address the challenges of model transferability, given this aspect of the general SDM framework remains poorly understood (Sequeira, et al., 2018; Yates et al., 2018).

6.2.8 Are the patterns of distributional shifts in insular species due to climate change related to their chorotypes and/or conservation status?

Although climate change studies on plant species are a common focus in the island SDM literature (Benavides et al., 2024), they have mostly been conducted at macroclimatic scales. It is only recently that the importance of shifting to microclimatic scales has been recognised as essential for understanding island plant responses to global environmental change (Patiño et al., 2023). A series of methodological challenges, previously identified and discussed in the earlier thesis chapters, have hampered this transition. However, the methodological groundwork laid out in this thesis has made it possible to assess the potential impacts of future climate change on plant species of high ecological and conservation significance, located in subtropical and tropical oceanic archipelagos.

The findings of Chapter 5 provide new evidence of the heightened risk of losing globally significant endemic species (Fordham et al., 2010; Harter et al., 2015; Manes et al., 2021) given that endemic species generally exhibit greater vulnerability to climate change, with suitability declines exceeding 50% across most archipelagos. This aligns with trends observed in other tropical island hotspots (Pouteau & Birnbaum, 2016; Tagliari et al., 2021). An exception to this was the

Revillagigedo Archipelago, where non-endemic species suffered higher losses, with some of these species disappearing entirely.

The spatial analysis of species loss further revealed the localised loss of both endemic and non-endemic species that play important ecological roles within these archipelagos, potentially destabilising ecosystems. These findings underscore the importance of using indicator species to assess climate change impacts (e.g., Bombi et al., 2021; Dickson et al., 2019), and they highlight the urgent need for targeted conservation efforts in insular biodiversity hotspots (Ghiloufi et al., 2020).

Interestingly, the analyses showed that climate change does not consistently exacerbate existing threats to vulnerable species. A notable disconnect was observed between species classified as "Least Concern" by the IUCN and their high vulnerability to climate change. Moreover, some species identified as threatened by climate change did not face imminent habitat loss in this assessment. This suggests that current IUCN criteria may not fully capture traits that increase vulnerability to climate change, emphasising the need to revise these criteria to better reflect climate-related risks (Akçakaya et al., 2006; Keith et al., 2014).

6.3 Limitations

There are several limitations associated with the research presented in the thesis. SDMs are often criticised for their correlative nature, as they rely on statistical associations between species occurrences and environmental variables without considering causal mechanisms, or other factors that influence species distributions (Wagner et al., 2023; Tourinho & Vale, 2023). In response to this limitation, researchers are exploring methods that integrate mechanistic models (see hybrid models: Dorman et al., 2012; Tourinho & Vale, 2023), that incorporate biotic interactions (Cabral, et al., 2017; Zurell, 2017) and consider evolutionary processes (Liu et al., 2022). These approaches complement traditional SDMs but they often require more detailed and extensive data (abiotic, biotic, and dispersal, microclimate), which can make their broader application challenging (Tourinho & Vale, 2023). Given the data challenges associated with these complementary approaches, SDMs remain a valuable tool for addressing ecological and biogeographical questions.

Recognising the relevance of SDMs but also the complexity of both selecting appropriate methods among numerous options and their implementation, due to the many decisions involved in the fitting process (Leroy, 2022), this thesis aimed to provide guidelines to simplify decision-making and enhance SDM predictions, tailored specifically to island environments. Default parameters were used for the tested methods, despite the potential benefits of parameter tuning in many scenarios (Valavi et al., 2021). This choice reflects the understanding that presence-only models generally benefit less from parameter tuning (Valavi et al., 2021), particularly when dealing with small occurrence datasets—a common issue in this study and in island studies more broadly (Benavides et al., 2024). Furthermore, while parameter tuning often aims to control model complexity and prevent overfitting, this study addressed these concerns by employing the Ensemble of Small Models (ESM) approach (Lomba et al., 2010; Breiner et al., 2015). The default ESM approach not only strengthens the guidelines but also makes them relatively simple and accessible, even for researchers who are not experts in SDMs.

In a more specific context, the model archipelago offered valuable insights but differs from other, more extensively studied volcanic archipelagos, which typically feature a larger number of islands and, consequently, more sampling opportunities. While increased replication can improve the robustness of findings, it also complicates the collection of high-quality data. Moreover, the relatively smaller number of islands and their limited area restricted the range of species that could be analysed. Despite these limitations, the study prioritised data quality over quantity to address targeted scientific questions effectively. This focus on high-quality data, particularly independent presence-absence data—rarely available but crucial to the research questions explored—ensured that the findings were robust. Although the scope was narrower, this emphasis on data quality strengthened the overall conclusions.

6.4 Conservation and management implications

The primary conservation goal of this thesis was to utilise SDMs to deepen our understanding of the unique and rich plant biodiversity found on tropical islands (Cayuela et al., 2009). Despite hosting a significant portion of global biodiversity, these regions are often underrepresented in both occurrence (Benavides et al., 2024) and environmental datasets (Deblauwe et al., 2016).

Improving SDM practices within the Revillagigedo Archipelago lays a strong foundation for informed conservation decision-making, significantly enhancing biodiversity management in the region. Currently, detailed geographical knowledge of species distribution patterns remains limited (Benavides et al., 2019) and, in some instances, outdated (Brattstrom, 1990; Arnaud et al., 1993; Rodriguez-Estrella et al., 1996). This is partly due to earlier studies being conducted prior to the successful eradication of two large introduced mammals: *Sus scrofa* on Clarion Island (Wanless et al., 2009) and *Ovis aries* on Socorro Island (Aguirre-Muñoz et al., 2011). Evidence suggests significant vegetation recovery following these eradications, and it is expected that associated taxa have reorganised in response to habitat restoration (Ortiz-Alcaraz et al., 2019). This presents a compelling opportunity to apply the insights from this thesis across a broader range of taxonomic groups, thereby increasing biogeographical knowledge of the archipelago through the use of SDMs.

Additionally, two medium-sized introduced mammals, *Oryctolagus cuniculus* on Clarion Island and *Felis catus* on Socorro Island, remain present in the archipelago. SDMs have proven valuable in assessing habitat use by invasive species in island ecosystems (Recio et al., 2015). Therefore, applying SDMs can support ongoing eradication efforts and facilitate post-eradication restoration by identifying suitable areas for the reintroduction of species that have been locally extirpated (Lentini et al., 2018). Furthermore, this thesis has expanded our understanding of plant species' habitat requirements in the archipelago, enabling SDMs to be used to help mitigate the impacts of climate change through predicting shifts in species distributions (Chapter 4).

The aim of applying these insights to local conservation efforts is particularly achievable due to existing partnerships with the NGO leading invasive species eradication and the management team of the Revillagigedo Archipelago National Park. Both parties are fully informed of the thesis objectives, and the findings will be shared to aid in shaping effective conservation strategies. Additionally, there are short-term plans to further collaborate with these organisations to develop SDMs for the identified management applications.

On a global scale, this thesis also provides valuable insights into the impacts of climate change on volcanic archipelagos in the tropical Pacific and Atlantic. Large-scale assessments are essential for understanding the broader effects of climate change on these vulnerable island environments (many of which are biodiversity hotspots) and for determining whether predicted species responses follow common patterns. Both local and macroecological patterns of response to global change are crucial for guiding conservation efforts (Rodríguez et al., 2007; Kueffer & Kinney, 2017).

6.5 Future Directions

Potential future directions for research on improving SDM practices in island environments include:

- Geographical and Taxonomical Extension: While this study offers significant insights into SDM practices for island environments, future research could explore whether the data requirements and proposed methodologies for obtaining fine-grained data are consistent across diverse spatial and taxonomical contexts. A more holistic approach might involve integrating multiple interacting taxonomic levels (e.g., plants and pollinators; nesting plants and birds; invasive herbivores and the plants they consume).
- Dissemination of Microclimatic Predictors: The development of microclimatic and environmental grids in this study was highly computationally demanding. Therefore, it was not feasible within the timeframe of this thesis to cover finer scales (<500m) in the other tropical archipelagos tested or to

encompass a broader range of islands. Future research should focus on continuing the development and testing of these datasets, with the support of experts in the biodiversity of the island study areas. Making these microclimatic grids more accessible to a larger number of researchers would significantly enhance the capacity to conduct high-resolution SDM studies.

- Linking Methodological Considerations with Applied Knowledge: Although SDMs will continue to require ongoing refinement, the guidelines provided in this thesis offer a solid foundation for applying the lessons learned to address specific scientific questions. During the field collection phase of this research, additional types of data were gathered to tackle future questions, such as the synergistic effects of invasive species and climate change, improving the detection of rare and endangered island species, and enhancing strategies for invasive species management to support conservation efforts. Additionally, collaborative opportunities that arose during this PhD project with other researchers could greatly benefit from the appropriate implementation of these tools, leveraging the expertise gained to advance the field further.

6.6 Concluding Remark

In summary, while this study has made significant advances in tailoring and enhancing SDM practices for island environments, the outlined future directions underscore the need for continued research and collaboration. Expanding the geographical and taxonomic scope, improving the accessibility of microclimatic predictors, and linking methodological advancements with practical applications will not only refine SDM methodologies but also strengthen their role in conservation efforts globally. By addressing these areas, future work can build on the foundation laid by this research, ultimately contributing to the development and implementation of more effective and informed environmental management strategies in island environments.

6.7 References

Acevedo, P., Jiménez-Valverde, A., Lobo, J. M., & Real, R. (2012). Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, *39*(8), 1383-1390.

Aguirre-Muñoz, A., Samaniego-Herrera, A., Luna-Mendoza, L., Ortiz-Alcaraz, A., Rodríguez-Malagón, M., Méndez-Sánchez, F., ... & Latofski-Robles, M. (2011). Island restoration in Mexico: ecological outcomes after systematic eradications of invasive mammals. *Island invasives: eradication and management*, 250-258.

Akçakaya, H. R., Butchart, S. H., Mace, G. M., Stuart, S. N., & Hilton-Taylor, C. R. A. I. G. (2006). Use and misuse of the IUCN Red List Criteria in projecting climate change impacts on biodiversity. *Global Change Biology*, *12*(11), 2037-2043.

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

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, *33*(10), 1677-1688.

Arnaud, G., Rodríguez, A., Ortega-Rubio, A., & Álvarez-Cárdenas, S. (1993). Predation by cats on the unique endemic lizard of Socorro Island (*Urosaurus auriculatus*), Revillagigedo, Mexico. *The Ohio Journal of Science, 93*(4), 101-104.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*(2), 327-338.

Barajas-Barbosa, M.P., Weigelt, P., Borregaard, M.K., Keppel, G. & Kreft, H. (2020). Environmental heterogeneity dynamics drive plant diversity on oceanic islands. *Journal of Biogeography*, *47*, 2248-2260.

Benavides, E., Kuethe, J. R., Ortiz-Alcaraz, A., & DE LA LUZ, J. L. (2019). *Oenothera resicum* (Onagraceae), a new species and the first record of the family from the Revillagigedo Archipelago, Mexico. *Phytotaxa*, *416*(1), 59-66.

Benavides, E., Sadler, J., Graham, L., & Matthews, T. J. (2024). Species distribution models and island biogeography: Challenges and prospects. *Global Ecology and Conservation*, e02943.

Booth, T. H. (2022). Checking bioclimatic variables that combine temperature and precipitation data before their use in species distribution models. *Austral Ecology*, *47*(7), 1506-1514.

Bombi, P., Salvi, D., Shuuya, T., Vignoli, L., & Wassenaar, T. (2021). Climate change effects on desert ecosystems: A case study on the keystone species of the Namib Desert *Welwitschia mirabilis*. *PloS one*, *16*(11), e0259767.

Borges, P. A., Cardoso, P., Kreft, H., Whittaker, R. J., Fattorini, S., Emerson, B. C., ... & Gabriel, R. (2018). Global Island Monitoring Scheme (GIMS): a proposal for the long-term coordinated survey and monitoring of native island forest biota. *Biodiversity and Conservation*, *27*, 2567-2586.

Boria, R. A., Olson, L. E., Goodman, S. M., & Anderson, R. P. (2014). Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, *275*, 73-77.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. (2002). Evaluating resource selection functions. *Ecological Modelling*, *157*(2-3), 281-300.

Brattstrom, B. H. (1990). Biogeography of the Islas Revillagigedo, Mexico. *Journal of Biogeography, 17*(2), 177-183.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, *6*(10), 1210-1218.

Cabral, J. S., Valente, L., & Hartig, F. (2017). Mechanistic simulation models in macroecology and biogeography: state-of-art and prospects. *Ecography*, *40*(2), 267-280.

Cayuela, L., Golicher, D. J., Newton, A. C., Kolb, M., De Alburquerque, F. S., Arets, E. J. M. M., ... & Pérez, A. M. (2009). Species distribution modeling in the tropics: problems, potentialities, and the role of biological data for effective species conservation. *Tropical Conservation Science*, *2*(3), 319-352.

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

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

Datta, A., Schweiger, O. and Kühn, I. (2020). Origin of climatic data can determine the transferability of species distribution models. *NeoBiota 59*, 61-76.

Deblauwe, V., Droissart, V., Bose, R., Sonké, B., Blach-Overgaard, A., Svenning, J. C., ... & Couvreur, T. L. P. (2016). Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. *Global Ecology and Biogeography*, *25*(4), 443-454.

Descombes, P., Chauvier, Y., Brun, P., Righetti, D., Wüest, R. O., Karger, D. N., ... & Zimmermann, N. E. (2022). Strategies for sampling pseudo-absences for species distribution models in complex mountainous terrain. *BioRxiv*, 2022-03.

Dickson, C. R., Baker, D. J., Bergstrom, D. M., Bricher, P. K., Brookes, R. H., Raymond, B., ... & McGeoch, M. A. (2019). Spatial variation in the ongoing and widespread decline of a keystone plant species. *Austral Ecology*, *44*(5), 891-905.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F. T., d'Amen, M., Randin, C., ... & Guisan, A. (2017). ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography*, *40*(6), 774-787.

Dormann, C. F., Schymanski, S. J., Cabral, J., Chuine, I., Graham, C., Hartig, F., ... & Singer, A. (2012). Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, *39*(12), 2119-2131.

Fernández-Palacios, J. M., Kreft, H., Irl, S. D., Norder, S., Ah-Peng, C., Borges, P. A., ... & Drake, D. R. (2021). Scientists' warning–The outstanding biodiversity of islands is in peril. *Global Ecology and Conservation*, *31*, e01847.

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

Fitzpatrick, M. C., Ellison, A. M., & Adamec, L. (2018). Estimating the exposure of carnivorous plants to rapid climatic change. *Carnivorous Plants: Physiology, Ecology and Evolution.* Oxford University Press, London.

Fitzpatrick, M. C., Weltzin, J. F., Sanders, N. J., & Dunn, R. R. (2007). The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, *16*(1), 24-33.

Flint, L. E., & Flint, A. L. (2012). Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, *1*, 1-15.

Fois, M., Bacchetta, G., Cuena-Lombraña, A., Cogoni, D., Pinna, M. S., Sulis, E., & Fenu, G. (2018). Using extinctions in species distribution models to evaluate and predict threats: a contribution to plant conservation planning on the island of Sardinia. *Environmental Conservation*, *45*(1), 11-19.

Fordham, D. A., & Brook, B. W. (2010). Why tropical island endemics are acutely susceptible to global change. *Biodiversity and Conservation*, *19*, 329-342.

Franklin, J., Davis, F. W., Ikegami, M., Syphard, A. D., Flint, L. E., Flint, A. L., & Hannah, L. (2013). Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Global Change Biology*, *19*(2), 473-483.

Gábor, L., Jetz, W., Lu, M., Rocchini, D., Cord, A., Malavasi, M., ... & Moudrý, V. (2022). Positional errors in species distribution modelling are not overcome by the coarser grains of analysis. *Methods in Ecology and Evolution, 13*(10), 2289-2302.

Gábor, L., Jetz, W., Zarzo-Arias, A., Winner, K., Yanco, S., Pinkert, S., ... & Moudrý, V. (2023). Species distribution models affected by positional uncertainty in species occurrences can still be ecologically interpretable. *Ecography*, e06358.

GBIF.org (2024), *GBIF Home Page*. Available from: https://www.gbif.org [August 2020].

Gillespie, R. G., Claridge, E. M., & Roderick, G. K. (2008). Biodiversity dynamics in isolated island communities: interaction between natural and human-mediated processes. *Molecular Ecology*, *17*(1), 45-57.

Goedecke, F., Marcenò, C., Guarino, R., Jahn, R., & Bergmeier, E. (2020). Reciprocal extrapolation of species distribution models between two islands–specialists perform better than generalists and geological data reduces prediction accuracy. *Ecological Indicators*, *108*, 105652.

Grimmett, L., Whitsed, R., & Horta, A. (2020). Presence-only species distribution models are sensitive to sample prevalence: Evaluating models using spatial prediction stability and accuracy metrics. *Ecological Modelling*, *431*, 109194.

Ghiloufi, W., & Chaieb, M. (2020). Predicted effects of climate change on a Mediterranean keystone plant species. *Folia Geobotanica*, *55*(3), 241-255.

Guisan, A., Graham, C. H., Elith, J., Huettmann, F., & Distri, N. S. (2007). Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions, 13*(3), 332–340.

Harter, D. E., Irl, S. D., Seo, B., Steinbauer, M. J., Gillespie, R., Triantis, K. A., ... & Beierkuhnlein, C. (2015). Impacts of global climate change on the floras of oceanic islands–Projections, implications and current knowledge. *Perspectives in Plant Ecology, Evolution and Systematics*, *17*(2), 160-183.

Heinänen, S., Erola, J., & von Numers, M. (2012). High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. *Landscape Ecology*, *27*(4), 545-555.

Hirzel, A. H., Hausser, J., Chessel, D., & Perrin, N. (2002). Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, *83*(7), 2027-2036.

Jeliazkov, A., Gavish, Y., Marsh, C. J., Geschke, J., Brummitt, N., Rocchini, D., ... & Henle, K. (2022). Sampling and modelling rare species: Conceptual guidelines for the neglected majority. *Global Change Biology*, *28*(12), 3754-3777.

Jiménez-Valverde, A., Acevedo, P., Barbosa, A. M., Lobo, J. M., & Real, R. (2013). Discrimination capacity in species distribution models depends on the representativeness of the environmental domain. *Global Ecology and Biogeography, 22*(4), 508-516.

Keith, D. A., Mahony, M., Hines, H., Elith, J., Regan, T. J., Baumgartner, J. B., ... & Akçakaya, H. R. (2014). Detecting extinction risk from climate change by IUCN Red List criteria. *Conservation Biology*, *28*(3), 810-819.

Kuebler, D., Hildebrandt, P., Guenter, S., Stimm, B., Weber, M., Mosandl, R., ... & Silva, B. (2016). Assessing the importance of topographic variables for the spatial distribution of tree species in a tropical mountain forest. *Erdkunde*, *70*(1), 19-47.

Kueffer, C., & Kinney, K. (2017). What is the importance of islands to environmental conservation? *Environmental Conservation*, *44*(4), 311-322.

Lannuzel, G., Balmot, J., Dubos, N., Thibault, M., & Fogliani, B. (2021). High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia. *Biodiversity and Conservation*, *30*(4), 963-990.

Lawson, C. R., Hodgson, J. A., Wilson, R. J., & Richards, S. A. (2014). Prevalence, thresholds and the performance of presence–absence models. *Methods in Ecology and Evolution*, *5*(1), 54-64.

Lentini, P. E., Stirnemann, I. A., Stojanovic, D., Worthy, T. H., & Stein, J. A. (2018). Using fossil records to inform reintroduction of the kakapo as a refugee species. *Biological Conservation*, *217*, 157-165.

Leroy, B. (2022). Choosing presence-only species distribution models. *Journal of Biogeography*, *50*(1),247-250.

Leroy, B., Delsol, R., Hugueny, B., Meynard, C. N., Barhoumi, C., Barbet-Massin, M., & Bellard, C. (2018). Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. *Journal of Biogeography*, *45*(9), 1994-2002.

Liu, C., Wolter, C., Courchamp, F., Roura-Pascual, N., & Jeschke, J. M. (2022). Biological invasions reveal how niche change affects the transferability of species distribution models. *Ecology*, *103*(8), e3719.

Liu, C., Newell, G., White, M., & Machunter, J. (in press). Improving the estimation of the Boyce index using statistical smoothing methods for evaluating species distribution models with presence-only data. DOI:10.22541/au.169941725.58752741/v1

Li, W., & Guo, Q. (2013). How to assess the prediction accuracy of species presence–absence models without absence data? *Ecography*, *36*(7), 788-799.

Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography, 33*(1), 103-114.

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.

Lomba, A., Pellissier, L., Randin, C., Vicente, J., Moreira, F., Honrado, J., & Guisan, A. (2010). Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. *Biological Conservation*, *143*(11), 2647-2657.

Manes, S., Costello, M. J., Beckett, H., Debnath, A., Devenish-Nelson, E., Grey, K. A., ... & Vale, M. M. (2021). Endemism increases species' climate change risk in areas of global biodiversity importance. *Biological Conservation*, *257*, 109070.

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

Matthews, T. J., Wayman, J. P., Cardoso, P., Sayol, F., Hume, J. P., Ulrich, W., ... & Triantis, K. A. (2022). Threatened and extinct island endemic birds of the world: Distribution, threats and functional diversity. *Journal of Biogeography*, *49*(11), 1920-1940.

McPherson, J. M., Jetz, W., & Rogers, D. J. (2006). Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions—possibilities and limitations. *Ecological Modelling*, *192*(3-4), 499-522.

McInerny, G. J., & Purves, D. W. (2011). Fine-scale environmental variation in species distribution modelling: regression dilution, latent variables and neighbourly advice. *Methods in Ecology and Evolution*, *2*(3), 248-257.

Mendes, P., Velazco, S. J. E., de Andrade, A. F. A., & Júnior, P. D. M. (2020). Dealing with overprediction in species distribution models: How adding distance constraints can improve model accuracy. *Ecological Modelling*, *431*, 109180.

Mesgaran, M. B., Lewis, M. A., Ades, P. K., Donohue, K., Ohadi, S., Li, C., & Cousens, R. D. (2016). Hybridization can facilitate species invasions, even without enhancing local adaptation. *Proceedings of the National Academy of Sciences*, *113*(36), 10210-10214.

Morera-Pujol, V., Mostert, P. S., Murphy, K. J., Burkitt, T., Coad, B., McMahon, B. J., ... & Ciuti, S. (2023). Bayesian species distribution models integrate presence-only and presence–absence data to predict deer distribution and relative abundance. *Ecography*, *2023*(2), e06451.

Moudrý, V., Keil, P., Gábor, L., Lecours, V., Zarzo-Arias, A., Barták, V., ... & Šímová, P. (2023). Scale mismatches between predictor and response variables in species distribution modelling: A review of practices for appropriate grain selection. *Progress in Physical Geography: Earth and Environment*, *47*(3), 467-482.

Moudrý, V., & Šímová, P. (2012). Influence of positional accuracy, sample size and scale on modelling species distributions: A review. *International Journal of Geographical Information Science*, *26*(11), 2083–2095.

Mousikos, A., Manolaki, P., Knez, N., & Vogiatzakis, I. N. (2021). Can distribution modeling inform rare and endangered species monitoring in Mediterranean islands? *Ecological Informatics*, *66*, 101434.

Olden, J. D., Jackson, D. A., & Peres-Neto, P. R. (2002). Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society, 131*(2), 329-336.

Ortiz-Alcaraz, A., Aguirre-Muñoz, A., Méndez-Sánchez, F., Rojas-Mayoral, E., Solís-Carlos, F., Rojas-Mayoral, B., ... & Ortega-Rubio, A. (2019). Ecological restoration of Socorro Island, Revillagigedo Archipelago, Mexico: the eradication of feral sheep and cats. *Island invasives: scaling up to meet the challenge. Proceedings of the international conference on island invasives, 62,* 267-273.

Osborne, O. G., Fell, H. G., Atkins, H., van Tol, J., Phillips, D., Herrera-Alsina, L., ... & Algar, A. C. (2022). Fauxcurrence: simulating multi-species occurrences for null models in species distribution modelling and biogeography. *Ecography*, *2022*(7), e05880.

Patiño, J., Whittaker, R. J., Borges, P. A., Fernández-Palacios, J. M., Ah-Peng, C., Araújo, M. B., ... & Emerson, B. C. (2017). A roadmap for island biology: 50 fundamental questions after 50 years of The Theory of Island Biogeography. *Journal of Biogeography*, *44*(5), 963-983.

Patiño, J., Collart, F., Vanderpoorten, A., Martin-Esquivel, J. L., Naranjo-Cigala, A., Mirolo, S., & Karger, D. N. (2023). Spatial resolution impacts projected plant responses to climate change on topographically complex islands. *Diversity and Distributions*, *29*(10), 1245-1262.

Petitpierre, B., Broennimann, O., Kueffer, C., Daehler, C., & Guisan, A. (2017). Selecting predictors to maximize the transferability of species distribution models: Lessons from cross-continental plant invasions*. Global Ecology and Biogeography,* *26*(3), 275-287.

Pouteau, R., Bayle, É., Blanchard, É., Birnbaum, P., Cassan, J. J., Hequet, V., ... & Vandrot, H. (2015). Accounting for the indirect area effect in stacked species distribution models to map species richness in a montane biodiversity hotspot. *Diversity and Distributions*, *21*(11), 1329-1338.

Qiao, H., Feng, X., Escobar, L. E., Peterson, A. T., Soberón, J., Zhu, G., & Papeş, M. (2019). An evaluation of transferability of ecological niche models. *Ecography*, *42*(3), 521-534.

Radomski, T., Beamer, D., Babineau, A., Wilson, C., Pechmann, J., & Kozak, K. H. (2022). Finding what you don't know: testing SDM methods for poorly known species. *Diversity and Distributions*, *28*(9), 1769-1780.

Raes, N., & ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, *30*(5), 727-736.

Rapacciuolo, G., Roy, D. B., Gillings, S., Fox, R., Walker, K., & Purvis, A. (2012). Climatic associations of British species distributions show good transferability in time but low predictive accuracy for range change. *PLoS One*, *7*(7), e40212.

Rodríguez, J. P., Taber, A. B., Daszak, P., Sukumar, R., Valladares-Padua, C., Padua, S., ... & Pearl, M. (2007). Globalization of conservation: a view from the south. *Science*, *317*(5839), 755-756.

Rodriguez-Estrella, R., de la Luz, J. L. L., Breceda, A., Castellanos, A., Cancino, J., & Llinas, J. (1996). Status, density and habitat relationships of the endemic terrestrial birds of Socorro Island, Revillagigedo Islands, Mexico. *Biological Conservation*, *76*(2), 195-202.

Rousseau, J. S., & Betts, M. G. (2022). Factors influencing transferability in species distribution models. *Ecography*, *2022*(7), e06060.

Segal, R. D., Massaro, M., Carlile, N., & Whitsed, R. (2021). Small-scale species distribution model identifies restricted breeding habitat for an endemic island bird. *Animal Conservation*, *24*(6), 959-969.

Sequeira, A. M., Bouchet, P. J., Yates, K. L., Mengersen, K., & Caley, M. J. (2018). Transferring biodiversity models for conservation: Opportunities and challenges. *Methods in Ecology and Evolution*, *9*(5), 1250-1264.

Sillero, N., & Barbosa, A. M. (2021). Common mistakes in ecological niche models. *International Journal of Geographical Information Science*, *35*(2), 213-226.

Syphard, A. D., & Franklin, J. (2009). Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. *Ecography*, *32*(6), 907-918.

Taylor, S., & Kumar, L. (2016). Global climate change impacts on pacific islands terrestrial biodiversity: a review. *Tropical Conservation Science*, *9*(1), 203-223.

Tourinho, L., & Vale, M. M. (2023). Choosing among correlative, mechanistic, and hybrid models of species' niche and distribution. *Integrative Zoology*, *18*(1), 93-109.

Tulloch, A. I., Sutcliffe, P., Naujokaitis-Lewis, I., Tingley, R., Brotons, L., Ferraz, K. M. P., ... & Rhodes, J. R. (2016). Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. *Biological Conservation*, *199*, 157-171.

Turvey, S. T., Kennerley, R. J., Hudson, M. A., Nuñez-Miño, J. M., & Young, R. P. (2020). Assessing congruence of opportunistic records and systematic surveys for predicting Hispaniolan mammal species distributions. *Ecology and Evolution*, *10*(11), 5056-5068.

van Proosdij, A. S. J., Sosef, M. S. M., Wieringa, J. J., & Raes, N. (2016). Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, *39*(6), 542-552.

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2021). Modelling species presence-only data with random forests. *Ecography*, *44*(12), 1731-1742.

Veron, S., Mouchet, M., Govaerts, R., Haevermans, T., & Pellens, R. (2019). Vulnerability to climate change of islands worldwide and its impact on the tree of life. *Scientific Reports*, *9*(1), 14471.

Wagner, T., Schliep, E. M., North, J. S., Kundel, H., Custer, C. A., Ruzich, J. K., & Hansen, G. J. (2023). Predicting climate change impacts on poikilotherms using physiologically guided species abundance models. *Proceedings of the National Academy of Sciences*, *120*(15), e2214199120.

Wanless, R. M., Aguirre-Muñoz, A., Angel, A., Jacobsen, J. K., Keitt, B. S., & McCann, J. (2009). Birds of Clarion Island, Revillagigedo Archipelago, Mexico. *The Wilson Journal of Ornithology*, *121*(4), 745-751.

Warren, D. L., Matzke, N. J., & Iglesias, T. L. (2020). Evaluating presence-only species distribution models with discrimination accuracy is uninformative for many applications. *Journal of Biogeography*, *47*(1), 167-180.

Whittaker, R. J., Fernández-Palacios, J. M., & Matthews, T. J. (2023). *Island biogeography: geo-environmental dynamics, ecology, evolution, human impact, and conservation*. Oxford University Press, Oxford.

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

Whitford, A. M., Shipley, B. R., & McGuire, J. L. (2024). The influence of the number and distribution of background points in presence-background species distribution models. *Ecological Modelling*, *488*, 110604.

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., ... & Sequeira, A. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology & Evolution*, *33*(10), 790-802.

Zurell, D. (2017). Integrating demography, dispersal and interspecific interactions into bird distribution models. *Journal of Avian Biology*, *48*(12), 1505-1516.

Appendices

Chapter 2

Appendix S2.1 Reviewed Literature

Aguilar, G. D., Blanchon, D. J., Foote, H., Pollonais, C. W., & Mosee, A. N. (2017). A performance based consensus approach for predicting spatial extent of the Chinese windmill palm (Trachycarpus fortunei) in New Zealand under climate change. *Ecological informatics, 39*, 130-139.

Aguilar, G., Blanchon, D., Foote, H., Pollonais, C., & Mosee, A. (2015). Queensland fruit fly invasion of New Zealand: predicting area suitability under future climate change scenarios. *Unitec e-Press Perspectives in Biosecurity Research Series*, 2: Auckland, New Zealand, 2015

Aguilar, G. D., Farnworth, M. J., & Winder, L. (2015). Mapping the stray domestic cat (Felis catus) population in New Zealand: Species distribution modelling with a climate change scenario and implications for protected areas. *Applied Geography, 63*, 146-154.

Alaniz, A. J., Carvajal, M. A., Smith-Ramírez, C., Barahona-Segovia, R. M., & Vieli, L. (2018). Habitat loss of a rainforest specialist pollinator fly as an indicator of conservation status of the South American Temperate Rainforests. *Journal of Insect Conservation, 22*, 745-755.

Alaniz, A. J., Soares, A. O., Vergara, P. M., de Azevedo, E. B., & Grez, A. A. (2021). The failed invasion of Harmonia axyridis in the Azores, Portugal: Climatic restriction or wrong population origin?. *Insect science, 28*(1), 238-250.

Algar, A. C., Mahler, D. L., Glor, R. E., & Losos, J. B. (2013). Niche incumbency, dispersal limitation and climate shape geographical distributions in a species-rich island adaptive radiation. *Global Ecology and Biogeography, 22*(4), 391-402.

Almadrones-Reyes, K. J., & Dagamac, N. H. A. (2018). Predicting local habitat suitability in changing climate scenarios: Applying species distribution modelling for Diderma hemisphaericum. *Current Research in Environmental & Applied Mycology, 8*(5), 492-500

Almarinez, B. J. M., Fadri, M. J. A., Lasina, R., Tavera, M. A. A., Carvajal, T. M., Watanabe, K., ... & Amalin, D. M. (2021). A bioclimate-based maximum entropy model for Comperiella calauanica Barrion, Almarinez and Amalin (Hymenoptera: Encyrtidae) in the Philippines. *Insects, 12*(1), 26.

Amin, R. J., Buettel, J. C., Fielding, M. W., Vaughan, P. M., & Brook, B. W. (2021). Hot, unpredictable weather interacts with land use to restrict the distribution of the Yellow-tailed Black-Cockatoo. *Emu-Austral Ornithology*, *121*(4), 323-332.

Ancillotto, L., Bosso, L., Smeraldo, S., Mori, E., Mazza, G., Herkt, M., ... & Russo, D. (2020). An African bat in Europe, Plecotus gaisleri: Biogeographic and ecological insights from molecular taxonomy and Species Distribution Models. *Ecology and Evolution, 10*(12), 5785-5800.

Angetter, L. S., Lötters, S., & Roedder, D. (2011). Climate niche shift in invasive species: the case of the brown anole*. Biological Journal of the Linnean Society, 104*(4), 943-954

Aoki, D., Sakamoto, H., Kitazawa, M., Kryukov, A. P., & Takagi, M. (2021). Migration-tracking integrated phylogeography supports long-distance dispersal-driven divergence for a migratory bird species in the Japanese archipelago. *Ecology and evolution, 11*(11), 6066-6079.

Aoki, K., Tamaki, I., Nakao, K., Ueno, S., Kamijo, T., Setoguchi, H., ... & Tsumura, Y. (2019). Approximate Bayesian computation analysis of EST-associated microsatellites indicates that the broadleaved evergreen tree Castanopsis sieboldii survived the Last Glacial Maximum in multiple refugia in Japan. *Heredity, 122*(3), 326-340.

Aranda, S. C., & Lobo, J. M. (2011). How well does presence-only-based species distribution modelling predict assemblage diversity? A case study of the Tenerife flora. Ecography, 34(1), 31-38.

Ariefiandy, A., Purwandana, D., Azmi, M., Nasu, S. A., Mardani, J., Ciofi, C., & Jessop, T. S. (2021). Human activities associated with reduced Komodo dragon habitat use and range loss on Flores. *Biodiversity and Conservation, 30*(2), 461-479.

Astuti, I. P., & Cropper, W. P. (2019). Comparing six different species distribution models with several subsets of environmental variables: Predicting the potential current distribution of Guettarda speciosa in Indonesia. *Biodiversitas Journal of Biological Diversity, 20*(8).

Bariotakis, M., Koutroumpa, K., Karousou, R., & Pirintsos, S. A. (2016). Environmental (in) dependence of a hybrid zone: Insights from molecular markers and ecological niche modeling in a hybrid zone of Origanum (Lamiaceae) on the island of Crete. *Ecology and Evolution, 6*(24), 8727-8739.

Barlow, M. M., Johnson, C. N., McDowell, M. C., Fielding, M. W., Amin, R. J., & Brewster, R. (2021). Species distribution models for conservation: identifying translocation sites for eastern quolls under climate change. *Global Ecology and Conservation, 29*, e01735.

Bazzichetto, M., Massol, F., Carboni, M., Lenoir, J., Lembrechts, J. J., Joly, R., & Renault, D. (2021). Once upon a time in the far south: Influence of local drivers and functional traits on plant invasion in the harsh sub-Antarctic islands. *Journal of Vegetation Science, 32*(4), e13057.

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

Benning, T. L., LaPointe, D., Atkinson, C. T., & Vitousek, P. M. (2002). Interactions of climate change with biological invasions and land use in the Hawaiian Islands: modeling the fate of endemic birds using a geographic information system. *Proceedings of the National Academy of Sciences, 99*(22), 14246-14249.

Bisconti, R., Canestrelli, D., Colangelo, P., & Nascetti, G. (2011). Multiple lines of evidence for demographic and range expansion of a temperate species (Hyla sarda) during the last glaciation. *Molecular Ecology, 20*(24), 5313-5327.

Bisconti, R., Canestrelli, D., Salvi, D., & Nascetti, G. (2013). A geographic mosaic of evolutionary lineages within the insular endemic newt Euproctus montanus. *Molecular Ecology, 22*(1), 143-156.

Bisi, F., Gagliardi, A., Cremonesi, G., Colombo, R., Mazzamuto, M. V., Wauters, L. A., ... & Martinoli, A. (2019). Distribution of wildlife and illegal human activities in the Lampi marine National Park (Myanmar). *Environmental Conservation, 46*(2), 163-170

Booth, T. H. (2016). Estimating potential range and hence climatic adaptability in selected tree species. *Forest Ecology and Management, 366*, 175-183.

Brown, J. L., Sillero, N., Glaw, F., Bora, P., Vieites, D. R., & Vences, M. (2016). Spatial biodiversity patterns of Madagascar's amphibians and reptiles. *PLoS One, 11*(1), e0144076.

Brown, J. L., & Yoder, A. D. (2015). Shifting ranges and conservation challenges for lemurs in the face of climate change. *Ecology and Evolution, 5*(6), 1131-1142.

Buckland, S., Cole, N. C., Aguirre-Gutierrez, J., Gallagher, L. E., Henshaw, S. M., Besnard, A., ... & Harris, S. (2014). Ecological effects of the invasive giant Madagascar day gecko on endemic Mauritian geckos: applications of binomial-mixture and species distribution models. *PLoS One, 9*(4), e88798.

Buerki, S., Callmander, M. W., Bachman, S., Moat, J., Labat, J. N., & Forest, F. (2015). Incorporating evolutionary history into conservation planning in biodiversity hotspots. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1662), 20140014.

Bulgarella, M., Trewick, S. A., Minards, N. A., Jacobson, M. J., & Morgan-Richards, M. (2014). Shifting ranges of two tree weta species (H emideina spp.): competitive exclusion and changing climate. *Journal of Biogeography, 41*(3), 524-535.

Cahyaningsih, R., Phillips, J., Brehm, J. M., Gaisberger, H., & Maxted, N. (2021). Climate change impact on medicinal plants in Indonesia. *Global Ecology and Conservation, 30*, e01752.

Casazza, G., Giordani, P., Benesperi, R., Foggi, B., Viciani, D., Filigheddu, R., ... & Mariotti, M. G. (2014). Climate change hastens the urgency of conservation for range-restricted plant species in the central-northern Mediterranean region. *Biological Conservation, 179*, 129-138.

Chiatante, G. (2022). Spatial distribution of an assemblage of an endemic genus of birds: an example from Madagascar. African Journal of Ecology, 60(1), 13-26.

Coello, A. J., Fernández-Mazuecos, M., García-Verdugo, C., & Vargas, P. (2021). Phylogeographic sampling guided by species distribution modeling reveals the Quaternary history of the Mediterranean–Canarian *Cistus monspeliensis* (Cistaceae). *Journal of Systematics and Evolution, 59*(2), 262-277.

Condro, A. A., Prasetyo, L. B., Rushayati, S. B., Santikayasa, I. P., & Iskandar, E. (2021). Predicting hotspots and prioritizing protected areas for endangered primate species in Indonesia under changing climate. *Biology, 10*(2), 154

Constandinou, S., Nikoloudakis, N., Kyratzis, A. C., & Katsiotis, A. (2018). Genetic diversity of Avena ventricosa populations along an ecogeographical transect in Cyprus is correlated to environmental variables. *PLoS One, 13*(3), e0193885.

Costa, H., Ponte, N. B., Azevedo, E. B., & Gil, A. (2015). Fuzzy set theory for predicting the potential distribution and cost-effective monitoring of invasive species. *Ecological Modelling, 316*, 122-132.

Cottee-Jones, H. E. W., Mittermeier, J. C., & Redding, D. W. (2013). The Moluccan Woodcock Scolopax rochussenii on Obi Island, North Moluccas, Indonesia: a 'lost'species is less endangered than expected. *Journal of Asian Ornithology,* (29), 88-93.

Croft, S., Ward, A. I., Aegerter, J. N., & Smith, G. C. (2019). Modeling current and potential distributions of mammal species using presence-only data: A case study on British deer. *Ecology and Evolution, 9*(15), 8724-8735.

Crystal-Ornelas, R., Lockwood, J. L., Cassey, P., & Hauber, M. E. (2017). The establishment threat of the obligate brood-parasitic Pin-tailed Whydah (*Vidua macroura*) in North America and the Antilles. *The Condor: Ornithological Applications, 119*(3), 449-458.

Cunningham, C. X., Perry, G. L., Bowman, D. M., Forsyth, D. M., Driessen, M. M., Appleby, M., ... & Johnson, C. N. (2022). Dynamics and predicted distribution of an irrupting 'sleeper'population: fallow deer in Tasmania. *Biological Invasions, 24*(4), 1131-1147.

Cursach, J., Far, A. J., & Ruiz, M. (2020). Geospatial analysis to assess distribution patterns and predictive models for endangered plant species to support management decisions: a case study in the Balearic Islands. Biodiversity and Conservation, 29, 3393-3410.

Dalgarno, S., Mersey, J., Gedalof, Z. E., & Lemon, M. (2017). Species-environment associations and predicted distribution of Black Oystercatcher breeding pairs in Haida Gwaii, British Columbia, Canada. *Avian Conservation and Ecology, 12*(2).

Da Re, D., Tordoni, E., De Pascalis, F., Negrín-Pérez, Z., Fernández-Palacios, J. M., Arévalo, J. R., ... & Bacaro, G. (2020). Invasive fountain grass (Pennisetum setaceum (Forssk.) Chiov.) increases its potential area of distribution in Tenerife island under future climatic scenarios. *Plant Ecology, 221*, 867-882.

Davies, K. F., & Melbourne, B. A. (1999). Statistical models of invertebrate distribution on Macquarie Island: a tool to assess climate change and local human impacts. *Polar Biology, 21*, 240-250.

De Lima, R. F., Sampaio, H., Dunn, J. C., Cabinda, G., Fonseca, R., Oquiongo, G., ... & Buchanan, G. M. (2017). Distribution and habitat associations of the critically endangered bird species of São Tomé Island (Gulf of Guinea). *Bird Conservation International, 27*(4), 455-469.

Dickson, C. R., Baker, D. J., Bergstrom, D. M., Bricher, P. K., Brookes, R. H., Raymond, B., ... & McGeoch, M. A. (2019). Spatial variation in the ongoing and widespread decline of a keystone plant species. *Austral Ecology, 44*(5), 891-905.

Dorrington, A., Joseph, L., Hallgren, W., Mason, I., Drew, A., Hughes, J. M., & Schmidt, D. J. (2020). Phylogeography of the blue-winged kookaburra Dacelo leachii across tropical northern Australia and New Guinea. *Emu-Austral Ornithology, 120*(1), 33-45.

Dubos, N., Montfort, F., Grinand, C., Nourtier, M., Deso, G., Probst, J. M., ... & Crottini, A. (2022). Are narrow-ranging species doomed to extinction? Projected dramatic decline in future climate suitability of two highly threatened species. *Perspectives in Ecology and Conservation, 20*(1), 18-28.

Duffy, G. A., Coetzee, B. W., Latombe, G., Akerman, A. H., McGeoch, M. A., & Chown, S. L. (2017). Barriers to globally invasive species are weakening across the Antarctic. *Diversity and Distributions, 23*(9), 982-996.

Edwards, W. M., Bungard, M. J., Rakotondrasoa, E. F., Razafindraibe, P., Andriantsimanarilafy, R. R., Razafimanahaka, J. H., & Griffiths, R. A. (2022). Predicted impact of climate change on the distribution of the Critically Endangered golden mantella (*Mantella aurantiaca*) in Madagascar. *The Herpetological Journal, 32*(1), 5-13.

Elias, R. B., Gil, A., Silva, L., Fernández-Palacios, J. M., Azevedo, E. B., & Reis, F. (2016). Natural zonal vegetation of the Azores Islands: characterization and potential distribution. *Phytocoenologia, 46*(2), 107-123.

Ellis-Soto, D., Blake, S., Soultan, A., Guézou, A., Cabrera, F., & Lötters, S. (2017). Plant species dispersed by Galapagos tortoises surf the wave of habitat suitability under anthropogenic climate change. *PLoS One, 12*(7), e0181333.

Enoki, T. (2003). Microtopography and distribution of canopy trees in a subtropical evergreen broad-leaved forest in the northern part of Okinawa Island, Japan. *Ecological Research, 18*, 103-113.

Fancourt, B. A., Bateman, B. L., VanDerWal, J., Nicol, S. C., Hawkins, C. E., Jones, M. E., & Johnson, C. N. (2015). Testing the role of climate change in species decline: is the eastern quoll a victim of a change in the weather?. *PLoS One, 10*(6), e0129420.

Fernández-Mazuecos, M., Jiménez-Mejías, P., Rotllan-Puig, X., & Vargas, P. (2014). Narrow endemics to Mediterranean islands: moderate genetic diversity but narrow climatic niche of the ancient, critically endangered Naufraga (Apiaceae). *Perspectives in Plant Ecology, Evolution and Systematics, 16*(4), 190-202.

Ferreira, M. T., Cardoso, P., Borges, P. A., Gabriel, R., de Azevedo, E. B., Reis, F., ... & Elias, R. B. (2016). Effects of climate change on the distribution of indigenous species in oceanic islands (Azores). *Climatic Change, 138*(3-4), 603-615

Ferreira, M. T., Cardoso, P., Borges, P. A., Gabriel, R., de Azevedo, E. B., & Elias, R. B. (2019). Implications of climate change to the design of protected areas: The case study of small islands (Azores). *PLoS One, 14*(6), e0218168.

Fois, M., Bacchetta, G., Cogoni, D., & Fenu, G. (2018). Current and future effectiveness of the Natura 2000 network for protecting plant species in Sardinia: a nice and complex strategy in its raw state?. *Journal of Environmental Planning and Management, 61*(2), 332-347.

Fois, M., Fenu, G., Lombrana, A. C., Cogoni, D., & Bacchetta, G. (2015). A practical method to speed up the discovery of unknown populations using Species Distribution Models. *Journal for Nature Conservation, 24*, 42-48.

Fleri, J. R., & Arcese, P. (2021). Predictive mapping to identify refuges for plant communities threatened by earthworm invasion. *Ecological Solutions and Evidence, 2*(2), e12064.

Fortini, L. B., Vorsino, A. E., Amidon, F. A., Paxton, E. H., & Jacobi, J. D. (2015). Large-scale range collapse of Hawaiian forest birds under climate change and the need 21st century conservation options. *PloS one, 10*(10), e0140389.

Fortini, L. B., Kaiser, L. R., Vorsino, A. E., Paxton, E. H., & Jacobi, J. D. (2017). Assessing the potential of translocating vulnerable forest birds by searching for novel and enduring climatic ranges. *Ecology and Evolution, 7*(21), 9119-9130.

Fois, M., Bacchetta, G., Cuena-Lombraña, A., Cogoni, D., Pinna, M. S., Sulis, E., & Fenu, G. (2018). Using extinctions in species distribution models to evaluate and predict threats: a contribution to plant conservation planning on the island of Sardinia. *Environmental Conservation, 45*(1), 11-19.

Foster, E. A., & Ackerman, J. D. (2021). Future changes in the distribution of two non-indigenous orchids and their acquired enemy in Puerto Rico. *Biological Invasions, 23*, 3545-3563.

Gehara, M., Summers, K., & Brown, J. L. (2013). Population expansion, isolation and selection: novel insights on the evolution of color diversity in the strawberry poison frog. *Evolutionary Ecology, 27*, 797-824.

Georgopoulou, E., Djursvoll, P., & Simaiakis, S. M. (2016). Predicting species richness and distribution ranges of centipedes at the northern edge of Europe. *Acta oecologica, 74*, 1-10.

Gibson, L. M., Mychajliw, A. M., Leon, Y., Rupp, E., & Hadly, E. A. (2019). Using the past to contextualize anthropogenic impacts on the present and future distribution of an endemic Caribbean mammal. *Conservation Biology, 33*(3), 500-510.

Puddu, G., & Maiorano, L. (2016). Combining multiple tools to provide realistic potential distributions: species distribution models, morphological spatial pattern analysis and circuit theory. *Hystrix*, *27*(1), 1-7.

Goedecke, F., Marcenò, C., Guarino, R., Jahn, R., & Bergmeier, E. (2020). Reciprocal extrapolation of species distribution models between two islands–specialists perform better than generalists and geological data reduces prediction accuracy. *Ecological Indicators, 108*, 105652.

Gonzalez, V. H., Koch, J. B., & Griswold, T. (2010). Anthidium vigintiduopunctatum Friese (Hymenoptera: Megachilidae): the elusive "dwarf bee" of the Galápagos Archipelago? *Biological Invasions, 12*, 2381-2383.

Hannah, L., Aguilar, G., & Blanchon, D. (2019). Spatial distribution of the Mexican daisy, Erigeron karvinskianus, in New Zealand under climate change. *Climate, 7*(2), 24.

Hodd, R. L., Bourke, D., & Skeffington, M. S. (2014). Projected range contractions of European protected oceanic montane plant communities: focus on climate change impacts is essential for their future conservation. *PloS one, 9*(4), e95147.

Hansen, M. F., Nawangsari, V. A., van Beest, F. M., Schmidt, N. M., Fuentes, A., Traeholt, C., ... & Dabelsteen, T. (2019). Estimating densities and spatial distribution of a commensal primate species, the long-tailed macaque (Macaca fascicularis). *Conservation Science and Practice, 1*(9), e88.

Heinänen, S., Erola, J., & von Numers, M. (2012). High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. *Landscape Ecology, 27*, 545-555.

Henareh Khalyani, A., Gould, W. A., Falkowski, M. J., Muscarella, R., Uriarte, M., & Yousef, F. (2019). Climate change increases potential plant species richness on Puerto Rican uplands. *Climatic Change, 156*, 15-30.

Herrera, J. P., Borgerson, C., Tongasoa, L., Andriamahazoarivosoa, P., Rasolofoniaina, B. R., Rakotondrafarasata, E. R., ... & Golden, C. D. (2018). Estimating the population size of lemurs based on their mutualistic food trees. *Journal of Biogeography, 45(*11), 2546-2563.

Hoover, J. D., Kumar, S., James, S. A., Leisz, S. J., & Laituri, M. (2017). Modeling hotspots of plant diversity in New Guinea. *Tropical Ecology, 58*(3).

Hsu, R. C. C., Tamis, W. L., Raes, N., de Snoo, G. R., Wolf, J. H., Oostermeijer, G., & Lin, S. H. (2012). Simulating climate change impacts on forests and associated vascular epiphytes in a subtropical island of East Asia. *Diversity and distributions, 18*(4), 334-347.

Hsu, R. C. C., & Wolf, J. H. (2013). A novel approach to simulate climate change impacts on vascular epiphytes: case study in Taiwan. *Treetops at risk: challenges of Global Canopy Ecology and Conservation*, 123-130.

Hsu, R. C. C., Wolf, J. H., & Tamis, W. L. (2014). Regional and elevational patterns in vascular epiphyte richness on an East Asian island. *Biotropica, 46*(5), 549-555.

Hsu, R. C. C., Wolf, J. H., Tsai, J. M., & Lin, Y. C. (2018). The long-term effect of typhoons on vascular epiphytes in Taiwan. *Journal of Tropical Ecology, 34*(5), 308-315.

Huang, B. H., Huang, C. W., Huang, C. L., & Liao, P. C. (2017). Continuation of the genetic divergence of ecological speciation by spatial environmental heterogeneity in island endemic plants. *Scientific reports, 7*(1), 1-13.

Hunter, D. O., Britz, T., Jones, M., & Letnic, M. (2015). Reintroduction of Tasmanian devils to mainland Australia can restore top-down control in ecosystems where dingoes have been extirpated. *Biological Conservation, 191*, 428-435.

Iannella, M., D'Alessandro, P., Longo, S., & Biondi, M. (2019). New records and potential distribution by Ecological Niche Modeling of Monoxia obesula in the Mediterranean area*. Bulletin of Insectology, 72*(1), 135-142.

Hunter, E. A., Gibbs, J. P., Cayot, L. J., & Tapia, W. (2013). Equivalency of Galápagos giant tortoises used as ecological replacement species to restore ecosystem functions*. Conservation Biology, 27*(4), 701-709.

Ibouroi, M. T., Cheha, A., Astruc, G., Dhurham, S. A. O., & Besnard, A. (2018). A habitat suitability analysis at multi-spatial scale of two sympatric flying fox species reveals the urgent need for conservation action. *Biodiversity and Conservation, 27*, 2395-2423.

Ibouroi, M. T., Hassane, N. A., Moindjié, S., Ombade, M., Mohamed, N., Saidou, M. H., ... & Chiffard, J. (2019). The first comprehensive survey of habitat suitability and population size for the endangered Grande Comoro Scops Owl (Otus pauliani): implications for its conservation. *Journal of Ornithology, 160*, 1121-1132.

Idris, N. A., Othman, N. L., & Abang, F. (2020). Distribution modeling of the Lamproptera species (Papilionidae: Leptocircini) in Borneo. *Serangaa, 25*(2). 43-55.

Irl, S. D., Obermeier, A., Beierkuhnlein, C., & Steinbauer, M. J. (2020). Climate controls plant life-form patterns on a high-elevation oceanic island. *Journal of Biogeography, 47*(10), 2261-2273.

Iryadi, R., & Kurniawati, F. (2021, April). Habitat suitability model of Agarwood in a changing climate. *In IOP Conference Series: Earth and Environmental Science ,724*(1), 12022.

Jarvie, S., Worthy, T. H., Saltré, F., Scofield, R. P., Seddon, P. J., & Cree, A. (2021). Using Holocene fossils to model the future: Distribution of climate suitability for tuatara, the last rhynchocephalian*. Journal of Biogeography, 48*(6), 1489-1502.

Jenkins, D. A., Lecomte, N., Andrews, G., Yannic, G., & Schaefer, J. A. (2020). Biotic interactions govern the distribution of coexisting ungulates in the Arctic Archipelago–A case for conservation planning. *Global Ecology and Conservation, 24*, e01239.

Jensen, R. A., Sunde, P., & Nachman, G. (2012). Predicting the distribution of Tawny Owl (Strix aluco) at the scale of individual territories in Denmark. *Journal of ornithology, 153*, 677-689.

Jiménez-Valverde, A., Diniz, F., de Azevedo, E. B., & Borges, P. A. (2009). Species distribution models do not account for abundance: the case of arthropods on Terceira Island. *Annales Zoologici Fennici, 46*(6):451-464

Jones, A. R., Jessop, T. S., Ariefiandy, A., Brook, B. W., Brown, S. C., Ciofi, C., ... & Fordham, D. A. (2020). Identifying island safe havens to prevent the extinction of the World’s largest lizard from global warming. *Ecology and evolution, 10*(19), 10492-10507.

Juiling, S., Leon, S. K., Jumian, J., Tsen, S., Lee, Y. L., Khoo, E., ... & Maycock, C. R. (2020). Conservation assessment and spatial distribution of endemic orchids in Sabah, Borneo. *Nature Conservation Research, 5*(Suppl. 1), 136-144

Kariyawasam, C. S., Kumar, L., & Ratnayake, S. S. (2019). Invasive plant species establishment and range dynamics in Sri Lanka under climate change. *Entropy, 21*(6), 571.

Kariyawasam, C. S., Kumar, L., Ratnayake, S. S., & Wijesundara, D. S. A. (2021). Potential risks of Invasive Alien Plant Species on native plant biodiversity in Sri Lanka due to climate change. *Biodiversity, 22*(1-2), 24-34.

Kariyawasam, C. S., & Ratnayake, S. S. (2019). Invasive ranges of Ulex europaeus (Fabaceae) in South Australia and Sri Lanka using species distribution modeling. Internationa *Journal of Scientific and Research Publication, 9*(3), 91-100.

Kassara, C., Gangoso, L., Mellone, U., Piasevoli, G., Hadjikyriakou, T. G., Tsiopelas, N., ... & Gschweng, M. (2017). Current and future suitability of wintering grounds for a long-distance migratory raptor. *Scientific Reports, 7*(1), 8798.

Kimura, M. K., Uchiyama, K., Nakao, K., Moriguchi, Y., San Jose-Maldia, L., & Tsumura, Y. (2014). Evidence for cryptic northern refugia in the last glacial period in Cryptomeria japonica. *Annals of botany, 114*(8), 1687-1700.

Kittle, A. M., Watson, A. C., Cushman, S. A., & Macdonald, D. W. (2018). Forest cover and level of protection influence the island-wide distribution of an apex carnivore and umbrella species, the Sri Lankan leopard (*Panthera pardus kotiya*). *Biodiversity and Conservation, 27*, 235-263.

Koch, J. B., & General, D. E. M. (2019). A preliminary assessment of bumble bee (Hymenoptera: Apidae) habitat suitability across protected and unprotected areas in the Philippines. *Annals of the Entomological Society of America, 112*(1), 44-49.

Koide, D., Higa, M., Nakao, K., Ohashi, H., Tsuyama, I., Matsui, T., & Tanaka, N. (2016). Projecting spatiotemporal changes in suitable climate conditions to regenerate trees using niche differences between adult and juvenile trees. *European journal of forest research, 135*, 125-136.

Koo, K. A., Kong, W. S., Park, S. U., Lee, J. H., Kim, J., & Jung, H. (2017). Sensitivity of Korean fir (Abies koreana Wils.), a threatened climate relict species, to increasing temperature at an island subalpine area. *Ecological Modelling, 353*, 5-16.

Kougioumoutzis, K., Kokkoris, I. P., Panitsa, M., Trigas, P., Strid, A., & Dimopoulos, P. (2020). Management implications based on diversity patterns under climate change scenarios in a continental island biodiversity hotspot. *bioRxiv*, 2020-03.

Kougioumoutzis, K., Kokkoris, I. P., Panitsa, M., Trigas, P., Strid, A., & Dimopoulos, P. (2020). Plant diversity patterns and conservation implications under climate-change scenarios in the Mediterranean: The case of Crete (Aegean, Greece). *Diversity, 12*(7), 270.

Kuo, H. C., Chen, S. F., Fang, Y. P., Flanders, J., & Rossiter, S. J. (2014). Comparative rangewide phylogeography of four endemic T aiwanese bat species*. Molecular Ecology, 23*(14), 3566-3586.

Kurita, T. (2013). Current status of the introduced common house gecko, Hemidactylus frenatus (Squamata: Gekkonidae), on Amamioshima Island of the Ryukyu Archipelago, Japan. *Current herpetology, 32*(1), 50-60.

La Marca, W., Elith, J., Firth, R. S., Murphy, B. P., Regan, T. J., Woinarski, J. C., & Nicholson, E. (2019). The influence of data source and species distribution modelling method on spatial conservation priorities. *Diversity and Distributions, 25*(7), 1060-1073.

Lamelas-López, L., Pardavila, X., Borges, P. A., Santos-Reis, M., Amorim, I. R., & Santos, M. J. (2020). Modelling the distribution of Mustela nivalis and M. putorius in the Azores archipelago based on native and introduced ranges. *Plos one, 15(*8), e0237216.

Lannuzel, G., Balmot, J., Dubos, N., Thibault, M., & Fogliani, B. (2021). High-resolution topographic variables accurately predict the distribution of rare plant species for conservation area selection in a narrow-endemism hotspot in New Caledonia. *Biodiversity and Conservation, 30*(4), 963-990.

Lentini, P. E., Stirnemann, I. A., Stojanovic, D., Worthy, T. H., & Stein, J. A. (2018). Using fossil records to inform reintroduction of the kakapo as a refugee species. *Biological Conservation, 217*, 157-165.

Lenton, S. M., Fa, J. E., & Del Val, J. P. (2000). A simple non-parametric GIS model for predicting species distribution: endemic birds in Bioko Island, West Africa. *Biodiversity & Conservation, 9*, 869-885.

Liao, C. C., & Chen, Y. H. (2021). Improving performance of species distribution model in mountainous areas with complex topography. *Ecological Research, 36*(4), 648-662.

Libeau, M., Meyer, J. Y., Taputuarai, R., & Pouteau, R. (2019). Predicting the invasion risk of *Miconia calvescens* in the Marquesas Islands (South Pacific): a modeling approach. *Pacific Science, 73*(1), 17-34.

Liew, T. S., Marzuki, M. E., Schilthuizen, M., Chen, Y., Vermeulen, J. J., & Mohd-Azlan, J. (2020). Molecular phylogenetics and evolutionary history of the endemic land snail genus *Everettia* in northern Borneo. *PeerJ, 8*, e9416.

Lin, C. T., & Chiu, C. A. (2018). The Relic Trochodendron aralioides Siebold & Zucc.(Trochodendraceae) in Taiwan: Ensemble distribution modeling and climate change impacts. *Forests, 10*(1), 7.

Lin, Y. P., Deng, D., Lin, W. C., Lemmens, R., Crossman, N. D., Henle, K., & Schmeller, D. S. (2015). Uncertainty analysis of crowd-sourced and professionally collected field data used in species distribution models of Taiwanese moths. *Biological conservation, 181,* 102-110.

Lin, W. C., Lin, Y. P., Lien, W. Y., Wang, Y. C., Lin, C. T., Chiou, C. R., ... & Crossman, N. D. (2014). Expansion of protected areas under climate change: an example of mountainous tree species in Taiwan. *Forests, 5*(11), 2882-2904.

Lloret, F., & González-Mancebo, J. M. (2011). Altitudinal distribution patterns of bryophytes in the Canary Islands and vulnerability to climate change. Flora-Morphology, Distribution, *Functional Ecology of Plants, 206*(9), 769-781.

Longcore, T., Noujdina, N., & Dixon, P. J. (2018). Landscape modeling of the potential natural vegetation of Santa Catalina Island, California. *Western North American Naturalist, 78*(4), 617-632.

Louca, M., Vogiatzakis, I. N., & Moustakas, A. (2015). Modelling the combined effects of land use and climatic changes: Coupling bioclimatic modelling with Markov-chain Cellular Automata in a case study in Cyprus. *Ecological Informatics, 30,* 241-249.

Lozano, V., Marzialetti, F., Carranza, M. L., Chapman, D., Branquart, E., Dološ, K., ... & Brundu, G. (2020). Modelling Acacia saligna invasion in a large Mediterranean island using PAB factors: A tool for implementing the European legislation on invasive species. *Ecological Indicators, 116*, 106516.

Manchanda, G., & Garg, N. (2011). Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology: Official Journal of the Societa Botanica Italiana.

Mangiacotti, M., Scali, S., Sacchi, R., Bassu, L., Nulchis, V., & Corti, C. (2013). Assessing the spatial scale effect of anthropogenic factors on species distribution. *PLoS One, 8*(6), e67573

Mercado Malabet, F., Peacock, H., Razafitsalama, J., Birkinshaw, C., & Colquhoun, I. (2020). Realized distribution patterns of crowned lemurs (*Eulemur coronatus*) within a human-dominated forest fragment in northern Madagascar. *American Journal of Primatology, 82*(4), e23125.

Mimura, M., Mishima, M., Lascoux, M., & Yahara, T. (2014). Range shift and introgression of the rear and leading populations in two ecologically distinct Rubus species. *BMC evolutionary biology, 14*(1), 1-13.

Moriguchi, S., Inoue, M. N., Kishimoto, T., Kameyama, T., Ito, F., & Goka, K. (2015). Estimating colonization and invasion risk maps for Linepithema humile, in Japan. *Journal of Asia-Pacific Entomology, 18*(2), 343-350.

Murienne, J., Guilbert, E., & Grandcolas, P. (2009). Species' diversity in the New Caledonian endemic genera Cephalidiosus and Nobarnus (Insecta: Heteroptera: Tingidae), an approach using phylogeny and species' distribution modelling. Biological *Journal of the Linnean Society, 97*(1), 177-184

Nakao, K., Higa, M., Tsuyama, I., Lin, C. T., Sun, S. T., Lin, J. R., ... & Tanaka, N. (2014). Changes in the potential habitats of 10 dominant evergreen broad-leaved tree species in the Taiwan-Japan archipelago. *Plant Ecology, 215*, 639-650.

Nania, D., Flecks, M., & Rödder, D. (2020). Continuous expansion of the geographic range linked to realized niche expansion in the invasive Mourning gecko Lepidodactylus lugubris (Duméril & Bibron, 1836). *PLoS One, 15*(7), e0235060.

Nursamsi, I., Partasasmita, R., Cundaningsih, N., & Ramadhani, H. S. (2018). Modeling the predicted suitable habitat distribution of Javan hawk-eagle Nisaetus bartelsi in the Java Island, Indonesia. *Biodiversitas Journal of Biological Diversity, 19*(4), 1539-1551.

O'Donnell, J., Gallagher, R. V., Wilson, P. D., Downey, P. O., Hughes, L., & Leishman, M. R. (2012). Invasion hotspots for non-native plants in a ustralia under current and future climates. *Global Change Biology, 18*(2), 617-629.

Oh, S. Y., Woo, J. J., & Hur, J. S. (2019). Distribution of foliicolous lichen Strigula and genetic structure of S. multiformis on Jeju Island, South Korea. *Microorganisms, 7(*10), 430.

Ogawa-Onishi, Y., Berry, P. M., & Tanaka, N. (2010). Assessing the potential impacts of climate change and their conservation implications in Japan: a case study of conifers. *Biological Conservation, 143*(7), 1728-1736.

Pang, S. E., De Alban, J. D. T., & Webb, E. L. (2021). Effects of climate change and land cover on the distributions of a critical tree family in the Philippines. *Scientific Reports, 11*(1), 276.

Papadopulos, A. S. T., Price, Z., Devaux, C., Hipperson, H., Smadja, C. M., Hutton, I., ... & Savolainen, V. (2013). A comparative analysis of the mechanisms underlying speciation on Lord Howe Island. *Journal of Evolutionary Biology, 26*(4), 733-745.

Patiño, J., Carine, M., Mardulyn, P., Devos, N., Mateo, R. G., González-Mancebo, J. M., ... & Vanderpoorten, A. (2015). Approximate Bayesian computation reveals the crucial role of oceanic islands for the assembly of continental biodiversity. *Systematic Biology, 64*(4), 579-589.

Perez-Correa, J., Carr, P., Meeuwig, J. J., Koldewey, H. J., & Letessier, T. B. (2020). Climate oscillation and the invasion of alien species influence the oceanic distribution of seabirds. *Ecology and evolution, 10*(17), 9339-9357.

Pertierra, L. R., Aragón, P., Shaw, J. D., Bergstrom, D. M., Terauds, A., & Olalla-Tárraga, M. Á. (2017). Global thermal niche models of two European grasses show high invasion risks in Antarctica. *Global Change Biology, 23*(7), 2863-2873.

Pertierra, L. R., Baker, M., Howard, C., Vega, G. C., Olalla-Tarraga, M. A., & Scott, J. (2016). Assessing the invasive risk of two non-native Agrostis species on sub-Antarctic Macquarie Island. *Polar Biology, 39*, 2361-2371.

Pertierra, L. R., Bartlett, J. C., Duffy, G. A., Vega, G. C., Hughes, K. A., Hayward, S. A., ... & Aragón, P. (2020). Combining correlative and mechanistic niche models with human activity data to elucidate the invasive potential of a sub-Antarctic insect. *Journal of Biogeography, 47*(3), 658-673.

Picanço, A., Rigal, F., & Borges, P. A. (2017). Area prioritization for insect pollinator communities on an Oceanic Island. *Arquipélago-Life and Marine Sciences, 34*, 85-104.

Piquet, J. C., Warren, D. L., Bolaños, J. F. S., Rivero, J. M. S., Gallo-Barneto, R., Cabrera-Pérez, M. Á., ... & López-Darias, M. (2021). Could climate change benefit invasive snakes? Modelling the potential distribution of the California Kingsnake in the Canary Islands. *Journal of Environmental Management, 294*, 112917.

Polce, C., Termansen, M., Aguirre-Gutiérrez, J., Boatman, N. D., Budge, G. E., Crowe, A., ... & Biesmeijer, J. C. (2013). Species distribution models for crop pollination: a modelling framework applied to Great Britain. *PloS one, 8*(10), e76308.

Porfirio, L. L., Harris, R. M., Stojanovic, D., Webb, M. H., & Mackey, B. (2016). Projected direct and indirect effects of climate change on the Swift Parrot, an endangered migratory species. *Emu-Austral Ornithology, 116*(3), 273-283.

Potter, S., Xue, A. T., Bragg, J. G., Rosauer, D. F., Roycroft, E. J., & Moritz, C. (2018). Pleistocene climatic changes drive diversification across a tropical savanna. *Molecular ecology, 27*(2), 520-532.

Pouteau, R., Bayle, É., Blanchard, É., Birnbaum, P., Cassan, J. J., Hequet, V., ... & Vandrot, H. (2015). Accounting for the indirect area effect in stacked species distribution models to map species richness in a montane biodiversity hotspot. *Diversity and Distributions, 21*(11), 1329-1338.

Pouteau, R., & Birnbaum, P. (2016). Island biodiversity hotspots are getting hotter: vulnerability of tree species to climate change in New Caledonia. *Biological Conservation, 201*, 111-119.

Pouteau, R., Hulme, P. E., & Duncan, R. P. (2015). Widespread native and alien plant species occupy different habitats. *Ecography, 38*(5), 462-471.

Pouteau, R., Munoz, F., & Birnbaum, P. (2019). Disentangling the processes driving tree community assembly in a tropical biodiversity hotspot (New Caledonia). *Journal of Biogeography, 46*(4), 796-806.

Pouteau, R., Trueba, S., Feild, T. S., & Isnard, S. (2015). New Caledonia: a Pleistocene refugium for rain forest lineages of relict angiosperms. *Journal of Biogeography, 42*(11), 2062-2077.

Proft, K. M., Bateman, B. L., Johnson, C. N., Jones, M. E., Pauza, M., & Burridge, C. P. (2021). The effects of weather variability on patterns of genetic diversity in Tasmanian bettongs. *Molecular Ecology, 30*(8), 1777-1790.

Rabara, R., Sotto, R. C., & SALAS, E. A. L. (2020). Species distribution modeling and phenotypic diversity reveals collection gap in *the Musa balbisiana* germplasm conservation in Philippines. *Asian Journal of Agriculture*, *4*(2), 60-71

Raes, N., Cannon, C. H., Hijmans, R. J., Piessens, T., Saw, L. G., van Welzen, P. C., & Slik, J. F. (2014). Historical distribution of Sundaland's Dipterocarp rainforests at Quaternary glacial maxima. *Proceedings of the National Academy of Sciences, 111*(47), 16790-16795.

Raes, N., Roos, M. C., Slik, J. F., Van Loon, E. E., & Steege, H. T. (2009). Botanical richness and endemicity patterns of Borneo derived from species distribution models. *Ecography, 32*(1), 180-192.

Rapacciuolo, G., Roy, D. B., Gillings, S., Fox, R., Walker, K., & Purvis, A. (2012). Climatic associations of British species distributions show good transferability in time but low predictive accuracy for range change. *PLoS One, 7*(7), e40212.

Recio, M. R., Seddon, P. J., & Moore, A. B. (2015). Niche and movement models identify corridors of introduced feral cats infringing ecologically sensitive areas in New Zealand. *Biological Conservation, 192*, 48-56.

Rizvanovic, M., Kennedy, J. D., Nogués-Bravo, D., & Marske, K. A. (2019). Persistence of genetic diversity and phylogeographic structure of three New Zealand forest beetles under climate change. *Diversity and Distributions, 25*(1), 142-153.

Richardson, K. M., & Ewen, J. G. (2016). Habitat selection in a reintroduced population: social effects differ between natal and post-release dispersal. *Animal Conservation, 19*(5), 413-421.

Rodríguez-Rodríguez, P., de Castro, A. G. F., & Sosa, P. A. (2018). The restoration of the endangered Sambucus palmensis after 30 years of conservation actions in the Garajonay National Park: genetic assessment and niche modeling. *PeerJ, 6*, e4985.

Rosauer, D. F., Blom, M. P. K., Bourke, G., Catalano, S., Donnellan, S., Gillespie, G., ... & Moritz, C. (2016). Phylogeography, hotspots and conservation priorities: an example from the Top End of Australia. *Biological Conservation, 204*, 83-93.

Rosenblad, K. C., Perret, D. L., & Sax, D. F. (2019). Niche syndromes reveal climate-driven extinction threat to island endemic conifers. *Nature Climate Change, 9*(8), 627-631.

Rotllan-Puig, X., & Traveset, A. (2016). Declining relict plants: climate effect or seed dispersal disruption? A landscape-scale approach. *Basic and Applied Ecology, 17*(1), 81-91.

Rutgrink, A. L., Visser, M., & van Welzen, P. C. (2018). Differences between the floras of the North and South Moluccas (Indonesia). J*ournal of Systematics and Evolution, 56*(6), 652-662.

Saito, M. U., Momose, H., Inoue, S., Kurashima, O., & Matsuda, H. (2016). Range-expanding wildlife: modelling the distribution of large mammals in Japan, with management implications. *International Journal of Geographical Information Science, 30*(1), 20-35.

Salvi, D., Pinho, C., & Harris, D. J. (2017). Digging up the roots of an insular hotspot of genetic diversity: decoupled mito-nuclear histories in the evolution of the Corsican-Sardinian endemic lizard Podarcis tiliguerta. *BMC evolutionary biology, 17*, 1-22.

Sandamal, S., Tennakoon, A., Meng, Q. L., Marambe, B., Ratnasekera, D., Melo, A., & Ge, S. (2018). Population genetics and evolutionary history of the wild rice species *Oryza rufipogon* and *O. nivara* in Sri Lanka. *Ecology and evolution, 8*(23), 12056-12065.

Santillan, M., & Santillan, J. (2014, October). Species distribution modeling to aid remote sensing of the starch-rich sago palm in the Philippines. In Proceedings of the 35th *Asian Conference on Remote Sensing (ACRS 2014)–Sensing for Reintegration of Society*. pp. 27-31.

Saputra, M. H., Sagala, S. A. H., & Lee, H. S. (2020, January). Species Distribution of *Styrax sumatrana* in North Sumatra Using Maxent Modelling Approach. *In Forum Geografi, 33*(2), 196-208.

Saputra, M. H., & Sunandar, A. D. (2020, June). Predicting the potential distribution of Taxussumatrana using Maximum Entropy Model. *In IOP Conference Series: Earth and Environmental Science* (Vol. 522, No. 1, p. 012013). IOP Publishing.

Saputra, M. H., & Lee, H. S. (2021). Evaluation of climate change impacts on the potential distribution of *Styrax sumatrana* in North Sumatra, Indonesia. *Sustainability, 13*(2), 462.

Scriven, S. A., Williams, S. H., Ghani, M. A., Agama, A. L., Benedick, S., Brodie, J. F., ... & Hill, J. K. (2020). Assessing the effectiveness of protected areas for conserving range-restricted rain forest butterflies in Sabah, Borneo. *Biotropica, 52*(2), 380-391.

Senan, A. S., Tomasetto, F., Farcomeni, A., Somashekar, R. K., & Attorre, F. (2012). Determinants of plant species invasions in an arid island: evidence from Socotra Island (Yemen). *Plant Ecology, 213*, 1381-1392.

Segal, R. D., Massaro, M., Carlile, N., & Whitsed, R. (2021). Small-scale species distribution model identifies restricted breeding habitat for an endemic island bird. Animal Conservation, 24(6), 959-969.

Senczuk, G., Harris, D. J., Castiglia, R., Litsi Mizan, V., Colangelo, P., Canestrelli, D., & Salvi, D. (2019). Evolutionary and demographic correlates of Pleistocene coastline changes in the Sicilian wall lizard Podarcis wagleriana. *Journal of Biogeography, 46(*1), 224-237

Setyawan, A. D., Supriatna, J., Darnaedi, D., ROKHMATULOH, R., SUTARNO, S., SUGIYARTO, S., ... & PRADAN, P. (2017). Impact of climate change on potential distribution of xero-epiphytic selaginellas (Selaginella involvens and S. repanda) in Southeast Asia. *Biodiversitas Journal of Biological Diversity, 18*(4), 1680-1695

Setyawan, A. D., Supriatna, J., Nursamsi, I., SUTARNO, S., SUGIYARTO, S., SUNARTO, S., ... & INDRAWAN, M. (2020). Anticipated climate changes reveal shifting in habitat suitability of high-altitude selaginellas in Java, Indonesia. *Biodiversitas Journal of Biological Diversity, 21*(11).

Setyawan, A. D., Supriatna, J., NISYAWATI, N., NURSAMSI, I., SUTARNO, S., SUGIYARTO, S., ... & INDRAWAN, M. (2020). Predicting potential impacts of climate change on the geographical distribution of mountainous selaginellas in Java, Indonesia. *Biodiversitas Journal of Biological Diversity, 21*(10).

Sheppard, C. S. (2013). How does selection of climate variables affect predictions of species distributions? A case study of three new weeds in New Zealand. *Weed Research, 53*(4), 259-268.

Sheppard, C. S. (2013). Potential spread of recently naturalised plants in New Zealand under climate change. *Climatic Change, 117*, 919-931.

Sheppard, C. S., Burns, B. R., & Stanley, M. C. (2014). Predicting plant invasions under climate change: are species distribution models validated by field trials?. *Global Change Biology, 20*(9), 2800-2814.

Shimazaki, M., Tsuyama, I., Nakazono, E., Nakao, K., Konoshima, M., Tanaka, N., & Nakashizuka, T. (2012). Fine-resolution assessment of potential refugia for a dominant fir species (Abies mariesii) of subalpine coniferous forests after climate change. *Plant Ecology, 213*, 603-612.

Silva, L. D., Costa, H., de Azevedo, E. B., Medeiros, V., Alves, M., Elias, R. B., & Silva, L. (2017). Modelling native and invasive woody species: a comparison of ENFA and MaxEnt applied to the Azorean forest. In *Modeling, Dynamics, Optimization and Bioeconomics II: DGS III, Porto, Portugal, February 2014, and Bioeconomy VII, Berkeley, USA, March 2014-Selected Contributions 3* (pp. 415-444). Springer International Publishing.

Silva, L., Dias, E. F., Sardos, J., Azevedo, E. B., Schaefer, H., & Moura, M. (2015). Towards a more holistic research approach to plant conservation: the case of rare plants on oceanic islands. *AoB Plants*, *7*, plv066.

Sodik, M., Pudyatmoko, S., Yuwono, P. S. H., Tafrichan, M., & Imron, M. A. (2020). Better providers of habitat for Javan slow loris (Nycticebus javanicus E. Geoffroy 1812): A species distribution modeling approach in Central Java, Indonesia. *Biodiversitas Journal of Biological Diversity, 21*(5), 1890-1900.

Suárez, N. M., Pestano, J., & Brown, R. P. (2014). Ecological divergence combined with ancient allopatry in lizard populations from a small volcanic island. *Molecular ecology, 23*(19), 4799-4812.

Sawyer, Y. E., MacDonald, S. O., Lessa, E. P., & Cook, J. A. (2019). Living on the edge: Exploring the role of coastal refugia in the Alexander Archipelago of Alaska. *Ecology and Evolution, 9*(4), 1777-1797

Spiers, J. A., Oatham, M. P., Rostant, L. V., & Farrell, A. D. (2018). Applying species distribution modelling to improving conservation based decisions: a gap analysis of Trinidad and Tobago's endemic vascular plants. *Biodiversity and Conservation, 27*, 2931-2949.

Sutton, L. J., De Roland, L. A. R., Thorstrom, R., & McClure, C. J. (2022). Distribution and habitat use of the Madagascar Peregrine Falcon: first estimates for area of habitat and population size. *Bird Conservation International, 32*(4), 624-640.

Suzuki-Ohno, Y., Morita, K., Nagata, N., Mori, H., Abe, S., Makino, T., & Kawata, M. (2017). Factors restricting the range expansion of the invasive green anole Anolis carolinensis on Okinawa Island, Japan. *Ecology and Evolution, 7*(12), 4357-4366.

Tabak, M. A., Poncet, S., Passfield, K., & del Rio, C. M. (2015). Modeling the distribution of Norway rats (Rattus norvegicus) on offshore islands in the Falkland Islands. *NeoBiota, 24*, 33-48.

Tabor, J. A., & Koch, J. B. (2021). Ensemble models predict invasive bee habitat suitability will expand under future climate scenarios in Hawai'i. *Insects, 12*(5), 443.

Talavera, G., Espadaler, X., & Vila, R. (2015). Discovered just before extinction? The first endemic ant from the Balearic Islands (Lasius balearicus sp. nov.) is endangered by climate change*. Journal of Biogeography, 42*(3), 589-601.

Tamar, K., Simó-Riudalbas, M., Garcia-Porta, J., Santos, X., Llorente, G., Vasconcelos, R., & Carranza, S. (2019). An integrative study of island diversification: insights from the endemic Haemodracon geckos of the Socotra Archipelago*. Molecular phylogenetics and evolution, 133*, 166-175.

Taylor, S., & Kumar, L. (2014). Climate Change and Weed Impacts on Small Island Ecosystems: Lantana camara L.(Magnoliopsida: Verbenaceae) Distribution in Fiji1. *Pacific Science, 68*(1), 117-133.

Tournebize, R., Manel, S., Vigouroux, Y., Munoz, F., De Kochko, A., & Poncet, V. (2017). Two disjunct Pleistocene populations and anisotropic postglacial expansion shaped the current genetic structure of the relict plant *Amborella trichopoda*. *Plos one, 12*(8), e0183412.

Tsuda, Y., Nakao, K., Ide, Y., & Tsumura, Y. (2015). The population demography of B etula maximowicziana, a cool-temperate tree species in Japan, in relation to the last glacial period: its admixture-like genetic structure is the result of simple population splitting not admixing. *Molecular Ecology, 24*(7), 1403-1418.

Tsuyama, I., Nakao, K., Matsui, T., Higa, M., Horikawa, M., Kominami, Y., & Tanaka, N. (2011). Climatic controls of a keystone understory species, Sasamorpha borealis, and an impact assessment of climate change in Japan. *Annals of Forest Science, 68(*4), 689-699.

Tsuyama, I., Nakao, K., Higa, M., Matsui, T., Shichi, K., & Tanaka, N. (2014). What controls the distribution of the Japanese endemic hemlock, Tsuga diversifolia? Footprint of climate in the glacial period on current habitat occupancy. *Journal of forest research, 19*(1), 154-165.

Turvey, S. T., Kennerley, R. J., Hudson, M. A., Nuñez-Miño, J. M., & Young, R. P. (2020). Assessing congruence of opportunistic records and systematic surveys for predicting Hispaniolan mammal species distributions. *Ecology and Evolution, 10*(11), 5056-5068.

Untalan, M. Z. G., Burgos, D. F. M., & Martinez, K. P. (2019). SPECIES DISTRIBUTION MODELLING OF TWO SPECIES ENDEMIC TO THE PHILIPPINES TO SHOW THE APPLICABILITY OF MAXENT. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 449-454.

Upson, R., Williams, J. J., Wilkinson, T. P., Clubbe, C. P., Maclean, I. M., McAdam, J. H., & Moat, J. F. (2016). Potential impacts of climate change on native plant distributions in the Falkland Islands. *PloS one, 11*(11), e0167026.

Sutomo, & Van Etten, E. (2017). Species distribution model of invasive alien species Acacia nilotica for Central-Eastern Indonesia using Biodiversity Climate Change Virtual Laboratory (BCCVL). *Tropical Drylands* **1**, 36-42.

Vasconcelos, R., Razgour, O., Tarroso, P., Fasola, M., Carranza, S., & Alves, P. C. (2018). Combining molecular and landscape tools for targeting evolutionary processes in reserve design: An approach for islands. Plos one, 13(7), e0200830.

Vences, M., Brown, J. L., Lathrop, A., Rosa, G. M., Cameron, A., Crottini, A., ... & Murphy, R. W. (2017). Tracing a toad invasion: lack of mitochondrial DNA variation, haplotype origins, and potential distribution of introduced Duttaphrynus melanostictus in Madagascar. *Amphibia-Reptilia, 38*(2), 197-207.

Vergilio, M., Fonseca, C., Calado, H., Borges, P. A., Elias, R. B., Gabriel, R., ... & Cardoso, P. (2016). Assessing the efficiency of protected areas to represent biodiversity: a small island case study. *Environmental Conservation, 43*(4), 337-349.

Vetter, V. M., Tjaden, N. B., Jaeschke, A., Buhk, C., Wahl, V., Wasowicz, P., & Jentsch, A. (2018). Invasion of a legume ecosystem engineer in a cold biome alters plant biodiversity. *Frontiers in Plant Science, 9*, 715.

Vieilledent, G., Cornu, C., Sanchez, A. C., Pock-Tsy, J. M. L., & Danthu, P. (2013). Vulnerability of baobab species to climate change and effectiveness of the protected area network in Madagascar: *Towards new conservation priorities. Biological Conservation, 166*, 11-22.

Villa-Machío, I., Fernandez de Castro, A. G., Fuertes-Aguilar, J., & Nieto Feliner, G. (2020). Colonization history of the Canary Islands endemic Lavatera acerifolia,(Malvaceae) unveiled with genotyping-by-sequencing data and niche modelling. *Journal of Biogeography, 47*(4), 993-1005.

Vollering, J., Schuiteman, A., de Vogel, E., van Vugt, R., & Raes, N. (2016). Phytogeography of New Guinean orchids: patterns of species richness and turnover. *Journal of Biogeography, 43*(1), 204-214.

Wang, W. C., Lo, N. J., Chang, W. I., & Huang, K. Y. (2012). Modeling spatial distribution of a rare and endangered plant species (Brainea insignis) in Central Taiwan. The International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences, 39*, 241-246.

Wan, J. N., Mbari, N. J., Wang, S. W., Liu, B., Mwangi, B. N., Rasoarahona, J. R., ... & Wang, Q. F. (2021). Modeling impacts of climate change on the potential distribution of six endemic baobab species in Madagascar. *Plant Diversity, 43*(2), 117-124.

Wulff, A. S., Hollingsworth, P. M., Ahrends, A., Jaffré, T., Veillon, J. M., L'Huillier, L., & Fogliani, B. (2013). Conservation priorities in a biodiversity hotspot: analysis of narrow endemic plant species in New Caledonia. *PLoS one, 8*(9), e73371.

Yamada, T., Kokubugata, G., Fujii, S., Chen, C. F., Asakawa, A., Ito, T., & Maki, M. (2021). Refugia during the last glacial period and the origin of the disjunct distribution of an insular plant. *Journal of Biogeography, 48*(6), 1460-1474

Yeh, H. T., Cheah, H. Y., Chiu, M. C., Liao, J. R., & Ko, C. C. (2021). Assessment of potential invasion for six phytophagous quarantine pests in Taiwan. *Scientific Reports, 11*(1), 10666

Yudaputra, A. N. G. G. A. (2020). Modelling potential current distribution and future dispersal of an invasive species Calliandra calothyrsus in Bali Island, *Indonesia. Biodiversitas, 21*(2), 674-682.

Yulia, E., & Iryadi, R. (2021). Kirinyuh (Chromolaena odorata): species distribution modeling and the potential use of fungal pathogens for its eradication. *In IOP Conference Series: Earth and Environmental Science, 762*(1), 12023.

Zhang, S. N., & Kubota, K. (2021). Dispersal constraints on the potential distribution of cold-adapted stag beetles (genus Platycerus) in Japan and the implications of climate change. *Insect Conservation & Diversity*, *14*(3).

Zozomová-Lihová, J., Melichárková, A., Svitok, M., & Španiel, S. (2020). Pleistocene range disruption and postglacial expansion with secondary contacts explain the genetic and cytotype structure in the western Balkan endemic Alyssum austrodalmaticum (Brassicaceae). *Plant systematics and evolution, 306*(2), 47.

Appendix S2.2 Supplementary tables and figures

Supplementary table S2.2.2 Key SDM implementation parameters form the reviewed literature

Study Area	Year	Number of Species Per Taxa	Taxa	Niche Truncation	Env Data Resolution	Modelling method	Modeling approach	Aim	Record Mean	Max Number of Records	Min Number of Records	Predictor N
Alexander archipelago	2018	3	MM	NA (Widespread)	4.5km	Maxent	Single	PHP	144	149	137	4
Antarctica and the Southern Ocean islands	2019	1	ART	IN (Widespread)	10km	DB	Single	IAS	15	15	15	5
Antarctica and the Southern Ocean islands	2017	3 / 24 / 3 / 3 / 14 / 2 / 47	AN/ART/BRD/FG/MM/RPT/VP	IN (Widespread)	27.75km	ANN / CTA / FDA / GAM / GBM / GKM / MARS / Maxent / RF / SER	Ensemble	SYN	25050	217124	13	5
Antarctica and the Southern Ocean islands	2017	2	VP	IN (Widespread)	10km	DB / ENFA / Maxent	Single	IAS	35000	35000	35000	7
Azores Islands	2016	67	VP	NT (Local)	100m	Maxent	Single	SHR	17	51	6	5
Azores Islands	2017	45	ART	EN-4/IN-7/NA-35	100m	Maxent	Single	LUC	11	37	3	11
Azores Islands	2015	1	VP	EN (Local)	100m	ENFA / Maxent	Single	MT	7	7	7	6
Azores Islands	2016	51 / 82 / 57	BRY/Mollu	NA (Widespread)	500m	Maxent	Single	PAM	229	548	20	-

			sca/VP									
Azores Islands	2019	2	MM	IN (Widespread)	1km	Maxent	Single	IAS	1226	3,396	24	12
Azores Islands	2015	1	VP	IN (Local)	100m	Maxent	Single	MT	530	530	530	11
Azores Islands	2009	48	ART	NA-43/IN-5	500m	ANN	Single	MT	240	590	69	5
Azores Islands	2021	1	ART	IN (Widespread)	100m	Maxent	Single	IAS	1260	2635	884	5
Azores Islands	2017	2	VP	IN (Local)	500m	ENFA / Maxent	Single	SHR	3885	12922	402	28
Azores Islands	2019	122 / 7 / 50	ART/BRY/VP	EN (Local)	100m	ANN / CTA / FDA / GAM / GBM / GLM / MARS / Maxent / RF / SER	Ensemble	PAM /CC	-	-	-	3
Azores Islands	2016	128/ 19/ 59	ART /BRY /VP	-	500m	BIOCLIM / GAM / GLM / RF	Ensemble	CC	-	-	-	-
Balearic Islands	2014	1	VP	EN (Local)	1km*	Maxent	Single	PHP	5	5	5	7
Balearic Islands	2015	1	ART	EN (Local)	1km	CTA / FDA / GAM / GBM / GLM / MARS / Maxent / RF	Ensemble	CC	139	139	139	5
Balearic Islands	2015	2	VP	EN (Local)	1km	Maxent	Single	SHR	255	478	32	11
Balearic Islands	2019	1	VP	EN (Local)	2m	Maxent	Single	SHR	1603	1603	1603	6
Bioko island	1999	45	BRD	EN (Local)	500m	PREDICT	Single	MT	-	-	-	-

Borneo	2020	17	NAI	EN (Local)	1km	Maxent	Single	PHP	42	187	4	7
Borneo	2020	2	ART	NA (Widespread)	1km*	Maxent	Single	SHR	40	42	38	14
Borneo	2020	47	VP	EN (Local)	200m	Maxent	Single	SHR	9	21	5	6
Borneo	2020	77	ART	EN (Local)	1km	GLM/Maxent/RF	Ensemble	PAM	-	-	-	5
Borneo	2009	2273	VP	NA/EN (Local)	4.5km	Maxent	Single	RP	-	-	-	14
Canadian High Arctic Archipalgo	2020	2	MM	NA (Widespread)	1km (577m final)	Maxent	Single	SHR	1027	1519	535	15
Canary Islands	2019	1	VP	EN (Local)	50m	ANN / GBM / GLM / MARS / RF	Ensemble	PHP	34	34	34	7
Canary Islands	2014	1	RPT	EN (Local)	1km	Maxent	Single	PHP	46	46	46	6
Canary Islands	2018	1	VP	EN (Local)	50m	ANN / GAM / GBM / GLM / MARS / RF	Ensemble	RST	331	441	56	5
Canary Islands	2020	1	VP	IN (Local)	500m	Maxent	Single	SYN	65	65	65	3
Canary Islands	2021	1	RPT	NT&IN	1km	BIOCLIM / DB / ENFA / Maxent	Ensemble	SYN	80	80	80	7
Canary Islands	2020	1	VP	IN (Local)	100m	GAM / Maxent / RF	Ensemble	IAS	207	241	172	10
Canary Islands	2011	401	BRY	NA (Local)	200m elevation bands	Net Diference	Single	CC	-	-	-	3
Canary Islands	2016	1	BRD	-	30m	DM	Single	LUC	-	-	-	-
Canary Islands	2011	841	VP	NA (Local)	500m	Maxent	Single	MT	-	-	-	7
Canary Islands	2020	NA	NA	NA (Local)	500m	GLM	Single	RP	976	-	-	12

Canary Islands	2020	1	VP	NA (Widespread)	1km &4.5Km (past)	Maxent	Single	PHP	60	60	60	7
Chagos Archipelago	2020	4	BRD	WD (Local)	4km	GBM	Single	SYN	173	255	113	-
Corsica–Sardinia island system	2013	1	AN	EN (Local)	1km	Maxent	Single	PHP	15	15	15	6
Corsica–Sardinia island system	2014	23	VP	NA (Local)	1km	Maxent	Single	CC	31	15	48	6
Corsica–Sardinia island system	2011	1	AN	EN (Local)	1km &4.5Km (past)	Maxent	Single	PHP	190	190	190	7/19
Corsica–Sardinia island system	2017	1	RPT	EN (Local)	1km &4.5Km (past)	Maxent	Single	PHP	480	480	480	5
Crete and Sicily	2020	12	VP	NT (Local)	1km	Maxent	Single	MT	180	452	20	10
Crete Island	2020	172	VP	EN (Local)	1km	CTA / MARS / RF	Ensemble	PAM	120	1792	4	7
Crete Island	2020	183	VP	EN (Local)	1km	CTA / MARS / RF	Ensemble	PAM	-	-	-	7
Cyprus Island	2018	1	VP	NA (Local)	1km	Maxent	Single	PHP	27	27	27	19
Cyprus Island	2015	2	VP	EN (Local)	1km/250m	CTA	Single	SYN	109	117	102	3_5
Elba Island	2017	6	VP	IN (Local)	250m	ANN / CTA / FDA / GAM / GBM / GLM / MARS /	Ensemble	IAS	122	150	95	8

						Maxent / RF / SER						
English Company Islands	2016	32	RPT	NA (Local)	1km	Maxent	Single	PHP	242	996	24	15
Falkland Islands	2016	8	VP	EN-3/NT-5 (Local)	100m	GAM / GBM / GLM / Maxent / RF	Ensemble	CC	32	58	20	1_5
Falkland Islands	2015	1	MM	IN (Local)	NE	GLM	Single	IAS	158	158	158	4
Fiji Islands	2014	1	VP	IN (Widespread)	18.5 km	CLIMEX	Single	SYN	1740	1740	1740	5
Finnish archipelago	2012	2	BRD	NA (Local)	10m	Maxent	Single	MT	722	1185	234	7
French sub-Antarctic islands (Possession Island, Crozet archipelago)	2021	6	VP	IN (Local)	30m	GLM	Single	SYN	427	439	79	2
Galapagos (Pinta island)	2013	2	RPT	IN (Widespread)	90m	Maxent	Single	RST	9	12	7	3
Galapagos (Santa Cruz Island)	2017	2	VP	IN (Local)	1km	Maxent	Single	SYN	1528	1621	1435	5
Galapagos archipelago	2009	1	ART	IN (Widespread)	1km	Maxent	Single	IAS	4	4	4	19
Global Islands	2019	55	VP	EN (Widespread)	1km	Maxent	Single	CC	40	140	5	4
Grande Comore	2017	2	MM	EN (Local)	25m	GLM	Single	SHR	42	53	33	2

Grande Comore	2019	1	BRD	EN (Local)	200m	DSM	Single	SHR	254	254	254	-
Grande Comore/Madagascar	2021	1	BRD	NA (Local)	1km	ENFA/GLM	Single	SHR	194	194	194	6
Great Britain	2013	8	MM	NA (Local)	50m & 100m	Maxent	Single	MT	254	769	44	13
Great Britain	2013	28	ART	NA (Local)	1km	Maxent	Single		867	4254	12	19
Great Britain	2019	6	MM	NA (Local)	1km*	BAYES	Single	MT	1411	4690	114	-
Great Britain	2012	53 / 183 / 1587	ART/BRD/VP	NA (Widespread)	-	ANN / BIOCLIM / CTA / GAM / GBM / GLM / MARS / Maxent / RF	Ensemble	MT	-	-	-	6
Haida Gwaii islands	2017	1	BRD	NA (Local)	100m	GBM	Single	SHR	166	166	166	8
Hawaii archipelago	2002	1	PRT	EN (Local)	30m	Net Diference	Single	SYN	-	-	-	2
Hawaii archipelago	2021	8	ART	IN (Widespread)	1km	ANN / CTA / FDA / GAM / GARP / GBM / GLM / MARS / Maxent / RF	Ensemble	SYN	5,709	50,640	28	7
Hawaii archipelago	2017	2	BRD	EN (Local)	500m	GBM / Maxent	Ensemble	RST	175	239	111	4
Hawaii archipelago	2017	1	BRD	IN (Widespread)	10km	Maxent	Single	IAS	2,329	2,329	2,329	21
Hawaii archipelago	2015	20	BRD	EN (Local)	1km/3km	GBM / Maxent	Ensemble	CC	-	-	-	4

Hispaniola island	2019	1	MM	EN (Local)	1km	Maxent	Single	LUC	233	447	19	8
Hispaniola island	2013	26	RPT	NA (Local)	1km	MARS	Single	SHR	-	-	-	-
Hispaniola island (Dominican republic)	2019	2	MM	EN (Local)	30m	Maxent	Single	SHR	91	135	48	5
Iceland	2018	1	VP	IN (Local)	1km	Maxent	Single	SYN	98	98	98	11
Indonesia Archipelago	2021	51	MM	NA (Widespread)	1km	Maxent	Single	PAM	48	417	10	6
Indonesia Archipelago	2019	1	VP	NA (Widespread)	-	BIOCLIM / DOMAIN / GLM / Maxent / RF / SVM	Ensemble	MT	98	98	98	8
Indonesia Archipelago	2017	2	VP	NA (Local)	1km	Maxent	Single	CC	563	935	192	7
Indonesia Archipelago	2020	1	VP	IN (Local)	1km	Maxent / RF	Single	SYN	507	507	507	13
Indonesia Archipelago	2017	1	VP	IN (Widespread)	4.5km	GLM	Single	SYN	232	232	232	7
Indonesia Archipelago	2021	1	VP	IN (Local)	-	Maxent	Single	SYN	5,389	5,389	5,389	19
Indonesia Archipelago	2013	1720	VP	NA (Widespread)	10km	Maxent	Single	RP	-	-	-	16
Indonesia Archipelago	2017	348	VP	NA (Widespread)	10km	Maxent	Single	RP	23	112	5	12
Indonesia Archipelago	2021	139	VP	NA (Local)	10km	Maxent	Single	CC	33	152	10	19

Indonesia Archipelago	2020	1	MM	EN (Local)	30m	Maxent	Single	SHR	36	36	36	8
Indonesia Archipelago	2018	1	BRD	EN (Local)	1km	Maxent	Single	SHR	48	48	48	16
Indonesia Archipelago	2013	1	BRD	EN (Local)	1km	Maxent	Single	SHR	60	60	60	21
Indonesia Archipelago	2019	1	MM	NA (Local)	15m	Maxent	Single	SHR	70	70	70	11
Indonesia Archipelago	2021	1	RPT	EN (Local)	NE	DB	Single	LUC	87	87	87	11
Indonesia Archipelago	2020	2	VP	NA (Local)	1km	Maxent	Single	CC	153	156	151	12
Indonesia Archipelago	2020	1	RPT	EN (Local)	1km	GAM / GBM / GLM / Maxent	Ensemble	CC	4,028	4,028	4,028	3
Indonesia Archipelago	2014	317	RPT	NA (Widespread)	10km	Maxent	Single	RP	-	-	-	8
Indonesia Archipelago	2020	4	VP	NA (Local)	1km	Maxent	Single	CC	-	-	-	12
Indonesia Archipelago	2020	1	VP	IN (Local)	5km	Maxent	Single	CC	-	-	-	22
Ireland	2014	14 / 16	BRY / VP	NA (Widespread)	10km	ANN / FDA / GAM / GBM / GLM / RF	Ensemble	PAM	-	-	-	-
Japan Archipelago	2015	1	ART	IN (Local)	1km	Maxent	Single	IAS	12	12	12	10
Japan Archipelago	2013	1	VP	EN (Local)	1km &4.5Km (past)	CTA / GAM	Ensemble	CC	339	1210	12	4
Japan Archipelago	2021	10	ART	EN (Local)	4.5km	Maxent	Single	CC	44	156	12	9

Japan Archipelago	2013	1	RPT	IN (Local)	1km	Maxent	Single	IAS	13	13	13	8
Japan Archipelago	2014	2	VP	NA-1/EN-1	1km &4.5Km (past)	CTA / GAM / GLM / GBM / MARS / Maxent / RF	Ensemble	PHP	224	436	13	6
Japan Archipelago	2010	25	VP	NA (Local)	1km	ANN / GAM / GBM / RF	Ensemble	CC	207	711	17	-
Japan Archipelago	2016	1	VP	EN (Local)	1km	GLM	Single	CC	1091	1983	200	-
Japan Archipelago	2017	1	RPT	IN (Widespread)	10km	Maxent	Single	IAS	824	824	824	21
Japan Archipelago	2014	1	VP	EN (Local)	4.5km	GAM	Single	PHP	761	761	761	4
Japan Archipelago	2019	1	VP	NA (Local)	4.5km	GAM	Single	PHP	3937	3937	3937	4
Japan Archipelago	2015	1	VP	EN (Local)	4.5km	RF	Single	PHP	4387	4387	4387	4
Japan Archipelago	2019	1	VP	NA (Local)	1km	CTA	Single	CC	9132	9,132	9132	-
Japan Archipelago	2011	1	VP	NA (Local)	1km	CTA	Single	CC	15,809	15,809	15,809	-
Japan Archipelago	2003	28	VP	NA (Local)	5m	GLM	Single	SHR	-	-	-	2
Japan Archipelago	2014	5	MM	NA-3/EN-2(Local)	100m	GLM	Single	LUC	-	-	-	7
Japan Archipelago	2011	1	VP	EN (Local)	50m	CTA	Single	CC	-	-	-	-
Japan Archipelago	2020	1	BRD	NA (Widespread)	4.5km	Maxent	Single	PHP	-	-	-	7
Jeju Island	2019	1	LCH	NA (Local)	1km	Maxent	Single	PHP	12	12	12	3

Jeju Island	2017	1	VP	NA (Widespread)	30m	GLM / MARS / RF	Ensemble	CC	420	420	420	5
Lampi island (Myanmar)	2019	19	MM	NA (Local)	1km	Maxent	Single	LUC	91	541	9	9
Lord Howe Island	2021	1	BRD	EN (Local)	10m	Maxent	Single	SHR	73	73	73	6
Lord Howe Island	2013	4	VP	NA (Local)	10m	Maxent	Single	PHP	288	150	404	8
Macaronesia (Azores, Canarias, Cabo verde)	2015	12	BRY	NA (Widespread)	5km	GBM / GLM / Maxent / RF	Ensemble	PHP	65	95	38	5
Macquarie Island	2019	1	VP	EN (Local)	5m	GLM	Single	MT	90	90	90	6
Macquarie Island	2015	2	VP	IN (Widespread)	1km	Maxent	Single	IAS	40122	80000	245	15
Macquarie Island	1999	1 / 9 / .2	AN/ART/MLC	NA (Local)	-	GLM	Single	SYN	-	-	-	-
Madagascar	2015	325. / 420	AN/RPT	EN (Local)	1km	Maxent	Single	SHR	8	100	3	24
Madagascar	2015	57	MM	EN (Local)	1km	ANN / CTA / FDA / GAM / GBM / GLM / MARS / Maxent / RF / SER	Ensemble	CC	64	988	6	-
Madagascar	2021	9	BRD	EN (Local)	1km	Maxent	Single	SHR	13.44	33	6	15
Madagascar	2018	19 / .11	MM/VP	EN-19/NA-11(Local)	1km	Maxent	Single	SHR	39	72	8	22
Madagascar	2021	6	VP	EN (Local)	4.5km	Maxent	Single	CC	41	108	9	7

Madagascar	2020	1	MM	EN (Local)	30m	Maxent	Single	SHR	12	12	12	10
Madagascar	2015	273	VP	EN (Local)	10km	GAM / GLM / MARS / Maxent / RF / SER	Ensemble	PHP	-	-	-	1 per 20 records
Madagascar	2017	1	AN	IN (Widespread)	1km	Maxent	Single	IAS	207	324	90	19
Madagascar	2021	1	AN	EN (Local)	250m	Maxent	Single	CC	98	98	98	8
Madagascar	2013	3	VP	EN (Local)	2km	GAM / GLM / Maxent	Ensemble	PAM	7549	15000	99	4
Madagascar	2017	1	BRD	NA (Widespread)	1km	Maxent	Single	SHR	5000	5000	5000	8
Madagascar/Mascarene archipelago (Reunion Island)	2021	1	RPT	EN (Local)	1km/ 15&30 m for landcover	ANN / CTA / FDA / GAM / GBM / GLM / MARS / Maxent / RF / SER	Ensemble	CC	57	101	14	4
Marquesas islands	2019	1	VP	IN (Widespread)	1km/10 m (Nuku Hiva/Fatu Hiva)	Maxent	Single	IAS	3000	3000	3000	9
Mascarene archipelago (Mauritius)	2014	5	RPT	IN-1/EN-4(Local)	900m	GARP / GLM / Maxent / RF	Ensemble	IAS	44650	89175	126	9
Mediterranean islands (Balearic	2019	1	ART	IN (Local)	1km	GBM / GLM / MARS / Maxent	Ensemble	IAS	20	20	20	2

islands, Malta, Sardinia)												
New Caledonia archipelago	2015	60	VP	NA (Local)	1km &4.5Km (past)	BIOCLIM / DOMAIN / SVM	Ensemble	RP	31	138	10	4
New Caledonia archipelago	2015	562	VP	NA (Local)	100m	Maxent	Single	RP	16	90	4	-
New Caledonia archipelago	2016	469	VP	NA (Local)	1km	Maxent	Single	CC	-	-	-	5
New Caledonia archipelago	2021	23	VP	EN (Local)	10m/ 30m	Maxent	Single	LUC	64	511	4	5
New Caledonia archipelago	2008	1	ART	EN (Local)	1km	Maxent	Single	RP	11	11	11	8
New Caledonia archipelago	2013	150	VP	EN (Local)	1km	Maxent	Single	SHR	297	516	86	6
New Caledonia archipelago	2017	1	VP	EN (Local)	4.5km	Maxent	Single	PHP	99	99	99	17_24
New Caledonia archipelago	2019	678	VP	EN-644/NA(33)	100m	ANN / GAM / GBM / GLM / MARS / Maxent / RF / SVM	Ensemble	RP	-	-	-	6
New guinea	2017	3000	VP	NA (Local)	1km	Maxent	Single	RP	-	-	-	10
New Guinea, the Bismarck Archipelago and surrounding islands.	2016	532	VP	NA (Local)	9.3km	Maxent	Single	RP	-	-	-	16

New guinea/Australia	2019	1	BRD	NA (Widespread)	4.5km	ANN / CTA / GLM / MARS /	Ensemble	PHP	18,092	18,092	18,092	12
New Zealand	2015	1	MM	IN (Local)	1km	Maxent	Single	SYN	827	5888	3	-
New Zealand	2013	3	VP	IN (Widespread)	4.5km	Maxent	Single	SYN	658	1741	81	6
New Zealand	2015	1	MM	IN (Local)	1km/200m	Maxent	Single	IAS	21	21	21	6
New Zealand	2015	1	BRD	EN (Local)	-	Maxent	Single	RST	34	34	34	5
New Zealand	2012	3	VP	IN (Widespread)	4.5km	Maxent	Single	MT	658	1741	81	6
New Zealand	2018	3	ART	EN (Local)	5km	ANN / CTA / FDA / GAM / GBM / GLM / MARS / RF	Ensemble	PHP	127	174	84	5
New Zealand	2017	1	VP	IN (Widespread)	1km	GLM / Maxent / RF / SVM	Ensemble	SYN	106	106	106	5
New Zealand	2015	1	ART	IN (Widespread)	4.5km	Maxent	Single	SYN	-	-	-	5
New Zealand	2018	1	BRD	EN (Local)	90m	Maxent	Single	RST	313	603	120	-
New Zealand	2014	1	ART	EN (Local)	25m	Maxent	Single	CC	458	458	458	5
New Zealand	2015	97	VP	NA-33/IN-64	10m	SVM	Single	IAS	-	-	-	11
New Zealand	2019	1	VP	IN (Widespread)	1km	Maxent	Single	SYN	680	1232	127	17
New Zealand	2014	3	VP	IN (Widespread)	4.5km	Maxent	Single	SYN	-	-	-	7
New Zealand	2021	1	RPT	EN(Widespread)	4.5km	Maxent	Single	CC	-	-	-	4
North islands of Australia	2017	4	RPT	NA (Widespread)	4.5km	Maxent	Single	PHP	110	199	37	10

Panama isthmus islands	2012	1	AN	NA (Local)	1km	Maxent	Single	PHP	38	38	38	19
Pantropical islands	2020	1	RPT	IN (Widespread)	4.5km	BIOCLIM	Single	IAS	1106	1106	1106	8
Philipines	2021	26	VP	NA (Widespread)	1km	Maxent	Single	SYN	47	111	12	11
Philipines	2018	2	VP	NA-1/EN-1(Local)	4.5km	Maxent	Single	SHR	73	78	69	7
Philipines	2020	1	ART	NA (Local)	1km	Maxent	Single	IAS	16	19	13	19
Philipines	2019	1. / 1	BRD/RPT	EN (Local)	4.5km	Maxent	Single	MT	392	771	14	19
Philipines	2020	1	VP	NA (Local)	1km	GBM / MARS / Maxent / RF	Single	PHP	93	93	93	6
Philipines	2018	1	FG	NA (Local)	4.5km	Maxent	Single	CC	46	46	46	19
Philipines	2014	1	VP	NA (Local)	1km	DOMAIN	Single	SHR	557	557	557	19
Puerto rico	2019	200	VP	NA (Local)	450m	ANN / CTA / FDA / GBM / GLM / MARS / RF / SER	Ensemble	CC	57	340	30	15
Puerto rico	2021	2	VP	IN (Local)	1km	GLM	Single	SYN	66	140	28	5
Santa Catalina Island	2018	11	VP	NT (Local)	1.5m	Maxent	Single	RST	704	2418	10	7
Sao Tomé	2017	3	BRD	EN (Local)	1km	Maxent	Single	SHR	117	363	13	2_12
Sardinia Island	2017	62	VP	EN-26/NA (Local)	1km	GLM / RF	Ensemble	SYN	198	31	3	8
Sardinia Island	2017	4	VP	EN-3/NA-1(Local)	1km	Maxent	Single	PAM	91	305	5	5
Sardinia Island	2015	1	VP	NA (Local)	250m	Maxent	Single	MT	30	8	58	7

Sardinia Island	2013	2	RPT	IN-1/EN-1(Local)	100m	Maxent	Single	LUC	323	355	291	7
Sardinia Island	2020	1	VP	IN (Local)	10m	GLM	Single	IAS	432	432	432	-
Sardinia Island	2017	1	MM	NA (Local)	40 m	Maxent	Single	MT	471	471	471	8
Sicily Channel Islands	2019	1	MM	NA (Widespread)	1km	GAM / GBM / GLM / Maxent / RF	Ensemble	PHP	25	25	25	21
Sicily Channel Islands	2018	1	RPT	EN (Local)	1km	GAM / GBM / GLM / Maxent	Ensemble	PHP	90	90	90	9
Sidney Island (Canada)	2020	1	NAI	IN (Local)	2m	BAYES / GBM / Maxent / RF	Ensemble	IAS	300	300	300	-
Sjælland (Denmark)	2011	1	BRD	NA (Local)	5km	GLM	Single	SHR	703	703	703	7
Socotra Island	2018	20	RPT	EN-17/NA-3 (Local)	100m	Maxent	Single	PHP	33	97	7	8
Socotra Island	2019	2	RPT	EN (Local)	100m	Maxent	Single	PHP	34	40	29	9
Socotra Island	2012	4	VP	IN (Local)	100m	RF	Single	IAS	90	153	30	9
Sri Lanka	2018	2	VP	NA (Local)	4.5km*	Maxent	Single	PHP	72	107	38	19
Sri Lanka	2020	4	VP	NA-2/IN-2	1km	Maxent	Single	CC	101	272	10	7
Sri Lanka	2019	14	VP	IN (Local)	1km	Maxent	Single	SYN	103	69	15	7
Sri Lanka	2018	1	MM	EN (Local)	100m	Maxent	Single	SHR	492	492	492	-
Sri Lanka	2021	14	VP	IN (Local)	1km	Maxent	Single	SYN	-	-	-	7
Sri Lanka/Australia	2019	1	VP	IN (Local)	1km	BIOCLIM / DOMAIN / GLM / Maxent	Single	IAS	154	154	154	6
Sumatra	2021	1	VP	EN (Local)	1km	Maxent	Single	CC	34	34	34	23
Sumatra	2020	1	VP	IN (Local)	1km	Maxent	Single	SHR	63	63	63	23
Sumatra	2020	1	VP	NA (Local)	1km	Maxent	Single	SHR	77	77	77	12

Taiwan Island	2021	1	VP	NA (Local)	50m	RF	Single	MT	90	160	20	14
Taiwan Island	2013	237	VP	NA (Local)	1km	Maxent	Single	CC	544	1083	5	21
Taiwan Island	2015	7	ART	NAT(Local)	90m	Maxent	Single	MT	28	39	21	18
Taiwan Island	2021	6	ART	IN (Widespread)	1km	Maxent	Single	IAS	72	136	25	7
Taiwan Island	2017	2	VP	EN (Local)	1km	Maxent	Single	PHP	76	78	75	5
Taiwan Island	2014	2	VP	EN (Local)	3km	Maxent	Single	CC	-	-	-	4
Taiwan Island	2014	4	MM	EN (Local)	4.5km	Maxent	Single	PHP	267	652	108	8
Taiwan Island	2012	1	VP	NA (Local)	5m	DB / GARP / GLM / Maxent	Ensemble	SHR	221	221	221	-
Taiwan Island	2018	1	VP	NA (Local)	1km*	ANN / CTA / FDA / GAM / GBM / GLM / MARS / Maxent / RF / SER	Ensemble	CC	3032	3032	3032	19
Taiwan Island	2014	156	VP	NA (Local)	1km	Maxent	Single	SHR	-	-	-	10
Taiwan Island	2014	156	VP	NA (Local)	1km	Maxent	Single	RP	-	-	-	5
Taiwan Island	2011	237	VP	NA (Local)	1km	Maxent	Single	CC	85	1083	5	16
Taiwan Island	2018	156	VP	NA (Local)	1km	Maxent	Single	CC	-	-	-	19
Taiwan Island/Japanese Archipelago	2020	1	ART	NA (Local)	4.5km	Maxent	Single	PHP	87	130	39	6
Taiwan Island/Japanese Archipelago	2014	10	VP	NA (Widespread)	1km	GAM	Single	CC	29 567	29 567	29 567	4
Tasmania	2021	1	MM	IN (Widespread)	4.5km	GBM / GLM / RF	Ensemble	IAS	3975	4,726	3,225	-
Tasmania	2016	01./01/.2	BRD/VP	NA-3/IN-1	10km	ANN / CTA / FDA / GAM /	Ensemble	SYN	1874	3634	68	3_ 5

						GBM / GLM / MARS / Maxent / RF / SER						
Tasmania	2021	1	MM	EN (Local)	250m	Maxent	Single	CC	254	254	254	8
Tasmania	2020	1	MM	NA (Widespread)	1km	GAM / GBM / GLM / RF	Ensemble	RST	996	996	996	4
Tasmania	2015	1	MM	NA (Local)	5km	Maxent	Single	CC	1590	1590	1590	8
Tasmania/Australia	2016	3	VP	NA (Widespread)	4.5km	BIOCLIM	Single	SHR	1131	2454	87	3
Tasmania/Australia	2015	1	MM	NA (Widespread)	-	Maxent	Single	RST	1469	1469	1469	1
Tasmania/Australia	2021	1	VP	EN (Local)	1km	GAM / GBM / GLM / RF	Ensemble	CC	4114	4114	4114	7
Tasmania/Australia	2012	72	VP	IN (Widespread)	10km	Maxent	Single	IAS	-	-	-	7
Tiwi Islands	2019	7	MM	NA (Local)	25m	GBM / GLM	Single	MT	175	327	24	15
Trinidad and Tobago Islands	2018	27	VP	EN (Local)	1km	Maxent	Single	RP	12	37	5	9

Taxa: AN-Anurans, ART-arthropods, BRD- birds, FNG- fungi, MAM- Mammals, NVP- non-vascular plants, NAI-non-arthropod invertebrates, RPT- reptiles and VP- vascular plants

Niche truncation (chorotypes): EN-Endemic, NA- Non-endemic Native, IN- Non-Native

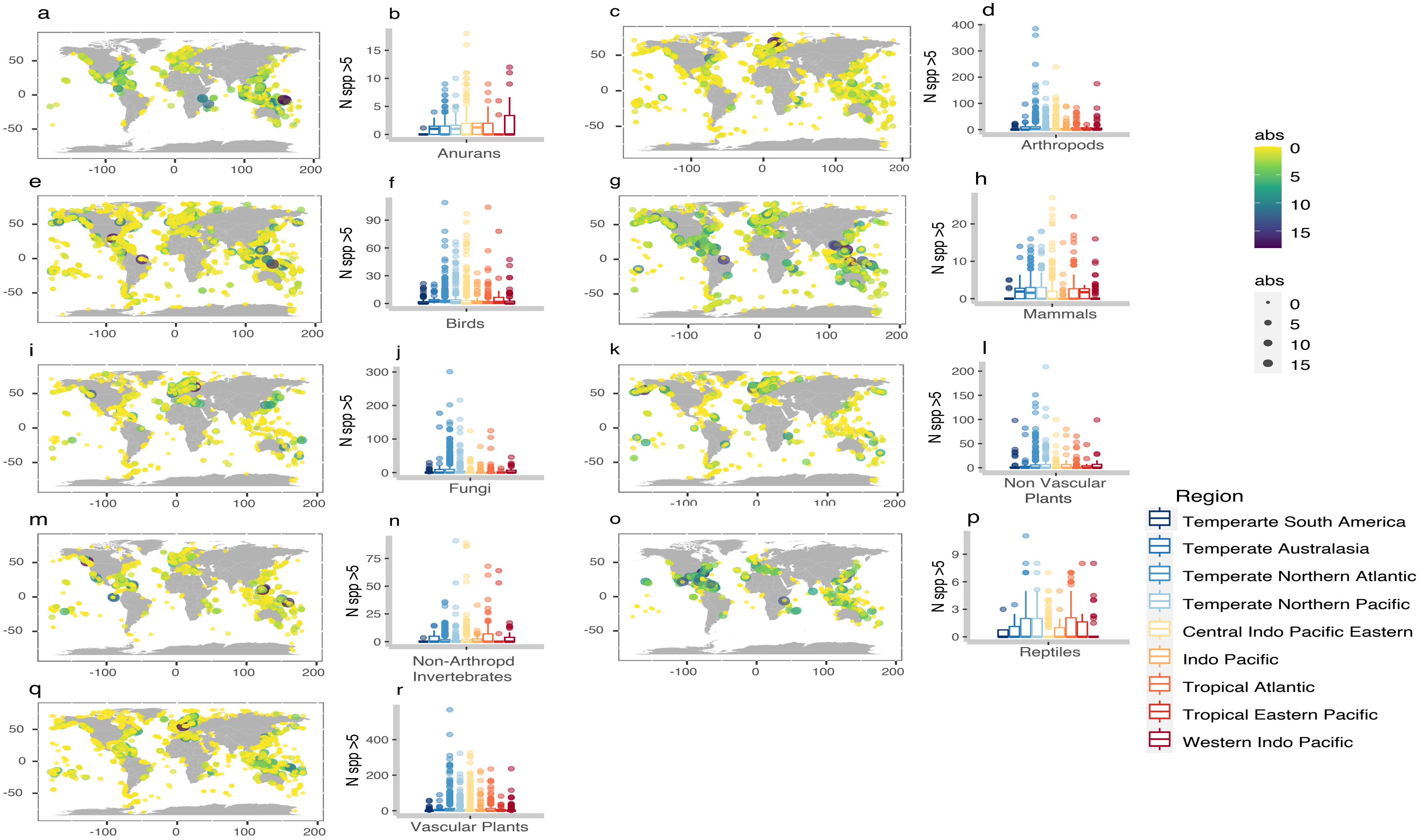
Study aims: CC-climate change, IAS-invasive alien species, LUC-land use change, SYN- synergy of extinction drivers, MT-Method testing, PAM-protected areas management, PHY- phylogeographic, RP- Richness patterns, RST-restoration, SHR- species habitat requirements.

Environmental data resolution: * are inferred resolutions based on predictors utilised

Modelling methods: ANN- Artificial Neural Networks, DB- Distance based, CTA - Classification tree analysis, FDA-Flexible Discriminant Analysis, GAM- Generalised additive model, GBM- Generalised Boosted Regression Model, GLM- Generalised Linear Models, MARS- multivariate adaptive regression splines, RF- Random forest, SVM- Support Vector Machines

Supplementary table S2.2 Summary of the number of records and modellability index (M index) per taxa for each island analysed, as obtained from the GBFI database.

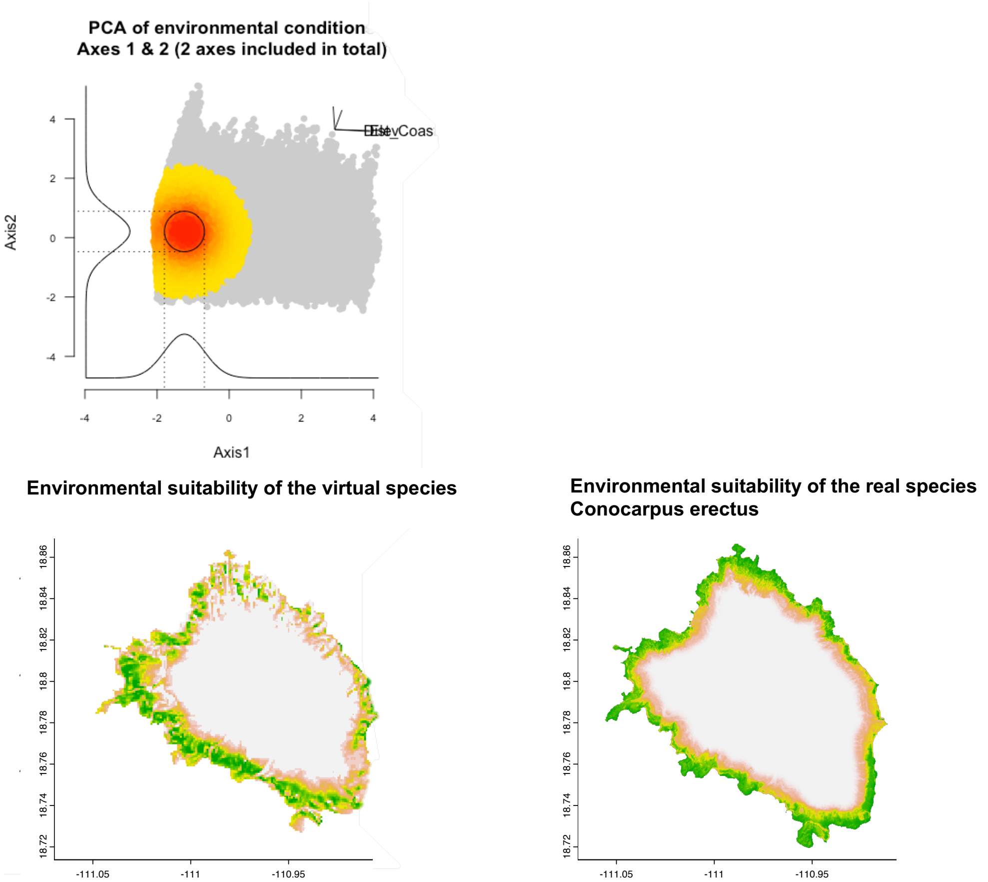
Taxa	**Total global records GBIF**	**Mean N Records per spp GBIF**	**N Islands with GBIF data**	**N islands with no modellable species**	**N islands M index [0.01-25] (%)**	**N islands M index [0.26-0.50] (%)**	**N islands M index [0.51-0.75] (%)**	**N islands M index [0.76-100] (%)**
Birds	6,069,146	33	2462	893 (36.3%)	444 (18%)	477 (19.4%)	531 (21.6%)	117 (4.7%)
Vascular	3,991,666	9	2275	1073 (47.2%)	796 (35%)	266 (11.6%)	128 (5.6%)	15 (0.6%)
Mammalia	185,402	15	1605	822 (51.2%)	253 (15.8%)	336 (21%)	122 (7.6%)	72 (4.4%)
Anura	42,402	9	948	469 (49.5%)	85 (9%)	150 (15.8%)	107 (11.3%)	137 (4.4%)
Reptilia	16,381	8	572	361 (63.1%)	36 (6.3%)	109 (19.1%)	32 (5.6%)	34 (5.9%)
Non-Arthropod invertebrates	56,079	4	1205	771 (64%)	283 (23.5%)	110 (9.1%)	27 (2.2%)	14 (1.2%)
Non-Vascular	748,916	7	1098	687 (62.6%)	311 (28.3%)	72 (6.5%)	23 (2.1%)	5 (0.5%)
Arthropoda	1,085,662	5	1872	1051 (56.1%)	697 (37.2%)	106 (5.8%)	12 (0.6%)	6 (0.3%)
Fungi	362,324	4	1214	768 (63.3%)	378 (31.1%)	62 (5.1%)	3 (0.25%)	3 (0.25%)



Supplementary figure S.2.2.3 Spatial distribution of the total number of species witn modelling potential (≥20 occurrence records, as derived from GBIF data) divided by taxa: a-b: Anurans, c-d; Arthropods, e-f; Birds, g-h; Mammals, i-j; Fungi, k-l; Non-Vascular plants, m-n; Non-Arthropod invertebrates, o-p; Reptiles, and q-r; Vascular Plants. Blue shades in boxplots indicate temperate environments, yellow to red shades indicate tropical environments.

Chapter 3

Appendix S3.1 Supplementary tables and figures



Supplementary figure S.3.1.1 Distribution of a plant species present in the Revillagigedo Archipelago (*Conocarpus erectus*, bottom right) used as a baseline for generating virtual species (bottom left) that closely mimic natural spatial distribution patterns using the PCA method (upper left) from the "virtualspecies" R package.

Supplementary Table S.3.1.2 Numerical results of the model: $\mathrm{D.index\_Narrow} \sim \mathrm{Psa/BG} \; * \; \mathrm{Sample\_size} \; * \; \mathrm{Method}$ (analysing model accuracy across various modelling conditions for narrow-ranged virtual species).

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.602	0.011	53.095	0.000
Psa/BG _per10	0.006	0.016	0.402	0.687
Psa/BG _per20	0.012	0.016	0.770	0.441
Psa/BG _per40	0.020	0.016	1.259	0.208
Sample_size10	0.018	0.016	1.120	0.263
Sample_size20	0.033	0.016	2.009	0.045
Sample_size50	0.058	0.017	3.376	0.001
MethodGAM	0.046	0.016	2.873	0.004
MethodGLM	0.032	0.016	2.002	0.045
MethodMXT	0.016	0.016	0.976	0.329
Psa/BG _per10: Sample_size10	-0.001	0.023	-0.029	0.977
Psa/BG _per20:Sample_size10	-0.007	0.023	-0.313	0.754
Psa/BG _per40: Sample_size10	-0.014	0.023	-0.623	0.533
Psa/BG _per10: Sample_size20	0.000	0.023	0.008	0.994
Psa/BG _per20: Sample_size20	-0.006	0.023	-0.244	0.807
Psa/BG _per40: Sample_size20	-0.026	0.023	-1.107	0.268
Psa/BG _per10: Sample_size50	0.006	0.024	0.257	0.797
Psa/BG _per20: Sample_size50	0.001	0.024	0.023	0.982
Psa/BG _per40: Sample_size50	-0.015	0.024	-0.609	0.543
Psa/BG _per10: MethodGAM	-0.008	0.023	-0.361	0.718
Psa/BG _per20: MethodGAM	-0.014	0.023	-0.628	0.530
Psa/BG _per40: MethodGAM	-0.030	0.023	-1.320	0.187
Psa/BG _per10: MethodGLM	0.015	0.023	0.656	0.512
Psa/BG _per20: MethodGLM	0.030	0.023	1.329	0.184
Psa/BG _per40: MethodGLM	0.038	0.023	1.678	0.093
Psa/BG _per10: MethodMXT	0.017	0.023	0.760	0.447
Psa/BG _per20: MethodMXT	0.035	0.023	1.532	0.126
Psa/BG _per40: MethodMXT	0.044	0.023	1.938	0.053
Sample_size10: MethodGAM	0.016	0.023	0.694	0.488
Sample_size20: MethodGAM	0.018	0.023	0.778	0.437
Sample_size50: MethodGAM	-0.002	0.024	-0.082	0.935
Sample_size10: MethodGLM	-0.039	0.023	-1.741	0.082
Sample_size20: MethodGLM	-0.073	0.023	-3.196	0.001
Sample_size50: MethodGLM	-0.127	0.024	-5.282	1.34e-07
Sample_size10: MethodMXT	-0.035	0.023	-1.541	0.123
Sample_size20: MethodMXT	0.036	0.023	1.559	0.119

Sample_size50: MethodMXT	0.015	0.024	0.623	0.533
Psa/BG10: Sample_size10: MethodGAM	0.015	0.032	0.455	0.649
Psa/BG20: Sample_size10: MethodGAM	0.025	0.032	0.774	0.439
Psa/BG40: Sample_size10: MethodGAM	0.041	0.032	1.279	0.201
Psa/BG10: Sample_size20: MethodGAM	0.009	0.032	0.264	0.792
Psa/BG20: Sample_size20: MethodGAM	0.021	0.032	0.662	0.508
Psa/BG40: Sample_size20: MethodGAM	0.048	0.033	1.483	0.138
Psa/BG10: Sample_size50: MethodGAM	0.015	0.034	0.427	0.670
Psa/BG20: Sample_size50: MethodGAM	0.029	0.034	0.858	0.391
Psa/BG40: Sample_size50: MethodGAM	0.056	0.034	1.635	0.102
Psa/BG10: Sample_size10: MethodGLM	0.004	0.032	0.114	0.909
Psa/BG20: Sample_size10: MethodGLM	0.009	0.032	0.287	0.774
Psa/BG40: Sample_size10: MethodGLM	0.022	0.032	0.685	0.494
Psa/BG10: Sample_size20: MethodGLM	0.002	0.032	0.061	0.952
Psa/BG20: Sample_size20: MethodGLM	0.010	0.032	0.309	0.758
Psa/BG40: Sample_size20: MethodGLM	0.023	0.033	0.704	0.482
Psa/BG10: Sample_size50: MethodGLM	0.002	0.034	0.062	0.951
Psa/BG20: Sample_size50: MethodGLM	0.012	0.034	0.349	0.727
Psa/BG40: Sample_size50: MethodGLM	0.034	0.034	1.010	0.313
Psa/BG10: Sample_size10: MethodMXT	0.001	0.032	0.040	0.968
Psa/BG20: Sample_size10: MethodMXT	0.005	0.032	0.146	0.884
Psa/BG40: Sample_size10: MethodMXT	0.013	0.032	0.393	0.694
Psa/BG10: Sample_size20: MethodMXT	-0.012	0.032	-0.367	0.714
Psa/BG20: Sample_size20: MethodMXT	-0.019	0.032	-0.596	0.551
Psa/BG40: Sample_size20: MethodMXT	-0.013	0.033	-0.394	0.693
Psa/BG10: Sample_size50: MethodMXT	-0.015	0.034	-0.440	0.660
Psa/BG20: Sample_size50: MethodMXT	-0.026	0.034	-0.748	0.454
Psa/BG40: Sample_size50: MethodMXT	-0.023	0.034	-0.666	0.505

Supplementary Table S.3.1.3 Numerical results of the model: $\mathrm{D.index\_Wide} \sim \mathrm{Psa/BG} * \mathrm{Sample\_size} * \mathrm{Method}$ (analysing model accuracy across various modelling conditions for wide-ranged virtual species).

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.847	0.008	105.028	0.000
Psa/BG_per10	0.004	0.011	0.375	0.707
BG_per20	0.018	0.011	1.557	0.119
BG_per40	0.042	0.011	3.719	0.000
Sample_size10	0.007	0.011	0.576	0.565
Sample_size20	0.015	0.011	1.311	0.190
Sample_size50	0.025	0.012	2.130	0.033
MethodGAM	-0.149	0.012	-12.890	0.000
MethodGLM	0.015	0.011	1.328	0.184
MethodMXT	0.036	0.012	3.080	0.002
Psa/BG_per10:Sample_size10	-0.003	0.016	-0.158	0.874
Psa/BG_per20:Sample_size10	-0.014	0.016	-0.847	0.397
Psa/BG_per40:Sample_size10	-0.022	0.016	-1.377	0.169
Psa/BG_per10:Sample_size20	-0.010	0.016	-0.597	0.550
Psa/BG_per20:Sample_size20	-0.022	0.016	-1.372	0.170
Psa/BG_per40:Sample_size20	-0.038	0.016	-2.363	0.018
Psa/BG_per10:Sample_size50	-0.005	0.016	-0.309	0.757
Psa/BG_per20:Sample_size50	-0.024	0.016	-1.460	0.144
Psa/BG_per40:Sample_size50	-0.052	0.016	-3.169	0.002
Psa/BG_per10:MethodGAM	-0.010	0.016	-0.644	0.520
Psa/BG_per20:MethodGAM	-0.036	0.016	-2.232	0.026
Psa/BG_per40:MethodGAM	-0.078	0.016	-4.783	0.000
Psa/BG_per10:MethodGLM	-0.003	0.016	-0.178	0.859
Psa/BG_per20:MethodGLM	-0.020	0.016	-1.211	0.226
Psa/BG_per40:MethodGLM	-0.049	0.016	-3.061	0.002
Psa/BG_per10:MethodMXT	-0.008	0.017	-0.463	0.643
Psa/BG_per20:MethodMXT	-0.026	0.017	-1.556	0.120
Psa/BG_per40:MethodMXT	-0.052	0.017	-3.108	0.002
Sample_size10:MethodGAM	0.090	0.016	5.497	0.000
Sample_size20:MethodGAM	0.127	0.016	7.852	0.000
Sample_size50:MethodGAM	0.150	0.016	9.170	0.000
Sample_size10:MethodGLM	0.010	0.016	0.607	0.544
Sample_size20:MethodGLM	0.004	0.016	0.239	0.811
Sample_size50:MethodGLM	-0.014	0.016	-0.833	0.405
Sample_size10:MethodMXT	0.009	0.017	0.549	0.583
Sample_size20:MethodMXT	-0.036	0.016	-2.201	0.028
Sample_size50:MethodMXT	-0.039	0.017	-2.333	0.020

Psa/BG_per10:Sample_size10:MethodGAM	0.007	0.023	0.313	0.754
Psa/BG_per20:Sample_size10:MethodGAM	0.019	0.023	0.821	0.412
Psa/BG_per40:Sample_size10:MethodGAM	0.038	0.023	1.661	0.097
Psa/BG_per10:Sample_size20:MethodGAM	0.011	0.023	0.491	0.624
Psa/BG_per20:Sample_size20:MethodGAM	0.030	0.023	1.332	0.183
Psa/BG_per40:Sample_size20:MethodGAM	0.061	0.023	2.667	0.008
Psa/BG_per10:Sample_size50:MethodGAM	0.011	0.023	0.466	0.641
Psa/BG_per20:Sample_size50:MethodGAM	0.040	0.023	1.733	0.083
Psa/BG_per40:Sample_size50:MethodGAM	0.079	0.023	3.428	0.001
Psa/BG_per10:Sample_size10:MethodGLM	0.006	0.023	0.281	0.779
Psa/BG_per20:Sample_size10:MethodGLM	0.022	0.023	0.949	0.343
Psa/BG_per40:Sample_size10:MethodGLM	0.032	0.023	1.409	0.159
Psa/BG_per10:Sample_size20:MethodGLM	0.015	0.023	0.662	0.508
Psa/BG_per20:Sample_size20:MethodGLM	0.032	0.023	1.417	0.156
Psa/BG_per40:Sample_size20:MethodGLM	0.064	0.023	2.812	0.005
Psa/BG_per10:Sample_size50:MethodGLM	0.016	0.023	0.697	0.486
Psa/BG_per20:Sample_size50:MethodGLM	0.045	0.023	1.955	0.051
Psa/BG_per40:Sample_size50:MethodGLM	0.085	0.023	3.703	0.000
Psa/BG_per10:Sample_size10:MethodMXT	0.009	0.024	0.366	0.715
Psa/BG_per20:Sample_size10:MethodMXT	0.030	0.023	1.284	0.199
Psa/BG_per40:Sample_size10:MethodMXT	0.032	0.023	1.386	0.166
Psa/BG_per10:Sample_size20:MethodMXT	0.013	0.023	0.573	0.567
Psa/BG_per20:Sample_size20:MethodMXT	0.023	0.023	0.998	0.318
Psa/BG_per40:Sample_size20:MethodMXT	0.045	0.023	1.908	0.056
Psa/BG_per10:Sample_size50:MethodMXT	0.009	0.023	0.383	0.701
Psa/BG_per20:Sample_size50:MethodMXT	0.029	0.023	1.221	0.222
Psa/BG_per40:Sample_size50:MethodMXT	0.052	0.023	2.224	0.026

Supplementary Table S.3.1.4 Numerical results of the model:
$\mathrm{D.index\_Narrow} \sim \mathrm{Psa/BG\_per} * \mathrm{Sample\_size} * \mathrm{Method}$
(analysing model accuracy across various modelling conditions for narrow-ranged real species).

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.90	0.02	41.37	0.00
Psa/BG_per10	0.04	0.03	1.36	0.18
Psa/BG_per20	0.06	0.03	1.80	0.08
Psa/BG_per40	0.05	0.03	1.59	0.12
MethodGAM	-0.03	0.03	-0.85	0.40
MethodGLM	0.00	0.03	-0.07	0.95
MethodMXT	0.07	0.03	2.18	0.03
Psa/BG_per10:MethodGAM	0.04	0.04	1.01	0.32
Psa/BG_per20:MethodGAM	-0.05	0.04	-1.10	0.28
Psa/BG_per40:MethodGAM	-0.04	0.04	-0.95	0.34
Psa/BG_per10:MethodGLM	0.03	0.04	0.70	0.49
Psa/BG_per20:MethodGLM	0.00	0.04	-0.07	0.94
Psa/BG_per40:MethodGLM	-0.04	0.04	-0.82	0.41
Psa/BG_per10:MethodMXT	-0.05	0.04	-1.05	0.30
Psa/BG_per20:MethodMXT	-0.07	0.04	-1.58	0.12
Psa/BG_per40:MethodMXT	-0.05	0.04	-1.26	0.21

Supplementary Table S.3.1.5 Numerical results of the model:
$\mathrm{D.index\_Wide} \sim \mathrm{Psa/BG\_per} * \mathrm{Sample\_size} * \mathrm{Method}$
(analysing model accuracy across various modelling conditions for wide-ranged real species).

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.94	0.01	107.45	0.00
Psa/BG_per10	0.00	0.01	-0.25	0.80
Psa/BG_per20	0.03	0.01	2.77	0.01
Psa/BG_per40	0.00	0.01	0.31	0.75
MethodGAM	0.03	0.01	2.10	0.04
MethodGLM	0.02	0.01	1.28	0.20
MethodMXT	0.02	0.01	1.23	0.22
Psa/BG_per10:MethodGAM	0.00	0.02	0.15	0.88
Psa/BG_per20:MethodGAM	-0.04	0.02	-2.36	0.02
Psa/BG_per40:MethodGAM	0.00	0.02	0.19	0.85
Psa/BG_per10:MethodGLM	0.01	0.02	0.34	0.73
Psa/BG_per20:MethodGLM	-0.04	0.02	-2.42	0.02
Psa/BG_per40:MethodGLM	0.00	0.02	-0.21	0.83
Psa/BG_per10:MethodMXT	0.01	0.02	0.52	0.61

Psa/BG_per20:MethodMXT	-0.05	0.02	-2.60	0.01
Psa/BG_per40:MethodMXT	0.00	0.02	0.13	0.90

Supplementary Table S.3.1.6 Stratified Pearson correlation between model accuracy (D-index) and AUC validation metric for Virtual Species SDMs

Range Size	Method	Sample size	PsA/BG percentage	Correlation
Narrow	BART	5n	5	-0.09
Narrow	BART	5n	10	-0.33
Narrow	BART	5n	20	-0.40
Narrow	BART	5n	40	-0.54
Wide	BART	5n	5	-0.59
Wide	BART	5n	10	-0.41
Wide	BART	5n	20	-0.36
Wide	BART	5n	40	-0.42
Narrow	BART	10n	5	0.16
Narrow	BART	10n	10	0.04
Narrow	BART	10n	20	-0.22
Narrow	BART	10n	40	-0.50
Wide	BART	10n	5	-0.57
Wide	BART	10n	10	-0.58
Wide	BART	10n	20	-0.63
Wide	BART	10n	40	-0.70
Narrow	BART	20n	5	0.41
Narrow	BART	20n	10	0.28
Narrow	BART	20n	20	0.16
Narrow	BART	20n	40	-0.13
Wide	BART	20n	5	-0.80
Wide	BART	20n	10	-0.63
Wide	BART	20n	20	-0.72
Wide	BART	20n	40	-0.85
Narrow	BART	50n	5	0.48
Narrow	BART	50n	10	0.42
Narrow	BART	50n	20	0.31
Narrow	BART	50n	40	0.03
Wide	BART	50n	5	-0.83
Wide	BART	50n	10	-0.83
Wide	BART	50n	20	-0.86
Wide	BART	50n	40	-0.96
Narrow	GAM	5n	5	0.24
Narrow	GAM	5n	10	0.18
Narrow	GAM	5n	20	0.12
Narrow	GAM	5n	40	-0.34
Wide	GAM	5n	5	-0.36
Wide	GAM	5n	10	-0.33

Wide	GAM	5n	20	-0.27
Wide	GAM	5n	40	-0.40
Narrow	GAM	10n	5	0.60
Narrow	GAM	10n	10	0.40
Narrow	GAM	10n	20	0.37
Narrow	GAM	10n	40	0.14
Wide	GAM	10n	5	-0.27
Wide	GAM	10n	10	-0.31
Wide	GAM	10n	20	-0.36
Wide	GAM	10n	40	-0.68
Narrow	GAM	20n	5	0.76
Narrow	GAM	20n	10	0.73
Narrow	GAM	20n	20	0.78
Narrow	GAM	20n	40	0.73
Wide	GAM	20n	5	-0.66
Wide	GAM	20n	10	-0.64
Wide	GAM	20n	20	-0.70
Wide	GAM	20n	40	-0.75
Narrow	GAM	50n	5	0.62
Narrow	GAM	50n	10	0.70
Narrow	GAM	50n	20	0.68
Narrow	GAM	50n	40	0.69
Wide	GAM	50n	5	-0.78
Wide	GAM	50n	10	-0.86
Wide	GAM	50n	20	-0.88
Wide	GAM	50n	40	-0.94
Narrow	GLM	5n	5	0.33
Narrow	GLM	5n	10	0.33
Narrow	GLM	5n	20	0.30
Narrow	GLM	5n	40	0.28
Wide	GLM	5n	5	-0.48
Wide	GLM	5n	10	-0.56
Wide	GLM	5n	20	-0.30
Wide	GLM	5n	40	-0.47
Narrow	GLM	10n	5	0.36
Narrow	GLM	10n	10	0.48
Narrow	GLM	10n	20	0.46
Narrow	GLM	10n	40	0.46
Wide	GLM	10n	5	-0.66
Wide	GLM	10n	10	-0.68
Wide	GLM	10n	20	-0.72
Wide	GLM	10n	40	-0.79
Narrow	GLM	20n	5	0.39
Narrow	GLM	20n	10	0.37

Narrow	GLM	20n	20	0.44
Narrow	GLM	20n	40	0.48
Wide	GLM	20n	5	-0.81
Wide	GLM	20n	10	-0.77
Wide	GLM	20n	20	-0.81
Wide	GLM	20n	40	-0.94
Narrow	GLM	50n	5	0.34
Narrow	GLM	50n	10	0.39
Narrow	GLM	50n	20	0.36
Narrow	GLM	50n	40	0.38
Wide	GLM	50n	5	-0.87
Wide	GLM	50n	10	-0.90
Wide	GLM	50n	20	-0.94
Wide	GLM	50n	40	-0.98
Narrow	MXT	5n	5	0.33
Narrow	MXT	5n	10	0.32
Narrow	MXT	5n	20	0.35
Narrow	MXT	5n	40	0.32
Wide	MXT	5n	5	-0.53
Wide	MXT	5n	10	-0.68
Wide	MXT	5n	20	-0.39
Wide	MXT	5n	40	-0.64
Narrow	MXT	10n	5	0.40
Narrow	MXT	10n	10	0.50
Narrow	MXT	10n	20	0.49
Narrow	MXT	10n	40	0.50
Wide	MXT	10n	5	-0.62
Wide	MXT	10n	10	-0.66
Wide	MXT	10n	20	-0.62
Wide	MXT	10n	40	-0.85
Narrow	MXT	20n	5	0.69
Narrow	MXT	20n	10	0.72
Narrow	MXT	20n	20	0.73
Narrow	MXT	20n	40	0.76
Wide	MXT	20n	5	-0.55
Wide	MXT	20n	10	-0.66
Wide	MXT	20n	20	-0.64
Wide	MXT	20n	40	-0.83
Narrow	MXT	50n	5	0.63
Narrow	MXT	50n	10	0.61
Narrow	MXT	50n	20	0.54
Narrow	MXT	50n	40	0.58
Wide	MXT	50n	5	-0.74
Wide	MXT	50n	10	-0.89

Wide	MXT	50n	20	-0.89
Wide	MXT	50n	40	-0.96

Supplementary Table S.3.1.7 Stratified Pearson correlation between model accuracy (D-index) and Boyce validation metric for Virtual Species SDMs

Range Size	Method	Sample size	PsA/BG percentage	Correlation
Narrow	BART	5n	5	0.54
Narrow	BART	5n	10	0.49
Narrow	BART	5n	20	0.51
Narrow	BART	5n	40	0.03
Wide	BART	5n	5	-0.34
Wide	BART	5n	10	0.00
Wide	BART	5n	20	0.22
Wide	BART	5n	40	0.01
Narrow	BART	10n	5	0.63
Narrow	BART	10n	10	0.53
Narrow	BART	10n	20	0.40
Narrow	BART	10n	40	-0.09
Wide	BART	10n	5	-0.43
Wide	BART	10n	10	-0.27
Wide	BART	10n	20	-0.01
Wide	BART	10n	40	0.10
Narrow	BART	20n	5	0.76
Narrow	BART	20n	10	0.69
Narrow	BART	20n	20	0.52
Narrow	BART	20n	40	-0.07
Wide	BART	20n	5	-0.71
Wide	BART	20n	10	-0.54
Wide	BART	20n	20	-0.31
Wide	BART	20n	40	0.11
Narrow	BART	50n	5	0.82
Narrow	BART	50n	10	0.81
Narrow	BART	50n	20	0.75
Narrow	BART	50n	40	-0.47
Wide	BART	50n	5	-0.69
Wide	BART	50n	10	-0.70
Wide	BART	50n	20	-0.68
Wide	BART	50n	40	-0.02
Narrow	GAM	5n	5	0.06
Narrow	GAM	5n	10	0.15
Narrow	GAM	5n	20	0.07

Narrow	GAM	5n	40	-0.23
Wide	GAM	5n	5	0.24
Wide	GAM	5n	10	0.33
Wide	GAM	5n	20	0.27
Wide	GAM	5n	40	0.03
Narrow	GAM	10n	5	0.76
Narrow	GAM	10n	10	0.59
Narrow	GAM	10n	20	0.47
Narrow	GAM	10n	40	-0.01
Wide	GAM	10n	5	0.19
Wide	GAM	10n	10	-0.03
Wide	GAM	10n	20	0.04
Wide	GAM	10n	40	-0.22
Narrow	GAM	20n	5	0.90
Narrow	GAM	20n	10	0.86
Narrow	GAM	20n	20	0.78
Narrow	GAM	20n	40	0.37
Wide	GAM	20n	5	0.00
Wide	GAM	20n	10	-0.02
Wide	GAM	20n	20	0.03
Wide	GAM	20n	40	-0.11
Narrow	GAM	50n	5	0.89
Narrow	GAM	50n	10	0.92
Narrow	GAM	50n	20	0.91
Narrow	GAM	50n	40	0.37
Wide	GAM	50n	5	-0.20
Wide	GAM	50n	10	-0.13
Wide	GAM	50n	20	-0.13
Wide	GAM	50n	40	0.07
Narrow	GLM	5n	5	0.53
Narrow	GLM	5n	10	0.51
Narrow	GLM	5n	20	0.46
Narrow	GLM	5n	40	0.40
Wide	GLM	5n	5	-0.13
Wide	GLM	5n	10	-0.23
Wide	GLM	5n	20	-0.15
Wide	GLM	5n	40	0.19
Narrow	GLM	10n	5	0.53
Narrow	GLM	10n	10	0.57
Narrow	GLM	10n	20	0.53

Narrow	GLM	10n	40	0.49
Wide	GLM	10n	5	-0.24
Wide	GLM	10n	10	-0.29
Wide	GLM	10n	20	-0.05
Wide	GLM	10n	40	0.66
Narrow	GLM	20n	5	0.53
Narrow	GLM	20n	10	0.52
Narrow	GLM	20n	20	0.51
Narrow	GLM	20n	40	0.69
Wide	GLM	20n	5	-0.37
Wide	GLM	20n	10	-0.33
Wide	GLM	20n	20	-0.27
Wide	GLM	20n	40	0.62
Narrow	GLM	50n	5	0.52
Narrow	GLM	50n	10	0.55
Narrow	GLM	50n	20	0.55
Narrow	GLM	50n	40	0.74
Wide	GLM	50n	5	-0.35
Wide	GLM	50n	10	-0.33
Wide	GLM	50n	20	-0.21
Wide	GLM	50n	40	0.79
Narrow	MXT	5n	5	0.01
Narrow	MXT	5n	10	-0.02
Narrow	MXT	5n	20	0.39
Narrow	MXT	5n	40	0.40
Wide	MXT	5n	5	0.45
Wide	MXT	5n	10	0.36
Wide	MXT	5n	20	0.16
Wide	MXT	5n	40	0.07
Narrow	MXT	10n	5	0.56
Narrow	MXT	10n	10	0.57
Narrow	MXT	10n	20	0.63
Narrow	MXT	10n	40	0.66
Wide	MXT	10n	5	-0.09
Wide	MXT	10n	10	-0.01
Wide	MXT	10n	20	-0.06
Wide	MXT	10n	40	-0.08
Narrow	MXT	20n	5	0.57
Narrow	MXT	20n	10	0.59
Narrow	MXT	20n	20	0.52

Narrow	MXT	20n	40	0.64
Wide	MXT	20n	5	-0.09
Wide	MXT	20n	10	-0.15
Wide	MXT	20n	20	-0.14
Wide	MXT	20n	40	-0.19
Narrow	MXT	50n	5	0.51
Narrow	MXT	50n	10	0.38
Narrow	MXT	50n	20	0.14
Narrow	MXT	50n	40	0.08
Wide	MXT	50n	5	-0.19
Wide	MXT	50n	10	-0.37
Wide	MXT	50n	20	-0.46
Wide	MXT	50n	40	-0.58

Supplementary Table S.3.1.8 Stratified Pearson correlation between model accuracy (D-index) and AUC validation metric for Real Species SDMs

Range Size	Method	PsA/BG percentage	Correlation
Narrow	BART	5	0.62
Narrow	BART	10	0.86
Narrow	BART	20	0.63
Narrow	BART	40	0.93
Wide	BART	5	0.74
Wide	BART	10	0.81
Wide	BART	20	0.95
Wide	BART	40	0.90
Narrow	GAM	5	0.72
Narrow	GAM	10	1.00
Narrow	GAM	20	1.00
Narrow	GAM	40	0.96
Wide	GAM	5	0.83
Wide	GAM	10	0.85
Wide	GAM	20	0.90
Wide	GAM	40	0.92
Narrow	GLM	5	0.71
Narrow	GLM	10	0.98
Narrow	GLM	20	0.66
Narrow	GLM	40	0.98
Wide	GLM	5	0.98
Wide	GLM	10	0.78
Wide	GLM	20	0.98
Wide	GLM	40	0.94
Narrow	MXT	5	0.79
Narrow	MXT	10	0.94
Narrow	MXT	20	0.7
Narrow	MXT	40	0.95
Wide	MXT	5	0.91
Wide	MXT	10	0.85
Wide	MXT	20	0.94
Wide	MXT	40	0.95

Supplementary Table S.3.1.9 Stratified Pearson correlation between model accuracy (D-index) and Boyce validation metric for Real Species SDMs

Range Size	Method	PsA/BG percentage	Correlation
Narrow	BART	5	0.60
Narrow	BART	10	0.78
Narrow	BART	20	0.69
Narrow	BART	40	0.86
Wide	BART	5	-0.87
Wide	BART	10	-0.21
Wide	BART	20	0.77
Wide	BART	40	0.29
Narrow	GAM	5	0.74
Narrow	GAM	10	0.96
Narrow	GAM	20	0.84
Narrow	GAM	40	0.86
Wide	GAM	5	0.61
Wide	GAM	10	0.68
Wide	GAM	20	0.83
Wide	GAM	40	0.96
Narrow	GLM	5	0.70
Narrow	GLM	10	0.89
Narrow	GLM	20	-0.52
Narrow	GLM	40	0.78
Wide	GLM	5	0.79
Wide	GLM	10	0.62
Wide	GLM	20	0.09
Wide	GLM	40	0.79
Narrow	MXT	5	0.22
Narrow	MXT	10	0.31
Narrow	MXT	20	0.27
Narrow	MXT	40	0.79
Wide	MXT	5	-0.58
Wide	MXT	10	-0.56
Wide	MXT	20	-0.69
Wide	MXT	40	-0.60

Supplementary Table S.3.1.10 Summary of 95th Percentile D-Index Ranges Across Methods by Niche Breadth (*Minimum Sample Sizes for Achieving Moderate >0.60 and High >0.70 Prediction overlap)

Niche Breadth	Method	Sample size	Mean	Y min	Y max	>0.60	>0.70
Narrow	BART	5n	0.59	0.57	0.72	-	-
Narrow	BART	10n	0.6	0.58	0.7	-	-
Narrow	BART	20n	0.61	0.6	0.76	*	-
Narrow	BART	50n	0.65	0.62	0.79	*	-
Narrow	GAM	5n	0.66	0.63	0.69	*	-
Narrow	GAM	10n	0.71	0.67	0.73	*	-
Narrow	GAM	20n	0.73	0.65	0.77	*	-
Narrow	GAM	50n	0.79	0.66	0.8	*	-
Narrow	GLM	5n	0.66	0.6	0.8	-	-
Narrow	GLM	10n	0.67	0.57	0.81	-	-
Narrow	GLM	20n	0.83	0.65	0.84	*	-
Narrow	GLM	50n	0.83	0.72	0.83	*	*
Narrow	MXT	5n	0.64	0.59	0.81	-	-
Narrow	MXT	10n	0.67	0.64	0.82	*	-
Narrow	MXT	20n	0.75	0.64	0.78	*	-
Narrow	MXT	50n	0.75	0.65	0.78	*	-
Wide	BART	5n	0.89	0.87	0.96	*	*
Wide	BART	10n	0.89	0.86	0.94	*	*
Wide	BART	20n	0.88	0.84	0.91	*	*
Wide	BART	50n	0.89	0.79	0.91	*	*
Wide	GAM	5n	0.78	0.77	0.78	*	*
Wide	GAM	10n	0.81	0.8	0.84	*	*
Wide	GAM	20n	0.86	0.8	0.88	*	*
Wide	GAM	50n	0.88	0.83	0.92	*	*
Wide	GLM	5n	0.87	0.82	0.91	*	*
Wide	GLM	10n	0.9	0.82	0.93	*	*
Wide	GLM	20n	0.92	0.89	0.94	*	*
Wide	GLM	50n	0.94	0.92	0.95	*	*
Wide	MXT	5n	0.9	0.79	0.94	*	*
Wide	MXT	10n	0.94	0.82	0.97	*	*
Wide	MXT	20n	0.89	0.83	0.91	*	*
Wide	MXT	50n	0.9	0.85	0.92	*	*

Supplementary table 3.1.11 Numerical results of the models:
$\mathrm{D.index\_Narrow} \sim \mathrm{Uncertainty} * \mathrm{Sample\_size} * \mathrm{Method}$
(analysing the the impact of varying levels of gridded uncertainty on model accuracy across different modelling conditions for narrow-ranged virtual species.

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.67	0.01	92.9	0
Uncert90	-0.04	0.01	-3.7	0
Uncert500	-0.05	0.01	-4.74	0
Uncert1000	-0.11	0.01	-10.98	0
Sample_sizeSmall	-0.04	0.01	-4.5	0
MethodGAM	0.06	0.01	5.5	0
MethodGLM	-0.06	0.01	-6.13	0
MethodMXT	0.04	0.01	3.83	0
Uncert90:Sample_sizeSmall	0.04	0.01	2.74	0.01
Uncert500:Sample_sizeSmall	0.04	0.01	2.78	0.01
Uncert1000:Sample_sizeSmall	0.05	0.01	3.68	0
Uncert90:MethodGAM	-0.01	0.01	-0.41	0.69
Uncert500:MethodGAM	0	0.01	-0.13	0.9
Uncert1000:MethodGAM	-0.01	0.01	-0.49	0.63
Uncert90:MethodGLM	-0.01	0.01	-0.44	0.66
Uncert500:MethodGLM	0.03	0.01	2.29	0.02
Uncert1000:MethodGLM	0.04	0.01	2.62	0.01
Uncert90:MethodMXT	0	0.01	-0.06	0.95
Uncert500:MethodMXT	0	0.01	-0.09	0.93
Uncert1000:MethodMXT	-0.01	0.01	-0.35	0.72
Sample_sizeSmall:MethodGAM	0.01	0.01	0.94	0.35
Sample_sizeSmall:MethodGLM	0.08	0.01	6.24	0
Sample_sizeSmall:MethodMXT	-0.03	0.01	-2.17	0.03
Uncert90:Sample_sizeSmall:MethodGAM	0.01	0.02	0.63	0.53
Uncert500:Sample_sizeSmall:MethodGAM	0.03	0.02	1.42	0.16
Uncert1000:Sample_sizeSmall:MethodGAM	0.02	0.02	1.14	0.26
Uncert90:Sample_sizeSmall:MethodGLM	0	0.02	0.08	0.94
Uncert500:Sample_sizeSmall:MethodGLM	-0.03	0.02	-1.69	0.09
Uncert1000:Sample_sizeSmall:MethodGLM	-0.04	0.02	-2	0.05
Uncert90:Sample_sizeSmall:MethodMXT	0	0.02	-0.1	0.92
Uncert500:Sample_sizeSmall:MethodMXT	0.01	0.02	0.29	0.77
Uncert1000:Sample_sizeSmall:MethodMXT	0	0.02	0	1

Supplementary table 3.1.12 Numerical results of the models:
$\mathrm{D.index\_Wide} \sim \mathrm{Uncertainty} * \mathrm{Sample\_size} * \mathrm{Method}$
(analysing the impact of varying levels of gridded uncertainty on model accuracy across different modelling conditions for wide-ranged virtual species.

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.87	0	199.21	0
Uncert90	0	0.01	-0.44	0.66
Uncert500	-0.02	0.01	-3.77	0
Uncert1000	0	0.01	-0.23	0.82
Sample_sizeSmall	-0.01	0.01	-1.19	0.23
MethodGAM	0	0.01	0.44	0.66
MethodGLM	0.02	0.01	3.29	0
MethodMXT	0	0.01	-0.23	0.82
Uncert90:Sample_sizeSmall	0	0.01	-0.29	0.77
Uncert500:Sample_sizeSmall	0.01	0.01	1.48	0.14
Uncert1000:Sample_sizeSmall	0	0.01	-0.14	0.89
Uncert90:MethodGAM	0	0.01	0.53	0.6
Uncert500:MethodGAM	0	0.01	-0.37	0.71
Uncert1000:MethodGAM	-0.05	0.01	-5.66	0
Uncert90:MethodGLM	0	0.01	-0.03	0.97
Uncert500:MethodGLM	0	0.01	-0.28	0.78
Uncert1000:MethodGLM	0	0.01	-0.23	0.82
Uncert90:MethodMXT	0	0.01	0.36	0.72
Uncert500:MethodMXT	0	0.01	0.51	0.61
Uncert1000:MethodMXT	0.03	0.01	2.99	0
Sample_sizeSmall:MethodGAM	-0.08	0.01	-8.96	0
Sample_sizeSmall:MethodGLM	0	0.01	0.21	0.83
Sample_sizeSmall:MethodMXT	0.04	0.01	4.94	0
Uncert90:Sample_sizeSmall:MethodGAM	0	0.01	-0.31	0.76
Uncert500:Sample_sizeSmall:MethodGAM	0	0.01	0.24	0.81
Uncert1000:Sample_sizeSmall:MethodGAM	0.01	0.01	0.92	0.36
Uncert90:Sample_sizeSmall:MethodGLM	0	0.01	-0.16	0.87
Uncert500:Sample_sizeSmall:MethodGLM	0	0.01	0.38	0.7
Uncert1000:Sample_sizeSmall:MethodGLM	-0.02	0.01	-1.95	0.05
Uncert90:Sample_sizeSmall:MethodMXT	0	0.01	-0.35	0.72
Uncert500:Sample_sizeSmall:MethodMXT	0	0.01	-0.25	0.8
Uncert1000:Sample_sizeSmall:MethodMXT	-0.09	0.01	-7.07	1.62E-12

Supplementary table 3.1.13 Numerical results of the models:
$\mathrm{D.index\_Narrow} \sim \mathrm{Uncertainty} * \mathrm{Sample\_size} * \mathrm{Method}$
(analysing the the impact of varying levels of gridded uncertainty on model accuracy across different modelling conditions for narrow-ranged real species).

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.95	0.03	30.73	0.00
Uncert90	-0.04	0.05	-0.79	0.43
Uncert500	-0.08	0.05	-1.65	0.10
Uncert1000	-0.10	0.05	-1.97	0.05
MethodGAM	-0.07	0.04	-1.55	0.13
MethodGLM	-0.04	0.04	-0.87	0.39
MethodMXT	0.01	0.04	0.29	0.78
Uncert90:MethodGAM	0.01	0.07	0.15	0.88
Uncert500:MethodGAM	-0.05	0.07	-0.70	0.49
Uncert1000:MethodGAM	-0.11	0.07	-1.49	0.14
Uncert90:MethodGLM	0.07	0.07	0.91	0.37
Uncert500:MethodGLM	0.03	0.07	0.37	0.72
Uncert1000:MethodGLM	0.04	0.08	0.53	0.60
Uncert90:MethodMXT	0.01	0.07	0.13	0.90
Uncert500:MethodMXT	-0.06	0.08	-0.66	0.51
Uncert1000:MethodMXT	-0.09	0.08	-1.23	0.23

Supplementary table 3.1.14 Numerical results of the models:
$\mathrm{D.index\_Wide} \sim \mathrm{Uncertainty} * \mathrm{Sample\_size} * \mathrm{Method}$
(analysing the impact of varying levels of gridded uncertainty on model accuracy across different modelling conditions for wide-ranged real species.

Term	Estimate	Std.error	Statistic	p.value
(Intercept)	0.94	0.01	83.08	0.00
Uncert90	-0.02	0.02	-0.94	0.35
Uncert500	-0.04	0.02	-2.15	0.04
Uncert1000	-0.04	0.02	-2.27	0.03
MethodGAM	0.03	0.02	1.82	0.07
MethodGLM	0.01	0.02	0.76	0.45
MethodMXT	0.02	0.02	1.09	0.28
Uncert90:MethodGAM	-0.04	0.03	-1.41	0.16
Uncert500:MethodGAM	-0.08	0.03	-3.06	0.00
Uncert1000:MethodGAM	-0.10	0.03	-3.73	0.00
Uncert90:MethodGLM	0.01	0.03	0.26	0.80
Uncert500:MethodGLM	0.03	0.03	0.92	0.36
Uncert1000:MethodGLM	0.02	0.03	0.68	0.50
Uncert90:MethodMXT	0.00	0.03	0.02	0.98
Uncert500:MethodMXT	0.01	0.03	0.21	0.83
Uncert1000:MethodMXT	0.01	0.03	0.36	0.72

Appendix S3.2 Method to obtain presence/absence plant data from the Revillagigedo Archipelago, Mexico.

For this study presence/absence points from vegetation transects conducted between 2018 and 2022 were used. Sampling focused on Socorro (132 km²) and Clarion (19.8 km²), the two largest islands in the archipelago, due to their ecological importance, being the hosts for most species in the archipelago.

In 2018, plant records were collected on Clarion Island during both the dry (May) and wet (November) seasons, given its limited prior study (Benavides et al., 2019). In 2022, data collection on Socorro Island took place at the start of the wet season (November) to cover a broad range of perennial and non-perennial species.

Transect lengths varied based on island size and habitat diversity, with 100m transects on Clarion and 500m to 1km transects on Socorro. The transects were selected systematically to reflect environmental heterogeneity, with efforts to avoid sampling bias through widespread distribution across the islands.

Chapter 4

Appendix S4.1 Supplementary tables and figures

Supplementary table 4.1.1 Model accuracy results for the AUC and Boyce Metrics

Island	Res	Species	Baseline	AUC	Boyce
Clarion	30	Dodonaea_viscosa	DEM	0.60	-0.10
Clarion	30	Brickellia_peninsularis	DEM	0.59	-0.17
Clarion	30	Caesalpinia_bonduc	DEM	0.80	0.24
Clarion	30	Euphorbia_anthonyi	DEM	0.61	0.30
Clarion	30	Euphorbia_californica	DEM	0.68	0.22
Clarion	30	Karwinskia_humboldtiana	DEM	0.72	0.24
Clarion	30	Tribulus_cistoides	DEM	0.56	0.13
Clarion	30	Waltheria_indica	DEM	0.61	0.08
Clarion	30	Zanthoxylum_fagara	DEM	0.58	0.15
Clarion	30	Nicotiana_stocktonii	DEM	0.72	0.05
Clarion	30	Teucrium_townsendii	DEM	0.60	-0.01
Clarion	30	Dodonaea_viscosa	CH	0.72	0.25
Clarion	30	Brickellia_peninsularis	CH	0.51	-0.03
Clarion	30	Caesalpinia_bonduc	CH	0.84	0.18
Clarion	30	Euphorbia_anthonyi	CH	0.56	0.24
Clarion	30	Euphorbia_californica	CH	0.65	0.10
Clarion	30	Karwinskia_humboldtiana	CH	0.57	0.10
Clarion	30	Tribulus_cistoides	CH	0.70	0.13
Clarion	30	Waltheria_indica	CH	0.59	0.14
Clarion	30	Zanthoxylum_fagara	CH	0.75	0.18
Clarion	30	Teucrium_townsendii	CH	0.82	0.11
Clarion	30	Dodonaea_viscosa	WC	0.62	0.04
Clarion	30	Brickellia_peninsularis	WC	0.58	0.26
Clarion	30	Caesalpinia_bonduc	WC	0.78	0.13
Clarion	30	Euphorbia_anthonyi	WC	0.54	0.20
Clarion	30	Euphorbia_californica	WC	0.64	0.14
Clarion	30	Karwinskia_humboldtiana	WC	0.58	0.13
Clarion	30	Tribulus_cistoides	WC	0.69	0.20
Clarion	30	Waltheria_indica	WC	0.70	0.26
Clarion	30	Zanthoxylum_fagara	WC	0.68	0.24
Clarion	30	Nicotiana_stocktonii	WC	0.96	0.33
Clarion	30	Teucrium_townsendii	WC	0.79	0.15
Clarion	90	Dodonaea_viscosa	DEM	0.54	-0.01
Clarion	90	Brickellia_peninsularis	DEM	0.58	-0.16
Clarion	90	Caesalpinia_bonduc	DEM	0.64	-0.08
Clarion	90	Euphorbia_anthonyi	DEM	0.58	0.22

Clarion	90	Euphorbia_californica	DEM	0.50	-0.01
Clarion	90	Karwinskia_humboldtiana	DEM	0.61	0.09
Clarion	90	Tribulus_cistoides	DEM	0.61	0.16
Clarion	90	Waltheria_indica	DEM	0.57	-0.10
Clarion	90	Zanthoxylum_fagara	DEM	0.65	0.17
Clarion	90	Nicotiana_stocktonii	DEM	0.64	0.11
Clarion	90	Teucrium_townsendii	DEM	0.63	0.04
Clarion	90	Dodonaea_viscosa	CH	0.65	0.15
Clarion	90	Brickellia_peninsularis	CH	0.54	0.01
Clarion	90	Caesalpinia_bonduc	CH	0.87	0.25
Clarion	90	Euphorbia_anthonyi	CH	0.64	0.30
Clarion	90	Euphorbia_californica	CH	0.69	0.17
Clarion	90	Karwinskia_humboldtiana	CH	0.71	0.17
Clarion	90	Tribulus_cistoides	CH	0.69	0.20
Clarion	90	Waltheria_indica	CH	0.64	0.29
Clarion	90	Zanthoxylum_fagara	CH	0.63	0.12
Clarion	90	Teucrium_townsendii	CH	0.88	0.22
Clarion	90	Dodonaea_viscosa	WC	0.54	-0.02
Clarion	90	Brickellia_peninsularis	WC	0.58	0.21
Clarion	90	Caesalpinia_bonduc	WC	0.80	0.14
Clarion	90	Euphorbia_anthonyi	WC	0.61	0.33
Clarion	90	Euphorbia_californica	WC	0.65	0.23
Clarion	90	Karwinskia_humboldtiana	WC	0.56	-0.01
Clarion	90	Tribulus_cistoides	WC	0.69	0.23
Clarion	90	Waltheria_indica	WC	0.65	0.22
Clarion	90	Zanthoxylum_fagara	WC	0.54	0.00
Clarion	90	Nicotiana_stocktonii	WC	0.84	0.22
Clarion	90	Teucrium_townsendii	WC	0.88	0.13
Clarion	500	Dodonaea_viscosa	DEM	0.56	0.06
Clarion	500	Brickellia_peninsularis	DEM	0.62	0.03
Clarion	500	Caesalpinia_bonduc	DEM	0.68	0.09
Clarion	500	Euphorbia_anthonyi	DEM	0.53	-0.11
Clarion	500	Euphorbia_californica	DEM	0.62	-0.01
Clarion	500	Karwinskia_humboldtiana	DEM	0.59	-0.09
Clarion	500	Tribulus_cistoides	DEM	0.74	0.14
Clarion	500	Waltheria_indica	DEM	0.64	0.01
Clarion	500	Zanthoxylum_fagara	DEM	0.63	-0.18
Clarion	500	Teucrium_townsendii	DEM	0.75	-0.04
Clarion	500	Dodonaea_viscosa	CH	0.64	0.09
Clarion	500	Brickellia_peninsularis	CH	0.54	-0.02
Clarion	500	Caesalpinia_bonduc	CH	0.80	0.19
Clarion	500	Euphorbia_anthonyi	CH	0.63	0.21
Clarion	500	Euphorbia_californica	CH	0.62	0.24
Clarion	500	Karwinskia_humboldtiana	CH	0.65	0.13

Clarion	500	Tribulus_cistoides	CH	0.71	0.26
Clarion	500	Waltheria_indica	CH	0.59	0.29
Clarion	500	Zanthoxylum_fagara	CH	0.89	0.51
Clarion	500	Nicotiana_stocktonii	CH	0.75	0.07
Clarion	500	Dodonaea_viscosa	WC	0.58	-0.05
Clarion	500	Brickellia_peninsularis	WC	0.60	0.23
Clarion	500	Euphorbia_anthonyi	WC	0.60	0.16
Clarion	500	Euphorbia_californica	WC	0.68	0.06
Clarion	500	Karwinskia_humboldtiana	WC	0.62	0.02
Clarion	500	Tribulus_cistoides	WC	0.68	0.22
Clarion	500	Waltheria_indica	WC	0.60	0.19
Clarion	500	Zanthoxylum_fagara	WC	0.66	-0.09
Clarion	500	Nicotiana_stocktonii	WC	0.60	-0.11
Socorro	30	Croton_masonii	DEM	0.69	0.24
Socorro	30	Conocarpus_erectus	DEM	0.86	0.27
Socorro	30	Bursera_epinnata	DEM	0.88	0.51
Socorro	30	Dodonaea_viscosa	DEM	0.66	0.30
Socorro	30	Ficus_cotinifolia	DEM	0.53	0.25
Socorro	30	Brickellia_peninsularis	DEM	0.59	0.08
Socorro	30	Guettarda_insularis	DEM	0.67	0.15
Socorro	30	Ilex_socorroensis	DEM	0.78	0.32
Socorro	30	Perityle_socorrosensis	DEM	0.76	0.39
Socorro	30	Psidium_socorrense	DEM	0.69	0.25
Socorro	30	Sideroxylon_socorrense	DEM	0.82	0.39
Socorro	30	Waltheria_indica	DEM	0.56	0.07
Socorro	30	Croton_masonii	CH	0.68	0.08
Socorro	30	Conocarpus_erectus	CH	0.86	0.18
Socorro	30	Bursera_epinnata	CH	0.96	0.45
Socorro	30	Dodonaea_viscosa	CH	0.61	0.19
Socorro	30	Ficus_cotinifolia	CH	0.62	0.11
Socorro	30	Brickellia_peninsularis	CH	0.60	0.24
Socorro	30	Guettarda_insularis	CH	0.63	0.07
Socorro	30	Ilex_socorroensis	CH	0.93	0.42
Socorro	30	Perityle_socorrosensis	CH	0.74	0.18
Socorro	30	Psidium_socorrense	CH	0.73	0.24
Socorro	30	Sideroxylon_socorrense	CH	0.90	0.57
Socorro	30	Waltheria_indica	CH	0.64	0.32
Socorro	30	Croton_masonii	WC	0.86	0.40
Socorro	30	Conocarpus_erectus	WC	0.87	0.25
Socorro	30	Bursera_epinnata	WC	0.97	0.31
Socorro	30	Dodonaea_viscosa	WC	0.70	0.31
Socorro	30	Ficus_cotinifolia	WC	0.67	0.29
Socorro	30	Brickellia_peninsularis	WC	0.69	0.46
Socorro	30	Guettarda_insularis	WC	0.75	0.32

Socorro	30	Ilex_socorroensis	WC	0.89	0.41
Socorro	30	Perityle_socorrosensis	WC	0.78	0.20
Socorro	30	Psidium_socorrense	WC	0.82	0.52
Socorro	30	Sideroxylon_socorrense	WC	0.89	0.44
Socorro	30	Waltheria_indica	WC	0.69	0.41
Socorro	90	Croton_masonii	DEM	0.66	0.26
Socorro	90	Conocarpus_erectus	DEM	0.87	0.35
Socorro	90	Bursera_epinnata	DEM	0.87	0.48
Socorro	90	Dodonaea_viscosa	DEM	0.66	0.26
Socorro	90	Ficus_cotinifolia	DEM	0.66	0.27
Socorro	90	Brickellia_peninsularis	DEM	0.66	0.32
Socorro	90	Guettarda_insularis	DEM	0.71	0.28
Socorro	90	Ilex_socorroensis	DEM	0.86	0.41
Socorro	90	Perityle_socorrosensis	DEM	0.73	0.31
Socorro	90	Psidium_socorrense	DEM	0.67	0.20
Socorro	90	Sideroxylon_socorrense	DEM	0.80	0.34
Socorro	90	Waltheria_indica	DEM	0.56	0.06
Socorro	90	Croton_masonii	CH	0.74	0.31
Socorro	90	Bursera_epinnata	CH	0.93	0.54
Socorro	90	Dodonaea_viscosa	CH	0.67	0.31
Socorro	90	Ficus_cotinifolia	CH	0.59	0.26
Socorro	90	Brickellia_peninsularis	CH	0.62	0.24
Socorro	90	Guettarda_insularis	CH	0.69	0.11
Socorro	90	Ilex_socorroensis	CH	0.90	0.24
Socorro	90	Perityle_socorrosensis	CH	0.76	0.16
Socorro	90	Psidium_socorrense	CH	0.82	0.47
Socorro	90	Sideroxylon_socorrense	CH	0.92	0.59
Socorro	90	Waltheria_indica	CH	0.69	0.40
Socorro	90	Croton_masonii	WC	0.83	0.39
Socorro	90	Conocarpus_erectus	WC	0.92	0.38
Socorro	90	Bursera_epinnata	WC	0.97	0.45
Socorro	90	Dodonaea_viscosa	WC	0.68	0.28
Socorro	90	Ficus_cotinifolia	WC	0.67	0.20
Socorro	90	Brickellia_peninsularis	WC	0.70	0.49
Socorro	90	Guettarda_insularis	WC	0.81	0.47
Socorro	90	Ilex_socorroensis	WC	0.93	0.57
Socorro	90	Perityle_socorrosensis	WC	0.85	0.45
Socorro	90	Psidium_socorrense	WC	0.82	0.50
Socorro	90	Sideroxylon_socorrense	WC	0.92	0.54
Socorro	90	Waltheria_indica	WC	0.69	0.46
Socorro	500	Croton_masonii	DEM	0.86	0.30
Socorro	500	Conocarpus_erectus	DEM	0.90	-0.05
Socorro	500	Bursera_epinnata	DEM	0.88	0.22
Socorro	500	Dodonaea_viscosa	DEM	0.75	0.42

Socorro	500	Ficus_cotinifolia	DEM	0.64	0.16
Socorro	500	Brickellia_peninsularis	DEM	0.71	0.30
Socorro	500	Guettarda_insularis	DEM	0.80	0.35
Socorro	500	Ilex_socorroensis	DEM	0.88	0.23
Socorro	500	Perityle_socorrosensis	DEM	0.82	0.13
Socorro	500	Psidium_socorrense	DEM	0.82	0.51
Socorro	500	Sideroxylon_socorrense	DEM	0.90	0.34
Socorro	500	Waltheria_indica	DEM	0.78	0.48
Socorro	500	Croton_masonii	CH	0.73	0.26
Socorro	500	Conocarpus_erectus	CH	0.88	0.13
Socorro	500	Bursera_epinnata	CH	0.91	0.31
Socorro	500	Dodonaea_viscosa	CH	0.74	0.30
Socorro	500	Ficus_cotinifolia	CH	0.63	0.01
Socorro	500	Brickellia_peninsularis	CH	0.69	0.30
Socorro	500	Guettarda_insularis	CH	0.66	0.05
Socorro	500	Ilex_socorroensis	CH	0.91	0.18
Socorro	500	Perityle_socorrosensis	CH	0.73	0.22
Socorro	500	Psidium_socorrense	CH	0.76	0.09
Socorro	500	Sideroxylon_socorrense	CH	0.90	0.42
Socorro	500	Waltheria_indica	CH	0.69	0.30
Socorro	500	Croton_masonii	WC	0.87	0.50
Socorro	500	Conocarpus_erectus	WC	0.94	0.31
Socorro	500	Dodonaea_viscosa	WC	0.71	0.40
Socorro	500	Ficus_cotinifolia	WC	0.58	-0.02
Socorro	500	Brickellia_peninsularis	WC	0.73	0.47
Socorro	500	Guettarda_insularis	WC	0.73	0.32
Socorro	500	Ilex_socorroensis	WC	0.92	0.25
Socorro	500	Perityle_socorrosensis	WC	0.74	0.21
Socorro	500	Psidium_socorrense	WC	0.83	0.32
Socorro	500	Sideroxylon_socorrense	WC	0.86	0.38
Socorro	500	Waltheria_indica	WC	0.76	0.52

Supplementary table 4.1.1 Model transferability results for the AUC Transferability Metric.

Island	Res	Species	Baseline	Method	AUC
Clarion	30	Brickellia_peninsularis	DEM	GAM	0.63
Clarion	30	Brickellia_peninsularis	CH	GAM	0.58
Clarion	30	Brickellia_peninsularis	WC	GAM	0.54
Clarion	30	Dodonaea_viscosa	DEM	GAM	0.55
Clarion	30	Dodonaea_viscosa	CH	GAM	0.56
Clarion	30	Dodonaea_viscosa	WC	GAM	0.58
Clarion	30	Perityle_socorrosensis	DEM	GAM	0.54
Clarion	30	Perityle_socorrosensis	CH	GAM	0.54
Clarion	30	Perityle_socorrosensis	WC	GAM	0.56
Clarion	30	Waltheria_indica	DEM	GAM	0.63
Clarion	30	Waltheria_indica	CH	GAM	0.55
Clarion	30	Waltheria_indica	WC	GAM	0.6
Clarion	90	Brickellia_peninsularis	DEM	GAM	0.54
Clarion	90	Brickellia_peninsularis	CH	GAM	0.54
Clarion	90	Brickellia_peninsularis	WC	GAM	0.53
Clarion	90	Dodonaea_viscosa	DEM	GAM	0.56
Clarion	90	Dodonaea_viscosa	CH	GAM	0.6
Clarion	90	Dodonaea_viscosa	WC	GAM	0.58
Clarion	90	Perityle_socorrosensis	DEM	GAM	0.6
Clarion	90	Perityle_socorrosensis	CH	GAM	0.53
Clarion	90	Perityle_socorrosensis	WC	GAM	0.56
Clarion	90	Waltheria_indica	DEM	GAM	0.56
Clarion	90	Waltheria_indica	CH	GAM	0.53
Clarion	90	Waltheria_indica	WC	GAM	0.64
Clarion	500	Brickellia_peninsularis	DEM	GAM	NA
Clarion	500	Brickellia_peninsularis	CH	GAM	0.56
Clarion	500	Brickellia_peninsularis	WC	GAM	0.55
Clarion	500	Dodonaea_viscosa	DEM	GAM	NA
Clarion	500	Dodonaea_viscosa	CH	GAM	0.58
Clarion	500	Dodonaea_viscosa	WC	GAM	0.6
Clarion	500	Perityle_socorrosensis	DEM	GAM	NA
Clarion	500	Perityle_socorrosensis	CH	GAM	0.52
Clarion	500	Perityle_socorrosensis	WC	GAM	0.56
Clarion	500	Waltheria_indica	DEM	GAM	0.53
Clarion	500	Waltheria_indica	CH	GAM	0.51
Clarion	500	Waltheria_indica	WC	GAM	0.64
Socorro	30	Euphorbia_anthonyi	DEM	GAM	0.84
Socorro	30	Euphorbia_anthonyi	CH	GAM	0.62
Socorro	30	Euphorbia_anthonyi	WC	GAM	0.74
Socorro	30	Tribulus_cistoides	DEM	GAM	0.63
Socorro	30	Tribulus_cistoides	CH	GAM	0.59

Socorro	30	Tribulus_cistoides	WC	GAM	0.6
Socorro	30	Nicotiana_stocktonii	DEM	GAM	0.92
Socorro	30	Nicotiana_stocktonii	CH	GAM	NA
Socorro	30	Nicotiana_stocktonii	WC	GAM	0.81
Socorro	90	Euphorbia_anthonyi	DEM	GAM	0.67
Socorro	90	Euphorbia_anthonyi	CH	GAM	0.64
Socorro	90	Euphorbia_anthonyi	WC	GAM	0.7
Socorro	90	Tribulus_cistoides	DEM	GAM	0.7
Socorro	90	Tribulus_cistoides	CH	GAM	0.58
Socorro	90	Tribulus_cistoides	WC	GAM	0.67
Socorro	90	Nicotiana_stocktonii	DEM	GAM	0.95
Socorro	90	Nicotiana_stocktonii	CH	GAM	0.71
Socorro	90	Nicotiana_stocktonii	WC	GAM	0.85
Socorro	500	Euphorbia_anthonyi	DEM	GAM	0.74
Socorro	500	Euphorbia_anthonyi	CH	GAM	0.84
Socorro	500	Euphorbia_anthonyi	WC	GAM	0.7
Socorro	500	Tribulus_cistoides	DEM	GAM	0.62
Socorro	500	Tribulus_cistoides	CH	GAM	0.66
Socorro	500	Tribulus_cistoides	WC	GAM	0.65
Socorro	500	Nicotiana_stocktonii	DEM	GAM	0.8
Socorro	500	Nicotiana_stocktonii	CH	GAM	0.54
Socorro	500	Nicotiana_stocktonii	WC	GAM	0.75

120

Chapter 5

Appendix S5.1.1 Supplementary tables and figures

Table S5.1.1 Species list and their criteria for inclusion in the climate change analysis, organized by archipelago. Categories include: Chr (Chorotype: END = Endemic, NE = Non-endemic native), IUCN (Conservation status: CR = Critically Endangered, EN = Endangered, VU= Vulnerable, LC = Least Concern, NA=Non assessed), Threats (Key threats to the species), Importance (Ecological or cultural significance), N (Initial sample size), N Thin (Sample size after thinning), and *Ref (References).

Island	Species	Chr	IUCN	Threats	Importance	N	N Thin	Ref*
Canaries	Echium acanthocarpum	END	CR	Land use change	Vegetation/ Rarity	22	14	1,2,3
Canaries	Ceropegia dichotoma	END	EN	Invasives/ Bottleneck	Vegetation/ Habitat stabilisation	124	64	1,2,3
Canaries	Sideroxylon canariense	END	EN	Invasives/ Climate change	Vegetation/ Edible fruits	21	20	1,2,3
Canaries	Aeonium gomerense	END	EN	Land use change / Invasives/ Diseases	Rarity/ Declining	32	13	1,2,3
Canaries	Dracaena draco	NE	EN	Land use change / Invasives/ Wildfires/ Diseases/ Wood	Vegetation/ Habitat stabilisation/ Edible fruits	465	199	1,2,3
Canaries	Pistacia lentiscus	NE	LC	Land use change	Vegetation/ Edible fruits	32	20	1,2,3
Canaries	Kleinia neriifolia	END	LC	Land use change	Vegetation/ Habitat stabilisation	2267	1120	1,2,3
Canaries	Tamarix canariensis	END	LC	Land use change	Vegetation/ Habitat stabilisation	192	105	1,2,3

Canaries	Cneorum pulverulentum	END	LC	Land use change	Vegetation/ Habitat stabilisation/ Floral visitors	107	67	1,2,3
Canaries	Morella faya	NE	LC	Land use change / Invasives	Vegetation/ Edible fruits	717	492	1,2,3
Canaries	Convolvulus floridus	END	LC	Land use change / Invasives	Vegetation/ Habitat stabilisation	326	174	1,2,3
Canaries	Phoenix canariensis	END	LC	Land use change / Invasives/ Wildfires	Vegetation/ Edible fruits	569	286	1,2,3
Canaries	Hypericum canariense	END	LC	Land use change / Invasives/ Wildfires/ Diseases	Vegetation/ Habitat stabilisation	488	298	1,2,3
Canaries	Euphorbia canariensis	END	LC	Land use change / Invasives/Wildfires/ Diseases	Vegetation/ Habitat stabilisation/ Declining	825	404	1,2,3
Canaries	Apollonias barbujana	NE	LC	Land use change / Invasives/Wildfires/ Diseases/Wood	Vegetation/ Habitat stabilisation	88	57	1,2,3
Canaries	Ocotea foetens	NE	LC	Land use change / Invasives/Wildfires/ Diseases/Wood	Vegetation/ Habitat stabilisation/ Edible fruits	90	30	1,2,3
Canaries	Echium wildpretii	END	NA	Bottleneck/Climate change	Rarity/ Floral visitors	454	81	2,3
Canaries	Argyranthemum maderense	END	NA	Land use change	Vegetation/ Habitat stabilisation	22	11	2,3

Canaries	Periploca laevigata	NE	NA	Land use change	Vegetation/ Habitat stabilisation	569	318	2,3
Canaries	Bystropogon canariensis	NE	NA	Land use change / Invasives	Vegetation/ Habitat stabilisation	99	64	2,3
Canaries	Descurainia bourgaeana	END	NA	Undetermined	Vegetation/ Habitat stabilisation/ Edible fruits	95	44	2,3
Canaries	Euphorbia balsamifera	NE	NA	Undetermined	Vegetation/ Habitat stabilisation/ Edible fruits/ Declining	1232	415	2,3
Canaries	Launaea arborescens	NE	NA	Undetermined	Vegetation/ Habitat stabilisation	1478	713	2,3
Canaries	Arbutus canariensis	END	NT	Invasives/Wildfires	Vegetation/ Edible fruits	255	108	2,3
Canaries	Salix canariensis	END	NT	Land use change / Invasives	Vegetation/ Habitat stabilisation	93	64	2,3
Canaries	Limonium arborescens	END	NT	Land use change / Invasives/ Diseases	Vegetation/Rarity	34	14	2,3
Canaries	Maytenus canariensis	END	NT	Undertermined	Vegetation/ Habitat stabilisation	61	58	2,3
Canaries	Euphorbia handiensis	END	VU	Land use change / Invasives	Rarity/ Declining	39	13	2,3
Canaries	Asparagus arborescens	END	VU	Land use change / Invasives	Vegetation/ Habitat stabilisation	46	44	2,3
Canaries	Heberdenia excelsa	NE	VU	Land use change / Invasives/Wildfires/ Diseases/Climate change	Vegetation/ Declining	26	24	2,3
Galapagos	Calandrinia galapagosa	END	CD	Hervibores	Rarity/ Declining	71	17	1,4,5

Galapagos	Gossypium darwinii	END	LC	Hervibores	Nesting material/ Vegetation	230	54	1,4,5
Galapagos	Scalesia affinis	END	LC	Land use change / Invasives	Declining/ Vegetation	73	26	1,4,5,6
Galapagos	Jasminocereus thouarsii	END	LC	Land use change / Invasives	Vegetation/ Edible fruits	269	64	1,4,5
Galapagos	Opuntia galapageia	END	LC	Land use change / Invasives	Vegetation/ Edible fruits	307	68	1,4,5
Galapagos	Laguncularia racemosa	NE	LC	Restricted range locally	Vegetation/ Habitat stabilisation	105	27	1,4,5
Galapagos	Avicennia germinans	NE	LC	Restricted range locally	Vegetation/ Habitat stabilisation	70	18	1,4,5
Galapagos	Rhizophora mangle	NE	LC	Restricted range locally	Vegetation/ Habitat stabilisation	169	40	1,4,5
Galapagos	Piscidia carthagenensis	END	LC	Wood	Declining/ Vegetation	104	19	1,4,5
Galapagos	Lecocarpus pinnatifidus	END	NA	Hervibores	Declining/ Vegetation/Rarity	26	6	4,5
Galapagos	Miconia robinsoniana	END	NA	Land use change / Invasives	Declining/ Vegetation	77	13	4,5
Galapagos	Galvezia leucantha	END	NA	Land use change / Invasives	Rarity /Declining	48	14	4,5
Galapagos	Trigonopterum laricifolium	END	NA	Undetermined	Vegetation/ Habitat stabilisation	25	14	4,5
Galapagos	Lippia rosmarinifolia	END	NA	Wood	Rarity / Declining	90	25	4,5
Galapagos	Darwiniothamnus tenuifolius	END	NT	Cascading effects	Rarity / Declining	207	136	1,4,5,6
Galapagos	Lecocarpus darwinii	END	NT	Hervibores	Rarity/ Declining	63	40	1,4,5,6
Galapagos	Scalesia villosa	END	VU	Land use change / Invasives	Declining/ Vegetation	23	13	1,4,5,7

Galapagos	Pleuropetalum darwinii	END	VU	Restricted range	Rarity / Declining	38	21	1,4,5,8
Galapagos	Scalesia pedunculata	END	VU	Wood	Declining/ Vegetation	161	124	1,4,5,9
Hawaii	Colubrina oppositifolia	END	CR	Herbivores/ Alien plants/Wildfires	Rarity/ Declining	16	16	1,7
Hawaii	Hibiscus brackenridgei	END	CR	Invasives/Wildfires	Rarity/ Declining	48	44	1,7
Hawaii	Abutilon menziesii	END	CR	Land use change / Invasives/Wildfires/ Diseases	Rarity/ Declining	17	15	1,7
Hawaii	Pritchardia martii	END	EN	Invasives/Wildfires	Vegetation/ Habitat stabilisation/ Cultural	58	54	1,7
Hawaii	Dicranopteris linearis	NE	LC	Invasives/Wildfires	Vegetation/ Habitat stabilisation	306	108	1,7
Hawaii	Dodonaea viscosa	NE	LC	Undetermined	Vegetation/ Habitat stabilisation	286	129	1,7
Hawaii	Machaerina angustifolia	END	NA	Alien plants	Vegetation/ Habitat stabilisation	100	45	7
Hawaii	Scaevola taccada	NE	NA	Alien plants	Vegetation/ Habitat stabilisation/ Cultural	1235	412	7
Hawaii	Pipturus albidus	END	NA	Diseases	Vegetation/ Habitat stabilisation/ Cultural	396	338	7
Hawaii	Syzygium sandwicense	END	NA	Herbivores/ Alien plants	Vegetation/ Edible fruits/ Cultural	119	109	7
Hawaii	Melicope clusiifolia	END	NA	Herbivores/ Alien plants	Vegetation/ Cultural	138	128	7
Hawaii	Sadleria pallida	END	NA	Herbivores/ Alien plants	Vegetation/ Habitat stabilisation	106	73	7

Hawaii	Freycinetia arborea	NE	NA	Herbivores/Alien plants	Vegetation/Habitat stabilisation	193	74	7
Hawaii	Microlepia strigosa	NE	NA	Herbivores/Alien plants	Vegetation/Habitat stabilisation/Cultural	62	35	7
Hawaii	Vaccinium calycinum	END	NA	Invasives	Vegetation/Edible fruits	127	47	7
Hawaii	Rubus hawaiensis	END	NA	Invasives	Vegetation/Edible fruits	241	220	7
Hawaii	Sophora chrysophylla	END	NA	Invasives	Vegetation/Edible fruits/Cultural	336	76	7
Hawaii	Cibotium menziesii	END	NA	Undetermined	Vegetation/Habitat stabilisation	31	19	7
Hawaii	Myoporum sandwicense	NE	NA	Undetermined	Vegetation/Habitat stabilisation/Cultural	237	83	7
Hawaii	Metrosideros polymorpha	END	VU	Cascading effects/Herbivores/Alien plants	Nesting material/Vegetation	314	129	1,7
Hawaii	Touchardia latifolia	END	VU	Herbivores/Alien plants	Vegetation/Edible fruits/Floral visitors/Declining	45	35	1,7
Hawaii	Acacia koa	END	VU	Invasives/Diseases	Vegetation/Habitat stabilisation/Declining	294	110	1,7
Hawaii	Diospyros sandwicensis	END	VU	Invasives/Wildfires	Vegetation/Edible fruits	79	77	1,7
Hawaii	Hibiscus kokio	END	VU	Invasives/Wildfires	Vegetation/Edible fruits	34	30	1,7
Revillagigedo	Guettarda insularis	END	CR	Global rarity	Vegetation/Habitat stabilisation/Edible fruits	298	43	1,8,9
Revillagigedo	Ilex socorroensis	END	CR	Global rarity/Diseas	Rarity/Declining/	0	0	1,8,9

				es/Volcanic activity	Vegetation/ Edible fruits			
Revillagigedo	Zanthoxylum fagara	NE	LC	Undetermined	Vegetation/ Habitat stabilisation	31	15	1,8,9
Revillagigedo	Oreopanax xalapensis	NE	LC	Undetermined	Vegetation/ Habitat stabilisation	14	9	1,8,9
Revillagigedo	Ficus cotinifolia	NE	LC	Undetermined	Vegetation/ Habitat stabilisation/ Edible fruits	103	35	1,8,9
Revillagigedo	Bursera epinnata	NE	LC	Undetermined	Vegetation/ Habitat stabilisation/Floral visitors	82	15	1,8,9
Revillagigedo	Karwinskia humboldtiana	NE	LC	Undetermined	Vegetation/ Habitat stabilisation/Nesting material	26	15	1,8,9
Revillagigedo	Hippomane mancinella	NE	LC	Undetermined	Vegetation/ Habitat stabilisation/Nesting material	36	15	1,8,9
Revillagigedo	Conocarpus erectus	NE	LC	Undetermined	Vegetation/ Habitat stabilisation/Nesting material	95	15	1,8,9
Revillagigedo	Nicotiana stocktonii	END	NA	Global rarity	Rarity/Floral visitors	55	32	8,9
Revillagigedo	Perityle socorrosensis	END	NA	Global rarity	Vegetation/ Habitat stabilisation	101	47	8,9
Revillagigedo	Croton masonii	END	NA	Global rarity	Vegetation/ Habitat stabilisation	277	37	8,9
Revillagigedo	Opuntia sp	END	NA	Global rarity/Diseases/Hervibory	Vegetation/ Habitat stabilisation/ Edible fruits	38	10	8,9
Revillagigedo	Psidium socorrense	END	NA	Global rarity/Invasives	Vegetation/ Habitat stabilisation/ Edible fruits	114	16	8,9

Revillagigedo	Teucrium townsendii	END	NA	Global rarity/Undetermined	Rarity/Floral visitors	31	15	8,9
Revillagigedo	Euphorbia californica	NE	NA	Undetermined	Vegetation/Habitat stabilisation	22	9	8,9
Revillagigedo	Brickellia peninsularis	NE	NA	Undetermined	Vegetation/Habitat stabilisation	256	60	8,9
Revillagigedo	Tribulus cistoides	NE	NA	Undetermined	Vegetation/Habitat stabilisation/Floral visitors	70	27	8,9
Revillagigedo	Sideroxylon socorrense	END	VU	Global rarity/Volcanic activity	Rarity/Vegetation/Habitat stabilisation/Edible fruits	118	27	1,8,9

(*): References

[1] IUCN. 2024. *The IUCN Red List of Threatened Species. Version 2024-1*. https://www.iucnredlist.org. Accessed on [August 2024].

[2] de Canaries, G. (2022). Banco de datos de biodiversidad de Canaries. Gobierno de Canaries, Canary Islands, Spain. https://www. biodiversidadCanaries. es/biota/. Accessed on [August 2024].

[3] Naranjo, E. R., & Armas, R. D. C. (2021). Etnobotánica y biodiversidad cultural canaria: el banco de saberes del jardín botánico canario "Viera y Clavijo"-UA CSIC. Botánica Macaronésica, 31, 145-164.

[4] Charles Darwin Foundation. (2024). Charles Darwin Foundation. https://www.darwinfoundation.org/en/ . Accessed on [August 2024].

[5] Boada, R. (2005). Insectos asociados con plantas amenzadas en las Islas Galápagos, Ecuador. Entomotrópica: Revista internacional para el estudio de la entomología *tropical, 20*(2), 77-88.

[6] Adsersen, H. (1989). The rare plants of the Galápagos Islands and their conservation. *Biological Conservation, 47*(1), 49-77.

[7] Department of Land and Natural Resources:Hawaii, (2024). https://dlnr.hawaii.gov/forestry/plants/. Accessed on [August 2024].

[8] CONANP. (2019). Programa de Manejo: Parque Nacional Revillagigedo. https://www.conanp.gob.mx/programademanejo/PMRevillagigedo.pdf. Accessed on [August 2024].

Table S5.12. Summary of species analysed for microclimatic-driven shifts in their distribution in the context of climate change across archipelagos. The table includes the following categories: Chr (Chorotype: END = Endemic, NE = Non-endemic native), SSP (Climate change tranjectory), Range change (Percentage change in distribution range), AUC (Area Under the Curve, indicating model performance), Loss Rank (Rank of predicted range loss: TL=Total loss,EL=Extreme loss, SVL = Severe Loss, ML = Moderate Loss, G = Gain), and Threat Rank (THR = Threatened, LC = Least Concern, NE=Non-assessed)

Island	Species	Chr	SSP	Range change	AUC	Loss Rank	Threat Rank
Canaries	Aeonium gomerense	END	ssp126	-18.1	0.87	ML	THR
Canaries	Aeonium gomerense	END	ssp370	0.6	0.87	G	THR
Canaries	Aeonium gomerense	END	ssp585	-4.7	0.87	ML	THR
Canaries	Apollonias barbujana	NE	ssp126	-47.1	0.71	ML	LC
Canaries	Apollonias barbujana	NE	ssp370	-58.2	0.71	SVL	LC
Canaries	Apollonias barbujana	NE	ssp585	-72.8	0.71	SVL	LC
Canaries	Arbutus canariensis	END	ssp126	-95.5	0.79	EL	THR
Canaries	Arbutus canariensis	END	ssp370	-95.5	0.79	EL	THR
Canaries	Arbutus canariensis	END	ssp585	-98.9	0.79	EL	THR
Canaries	Argyranthemum maderense	NE	ssp126	11.8	0.95	G	THR
Canaries	Argyranthemum maderense	NE	ssp370	-23.4	0.95	ML	THR
Canaries	Argyranthemum maderense	NE	ssp585	-14.8	0.95	ML	THR
Canaries	Asparagus arborescens	END	ssp126	-11.0	0.75	ML	THR
Canaries	Asparagus arborescens	END	ssp370	-28.8	0.75	ML	THR
Canaries	Asparagus arborescens	END	ssp585	-36.4	0.75	ML	THR
Canaries	Bystropogon canariensis	NE	ssp126	-56.2	0.72	SVL	THR
Canaries	Bystropogon canariensis	NE	ssp370	-61.8	0.72	SVL	THR
Canaries	Bystropogon canariensis	NE	ssp585	-66.5	0.72	SVL	THR
Canaries	Ceropegia dichotoma	END	ssp126	0.5	0.75	G	THR
Canaries	Ceropegia dichotoma	END	ssp370	-2.8	0.75	ML	THR
Canaries	Ceropegia dichotoma	END	ssp585	-2.0	0.75	ML	THR
Canaries	Cneorum pulverulentum	END	ssp126	-8.5	0.75	ML	LC

Canaries	Cneorum pulverulentum	END	ssp370	-21.4	0.75	ML	LC
Canaries	Cneorum pulverulentum	END	ssp585	-37.8	0.75	ML	LC
Canaries	Convolvulus floridus	END	ssp126	-33.5	0.78	ML	LC
Canaries	Convolvulus floridus	END	ssp370	-36.8	0.78	ML	LC
Canaries	Convolvulus floridus	END	ssp585	-61.2	0.78	SVL	LC
Canaries	Descurainia bourgaeana	NE	ssp126	-40.0	0.93	ML	THR
Canaries	Descurainia bourgaeana	NE	ssp370	-49.6	0.93	ML	THR
Canaries	Descurainia bourgaeana	NE	ssp585	-61.1	0.93	SVL	THR
Canaries	Dracaena draco	NE	ssp126	-53.9	0.80	SVL	THR
Canaries	Dracaena draco	NE	ssp370	-49.6	0.80	ML	THR
Canaries	Dracaena draco	NE	ssp585	-61.2	0.80	SVL	THR
Canaries	Echium acanthocarpum	END	ssp126	-47.2	0.93	ML	THR
Canaries	Echium acanthocarpum	END	ssp370	-24.1	0.93	ML	THR
Canaries	Echium acanthocarpum	END	ssp585	-25.3	0.93	ML	THR
Canaries	Echium wildpretii	NE	ssp126	-77.0	0.89	SVL	THR
Canaries	Echium wildpretii	NE	ssp370	-86.5	0.89	EL	THR
Canaries	Echium wildpretii	NE	ssp585	-90.0	0.89	EL	THR
Canaries	Euphorbia balsamifera	NE	ssp126	-42.5	0.82	ML	THR
Canaries	Euphorbia balsamifera	NE	ssp370	-38.8	0.82	ML	THR
Canaries	Euphorbia balsamifera	NE	ssp585	-35.9	0.82	ML	THR
Canaries	Euphorbia canariensis	END	ssp126	-20.2	0.85	ML	LC
Canaries	Euphorbia canariensis	END	ssp126	-20.2	0.85	ML	LC
Canaries	Euphorbia canariensis	END	ssp370	-36.1	0.85	ML	LC
Canaries	Euphorbia canariensis	END	ssp370	-36.1	0.85	ML	LC
Canaries	Euphorbia canariensis	END	ssp585	-38.8	0.85	ML	LC
Canaries	Euphorbia canariensis	END	ssp585	-38.8	0.85	ML	LC
Canaries	Euphorbia handiensis	END	ssp126	-23.6	0.80	ML	THR
Canaries	Euphorbia handiensis	END	ssp370	-37.0	0.80	ML	THR
Canaries	Euphorbia handiensis	END	ssp585	-39.2	0.80	ML	THR
Canaries	Heberdenia excelsa	NE	ssp126	-1.9	0.82	ML	THR
Canaries	Heberdenia excelsa	NE	ssp370	-1.4	0.82	ML	THR
Canaries	Heberdenia excelsa	NE	ssp585	-0.9	0.82	ML	THR
Canaries	Hypericum canariense	END	ssp126	-90.2	0.74	EL	LC
Canaries	Hypericum canariense	END	ssp370	-78.1	0.74	SVL	LC
Canaries	Hypericum canariense	END	ssp585	-83.3	0.74	EL	LC
Canaries	Kleinia neriifolia	END	ssp126	-21.7	0.68	ML	LC
Canaries	Kleinia neriifolia	END	ssp370	-25.1	0.68	ML	LC
Canaries	Kleinia neriifolia	END	ssp585	-36.7	0.68	ML	LC

Canaries	Launaea arborescens	NE	ssp126	2.2	0.76	G	THR
Canaries	Launaea arborescens	NE	ssp370	-0.7	0.76	ML	THR
Canaries	Launaea arborescens	NE	ssp585	-0.4	0.76	ML	THR
Canaries	Limonium arborescens	END	ssp126	-56.7	0.74	SVL	THR
Canaries	Limonium arborescens	END	ssp370	-59.4	0.74	SVL	THR
Canaries	Limonium arborescens	END	ssp585	-79.3	0.74	SVL	THR
Canaries	Maytenus canariensis	END	ssp126	-5.9	0.73	ML	THR
Canaries	Maytenus canariensis	END	ssp370	-3.6	0.73	ML	THR
Canaries	Maytenus canariensis	END	ssp585	-7.7	0.73	ML	THR
Canaries	Morella faya	NE	ssp126	-87.6	0.87	EL	LC
Canaries	Morella faya	NE	ssp370	-84.3	0.87	EL	LC
Canaries	Morella faya	NE	ssp585	-84.5	0.87	EL	LC
Canaries	Ocotea foetens	NE	ssp126	-29.2	0.73	ML	LC
Canaries	Ocotea foetens	NE	ssp370	-37.5	0.73	ML	LC
Canaries	Ocotea foetens	NE	ssp585	-39.0	0.73	ML	LC
Canaries	Periploca laevigata	NE	ssp126	-56.7	0.82	SVL	THR
Canaries	Periploca laevigata	NE	ssp370	-60.7	0.82	SVL	THR
Canaries	Periploca laevigata	NE	ssp585	-77.3	0.82	SVL	THR
Canaries	Phoenix canariensis	END	ssp126	-63.1	0.74	SVL	LC
Canaries	Phoenix canariensis	END	ssp370	-44.4	0.74	ML	LC
Canaries	Phoenix canariensis	END	ssp585	-70.3	0.74	SVL	LC
Canaries	Pistacia lentiscus	NE	ssp126	-31.7	0.72	ML	LC
Canaries	Pistacia lentiscus	NE	ssp370	-47.5	0.72	ML	LC
Canaries	Pistacia lentiscus	NE	ssp585	-53.9	0.72	SVL	LC
Canaries	Sideroxylon canariense	END	ssp126	-15.3	0.82	ML	THR
Canaries	Sideroxylon canariense	END	ssp370	-11.9	0.82	ML	THR
Canaries	Sideroxylon canariense	END	ssp585	-12.5	0.82	ML	THR
Canaries	Tamarix canariensis	END	ssp126	10.1	0.83	G	LC
Canaries	Tamarix canariensis	END	ssp370	8.9	0.83	G	LC
Canaries	Tamarix canariensis	END	ssp585	8.5	0.83	G	LC
Galapagos	Avicennia germinans	NE	ssp126	-23.5	0.73	ML	LC
Galapagos	Avicennia germinans	NE	ssp370	-24.0	0.73	ML	LC
Galapagos	Avicennia germinans	NE	ssp585	-23.6	0.73	ML	LC
Galapagos	Conocarpus erectus	NE	ssp126	-54.6	0.72	SVL	LC
Galapagos	Conocarpus erectus	NE	ssp370	-51.4	0.72	SVL	LC
Galapagos	Conocarpus erectus	NE	ssp585	-52.8	0.72	SVL	LC
Galapagos	Darwiniothamnus tenuifolius	END	ssp126	-4.1	0.69	ML	THR
Galapagos	Darwiniothamnus tenuifolius	END	ssp370	-11.9	0.69	ML	THR

Galapagos	Darwiniothamnus tenuifolius	END	ssp585	-14.6	0.69	ML	THR
Galapagos	Galvezia leucantha	NE	ssp126	-25.6	0.73	ML	THR
Galapagos	Galvezia leucantha	NE	ssp370	-25.8	0.73	ML	THR
Galapagos	Galvezia leucantha	NE	ssp585	-26.6	0.73	ML	THR
Galapagos	Gossypium darwinii	END	ssp126	-53.8	0.73	SVL	LC
Galapagos	Gossypium darwinii	END	ssp370	-53.8	0.73	SVL	LC
Galapagos	Gossypium darwinii	END	ssp585	-53.6	0.73	SVL	LC
Galapagos	Jasminocereus thouarsii	END	ssp126	-66.1	0.75	SVL	LC
Galapagos	Jasminocereus thouarsii	END	ssp370	-58.5	0.75	SVL	LC
Galapagos	Jasminocereus thouarsii	END	ssp585	-59.3	0.75	SVL	LC
Galapagos	Laguncularia racemosa	NE	ssp126	-52.8	0.81	SVL	LC
Galapagos	Laguncularia racemosa	NE	ssp370	-49.6	0.81	ML	LC
Galapagos	Laguncularia racemosa	NE	ssp585	-51.7	0.81	SVL	LC
Galapagos	Lecocarpus darwinii	END	ssp126	-66.4	0.83	SVL	THR
Galapagos	Lecocarpus darwinii	END	ssp370	-100.0	0.83	TL	THR
Galapagos	Lecocarpus darwinii	END	ssp585	-100.0	0.83	TL	THR
Galapagos	Lecocarpus pinnatifidus	NE	ssp126	-25.6	0.71	ML	THR
Galapagos	Lecocarpus pinnatifidus	NE	ssp370	-7.3	0.71	ML	THR
Galapagos	Lecocarpus pinnatifidus	NE	ssp585	-8.0	0.71	ML	THR
Galapagos	Lippia rosmarinifolia	NE	ssp126	-19.7	0.84	ML	THR
Galapagos	Lippia rosmarinifolia	NE	ssp370	-31.8	0.84	ML	THR
Galapagos	Lippia rosmarinifolia	NE	ssp585	-39.0	0.84	ML	THR
Galapagos	Miconia robinsoniana	NE	ssp126	-25.9	0.75	ML	THR
Galapagos	Miconia robinsoniana	NE	ssp370	-48.1	0.75	ML	THR
Galapagos	Miconia robinsoniana	NE	ssp585	-51.7	0.75	SVL	THR
Galapagos	Opuntia galapageia	END	ssp126	0.1	0.71	G	LC
Galapagos	Opuntia galapageia	END	ssp370	-8.0	0.71	ML	LC
Galapagos	Opuntia galapageia	END	ssp585	-9.4	0.71	ML	LC
Galapagos	Piscidia carthagenensis	END	ssp126	-55.1	0.71	SVL	LC
Galapagos	Piscidia carthagenensis	END	ssp370	-49.1	0.71	ML	LC
Galapagos	Piscidia carthagenensis	END	ssp585	-49.5	0.71	ML	LC
Galapagos	Pleuropetalum darwinii	END	ssp126	-11.2	0.72	ML	THR
Galapagos	Pleuropetalum darwinii	END	ssp370	-24.1	0.72	ML	THR
Galapagos	Pleuropetalum darwinii	END	ssp585	-11.7	0.72	ML	THR
Galapagos	Rhizophora mangle	NE	ssp126	-36.5	0.74	ML	LC
Galapagos	Rhizophora mangle	NE	ssp370	-41.0	0.74	ML	LC
Galapagos	Rhizophora mangle	NE	ssp585	-38.4	0.74	ML	LC
Galapagos	Scalesia affinis	END	ssp126	-68.6	0.69	SVL	LC
Galapagos	Scalesia affinis	END	ssp370	-63.0	0.69	SVL	LC

Galapagos	Scalesia affinis	END	ssp585	-65.0	0.69	SVL	LC
Galapagos	Scalesia pedunculata	END	ssp126	1.4	0.92	G	THR
Galapagos	Scalesia pedunculata	END	ssp370	-0.4	0.92	ML	THR
Galapagos	Scalesia pedunculata	END	ssp585	0.0	0.92	G	THR
Galapagos	Zanthoxylum fagara	END	ssp126	-20.3	0.69	ML	LC
Galapagos	Zanthoxylum fagara	END	ssp370	-28.1	0.69	ML	LC
Galapagos	Zanthoxylum fagara	END	ssp585	-25.4	0.69	ML	LC
Hawaii	Abutilon menziesii	END	ssp126	2.0	0.61	G	THR
Hawaii	Abutilon menziesii	END	ssp370	2.1	0.61	G	THR
Hawaii	Abutilon menziesii	END	ssp585	2.1	0.61	G	THR
Hawaii	Acacia koa	END	ssp126	-13.6	0.78	ML	THR
Hawaii	Acacia koa	END	ssp370	-24.4	0.78	ML	THR
Hawaii	Acacia koa	END	ssp585	-21.1	0.78	ML	THR
Hawaii	Cibotium menziesii	NE	ssp126	-14.4	0.78	ML	THR
Hawaii	Cibotium menziesii	NE	ssp370	-22.2	0.78	ML	THR
Hawaii	Cibotium menziesii	NE	ssp585	-23.5	0.78	ML	THR
Hawaii	Colubrina oppositifolia	END	ssp126	-69.8	0.69	SVL	THR
Hawaii	Colubrina oppositifolia	END	ssp370	-81.1	0.69	EL	THR
Hawaii	Colubrina oppositifolia	END	ssp585	-95.9	0.69	EL	THR
Hawaii	Dicranopteris linearis	NE	ssp126	-28.0	0.77	ML	LC
Hawaii	Dicranopteris linearis	NE	ssp370	-36.1	0.77	ML	LC
Hawaii	Dicranopteris linearis	NE	ssp585	-26.3	0.77	ML	LC
Hawaii	Diospyros sandwicensis	END	ssp126	-23.8	0.78	ML	THR
Hawaii	Diospyros sandwicensis	END	ssp370	-28.7	0.78	ML	THR
Hawaii	Diospyros sandwicensis	END	ssp585	-28.3	0.78	ML	THR
Hawaii	Dodonaea viscosa	NE	ssp126	-5.4	0.71	ML	LC
Hawaii	Dodonaea viscosa	NE	ssp370	-10.4	0.71	ML	LC
Hawaii	Dodonaea viscosa	NE	ssp585	-5.2	0.71	ML	LC
Hawaii	Freycinetia arborea	NE	ssp126	-57.3	0.83	SVL	THR
Hawaii	Freycinetia arborea	NE	ssp370	-57.9	0.83	SVL	THR
Hawaii	Freycinetia arborea	NE	ssp585	-60.7	0.83	SVL	THR
Hawaii	Hibiscus brackenridgei	END	ssp126	-79.1	0.60	SVL	THR
Hawaii	Hibiscus brackenridgei	END	ssp370	-86.5	0.60	EL	THR
Hawaii	Hibiscus brackenridgei	END	ssp585	-99.4	0.60	EL	THR
Hawaii	Hibiscus kokio	END	ssp126	-4.4	0.75	ML	THR
Hawaii	Hibiscus kokio	END	ssp370	-8.7	0.75	ML	THR
Hawaii	Hibiscus kokio	END	ssp585	-11.3	0.75	ML	THR
Hawaii	Machaerina angustifolia	NE	ssp126	-27.5	0.80	ML	THR
Hawaii	Machaerina angustifolia	NE	ssp370	-40.1	0.80	ML	THR
Hawaii	Machaerina angustifolia	NE	ssp585	-52.3	0.80	SVL	THR

Hawaii	Melicope clusiifolia	NE	ssp126	-1.3	0.85	ML	THR
Hawaii	Melicope clusiifolia	NE	ssp370	-0.2	0.85	ML	THR
Hawaii	Melicope clusiifolia	NE	ssp585	5.2	0.85	G	THR
Hawaii	Metrosideros polymorpha	END	ssp126	-34.6	0.81	ML	THR
Hawaii	Metrosideros polymorpha	END	ssp370	-54.3	0.81	SVL	THR
Hawaii	Metrosideros polymorpha	END	ssp585	-58.4	0.81	SVL	THR
Hawaii	Microlepia strigosa	NE	ssp126	-1.9	0.62	ML	THR
Hawaii	Microlepia strigosa	NE	ssp370	-0.9	0.62	ML	THR
Hawaii	Microlepia strigosa	NE	ssp585	2.2	0.62	G	THR
Hawaii	Myoporum sandwicense	NE	ssp126	2.8	0.82	G	THR
Hawaii	Myoporum sandwicense	NE	ssp370	-0.4	0.82	ML	THR
Hawaii	Myoporum sandwicense	NE	ssp585	-0.7	0.82	ML	THR
Hawaii	Pipturus albidus	NE	ssp126	-39.0	0.73	ML	THR
Hawaii	Pipturus albidus	NE	ssp370	-63.3	0.73	SVL	THR
Hawaii	Pipturus albidus	NE	ssp585	-59.6	0.73	SVL	THR
Hawaii	Pritchardia martii	END	ssp126	0.8	0.93	G	THR
Hawaii	Pritchardia martii	END	ssp370	0.8	0.93	G	THR
Hawaii	Pritchardia martii	END	ssp585	0.9	0.93	G	THR
Hawaii	Rubus hawaiensis	NE	ssp126	-24.6	0.84	ML	THR
Hawaii	Rubus hawaiensis	NE	ssp370	-30.6	0.84	ML	THR
Hawaii	Rubus hawaiensis	NE	ssp585	-27.3	0.84	ML	THR
Hawaii	Sadleria pallida	NE	ssp126	-4.4	0.76	ML	THR
Hawaii	Sadleria pallida	NE	ssp370	-4.2	0.76	ML	THR
Hawaii	Sadleria pallida	NE	ssp585	0.8	0.76	G	THR
Hawaii	Scaevola taccada	NE	ssp126	6.5	0.92	G	THR
Hawaii	Scaevola taccada	NE	ssp370	6.1	0.92	G	THR
Hawaii	Scaevola taccada	NE	ssp585	6.5	0.92	G	THR
Hawaii	Sophora chrysophylla	NE	ssp126	-53.9	0.81	SVL	THR
Hawaii	Sophora chrysophylla	NE	ssp370	-62.6	0.81	SVL	THR
Hawaii	Sophora chrysophylla	NE	ssp585	-67.1	0.81	SVL	THR
Hawaii	Syzygium sandwicense	NE	ssp126	-40.0	0.74	ML	THR
Hawaii	Syzygium sandwicense	NE	ssp370	-43.4	0.74	ML	THR
Hawaii	Syzygium sandwicense	NE	ssp585	-39.6	0.74	ML	THR
Hawaii	Touchardia latifolia	END	ssp126	1.4	0.70	G	THR

Hawaii	Touchardia latifolia	END	ssp370	1.2	0.70	G	THR
Hawaii	Touchardia latifolia	END	ssp585	1.3	0.70	G	THR
Hawaii	Vaccinium calycinum	NE	ssp126	-46.8	0.85	ML	THR
Hawaii	Vaccinium calycinum	NE	ssp370	-71.9	0.85	SVL	THR
Hawaii	Vaccinium calycinum	NE	ssp585	-73.8	0.85	SVL	THR
Revillagigedo	Brickellia peninsularis	NE	ssp126	-44.8	0.70	ML	THR
Revillagigedo	Brickellia peninsularis	NE	ssp370	-48.6	0.70	ML	THR
Revillagigedo	Brickellia peninsularis	NE	ssp585	-56.2	0.70	SVL	THR
Revillagigedo	Bursera epinnata	NE	ssp126	-45.9	0.76	ML	LC
Revillagigedo	Bursera epinnata	NE	ssp370	-20.1	0.76	ML	LC
Revillagigedo	Bursera epinnata	NE	ssp585	-18.4	0.76	ML	LC
Revillagigedo	Conocarpus erectus	NE	ssp126	-42.3	0.76	ML	LC
Revillagigedo	Conocarpus erectus	NE	ssp370	-42.3	0.76	ML	LC
Revillagigedo	Conocarpus erectus	NE	ssp585	-42.3	0.76	ML	LC
Revillagigedo	Croton masonii	NE	ssp126	-16.7	0.65	ML	THR
Revillagigedo	Croton masonii	NE	ssp370	-16.7	0.65	ML	THR
Revillagigedo	Croton masonii	NE	ssp585	-16.7	0.65	ML	THR
Revillagigedo	Euphorbia californica	NE	ssp126	-100.0	0.72	TL	THR
Revillagigedo	Euphorbia californica	NE	ssp370	-100.0	0.72	TL	THR
Revillagigedo	Euphorbia californica	NE	ssp585	-100.0	0.72	TL	THR
Revillagigedo	Hippomane mancinella	NE	ssp126	-15.6	0.79	ML	LC
Revillagigedo	Hippomane mancinella	NE	ssp370	-15.6	0.79	ML	LC
Revillagigedo	Hippomane mancinella	NE	ssp585	-15.6	0.79	ML	LC
Revillagigedo	Karwinskia humboldtiana	NE	ssp126	-100.0	0.89	TL	LC
Revillagigedo	Karwinskia humboldtiana	NE	ssp370	-100.0	0.89	TL	LC
Revillagigedo	Karwinskia humboldtiana	NE	ssp585	-100.0	0.89	TL	LC
Revillagigedo	Nicotiana stocktonii	NE	ssp126	7.1	0.63	G	THR
Revillagigedo	Nicotiana stocktonii	NE	ssp370	7.1	0.63	G	THR
Revillagigedo	Nicotiana stocktonii	NE	ssp585	7.1	0.63	G	THR
Revillagigedo	Opuntia sp	NE	ssp126	8.8	0.62	G	THR
Revillagigedo	Opuntia sp	NE	ssp370	8.8	0.62	G	THR
Revillagigedo	Opuntia sp	NE	ssp585	8.8	0.62	G	THR
Revillagigedo	Oreopanax xalapensis	NE	ssp126	14.5	0.87	G	LC
Revillagigedo	Oreopanax xalapensis	NE	ssp370	14.5	0.87	G	LC
Revillagigedo	Oreopanax xalapensis	NE	ssp585	14.5	0.87	G	LC
Revillagigedo	Perityle socorrosensis	NE	ssp126	-66.4	0.62	SVL	THR
Revillagigedo	Perityle socorrosensis	NE	ssp370	-66.4	0.62	SVL	THR

Revillagigedo	Perityle socorrosensis	NE	ssp585	-66.4	0.62	SVL	THR
Revillagigedo	Psidium socorrense	NE	ssp126	17.4	0.76	G	THR
Revillagigedo	Psidium socorrense	NE	ssp370	17.4	0.76	G	THR
Revillagigedo	Psidium socorrense	NE	ssp585	17.4	0.76	G	THR
Revillagigedo	Sideroxylon socorrense	END	ssp126	17.1	0.92	G	THR
Revillagigedo	Sideroxylon socorrense	END	ssp370	17.1	0.92	G	THR
Revillagigedo	Sideroxylon socorrense	END	ssp585	17.1	0.92	G	THR
Revillagigedo	Teucrium townsENDii	NE	ssp126	-87.3	0.63	EL	THR
Revillagigedo	Teucrium townsENDii	NE	ssp370	-100.0	0.63	TL	THR
Revillagigedo	Teucrium townsENDii	NE	ssp585	-100.0	0.63	TL	THR
Revillagigedo	Tribulus cistoides	NE	ssp126	-100.0	0.80	TL	THR
Revillagigedo	Tribulus cistoides	NE	ssp370	-100.0	0.80	TL	THR
Revillagigedo	Tribulus cistoides	NE	ssp585	-100.0	0.80	TL	THR
Revillagigedo	Zanthoxylum fagara	NE	ssp126	-73.6	0.78	SVL	LC
Revillagigedo	Zanthoxylum fagara	NE	ssp370	-73.6	0.78	SVL	LC
Revillagigedo	Zanthoxylum fagara	NE	ssp585	-73.6	0.78	SVL	LC
Revillagigedo	Ilex socorroensis	END	ssp126	18.3	0.80	G	THR
Revillagigedo	Ilex socorroensis	END	ssp370	15.3	0.80	G	THR
Revillagigedo	Ilex socorroensis	END	ssp585	4.2	0.80	G	THR

Table S5.13. Model results for the linear model:

$$Net\ Range\ Change \sim Archipelago\ *\ Fut\ Model\ *\ SSP$$

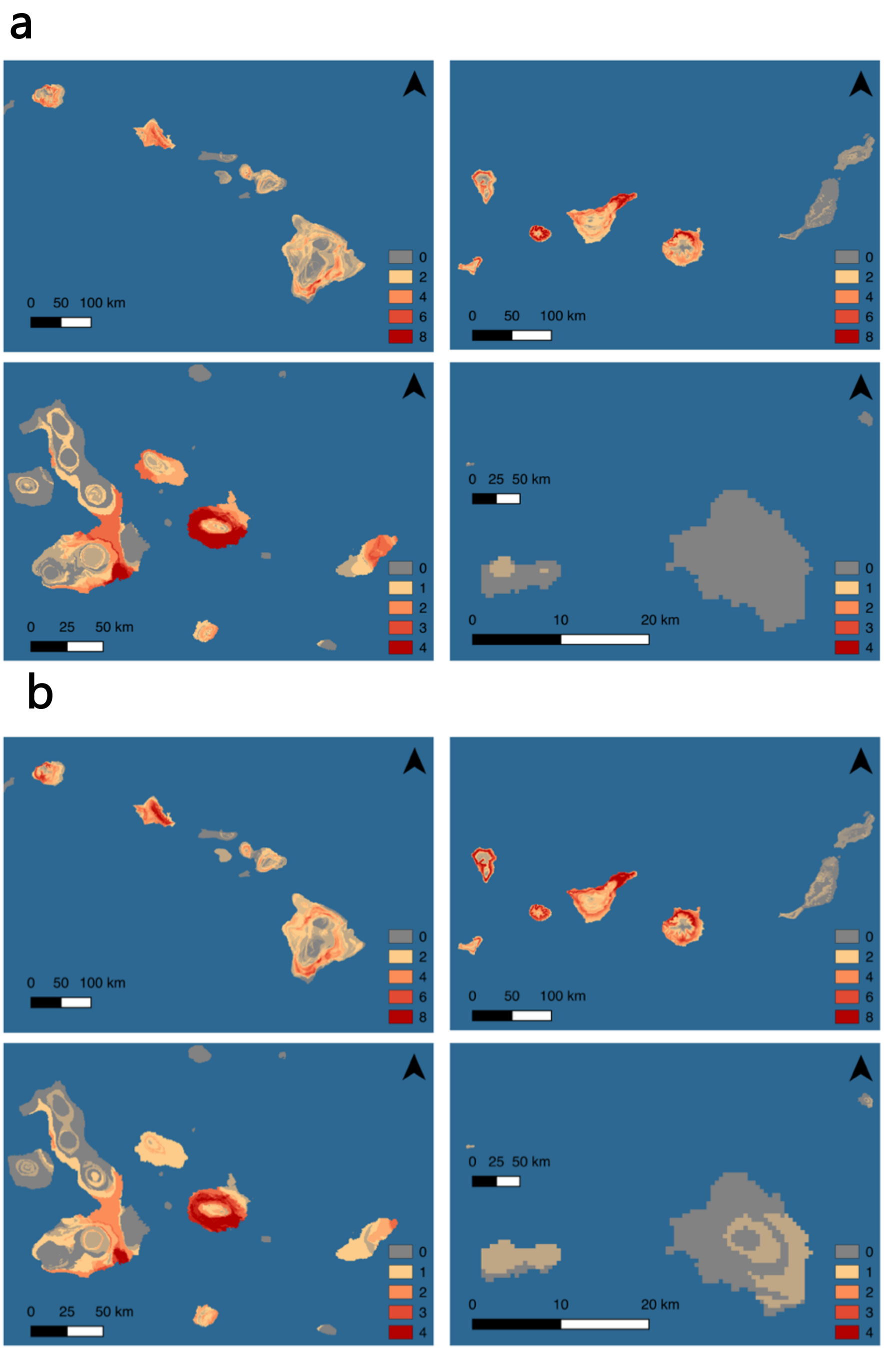
Term	Estimate	Std.error	Statistic	p.value
(Intercept)	-32.524	4.065	-8.001	2.098e-15
Island_Galapagos	3.786	6.675	0.567	0.571
Island_Hawaii	15.754	6.109	2.579	0.010
Island_Revillagigedo	-3.836	7.103	-0.540	0.589
Fut Model_UK	-3.349	5.762	-0.581	0.561
Trajectory_ssp370	-8.950	5.762	-1.553	0.120
Trajectory_ssp585	-15.552	5.762	-2.699	0.007
islandGalapagos:Fut Model_UK	-5.088	9.447	-0.539	0.590
islandHawaii:Fut Model_UK	-8.553	8.647	-0.989	0.323
islandRevillagigedo:Fut Model_UK	8.535	10.052	0.849	0.396

islandGalapagos:Trajectory_ssp370	1.717	9.447	0.182	0.856
islandHawaii:Trajectory_ssp370	-3.376	8.647	-0.390	0.696
islandRevillagigedo:Trajectory_ssp370	12.573	10.052	1.251	0.211
islandGalapagos:Trajectory_ssp585	9.788	9.447	1.036	0.300
islandHawaii:Trajectory_ssp585	3.742	8.647	0.433	0.665
islandRevillagigedo:Trajectory_ssp585	19.404	10.052	1.930	0.054
Fut ModelUK:Trajectory_ssp370	11.796	8.157	1.446	0.148
Fut ModelUK:Trajectory_ssp585	12.283	8.157	1.506	0.132
islandGalapagos:Fut Model_UK:Trajectory_ssp370	-3.756	13.365	-0.281	0.779
islandHawaii:Fut Model_UK:Trajectory_ssp370	-0.439	12.235	-0.036	0.971
islandRevillagigedo:Fut ModelUK:Trajectory_ssp370	-17.710	14.221	-1.245	0.213
islandGalapagos:Fut ModelUK:Trajectory_ssp585	-12.233	13.913	-0.879	0.379
islandHawaii:Fut ModelUK:Trajectory_ssp585	-2.810	12.235	-0.230	0.818

Table S5.1.4. ANOVA results for the linear model:

$$Net\ Range\ Change \sim Archipelago\ *\ Fut\ Model\ *\ SSP$$

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Island	3	39142.53	13047.51	6.75	0.00
Fut Model	1	262.86	262.86	0.14	0.71
Trajectory	2	13371.63	6685.82	3.46	0.03
Island:Fut Model	3	9966.97	3322.32	1.72	0.16
Island:Trajectory	6	4806.95	801.16	0.41	0.87
Fut Model:Trajectory	2	5511.53	2755.76	1.43	0.24
Island:Fut Model:Trajectory	6	6122.47	1020.41	0.53	0.79
Residuals	1963	3795673.51	1933.61	-	-



Supplementary Figure S5.1.5 Spatial patterns of species loss by archipelago for the a) SSP 126 (optimistic scenario) and b)SSP 585 scenarios (pessimistic scenario) for: Hawaii (Top left), Canaries (Top Right), Galapagos (Bottom left) and Revillagigedo (Bottom Right). For Revillagigedo, the top section shows the actual distance and size of the islands, while the bottom section provides a magnified view. Increased colour intensity represents a higher number of species lost measured as switching favourability within grid squares (500m^2).

Appendix S.5.2.1 Climatic-driven spatial patterns of species loss

In the Canaries we see, under the SSP126 trajectory, an average (i.e., the average across the IP and UK models) of 26.67% of grid squares experiences no species loss, while 18.50% lose at least one species, and 16.46% with ≥3 species loss per grid affecting only 0.09%. The negative effect increases progressively in the other two SSP scenarios, Under the SSP585 trajectory, a further reduction on average in the grid squares experiencing no species loss with 22.16%, while 17.91% lose at least one species, and 11.26% lose ≥3 species, with species losses up to 12 species affecting about 0.01% of the total grid squares.

In the Galápagos, under the SSP126 trajectory, an average of 40.90% of the grid squares experience no species loss, while 28.80% lose at least one species, and 12.96% ≥2 species. Under the SSP585 trajectory, 41.92% of grid squares experience no species loss, while 37.24% lose at least one species, and 8.30% lose ≥2 species.

In Hawaii, for the SSP126 trajectory, an average of 21.73% of grid squares experience no species loss, while 35.10% lose at least one species, and 0.11% lose >1, up to seven species. Under the SSP585 trajectory, 15.89% of the area experiences no species loss, while 28.94% lose at least one species, and 13.52% lose ≥3 species.

The Revillagigedo archipelago, under the SSP126 trajectory, an average of 98.54% of the area experiences no species loss. Under the SSP585 trajectory, the percentage drops slightly, with 80.58% of the area experiencing no species loss, while 19.42% lose at least one species.