CONSERVATION PRIORITIZATION AND ECOLOGY OF DATA-POOR MARINE SPECIES UNDER HUMAN PRESSURES

by

Xiong Zhang

B.S., Chongqing University, 2010

M.Sc., The University of Chinese Academy of Sciences, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Zoology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

January 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Conservation Prioritization and Ecology of Data-poor Marine Species under Human Pressures

submitted by Xiong Zhang in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

In The Faculty of Graduate and Postdoctoral Studies (Zoology)

Examining Committee:

Dr. Amanda Vincent
Supervisor

Dr. William Cheung
Supervisory Committee Member

Dr. Tara Martin
University Examiner

Dr. Mary O'Connor
University Examiner

Additional Supervisory Committee Members:

Dr. Christopher Harley
Supervisory Committee Member

Dr. Sarah Gergel
Supervisory Committee Member

Abstract

Securing a healthy and biodiverse ocean is vital to our human wellbeing. However, marine conservation is challenging, especially for data-poor species, whose habitats and threats are understudied. My dissertation explored how to address such challenges at two large spatial scales (in China and globally), with a focus on the little studied seahorses (Hippocampus spp.).

In the first two data chapters, I explored the utility of various sources of information about species and habitat covariates in species distribution models (SDMs). My results from the first chapter showed that local ecological knowledge provided useful biogeographic data of five Chinese seahorse species to predict their distributions, which were mainly associated with ocean temperature. My second chapter at the global scale indicated that integrating citizen sciences, museum collections, and research-grade data with continuous predictors derived the best SDM models; these models predicted reliable habitat maps for 33 out of 42 species that were primarily associated with depths, proximity to macrohabitats (e.g., sponges), pH, and ocean temperature.

In the third analytical chapter, I explored global threat patterns and conservation status for 42 seahorse species with two cumulative-human-impact (CHI) models (spatial and non-spatial) and random forest (RF) models. I found that human-impact indices (from the CHI models) can be used to predict conservation status at high accuracies (87% and 96%) in RF models. Applying a non-spatial CHI model derived indices better predicted conservation status, while using a spatial CHI model identified distribution patterns of threats.

In the fourth data chapter, I integrated the derived biogeographic and threat maps in a novel framework to set conservation priorities for seahorse habitats in China and globally, using Marxan software. I found that the two major outputs of Marxan (i.e., selection frequency and best solution) were useful to determine feasible priority solutions at large spatial scales.

My results identified valuable datasets and approaches to advancing ecological and conservation knowledge for data-poor marine species, an essential precursor to action for the ocean.

Lay Summary

Managing and protecting marine life requires an understanding of where species live, what threatens them and their habitats, and what we might do to help them. Seahorses are particularly charming examples of the many marine species that are little understood. I set off to study seahorses in China, where they are heavily fished, and around the world. I found we could learn most by integrating information from diverse sources, including local fishers' knowledge, citizen science (e.g., divers' observations), museum collections, and scientists' knowledge and data. We then applied novel computer techniques (e.g. machine learning models and decision-making software) and considered our results from both natural and social science perspectives. This helped us find out what seahorses are threatened and allowed us to plan possible conservation action, particularly for deciding where to urge creation of protected areas. Our new methods are of value to many species and spaces.

Preface

This dissertation contains four scientific papers of which I am or will be the first author. I designed the framework, scope, and question of each research chapter with my supervisor, Dr. Amanda Vincent. I was primarily responsible for conducting the research, including project planning, field work and data gathering, development of methodology, data analysis, interpretation of results, and manuscript writing. My supervisor, Dr. Amanda Vincent, guided me through all these processes, especially helping me with grant applications, project management, field work, and drafting and refining the manuscripts. My committee members, Dr. William Cheung, Dr. Sarah Gergel, and Dr. Christopher Harley also provided intellectual feedback that helped me refine my research methods and manuscript writing. For Chapter 2, Dr. William Cheung introduced me to local colleagues in China who facilitated my field work. This field work was assisted by many Chinese colleagues (see Acknowledgements). For Chapter 3, some undergraduate and graduate students at UBC worked with me to gather seahorse data from peer-reviewed literature (see Acknowledgements).

My Chapters 2 and 3 have resulted in two publications:

- Chapter 2: Zhang, X. and Vincent, A.C. (2017) Integrating multiple datasets with species distribution models to inform conservation of the poorly-recorded Chinese seahorses. *Biological Conservation*, 211, 161-171. I conceived the preliminary idea, conducted the field work, analyzed the data, and wrote the manuscript. Dr. Amanda Vincent helped me refine the idea and the manuscript.
- Chapter 3: Zhang, X. and Vincent, A.C. (2018) Predicting distributions, habitat preferences and associated conservation implications for a genus of rare fishes, seahorses (*Hippocampus* spp.). *Diversity and Distributions*. https://doi.org/10.1111/ddi.12741. I conceived the preliminary idea, conducted the research and wrote the manuscript. Dr. Amanda Vincent helped me refine the idea and the manuscript.

Chapters 4 and 5 will be submitted to leading journals as two co-authored papers with Dr. Amanda Vincent.

The field work in my dissertation was approved by UBC’s Animal Care Committee Office of Research Service (A12-0288-A011) and UBC’s Human Behavioural Research Ethics Board (H12-02731).

Table of Contents

List of Tables

List of Figures

List of Symbols

CHI_j – Cumulative human impact of all stressors upon species $\boldsymbol{j}$

$CHI_{j,m}$ – Cumulative human impact of all stressors on species $\boldsymbol{j}$ in location $\boldsymbol{m}$

$ED_{i,j}$ – Exposure degree of species $\boldsymbol{j}$ to stressor $\boldsymbol{i}$

$ED_{i,j,m}$ – Exposure degree of stressor $\boldsymbol{i}$ to species $\boldsymbol{j}$ in location $\boldsymbol{m}$

$F_{i,j}$ – Frequency of the stressor $\boldsymbol{j}$ occurring to species $\boldsymbol{i}$

$HI_{i,j}$ – Human-impact index of stressor $\boldsymbol{i}$ upon species $\boldsymbol{j}$

$HI_{i,j,m}$ – Human impact of stressor $\boldsymbol{i}$ on species $\boldsymbol{j}$ in location $\boldsymbol{m}$

$I_{i,j}$ – Intensity of stressor $\boldsymbol{i}$ across species $\boldsymbol{j}$'s habitat

$I_{i,j,m}$ – Intensity of stressor $\boldsymbol{i}$ to species $\boldsymbol{j}$ in location $\boldsymbol{m}$

$P_{j,m}$ – Presence probability of species $\boldsymbol{j}$ in location $\boldsymbol{m}$

$PS_{i,j}$ – Proportion of the species $\boldsymbol{j}$'s habitats experiencing the stressor $\boldsymbol{i}$

$S_{i,j}$ – Sensitivity (degree of intolerance and incapability of recovery) of species $\boldsymbol{j}$ to stressor $\boldsymbol{i}$

List of Abbreviations

AOO – Area of occupancy

AUC – Area under the curve of the receiver-operating characteristic plot

CBD – Convention on Biological Diversity

CD – Coastal development

CHI – Cumulative human impact

CITES – Convention on International Trade in Endangered Species

CS – Citizen sciences

GPAs – Greater-protection areas

IUCN – International Union for Conservation of Nature

LEK – Local ecological knowledge

LPAs – Lower-protection areas

MC – Museum collections

MCP – Marine Conservation prioritization

MPAs – Marine protected areas

PRL – Peer-reviewed literature

RG – Research grade data

SDMs – Species distribution models

SS – Species sightings

TCM – Traditional Chinese Medicine

Acknowledgements

I offer my most sincere and enduring gratitude to my PhD supervisor, Dr. Amanda Vincent, who have guided me through the whole journey of my growth over the past five years. It is under the shining of her wisdom and her supervision that I have become a better scientist as well as a better conservationist. I am so lucky and honored to have been your student. Thank you to my committee members – Dr. Sarah Gergel, Dr. William Cheung, and Dr. Christopher Harley. Your wisdom and intellectual support to my thesis helped me grow as a scholar. I am also grateful to IOF and UBC for providing such a beautiful environment and diverse society that inspired my research and colored my life as an international student.

I received financial support from Disney Conservation Fund and Guylian Chocolates to conduct my field work and conservation outreach in China. I owe particular thanks to the China Scholarship Council for offering me a four-year scholarship that supported my living in Vancouver.

I am very grateful to the wonderful Chinese colleagues who have provided great help for my research and conservation outreach in China. I would like to thank Dr. Shan Xiujuan, Dr. Du Jianguo, Dr. Wang Jiangyong, and Dr. Lin Qiang for helping me select sampling sites and familiarize local context. I thank Dr. Jiang Jingzhe, Dr. Wang Jiangyong, Dr. Wang Shifeng, Mr. Lin Haisheng for introducing me to local fishers. I sincerely thank Dr. Zhang Chunguang, Dr. Liu Jing, Dr. Chen Zuozhi, Dr. Yan Bing, Ms. Chen Zhu for providing local data about seahorse specimens, marine fisheries and planning. I thank Dr. Guo Huayang, Mr. Dai Fangqun in assisting interviews with fishers. I also thank Dr. Qing Gen for hosting my visits and exchanging research ideas on seahorse populations in China.

My sincere thanks to current and previous colleagues in Project Seahorse and Institute for the Oceans and Fisheries (IOF), and to my research assistants at UBC. Dr. Lindsay Aylesworth, Dr. Jennifer Selgrath, Dr. Ting-Chun Kuo, thank you so much for sharing your experiences in doing a PhD and helping me review my papers with tremendous constructive comments. Dr. Sarah Foster, thank you for helping me conduct conservation

outreach in China and being so nice to support my PhD study in many ways. Tanvi Vaidyanathan, my best buddy, thank you for cheering me up with your friendship, optimistic spirit and incredible humor. My life could have lost lots of joy without you being aside. Thank you to other Project Seahorse fellows and IOF colleagues, including Kyle Gillespie, Clayton Manning, Emilie Stump, Allison Stocks, Ravi Maharaj, Muhammed Oyinlola, Tayler Clarke, Dr. Oai Li Chen, Dr. Colette Wabnitz, for their help with reviewing my manuscripts and developing research methods. I also thank the Project Seahorse Team, previous and current, including Scott Finestone, Gina Bestbier, Riley Pollom, and Lily Stanton. My research could take much longer time without the volunteering assistants from Francesca Fogliata, Kelly Wang, Marie McCallum, Myra Lam, Naman Siddique, Sammi Conard, and Xinxin Xue.

Special and deepest thanks are owed to my grandparents and parents, who have supported me throughout my years of oversea study, both morally and financially. Thank you to my younger sister for taking my portion of responsibility on her shoulders to take care of my grandparents. Thank you to my uncle for supporting me to study abroad and always being my role model. Thank you to my girlfriend, Doctor Li Na, who always supports me with her whole heart and brings sunshine into my life!

Dedication

With deepest love and gratitude, this dissertation is dedicated to my parents, Zhang Xianmei and Zhang Yuan, my grandparents, Zhang Guishui and Shi Yinjiao. Your unwavering love gave me the strength to make this dissertation possible. Thank you for encouraging me to follow my passion and dream to become a scientist.

Chapter 1: Introduction

1.1 Rationale

As oceans face unprecedented pressures, conservationists urgently need to know about organisms whose populations, habitats, and threats are understudied. Over the past few centuries, intensified human activities have imposed enormous pressures on the oceans (Halpern et al. 2008, McCauley et al. 2015). These pressures have gradually degraded the health of the oceans, and thus depleted the benefits (e.g., food provision) that oceans have given us since time immemorial (Halpern et al. 2012). For instance, we have overfished 31% of commercially important fish stocks in the ocean (FAO 2016). Non-selective fishing gears (e.g., bottom trawlers, purse sines) have been particularly problematic because of their high bycatch of non-target organisms, such as marine turtles, fishes, and many other organisms (Alverson 1994, Lewison et al. 2004). In addition to fishing, pollutions, habitat destruction, invasive species, and climate change are collectively making marine conservation very challenging. Currently, thousands of marine species can be considered as data poor given the few data available about their populations, habitats, or threats that can help determine conservation status or action plans (McCauley et al. 2015). Such data paucity has long been a notable concern in marine conservation.

In this thesis, I illustrate new ways to explore species distributions, threats, and priority habitats for data-poor marine organisms, with ecological and conservation perspectives. Understanding species distributions and habitat/environmental drivers is key to ecological studies and wildlife conservation (Brooks et al. 2002, Hanski 2011). Determining threatened status and underlying anthropogenic stressors is also fundamental for conservation sciences (Halpern et al. 2008, Davidson et al. 2012). However, revealing the distribution and threat patterns demands considerable data and is generally very challenging for understudied species. To address this challenge, I examine the utility of multiple types of datasets (including stakeholders' knowledge) in analyzing species distribution patterns and threatened status for data-poor species. To explore the value of these studies for guiding conservation practices, I apply my newly acquired knowledge to identify priority areas/habitats for data-poor organisms. It is my intention that such prioritization can be used to help reconcile marine conservation and economic use (Margules and Pressey 2000, Aswani et al. 2018).

I am particularly interested in case studies of seahorse species (*Hippocampus* spp.) at large spatial scales. Seahorses are flagship organisms for marine conservation (Foster and Vincent 2004). They are a genus of sedentary fishes living in coastal waters, which are at the front line of exposure to human impacts. Although conservation actions are urgently needed, seahorse populations are understudied in most regions (Vincent et al. 2011), with no papers published on the ecology or distribution of most species. In my thesis, I first conduct a national scale biogeographic study on seahorse populations in China (Chapter 2), where I come from. I then zoom out to the global scale to examine distribution patterns (Chapter 3) and threat patterns (Chapter 4) for different seahorse species. Such large geographic coverage is not only essential to generalizing ecological patterns (e.g., species-habitat relationships) (MacArthur 1984, Rahbek 2005), but also vital to obtaining a holistic profile of human impacts upon different populations/species. Based on these studies, I then set conservation priorities for Chinese seahorse populations and for all seahorse species around the world (Chapter 5).

1.2 Background

To stem biodiversity loss with constrained resources, conservation prioritization (a.k.a., conservation planning) is an essential approach that is used worldwide (Margules and Pressey 2000). Human activities (e.g., overexploitation) are causing a biodiversity crisis with species extinction rates up to 1000 times higher than background (Pimm et al. 1995). Populations of plants (e.g., tropical forests) and animals (e.g., amphibians, mammals) continue to decline both in terrestrial and aquatic ecosystems (Bland et al. 2015, McCauley et al. 2015). Facing this challenge, scientists have developed various approaches and advice to guide decision makers (e.g., government leaders) to take actions (Gullison et al. 2000, Wilson et al. 2009). As conservation resources (e.g. funds) are usually limited (James et al. 1999, Waldron et al. 2013), resource-use prioritizing is crucial in many conservation initiatives. This prioritization includes selecting populations/species (e.g., umbrella species, flagship species) under threat to be preferentially protected, and choosing priority habitats to protect species *in situ* ((Roberge and Angelstam 2004, Brooks et al. 2006). My dissertation focus on the latter, which is known as conservation prioritization/planning. It is a vital approach in designing protected areas to meet a given conservation target (e.g., a certain amount of habitats) with minimum costs (Margules and Pressey 2000).

To inform conservation planning, we first need to address at least four essential questions: (1) Which species are threatened and where are they? (2) Which human-induced stressors drive the species to be threatened? (3) Where are the populations impacted the most and the least? (4) How can we integrate socioeconomic costs in conservation prioritization? The first question may emerge from field observations (e.g., population decline) and is fundamental to determining conservation targets (Joseph et al. 2009). The second question then arises to guide conservation measures (e.g., threat mitigation programs) for threatened species (Maxwell et al. 2013). To better do so, we should know where populations are the most and the least threatened (i.e., the third question). Locating the most threatened populations could inform where conservation measures are urgently needed (Maxwell et al. 2013). Often the least threatened populations may be preferred in conservation planning as they imply higher probability of population recovery with potentially minimum cost (Klein et al. 2013). However, other socioeconomic costs (e.g., land/sea uses) should also be considered in conservation prioritization to ensure that the proposed priorities will be well supported by stakeholders (Ban and Klein 2009).

Addressing the above four questions is challenging but urgently needed for many organisms, whose populations and conservation context are understudied. First, we need solid data to assess threatened status (a.k.a., conservation status). These data may include population’s demographic information (e.g. population size and trend), species distribution maps, and knowledge about potential threats (Mace et al. 2008, Davidson et al. 2012). Such data are currently unavailable for 12,464 animal species (i.e., Data Deficient species) on the IUCN Red List – 18.5% of all assessed animals (IUCN 2018). Second, we need data to understand why certain species are threatened. Different species can be threatened because of a variety of human-induced stressors, including overexploitation, habitat destruction, pollution, climate change, and invasive species (Gibbons et al. 2000, Halpern et al. 2008).

Third, we need data about the distribution of species and stressors – and about how species respond to stressors - to understand where species are most/least threatened (Crain et al. 2008). Such studies have only been done for certain species in some data-abundant regions, such as Australian continent (Evans et al. 2011), and Mediterranean Sea (Coll et al. 2012). Forth, we need good spatial data to estimate socioeconomic costs if conservation prioritization is to be meaningful (Ban et al. 2009). This is especially true when the spatial scale is very large and data quality may vary from place to place (Agardy et al. 2011). Nevertheless, most data-poor species are predicted to be threatened, creating a huge need for conservation actions in many data poor situations (Bland et al. 2015).

Facing the challenges of data paucity, scientists have been learning how to tap into multiple data sources and modeling approaches. These data sources may include local/traditional ecological knowledge (Huntington 2000, Drew 2005), citizen science (Silvertown 2009, Bonney et al. 2014), museum collections, unpublished data, and expert knowledge (Martin et al. 2012, Bennett et al. 2017). Local ecological knowledge refers to the knowledge acquired by local citizens (e.g., hunter, fisher) through interactions with their environment and wildlife therein (Huntington 2000). It can be very useful in monitoring local population abundance (Anadon et al. 2009, Parry and Peres 2015), and mapping habitats (Bergmann et al. 2004, Aylesworth et al. 2017). Citizen science is the scientific research conducted partly or wholly by people who are generally not professional scientists (Silvertown 2009). Citizen-science initiatives (e.g., iNaturalist) can be cost-effective for collecting biogeographic data both on land and in the ocean (Silvertown 2009, Thiel et al. 2014), assuming careful attention to data quality. Museum collections and unpublished data can provide historical and up-to-date information for data-poor species (Suarez and Tsutsui 2004, Tweh et al. 2015). Expert knowledge can be vital to evaluating species vulnerability to human-induced stressors for data-poor species (Cheung et al. 2005, Maxwell et al. 2013). On the other hand, computer modeling techniques (e.g., machine learning) can empower scientists to deal with data-poor situations (Elith et al. 2006, Phillips et al. 2006). For instance, species distribution models can be very useful to derive spatially-explicit maps for species with limited occurrences (Franklin 2010).

Marine conservation planning lags far behind its terrestrial counterpart and is more challenging (Turpie et al. 2000, Lourie and Vincent 2004). Marine ecosystems – especially the near-shore regions – have been facing increasing threats from land- and ocean-based human pressures (Halpern et al. 2008, Halpern et al. 2015). However, we generally know little even about many of the abundant marine species. Only recently (since ~ 1960s) have we gained technological capabilities to observe marine organisms in the oceans, and to obtain essential knowledge about their populations and habitats (Lourie and Vincent 2004, Selig et al. 2014). Human-derived threats to marine ecosystems and ocean health are increasingly recognized and evaluated (Halpern et al. 2012, Martin et al. 2012). To date, however, initiatives towards setting marine priorities have only been conducted for some well-known marine animals, such as commercial fishes (Turpie et al. 2000, Fox and Beckley 2005), sea turtles (Wallace et al. 2011), and sharks (Lucifora et al. 2011) at the global or regional scales. There is little prioritization research specifically focused on data-poor marine species. Although ecosystem-based management is becoming the notional mainstream of fisheries management and conservation, conservation planning for focal species (e.g., flagship species) deserves no less concern than community- or ecosystem-level research (Olsoy et al. 2016).

As a Chinese national, I feel that China tends to face greater challenges than many developed maritime nations in marine conservation. First, marine biological research only began in the 1920s in China (Wang and Nie 1932, Xiang 2003), while in the Europe this can be traced back at least two centuries earlier to the time of Captain James Cook (1728 – 1779). Moreover, marine biogeographic data are still rarely documented or published in China (Liu 2013), while such information are often much better recorded and open-access in many developed countries (Boakes et al. 2010). Second, compared with the developed western countries, marine conservation in China can be more severely challenged by the socioeconomic context. This includes the lower social-wellbeing status, higher demand for economic growth, and higher livelihood dependency of local communities on fisheries (Caldwell and Vincent 2013, Han et al. 2017). As a result, China is facing great challenges in marine conservation initiatives such as the implementation of marine protected areas (Qiu et al. 2009).

Within this thesis, I address the four conservation questions for marine data-poor organisms in China and at the global scale. More specifically, I conduct studies to explore the utility of various biogeographic datasets (e.g., local ecological knowledge, citizen science) in predicting species distributions and species-habitat relationships. I then integrate expert knowledge and spatial data to identify major stressors and map cumulative human impacts for data-poor species. These results are used to set conservation priorities for data-poor species in different socioeconomic scenarios. My goal is to generate both ecological and practical knowledge to guide marine conservation planning for data-poor species.

1.3 Case Study

My thesis uses seahorses (*Hippocampus* spp.) to examine pragmatic approaches in deriving conservation knowledge for data-poor marine species. At the beginning of my research (in 2016), only 42 species were considered as valid (Lourie et al. 2016; Zhang et al. 2016), thus this number was kept consistently throughout the thesis. But it should be noted that a new species was published in 2017 (Han et al. 2017), and there are still debates on taxonomy of serval species complex (e.g., *Hippocampus kuda*, see details in Lourie et al. 2016) and new species might emerge in the future. As a case study, I investigate Chinese seahorse populations at a national scale and the 42 species at the global scale.

Seahorses are a genus of data-poor fishes, whose populations, habitats, and threats are generally understudied for most species (Vincent et al. 2011). Seahorse species live in the world’s shallow seas (typically < 30 m in depth) from the temperate to the tropical zones (Foster and Vincent 2004). These small sedentary fishes are masters of camouflage and locally rare with low population density (Foster and Vincent 2004; Vincent et al. 2011). Such biological traits rendered them difficult to find or study in the field (Aylesworth et al. 2017). As a result, population demographic data (e.g., population size, growth rate) and habitat knowledge were only known for a handful of species at small spatial scales (e.g., a lagoon or an estuary) (Vincent et al. 2011, Caldwell and Vincent 2013, Cohen et al. 2017). There were no spatially-explicit maps of distributions for seahorse species, although we roughly inferred their distribution ranges (see geographic range maps in Lourie et al. 2016). Among the 42 species, 14 species were considered as threatened, 11 species were not threatened, but 17 species were “Data Deficient” according to the IUCN Red List assessments (IUCN 2018). Major known threats included fishing using non-selective gears (e.g., bottom trawlers, purse seines) and habitat degradation (Vincent et al. 2011; IUCN 2018). Nevertheless, there were no maps to demonstrate human impacts for each species.

Using seahorses as the case study is meaningful to advancing marine conservation. First, seahorses' biological traits render them flagship species in promoting marine conservation around the world (Vincent et al. 2011, Harasti et al. 2014). Seahorses display a fantastic diversity of attractive looks, mate fidelity (for most species), and unique male pregnancy (Foster and Vincent 2004; Lourie et al. 2016). Their low fecundity, mate fidelity, small home range, low population density, and inshore residence can render them susceptible to human activities (Foster and Vincent 2004). Seahorses are among the most attractive denizens of many vital marine ecosystems such as seagrass beds, estuaries, mangroves, and coral reefs (Foster and Vincent 2004). Protecting seahorse might benefit many other species living in the same habitats (Vincent et al. 2011). Second, there is some suggestion that seahorse populations may be able to recover well if the impacts are not severe and removed, given their rapid growth rate, early age at maturity, and short generation time (Curtis and Vincent 2006). Such an optimistic outlook may encourage marine conservation actions. Third, seahorses are found, fished and traded around the world for traditional medicines, curios, and aquarium fishes (Vincent et al. 2011, Foster et al. 2016). Global concerns about the unsustainable trade led to seahorses being the first fully marine fishes listed on Appendix II of the Convention on International Trade in Endangered Species (CITES) in 2002 (effective since 2004) (Vincent et al. 2011). Such a worldwide visibility and relevance makes seahorses an excellent 'story-teller' for global marine conservation.

China faces enormous challenges in marine conservation and fisheries management along with its rapid development. China has grown to be the second largest economic power (GDP ~ 11 trillion USD in 2016; data.worldbank.org) through unprecedented economic development over past four decades. However, such a long-term rapid growth has included overexploitation of China's marine resources and accelerated the degradation of coastal ecosystems (He et al. 2014). Nowadays China's marine systems provide ecosystem services to approximately half of the nation's population and 45 of the 60 major cities (He et al. 2014), while they are also the home to over 22,000 known marine species (Liu 2013).

Nearly 6% of the coastal provinces' GDP (2010) came directly from marine industries including fisheries, tourism, transportation, and oil/gas production (SOA 2010). Approximately half of the nation's domestic marine catch comes from trawlers (Goldstein 2013, Shen and Heino 2014), which have been blamed for the decline of body sizes and trophic levels of marine fish (He et al. 2014). Water pollution has caused harmful algal blooms (e.g. red tides) which are nearly ten times more frequent than in the 1980s (He et al. 2014). Coral coverage in the South China Sea has decreased to less than 15% of that four decades ago (Zhao et al. 2012, He et al. 2014). Facing these and other challenges, China has gradually established many well-intentioned laws and policies to stem marine overexploitation and ecosystem degradation. These include fisheries management policies (e.g. 'Zero Growth') to sustain marine catch (see a review by Shen and Heino 2014), marine protected areas and conservation planning (e.g., ecological redline) to prevent overuse of coastal habitats (Qiu et al. 2009, Lu et al. 2015). However, a lack of local political will – aided by a dearth of scientific information - has limited China's performance in marine conservation (Liu 2013; Lu et al. 2015).

Chinese seahorses were among the most poorly known seahorse populations around the world, despite being heavily exploited. China is considered as the largest consumer of dried seahorses (Vincent et al. 2011). Seahorses are distinguished by their heavy use (in trade amounting to tens of millions of individual seahorses per year) in Traditional Chinese Medicine (TCM), which accounts for about 95% of seahorses in trade (Vincent et al. 2011, Lawson et al. 2017). The use of seahorses in TCM can be traced back to about 2,000 years ago (Tang 1987, Zhu and Woerdenbag 1995), and is still quite common in China especially the southern coastal areas (e.g., Hong Kong and Guangdong). However, formal seahorse population and conservation studies were rare in China. Most studies focus on seahorse evolution, medical functions, and aquaculture. Six seahorse species were purportedly present, and all were considered as threatened by China's Red List based on the same criteria developed by IUCN (Wang and Xie 2009). One of them, *H. kelloggi* (Great seahorse), is on China's List of Wildlife under National Protection (MEP 2002). The other four have been proposed for addition to the List (Zhang Chun-Guang, *per. comm.*).

1.4 Context and Collaborators

My study is supported by the Project Seahorse (www.projectseahorse.org) team, based at the University of British Columbia and the Zoological Society of London (UK). Project Seahorse is an interdisciplinary and international organization committed to conservation and sustainable use of the world's coastal marine ecosystems. We engage in connected research and management at scales ranging from community initiatives to international accords. Collaborating with stakeholders and partners, we use seahorses to focus our efforts in finding marine conservation solutions. We serve as the IUCN Seahorse, Pipefish and Stickleback Specialist Group (www.iucn-seahorse.org). Project Seahorse has engaged in China intermittently since 1993.

In China, my research was supported by many organizations, including academic institutions, local governments, fishing communities, NGOs, etc. Colleagues from the South China Sea Fisheries Research Institute, Chinese Academy of Fisheries Sciences (CAFS) provided the major support to help me reach to many local fishers in South China Sea. Colleagues from the Yellow Sea Fisheries Research Institute, CAFS helped me establish contacts with local fishers in the East China Sea and Yellow Sea. I provided more detailed information about these collaborators in thesis Chapter 2 and acknowledged their contributions to my research.

1.5 Research Questions

In this thesis, I address four major questions related to conservation research on data-poor marine species. These questions are as follows:

1) How is local fishers' ecological knowledge useful for mapping distributions for data-poor marine species? (Chapter 2)
2) How can we derive useful species occurrences and habitat variables for identifying suitable habitats of data-poor marine species at the global scale? (Chapter 3)
3) How can we estimate cumulative human impacts and threatened status for data-poor marine species? (Chapter 4)

4) How can we identify priority areas for data-poor marine species at large spatial scales? (Chapter 5)

1.6 Thesis Outline

This dissertation contains the introduction (Chapter 1) and four research chapters, followed by a general discussion about conservation implications and research suggestions for data-poor marine species. The four research chapters are highly organized to address the four research questions (Fig. 1.1).

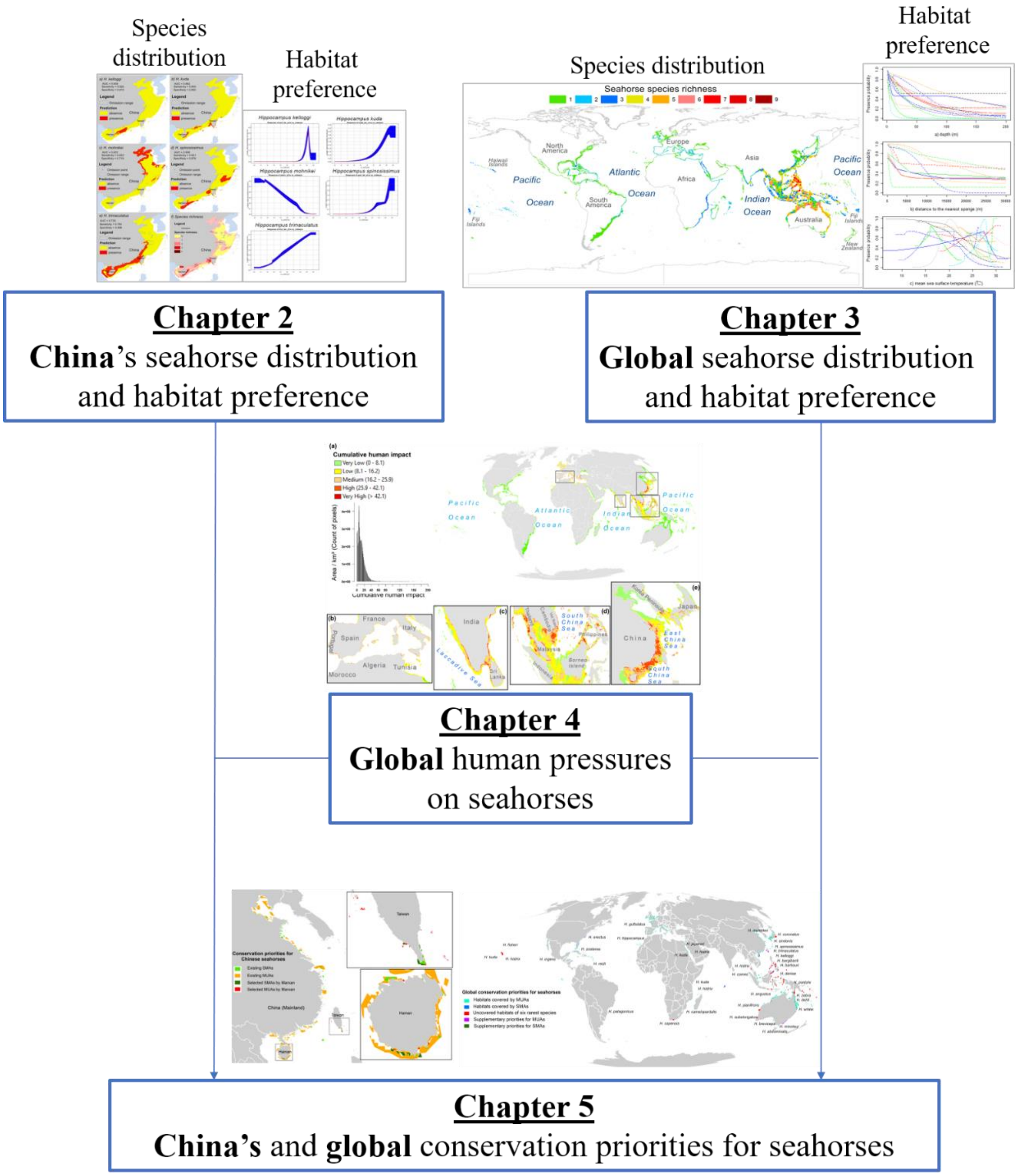


Figure 1. 1 Schematic outline of this thesis.

1.6.1 Chapter 2: Integrating multiple datasets with species distribution models to inform conservation of the poorly-recorded Chinese seahorses

In Chapter 2, I examine the value of local ecological knowledge in determining distributions of Chinese seahorse populations. My interest is to answer four specific questions: 1) How different is local fishers' knowledge compared with researchers' observations? 2) How can we integrate coarse-grain species data in species distribution models? 3) How does the distribution pattern and species-habitat relationship differ among seahorse species? To answer these, I gather species data from global sources, Chinese peer-reviewed literature, and local fishers' ecological knowledge. I use species distribution models (SDMs) to examine their quality in predicting presences for each species. Among the many marine species poorly studied in China, seahorses are highly valued and used as traditional medicine in large numbers by Chinese people. However, we still know little about the species composition and distribution ranges of these valuable species in China's seas. In this chapter, I undertake the first systematic study to understand the species composition, geographic distributions, and ecological traits of Chinese seahorse populations. To fill knowledge gaps, I deploy semi-structured interviews to gather local ecological knowledge from Chinese fishers – an approach that has rarely been used in China.

1.6.2 Chapter 3: Predicting distributions, habitat preferences, and associated conservation implications for a genus of data-poor species, seahorses (*Hippocampus* spp.)

In this chapter, I examine suitable habitats and habitat preferences for seahorse species at the global scale. Specifically, I address four questions: 1) How does citizen science and museum collections contribute to map suitable habitats for data-poor species? 2) What are the useful habitat variables correlated with seahorse species? 3) How can we apply species distribution models in assessing conservation status for data-poor species? Although data-poor species are by its definition understudied, there are some piecemeal occurrences hidden in existing databases such as the Ocean Biogeographic Information System, unpublished studies, peer-reviewed literatures, and citizen sciences. Collating species data from such sources is needed to maximize data availability and minimize errors, although their quality might vary in resolution and accuracy. Moreover, habitat data (e.g. locations of coral reefs) might be essential predictors for modeling distribution of marine species. The appropriate use of both species and habitat data can be crucial to model performance and species-habitat relations. My research in this chapter represents the first attempt to explore the above issues for a genus of data-poor marine fish species (i.e. seahorses).

1.6.3 Chapter 4: Cumulative human impact models reveal threat patterns of seahorses (*Hippocampus* spp.)

In Chapter 4, I develop quantitative models to evaluate cumulative human impact (CHI) for seahorses, and use the result to estimate their threatened status at the global scale. In undertaking this research, I am interested in answering three major questions: 1) Can linear-additive models be useful to quantify CHI for data-poor seahorses? 2) Are human-impact indices based on CHI models useful to predict conservation status for seahorses? 3) What are the major stressors and where seahorses are threatened the most/least? Threat status is poorly known for thousands of data-poor marine species, despite the IUCN's efforts. Studies on cumulative human impact (CHI) might provide a useful approach to infer threats for such data-poor marine organisms. The CHI research assesses cumulative human impact based on the intensity of human stressors and expert knowledge about the vulnerability of the species to each stressor. Previous CHI studies have mainly focused on mapping human impacts on marine habitats, with litter attention given to score CHI for particular species. Theoretically, species with higher CHI scores are more likely threatened by human activities. However, assessing CHI is challenging since species' responses to stresses may be non-linear and interactions between stressors are common in nature (i.e., additive, synergistic, antagonistic effects). Given these are difficult to determine for most marine species, previous studies often utilize linear-additive models. For this chapter, I develop linear-additive models to quantify CHI to address the three research questions for seahorses.

1.6.4 Chapter 5: Conservation prioritization at large spatial scales for seahorses (*Hippocampus* spp.)

In Chapter 5, I combine the research results from Chapters 2 – 4 to identify conservation priorities for seahorses in China and around the world. I am mainly interested in exploring three questions at large spatial scales: 1) How can we incorporate socioeconomic data in marine conservation planning? 2) How can we derive priority solutions to inform conservation actions? 3) Where are the priority areas for global and Chinese seahorses? Marine protected areas (MPAs) play a vital role in addressing fishing pressures on marine ecosystems and species. However, MPAs are frequently challenged by local communities because of the dearth of socioeconomic considerations included in designing MPAs. Conservation prioritization or planning at large spatial scales is challenging, although there is an emergent demand to create large-scale networks of MPAs. There is also a geographic bias on marine conservation prioritization at regional scales, with little research has been conducted in developing countries such as China, which faces one of the highest levels of cumulative human impact around the world. The extent of spatial gaps between present MPAs and priority areas for marine species almost remains unknown, although both areas are expanding. In the final research chapter, I develop a conservation prioritization framework for marine species at large spatial scales. Based on this framework, I incorporate socioeconomic costs with seahorse and human-impact maps (from the above chapters) to conduct conservation prioritization in China and globally.

1.6.5 Chapter 6: Discussion

In this final chapter, I end with a general discussion of findings from my thesis, their conservation implications for data-poor marine species, and future research directions.

Chapter 2: Integrating multiple datasets with species distribution models to inform conservation of the poorly-recorded Chinese seahorses

2.1 Summary

Modeling and mapping species distributions are vital to biodiversity conservation but are challenging for data-limited species whose localities are poorly recorded. Here I examine the utility of three datasets and species distribution models in conservation of seahorses (*Hippocampus spp.*), a genus of poorly-recorded marine fishes. I collated occurrences from field data of species sightings (SS), peer-reviewed literature (PRL), and fishers' local ecological knowledge (LEK) for five seahorse species in China. I modelled seahorse distributions using different combinations of these datasets. I first compared model performance and predictions between PRL and LEK, and then evaluated the impact of adding LEK and/or PRL to SS. My results indicated that LEK provided higher-resolution maps than PRL and tended to generate slightly better models. There is more predictive consistency between LEK and PRL on presence-probability maps than on presence/absence maps. Adding LEK and/or PRL to SS improved model performance across species. My study suggests that integrating LEK (and PRL) and limited SS with species distribution models can usefully inform conservation for poorly-recorded species.

2.2 Introduction

Species distribution maps are vital to biodiversity conservation (Pimm et al. 2014). Anthropogenic activities have driven incredible biodiversity loss and ecosystem degradation, which in turn has significant impact on human society. To protect the threatened wildlife, we need biogeographic information to assess their conservation status (Mace et al. 2008), and design nature reserves (Lourie and Vincent 2004, Micheli et al. 2013). Wildlife habitat maps are also indispensable for resource management, as new development projects expand across land and the sea (McShane et al. 2011, Reis et al. 2012).

Mapping species distributions is challenging for poorly-recorded species, whose population localities are poorly documented in peer-reviewed literature or other sources. This difficulty often necessitates the use of multiple datasets, including new field data. Fine-resolution (e.g. 10 x 10 m^2) species sightings (SS, in the form of GPS coordinates) from natural history collections or other sources (e.g. citizen science) are the most frequently-used datasets. But SS collection is often biased towards easily-accessed regions and common taxa (Phillips et al. 2009, Robinson et al. 2011). Peer-reviewed literature (PRL) can be a second dataset, but it may only contribute coarse range maps for poorly-recorded species. A third source of species data is local ecological knowledge (LEK), which refers to the knowledge system learnt by people through interactions with their local environment (Berkes 1993). Compared with traditional surveys (e.g. transect sampling), interview-based LEK research can generate cost-effective but often coarse-resolution (e.g. 10 x 10 km^2) datasets (Carter and Nielsen 2011, Laze and Gordon 2016).

Species distribution models (SDMs), which predict presence probability of focal species based on limited species presences/absences and environmental data, might provide a powerful way to derive spatially-explicit maps and to inform conservation for poorly-recorded species (Guisan and Thuiller 2005, Franklin 2010). The predictive maps based on SDMs have facilitated population surveys for rare species (Guisan et al. 2006, Stirling et al. 2016), and are useful for conservation planning (Guisan et al. 2013). Some SDMs contain techniques to examine species-habitat relationships, which are central to ecology (Guisan and Thuiller 2005). In literature, there are basically two types of SDMs regarding the availability of species-absence data: presence-absence models, and presence-only models (see Franklin 2010 for a review). Presence-only models are more suitable to poorly-recorded species since their absences are hard to determine.

Mapping and modeling species distributions is particularly challenging for poorly-recorded marine species. Marine biota and environmental surveys have historically fallen behind the terrestrial counterparts (Costello et al. 2010). Scuba-diving has only been used for collecting site-level species data since ~ 1960s (Caddy 1968), and remote sensing techniques have only contributed spatial data for marine environments since 1980s (Bernstein 1982, Wentz and Schabel 2000). The utility of survey techniques (e.g.

underwater visual census) can be restricted by the unique features of marine environment (water clarity, depth, etc.). These characteristics of ocean systems make it more difficult to study geographic distributions of marine organisms.

Seahorses (*Hippocampus* spp.) provide a typical example of rarely-recorded marine organisms whose distributions are difficult to determine. These relatively rare, cryptic, and small fishes are difficult to detect or survey (Vincent et al. 2011; Aylesworth et al. 2017). Additionally, seahorses can raft with holdfasts (e.g. seaweed) and disperse over long distances, although they are generally stationary (Lourie et al. 2005, Caldwell and Vincent 2013). Our knowledge about their distribution ranges is still developing. About 15% of the current total sightings from our citizen science database (iSeahorse, iseahorse.org) are located beyond the ranges that we previously knew. To date, seahorse localities are poorly-recorded in many regions.

China is among the countries where seahorses are poorly-documented and threatened. Seahorses are distinguished by their heavy use in Traditional Chinese Medicine (TCM). Every year, millions of dried seahorses are used in TCM by Chinese people (Vincent et al. 2011). To date, formal seahorse biogeographic research is rare in China. Six seahorse species are purportedly present, and probably all are threatened (Wang and Xie 2009). One of these species, *H. kelloggi* (great seahorse), is on China's List of Wildlife under National Protection, mandating a nationwide ban on its catch and trade by law (MEP 2002). The other five species have been proposed to be added to the List, which is under review (Zhang Chun-Guang, per. comm.). However, the lack of distribution knowledge of seahorse populations in China's vast marine territory impedes the protection of these poorly-known animals.

Here I present the first biogeographic study of seahorses in China, with an aim to inform their conservation. I collate multiple species datasets (i.e. SS, PRL, and LEK) and environmental data to build and compare species distribution models. I test whether species data from PRL and LEK can generate similar predictions of seahorse distributions. I examine if adding information from LEK and PRL to SS can improve model performance and predictions. By doing so, my study provides insights on species data collection and analyzing techniques for distribution modeling studies on poorly-recorded species.

2.3 Methods

2.3.1 Study area

My study area spans China's coastal waters (17° to 41°N; 106° to 125°E, Fig. 2.1), which are fringed by the Bohai Sea, Yellow Sea, East China Sea, and the northern South China Sea. The coastline stretches across 18,000 km from temperate to tropical zones (see details in Liu 2013).

2.3.2 Species distribution model

I used a typical presence-only model, maximum entropy (Maxent, Phillips et al. 2006), to analyze my data and to predict seahorse distributions. Maxent produces a habitat suitability map for the focal species based on a set of related variables (model predictors) and a set of georeferenced occurrences. Maxent is considered as one of the most powerful modeling techniques (Hernandez et al. 2006, Phillips and Dudík 2008), as it is 1) robust to positional uncertainty/errors in species occurrences (Graham et al. 2008, Fernandez et al. 2009), 2) suitable for limited occurrences (e.g. SS dataset in my case), and 3) reliable for deriving predictive maps with coarse-grain data (Osborne and Leitao 2009).

2.3.3 Model predictors

I compiled data for twenty-one variables belonging to three categories: 1) climate and geophysical suitability (Tyberghein et al. 2012), 2) food availability, and 3) macro-habitat availability from online databases (Table A.1 in Appendix A). Original data were interpolated with resolution of 1/12 degree in latitude and longitude (~ 10 km) using Inverse Distance Weighting in an ArcMap (Cheung et al. 2009). I chose 1/12 degree as my standard resolution because the majority of the original data were at this resolution, and it also represents cells explicit enough for mapping seahorses at the broad spatial scale of my study area. Since seahorses are typically found in shallow waters, I used a 200-m depth envelope (commonly considered to be continental shelf) as the geographic boundary for all environmental data. By doing so, I can prevent model over-prediction. I then used Pearson correlation coefficients to identify and exclude highly correlated variables (|r| > 0.7), which were not used in the model.

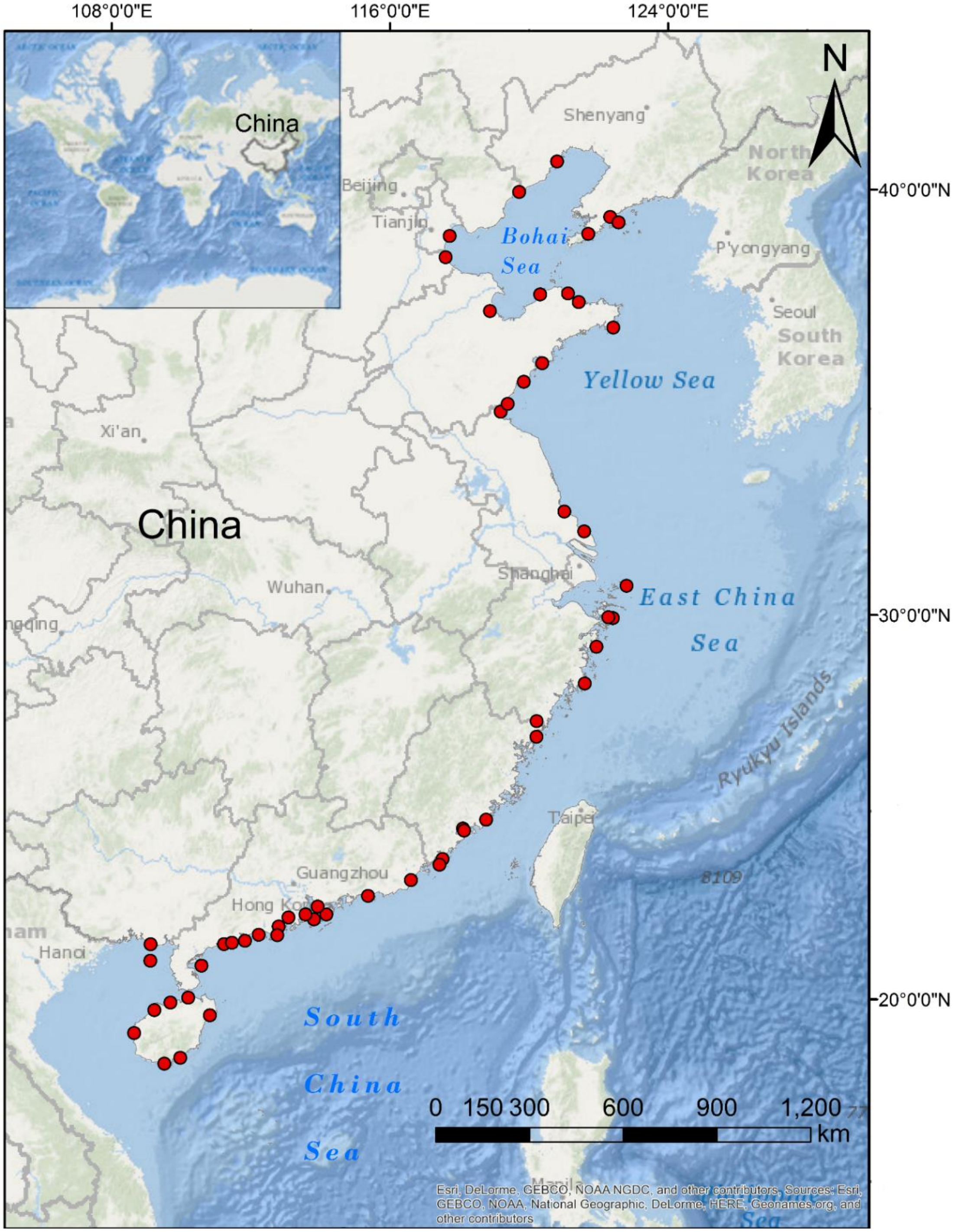


Figure 2. 1. Study area displaying the sampled fishing ports (red points) along the coast (including islands) of the People's Republic of China.

2.3.4 Species data

Species sightings (SS)

I first obtained a total of 33 species sightings (SS) from five online databases: Global Biodiversity Information Faculty (GBIF, www.gbif.org), Oceanic Biodiversity Information System (OBIS, www.iobis.org), FishNet2 (www.fishnet2.net), FishBase (www.fishbase.org), and iSeahorse (www.iseahorse.org). I then obtained new sightings records of seahorses from Chinese colleagues, divers, and fishers during my interview-based research in China (see next paragraph of local ecological knowledge). I validated the species identification for all records by checking specimens where possible, using a standard identification textbook (Lourie et al. 2004). To ensure data quality, sightings located on land or out of my defined range (i.e. 200-m depth of China's seas) were not used.

Peer-reviewed literature (PRL)

I extracted data from peer-reviewed literature (PRL) drawn from the China Knowledge Resource Integrated Database (www.eng.oversea.cnki.net, see Appendix A), having found little information in western literature. I emailed authors to request photos of the specimens to validate their identifications. If specific localities were not documented, I included the entire study/sampling area described in the paper as part of the species' range. All species maps from the validated records in literature were digitalized in an ArcMap.

Local ecological knowledge (LEK)

To derive local ecological knowledge (LEK), I conducted semi-structured interviews (Huntington 2000) at 79 fishing ports (Fig. 2.1) along the entire coast of China from April to September 2015 (see protocol in Appendix A). The choice of these sites was based on comprehensive consultation with four Chinese colleagues and 28 fishers in the field. At each fishing port, I first chose participants recommended by local fisheries scientists, community leaders, and interviewed fishers. I also haphazardly reached out to other fishers who were available and knowledgeable (e.g. skippers). I conducted each semi-structured interview on board a vessel allowing all fishers working on the boat to participate. This group setting allowed me to cross-validate data among the fishers. My interviews covered fishers using different types of fishing gears (n = 10) *in situ*.

In each interview, I first identified the seahorses (Fig. 2.2). I evaluated available specimens *in situ* then presented a collection of seahorse photographs to help participants recall seahorses that they had sighted. After the interview, I validated the interview data by checking specimens from other sources at the same site. These sources included other participants, local seafood landings and markets, and stores at the same fishing port.

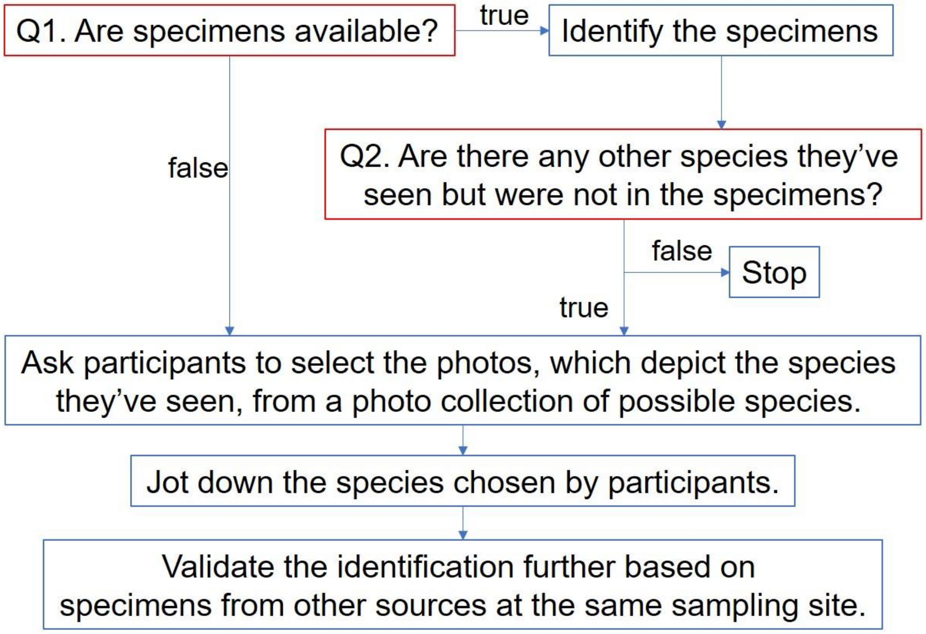


Figure 2. 2. Process of seahorse taxonomic data collection and validation based on fishers' local ecological knowledge.

After the taxonomic portion of the interview, I worked with participants to generate distribution maps of each species (Fig. 2.3). Local commercial fishers often use China's fishing-zone maps (Fig. A.1 in Appendix A), sometimes in a digital version, to guide their fishing activities, such as locating fishing grounds. Therefore, these maps were ideal tools to help fishers describe (without drawing) species locations. I also asked fishers to recall habitat (depth, substrate) or geographic (distance to landmarks) features of each location if they could. These additional data were checked against nautical charts prior to analyses as a manner to test the reliability of fishers' knowledge. If fishers were not familiar with fishing-zone maps, I presented nautical charts instead for them to recall species distributions. I used an iPad with iGIS software to facilitate fishers' mapping in the field, and digitalized their narrative data in an ArcMap later. As a final step to ensure data quality, I overlaid all fishers' maps of the same species and only retained areas that included at least two observations.

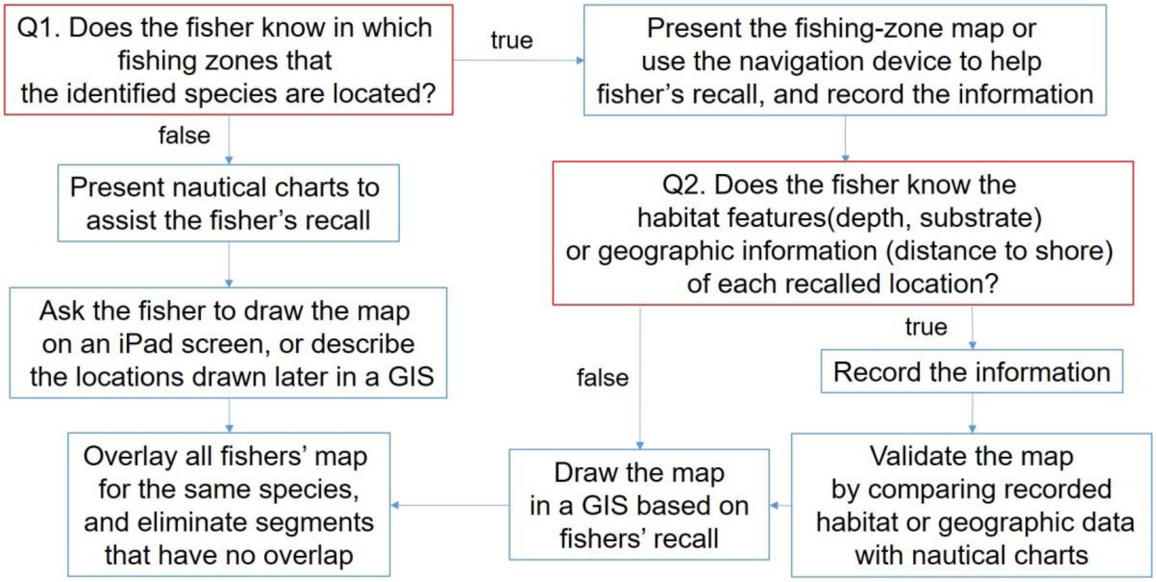


Figure 2. 3. Processes of mapping and validation of seahorse species distributions based on fishers' local ecological knowledge.

Occurrence-points sampling from PRL and LEK coarse maps

Given that species distribution models can only use species point data rather than polygon maps (PRL and LEK original datasets), I systematically sampled presence points from the PRL and LEK maps. To do so, I first refined the original irregular polygons (i.e. PRL or LEK maps) to range maps consisting of cells with my standard resolution (1/12 degree) in an ArcMap (see Appendix A).

To generate occurrence points from the coarse range maps, I adapted a probability-based sampling approach based on habitat suitability (Niamir et al. 2011). This approach performed better than another technique, random sampling, in an initial trial (see Appendix A). The probability-based approach assigns denser occurrence points to more suitable habitats within the range map, based on the ecological principle that higher abundance was expected in more suitable areas. The required habitat suitability was derived from the Maxent model based on the SS dataset (see 1st scenario in 2.3.5). I then employed the spatially-balanced sampling in an ArcMap to generate occurrences from the range maps of LEK and PRL for each dataset. This probability-based sampling tool considers sampling probability (i.e. habitat suitability, in this case) and minimizes spatial autocorrelations among the generated points (Theobald et al. 2007).

2.3.5 Model scenarios and settings

I executed three model scenarios based on different datasets (and combinations) for each species (Table 2.1): 1) a model with only SS data (1st scenario), 2) three models respectively based on LEK, PRL, and their combination (LEK&PRL, 2nd scenario), and 3) three models separately with SS&LEK, SS&PRL, and ALL (i.e. all sources of datasets, 3rd scenario).

To generate pseudo-absence points for each model, I created explicit bias files to determine sampling background for Maxent (Phillips et al. 2006). For SS datasets, I defined a buffer zone around sightings of all species as the common sampling background for each species. By doing so, the model will generate pseudo-absences only within a certain distance from seahorse presence points. This can advance model’s ability to discriminate “highly suitable” from “suitable” habitats (i.e. minimizing over-prediction) (Mateo et al. 2010), which is important for mapping distributions for rare species like seahorses (Zarnetske et al. 2007). I chose ¼ degree (~ 15 nautical miles) as the buffer size, as it produced models with the smallest variation among model parameter estimates using different buffer sizes in an initial trial. Similarly, for LEK (or PRL) datasets, I used the union of range maps from LEK (or PRL) of all species as the same sampling background for each species. For each combination of different datasets, I overlaid the sampling backgrounds from the member datasets as the new sampling background. Then I created the bias file based on the sampling background for each model. All the above processes were done by using the SMDtoolbox (Brown 2014).

I conducted all models by using the Maxent software (version 3.3.3k, Phillips et al. 2006). For each model, the number of pseudo-absence points were kept at default (n < 10,000). I randomly divided the species data into training and test subsets (75% and 25% respectively) and replicated the randomizations 15 times. I applied the regularization multiplier to control for over-parameterization (Crall et al. 2015).

Table 2. 1. The three model scenarios tested in the study. Abbreviation: SS, species sightings; LEK, local ecological knowledge; PRL, peer-reviewed literature; ALL, all data combined (SS, LEK, and PRL); AUC, Area under the curve.

Model scenario	Dataset	Model measure	Statistical analysis	Aim
1st Scenario	SS	AUC, Sensitivity, Specificity	-	To derive probability map; to examine effects of adding LEK and/or PRL data to SS when combined with the 3rd Scenario
2nd Scenario	LEK, PRL, LEK&PRL	AUC, Sensitivity, Specificity; I similarity statistic, presence agreement, absence agreement	Quade test	To compare model performance and results between LEK and PRL
3rd Scenario	SS&LEK, SS&PRL, ALL	AUC, Sensitivity, Specificity	Quade test	To examine effects of adding LEK and/or PRL data to SS when combined with the 1st Scenario

2.3.6 Model evaluation and statistics analyses

I used the mean and standard deviation of three statistics as measures of model performance and variability. The first statistic was the area under the curve (AUC) of the receiver-operating characteristic (ROC) plot (Hanley and McNeil 1982), which is one of outputs of the Maxent (i.e. test AUC). The ROC plot demonstrates presence/absence prediction accuracy with all possible thresholds of the probability value predicted by the model. The AUC of the ROC plot measures model's general accuracy of both presence and absence predictions. Although the use of AUC has been criticized (Lobo et al. 2008), it is reliable for evaluating presence-only models without applying any probability threshold (Lawson et al. 2014). The second and third statistics were Sensitivity (true presence rate) and Specificity (true absence rate) (Altman and Bland 1994). These two are threshold-dependent and could be better indicators of a model's discriminatory power than the AUC (Lobo et al. 2008). I used the probability threshold at which training Sensitivity plus Specificity was maximized (hereafter SSM threshold) to calculate these two statistics (Liu et al. 2005). In the 2nd scenario of comparison between LEK and PRL, I used SS datasets as independent test data to estimate Sensitivity and Specificity. The calculation was realized in R (R Core Team 2016) based on the output data from Maxent.

I used Quade tests (Quade 1979) to examine the effect of using different datasets (LEK, PRL, LEK&PRL) on model performance and variability by controlling for species (n = 5) in the 2nd scenario. For each species, I employed the I similarity statistic (ISS) to measure the strength of agreement on presence-probability maps between LEK and PRL (Warren et al. 2008). I derived the presence/absence maps (SSM thresholds, see above) for LEK and PRL to calculate an agreement ratio (spatial overlap to spatial union) respectively for the predicted presence and absence. Both ISS and the two agreement ratios range from 0 (unmatched) to 1 (fully matched). I then examine the effect of adding LEK and/or PRL to SS on model performance and variability by Quade tests based on model results of the 1st and the 3rd scenarios (Table 2.1).

2.3.7 Model prediction and ecological interpretation

I followed three steps to determine the presence/absence map for each species. First, I only considered the four models in the first and third scenarios based on datasets including SS, given that SS were more precise than those points downscaled from LEK and PRL maps. Second, I rejected poorly fitting models (AUC lower than 0.7) and derived the presence/absence maps (by SSM threshold) for the remaining models (Manel et al. 2001). Third, I derived an original-data map by overlaying original SS, LEK, and PRL datasets, and then compared the original-data map with each of the predictive maps. I finally selected the predictive map that had a higher number of overlaps with the original-data map and covered less area. If some of the original sightings/ranges were not represented by the selected map, I added the cells occupied by these sightings/ranges to the "selected" map and labeled them as "omission points/ranges".

I derived a species-richness map by stacking the presence-probability maps instead of the binary presence/absence maps to avoid overprediction on species richness (Calabrese et al. 2014). The presence-probability map for each species was from the logistic-probability predictions of the model that generated the "selected" map (hereafter, selected model). The species-richness map was derived by overlaying these probability maps in an ArcMap.

I applied the selected model's permutation importance to identify key environmental factors (Searcy and Shaffer 2016), and the partial response curves to interpret species ecological niches (Stirling et al. 2016).

2.4 Results

2.4.1 Species data

I obtained a total of 55 species sightings (SS), 463 fishers' maps (LEK), and 42 literature maps (PRL) across the same five species (Table 2.2, see Fig. B.1 – B.3 in Appendix B). Four other species were reported but not validated (see details in Appendix B and Table B.1 and B.2 therein). I excluded these species and their maps (21% of total maps) from model datasets. Among the valid species, *H. trimaculatus* was the most frequently sighted in SS dataset and the most frequently reported in the LEK dataset. *Hippocampus mohnikei* was the most frequently recorded in PRL dataset. All LEK and PRL data were range maps coarser than my mapping resolution, but LEK were finer (LEK: 1,177 ± 1,115 km^2, PRL: 12,453 ± 11,630 km^2). I generated occurrence points (n = 50 – 200, Table 2.2) from the coarse-grain maps (LEK and PRL) by PBS approach. The total number of points differed among species given the different sizes of total area covered by original maps. A minimum of 50 points and a maximum of 200 points were chosen to build robust models and minimize spatial autocorrelation after initial trials.

Table 2.2. Summary of five Chinese seahorse species with the frequency of records (Frequency), the total number of maps (Maps), and the total amount of points (Points) from three sources: species sightings, local ecological knowledge, and peer-reviewed literature. For the latter two, the points were generated by the downscaling technique based on original maps.

Seahorse species	Species sightings		Local ecological knowledge			Peer-reviewed literature		
	Frequency	Points	Frequency	Maps	Points	Frequency	Maps	Points
H. kelloggi	10.9%	6	13.4%	37	100	11.8%	4	100
H. kuda	27.3%	15	11.2%	31	50	5.9%	2	50
H. mohnikei	14.5%	8	37.1%	103	100	82.4%	28	100
H. spinosissimus	9.1%	5	18.3%	51	100	2.9%	1	100
H. trimaculatus	38.2%	21	87.0%	241	200	8.8%	3	200

2.4.2 Local ecological knowledge (LEK) vs. peer-reviewed literature (PRL)

I found that although LEK tended to produce better models than PRL and LEK&PRL (Fig. 2.4), the differences were not statistically significant. The highest AUC was obtained by LEK for all species except *H. trimaculatus* (Fig. B.4 in Appendix B). The highest values of true presence rate (Sensitivity) were derived from LEK for all species but *H. kelloggi* (Fig. B.4). The highest values of true absence rate (Specificity) were also generated by LEK datasets for all species except *H. mohnikei* and *H. kuda* (Fig. B.4). All model performance measures (AUC mean and SD, Sensitivity, and Specificity) across species were not statistically different among the three datasets (LEK vs. PRL vs. LEK&PRL, Quade tests, all $p > 0.05$).

My result indicated that LEK and PRL were generally consistent on model predictions when tested on probability maps, but they did not match well on predicted distributions. The I similarity statistics were medium to high across the five species (0.571 – 0.853, Table 2.3), suggesting a roughly good agreement on the predictive probabilities between LEK and PRL. In contrast, the Presence agreements between model results of the two datasets were generally low (Mean = 21.3%), although the Absence agreements were relatively high (Mean = 71.9%, Table 2.3).

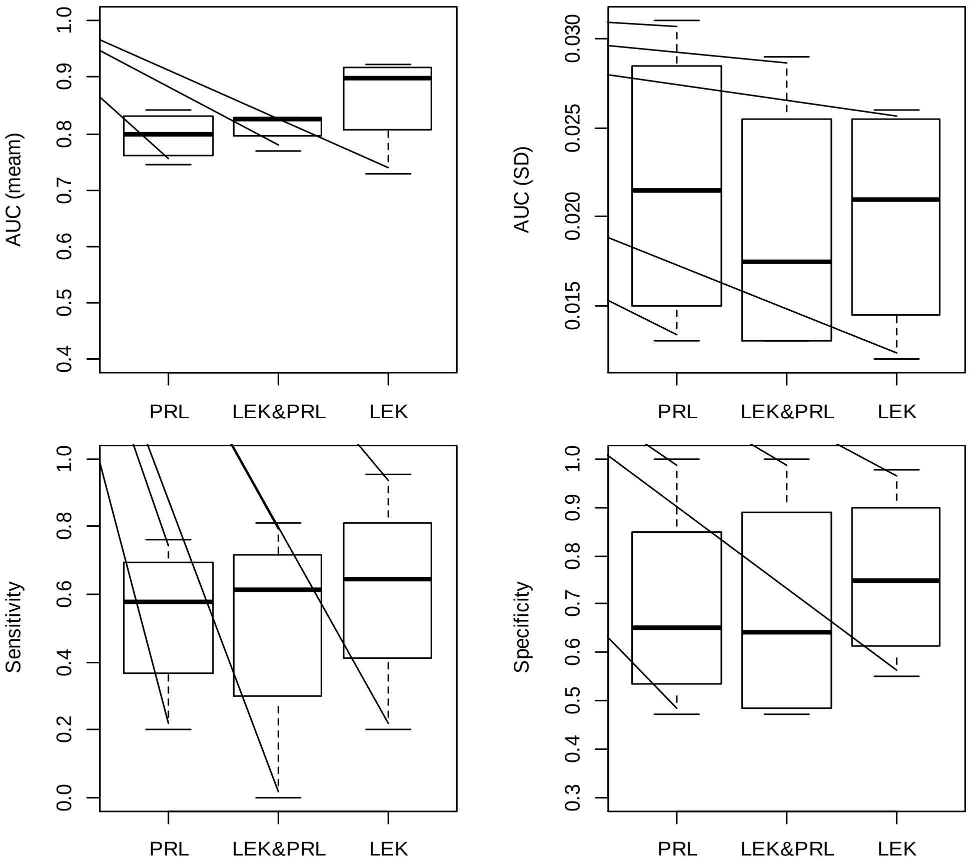


Figure 2. 4. Boxplots of model performance and variability of Maxent models generated from three different datasets (PRL, LEK, LEK&PRL) based on four measures: AUC (area under the curve) mean and SD (standard deviation), Sensitivity (true presence rate), and Specificity (true absence rate). PRL, peer-reviewed literature data; LEK, local ecological knowledge data; LEK&PRL, the combination of the two.

Table 2. 3. Estimates of agreements on predictive maps between models of local ecological knowledge and models of peer-reviewed literature across the five seahorse species, based on three different measures: I similarity statistic, Presence agreement, and Absence agreement.

Species	**I similarity statistic**	**Presence agreement**	**Absence agreement**
H. kelloggi	0.571	12.0%	64.7%
H. kuda	0.804	7.1%	86.1%
H. mohnikei	0.681	26.7%	54.0%
H. spinosissimus	0.778	16.6%	75.6%
H. trimaculatus	0.853	44.3%	78.9%
Mean	0.737	21.3%	71.9%

2..4.3 Effects of adding species locations from coarse-grain data (LEK and PRL) to species sightings (SS)

I found that adding LEK and/or PRL to SS generally derived better predictive models than using SS alone (Fig. 2.5 left), although the effects differed among the measures. I detected significant differences on Sensitivity (true presence rate) among datasets (Quade tests, $p < 0.05$, $n = 5$), although not on AUC ($p = 0.115$) or Specificity (true absence rate; $p = 0.108$). Sensitivity mean value for SS dataset was significantly lower than those for SS&PRL and SS&LEK (both $p < 0.05$), but not for ALL ($p = 0.07$). Therefore, adding LEK or PRL, but not both, to SS could significantly improve presence prediction. Model general performance (AUC mean) tended to improve by adding LEK and/or PRL (Fig. 2.5 left), while absence prediction (Specificity mean) only tended to improve by adding LEK alone (Fig. 2.5 left).

For model variability, I obtained similar but more consistent results on different measures (Fig. 2.5 right). I examined statistical differences on standard deviations of all measures (AUC, Sensitivity, and Specificity) among the compared datasets (Quade tests, all $p < 0.05$, $n = 5$). The standard deviations of AUC and Sensitivity for SS dataset were significantly higher than those for the others (posthoc-Quade test, all $p < 0.05$). The standard deviation of Specificity for SS dataset was significantly higher than SS&PRL and ALL (both $p < 0.05$), but not SS&LEK ($p = 0.06$). These results revealed that adding LEK and/or PRL reduced model variability on predicting presence; while only adding PRL or LEK&PRL data decreased model variability on predicting absence.

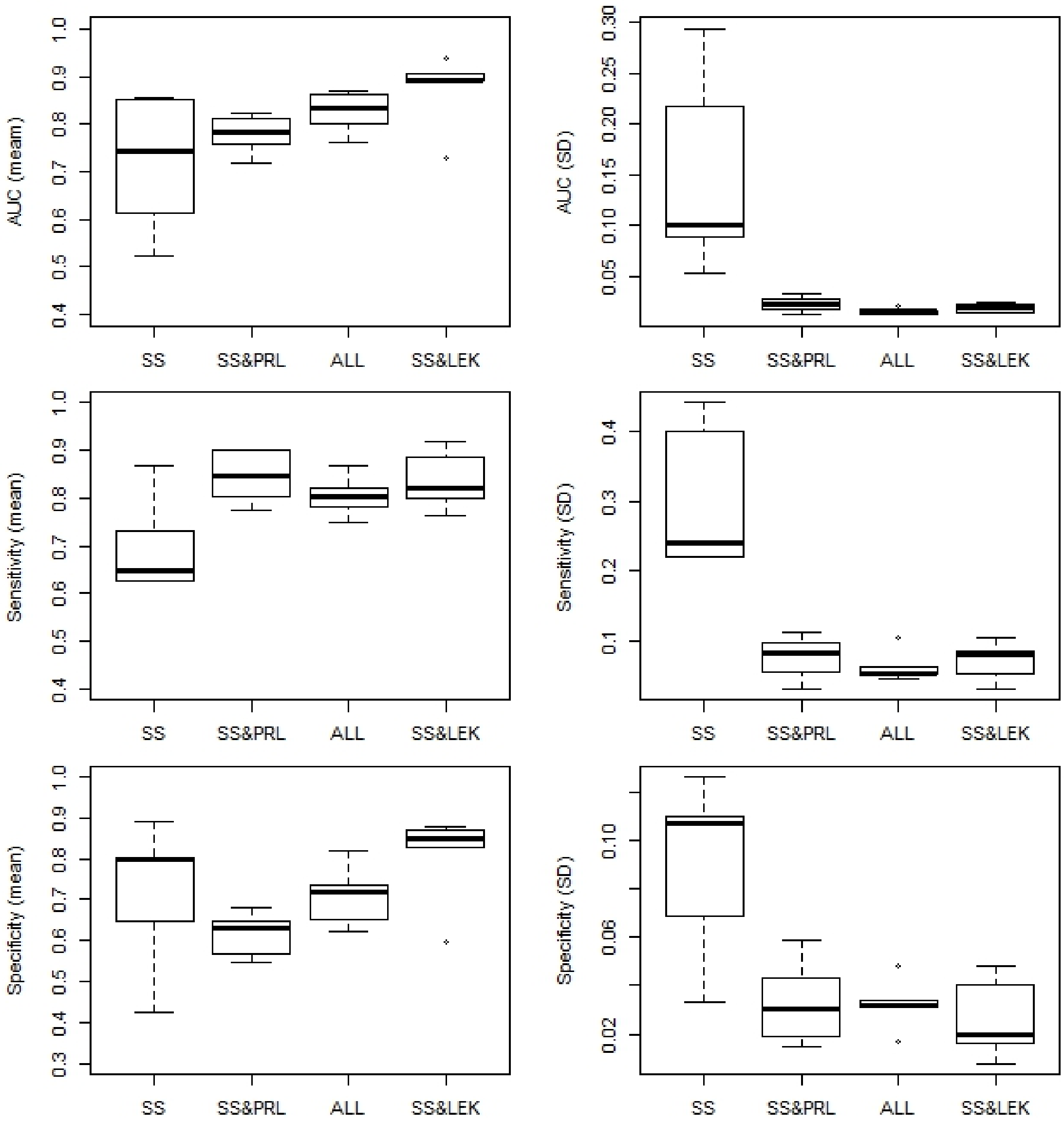


Figure 2. 5. Boxplots of model performance (mean, left panel) and variability (SD, standard deviation; right panel) measured with the AUC (area under the curve), Sensitivity (true presence rate), and Specificity (true absence rate) among all five seahorse species, based on four different datasets: SS, species sightings; SS&PRL, sightings plus peer-reviewed literature; ALL, sightings plus peer-reviewed literature plus local ecological knowledge; and SS&LEK, sightings plus local ecological knowledge.

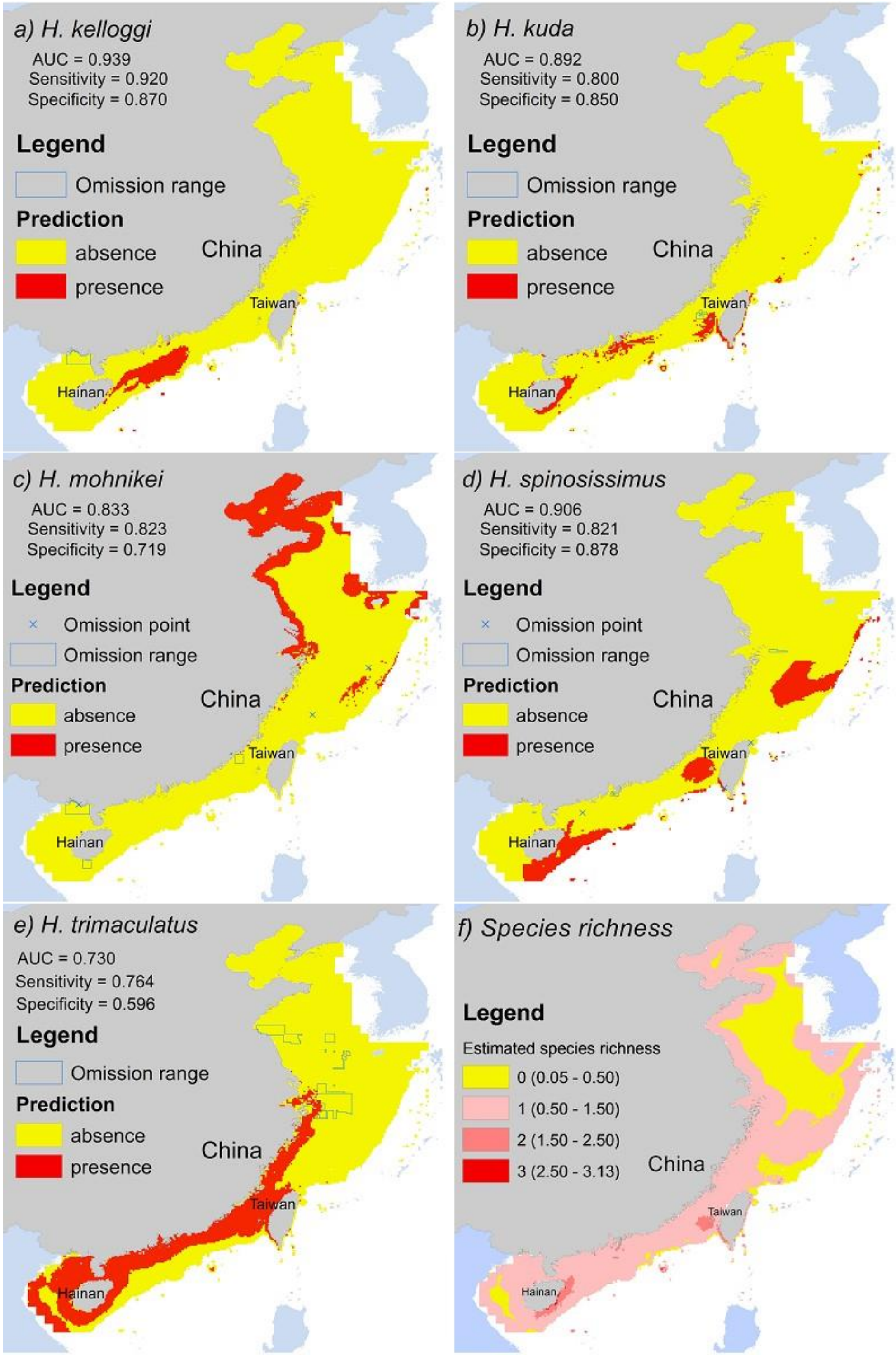


Figure 2. 6. Predicted distribution maps of a) *H. kelloggi*, b) *H. kuda*, c) *H. mohnikei*, d) *H. spinosissimus*, e) *H. trimaculatus*, and f) species richness of seahorses in China. Omission range or point represent model predictive errors. AUC, area under the curve; Sensitivity, true presence rate; Specificity, true absence rate.

2.4.4 Model prediction

The predictive maps indicated that these five seahorse species were generally divergent in spatial distributions, with more species located in the south (Fig. 2.6). I derived presence/absence maps (Fig. 2.6 a - e) and presence-probability maps (Fig. B.5 in Appendix B) based on SS&LEK datasets for all species but *H. mohnikei*, for which ALL dataset was used. Stacking the presence-probability maps derived a species richness ranging from 0.05 to 3.13, which were then rounded to the nearest integer (i.e. 0 to 3, Fig. 2.6 f). That resulted in an area of 70.3% of my defined shallow seas was probably occupied by only one seahorse species, 2.1% by two species together, and 0.1% by three species together. Different seahorse species mainly co-existed in the South, especially the eastern Hainan Island and the Penghu Archipelago (Taiwan Province, Fig. 2.6 f).

2.4.5 Parameter estimation and species ecological niches

I selected eight predictors from the original twenty-one factors (Table 2.4; Table A.1 in Appendix A). The rest thirteen predictors were removed since they were highly correlated with one of remaining eight predictors (Pearson correlation test, $|r| > 0.7$; see details in Table A.1). The ecological niches of different species were reflected in the response curves on the selected environmental predictors (Fig. B.6 – B.13 in Appendix B). Spatial autocorrelation was low (Moran's I = 0.01 to 0.09) in the model datasets for all species, which justified the use the model results to interpret predictor importance and species ecological niches.

The importance of different predictors varied among species, with sea surface temperature (SST mean) generally the most influential variable (Table 2.4). *Hippocampus mohnikei* was the only species that was more likely to occur in colder waters (Fig. B.6), which might explain the significant spatial separation between *H. mohnikei* and the other four species. Species niche separation was also reflected on other predictors. For example, compared with other species, the presence probabilities of *H. kuda*, *H. mohnikei, H. trimaculatus* were higher in shallower water closer to the shore (Fig. B.7 and B.8).

Table 2. 4. Selected model predictors and their relative importance (%) for the model of each species, based on analysis of variable contributions in the Maxent model. Predictors were ranked by the average importance across the five species.

Predictors	Predictor importance % in modeling for each species					Mean
	H. kelloggi	*H. kuda*	*H. mohnikei*	*H. spinosissimus*	*H. trimaculatus*	
sea surface temperature	**65.6**	**46.2**	**72.2**	**47.6**	**58.2**	58.0
distance to shore	2.4	31	16.8	2.2	17.6	14.0
calcite concentration	11.3	14.3	2	20.3	7.4	11.1
silicate concentration	16.1	4.4	2.8	10.7	5.1	7.8
depth	0.7	1.9	4.4	1.1	7	3.0
pH	1.9	1.3	0.5	10.6	0.3	2.9
primary productivity	2.2	0.3	0.6	6.7	3.1	2.6
macro-habitat	0.6	0.6	0.8	0.9	1.3	0.8

2.5 Discussion

My study demonstrated that spatially-explicit population-distribution maps of poorly-documented species can be derived by integrating readily available data with species distribution models. Lacking explicit occurrence maps is a common challenge for conservation planning for poorly-recorded organisms (Rondinini et al. 2006, Levin et al. 2014). I indicated that valuable species data could be derived from local ecological knowledge (LEK) and peer-reviewed literature (PRL), when species sightings (i.e., SS) were rare. By integrating different species datasets with a presence-only model (i.e. Maxent), I illustrated that it is beneficial to add LEK and/or PRL to the limited SS. These findings are encouraging given the need to inform conservation actions for rarely-studied species (Rondinini et al. 2006), but the financial and temporal constraints on data collection (Anadón et al. 2009).

I demonstrated that fishers can provide species-level maps for seahorse species in China, a capacity which is not universally held by fishers (Aylesworth et al. 2017). The divergence might relate to fisher interest, seahorse morphology, and overlaps in seahorse distributions. For instance, in Thailand, fishers were poor at distinguishing species and may have lacked incentives to care (Aylesworth et al. 2017). In contrast, Chinese fishers could benefit financially from species identification, for large or smooth seahorses fetched higher prices according to respondents. In Thailand, morphological divergences among the seven species may be difficult to differentiate, whereas the five Chinese species have apparent differences on species-specific features, body size, and smoothness (see details in Appendix B and Fig. B.14 and B.15 therein, and Lourie et al. 2004). In the Thai study, there was likely considerable overlap in seven seahorse species' distributions across the five degrees of latitude (5-10°N) in a tropic zone, whereas the five species in China were apparently more dispersed across the 24 degrees of latitude (17° to 41°N). Fishers in my study only mentioned a maximum of two species each, suggesting little overlap. The uncertainty of the rest four un-validated species may, however, reflects that not all fishers are knowledgeable in seahorse identifications in China. Identifying the key informants is indeed important for using local ecological knowledge.

2.5.1 Seahorse species distributions

The distribution patterns for the five species in China are generally consistent with counterparts in other regions. As in Peninsular Malaysia (Choo and Liew 2003), I found that *H. kelloggi* were more likely to occur in deep (> 30 m) offshore waters, and *H. kuda* was patchily restricted to shallow inshore waters in China. *Hippocampus mohnikei,* was largely clumped in China's temperate zone, as in Japan (Lourie et al. 2004), and Korea (Choi et al. 2012). But this species could occasionally occur in China's warmer regions extending to Southeast Asian countries (see Aylesworth et al. 2016 and references therein). *Hippocampus spinosissimus* and *H. trimaculatus* were more likely to be sympatric, and less patchy than the other three species (Choo and Liew 2003, Lawson et al. 2015). *Hippocampus trimaculatus* also had the widest habitat and was relatively more abundant than the other species in bycatch (Choo and Liew 2003). I noted that Chinese seahorse populations were not likely to be found in estuaries, in contrast to its Malaysian counterparts. A possible explanation might be the relatively higher environmental stress (e.g., water pollution) in China's estuaries (SOA 2012).

My study indicated that seahorse distributions are highly correlated with ocean temperature at a large spatial scale, in line with many other marine taxa (Tittensor et al. 2010). As ectothermic fish, seahorses are expected to require suitable water temperatures to sustain their metabolism and reproduce (Beitinger and Fitzpatrick 1979). This temperature suitability could vary among species as indicated in my study and reported from studies in seahorse aquaculture (Koldewey and Martin-Smith 2010). Given China's seas are among the world's most rapid warming zones (Belkin 2009), a further study to explore the impact of global warming on seahorse distributions in China would be interesting and meaningful.

2.5.2 Limitations

The low modeling importance of macro-habitat and the derived species-habitat relations should be viewed with caution, given the constraints on my data availability and spatial scale. First, the macro-habitat datasets I used here were largely drawn from global sources, and thus they might not have good coverage or resolutions in China's seas. Second, I learnt from fishers that Chinese seahorses use various other macro-habitats, including macro-algae, sea fans, and artisanal structures (e.g. mussel farms) (Aylesworth et al. 2016). However, I was unable to obtain these habitat data for China. On the other hand, the importance of macro-habitats might be more prominent at a smaller spatial scale (e.g. a lagoon), as found in European seahorses (Curtis and Vincent 2005) and reef fishes (Komyakova et al. 2013). To clarify the truth, a study in multiple spatial scales with more comprehensive and explicit macro-habitat variables could be helpful.

2.5.3 Implications for conservation and management in China

My new spatially-explicit maps for five seahorse species across China can be used for conservation and management in at least four ways. First, the predictive maps can guide local authorities and researchers to identify more specific locations of seahorse populations for further conservation actions (e.g. protected areas). Second, Chinese authorities can use my maps to protect seahorses through fisheries management. For instance, fisheries officers can explore overlap between seahorse distribution maps and fishing-zone maps to identify locations, where reducing or relocating fishing activities could maximize the benefit for seahorse conservation (Foster and Arreguin-Sánchez 2014). Third, local law enforcement officers can identify localities (e.g. fishing ports) that are close to seahorse habitats to effectively strike against illegal fishing and trade in seahorses. Fourth, my maps come timely to serve China's ongoing coastal ecological-redline planning, which aims to protect coastal waters with significant ecological importance (Peng et al. 2016). Given that seahorses are found in various coastal ecosystems, protecting seahorses means protecting these critical waters. I think a planning includes seahorse habitats could be ecologically meaningful and beneficial to China's marine systems.

2.5.4 Integrating multiple datasets in SDM research

Among the approaches used in this study, I adapted a technique especially applicable for poorly-recorded species to address the coarse-grain maps in SDM research. The use of LEK in modeling distributions is a common challenge in both terrestrial and marine systems. Local citizens could only provide coarse-range maps for the focal organisms (Bergmann et al. 2004; Carter and Nielsen 2011; Laze and Gordon, 2016; Aylesworth et al. 2017). Downscaling techniques are thus vital to generate finer-resolution data from coarse-grain maps like LEK. I adapted the downscaling technique of Niamir et al. (2011), which used expert knowledge to determine habitat suitability for each species. Expert knowledge may be only available for well-known populations (Murray et al. 2009, Niamir et al. 2011), and not for rarely-studied ones like Chinese seahorses. To fill the gap, I used the Maxent model to generate habitat suitability, as it only requires a small number of species sightings ($n \geq 5$). This advantage may make my technique more applicable to poorly-documented species, which was rarely addressed in the literature.

My study indicated that integrating multiple datasets in predicting distributions for poorly-recorded species is beneficial, but it should be done with appropriate datasets combinations. I demonstrated that integrating LEK with PRL should be taken with caution, as it may not necessarily improve model performance. In my study, both LEK and PRL occurrences were sourced from coarse-grain maps, which were not very consistent in space. Previous studies of comparing LEK with data from traditional biological surveys have also shown similar mismatch on other species (Thornton and Scheer 2012). Combining species data from "inconsistent" sources might cumulate spatial variances or variances in tuning species niche to environmental factors. Besides, given both LEK and PRL are likely coarse maps and opportunistic observations, the derived occurrences could include some spatial errors and biases. This is especially true for PRL maps in my study (10 times coarser than LEK maps, on average). Both reasons above could increase model residual and degrade model performance (Graham et al. 2008). I suggest future studies, which aim to integrate datasets from different sources (especially coarse-grain data), better compare different scenarios as I executed to identify the best datasets combinations and model predictions.

I highlight the possible utility of LEK in species distribution modeling for poorly-recorded marine species, in line with similar studies on terrestrial counterparts (Irvine et al. 2009, Anadón et al. 2010, Laze and Gordon 2016). As acknowledged in terrestrial studies (Anadón et al. 2010), I think that LEK can be a cost-effective data source for modeling poorly-recorded species in marine systems. In addition, LEK can provide information to potentially improve model predictions. For instance, with the key information about physical features (e.g. fence, roads) and weightings of environment factors, Irvine et al. (2009) significantly improved model predictions of deer (*Cervus elaphus*). In my study, local fishers have sighted seahorses often co-occurring with sea fans in bycatch, and clumps of juvenile seahorses drifting in water column with branches of macro-algae (Fig. B.16 in Appendix B). These LEK data suggest that other related macro-habitats variables if available might derive better predictions.

2.5.5 The importance of spatially-explicit maps of poorly-recorded species

Spatially-explicit biogeographic maps for poorly-recorded species can help inform conservation actions. In particular, greater detail in spatially-explicit biogeographic maps can improve conservation planning (Rondinini et al. 2006). For instance, species of concern can benefit from fine-grain resolution when exploring overlaps between their distributions and protected areas (Rondinini et al. 2006; Pimm et al. 2014). The corollary is that fine-resolution distribution maps of focal species can inform us of conservation gaps, where actions should be taken. The ability to develop such spatially-explicit maps for poorly-known species should allow protective measures even as knowledge is being improved. Such potential may be particularly important in the many biodiverse countries (e.g. China) with limited biogeographic data and resources (Liu 2013). In such instances tapping into diverse sources of information (including local ecological knowledge) can create valuable species distribution models and predictive maps.

Chapter 3: Predicting distributions, habitat preferences, and associated conservation implications for a genus of rare fishes, seahorses (*Hippocampus* spp.)

3.1 Summary

Using species distribution models to determine global distributions of rare species and their species-habitat relationships is vital to biodiversity conservation, but often challenged by the shortage of data. My study provides guidance for identifying useful sources of species data and instrumental habitat variables to build robust species distribution models for rare marine fishes, using seahorses (*Hippocampus* spp.) as the case study. My study demonstrated that using "proximity to macrohabitats" and integrating all datasets of species occurrences derived models with the highest accuracies among all dataset variations. This finding highlighted that using proper habitat variables is crucial to determine distributions and habitat preferences for rare and habitat-dependent marine fishes; collating and integrating quality-unknown occurrences (e.g. citizen science and museum collections) with quality research data is meaningful for building SDMs for rare species. I also encourage the application of SDMs to estimate area of occupancy for rare organisms to facilitate their conservation status assessment.

3.2 Introduction

Understanding global distribution and habitat preference of rare animals is key to ecology, and wildlife conservation (Brooks et al. 2002, Hanski 2011). Over the past few centuries, anthropogenic activities have caused astonishing biodiversity loss, with a detrimental impact on human society (Hooper et al. 2012). Rare species, with low densities or small ranges, are more sensitive to human disturbance and usually have higher extinction risks than common species (Reynolds et al. 2005). Estimating threatened status of rare species usually requires a good knowledge about their geographical distributions (Gaston and Fuller 2009). Their distribution maps are also helpful for identifying global biodiversity hotspots (high species-richness sites that are threatened), where limited conservation resources should be allocated (Myers et al. 2000, Mittermeier et al. 2011). In addition to distribution maps, habitat-preference knowledge is essential to find species at local scales, where conservation actions are usually taken (Harris et al. 2005).

Species distribution models (SDMs) are useful for analyzing species distributions and habitat preferences for rare species conservation (Franklin 2010, Marcer et al. 2013). These techniques are built on a variety of algorithms that correlate species occurrences and ecological covariates (i.e. model predictors) based on the "ecological niche" concept (Whittaker et al. 1973). This concept suggests that species choose their habitats based on their fitness to various surrounding factors. The model approach includes quantitative description of the relationship between species occurrences and ecological covariates. By modeling this relationship, biologists can predict species distributions in un-surveyed regions and generate a global view of species distribution patterns (Franklin 2010). Moreover, integrating habitat variables in the model allows identification of critical habitats based on the parameter estimation function. The spatially-explicit maps derived from SDMs can be used to estimate area of occupancy (AOO), which is a critical measure for assessing threatened status (IUCN Standards and Petitions Subcommittee 2017).

Identifying useful species data and habitat-related predictors is essential for building robust SDMs for rare species (Aubry et al. 2017). Given the difficulty in studying rare organisms, scientists usually need to collate distribution data from quality-unknown sources (e.g. citizen science) to supplement the limited high-accuracy data (e.g. research data). Applying unknown-quality data can be problematic and might, through errors, generate different results than quality data (Graham et al. 2004, Aubry et al. 2017). Studies comparing different data sources are rare in the literature (Jackson et al. 2015). Furthermore, including appropriate habitat covariates in SDMs can be critical for rare species whose presences are correlated with important resources (e.g. food, shelters) within a particular habitat (Rainho and Palmeirim 2011). These issues above can degrade model accuracy and distort species-habitat relationships if they are not addressed appropriately (Aubry et al. 2017).

Seahorses are rare animals, whose conservation is of global concern (Vincent et al. 2011). Seahorses are a genus (*Hippocampus* Rafinesque, 1810) of small, cryptic, and sedentary marine fishes in the family Syngnathidae, well-known for their male pregnancy and charismatic appearance. They are usually found at low population densities (0 – 0.51 individuals /m^2, Foster and Vincent 2004). Because of this and other biological traits (e.g. low fecundity, extensive paternal care and, often, mate fidelity), seahorses are vulnerable to various human activities, especially poorly managed fisheries and habitat degradation (Foster and Vincent 2004). Wild seahorses are often caught in fisheries and traded worldwide, mainly for traditional medicines. In 2002, seahorses became the first marine fishes to be listed on Appendix II of the Convention on International Trade in Endangered Species (CITES). This listing mandates 183 Parties to CITES to ensure that their exports do not threaten wild seahorse populations. The recent estimate of annual seahorse catches in 22 countries, totaling at least 37 million individuals (Lawson et al. 2017), emphasizes the importance of global actions to conserve these rare species.

Global distribution, habitat preference, and conservation status are poorly known for seahorses. Studies of seahorse ecology have only centered on a small fraction of species (Cohen et al. 2017). Available species-range maps are not spatially explicit enough to inform global and local conservation actions (Lourie et al. 2016). Comprehensive habitat knowledge is lacking for most species, though site-level or regional-scale habitat studies are available for some populations (Gullison et al. 2000, Harasti et al. 2014, Aylesworth et al. 2015). Currently, 14 species are considered as threatened (Endangered or Vulnerable), but 17 species are still "Data Deficient" according to the latest IUCN Red List assessment (IUCN 2018; Pollom et al. in preparation).

My study aims to apply SDMs to advance global conservation for all seahorse species. I am interested in examining the utility of quality-unknown species data and identifying proper habitat variables to build robust SDMs for rare species. In undertaking this work, I collated species-level seahorse occurrences and ecological covariates including habitat variables. I then identified the best species datasets and habitat variables deriving the most accurate models. I used these models to predict seahorse distributions, to understand species habitat preferences, and to inform their conservation-status assessment.

3.3 Methods

3.3.1 Species distribution model

I used maximum entropy (Maxent; Phillips et al. 2006) to build species distribution models. Maxent is one of the most powerful and popular SDMs (Hernandez et al. 2006; Phillips and Dudík 2008). It estimates presence probability by finding a distribution with maximum entropy (i.e. closest to uniform), subject to constraints defined by conditions at known occurrence locations (Phillips et al. 2004). Earlier research indicated that Maxent is robust to positional uncertainty/errors in species occurrences (Graham et al. 2008; Fernandez et al. 2009), and particularly suitable for rare species with limited occurrences (Pearson et al. 2007). I developed the models with the latest Maxent software (Version 3.4.1) (Phillips et al. 2017b).

3.3.2 Study species

My study focused on 42 valid species of the genus *Hippocampus* (Lourie et al. 2016; Zhang et al. 2016). Seahorses were typically found in shallow waters (depth < 200 m) from tropic to temperate zones. They are cryptic, sedentary, small fishes (usually body height < 35 cm) that ambush zooplankton and benthic organisms (e.g. Crustacea and Amphipoda) (Manning 2017). Their predators are believed rare, although they were occasionally found in diets of various marine species such as larger fishes and birds (Kleiber et al. 2011). Identification of seahorses at the species level is sometimes challenging because of morphological similarity across species and individual phenotypic plasticity within species (Curtis 2006, Roos et al. 2011). I used work by Lourie et al. (2016), which clarified seahorse nomenclature and revealed many synonyms, as the basis for modeling species-level distributions in my study.

3.3.3 Data collection

I gathered data for model predictors that have ecological relevance and available information. I first derived data for 12 variables related to seahorse physiological suitability and primary productivity (Foster and Vincent, 2004; Table 3.1). I selected seven variables from these twelve covariates to minimize collinearity based on Pearson correlation tests ($|r| > 0.7$) (Dormann et al. 2013). The resulting seven predictors were depth, pH, salinity, sea-surface-temperature mean and range, and chlorophyll-a mean and range.

I then derived datasets of nine categories of potentially key habitats (e.g. coral reefs) from online biogeographic databases (Table 3.1, Table C.1, Supporting Information Appendix C), and generated a binary and a continuous variable for each habitat category. The binary variable was 'habitat presence/absence (1/0)', and the continuous variable was the 'distance to the nearest location of each habitat'. I generated all these model predictors as global maps with a resolution of 1 km^2 (Cylindrical Equal-area Projection), constrained within the 200-m depth range in ArcMap (version 10.2.2). I predicted that continuous habitat variables could be more useful than binary ones, as the former would be resilient to some extent of mismatch between species occurrences and habitat locations. This is crucial as I found many species occurrences and sightings of habitat-forming organisms have no location-certainty information or have coarse resolutions (e.g. 10×10 km^2). This coarse-resolution issue may distort the modeled species-habitat relations, as the recorded species locations and its habitat locations might not overlap in space as they should be.

I collected species-level (presence-only) locations of seahorses from multiple databases including online biogeographic databases, published peer-reviewed literature, unpublished research data (from Project Seahorse), and iSeahorse (www.iseahorse.org) – a global citizen-science platform for gathering seahorse occurrences. I collated the species location data by checking nomenclature and spatial errors. Based on the collated geo-referenced species occurrences, I determined the potential geographic range (i.e. modeling envelope, Fig. 3.1) for each species to constrain model prediction. I predicted that the quality of multiple sources of species data might differ. Therefore, I divided all occurrences of each species into three commonly-used categories: research grade, citizen science, and museum collection. A more detailed description of this section is provided in Supporting Information Appendix C.

Table 3. 1. Original data resolutions and sources for seahorse ecological variables. '*' indicates the seven selected predictors from the original twelve covariates of physiological suitability and primary productivity. SST, sea surface temperature; GEBCO, General Bathymetric Chart of the Oceans; UNEP-WCMC, United Nations Environmental Programme World Conservation Monitoring Centre; GBIF, Global Biodiversity Information Facility; OBIS, Ocean Biogeographic Information System.

Categories	Parameters	Resolutions (arc-degree)	Data Sources
Physiological suitability	depth*	0.00833	GEBCO 2014
	pH*	0.0833	Tyberghein et al. 2012
	salinity*	0.0833	Tyberghein et al. 2012
	dissolved oxygen	0.0833	Tyberghein et al. 2012
	SST mean *	0.0833	Tyberghein et al. 2012
	SST range *	0.0833	Tyberghein et al. 2012
	SST maximum	0.0833	Tyberghein et al. 2012
	SST minimum	0.0833	Tyberghein et al. 2012
Primary productivity	Chlorophyll a (mean) *	0.0833	Tyberghein et al. 2012
	Chlorophyll a (range) *	0.0833	Tyberghein et al. 2012
	Chlorophyll a (maximum)	0.0833	Tyberghein et al. 2012
	Chlorophyll a (minimum)	0.0833	Tyberghein et al. 2012
Habitat availability	coral reefs	0.0003 – 0.01	UNEP-WCMC 2010
	seagrass beds	0.0026	UNEP-WCMC 2016
	mangroves	0.0003	UNEP-WCMC 2011
	estuary	-	UNEP-WCMC Global Estuary Database
	soft corals	-	GBIFandOBIS
	seaweed	-	GBIFandOBIS
	sponge	-	GBIFandOBIS
	sea pen	-	GBIFandOBIS
	hydrozoa	-	GBIFandOBIS

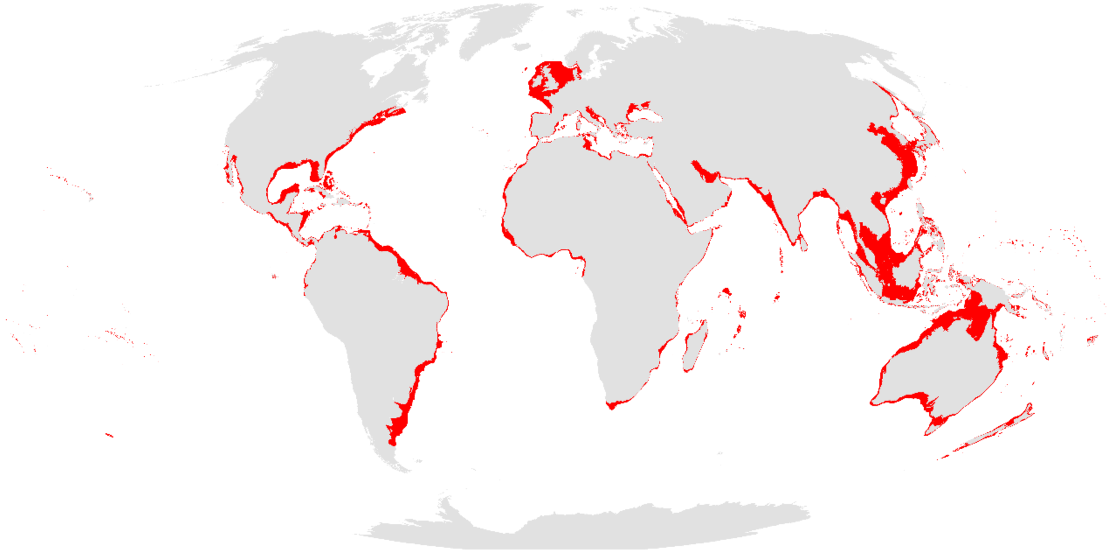


Figure 3. 1. A map of the potential distribution ranges of all seahorse species combined (polygons in red), that we used in the model to constrain model prediction.

Table 3. 2. Description of model variations used to test the utility of different types of macrohabitat variables (Group 1) and different sources of seahorse data (Group 2). Note that one type of model variation (based on all occurrences and predictor Dataset 3) was used in both groups.

Model group	Species dataset	Predictor dataset	Model statistics
Group1: Testing the utility of different habitat variables *(3 models per species).*	All occurrences	Each model uses one of the **three** types: ***Dataset 1****: seven selected variables;* ***Dataset 2****: Dataset 1 + nine binary macrohabitat variables;* ***Dataset 3****: Dataset 1 + nine distance-based macrohabitat variables*	Area under the curve (AUC), Habitat variable importance (HVI), Predicted area ratio (PAR)
Group 2: Testing the utility of different sources of species data. *(7 models per species)*	Each model uses one of the **seven** types: *Research grade (**RG**), citizen science (**CS**), museum collection (**MC**), **CS+MC**, **RG+CS**, **RG+MC**, and all occurrences (**ALL**)*	Dataset 3 as described above	Sensitivity, I Similarity Statistic, Presence Agreement

3.3.4 Model description and evaluation

I executed two groups of model variations (Table 3.2) to examine my datasets. To do so, I created different predictor datasets and species datasets. First, I developed three different predictor datasets which represented the three methods of habitat data usage in the model (none, binary, continuous; Table 3.2): Dataset 1 – Seven selected predictors reflecting species' physiological suitability and primary productivity (Table 3.1); Dataset 2 – Combination of Dataset 1 and nine *binary* habitat variables (see Section *3.3.3*); and Dataset 3 – Combination of Dataset 1 and nine *continuous* habitat variables (see Section *3.3.3*). Secondly, I generated seven different datasets of species occurrences (Table 3.2): 1) research grade (RG), 2) citizen science (CS), 3) museum collections (MC), 4) CS plus MC, 5) RG plus CS, 6) RG plus MC, and 7) ALL (i.e. RG plus CS plus MC).

Model Group 1: To examine the utility of habitat variables, I built models for each species based on the species dataset ALL and each of the three predictor datasets (Datasets 1 – 3). I then compared the three models across species using three statistics: area under the curve of the receiver operating characteristic (AUC; Hanley and McNeil 1982), Habitat Variable Importance (HVI), and Predicted Area Ratio (PAR). The AUC is a standard technique for measuring models' omission and commission errors, producing a score (0 – 1) for general predictive accuracy. Score '1' means no errors of commission or omission, and '$\text{AUC} \leq 0.5$' means the model prediction is no better than random selection. I defined HVI as the cumulative permutation importance of all habitat variables, and used it as an index of species habitat dependency. The PAR is the proportion of the predicted area to the modeling envelope of each species (see Section *3.3.3*). Both HVI and PAR can be easily derived from Maxent's summary file (i.e., 'maxentResults.csv').

Model Group 2: To examine the utility of different sources of species data, I conducted models for each species based on the 'best performing' predictor dataset (derived from Model Group 1) and each of the seven datasets of seahorse occurrences: RG, CS, MC, CS plus MC, RG plus CS, RG plus MC, and ALL (i.e. RG plus MC plus CS). I compared the accuracy in predicting presences (i.e. Sensitivity; Altman and Bland, 1994) among the models, and estimated prediction agreement between models of individual species-data sources (RG, CS, and MC) based on I Similarity Statistic (Warren et al. 2008) and Presence Agreement (see results from Chapter 2). The former was used to estimate similarity between two presence-probability maps, while the latter evaluated overlapping rate between two predictive-presence maps.

To ensure robust statistical analyses, I only built models on species with a total sample size (i.e. the number of occurrences) ≥ 50 in Group 1, and executed models on species with a sample size ≥ 30 from each source (i.e. RG, CS, and MC) in Group 2. Detailed methods of Section 3.3.4 can be found in Supporting Information Appendix D.

3.3.5 Model prediction and interpretation

I applied Maxent to generate the presence-probability map for every species that had at least 5 occurrences (Pearson et al. 2007), based on the best predictor and species datasets identified above. To estimate species richness, I stacked the presence-probability maps of all species that had fair to excellent model performance (i.e. AUC ≥ 0.7; Calabrese et al. 2014). For species that I were unable to derive acceptable predictive maps or that had few occurrences (n < 5), I added their collated occurrences directly to corresponding pixels of the stacked map in ArcMap.

I applied Maxent's variable permutation importance and marginal response curve to interpret species-habitat relationships (Searcy and Shaffer 2016; Stirling et al. 2016). To this end, I first converted permutation importance values into ranks, with rank 1 assigned to the largest value (Supporting Information Appendix D). I then used the mean rank (of each variable across species) as an index of the general importance of that variable in determining seahorse distributions. I also calculated Habitat Variable Importance (HVI) for each seahorse species. I identified habitats that were positively correlated with species presence probability through examining the marginal response curves, and ranked them by permutation importance. This allowed me to determine the rank of importance of each habitat for each species.

3.3.6 Identifying potentially threatened species

With the above distribution information, I calculated and compared geographic metrics (extent of occurrence, area of occupancy) against the IUCN threatened thresholds for geographic range (Criteria B and D), and number of locations (Criterion D) to identify potentially threatened species, following the latest IUCN Red List Guidelines (IUCN Standards and Petitions Subcommittee 2017; Supporting Information Appendix D).

3.4 Results

I obtained a total of 6,128 unique occurrences (collated from the original 6,316 occurrences) for 42 species, with a wide range of sample size (1 – 1,990 occurrences per species). The temporal range spanned from 1828 to 2016 (96% from 1950 to 2016). Most occurrences had no location-precision information, except 448 records (precision = 727 ± 609 m). Most seahorse occurrences were sourced from MC (75%), followed by RG (18%) and CS (7%). In general, seahorses were found in a very wide geographic range (46.73° S to 54.75° N, 160.62° W to 179.12° E) and diverse environmental spaces (Table E.1 in Supporting Information Appendix E). According to the sample-size requirement to build robust models (see Section *3.3.4*), a total of 16 species were used in Model Group 1, and two species (*H. erectus* and *H. kuda*) were examined in Model Group 2.

3.4.1 Group 1: Utility of habitat predictors

My results indicated that including habitat variables had statistically significant improvements on model predictive accuracy, with models using continuous habitat predictors (Dataset 3) having the best performance (Fig. 3.2a). The AUC values (model performance) of Dataset 3 and Dataset 2 were significantly higher than those of Dataset 1 ($n = 16$, paired Wilcoxon tests; Dataset 3 vs. 1, $p < 0.005$; Dataset 2 vs. 1, $p < 0.01$), and Dataset 3 derived higher performance than Dataset 2 ($n = 16$, paired Wilcoxon test, $p < 0.005$).

I didn’t find statistically significant difference between Predicted Area Ratios estimated from different predictor datasets ($n = 16$, paired Wilcoxon tests, $p = 0.12, 0.12, 0.46$). But models based on Dataset 3 consistently derived the smallest predicted areas (Fig. 3.2b).

I found that Dataset 3 was more informative given it revealed the importance of habitats to seahorse distributions (Fig. 3.2c). On average, models using Dataset 3 derived high values of Habitat Variable Importance (46.2 ± 20.3), which were significantly larger than those using binary counterparts in Dataset 2 (9.2 ± 8.9; n = 16, paired Wilcoxon test, p < 0.001).

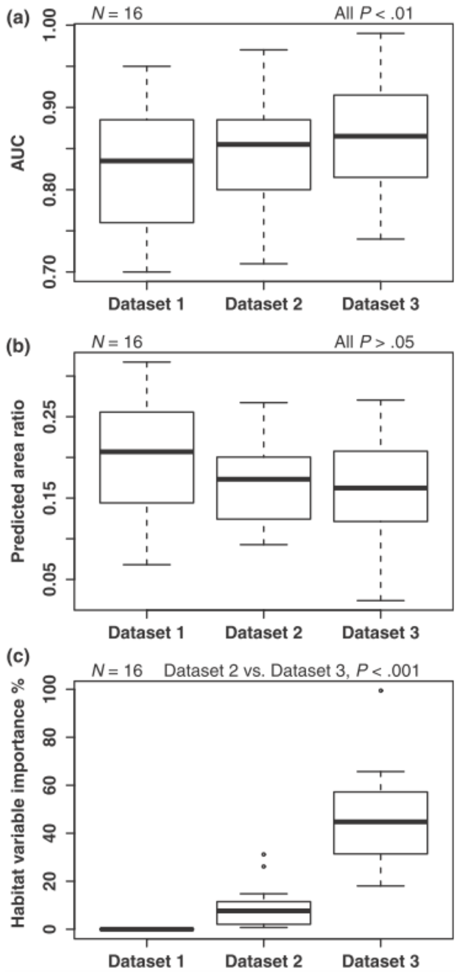


Figure 3. 2. Maxent models compared among three predictor datasets (Dataset 1, 2, and 3) across 16 species using a) AUC (model predictive accuracy), and b) predicted area ratio (proportion of predictive presences to the potential range), and c) habitat-variable importance (cumulative importance of all habitat variables used in the model). Dataset 1, using model predictors excluding macro-habitat variables; Dataset 2, using binary macrohabitat predictors (macrohabitat presence/absence) and Dataset 1; Dataset 3, using continuous macrohabitat predictors (distance to the nearest macrohabitat) and Dataset 1.

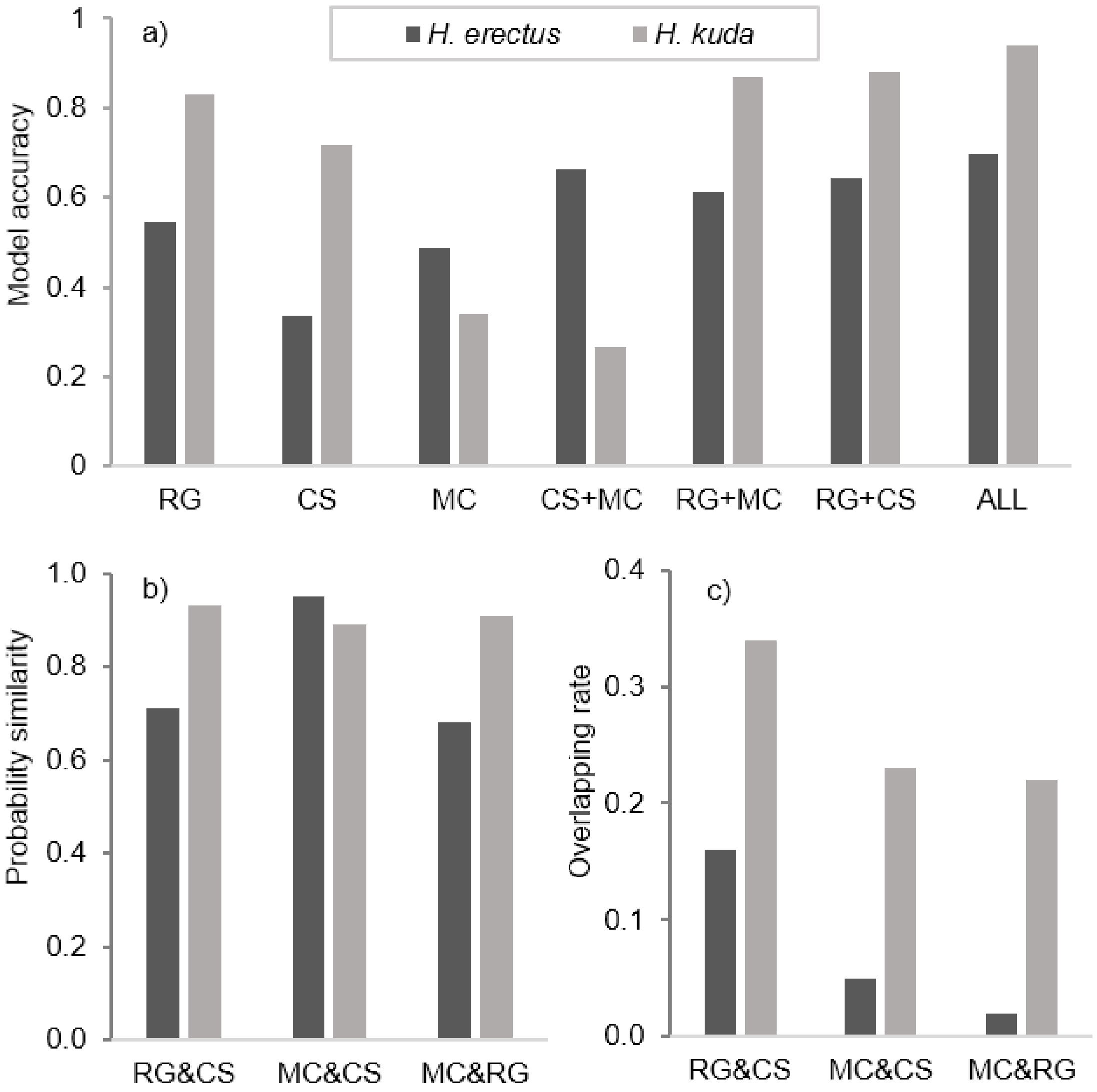


Figure 3. 3. Comparisons among Maxent models based on three different sources of seahorse occurrences, i.e. research grade (RG), citizen science (CS), and museum collections (MC), and their combinations. They were compared on two species (*H. erectus*, *H. kuda*) with sufficient sightings from each source (N ≥ 30). Panel a) depicts the model accuracy of different sources (RG, CS, MC) and combinations: CS+MC, RG+MC, RG+CS, and ALL (all three datasets combined). Panels b) and c) respectively demonstrates the probability similarity and the overlapping rate between each pair of models (RG&CS, MC&CS, MC&RG).

3.4.2 Group 2: Utility of different sources of species data

My comparisons indicated that CS and MC data derived less accurate models than RG data and integrating all three derive the best models (Fig. 3.3a). I found that RG consistently derived the most accurate models (Sensitivity = 0.55 on *H. erectus* and 0.83 on *H. kuda*) among the individual datasets (Fig. 3.3a). When individual datasets were combined, adding MC or CS to RG consistently improved model performance (Fig. 3.3a). Moreover, using all occurrences derived the most accurate models (Sensitivity = 0.69 and 0.94, respectively; Fig. 3.3a). Interestingly, while "CS plus MC" derived a model with the 2nd highest accuracy on *H. erectus* (0.66), it resulted in the lowest-accurate model on *H. kuda*. (0.27; Fig. 3.3a), suggesting that the reliability of this type of combination varies between species.

There were moderate to high consistencies (I Similarity = 0.68 – 0.95) in predicting presence probability, but low agreements in predictive presences (PA = 2% to 34%) between different species data sources (Fig. 3.3bandc). The I Similarity Statistics (i.e. presence-probability similarity) suggested that the agreement between RG models and CS models (0.71 and 0.93, respectively) were slightly higher than those between RG models and MC models (0.68 and 0.91, respectively; Fig. 3.3b). This difference was more prominent in the measure of Presence Agreement (i.e. overlapping rate, Fig. 3.3c).

3.4.3 Seahorse global distributions and biodiversity epicenters

Based on the above results, I chose Dataset 3 (the best predictor dataset) and dataset ALL (the best species dataset) to construct models for the 34 species with a least 5 occurrences (Tables E.2 and E.3). The remaining eight species with fewer occurrences were not modeled (Table E.4). Finally, I derived acceptable models for 33 species, with excellent performance (AUC = 0.90 – 1) for ten species, good performance (AUC = 0.80 – 0.89) for twelve species, and fair performance (AUC = 0.70 – 0.78) for eleven species (Table E.3). The model was poor (AUC = 0.56) and rejected for one species (*H. casscsio*, endemic to China and only recently described by Zhang et al. 2016).

The predicted biodiversity map demonstrated that locations with high species richness (value = 4 – 9) of seahorses were largely concentrated in tropical shallow waters of the central Indo-Pacific, with the epicenters mainly in the Philippines (Fig. 3.4 and 3.5, also see Fig. E.1 - E.13 in Supporting Information Appendix E). Other biodiversity epicenters near the central Indo-Pacific were in southern India and Sri Lanka (Fig. E.3), Taiwan (China, Fig. E.4), subtropical Japan (Fig. E.4), and Sydney and Melbourne in Australia (Fig. E.6). I also predicted moderate species richness (value = 3) in southern Florida (Fig. E.1), northern and central Red Sea (Fig. E.2), southeast Africa (Fig. E.2), Hawaii and Fiji (Fig. E.7).

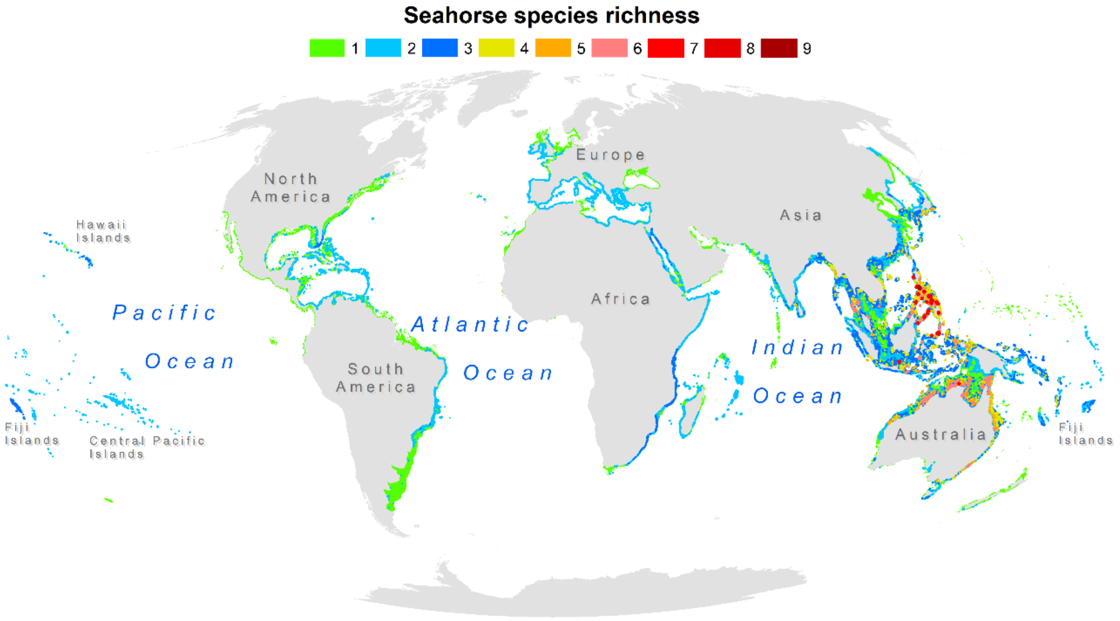


Figure 3. 4. Global map of biodiversity distributions of seahorse species (n = 42).

In total, the predicted suitable area for seahorses was 9 million km^2 (2.5% of the ocean's surface), with large extents of geographic separation among species. 84% of the 'potential range' (Fig. 3.1) was either unsuitable to seahorses (40.8%) or suitable to one species alone (43.2%). 15.5% of the potential range was suitable to two (13.3%) or three (2.3%) species together, and 0.4% was fit for more than three species together.

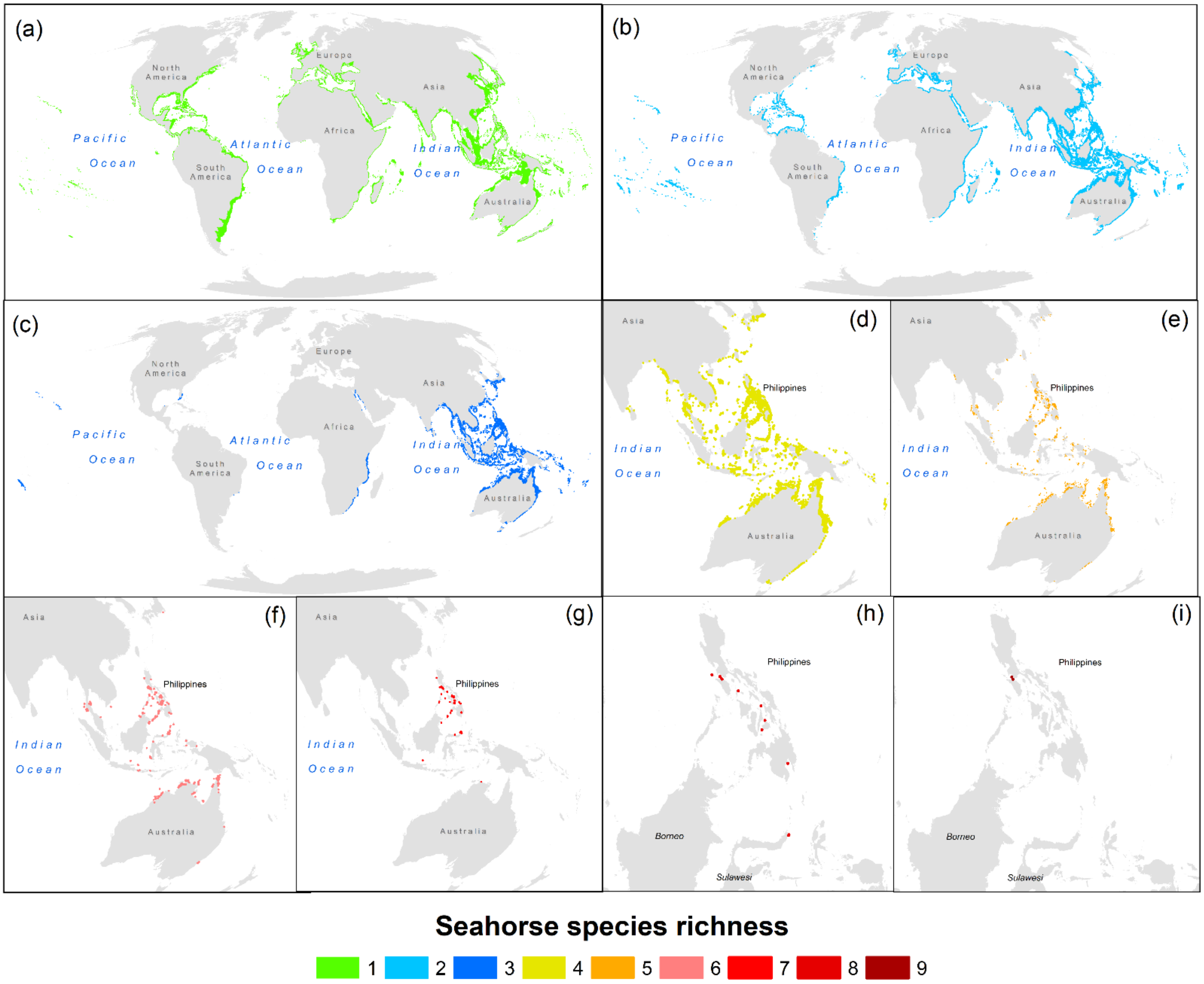


Figure 3. 5. Separated maps of different levels of seahorse species richness.

3.4.4 Key predictors and habitats

The ranks of predictor importance suggested that all 16 predictors provided valuable information in my models, although the most important factors varied among species (Fig. 3.6). On average, depth, distance to the nearest sponge, distance to the nearest macroalgae, pH, and ocean temperature (range and mean) were the most influential predictors (Fig. 3.6, Table E.5). The marginal response curves revealed that generally seahorses were more likely to live in shallower waters (< 50 m; Fig. 3.7 a); many species tended to live close to sponges but have divergent adaptations to other factors like ocean temperature (Fig. 3.7 b and 2.7 c). Habitat variable importance (HVI = 3.9 – 99%) and key habitat types differed largely among seahorse species (Table E.2). Two pygmy seahorses specializing in gorgonian corals, *H. denise* and *H. bargibanti,* had the highest HVI values.

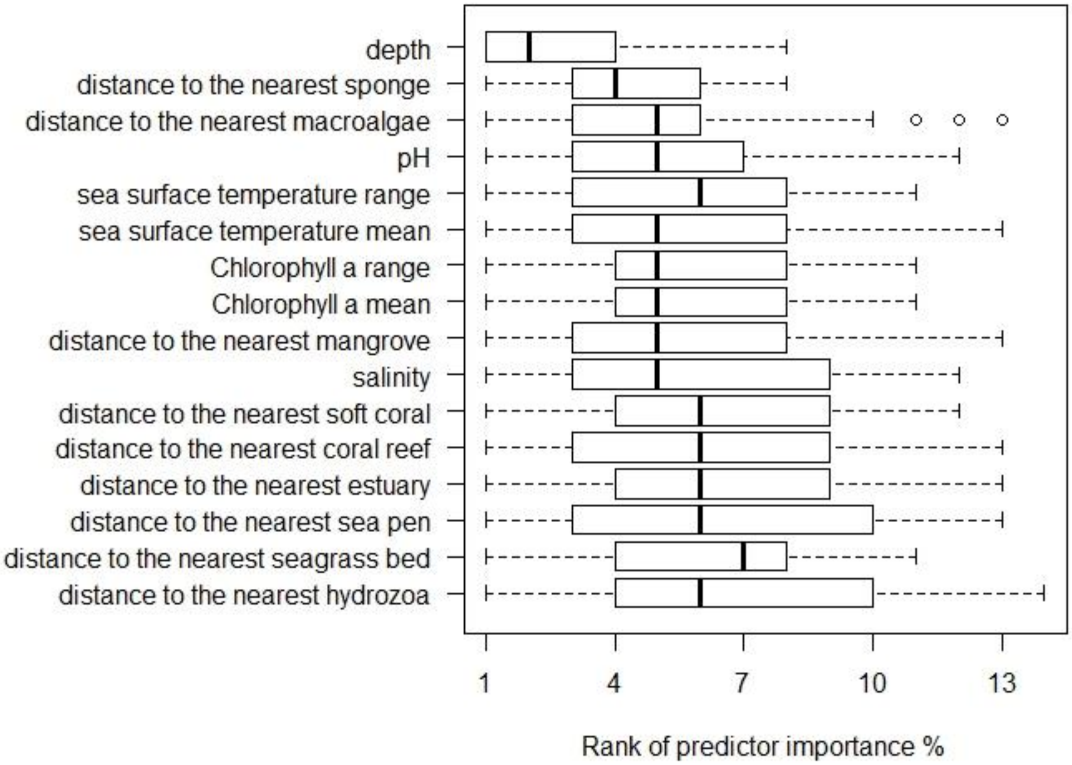
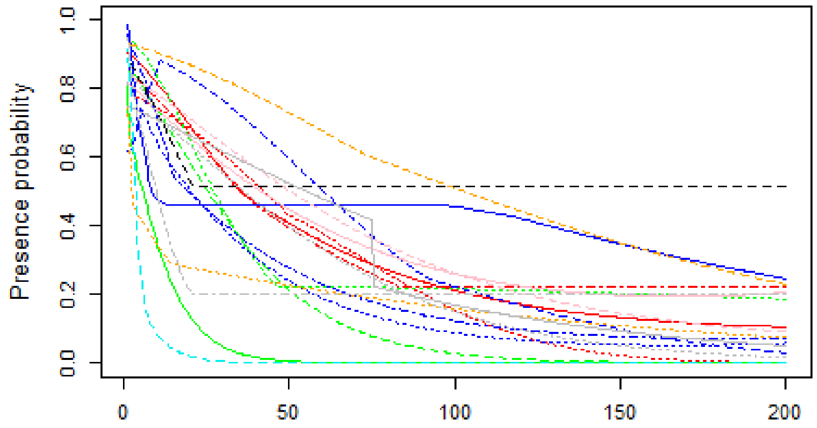
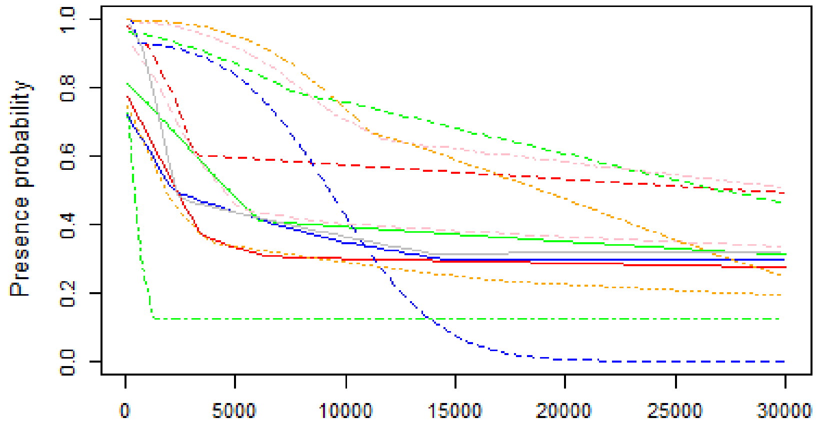


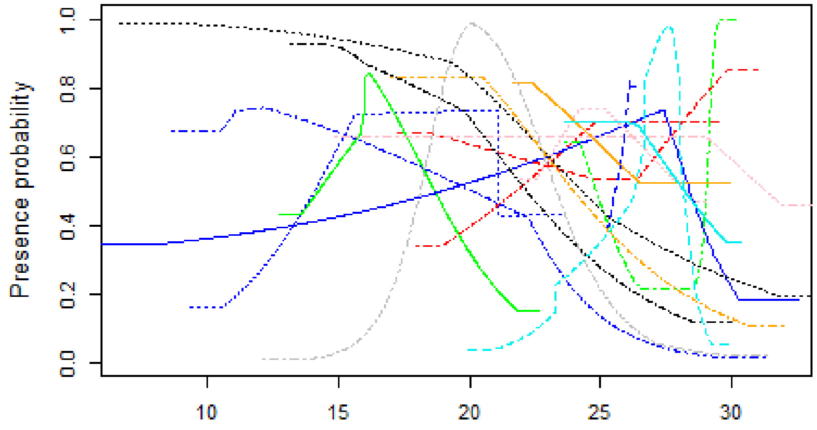
Figure 3. 6. Ranks of predictor importance across 33 seahorse species with acceptable models in our study. The boxplots are shown with the rank on the x-axis, and predictors on the y-axis ordered by the mean rank (not the median) from the highest (up) to the lowest (bottom). Note that lower numbers indicate higher ranks.



a) depth (m)



b) distance to the nearest sponge (m)



c) mean sea surface temperature (℃)

Figure 3. 7. Species response curves of a) depth, b) distance to the nearest sponge, and c) mean sea surface temperature, with different colors and types of line for different species. Note that not all species were shown in the curves, either because the importance of the variables was low (< 1%) or the species was not sensitive to the change of the variables.

3.4.5 Potentially threatened species based on IUCN Criteria

My results suggested that four species met the “Endangered” thresholds, and two met the “Vulnerable” thresholds based on geographic range alone (IUCN Criterion B; Table E.3). Among the six species, four met the IUCN thresholds based on Area of Occupancy (AOO) but not Extent of Occurrence (EOO), and the other two met thresholds for both AOO and EOO. However, information about population fragmentation, fluctuation or decline are required to fully justify these categories under Criterion B of IUCN. I also identified seven species with fewer than five known locations (Table E.4) that might be considered as “Vulnerable” based on IUCN Criterion D2.

3.5 Discussion

My study demonstrates that appropriate integration of multiple sources of species occurrences and habitat datasets is vital to derive robust SDMs to inform rare-species conservation. I provide global-scale, spatially-explicit maps and conservation knowledge that are urgently needed for a group of rare and data-poor marine fishes (i.e. seahorses). My analyses highlight that ‘proximity to habitats’ is more informative than ‘habitat presence/absence’ for improving model accuracy and detecting key habitats. Meanwhile, I indicate that it is better to combine CS and/or MC with RG, whenever available, to improve model accuracy. My study also demonstrates that SDM-based predictive maps can help to identify potentially threatened species with small Area of Occupancy.

3.5.1 Global biogeographic pattern and habitat associations of seahorses

The predicted biogeographic pattern of seahorse species from my study is consistent with a previous analysis of coastal fishes in general (Tittensor et al. 2010). The latitudinal gradient that generally more species live in the tropics may be largely shaped by temperature (Tittensor et al. 2010). Temperature variables can influence food availability and climate suitability (Willig et al. 2003), and had high importances in my models. But like other shore fishes, more species of seahorses are predicted to occur in the central Indo-Pacific (especially the Philippines) versus other tropical regions (e.g. Caribbean; Tittensor et al. 2010). This longitudinal divergence matches well with the hypothesized footprints of seahorse origin (probably in northeastern Australia), dispersal (by rafting), segregation (e.g. the closure of the Isthmus of Panama), and evolution over past 20 - 30 million years (Casey et al. 2004, Teske et al. 2004, Boehm et al. 2013). Higher availability and heterogeneity of shallow-water habitats in the central Indo-Pacific might facilitate species immigration and diversification (Sanciangco et al. 2013), as I find that seahorses tend to live in shallow depths and have diverse habitat preferences among species.

My study provides comprehensive, global evidence to support the hypothesis that seahorse species have different levels of habitat reliance and species-specific habitat preferences (Curtis and Vincent 2005). I am the first to provide quantitative evidence that habitat dependency varies largely among seahorse species. I demonstrate global divergences on distributions and habitat associations among species, which has been observed in local studies (Curtis and Vincent 2005; Lourie et al. 2005; also see Chapter 2). Similar divergence patterns among congeneric species are also common in pipefishes (Malavasi et al. 2007), and other fishes (Lombarte et al. 2000, Fairclough et al. 2008). Such segregation and difference may be important to minimize competition among closely-related species (Fairclough et al. 2008). My evidence suggests that habitat-forming sponges are important to many seahorse species at large spatial scales. These sponges may provide various functional roles (e.g. shelter, prey sources) in seahorse's life (Bell 2008). Further discussion about seahorse-habitat/environment relationships are presented in Supporting Information Appendix S4.

3.5.2 Conservation implications for seahorses

The spatially-explicit maps and habitat-preference knowledge derived from my models may inform both global- and local-scale seahorse conservation. For instance, to minimize the impact of international trade, CITES Parties can use my maps to locate focus areas rich in those heavily-traded species at the global scale (Vincent et al. 2011). These maps can further be used with threat maps if available to determine global priority areas for seahorse conservation. Conservation programs can use and validate my maps and habitat knowledge to study and protect seahorse populations at local scales.

My study provided new geographic information to assess conservation status for seahorses. IUCN requires assessors to evaluate species against all criteria (A – E) with available data, and to assign the most severe category to the species (IUCN Standards and Petitions Subcommittee 2017). Previous assessments of seahorses largely applied Criterion A (i.e. population decline rate). Only one case (*H. capensis*) used Criterion B1 (Extent of Occurrence). I identified six species that met threatened thresholds of Criterion B2 (Area of Occupancy, AOO). Currently, one (*H. capensis*) has been evaluated as Endangered, but other five were either “Least Concern” or “Data Deficient” (Table E.3, IUCN 2018). Although the AOO is estimated based on predictive maps and thus contain uncertainties and require further improvement (Guisan et al. 2013), it is likely that I overestimated the AOO. If I include other constraint factors (e.g. anthropogenic impacts) in my models, the AOO might become smaller. Given that, these five species can still justify the IUCN ‘Vulnerable’ threshold of AOO (< 2,000 km^2) for Criterion B2.

3.5.3 Utility of distance-based habitat predictors in SDMs

I am the first to demonstrate that ‘proximity to macrohabitats’ is more informative than ‘habitat presence/absence’ to predict distributions of low-mobility organisms. Proximity to macrohabitats has been employed and proved useful in SDMs for high-mobility animals, including reef fishes (Shelton et al. 2014), Bonelli’s eagles (Balbontín 2005), and bats (Rainho and Palmeirim 2011). One underlying assumption is that high-mobility animals choose to live close to important resource patches (e.g. feeding grounds). But this assumption might be questionable for low-mobility species, as they were believed unlikely able to select habitats at large spatial scales. Instead, site-level habitat characteristics (e.g. habitat presence/absence) were used in low-mobility species including Madagascar geckos (Pearson et al. 2007) and Juliana’s golden mole (Jackson and Robertson 2011). Although a few studies have used proximity to macrohabitats for low-mobility species (Dorrough and Ash 1999, Dillard et al. 2008), mine is the first to indicate that proximity to macrohabitats can be more useful than habitat presence/absence for low-mobility animals.

The correlations between sedentary animals and the proximity to habitats may result from two factors: 1) behavior and ecology of the animal, and 2) coarse resolution of original data. First, low-mobility organisms can disperse across large spatial scales through natural disturbance (e.g. ocean currents) and hitchhiking (Luiz et al. 2015). They may then choose 'stops' in preferred habitats or be 'dropped' in unsuitable ones. In the second case, the animal may move over a distance longer than its home range to find suitable habitats (Matthews 1990, Caldwell and Vincent 2013). Some sightings of the species might be recorded during this 'habitat-finding' process and thus distorted the species-habitat spatial relations. Second, there might be spatial mismatch between species occurrences and habitat locations due to coarse resolutions of the original records. As a result, some species occurrences might not overlap with suitable habitats, but they are still close in space.

3.5.4 Utility of different sources of species data in SDMs

This study is among the first to demonstrate that adding citizen science (CS) and museum collections (MC) to research-grade data (RG) can help to derive more accurate SDMs. I encourage the integration of MC and/or CS data with RG data as MC and CS can be more sufficient and helpful for improving model accuracy. MC are usually the most data-rich source for many organisms (Ponder et al. 2001), as is also shown in my study of seahorses. MC have been commonly used in SDM research (Newbold 2010), although I demonstrate that, compared with RG, MC may derive lower quality models with quite different predictions (Aubry et al. 2017). This is also true for citizen science. My results suggest that combining quality-unknown data (MC and CS) without RG is risky, as this may accumulate errors and result in lower-accuracy models (see Discussion in Chapter 2).

Additionally, compared with MC, CS might have some advantages. First, CS provide more recent geo-referenced information than historical collections to reflect the current distributions of species. Second, validating CS data might be easier than checking the historical specimens from worldwide museums. These advantages may partly explain why CS derived more similar predictions to those of RG data than MC did in my study. A recent study on a rare snow quail species also indicated that CS data could derive similar predictions to those based on RG data (Jackson et al. 2015), encouraging the use of citizen-science data sources.

3.5.5 Applying SDMs to IUCN Red List assessment

My study highlights that applying SDMs can derive useful geographic information for the assessment of conservation status for rare species (Aubry et al. 2017). Rare species (especially habitat specialists) that have low population density are often patchily distributed along with their key habitats (Marcer et al. 2013). Therefore, even though they may have large extent of occurrence (EOO, IUCN Criterion B1), their area of occupancy (AOO, IUCN Criterion B2) could be very small. For example, in my study, *H. fisheri* (a species endemic to Hawaii) has an EOO that does not meet a threatened threshold, but its AOO is smaller than the "Endangered" threshold of AOO according to IUCN Criteria B2. Therefore, estimating AOO based on SDMs like Maxent may add essential information to assess conservation status for rare species.

Chapter 4: Cumulative human impact models reveal threats for seahorses (*Hippocampus* spp.)

4.1 Summary

Understanding cumulative human impact (CHI) on marine organisms is vital, given that they are threatened by multiple anthropogenic pressures. Here we provide a global-scale study on human impacts on, and conservation status of, a genus of rare marine fishes, seahorses (*Hippocampus* spp.). We developed species-level models to assess and map the cumulative impact of 12 anthropogenic stressors on 42 seahorse species, based on expert knowledge and spatial datasets. I then compared the CHI estimates between 'threatened' and 'un-threatened' species listed on the IUCN Red List. To predict conservation status for 'Data Deficient' species, I built random forest models based on the derived human impact indices from the CHI models. I mapped the CHI on seahorses and compared it with CHI on marine ecosystems. The results indicate that my CHI estimates for 'threatened' species are significantly higher than counterparts for 'un-threatened' species. The random forest models suggest that five of the 19 'Data Deficient' species are 'threatened'. The major stressors that determine conservation status are demersal fishing with high bycatch and (ocean and nutrient) pollutions. The threat epicenters with high CHIs on seahorses concentrate in the East and South China Seas, Southeast Asian waters, and European waters. I show that impacts on seahorses are more likely higher in shallower inshore waters compared with previous estimates on marine ecosystems, with only a medium correlation between them. My study highlights the importance of developing species-level CHI models to better estimate and map threats on specific organisms. I provide useful maps to guide threat management on seahorse species. My approaches might be useful to analyze threats and conservation status for other marine species, especially data-poor fishes.

4.2 Introduction

Understanding which marine species are most threatened and where threat epicenters are located has long been a concern in marine conservation. The past few centuries has witnessed an increasing intensity of human disturbances in the ocean, which caused prominent ecosystem degradations and population depletions of many marine organisms (Halpern et al. 2008, McCauley et al. 2015). To address this pressing issue, we need first to identify those threatened species and understand the underlying drivers. Such work has been mainly done by the International Union for Conservation of Nature (IUCN) Red List assessments (www.iucn.redlist). While evaluating threatened status, we also want to know where species are most impacted so that conservation actions may be guided towards to these places. Such spatial knowledge has long been required (Gundlach and Hayes 1978), but a systematic methodology only occurred ten years ago with a study of global cumulative human impact (CHI) on marine ecosystems (hereafter, ecosystem-scale CHI; Halpern et al. 2008). This study has since motivated an increasing amount of assessments of human impacts on various marine ecosystem components, including specific organisms and habitats (see a review by Korpinen and Andersen 2016).

Despite the IUCN’s efforts, the conservation status of many marine taxa is still poorly known (McCauley et al. 2015). To date, the majority of marine species are not evaluated by IUCN and thousands of the evaluated species are ‘Data Deficient’ (IUCN 2018). For instance, more than half of the marine fish species are still not evaluated by IUCN; among the 8,905 evaluated ones (Actinopterygii, Cephalaspidomorphi, and Chondrichthyes), nearly 19% are ‘Data Deficient’ (IUCN 2018). One major reason is the difficulty in collecting sufficient data in the ocean (McClenachan et al. 2012). Given this challenge, scientists have started to use alternative approaches to assess extinction risks of data-poor marine species. For instance, some biologists have identified extinction-prone traits such as body size to evaluate vulnerability or predict extinction risks for marine organisms (Dulvy et al. 2004, Reynolds et al. 2005).

Studies on cumulative human impact (CHI) at species levels might provide a promising approach to quantify threats for data-poor marine organisms. Previous CHI models usually combine the intensity of anthropogenic stressors and the vulnerability of ecosystem components (e.g. habitats/species) to different stressors (Korpinen and Andersen 2016). Therefore, if I derive a CHI index for the focal species based on the model, such index might be useful predictors of species threatened status. Some colleagues have already examined this question but might not use appropriate CHI data (Davidson et al. 2012). In their study, a CHI index (measured for marine ecosystems) and species intrinsic factors (e.g. body mass at weaning) were used to predict threatened status for marine mammals. They found that the CHI index had very little significance in predicting threatened status for these mammals (Davidson et al. 2012). However, the conclusion might be different if the used CHI index was estimated based on the vulnerability of the focal mammals, rather than ecosystems, to anthropogenic stressors. The vulnerability measures might be highly different when it was compared between a given ecosystem and a species (Maxwell et al. 2013). However, current CHI studies mainly focus on mapping impact on ecosystems rather than deriving impact indices to predict status for particular species (Korpinen and Andresen 2016).

Seahorses (*Hippocampus* spp.) are data-poor marine fishes that are naturally susceptible to various human activities (Vincent et al. 2011). Seahorses are rare and sedentary fishes living in shallow waters (generally < 50 m, (Foster and Vincent 2004). Unlike many other marine fish, seahorses have relatively low fecundity, mate fidelity, and lengthy parental care (Foster and Vincent 2004). These characteristics can make them highly exposed and sensitive to human activities including demersal high-bycatch fisheries such as bottom trawling. A recent estimate showed that annually about 37 million seahorses were caught as bycatch from 22 countries alone (Lawson et al. 2017). By far, 17 seahorse species are Data Deficient on the IUCN Red List, while 14 are ‘threatened’ (2 Endangered and 12 Vulnerable) and another 11 are ‘non-threatened’ (10 “Least Concern” and one “Near Threatened”; IUCN 2018).

This study aims to 1) create species-level models to estimate and map cumulative human impact (CHI) for seahorses, 2) examine whether the model outputs can be used to predict threatened status for Data Deficient species, 3) identify regions where seahorse populations are under high impacts, and 4) distinguish major stressors. To do so, I build CHI models based on various spatial datasets and expert knowledge on species vulnerability to various human activities. I compared the derived CHI indices between threatened and unthreatened species according to IUCN Red List assessments. By doing so, I examined whether my CHI indices were consistent with IUCN assessments. I then utilized machine-learning models to correlate threatened status with human impact indices. I explored the importance of each stressor and identified the most-accurate models to predict threatened status for Data Deficient species.

4.3 Methods

4.3.1 Species-level Cumulative Human Impact (CHI) Models

Based on CHI studies for other marine species (Maxwell et al. 2013), I developed two 'linear-additive' models to estimate and map CHI at the species level for seahorses (Table 4.1). One is a 'non-spatial' model, which was used to derive a CHI index for each species. The other is a 'spatial' model, which was developed for mapping CHI for each species at 1 km by 1 km resolution.

Non-spatial CHI model

I calculated the human-impact index ($HI_{i,j}$) of stressor ***i*** upon species ***j*** with the following model:

$$HI_{i,j} = I_{i,j} \times ED_{i,j} \times S_{i,j} = I_{i,j} \times (PS_{i,j} \times F_{i,j}) \times S_{i,j} \qquad (1)$$

where $I_{i,j}$ is the intensity of stressor ***i*** across species ***j***'s habitat; $ED_{i,j}$ is the exposure degree of species ***j*** to stressor ***i***; $S_{i,j}$ is the sensitivity (degree of intolerance and incapability of recovery) of species ***j*** to stressor ***i***; $PS_{i,j}$ is the proportion of the species ***j***'s habitats experiencing the stressor ***i***; and $F_{i,j}$ is the corresponding frequency (i.e., how often the stressor ***j*** occurs to species ***i***). Here $PS_{i,j}$ and $F_{i,j}$ were used to measure exposure degree spatially and temporally.

The cumulative human impact (CHI_j) of all stressors upon species ***j*** was estimated by adding up $HI_{i,j}$ across stressors:

$$CHI_j = \sum_i HI_{i,j} = \sum_i I_{i,j} \times (PS_{i,j} \times F_{i,j}) \times S_{i,j} \qquad (2)$$

Table 4. 1. Summary of the two cumulative-human-impact models used to estimate and map human impacts for each seahorse species. $HI_{i,j}$, human impact of stressor i upon species j; $HI_{i,j,m}$, human impact of stressor i on species j in location m; $I_{i,j}$, the intensity of stressor i across species j's habitat; $ED_{i,j}$, the exposure degree of species j to stressor i; $I_{i,j,m}$ and $ED_{i,j,m}$, respectively the intensity and exposure degree of stressor i to species j in location m; $P_{j,m}$, the presence probability of species j in location m; $S_{i,j}$, the sensitivity (degree of intolerance and incapability of recovery) of species j to stressor i; $PS_{i,j}$, the proportion of the species j's habitats experiencing the stressor i ($PS_{i,j}$); $F_{i,j}$, the corresponding frequency (i.e., how often the stressor j occurs to species i); CHI_j, cumulative human impact on species j; $CHI_{j,m}$, cumulative human impact on species j at location m.

Model	**Purpose**	**Model equations**	**Variable explanations**
Non-spatial CHI model	Estimate human impact indexes (human impact of individual stressor, and cumulative human impact of all stressor) at species level.	$HI_{i,j} = I_{i,j} \times ED_{i,j} \times S_{i,j}$ $= I_{i,j} \times (PS_{i,j} \times F_{i,j}) \times S_{i,j}$ $CHI_j = \sum_i HI_{i,j} = \sum_i I_{i,j} \times (PS_{i,j} \times F_{i,j}) \times S_{i,j}$	Only require estimates of stressors' intensities ($I_{i,j}$), species exposure degree ($PS_{i,j} \times F_{i,j}$) and sensitivities ($S_{i,j}$) to the stressors.
Spatial CHI model	Map human impact indexes and can also calculate human impact indexes across species habitat.	$HI_{i,j,m} = I_{i,j,m} \times ED_{i,j,m} \times S_{i,j}$ $= I_{i,j,\mathrm{m}} \times (P_{j,m} \times F_{i,j}) \times S_{i,j}$ $CHI_{j,m} = \sum_i HI_{i,j,m} = \sum_i I_{i,j,m} \times (P_{j,m} \times F_{i,j}) \times S_{i,j}$	Only stressor's intensity ($I_{i,j,m}$) and species presence-probability ($P_{j,m}$) require spatial datasets; other variables are the same as non-spatial model.

Spatial CHI model

Correspondingly, I defined a spatial CHI model (adapted from Maxwell et al. 2013) to estimate the human impact ($HI_{i,j,m}$) of stressor $\boldsymbol{i}$ on species $\boldsymbol{j}$ in location $\boldsymbol{m}$ (here, a 1 km by 1 km cell in species habitat):

$$HI_{i,j,m} = I_{i,j,m} \times ED_{i,j,m} \times S_{i,j} = I_{i,j,m} \times (P_{j,m} \times F_{i,j}) \times S_{i,j} \qquad (3)$$

where $I_{i,j,m}$ and $ED_{i,j,m}$ are respectively the intensity and exposure degree of stressor $\boldsymbol{i}$ to species $\boldsymbol{j}$ in location $\boldsymbol{m}$; $P_{j,m}$ is the presence probability of species $\boldsymbol{j}$ in location $\boldsymbol{m}$; while $F_{i,j}$ and $S_{i,j}$ are the same as in the non-spatial model.

I then estimated the cumulative human impact ($CHI_{j,m}$) on species $\boldsymbol{j}$ in location $\boldsymbol{m}$ with the following model:

$$CHI_{j,m} = \sum_i HI_{i,j,m} = \sum_i \left[I_{i,j,m} \times (P_{j,m} \times F_{i,j}) \times S_{i,j} \right] \qquad (4)$$

I also generated human impact indices from the spatial model, with the purpose to compare the indices between the spatial and non-spatial models. The overall human impact of stressor $\boldsymbol{i}$ upon species $\boldsymbol{j}$ across its habitat was estimated by the following model:

$$HI_{i,j} = \frac{\sum_m (I_{i,j,m} \times ED_{i,j,m} \times S_{i,j})}{\sum_m P_{j,m}} = \frac{\sum_m [I_{i,j,m} \times (P_{j,m} \times F_{i,j}) \times S_{i,j}]}{\sum_m P_{j,m}} \qquad (5)$$

The cumulative human impact of all stressors ($\boldsymbol{i}$ = 1, 2, …, n) upon species j across the habitat was calculated by the following model:

$$CHI_j = \sum_i HI_{i,j} = \sum_i \frac{\sum_m [I_{i,j,m} \times (P_{j,m} \times F_{i,j}) \times S_{i,j}]}{\sum_m P_{j,m}} \qquad (6)$$

In Eq. 5 and 6, I used a numerator ($\sum_m P_{j,m}$) to eliminate the influence of species' suitable-habitat size on the value of overall impact. This numerator was the sum of habitat suitability ($P_{j,m}$) multiplied by the pixel size (1 km^2) across species' geographic range. By doing so, I was able to compare these impact indices ($HI_{i,j}$ and CHI_j) among species.

Table 4. 2. Summary of 12 anthropogenic stressors on seahorses. All datasets are from http://ohi-science.org/data/.

Category	Measures	Description	Original resolution
Fishing	Demersal destructive fishing (DDF)	bomb/chemical (all types), bottom trawl (all types), dredge (all types)	half-degree
	Demersal, non-destructive, high-bycatch fishing (DNHF)	gillnet fixed, gillnet encircling, seine, beach seine, boat seine, trammel (all types), trap (all types)	half-degree
	Pelagic high-bycatch fishing (PHF)	driftnets, midwater trawl	half-degree
	Artisanal fishing (AF)	an estimate that cannot distinguish between methods that do and do not modify habitat.	modeled 1 km^2
Habitat change	Habitat destruction caused by fishing (HDF)	bomb/chemical (all types), bottom trawl (all types), dredge (all types)	half-degree
	Coastal development (CD)	estimated by population density + oil rigs	modeled 1 km^2
Pollution	Nutrient pollution (NuP)	modeled based on fertilizer data	modeled 1 km^2
	Ocean pollution (OP)	mainly oil spills estimated by commercial shipping intensity plus port volume	modeled 1 km^2
	Noise pollution (NoP)	estimated by the intensity of shipping + all kinds of fishing + oil rigs + population density	modeled 1 km^2
Invasive species	Invasive species (IS)	the amount of cargo at ports	modeled 1 km^2
Climate change	SST abnormality (SSTA)	frequency of SST passes the long-term mean SST (i.e. threshold)	~21 km^2
	Ocean acidification (OA)	differences on the aragonite saturation state of the ocean between pre-industrial (circa 1870) and modern times (2000 - 2009)	1 degree

Non-spatial and spatial models were similar in structure (Eq. 1 vs. Eq. 3, Eq. 2 vs. Eq. 4) and share some variables ($F_{i,j}$ and $S_{i,j}$). In the spatial CHI model, the stressor's intensity ($I_{i,j,m}$) and species presence-probability ($P_{j,m}$) require spatial datasets, which could constrain its utility for data-poor species. A further model explanation including some assumptions is given in Appendix G.

To construct the above CHI models, I identified 12 anthropogenic stressors (Table 4.2) on the 42 valid seahorse species (Lourie et al. 2016, Zhang et al. 2016) (see a detailed description in Appendix G). These stressors covered five dimensions of human-derived threats commonly found on marine species: fishing (here, high by-catch fisheries), habitat change (i.e., degradation/loss), pollution, invasive species, and climate change (Halpern et al. 2008).

4.3.2 Data collection for model variables

I first gathered the spatial datasets to separately estimate $I_{i,j,m}$, $I_{i,j}$, and $P_{j,m}$. The original intensity dataset of all 12 stressors were extracted from a global-scale study (Halpern et al. 2015) (see a description in Appendix G). I then derived $I_{i,j,m}$ from this global-scale dataset for the spatial CHI model. This was done by extracting the original intensity maps of stressor *i* into the distribution range of species j (see Methods in Chapter 3). I then averaged the intensities of stressor *i* across species j's distribution range as the estimate of $I_{i,j}$ for the non-spatial CHI model. The dataset of $P_{j,m}$ for the spatial CHI model was from Chapter 3. This study produced robust presence-probability maps for 33 seahorse species based on species distribution models. I was unable to derive similar maps for the other nine species mainly because of data paucity, and therefore they were not assessed in the spatial CHI model. But all 42 species were involved in the non-spatial CHI model. The intensity ($I_{i,j,m}$) and presence probability ($P_{j,m}$) datasets were projected at 1 km^2 resolution under an equal area projection (WGS84 Mollweide) to prevent the distortion of the "area" property (Halpern et al. 2008).

I estimated the remaining vulnerability variables, i.e., exposure degree and sensitivity ($PS_{i,j}$, $F_{i,j}$, $S_{i,j}$), by eliciting expert knowledge. I conducted expert-opinion surveys among seahorse experts around the world through emails, skypes and phones (Halpern et al. 2007, Teck et al. 2010) (see details in Appendix H1). In the survey, I asked experts to score each vulnerability variable for each species-stressor pair. The scores of the exposure degree measures (i.e. $PS_{i,j}$ and $F_{i,j}$) were estimated into five levels ranging from 0 to 4 ($PS_{i,j}$) and from 1 to 5 ($F_{i,j}$), respectively. Higher scores represented higher exposure degrees. The score of sensitivity ($S_{i,j}$) was assigned to three levels: low, 1; medium, 3; and high, 5. For species that lacked expert evaluations, I assessed them based on the literature review and experts' knowledge of other species which share the similar habitats (see a review in Appendix G and Vincent et al. 2011). For each estimation, experts were asked to provide a certainty index to reflect their confidence levels: 1, not sure; 2, likely; 3, confident; 4, very confident. I then used the certainty-weighted average to estimate each vulnerability variable for each species-stressor pair (Halpern et al. 2007; see Appendix H).

4.3.3 CHI Model computations and comparisons

I calculated species-level impact indices and plotted individual and cumulative human impacts ($HI_{i,j}$ and CHI_j) based on the equations of both non-spatial and spatial CHI models. I derived the CHI index from the non-spatial model for each of the 42 species (hereafter, CHI_n). Given that only 33 species had presence-probability maps ($P_{j,m}$, see Results from Chapter 3), I only calculated CHI indices for these species from the spatial model (hereafter, CHI_s). I also derived an 'ecosystem-scale' CHI index for these 33 species (hereafter, CHI_e). This index was estimated based on the global map of CHI on marine ecosystems (from Halpern et al. 2015) and species presence-probability maps for the 33 species (see Appendix G). The global CHI map was based on a different set of vulnerability measures of marine ecosystems (rather than seahorse species) to a total of 18 stressors (rather than 12 stressors). I aimed to examine whether the species-level CHIs would differ from its counterpart estimated at the ecosystem level (CHI_e) in seahorse habitats. To this end, I examined their correlations using Spearman rank correlation and simple linear regression models.

I compared the above CHI indices (CHI_n, CHI_s, and CHI_e) with IUCN Red List assessments. This was done by examining whether CHI indices of "threatened" species were significantly higher from the counterparts of "non-threatened" species (Mann-Whitney U tests). Seahorse species with IUCN threatened categories (here, Vulnerable and Endangered) were considered as "threatened" (n = 14), species with Near Threatened or Least Concern categories were considered as "non-threatened" (n = 18; Davidson et al. 2012).

4.3.4 Predicting extinction risks for Data Deficient species

I correlated the impact indices (individual impact and CHI) and species status (threatened/non-threatened) using a random forest (RF) model (R package *randomForest*) of 500 classification trees (Liaw and Wiener 2002, Cutler et al. 2007). I applied the Boruta algorithm (R package *Boruta*) to select important predictors by iteratively comparing original predictors' importance with importance achievable at random using permuted copies of all predictors (a.k.a., shadow attributes) (Kursa and Rudnicki 2010). The predictor's importance was determined by using the Z score, which was the mean accuracy loss resulting from permutation of the focal variable divided by its standard deviation across the trees in the random forest. To identify the 'minimal-optimal' predictors which could derive the most accurate model, I then rerun the model by removing the least important predictor stepwise. Model accuracy was measured with the percentage of species correctly classified based on model predictions (Cutler et al. 2007, Davidson et al. 2012), and Cohen's kappa statistic (function kappa2 in R package *irr*) (Fielding and Bell 1997, Gamer et al. 2007). The latter measures the agreement between predicted and actual status, while correcting for random effect (Prasad et al. 2006, Cutler et al. 2007). I equally weighted classification costs in the RF models.

All above processes were done respectively for the impact indices (i.e., model predictors) derived from my spatial and non-spatial CHI models. Therefore, two best RF models with optimal predictors were derived separately. I then applied these two best RF models to predict threatened status for Data Deficient species.

4.3.5 Identifying threat patterns: threat epicenters and major drivers

I identified threat epicenters and major drivers for the threatened seahorse species based on my spatial CHI model outputs. I first plotted the CHI of threatened species (including the predicted ones by this study) by overlaying their CHI maps in ArcMap, with the sums assigned to overlapping cells (Halpern et al. 2008, Maxwell et al. 2013). I classified the ultimate CHI values to five levels (Very Low, Low, Medium, High, and Very High) based on the Jenks Natural Breaks algorithm in ArcMap (De Smith et al. 2007). This classification algorithm was based on the natural grouping inherent in the data, which maximized the difference between classes and similarity within each class. For reporting purpose, locations with ‘High’ to ‘Very High’ impacts (i.e., threat epicenters) were highlighted. I then plotted intensity maps of the identified truly important stressors across these epicenters.

I conducted sensitivity analyses for my methods of determining CHI’s spatial patterns. I explored the effects of using different versions of species groups (threatened-only vs. all 33 species) and CHI algorithms (additive vs. average) for overlapping pixels. The effects were tested on the derived CHI’s spatial patterns and its correlation to species richness (see details in Appendix G). I also compared the additive CHI map of all seahorse species against the additive CHI map of marine ecosystems (extracted from Halpern et al. 2015) to explore how they might differ in space.

4.4 Results

4.4.1 Cumulative human impact indices

I obtained wide ranges of CHI indices from the non-spatial model (CHI_n = 1.22 – 63.40, n = 42) and the spatial model (CHI_s = 0.78 – 16.23, n = 33). The highest values of the two models were given to a Vulnerable species, *H. kelloggi*. The lowest CHI_s was given to a Data Deficient species (*H. camelopardalis*), and the lowest CHI_n to another Data Deficient species (*H. pusillus*). The CHI index based on vulnerability measures of marine ecosystems (CHI_e) had a much wider range of values (0.17 – 8103.24, n = 33), with the highest value given to an Endangered species (*H. capensis*) but the lowest value to another Endangered species (*H. whitei*).

I found a strong rank-order correlation between the indices derived from my two CHI models (CHI_n vs. CHI_s: r = 0.91, P < 0.001), but not between the ecosystem-scale index and the two species-level indices (CHI_e vs. CHI_n, r = - 0.09, P = 0.66; CHI_e vs. CHI_s, r = - 0.21, P = 0.33). I found that CHI_n and CHI_s had a high linear correlation ($R^2 = 0.78$; Fig. I.1 in Appendix I). The impact indices derived from my species-level CHI models were significantly higher in the 'threatened' species (CHI_n = 23.5 ± 18.2, CHI_s = 7.7 ± 3.7) than those in the 'non-threatened' species (CHI_n = 8.2 ± 6.7, CHI_s = 4.7 ± 3.1; Mann-Whitney U tests, both $P < 0.01$; Fig. 4.1). However, the ecosystem-scale index (CHI_e), however, didn't differ between the 'threatened' species and 'non-threatened' species (Mann-Whitney U tests, $P = 0.71$).

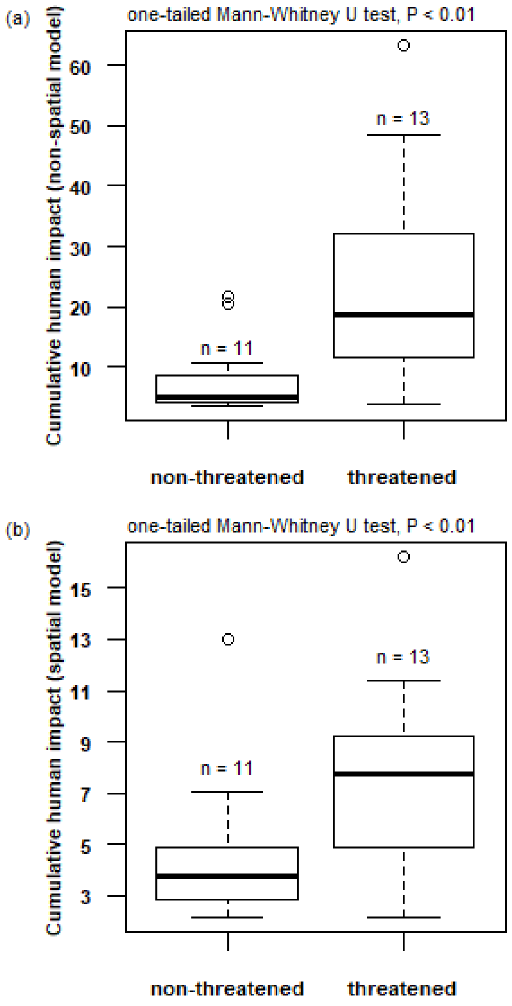


Figure 4. 1. Comparisons of cumulative human impacts based on (a) a non-spatial model and (b) a spatial model for 'non-threatened' species and 'threatened' species.

4.4.2 Predictions on conservation status of ‘Data Deficient’ species

I obtained high-accuracy RF models based on datasets from both CHI models (non-spatial and spatial), with some differences on predictor’s importance. Generally, five stressors were important to the RF models based on both datasets (Fig. 4.2): 1) demersal, non-destructive, high-bycatch fishing (DNHF), 2) nutrient pollution (NuP), 3) ocean pollution (OP), 4) demersal destructive fishing (DDF), and 5) habitat change caused by fishing (HCF). More specifically, nine stressors were important to determining conservation status, when impact indices from the non-spatial CHI model were used (Fig. 4.2a). The minimal-optimal predictors only contained three most important variables: DNHF, NuP, and CHI_n. They derived the best RF model that classified species on the IUCN Red List with 95.7% accuracy (Cohen’s kappa = 0.911, $P < 0.0001$). Only seven stressors were important when impact indices from the spatial CHI model were applied (Fig. 4.2b). Three most important variables composed the minimal-optimal predictors: OP, DNHF, and NuP, in the decreasing order of importance. They derived the best RF model which classified species on the IUCN Red List with 87% accuracy (Cohen’s kappa = 0.738, $P < 0.001$).

Predictions on Data Deficient species were highly consistent between the two optimal RF models (based on indices from non-spatial and spatial CHI models). The optimal RF model, based on the indices from my non-spatial CHI model, predicted that four species were ‘threatened’ and 13 ‘non-threatened’ (Table 4.3). Given data limitations, the optimal RF model, based on the indices from my spatial CHI model, only predicted status for nine species. It predicted that four species were ‘threatened’ and five ‘non-threatened’ (Table 4.3), which were consistent with the those based on indices from the non-spatial CHI model, except for one species (*H. camelopardalis*). This species was predicted to be ‘threatened’ by the RF model based on the spatial CHI dataset, but ‘non-threatened’ by the non-spatial CHI dataset. It also had apparently very low values of CHI_s (0.78) and CHI_n (2.35). Therefore, I determined that *H. camelopardalis* were likely not threatened. Another species (*H. debelius*) was predicted to be ‘non-threatened’ by the RF model based on the non-spatial CHI dataset, but it had relatively higher values of CHI_n (17.9) than other ‘non-threatened’ species and was only reported from a few locations (see Results from Chapter 3). I therefore decided its status was ‘threatened’.

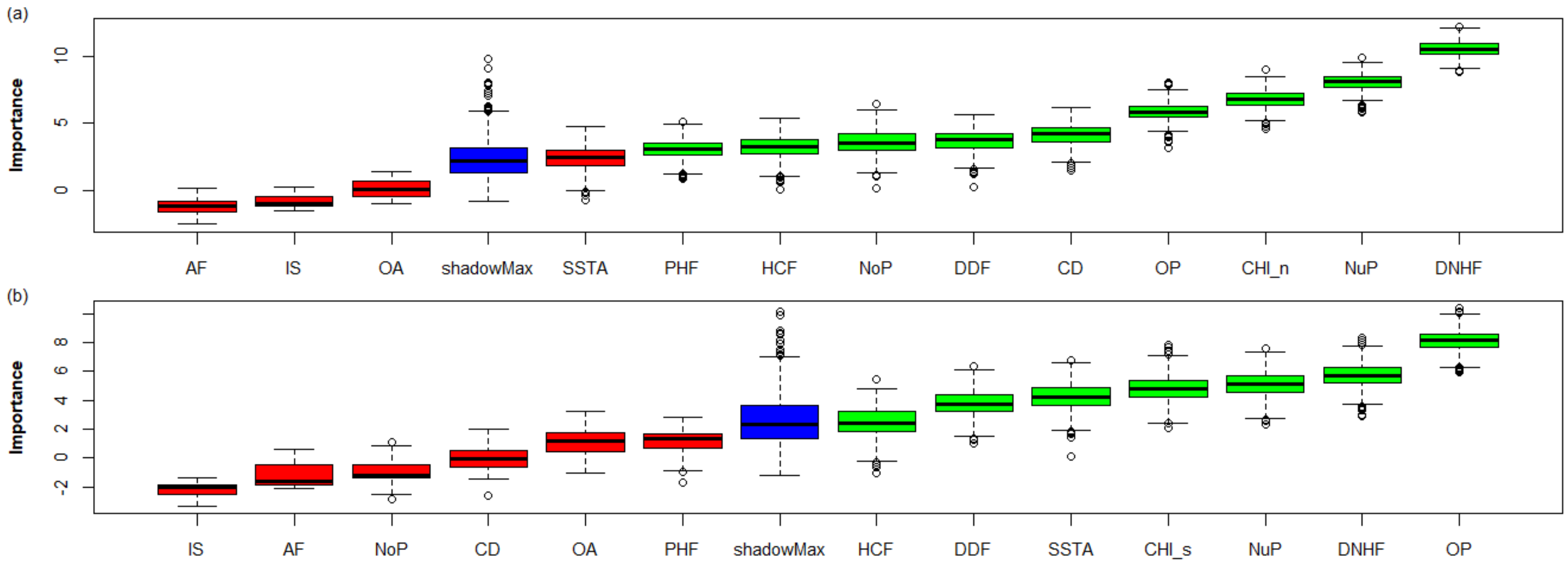


Figure 4. 2. Predictors' relative importance (measured by the normalized mean decrease in model accuracy after permutating data of the predictor) for stressors from (a) the non-spatial CHI model, and (b) the spatial CHI model. AF, artisanal fishing; CD, coastal development; CHI_n, cumulative human impact derived from the non-spatial model; CHI_s, cumulative human impact derived from the spatial model; DDF, demersal destructive fishing; DNHF, demersal non-destructive high-bycatch fishing; HCF, habitat change caused by fishing; IS, invasive species; NoP, noise pollution; NuP, nutrient pollution; OA, ocean acidification; OP, ocean pollution; PHF, pelagic high-bycatch fishing; SSTA, sea-surface-temperature abnormality; shadowMax, the maximum importance of 'shadow' variables in each permutation. The shadow variables' values are derived by shuffling values of the original attribute across objects. Each time, the minimum, mean, and maximum importance value of all shadow variables were calculated. This shuffling was performed multiple times (here, <800) to obtain statistically valid results. The maximum importance of the shadow variables was used as a reference for detecting important attributes. Attributes with green boxes were important, while attributes with red boxes were not important.

Table 4. 3. Summary of threatened status assessments and predictions, cumulative human impact, and expert evaluation confidence for 17 'Data Deficient' seahorse species according to IUCN Red List. CHI_s, cumulative human impact based on our spatial evaluation model; CHI_n, cumulative human impact based on our non-spatial evaluation model; Confidence, average confidence level for the estimates of CHI_s and CHI_n (4, very confident; 3, confident; 2, likely; 1, not sure); ORF_s model, optimal random forest model for predictor datasets derived from our spatial evaluation model; ORF_n model, optimal random forest model for predictor datasets derived from our non-spatial evaluation model; T, threatened; NoT, non-threatened.

Species	CHI_s	CHI_n	Confidence	ORF_s model prediction	ORF_n model prediction
H. camelopardalis	0.78	2.35	3	T*	NoT
H. casscsio	-	46.612	3	-	T
H. coronatus	10.96	23.18	2	T	T
H. guttulatus	10.57	35.75	2	T	T
H. hippocampus	14.56	42.8	2	T	T
H. bargibanti	2.49	2.95	2	NoT	NoT
H. colemani	-	2.09	2	-	NoT
H. debelius	-	17.9	2	-	NoT**
H. denise	2.59	7.2	2	NoT	NoT
H. jugumus	-	1.54	2	-	NoT
H. minotaur	3.35	6.66	2	NoT	NoT
H. paradoxus	-	6.37	2	-	NoT
H. pusillus	-	1.22	2	-	NoT
H. satomiae	-	6.51	2	-	NoT
H. subelongatus	5.36	5.16	2	NoT	NoT
H. tyro	-	3.29	2	-	NoT
H. zebra	3.35	7.6	2	NoT	NoT

* Corrected to 'non-threatened' given the low CHI_s and CHI_n values; ** Corrected to 'threatened' given the high CHI_n value.

4.4.3 Threat epicenters for seahorse species and major stressors therein

The cumulative human impacts on the 'threatened' seahorse species (including my inferred ones) demonstrated prominent spatial heterogeneity (Fig. 4.3). Generally, cumulative human impacts were low in the Americas, Africa, and Australia (Fig. 4.3 a). Some small and patchy areas of ‘High’ impacts occurred in Europe, especially the Mediterranean (Fig. 4.3b). The threat epicenters (‘High’ and ‘Very High’ impacts) mainly concentrated in Asia, with the largest areas in the East China Sea and northern South China Sea (Fig. 4.3). Other threat epicenters were found in the southeastern Korea Peninsula, southern Japan, southern Viet Nam and Cambodia, Peninsula Thai-Malaysia-Myanmar, northeastern Borneo, central Philippines, India, and Sri Lanka (Fig. 4.3c – e). These threat epicenters accounted for 8.3% of the total habitat area of all ‘threatened’ seahorse species, with a smaller fraction (1.1%) but relatively large area (63,261 km^2) experiencing ‘Very High’ impacts ($CHI\_s > 42.1$). In these threat epicenters, the identified major stressors were high and widespread, except nutrient pollution which was largely confined to inshore waters (Fig. 4.4 and Fig. I.2 in Appendix I). The contributions of stressors to the CHI varied among threatened species, with demersal non-selective fishing and/or pollution-related activities representing the major stressors (Fig. I.3).

My sensitivity analyses indicated that the distribution patterns of CHI were very similar between the two groups (19 threatened species vs. all 33 species; Fig. 4.5). I found that the standardized CHI map for all 33 seahorse species only had a medium positive correlation with the counterpart for marine ecosystems ($r = 0.50$, Pearson’s correlation coefficient; Fig. I.4). The latter also overestimated CHI on seahorses in almost all (99.96%) habitat pixels, with a tendency to have higher estimates than the former in offshore deeper waters (Fig. I.5). The CHI maps based on two different algorithms (additive vs. average for overlapping cells among species) were highly correlated ($r = 0.74$ and 0.75; Pearson correlation coefficients) but showed some divergences in spatial distribution patterns (Fig. 4.5). Compared with the CHI map by additive algorithm (Fig. 4.5c and 4.5d), the CHI map by average algorithm showed that habitats with high CHI values were more widely spread and included some patches in North and South America (Fig. 4.5e and 4.5f). I found medium positive correlations between species richness and CHI by additive algorithm ($r = 0.46$, 0.64) but no significant correlations between species richness and CHI by average algorithm ($r = -0.09$, 0.08).

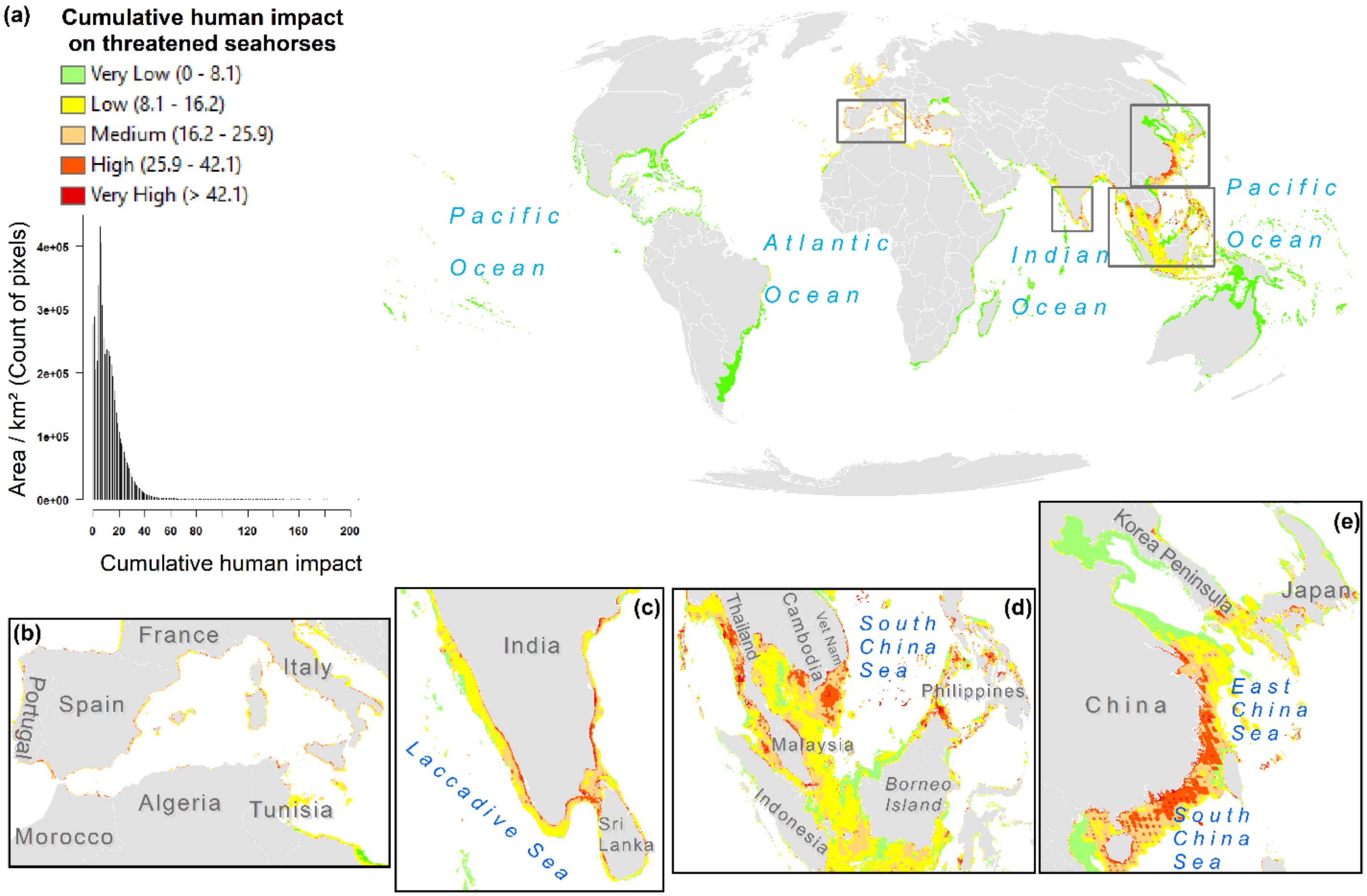


Figure 4. 3. Global map of cumulative human impact on seahorse species known or inferred to be threatened. The classification of cumulative human impact (from Very Low to Very High) is based on the Jenks Natural Breaks algorithm, which maximizes the difference between classes and similarity within each class. (Note that the spatial map might be inaccurate in some regions, such as the Gulf of Thailand, because of data limitation).

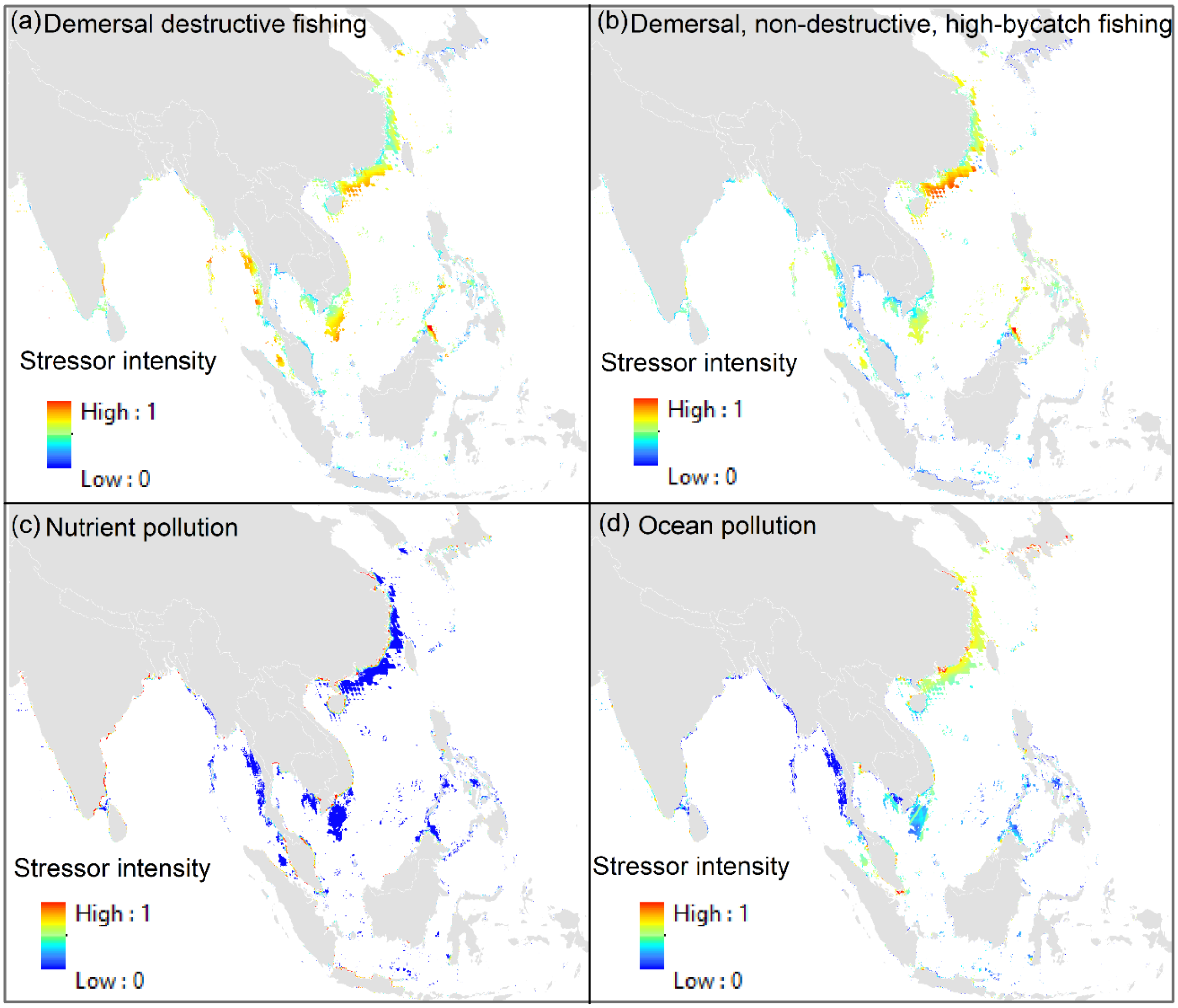


Figure 4. 4. Intensity map of four important anthropogenic stressors in threat epicenters of threatened seahorses: (a) demersal destructive fishing (DDF), (b) demersal, non-destructive, high-bycatch fishing (DNHF), (c) nutrient pollution (NuP), and (d) ocean pollution (OP). Note that another important stressor, habitat change caused by fishing was not displayed as its intensity was estimated by demersal destructive fishing (i.e., same as (a)).

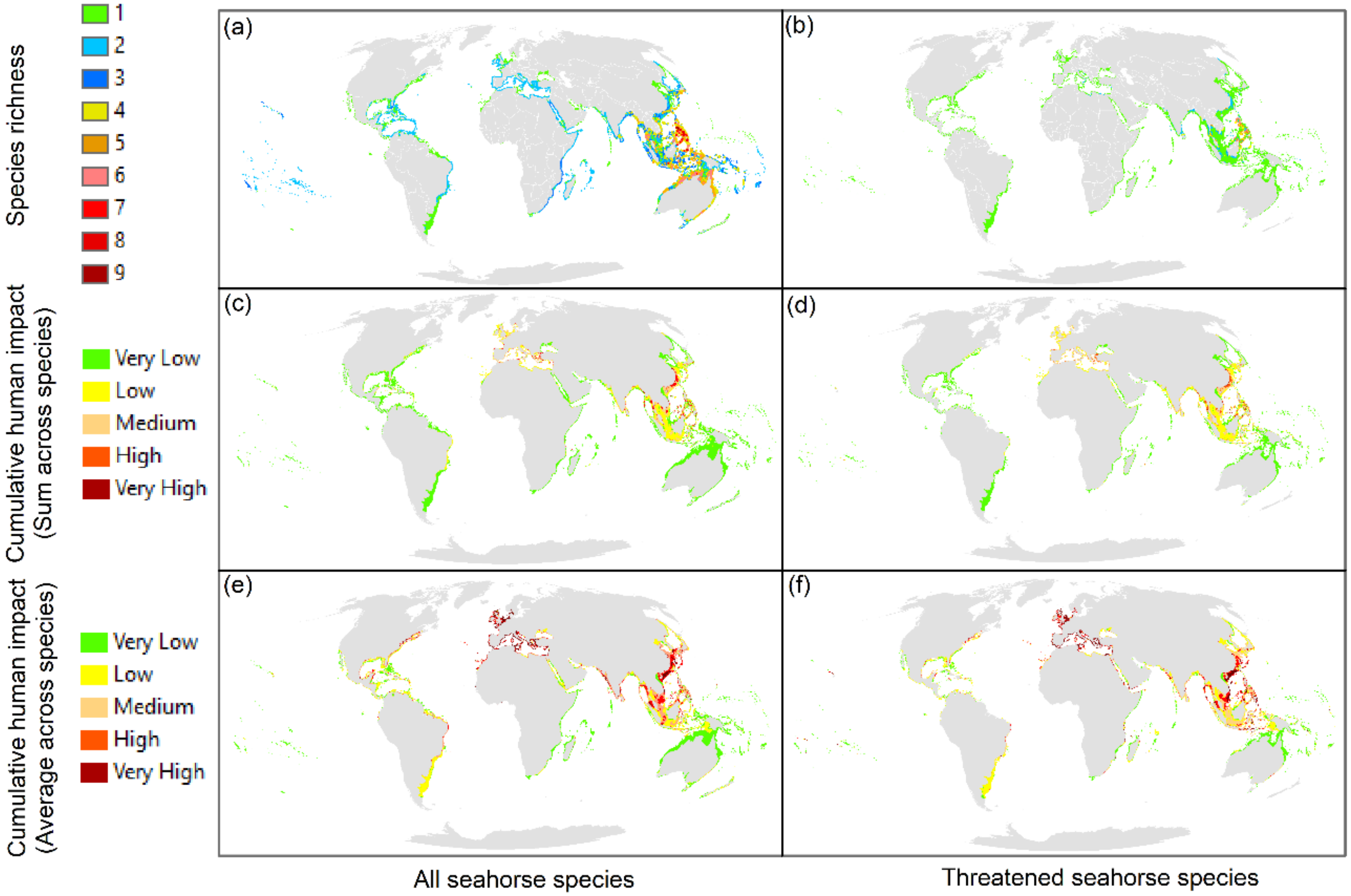


Figure 4. 5. Spatial distributions of species richness and two versions of cumulative human impact (sum across species vs. average across species) for all seahorse species (n = 33) and threatened species only (n = 19). The classification of cumulative human impact (from Very Low to Very High) is based on the Jenks Natural Breaks algorithm, which maximizes the difference between classes and similarity within each class.

4.5 Discussion

I find that threat patterns of data-poor marine species can be identified by integrating expert knowledge and spatial datasets with modeling techniques. My two CHI models demonstrate good consistency with IUCN Red List assessments in estimating species' status. I also show that CHI models can derive useful human-impact indices to predict threatened status for Data Deficient species. This finding is encouraging as it might provide an alternative approach to assessing the many Data Deficient marine taxa (especially fishes) on the IUCN Red List (McCauley et al. 2015, IUCN 2018). My study also provides a global spatially-explicit map of CHI on marine species and identifies major stressors. These results are useful for coastal management and conservation (Halpern et al. 2008). Although my study was initiated for seahorses, the general framework and techniques have the potential to be used for many other marine taxa.

4.5.1 Threat patterns and their conservation implications for seahorses

My study contributes new and quantitative knowledge about human impacts for the whole genus of 42 seahorse species. I am the first to evaluate cumulative human impact for seahorse species. I predicted five 'threatened' and 14 'non-threatened' species among the Data Deficient species on the IUCN Red List (IUCN 2018). This finding, along with current IUCN assessments, suggests that a total of 19 (45%) seahorse species are globally threatened by anthropogenic pressures. All seahorse species have been listed on the Appendix II of the Convention on International Trade in Endangered Species (CITES) since 2002, with implementation in 2004 (Vincent et al. 2014). This listing requires CITES Parties to ensure their exports of seahorses do not damage wild populations. I found that all 19 'threatened' species are traded internationally (Foster et al. 2014) and need focused attention by CITES. I identify the major threat epicenters as concentrated in China and Southeast Asian countries, where marine ecosystems are also highly impacted (Halpern et al. 2008, Halpern et al. 2015). Among them, China is the largest consumer and importer of seahorses, and Southeast Asian countries are the major exporters (Vincent et al. 2011, Foster et al. 2016).

My findings highlight the importance of mitigating demersal, non-destructive, high-bycatch fishing (DNHF, e.g., purse seines), ocean pollution (OP), and nutrient pollution (NuP) for seahorse conservation. Previous studies often considered demersal destructive fishing (DDF) such as bottom trawling as the major threat on seahorses, as it is the major source of traded seahorses and damages seahorse habitats (Vincent et al. 2011). Only recently have studies indicated that DNHF (e.g. purse seines) may catch similarly large numbers of seahorses as DDF (Lawson et al. 2017). As well, ocean and nutrient pollutions (OP and NuP) have usually been considered a secondary concern for seahorses, although they can cause eutrophication, hypoxia, and degradation of seahorse habitats (e.g., seagrass beds; Vincent et al. 2011). My models suggest that DNHF, OP, and NuP may be more important than DDF in determining seahorse threatened status, perhaps because the impacts of the first three stressors (DNHF, OP, NuP) had greater divergences between ‘threatened’ species and ‘non-threatened’ species than counterparts of DDF. My results thus suggest that we might have previously underestimated the influence of DNHF, OP, and NuP in determining extinction risks of seahorses.

4.5.2 Cumulative human impacts on marine organisms and ecosystems

I highlight the importance of developing species-level CHI models to estimate cumulative impacts for particular marine taxa (Maxwell et al. 2013). My models are meaningful as they derived significantly higher CHI indices (CHI_n and CHI_s) for threatened species than non-threatened species. In contrast, such difference didn’t exist for the CHI index (CHI_e) based on marine ecosystems’ vulnerability data, suggesting that the vulnerability measures for marine ecosystems differ from counterparts for specific organisms (Maxwell et al. 2013). The spatial difference between my CHI map for seahorses and the CHI map for marine ecosystems might result from seahorses’ greater distributions in shallower waters (see Results from Chapter 3). Therefore, higher CHI values for seahorses were concentrated in shallower waters, compared to the CHI for marine ecosystems.

Building species-level CHI models is helpful for generating conservation plans for different focal organisms. My results suggest that some leading threats (e.g., climate change) on marine offshore ecosystems may not be the leading stressors to inshore fish species (Halpern et al. 2008, Halpern et al. 2015). Given that importance of each stressor varies among species, there is clearly a need to tailor threat-mitigation plans to particular organisms (Maxwell et al. 2013). Given that some major pressures on seahorses (e.g., non-selective fishing and pollutions) are also leading stressors on marine mammals, sea birds and turtles (Maxwell et al. 2013, Avila et al. 2018), mitigating impacts of these stressors may benefit a large variety of marine animals (Maxwell et al. 2013). Identifying such stressors could be beneficial to achieve broad conservation goals.

My research suggests that the actual cumulative human impact is more complex than a linear-additive model can explain. As with many other CHI studies, my CHI models are based on the ‘linear additive’ assumptions (Halpern and Fujita 2013), which might be not valid in marine ecosystems (Crain et al. 2008, Coll et al. 2012). Therefore, researchers are highly recommended to validate the modeled CHI with observed status of the ecosystem components (Korpinen and Andresen, 2016). The few studies have done so report both positive and negative findings (Andersen et al. 2015, Clark et al. 2016, Korpinen and Andersen 2016). One other study showed the relationship between the CHI and degradation conditions of Mediterranean coralligenous outcrops was best fitted by a log-log model rather than linear model (Bevilacqua et al. 2018). I correlated my CHI estimates with IUCN Red List assessments and found a generally good consistency between them. However, threatened and non-threatened species cannot be clearly split by CHI indices alone (CHI_s or CHI_n). My random forest models demonstrated that impacts of several individual stressors were more important than the CHI (of all stressors) in predicting threatened status.

I highlight that distribution patterns of CHI can be influenced by using different algorithms (sum vs. average) to derive CHIs in overlapping cells of multiple species. Halpern et al. (2008) showed that the average algorithm derived a CHI map very similar to that by the additive algorithm. I displayed that although these two algorithms also derived similar results for seahorses, the differences were also notable and further influenced the relationship between CHI and species richness. Previous studies have shown significant positive correlations between CHI (based on the additive algorithm) and species richness. For instance, a substantial overlap between regions with high species richness and high cumulative threats was found in the Mediterranean Sea (Coll et al. 2012). A global-scale study also indicated that species richness is positively correlated to cumulative human impacts on marine ecosystems (Lindegren et al. 2018). I also showed such correlation exist in seahorses. However, this correlation became insignificant when the average algorithm was used for CHI in my case. This finding suggests that using additive algorithm could fake the relationship between species richness and human impacts. Another potential risk of using additive algorithm is that species-poor regions could be overshadowed by species-rich ones as shown in my results. However, if one aims to estimate cumulative impact on all species in the same location, using the additive algorithm will be more reasonable than using the average algorithm.

4.5.3 Challenges and limitations

My study was subject to several limitations such that some results should be interpreted with caution. First, the number of expert evaluations per species and the confidence levels of most evaluations were not high, especially for impacts from climate change (see Appendix H). These two concerns reflect the paucity of experts and studies on impacts of anthropogenic stressors upon seahorses (Vincent et al. 2011, Cohen et al. 2017). Further species-stressor studies might change current threat patterns. Second, I may not involve or estimate well some important stressors, as their spatial intensities are difficult to model. This may particularly explain the unexpected low CHI values of the two Endangered seahorse species, *H. capensis* and *H. whitei*. *Hippocampus capensis* lives in a narrow range (several estuaries in South Africa) and is thought to be negatively affected by a pressure not included in my stressor data set: storm-water runoff with heavy loads of sediments (Pollom 2017). *Hippocampus whitei* inhabits estuaries and areas with habitat-forming biota (e.g. seagrasses) in Australia. It is mainly threatened by habitat change caused by coastal development (Harasti and Pollom 2017), which was mainly inferred from population density in my study. Further analyses that better map intensities of these and other stressors are needed.

4.5.4 Applying CHI models to guide marine conservation

I am the first to illustrate that human-impact indices derived from CHI models can be used to estimate conservation status for data-poor species. I found that the ecosystem-scale CHI index did little to predict conservation status of seahorses, just as was previously found for marine mammals (Davidson et al. 2012). In contrast, I found that human-impact indices derived from species-level CHI models are highly useful in conservation-status prediction. This supports my hypothesis that species-level models can derive human-impact indices more relevant to conservation status than does an ecosystem-scale model.

I also indicate that a non-spatial CHI model can derive impact indices to better predict threatened status than a spatial model. This is encouraging for many Data Deficient marine species that lack spatial maps for applying a spatial CHI model (McCauley et al. 2015). In general, the non-spatial CHI model relies more on expert's knowledge. The better predictions based on indices from my non-spatial model might be explained by the fact that eight out of my 22 experts have also participated in IUCN Red List assessments, with five as Assessors, two as Reviewers, and one as the Contributor for a total of 14 seahorse species. Interestingly, in a post-hoc survey, almost all experts reported that they directly used their knowledge or data rather than IUCN Red List assessments in evaluating my measures. Given my CHI system is different from the IUCN Red List system, the high consistency between the two reflects the utility of my approach to supplement IUCN system in evaluating Data Deficient species.

As shown in my study, a spatial CHI model can also derive results generally consistent with IUCN Red List assessments. The merit of a spatial CHI model is to provide spatially-explicit maps of human impacts on focal species (Maxwell et al. 2013). Currently, we need such maps to better understand the patterns of human impacts on many marine organisms (Maxwell et al. 2013). Understanding these threat patterns is vital to develop cost-effective management plans for ameliorating their impacts (Lewison et al. 2004, Crain et al. 2009). The impact map can be further used to deduce conservation costs in marine conservation planning and prioritization (Ban and Klein 2009, da Luz Fernandes et al. 2017).

Chapter 5: Conservation prioritization at large spatial scales for seahorses (*Hippocampus* spp.)

5.1 Summary

To ensure that marine protected areas (MPAs) are supported by multiple stakeholders, planners must integrate scientific evidence and socioeconomic costs into conservation prioritization. However, such practices are challenging at large spatial scales. To explore possibilities, I initiated a framework of large-scale conservation prioritization for marine species. My framework combines a gap analysis of the conservation targets against current MPAs and a further prioritization with consideration of different socioeconomic costs. I illustrated this with case studies of seahorses (*Hippocampus* spp.) at the global scale and at a national scale in China. At both scales, I built a conservation scenario and four alternative scenarios to examine the effect of socioeconomic concerns on priority selection with Marxan software. I set a minimum conservation target to protect 2,000 km^2 area of occupancy (AOO, *sensu* IUCN) for each species, representing a notional bottom line for protecting marine species with limited conservation budgets. I examined two approaches based on Marxan outputs (best solution vs. selection frequency) to deriving priority integrated solutions. My results at both scales indicated that current MPA coverage meets this notional bottom-line target for most seahorse species, but more habitats were covered by lower-protection MPAs than greater-protection MPAs (i.e., no-take reserves). Nine species globally and one species in China did not even meet the very minimal target for MPA coverage. Global priorities for enhanced protection include habitats in Hawaii, South Africa, Australia, Southeast Asia, and Japan. New priority areas for China were mainly located in southern Hainan and Taiwan provinces. I found that the selection-frequency approach performed better than the best-solution approach in two of the three prioritization cases. My study highlights the complexity of selecting conservation priorities under different socioeconomic constraints at large spatial scales and provides a template for such an approach.

5.2 Introduction

Marine protected areas (MPA) have served as a major cornerstone in marine conservation, but they are often challenged by socioeconomic interests (Klein et al. 2013, Costello and Ballantine 2015). Globally, MPAs are widely used to safeguard marine biodiversity and threatened ecosystems (Klein et al. 2013). Well-managed MPAs not only protect the conservation targets but also deliver socioeconomic benefits and ecosystem services (Murray et al. 1999, Cinner et al. 2005, McCook et al. 2010). Many countries have set up MPAs to meet their obligations under the Convention on Biological Diversity (CBD) and the Aichi Target 11 (protecting 10% of the ocean by 2020) (Costello and Ballantine 2015). However, the implementation of MPAs has often encountered challenges because of noncompliance of stakeholders, especially fishers (Gill et al. 2017). One reason is the lack of clarity in fisheries benefits from MPAs. Another reason is the failure to embrace socioeconomic factors in setting up MPAs (Green et al. 2009, Weeks et al. 2010, Cisneros-Montemayor and Vincent 2016).

Conservation prioritization/planning is a vital approach in identifying potential MPAs that can meet a given conservation target with minimum socioeconomic costs (Margules and Pressey 2000, Smith et al. 2006). Usually the conservation target is to protect a certain amount of habitats that could secure the focal biodiversity (Brooks et al. 2006, Drummond et al. 2010). The socioeconomic costs mainly include three components (Naidoo et al. 2006, Ban and Klein 2009): (1) management costs from enforcing and maintaining the priority habitats (Balmford et al. 2004); (2) transaction costs resulting from negotiating protection (Naidoo et al. 2006); and (3) opportunity costs from foregone revenues (e.g., fisheries value) (Cameron et al. 2008). Additionally, human impacts have also been used as a type of cost in prioritization studies (Tallis et al. 2008, Ban and Klein 2009, Klein et al. 2013), based on the assumption that protecting highly-impacted habitats/populations leads to higher costs (Klein et al. 2013). Integrating these socioeconomic costs into prioritization is crucial for understanding stakeholders' interests and setting up feasible MPAs (Scholz et al. 2004, Carwardine et al. 2008, Klein et al. 2008).

Marine conservation prioritization (MCP) at large spatial scales (nationally or internationally) is challenging, although increasingly popular. First, few spatially-explicit data for socio-economic costs exist for the oceans at large scales (Balmford et al. 2004, Ban and Klein 2009, Freudenberger et al. 2013). Second, only a few prioritization tools (e.g., Marxan) can incorporate socio-economic costs especially for large-scale datasets (Stelzenmüller et al. 2013, Pınarbaşı et al. 2017). Third, few prioritization tools provide exact priority solutions, as they were normally designed to produce multiple choices for local stakeholders' negotiations (Ban and Klein 2009, Pınarbaşı et al. 2017). At large spatial scales, such negotiation can be much more formidable than at local scales. Techniques to address this gap are understudied in the literature (Solovyev et al. 2017). On the other hand, there is an emergent need for large-scale MCP for the development of large MPA networks (Douvere 2008, Mazor et al. 2014). Currently, MCP studies at the multinational and global scales are rare and dominated by the 'biodiversity hotspot' schema (Myers et al. 2000). Most studies have only used area and/or human impacts as the surrogates for conservation costs (Selig et al. 2014, Asaad et al. 2018, Lindegren et al. 2018).

Large-scale MCP studies can derive large networks of MPAs that may better protect marine species and ecosystems. Conservation initiatives should be based on a clear understanding of the scale at which they are working (Lourie and Vincent 2004). Marine species usually have large geographic ranges (e.g., over 100,000 km^2) and may migrate to different habitats throughout their life history (Hooker et al. 2002). As well, numerous anthropogenic activities (e.g., fishing) often occur at large spatial scales. Such ecological and social factors require large-scale conservation planning to identify and protect all priority habitats across the species ranges (Agardy et al. 2011). Marine species covered by large networks of MPAs are likely more resilient to large-scale disturbances. Such resilience is vital for organisms to survive and adapt in a changing ocean environment (Hazen et al. 2013). Ignoring the importance of large-scale MCP could cause conservation failures (Agardy et al. 2011).

Seahorses (*Hippocampus* spp.) are a genus of flagship fishes whose conservation prioritization may benefit many other marine species. Given their charismatic appearance and threatened status, seahorses have been used as ‘flagships’ for promoting marine conservation around the word (Vincent et al. 2011, Harasti et al. 2014). Nearly 45% of seahorse species are known/predicted as being threatened (IUCN 2018; see Chapter 3), calling for urgent conservation measures such as new marine reserves to protect them. Distributing around the world’s shallow seas from the temperate to the tropic, seahorses are found in many of the species-rich habitats such as seagrass beds, estuaries, mangroves, and coral reefs (Foster and Vincent 2004). Such geographic and habitat traits imply that protecting seahorse habitats may benefit many other species therein. Seahorses are traded around the world for traditional medicines, curios, and aquarium fishes (Vincent et al. 2011, Foster et al. 2014). As the first fully marine genus being listed on Appendix II of the Convention on International Trade in Endangered Species (CITES) (effective since 2004; Vincent et al. 2011), seahorse conservation planning may help CITES parties better fulfil their obligations and set an example for other listed species such as sharks and rays.

My study seeks approaches to marine conservation planning at large spatial scales, using seahorses as the case study. I here mainly examine three questions: 1) How do we incorporate socioeconomic costs in marine conservation planning at large spatial scales? 2) How do we derive priority solutions from prioritization tools? 3) Where are the priority areas for seahorses? To address the above questions, I developed a new prioritization framework. I collated habitat and socioeconomic-cost maps for seahorse species, and global and regional MPA maps. I used this framework to set priorities for seahorses at the global and national scales. To do so, I first examined current MPA coverage of seahorse habitats, and then examined two novel approaches to selecting priority habitats. This work is directed at developing a framework with useful techniques to address current challenges in large-scale MCP for marine species.

5.3 Methods

5.3.1 Prioritization framework for conservation and management

I developed a framework to set conservation and management priorities based on previous schemas of systematic conservation planning (Fig. 5.1) (Margules and Pressey 2000, McIntosh et al. 2017). My framework considered the focal species as the conservation features. I set the conservation target to protect at least 2,000 km^2 of ‘area of occupancy (AOO)’ for each species. Here, the AOO is defined by the guidelines for using the IUCN Red List criteria (IUCN 2005). I chose 2,000 km^2 because this is the threshold of AOO (Criterion B2) below which the IUCN considers species to be threatened, starting with Vulnerable. This threshold is not conservative in identifying threatened marine species and sets aside at least three possible concerns. First, conservation status of most marine species (including seahorses) on the IUCN Red List has been evaluated against population decline rate (Criterion A) rather than the AOO (Criterion B2), Second, Red Listing under Criterion B2 requires more than just AOO in deciding on conservation status based on geographic criteria (e.g., evidence about population fragmentation or highly fluctuated; IUCN 2005). Third, one reason that Criterion B2 is seldom used for marine species is that 2,000 km^2 is very low as a target for protecting marine species, where AOO of even very threatened species could easily exceed this threshold. However, these caveats notwithstanding, I used this ‘bottom-line’ target of 2,000 km^2 in the absence of better advice. I first compared it against current MPAs and related conservation plans. I then identified supplementary priorities for conservation and management based on multiple socioeconomic scenarios.

Ideally, I would like to protect species AOO with the least cumulative human impact (CHI) and fewest management costs (hereafter, CHI-cost scenario), since such areas would likely have the highest probability of saving the species with least effort. But using least CHI does not minimize costs to specific stakeholders for CHI is a composite of individual impacts (Ban and Klein, 2009). Therefore, alternative scenarios were applied to identify priorities that could minimize the opportunity costs (e.g., fisheries) for different stakeholders in using species habitats. I compared the CHI-cost scenario with each alternative scenario to explore spatial convergences and divergences. I considered the convergences as potential priorities for no-take reserves, and the divergences as priorities for sustainable use multizone MPAs. To illustrate the approach, I applied the framework to identify priorities for seahorses in two spatial scales: nationally in China and globally.

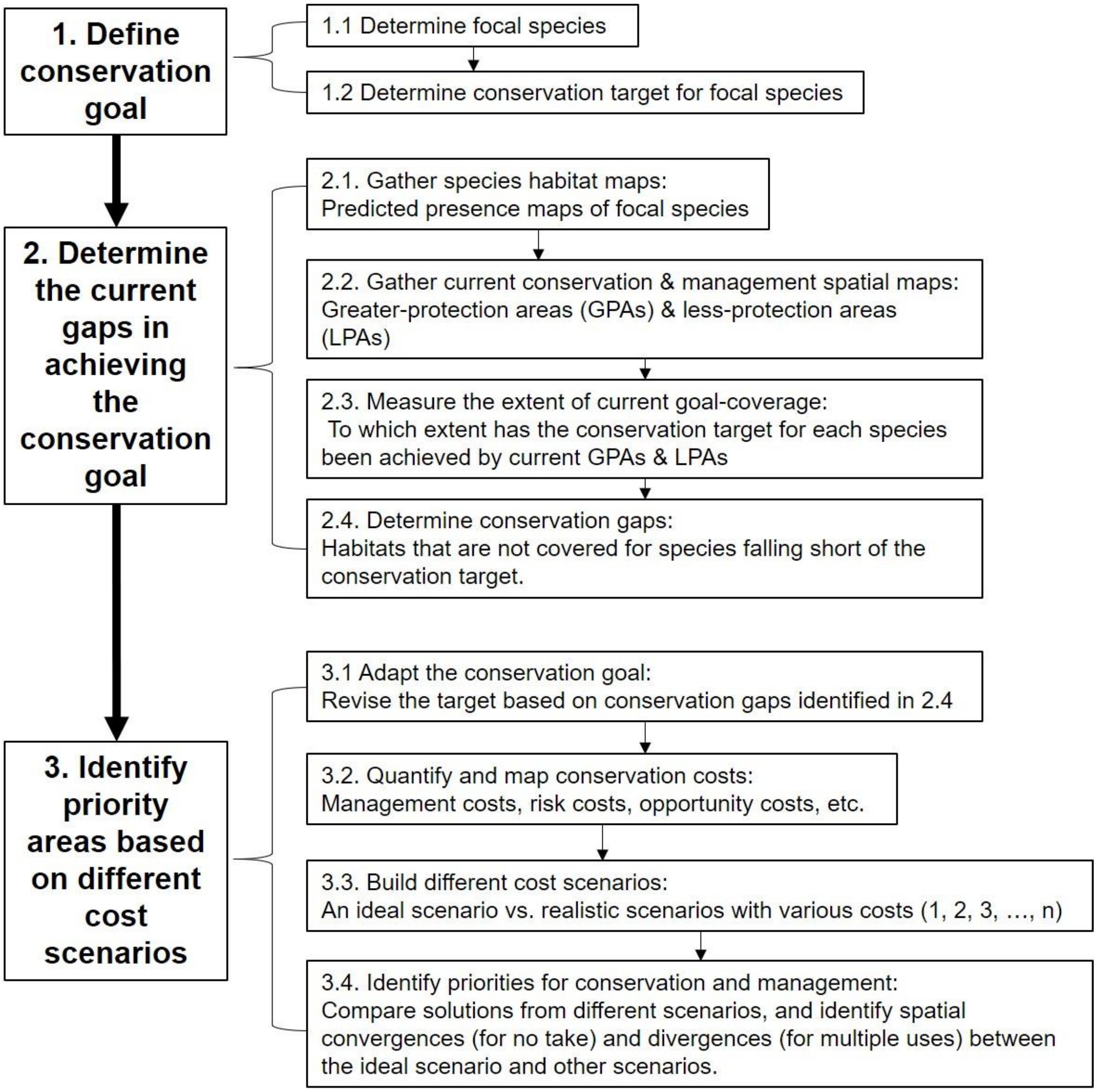


Figure 5.1. The framework of conservation prioritization for marine species at large spatial scales.

5.3.2 The national-scale case study on seahorse populations in China

I aimed to protect at least 2,000 km^2 AOO for each of the six Chinese seahorse species whose AOO maps were available. Approximately 14 seahorse species inhabit China's waters (including Hong Kong and Taiwan) based on my study in Chapter 2 and Chinese colleagues' records (see details in Table J.1 in Appendix J). Currently, only six of them have habitat maps in China's waters from my studies in Chapter 2 and 3, so could be used for this study. Among them, the habitat map of the Spiny seahorse (*Hippocampus histrix*) came from my global-scale study (Chapter 3) and the others from my national-scale study (Chapter 2). All these six species are threatened in China (Wang and Xie 2009), as well as globally (IUCN 2018).

I gathered China's MPA and marine-spatial-planning data from local sources and categorized them as greater-protection areas (GPAs) and lower-protection areas (LPAs). China's coastal provinces/municipalities have initiated 'marine functional zoning' and 'ecological-redline' planning in recent years (Lu et al. 2015). These plans have defined marine reserves as no-take zones (here, as GPAs) and other important zones for sustainable multiple uses (here, as LPAs). The latter allow certain types of human uses (including fisheries) but prohibit constructions (e.g., sea filling/reclamation) and pollutions (Lu et al. 2015). I gathered these maps for coastal regions from local governments' websites and Chinese colleagues (see Table J.2 in Appendix J). I then estimated the proportion of each seahorse's habitats covered by GPAs and LPAs separately, as well as all protected areas together (i.e., the union of GPAs and LPAs).

For those species whose MPA coverage was less than my target, I conducted conservation planning in five scenarios (Table 5.1). The planning units were the AOO cells (1 x 1 km^2 pixels) of each species that were not protected by existing MPAs. The first scenario was composed of the management cost and the cumulative human impact (i.e., CHI-cost scenario; Table 1). Four alternative scenarios replaced the CHI cost with one of the socioeconomic surrogates (Table 1), respectively: (1) commercial fisheries, (2) artisanal fisheries, (3) shipping, and (4) (land-based) nutrient pollution. I used the total boundary length of the selected planning units as the surrogate of management cost (Possingham et al. 2000, Balmford et al. 2004, Ban and Klein 2009). This management cost was derived with the ArcMarxan Toolbox based on the default setting in an ArcMap. The other five types of cost were derived from Halpern et al. (2015) and my Chapter 3 (see detailed descriptions in Appendix J). Among them, commercial fisheries were model estimates of catch from demersal non-selective fisheries (e.g., bottom trawling and purse seine). I only included these types of fisheries as such active gears exerted the major pressures on seahorses which live close to the seafloor (see Results in Chapter 3).

Table 5. 1. Five cost scenarios for conservation and management planning for seahorses.

Cost surrogates	**Conservation cost scenarios**				
	CHI-cost scenario	AF scenario	CF scenario	SP scenario	NP scenario
Total boundary length	√	√	√	√	√
Cumulative human impact (CHI)	√				
Artisanal fisheries (AF)		√			
Commercial fisheries (CF)			√		
Shipping (SP)				√	
Nutrient pollution (NP)					√

I used Marxan software (version 2.3.4; Possingham et al. 2000) to configure priorities that minimize conservation costs in the above five scenarios. Marxan is one the most widely used prioritization tools that allows researchers to incorporate socioeconomic constraints in planning (Possingham et al. 2000). Compared with other tools, Marxan also has capacity to handle large datasets (Leslie et al. 2003). The simulated annealing algorithm and iterative improvement features were chosen in Marxan. In each scenario, I ran Marxan with 500 replicates at 1 million or 10 million iterations for each repeat after initial trials of calibration (different times of iterations used in different species cases) (see guideline by Game and Grantham, 2008). This resulted in 500 solutions that satisfied my goal and demonstrated how frequently the planning units contributed to my goal. Marxan uses two weighting factors to balance the conservation goal and total boundary length (here, management cost): species penalty factor (SPF), and boundary length modifier (BLM). By following previous guides, we determined the values for these factors to ensure my conservation targets were always met and priorities were compact and cost-effective (Stewart and Possingham 2005, Game and Grantham 2008, Ardron et al. 2010). See a detailed description in Appendix J.

I derived priority solutions based on two approaches and compared solutions of the CHI-cost scenario against the other four scenarios. To do so, I plotted two Marxan outputs, the best selection and solution frequency in each scenario. The best solution is an optimal solution with the least cost among the 500 replicates in determining whether a planning unit should be selected as priority site (1, selected; 0, not selected). This best solution does not necessarily ensure the best priority system and it may only marginally better than the other solutions (Game and Grantham 2008). The solution frequency records the number of times (0 – 500) of each planning unit being selected among the 500 replicates. I examined the Spearman’s rank correlations on the selection frequency (0 – 500) between each scenario pair. For each scenario, I determined the priority units based on the solution frequency by using a frequency threshold, which ensured the conservation targets were met with the minimum number of planning units (see detailed methods in Appendix J). By doing so, I converted the solution frequency into a binary map (‘priority’ and ‘non-priority’, 1 and 0), same as the best-solution map. I then compared the two types of priority solutions (solution frequency vs. best solution) using Cohen’s kappa statistic (function kappa2 in R package irr) (Fielding and Bell 1997, Gamer et al. 2007). This measures the agreement between two different solutions, while correcting for random effects (Prasad et al. 2006, Cutler et al. 2007). I also used this statistic to examine the extent of spatial convergence in priority solutions between the CHI-cost scenario and other cost scenarios respectively. In this case, both priority solutions (selection frequency vs. best solution) were tested individually to examine the effect of different solution methods on the extent of spatial convergence.

Finally, I identified priorities based on the above comparisons. I decided to use the solution (selection frequency vs. best solution) that derived more spatial convergences between the CHI-cost scenario and alternative scenarios. Based on such a solution, I identified the planning units that were agreed by all scenarios (i.e., spatial convergences) and units that were not (i.e., spatial divergences). The former represented the units with the lowest values for my full set of human uses, and thus were considered as priorities for greater-protection areas (GPAs): prohibiting demersal non-selective fishing, artisanal fishing, shipping and nutrient pollution. In contrast, the latter represented conflict zones and was considered as priority sites for lower-protection areas (LPAs, e.g., multiple-use zones). The permitted types of human uses could be determined by which socioeconomic factors were influential.

5.3.3 The global-scale study on seahorse species

At the global scale, I aimed to protect at least 2,000 km^2 habitat for each of the 33 species whose AOO maps were available from my study in Chapter 3.

To achieve this goal, I first explored the extent of AOO for each species fell within current MPA coverage (unit: km^2). I gathered the MPA data from the World Database on Protected Areas (UNEP-WCMC and IUCN 2018). I then categorized them into two groups as I did in China: greater-protection areas (GPAs) and lesser-protection areas (LPAs). The GPAs (n = 3313) contained all MPAs with IUCN Categories I – IV and MPAs without IUCN Categories that were claimed to be entirely 'no-take'. Protected areas with IUCN Categories I – IV are commonly considered as strict marine reserves (Soutullo et al. 2008, Jenkins and Van Houtan 2016), although their protection efficacy can be variable. The LPAs (n = 6520) contained all remaining MPAs, including those with IUCN Categories V and VI, which were created for sustainable multiple uses (e.g. recreation and tourism) (Dudley 2008, Kenchington 2010). The GPAs and LPAs maps from China were also added to the global categories respectively. I then calculated 1) the AOO of each species covered by GPAs and LPAs separately, and 2) the difference between the protection coverage (by all MPAs) and my conservation target. I then applied Marxan to identify priority areas for species whose habitats were not sufficiently protected following the same approach described in the case study of Chinese seahorses.

5.4 Results

5.4.1 National-scale MPA coverage in the areas of occupancy of Chinese seahorse populations

I found that the conservation targets of five Chinese seahorse species were met by current MPAs and spatial planning, with only *H. histrix* (protected AOO < 2,000 km^2) remained for further prioritization in Marxan at the national level (Table 5.2). The Japanese seahorse (*H. mohnikei*) and the Three-spot seahorse (*H. trimaculatus*) had the largest values of AOO covered by both greater- and lower-protection areas. In contrast, the ratio of protected AOO to all AOO (per species) was low (4 – 11%) for five Chinese seahorses, but moderate (50%)

for *H. histrix* (Table 5.2). The AOO covered by greater-protection areas (GPAs, 10,380 km^2) only accounted for 16% of the overall protected AOO. The unprotected AOO units of *H. histrix* were located in coastal waters of Hainan and Taiwan Provinces, where I further set priorities.

Table 5. 2. Summary of conservation coverage in the area of occupancy (AOO) of six seahorse species in China, sorted by AOO. The protected area (Area / km^2) and main region of the protection for each species were summarized for two types of MPAs: greater-protection areas and lower-protection areas. All MPAs (km^2): the overall protected areas for each species by all MPAs. Protection ratio: the ratio of protected AOO to all AOO for each species.

Species	Greater-protection areas		Lower-protection areas		All MPAs	Area of occupancy	
	Area (km^2)	Main region	Area (km^2)	Main region	(km^2)	Area (km^2)	Protection ratio
H. histrix	175	Pingtung (Taiwan)	1,159	Wenchang (Hainan)	1,334	2,622	0.51
H. kelloggi	212	Beihai (Guangxi)	4,758	Beibuwan Bay (Guangxi)	4,970	83,396	0.06
H. kuda	1,662	Dongsha Islands	7,499	Hainan	9,161	93,323	0.10
H. spinosissimus	1,035	Dongsha Islands	4,452	Hainan	5,487	128,365	0.04
H. mohnikei	6,802	Lianyungang (Jiangsu)	34,609	Zhoushan Islands (Zhejiang)	41,411	351,795	0.12
H. trimaculatus	5,036	Zhoushan Islands (Zhejiang)	31,115	Hainan	36,151	387,582	0.09

5.4.2 National-scale priorities for Chinese seahorse populations

For *H. histrix*, the CHI-cost scenario (Fig. 5.2a and 5.2f) had high correlations in terms of selection frequency with all alternative scenarios (Table 5.3), with the highest correlation with the artisanal-fisheries scenario (r = 0.95, Fig. 5.2b and 5.2g), followed by the nutrient-pollution (r = 0.81, Fig. 5.2c and 5.2h), shipping (r = 0.77, Fig. 5.2d and 5.2i), and commercial-fisheries scenarios (r = 0.76; Fig. 5.2e and 5.2j).

Table 5. 3. Comparisons on prioritization results between the CHI-cost scenario and other four scenarios: artisanal fisheries, commercial fisheries, nutrient pollution, and shipping. The prioritization was done in three regions: China (for *H. histrix*), Southeast Asia (for *Hippocampus barbouri*), and Australia (for *H. abdominalis* and *H. minotaur*). r, Spearman rank correlation coefficient of the pairwise selection frequencies; Kappa-f, Cohen's kappa statistic of the pairwise binary solutions based on the selection frequency data; Kappa-b, Cohen's kappa statistic of the pairwise binary solutions based on the best solution data.

Alternative scenarios	**CHI-cost scenario (China)**			**CHI-cost scenario (Southeast Asia)**			**CHI-cost scenario (Australia)**		
	r	Kappa-f	Kappa-b	r	Kappa-f	Kappa-b	r	Kappa-f	Kappa-b
Artisanal fisheries	0.950	0.860	0.715	0.606	0.630	0.433	0.307	0.724	0.484
Commercial fisheries	0.760	0.800	0.639	0.446	0.280	0.611	-0.009*	0.553	0.372
Nutrient pollution	0.810	0.600	0.515	0.426	0.520	0.491	0.346	0.625	0.443
Shipping	0.770	0.710	0.559	0.526	0.250	-0.027	-0.026	0.365	0.458

*NOT statistically significant (P > 0.05).

Table 5. 4. Measure of cumulative human impact (CHI), total boundary length (TBL), total number of patches (TP), and ratio of greater-protection areas (GPAs) to lower-protection areas (LPAs, value shown as area by km^2) of the priority solutions based on two approaches (selection frequency vs. best solution) for three regions (and species): Southeast Asia (*H. barbouri*), Australia (*H. abdominalis* and *H. minotaur*) and China (*H. histrix*).

Solution approach	China				Southeast Asia				Australia			
	CHI	TBL	TP	GPAs : LPAs	CHI	TBL	TP	GPAs : LPAs	CHI	TBL	TP	GPAs : LPAs
Selection frequency	197.7	740	41	447:219	343.7	1016	30	828:844	91.5	704	13	0:1415
Best solution	199.9	776	50	407:260	337	1044	39	649:1021	97.7	670	10	0:1415

The selection-frequency approach performed better than the best-solution approach in determining priority solutions from Marxan outputs. Compared with the best solution, the selection frequency derived priorities with higher convergence between CHI-cost and alternative scenarios (Kappa-f > Kappa-b in Table 5.3), except for the shipping scenario. Consequentially, more priorities were selected as greater-protection areas (GPAs) based on the selection frequency than based on the best solution (Table 5.4, Fig. K.1 and K.2 in Appendix K). The priority areas based on the selection frequency also contained less cumulative human impacts, shorter total boundary length, and less patches than based on the best solution (Table 5.4). There were medium to high extents of consistency between the two types of priority solutions (selection frequency vs. best solution) within each scenario (Cohen's kappa = 0.645 – 0.806, all $P < 0.001$).

Given the above comparison, I chose the priorities based on the selection-frequency approach for *H. histrix*. A total of 447 km^2 AOO (47% out of the 944 planning units) were selected as GPAs, and 219 km^2 AOO as LPAs in Hainan and Taiwan (Fig. 5.3 and Fig. K.2).

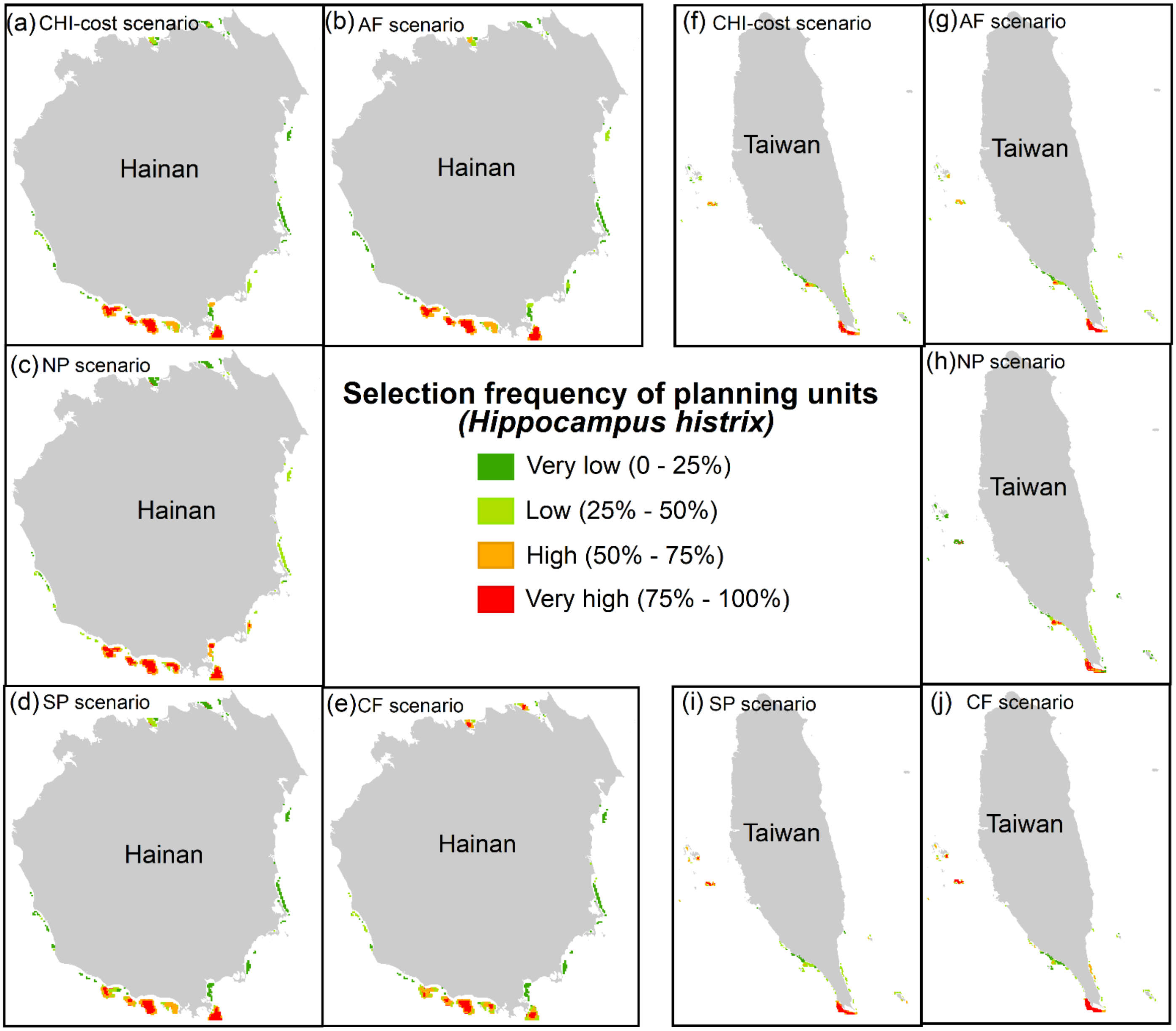


Figure 5. 2. Selection frequency of planning units for the Chinese seahorse population of *Hippocampus histrix* in Hainan (a – e) and Taiwan (f – j) for five cost scenarios: cumulative-human-impact cost (CHI-cost scenario, (a) and (f)), artisanal fisheries (AF scenario, (b) and (g)), nutrient pollution (NP scenario, (c) and (h)), shipping (SP scenario, (d) and (i)), and commercial fisheries (CF scenario, (e) and (j)). Selection frequency was classified into four categories based on the quantiles in the distribution of the selection frequency.

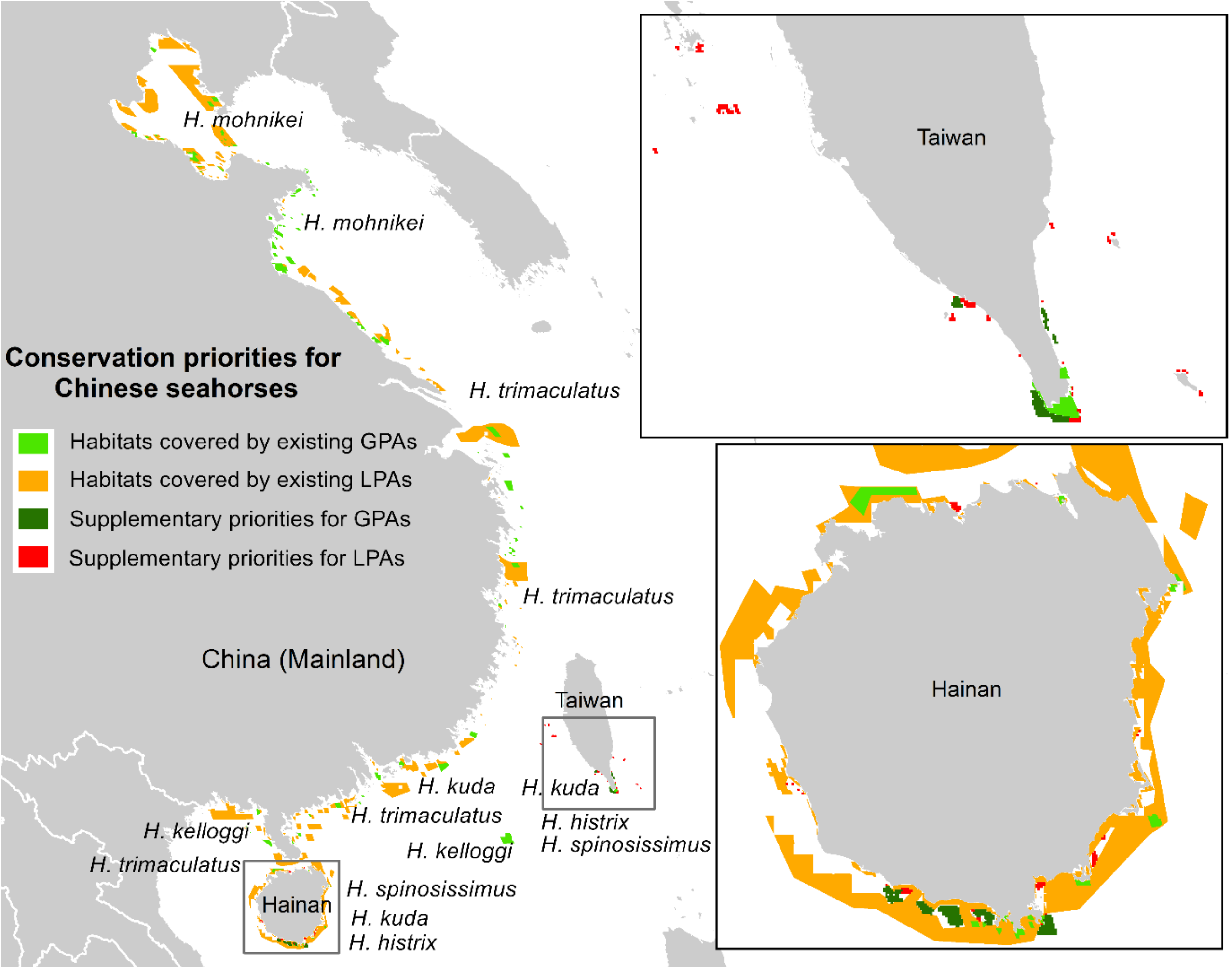


Figure 5. 3. Conservation priorities for Chinese seahorses including habitats covered by current marine protected areas, i.e., greater-protection areas (GPAs) and lower-protection areas (LPAs), and selected priorities for GPAs and LPAs based on the selection frequency in Marxan.

5.4.3 Global MPA coverage in the areas of occupancy (AOO) of seahorses

I found that all the 33 species have some part (10 – 128,299 km^2, or 2 – 85%) of their AOO covered by existing MPAs, but nine species fell short of my already very low target (MPA coverage < 2,000 km^2, Table 5.5). An Australian species currently assessed as ‘Least Concern’ (*H. angustus*) benefited from the highest MPA coverage, while a South African species assessed as ‘Endangered’ (*H. capensis*) was afforded the least MPA coverage. Generally, significantly more AOO was covered by lower-protection areas (LPAs) than by greater-protection areas (GPAs; paired Wilcoxon test, $P < 0.001$; Fig. 5.4a). Indeed, GPAs (206,385 km^2) only accounted for 35% of the overall protected AOOs of seahorses. Furthermore, the protected AOO largely consisted of small patches (area < 10 km^2), itself a conservation concern. I found that threatened species benefited from significantly more protected AOO (absolute value per species) than did non-threatened species (Wilcoxon test, $P < 0.05$). Further tests revealed that, although AOO protected by GPAs (i.e., GPA coverage) did not significantly vary between the two groups (threatened vs. non-threatened, Wilcoxon test, $P = 0.059$; Fig. 5.4b), AOO protected by LPAs (i.e., LPA coverage) was significantly higher for threatened species than for non-threatened species (Wilcoxon test, $P < 0.05$; Fig. 5.4c). However, when the ratio of the MPA coverage to all AOO (per species) was used instead of the absolute value, I found that the MPA-coverage ratio for threatened species was in fact significantly lower than that for non-threatened species (Wilcoxon test, $P < 0.05$).

Among the nine species falling short of conservation target, six rarest species had AOO smaller than 2,000 km^2 based on results from my Chapter 3 (Table 5.4). It was, therefore, important to include all their habitats (i.e., AOOs) as conservation priorities for greater protections. That left only three species for habitat prioritization analyses in Marxan: one Southeast Asian species (*H. barbouri*) and two sympatric Australian species (*H. abdominalis* and *H. minotaur*). I conducted the prioritization for *H. barbouri* alone as it does not share habitats with the other two species, whose prioritization was done together. In the end, I derived a total of 20 priority solutions (2 regions x 5 cost scenarios x 2 priority-selection approaches).

Table 5.5. Summary of marine protected area (MPA) coverage in the areas of occupancy (AOO) of each seahorse species at the global scale, ranked by AOO. Threatened status based on the IUCN Red List: EN, Endangered; VU, Vulnerable; NT, Near Threatened; LC, Least Concern; DD, Data Deficient. Threatened status of Data Deficient species predicted by my Chapter 3: T, threatened; NoT, non-threatened. The protected area (km^2) and main region of the protection for each species (Main region) were summarized for two types of MPAs: greater-protection areas and lower-protection areas. All MPAs (km^2): the overall protected areas for each species by all MPAs. Protection ratio: the ratio of protected AOO to all AOO for each species.

Species	**Threatened status**	**Greater-protection areas**		**Lower-protection areas**		**All MPAs**	**Area of occupancy**	
		Area (km^2)	Main region	Area (km^2)	Main region	(km^2)	Area (km^2)	Protection ratio
H. capensis	EN	-	-	10	South Africa	10	64	0.16
H. fisheri	LC	169	United States	-	-	169	280	0.60
H. sindonis	LC	-	-	155	Japan	155	376	0.41
H. subelongatus	DD, NoT	-	-	58	Australia	58	396	0.15
H. planifrons	LC	323	Australia	4	Australia	327	560	0.58
H. denise	DD, NoT	64	Australia	323	Indonesia	387	832	0.47
H. barbouri	VU	116	Philippines	214	Philippines	330	4,152	0.08
H. coronatus	DD, T	373	South Korea	4,033	Japan	4,406	5,968	0.74
H. minotaur	DD, NoT	413	Australia	172	Australia	585	6,824	0.09
H. breviceps	LC	2,179	Australia	1,949	Australia	4,128	12,360	0.33
H. jayakari	LC	3,096	Saudi Arabia	1,949	Saudi Arabia	5,045	13,036	0.39
H. dahli	LC	3,382	Australia	6,425	Australia	9,807	13,836	0.71
H. zebra	DD, NoT	4,003	Australia	10,259	Australia	14,262	16,708	0.85
H. camelopardalis	DD, NoT	35	Mozambique	4,014	Mozambique	4,049	19,716	0.21
H. abdominalis	LC	509	Australia	672	Australia	1,181	21,776	0.05
H. bargibanti	DD, NoT	4,082	Australia	8,305	Australia	12,387	24,804	0.50
H. zosterae	LC	2,953	United States	8,181	United States	11,134	25,784	0.43

Species	**Threatened status**	**Greater-protection areas**		**Lower-protection areas**		**All MPAs**	**Area of occupancy**	
		Area (km^2)	Main region	Area (km^2)	Main region	(km^2)	Area (km^2)	Protection ratio
H. comes	VU	1,525	Thailand	1,845	Indonesia	3,370	29,880	0.11
H. ingens	VU	2,809	Mexico	13,220	Mexico	16,029	52,560	0.30
H. whitei	EN	12,553	Australia	24,909	Australia	37,462	57,356	0.65
H. pontohi	LC	4,992	Indonesia	13,895	Indonesia	18,887	134,948	0.14
H. guttulatus	DD, T	6,855	Italy	28,969	Spain	35,824	147,056	0.24
H. patagonicus	VU	522	Uruguay	2,171	Argentina	2,693	176,284	0.02
H. erectus	VU	80,008	United States	30,600	United States	110,608	204,396	0.54
H. hippocampus	DD, T	12,592	Netherlands	54,838	United Kingdom	67,430	226,100	0.30
H. reidi	NT	27,813	Cuba	33,331	United States	61,144	398,832	0.15
H. spinosissimus	VU	24,177	Australia	54,949	Australia	79,126	569,124	0.14
H. angustus	LC	29,101	Australia	99,198	Australia	128,299	573,012	0.22
H. histrix	VU	17,503	Indonesia	51,872	Indonesia	69,375	652,468	0.11
H. kelloggi	VU	4,471	Thailand	13,997	China	18,468	667,876	0.03
H. mohnikei	VU	12,975	China	62,314	China	75,289	909,216	0.08
H. kuda	VU	52,257	British Indian Ocean Territory	84,941	Australia	137,198	957,420	0.14
H. trimaculatus	VU	24,695	Indonesia	84,916	China	109,611	1,928,620	0.06

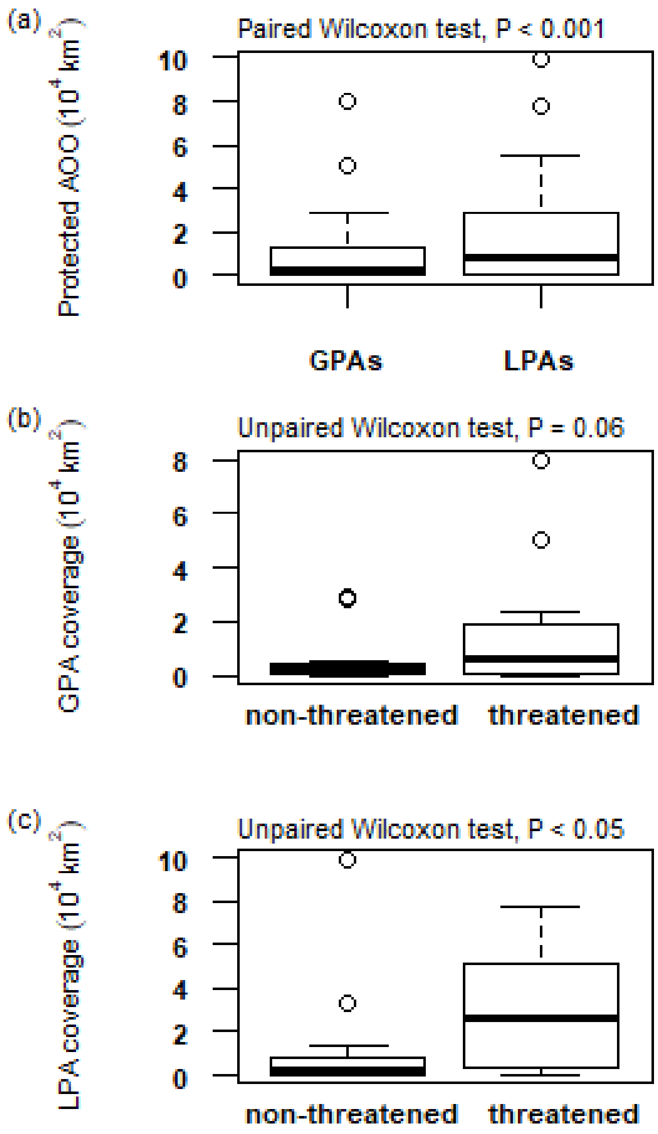


Figure 5. 4. Boxplots of (a) area of occupancy (AOO) covered by marine protected areas estimated for each seahorse species (n = 33), (b) AOO covered by greater-protection areas (GPAs) estimated for non-threatened vs. threatened species, and (c) AOO covered by lower-protection areas (LPAs) estimated for non-threatened vs. threatened species.

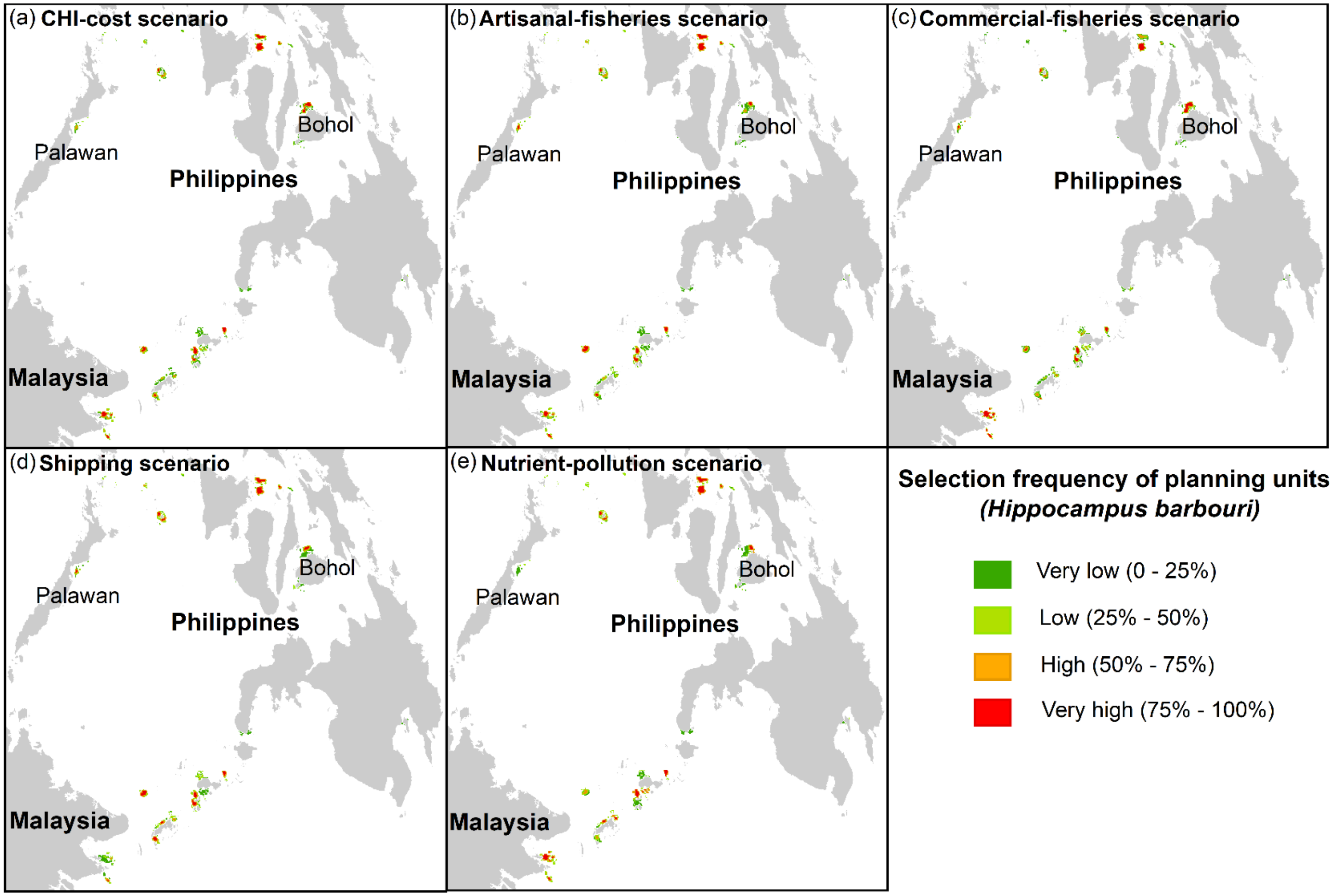


Figure 5. 5. Selection frequency of planning units for a Southeast Asian species (*Hippocampus barbouri*) in five scenarios: (a) cumulative human impact (CHI) cost, (b) artisanal-fisheries cost, (c) commercial-fisheries cost, (d) shipping cost, and (e) nutrient-pollution cost. Selection frequency was classified into four categories based on the quartiles in the distribution of the selection frequency.

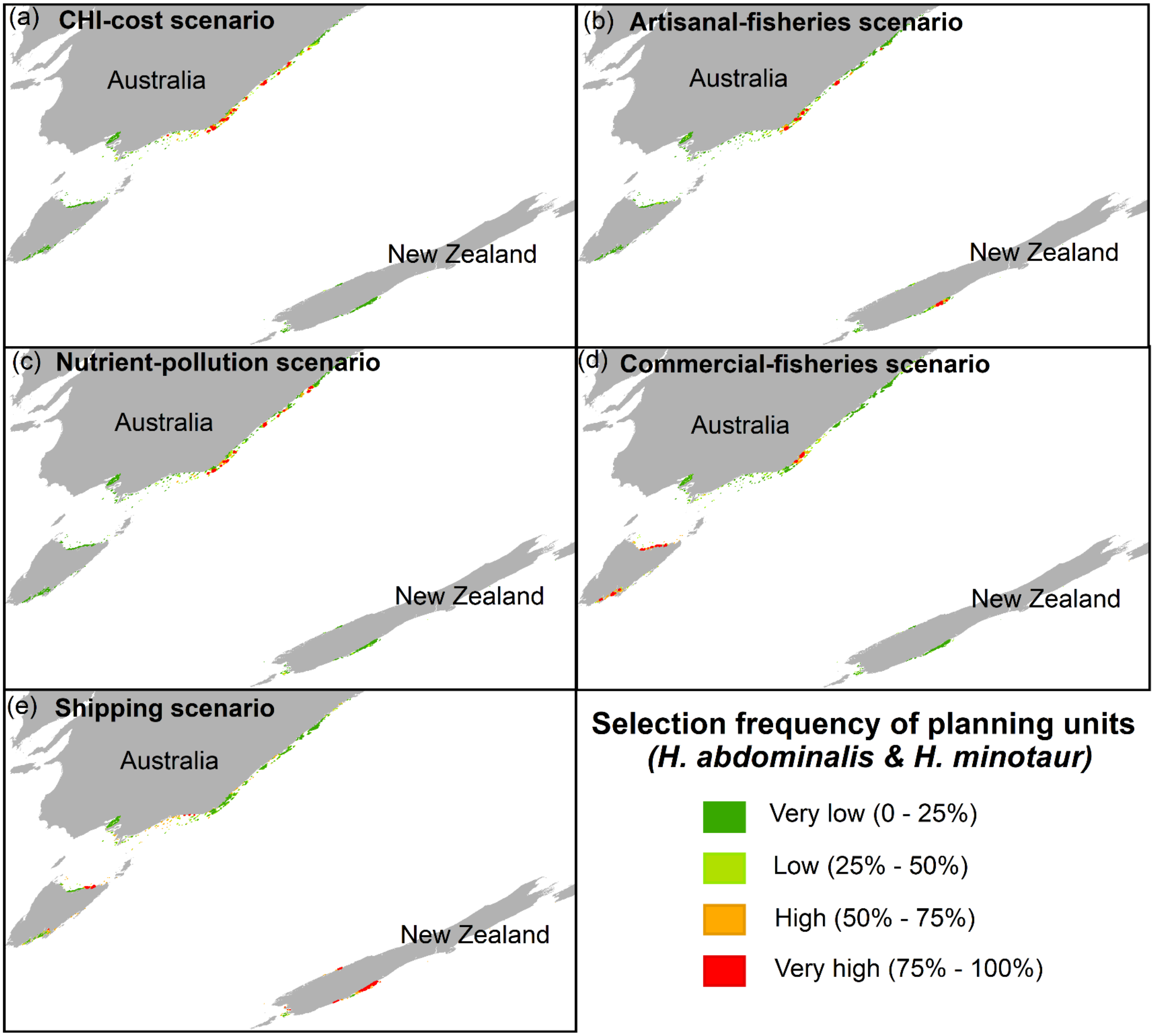


Figure 5. 6. Selection frequency of planning units for the two Australian species (*Hippocampus abdominalis* and *H. minotaur*) in five cost scenarios: (a) cumulative human impact (CHI) cost, (b) artisanal-fisheries scenario, (c) nutrient-pollution scenario, (d) commercial-fisheries scenario, and (e) shipping scenario. Selection frequency was classified into four categories based on the quartiles in the distribution of the selection frequency.

5.4.4 Conservation priorities for three focal species at the global scale

I found that the CHI-cost scenario and alternative scenarios can derive either very similar or very different priorities (Table 5.3). In the case of *H. barbouri*, selection frequency of CHI-cost scenario (Fig. 5.5a) was generally similar to that of artisanal-fisheries (Fig. 5.5b), commercial-fisheries (Fig. 5.5c), shipping (Fig. 5.5d), and nutrient-pollution scenarios (Fig. 5.5e), in decreasing order of similarity ($r = 0.86 – 0.60$). In the case of two Australian seahorses (*H. abdominalis* and *H. minotaur*), selection frequency of CHI-cost scenario (Fig. 5.6a) was generally different from that of artisanal-fisheries (Fig. 5.6b), nutrient-pollution (Fig. 5.6c), commercial-fisheries (Fig. 5.6d), and shipping scenarios (Fig. 5.6e), in decreasing order of similarity ($r = 0.63 – 0.25$).

The priority solution (i.e., selected priority units) based on the selection frequency performed better than that based on the best solution for *H. barbouri*, but not for the two Australian species. Compared with the best solution, the selection frequency derived generally higher convergences (i.e., higher Cohen's kappa statistics) between CHI-cost scenario and alternative scenarios (Kappa-f vs. Kappa-b in Table 5.3). For *H. barbouri*, the selection frequency derived more priority units agreed by all scenarios than did the best solution (Table 5.4, also see Fig. K.3 vs. K.4 in Appendix K). The priority solution based on the selection frequency also contained lower CHI cost, shorter total boundary length, and fewer patches than did the best solution (Table 5.4). For the two Australian species, no priorities were agreed by all five scenarios based on either the selection frequency or the best solution (Table 5.4, also see Fig. K.5 vs. K.6 in Appendix K). The priority solution based on the selection frequency contained slightly lower CHI cost than the best solution (Table 5.4). However, the former contained longer total boundary length and more patches than the latter for the Australian species (Table 5.4). I found medium to high extents of consistency between the two types of priority solutions (selection frequency vs. best solution) within each scenario (Cohen's kappa = 0.622 – 0.927, all $P < 0.001$; Table K.1).

Based on the above comparison, I chose the selection-frequency approach to set priorities for *H. barbouri*, and the best-solution approach for the two Australian species (*H. abdominalis* and *H. minotaur*). For *H. barbouri*, a total of 1672 km^2 area (i.e., 53% out of the 3,156 planning units) in the Philippines and northeast Malaysia were selected as priorities by the CHI-cost scenario (Fig. 5.7). Within these priorities, 828 km^2 area were also selected by all four socioeconomic scenarios (i.e., spatial convergence) and thus were considered as priorities for greater-protection areas (e.g., no-take reserves, Fig. 5.7). The remaining 844 km^2 habitats selected by CHI-cost scenario were not consistently selected by the four alternative scenarios, and thus were considered as priorities for lower-protection areas (e.g., multiple-use zones, Fig. 5.7 and Fig. K.4). For *H. abdominalis* and *H. minotaur*, a total of 1415 km^2 (13% out of the 10,685 planning units) in the southeast Australia were selected by CHI-cost scenario but not agreed by alternative scenarios (Fig. 5.7 and Fig. K.6). These habitats were then all considered as priorities for lower-protection areas. I included the unprotected habitats of the six rarest species in the global priority map (Fig. 5.8). These unprotected habitats included shallow waters in Hawaii, George to Plettenberg Bay (South Africa), Shark Bay (Western Australia), Mandurah to Perth (Western Australia), the Coral Triangle region, and Shizuoka (Japan) (see Fig. K.7 – K.13 in Appendix K).

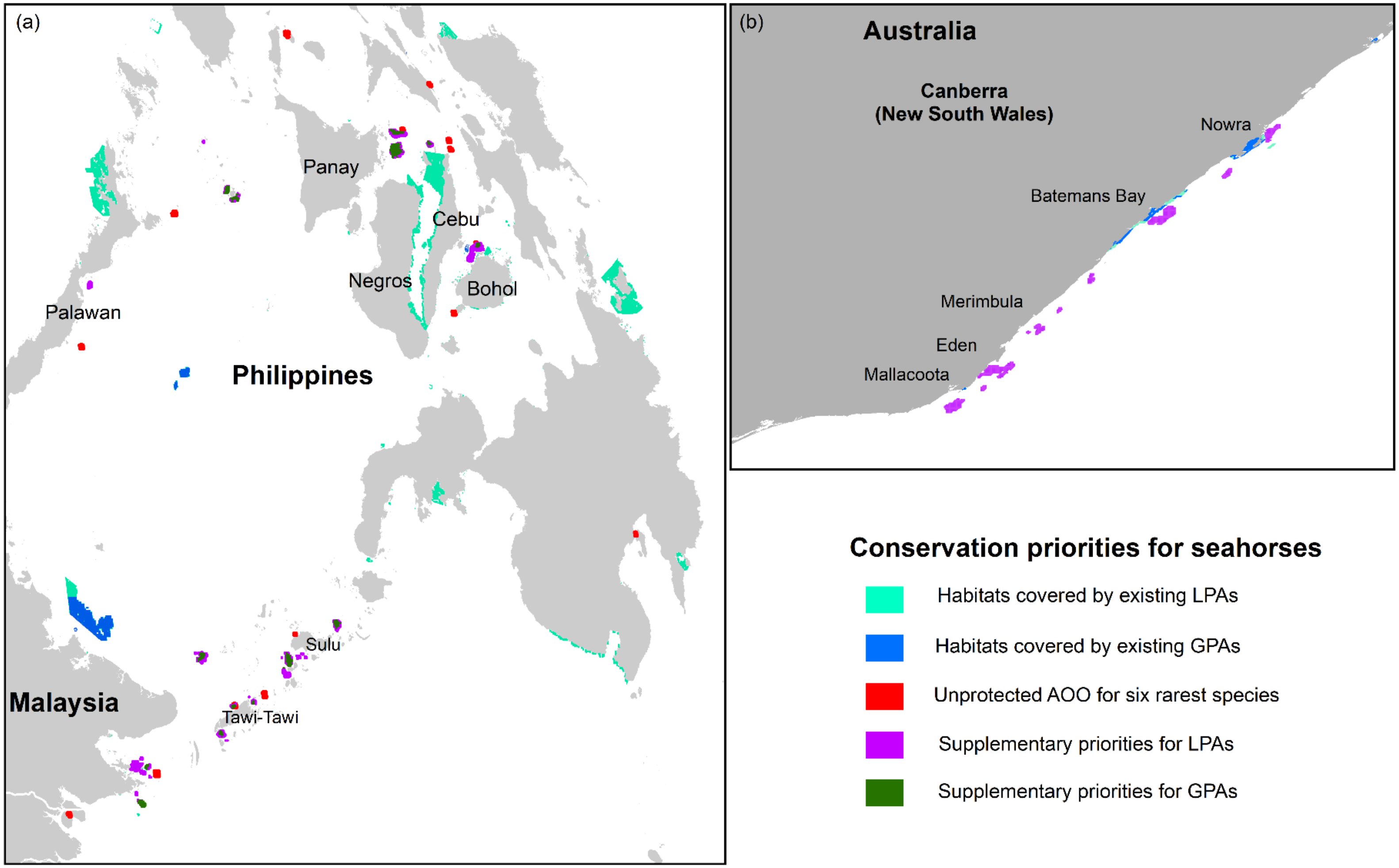


Figure 5. 7. Conservation priorities for the seahorse species in (a) the Philippines and (b) southeastern Australia, including habitats covered by existing lower-protection areas (LPAs) and greater-protection areas (GPAs), habitats not yet in MPAs for six rarest species (here in (a), *H. denise*), and identified priorities for LPAs and GPAs in Marxan.

Figure 5. 8. Global conservation priorities for 33 seahorse species, including habitats covered by lower-protection areas (LPAs) and greater-protection areas (GPAs), habitats falling outside MPAs for six rarest species, supplementary priorities for LPAs and GPMs. Note that conservation priorities in China are included in the map.

5.5 Discussion

My study demonstrates that setting conservation priorities for marine species at large spatial scales can be achieved by integrating multiple socioeconomic costs in my novel prioritization framework. Like its antecedents, my prioritization framework includes a ‘gap analysis’ by assessing the coverage of conservation targets by current protection measures (Margules and Pressey 2000, McIntosh et al. 2017). This analysis helps me narrow down my focus from many seahorse species to just a few least protected species, and from large spatial scales to much smaller scales. Unlike previous frameworks, however, I distinguish among MPAs in terms of their proposed management strictness (GPAs vs. LPAs). This helps us further understand that the protected habitats of seahorses are mainly located in the less-protection MPAs (i.e., LPAs), flagging the uncertainty of true protection effectiveness. The prioritization analysis in different cost scenarios reveals that incorporating different socioeconomic costs can profoundly influence the selection of priorities (Mazor et al. 2014). The level of convergence on priorities between the CHI-cost scenario and other socioeconomic scenarios depends on the focal species/regions and the approaches chosen for priority solution. I find that both the selection-frequency approach and the best-solution approach can derive priority solutions from Marxan, although the former performed better than the latter in two out of the three cases. These approaches may help planners derive practical priority solutions at large spatial scales, which has rarely been addressed in the literature (Ban and Klein 2009). My framework provides a platform for future expansion of prioritization for marine conservation and management at large spatial scales.

5.5.1 Conservation and management priorities for seahorses

Although my study indicates that many seahorse species have a very minimal AOO (that I set) covered by existing MPAs, this finding should not be interpreted as a success for seahorse conservation. In this exercise, I used only on a bottom-line requirement for preventing extinction of some species: area of occupancy (AOO) > 2,000 km^2, according to the IUCN Red List Criterion B2 (IUCN 2005). Such a limit is seldom invoked for marine species whose populations are likely to have plummeted considerably, qualifying them as threatened under Criterion A, long before they qualify under Criterion B, associated with range. Except for the six narrow-ranged species (AOO < 2,000 km^2), reducing the wide-spread global habitats of most seahorse species to just two thousand km^2 would likely be recipe for disaster. My attempt here is purely to identify the species that most require habitat protection, and to determine their priority habitats within this bottom-line requirement, as a starting point for application of this approach.

My research reveals some new insights about the role of existing MPAs to seahorse conservation. Given at least 45% seahorse species are threatened primarily by non-selective bottom fishing (IUCN 2018; see my Chapter 3), MPAs that prevent these fishing practices (e.g., no-take reserves) may play a vital role in safeguarding seahorses. However, my study indicates that this might be not the case by far, given such greater-protection MPAs remain relatively small in covering seahorse habitats in China and worldwide. Moreover, I found that the threatened seahorses may require more habitat protection, since their MPA-coverage rates were lower than those of non-threatened species. In China and worldwide, protecting seahorses is very challenging, since they are experiencing great pressure from anthropogenic activities especially demersal non-selective fishing (see Chapters 2 and 3). I found that China's marine spatial planning and existing MPAs include habitats of these species, suggesting that these MPAs and plans might be meaningful in terms of protecting seahorses and related habitats (e.g., coral reefs and mangroves). However, the effectiveness of implementation for both China's and global MPAs are quite uncertain and may be generally low (Ma et al. 2013, Gill et al. 2017, Zhang et al. 2017). China has historically engaged in little or no social consultation before implementing MPAs (Qiu et al. 2009). Effective implementation will require collaboration between MPA managers and fisheries managers (Hilborn 2016), as well as the engagement of local stakeholders (e.g. fishers) who benefit from seahorse trade (Vincent et al. 2011). In addition, MPAs need good staffing and resourcing to ensure their effectiveness and viable conservation contributions (Gill et al. 2017).

The identified priorities for the Spiny seahorse *H. histrix* in China, especially in Hainan Province, highlight the potential gaps in local government's conservation plans (Lu et al. 2015, Peng et al. 2016). This Vulnerable species is found in various habitats including weedy rocky reefs, sponges, soft corals, and seagrass beds (Lourie et al. 2004, also see my Chapter 3). Their habitats in China were mainly distributed in the southern coastal waters of Hainan. Although there are existing marine reserves for coral reefs (in Sanya) and seagrasses (in Xincun) near the supplementary priorities, these reserves are small and have long been threatened by human activities especially fishing and aquaculture (Yang and Yang 2009, Zhao et al. 2012). Given my study used modeled fishing data and didn't include aquaculture information, ground truthing the feasibility of my identified priorities would be interesting and necessary. This can be done through surveys on local stakeholders such as fishers and farmers (Hamel et al. 2018).

My global-scale study of conservation priorities for seahorses revealed priority species and habitats for future conservation efforts. At the global scale, nine species' AOO are insufficiently covered by current MPAs even against the minimal bottom-line target, indicating that they deserve a high priority for further habitat protection. The priorities for *H. barbouri* contain large areas of spatial convergence that are potential for GPAs (e.g., eastern water of the Panay Island, the Philippines), suggesting likely high compliance of local communities in protecting these waters. However, again this should be examined in the field given my socioeconomic data were model estimates with few studies in the literature to validate them. In contrast, the selected priorities for the two Australian species have no spatial match among all five scenarios, suggesting intensive conflicts between conservation and human activities (e.g., fishing and shipping) in these waters (e.g., New South Wales) (Gladstone 2014, Jordan et al. 2016). The six rarest species (AOO < 2,000 km^2) are currently not assessed as threatened (see Chapters 3 and 4), except the 'Endangered' *H. capensis* (IUCN 2018). Local conservationists need to focus on the AOO of these species, which are currently unprotected, with management strategies that can anticipate and preclude future human impacts. For instance, the pygmy seahorse *H. denise* is highly specialized to live in gorganian corals (e.g., *Annella reticulata*) (Lourie and Randall 2003). Demersal destructive fishing and pollution should be eliminated near these coral habitats within the unprotected AOO. In some instances, the use of permanent artificial habitats (e.g., jetty pilings, swimming nets) may benefit some species such as *H. subelongatus* (Clynick 2008, Hellyer et al. 2011). This might be also true of *H. capensis*, since it has been found in higher density in Reno mattresses within Thesen Islands Marina than in natural habitats (Claassens et al. 2018).

5.5.2 Deriving priority solutions from Marxan’s outputs: selection frequency vs. best solution

I provide a novel approach to determine priority solutions from Marxan’s outputs, which has been understudied in literature (Ban and Klein 2009, Mazor et al. 2014). Previous studies have determined priorities from the selection frequency or from the best solution alone (Ardron et al. 2010, Hamel et al. 2018), and rarely compared the two. I am among the first to do the comparison and demonstrate that using my selection-frequency approach might be generally, although not universally, more meaningful than the best solution in setting conservation priorities. When the selection frequency was used, I recommend my approach to determining the frequency threshold which could meet conservation targets at minimum costs. Previous studies using the selection frequency often derive priority solutions based on arbitrary frequency thresholds (e.g., frequency > 50%) (Mazor et al. 2014, Solovyev et al. 2017), which might not ensure that all conservation targets be met or might result in an over-represented reserve system.

Future researchers should be aware of a few issues about using my selection-frequency and the best-solution approaches. First, my study indicates that using these two approaches can derive very similar or different priority solutions in various scenarios. Given this, I recommend future research compare the two approaches to identify a better solution for specific cases based on methods illustrated in my study. Second, it should be noted that using the single best solution from multiple runs can be risky (Game and Grantham 2008, Ardron et al. 2010). Another set of runs may derive a different version of best solution, given the commonly-used algorithm of Marxan (i.e., simulated annealing) is finding near-optimal rather than absolutely the best solution (Possingham et al. 2000). Third, using selection frequency to determine an exact priority solution is more meaningful when Marxan is often finding near-optimal solutions with appropriate setting of parameters as I did in my study (e.g. high SPF value; Ardron et al. 2010).

5.5.3 Marine spatial prioritization at large spatial scales

I am among the first to set priorities for both strict management (i.e., greater protection) and multiple uses (i.e., less protection) for focal marine species at large spatial scales (Klein et al. 2010; Pompa et al. 2011). It is important to identify an optimal set of priority habitats for wildlife that is associated with minimum costs in conservation planning (Ban and Klein 2009), but this is very challenging at large spatial scales (Solovyev et al. 2017). Usually, marine conservation needs strict marine reserves to constrain fishing pressures upon focal species (Costello and Ballantine 2015). It also requires managers to set up sustainable multiple-use zones where conflicts of interest occur, and to make these zones work together with no-take reserves (Hilborn 2016). By setting priorities for both areas during the planning process, one should be able to establish a network of protected areas that works for both the species and the people that rely on marine resources. The new version of Marxan software, Marxan with Zones, can do comparable work if specific targets for socioeconomic values are considered in the prioritization (Watts et al. 2009).

To identify conservation priorities at large scales, it will work best if planners present an ideal conservation scenario against different socioeconomic scenarios (Cameron et al. 2008, Ban and Klein 2009, Mazor et al. 2014), as I did in my study. This enables planners to explore their commonalities and differences with stakeholders. Previous studies at smaller spatial scales have also used alternative-scenario approaches to help stakeholders better understand and reconcile each other’s interest, with an expectation of finding feasible priority solutions (Ban and Klein 2009). Unlike these local-scale studies, I identified priority solutions based on spatial convergences and divergences among different cost scenarios, which might provide a practical and clear comparison among different stakeholder’s interests at large spatial scales. The map of convergence and divergences based on this comparison demonstrated clear agreements and conflicts between conservation and human uses in the sea. In other conservation circumstances, people might prefer to combine multiple costs into one cost for ease of the planning. But such a combination is very challenging, because weighting each cost appropriately is often difficult or contentious (Cameron et al. 2008, Ban and Klein, 2009). Using my approach could avoid such problem.

Chapter 6: Conclusion

6.1 Overview

My dissertation sought novel approaches in ecological and conservation studies for data-poor marine species at large spatial scales, using seahorses as a case study. Large-scale ecological and conservation research demands a good deal of data, which can be very challenging for understudied species. In Chapter 2, I initiated the first nationwide biogeographic study of Chinese seahorses to examine the application of local ecological knowledge (LEK) in determining species distributions. In Chapter 3, I conducted a global-scale study of 42 seahorse species to explore the appropriate approaches to integrating biogeographic datasets and habitat variables in models of species-habitat relationships. In Chapter 4, I explored modeling techniques to estimate cumulative human impact (CHI) for 42 seahorse species at the global scale, and revealed conservation status for data-poor species. With my final data chapter (Chapter 5), I used results from previous chapters to set conservation priorities for seahorses in China and globally based on a novel framework and priority-solution approaches.

Here I first highlight a few general contributions of my thesis, and then review my four research questions and summarize my corresponding findings in each chapter. I then talk about some of the limitations in my studies with recommendations for future research. I finally conclude the thesis with highlights of the possible value of my studies for current conservation thinking and action.

6.2 Research Contributions

My research offers a rare comprehensive study to address conservation knowledge gaps for data-poor marine species at large spatial scales. We are living in a complex world: On one hand, many species are data-poor and potentially threatened, demanding expeditious conservation actions (Jaric et al. 2016); On the other hand, human development generally draws on ever more natural resources and conservation budgets are usually limited (Margules and Pressey 2000, Spalding et al. 2013). Such complexity usually overwhelms decision makers and undermines global conservation actions (Spalding et al. 2013). Facing this challenge, I identified four critical and entangled questions that can be encountered in conservation planning (Chapter 1). To address these questions, I then conducted a diverse research in species-habitat relationships, cumulative human impacts, threatened status, and conservation planning for data-poor marine species.

My research is of value to the current conservation literature for at least two reasons. First, it presents a rare example of how to integrate multiple data sources and approaches to derive conservation knowledge at large spatial scales. Ecological and conservation knowledge are always limited for many species (Devictor et al. 2010, Agardy et al. 2011, Martin et al. 2012). It is thus crucial for conservationists to know which data sources and approaches are potentially useful. My thesis addressed this important question in different chapters. Second, my research is among the first to address these issues in-depth at large spatial scales (globally and nationally). Although there are many conservation studies for marine species at local spatial scales (< 100 km^2), publications focused on large spatial scales are limited (Velasco et al. 2015). Drawing a large-scale picture could expand our understanding and enhance conservation of marine species, which usually have wide geographic ranges. Local studies may reveal population knowledge in small areas, while a holistic understanding of species distributions requires global-scale studies. When protected areas are proposed to protect marine species, ignoring regional or global priorities may lead to protecting less important habitats in local areas (Mills et al. 2010, Agardy et al. 2011).

The approaches and techniques in my thesis provide a template that can be used for conservation studies on other data-poor marine species. My interview approach and down-scaling technique (Chapter 2) can be useful for commercially valuable or charismatic fishes (e.g., dogtooth grouper, whale shark, small-eye stingray) (Moore 2010, Stacey et al. 2012, Boltachev et al. 2013), marine mammals (e.g., killer whales) (Higdon et al. 2014), and sea turtles (Moore et al. 2010). My novel approaches to selecting useful data in modeling species-habitat relationships (Chapter 3) can be applicable for widespread but locally rare animals, such as sea turtles (Cornwell and Campbell 2012), whale sharks (Araujo et al. 2017), and humpback whales (Bruce et al. 2014). My cumulative-human-impact (CHI) approaches (Chapter 4) can be used to understand threat patterns and conservation status for thousands of Data Deficient marine species on IUCN Red List (IUCN 2018). Such species-level knowledge (e.g., CHI maps) will supplement the ecosystem-level data to better guide threat management and global biodiversity conservation in the marine realm (Maxwell et al. 2013). The prioritization framework and techniques developed in Chapter 5 can further help to set large-scale conservation priorities for marine species or habitats. These large-scale priorities may facilitate nations to meet their obligations for Aichi Biodiversity Target 11.

I now comment on how my findings contribute to address the research question in each analytical chapter.

6.2.1 How is local fishers' ecological knowledge useful for mapping distributions for data-poor marine species? (Chapter 2)

My Chapter 2 provides a rare example of integrating local ecological knowledge (LEK) in species distribution models (SDMs) for data-poor marine species. Previously, LEK has been used in both terrestrial and marine studies on a variety of ecological and conservation issues, including mapping species habitats (Drew 2005, Anadon et al. 2009). However, the application of LEK remains controversial. For instance, LEK-based maps are usually very coarse in resolution (Aylesworth et al. 2017, Laze and Gordon 2016, Selgrath et al. 2016) – as I found in my Chapter 2, and thus might be imprecise in guiding conservation practices such as conservation planning and other activities (Gilchrist et al. 2005). My study indicates that such problem can be addressed by applying a downscaling technique and integrating LEK with quality data (i.e., fine-resolution occurrence) in SDMs. Although similar studies have been done in terrestrial systems (Anadon et al. 2010, Laze and Gordon 2016), I am among the first to take this more comprehensive approach in marine systems. Such technique may extend the capability of future researchers to use LEK in predicting species distributions for data-poor marine species. Additionally, I pointed out that combining two coarse-grain datasets can degrade model performance of SDMs.

My results from Chapter 2 indicate that LEK can provide useful information for data-poor marine fishes, in line with the limited studies of other taxa in China (e.g., amphibians and mammals) (Turvey et al. 2013, Pan et al. 2016, Liu et al. 2017). Although LEK has been studied over three decades and increasingly recognized as valuable around the world (Johannes 1989, Brook and McLachlan 2008, Beaudreau and Levin 2014, Martinez-Levasseur et al. 2017, Bélisle et al. 2018), similar research only has emerged very recently in China (Turvey et al. 2013, Pan et al. 2016, Liu et al. 2017). A relevant obstacle is likely that many colleagues, especially in less developed countries, lack the training in applying LEK (e.g., interview approaches) or the awareness of LEK. This is especially true in China based on my own observation in conducting this study. Like the previous few cases, my research on China's seahorses present an evidence that LEK can provide instrumental information for data-poor species. Moreover, I demonstrated that it is possible to derive species-level distribution maps from fishers under certain circumstances, which is rare in the literature (Aylesworth et al. 2017). Such finding is encouraging to colleagues in countries like China, where marine biogeographic data are often limited (Liu 2013, Costello et al. 2010).

6.2.2 How can we derive useful species occurrences and habitat variables for identifying suitable habitats of data-poor marine species at the global scale? (Chapter 3)

My Chapter 3 provides a rare study on modeling species-habitat relationships for data-poor marine species. Species-habitat relations are the cornerstone in many of the ecological studies (Grand and Cushman 2004, Rice 2005, Stuber et al. 2017), as well as in predicting suitable habitats for species-conservation purpose (Rice 2005, Franklin 2010). Understanding how marine organisms use their habitats, and how they interact with or respond to the change of the ocean environment is an urgent need, because many marine ecosystems are threatened by human activities (Halpern et al. 2008, Halpern et al. 2015). My research contributes to the current literature by addressing seahorses at the global scale, which are difficult because of data paucity. I demonstrated that this can be achieved through mining various data sources, both the historical records and current observations, both quality-unknown sources and quality research datasets. Moreover, I compared modeling results based on each dataset and various ways of integrating these datasets, which is rare in the literature. Such data comparison and integration approaches provide useful insights for future relevant studies on data-poor marine species at global scales.

The results show that collating and adding data of unknown quality such as from citizen sciences (CS) and museum collections (MC) into limited research-grade data (RG) can allow better models of species-habitat relationships, which is encouraging to many data-poor species. Unlike Chapter 2, CS and MC constitute the datasets with unknown quality at the global scale in Chapter 3. I demonstrated that (1) the quality-unknown CS and MC can provide more abundant occurrences than RG for data-poor species, and (2) adding MC and/or CS to RG can significantly improve model accuracy. Although different data sources, including MC and RG, have already been used in other marine SDM studies such as the ocean sunfishes (Phillips et al. 2017a), few studies have focused on the appropriate way of combining these different datasets. Additionally, CS have rarely been applied in SDMs for marine species, although they are used for other purposes such as monitoring population abundance (Thiel et al. 2014). My research in Chapter 3 highlights the utility of CS in predicting suitable habitats for data-poor marine species. Moreover, as found in Chapter 2, I demonstrate that it is more sensible to combine quality-unknown data (e.g., MC and CS) with quality data to derive better SDMs, but integrating multiple datasets with unknown quality is not encouraged because of potential error accumulation.

I highlight the importance of using appropriate macrohabitat variables to derive habitat knowledge for sedentary and locally-rare marine species, which is rarely addressed in the literature. Selecting appropriate predictors is vital to model species-habitat relations in SDMs (Petitpierre et al. 2017). For locally rare species whose distributions may highly depend on scarce patchy habitats and resources therein, it is important to understand which habitats they might prefer. My research indicates that the proximity to habitat-forming benthos (i.e., continuous variable) can be a more useful predictor than presence/absence of habitat-forming benthos (i.e., binary variable) in SDMs for rare marine species, even for those with low mobility. Such finding contrasts with the previous assumption in terrestrial animals: Distributions of low-mobility species are more related with macrohabitat presence/absence than proximity to macrohabitats, given they cannot move far away from macrohabitats (Jackson and Robertson 2011). My results suggest that low-mobility species can still move among macrohabitats through external forces, which is especially common in the ocean given that the dispersal of many marine species (e.g., fish with pelagic larval stages) relies on ocean currents and buoyant agents (e.g., seagrass/macroalgae debris) (Kinlan and Gaines 2003, Cowen and Sponaugle 2009).

6.2.3 How can we estimate cumulative human impacts and threatened status for data-poor marine species? (Chapter 4)

In Chapter 4, I addressed gaps in our knowledge about pressures on data-poor marine species. Over past decade, CHI has been mapped for a variety of marine species, including mammals, sea birds, sea turtles, and commercial fishes (Davidson et al. 2012, Maxwell et al. 2013, Korpinen and Andresen 2016). My research is among the first to study a group of data-poor and rare marine fish (i.e., seahorses) at the global scale. A unique feature of my research is that I have not only created a spatial model to map CHI, but also built a non-spatial model to estimate human-impact indices. The latter is especially important to many data-poor species, for which spatial data such as presence-probability maps are rarely available for application of the spatial CHI model. Moreover, I illustrated that human-impact indices derived from the non-spatial model are more useful than the spatial model in predicting species threatened status. This finding encourages the use of more expert knowledge (with the non-spatial model) in estimating threatened status for data-poor species. However, the benefit of the spatial CHI model is also apparent if the goal is to identify threat patterns in species habitats (Maxwell et al. 2013). Therefore, I recommend future researchers use both spatial and non-spatial models to derive different types of threat knowledge for data-poor species.

My results suggest that CHI is more complex than a linear-additive model can explain, although the latter can derive estimates generally consistent with IUCN Red List assessments. Current estimations of CHI on marine ecosystems are largely based on the assumptions that species' responses to impacts are linear and interactions among impacts are additive (Halpern and Fujita 2013). Although it is well documented that threshold responses and synergistic and mitigative interactions commonly exist in marine ecosystems

(Crain et al. 2008), there is insufficient data to estimate relations in CHI models (Halpern et al. 2015). Linear-additive models remain popular as they are simple and transparent. My research is among the first attempts of using linear-additive models to estimate CHI for a genus of data-poor marine species at the global scale. The results demonstrate that although a linear-additive model can derive useful CHI estimates, the realistic cumulative impact might be much more complicated, in line with previous findings (Andersen et al. 2015, Bevilacqua et al. 2018). However, when little is known about species responses and impact interactions, which is likely common in data-poor species, my finding encourages the use of linear-additive models. But the results should be validated with relevant observations/assessments (e.g., IUCN Red List assessments).

I am among the first to indicated that human-impact indices have the potential to be used for predicting conservation status. Although IUCN Red List Categories and Criteria have been widely used to quantitatively assess species threatened status, already more than 2,000 marine species of those that have been evaluated (and most have never been evaluated) are listed as Data Deficient (IUCN 2018). Lack of sufficient population data to apply IUCN criteria is a major obstacle in the assessments. My research indicates that this data-paucity problem can be addressed with human-impact indices and machine learning techniques (e.g., random forest). Although previous studies have also applied similar techniques to predict conservation status both on land and in the ocean, they have been based mostly on biological (e.g., life history traits) and geographical/environmental factors (Reynolds et al. 2005, Jetz and Freckleton 2015, Luiz et al. 2016), and only a few studies also used threat information such as human-population density (Bland 2017) or CHI on marine ecosystems (Davidson et al. 2012). My research thus highlights the utility of human-impact indices, which integrated threat intensity and species vulnerability (based on species-level CHI models), in predicting threatened status for Data Deficient species.

6.2.4 How can we identify priority areas for data-poor marine species at large spatial scales? (Chapter 5)

In Chapter 5, I create a novel framework to set conservation priorities for marine species at large spatial scales. Marine conservation planning at large spatial scales is challenging, despite a rising demand for large MPA networks around the world (Douvere 2008, Mazor et al. 2014). My research is timely in addressing the issue by proposing a prioritization framework. It embraces the ‘gap analysis’ concept from systematic conservation planning (Margules and Pressey 2000, McIntosh et al. 2017), and discriminates priorities between strictly-managed reserves and multiple-use zones, keeping in mind the frequent need to reconcile conservation and socioeconomic use (Gill et al. 2017). The framework also incorporates one part of IUCN criterion B2 (i.e., area of occupancy < 2,000 km^2) in defining bottom-line conservation targets (Maze et al. 2008), rather than using arbitrary values (e.g., 10% of each habitat) as in previous studies (Klein et al. 2013, Mazor et al. 2014). That said, this Criterion B has seldom been used for marine species, being less relevant than Criterion A (decline trends) and may represent only a notional bottom line. It is also rare to include large-scale socioeconomic costs in prioritization frameworks (Klein et al. 2013), as did in my thesis. Such a framework may lay a solid ground for further methodology expansion in conservation prioritization at large spatial scales.

My research indicates that the framework and associated prioritization methods are useful in conservation prioritization for seahorse species. I demonstrate that conservation prioritization for data-poor species can be done based on the inferred habitats (i.e., area of occupancy; Chapter 2 and 3), estimated cumulative human impacts (Chapter 4), and socioeconomic costs (Chapter 4). The use of ‘gap analysis’ in estimating MPA coverages can largely reduce the dataset size, thus making my planning tools (i.e., Marxan software) more applicable. This is encouraging to future prioritization studies aiming at large-scale planning. By comparing an ideal conservation scenario with alternative socioeconomic scenarios, I then identified priorities for strict management (i.e., greater protection) and multiple uses (i.e., lower protection). The results suggest that the two major outputs of Marxan (i.e., selection frequency vs. best solution) can be used to determine priority solutions, with the selection-frequency approach tending to derive better priority networks than the best-solution approach. Such finding is helpful to address the methodology gap of deriving priority solutions from Marxan, especially in large-scale studies (Ardron et al. 2010, Hamel et al. 2018).

6.3 Caveats and Recommendations for Future Research

In conducting this research, I encountered some challenges related to each research chapter. Some of these challenges have been discussed in the above chapters. Here I highlight several other issues that were not covered sufficiently elsewhere. I then shed a light on the directions for future studies on data-poor marine species.

I found that the major challenges in Chapter 2 were the availability of macro-habitat data (e.g., distributions of soft corals and macroalgae) and the coarse-grain nature of fishers' maps. I learned from local fishers that seahorses are usually found with various habitat-forming benthos (e.g., macroalgae, sea fans) in China's seas. However, given data paucity of these habitat-forming organisms in China, I could not examine well the species-habitat relations for Chinese seahorse populations. Such a barrier calls for Chinese biologists to document and publish biogeographic data for their marine life including vital habitat-forming creatures (Liu 2013). On the other hand, the use of coarse-grain biogeographic maps from fishers can be a common challenge elsewhere (Laze and Gordon 2016). To solve this problem, I developed a downscaling technique that requires a minimum of five geo-referenced occurrences to generate a preliminary suitability map. However, such data requirement could not be met elsewhere, especially in poorly-surveyed regions such as in the deep seas (Costello et al. 2010). I then recommend future researchers use alternative sources of information including citizen sciences and expert knowledge to fill the gap: either to gather fine-resolution occurrences or to elicit species-habitat relations.

My research in Chapter 3 was challenged by the limitation of both habitat data and species occurrences. As for Chapter 2, habitat data included in Chapter 3 were not exclusive as well, although there were abundant geo-referenced data of habit-forming benthos such as soft corals and sponges at the global scale from open-access database (e.g., Ocean Biogeographic Information System). Other habitats used by seahorses, such as sandy or muddy open bottoms and artificial habitats (e.g., harbor swimming nets, Reno mattress; Foster and Vincent, 2004, Clynick 2008, Claassens et al. 2018), were however not globally available for my research. It is likely that this limitation will last at large spatial scales but can be more easily addressed in local scales where *in situ* surveys can be carried out in reasonable time periods. Citizen science programs that engage participants (e.g., divers) to record underwater natural and artificial habitats in coastal waters might be a promising approach to fill this information gap at large spatial scales in the future. This is also true to address another challenge in this chapter: lacking occurrences for some understudied species. Thanks to the development of internet and smartphones, more and more citizens are participating in citizen-science programs around the world (Crain et al. 2014). I expect that citizen sciences will play a more vital role in future marine conservation, especially for the vast numbers of data-poor species (Devictor et al. 2010).

My study on cumulative human impact (CHI) on seahorse species in Chapter 4 is subject to several limitations that may be encountered elsewhere. First, given that our CHI models rely on expert knowledge to derive vulnerability measures for each species-stressor pair, I found it was challenging to find and engage experts to meet this end. I highly recommend that future researchers take the opportunities of international meetings/conferences to reach out to the experts. It is also important to make the evaluation protocol concise and easy to follow for experts, and it may be better to gather their evaluations through interviews than questionnaires or online surveys. A further concern is the very small pool of experts on many marine taxa, which meant that there was some overlap in those who conducted IUCN conservation assessments and those who offered expert opinion on threats; convergence in findings may not, therefore, be particularly surprising. Second, future researchers should be aware of other approaches to eliciting expert knowledge that were not applied in my research. These alternative but more complex approaches include fuzzy logic and multi-criteria decision models (Cheung et al. 2005, Teck et al. 2010). Third, I did not include life-history variables or geographic factors in predicting the threatened status for seahorse species, given that my focus lay in examining the utility of human-impact indices and life-history and geographical data are only available for some seahorse populations. However, these biological and geographical data might be available in other cases and should then be used in predicting threatened status (Davidson et al. 2012).

Future researchers should take caution with at least two issues in my research in Chapter 5. First, it should be noted that robust conservation planning relies on the good mapping of conservation cost to stakeholders, which is a real challenge in marine conservation planning (Mazor et al. 2014). For instance, mapping the cost of fisheries, a major type of human uses of the oceans, is still in its infancy (Hamel et al. 2018). This is partly because of gaps in coverage or accessibility of spatially-explicit and precise fisheries data, especially for very large spatial scales (e.g., nationally or globally; Watson et al. 2004, Kroodsma et al. 2018). In Chapter 5, I used the modeled catch dataset from previous studies (Halpern et al. 2015), as this was the most available and spatially-explicit data at the beginning of my initiative. Recently, the Global Fishing Watch program has provided another valuable database for mapping global industrial fishing effort on the oceans (Kroodsma et al. 2018). Another obstacle of incorporating fisheries value in MCP is the difficulty of assigning monetary values to different fisheries in space (Sala et al. 2002, O'Higgins et al. 2010, Guerry et al. 2015). This challenge can be more crucial if values of multiple fisheries sectors are integrated to a single monetary value. Second, my conservation prioritization was done only for seahorses given that they can be flagship and indicator species for important marine habitats. However, other data-poor marine species may not have such characteristics. In such cases, their conservation planning may need to be embedded in a holistic project for the whole communities or ecosystems (Norris 2012).

6.4 Conclusions

Human wellbeing relies on a biodiverse and healthy ocean, which has been profoundly degraded due to centuries of intensified anthropogenic activities. Although we have taken some remedial measures around the world, reversing this trend undoubtedly urgently requires many more initiatives to address the remaining gaps. Marine biologists play an essential role by empowering decision makers with scientific knowledge about the species, their habitats and major threats, and priority solutions. However, such data are lacking for thousands of data-poor species, which is especially true in many developing countries (e.g., China). My thesis addressed these knowledge gaps related to data-poor marine species at large spatial scales. From my research, it is evident that these knowledge gaps can be filled with a variety of datasets and techniques: local ecological knowledge, citizen sciences, museum collections, peer-reviewed literature, expert knowledge, habitat-feature datasets, interviews, model estimates, GIS, and decision-making tools. Based on these data and approaches, I derived scientific knowledge for a genus of data-poor marine fish – seahorses

(*Hippocampus* spp.): suitable habitats, threatened status, cumulative human impacts, major stressors, and conservation priorities. Such information may contribute to guide conservation actions for these species with global concern. Although my thesis used seahorses as the case study, the approaches and findings are relevant to addressing similar knowledge gaps for many other data-poor species in the marine realm. It is my expectation that this research can help and stimulate more conservation studies and actions for data-poor marine species.

References

Agardy, T., G. N. di Sciara, and P. Christie. 2011. Mind the gap Addressing the shortcomings of marine protected areas through large scale marine spatial planning. Marine Policy 35(2):226-232.

Altman, D. G. and J. M. Bland. 1994. Diagnostic tests. 1: Sensitivity and specificity. BMJ: British Medical Journal 308(6943):1552.

Alverson, D. L. 1994. A global assessment of fisheries bycatch and discards. Food & Agriculture Org.

Anadón, J. D., A. Giménez, and R. Ballestar. 2010. Linking local ecological knowledge and habitat modelling to predict absolute species abundance on large scales. Biodiversity and Conservation 19(5):1443-1454.

Anadon, J. D., A. Gimenez, R. Ballestar, and I. Perez. 2009. Evaluation of Local Ecological Knowledge as a Method for Collecting Extensive Data on Animal Abundance. Conservation Biology 23(3):617-625.

Anadón, J. D., A. Giménez, R. Ballestar, and I. Pérez. 2009. Evaluation of local ecological knowledge as a method for collecting extensive data on animal abundance. Conservation Biology 23(3):617-625.

Andersen, J. H., B. S. Halpern, S. Korpinen, C. Murray, and J. Reker. 2015. Baltic Sea biodiversity status vs. cumulative human pressures. Estuarine, Coastal and Shelf Science 161:88-92.

Araujo, G., S. Snow, C. L. So, J. Labaja, R. Murray, A. Colucci, and A. Ponzo. 2017. Population structure, residency patterns and movements of whale sharks in Southern Leyte, Philippines: results from dedicated photo-ID and citizen science. Aquatic Conservation: Marine and Freshwater Ecosystems 27(1):237-252.

Ardron, J. A., H. P. Possingham, and C. J. Klein. 2010. Marxan good practices handbook. Victoria, BC.

Asaad, I., C. J. Lundquist, M. V. Erdmann, and M. J. Costello. 2018. Delineating priority areas for marine biodiversity conservation in the Coral Triangle. Biological Conservation 222:198-211.

Aswani, S., X. Basurto, S. Ferse, M. Glaser, L. Campbell, J. E. Cinner, T. Dalton, L. D. Jenkins, M. L. Miller, and R. Pollnac. 2018. Marine resource management and conservation in the Anthropocene. Environmental conservation 45(2):192-202.

Aubry, K. B., C. M. Raley, and K. S. McKelvey. 2017. The importance of data quality for generating reliable distribution models for rare, elusive, and cryptic species. PLoS One 12(6):e0179152.

Avila, I. C., K. Kaschner, and C. F. Dormann. 2018. Current global risks to marine mammals: Taking stock of the threats. Biological Conservation 221:44-58.

Aylesworth, L., J. Lawson, P. Laksanawimol, P. Ferber, and T. L. Loh. 2016. New records of the Japanese seahorse *Hippocampus mohnikei* in Southeast Asia lead to updates in range, habitat and threats. Journal of fish biology 88(4):1620-1630.

Aylesworth, L., R. Phoonsawat, P. Suvanachai, and A. C. J. Vincent. 2017. Generating spatial data for marine conservation and management. Biodiversity and conservation 26(2):383-399.

Aylesworth, L., J. Xavier, T. Oliveira, G. Tenorio, A. Diniz, and I. Rosa. 2015. Regional-scale patterns of habitat preference for the seahorse *Hippocampus reidi* in the tropical estuarine environment. Aquatic ecology 49(4):499-512.

Balbontín, J. 2005. Identifying suitable habitat for dispersal in Bonelli’s eagle: an important issue in halting its decline in Europe. Biological Conservation 126(1):74-83.

Balmford, A., P. Gravestock, N. Hockley, C. J. McClean, and C. M. Roberts. 2004. The worldwide costs of marine protected areas. Proceedings of the National Academy of Sciences of the United States of America 101(26):9694-9697.

Ban, N. C., G. J. A. Hansen, M. Jones, and A. C. J. Vincent. 2009. Systematic marine conservation planning in data-poor regions: Socioeconomic data is essential. Marine Policy 33(5):794-800.

Ban, N. C. and C. J. Klein. 2009. Spatial socioeconomic data as a cost in systematic marine conservation planning. Conservation Letters 2(5):206-215.

Beaudreau, A. H. and P. S. Levin. 2014. Advancing the use of local ecological knowledge for assessing data-poor species in coastal ecosystems. Ecological Applications 24(2):244-256.

Beitinger, T. L. and L. C. Fitzpatrick. 1979. Physiological and ecological correlates of preferred temperature in fish. American Zoologist 19(1):319-329.

Bélisle, A. C., H. Asselin, P. LeBlanc, and S. Gauthier. 2018. Local knowledge in ecological modeling. Ecology and Society 23(2):14.

Belkin, I. M. 2009. Rapid warming of large marine ecosystems. Progress in Oceanography 81(1):207-213.

Bennett, N. J., R. Roth, S. C. Klain, K. M. A. Chan, D. A. Clark, G. Cullman, G. Epstein, M. P. Nelson, R. Stedman, T. L. Teel, R. E. W. Thomas, C. Wyborn, D. Curran, A. Greenberg,

J. Sandlos, and D. Verissimo. 2017. Mainstreaming the social sciences in conservation. Conservation Biology 31(1):56-66.

Bergmann, M., H. Hinz, R. E. Blyth, M. J. Kaiser, S. I. Rogers, and M. Armstrong. 2004. Using knowledge from fishers and fisheries scientists to identify possible groundfish 'Essential Fish Habitats'. Fisheries Research 66(2-3):373-379.

Berkes, F. 1993. Traditional ecological knowledge: Concepts and cases. Page 1 *in* Inglis, J. editor. Traditional ecological knowledge in perspective. International Program on Traditional Ecological Knowledge and International Development Research Centre, P.O. Box 3443, Station D, Ottawa, Ontario KIP 6P4.

Bernstein, R. 1982. Sea surface temperature estimation using the NOAA 6 satellite advanced very high resolution radiometer. Journal of Geophysical Research: Oceans 87(C12):9455-9465.

Bevilacqua, S., G. Guarnieri, G. Farella, A. Terlizzi, and S. Fraschetti. 2018. A regional assessment of cumulative impact mapping on Mediterranean coralligenous outcrops. Scientific reports 8(1):1757.

Bland, L. M. 2017. Global correlates of extinction risk in freshwater crayfish. Animal Conservation 20(6):532-542.

Bland, L. M., B. Collen, C. D. L. Orme, and J. Bielby. 2015. Predicting the conservation status of data-deficient species. Conservation Biology 29(1):250-259.

Boakes, E. H., P. J. McGowan, R. A. Fuller, D. Chang-qing, N. E. Clark, K. O'Connor, and G. M. Mace. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. PLoS biology 8(6):e1000385.

Boehm, J., L. Woodall, P. R. Teske, S. A. Lourie, C. Baldwin, J. Waldman, and M. Hickerson. 2013. Marine dispersal and barriers drive Atlantic seahorse diversification. Journal of biogeography 40(10):1839-1849.

Boltachev, A. R., E. P. Karpova, S. N. Brodskij, and M. P. Kirin. 2013. First registration of the dogtooth grouper Epinephelus caninus (Osteichthyes, Serranidae) in the Black Sea. Mors'kyji ehkologichnyji zhurnal/Marine ecological journal 12(3):22-23.

Bonney, R., J. L. Shirk, T. B. Phillips, A. Wiggins, H. L. Ballard, A. J. Miller-Rushing, and J. K. Parrish. 2014. Next Steps for Citizen Science. Science 343(6178):1436-1437.

Brook, R. K. and S. M. McLachlan. 2008. Trends and prospects for local knowledge in ecological and conservation research and monitoring. Biodiversity and conservation 17(14):3501-3512.

Brooks, T. M., R. A. Mittermeier, G. A. da Fonseca, J. Gerlach, M. Hoffmann, J. F. Lamoreux, C. G. Mittermeier, J. D. Pilgrim, and A. S. Rodrigues. 2006. Global biodiversity conservation priorities. Science 313(5783):58-61.

Brooks, T. M., R. A. Mittermeier, C. G. Mittermeier, G. A. Da Fonseca, A. B. Rylands, W. R. Konstant, P. Flick, J. Pilgrim, S. Oldfield, and G. Magin. 2002. Habitat loss and extinction in the hotspots of biodiversity. Conservation Biology 16(4):909-923.

Brown, J. L. 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods in Ecology and Evolution 5(7):694-700.

Bruce, E., L. Albright, S. Sheehan, and M. Blewitt. 2014. Distribution patterns of migrating humpback whales (Megaptera novaeangliae) in Jervis Bay, Australia: A spatial analysis using geographical citizen science data. Applied Geography 54:83-95.

Caddy, J. 1968. Underwater observations on scallop (Placopecten magellanicus) behaviour and drag efficiency. Journal of the Fisheries Board of Canada 25(10):2123-2141.

Calabrese, J. M., G. Certain, C. Kraan, and C. F. Dormann. 2014. Stacking species distribution models and adjusting bias by linking them to macroecological models. Global Ecology and Biogeography 23(1):99-112.

Caldwell, I. R. and A. C. Vincent. 2013. A sedentary fish on the move: effects of displacement on long-snouted seahorse (*Hippocampus guttulatus* Cuvier) movement and habitat use. Environmental Biology of Fishes 96(1):67-75.

Cameron, S. E., K. J. Williams, and D. K. Mitchell. 2008. Efficiency and concordance of alternative methods for minimizing opportunity costs in conservation planning. Conservation Biology 22(4):886-896.

Carter, B. T. and E. A. Nielsen. 2011. Exploring ecological changes in Cook Inlet beluga whale habitat though traditional and local ecological knowledge of contributing factors for population decline. Marine Policy 35(3):299-308.

Carwardine, J., K. A. Wilson, M. Watts, A. Etter, C. J. Klein, and H. P. Possingham. 2008. Avoiding costly conservation mistakes: the importance of defining actions and costs in spatial priority setting. PLoS One 3(7):e2586.

Casey, S. P., H. J. Hall, H. F. Stanley, and A. C. Vincent. 2004. The origin and evolution of seahorses (genus *Hippocampus*): a phylogenetic study using the cytochrome b gene of mitochondrial DNA. Molecular phylogenetics and evolution 30(2):261-272.

Cheung, W. W., V. W. Lam, J. L. Sarmiento, K. Kearney, R. Watson, and D. Pauly. 2009. Projecting global marine biodiversity impacts under climate change scenarios. Fish and fisheries 10(3):235-251.

Cheung, W. W. L., T. J. Pitcher, and D. Pauly. 2005. A fuzzy logic expert system to estimate intrinsic extinction vulnerabilities of marine fishes to fishing. Biological Conservation 124(1):97-111.

Choi, Y.-U., S. Rho, H.-S. Park, and D.-H. Kang. 2012. Population characteristics of two seahorses, *Hippocampus coronatus* and *Hippocampus mohnikei*, around seagrass beds in the southern coastal waters of Korea. Ichthyological research 59(3):235-241.

Choo, C. and H. Liew. 2003. Spatial distribution, substrate assemblages and size composition of sea horses (Family Syngnathidae) in the coastal waters of Peninsular Malaysia. Journal of the Marine Biological Association of the UK 83(02):271-276.

Cinner, J. E., M. J. Marnane, and T. Mcclanahan. 2005. Conservation and community benefits from traditional coral reef management at Ahus Island, Papua New Guinea. Conservation Biology 19(6):1714-1723.

Cisneros-Montemayor, A. and A. Vincent. 2016. Science, society, and flagship species: social and political history as keys to conservation outcomes in the Gulf of California. Ecology and Society 21(2).

Claassens, L., A. J. Booth, and A. N. Hodgson. 2018. An endangered seahorse selectively chooses an artificial structure. Environmental Biology of Fishes 101(5):723-733.

Clark, D., E. Goodwin, J. Sinner, J. Ellis, and G. Singh. 2016. Validation and limitations of a cumulative impact model for an estuary. Ocean & Coastal Management 120:88-98.

Clynick, B. G. 2008. Harbour swimming nets: a novel habitat for seahorses. Aquatic Conservation-Marine and Freshwater Ecosystems 18(5):483-492.

Cohen, F. P., W. C. Valenti, M. Planas, and R. Calado. 2017. Seahorse aquaculture, biology and conservation: knowledge gaps and research opportunities. Reviews in Fisheries Science & Aquaculture 25(1):100-111.

Coll, M., C. Piroddi, C. Albouy, F. Ben Rais Lasram, W. W. Cheung, V. Christensen, V. S. Karpouzi, F. Guilhaumon, D. Mouillot, and M. Paleczny. 2012. The Mediterranean Sea under siege: spatial overlap between marine biodiversity, cumulative threats and marine reserves. Global Ecology and Biogeography 21(4):465-480.

Cornwell, M. L. and L. M. Campbell. 2012. Co-producing conservation and knowledge: Citizen-based sea turtle monitoring in North Carolina, USA. Social Studies of Science 42(1):101-120.

Costello, M. J. and B. Ballantine. 2015. Biodiversity conservation should focus on no-take Marine Reserves: 94% of Marine Protected Areas allow fishing. Trends in Ecology & Evolution 30(9):507-509.

Cowen, R. K. and S. Sponaugle. 2009. Larval Dispersal and Marine Population Connectivity. Annual Review of Marine Science 1(1):443-466.

Crain, C. M., B. S. Halpern, M. W. Beck, and C. V. Kappel. 2009. Understanding and managing human threats to the coastal marine environment. Annals of the New York Academy of Sciences 1162(1):39-62.

Crain, C. M., K. Kroeker, and B. S. Halpern. 2008. Interactive and cumulative effects of multiple human stressors in marine systems. Ecology letters 11(12):1304-1315.

Crain, R., C. Cooper, and J. L. Dickinson. 2014. Citizen Science: A Tool for Integrating Studies of Human and Natural Systems. Annual Review of Environment and Resources 39(1):641-665.

Crall, A. W., C. S. Jarnevich, N. E. Young, B. J. Panke, M. Renz, and T. J. Stohlgren. 2015. Citizen science contributes to our knowledge of invasive plant species distributions. Biological Invasions 17(8):2415-2427.

Curtis, J. 2006. A case of mistaken identity: skin filaments are unreliable for identifying *Hippocampus guttulatus* and *Hippocampus hippocampus*. Journal of fish biology 69(6):1855-1859.

Curtis, J. and A. Vincent. 2006. Life history of an unusual marine fish: survival, growth and movement patterns of *Hippocampus guttulatus* Cuvier 1829. Journal of Fish Biology 68(3):707-733.

Curtis, J. M. and A. C. Vincent. 2005. Distribution of sympatric seahorse species along a gradient of habitat complexity in a seagrassdominated community. Marine Ecology Progress Series 291:81-91.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. Ecology 88(11):2783-2792.

da Luz Fernandes, M., T. C. Esteves, E. R. Oliveira, and F. L. Alves. 2017. How does the cumulative impacts approach support Maritime Spatial Planning? Ecological Indicators 73:189-202.

Davidson, A. D., A. G. Boyer, H. Kim, S. Pompa-Mansilla, M. J. Hamilton, D. P. Costa, G. Ceballos, and J. H. Brown. 2012. Drivers and hotspots of extinction risk in marine mammals. Proceedings of the National Academy of Sciences 109(9):3395-3400.

De Smith, M. J., M. F. Goodchild, and P. Longley. 2007. Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Troubador Publishing Ltd.

Devictor, V., R. J. Whittaker, and C. Beltrame. 2010. Beyond scarcity: citizen science programmes as useful tools for conservation biogeography. Diversity and Distributions 16(3):354-362.

Dillard, L. O., K. R. Russell, and M. W. Ford. 2008. Macrohabitat models of occurrence for the threatened Cheat Mountain salamander, Plethodon nettingi. Applied Herpetology 5(3):201-224.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, and P. J. Leitão. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36(1):27-46.

Dorrough, J. and J. E. Ash. 1999. Using past and present habitat to predict the current distribution and abundance of a rare cryptic lizard, Delma impar (Pygopodidae). Austral Ecology 24(6):614-624.

Douvere, F. 2008. The importance of marine spatial planning in advancing ecosystem-based sea use management. Marine Policy 32(5):762-771.

Drew, J. A. 2005. Use of traditional ecological knowledge in marine conservation. Conservation Biology 19(4):1286-1293.

Drummond, S. P., K. A. Wilson, E. Meijaard, M. Watts, R. Dennis, L. Christy, and H. P. Possingham. 2010. Influence of a Threatened-Species Focus on Conservation Planning. Conservation Biology 24(2):441-449.

Dudley, N. 2008. Guidelines for Applying Protected Area Management Categories, Gland, Switzerland: IUCN. x + 86pp.

Dulvy, N. K., J. R. Ellis, N. B. Goodwin, A. Grant, J. D. Reynolds, and S. Jennings. 2004. Methods of assessing extinction risk in marine fishes. Fish and Fisheries 5(3):255-276.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29(2):129-151.

Evans, M. C., J. E. M. Watson, R. A. Fuller, O. Venter, S. C. Bennett, P. R. Marsack, and H. P. Possingham. 2011. The Spatial Distribution of Threats to Species in Australia. BioScience 61(4):281-289.

Fairclough, D., K. Clarke, F. Valesini, and I. Potter. 2008. Habitat partitioning by five congeneric and abundant Choerodon species (Labridae) in a large subtropical marine embayment. Estuarine, Coastal and Shelf Science 77(3):446-456.

FAO. 2016. The state of world fisheries and aquaculture. Fisheries and Aquaculture Department, Food and Agriculture Organization of the United Nations, Rome.

Fernandez, M., S. Blum, S. Reichle, Q. Guo, B. Holzman, and H. Hamilton. 2009. Locality uncertainty and the differential performance of four common niche-based modeling techniques. Biodiversity Informatics 6(1):36-52.

Fielding, A. H. and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental conservation 24(1):38-49.

Foster, S. and F. Arreguin-Sánchez. 2014. Using distribution patterns of small fishes to assess small fish by-catch in tropical shrimp trawl fisheries. Animal Conservation 17(3):217-224.

Foster, S., S. Wiswedel, and A. Vincent. 2016. Opportunities and challenges for analysis of wildlife trade using CITES data–seahorses as a case study. Aquatic Conservation: Marine and Freshwater Ecosystems 26(1):154-172.

Foster, S. a. and A. Vincent. 2004. Life history and ecology of seahorses: implications for conservation and management. Journal of Fish Biology 65(1):1-61.

Fox, N. J. and L. E. Beckley. 2005. Priority areas for conservation of Western Australian coastal fishes: A comparison of hotspot, biogeographical and complementarity approaches. Biological Conservation 125(4):399-410.

Franklin, J. 2010. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press. pp. 11-12.

Freudenberger, L., P. Hobson, M. Schluck, S. Kreft, K. Vohland, H. Sommer, S. Reichle, C. Nowicki, W. Barthlott, and P. L. Ibisch. 2013. Nature conservation: priority-setting needs a global change. Biodiversity and conservation 22(5):1255-1281.

Game, E. T. and H. S. Grantham. 2008. Marxan User Manual: For Marxan version 1.8.10. University of Queensland, St. Lucia, Queensland, Australia, and Pacific Marine Analysis and Research Association, Vancouver, British Columbia, Canada. .

Gamer, M., J. Lemon, and I. Fellows. 2007. IRR: Various coefficients of interrater reliability and agreement. R package version 0.70. Zugriff am 19:2009.

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

Gibbons, J. W., D. E. Scott, T. J. Ryan, K. A. Buhlmann, T. D. Tuberville, B. S. Metts, J. L. Greene, T. Mills, Y. Leiden, S. Poppy, and C. T. Winne. 2000. The global decline of reptiles, Deja Vu amphibians. BioScience 50(8):653-666.

Gilchrist, G., M. Mallory, and F. Merkel. 2005. Can Local Ecological Knowledge Contribute to Wildlife Management? Case Studies of Migratory Birds. Ecology and Society 10(1):20.

Gill, D. A., M. B. Mascia, G. N. Ahmadia, L. Glew, S. E. Lester, M. Barnes, I. Craigie, E. S. Darling, C. M. Free, J. Geldmann, S. Holst, O. P. Jensen, A. T. White, X. Basurto, L. Coad,

R. D. Gates, G. Guannel, P. J. Mumby, H. Thomas, S. Whitmee, S. Woodley, and H. E. Fox. 2017. Capacity shortfalls hinder the performance of marine protected areas globally. Nature 543(7647):665-682.

Gladstone, W. 2014. Criticisms of science, social impacts, opinion leaders, and targets for no-take zones led to cuts in New South Wales'(Australia) system of marine protected areas. Aquatic Conservation: Marine and Freshwater Ecosystems 24(3):287-296.

Goldstein, L. J. 2013. Chinese fisheries enforcement: Environmental and strategic implications. Marine Policy 40:187-193.

Graham, C. H., J. Elith, R. J. Hijmans, A. Guisan, A. Townsend Peterson, and B. A. Loiselle. 2008. The influence of spatial errors in species occurrence data used in distribution models. Journal of Applied Ecology 45(1):239-247.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in ecology & evolution 19(9):497-503.

Grand, J. and S. A. Cushman. 2004. A multi-scale analysis of species-environment relationships: breeding birds in a pitch pine-scrub oak (Pinus rigida-Quercus ilicifolia) community (vol 112, pg 307, 2003). Biological Conservation 115(1):173-173.

Green, A., S. E. Smith, G. Lipsett-Moore, C. Groves, N. Peterson, S. Sheppard, P. Lokani, R. Hamilton, J. Almany, and J. Aitsi. 2009. Designing a resilient network of marine protected areas for Kimbe Bay, Papua New Guinea. Oryx 43(4):488-498.

Guerry, A. D., S. Polasky, J. Lubchenco, R. Chaplin-Kramer, G. C. Daily, R. Griffin, M. Ruckelshaus, I. J. Bateman, A. Duraiappah, T. Elmqvist, M. W. Feldman, C. Folke, J. Hoekstra, P. M. Kareiva, B. L. Keeler, S. Li, E. McKenzie, Z. Ouyang, B. Reyers, T. H. Ricketts, J. Rockström, H. Tallis, B. Vira, u. Stockholms, f. Naturvetenskapliga, and C. Stockholm Resilience. 2015. Natural capital and ecosystem services informing decisions: From promise to practice. Proceedings of the National Academy of Sciences 112(24):7348-7355.

Guisan, A., O. Broennimann, R. Engler, M. Vust, N. G. Yoccoz, A. Lehmann, and N. E. Zimmermann. 2006. Using niche-based models to improve the sampling of rare species. Conservation Biology 20(2):501-511.

Guisan, A. and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. Ecology letters 8(9):993-1009.

Guisan, A., R. Tingley, J. B. Baumgartner, I. Naujokaitis-Lewis, P. R. Sutcliffe, A. I. Tulloch, T. J. Regan, L. Brotons, E. McDonald-Madden, and C. Mantyka-Pringle. 2013. Predicting species distributions for conservation decisions. Ecology letters 16(12):1424-1435.

Gullison, R., R. Rice, and A. Blundell. 2000. 'Marketing'species conservation. Nature 404(6781):923.

Gundlach, E. R. and M. O. Hayes. 1978. Vulnerability of coastal environments to oil spill impacts. Marine technology society Journal 12(4):18-27.

Halpern, B. S., M. Frazier, J. Potapenko, K. S. Casey, K. Koenig, C. Longo, J. S. Lowndes, R. C. Rockwood, E. R. Selig, and K. A. Selkoe. 2015. Spatial and temporal changes in cumulative human impacts on the world's ocean. Nature communications 6:7615.

Halpern, B. S. and R. Fujita. 2013. Assumptions, challenges, and future directions in cumulative impact analysis. Ecosphere 4(10):1-11.

Halpern, B. S., C. Longo, D. Hardy, K. L. McLeod, J. F. Samhouri, S. K. Katona, K. Kleisner, S. E. Lester, J. O'Leary, and M. Ranelletti. 2012. An index to assess the health and benefits of the global ocean. Nature 488(7413):615.

Halpern, B. S., K. A. Selkoe, F. Micheli, and C. V. Kappel. 2007. Evaluating and ranking the vulnerability of global marine ecosystems to anthropogenic threats. Conservation Biology 21(5):1301-1315.

Halpern, B. S., S. Walbridge, K. A. Selkoe, C. V. Kappel, F. Micheli, C. D'agrosa, J. F. Bruno, K. S. Casey, C. Ebert, and H. E. Fox. 2008. A global map of human impact on marine ecosystems. Science 319(5865):948-952.

Hamel, M. A., R. L. Pressey, L. S. Evans, and S. Andrefouet. 2018. The Importance of Fishing Grounds as Perceived by Local Communities Can be Undervalued by Measures of Socioeconomic Cost Used in Conservation Planning. Conservation Letters 11(1):1-9.

Han, S.-Y., J.-K. Kim, Y. Kai, and H. Senou. 2017. Seahorses of the *Hippocampus coronatus* complex: taxonomic revision, and description of *Hippocampus haema*, a new species from Korea and Japan (Teleostei, Syngnathidae). ZooKeys(712):113.

Hanley, J. A. and B. J. McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29-36.

Hanski, I. 2011. Habitat loss, the dynamics of biodiversity, and a perspective on conservation. AMBIO: A Journal of the Human Environment 40(3):248-255.

Harasti, D., K. Martin-Smith, and W. Gladstone. 2014. Does a no-take marine protected area benefit seahorses? PLoS One 9(8):e105462.

Harasti, D. and R. A. Pollom. 2017. *Hippocampus whitei*. The IUCN Red List of Threatened Species 2017: e.T10088A46721312. http://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T10088A46721312.en. Downloaded on 22 May 2018.

Harris, G. M., C. N. Jenkins, and S. L. Pimm. 2005. Refining biodiversity conservation priorities. Conservation Biology 19(6):1957-1968.

Hazen, E. L., S. Jorgensen, R. R. Rykaczewski, S. J. Bograd, D. G. Foley, I. D. Jonsen, S. A. Shaffer, J. P. Dunne, D. P. Costa, and L. B. Crowder. 2013. Predicted habitat shifts of Pacific top predators in a changing climate. Nature Climate Change 3(3):234.

He, Q., M. D. Bertness, J. F. Bruno, B. Li, G. Chen, T. C. Coverdale, A. H. Altieri, J. Bai, T. Sun, and S. C. Pennings. 2014. Economic development and coastal ecosystem change in China. Scientific Reports 4:5995.

Hellyer, C. B., D. Harasti, and A. G. B. Poore. 2011. Manipulating artificial habitats to benefit seahorses in Sydney Harbour, Australia. Aquatic Conservation-Marine and Freshwater Ecosystems 21(6):582-589.

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

Higdon, J. W., K. H. Westdal, and S. H. Ferguson. 2014. Distribution and abundance of killer whales (Orcinus orca) in Nunavut, Canada-an Inuit knowledge survey. Journal of the Marine Biological Association of the United Kingdom 94(6):1293-1304.

Hilborn, R. 2016. Marine biodiversity needs more than protection. Nature 535(7611):224.

Hooker, S. K., H. Whitehead, and S. Gowans. 2002. Ecosystem consideration in conservation planning: energy demand of foraging bottlenose whales (Hyperoodon ampullatus) in a marine protected area. Biological Conservation 104(1):51-58.

Hooper, D. U., E. C. Adair, B. J. Cardinale, J. E. Byrnes, B. A. Hungate, K. L. Matulich, A. Gonzalez, J. E. Duffy, L. Gamfeldt, and M. I. O’Connor. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. Nature 486(7401):105-108.

Huntington, H. P. 2000. Using traditional ecological knowledge in science: methods and applications. Ecological Applications 10(5):1270-1274.

Irvine, R., S. Fiorini, S. Yearley, J. McLeod, A. Turner, H. Armstrong, P. White, and R. Van Der Wal. 2009. Can managers inform models? Integrating local knowledge into models of red deer habitat use. Journal of Applied Ecology 46(2):344-352.

IUCN. 2018. The IUCN Red List of Threatened Species. Version 2018-2. <http://www.iucnredlist.org>. Downloaded on 14 November 2018.

IUCN Species Survial Committee. 2005. Guidelines for using the IUCN Red List categories and criteria. IUCN, Gland, Switzerland.

Jackson, C. R. and M. P. Robertson. 2011. Predicting the potential distribution of an endangered cryptic subterranean mammal from few occurrence records. Journal for Nature Conservation 19(2):87-94.

Jackson, M. M., S. E. Gergel, and K. Martin. 2015. Citizen science and field survey observations provide comparable results for mapping Vancouver Island White-tailed Ptarmigan (Lagopus leucura saxatilis) distributions. Biological Conservation 181:162-172.

James, A. N., K. J. Gaston, and A. Balmford. 1999. Balancing the Earth's accounts. Nature 401(6751):323.

Jaric, I., F. Courchamp, J. Gessner, and D. L. Roberts. 2016. Potentially threatened: a Data Deficient flag for conservation management. Biodiversity and conservation 25(10):1995-2000.

Jenkins, C. N. and K. S. Van Houtan. 2016. Global and regional priorities for marine biodiversity protection. Biological Conservation 204(2016):333-339.

Jetz, W. and R. P. Freckleton. 2015. Towards a general framework for predicting threat status of data-deficient species from phylogenetic, spatial and environmental information. Philosophical Transactions of the Royal Society B-Biological Sciences 370(1662):20140016.

Johannes, R. E. 1989. Traditional ecological knowledge: a collection of essays. 2880329981;9782880329983, IUCN, World Conservation Union, Gland, Switzerland.

Jordan, A., S. Fairfull, and B. Creese. 2016. Managing threats to the marine estate in New South Wales (Australia) to maximise community wellbeing. Journal of Coastal Research 75(sp1):642-646.

Joseph, L. N., R. F. Maloney, and H. P. Possingham. 2009. Optimal allocation of resources among threatened species: a project prioritization protocol. Conservation Biology 23(2):328-338.

Kenchington, R. 2010. Strategic roles of marine protected areas in ecosystem scale conservation. Bulletin of Marine Science 86(2):303-313.

Kinlan, B. P. and S. D. Gaines. 2003. Propagule Dispersal in Marine and Terrestrial Environments: A Community Perspective. Ecology 84(8):2007-2020.

Kleiber, D., L. Blight, I. Caldwell, and A. Vincent. 2011. The importance of seahorses and pipefishes in the diet of marine animals. Reviews in fish biology and fisheries 21(2):205-223.

Klein, C., A. Chan, L. Kircher, A. Cundiff, N. Gardner, Y. Hrovat, A. Scholz, B. Kendall, and S. Airame. 2008. Striking a balance between biodiversity conservation and socioeconomic viability in the design of marine protected areas. Conservation Biology 22(3):691-700.

Klein, C. J., V. J. Tulloch, B. S. Halpern, K. A. Selkoe, M. E. Watts, C. Steinback, A. Scholz, and H. P. Possingham. 2013. Tradeoffs in marine reserve design: habitat condition, representation, and socioeconomic costs. Conservation Letters 6(5):324-332.

Koldewey, H. J. and K. M. Martin-Smith. 2010. A global review of seahorse aquaculture. Aquaculture 302(3):131-152.

Komyakova, V., P. L. Munday, and G. P. Jones. 2013. Relative importance of coral cover, habitat complexity and diversity in determining the structure of reef fish communities. PLoS One 8(12):e83178.

Korpinen, S. and J. H. Andersen. 2016. A global review of cumulative pressure and impact assessments in marine environments. Frontiers in Marine Science 3:153.

Kursa, M. B. and W. R. Rudnicki. 2010. Feature selection with the Boruta package. J Stat Softw 36(11):1-13.

Lawson, C. R., J. A. Hodgson, R. J. Wilson, and S. A. Richards. 2014. Prevalence, thresholds and the performance of presence–absence models. Methods in Ecology and Evolution 5(1):54-64.

Lawson, J., S. Foster, A. Lim, V. Chong, and A. Vincent. 2015. Novel life-history data for threatened seahorses provide insight into fishery effects. Journal of fish biology 86(1):1-15.

Lawson, J. M., S. J. Foster, and A. C. Vincent. 2017. Low bycatch rates add up to big numbers for a genus of small fishes. Fisheries 42(1):19-33.

Laze, K. and A. Gordon. 2016. Incorporating natural and human factors in habitat modelling and spatial prioritisation for the Lynx lynx martinoi. Web Ecology 16(1):17.

Leslie, H., M. Ruckelshaus, I. R. Ball, S. Andelman, and H. P. Possingham. 2003. Using siting algorithms in the design of marine reserve networks. Ecological Applications 13(1):S185-S198.

Levin, N., M. Coll, S. Fraschetti, G. Gal, S. Giakoumi, C. Göke, J. J. Heymans, S. Katsanevakis, T. Mazor, and B. Öztürk. 2014. Biodiversity data requirements for systematic conservation planning in the Mediterranean Sea. Marine Ecology Progress Series 508:261-281.

Lewison, R. L., L. B. Crowder, A. J. Read, and S. A. Freeman. 2004. Understanding impacts of fisheries bycatch on marine megafauna. Trends in Ecology & Evolution 19(11):598-604.

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. R news 2(3):18-22.

Lindegren, M., B. G. Holt, B. R. MacKenzie, and C. Rahbek. 2018. A global mismatch in the protection of multiple marine biodiversity components and ecosystem services. Scientific reports 8(1):4099.

Liu, C., P. M. Berry, T. P. Dawson, and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28(3):385-393.

Liu, J. 2013. Status of marine biodiversity of the China Seas. PLoS One 8(1):e50719.

Liu, M., M. Lin, S. T. Turvey, and S. Li. 2017. Fishers' knowledge as an information source to investigate bycatch of marine mammals in the South China Sea. Animal Conservation 20(2):182-192.

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

Lombarte, A., L. Recasens, M. González, and L. G. de Sola. 2000. Spatial segregation of two species of Mullidae (Mullus surmuletus and M. barbatus) in relation to habitat. Marine Ecology Progress Series 206:239-249.

Lourie, S., D. Green, and A. Vincent. 2005. Dispersal, habitat differences, and comparative phylogeography of Southeast Asian seahorses (Syngnathidae: *Hippocampus*). Molecular ecology 14(4):1073-1094.

Lourie, S. A., S. J. Foster, E. W. Cooper, and A. C. Vincent. 2004. A guide to the identification of seahorses. Project Seahorse and TRAFFIC North America Project Seahorse and TRAFFIC North America, University of British Columbia and World Wildlife Fund, Washington, DC.

Lourie, S. A., R. A. Pollom, and S. J. Foster. 2016. A global revision of the Seahorses *Hippocampus* Rafinesque 1810 (Actinopterygii: Syngnathiformes): Taxonomy and biogeography with recommendations for further research. Zootaxa 4146(1):1-66.

Lourie, S. A. and J. E. Randall. 2003. A new pygmy seahorse, *Hippocampus denise* (Teleostei : Syngnathidae), from the Indo-Pacific. Zoological Studies 42(2):284-291.

Lourie, S. A. and A. C. J. Vincent. 2004. Using biogeography to help set priorities in marine conservation. Conservation Biology 18(4):1004-1020.

Lu, W.-H., J. Liu, X.-Q. Xiang, W.-L. Song, and A. McIlgorm. 2015. A comparison of marine spatial planning approaches in China: Marine functional zoning and the marine ecological red line. Marine Policy 62:94-101.

Lucifora, L. O., V. B. Garcia, and B. Worm. 2011. Global Diversity Hotspots and Conservation Priorities for Sharks. PLoS One 6(5): e19356.

Luiz, O. J., A. P. Allen, D. R. Robertson, S. R. Floeter, and J. S. Madin. 2015. Seafarers or castaways: ecological traits associated with rafting dispersal in tropical reef fishes. Journal of biogeography 42(12):2323-2333.

Luiz, O. J., R. M. Woods, E. M. P. Madin, and J. S. Madin. 2016. Predicting IUCN Extinction Risk Categories for the World's Data Deficient Groupers (Teleostei: Epinephelidae). Conservation Letters 9(5):342-350.

Ma, C., X. Zhang, W. Chen, G. Zhang, H. Duan, M. Ju, H. Li, and Z. Yang. 2013. China's special marine protected area policy: Trade-off between economic development and marine conservation. Ocean & Coastal Management 76:1-11.

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

Mace, G. M., N. J. Collar, K. J. Gaston, C. Hilton-Taylor, H. R. Akcakaya, N. Leader-Williams, E. J. Milner-Gulland, and S. N. Stuart. 2008. Quantification of Extinction Risk: IUCN's System for Classifying Threatened Species. Conservation Biology 22(6):1424-1442.

Malavasi, S., A. Franco, F. Riccato, C. Valerio, P. Torricelli, and P. Franzoi. 2007. Habitat selection and spatial segregation in three pipefish species. Estuarine, Coastal and Shelf Science 75(1):143-150.

Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. Journal of Applied Ecology 38(5):921-931.

Manning, C. G. 2017. How is a family of sedentary marine fishes shaped by its habitats, prey, and predators? University of British Columbia.

Marcer, A., L. Sáez, R. Molowny-Horas, X. Pons, and J. Pino. 2013. Using species distribution modelling to disentangle realised versus potential distributions for rare species conservation. Biological Conservation 166:221-230.

Margules, C. R. and R. L. Pressey. 2000. Systematic conservation planning. Nature 405(6783):243.

Martin, T. G., M. A. Burgman, F. Fidler, P. M. Kuhnert, S. Low-Choy, M. McBride, and K. Mengersen. 2012. Eliciting Expert Knowledge in Conservation Science. Conservation Biology 26(1):29-38.

Martinez-Levasseur, L. M., C. M. Furgal, M. O. Hammill, and G. Burness. 2017. Challenges and strategies when mapping local ecological knowledge in the Canadian Arctic: the importance of defining the geographic limits of participants' common areas of observations. Polar Biology 40(8):1501-1513.

Mateo, R. G., T. B. Croat, Á. M. Felicísimo, and J. Munoz. 2010. Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. Diversity and Distributions 16(1):84-94.

Matthews, K. R. 1990. A telemetric study of the home ranges and homing routes of copper and quillback rockfishes on shallow rocky reefs. Canadian Journal of Zoology 68(11):2243-2250.

Maxwell, S. M., E. L. Hazen, S. J. Bograd, B. S. Halpern, G. A. Breed, B. Nickel, N. M. Teutschel, L. B. Crowder, S. Benson, and P. H. Dutton. 2013. Cumulative human impacts on marine predators. Nature communications 4:2688.

Mazor, T., S. Giakoumi, S. Kark, and H. P. Possingham. 2014. Large-scale conservation planning in a multinational marine environment: cost matters. Ecological Applications 24(5):1115-1130.

McCauley, D. J., M. L. Pinsky, S. R. Palumbi, J. A. Estes, F. H. Joyce, and R. R. Warner. 2015. Marine defaunation: animal loss in the global ocean. Science 347(6219):1255641.

McClenachan, L., F. Ferretti, and J. K. Baum. 2012. From archives to conservation: why historical data are needed to set baselines for marine animals and ecosystems. Conservation Letters 5(5):349-359.

McCook, L. J., T. Ayling, M. Cappo, J. H. Choat, R. D. Evans, D. M. De Freitas, M. Heupel, T. P. Hughes, G. P. Jones, and B. Mapstone. 2010. Adaptive management of the Great Barrier Reef: a globally significant demonstration of the benefits of networks of marine reserves. Proceedings of the National Academy of Sciences 107(43):18278-18285.

McIntosh, E. J., R. L. Pressey, S. Lloyd, R. J. Smith, and R. Grenyer. 2017. The Impact of Systematic Conservation Planning. Annual Review of Environment and Resources 42:677-697.

McShane, T. O., P. D. Hirsch, T. C. Trung, A. N. Songorwa, A. Kinzig, B. Monteferri, D. Mutekanga, H. Van Thang, J. L. Dammert, and M. Pulgar-Vidal. 2011. Hard choices: making trade-offs between biodiversity conservation and human well-being. Biological Conservation 144(3):966-972.

MEP. 2002. National List of Wildlife with Protection Priority. http://sts.mep.gov.cn/swdyx_1/swwzzybh/200211/t20021118_83384.htmHigher Education Press. pp. 746.

Micheli, F., N. Levin, S. Giakoumi, S. Katsanevakis, A. Abdulla, M. Coll, S. Fraschetti, S. Kark, D. Koutsoubas, and P. Mackelworth. 2013. Setting priorities for regional conservation planning in the Mediterranean Sea. PLoS One 8(4):e59038.

Mills, M., R. L. Pressey, R. Weeks, S. Foale, and N. C. Ban. 2010. A mismatch of scales: challenges in planning for implementation of marine protected areas in the Coral Triangle. Conservation Letters 3(5):291-303.

Mittermeier, R. A., W. R. Turner, F. W. Larsen, T. M. Brooks, and C. Gascon. 2011. Global biodiversity conservation: the critical role of hotspots. Pages 3-22 Biodiversity hotspots. Springer.

Moore, A. B. M. 2010. The Smalleye Stingray Dasyatis microps (Myliobatiformes: Dasyatidae) in the Gulf: previously unreported presence of a large, rare elasmobranch. ZOOLOGY IN THE MIDDLE EAST 49(Journal Article):101-103.

Moore, J. E., T. M. Cox, R. L. Lewison, A. J. Read, R. Bjorkland, S. L. McDonald, L. B. Crowder, E. Aruna, I. Ayissi, P. Espeut, C. Joynson-Hicks, N. Pilcher, C. N. S. Poonian, B. Solarin, and J. Kiszka. 2010. An interview-based approach to assess marine mammal and sea turtle captures in artisanal fisheries. Biological Conservation 143(3):795-805.

Murray, J. V., A. W. Goldizen, R. A. O’Leary, C. A. McAlpine, H. P. Possingham, and S. L. Choy. 2009. How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? A case study using brush-tailed rock-wallabies Petrogale penicillata. Journal of Applied Ecology 46(4):842-851.

Murray, S. N., R. F. Ambrose, J. A. Bohnsack, L. W. Botsford, M. H. Carr, G. E. Davis, P. K. Dayton, D. Gotshall, D. R. Gunderson, and M. A. Hixon. 1999. No-take reserve networks: sustaining fishery populations and marine ecosystems. Fisheries 24(11):11-25.

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

Naidoo, R., A. Balmford, P. J. Ferraro, S. Polasky, T. H. Ricketts, and M. Rouget. 2006. Integrating economic costs into conservation planning. Trends in Ecology & Evolution 21(12):681-687.

Newbold, T. 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. Progress in Physical Geography 34(1):3-22.

Niamir, A., A. K. Skidmore, A. G. Toxopeus, A. R. Muñoz, and R. Real. 2011. Finessing atlas data for species distribution models. Diversity and Distributions 17(6):1173-1185.

Norris, K. 2012. Biodiversity in the context of ecosystem services: the applied need for systems approaches. Philosophical Transactions: Biological Sciences 367(1586):191-199.

O’Higgins, T. G., S. P. Ferraro, D. D. Dantin, S. J. Jordan, and M. M. Chintala. 2010. Habitat Scale Mapping of Fisheries Ecosystem Service Values in Estuaries. Ecology and Society 15(4):7.

Olsoy, P. J., K. A. Zeller, J. A. Hicke, H. B. Quigley, A. R. Rabinowitz, and D. H. Thornton. 2016. Quantifying the effects of deforestation and fragmentation on a range-wide conservation plan for jaguars. Biological Conservation 203:8-16.

Osborne, P. E. and P. J. Leitao. 2009. Effects of species and habitat positional errors on the performance and interpretation of species distribution models. Diversity and Distributions 15(4):671-681.

Pan, Y., G. Wei, A. A. Cunningham, S. Li, S. Chen, E. J. Milner-Gulland, and S. T. Turvey. 2016. Using local ecological knowledge to assess the status of the Critically Endangered Chinese giant salamander Andrias davidianus in Guizhou Province, China. Oryx 50(2):257-264.

Parry, L. and C. A. Peres. 2015. Evaluating the use of local ecological knowledge to monitor hunted tropical-forest wildlife over large spatial scales. Ecology and Society 20(3):15.

Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A. Townsend Peterson. 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.

Peng, Z., L. Wenhu, S. Jun, S. Pearson, and Y. Hongsheng. 2016. Natural coast protection and use in China: implications of resource protection “Redline” policies. Coastal Management 44(1):21-35.

Petitpierre, B., O. Broennimann, C. Kueffer, C. Daehler, A. Guisan, and A. Jiménez-Valverde. 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, N. D., N. Reid, T. Thys, C. Harrod, N. L. Payne, C. A. Morgan, H. J. White, S. Porter, and J. D. R. Houghton. 2017a. Applying species distribution modelling to a data poor, pelagic fish complex: the ocean sunfishes. Journal of biogeography 44(10):2176-2187.

Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, and M. E. Blair. 2017b. Opening the black box: an open-source release of Maxent. Ecography 40(7):887-893.

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

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

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

Phillips, S. J., M. Dudík, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Page 83 *in* Proceedings of the twenty-first international conference on Machine learning. ACM.

Pimm, S. L., C. N. Jenkins, R. Abell, T. M. Brooks, J. L. Gittleman, L. N. Joppa, P. H. Raven, C. M. Roberts, and J. O. Sexton. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. Science 344(6187):1246752.

Pimm, S. L., G. J. Russell, J. L. Gittleman, and T. M. Brooks. 1995. The future of biodiversity. Science 269(5222):347-350.

Pınarbaşı, K., I. Galparsoro, Á. Borja, V. Stelzenmüller, C. N. Ehler, and A. Gimpel. 2017. Decision support tools in marine spatial planning: Present applications, gaps and future perspectives. Marine Policy 83:83-91.

Pollom, R. A. 2017. *Hippocampus capensis*. The IUCN Red List of Threatened Species 2017: e.T10056A54903534. http://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T10056A54903534.en. Downloaded on 22 May 2018.

Ponder, W. F., G. Carter, P. Flemons, and R. Chapman. 2001. Evaluation of museum collection data for use in biodiversity assessment. Conservation Biology 15(3):648-657.

Possingham, H., I. Ball, and S. Andelman. 2000. Mathematical methods for identifying representative reserve networks. Pages 291-306 *in* Quantitative methods for conservation biology.

Prasad, A. M., L. R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems 9(2):181-199.

Qiu, W., B. Wang, P. J. Jones, and J. C. Axmacher. 2009. Challenges in developing China's marine protected area system. Marine Policy 33(4):599-605.

Quade, D. 1979. Using weighted rankings in the analysis of complete blocks with additive block effects. Journal of the American Statistical Association 74(367):680-683.

Rahbek, C. 2005. The role of spatial scale and the perception of large-scale species-richness patterns. Ecology letters 8(2):224-239.

Rainho, A. and J. M. Palmeirim. 2011. The importance of distance to resources in the spatial modelling of bat foraging habitat. PLoS One 6(4):e19227.

Reis, E., G. M. López-Iborra, and R. T. Pinheiro. 2012. Changes in bird species richness through different levels of urbanization: Implications for biodiversity conservation and garden design in Central Brazil. Landscape and Urban Planning 107(1):31-42.

Reynolds, J. D., N. K. Dulvy, N. B. Goodwin, and J. A. Hutchings. 2005. Biology of extinction risk in marine fishes. Proceedings of the Royal Society of London B: Biological Sciences 272(1579):2337-2344.

Rice, J. C. 2005. Understanding fish habitat ecology to achieve conservation. Journal of Fish Biology 67(Journal Article):1-22.

Roberge, J.-M. and P. Angelstam. 2004. Usefulness of the umbrella species concept as a conservation tool. Conservation Biology 18(1):76-85.

Robinson, L., J. Elith, A. Hobday, R. Pearson, B. Kendall, H. Possingham, and A. Richardson. 2011. Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. Global Ecology and Biogeography 20(6):789-802.

Rondinini, C., K. A. Wilson, L. Boitani, H. Grantham, and H. P. Possingham. 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. Ecology letters 9(10):1136-1145.

Roos, G., S. Van Wassenbergh, P. Aerts, A. Herrel, and D. Adriaens. 2011. Effects of snout dimensions on the hydrodynamics of suction feeding in juvenile and adult seahorses. Journal of theoretical biology 269(1):307-317.

Sala, E., O. Aburto-Oropeza, G. Paredes, I. Parra, J. C. Barrera, and P. K. Dayton. 2002. A General Model for Designing Networks of Marine Reserves. Science 298(5600):1991-1993.

Sanciangco, J. C., K. E. Carpenter, P. J. Etnoyer, and F. Moretzsohn. 2013. Habitat availability and heterogeneity and the Indo-Pacific warm pool as predictors of marine species richness in the tropical Indo-Pacific. PLoS One 8(2):e56245.

Scholz, A., K. Bonzon, R. Fujita, N. Benjamin, N. Woodling, P. Black, and C. Steinback. 2004. Participatory socioeconomic analysis: drawing on fishermen's knowledge for marine protected area planning in California. Marine Policy 28(4):335-349.

Searcy, C. A. and H. B. Shaffer. 2016. Do ecological niche models accurately identify climatic determinants of species ranges? The American Naturalist 187(4):423-435.

Selig, E. R., W. R. Turner, S. Troeng, B. P. Wallace, B. S. Halpern, K. Kaschner, B. G. Lascelles, K. E. Carpenter, and R. A. Mittermeier. 2014. Global Priorities for Marine Biodiversity Conservation. PLoS One 9(1): e82898.

Shelton, A. O., J. T. Thorson, E. J. Ward, and B. E. Feist. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Canadian journal of fisheries and aquatic sciences 71(11):1655-1666.

Shen, G. and M. Heino. 2014. An overview of marine fisheries management in China. Marine Policy 44:265-272.

Silvertown, J. 2009. A new dawn for citizen science. Trends in Ecology & Evolution 24(9):467-471.

Smith, R. J., P. S. Goodman, and W. S. Matthews. 2006. Systematic conservation planning: a review of perceived limitations and an illustration of the benefits, using a case study from Maputaland, South Africa. Oryx 40(4):400-410.

SOA. 2010. Bulletin of China Marine Economy 2010 (2011). SOA (State Oceanic Administration China). http://www.soa.gov.cn/zwgk/hygb/.

SOA. 2012. Bulletin of Marine Environmental Quality of China in 2011. China Ocean Press, Beijing, China (in Chinese).

Solovyev, B., V. Spiridonov, I. Onufrenya, S. Belikov, N. Chernova, D. Dobrynin, M. Gavrilo, D. Glazov, Y. Krasnov, and S. Mukharamova. 2017. Identifying a network of priority areas for conservation in the Arctic seas: Practical lessons from Russia. Aquatic Conservation: Marine and Freshwater Ecosystems 27(S1):30-51.

Soutullo, A., M. De Castro, and V. Urios. 2008. Linking political and scientifically derived targets for global biodiversity conservation: implications for the expansion of the global network of protected areas. Diversity and Distributions 14(4):604-613.

Stacey, N., J. Karam, M. Meekan, S. Pickering, and J. Ninef. 2012. Prospects for whale shark conservation in Eastern Indonesia through bajo traditional ecological knowledge and community-based monitoring. Conservation and Society 10(1):63-75.

Stelzenmüller, V., J. Lee, A. South, J. Foden, and S. I. Rogers. 2013. Practical tools to support marine spatial planning: a review and some prototype tools. Marine Policy 38:214-227.

Stewart, R. R. and H. P. Possingham. 2005. Efficiency, costs and trade-offs in marine reserve system design. Environmental Modeling & Assessment 10(3):203-213.

Stirling, D. A., P. Boulcott, B. E. Scott, and P. J. Wright. 2016. Using verified species distribution models to inform the conservation of a rare marine species. Diversity and Distributions 22(7):808-822.

Stuber, E. F., L. F. Gruber, and J. J. Fontaine. 2017. A Bayesian method for assessing multi-scale species-habitat relationships. Landscape Ecology 32(12):2365-2381.

Suarez, A. V. and N. D. Tsutsui. 2004. The value of museum collections for research and society. BioScience 54(1):66-74.

Subcommittee, I. S. a. P. 2017. Guidelines for Application of IUCN Red List Criteria at Regional and National Levels. Version 4.0.

Tallis, H., Z. Ferdana, and E. Gray. 2008. Linking terrestrial and marine conservation planning and threats analysis. Conservation Biology 22(1):120-130.

Tang, W. C. 1987. Chinese medicinal materials from the sea. Abstracts of Chinese Medicines 1(4):571-600.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Teck, S. J., B. S. Halpern, C. V. Kappel, F. Micheli, K. A. Selkoe, C. M. Crain, R. Martone, C. Shearer, J. Arvai, and B. Fischhoff. 2010. Using expert judgment to estimate marine ecosystem vulnerability in the California Current. Ecological Applications 20(5):1402-1416.

Teske, P. R., M. I. Cherry, and C. A. Matthee. 2004. The evolutionary history of seahorses (Syngnathidae: *Hippocampus*): molecular data suggest a West Pacific origin and two invasions of the Atlantic Ocean. Molecular phylogenetics and evolution 30(2):273-286.

Theobald, D. M., D. L. Stevens Jr, D. White, N. S. Urquhart, A. R. Olsen, and J. B. Norman. 2007. Using GIS to generate spatially balanced random survey designs for natural resource applications. Environmental Management 40(1):134-146.

Thiel, M., M. A. Penna-Diaz, G. Luna-Jorquera, S. Salas, J. Sellanes, and W. Stotz. 2014. CITIZEN SCIENTISTS AND MARINE RESEARCH: VOLUNTEER PARTICIPANTS, THEIR CONTRIBUTIONS, AND PROJECTION FOR THE FUTURE. Pages 257-314 *in* R. N. Hughes, D. J. Hughes, and I. P. Smith, editors. Oceanography and Marine Biology: An Annual Review, Vol 52.

Thornton, T. F. and A. M. Scheer. 2012. Collaborative engagement of local and traditional knowledge and science in marine environments: a review. Ecology and Society 17(3):8.

Tittensor, D. P., C. Mora, W. Jetz, H. K. Lotze, D. Ricard, E. V. Berghe, and B. Worm. 2010. Global patterns and predictors of marine biodiversity across taxa. Nature 466(7310):1098-1101.

Turpie, J. K., L. E. Beckley, and S. M. Katua. 2000. Biogeography and the selection of priority areas for conservation of South African coastal fishes. Biological Conservation 92(1):59-72.

Turvey, S. T., C. L. Risley, J. E. Moore, L. A. Barrett, H. Yujiang, Z. Xiujiang, Z. Kaiya, and W. Ding. 2013. Can local ecological knowledge be used to assess status and extinction drivers in a threatened freshwater cetacean? Biological Conservation 157:352-360.

Tweh, C. G., M. M. Lormie, C. Y. Kouakou, A. Hillers, H. S. Kuhl, and J. Junker. 2015. Conservation status of chimpanzees Pan troglodytes verus and other large mammals in Liberia: a nationwide survey. Oryx 49(4):710-718.

Tyberghein, L., H. Verbruggen, K. Pauly, C. Troupin, F. Mineur, and O. De Clerck. 2012. Bio-ORACLE: a global environmental dataset for marine species distribution modelling. Global Ecology and Biogeography 21(2):272-281.

UNEP-WCMC and IUCN. 2018. Protected Planet: [insert name of component database; The World Database on Protected Areas (WDPA)/The Global Database on Protected Areas Management Effectiveness (GD-PAME)] [On-line], [insert month/year of the version downloaded], Cambridge, UK: UNEP-WCMC and IUCN. Available at: www.protectedplanet.net.

Velasco, D., M. García-Llorente, B. Alonso, A. Dolera, I. Palomo, I. Iniesta-Arandia, and B. Martín-López. 2015. Biodiversity conservation research challenges in the 21st century: A review of publishing trends in 2000 and 2011. Environmental Science & Policy 54:90-96.

Vincent, A. C., S. Foster, and H. Koldewey. 2011. Conservation and management of seahorses and other Syngnathidae. Journal of Fish Biology 78(6):1681-1724.

Vincent, A. C., Y. J. Sadovy de Mitcheson, S. L. Fowler, and S. Lieberman. 2014. The role of CITES in the conservation of marine fishes subject to international trade. Fish and Fisheries 15(4):563-592.

Waldron, A., A. O. Mooers, D. C. Miller, N. Nibbelink, D. Redding, T. S. Kuhn, J. T. Roberts, and J. L. Gittleman. 2013. Targeting global conservation funding to limit immediate biodiversity declines. Proceedings of the National Academy of Sciences 110(29):12144-12148.

Wallace, B. P., A. D. DiMatteo, A. B. Bolten, M. Y. Chaloupka, B. J. Hutchinson, F. A. Abreu-Grobois, J. A. Mortimer, J. A. Seminoff, D. Amorocho, K. A. Bjorndal, J. Bourjea, B. W. Bowen, R. B. Duenas, P. Casale, B. C. Choudhury, A. Costa, P. H. Dutton, A. Fallabrino, E. M. Finkbeiner, A. Girard, M. Girondot, M. Hamann, B. J. Hurley, M. Lopez-Mendilaharsu, M. A. Marcovaldi, J. A. Musick, R. Nel, N. J. Pilcher, S. Troeng, B. Witherington, and R. B. Mast. 2011. Global Conservation Priorities for Marine Turtles. PLoS One 6(9).

Wang, C. C. and D. Nie. 1932. A survey of the marine protozoa of Amoy. The Science Society of China, Nanking, China.

Wang, S. and Y. Xie. 2009. China species red list. In: Vertebrates. Vol. 2 Higher Education.

Warren, D. L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62(11):2868-2883.

Watts, M. E., I. R. Ball, R. S. Stewart, C. J. Klein, K. Wilson, C. Steinback, R. Lourival, L. Kircher, and H. P. Possingham. 2009. Marxan with Zones: Software for optimal conservation based land- and sea-use zoning. Environmental Modelling & Software 24(12):1513-1521.

Weeks, R., G. R. Russ, A. A. Bucol, and A. C. Alcala. 2010. Shortcuts for marine conservation planning: The effectiveness of socioeconomic data surrogates. Biological Conservation 143(5):1236-1244.

Wentz, F. J. and M. Schabel. 2000. Precise climate monitoring using complementary satellite data sets. Nature 403(6768):414-416.

Whittaker, R. H., S. A. Levin, and R. B. Root. 1973. Niche, habitat, and ecotope. The American Naturalist 107(955):321-338.

Willig, M. R., D. Kaufman, and R. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. Annual Review of Ecology, Evolution, and Systematics 34(1):273-309.

Wilson, K. A., M. Cabeza, and C. J. Klein. 2009. Fundamental concepts of spatial conservation prioritization. Spatial Conservation Prioritization: quantitative methods and computational tools. Oxford, UK.

Xiang, J. 2003. Development History of Chinese Marine Biology. Pages 7 - 8 *in* J. Xiang, editor. Marine Biology. Science Press, Beijing, China.

Yang, D. and C. Yang. 2009. Detection of seagrass distribution changes from 1991 to 2006 in Xincun Bay, Hainan, with satellite remote sensing. Sensors 9(2):830-844.

Zarnetske, P. L., T. C. Edwards, and G. G. Moisen. 2007. Habitat classification modeling with incomplete data: pushing the habitat envelope. Ecological Applications 17(6):1714-1726.

Zhang, L., Z. Luo, D. Mallon, C. Li, and Z. Jiang. 2017. Biodiversity conservation status in China's growing protected areas. Biological Conservation 210:89-100.

Zhang, Y.-H., G. Qin, X. Wang, and Q. Lin. 2016. A new species of seahorse (Teleostei: Syngnathidae) from the South China Sea. Zootaxa 4170(2):384-392.

Zhao, M., K. Yu, Q. Zhang, Q. Shi, and G. J. Price. 2012. Long-term decline of a fringing coral reef in the northern South China Sea. Journal of Coastal Research 28(5):1088-1099.

Zhu, Y.-P. and H. J. Woerdenbag. 1995. Traditional Chinese herbal medicine. Pharmacy World and Science 17(4):103-112.

Appendix A Supplementary methods for Chapter 2

This appendix contains approaches and techniques related to the following contents: semi-structured interviews and data validation of fishers' local ecological knowledge (LEK), literature mining, spatial data editing (e.g. polygon to raster, resolution standardization), comparison between probability-based sampling and random sampling, original predictor data collection and creation, and macro-habitat data editing. Table A.1 demonstrates the original predictor data of twenty-one parameters collected from various online databases. Figures A.1 show the fishing-zone map used for collecting LEK data.

Table A.1. Original predictors collected for modeling seahorse distributions in China. Those in bold represent the parameters selected by Pearson correlation test. Other predictors (marked with superscripts) were highly correlated (|r| > 0.7) with the selected ones: 1, sea surface temperature (mean); 2, calcite concentration; 3, depth (mean); 4, primary productivity (Dec.).

Category	Predictors	Resolution (arc-degree)	Units	Manipulation	Source
Climate and geophysical suitability	**Sea surface temperature (mean)**	0.0833	℃	Temporal mean from monthly climatologies (2002-2009)	Aqua-MODIS
	Sea surface temperature (maximum)[1]	0.0833	℃	Temporal maximum from monthly climatologies (2002-2009)	Aqua-MODIS
	Sea surface temperature (range)[1]	0.0833	℃	Temporal range from monthly climatologies (2002-2009)	Aqua-MODIS
	Sea surface temperature (minimum)[1]	0.0833	℃	Temporal minimum from monthly climatologies (2002-2009)	Aqua-MODIS
	Chlorophyll A (maximum)[2]	0.0833	mg/m^3	Temporal maximum from monthly climatologies (2002-2009)	Aqua-MODIS
	Chlorophyll A (minimum)[2]	0.0833	mg/m^3	Temporal minimum from monthly climatologies (2002-2009)	Aqua-MODIS
	Chlorophyll A (mean)[2]	0.0833	mg/m^3	Temporal mean from monthly climatologies (2002-2009)	Aqua-MODIS
	Chlorophyll A (range)[2]	0.0833	mg/m^3	Temporal range from monthly climatologies (2002-2009)	Aqua-MODIS
	Dissolved oxygen[1]	0.0833	ml/l	DIVA interpolation of in-situ measurements	WOD 2009
	Nitrate[1]	0.0833	µmol/l	DIVA interpolation of in-situ measurements	WOD 2009
	pH	0.0833	-	DIVA interpolation of in-situ measurements	WOD 2009
	Silicate	0.0833	µmol/l	DIVA interpolation of in-situ measurements	WOD 2009

Category	Predictors	Resolution (arc-degree)	Units	Manipulation	Source
Food availability	**Calcite concentration**	0.0833	mol/m^3	DIVA interpolation of in-situ measurements	WOD 2009
	Salinity[1]	0.0833	pps	DIVA interpolation of in-situ measurements	WOD 2009
	Depth (mean)	0.00833	m	Interpolation with in-situ sounding data guided by satellite-derived gravity data	GEBCO 2014
	Depth (minimum)[3]	0.00833	m	Interpolation with in-situ sounding data guided by satellite-derived gravity data	GEBCO 2014
	Depth (maximum)[3]	0.00833	m	Interpolation with in-situ sounding data guided by satellite-derived gravity data	GEBCO 2014
	Distance to shore	0.01	km	Interpolation	NOAA 2009
	Primary productivity (Dec.)	0.18	mg C/(㎡*day)	VGPM algorithm with satellite images (Modis, SeaWiFS) from 2003-2007	Kershaw 2008
	Primary productivity (Jun.)[4]	0.18	mg C/(㎡*day)	VGPM algorithm with satellite images (Modis, SeaWiFS) from 2003-2007	Kershaw 2008
Macro-habitat	**Coral reefs**	0.0003 – 0.01	-	Field survey 1954-2009	UNEP-WCMC 2010
	Seagrass beds	0.0026	-	Field survey and expert data 1934-2004	UNEP-WCMC 2005
	Mangroves	0.0003	-	satellite imagery 1997-2000	UNEP-WCMC 2011
	Island waters	0.0833	-	see description	This study

Semi-structured interviews and data validation

Our interviews were conducted during fishers' breaks and followed UBC's Human Ethics. We approached to fishers with some ice-breaking questions such as 1) How was your fishing today? 2) How many fishes have you caught? 3) Where is your hometown? We then identify if they have caught seahorses. We then started our interviews with those knowledgeable persons with an introduction about our purpose to build trust. Each interview lasted from 15 to 60 min depending on the amount of data provided by fishers or their time availability. Given budget and time constraints, we sampled at least five vessels per type of fishing gear if available at each port. Sampling efforts varied locally but we normally spent 3 - 7 days at each fishing port, depending on the size of local fisheries and budget constraints. We continue interviewing fishers at the fishing port until little new information was reported (a.k.a. data saturation, Glaser and Strauss 1967, Francis *et al.* 2010), or no more fishers could be interviewed without our sampling efforts.

During the interview, we took the following steps to derive and validate taxonomy data from fishers. Respondents were first asked to display seahorse specimens (either dry or fresh) or photos (in cellphones) of seahorses they had caught. If they have specimens at hand, we then identified the species of the presented specimens.

If they didn't stored specimens or they claimed that there were other species not included in the presented specimens, we used both fresh and dried seahorse photographs to prompt fishers' recalls. The fresh-specimen photos included different colors and sizes of same species, and covered 13 Southeast Asian seahorse species and 3 Japanese seahorse species collected from our iSeahorse database (iseahorse.org) and photos of fresh specimens taken in this field study. Photos of dried seahorses were taken from a collection of specimens that had been collected and identified by Project Seahorse, and from the specimens sighted during this field study. Fishers were asked to draw on differences among the constant body characteristics (e.g. spiny or not, snout length, body size, etc.) rather than using color which is an unreliable trait. Given fishers working on the same vessel tend to have the same knowledge profile of seahorses and could cross-validate each other's information, we interviewed more than just one fisher (if available) on each vessel to identify seahorse species and then map their distributions (See below).

Additionally, we used specimens sighted at the same fishing port from other fishers, local seafood landings or markets, or stores to validate the identification data from fishers' recalls based on our photographs. To be conservative, only those recalls that agreed with sighted specimens were considered as reliable.

To derive distribution maps for each identified species, we designed the following logical flow. First, we asked the respondents if they knew which fishing zones (Fig. A1) have the specific species. For management convenience, China has published fishing-zone maps that evenly divide its marine territory into hundreds of coded fishing zones (resembling a fishnet, Fig. A.1). Each zone spans 1/2 degree in latitude and longitude and consists of nine coded cells (1/6 degree).

For negative answer, we presented fishers with regional nautical charts (scale: 1 : 500,000 to 1,000,000, Fig. A.1) with isobaths, depth points, and substrates. Participants on each vessel were asked to collectively draw a digital distribution map for each species they nominated on an iPad screen with iGIS tools (http://www.geometryit.com/igis/). For fishers who preferred to provide distribution information orally based on nautical charts or their own navigation devices, the field researcher took detailed notes to draw the map later in a GIS. Commercial vessels were equipped with navigation devices that demonstrated their GPS positions against the digital nautical chart and the government fishing-zone map on a screen. This facility helped participants to recall or draw seahorse distribution maps.

For positive answer, we presented the fishing-zone map and asked the respondents to point out the locations. We then further asked habitat (depth range, substrate) and geographic information (longitude and latitude ranges) about the locations. If the respondents didn't know this information, we will stop the interview and draw the map later in a GIS based on the fishing-zone data provided by the respondents. Otherwise, we jotted down the habitat and/or geographic data provided by the fishers, and validate the information with nautical charts later in a GIS. We only remained the maps that matched the nautical charts and refined the map based on the depth range and substrate information.

Finally, we overlaid all polygon maps for each species to seek out "suspicious" distribution data. We assume that distribution maps from one vessel should be cross-validated with another. Therefore, by overlaying all fishers' maps and eradicating non-overlapped portions, we can generate a map that only contains overlapped polygon segments. This process can be realized in a GIS by using a set of tools. We first covert the polygon maps to raster maps by using Conversion Toolbox and then detect overlaps by using the Overlay in the Spatial Analyst Toolbox. Raster Calculator was then used to delete the non-overlapped cells from the map.

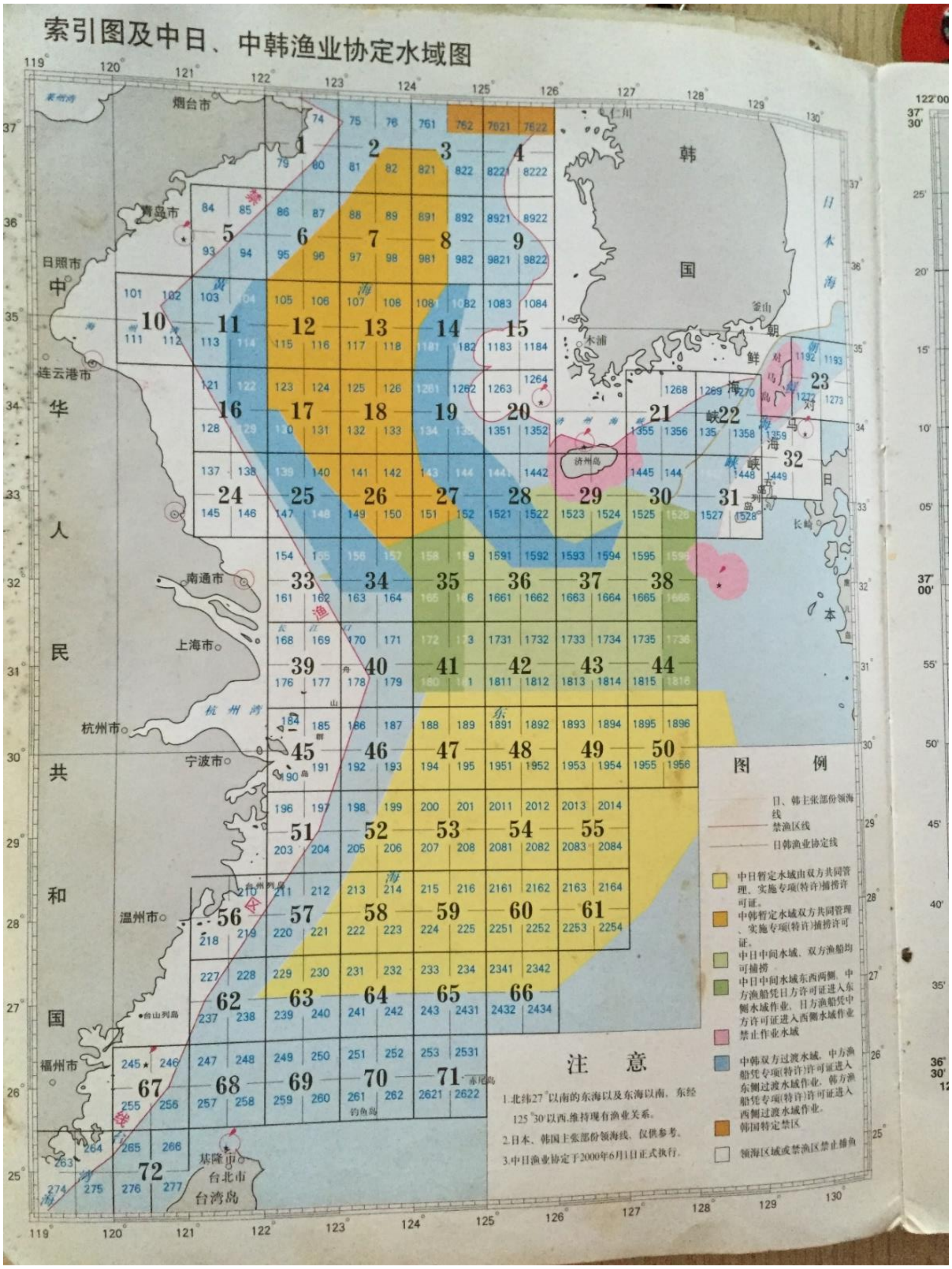


Figure A.1. Part of China's fishing-zone map, including Yellow Sea and East China Sea.

Literature mining for peer-reviewed literature data (PRL)

We mined three sets of Chinese publications: 1) China Academic Journals Full-text Database (1994 onwards), 2) China Doctoral Dissertations Full-text Database (1984 onwards), 3) China Masters' Theses Full-text Database (1984 onwards). We used "*Hippocampus*" and "海马" (meaning seahorses in Chinese) as the "Topic" to search articles. We excluded brain, aquaculture, molecular, and medical studies; the *Hippocampus* is also a part of the brain ("海马体", in Chinese). For all remaining papers, we checked their full-texts to extract taxonomy and distribution information of seahorses in China.

Convert LEK and PRL polygons to regular cells

We first generated a grid (cell size: 0.08333 degree, same as environmental predictors) by using the fishnet tool in a GIS. We then overlaid the fishnet with each polygon map from LEK or PRL. We used the grid cells to replace their overlaid polygon segments to represent distribution data with regular cells. Within each grid cell, the polygon segment with an area less than 5% of the grid cell were eliminated under the assumption that these segments were too small to be replaced by the defined cell and could be treated as noise.

Comparison between Probability-based sampling and random sampling

We followed Niamir et al. (2011) to conduct the comparison. We arbitrarily chose one species (*H. kelloggi*) as the test example, given that the computation complexity and the comparison results are more likely dependent on map resolution rather than species (Niamir et al. 2011). We conducted 30 replicates for both PBS and RS techniques on each dataset (LEK and PRL) to calculated model performance statistics (See Methods and Materials 2.6). We then used the Mann-Whitney-Wilcoxon test to compare model performances between the two techniques.

We found the probability-based sampling (PBS) generally produced more robust models than the random sampling (RS). When tested with LEK data, PBS generated statistically better models than RS by measure of Specificity (true presence rate; Wilcoxon test, $p < 0.05$, $n = 30$), but not AUC (mean or SD) or Sensitivity (true absence rate; all $p > 0.05$, $n = 30$). When tested on PRL data, PBS produced statistically better models than RS by measure of AUC (higher mean and lower SD, Wilcoxon test, $p < 0.01$ for both, $n = 30$), but not Specificity or Sensitivity ($p > 0.05$ for both, $n = 30$). Additionally, the standard deviations of both Sensitivity and Specificity of PBS models were lower than those of RS models, indicating PBS could generate models with more consistent performance than RS.

Original model predictors collected from various sources

We collected a total of 21 original predictors and placed them into three categories.

The first category was climate and geophysical parameters. We selected thirteen climatic predictors that have ecological relevance to seahorses from a global marine environmental database (Tyberghein *et al.* 2012), one parameter of depth from the global bathymetry data (GEBCO_2014 Grid, http://www.gebco.net), and one predictor from the global dataset of distance to the nearest coast (http://oceancolor.gsfc.nasa.gov). For the bathymetry data, over 12,200 soundings were provided with over 8,570 in waters of a depth of 200 m or shallower in the South China Sea provided by the East Asia Hydrographic Commission. The data set was also built on bathymetric contours of the whole China's seas. We then created three variables (mean, minimum, and standard deviation) of depth based on the original bathymetry data.

The second category contained the mean-sea-surface primary productivity in June and December (UNEP-WCMC, http://data.unep-wcmc.org/datasets), which was used as an index of food availability in determining seahorse distributions. The third category was a raster layer, macro-habitat, derived from five habitat types (coral reefs, mangroves, seagrass beds, estuaries, island waters). We explain below how it was generated.

Generating macro-habitat data

We downloaded polygon maps of the global coral reefs, seagrass beds, and mangroves from the UNEP-WCMC Ocean Dataset (http://data.unep-wcmc.org/datasets), and updated the seagrass data with latest published information on China's seagrass distributions (Zheng *et al.* 2013). The polygon map of world estuaries was from the UNEP-WCMC Global Estuary Database (UBC-003-SAU-EstuarieK003-Polygons). For mangrove polygon layer, we assigned the nearest cells to represent its geographic occurrence since most mangrove polygons were onshore. Given there are many small islands offshore in China, we created a raster map for islands smaller than the resolution size of the mapping cell (< 80 sq km). All these habitat data were first trimmed to fit the study region by using the 200-meter depth-range envelope, and then rasterized as binary data (0 for absence, 1 for presence) with the same resolution (0.08333 degrees) in a GIS to generate five categorical predictors.

We then combined these five raster layers into a categorical variable, macro-habitat, with a value ranging from 0 to 29. This process was realized by using Raster Calculator in the Spatial Analyst Toolbox following the below formula:

Macro-habitat = "coral reef" + "estuary" * 2 + "islandwater" * 4 + "mangrove" * 8 + "seagrass" * 16

Macro-habitat codes resulting from the above formula were given below:

0: absence; 1: coral reef; 2: estuary; 3: coral reef+ estuary; 4: islandwater; 5: coral reef + islandwater; 6: estuary + islandwater; 7: coral reef + estuary + islandwater; 8: mangrove; 9: mangrove + coral reef; 10: mangrove + estuary; 11: mangrove + coral reef+ estuary; 12: mangrove + islandwater; 13: mangrove + coral reef + islandwater; 14: mangrove + estuary + islandwater; 15: mangrove + coral reef + estuary + islandwater; 16: seagrass; 17: seagrass + coral reef; 18: seagrass + estuary; 19: seagrass + coral reef + estuary; 20: seagrass + islandwater; 21: seagrass + coral reef +islandwater; 22: seagrass + estuary +islandwater; 23: seagrass + coral reef + estuary + islandwater; 24: seagrass + mangrove; 25: seagrass + mangrove + coral reef; 26: seagrass + mangrove + estuary; 27: seagrass + mangrove + coral reef + estuary; 28: seagrass + mangrove + islandwater; 29: seagrass + mangrove + coral reef + islandwater.

References cited in Appendix A

Francis, J. J., M. Johnston, C. Robertson, L. Glidewell, V. Entwistle, M. P. Eccles, and J. M. Grimshaw. 2010. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychology and Health 25(10):1229-1245.

Glaser, B. G. and A. L. Strauss. 1967. The discovery of grounded theory: Strategies for qualitative research. Piscataway. NJ: Transaction.

Niamir, A., Skidmore, A.K., Toxopeus, A.G., Muñoz, A.R., Real, R., 2011. Finessing atlas data for species distribution models. Diversity and Distributions 17, 1173-1185.

Tyberghein, L., H. Verbruggen, K. Pauly, C. Troupin, F. Mineur, and O. De Clerck. 2012. Bio-ORACLE: a global environmental dataset for marine species distribution modelling. Global Ecology and Biogeography 21(2):272-281.

Zheng, F., G. Qiu, H. Fan, and W. Zhang. 2013. Diversity, Distribution and Conservation of Chinese Seagrass Species. Biodiversity Science 21(5):517-526.

Appendix B Supplementary results for Chapter 2

This appendix contains five supplementary results. Table B.1 refers to the four species that were reported by fishers by unable to be validated in the field. Table B.2 is a list of literature occurrences of seahorse species. Figures B.1 – B.3 show the original maps of species sightings, fishers knowledge, and literature. Figure B.4 depicts the model performance details for each species based on each dataset. Figure B.5 illustrates the predicted probability of seahorse distributions. Figure B.6 – B.13 show species response curves of the eight predictors. Figures B.14 & B.15 demonstrate morphological differences of the validated five species and the un-validated (suspicious) four species. Figure B.16 illustrates sighted seahorse with specific holdfast organisms. Figure B.17 depicts a sighted specimen in the field. The following paragraphs are information about morphological differences among species, species taxonomic errors and uncertainties of the literature occurrences, and the preliminary model performance using the species sightings to derive habitat suitability for downscaling process.

Table B.1. Suspicious seahorse species identified by local fishers in China.

Species	Percentage of reports	Number of reports
Hippocampus histrix	9.4%	21
Hippocampus barbouri	2.7%	6
Hippocampus comes	1.8%	4
Hippocampus coronatus	1.3%	3

Table B.2. Seahorse records in Chinese academic literature.

Record name	Valid name	References	Source type	Institution code
H.coronatus	-	Liu and Ning 2011	PRJ	IOCAS
H.coronatus	-	Song *et al.* 2010	PRJ	MFRISD
H.coronatus	-	Zhu 1999	PRJ	PLA
H.coronatus	-	Han 2013	MscT	IZCAS
H.histrix	*H. spinosissimus*	Han 2013	MscT	IZCAS
H.histrix	*H. spinosissimus*	Liu 2011	MscT	SOU
H.histrix	-	Jiang *et al.* 2014	PRJ	ECSFRI
H.histrix	-	Zhang *et al.* 2013	MscT	JMU
H.histrix	-	Zhao and Zhong 2005	PRJ	ZJOU, SOU
H.kelloggi	*H. kelloggi*	Han 2013	MscT	IZCAS
H.kelloggi	*H. kelloggi*	Liu 2011	MscT	SOU
H.kelloggi	-	Liu and Ning 2011	PRJ	IOCAS
H.kelloggi	-	Zhang *et al.* 2008	PRJ	MFRIZJ, ZJOU
H.kelloggi	-	Zhu 1999	MscT	ZJOU
H.kelloggi	-	Li *et al.* 1987	PRJ	GXOI
H.kelloggi	*H. kelloggi*	Yang 1989	PRJ	XMU
H.kelloggi	-	Zhao and Zhong 2005	PRJ	ZJOU, SOU
H.kelloggi	-	Zhao *et al.* 2012	PRJ	ZJOU
H.kelloggi	*H. kelloggi*	Lian 1990	PRJ	MFRIFJ
H.kuda	-	Han 2013	MscT	IZCAS
H.kuda	*H. kuda*	Liu 2011	MscT	SOU
H.kuda	-	Liu and Ning 2011	PRJ	IOCAS
H.kuda	-	Liu and Tian 1995	PRJ	IOCAS
H.kuda	-	Zhu 1999	PRJ	PLA
H.kuda	*H. kuda*	Li *et al.* 1987	PRJ	GXOI
H.mohnikei	*H. mohnikei*	Wan and Jiang 1998	PRJ	YSFRI
H.mohnikei	*H. mohnikei*	Liu and Ning 2011	PRJ	IOCAS
H.mohnikei	*H. mohnikei*	Liao *et al.* 2014	PRJ	SOA
H.mohnikei	*H. mohnikei*	Zhang *et al.* 2014	PRJ	MEDL
H.mohnikei	*H. mohnikei*	Zhang *et al.* 2010	PRJ	ECSFRI
H.mohnikei	*H. mohnikei*	Zhu 1999	PRJ	PLA

Record name	Valid name	References	Source type	Institution code
H.mohnikei	*H. mohnikei*	Piao 2005	MscT	COU
H.mohnikei	*H. mohnikei*	Li *et al.* 2013	PRJ	MFRISD
H.mohnikei	*H. mohnikei*	Li *et al.* 2014	PRJ	MFRILN,MBEL
H.mohnikei	*H. mohnikei*	Li 1987	PRJ	NKU
H.mohnikei	*H. mohnikei*	Li et al. 1987	PRJ	GXOI
H.mohnikei	*H. mohnikei*	Yang 1989	PRJ	XMU
H.mohnikei	*H. mohnikei*	Mei *et al.* 2010	PRJ	COU
H.mohnikei	*H. mohnikei*	Jiang and Chen 1993	PRJ	XMU
H.mohnikei	*H. mohnikei*	Wang *et al.* 2006	PRJ	ECSFRI
H.mohnikei	*H. mohnikei*	Wang 2013	MscT	COU
H.mohnikei	*H. mohnikei*	Wang 2009	MscT	COU
H.mohnikei	*H. mohnikei*	Zhai *et al.* 2014	PRJ	COU
H.mohnikei	*H. mohnikei*	Zhao and Zhong 2005	PRJ	ZJOU, SOU
H.mohnikei	*H. mohnikei*	Zhao et al. 2012	PRJ	ZJOU
H.mohnikei	*H. mohnikei*	Zheng 2014	MscT	SOU
H.mohnikei	*H. mohnikei*	Qi 2013	MscT	COU
H.mohnikei	*H. mohnikei*	Han 2013	MscT	IZCAS
H.trimaculatus	*H. trimaculatus*	Han 2013	MscT	IZCAS
H.trimaculatus	*H. trimaculatus*	Liu 2011	MscT	SOU
H.trimaculatus	*H. trimaculatus*	Li et al. 1987	PRJ	GXOI
H.trimaculatus	*H. trimaculatus*	Xiao *et al.* 2013	PRJ	SOA

[1] Source type: PRJ, peer-reviewed journal; MscT, master thesis.

[2] Institution: OUC, Ocean University of China; ECSFRI, East China Sea Fisheries Research Institute, Chinese Academy of Fisheries Sciences; GXOI, Guangxi Oceanography Institute; IOCAS, Institute of Oceanography, Chinese Academy of Sciences; IZCAS, Institute of Zoology, Chinese Academy of Sciences; JMU, Jimei University; MBEL, Marine Biological Resources and Ecology Key Lab of Liaoning Province; MEDL, Marine Ecological Environment and Disaster Prevention and Reduction Key Lab of Shangdong Province; MFRIFJ, Marine Fisheries Research Institute of Fujian Province; MFRILN, Academy of Marine Fisheries Sciences of Liaoning Province; MFRISD, Marine Fisheries Research Institute of Shandong Province; MFRIZJ, Marine Fisheries Research Institute of Zhejiang Province; NKU, Nankai University; PLA, Penglaige Aquarium; SOA, State

Oceanic Administration; SOU, Shanghai Ocean University; XMU, Xiamen University; YSFRI, Yellow Sea Fisheries Research Institute; ZJOU, Zhejiang Ocean University.

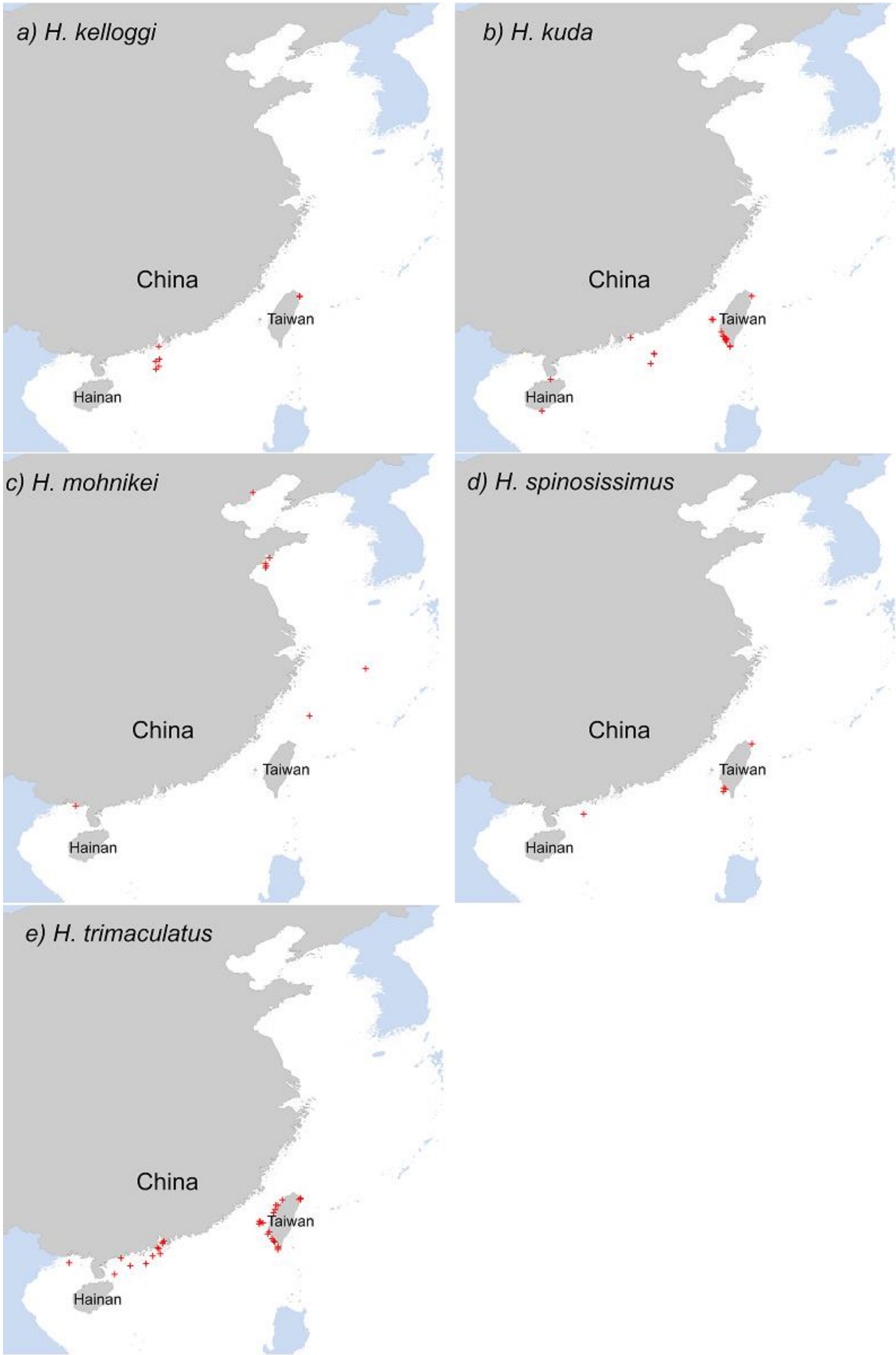


Figure B.1. Map of the original species sightings ("+") of a) *H. kelloggi*, b) *H. kuda*, c) *H. mohnikei*, d) *H. spinosissimus*, e) *H. trimaculatus.*

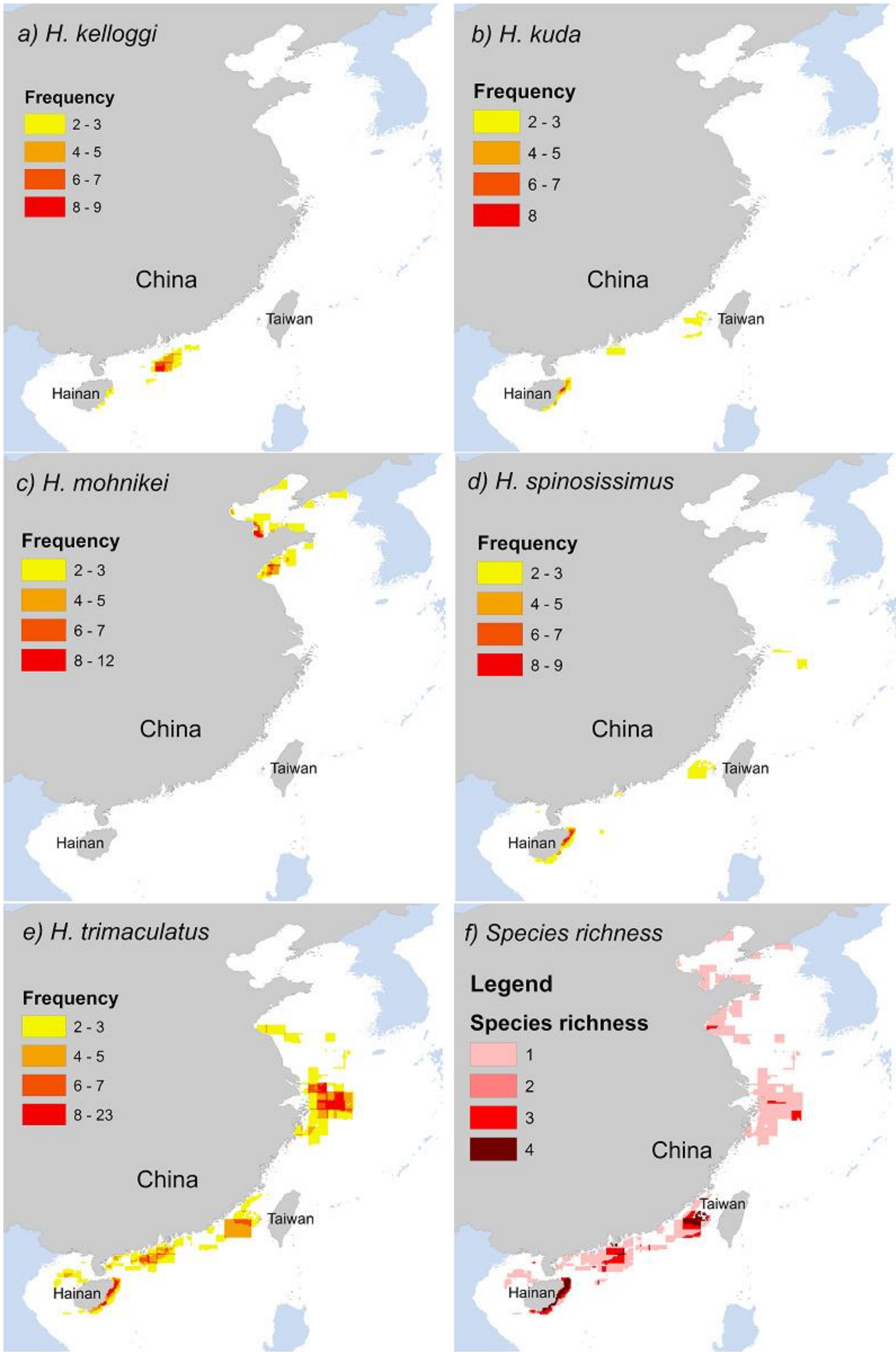


Figure B.2. Map of the original species distributions of a) *H. kelloggi*, b) *H. kuda*, c) *H. mohnikei*, d) *H. spinosissimus*, and e) *H. trimaculatus*, based on fishers' local ecological knowledge. Frequency indicates the number of visited vessels that agreed on the range maps. Only frequency larger than two (i.e. maps from more than two vessels overlapped at the same area) was considered as valid.

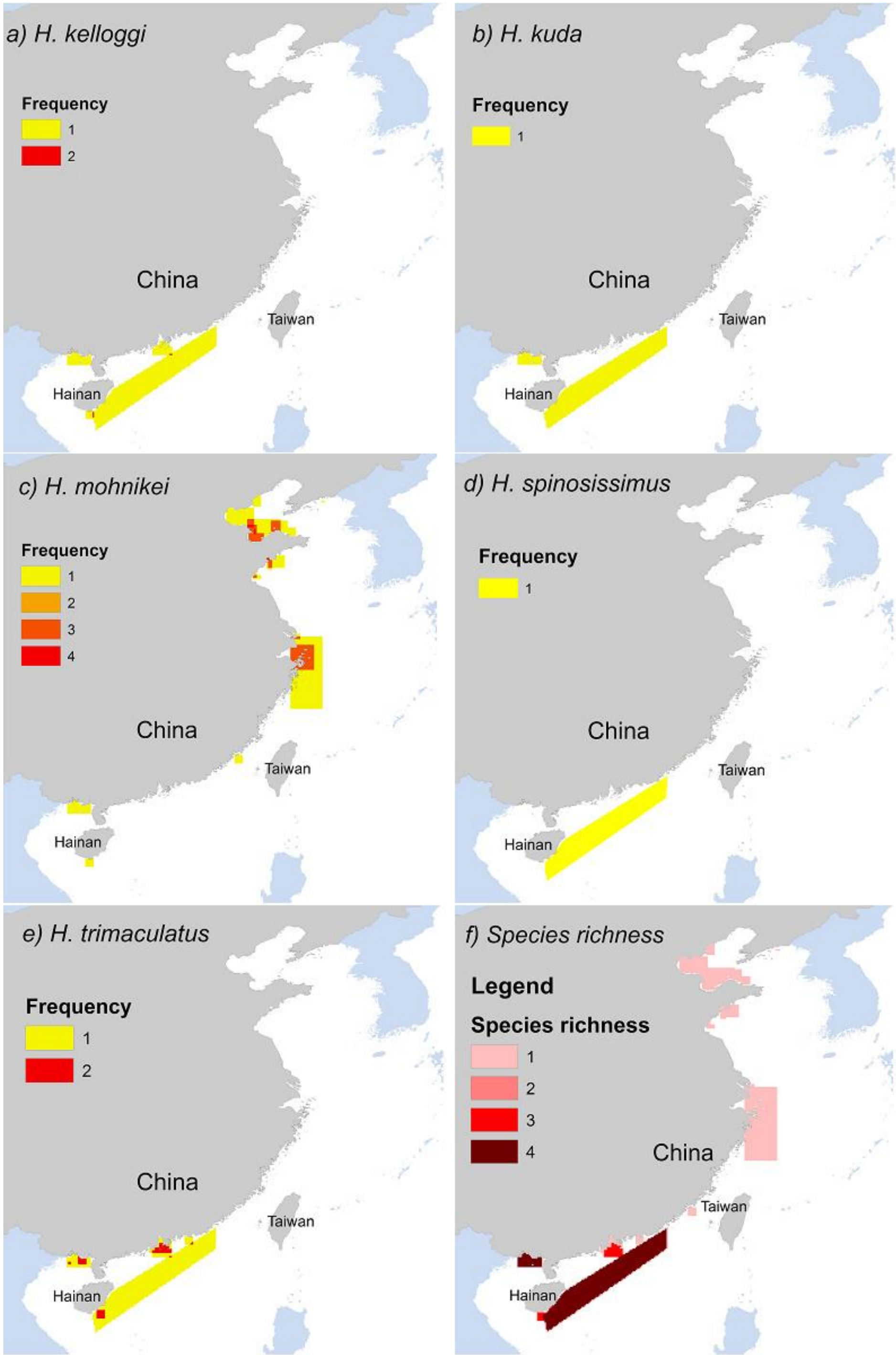


Figure B.3. Map of the original species distributions of a) *H. kelloggi*, b) *H. kuda*, c) *H. mohnikei*, d) *H. spinosissimus*, and e) *H. trimaculatus*, based on peer-reviewed literature. Frequency indicates the number of articles that agreed on the range map.

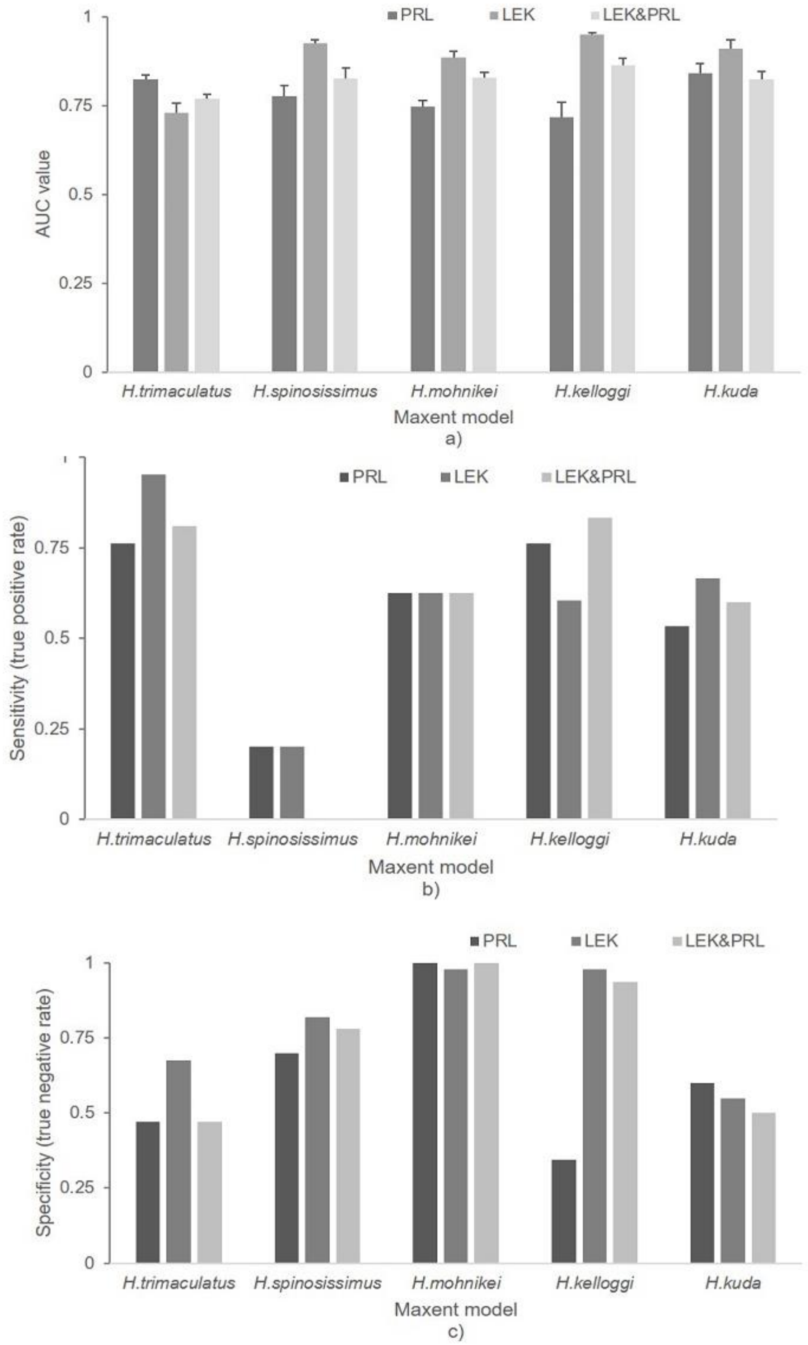


Figure B.4. Model performance measured with the a) AUC, b) Sensitivity, and c) Specificity of each Maxent model for each seahorse species based on three different datasets: PRL, peer-reviewed literature; LEK, local ecological knowledge; LEK&PRL, them both. The error bar for AUC on a) indicated standard deviations across 15 replicates.

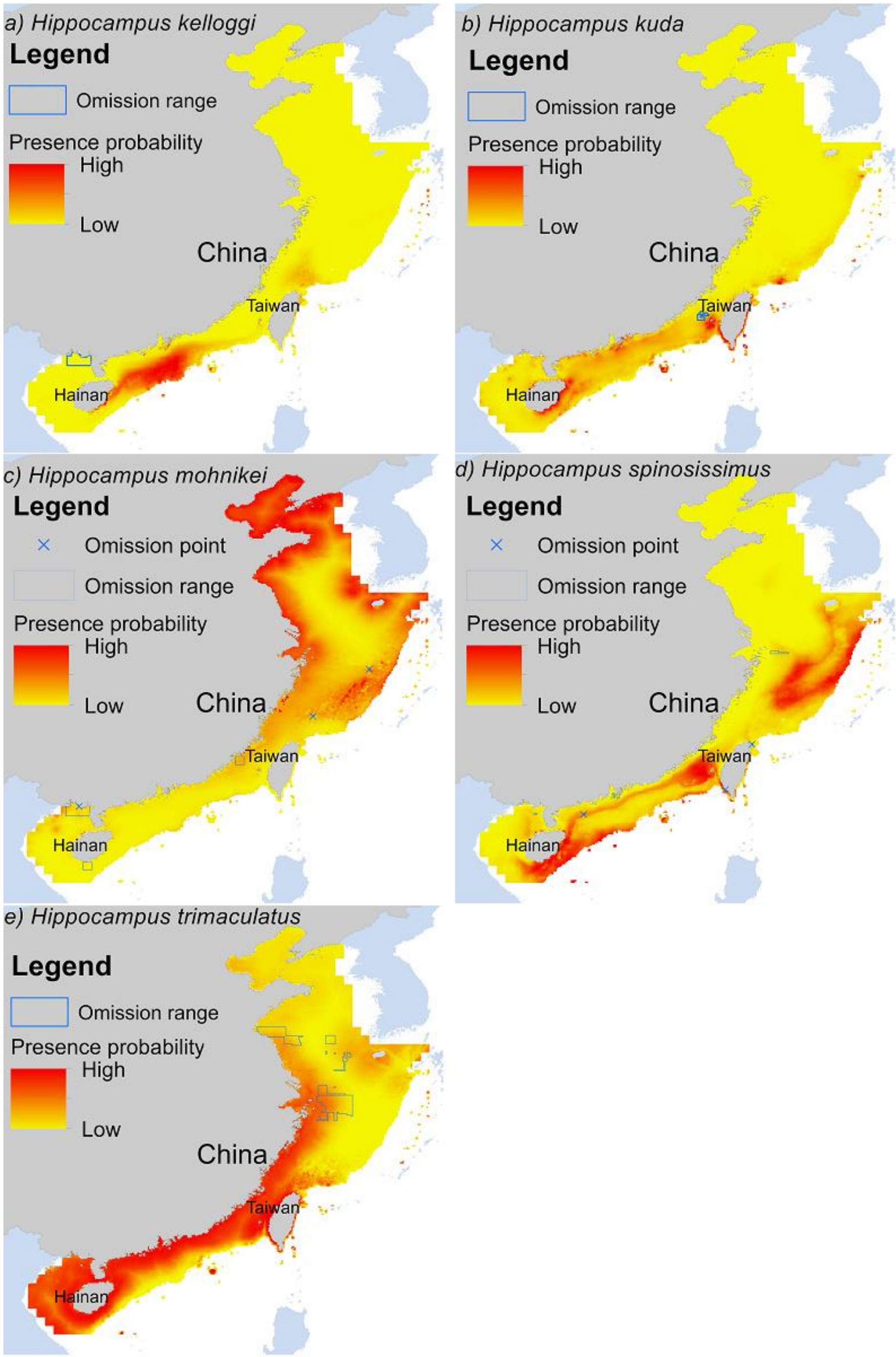


Figure B.5. Presence-probability map of a) *H. kelloggi, b) H. kuda, c) H. mohnikei, d) spinosissimus, and e) H. trimaculatus* in China. Omission range or point refers to the range or point that was not represented in the derived presence/absence map, but was shown in original data.

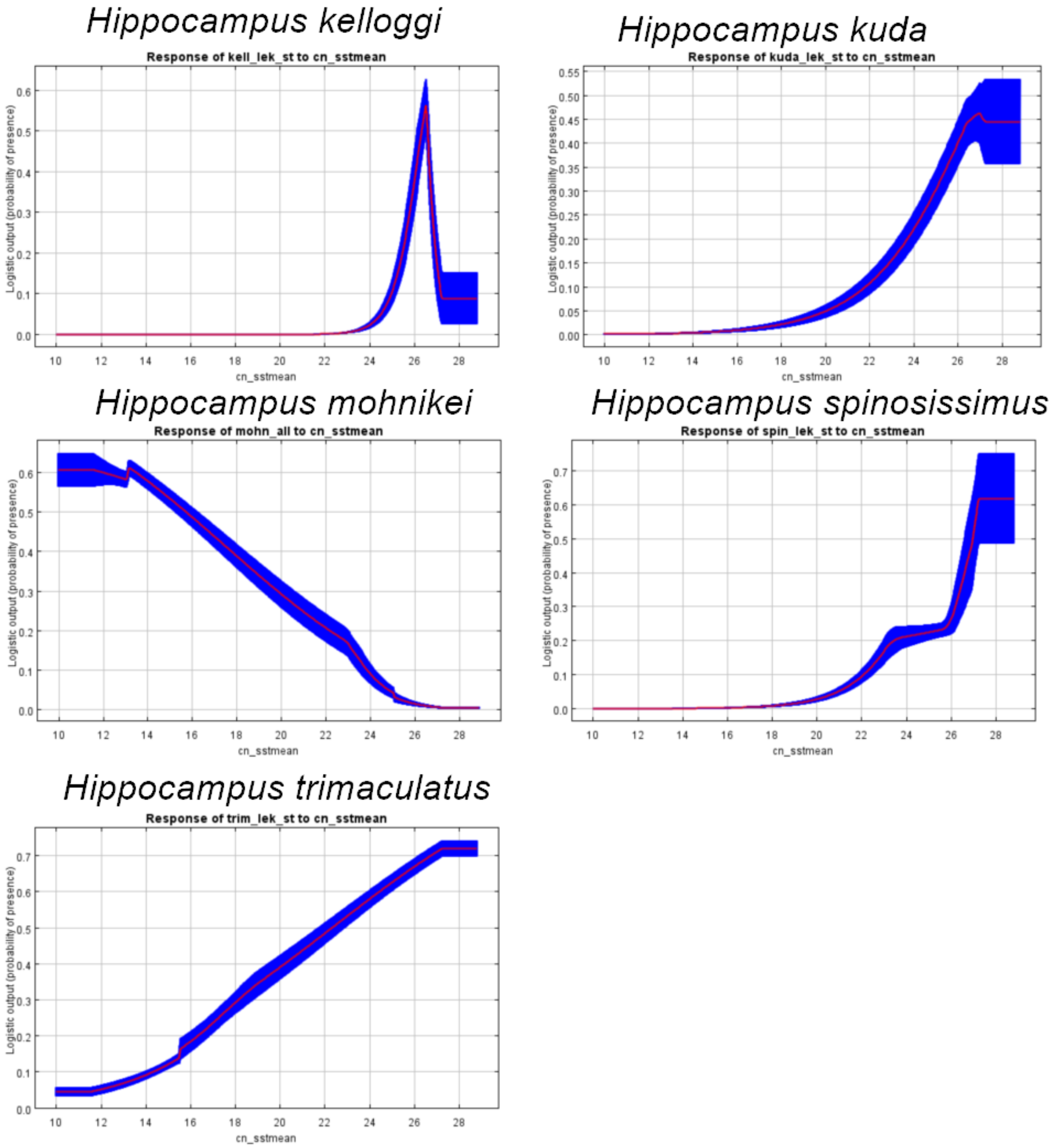


Figure B.6. Species-temperature (mean sea surface temperature) response curves for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and temperature on x-axis.

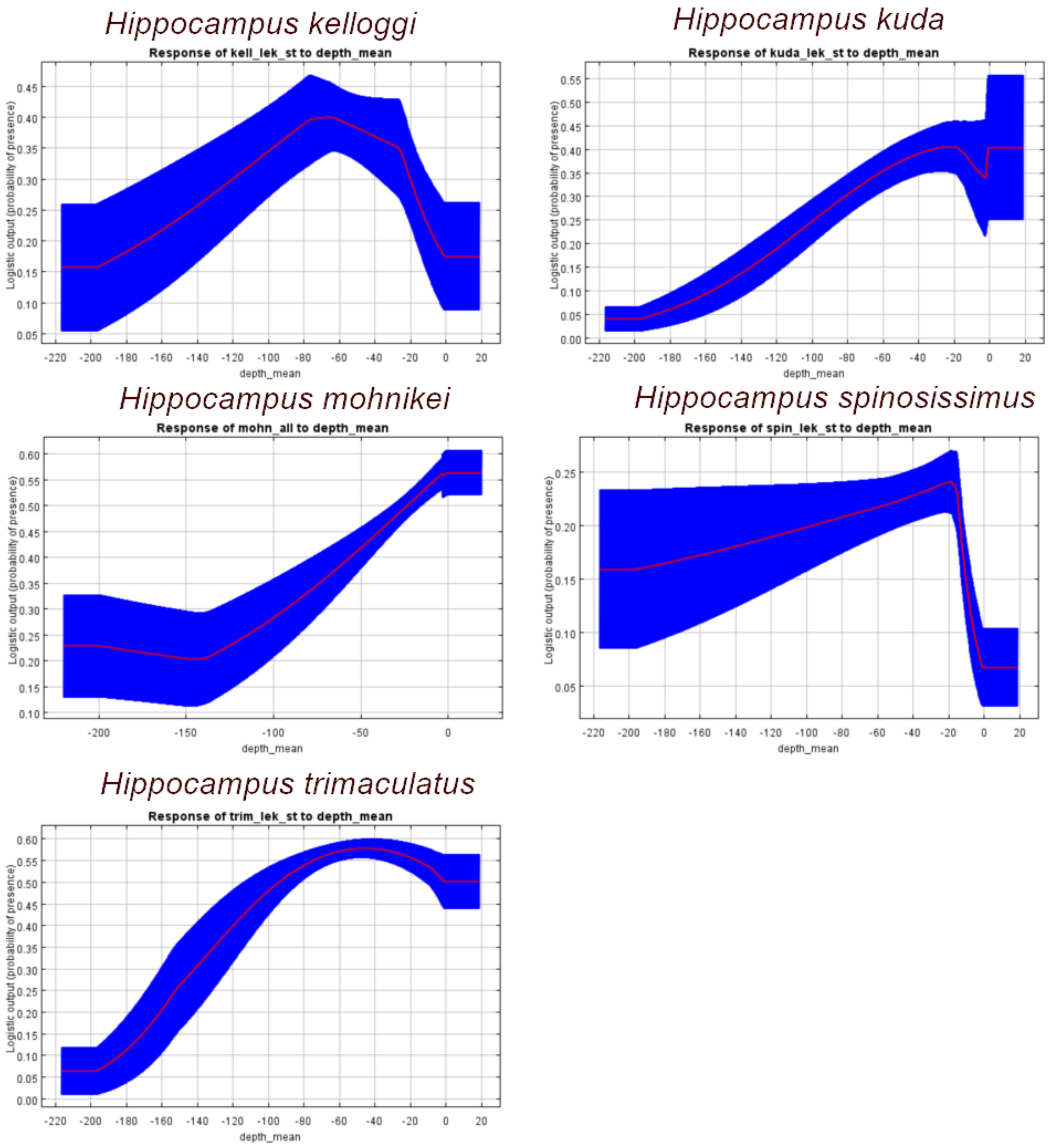


Figure B.7. Species-depth response curves for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and depth on x-axis.

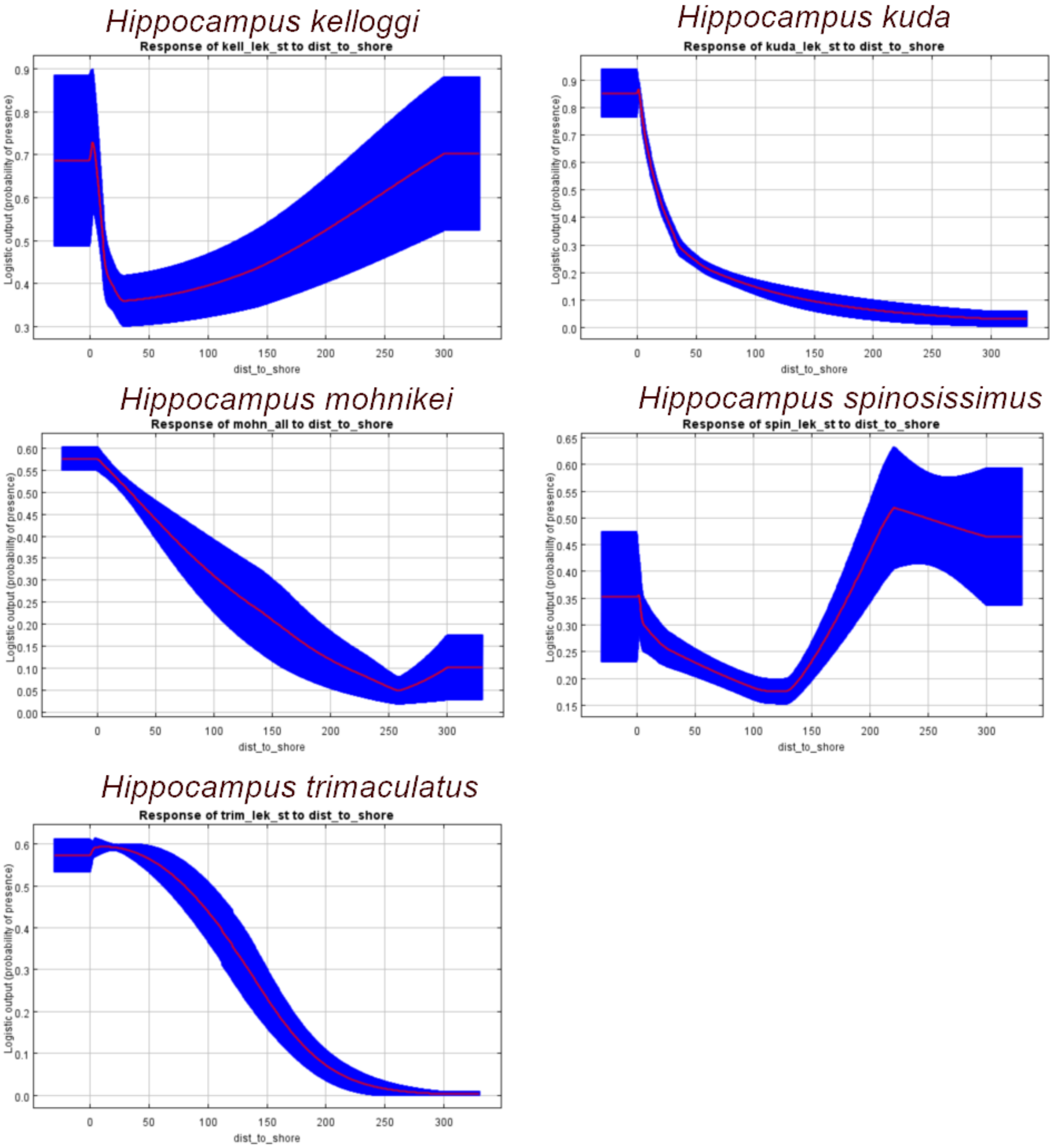


Figure B.8. Species-distance (distance to the nearest shore) response curves for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and distance on x-axis.

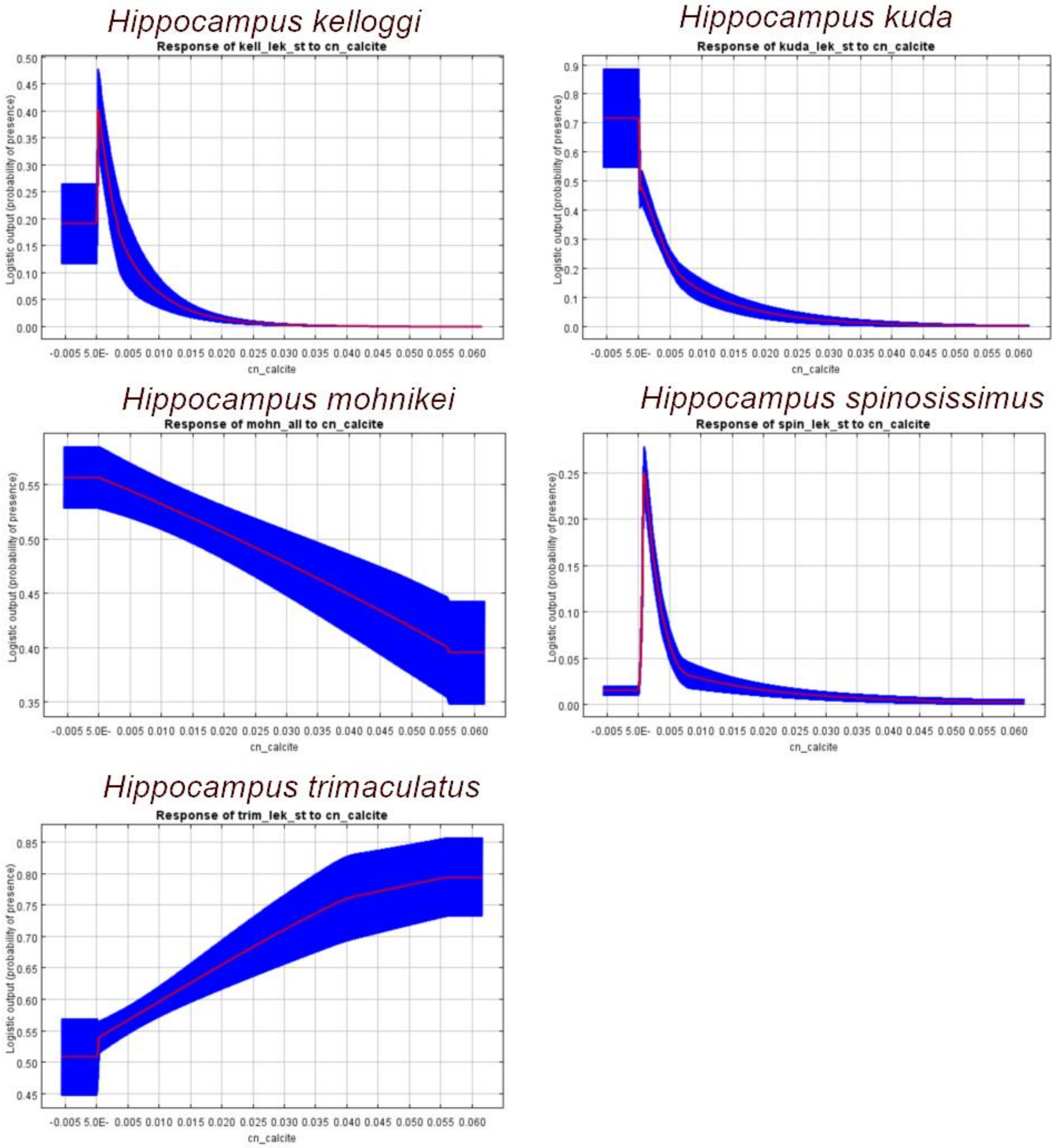


Figure B.9. Species-calcite response curves for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and calcite concentration on x-axis.

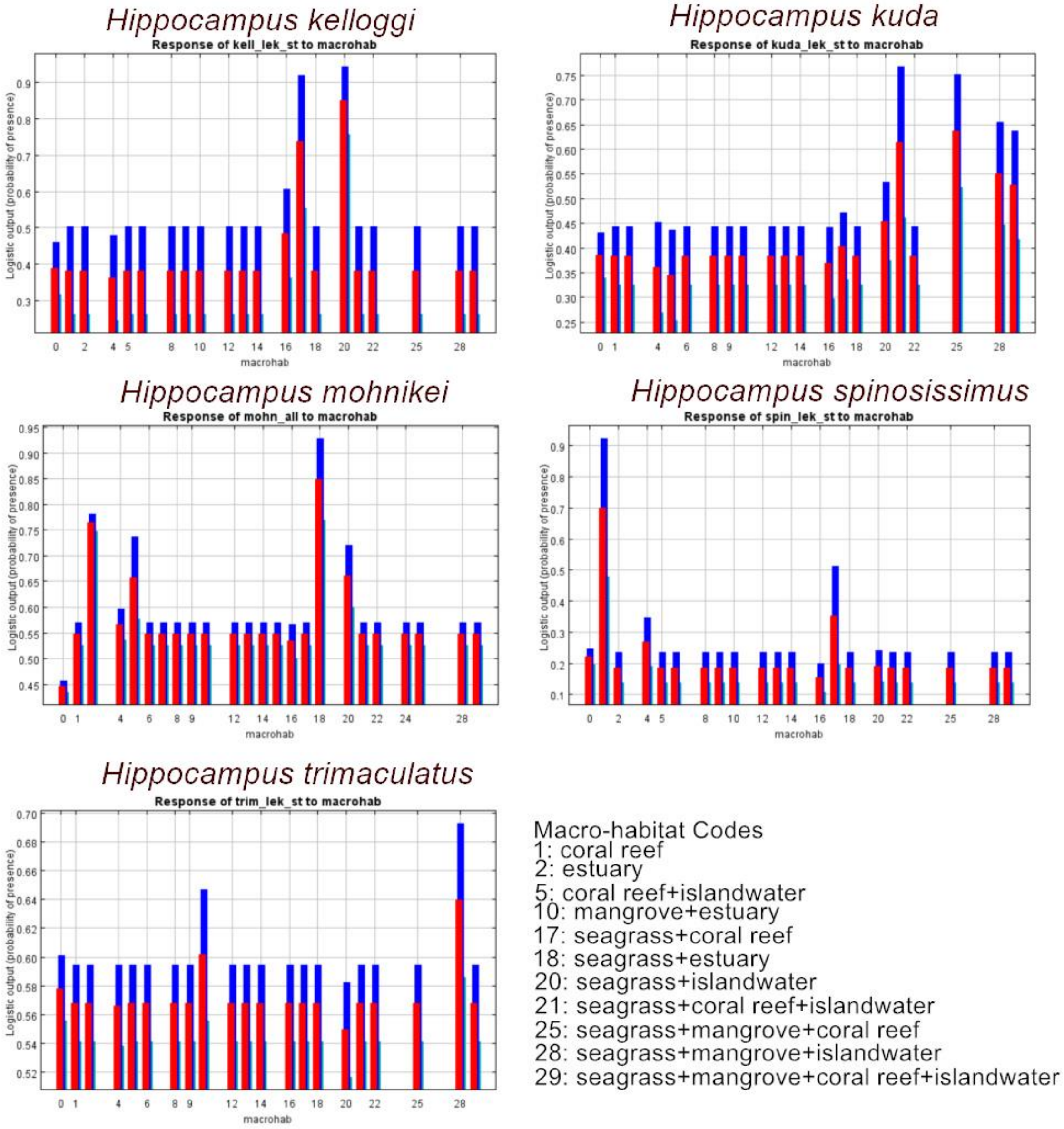


Figure B.10. Species-macrohabitat relations for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and macrohabitat categories on x-axis.

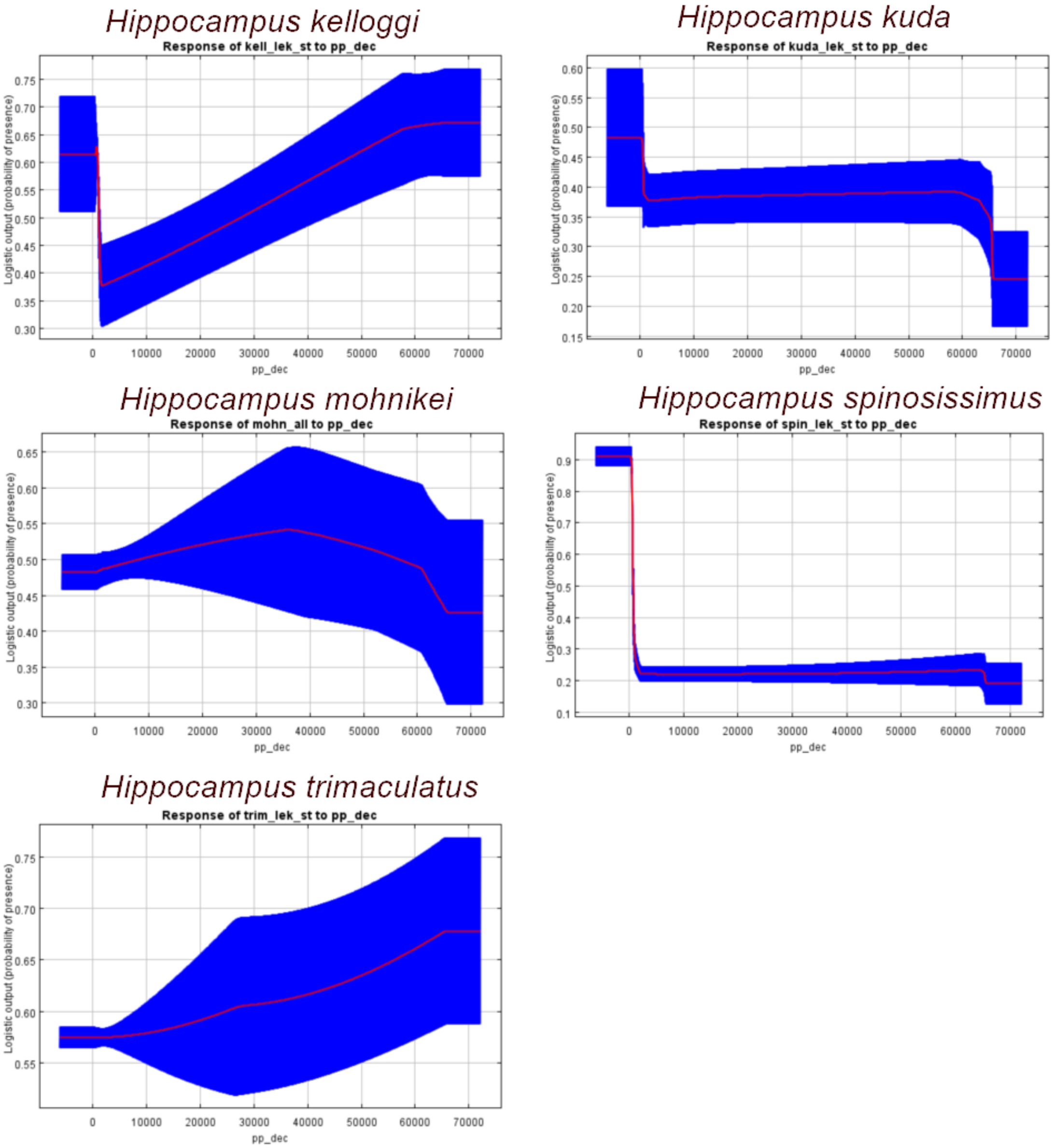


Figure B.11. Species-productivity (mean primary productivity in December) response curves for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and primary productivity on x-axis.

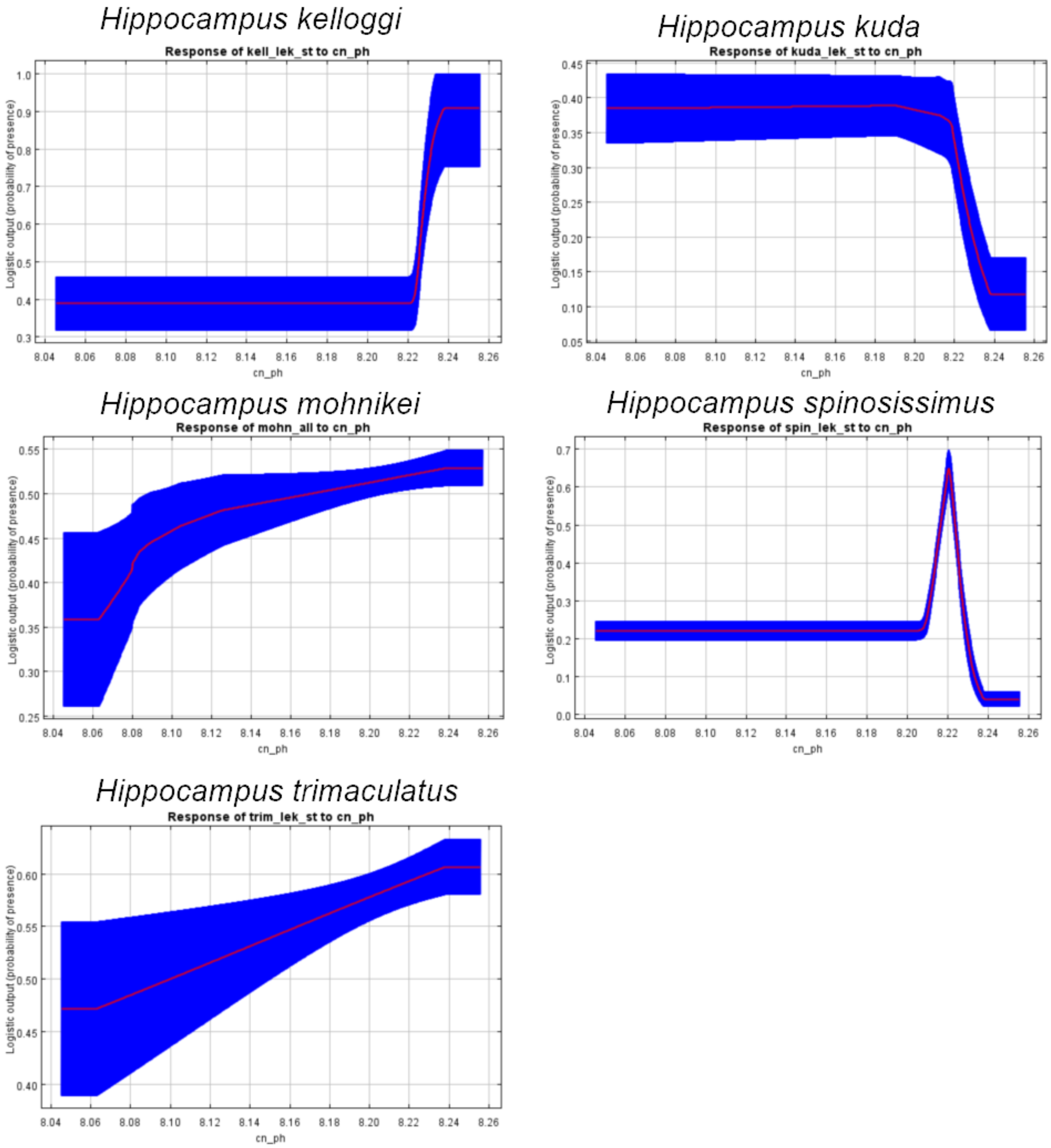


Figure B.12. Species-pH response curves for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and pH on x-axis.

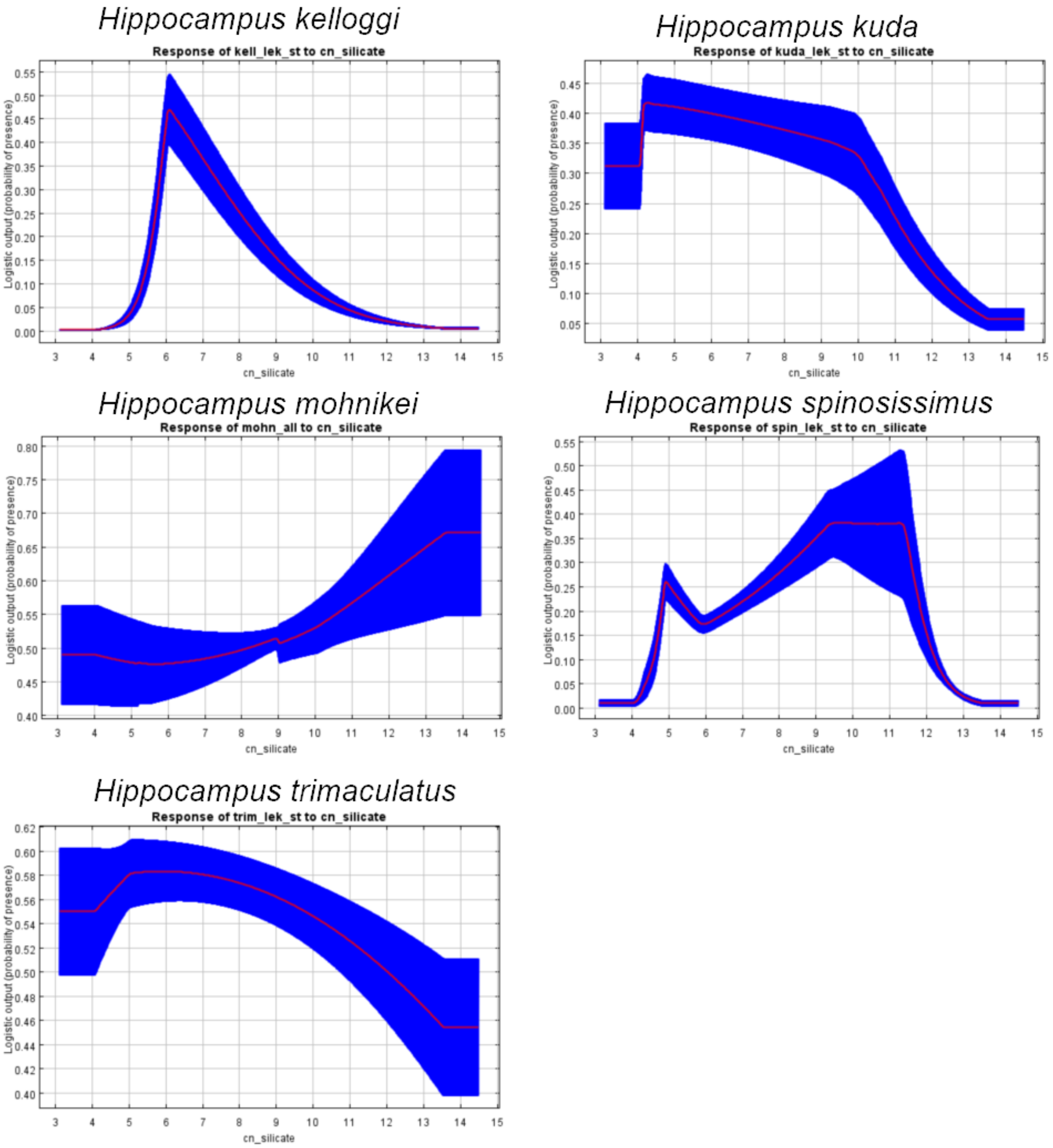


Figure B.13. Species-silicate response curves for the five seahorse species (*H. kelloggi, H. kuda, H. mohnikei, H. spinosissimus, H. trimaculatus*), with habitat suitability on y-axis and silicate concentration on x-axis.

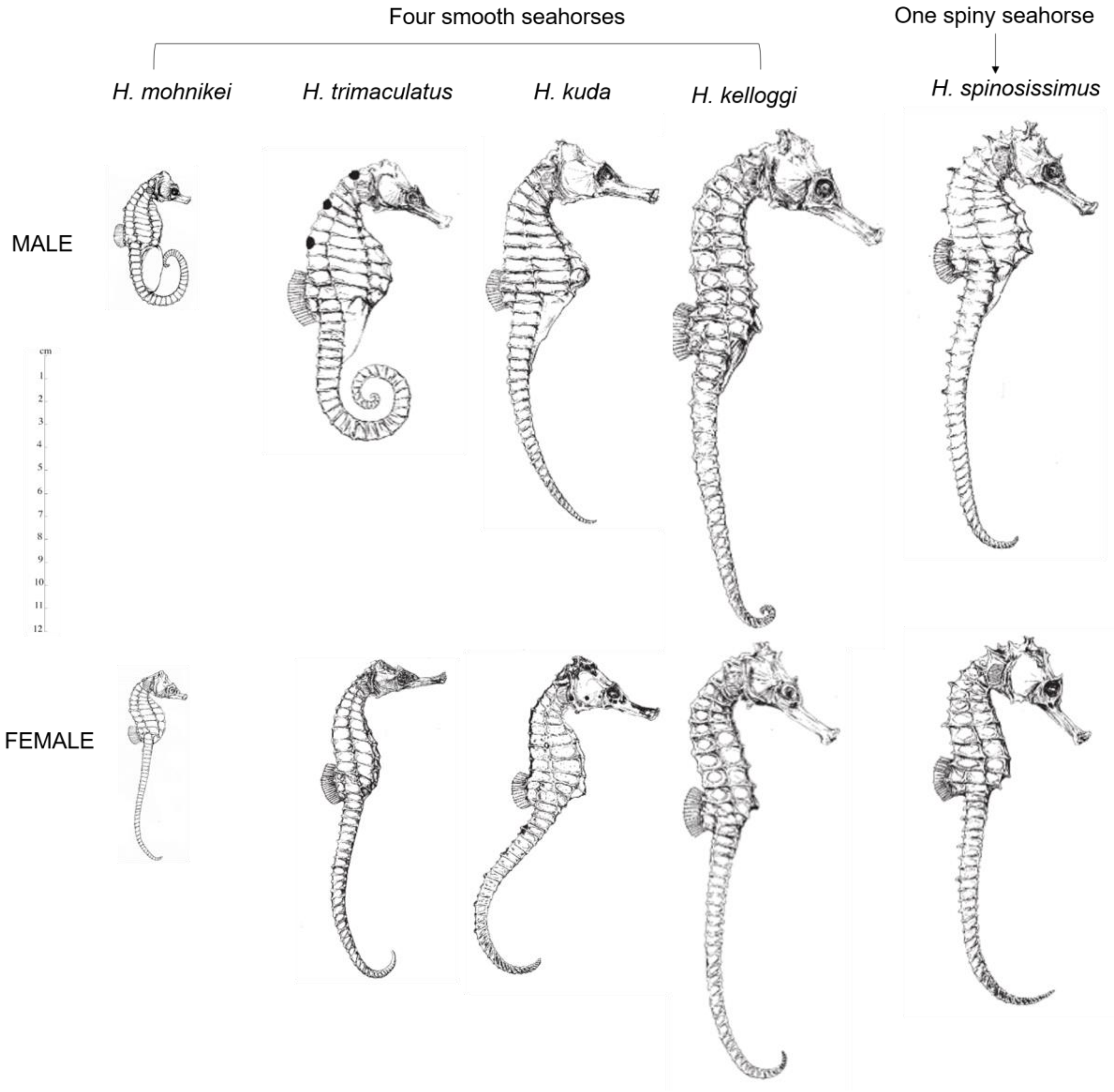


Figure B.14. A sketch figure depicting five confirmed seahorse species (drawn to scale based on their common heights, pictures were from Lourie *et al.,* 2004)

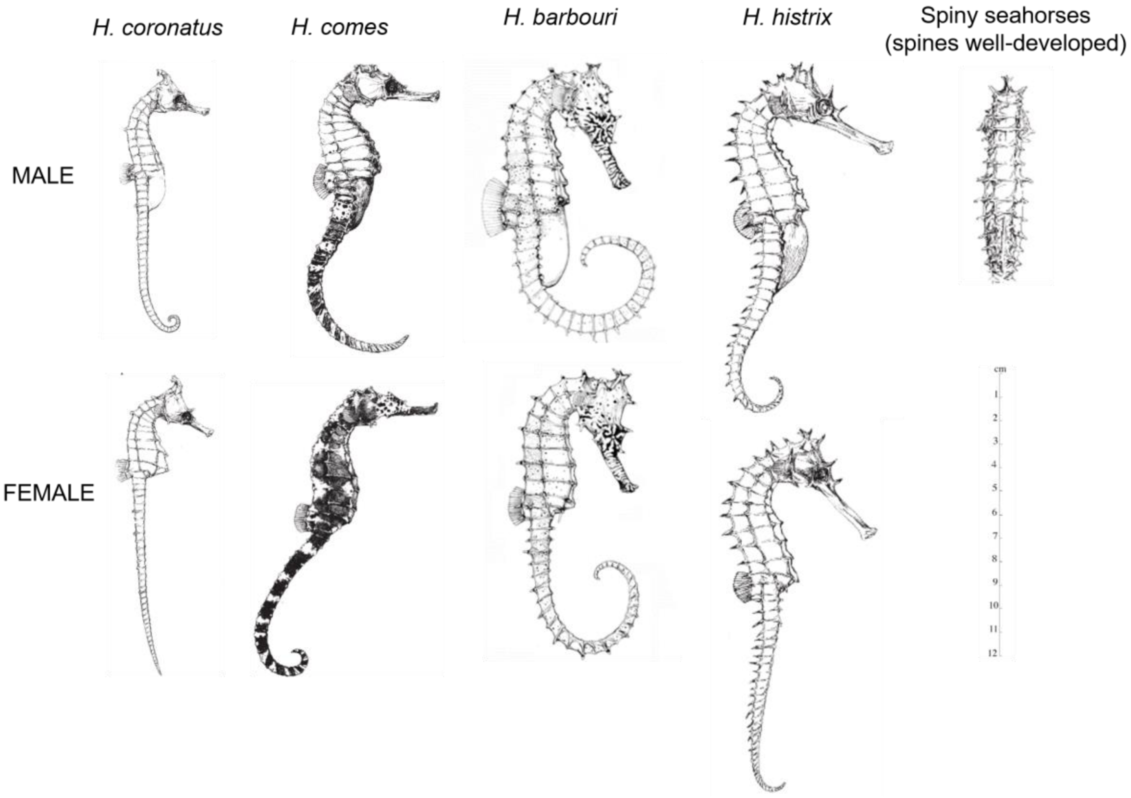


Figure B.15. A sketch map showing the four suspicious seahorse species claimed by some fishers (drawn to scale based on Lourie *et al.*, 2004)

Figure B.16. Photos of a) a juvenile seahorse holding a branch of macro-algae, b) a sea fan shown by a fisher, and c) a clump of juvenile seahorses caught by a surface-trawler.

Figure B.17. Sighted specimens of *Hippocampus kuda* (a dried sample, male, height length ~ 11 cm) at Nan’ao Fishing Port, Hong Kong SAR.

Morphological differences among the five confirmed species

Hippocampus mohnikei is the smallest seahorse species among the identified ones, its snout length is also the shortest (relative to head length). This species was mostly found alone in the north. Local fishers often call it the "small seahorse" – "Xiao Hai Ma" (Chinese pinyin). *Hippocampus trimaculatus* has three iconic black spots on the back (more apparent on males) and the mostly caught seahorses in fisheries in the south, and local fishers call it the "spotted seahorses" – "Ban Hai Ma" (Chinese pinyin). *Hippocampus kuda* has relatively thick trunk and large body size but rarely occurs in fisheries especially commercial fisheries. Local fishers call it the "large seahorses" – "Da Hai Ma" (Chinese pinyin). *Hippocampus kelloggi* has the significantly largest body size and usually occur in deeper waters as told by fishers. We found this species usually has a gold-to-white color on females, and brown color on males. Local fishers call it the "golden seahorses" or "deep-water seahorses" – "Huang Jing Hai Ma" or "Shen Shui Hai Ma" (Chinese pinyin). *Hippocampus spinosissimus* is the only spiny species among the four and local fishers call it "spiny seahorses" – "Ci Hai Ma" (Chinese pinyin).

Misidentification of *Hippocampus spinosissimus* as *Hippocampus histrix*

Hippocampus histrix was first recorded in China in the *Ichthyography of South China Sea* dating back to 1962 (Zhu and Cheng 1962), and currently in two PRL articles and three Master theses (Zhao and Zhong 2005, Liu 2011, Han 2013, Zhang 2013, Jiang et al. 2014). The recent identifications of *H. histrix* were all based on the above *Ichthyography*. We found that this is a misidentification of *H. spinosissimus* (another thorny seahorse species) by checking the specimen photos provided by two colleagues. The name *H. histrix* has been indiscriminately used for many spiny seahorses in history including *H. spinosissimus*, and the morphological differences between the two have only been described recently (Lourie et al. 2004), which were unaware to Chinese colleagues.

Uncertainty of Hippocampus coronatus

Hippocampus coronatus was first documented in China in the *Fish of Yellow Sea and Bohai Sea* (Zhang et al. 1994) and then in three academic articles and one Master Thesis (Zhu 1999, Liu and Ning 2011, Song *et al.* 2010, Han 2013). All authors identified specimens based on Zhang *et al.* (1994), but did not reserve the specimens. We only obtained a photo of dried specimen of this species collected by the first recorder (Dr. Zhang Chun-ling, died in 1962) in the Bohai Sea with unclear source. This species hasn't been sighted in Bohai Sea in recent years by a Chinese ichthyologist (Dr. Zhang Chun-Guang, *pers. comm.*). We did not sight this species at nine fishing ports we visited around the Bohai Sea, although some fishers claimed rare historic occurrence of this species.

Preliminary model performance based on SS dataset

Our model produced a large range of test AUC mean values (0.522 – 0.854) based on SS dataset alone for the identified five seahorse species in China. Among them, three species obtained a moderate model performance (0.7 < AUC < 0.9), while the other two derived low model performances (0.5 < AUC < 0.7). *Hippocampus spinosissimus*, which had the lowest number of sightings (n = 5), gained a model close to random performance when measured by testing AUC (0.522 ± 0.294). The variability in predictive ability (AUC standard deviation) for all models was high (0.057 – 0.275).

References cited in the Appendixes B

Han, S. 2013. Taxonomy, resources, utilization and conservation of seahorses in China. Master thesis. Guangxi Normal University (In Chinese).

Jiang, S. and F. Chen. 1993. Studies on Estuarine Ecology of Ichthyoplankton of Jiulong River. Journal of Oceanography in Taiwan Strait 12(4):351-358.

Jiang, Y., N. Lin, X. Yuan, H. Jiao, and S. Li. 2014. Functional group composition and functional diversity of nekton community in the Xiangshan Bay. Oceanologia et Limnologia Sinica 45(1):108-114.

Li, F., H. Zhang, Z. Lu, B. Xu, and Z. Liang. 2013. Species composition and community diversity of nekton in Laizhou Bay, China. Biodiversity Science 21(5):537-546.

Li, H., X. Liu, M. Zhang, Q. Wang, X. Wang, C. Fu, and D. Li. 2014. The Investigation on Nekton of Southwestern Dalu Island in Spring and Autumn. Fishery Modernization 41(3):48-52.

Li, M. 1987. Distributions and Reproduction Characteristics of Fishes in the Beitang Estuary. Acta Oceanologica Sinica 9(2):227-238.

Li, X.-S., Z.-H. Liang, and M.-X. Jiang. 1987. A Preliminary Survey of the Fish Fauna in the North of the Beibuwan Bay, China. Journal of the Guangxi Academy of Sciences 3(2):95-116.

Lian, Z.-S. 1990. A Study on Fish Fauna in the South of Quanzhou Bay, China. Fujian Fisheries 1:8-11.

Liao, J., X. Zheng, J. Du, B. Chen, Z. Ma, and W. Hu. 2014. Biodiversity and trophic level characteristics of fishes captured by set nets in Tong'an Bay, Xiamen. Biodiversity Science 22(5):624-629.

Liu, J. and P. Ning. 2011. Species composition and faunal characteristics of fishes in the Yellow Sea. Biodiversity Science 19(6):764-769. In Chinese.

Liu, J. and M.-C. Tian. 1995. A preliminary study and review of the coral reef fishes Hainan Island. Marine Science 5:28-32.

Liu, W. 2011. Community structure dynamics of nekton in the continental shelf of northern South China Sea. Master thesis. Shanghai Ocean University. In Chinese.

Lourie, S. A., S. J. Foster, E. W. Cooper, and A. C. Vincent. 2004. A guide to the identification of seahorses. Project Seahorse and TRAFFIC North America Project Seahorse and TRAFFIC North America, University of British Columbia and World Wildlife Fund, Washington, DC.

Mei, C., B. Xu, Y. Xue, Y. Ren, and X. Zan. 2010. A Study on Diversity and Community of Fish Assemblages in the Middle Region of Jiaozhou Bay during Autumn and Winter. JOURNAL OF FISHERY SCIENCES OF CHINA(1):110-118.

Piao, C. 2005. The Geographical Features and Fish Biodiversity Investigation in the Jiaozhou Bay and its Neighboring Waters [D]. China Ocean University.

Qi, G. 2013. Catch Composition and Effects on Fish Resources of the Set-net along the Coastal Water of Haizhou Bay. China Ocean University.

Song, X., A. Liu, Y. Yang, J. Yang, L. Ren, L. Liu, G. Sun, and X. Liu. 2010. Distribution of Ichthyoplankton and the correlation with environmental factors in Laizhou Bay. Oceanologia et Limnologia Sinica 41(3):378-385. In Chinese.

Wan, R. and Y. Jiang. 1998. THE DISTRIBUTION AND VARIATION OF EGGS AND LARVAE OF OSTEICHTHYES IN THE BOHAI SEA [J]. JOURNAL OF FISHERY SCIENCES OF CHINA 5(1):43-50.

Wang, A. 2009. A Preliminary Study on Fish Community Structure of Spring Eggs and Larvae and Their Relationships with Environmental Factors in Laizhou Bay, Bohai Sea. China Ocean University.

Wang, X. 2013. Spatio-temporal Dynamics of Fish Community Structure in Haizhou Bay and Neighboring Waters. China Ocean University.

Wang, Y., Y. Ni, C.-s. Li, J.-x. Xhu, P. Zhuang, and T. Zhang. 2006. A new record on the fish fauna in Shanghai - *Hippocampus japonicus* Kaup. Marine Fisheries 28(1):87-88.

Xiao, Y.-z., R. Wang, Y.-j. Zheng, and W. He. 2013. Species Composition and Abundance Distribution of Ichthyoplankton in the Pearl River Estuary. Journal of Tropical Oceanography 32(6):80-87.

Yang, S.-Y. 1989. Characteristics of the Fisheries Resources and Fisheries Production of the Luoyuanwan Bay, China. Fujian Fisheries 64(1):23-32.

Zhai, L., D. Han, D. Fu, C. Zhang, and Y. Xue. 2014. Fish community structure and the relationship with environmental factors in Jiaozhou Bay and adjacent waters JOURNAL OF FISHERY SCIENCES OF CHINA 21(4):810-821.

Zhang, C., Q. Cheng, and B. Zheng. 1994. Fish of Yellow Sea and Bohai Sea. Taiwan: Aquaculture Press. Page 83 - 84.

Zhang, H.-l., K.-d. XU, Z.-t. HE, Z.-f. LIU, and Y.-d. ZHOU. 2008. Analysis on the condition of fishery resources in the sea area around the Jiushan Archipelago [J]. Marine Fisheries 2:003.

Zhang, J., D. R. Wang, T. Jennerjahn, and L. Dsikowitzky. 2013. Land–sea interactions at the east coast of Hainan Island, South China Sea: A synthesis. Continental Shelf Research 57:132-142.

Zhang, L., J. Wang, R. Ren, N. Zhang, K. Su, W. Qu, H. Tao, and Z. Lu. 2014. Fishery Resources in Northern Haizhou Bay in Spring. PROGRESS IN FISHERY SCIENCES 35(5):1-7.

Zhang, X., X. Zhang, and T. Gao. 2010. Comparative analysis on catch composition with two fishing gears at Yellow River estuary in spring. South China Fisheries Science 6(1):59-67.

Zhao, S.-l., J. Chen, and F.-j. Yu. 2012. Study on Species Composition, Distribution and Resources of Zhejiang Marine Fishes [J]. Journal of Zhejiang Ocean University (Natural Science) 31(7):1-11.

Zhao, S. and J. Zhong. 2005. New checklist of marine fishes around Zhoushan Island. Journal of Zhejiang Ocean University (Natural Science) 24(4):364-376.

Zheng, L. 2014. A Preliminary Study on Fish Community Structure in Yellow River Estuary. Shanghai Ocean University.

Zhu, Y., C. Zhang, and Q. Cheng. 1962. Ichthyography of South China Sea. Science Press, Beijing.

Zhu, L. 1999. A Preliminary Survey of Fishes in the Littoral Waters of Penglai Shandong. Journal of Zhanjiang Ocean University 19(1):22-29.

Appendix C Data collection and collation for Chapter 3

This appendix presents a detailed description about data collection and collation process. It contains the following information: 1) Model predictor data, 2) Seahorse location data, 3) Occurrence-data quality control, 4) Table C.1, and 5) References.

Model predictor data

I selected seven variables from twelve original ecological covariates relating to physiological suitability and primary productivity. Physiological suitability defines the physiological fitness of seahorses to its surroundings. I collected data for five environmental factors: depth, sea surface temperature (SST), salinity, dissolved oxygen, and pH. Primary productivity was used as an index of zooplankton availability for seahorse since the latter could not be obtained. I used surface chlorophyll a (Chlo-a) to estimate primary productivity, as they are usually highly correlated (Antoine et al., 1996). I only extracted data within 200 m depth as seahorses commonly live in shallow waters. I converted all these covariates to maps with a resolution of 1 km (using the Cylindrical Equal-area Projection). For datasets with coarser resolutions, I assigned values to unknown pixels with weighted mean values at known pixels based on the Inverse Distance Weighting (Cheung et al., 2009). All these processes were manipulated in an ArcMap. I selected seven variables from these twelve covariates to minimize collinearity based on Pearson correlation tests ($|r| > 0.7$, Dormann et al., 2013). These seven predictors were depth, pH, salinity, sea-surface-temperature mean, sea-surface-temperature range, chlorophyll a mean, and chlorophyll a range. I extracted depth from the latest global bathymetry data (GEBCO_2014 Grid, www.gebco.net) and derived other data from Bio-ORACLE, a global marine environmental dataset (Tyberghein et al., 2012).

I then derived datasets of nine categories of macrohabitat. I first gathered existing global maps (polygons) of four macrohabitats - seagrass beds, mangroves, coral reefs, and estuaries – from the United Nations Environment Programme World Conservation Monitoring Centre (UNEP-WCMC, http://data.unep-wcmc.org/). I then gathered occurrence data (points) of five other habitats – soft corals, macro-algae, sponges, sea pens, and hydrozoa – from Global Biodiversity Information Facility (GBIF, www.gbif.org) and Oceanic Biodiversity Information System (OBIS, www.iobis.org). I chose these habitat-forming organisms based on the literature (Foster & Vincent, 2004) and information from iSeahorse (www.iseahorse.org) – a global citizen-science platform for gathering seahorse sightings and their habitat information. To control quality in location uncertainty/errors, I only used data that had been collected since 1950 for species (of the five taxa) with tube- or branch-like features (see details in Table C.1). I chose 1950 as it corresponds to the time when SCUBA diving became effectively used for underwater observations. This time period (1950s afterward) is also approximately consistent with other model-predictor datasets and the collected seahorse occurrences.

For each of the nine macrohabitat categories, I created two types (binary vs. continuous) of variables. The binary variable was distributions of presences/absences (1/0) of each habitat. The continuous variable was transformed from binary predictor by using the Euclidean distance algorithm – distance to the nearest habitat. All these processes were conducted in an ArcMap.

Theoretically, compared with the binary variables, continuous variables could be more resilient to some extent of mismatch between species occurrences and habitat locations due to poor resolution or locational uncertainty. For example, suppose one species was correlated with habitat A, but due to resolution issues many of their (i.e. species and habitats) occurrences were not matched in the same pixels (but very close to each other). Then the model based on binary predictors might detect no relationship or even a false-negative relationship between the species and habitat A; while using the continuous predictors could still detect a tendency of species locating close to habitat A. Therefore, I expected that continuous variables would better inform species-habitat relationships.

Seahorse location data

I collected geo-referenced sightings of seahorses from eight databases: 1) GBIF, 2) OBIS, 3) iSeahorse, 4) FishNet2 (www.fishnet2.net), 5) FishBase (www.fishbase.org), 6) Reef Life Survey (www.reeflifesurvey.com), 7) unpublished data from our Project Seahorse team’s previous studies, and 8) data in peer-reviewed literature (PRL) drawn from the Zoological Record, Web of Science.

Data in GBIF and OBIS came originally from various common sources, including museum collections and human observations. These human observations included data from citizen science databases, such as iNaturalist (www.inaturalist.org) and Diveboard (www.diveboard.com). The FishNet2 is a global database of fish collections in natural history museums, universities, and other institutions. FishBase is a global biodiversity information system on finfishes and sometimes also contains occurrences derived from other sources. iSeahorse is a global citizen-science platform developed specifically for collecting seahorse sightings, mainly from recreational divers and researchers. Reef Life Survey is another citizen-science database for collecting distribution data of reef species from recreational divers.

To collect seahorse occurrence data from peer-reviewed literature, I searched published papers related to seahorses in Zoological Record database in the Web of Science (http://apps.webofknowledge.com/). I used “seahorse” OR “sea horse” OR “Hippocampus” NOT “brain” NOT “neuro” NOT “cell” as the Topic words to search in publications within the default year range (1864 – 2016). This search was conducted on August 30, 2016, and it resulted in a total of 994 articles. I then get access to the full text of each article and read the title and abstract first to quickly to get an idea if there is a possibility to find seahorse occurrences in the article. Papers about underwater surveys, species checklists, new records, and related fisheries studies have higher chances to contain detailed information about locations than studies focusing on genetics, aquaculture, medical use and trade. To facilitate the mining process in each paper, I used the key words “longitude”, “latitude”, “location”, “study site”, and “study area” to locate information about seahorse occurrences. I also checked the study-area map to find if there were any seahorse occurrence points that can be georeferenced from the map. These approaches were proved as helpful to save the time.

To examine the utility of data from different sources, I then divided speices occurrences into three subsets: research-grade data, citizen-science data, and museum-collection data.

Research-grade (RG) data were those sourced from biologists, including (i) Project Seahorse's unpublished data, (ii) PRL, and (iii) 'research-grade' data in iSeahorse. The 'research-grade' data in iSeahorse are those occurrences validated by our Project Seahorse or provided directly from seahorse researchers. Citizen-science (CS) data were derived from (i) Reef Life Survey, (ii) GBIF data sourced from Diveboard (www.diveboard.com), and (iii) 'casual observations' in iSeahorse, which were from recreational divers. I found that museum collection (MC) data had ambiguous quality and could not easily be categorized as either RG or CS data. Instead, I treated them as an independent dataset and aimed to compare its utility with RG and CS data.

Occurrence-data quality control

The species data derived from multiple sources inevitably included some errors, sampling bias, and spatial autocorrelations (Guisan & Thuiller, 2005; Dambach & Rödder, 2011). For instance, data collected by citizen scientists could be biased to tourism sites. Taxonomic identification could be incorrect if the person who examined the specimen were not well-trained. Spatial errors or location uncertainty could also exist if the organisms were not well documented as, for example, when seahorses were not the target species for research or monitoring.

I screened the collected species data by checking nomenclature and spatial errors and reduced spatial autocorrelations by spatial filtering for each species. First, I carefully checked species identifications of all iSeahorse sightings based on the available online photographs provided by the citizen scientists. To validate data from other databases, I adapted the instruction from Graham et al. 2004. I visually compared occurrence positions in an ArcMap to identify 'outliers' that were apparently isolated from other localities or far beyond the known geographic ranges based on the authors' knowledge and Lourie et al. (2016). For the suspected outliers, I then checked the nomenclature and positions by tracking back to their original sources (e.g. museum specimens), where possible. Due to resolution issues of the bathymetry map, some occurrences may fall in 'no-data' cells of the bathymetry map, but they are still located in the sea. I filled these gaps in the bathymetry map (and all other predictor datasets) by assigning the value of their nearest cells or the mean if two or more cells were equidistant (Marcer et al., 2013). In other cases, some occurrences may fall on the coast close to sea. Given the rarity of each sighting, I moved those occurrences within 5 km of the shore to the nearest pixel of bathymetry map. I designated occurrences that were farther inland as errors and deleted them.

Table C.1. Description of the five biotic habitats used in modeling seahorse distributions. Note that only those taxa with a 'tube shape' or 'branch shape' were used to build the dataset, given seahorses require holdfasts to anchor themselves on the seafloor.

Biotic habitats	Description
soft corals	occurrences of Alcyonacea (sea fans), Isididae, Helioporacea (blue corals), Pseudopterogorgia and Antillogorgia (sea plumes).
macroalgae	occurrences of *Ahnfeltia plicata*, *Pyropia columbina*, Porphyra, Caulerpa, Ceramium, Codium, Corallina, Thalassodendron, Cymodocea, Furcellaria, Gelidiella, Chondrus, Osmundea, Laurencia, Kappaphycus, and Eucheuma.
sponges	occurrences of Aiolochroia crassa, Aplysina, Suberea spp., Agelas, Ircinia felix, Smenospongia aurea, Cervicornia cuspidifera, Axinella, Ptilocaulis, Svenzea zeai, Callyspongia, Haliclona, Amphimedon, Niphates erecta, Neopetrosia subtriangularis, Monanchora arbuscula, Desmapsamma anchorata, Clathria (Thalysias) curacaoensis, and Leucosolenia variabilis.
sea pens	occurrences of Pennatulacea.
hydrozoa	occurrences of Obelia, Tubularia, Sertularia argentea, Kirchenpaueriidae, Pennaria disticha, Aglaophenia pluma, and Thyroscyphus ramosus.

References cited in Appendix C

Antoine, D., André, J.M. & Morel, A. (1996) Oceanic primary production: 2. Estimation at global scale from satellite (coastal zone color scanner) chlorophyll. Global biogeochemical cycles, 10, 57-69.

Cheung, W.W., Lam, V.W., Sarmiento, J.L., Kearney, K., Watson, R. & Pauly, D. (2009) Projecting global marine biodiversity impacts under climate change scenarios. Fish and fisheries, 10, 235-251.

Dambach, J. & Rödder, D. (2011) Applications and future challenges in marine species distribution modeling. Aquatic Conservation: Marine and Freshwater Ecosystems, 21, 92-100.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . Leitão, P.J. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36, 27-46.

Foster, S. & Vincent, A. (2004) Life history and ecology of seahorses: implications for conservation and management. Journal of fish biology, 65, 1-61.

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

Lourie, S.A., Pollom, R.A. & Foster, S.J. (2016) A global revision of the Seahorses *Hippocampus* Rafinesque 1810 (Actinopterygii: Syngnathiformes): Taxonomy and biogeography with recommendations for further research. Zootaxa, 4146, 1-66.

Marcer, A., Sáez, L., Molowny-Horas, R., Pons, X. & Pino, J. (2013) Using species distribution modelling to disentangle realised versus potential distributions for rare species conservation. Biological Conservation, 166, 221-230.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F. & De Clerck, O. (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. Global Ecology and Biogeography, 21, 272-281.

Appendix D Supplementary methods for Chapter 3

This appendix presents a detailed description about model construction and data analyses process. It contains the following information: 1) Model settings, 2) Model comparisons, 3) Model Group 1, 4) Model Group 2, 5) Ranking permutation importance, 6) Calculating geographic metrics for IUCN Red List assessment, and 7) References.

Model settings

In model-comparison group 1, for each species, I used the defined geographic range (i.e. modeling envelope, see 2.3 in the manuscript) to clip model predictor datasets. Given H. denise and H. bargibanti were known to be highly associated with particular taxa of gorgonian corals (Gomon, 1997; Lourie & Randall, 2003), I used a predictor based on such corals to replace the soft-coral predictor in their models. Highly correlated ($|r| > 0.7$) predictors in each dataset were excluded based on the Pearson correlation test. I chose 10-fold cross-validation to evaluate each model (Merow et al., 2013). The cross-validation split species data into 10 independent subsets and the model was then trained and evaluated 10 times. For each time, one of the 10 subsets was held out in turn as test data and the model was trained with the rest 9 subsets. I did group 1 only on species which had at least 50 sightings. This manipulation ensured at least 5 occurrences were included each time in the test dataset for cross-validation, as a manner to derive reliable statistics.

In model-comparison group 2, I used only data from species which have relatively sufficient sightings ($n \geq 30$) from each category (CS/RG/MC) to construct models. I used the same amount of species data from each category to train each model, as a manner to eliminate the impact of data availability on model performance (Wisz et al., 2008). The number of training data was determined by data availability of CS of the modeled species, as CS had the lowest number of sightings. For the modeling group based on CS data, I used all sightings to train the models (hereafter CS models). For the modeling groups of RG and MC data, I randomly selected a subset of sightings, as many as CS, to train the models (hereafter RG models and MC models). The remaining set of RG data was used to test performance (i.e. presence predictive accuracy) of each model including those based on dataset combinations, given their known quality.

To eliminate the sampling bias in presence data, I derived pseudo-absences with the same sampling bias. To do so, I created a 'bias file' for each species based on a 'congeneric-species (or target-group) absences' approach (Phillips et al., 2009; Mateo et al., 2010; Chapter 2). This approach limits Maxent to randomly select pseudo-absences within a buffer area (radius = 30 km) centered around the presences of all congeneric species (i.e. all seahorse species) located in the defined geographic range (Chapter 2). By doing so, the pseudo-absences would contain the same bias as in the presences. Therefore, the model will predict species presences in the locations that are truly suitable instead of areas that are heavily sampled. To diminish spatial autocorrelations among occurrences without sacrificing data too much, I rarefied the sightings with a threshold distance of 2 km between every two points. I used the SDMtoolbox in an ArcMap to fulfill the above process (Brown, 2014).

To set up each model, I used the 'Cloglog' transformation as the output format, and selected all features to fit the model. To eliminate the sampling bias in presence data, I selected the 'bias file' as I generated above in ArcMap to derive pseudo-absences with the same sampling bias (Philips et al., 2009; Mateo et al., 2010; Chapter 2). I picked the regularization multiplier to control for over-parameterization (Crall et al., 2015; Chapter 2). Other default settings in Maxent were not changed as they have performed well (Phillips et al., 2006).

Model comparisons

I used two statistics to measure model performance (i.e. ability to discriminate presences and absences) respectively for the above two comparison groups. Model performance in comparison group 1 was measured with the area under the curve (AUC) of the receiver-operating characteristic (ROC) plot (Hanley & McNeil, 1982), which is one of outputs of the Maxent (i.e. test AUC). The AUC of the ROC plot measures model's general accuracy of both presence and absence predictions. The value of AUC ranges from 0 to 1 with a qualitative description of model accuracy: 0.90 – 1.00 excellent, 0.80 – 0.90 good, 0.70 – 0.80 fair, 0.60 – 0.70 poor, 0.50 – 0.60 fail (Swets, 1988). The second statistic was Sensitivity (accuracy of predicting presences, Altman and Bland, 1994), which is a threshold-dependent measure of prediction accuracy. I used the threshold that maximized the sum of training sensitivity and specificity (hereafter, MSS threshold) to derive Sensitivity (Liu et al., 2013). I used Sensitivity rather than AUC to measure model performance in comparison group 2, given that model predictions were tested with only species occurrences.

Model Group 1: the effect of habitat predictors

To examine the utility of habitat predictors in comparison group 1, I calculated AUC and another two statistics: Habitat Variable Importance (HVI) and Predicted Area Ratio (PAR). They are all Maxent's outputs averaged across 10 cross-validations. Habitat Variable Importance (HVI) was the sum of permutation importance (range from 0 to 100) of each habitat covariate included in the model. Predicted Area Ratio (PAR) was the proportion (0 - 1) of the predicted area (based on the MSS threshold) to the modeling envelope. I used paired Wilcoxon-rank-sum tests to compare AUC and PAR among the three predictor datasets across species, and compare HVI between Dataset 2 and Dataset 3. HVI was not applicable for Dataset 1 as it didn't include habitat covariates. My expectations were: 1) models based on Dataset 2 and 3 (which included habitat predictors) could derive higher AUC than models based on Dataset 1 (non-habitat predictors), 2) Dataset 3 (using continuous habitat predictors) would derive higher AUC than Dataset 2 (using binary habitat predictors), 3) models integrated habitat predictors would derive smaller PAR than models only integrated non-habitat predictors, 4) models based on Dataset 3 would result in higher HVI than models based on Dataset 2.

Model Group 2: the effect of species-data sources

I compared model performance and predictions among research-grade (RG), citizen-science (CS) and museum-collection (MC) datasets. Model performance was measured with Sensitivity. Model prediction similarity were measured with two spatial agreement statistics: I Similarity Statistic (Warren et al., 2008) and Presence Agreement (Chapter 2). The I Similarity is a metric usually applied to measure the absolute spatial agreement between two probability predictions. The Presence Agreement was the ratio of overlapping area of two presence predictions to their spatial union. Both statistics range from 0 (no overlap) to 1 (identical). I calculated the I Similarity Statistic with the package 'ENMeval' in R (Muscarella et al., 2014). I derived the presence predictions from Maxent probability outputs based on the MSS threshold, and calculated Presence Agreement based on the analyses in an ArcMap. In addition, I compared Sensitivity among four models: RG model, RG+CS model, RG+MC model, and ALL model. My expectations were that 1) models based on different datasets could differ on model performance and predictions, 2) RG models would have higher performance (i.e. Sensitivity) than CS and MC models, and 3) adding CS and/or MC to RG would improve model performance.

Ranking permutation importance

The ranking of permutation importance of variables starts from 1 (the highest value) and followed with continuous integers (2, 3, 4, …). For variables with same permutation importance, I assigned the same rank to them. For instance, suppose I have five variables, and their permutation importance are 30, 20, 20, 15, 15. Then, the ranks of these variables are respectively 1, 2, 2, 3, 3. Given some habitat types (e.g. hydrozoa) were absent in the modeling envelope of some species, these predictors were then not used in these models. I used zero to reflect the importance of these habitats, and thus assigned the lowest rank (i.e. largest integer) to them. For the predictor that was excluded because of collinearity, I assigned it the same rank of its covariate, which was most correlated with it (i.e. the highest Pearson correlation coefficient) and was included in the model. For instance, if sea surface temperature mean and its range are highly correlated, and only the mean is used in the model, I then assign the same rank of the mean to its range. I did this based on my initial modeling trials in which I replaced the variable with its highly-correlated one and derived almost the same rank for the latter. This is understandable as highly-correlated variables may contribute very similar information to the model. Therefore, I think it is cost-effective to directly assign the same rank to the highly correlated variables.

Calculating geographic metrics for IUCN Red List assessment

To evaluate conservation status, I first created the minimum convex polygon for each species based on species occurrences to derive the extent of occurrence (EOO, IUCN Criterion B1) at the 1 km resolution. I then applied the threshold that maximized the sum of training sensitivity and specificity to derive predictive presences from the presence-probability map of each species (Liu et al., 2013). This allowed models to cut off the continuous probability into presence/absence map with the threshold at which true presence rate and true absence rate are maximized, at the sacrifice of some presence omissions. I think it is a reasonable choice as there might be spatial uncertainties/errors hidden in my data. I then derived the area of occupancy (AOO, IUCN Criterion B2) by extracting the predicted presences from the EOO map (Marcer et al., 2013). The AOO was calculated at 2 km resolution as recommended by IUCN (IUCN Standards and Petitions Subcommittee, 2017). I undertook the above analyses in an ArcMap following the IUCN's guidelines (IUCN Standards and Petitions Subcommittee, 2017).

References cited in Appendix D

Brown, J.L. 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods in Ecology and Evolution, 5, 694-700.

Crall, A.W., Jarnevich, C.S., Young, N.E., Panke, B.J., Renz, M. & Stohlgren, T.J. 2015. Citizen science contributes to our knowledge of invasive plant species distributions. Biological Invasions, 17, 2415-2427.

Gomon, M.F. 1997. A remarkable new pygmy seahorse (Syngnathidae: *Hippocampus*) from Southeastern Australia, with a redescription of H. bargibanti Whitley from New Caledonia. Memoirs of the Museum of Victoria, 56, 245-253.

Hanley, J.A. & McNeil, B.J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology, 143, 29-36.

IUCN Standards and Petitions Subcommittee. 2017. Guidelines for Application of IUCN Red List Criteria at Regional and National Levels. Version 4.0.

Liu, C., White, M. & Newell, G. 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. Journal of biogeography, 40, 778-789.

Lourie, S.A. & Randall, J.E. 2003. A new pygmy seahorse, *Hippocampus denise* (Teleostei Syngnathidae) from the Indo-Pacific. ZOOLOGICAL STUDIES-TAIPEI-, 42, 284-291.

Mateo, R.G., Croat, T.B., Felicísimo, Á.M. & Munoz, J. 2010. Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. Diversity and Distributions, 16, 84-94.

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

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

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

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

Swets, J.A. 1988. Measuring the accuracy of diagnostic systems. Science, 240, 1285-1293.

Warren, D.L., Glor, R.E. & Turelli, M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution, 62, 2868-2883.

Wisz, M.S., Hijmans, R., Li, J., Peterson, A.T., Graham, C. & Guisan, A. 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions, 14, 763-773.

Appendix E Supplementary results for Chapter 3

This appendix presents supplementary tables and figures for the Results section in the manuscript. It contains a total of 5 tables (Tables E.1 – 3.5) and 13 figures (Figures E.1 – 3.13).

Table E.1. Summary statistics of the 16 model predictors across all seahorse species globally.

Predictor	Minimum	Maximum	Mean	Median
depth (m)	1	199	25	13
pH	7.127	8.461	8.191	8.21
salinity (PSS)	16.5	41.41	34.69	35.08
sea surface temperature mean (°C)	7.852	31.16	24.653	26.076
sea surface temperature range (°C)	0.654	25.885	7.476	7.052
Chlorophyll a mean (mg/cm^3)	0.072	50.287	3.02	1.155
Chlorophyll a range (mg/cm^3)	0.025	38.602	3.303	1.3
distance to nearest coral reef (km)	0	5815.85	620.14	643.03
distance to nearest estuary (km)	0	2169.51	128.66	62.46
distance to nearest hydrozoa (km)	0	6889.8	1182.3	662.3
distance to nearest macroalgae (km)	0	1281.07	114.02	50.17
distance to nearest mangrove (km)	0	4571.54	360.46	18.76
distance to nearest sea pen (km)	0	2547.57	316.99	182.92
distance to nearest soft coral (km)	0	1615.95	50.898	19.966
distance to nearest sponge (km)	0	321.844	18.03	6.646
distance to nearest seagrass bed (km)	0	2662.74	74.9	21.4

Table E.2. A summary of area under the curve (AUC, model predictive measure), habitat-variable importance (HVI, cumulative permutation importance of all macrohabitat variables), and key macrohabitats (positively correlated with the seahorse species, listed in order of permutation importance), ranked by HVI.

Species	AUC	HVI	Key macrohabitats (in the order of importance)
H. denise	1.00	99.8	soft coral (gorgonians)
H. bargibanti	0.99	99.4	soft coral (Muricella)
H. zebra	0.71	97.2	macroalgae, coral reef, soft coral
H. plantifrons	0.71	86.0	macroalgae, seagrass
H. sindonis	0.92	84.3	sponge, soft coral, hydrozoa
H. minotaur	0.87	81.4	hydrozoa, sponge, mangrove
H. kelloggi	0.78	80.4	sponge, coral reef
H. capensis	0.99	77.4	estuary, macroalgae
H. erectus	0.76	64.7	sponge, hydrozoa, soft coral, mangrove
H. comes	0.89	61.4	sponge, soft coral, macroalgae
H. jayakari	0.81	60.7	seagrass
H. kuda	0.82	59.3	sponge, sea pen, seagrass, soft coral
H. dahli	0.73	58.7	sponge, soft coral
H. abdominalis	0.80	58.4	estuary, sponge, hydrozoa, macroalgae
H. ingens	0.78	53.6	sponge, hydrozoa, coral reef, mangrove
H. subelongatus	0.95	52.9	sea pen/coral reef, sponge/soft coral/macroalgae
H. trimaculatus	0.78	52.4	sea pen, soft coral
H. fisheri	0.70	51.0	sponge
H. angustus	0.73	37.2	sponge, sea pen, macroalgae, mangrove
H. camelopardalis	0.75	36.8	mangrove
H. pontohi	0.91	33.8	sponge
H. hippocampus	0.87	31.4	macroalgae, sponge, soft coral
H. spinosissimus	0.84	30.6	sponge, soft coral, coral reef, sea pen
H. coronatus	0.90	23.6	hydrozoa

Species	AUC	HVI	Key macrohabitats (in the order of importance)
H. whitei	0.92	21.9	mangrove, sponge
H. histrix	0.87	21.7	macroalgae, soft coral, sponge, coral reef, seagrass
H. barbouri	0.83	20.3	soft coral
H. guttulatus	0.83	19.6	macroalgae
H. reidi	0.89	19.5	sponge, macroalgae, sea pen
H. zosterae	0.97	15.1	mangrove, macroalgae, seagrass
H. breviceps	0.93	14.3	macroalgae, sponge, seagrass
H. patagonicus	0.78	6.2	macroalgae
H. mohnikei	0.89	3.9	sponge
H. casscsio	0.56	-	-

Table E.3. IUCN assessments and my estimate for 34 seahorse species that could be evaluated with geographic-range size. Note that my study only evaluated species against geographic-range size, and more information about fragmentation, fluctuation or decline were required to justify these categories (in bold). AUC, Area Under the Curve (i.e. model predictive accuracy); AOO, area of occupancy; EOO, extent of occurrence. IUCN categories: EN, Endangered; VU, Vulnerable; NT, Near Threatened; LC, Least Concern. The IUCN Criteria used to determine the conservation status were provided in the brackets (i.e. B1; B2): B1 refers to EOO thresholds, B2 refers to AOO thresholds. EOO ≤20,000 or AOO ≤2,000 km^2 corresponds to VU, EOO ≤5,000 or AOO ≤500 km^2 corresponds to EN.

Species	IUCN assessment	My geographic-range estimate			AUC
	Current category	Potential category	AOO/km^2	EOO/km^2	
H. capensis	EN	EN (B1; B2)	64	558	0.99
H. whitei	EN	-	57,356	9,799,059	0.92
H. barbouri	VU	-	4,152	1,976,877	0.83
H. comes	VU	-	29,880	5,297,952	0.89
H. ingens	VU	-	52,560	5,531,471	0.78
H. patagonicus	VU	-	176,284	1,549,217	0.78
H. erectus	VU	-	204,396	48,686,396	0.76
H. spinosissimus	VU	-	569,124	29,300,225	0.84
H. histrix	VU	-	652,468	208,501,354	0.87
H. kelloggi	VU	-	667,876	15,273,083	0.78
H. mohnikei	VU	-	909,216	14,968,166	0.89
H. kuda	VU	-	957,420	169,772,887	0.82
H. trimaculatus	VU	-	1,928,620	40,659,171	0.78
H. reidi	NT	-	398,832	35,618,844	0.89
H. fisher	LC	**EN** (B2)	280	40,382	0.70
H. sindonis	LC	**EN** (B2)	376	84,374	0.92
H. plantifrons	LC	**VU** (B2)	560	180,716	0.71
H. breviceps	LC	-	12,360	2,068,054	0.93

Species	IUCN assessment	My geographic-range estimate			AUC
	Current category	Potential category	AOO/km^2	EOO/km^2	
H. jayakari	LC	-	13,036	110,929	0.81
H. dahli	LC	-	13,836	266,333	0.73
H. abdominalis	LC	-	21,776	2,995,183	0.8
H. zosterae	LC	-	25,784	1,669,764	0.97
H. pontohi	LC	-	134,948	10,015,702	0.91
H. angustus	LC	-	573,012	9,045,917	0.73
H. subelongatus	DD	**EN** (B1; B2)	396	428	0.95
H. denise	DD	**VU** (B2)	832	10,483,181	1
H. coronatus	DD	-	5,968	174,828	0.9
H. minotaur	DD	-	6,824	828,499	0.87
H. zebra	DD	-	16,708	2,378,512	0.71
H. camelopardalis	DD	-	19,716	1,693,859	0.75
H. bargibanti	DD	-	24,804	14,882,812	0.99
H. guttulatus	DD	-	147,056	22,053,816	0.84
H. hippocampus	DD	-	226,100	17,401,651	0.86
H. casscsio	DD	-	-	79,952	0.56

Table E.4. IUCN assessments and my estimate on locations for 8 rarely-sighted seahorse species. Note that my study only evaluated species against the number of locations, and more information about population threats were required to justify these categories (in bold). IUCN categories: VU, Vulnerable; DD, Data Deficient. IUCN Criterion D2 here refers to 'number of locations ≤ 5'.

Species	IUCN assessment	My estimate on locations	
	Current category	Potential category	No. of locations
H. algiricus	VU	-	> 5
H. colemani	DD	VU (D2)	<5
H. debelius	DD	VU (D2)	<5
H. jugumus	DD	VU (D2)	<5
H. paradoxus	DD	VU (D2)	<5
H. pusillus	DD	VU (D2)	<5
H. satomiae	DD	VU (D2)	<5
H. tyro	DD	VU (D2)	<5

Table E.5. Matrix of permutation-importance rank of each predictor for the 33 seahorse species. Rank 1 was given to the highest value.

Species	Depth	sponge	SST mean	pH	macro -algae	SST range	man-grove	Chlo-a range	Chlo-a mean	coral reef	salinity	soft coral	sea pen	estuary	sea-grass	hydrozoa
H. abdominalis	1	4	9	9	5	7	9	6	8	9	9	2	11	3	10	5
H. angustus	5	2	8	1	6	8	7	8	8	8	8	8	3	4	7	8
H. barbouri	1	5	5	5	5	5	5	3	3	5	5	2	5	5	4	5
H. bargibanti	3	6	6	6	6	6	2	5	5	2	6	1	6	6	4	4
H. breviceps	1	3	6	6	2	4	10	8	10	10	9	7	10	6	5	10
H. camelopardalis	1	4	2	4	4	4	2	3	3	2	4	4	2	4	4	4
H. capensis	3	5	1	2	4	5	6	5	5	6	3	5	1	1	6	6
H. comes	2	1	10	6	4	8	10	5	5	10	9	3	11	7	7	11
H. coronatus	1	5	7	7	7	7	7	3	3	7	6	5	2	4	7	2
H. dahli	6	1	4	4	6	6	5	2	2	6	3	3	6	6	5	6
H. denise	3	4	4	2	4	4	4	4	4	4	4	1	4	4	4	5
H. erectus	2	1	6	10	10	10	6	8	8	6	4	5	7	9	10	3
H. fisher	1	2	4	3	6	7	5	7	7	3	2	7	2	2	3	5
H. guttulatus	1	4	2	3	8	6	13	10	9	13	11	12	13	5	7	2
H. hippocampus	1	4	2	10	3	2	2	7	8	11	9	6	11	5	8	2
H. histrix	2	8	9	1	4	3	13	6	6	10	5	7	10	12	11	14
H. ingens	1	3	12	11	13	4	9	10	10	7	11	12	2	8	6	5

Species	Depth	sponge	SST mean	pH	macro-algae	SST range	mangrove	Chlo-a range	Chlo-a mean	coral reef	salinity	soft coral	sea pen	estuary	seagrass	hydrozoa
H. jayakari	4	4	3	3	3	3	3	2	2	4	3	4	4	4	1	4
H. kelloggi	6	1	5	7	3	5	8	4	4	8	9	9	10	2	9	10
H. kuda	1	3	11	10	12	8	4	9	9	11	8	7	5	13	6	2
H. minotaur	2	6	9	9	9	8	3	4	5	9	3	9	3	9	7	1
H. mohnikei	4	6	3	2	11	3	3	1	1	3	5	9	7	10	8	11
H. patagonicus	5	6	1	4	1	1	1	6	2	1	1	6	6	3	1	7
H. plantifrons	5	5	4	4	1	3	5	5	5	2	5	5	5	4	2	6
H. pontohi	4	6	8	5	3	10	6	7	7	9	1	9	12	2	11	12
H. reidi	3	2	5	12	6	1	5	4	4	5	10	9	7	11	8	13
H. sindonis	8	1	3	3	6	3	3	5	5	3	2	4	7	10	9	3
H. spinosissimus	2	3	10	9	7	10	11	6	6	5	1	4	6	11	8	12
H. subelongatus	2	3	1	1	3	2	6	4	4	1	5	3	1	5	5	6
H. trimaculatus	3	5	13	2	1	11	9	10	10	13	12	7	4	6	8	14
H. whitei	1	7	2	6	5	9	3	4	4	2	11	11	10	8	11	12
H. zebra	8	7	6	6	1	8	8	8	8	4	8	5	3	8	2	9
H. zosterae	1	7	6	4	3	6	2	11	11	6	8	10	11	9	5	6
Mean rank	2.8	4.1	5.7	5.4	5.2	5.7	5.9	5.8	5.8	6.2	6.1	6.1	6.3	6.2	6.3	6.8

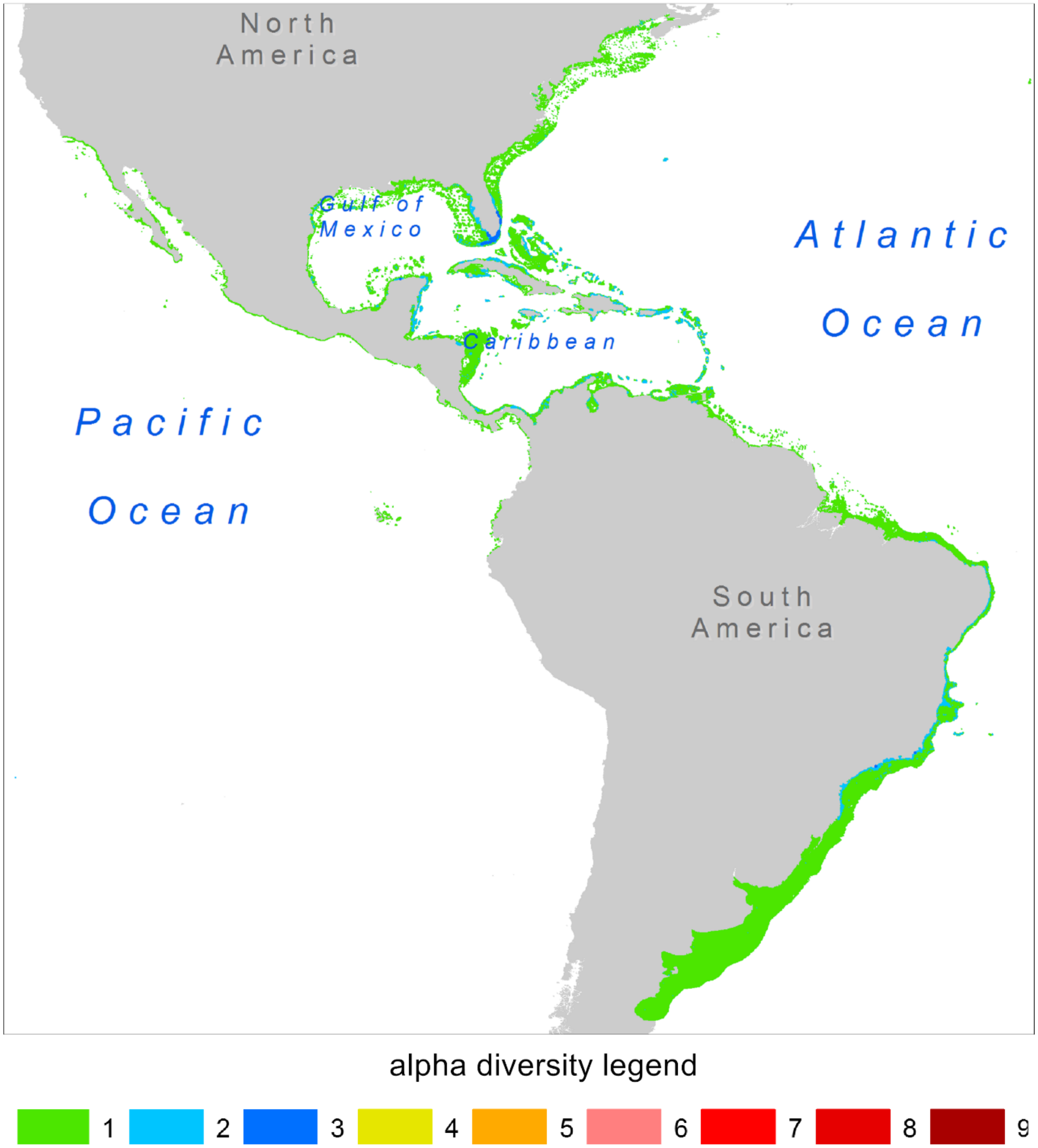


Figure E.1 Distribution map of species richness of seahorse species in America.

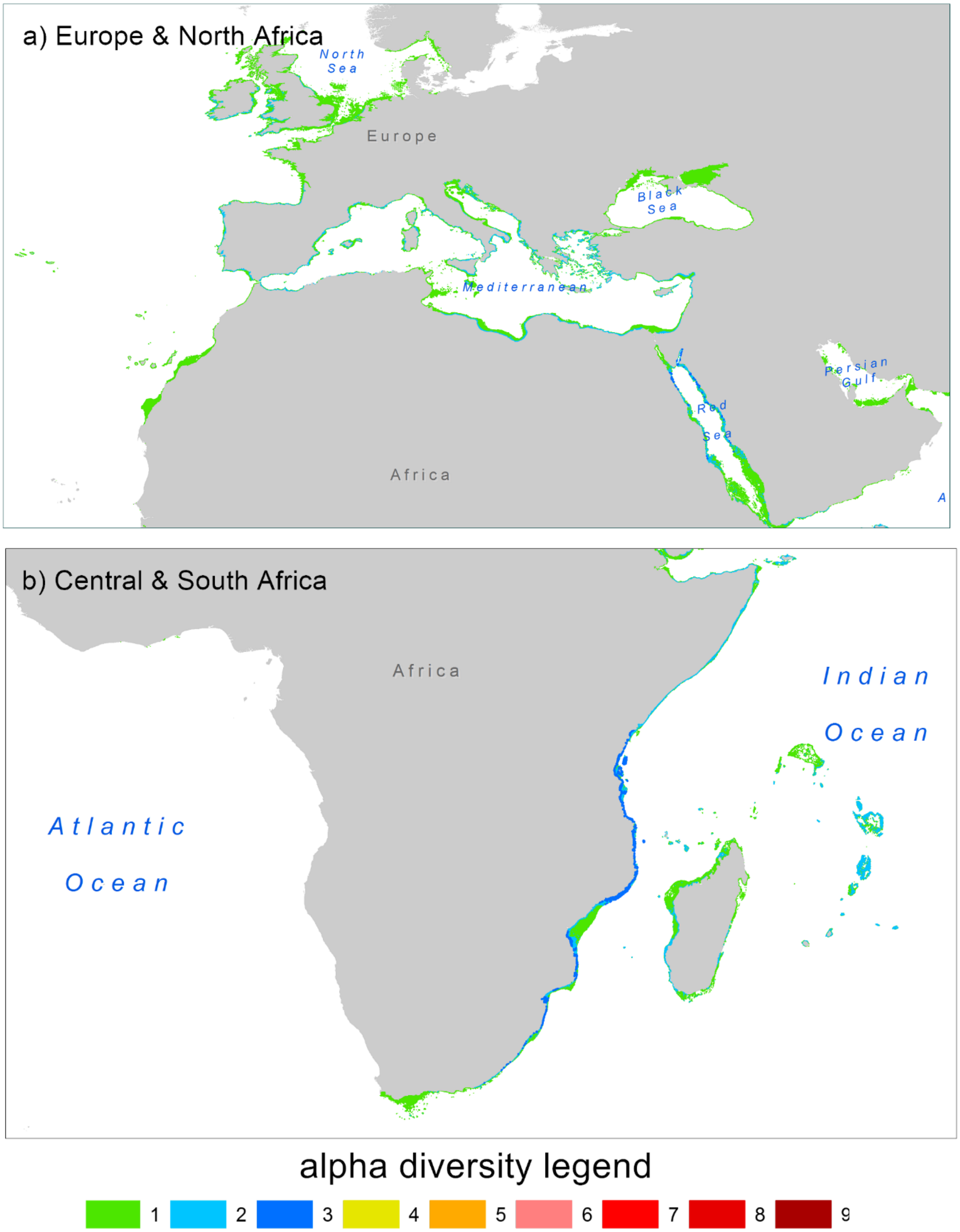


Figure E.2 Distribution map of species richness of seahorse species in Europe and Africa.

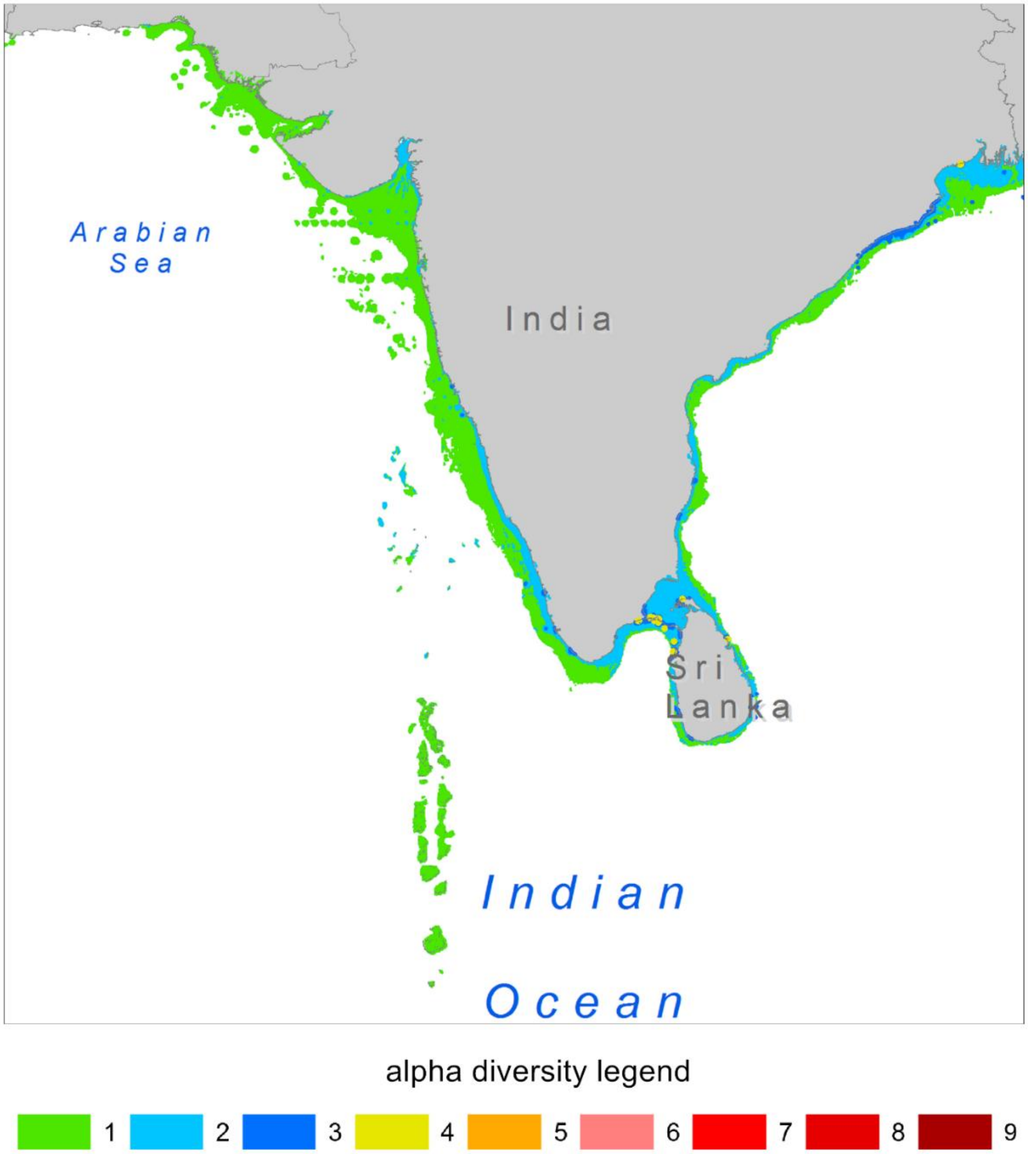


Figure E.3 Distribution map of species richness of seahorse species in India and its surrounding countries.

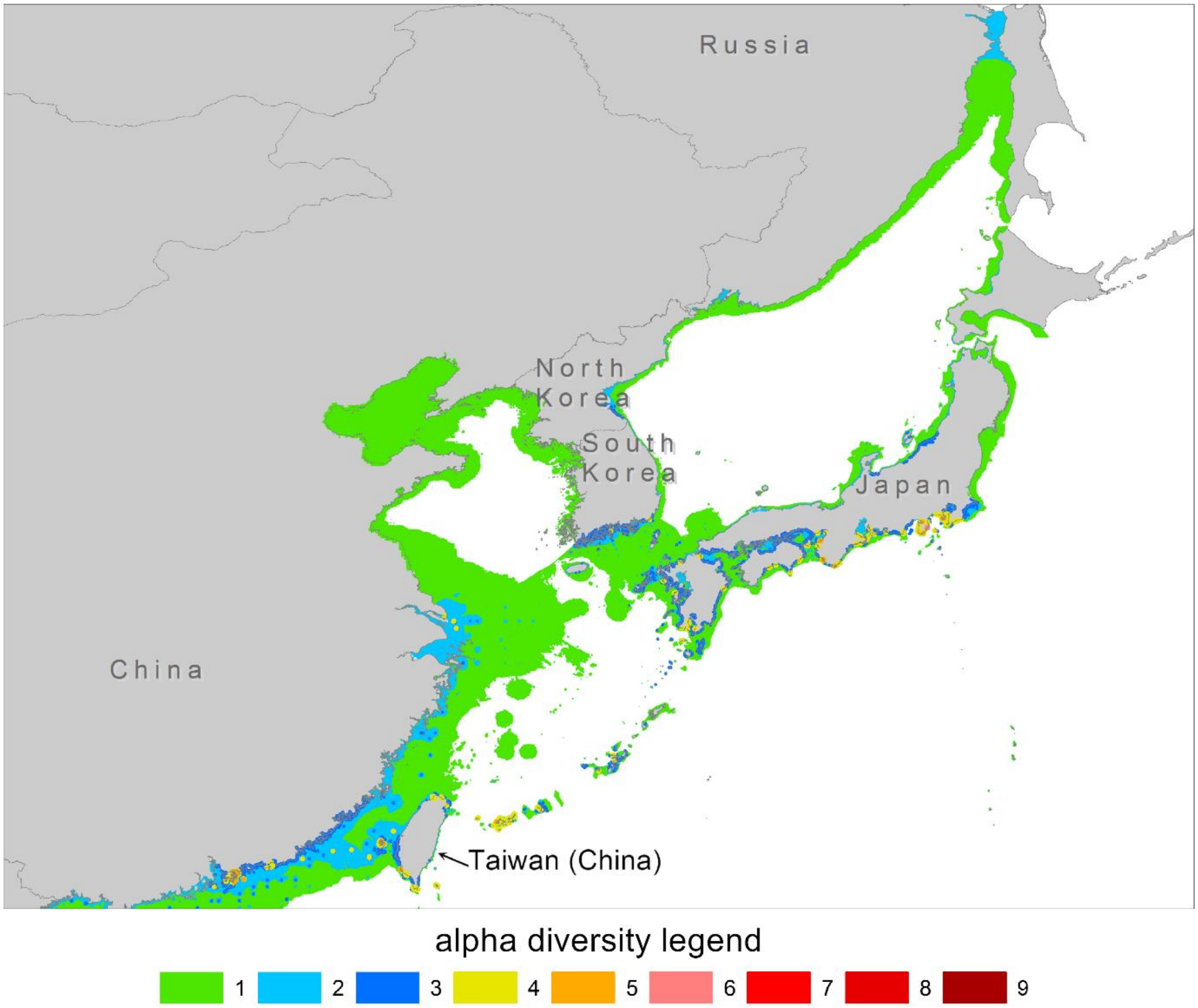


Figure E.4 Distribution map of species richness of seahorse species in East and North Asia.

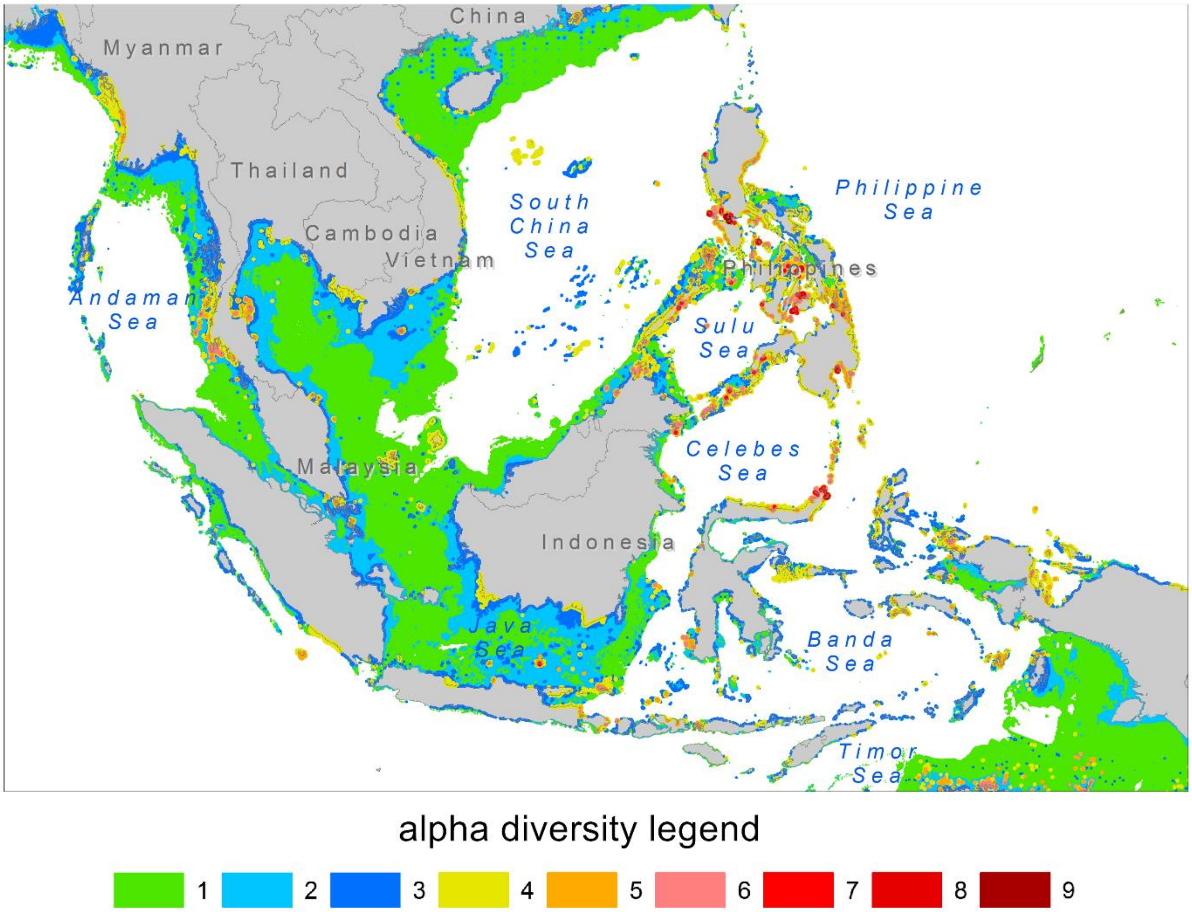


Figure E.5 Distribution map of species richness of seahorse species in Southeast Asia.

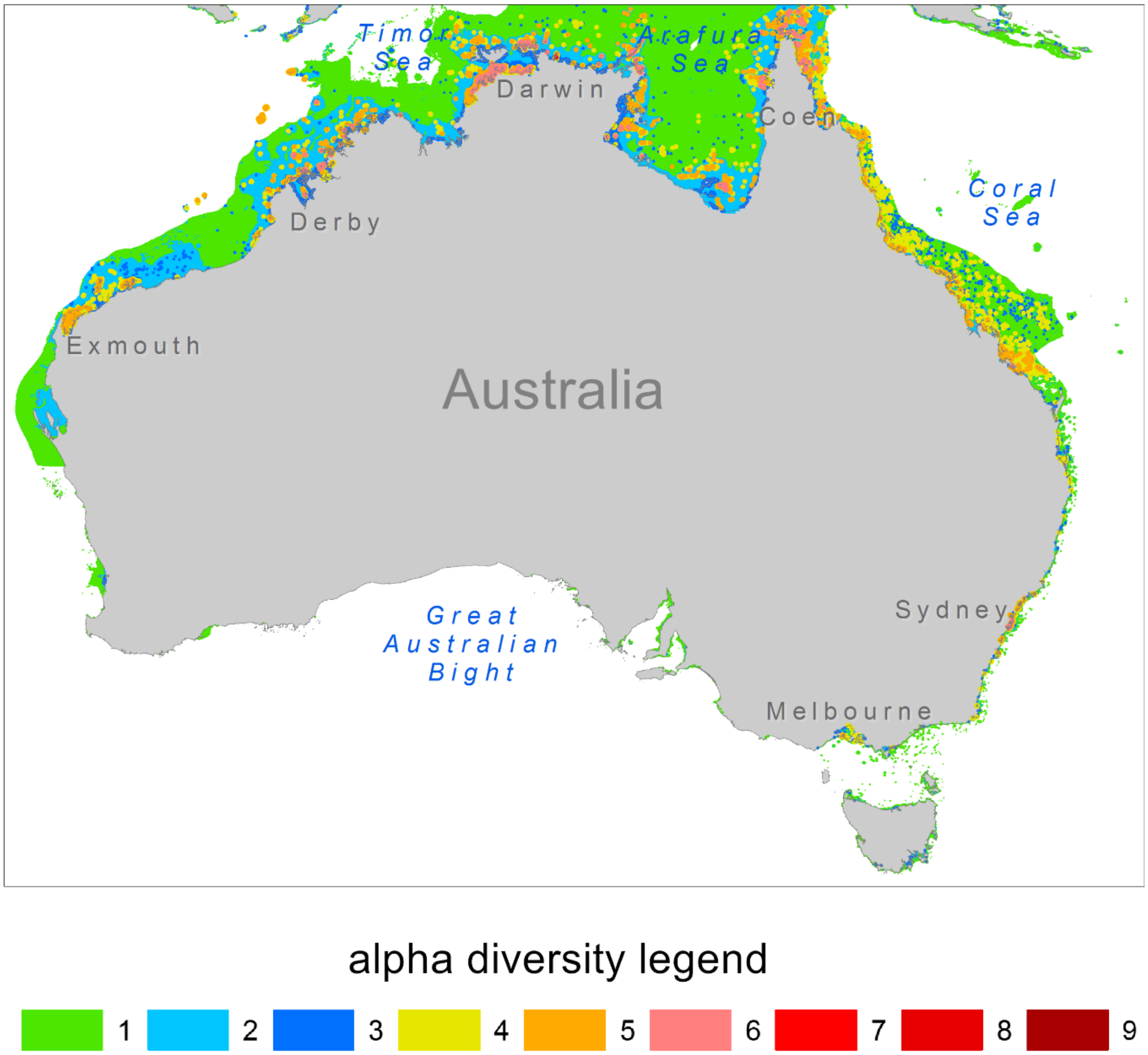


Figure E.6 Distribution map of species richness of seahorse species in Australia.

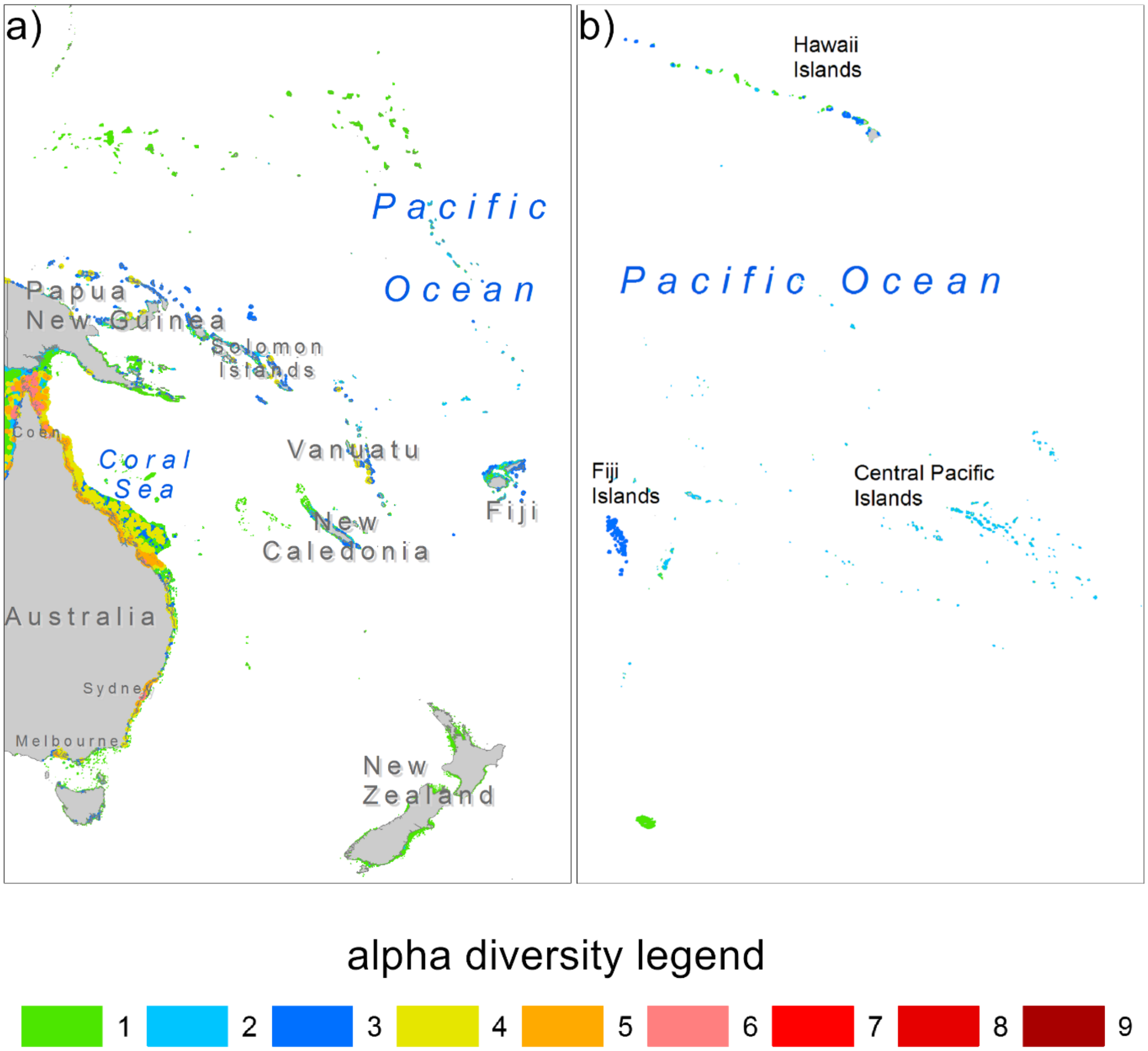


Figure E.7 Distribution map of species richness of seahorse species in Papua New Guinea, and the east of Australia (including Vanuatu, New Caledonia, New Zealand, and Fiji).

Figure E.8 Distribution map of seahorse biodiversity hotspots of species richness = 4.

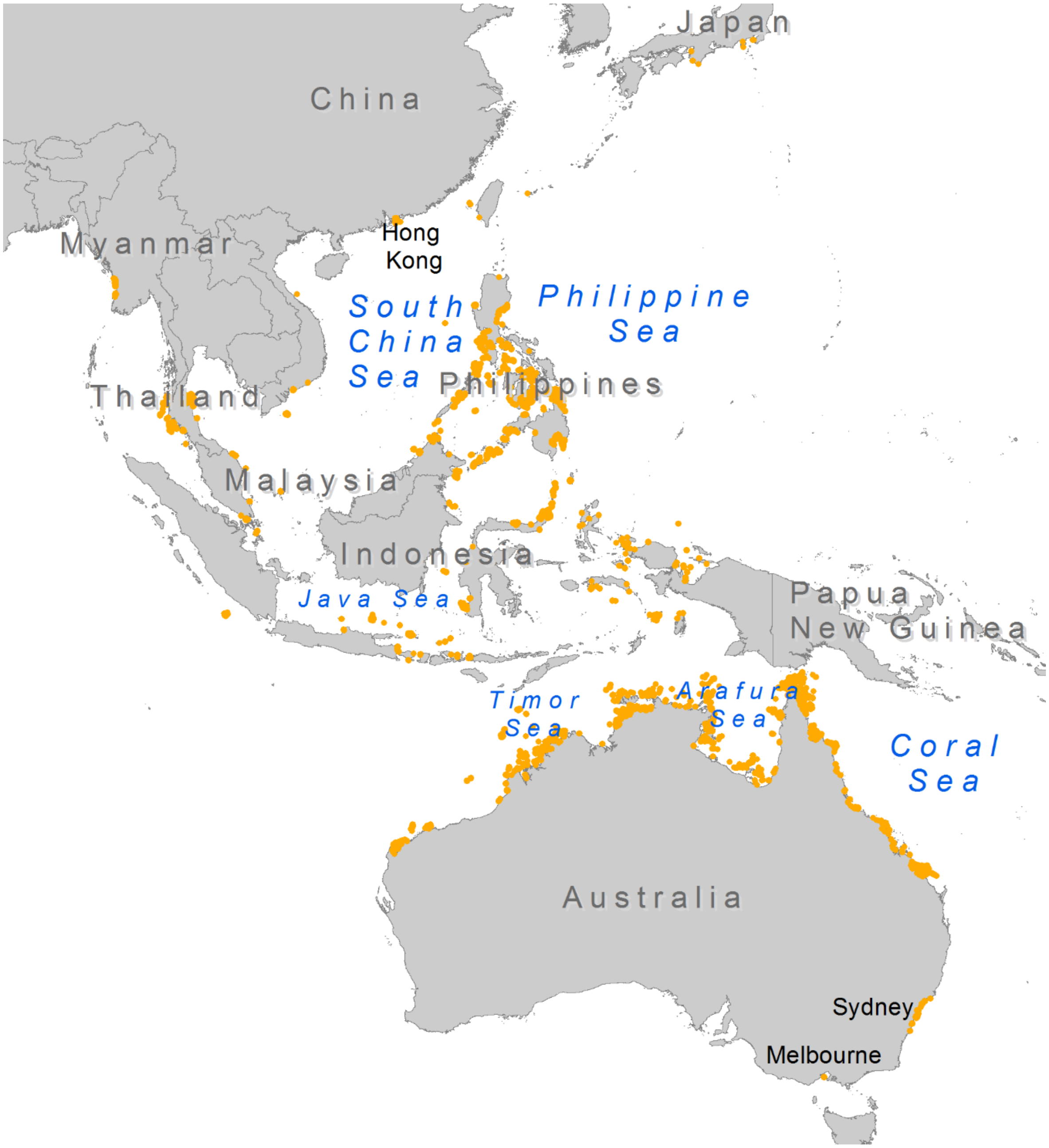


Figure E.9 Distribution map of seahorse biodiversity hotspots of species richness = 5.

Figure E.10 Distribution map of seahorse biodiversity hotspots of species richness = 6.

Figure E.11 Distribution map of seahorse biodiversity hotspots of species richness = 7.

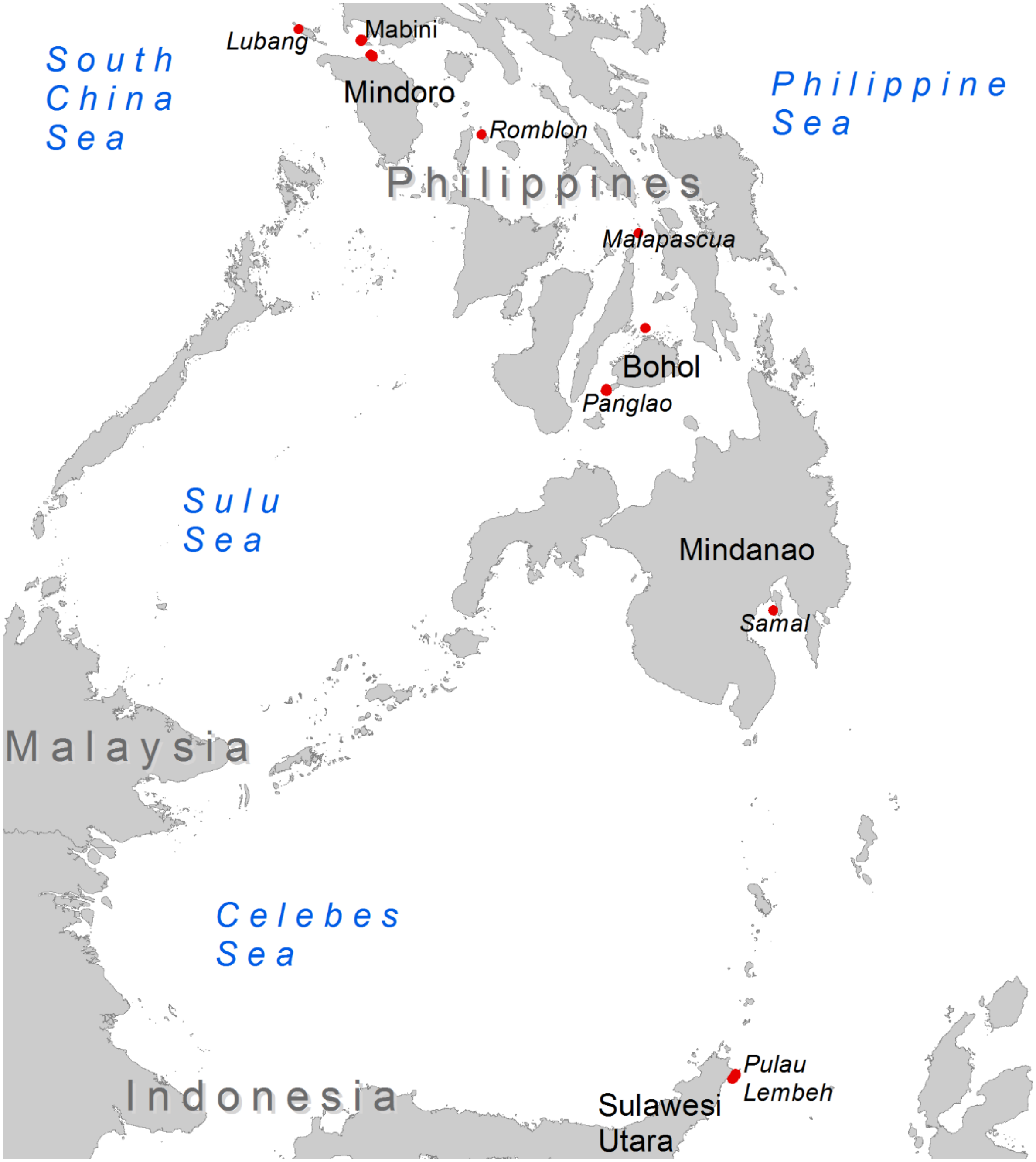


Figure E.12 Distribution map of seahorse biodiversity hotspots of species richness = 8.

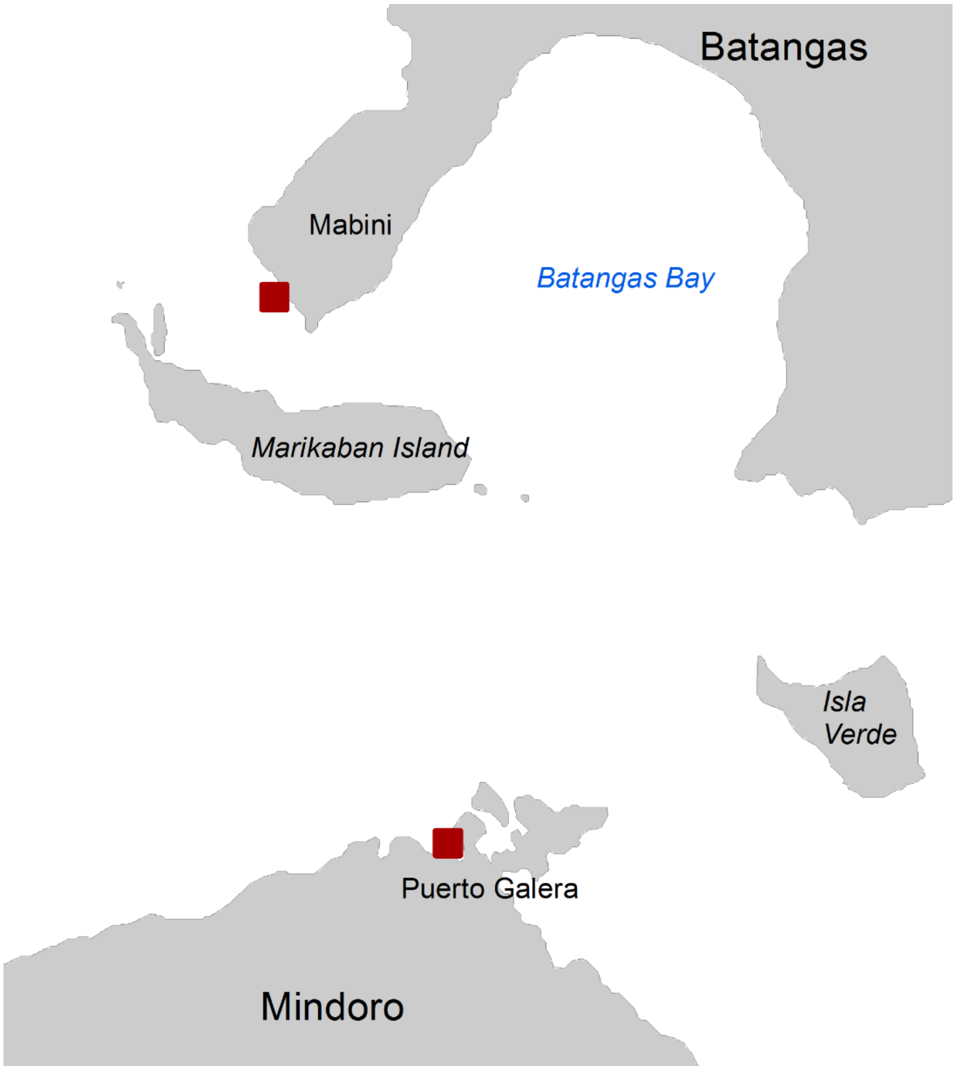


Figure E.13 Distribution map of seahorse biodiversity hotspots of species richness = 9.

Appendix F Supplementary discussion for Chapter 3

This appendix presents a supplementary discussion about species-habitat relations and the limitation of using these habitat data in species distribution models in the present study. A list of references is also provided.

Explanation for relationships between seahorse species and ecological variables

Interestingly, I found that habitat-forming sponges (i.e. those with branches or pipe-like shapes) were generally more vital than other habitat categories in predicting presences of seahorses. I think this result could be explained by at least three reasons. First, unlike other habitats (e.g. seagrass), these sponges are more widely-spread, structurally robust, and long-lived. These advantages may ensure them much more constantly available, which are vital for sedentary fishes like seahorses. Second, these sponges are structurally suitable for seahorses to use as holdfasts, shelters, and camouflage. Third, sponges might also be rich in prey provision for seahorses. For example, sponges are particularly rich in caridean shrimps of genus Synalpheus (Duffy, 1996), which have been commonly found in the diet of adult seahorses, including *H. abodominalis* (Woods, 2002) and *H. reidi* (Castro et al., 2008).

On the other hand, the generally low importance of other model predictors should be explained with caution. I found that the importance of hydrozoa was the lowest which was not surprising as hydrozoa were only present in the modeling envelopes of 13 out of the 33 species and thus likely less informative than many other predictors. The same reason might also explain the low rank of distance to the nearest sea pen, which was absent in 7 species' ranges. In addition, seagrass beds, estuaries, coral reefs, and soft corals were more widely present, but their importance ranks were low. This may suggest that these low-rank habitats were less important than those sponges for the 33 species, or they were superabundant and thus not the limiting factors for seahorse distributions (Aarts et al., 2013).

Two reasons might explain the results that most species were weakly or even negatively correlated with chlorophyll a mean. For one thing, the sea-surface primary productivity, as represented by chlorophyll a mean, might be weakly and sometimes even negatively correlated with the benthic food availability for seahorses. For another, food availability might not be a limiting factor for seahorses, given benthic zooplanktons are not likely scarce resources.

Limitations of including habitat data

It should be noted that the habitat categories used in my study were not exclusive. For instance, sometimes seahorses are also found in (sandy or muddy) open-bottom habitats, and some artificial features including harbor swimming nets and aquaculture farms (Dias et al., 2002; Foster & Vincent, 2004; Clynick, 2008; Aylesworth et al., 2016). I didn't include these habitats since they were unavailable for my global-scale modeling. But I encourage future SDM studies on seahorses at local or reginal scales to involve these habitat data if available to identify seahorse species-habitat relationships.

References cited in Appendix F

Aarts, G., Fieberg, J., Brasseur, S. & Matthiopoulos, J. 2013. Quantifying the effect of habitat availability on species distributions. Journal of Animal Ecology, 82, 1135-1145.

Aylesworth, L., Lawson, J., Laksanawimol, P., Ferber, P. & Loh, T.L. 2016. New records of the Japanese seahorse *Hippocampus mohnikei* in Southeast Asia lead to updates in range, habitat and threats. Journal of fish biology, 88, 1620-1630.

Castro, A.L.d.C., Diniz, A.d.F., Martins, I.Z., Vendel, A.L., Oliveira, T.P.R.d. & Rosa, I.M.d.L. 2008. Assessing diet composition of seahorses in the wild using a non-destructive method: *Hippocampus reidi* (Teleostei: Syngnathidae) as a study-case. Neotropical Ichthyology, 6, 637-644.

Clynick, B. 2008. Harbour swimming nets: a novel habitat for seahorses. Aquatic Conservation: Marine and Freshwater Ecosystems, 18, 483-492.

Dias, T.L., Rosa, I.L. & Baum, J.K. 2002. Threatened fishes of the world: *Hippocampus erectus* Perry, 1810 (Syngnathidae). Environmental biology of fishes, 65, 326-326.

Duffy, J.E. 1996. Species boundaries, specialization, and the radiation of sponge-dwelling alpheid shrimp. Biological Journal of the Linnean Society, 58, 307-324.

Foster, S. & Vincent, A. 2004. Life history and ecology of seahorses: implications for conservation and management. Journal of fish biology, 65, 1-61.

Woods, C.M. 2002. Natural diet of the seahorse *Hippocampus abdominalis*. New Zealand Journal of Marine and Freshwater Research, 36, 655-660.

Appendix G Supplementary methods for Chapter 4

This Appendix contains the following supplementary information for the methods of the manuscript: 1) further explanations of the cumulative-human-impact (CHI) models; 2) description of anthropogenic pressures and how their intensities were estimated; 3) calculating ecosystem-scale CHI for seahorse species, and 4) sensitivity analyses.

Further explanations of the CHI models

Based on literature review, I found that the human impact of a given stressor i upon a given species j can be estimated by three commonly used variables: (1) the intensity ($I_{i,j}$) of stressor i across species j's habitat (Halpern et al., 2008); (2) the exposure degree ($ED_{i,j}$) of species j to stressor i (Certain et al., 2015), and (3) the sensitivity ($S_{i,j}$) of species j to stressor i (degree of intolerance and incapability of recovery; Eno et al., 2013; Certain et al., 2015). Here, I defined the 'exposure degree ($ED_{i,j}$)' spatially and temporally: the proportion of the species' habitats experiencing the impact ($PS_{i,j}$), and the frequency of the impact ($F_{i,j,}$ how often the stressor j occurs to species i). The 'sensitivity ($S_{i,j}$)' was determined by the nature of the impact (e.g., direct vs. indirect) and how species would respond to it (e.g., directly death vs. consistent stress).

For the non-spatial model, I assumed that 1) there would be no human impact if any of the above variables (i.e., $I_{i,j}$, $PS_{i,j}$, $F_{i,j}$, $S_{i,j}$) became zero, 2) species' response to the impact was 'linear', and 3) interactions between different stressors' impacts were 'additive'. These assumptions were also common in previous CHI studies (Halpern and Fujita, 2013). I acknowledged that these assumptions might not be robust as 'non-linear' responses and 'non-additive' interactions commonly exist in nature (Crain et al., 2008; Coll et al., 2016), but this was understudied for seahorses. I made these assumptions to simplify the model and to examine whether a linear additive model is justifiable.

I then developed a spatially-explicit CHI model to map the cumulative human impacts similar to a previous study on marine mammals (Maxwell et al., 2013). I defined that, in a given location $\boldsymbol{m}$ (here, a 1 km by 1 km pixel), the impact of stressor $\boldsymbol{i}$ on species $\boldsymbol{j}$ was determined by three components: (1) the intensity ($I_{i,j,m}$) of stressor $\boldsymbol{i}$ in species $\boldsymbol{j}$'s habitat location $\boldsymbol{m}$, (2) the exposure degree ($ED_{i,j,m}$) of species ***j*** to stressor ***i*** in location ***m***, and (3) the species sensitivity ($S_{i,j}$, same as the non-spatial model). I defined the 'exposure degree' ($ED_{i,j,m}$) as the product of the presence probability of species ***j*** in location ***m*** ($P_{j,m}$) and the frequency of stressor ***i*** occurring upon species ***j*** in location ***m*** ($F_{i,j,m}$). Given that spatially-explicit maps of the frequency was unavailable in my study, I used the same frequency ($F_{i,j}$) as in the non-spatial model instead. By doing so, I assumed that the impact's frequency of a stressor would not differ spatially. This assumption seemed to be very unlikely but I made it for the merit of simplicity. I also made the other assumptions as I did for the non-spatial model.

Description of anthropogenic pressures and how their intensities were estimated in the study

I identified **12 different anthropogenic stressors** on seahorses. These stressors are: 1) demersal destructive fishing, 2) demersal, non-destructive, high-bycatch fishing, 3) pelagic high-bycatch fishing, 4) artisanal fishing, 5) habitat destruction caused by fishing, 6) coastal development, 7) (land-based) nutrient pollution, 8) ocean pollution, 9) noise pollution, 10) invasive species, 11) sea surface temperature abnormality, and 12) ocean acidification. They were derived from previous global-scale CHI study on marine ecosystems (Halpern et al. 2015). The dataset was modeled estimates and the normalized and log(X+1)-transformed values of intensity, rescaled from 0 – 1 by the variable's maximum value (in 2013, the latest version; Halpern et al. 2015). Such transformation and rescaling were commonly applied to mitigate the effects of outliers on impact estimations and to make different stressors comparable (Halpern and Fujita, 2013; Geldmann et al., 2014; Andersen et al., 2017).

These pressures were described under the framework of five major threats (Vincent et al., 2011). Detailed description is provided as follows.

Fishing

Fishing, or more precisely overfishing, is a major threat that caused the population decline of seahorses globally. Seahorses are typically found as bycatch in various commercial fishing gears including bottom (shrimp) trawls, gill and entangling nets, seines, and some small-scale fishing methods such as scoop net and traps (see a review in Vincent et al., 2011 and Lawson et al., 2017).

Here I estimated Fishing based on four types of anthropogenic pressures: 1) demersal destructive fishing, 2) demersal, non-destructive, high-bycatch fishing, 3) pelagic high-bycatch fishing, 4) artisanal fishing. *Demersal destructive fishing* contains bomb/chemical (all types), bottom trawl (all types), and dredge (all types). *Demersal, non-destructive, high-bycatch fishing* contains gillnet fixed, gillnet encircling, seine, beach seine, boat seine, trammel (all types), and trap (all types). *Pelagic high-bycatch fishing* contains driftnets and midwater trawl. *Artisanal fishing* refers to an estimate of traditional, small-scale fisheries that cannot distinguish between methods. I aim to use these four categories to estimate the vulnerability of seahorses to Fishing, as they represent the major fishing practices that can catch seahorses and have available data from Halpern et al. (2015).

Habitat change

Seahorses live in some of the world's most threatened habitats near shore. These habitats include seagrass beds, mangroves, coral reefs, sponge gardens, and macroalgae beds (Foster and Vincent 2004; Chapter 3). Habitat degradation/loss can certainly affect seahorse populations given that these habitats provide holdfasts, shelters, and food for them (see a review in Vincent et al., 2011). Human activities that can cause seahorse habitat change includes destructive fishing and coastal development, such as coastal engineering (e.g. marina development), boat anchoring, moorings, etc.

I use two datasets to represent habitat change. The first one is habitat destruction caused by fishing. The second one is coastal development, which is an integration of two sub-datasets – modelled coastal population density and intensity of oil rigs (Halpern et al. 2015). The habitat destruction caused by fishing was measured by the intensity data of demersal destructive fishing (Halpern et al. 2015). The population density dataset was assessed by calculating the population size within a moving circular window around an arbitrary focal coastal cell of radius 10 km on the basis of 2011 LandScan data (Halpern et al. 2015). The oil rigs dataset was estimated based on the stable lights at night in the seas from NOAA's National Geophysical Data Center, with ephemeral sources of lights (e.g. fires, mobile structures) eliminated (Halpern et al., 2015). Non-flare areas were also masked out from the oil rigs layers to correct the errors (Halpern et al., 2015).

Note: The impact of <u>demersal destructive fishing</u> SHOULD therefore be assessed TWICE based on the direct mortality of seahorses and its collateral damage on seahorse habitats, respectively.

Pollution

At least some seahorse species can be affected by pollution including chemical pollutants, eutrophication, hypoxia, and noise pollution (see a review in Vincent et al., 2011). Although examples of these impacts are rare in the field, there are experimental observations in aquaculture environment. For instance, lower light levels reduced prey capture rates by *H. erectus* (James and Heck, 1994). Hypoxic conditions arising from excessive fertilizer use in Chesapeake Bay, U.S.A., led to reductions in feeding by northern and dusky pipefishes (*S. fuscus* and *S. floridae*, respectively) that were predicted to affect health, growth and reproduction (Ripley & Foran, 2007). There are also evidences of hypoxia causing massive death of marine life such as in the Southeast Asia (Todd et al., 2010). Although seahorses may have high tolerance to hypoxia, there is a limit of tolerance for every species. Noise pollution may also affect seahorses, resulting from physiological stress effects (Anderson, 2009). Current studies have demonstrated that noise can increase respiration rate and reduce growth rate but may not necessarily reduce reproductive or feeding rates (Anderson, 2009; Magalhães, 2016).

To reflect the impact of pollution, I gathered data for two major types of pressures: 1) nutrient pollution (i.e. fertilizers), and 2) noise pollution (estimated by shipping intensity, fishing intensity (all kinds), oil rigs, plus human population density) from Halpern et al. (2015).

Invasive species

Invasive species may threaten seahorses through changing their habitats or competition and predation. For instance, fewer or no syngnathids were found in beds of invasive *C. taxifolia* than those in adjacent seagrass beds in estuaries in New South Wales, Australia (York et al., 2006). So far, the evidence of impact of invasion on seahorses is only speculated based on the suggested relation between the population decline of big-belly seahorses (*H. abdominalis*) and the invasion of invertebrates in Tasmania, Australia (Ross et al., 2004; Martin-Smith and Vincent, 2005).

I extracted the spatial data of invasive species from Halpern et al. (2015), which is an estimate based on cargo traffic at a port rather than the real distribution of invasive species.

Climate change

Climate change is an increasingly recognized threat on seahorses, as well as many other marine species. More frequent events of extreme temperatures and ocean acidification driven by climate change may impose instant and acute threats on seahorses. A combination of flooding and high littoral water temperatures of up to 32 ˚C resulted in the death of at least 3×10^3 *H. capensis* in the marginal areas of the Swartvlei Estuary in South Africa (Russell, 1994). Although some species (e.g. *H. erectus* and *H. whitei*) grew more rapidly at warmer water temperatures, there are presumably limits on seahorses' capacity to cope (Wong & Benzie, 2003; Lin et al.,2008).

The tolerance to high temperature may also vary among species and life-history stages (Aurelio et al., 2013; Chapter 2). For instance, exposing to a future-scenario temperature of 30 ˚C has induced apparently less expression of heat-shock proteins and higher mortality rates on *H. erectus* compared with current temperature scenario in lab experiments (Johnson, 2017), while similar impacts have not been found on *H. guttulatus* (Aurelio et al., 2013). Moreover, newborn juveniles of *H. guttulatus* have shown to be more sensitive than the adults to ocean warming via heat-induced hypermetabolism (Aurelio et al., 2013).

On the other hand, ocean acidification may cause hypercapnia that could reduce activity level, feeding and ventilation rates of *H. guttulatus* (Faleiro et al. 2015).

I used two measures to estimate the impact of climate change on seahorses: 1) sea surface temperature abnormality, and 2) ocean acidification. Both datasets were derived from Halpern et al. (2015).

Calculating ecosystem-scale CHI for seahorse species

I downloaded the cumulative-human-impact map (year 2013) for global marine ecosystems from the online database (Halpern et al., 2015): http://ohi-science.org/data/. I then extracted the CHI map for each species from the global data layer by using species distribution-range map as the mask (from Chapter 3). This extracted CHI layer was then multiplied with the presence-probability layer, and the sum of all grid values of the output layer was calculated. I then divided this sum value by the sum of all grid values of the presence-probability layer to derive the CHI_e for each species. This operation corresponded to the equation 6 in the spatial model, and made the CHI_e comparable among species. All the above steps were conducted in ArcMap with the ‘Extraction’, ‘Raster Calculator’, and ‘Zonal Statistics’ tools.

Sensitivity analyses

Previously the additive algorithm was used for the value of CHI in an overlapping pixel where different marine ecosystems or multiple species co-existed (Halpern et al., 2008; 2015; Coll et al., 2012; Maxwell et al., 2013). Although the final map was suggested to be very similar to a map using the mean algorithm for the ecosystem-level study (Halpern et al., 2008), the influence of choosing different algorithms was not examined for species-level CHI models. Previous studies have displayed that species-rich areas were often overlaid with highly impacted regions (Coll et al., 2012; Lindegren et al., 2018). However, given the CHIs used in these studies was the sum of the CHIs of all species coexisting in this pixel, it would be not surprising if this pixel derived higher CHI than the other. Without excluding this artefact of methods for calculating CHI, one cannot conclude that high species richness is sympatric with high CHI. To address this issue, the mean algorithm can be used instead of the sum for overlapping cells.

On the other hand, if additive algorithm was applied, the spatial patterns of CHI may be also sensitive to the number of species involved. In my case, the total number of threatened species is 19, and only 16 species with spatial CHI map was involved in the mapping. Beside these 16 threatened species, I also have CHI maps for another 17 non-threatened species. One might argue that the if these 17 species were also involved to represent a case for all species, the threat patterns might be different. Therefore, it was also necessary to examine this effect.

I conduct the following sensitivity analyses to test the influence of species groups (threatened-only vs. all species) and pixel algorithms (sum vs. average). I examined spatial relationships between seahorse species richness and their cumulative human impacts separately for the two groups of species. Two types of cumulative human impacts were generated and used in the comparisons. One is the sum of cumulative human impacts for all species (or threatened only) in a given pixel (hereafter, CHI_sum). I then divided this CHI_sum by species richness in each pixel to derive an average CHI value (hereafter, CHI_avg). I calculated Pearson’s correlation coefficients among the CHI_sum, CHI_avg, and species richness to examine whether they were highly correlated.

References cited in Appendix G

Andersen, J.H., Berzaghi, F., Christensen, T., Geertz-Hansen, O., Mosbech, A., Stock, A., . . . Wisz, M.S. (2017) Potential for cumulative effects of human stressors on fish, sea birds and marine mammals in Arctic waters. Estuarine, Coastal and Shelf Science, 184, 202-206.

Anderson, P.A. (2009) *The functions of sound production in the lined seahorse, Hippocampus erectus,* and effects of loud ambient noise on its behavior and physiology in captive environments. University of Florida.

Aurélio, M., Faleiro, F., Lopes, V.M., Pires, V., Lopes, A.R., Pimentel, M.S., . . . Rosa, R. (2013) Physiological and behavioral responses of temperate seahorses (*Hippocampus guttulatus*) to environmental warming. Marine biology, 160, 2663-2670.

Belkin, I.M. (2009) Rapid warming of large marine ecosystems. Progress in Oceanography, 81, 207-213.

Burke, L., Reytar, K., Spalding, M. & Perry, A. (2011) Reefs at risk revisited.

Certain, G., Jørgensen, L.L., Christel, I., Planque, B. & Bretagnolle, V. (2015) Mapping the vulnerability of animal community to pressure in marine systems: disentangling pressure types and integrating their impact from the individual to the community level. ICES journal of marine science, 72, 1470-1482.

Coll, M., Piroddi, C., Albouy, C., Ben Rais Lasram, F., Cheung, W.W., Christensen, V., . . . Paleczny, M. (2012) The Mediterranean Sea under siege: spatial overlap between marine biodiversity, cumulative threats and marine reserves. Global Ecology and Biogeography, 21, 465-480.

Coll, M., Shannon, L., Kleisner, K., Juan-Jordá, M., Bundy, A., Akoglu, A., . . . Cook, A. (2016) Ecological indicators to capture the effects of fishing on biodiversity and conservation status of marine ecosystems. Ecological Indicators, 60, 947-962.

Crain, C.M., Kroeker, K. & Halpern, B.S. (2008) Interactive and cumulative effects of multiple human stressors in marine systems. Ecology letters, 11, 1304-1315.

Eno, N., Frid, C., Hall, K., Ramsay, K., Sharp, R., Brazier, D., . . . Paramor, O. (2013) Assessing the sensitivity of habitats to fishing: from seabed maps to sensitivity mapsa. Journal of Fish Biology, 83, 826-846.

Faleiro, F., Baptista, M., Santos, C., Aurélio, M.L., Pimentel, M., Pegado, M.R., . . . Rosa, R. (2015) Seahorses under a changing ocean: the impact of warming and acidification on the behaviour and physiology of a poor-swimming bony-armoured fish. Conservation physiology, 3

Foster, S. & Vincent, A. (2004) Life history and ecology of seahorses: implications for conservation and management. Journal of fish biology, 65, 1-61.

Geldmann, J., Joppa, L.N. & Burgess, N.D. (2014) Mapping change in human pressure globally on land and within protected areas. Conservation Biology, 28, 1604-1616.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'agrosa, C., . . . Fox, H.E. (2008) A global map of human impact on marine ecosystems. Science, 319, 948-952.

Halpern, B.S. & Fujita, R. (2013) Assumptions, challenges, and future directions in cumulative impact analysis. Ecosphere, 4, 1-11.

Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., . . . Selkoe, K.A. (2015) Spatial and temporal changes in cumulative human impacts on the world/'s ocean. Nature communications, 6

James, P.L. & Heck, K.L. (1994) The effects of habitat complexity and light intensity on ambush predation within a simulated seagrass habitat. Journal of Experimental Marine Biology and Ecology, 176, 187-200.

Johnson, C. (2017) Thermal Response of Behavior and Gene Expression of Heat Shock Proteins in the Lined Seahorse, *Hippocampus erectus*.

Lawson, J.M., Foster, S.J. & Vincent, A.C. (2017) Low bycatch rates add up to big numbers for a genus of small fishes. Fisheries, 42, 19-33.

Lin, Q., Lin, J. & Zhang, D. (2008) Breeding and juvenile culture of the lined seahorse, *Hippocampus erectus* Perry, 1810. Aquaculture, 277, 287-292.

Lindegren, M., Holt, B.G., MacKenzie, B.R. & Rahbek, C. (2018) A global mismatch in the protection of multiple marine biodiversity components and ecosystem services. Scientific reports, 8, 4099.

Lourie, S.A., Foster, S.J., Cooper, E.W. & Vincent, A.C. (2004) A guide to the identification of seahorses. Project Seahorse and TRAFFIC North America Project Seahorse and TRAFFIC North America, University of British Columbia and World Wildlife Fund, Washington, DC.

Magalhães, M.R.G. (2016) The effect of anthropogenic noise as a source of acoustic stress in wild populations of *Hippocampus guttulatus* in the Ria Formosa.

Martin-Smith, K.M. & Vincent, A.C. (2005) Seahorse declines in the Derwent estuary, Tasmania in the absence of fishing pressure. Biological Conservation, 123, 533-545.

Maxwell, S.M., Hazen, E.L., Bograd, S.J., Halpern, B.S., Breed, G.A., Nickel, B., . . . Dutton, P.H. (2013) Cumulative human impacts on marine predators. Nature communications, 4, 2688.

MEA, M.E.A. (2005) Ecosystems and human well-being: synthesis. Island, Washington, DC,

Negreiros, L.A., Silva, B.F., Paulino, M.G., Fernandes, M.N. & Chippari-Gomes, A.R. (2011) Effects of hypoxia and petroleum on the genotoxic and morphological parameters of *Hippocampus reidi*. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, 153, 408-414.

Orr, J.C., Fabry, V.J., Aumont, O., Bopp, L., Doney, S.C., Feely, R.A., . . . Joos, F. (2005) Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. Nature, 437, 681-686.

Ripley, J.L. & Foran, C.M. (2007) Influence of estuarine hypoxia on feeding and sound production by two sympatric pipefish species (Syngnathidae). Marine Environmental Research, 63, 350-367.

Ross, D., Johnson, C., Hewitt, C. & Ruiz, G.M. (2004) Interaction and impacts of two introduced species on a soft-sediment marine assemblage in SE Tasmania. Marine biology, 144, 747-756.

Russell, I. (1994) MASS MORTALITY OF MARINE AND ESTUARINE FISH IN THE SWARTVLEI AND WILDERNESS LAKE SYSTEMS, SOUTHERN CAPE. Southern African Journal of Aquatic Science, 20, 93-96.

Todd, P.A., Ong, X. & Chou, L.M. (2010) Impacts of pollution on marine life in Southeast Asia. Biodiversity and Conservation, 19, 1063-1082.

Vincent, A.C.J., Foster, S.J. & Koldewey, H.J. (2011) Conservation and management of seahorses and other Syngnathidae. Journal of fish biology, 78, 1681-1724.

Wong, J. & Benzie, J. (2003) The effects of temperature, Artemia enrichment, stocking density and light on the growth of juvenile seahorses, *Hippocampus whitei* (Bleeker, 1855), from Australia. Aquaculture, 228, 107-121.

York, P., Booth, D., Glasby, T. & Pease, B. (2006) Fish assemblages in habitats dominated by Caulerpa taxifolia and native seagrasses in south-eastern Australia. Marine Ecology Progress Series, 312, 223-234.

Appendix H Expert knowledge survey for Chapter 4

This Appendix contains the following supplementary information for the methods of the manuscript:1) the approach to identifying and contacting seahorse experts; 2) expert-opinion survey protocol; 3) certainty-weighted average model to estimate exposure-degree and sensitivity variables; and 4) results of the expert-opinion surveys.

The approach to identifying and contacting seahorse experts

Experts were identified by searching the Web of Science for literature on biology, distribution, ecology, and threats for each seahorse species over the past decade (2008 – 2017). We also identified additional experts who studied seahorses in relevant fiends from the attendees to SyngBio 2017 (May 14 – 19th, 2017) – the third meeting of researchers and other professionals working to understand the biology and threats of Syngnathid fishes (seahorses, pipefishes, pipehorses, and seadragons). We also requested these experts to pass our survey on to other colleagues. Finally, we invited a total of 72 seahorse biologists from 23 countries to participate through emails and skypes (27th Oct. 2017 to 30th Jan. 2018). For those contacted by email at the early stage of the survey, each biologist was asked to read through an evaluation protocol and then conduct the evaluation in specifically-designed spreadsheets accordingly. Each person was given 1.5 months to complete the evaluation, and a reminder was sent around one week ahead of the deadline. Given this didn’t work well, we then ask those who didn’t reply to our emails to accept an interview invitation on skype or phone. On skype or phone (~ 30 to 45 mins), we explained our protocol directly, and asked them questions related to threats on seahorse to help them be better prepared for the evaluation later after the skype or phone. In this case, we asked participants to send back the evaluations in a week.

Expert-opinion survey protocol

This document (including four tables) describes an easy-following protocol for assessing seahorse exposure degree and sensitivity to different anthropogenic pressures. For this survey, we will first ask you about your experience on studying seahorses (Part I). We then ask you to read the description of the exposure degree & sensitivity measures, and anthropogenic stressors we selected (Part II). Finally, we ask you to follow this protocol to do the evaluation for each pair of species and pressure (that you are familiar with) in the excel table (sent along with this protocol).

Please note that we are asking for your best estimates about the human impacts on seahorses within the past five years. Therefore, it should be under present-day context, rather than the projected impacts in the future. The survey is confidential, and your responses will be stored by us securely. The list of participants will be shared only among the two researchers involved in the project.

Part I. Your experience about seahorses

We ask you to fill a form regarding your experience about seahorse species as shown below. It can inform you and us of which species you are able to evaluate. Please use the excel table ("E&S evaluation table", see Table H.1) we sent to you along this protocol! Please do not miss any species you have studied or are familiar with.

Table H.1. Expert background information table

Species you have studied or are familiar with **(e.g. *Hippocampus erectus*)**	**For how long?** **(e.g. 5 years)**	**Research field you know/study** **(e.g. biology; can be more than one field)**	**Are you familiar with human impacts on this species? (e.g. Yes/No)**
Hippocampus			
Hippocampus			
Hippocampus			

Part II. Exposure degree and sensitivity measures (10 minutes)

We defined exposure degree as the degree of exposure of the species to a given pressure; and sensitivity as the degree of intolerance and incapability to recover.

Please read through Tables H.2, H.3 & H.4. Here we describe how to estimate exposure degree and sensitivity of seahorses to 12 human stressors.

Table H.2 contains two exposure degree measures: A. Frequency of impact; B. Proportion of population under impact. For each measure there are five levels of category. Meanwhile, for each evaluation, we ask you to provide a certainty category: very confident, confident, likely, or not sure.

Table H.3 contains only one sensitivity index – the severity of the impact. However, the sensitivity can vary between adults and juveniles as they use different habitats. We discriminate the differences if it exists, and use three levels of severity: low, medium, and high. Again, we ask you to provide a certainty category for each evaluation.

Table H.4 describes the 12 human-derived stressors we used for the evaluation on seahorse exposure degree and sensitivity.

We provided **an example of evaluation in Table H.5**. Please use it as a reference to fill in your information in the excel.

Table H.2. Exposure degree measures for seahorses.

Exposure degree measure	Category	Descriptive notes
A. Frequency of impact: **How often does the impact occur on the population?**	never	-
	Rare	infrequent enough to affect long-term dynamics of a given population or location
	occasional	frequent but irregular in nature
	annual or regular	frequent and often seasonal or periodic in nature
	persistent	more or less constant year-round, lasting through multiple years or longer
*certainty		
B. Proportion of population under impact: **To what spatial extent does the impact distribute across the species' habitats?**	very small	extent ≤ 1 km^2 (<u>or</u> 10% of the habitats)
	small	1 km^2 (<u>or</u> 10% of the habitats) < extent ≤ 10 km^2 (<u>or</u> 30% of the habitats)
	medium	10 km^2 (<u>or</u> 30% of the habitats) < extent ≤100 km^2 (<u>or</u> 50% of the habitats)
	large	100 km^2 (<u>or</u> 50% of the habitats) < extent ≤ 1,000 km^2 (<u>or</u> 80% of the habitats)
	very large	extent > 1,000 km^2 (<u>or</u> 80% of the habitats)
*certainty		

* certainty: very confident, confident, likely, or not sure.

Table H.3. Sensitivity metrics for seahorses.

Sensitivity metric	Category	Descriptive notes
Does the impact differ among life-history stages?	A. Mainly affects adults	Some stressors mainly affect adult fishes
	B. Mainly affects juveniles	Some stressors mainly affect juvenile fishes
	C. Affects all individuals equally	Some stress may equally impact each individual if contacted
If yes, **then** to what extent of severity **does the pressure** usually **impose on the adults and juveniles separately?**		
For adults	Low	only cause low-level physiological stress, like increased ventilation, but no significant impact on growth
	Medium	induce significant decrease on growth or reproduction, but not lead to death
	High	direct death
*certainty		
For juveniles	Low	only cause low-level physiological stress, like increased ventilation, but no significant impact on growth
	Medium	induce significant decrease on growth, but not lead to death
	High	direct death
*certainty		
If no, to what extent of severity **does the pressure impose on the population?**	Low	only cause low-level physiological stress, like increased ventilation, but no significant impact on growth
	Medium	induce significant decrease on growth or reproduction, but not lead to death
	High	direct death
*certainty		

* certainty: very confident, confident, likely, or not sure.

Table H.4. A list of human activities used in the evaluation on seahorse exposure degree & sensitivity.

Dimension	**Category**	**Description (Estimated by)**	**Justification**
Fishing	**Demersal destructive fishing**	bomb/chemical (all types), bottom trawl (all types), dredge (all types)	typically cause direct mortality especially adults
	Demersal, non-destructive, high-bycatch fishing	gillnet fixed, gillnet encircling, seine, beach seine, boat seine, trammel (all types), trap (all types)	typically cause direct mortality
	Pelagic high-bycatch fishing	driftnets, midwater trawl	typically cause direct mortality of juveniles and larvae
	Artisanal fishing	an estimate that cannot distinguish between methods that do and do not modify habitat.	assumed to cause overfishing in inshore shallow waters
Habitat change	**Habitat destruction caused by fishing**	bomb/chemical (all types), bottom trawl (all types), dredge (all types)	typically cause direct bottom habitat change
	Coastal development	port volume, population density	Coastal development can cause damage on seahorse habitats
Pollution	**Nutrient pollution**	modeled based on fertilizer data	induce hypoxia or derive toxic pollutants
	Ocean pollution	mainly oil spills estimated by commercial shipping intensity plus port volume	induce hypoxia or derive toxic pollutants
	Noise pollution	estimated by the intensity of shipping & human population along the coast	induce metabolic change and stress

Dimension	Category	Description (Estimated by)	Justification
Invasive species	**Invasive species**	the amount of cargo at ports	change habitat, competition and predation
Climate change	**SST abnormality**	frequency of SST passes the long-term mean SST (i.e. threshold)	increase metabolic rate
	Ocean acidification	differences on the aragonite saturation state of the ocean between pre-industrial (circa 1870) and modern times (2000 - 2009)	reduce activity level, feeding and ventilation rates， affect food availability

Table H.5. An example of evaluation table for exposure degree & sensitivity assessment on H. kuda (faked), with evaluations in red and certainties in blue.

Species: ***H. kuda***	**Demersal destructive fishing**	**Demersal, non-destructive, high-bycatch fishing**	**Pelagic high-bycatch fishing**	**Artisanal fishing**	**Habitat degradation or loss caused by fishing**	**Coastal non-fishing activities**	**Nutrient pollution**	**Ocean pollution**	**Noise pollution**	**Invasive species**	**SST abnormality**	**Ocean acidification**
Frequency of the impact	Annual or regular	Occasional	Annual or regular	Occasional	Annual or regular	Annual or regular	Annual or regular	Annual or regular	rare	Annual or regular	Annual or regular	Never occur
Certainty	Confident	Confident	Confident	Confident	Confident	Confident	Likely	Likely	Confident	Confident	Confident	Confident
Proportion of population that suffer	Very small	Small	Very small	Small	Very small	Very large	Large	Large	Very small	Medium	Small	Very small
Certainty	Confident	Confident	Confident	Confident	Confident	Likely	Likely	Likely	Confident	Likely	Confident	Confident
*** Does impact differ among life-history stages?**	A. Mainly affects adults	A. Mainly affects adults	B. Mainly affects juveniles	C. Affects all individuals equally	A. Mainly affects adults	C. Affects all individuals equally	C. Affects all individuals equally	C. Affects all individuals equally	C. Affects all individuals equally	C. Affects all individuals equally	B. Mainly affects juveniles	C. Affects all individuals equally
If A or B, severity for adult	High	High	Low		High						Low	
Certainty	Very confident	Confident	Confident		Confident						Confident	
If A or B, severity for juveniles	Low	Low	High		Low						Medium	
Certainty	Confident	Confident	Confident		Confident						Confident	
If C, severity for the population				Low		Medium	Low	Low	Low	Medium		Medium
Certainty				Likely		Likely	Likely		Confident	Likely		Confident

Certainty-weighted average model to estimate exposure-degree and sensitivity variables

Model equations:

$$F_{i,j} = \frac{\sum_n F_{i,j,n} \times C\ (F_{i,j,n})}{\sum_n C\ (F_{i,j,n})} \tag{7}$$

$$PS_{i,j} = \frac{\sum_n PS_{i,j,n} \times C\ (PS_{i,j,n})}{\sum_n C\ (PS_{i,j,n})} \tag{8}$$

$$S_{i,j} = \frac{\sum_n S_{i,j,n} \times C\ (S_{i,j,n})}{\sum_n C\ (S_{i,j,n})} \tag{9}$$

$$S_{i,j,n} = \begin{cases} \frac{AS_{i,j,n} \times C\ (AS_{i,j,n}) + JS_{i,j,n} \times C\ (JS_{i,j,n})}{C\ (AS_{i,j,n}) + C\ (JS_{i,j,n})}, if\ severity\ differs\ between\ adults\ and\ juveniles \\ SI_{i,j,n}, if\ severity\ doesn't\ differs\ among\ life\ history\ stages \end{cases}$$

$$C\ (S_{i,j,n}) = \begin{cases} \frac{C\ (AS_{i,j,n}) + C\ (JS_{i,j,n})}{2}, if\ severity\ differs\ between\ adults\ and\ juveniles \\ C\ (SI_{i,j,n}), if\ severity\ doesn't\ differs\ among\ life\ history\ stages \end{cases}$$

where $F_{i,j,n}$, $PS_{i,j,n}$, and $S_{i,j,n}$ were subsequently the specific score of the three measures from a given expert *n*; and $C\ (F_{i,j,n})$, $C\ (PS_{i,j,n})$, and $C\ (S_{i,j,n})$ were respectively the certainty index of these scores given by expert *n*. Given species sensitivity to a given stressor may vary between adults and juveniles (Aurelio et al., 2013), we asked experts to evaluate adults and juveniles respectively if the sensitivity differ. When sensitivity was inconsistent and assessed for adults and juveniles separately, $AS_{i,j,n}$ (adults' sensitivity to the impact) and $JS_{i,j,n}$ (juveniles' sensitivity to the impact) were used with their certainty index: $C\ (AS_{i,j,n})$ and $C\ (JS_{i,j,n})$. When sensitivity was consistent and evaluated for the species collectively, $SI_{i,j,n}$ (the sensitivity to all individuals) and its certainty index $C\ (SI_{i,j,n})$ were used.

To understand the impacts of which anthropogenic stressors were poorly known by our experts, we compared the confidence index of expert evaluation (average of $C\ (F_{i,j,n})$, $C\ (PS_{i,j,n})$, and $C\ (S_{i,j,n})$) on each stressor across the species (n = 42).

Results of the expert knowledge surveys

We obtained E&S evaluation results for a total of 24 species from 22 participants, who have conducted related seahorse studies on one or more species in 15 countries for 10±6 years per species (**Fig. H.1**). Only two participants have also conducted the IUCN Red List assessment for seahorses. The mean certainty indexes of E&S evaluations differed significantly among different stressors (Kruskal-Wallis test $P < 0.001$), with demersal-destructive fishing derived the highest mean value (3.0 ± 0.5) and climate change stressors (ocean acidifications and SST abnormality) obtained the lowest mean values (1.6 ± 0.5, 1.9 ± 0.5, respectively; **Fig. H.2**; see details about pairwise comparisons between stressors in **Table H.6**). The number of experts contributed evaluations for each species ranged from 1 to 5, with averaged confidence levels from 2 to 4 ('likely' to 'very confident').

Table H.6. Statistics significances (*P*-values) of Post-hoc pairwise Wilcox tests between 12 anthropogenic stressors: AF, artisanal fishing; CD, coastal development; DDF, demersal destructive fishing; DNHF, demersal non-destructive high-bycatch fishing; HCF, habitat change caused by fishing; IS, invasive species; NUP, nutrient pollution; OA, ocean acidification; OP, ocean pollution; PHF, pelagic high-bycatch fishing.

	AF	CD	DDF	DNHF	HCF	IS	NoP	NuP	OA	OP	PHF
CD	1.000	-	-	-	-	-	-	-	-	-	-
DDF	0.454	*	-	-	-	-	-	-	-	-	-
DNHF	1.000	1.000	0.223	-	-	-	-	-	-	-	-
HCF	0.903	0.057	1.000	0.177	-	-	-	-	-	-	-
IS	*	0.231	****	1.000	****	-	-	-	-	-	-
NoP	*	***	****	0.310	****	1.000	-	-	-	-	-
NuP	0.074	0.618	****	1.000	****	1.000	0.310	-	-	-	-
OA	****	****	****	****	****	****	***	****	-	-	-
OP	****	****	****	***	****	0.732	1.000	****	0.085	-	-
PHF	0.231	0.395	****	***	****	1.000	1.000	1.000	***	0.134	-
SSTA	****	****	****	*	****	*	0.811	***	*	1.000	0.535

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$, **** $P < 0.001$.

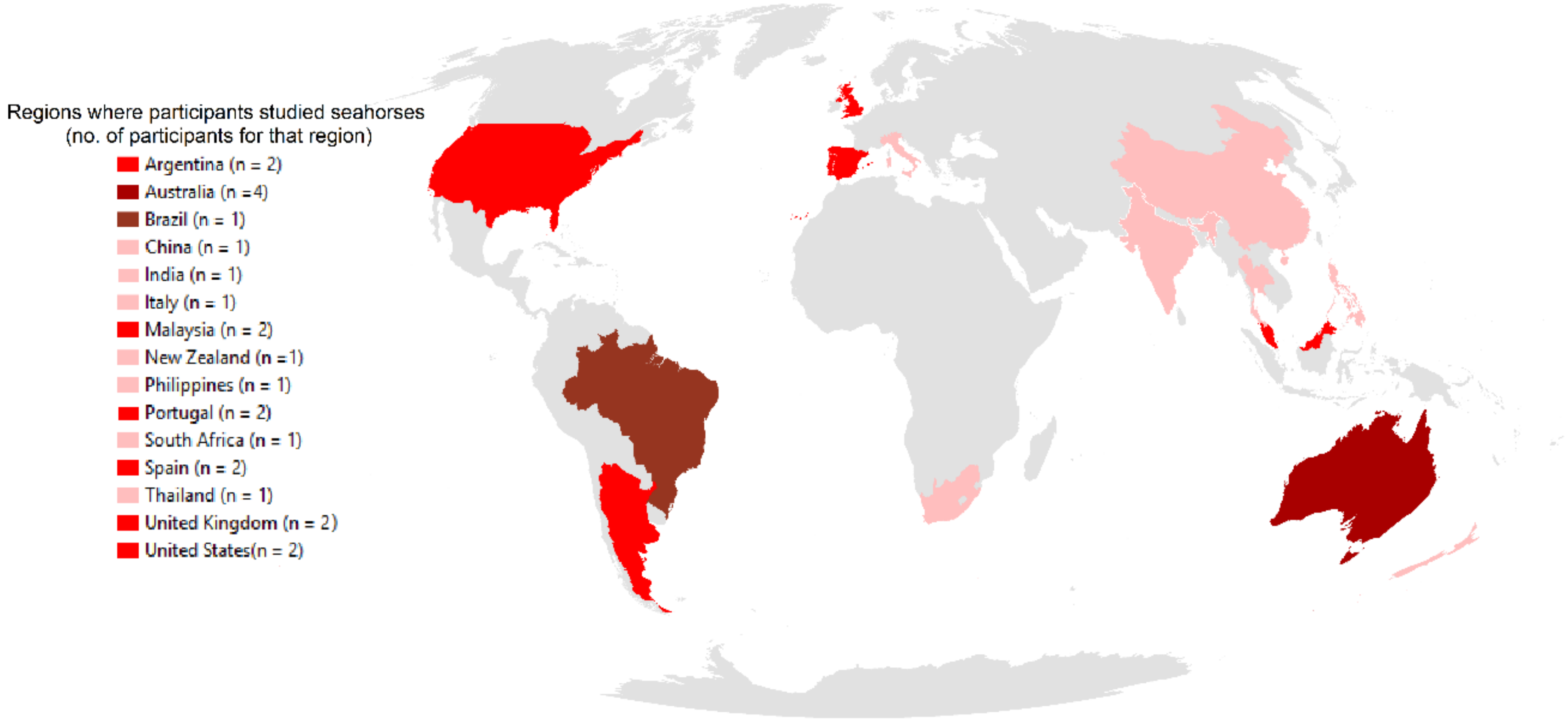


Figure H.1. Regions where participants studied seahorses with the number of participants for each region presented in parentheses.

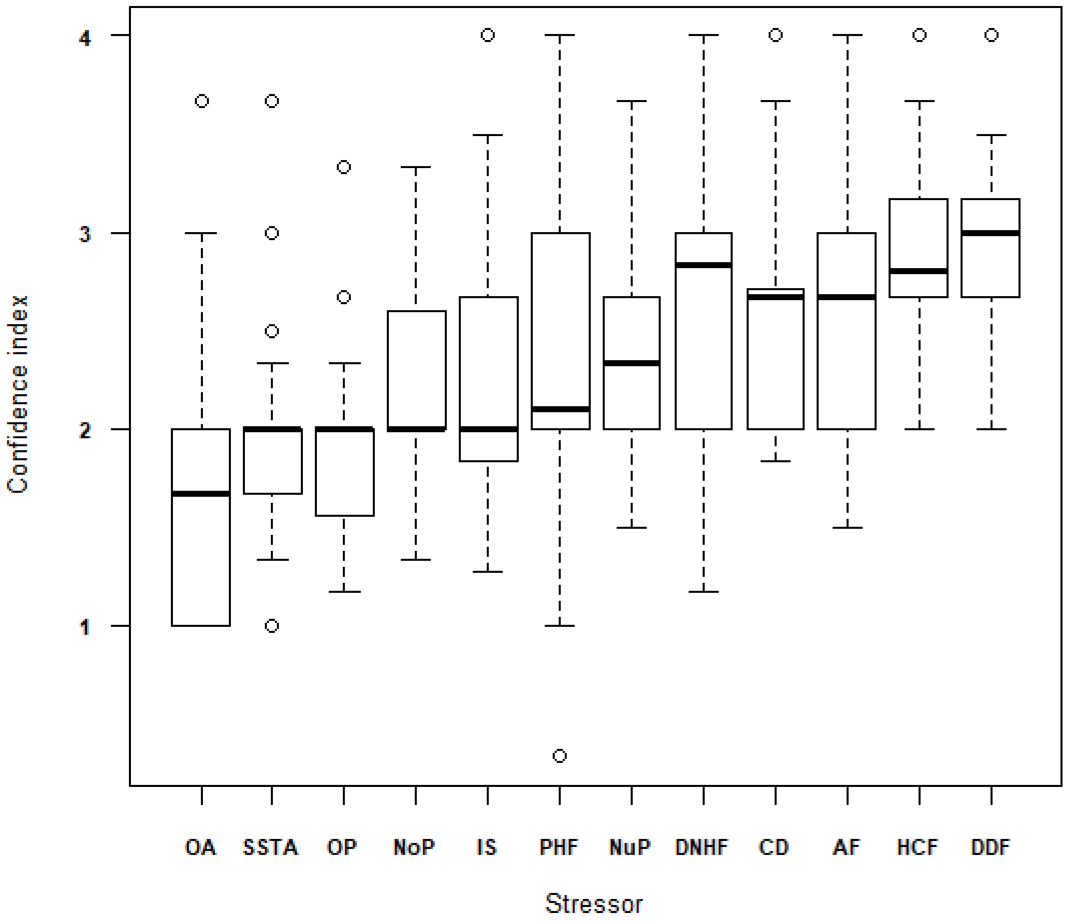


Figure H.2. Confidence index of experts knowledge on the 12 anthropogenic stressors: AF, artisanal fishing; CD, coastal development; CHI_n, cumulative human impact derived from the non-spatial model; CHI_s, cumulative human impact derived from the spatial model; DDF, demersal destructive fishing; DNHF, demersal non-destructive high-bycatch fishing; HCF, habitat change caused by fishing; IS, invasive species; NoP, noise pollution; NuP, nutrient pollution; OA, ocean acidification; OP, ocean pollution; PHF, pelagic high-bycatch fishing; SSTA, sea-surface-temperature abnormality. The stressors were ranged by the mean value across all species from lowest (left) to the highest (right).

References cited in Appendix H

Aurélio, M., Faleiro, F., Lopes, V.M., Pires, V., Lopes, A.R., Pimentel, M.S., . . . Rosa, R. 2013. Physiological and behavioral responses of temperate seahorses (*Hippocampus guttulatus*) to environmental warming. *Marine Biology*, 160, 2663-2670.

Appendix I Supplementary results for Chapter 4

This appendix contains supplementary information of results of the manuscript: 1) Fig. I.1. Scatterplot and linear model between the two cumulative human impact indexes derived from different models (spatial vs. non-spatial); 2) Fig. I.2. Global intensity maps of four major anthropogenic stressors; 3) Fig. I.3. Proportion of individual impact of eight major stressors to cumulative human impact (based on the non-spatial CHI model) on 19 threatened seahorse species; 4) Fig. I.4. Spatial distributions of species richness and two versions of cumulative human impact (sum across species vs. average across species) for all seahorse species and threatened species only; 5) Fig. I.5. Cumulative human impacts on a) all seahorse species and b) marine ecosystems (across seahorse habitats); and 6) Fig. I.6. Difference on cumulative human impacts between the ecosystems and seahorse species.

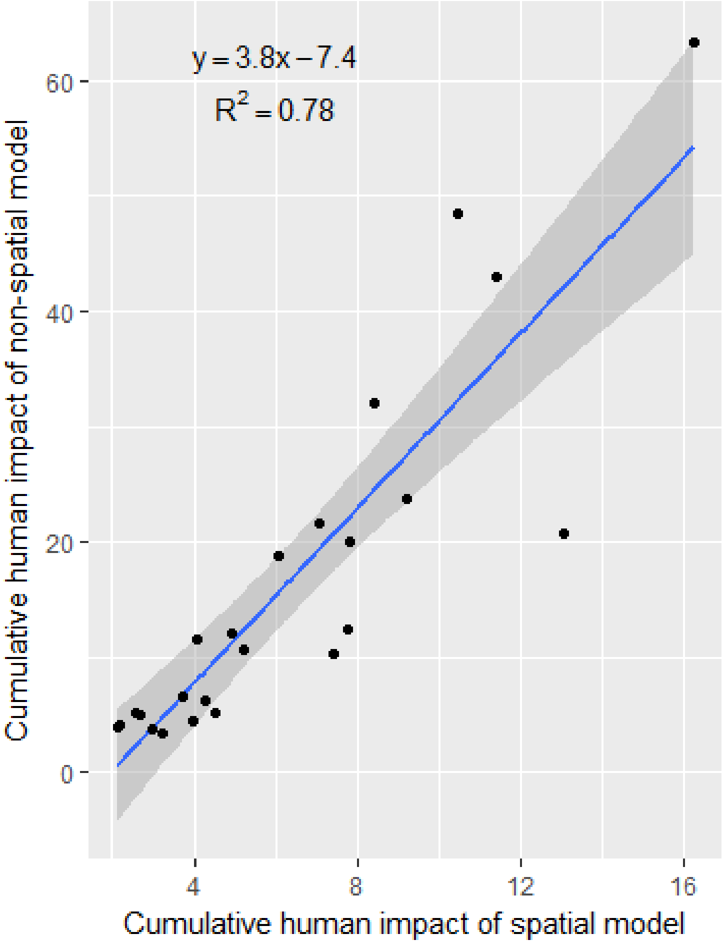


Figure I.1. Scatterplot and linear model between the two cumulative human impact indexes derived from different models (spatial vs. non-spatial).

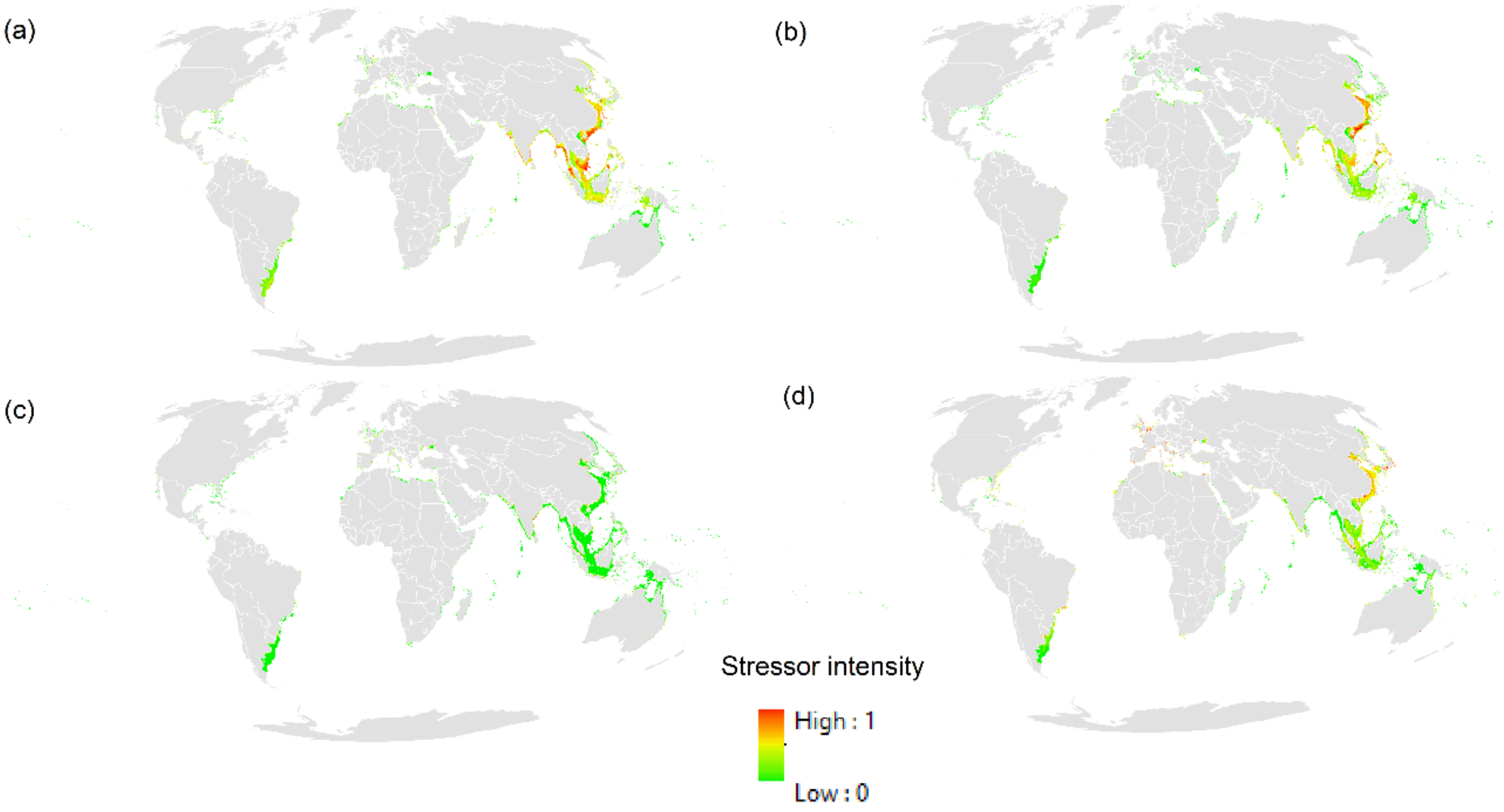


Fig. I.2. Global intensity maps of four major anthropogenic stressors: a) demersal destructive fishing, b) demersal, non-destructive, high-bycatch fishing, c) nutrient pollution, and d) ocean pollution.

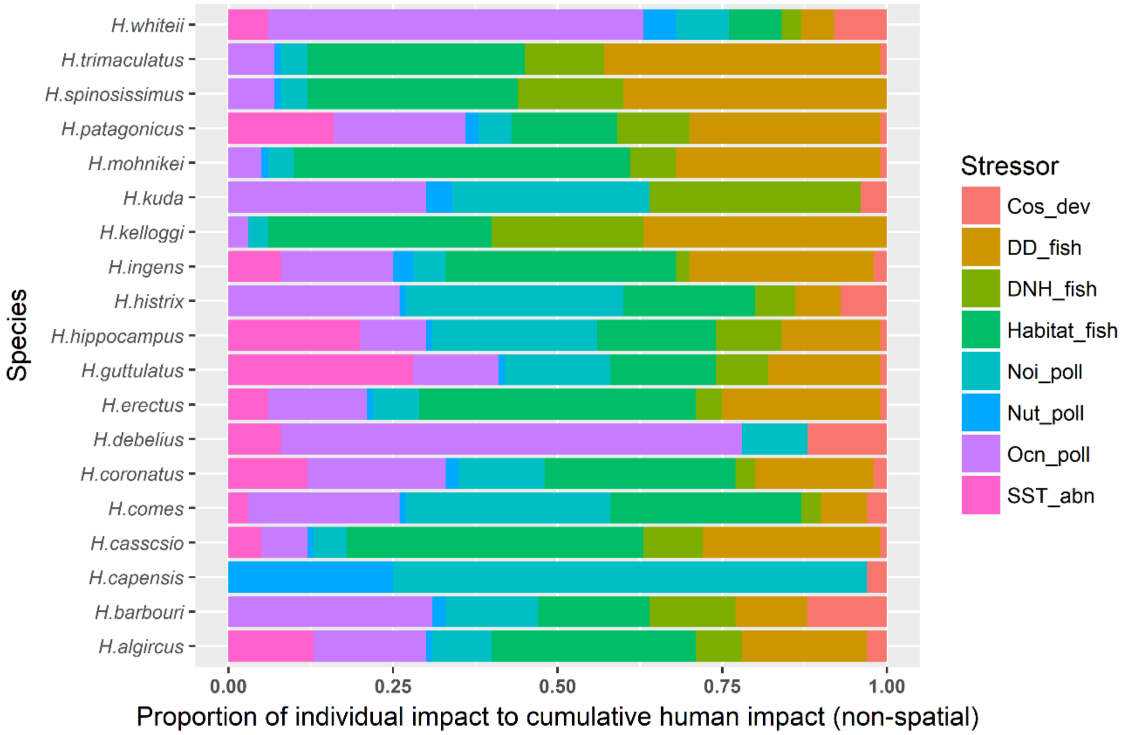


Fig. I.3. Proportion of individual impact of eight major stressors to cumulative human impact (based on the non-spatial CHI model) on 19 threatened seahorse species. Cos_dev, coastal development; DD_fish, demersal destructive fishing; DNH_fish, demersal, non-destructive, high-bycatch fishing; Habitat_fish, habitat change caused by fishing; Noi_poll, noise pollution; Nut_poll, nutrient pollution; Ocn_poll, ocean pollution; SST_abn, sea-surface-temperature abnormality. Other four stressors were not included as their impacts were consistently low (impact index < 0.3) and not truly important to species threatened status. These four stressors are pelagic high-bycatch fishing, artisanal fishing, invasive species, and ocean acidification.

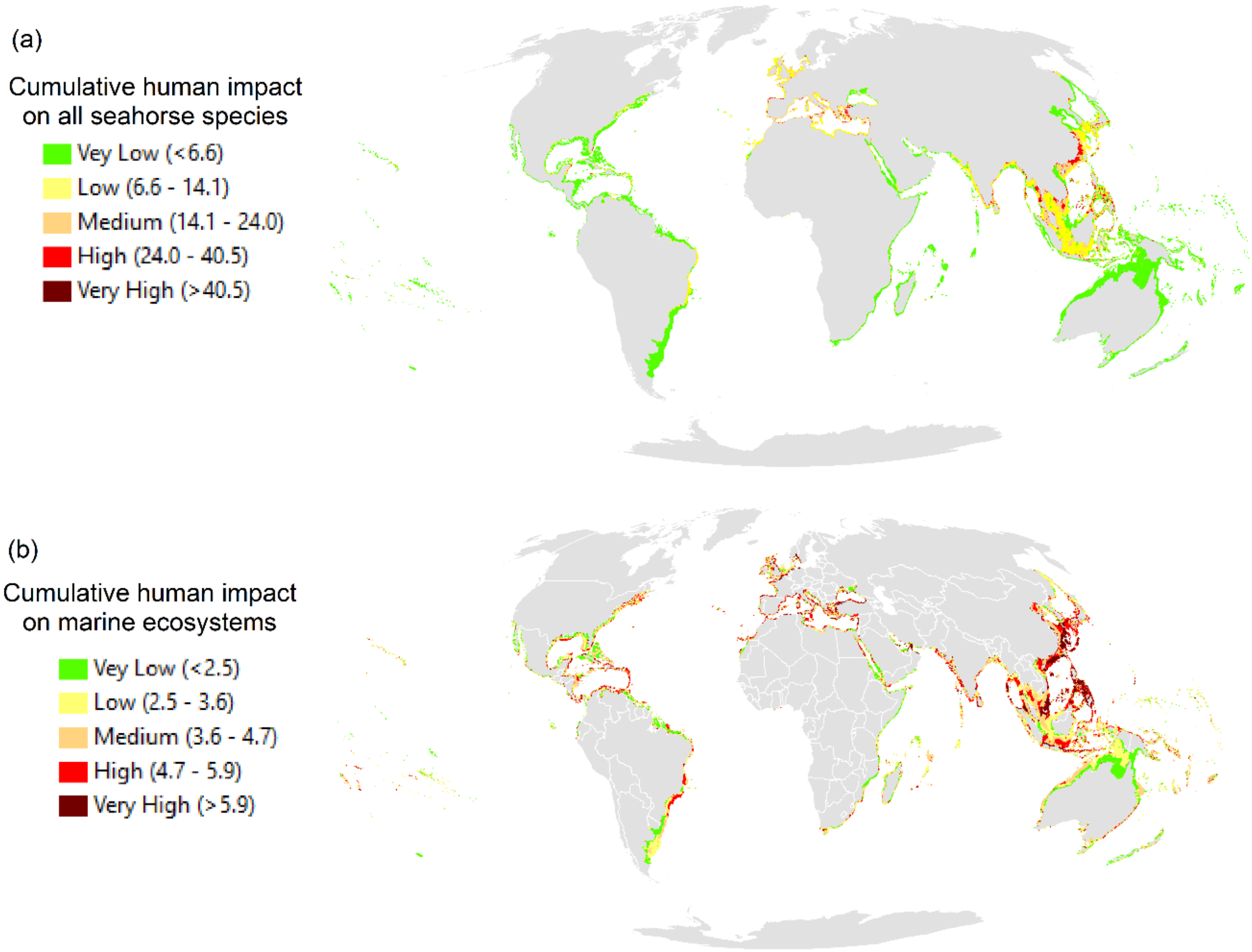


Fig. I.4. Cumulative human impacts on a) all seahorse species and b) marine ecosystems (across seahorse habitats). The impact values were classified into five levels based on the Jenks Natural Breaks algorithm in ArcMap. Pearson's correlation coefficient between the two maps was 0.50.

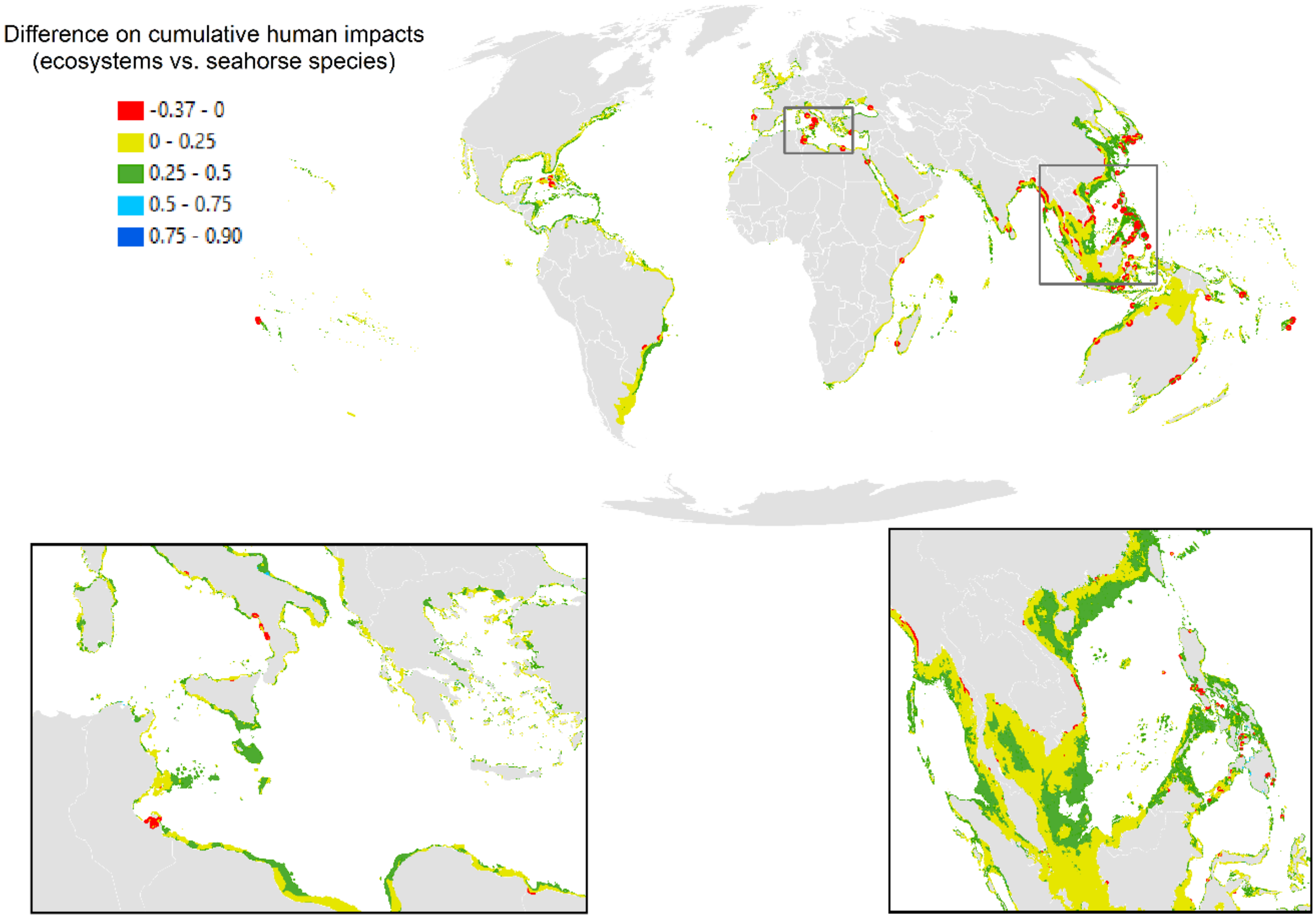


Fig. I.5. Difference on cumulative human impacts between the ecosystems and seahorse species. The value was derived by deducting the CHI values on marine ecosystems by the CHI values on all seahorse species across seahorse habitats (1 km by 1 km pixels). Both CHIs (ecosystems vs. seahorse species) were rescaled from 0 – 1 by the maximum value before the comparison.

Appendix J: Supplementary information for Chapter 5

This appendix contains supplementary information for Chapter 5: 1) Prepare four types of opportunity cost data, 2) Set appropriate value for species penalty factor and boundary length modifier, 3) Threshold for determining priorities based on the solution frequency, 4) Table J.1, and 5) Table J.2.

Prepare four types of opportunity cost data

My five types of cost data were derived from Halpern et al. (2015) and my Chapter 4. Among them, cumulative-human-impact (CHI) cost was derived from the CHI model developed by Chapter 4. This model measured the human impact of each species-stressor pair in seahorse habitats (pixel size: 1 x 1 km2) and summed them across 12 anthropogenic stressors including fishing, habitat destruction, pollution, invasion, and climate change. For habitat pixels shared by multiple species, I used the average of the CHI value of these species. For commercial fisheries cost, I used two types of fishing categories to estimate its intensity: demersal destructive fishing (e.g., bottom trawling and dredging), and demersal non-destructive fishing (e.g., seines and traps). These two fishing activities were major threats to seahorses suggested by Chapter 4. I obtained raw intensity data of each category from Halpern et al. (2015) and extracted the maps into my planning units and in ArcMap. These data were model estimates of annual catch (by weight) in each pixel (1 km by 1 km) based on FAO's statistics and Sea Around Us Project's estimates (see detailed description in the supplement information of Halpern et al., 2008 & 2015). I then added the intensity values of the two layers in each pixel in ArcMap. I then log(1+X)-transformed the value and rescaled the value from 0 – 1 by dividing each value by the maximum value across the planning units. For artisanal fishing, ocean pollution, and nutrient pollution, I derived the intensity estimates from Halpern et al. (2015). Again, these were model estimates based on relevant activities rather than observations. For instance, ocean-based pollution is a combination of the shipping and port volume. I extracted these intensity data into my planning units and then rescaled them from 0 – 1 respectively.

Set appropriate values for species penalty factor and boundary length modifier

Species penalty factor (SPF) and boundary length modifier (BLM) are crucial to getting good results from Marxan. Too high SPF would restrict Marxan's performance, while too low SPF would allow Marxan to sacrifice the conservation goal which is unacceptable. On the other hand, BLM determines how much emphasis to be placed on the connectivity of resulting planning units. If a BLM was too high, then the Marxan would more likely to derive a priority system that were more compact but also more expensive. If a BLM was too low, the resulting priorities would be very scattered in space, which would also not acceptable. To find good values for the two factors, I conducted the following process based on previous guides.

I first controlled my BLM as a generally reasonable value (= 1), which makes the product of BLM and my boundary length (here, single value = 1) equal to the largest opportunity cost (= 1). I then used a set of trial values of SPF (here, start from 10 and decrease by 2 each time) to run the Marxan for 100 times per trial. I examined the penalty value for each trial until it became positive – conservation goal was not fully met. I then increased the SPF values slightly to find a good value that could meet conservation goal. I then identified this SPF value, and did a set of trails to find appropriate BLM value. Again, I run Marxan for 100 times per trial. I start BLM from 0 and then increase the value by some increments (here I used 0.1, 0.5, 1, 10, 100) for each run. I collected the average cost and total boundary length from the output and plotted them to identify a tipping point where the cost increase dramatically.

Finally, I chose a BLM = 0.5 and a SPF = 5 for the prioritization scenarios of *Hippocampus barbouri*, and a BLM = 0.3 and a SPF = 5 for *H. abdominalis* and *H. minotaur*, a BLM = 0.7 and a SPF = 3 for *H. histrix*.

Threshold for determining priorities based on the solution frequency

I determined the priority units based on the solution frequency by using a threshold after multiple trials, ensuring the conservation targets were met with the minimum number of planning units. The first trial used the threshold that derives a network with equal number of the selected units to the best-solution map. For instance, if the best-solution map selected 400 planning units, the initial threshold was the value at the top 400 in the solution-frequency list. By doing so, I could derive a roughly ideal amount of planning units suggested by the best solution. I then check the derive solution against the conservation target to see if the target was met. If the target was not met, I then gradually decrease the threshold (i.e., smaller frequency) and repeat the checking until I found a minimum amount of planning units that met the target. In contrast, if the initial threshold generated a solution overpass the target, I then gradually increase the threshold (i.e., larger frequency) and repeat the same process.

Table J.2. Summary of sources of China's seahorse species records.

Chinese name	Scientific name	*Reference code
三斑海马	*Hippocampus trimaculatus*	1,2,7
日本海马	*Hippocampus mohnikei*	1,3,7,8
克氏海马	*Hippocampus kelloggi*	1,2,7,8
库达海马	*Hippocampus kuda*	1,2,7
棘海马	*Hippocampus spinosissimus*	1,2
刺海马	*Hippocampus histrix*	2,7
科氏海馬	*Hippocampus colemani*	2
彭氏海马	*Hippocampus pontohi*	2
花海马	*Hippocampus sindonis*	2
巴氏豆丁海马	*Hippocampus bargibanti*	2
冠海马	*Hippocampus coronatus*	3
北部湾海马	*Hippocampus casscsio*	4
鲍氏海马	*Hippocampus barbouri*	5，6
虎尾海马	*Hippocampus comes*	6

*Reference Code: 1, Chapter 2 of this thesis; 2, The Fish Database of Taiwan; 3, Zhang et al. 1994; 4, Zhang et al. 2016; 5, my own observations 2016, unpublished; 6, Dr. Lin Qiang 2017, South China Sea Institute of Oceanography, unpublished; 7, Zhu et al. 1962; 8, Zhu et al. 1963.

Table J.2. Summary of sources of China's marine ecological redline maps.

Region	source	Website
Tianjin	China's Economics	http://district.ce.cn/newarea/roll/201210/17/t20121017_23761214.shtml
Fujian	Urban planning information	http://chinaup.info/2011/07/1433.html
Zhejiang	Zhejiang News	https://zj.zjol.com.cn/news/671033.html
Hainan	South China Inspection Bureau, Ministry of Environmental Protection	http://hndc.mep.gov.cn
Shandong	Qingdao Geo-Marine Engineering Survey	http://www.qgmes.com.cn/Project.asp?BigClassId=5&SmallClassId=21
Guangdong	Chinese colleague (Anonymity)	-
Guangxi	Chinese colleague (Anonymity)	-
Hebei	Sina News	http://news.sina.com.cn/o/2018-03-09/doc-ifyscerk9155497.shtml
Jiangsu	China News	http://www.chinanews.com/gn/2017/04-12/8197565.shtml
Liaoning	Wangyi Blog	http://blog.163.com/special/0012sp/disappearcoastline.html

References cited in Appendix J

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'agrosa, C., . . . Fox, H.E. 2008. A global map of human impact on marine ecosystems. *Science*, 319, 948-952.

Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., . . . Selkoe, K.A. 2015. Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nature communications*, 6, 7615.

Zhang, C., Q. Cheng, and B. Zheng. 1994. Fish of Yellow Sea and Bohai Sea. Taiwan: Aquaculture Press. Page 83 - 84.

Zhang, Y.-H., G. Qin, X. Wang, and Q. Lin. 2016. A new species of seahorse (Teleostei: Syngnathidae) from the South China Sea. Zootaxa 4170(2):384-392.

Zhu, Y., C. Zhang, and Q. Cheng. 1962. Ichthyography of South China Sea. Science Press, Beijing.

Zhu, Y., C. Zhang, and Q. Cheng. 1963. Ichthyography of East China Sea. Science Press, Beijing.

Appendix K: Supplementary results for Chapter 5

This appendix contains supplementary results from Chapter 5: Table K.1 and Figures K.1 – K.12.

Table K.1. Consistency test (Cohen's kappa statistic) between two types of priority solutions (solution frequency vs. best solution) within each of the five scenarios: AF-cost, conservation cost measured with artisanal fisheries and total boundary length; CF-cost, conservation cost measured with commercial fisheries (here, demersal non-selective fishing) and total boundary length; CHI-cost, conservation cost measured with cumulative human impact and total boundary length; NP-cost, conservation cost measured with nutrient pollution and total boundary length; SP-cost, conservation cost measured with shipping intensity and total boundary length. The kappa value ranges from -1 to 1, with higher value indicating higher consistency. Here, all P-values < 0.001, suggesting the consistency between the two methods was not random.

	Kappa statistics (solution frequency vs. best selection)		
	Australia	Southeast Asia	China
AF-cost	0.862	0.805	0.680
CF-cost	0.861	0.622	0.796
CHI-cost	0.730	0.720	0.645
NP-cost	0.758	0.793	0.806
SP-cost	0.927	0.840	0.806

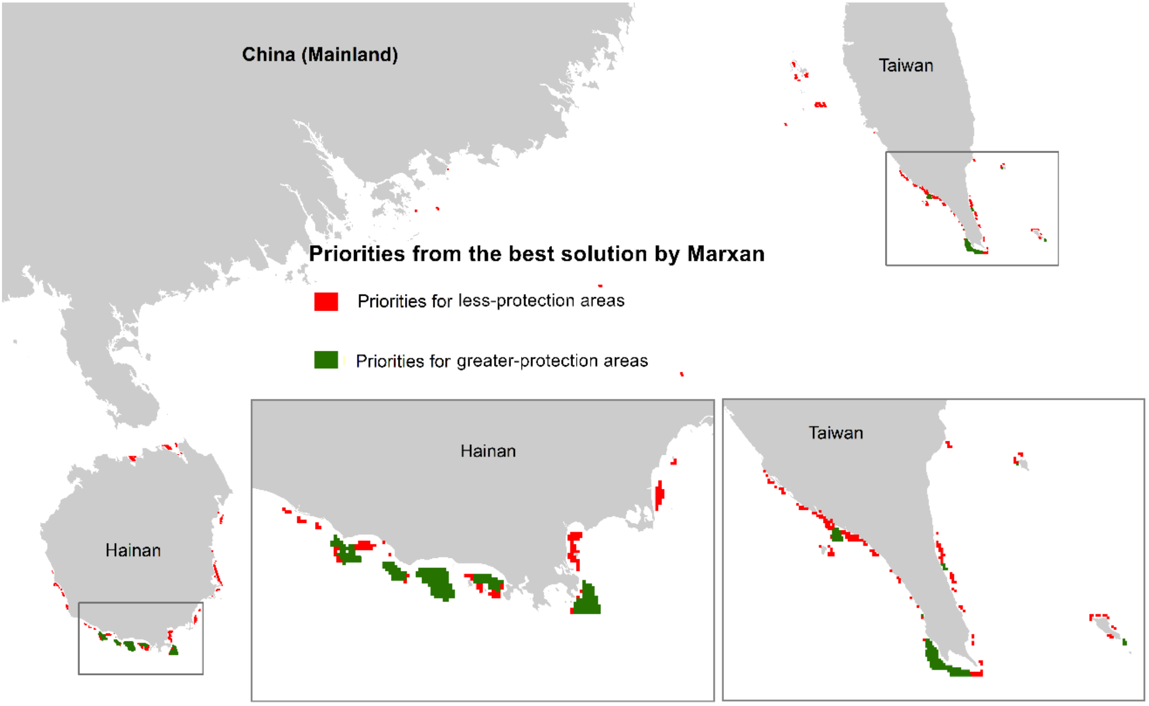


Figure K.1. Priorities from the best solution derived from the Marxan software for the *Hippocampus histrix* in China. LPAs, lower-protection areas (spatial disagreements between the biological conservation and diverse human uses); GPAs, greater-protection areas (spatial agreements between the biological conservation and diverse human uses). The detailed information about multiple-use areas were not shown because there were many classifications and difficult to illustrate by colors. Specific information can be found from the figshare database under the project folder "Global Priorities for Seahorse Species".

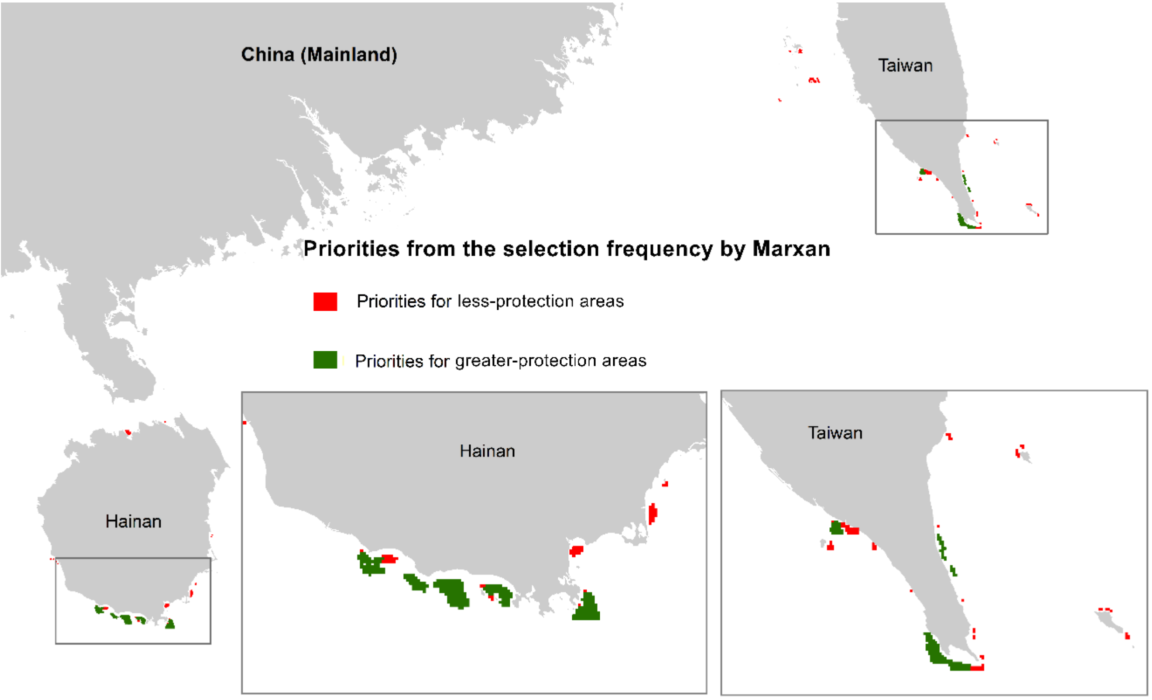


Figure K.2. Priorities from the selection frequency derived from the Marxan software for the *Hippocampus histrix* in China. LPAs, lower-protection areas (spatial disagreements between the biological conservation and diverse human uses); GPAs, greater-protection areas (spatial agreements between the biological conservation and diverse human uses). The detailed information about multiple-use areas were not shown because there were many classifications and difficult to illustrate by colors. Specific information can be found from the figshare database under the project folder "Global Priorities for Seahorse Species".

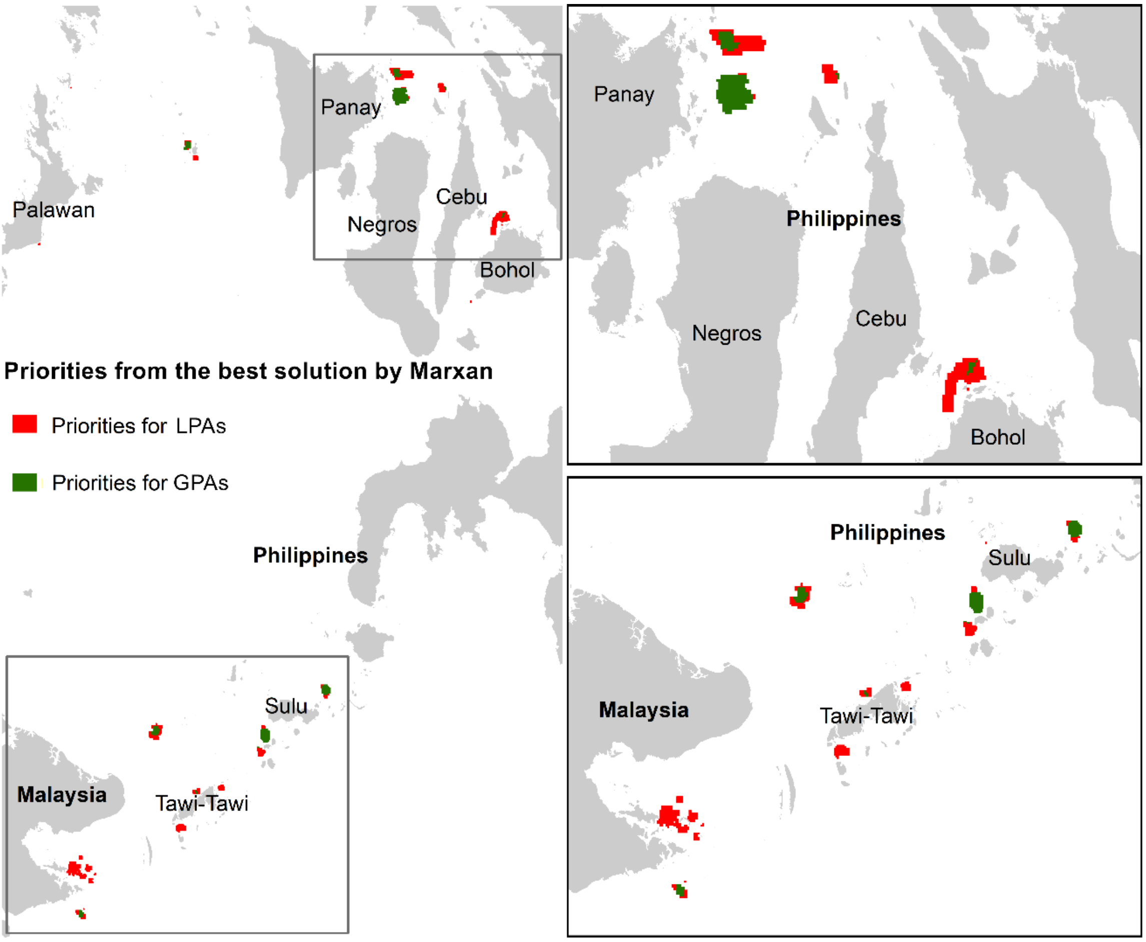


Figure K.3. Priorities from the best solution derived from the Marxan software for the Southeast Asia seahorse species (*Hippocampus barbouri*). LPAs, lower-protection areas (spatial disagreements between the biological conservation and diverse human uses); GPAs, greater-protection areas (spatial agreements between the biological conservation and diverse human uses). The detailed information about multiple-use areas were not shown because of there were many classifications and difficult to illustrate by colors. Specific information can be found from the figshare database under the project folder "Global Priorities for Seahorse Species".

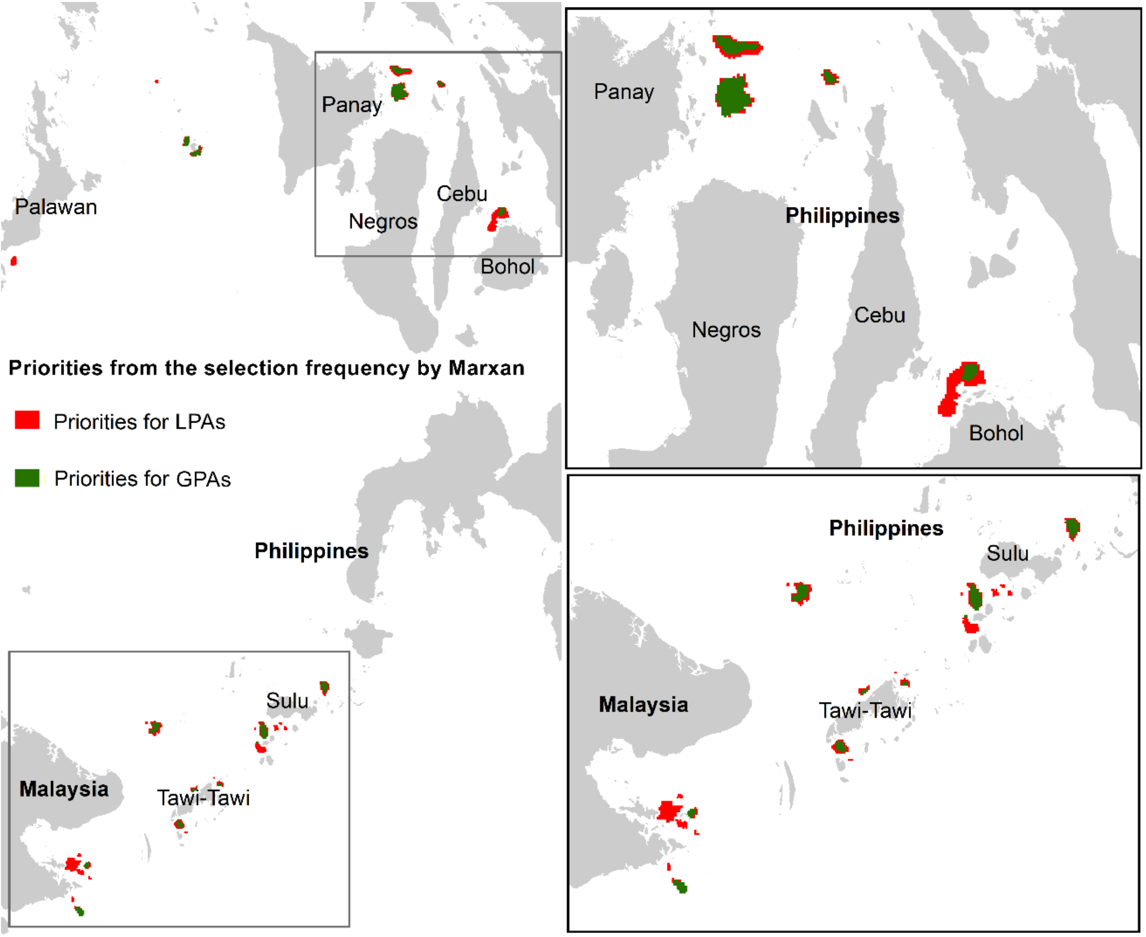


Figure K.4. Priorities from the selection frequency derived from the Marxan software for the Southeast Asia seahorse species (*Hippocampus barbouri*). LPAs, lower-protection areas (spatial disagreements between the biological conservation and diverse human uses); GPAs, greater-protection areas (spatial agreements between the biological conservation and diverse human uses). The detailed information about multiple-use areas were not shown because there were many classifications and difficult to illustrate by colors. Specific information can be found from the figshare database under the project folder "Global Priorities for Seahorse Species".

Figure K.5. Priorities for lower-protection areas (LPAs) from the best solution derived from the Marxan software for the two Australian seahorse species (*Hippocampus abdominalis & H. minotaur*). Lower-protection areas refer to spatial disagreements between the biological conservation and diverse human uses.

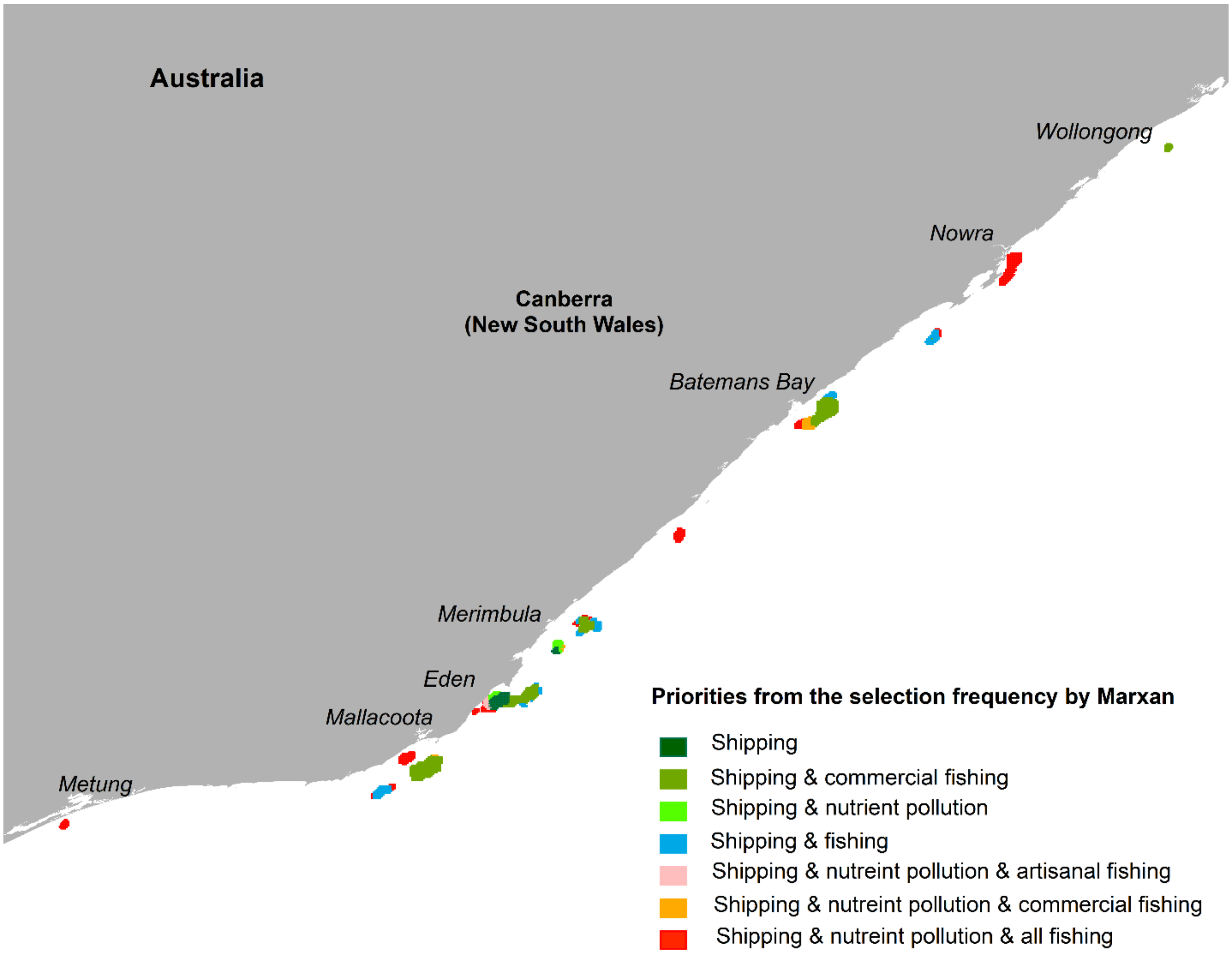


Figure K.6. Priorities for lower-protection areas (LPAs) from the selection frequency derived from the Marxan software for the two Australian seahorse species (*Hippocampus abdominalis & H. minotaur*). Lower-protection areas refer to spatial disagreements between the biological conservation and diverse human uses.

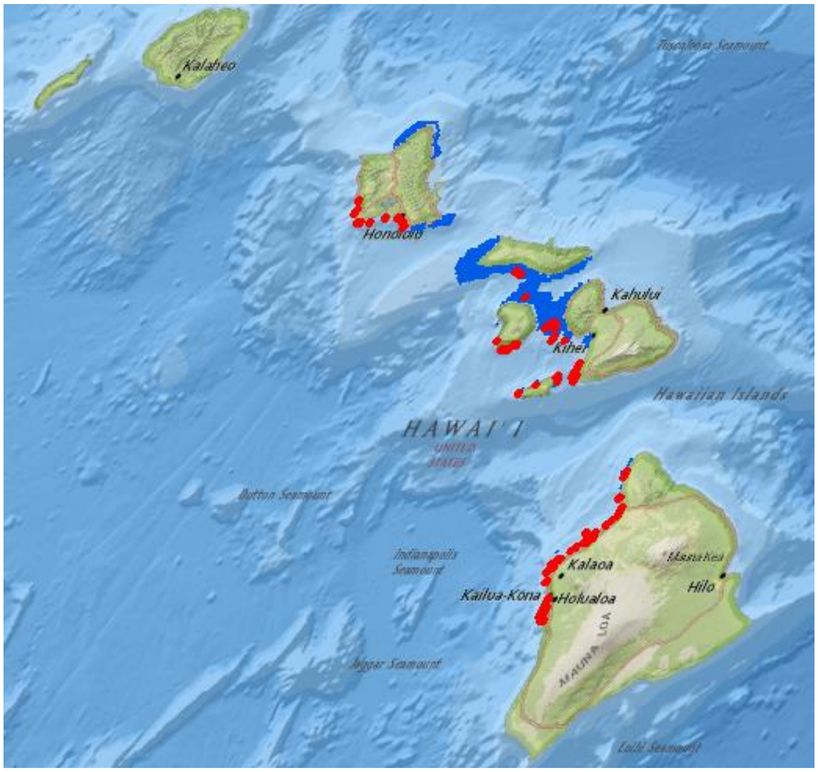


Figure K.7. Uncovered habitats (red) for the seahorse species endemic to Hawaii (*Hippocampus fisheri*), demonstrated with covered habitats of this and other species (*H. histrix* and *H. kuda*) by strict management areas (dark blue).

Figure K.8. Uncovered habitats (red) for the seahorse species endemic to South Africa (*Hippocampus capensis*), demonstrated with covered habitats by strict management areas (dark blue).

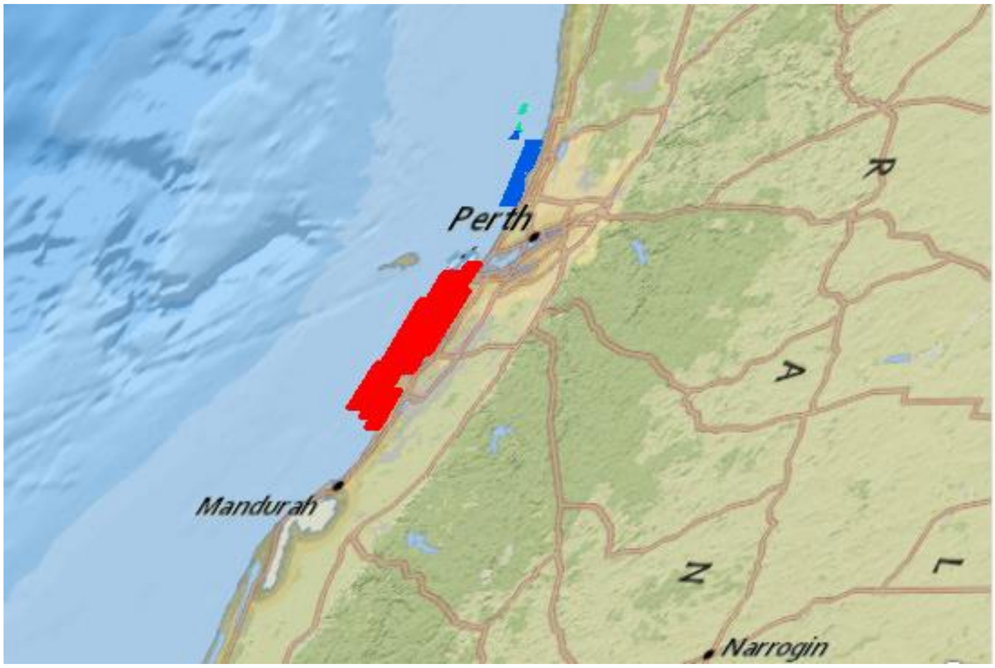


Figure K.9. Uncovered habitats (red) for the seahorse species endemic to West Australia (*Hippocampus subelongatus*), demonstrated with covered habitats by strict management areas (dark blue).

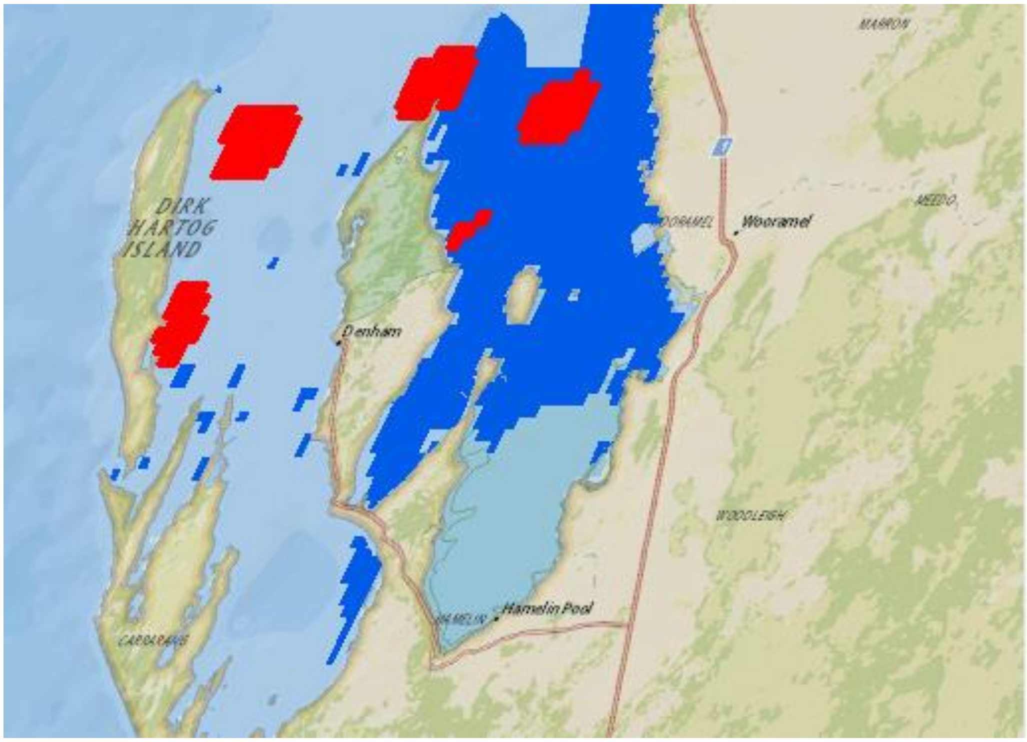


Figure K.10. Uncovered habitats (red) for the seahorse species endemic to Australia (*Hippocampus planifrons*), demonstrated with covered habitats by strict management areas (dark blue) of this and other species (e.g. *H. angustus*).

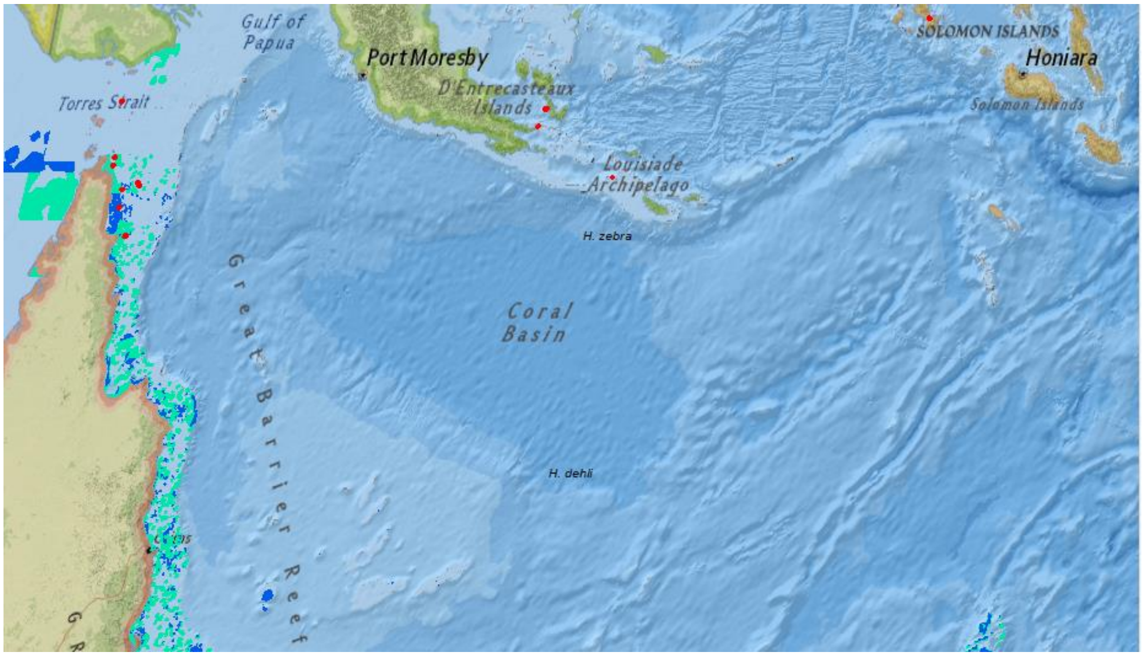


Figure K.11. Uncovered habitats (red) for the *Hippocampus angustus* and *H. denise*, demonstrated with covered habitats (by strict management areas in dark blue and by multiple-use areas in light green).

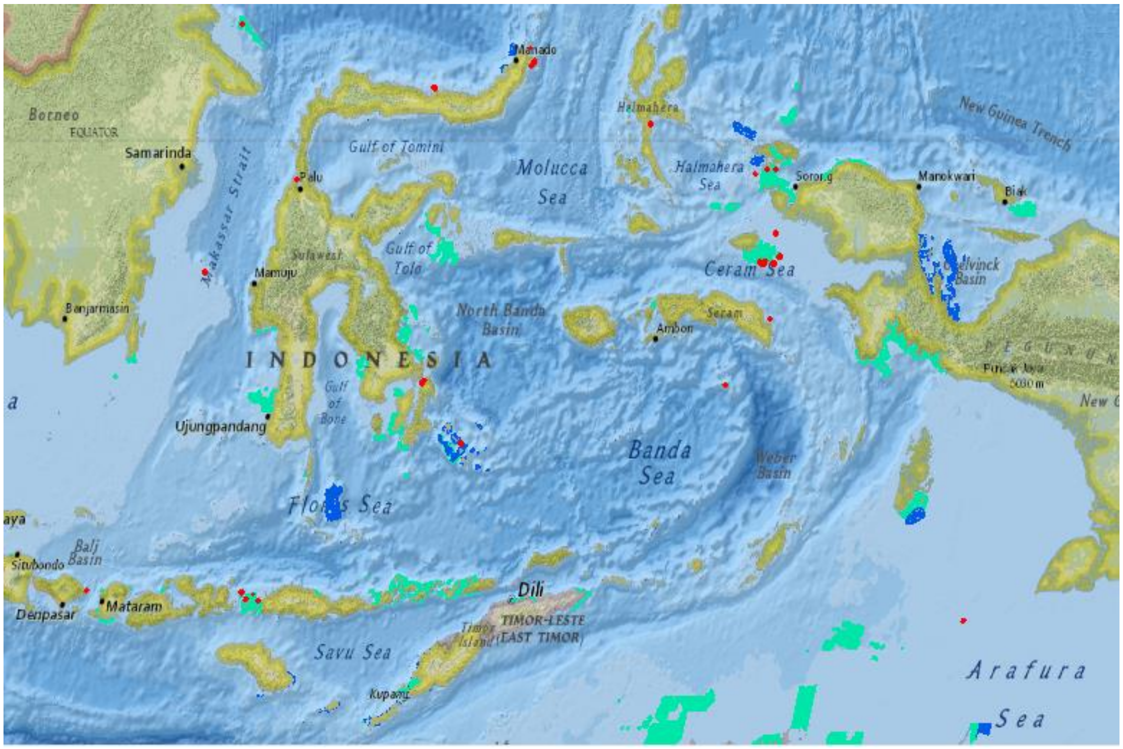


Figure K.12. Uncovered habitats (red) for the seahorse species endemic to Southeast Asia (*Hippocampus denise*), demonstrated with covered habitats (by strict management areas in dark blue and by multiple-use areas in light green).

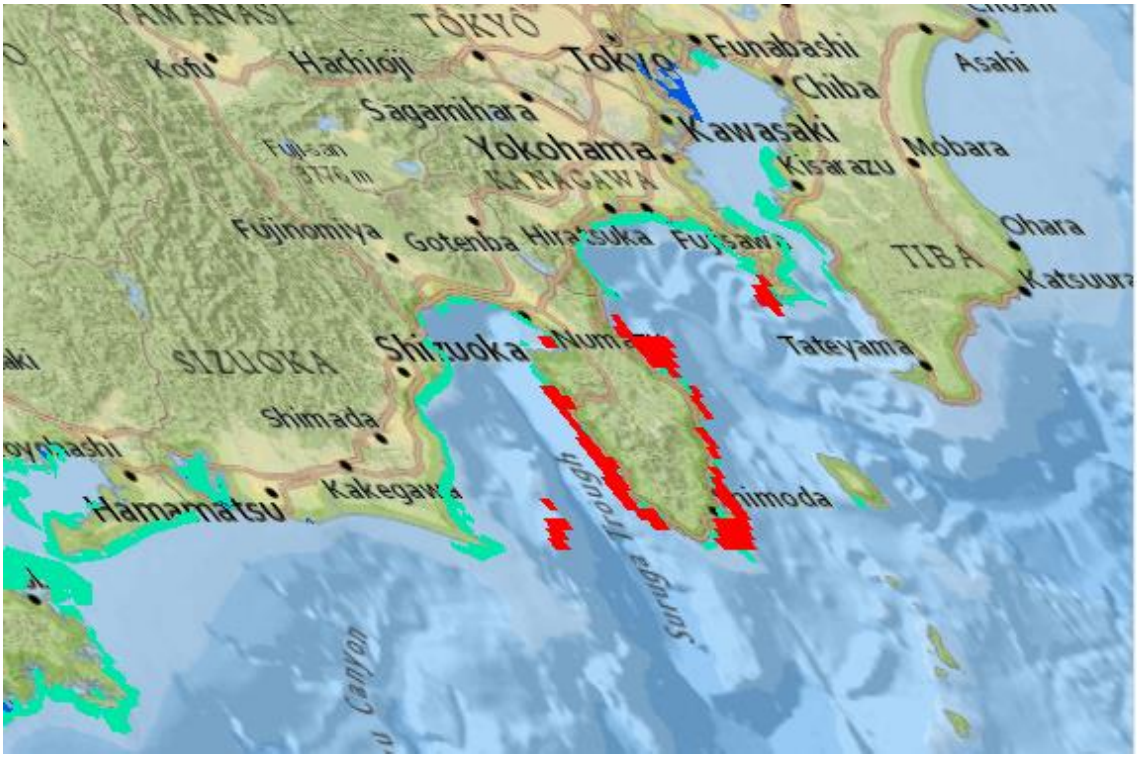


Figure K.13. Uncovered habitats (red) for the seahorse species mainly found in Japan (*Hippocampus sindonis*), demonstrated with covered habitats (by strict management areas in dark blue and by multiple-use areas in light green) of this species and other species (e.g. *H. coronatus*).