Mapping deep-sea biodiversity and good environmental status in the Azores: assisting with the implementation of EU Marine Strategy Framework Directive

Tese de Doutoramento

Gerald Hetcher Taranto

Doutoramento em

CIÊNCIAS DO MAR

Horta

2022

Mapping deep-sea biodiversity and good environmental status in the Azores: assisting with the implementation of EU Marine Strategy Framework Directive

Tese de Doutoramento

Gerald Hetcher Taranto

Orientadores

Doutor Telmo Alexandre Fernandes Morato Gomes, Okeanos Instituto de Investigação em Ciências do Mar da Universidade dos Açores

Professor doutor Gui Manuel Machado Menezes, Okeanos Instituto de Investigação em Ciências do Mar da Universidade dos Açores

Tese especialmente elaborada para obtenção do grau de Doutor em Ciências do Mar

“If we have learned anything over the last few decades of deep-sea exploration, it is that we still have much to learn” – Mcclain, 2010

Esta bolsa de doutoramento com a referência M3.1.a/F/052/2015 foi atribuída pelo Fundo Regional de Ciência e Tecnologia (FRCT) do Governo Regional dos Açores, cofinanciada pelo programa operacional regional dos Açores – Fundo Social Europeu, programas de formação avançada: doutoramentos e pós-doutoramentos, com o nº da operação FSE: ACORES-10-5369-FSE-000002. O trabalho de investigação apresentado foi co-financiado pelos projetos europeus ATLAS – "ATLAS "A Trans-Atlantic Assessment and deep-water ecosystem-based spatial management plan for Europe (H2020/project 678760), iAtlantic "Integrated Assessment of Atlantic Marine Ecosystems in Space and Time" (H2020/project 818123) financiados pelo European Union's Horizon 2020 research and innovation programme e MapGES "Mapping deep-sea biodiversity and "Good Environmental Status" in the Azores: assisting with the implementation of EU Marine Strategy Framework Directive (Acores-01-0145-FEDER-000056) no ambito dos projetos aprovados no programa operacional dos açores 2020.

Host Institution:

Institute of Marine Research, Azores (IMAR)

Marine and Environmental Sciences Centre (MARE)

OKEANOS Instituto de Investigação em Ciências do Mar da Universidade dos Açores

Acknowledgments

This long journey called Phd would not have been possible without the help of my collegues and supervisors in Horta and without the crews of the reseach vessels that helped collecting the data I used. The support of the many people I love including my Italian and my Azorean families and friends was essential to complete this journey. You are too many to list without forgetting some. I am very grateful to have you all.

Table of Contents

Abstract

One of the major shortfall of biodiversity knowledge steams from an incomplete description of the geographical distribution of species. Overcoming this shortfall is essential for conserving nature and its services and it is a required first step to tackle more complex ecological processes (e.g. dispersal, speciation, disturbance, biotic interactions, etc.) in remote and poorly studied regions such as the deep sea. In a region such as the Azores (NE Atlantic), where the deep sea represents a dominant component of the seascape, it is essential to characterize patterns and processes of deep-sea biodiversity. In fact, only by understanding how species and marine resources distribute it is possible to correctly inform area- and ecosystem-based management and achieve the goals of policies aiming at reversing the cycle of decline in ocean health. In particular, the European Commission has adopted a number of policies to grant a sustainable use of nature space and resources which include the Marine Strategy Framework Directive (MFSD) and the Maritime Spatial Planning Directive (MSPD). The overall goal of this thesis is to bring together existing and new biodiversity data from recent scientific surveys to deepen our understanding of biodiversity and biogeographic patterns of deep-sea Vulnerable Marine Ecosystems (VMEs) indicator taxa. The focus is on deep-sea hard-substrate communities of the Azores and, in particular, on ecosystem engineer species of the Phyla Cnidaria and Porifera. Four major environmental drivers of deep-sea benthic engineer species are recognized in the Azores: (*i*) a latitudinal gradient in primary production strongly influenced by the Azores Current-Azores Front (AzC-AzF) system; (*ii*) the depth-wise succession of the regional water masses and their stratification into different isopycnal (vertical) layers; (*iii*) the spatial distribution of prominent geomorphic features such as seamounts ridges and island slopes; (*iv*) the availability of hard substrate for attachment. The recognition of these environmental drivers sets an interesting background for future ecological research, ecosystem-based management and spatial monitoring. The response of deep-sea species to these environmental drivers is explored in detail in the different chapters of the present manuscript.

Keywords:

Atlantic, biogeography, deep sea, species turnover, vunerable marine ecosystems, zonation

Resumo

Um dos problemas fundamentais do conhecimento da biodiversidade deriva de uma descrição incompleta da distribuição geográfica das espécies. Superar esse déficit é essencial para a conservação da natureza e dos seus serviços e é um primeiro passo necessário para descrever processos ecológicos mais complexos (por exemplo, dispersão, especiação, perturbação, interações bióticas, etc.) em regiões remotas e pouco estudadas, como o mar profundo. Numa região como os Açores (Atlântico NE), onde o mar profundo representa uma componente dominante da paisagem marítima, é essencial caracterizar padrões e processos que influenciam a biodiversidade do mar profundo. De fato, somente entendendo como as espécies e os recursos marinhos se distribuem é possível informar corretamente a gestão baseada em ecossistemas e em áreas marinhas protegidas e atingir os objetivos das políticas que visam reverter o ciclo de declínio da saúde dos oceanos. Em particular, a Comissão Europeia adotou uma série de diretivas para garantir um uso sustentável do espaço maritimo e dos recursos naturais que incluem a Diretiva-Quadro de Estratégia Marinha e a Diretiva de Planejamento Espacial Marítimo. O objetivo geral desta tese é reunir dados de biodiversidade existentes e de novas pesquisas científicas para aprofundar a nossa compreensão da biodiversidade e dos padrões biogeográficos de táxons indicadores de Ecossistemas Marinhos Vulneráveis de águas profundas. O foco está nas comunidades de substrato duro do fundo do mar dos Açores e, em particular, nas espécies engenheiras dos Filos Cnidaria e Porifera. São reconhecidos nos Açores quatro grandes fatores ambientais que influenciam a distribuição das espécies engenheiras bentónicas de profundidade: (i) um gradiente latitudinal na produção primária fortemente influenciado pelo sistema Corrente-Frente dos Açores (AzC-AzF); (ii) a sucessão vertical das massas de água regionais e sua estratificação em diferentes camadas isopicnais (verticais); (iii) a distribuição espacial de feições geomórficas proeminentes, como cordilheiras de montes submarinos e encostas de ilhas; (iv) a disponibilidade de substrato duro para fixação. O reconhecimento desses fatores ambientais estabelece um cenário interessante para futuras pesquisas ecológicas, gestão baseada em ecossistemas e monitoramento espacial. A resposta das espécies de águas profundas a esses fatores ambientais é explorada em detalhes nos diferentes capítulos do presente manuscrito.

Chapter 1
General Introduction

1.1 Overview

The deep seabed beyond continental shelf depths (≈200 m) is one of the largest biome on Earth covering an area of over 400 million km^2 (Danovaro et al., 2014). Its remoteness and immensity has both fascinated and puzzled many ecologists whose attempts to describe natural processes and patterns have been frustrated by a chronic scarcity of data (Howell et al., 2020; McClain and Hardy, 2010; Ramirez-Llodra et al., 2010). Over the last decade, technological advances in seafloor mapping and biological sampling have provided a better understanding of the importance of deep-sea biodiversity and of the ecological services it provides (Danovaro et al., 2017, 2014; Thurber et al., 2014). There is now a great awareness that the deep sea is under multiple stressors including climate change, resource extraction and pollution (Halpern et al., 2019; Levin et al., 2019; Sweetman et al., 2017). These threats can be particularly severe for key ecosystem engineers such as deep-sea corals and sponges generally regarded as components of vulnerable marine ecosystems (VMEs) (FAO, 2009; Hogg et al., 2010; OSPAR, 2010; Ragnarsson et al., 2017).

Multiple international organizations have proposed a number of policies to reverse the cycle of decline in ocean health. The United Nations General Assembly proclaimed the Decade of Ocean Science for Sustainable Development (2021-2030) (United Nations, 2018). The European Commission has adopted a number of policies to grant a sustainable use of nature space and resources. These policies include the Maritime Spatial Planning Directive (MSPD) (EU Directive 2014/89/EU) and the Marine Strategy Framework Directive (MSFD) (EU Directive 2008/56/EC). In particular, the MSFD was adopted to achieve a Good Environmental Status (GES) in the EU's marine waters and to protect resources of socio-economic interest (e.g. Danovaro et al., 2020). It includes the deep-sea waters, seafloor and sub-seafloor of the Exclusive Economic Zones (EEZ). In general, these policies call on the scientific community to conduct research that advances understanding of marine ecosystems facilitating in this way the adoption of area-based/ecosystem-based management approaches (Howell et al., 2020; Kirkfeldt et al., 2022). Although the exploration of the deep sea now proceeds at a fast pace (e.g. Dominguez-Carrió et al., 2021; Swanborn

et al., 2022; Wölfl et al., 2019), the data about deep-sea communities still cover a limited portion of the seafloor constraining adequate ecosystem-based management approaches (Folkersen et al., 2019; Howell et al., 2020; Townsend et al., 2018). One of the strategy to partially overcome this issue and to promote and direct new collections of samples is the use of ecological maps. Maps represent a powerful tool to synthetize and communicate information about a geographic space (Coetzee et al., 2021). In the context of marine ecology, they are widely used to capture the environmental diversity and the distribution of species and communities (Buhl-Mortensen et al., 2015; Hogg et al., 2016; Howell, 2010; Ismail et al., 2015; Longhurst, 2007; Sayre et al., 2017; Sonnewald et al., 2020; Watling et al., 2013; Woolley et al., 2020). In general, ecological maps can be defined as a visual model (simplification) of complex ecological phenomena. They rely heavily on physical, spatial, or biological surrogates to describe biodiversity assuming that the biophysical patterns and processes observed for certain taxonomic groups or within particular environmental structures have a general relevance for ecological systems (Brown et al., 2011). The use of continuous coverage data layers is essential as they provide a mean to interpret patchy in situ observations, correlate the presence or the intensity of ecological phenomena with certain conditions and project such interpretations over large geographic areas (Brown et al., 2011).

The current thesis aims at describing and mapping the distributional patterns and the environmental drivers of deep-sea hard-substrate communities in the Azores (NE Atlantic) with a special focus on ecosystem engineer species of the Phyla Cnidaria and Porifera. These species are key components of regional deep-sea communities that fall within the definition of vulnerable marine ecosystems (VMEs) (FAO, 2009; Hogg et al., 2010; OSPAR, 2010). Understanding how these important species distribute is of pivotal importance for effective area-based management and for the implementation of national and international policies aiming to reverse the cycle of decline in ocean health (e.g. Combes et al., 2021; Galparsoro et al., 2015; Gonçalves, 2021; Grehan et al., 2017; Kazanidis et al., 2020; Rice et al., 2012; Rueda et al., 2016; Tunnicliffe et al., 2020). In fact, these species have the characteristics of benthic foundation organisms (Crotty et al., 2019; Ellison, 2019). Acting as foundation species these taxa can be expected to mediate a crucial suite of biotic interactions essential to structure deep-sea benthic communities (Kikvidze and Callaway, 2009). Forming vulnerable marine ecosystems, they are, by definition, greatly affected by human induced disturbance (Carreiro-Silva et al., 2013; Clark and Tittensor, 2010; Morato et al., 2020c; Ragnarsson et al., 2017; Vieira et al., 2020). In the Azores human-induced stressors can be particular severe. Climate change is expected to be particularly intense at the latitudes of the Azores (Levin et al., 2020; Morato et al., 2020c). The incredible amount of lost fishing gears that

were observed in recent surveys suggest a high fishing pressure (unpublished data from Dominguez-Carrió et al., 2021) even at remote locations (Morato et al., 2021b). Finally, deep-sea mining will likely start in the region with detrimental consequences for benthic organisms (Lopes et al., 2019; Ribeiro et al., 2020; Van Dover et al., 2017).

1.2 Study area and environmental patterns

The Archipelago of the Azores consists of nine volcanic islands located close to the Mid-Atlantic Ridge (MAR) at the junction of the Eurasian, African, and North American plates (Figure 1.1) (Miranda et al., 2018; Santos et al., 1995). Regional waters span over 1 million km^2 and hold a range of seascapes that are quite unique when compared to other European seascapes. This region is characterized by a huge area of deep seafloor scattered across a complex array of plains, ridges, seamounts and island slopes (Mitchell et al., 2018; Peran et al., 2016; Tempera et al., 2012a). In general, these topographic features present large fractions of exposed bedrock and complex interactions with steady or variable flows, a setting that provides favorable conditions for thriving benthic communities dominated by filter feeders (Morato et al., 2021b; Rogers, 2018). As a result, the Azores host a great diversity of deep-sea corals and sponges that form important and vulnerable communities (Braga-Henriques et al., 2013; Carreiro-Silva et al., 2013; Morato et al., 2020c; Sampaio et al., 2019a; Somoza et al., 2020; Xavier et al., 2021). Ocean circulation patterns in the Azores are complex (Bashmachnikov et al., 2004; Caldeira and Reis, 2017; Santos et al., 1995). Large-scale circulation is dominated by eastward-flowing branches of the Gulf Stream: Gulf Stream and Gulf Stream Bifurcations in the north and the Azores Current (AzC) in the south. Eddies systems associated with these currents play an important role in the region (Caldeira and Reis, 2017; Sala et al., 2016). These flows interact with the regional topography and with local winds altering surface temperatures and primary production (Amorim et al., 2017; Bashmachnikov et al., 2004; Caldeira and Reis, 2017). Warmer and less productive waters are located south-east, colder and more productive waters are located north-west of the archipelago. These differences suggest that the Azores is a transitional region between subtropical and temperate waters. The latitudinal gradient in primary production appears to be strongly influenced by the Azores Current-Azores Front (AzC-AzF) system (Caldeira and Reis, 2017; Frazão et al., 2022; Fründt and Waniek, 2012; Riou et al., 2016; Schiebel et al., 2011).

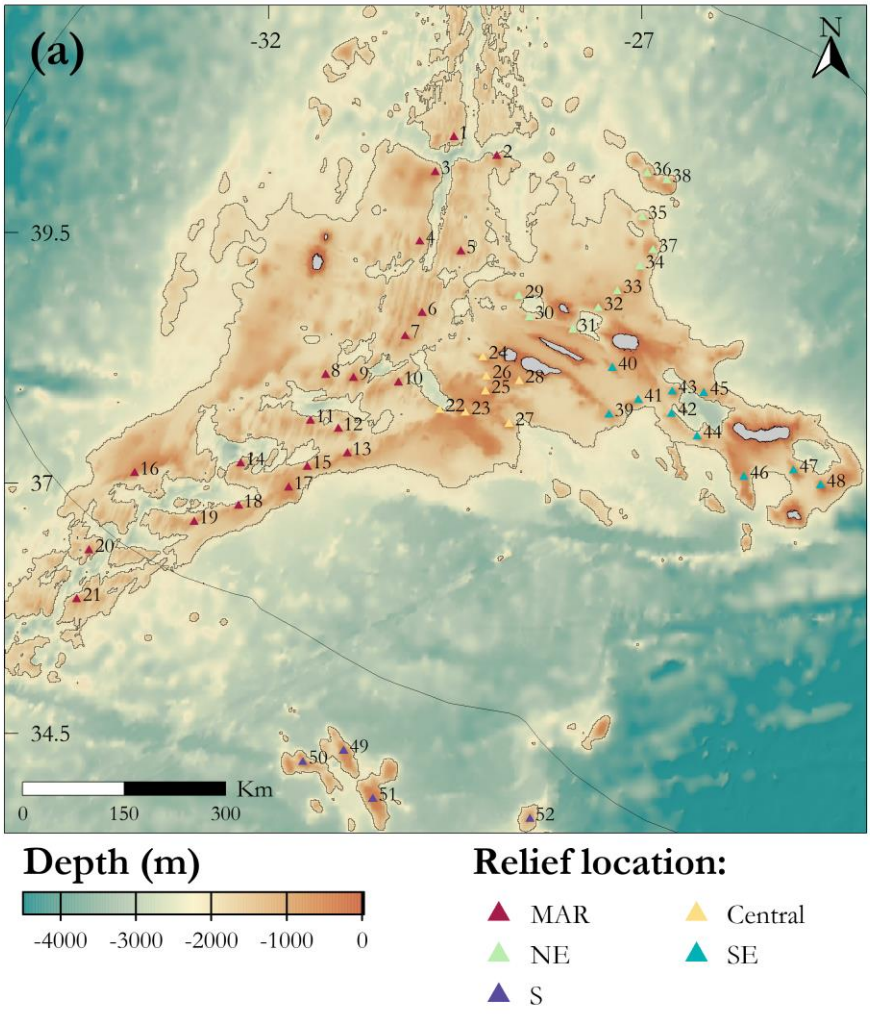

Figure 1.1 – Azores study area. The islands from west to east are: Flores, Corvo, Faial, Pico, São Jorge, Graciosa, Terceira, São Miguel and Santa Maria. Only seafloor locations shallower than 2000 m are considered in the present manuscript (identified by the contour lines). Relief location: Mid-Atlantic Ridge (MAR), Central Seamounts (Central), Northeastern Seamounts (NE), Southeastern Seamounts (SE), Southern Seamounts (S). Important regional relieves: (1) Kurchatov N, (2) Kurchatov SE, (3) Kurchatov SW, (4) Oscar W, (5) Oscar, (6) G127, (7) Gigante, (8) Beta, (9) Cavala, (10) A6, (11) Picoto, (12) Alpha, (13) Voador, (14) A3, (15) Voador W, (16) Sarda, (17) Monte Alto, (18) Cavalo, (19) Cavalo W, (20) SW1, (21) SW2, (22) Princess Alice W, (23) Princess Alice, (24) Condor, (25) Açor, (26) Condor de Fora, (27) De Guerne, (28) Baixo São Mateus, (29) Graciosa West, (30) Perestrelo Hill, (31) Ridge SE Graciosa, (32) Mar da Fortuna, (33) Serreta, (34) João Leonardes, (35) Gaillard, (36) Sedlo, (37) Borda, (38) Esperance, (39) Albatroz, (40) São Jorge de Fora, (41) Agostinho, (42) Dom João de Castro S, (43) Dom João de Castro, (44) Mar da Prata W, (45) Alcatraz, (46) Mar da Prata, (47) Fouque, (48) Formigas, (49) Atlantis N, (50) Atlantis NW, (51) Atlantis, (52) Tryo.

Based on the work of Liu and Tanhua (2021), it is possible to identify the principal water masses around the Azores and approximate their vertical extent. In the Atlantic the different water masses are distributed in four main isopycnal (vertical) layers: upper, (ii) intermediate, (iii) deep and overflow and (iv) bottom layer. In the Azores, the influence of upper water masses goes down to about 1100 m. The Eastern North Atlantic Central Water (ENACW) has the widest vertical extent with its core at about 500 m depth. The Western North Atlantic Central Water (WNACW) and the Western South Atlantic Central Water (WSACW) are present within the upper 500 m of the water column. The intermediate layer is characterized by the Mediterranean Water (MW) and the Antarctic Intermediate Water (AAIW). Their core is at about 800 m depth. A third intermediate water mass, the Subarctic Intermediate Water (SAIW) has been reported in the northwestern portion of the study area (Bashmachnikov et al., 2015), however, given the lack of sampling stations in this area it was almost absent in the data provided by Liu and Tanhua (2021). In the deep and overflow layer, down to 2000 m, the Upper North Atlantic Deep Water (uNADW) is the predominant water mass with its core located at about 1600 m depth. It occupies a wide depth range in the north (500-2000+ m), while in the south it was observed only at depths below 1000 m. The Lower North Atlantic Deep Water (lNADW) and the Northeast Atlantic Bottom Water (NEABW) characterize the deepest portion of the considered bathyal range (200-2000 m) and extend well below 2000 m.

Given the scarcity of biological sampling below 2000 m, **this work focuses only on seafloor areas shallower than 2000 m**.

1.3 Hard-substrate benthic communities

This work focuses on large deep-sea habitat-forming suspension feeders growing on hard substrate (Phyla Cnidaria and Porifera) characteristic of vulnerable marine ecosystems (VMEs) and builds on the many scientific surveys that have explored the deep sea of the Azores in recent years (e.g. Dominguez-Carrió et al., 2021; Morato et al., 2021a, 2020b, 2019, 2018b; Morato and Taranto, 2019; Orejas et al., 2017). The first part of the manuscript gives particular attention to cold-water coral (CWCs) taxa (Figure 1.2). Since the OSPAR convention adopted a list of threatened and/or declining species and habitats (Agreement 2008-6), and the United Nation General Assembly together with the Food and Agriculture Organization endorsed the concept of Vulnerable Marine Ecosystems (VMEs) (FAO, 2009), habitat-forming CWCs emerged as an important marine conservation target. CWCs belong to a diverse polyphyletic group of sessile cnidarians (classes Anthozoa and Hydrozoa) that generally inhabit deep, aphotic and cold marine waters. They are azooxanthellate (i.e., lacking symbiotic dinoflagellates), heterotrophs and characterized

by the production of a calcium carbonate or proteinaceous supporting axis (Cairns, 2007; Roberts et al., 2009). CWCs are distributed globally down to depths of several thousands of meters. Some species can form large and long lasting colonies, build extensive reefs or occur in dense 'deep-sea gardens'. The fascinating variety of CWC shapes and colors, together with their sensitivity to human disturbance, has put the deep sea in the spotlight, focusing society's attention, research funds and conservation actions (Armstrong et al., 2019). Increasing evidence suggests that CWC aggregations can influence local current regimes (Cyr et al., 2016; Juva et al., 2020; Mienis et al., 2019; Soetaert et al., 2016), regulate nutrient flows (Cathalot et al., 2015; Georgian et al., 2016; Middelburg et al., 2016; Rix et al., 2016), and provide habitat to other benthic and demersal species (Buhl-Mortensen et al., 2010, 2017; Klompmaker et al., 2016; Linley et al., 2017). Such roles in ecological facilitation, community development and promotion of local biodiversity have led to CWCs being considered foundation species central to deep-sea benthic communities (Crotty et al., 2019; Ellison, 2019; Orejas and Jiménez, 2017). Finally, CWCs likely contribute to regulating, supporting and provisioning ecosystem services (Thurber et al., 2014).

To date, 5.6% of the 3300+ CWC species estimated to exist worldwide (Roberts and Cairns, 2014) have been reported in the Azores, which occupies 0.25% of the world's ocean (Braga-Henriques et al., 2013; Sampaio et al., 2019a). Such high species diversity establishes this region as an important biodiversity hotspot of CWCs (Sampaio et al., 2019a). Most habitat-forming species belong to the orders Scleractinia (stony corals) and Antipatharia (black corals), the subclass Octocorallia (soft corals, gorgonians and sea pens) and the family Stylasteridae (lace corals) (Roberts et al., 2009). In the Azores the subclass Octocorallia reaches the highest species diversity recorded so far in the North Atlantic (Sampaio et al., 2019a). Recent video surveys performed with a portable drift cam system and with other underwater vehicles (Dominguez-Carrió et al., 2021; Morato et al., 2021a, 2020b, 2019, 2018b, 2018a; Orejas et al., 2017; Somoza et al., 2020) have reported that Octocorallia and Antipatharia species form dense aggregations both on island slopes and on offshore relieves (e.g. Morato et al., 2021b, 2020b; Somoza et al., 2020; Tempera et al., 2012b). Octocorallia and Antipatharia include some of the most important habitat-forming corals existing in the region. Reef-forming Scleractinia species have also been reported in the Azores (*Lophelia pertusa*, *Madrepora oculata* and *Solenosmilia variabilis*) (Braga-Henriques et al., 2013; Sampaio et al., 2012).

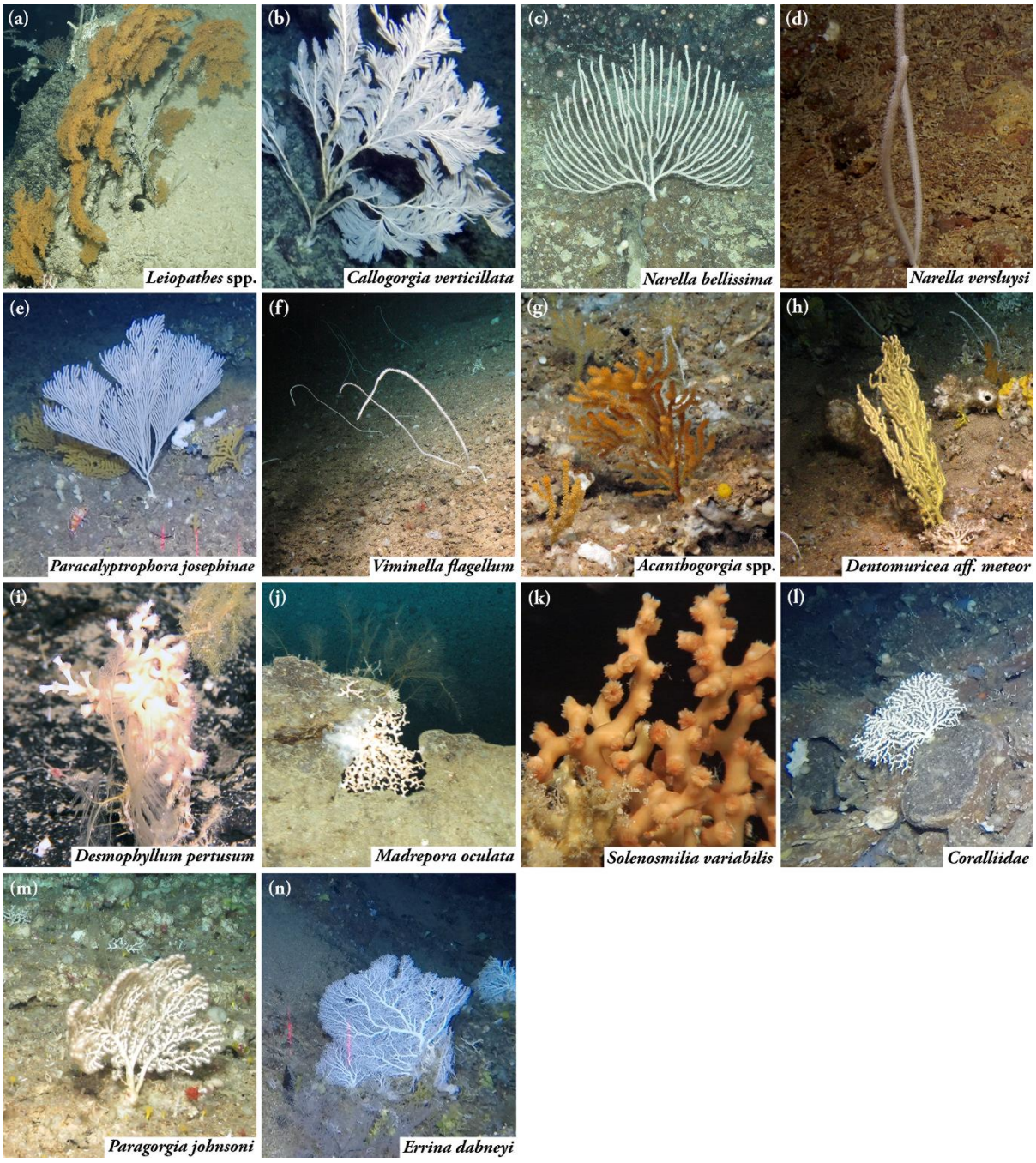

Figure 1.2 – Cold water coral taxa selected to develop habitat suitability models and to define the regional ecoscapes of the Azores. These species were selected considering their vulnerability to human activities and their potential as foundation species. © MEDWAVES, ATLAS project (*Leiopathes* cf. *expansa*, *N. bellissima*, *N. versluysi*, *D. pertusum* and *M. oculata*); © IMAR/OKEANOS-UAz, Drift camera (*C. verticillata*, *P. josephinae*, Coralliidae sp., *E. dabneyi*); © ROV Luso / EMEPC / 2018 Oceano Azul Expedition, organized by Oceano Azul Foundation & partners (*V. flagellum*, *Acanthogorgia* sp. and *D.* aff. *meteor*); © EMEPC/Luso/Açores/2013 (*S. variabilis*).

The video footage currently available shows that in the region, down to 2000 m, these species mostly exist as single colonies, resembling the categories II and III of Flögel et al. (2014). However, they are not in complete isolation, they rather tend to be embedded within multi-species coral gardens. Finally, in the Stylasteridae family there is the only habitat-forming coral species that appears to be endemic to the Azores, *Errina dabneyi* (Braga-Henriques et al., 2011; Zibrowius and Cairns, 1992).

The second part of the manuscript integrates CWC positional data with deep-sea sponge (Phylum Porifera) positional data. Species of the Phyla Cnidaria and Porifera represent one of the most important constituents of regional hard-substrate communities. Sponges are common members of many marine benthic communities and they may form dense aggregations (i.e. sponge grounds) that can be constituted by single or mixed species assemblages (Hogg et al., 2010; Maldonado et al., 2015). Sponge grounds contributes to the provision of various ecosystem goods and services (e.g. habitat provision, benthic–pelagic coupling, biogeochemical cycling, and biotechnological potential) (Afoullouss et al., 2022; Bart et al., 2021; Beazley et al., 2013; Henderson et al., 2020; Hogg et al., 2010; Maldonado et al., 2015). Thus far, the characterization of deep-sea sponges in the Azores is still in a preliminary phase and much of the work has focused on the order Lithistida (Carvalho et al., 2020; Xavier et al., 2021). However, great efforts have been placed in the taxonomic characterization of this group and new works should soon be available in literature.

1.4 Bioregionalization

The distribution of species is a dynamic process influenced by several interacting drivers operating at different scales (D'Amen et al., 2017; Hortal et al., 2015; Myers and Giller, 1988). Understanding how these drivers operate and how biodiversity changes over space and time is a central theme in ecology and biogeography and it is essential for nature management. This task is quite challenging in the context of the deep sea, the most remote and the largest ecosystem on Earth (Folkersen et al., 2018; Glover et al., 2018; Morato et al., 2013; Smith et al., 2020). Species-environment relationships tend to be easier to measure at regional scales and are often the only proxy readily available to estimate the distribution of species. In this study several bioregionalization approaches are used to understand how key species of hard-substrate communities respond to regional environmental gradients. Bioregionalization is the process of identifying bioregions which are defined by Woolley et al. (2020) as "a model of the true distribution of multiple species that share a similar ecological and abiotic preference and sometimes an evolutionary history". A bioregion covers a particular geographic area mapping simultaneously biological

and physical phenomena. It defines a meaningful spatial unit that can be used as a basis for research, ecosystem-based management and spatial monitoring (Howell et al., 2020; Rice et al., 2011; Whittaker et al., 2005). The bioregionalization approaches used are: habitat suitability models, species archetype models and an expert-driven ecological zonation.

1.5 Objectives and thesis outline

The main objective of this thesis is to understand how regional environmental drivers influence deep-sea benthic communities. The focus is on habitat-forming benthic organisms characteristics of hard-substrate communities. Many of the described patterns are of general interest as they seem to be caused by environmentally driven processes that most likely affect the functioning and the structure of several components of deep-sea ecosystems in the Azores. Understanding the distributional patterns of foundation and vulnerable taxa is of pivotal importance to adequately inform ecosystem-based management. In addition, by understanding the major distributional patterns of these species it is possible to better design monitoring strategies and to better frame new ecological questions.

This thesis is organized in several chapters:

Chapter 2 employs habitat suitability models (HSMs) to estimate the regional distribution of the most common and vulnerable habitat-forming CWCs of the Azores. Geographic and bathymetric distributional gradients are explored in the light of regional environmental patterns.

Chapter 3 represents one of the first statistical bioregionalization of deep-sea benthic communities in the Azores (NE Atlantic). It focuses on habitat-forming filter feeder species (Phyla Cnidaria and Porifera) known to represent a critical component of hard-substrate communities in the region. Species archetype models were developed to simultaneously analyze and combine biological and environmental data into meaningful bioregions following a one-stage classification approach.

Chapter 4 presents a new R package *scapesClassification* designed to translate users' perceptions of landscapes and seascapes into computer representation. Particular points of views are encoded in a transparent and explicit manner, can be reproduced on different machines and, ultimately, can be considered as task- or domain-oriented ontologies.

Chapter 5 provides a synthesis of the distributional patterns and of the environmental drivers of deep-sea hard-substrate communities in the Azores (NE Atlantic) with a special focus on ecosystem engineer species of the Phyla Cnidaria and Porifera. It builds on the results of Chapter 2 and 3. The software developed in

Chapter 4 is used to perform a geomorphological classification of prominent seafloor features (i.e. seamounts, ridges, island slopes). It investigates the role of the Azores Current-Azores Front (AzC-AzF) system in structuring deep-sea biodiversity patterns.

Finally, Chapter 6 provides a general overview of the main results obtained by the present work.

Chapter 2
Infer bathymetric and geographic gradients of key cold-water corals in the Azores (NE Atlantic): strengths and caveats of habitat suitability models

2.1 Abstract

Habitat-forming cold-water corals (CWCs) represent a key component of deep-sea communities and a priority target for conservation. Although mounting efforts are trying to identify drivers of CWC distributions much remain unknown because of the scarcity of available ecological data. The present work employs habitat suitability models (HSMs) to estimate the regional distribution of the most common and vulnerable habitat-forming CWCs of the Azores, a hotspot of coral diversity in the Atlantic. Geographic and bathymetric distributional gradients are explored in the light of regional environmental patterns. Our results suggest that the Azores Current could define two ecological provinces in the study area. The depth zonation of CWCs is shaped by regional water masses. In particular, the modelled CWCs form four groups of species: species restricted to upper waters, species extending down from upper waters, species restricted to intermediate waters and species extending up from deep waters. These results have important implications for the regional management of deep-sea benthic communities and, in particular, for the design of representative networks of protected areas. A careful assessment of our models provided four take home messages that can help in the design and in the interpretation of future models targeting deep-sea species: (i) models can have a different capacity to correctly detect distributional patterns along geographic and bathymetric gradients; (ii) the lack of readily available variables acting as a meaningful proxy of the pelagic–benthic coupling hinders the detection of important geographic patterns; (iii) unconventional variables mimicking known regional structures or processes can enhance model predictions; (iv) variables with low model importance can be of great importance in the interpretation of model results.

2.2 Introduction

Data about the distribution of species, especially vulnerable and foundation species, are important to inform marine management strategies and to better understand ecological and evolutionary processes. However, in the deep sea distributional data are scarce. Thus, habitat suitability models (HSMs) are often used to estimate geographic distributions by correlating available distributional data to environmental conditions (Elith and Leathwick, 2009). Although the relation of organisms with the abiotic environment is only one of the factors shaping species ranges (D'Amen et al., 2017), species environment relationships tend to be easier to measure at regional scales and are often the only proxy readily available to estimate the distribution of species in space and time. The abiotic environment constitutes one of the ecological filters driving community assembly (D'Amen et al., 2017). At a regional scale, environmental filtering can be considered as an intermediate process between those defining the regional pool of species and those structuring realized communities (D'Amen et al., 2017; Kraft et al., 2015). In particular, the term environmental filtering, *sensu stricto*, refers to abiotic conditions that prevent a species from establishing and surviving (Kraft et al., 2015). Only species able to tolerate similar environmental conditions can coexist at a site. Although it is challenging to distinguish environmental filtering from other ecological and evolutionary processes using observational data (Cadotte and Tucker, 2017), it should be possible to deduce a direct or an indirect role of the environment in structuring regional biodiversity if the environmental conditions of a region show a clear structure to which distinct pools of species correspond. This correlation should be easier to detect considering sessile species such as cold-water corals (CWCs).

Since the OSPAR convention adopted a list of threatened and/or declining species and habitats (Agreement 2008-6), and the United Nation General Assembly together with the Food and Agriculture Organization endorsed the concept of Vulnerable Marine Ecosystems (VMEs) (FAO, 2009), habitat-forming CWCs emerged as an important marine conservation target. CWCs belong to a diverse polyphyletic group of sessile cnidarians (classes Anthozoa and Hydrozoa) that generally inhabit deep, aphotic and cold marine waters. They are azooxanthellate (i.e., lacking symbiotic dinoflagellates), heterotrophs and characterized by the production of a calcium carbonate or proteinaceous supporting axis (Cairns, 2007; Roberts et al., 2009). CWCs are distributed globally down to depths of several thousands of meters. Some species can form large and long lasting colonies, build extensive reefs or occur in dense 'deep-sea gardens'. The fascinating variety of CWC shapes and colors, together with their sensitivity to human disturbance, has put the deep sea in the spotlight, focusing society's attention, research funds and conservation actions (Armstrong et al., 2019). Increasing evidence suggests that CWC

aggregations can influence local current regimes (Cyr et al., 2016; Juva et al., 2020; Mienis et al., 2019; Soetaert et al., 2016), regulate nutrient flows (Cathalot et al., 2015; Georgian et al., 2016; Middelburg et al., 2016; Rix et al., 2016), and provide habitat to other benthic and demersal species (Buhl-Mortensen et al., 2010, 2017; Klompmaker et al., 2016; Linley et al., 2017). Such roles in ecological facilitation, community development and promotion of local biodiversity have led to CWCs being considered foundation species central to deep-sea benthic communities (Crotty et al., 2019; Ellison, 2019; Orejas and Jiménez, 2017). Finally, CWCs likely contribute to regulating, supporting and provisioning ecosystem services (Thurber et al., 2014).

To date, 5.6% of the 3300+ CWC species estimated to exist worldwide (Roberts and Cairns, 2014) have been reported in the Azores, which occupies 0.25% of the world's ocean (Braga-Henriques et al., 2013; Sampaio et al., 2019a). Such high species diversity establishes this region as an important biodiversity hotspot of CWCs (Sampaio et al., 2019a). Most habitat-forming species belong to the orders Scleractinia (stony corals) and Antipatharia (black corals), the subclass Octocorallia (soft corals, gorgonians and sea pens) and the family Stylasteridae (lace corals) (Roberts et al., 2009). In the Azores the subclass Octocorallia reaches the highest species diversity recorded so far in the North Atlantic (Sampaio et al., 2019a). Recent video surveys performed with a portable drift cam system and with other underwater vehicles (Dominguez-Carrió et al., 2021; Morato et al., 2021a, 2020b, 2019, 2018b, 2018a; Orejas et al., 2017; Somoza et al., 2020) have reported that Octocorallia and Antipatharia species form dense aggregations both on island slopes and on offshore relieves (e.g. Morato et al., 2021b, 2020b; Somoza et al., 2020; Tempera et al., 2012b). Octocorallia and Antipatharia include some of the most important habitat-forming species existing in the region. Reef-forming Scleractinia species have also been reported in the Azores (*Lophelia pertusa*, *Madrepora oculata* and *Solenosmilia variabilis*) (Braga-Henriques et al., 2013; Sampaio et al., 2012). The video footage currently available shows that in the region, down to 2000 m, these species mostly exist as sparse colonies, resembling the categories II and III of Flögel et al. (2014). However, they are not in complete isolation, they rather tend to be embedded within multi-species coral gardens. Finally, in the Stylasteridae family there is the only habitat-forming coral species that appears to be endemic to the Azores, *Errina dabneyi* (Braga-Henriques et al., 2011; Zibrowius and Cairns, 1992).

The present work employs habitat suitability models (HSMs) to estimate the regional distribution of the most common and vulnerable habitat-forming CWCs of the Azores. Geographic and bathymetric distributional gradients are explored in the light of regional environmental structures. Since the modelled species represent a structuring and a prevalent element of the benthic communities in the region,

understanding how they distribute and how their composition change along environmental gradients will improve policies attempting to preserve vulnerable marine ecosystems (VMEs) and will help defining new ecological questions. Based on a careful review of model predictions several take home messages are provided. Hopefully, these messages will help in the design and in the interpretation of future models targeting deep-sea species. To our knowledge, this the first work to specifically model key cold-water corals in the Azores and the first work that specifically modelled the octocorals *Acanthogorgia* spp., *Dentomuricea* aff. *meteor*, *Narella bellissima*, *Narella versluysi*, *Paracalyptrophora josephinae* and *Paragorgia johnsoni*, the black coral *Leiopathes* cf. *expansa* and the lace coral *Errina dabneyi*.

2.3 Materials and Methods

2.3.1 Study area

Spanning over 1 million km^2 on the northern Mid-Atlantic Ridge (MAR) (Figure 2.3 a-b), the Azores hold a range of seascapes unique to European waters. Volcanic islands, ridges and seamounts rise from plains several kilometers deep and support different habitats across all the regional bathymetric range (Mitchell et al., 2018; Peran et al., 2016; Tempera et al., 2012b). Oceanography in the Azores is influenced by two eastward currents branching from the Gulf Stream: The North Atlantic Current in the north and the Azores Current in the south (Bashmachnikov et al., 2009; Caldeira and Reis, 2017). Eddies systems associated with these currents play an important role in the region (Caldeira and Reis, 2017; Sala et al., 2016). The permanent influence of the Azores high pressure system and the consequent sustained Elkman transport determine a confluence zone with different regimes northeast and southwest of the Archipelago (Caldeira and Reis, 2017) with potential implication for larval dispersal, marine colonization and speciation (Sala et al., 2016). Surface waters present a northwest-southeast gradient in productivity and other environmental parameters, with cooler and more productive waters located north and warmer less productive waters located south of the archipelago (Amorim et al., 2017; Caldeira and Reis, 2017; Palma et al., 2012). This environmental gradient is correlated to the spatial distribution of distinct planktonic assemblages in the region (Silva et al., 2013). Given the scarcity of biological sampling below 2000 m, this work focuses only on seafloor areas shallower than 2000 m (Figure 2.3c).

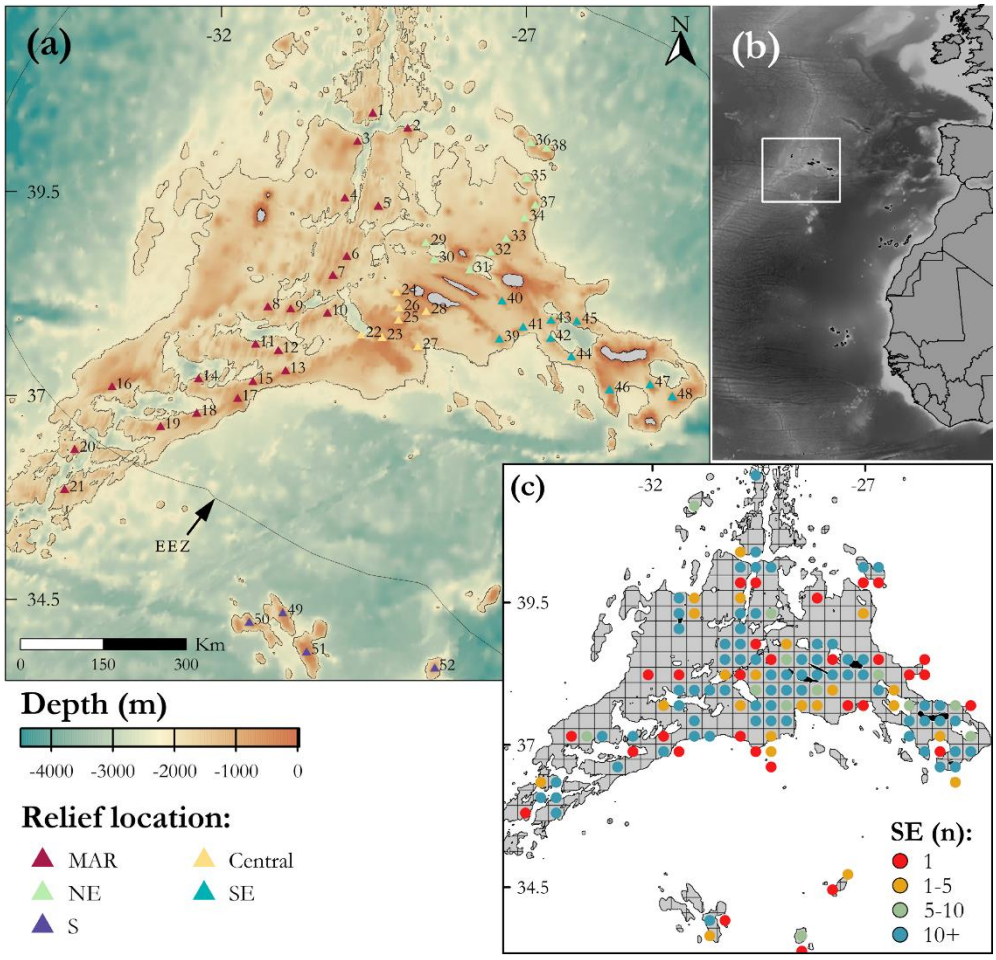

Figure 2.3 – Azores study area (a). The islands from west to east are: Flores, Corvo, Faial, Pico, São Jorge, Graciosa, Terceira, São Miguel and Santa Maria. Habitat suitability models consider only seafloor locations shallower than 2000 m (identified by the contour lines). Relief location: Mid-Atlantic Ridge (MAR), Central Seamounts (Central), Northeastern Seamounts (NE), Southeastern Seamounts (SE), Southern Seamounts (S). (b) Position of the Azores in the NE Atlantic. (c) Number of sampling events (SE) within the modelled area aggregated on a 30x30 km grid. Sampling events include video transects, longline surveys and bycatch data from fishing observers. Important regional relieves: (1) Kurchatov N, (2) Kurchatov SE, (3) Kurchatov SW, (4) Oscar W, (5) Oscar, (6) G127, (7) Gigante, (8) Beta, (9) Cavala, (10) A6, (11) Picoto, (12) Alpha, (13) Voador, (14) A3, (15) Voador W, (16) Sarda, (17) Monte Alto, (18) Cavalo, (19) Cavalo W, (20) SW1, (21) SW2, (22) Princess Alice W, (23) Princess Alice, (24) Condor, (25) Açor, (26) Condor de Fora, (27) De Guerne, (28) Baixo São Mateus, (29) Graciosa West, (30) Perestrelo Hill, (31) Ridge SE Graciosa, (32) Mar da Fortuna, (33) Serreta, (34) João Leonardes, (35) Gaillard, (36) Sedlo, (37) Borda, (38) Esperance, (39) Albatroz, (40) São Jorge de Fora, (41) Agostinho, (42) Dom João de Castro S, (43) Dom João de Castro, (44) Mar da Prata W, (45) Alcatraz, (46) Mar da Prata, (47) Fouque, (48) Formigas, (49) Atlantis N, (50) Atlantis NW, (51) Atlantis, (52) Tryo.

2.3.2 Regional water masses

The different water masses existing in the Atlantic Ocean have been recently described by Liu and Tanhua (2021) using conservative (temperature and salinity) and non-conservative (oxygen, silicate, phosphate and nitrate) water properties. Based on their work, at least five water masses can be identified in the Azores (Appendix A.1.1, Figure A.1). Regional water masses present a complex stratification and co-occur within specific depth intervals. They can be divided into different layers: upper, intermediate and deep layer. Within the upper 400 m of the water column, the upper water masses are prevalent. Below this depth, their importance decrease slowly down to a depth of 1200 m where their signature disappears. Intermediate water masses roughly occupy depths between 400 and 1600 m. Their core is at about 800 m depth. The core of the deep water masses is located at about 1600 m depth and they extend from the deepest portion of the study area (2000 m) up to depths of 400 m.

2.3.3 Selection of CWC species and occurrence data

Traits characterizing relative vulnerability and ecological importance were used to define the final list of modelled CWCs (Table 2.1 and Figure A.2) following the approach described in Appendix A.1. The modelled species represent the most common coral engineers of the region. Preliminary analysis of underwater videos showed that all the selected species form monospecific or mixed coral gardens in the Azores (e.g. Dominguez-Carrió et al., 2021; Morato et al., 2021b, 2018b). In addition, they are vulnerable to human activities (Carreiro-Silva et al., 2013; Morato et al., 2021b), appear to influence nutrient regimes (Rovelli et al., 2022, 2015), provide essential habitat to fish species (Gomes-Pereira et al., 2017) and one species (*Errina dabneyi*) seems to be endemic to the region (Braga-Henriques et al., 2011; Zibrowius and Cairns, 1992). Models were built at the genus or at the family level when species identifications were uncertain (see Appendix A.1).

Presence records were obtained from the Azores' Marine Biological Reference Collection and database (COLETA) which compiles information from three main sources: (i) bycatch obtained from regional longline surveys on-board the N/I "Arquipélago"; (ii) bycatch collected under regional observer programs; (iii) species annotations from underwater imagery (Table A.2) (Sampaio et al., 2019b). Data from literature compilations (Braga-Henriques et al., 2013) were excluded to reduce positional errors (Araújo et al., 2019). Most of the presence records came from video annotation (76%), the remaining from longline and handline bycatch records (24%) (Figure A.3 and Figure A.4). The georeferenced quality of bycatch data was increased by: (i) retaining only presence records associated to georeferenced segments shorter than

1.5 km so that positional errors were comparable to the resolution of the model grid (1.13 x 1.11 km) and (ii) discarding bycatch records with no reported depth or with a depth mismatching our reference bathymetric grid by more than 200 m. After removing duplicate records falling in the same model grid cell, the number of records ranged from 23 (*Errina dabneyi*) to 161 (*Viminella flagellum* and *Acanthogorgia* spp.) (Table 2.1 and Figure A.3 and Figure A.4). For most species the number of records was well above 25, a conservative minimum threshold for the development of HSMs (van Proosdij et al., 2016). An estimate of the regional sampling effort was produced combining all the scientific surveys that occurred in the region (Figure 2.3).

Table 2.1 – Cold-water coral taxa selected for the modelling exercise. Relative abundance, 3D-structure and relative vulnerability were the three criteria used to select species with high vulnerability and with the characteristics of a foundation species. Evaluations were expert-driven, relative to each other and referred only to colonies occurring in the Azores. N: Number of records.

Group	Taxon	Relative Abundance	3D-structure	Relative Vulnerability	N
Antipatharia	*Leiopathes* cf. *expansa*	Medium	Large	High	46
Calcaxonia	*Callogorgia verticillata*	Medium	Large	High	98
Calcaxonia	*Narella bellissima*	High	Medium	Low	42
Calcaxonia	*Narella versluysi*	High	Small	Low	58
Calcaxonia	*Paracalyptrophora josephinae*	Medium	Large	High	24
Calcaxonia	*Viminella flagellum*	High	Small	Low	161
Holaxonia	*Acanthogorgia* spp. [1]	High	Medium	Low	161
Holaxonia	*Dentomuricea* aff. *meteor*	Medium	Large	High	51
Scleractinia	*Desmophyllum pertusum* [2]	Low	Large	High	31
Scleractinia	*Madrepora oculata*	Low	Large	High	63
Scleractinia	*Solenosmilia variabilis*	Low	Medium	High	24
Scleraxonia	Coralliidae [3]	Medium	Medium	High	128
Scleraxonia	*Paragorgia johnsoni*	High	Large	High	51
Stylasteridae	*Errina dabneyi*	Medium	Medium	High	23

[1] In the Azores it includes *A. armata* and *A. hirsuta*.
[2] Resulted from a recent combination of the genus *Desmophyllum* and *Lophelia* and is used as a synonym of the well-kwon reef-forming *Lophelia pertusa*.
[3] In the Azores it includes *Pleurocorallium johnsoni*, *Hemicorallium niobe* and *Hemicorallium tricolor*.

2.3.4 Environmental predictors

The candidate environmental predictors considered in this study are presented in Table 2.2. Resource and direct predictors were preferred as explanatory variables (Araújo et al., 2019), but indirect predictors were considered when the former were not available. A cell size of a 1.13 x 1.11 km (UTM zone 26N) was adopted for the analysis. This resolution was considered a good compromise between the original resolution of environmental data and our capacity to resolve suitable and unsuitable areas within the same geomorphological feature using model predictions. Grid harmonization required the downscale of depth-derived variables and the upscale of environmental variables using bilinear interpolation.

Following the approach adopted in similar studies (Anderson et al., 2016; Davies and Guinotte, 2011; Yesson et al., 2012), depth-derived variables were computed using the Benthic Terrain Modeler toolbox (Walbridge et al., 2018) and were considered as environmental predictors. Nutrients (silicate, phosphate and nitrate) and carbonate compounds (aragonite and calcite) are often used to model the distribution of CWCs (Davies and Guinotte, 2011; Yesson et al., 2012). In our case they were all highly correlated (Figure A.5). Principal Component Analysis (PCA) was used to reduce the high collinearity (Dormann et al., 2013). However, when two PCA were performed independently for nutrients and carbonate compounds their first principal components remained highly correlated ($r > 0.9$). Based on the fact that different water masses present different nutrient and carbonate concentrations (e.g. Azetsu-Scott et al., 2010; Fontela et al., 2020; González-Dávila et al., 2011; Liu and Tanhua, 2021), it was decided to join these variables into the same PCA. The first principal component (named 'seawater chemistry') explained 90% of the variance. It was used as a predictor variable with the intention to detect an eventual correlation existing between the distribution of CWCs and the distribution of different water masses. Figure A.6 shows the relations of 'seawater chemistry' with nutrient concentrations and with aragonite and calcite saturation states. Undersaturated conditions are located below the modelled depth range (0-2000 m) (Figure A.6e). The particulate organic carbon (POC) flux was computed based on Lutz et al. (2007) using the net primary productivity values reported in Amorim et al. (2017).

Correlation and importance in preliminary models provided the criteria to identify the final set of variables used in the HSMs (Table 2.2). Correlation among predictors (Figure A.5) was maintained at acceptable levels ($r \leq 0.7$) (Dormann et al., 2013) by removing those variables less meaningful for coral distribution, as inferred from the literature. Temperature and 'seawater chemistry' ($r = 0.72$) were the exceptions as they both appeared to be relevant in preliminary model outputs. Depth was excluded because it represented an indirect predictor showing a high correlation with other variables (Figure A.5).

Table 2.2 – Environmental layers considered in this study. The final set of layers used to develop habitat suitability models is presented in bold. Mean and SD refer to seafloor values at locations shallower than 2000 m. (‡) Depth-derived layers computed in ArcGIS using the Benthic Terrain Modeler toolbox (Walbridge et al., 2018); (†) Seafloor conditions derived from the VIKING20 oceanographic model computed as the mean of monthly values for the period 1989-2009 (Böning et al., 2016). BPI is the bathymetric position index (*sensu* Walbridge et al., 2018); POC is the particulate organic carbon flux computed based on Lutz et al. (2007). All layers from Amorim et al. (2017), Böning et al. (2016) and Morato et al. (2020c) refers to seafloor conditions. BPI is bathymetric position index, POC is particulate organic carbon.

Available predictors	Original resolution	Mean and SD	Units	Source
Depth	115 m	-1507 ± 415	m	EMODnet, 2018
BPI (5 km radius)‡	115 m	29.6 ± 107.5	-	-
BPI (20 km radius) ‡	115 m	158.4 ± 316.3	-	-
Slope‡	115 m	6.2 ± 4.9	°	-
Ruggedness (5km radius)‡	115 m	0.0016 ± 0.0025	-	-
Eastness‡	115 m	0.00 ± 0.55	-	-
Northness‡	115 m	0.01 ± 0.56	-	-
Temperature†	0.05°	5.6 ± 2.5	°C	Böning et al., 2016
Salinity†	0.05°	26.40 ± 0.19	PSU	Böning et al., 2016
Current speed†	0.05°	0.01 ± 0.008	m s^{-1}	Böning et al., 2016
Aragonite	1°	1.33 ± 0.27	Ω	Morato et al., 2020
Calcite	1°	2.07 ± 0.43	Ω	Morato et al., 2020c
Nitrate	1°	17.74 ± 1.90	µmol/l	Amorim et al., 2017
Phosphate	1°	1.14 ± 0.13	µmol/l	Amorim et al., 2017
Silicate	1°	12.64 ±1.85	µmol/l	Amorim et al., 2017
Dissolved Oxygen	1°	5.58 ± 0.51	ml/l	Amorim et al., 2017
Oxygen Saturation	1°	79.40 ± 4.76	%	Amorim et al., 2017
Oxygen Utilization	1°	1.44 ± 0.29	ml/l	Amorim et al., 2017
Seawater chemistry	1°	3.01 ± 1.74	-	-
POC flux	9 km	14.92 ± 1.52	mg C_{org} m^{-2} d^{-1}	-

Based on preliminary models, we excluded predictors that consistently showed little importance (*sensu* Thuiller et al., 2009) (eastness, northness and ruggedness). The final set of predictors included: bathymetric position index – a measure of the relative elevation of a location above the seafloor – computed with two different radius (5 km and 20 km) (see Lundblad et al., 2006; Walbridge et al., 2018), seafloor current speed, seawater chemistry, seafloor oxygen saturation, POC flux, slope and seafloor temperature (Table 2.2). These layers are plotted in Figure A.7.

2.3.5 Software

All analyses were performed in the R environment (version 4.2.0) (R Core Team, 2022).

2.3.6 Habitat suitability models (HSMs)

Habitat suitability models (HSMs) were built using two approaches: generalized additive models (GAMs) (Hastie and Tibshirani, 1990) and maximum entropy algorithms (Maxent) (Phillips et al., 2017). Models were built using the functions 'gam' from the R package *mgcv* (Wood, 2011) and 'maxent' from the R package *dismo* (Hijmans et al., 2022). GAMs were computed using the binomial family and a maximum basis dimension (k) equal to 4. An attempt to change the default parameter of the function 'maxent' was made, however, no improvement in model predictions could be detected. Therefore, default parameters were used. The general workflow followed to develop HSMs and combine their results is synthetized in Figure 2.4. Study area and model background were limited to depths shallower than 2000 m where most of the sampling events took place (Figure 2.3) so that the sampling-bias could be mitigated (Phillips et al., 2009).

***Pseudo-absences and background points*.** Models were trained using a sampling-bias corrected set of 10,000 pseudo-absences and background points (following Barbet-Massin et al., 2012; Hertzog et al., 2014; Phillips and Dudík, 2008). As the sampling-bias was mostly depth related, raster locations were grouped into five depth strata and the probability of selection of pseudo-absences and background points was correlated to the sampling effort that occurred in each depth stratum. The geographic region for the generation of GAM pseudo-absences was further limited using environmental profiling based on presence data (function 'OCSVMprofiling' from the R package *MOPA*, Iturbide et al. 2018). The nu (or v) parameter of 'OCSVMprofiling' controls the error percentage of suitable locations misclassified as unsuitable by a support vector machine classifier – the higher the value of nu, the higher is the error tolerance (Drake et al., 2006; Schölkopf et al., 2001). The nu parameter was set to a value of 0.9 because with lower nu values pseudo-absences were not generated over large regions and this caused less plausible model outputs.

Predictor variables and response curves. From the final set of predictor variables (Table 2.2), only a subset was used to build the GAMs, while all predictor variables were used for Maxent models. For each CWC taxa, GAMs were fitted using all the possible combinations of the predictor variables. The model presenting the lowest Akaike Information Criterion (AIC, Symonds and Moussalli, 2011) was selected (function dredge in the R package MuMIn, Bartoń, 2022). Variable importance was estimated following the permutation approach (n = 100) implemented by the function 'variables_importance' (R package *Biomod2*) described in Thuiller et al. (2022, 2009). Response curves were built considering changes in habitat suitability indices (HSI) when a focal predictor varied at 100 points across its range while all other predictors were kept at their mean value (Elith et al., 2005). Finally, we tried to assess if similar pools of variables contributed the most to individual GAM and Maxent models. Subsets of variables were defined based on two conditions: (i) they ranked within the top four and (ii) they had a permutation importance greater than 10%. Shared important variables (SIVs) corresponded to those present in both Maxent and GAM subsets.

5-seed 5-fold spatial block cross validation. Spatial block cross-validation is considered the best practice when no independent datasets are available to validate model predictions (Araújo et al., 2019). Presence and pseudo-absences/background points were assigned to different spatial bocks using the function 'spatialBlock' (R package *blockCV*, Valavi et al., 2019). The function 'spatialAutoRange' was used to define the block size (Valavi et al., 2019). Each block was then assigned randomly to one of five folds. At turns, each fold is used for cross-validation and the remaining four are used for model fitting. The random grouping of spatial blocks into folds can influence the cross validation. In order to have more conservative evaluation scores, the entire process was repeated five times setting different random seeds (5-seed 5-fold spatial block cross-validation). Averaged values of area under the curve (AUC) and of true skill statistics (TSS) provided estimates of the overall confidence of model predictions (Allouche et al., 2006). According to these statistics, models can be classified as good (AUC > 0.8; TSS > 0.6), fair ($0.7 \leq AUC \leq 0.8$; $0.4 \leq TSS \leq 0.6$) or poor (AUC < 0.7; TSS < 0.4) (Landis and Koch, 1977). Final models were built using all presence records and produced binary predictions based on the sensitivity-specificity sum maximization (MSS) threshold (Liu et al., 2005).

Spatial autocorrelation. Spatial autocorrelation in model residuals was investigated using the global Moran's I coefficient (Dormann et al., 2007) computed with the R package *ape* (Paradis and Schliep, 2019). All Moran's I coefficients were close to zero and non-significant (p-value > 0.05).

Bootstrapping. A bootstrap process (n = 100) evaluated the local confidence of model predictions. Each bootstrap iteration sampled with replacement the input data, fitted GAM and Maxent models and produced binary suitability maps based on the MSS threshold (Liu et al., 2005). Local suitability confidence was then estimated by the number of times individual raster cells were classified as suitable: low [1-30%), medium [30 70%) or high [70 100%] confidence cell.

Fuzzy Matching and Fuzzy Kappa. The level of similarity between the spatial distribution of GAM and Maxent binary predictions (low, medium and high confidence cells) was measured using both a local (Fuzzy Matching) and an overall (Improved Fuzzy Kappa) measure of similarity (Hagen-Zanker, 2009; Hagen, 2003). Measures of local similarity specify for each raster cell the degree of similarity between GAM and Maxent predictions on a scale of 0 (totally different) to 1 (identical). These measures were based on two membership functions (Hagen, 2003): category similarity, which assumed that some categories were more similar than others (Table A.3); distance decay, which defined the fuzzy similarity of two cells as (i) identical if they matched perfectly, (ii) linearly decreasing with distance if the matching category was found within a 2-cell radius (~2 km) or (iii) totally different when no matching category was found within a 2-cell radius. Hagen (2003) provides the details to combine membership functions and compute local fuzzy similarity. The overall similarity of GAM and Maxent predictions was measured using two Improved Fuzzy Kappa scores (Hagen-Zanker, 2009): (i) the first measured the correspondence of all GAM and Maxent predictions (F_k); (ii) the second measured only the correspondence of the high-confidence cells (F_{kHC}). Areas consistently predicted as suitable with high confidence by both modelling approaches were considered the best suited to characterize and describe CWC habitats. Therefore, we decided to use the F_{kHC} index to measure the overall level of similarity of model predictions. Following the classification adopted for kappa type statistics, the overall similarity of habitat suitability predictions was considered as good ($F_{kHC} > 0.6$), fair ($04 \leq F_{kHC} \leq 0.6$) or poor ($F_{kHC} < 0.4$) (Landis and Koch, 1977).

Combined suitability maps and overall confidence in habitat suitability estimates. GAM and Maxent predictions were combined and each raster cell predicted as suitable was classified as follow:

i. high confidence suitable cell, raster cell predicted as suitable with high confidence by both GAM and Maxent models;
ii. medium confidence suitable cell, raster cell predicted as suitable with medium or high confidence by GAM, Maxent or both and with a local fuzzy similarity greater than 0.5;
iii. low confidence suitable cell, any other cell predicted as suitable by GAM or Maxent.

The overall confidence of combined suitability maps was defined as good, fair or poor based on the lowest score among AUC_{GAM}, TSS_{GAM}, AUC_{Maxent}, TSS_{Maxent} and F_{kHC} (Figure 2.4).

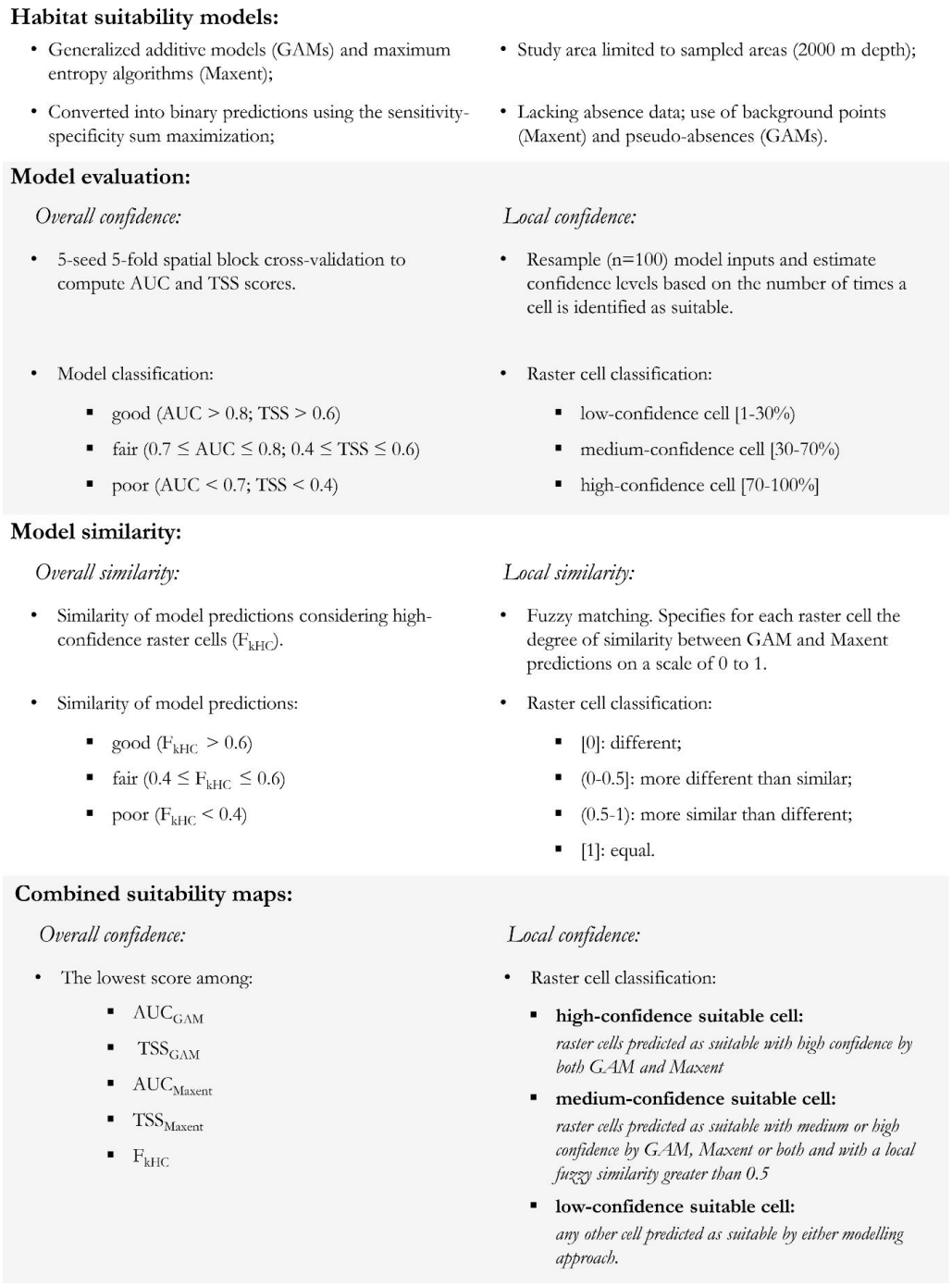

Figure 2.4 – Habitat suitability models design and evaluation. Area under the curve (AUC); true skill statistics (TSS); improved Fuzzy Kappa for high-confidence suitable cells (F_{kHC}).

2.3.7 Niche equivalence and similarity

Pair-wise niche overlaps were measured using the Schoener's D which ranges from 0 (no overlap) to 1 (complete overlap) (Warren et al., 2008). This index evaluates the degree of shared environmental niche space among distinct species or populations (e.g. Aguirre-Gutiérrez et al., 2015). In the present case it is used to identify niche equivalences and similarities based on the outputs of HSMs. Niches were quantified considering only the environmental conditions occurring at the high-confidence suitable cells of the combined suitability maps. Only variables describing water properties were used to quantify the niche space (i.e. seawater chemistry, POC flux, seafloor temperature, seafloor oxygen saturation and seafloor current speed). The idea was to verify if distinct pools of CWCs could be identified based on water properties alone.

All niche analyses were implemented using the R package *humboldt* (Brown and Carnaval, 2019). This package builds on the analysis proposed by Warren et al. (2008) and Broennimann et al. (2012). The Schoener's D was computed in the environmental space to reduce eventual geographic biases in habitat availability (Broennimann et al., 2012; Brown and Carnaval, 2019). The total environmental space was bounded by the minimum and maximum environmental values found across the suitable cells of all the modelled CWC taxa. The environmental dimensions were reduced using the principal component analysis and the first two components were considered to quantify the niches occupied by different species. The principal components were projected on a grid of 100 x 100 cells and a kernel density function was used to determine the smoothed density of occurrences in each cell of the grid (Broennimann et al., 2012; Brown and Carnaval, 2019). The analysis corrected the occurrence densities of each species by the prevalence of the environments in their range (Brown and Carnaval, 2019) and removed duplicate suitable cells in buffers of 5 km (distance chosen based on the spatial autocorrelation ranges of the environmental layers, function 'spatialAutoRange' from the R package *blockCV*, Valavi et al., 2019). The corrected densities of occurrences were used to measure the niche overlap of different species (Schoener's D). We used the equivalence statistic to assess whether the ecological niches of pairs of CWCs were significantly different from each other. The observed Schoener's D value was compared to values obtained when the occurrences of the two species were resampled (200 times). Significance was determined by the number of times the observed overlap was greater than the overlap of the reshuffled datasets. The hypothesis of niche equivalency was rejected with p-values < 0.05. We used the background statistic to assess if the niche overlap of two species was more similar than expected by chance given the underlying environmental conditions. The statistic compares the similarity of the niches of species 1 and 2 to the

similarity of species 1 and the random shifting of the spatial distribution of species 2 (and vice versa). The process is repeated 200 times and if the simulated overlap is consistently lower than the observed overlap (p value < 0.05) then the background test is significant. See Brown and Carnaval (2019) for more details about the equivalence and the background tests.

2.4 Results

2.4.1 Model evaluation

Evaluation scores are shown in Table 2.3. The overall confidence score was good for 8 taxa and fair for 6. No model showed a low evaluation score nor a low similarity score. Evaluation scores were within the fair range only for *Desmophyllum pertusum* (AUC_{GAM} = 0.78; AUC_{Maxent} = 0.79; TSS_{GAM} = 0.58; TSS_{Maxent} =0.59), *Leiopathes* cf. *expansa* (AUC_{GAM} = 0.76) and *Madrepora oculata* (TSS_{Maxent} = 0.59). Maxent models and GAMs showed very similar performances with Maxent models being only slightly better. Five species showed similarity scores within the fair range: *Errina dabneyi* (F_{kHC} = 0.58), *L*. cf. *expansa* (F_{kHC} = 0.58), *M. oculata* (F_{kHC} = 0.56), *Paracalyptrophora josephinae* (F_{kHC} = 0.54) and *Solenosmilia variabilis* (F_{kHC} = 0.54). On average, the overall similarity of model predictions increased when only high-confidence cells were considered (average F_{kHC} = 0.63; average F_k = 0.5). This suggests that similar areas are consistently predicted as suitable by both modelling approaches. The only exceptions were *E. dabneyi* and *S. variabilis* (F_{kHC} and F_k almost equivalent) and *P. josephinae* (F_{kHC} = 0.54; F_k = 0.63).

The local confidence of GAM and Maxent models for Octocorallia appeared to be between good and fair (Figure A.8 and Figure A.9). In general, more than half of the areas identified as suitable were predicted as suitable consistently throughout the bootstrap process. The remaining areas were predicted as suitable fairly consistently. The main exceptions were *Callogorgia verticillata* and *Dentomuricea* aff. *meteor* in GAMs (Figure A.8 a and g) and Coralliidae in Maxent (Figure A.9h). These were also the species with the lowest local agreement between models predictions (Figure A.10 a, g and h). The agreement between the two modelling approaches seemed to be good for the remaining taxa (Figure A.10). Interestingly, the overall similarity score for high confidence cells resulted quite low for *P. josephinae* (F_{kHC} = 0.54; F_k = 0.63) while the local fuzzy matching showed a good overlap in model predictions (Figure A.10d). The high fuzzy matching values show that the suitable areas predicted by the two models largely overlap, however the low F_{kHC} shows that the uncertainties of the two models are distributed differently.

Suitable habitats estimated by GAMs for Scleractinia, Antipatharia and Anthoathecata showed a proportion of high, fair and low confidence cells similar to the results obtained for the Octocorallia (Figure A.11). The only exception was *S. variabilis*, which presented a high percentage of fair confidence cells (Figure A.11c). Maxent models appeared to be more influenced by the resampling process (Figure A.12) and to overestimate the suitable habitat of *L.* cf. *expansa*, *M. oculata* and *S. variabilis* when compared to the suitability maps produced by GAMs (Figure A.11 and Figure A.12 panels b, c and e). Habitat suitability predictions had the lowest overlap for *L.* cf. *expansa* and *M. oculata* (Figure A.13). *E. dabneyi*, showed good local fuzzy matching (Figure A.13d) and poor fuzzy kappa scores (both F_k and F_{kHC}). This disagreement is likely caused by the fact that the two modelling approaches predict an important fraction of non-overlapping suitable cells that are in close proximity. While this influences negatively the overall fuzzy kappa scores, local measurements of similarity can still be high if suitable cells are within the distance considered by the distance decay component of the fuzzy membership function.

Table 2.3 – Area under the curve (AUC), true skill statistics (TSS) and fuzzy kappa scores. Highest scores between GAM and Maxent models in bold. N: number of records; F_k: improved fuzzy kappa; F_{kHC}: improved fuzzy kappa computed for high-confidence cells; OC: overall confidence score of combined habitat suitability maps with levels good (G), fair (F) or poor (P). The lowest score among AUC_{GAM}, TSS_{GAM}, AUC_{Maxent}, TSS_{Maxent} and F_{kHC} determines the overall confidence score.

TAXA	N	TSS_{GAM}	AUC_{GAM}	TSS_{Maxent}	AUC_{Maxent}	F_k	F_{kHC}	OC
Acanthogorgia spp.	161	0.69 ± 0.12	0.90 ± 0.07	0.69 ± 0.12	0.90 ± 0.06	0.66	0.74	G
C. verticillata	98	0.68 ± 0.09	0.90 ± 0.04	**0.69 ± 0.09**	0.90 ± 0.04	0.42	0.75	G
Coralliidae spp.	128	**0.75 ± 0.10**	**0.92 ± 0.05**	0.74 ± 0.10	0.91 ± 0.05	0.4	0.71	G
D. aff. *meteor*	51	0.76 ± 0.20	0.88 ± 0.14	**0.80 ± 0.14**	**0.91 ± 0.09**	0.38	0.64	G
D. pertusum	31	0.58 ± 0.21	0.78 ± 0.13	**0.59 ± 0.17**	**0.79 ± 0.10**	0.5	0.61	F
E. dabneyi	23	0.72 ± 0.23	0.85 ± 0.14	**0.74 ± 0.19**	0.85 ± 0.12	0.59	0.58	F
L. cf. *expansa*	46	0.60 ± 0.26	0.76 ± 0.24	**0.62 ± 0.22**	**0.80 ± 0.17**	0.32	0.58	F
M. oculata	63	**0.61 ± 0.15**	**0.82 ± 0.10**	0.59 ± 0.11	0.81 ± 0.09	0.42	0.56	F
N. bellissima	42	0.74 ± 0.17	0.89 ± 0.08	0.74 ± 0.16	**0.90 ± 0.07**	0.59	0.67	G
N. versluysi	58	0.73 ± 0.17	**0.88 ± 0.08**	0.74 ± 0.17	**0.89 ± 0.08**	0.53	0.62	G
P. josephinae	24	**0.86 ± 0.14**	0.93 ± 0.07	0.85 ± 0.13	0.93 ± 0.08	0.63	0.54	F
P. johnsoni	51	**0.83 ± 0.09**	**0.94 ± 0.04**	0.79 ± 0.08	0.93 ± 0.04	0.53	0.66	G
S. variabilis	24	0.66 ± 0.20	0.80 ± 0.13	**0.69 ± 0.20**	**0.82 ± 0.11**	0.47	0.46	F
V. flagellum	161	0.76 ± 0.07	0.93 ± 0.03	**0.79 ± 0.06**	**0.94 ± 0.03**	0.55	0.71	G
Average	69	0.71	0.87	0.72	0.88	0.5	0.63	-

2.4.2 Variable importance and response curves

The importance of variables for GAMs and Maxent models is shown in Figure 2.5. The most important variable was 'seawater chemistry' for both modelling approaches with a median permutation importance of 29.6% for GAMs and 36.5% for Maxent models. It ranked among the top variables for all models with the exception of the *Paracalyptrophora josephinae* and *Solenosmilia variabilis* (GAMs). Among the depth-derived layers the most relevant variables were BPI 20 (bathymetric position index) and slope. BPI 20 ranked among the top four variables with a permutation importance greater than 10% five times for GAMs and eight times for Maxent models, 'slope' five times for GAMs and 12 for Maxent models. Besides 'seawater chemistry', among the variables describing water conditions, particulate organic carbon (POC) resulted important for most Maxent models, albeit it was relevant only for three GAMs: *Desmophyllum pertusum*, *Madrepora oculata* and *Leiopathes* cf. *expansa*. Seafloor water temperature also seemed to be quite relevant for nine of the modelled species: Coralliidae (both HSMs), *D. pertusum* (both HSMs), *Dentomuricea* aff. *meteor* (both HSMs), *Errina dabneyi* (both HSMs), *L.* cf. *expansa* (Maxent), *Narella bellissima* (GAM), *P. josephinae* (GAM), *Paragorgia johnsoni* (Maxent), *Viminella flagellum* (both HSMs). The least important variables were BPI 5, current speed and oxygen saturation.

Response curves are shown in Appendix A.3, Figure A.14 to Figure A.21. In general, GAM and Maxent models showed comparable responses. Responses to the variable 'seawater chemistry' can be divided into two groups: 'group 1' associated to high 'seawater chemistry' values (*D.* aff. *meteor*, *E. dabneyi*, *P. josephinae*, *V. flagellum* and *Callogorgia verticillata* – Figure A.14a-e); 'group 2' associated to low 'seawater chemistry' values (*N. bellissima*, *Narella versluysi*, *P. johnsoni*, *L. expansa*, *D. pertusum*, *S. variabilis* and *M. oculata* – Figure A.14f-n). *Acanthogorgia* spp. and Coralliidae spp. seemed to have responses in between these two groups. In GAMs, the responses to temperature showed that three species of 'group 1' preferred a narrow temperature range (Figure A.15a-c): *D.* aff. *meteor*, *E. dabneyi* and *P. josephinae*. In general, Maxent responses showed wider tolerance ranges. *C. verticillata* appeared to have its optimal habitat slightly below 10 ºC but showed wider tolerance ranges than to the other species of 'group 1' (Figure A.15d). The optimal temperature for *V. flagellum* appeared to at about 8 ºC (Figure A.15e). The optimal temperature for the octocorals and the black coral of 'group 2' was roughly between 5 and 10 ºC. The scleractinian species either had an unimportant (*S. variabilis* and *M. oculata*) or a positive relationship with temperature (*D. pertusum*) (Figure A.15l-n). Particulate organic carbon (POC) flux was a shared important variable for three species: *L.* cf. *expansa*, *D. pertusum* and *M. oculata* (all showing a positive relationship with POC flux) (Figure A.16k-m). This variable had little importance for the remaining

GAMs while it was relevant for most Maxent models. In general, all taxa of 'group 2' avoided areas of low carbon fluxes (Figure A.16f-n) as well as *D.* aff. *meteor* of 'group 1' (Figure S3.15a). *P. josephinae* and *V. flagellum* had a bell-shaped response curve centered on low and intermediate flux values, respectively (Figure A.15c, e). POC flux was not relevant for the remaining models. The taxon Coralliidae was the only one to show a significant response to oxygen saturation in both HSMs, with suitability indices inversely correlated to saturation values (Figure A.17g). Oxygen saturation was the only variable with high permutation importance for the GAM model of *S. variabilis* (Figure A.17n). The Maxent model showed a negative response of *V. flagellum* to oxygen saturation (Figure A.17e). The habitat suitability index of almost all species was positively correlated with slope and fine-scale BPI values (Figure A.18 and Figure A.19). Most species preferred high coarse-scale BPI values (Figure A.20) with the exception of Coralliidae, *P. johnsoni*, and *M. oculata* associated both to high and low BPI values (Figure A.20 g, j, m). Bottom current speed appeared to have some relevance only for D. pertusum (Figure A.21l).

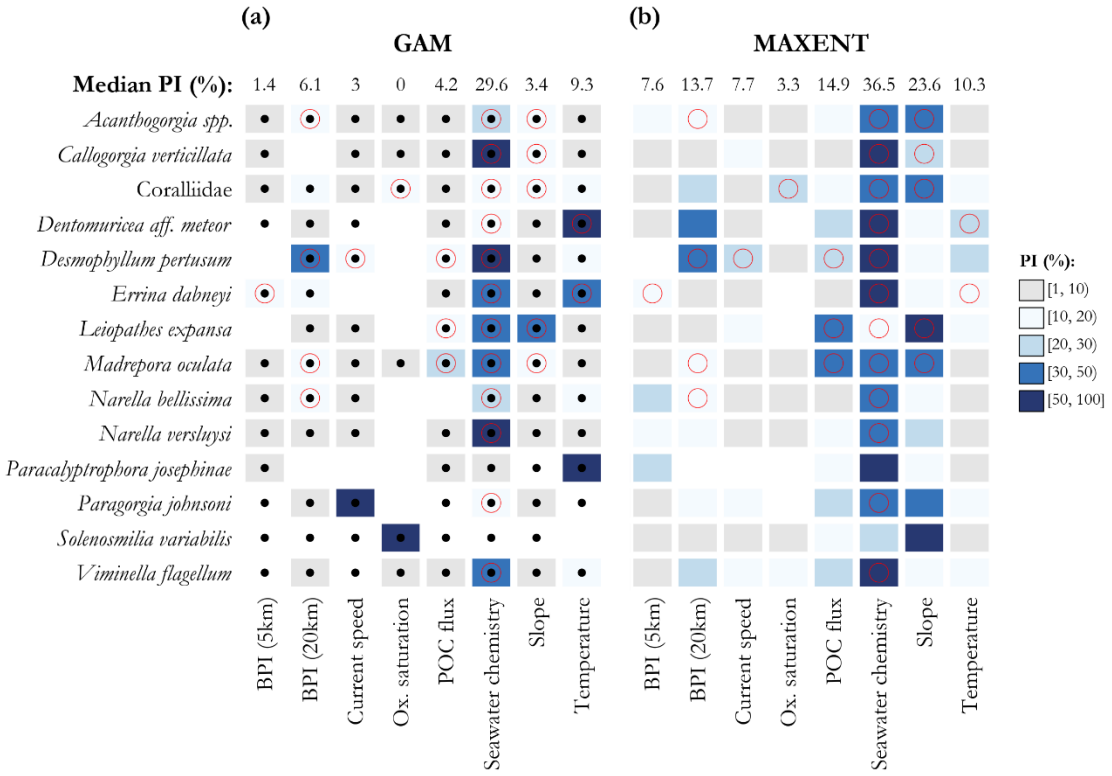

Figure 2.5 – Variable permutation importance (PI) for (a) GAM and (b) Maxent models. Black dots in (a) identify the variables used to build GAMs; all variables were used for Maxent models. Red circles identify the shared important variables: variables that ranked within the top four and that had a permutation importance greater than 10% in both GAM and Maxent models.

2.4.3 Geographic distribution

Considering high-confidence suitable cells of the combined suitability maps (Figure 2.6 and Figure 2.7), the potential habitat of the modelled CWCs covered about 23,000 km^2, the 11% of the modelled area. Most of this habitat was in waters shallower than 1500 m (Figure 2.8b) and it was always associated to major geological relieves (islands, seamounts or ridges) (Figure 2.6 and Figure 2.7). All species had patches of suitable habitat on both sides of the Mid-Atlantic Ridge (MAR), even though habitat suitable for *E. dabneyi* was almost absent west of the MAR (Figure 2.7a). The scleractinian corals *D. pertusum*, *M. oculata* and the black coral *L.* cf. *expansa* showed some latitudinal gradient with high-confidence suitable cells mostly located in the north-northwestern portion of the study area (Figure 2.7b-d and Figure 2.8a). Raster cells suitable for 9 or more species were rare and scattered between the MAR (on most of the features highlighted in Figure 2.3a), the Central Seamounts (especially Condor, Baixo de São Mateus, Princess Alice W and De Guerne) and the South Eastern Seamounts (especially São Jorge de Fora, Dom João de Castro, Mar da Prata and Mar da Prata W) (Figure A.22). Overall, the Northeastern and the Southern Seamounts did not appear to have the potential to aggregate many of the modelled CWCs. The only exceptions were Graciosa West and the Ridge SE Graciosa (Northeastern Seamounts) (see Figure 2.3a for the features' names). On island slopes, multi-species aggregations were more likely west of Corvo, southwest of Flores, west of Faial, Graciosa and Terceira, the westernmost and the easternmost tips of São Jorge, south and east of Pico and east and west of São Miguel. Santa Maria seemed to be the island less likely to aggregate the modelled CWCs.

Based on high-confidence suitable cells, the modelled taxa could be divided into four groups (1, 1.5, 2, 2.5). The first group (group 1) included the species presenting the smallest estimated distributions: *E. dabneyi*, *P. josephinae* and *D.* aff. *meteor* (Figure 2.6a-b and Figure 2.7a). The species showing the smallest distribution was *E. dabneyi*. It was mostly limited to the central group of seamount and islands and to the Southeastern Seamounts. Also *P. josephinae* models predicted a low proportion of suitable habitat, however it covered a greater area in the southern portion of the MAR and in the Southeastern Seamounts. It also covered part of the Southern Seamounts. The potential distribution of *D.* aff. *meteor* was similar to *P. josephinae* (Figure 2.6a), however it had larger patches of suitable habitats associated to island slopes, especially around São Miguel, Terceira, Graciosa, São Jorge, Flores and Corvo. It appeared to be absent in the Southern Seamounts and in Santa Maria. These three species were virtually absent from the Northeastern Seamounts and from the northernmost portion of the MAR.

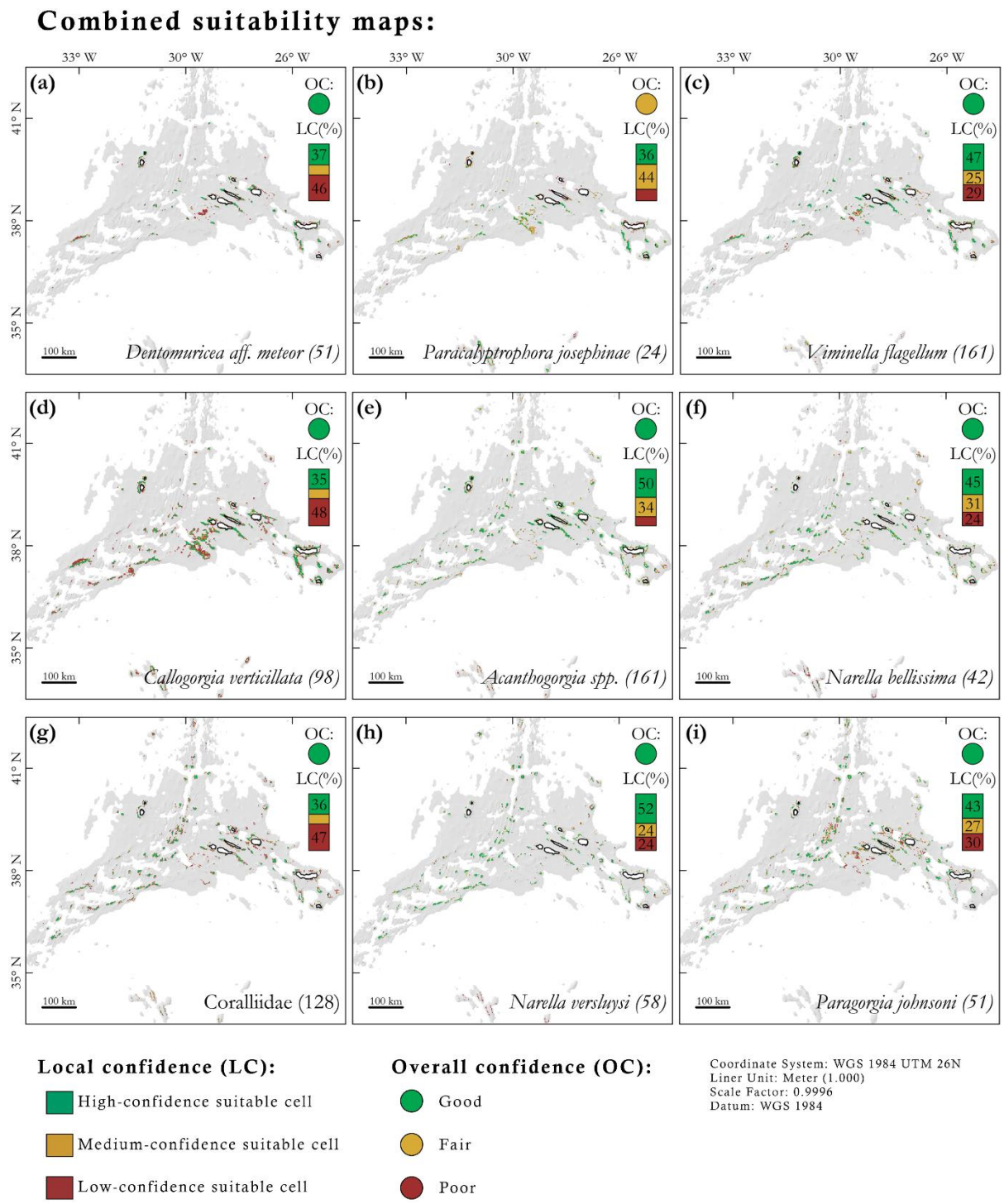

Figure 2.6 – Combined suitability maps for the selected octocoral (Alcyonacea) taxa. Local confidence (LC): the percentage of cells classified as high, medium or low confidence. Overall confidence (OC): determined by the lowest score among AUC_{GAM}, TSS_{GAM}, AUC_{Maxent}, TSS_{Maxent} and F_{kHC}. Area under the curve (AUC); true skill statistics (TSS); improved fuzzy kappa for high-confidence suitable cells (F_{kHC}). In parenthesis the number of records used to fit the models. Species are ordered according to the mean depth of their estimated suitable habitat.

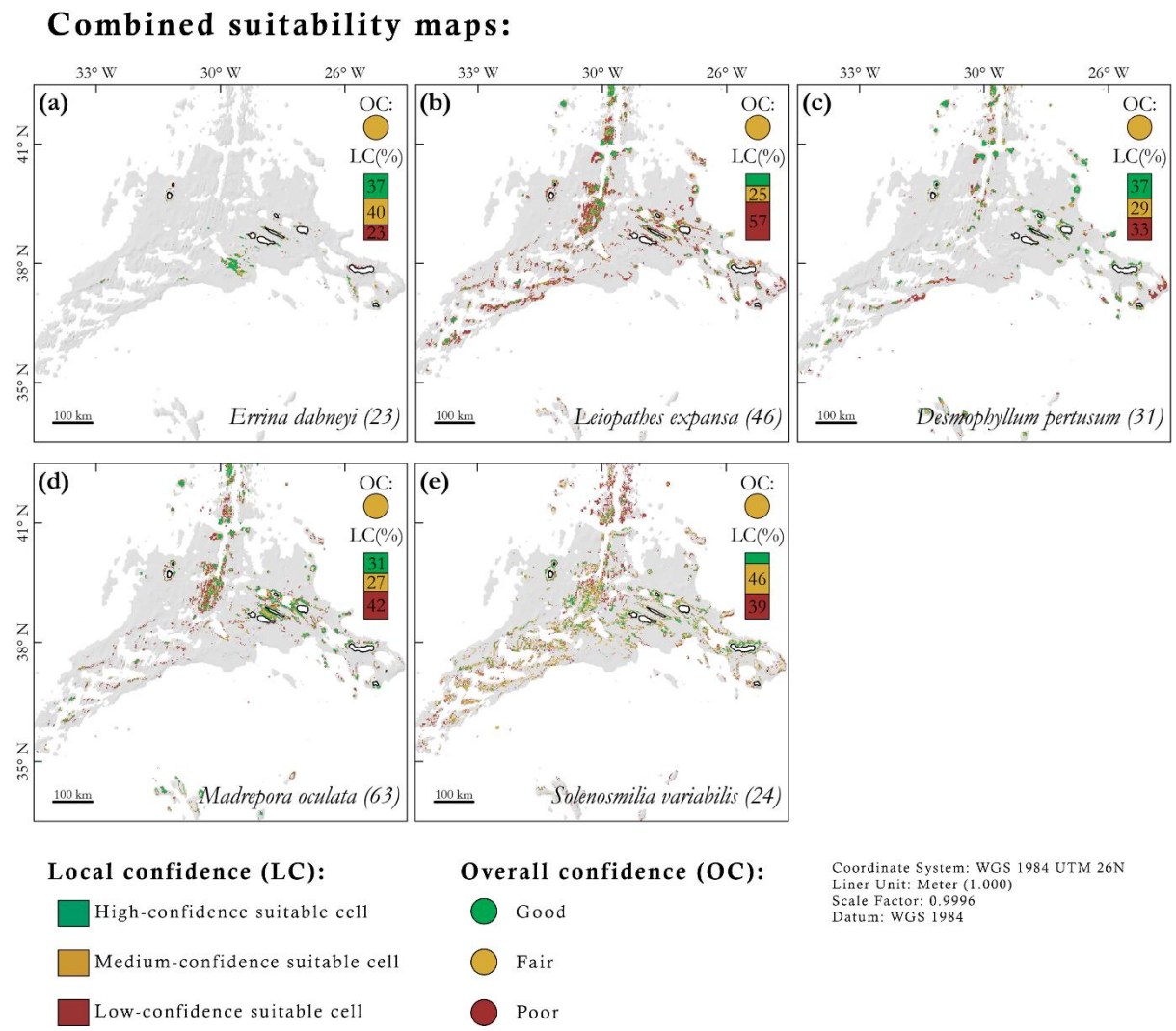

Figure 2.7 – Combined suitability maps for the selected Anthoathecata (Stylasteridae) (a), Antipatharia (b) and Scleractinia (c-e). Local confidence (LC): the percentage of cells classified as high, medium or low confidence. Overall confidence (OC): determined by the lowest score among AUC_{GAM}, TSS_{GAM}, AUC_{Maxent}, TSS_{Maxent} and F_{kHC}. Area under the curve (AUC); true skill statistics (TSS); improved fuzzy kappa for high-confidence suitable cells (F_{kHC}). In parenthesis the number of records used to fit the models. Species are ordered according to the mean depth of their estimated suitable habitat.

A second group of taxa (group 2) included: *N. bellissima*, *N. versluysi*, Coralliidae spp., and *P. johnsoni* (Figure 2.6f-i). These species presented very similar habitat distributions. The biggest differences were that the habitat of *N. bellissima* covered a larger area in the Central Seamounts and was almost absent north of Corvo Island while the habitat of *P. johnsoni* was absent from the Southern Seamounts and covered a more limited area than each of the other thee taxa. These species presented little overlap with the species of group 1. The most relevant zones of potential co-occurrence were: Gigante and the

associated ridges, Cavala, Voador and Sarda on the MAR; Condor and De Guerne in the Central Seamounts; Mar da Prata and Fouque in the Southeastern Seamounts; slopes west of São Miguel, Terceira and Corvo Islands and the tips of São Jorge Island.

A third group of taxa (group 1.5) included *V. flagellum*, *C. verticillata* and *Acanthogorgia* spp. (Figure 2.6c-e). This group appeared to lay in-between group 1 and group 2. In fact, the combined suitable habitat of *V. flagellum*, *C. verticillata* and *Acanthogorgia* spp. could be divided in two parts. The first part overlapped and included almost entirely the combined habitat of the species of group 1 (with the exception of the flat areas of Princess Alice). The remaining portion of suitable habitat overlapped almost entirely with the combined habitat of the taxa of group 2, even though the taxa of group 2 could inhabit areas unsuitable for species of group 1.5. Considering each species individually, the habitat distribution of *V. flagellum* was more similar to the habitat distribution of species of group 1, however *V. flagellum* appeared to be absent from Santa Maria. The habitat of *C. verticillata* showed a distribution similar to *V. flagellum*, but it was more common in the Central and in the Southern Seamounts and it was present in Santa Maria. The habitat distribution of *Acanthogorgia* spp. seemed to be confined to areas having steeper slopes (Figure 2.8d), to be more common on the MAR and to present a distribution more similar to the distribution of the taxa of group 2.

Finally, *D. pertusum*, *M. oculata* and *L.* cf. *expansa* were assigned to group 2.5. Their predicted habitat almost never overlapped with the combined habitat of group 1 and very rarely with species of group 1.5. The most evident locations where it happened were: Gigante and associated ridges, Cavalo W, SW1 and SW2 on the MAR; Mar da Fortuna in the Northeastern Seamounts; Mar da Prata, Fouque, Dom João de Castro Alcatraz in the Southeastern Seamounts; the island slopes around São Miguel, Terceira, Graciosa and Corvo. Most of the habitat overlap was with *Acanthogorgia* spp. Overlapping habitats were more common among species of group 2 and 2.5. The habitats of *D. pertusum*, *M. oculata* and L. cf. *expansa* were very scarce in the Central Seamount Group and on the slopes south of Pico and Faial while they appeared on most of the remaining geomorphic features highlighted in Figure 2.3a.

Predictions and responses of *Solenosmilia variabilis*, whose models were essentially fitted using bycatch data (Figure A.4c), seemed to be the least trustworthy, thus, were not associated to any particular group.

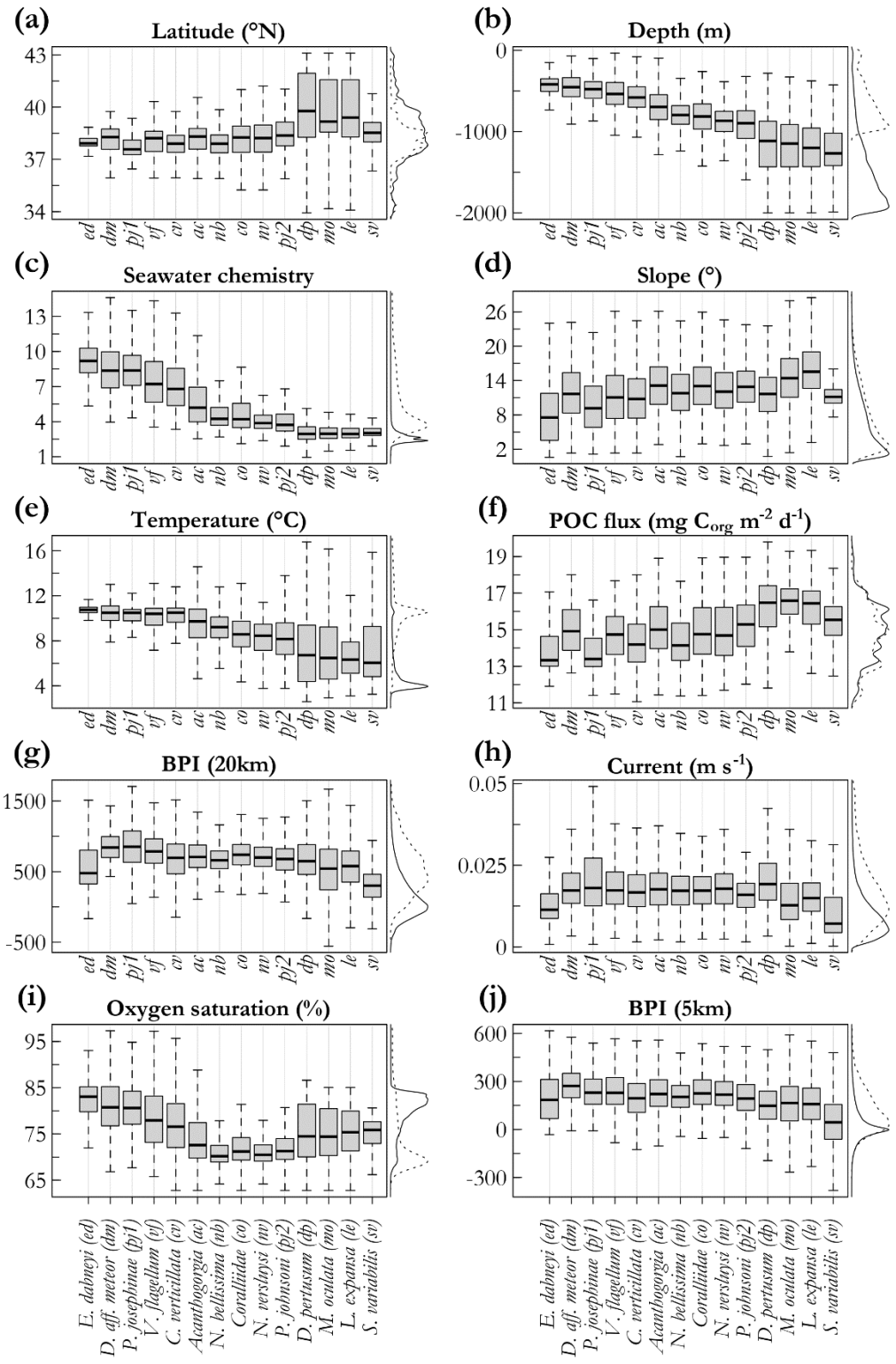

Figure 2.8 – Variable ranges within high-confidence cells of the combined suitability maps. (a) Latitude; (b) depth; (c-j) model predictors ordered by importance. Lines on the right side of the plots show the density distribution of each environmental predictor considering both suitable and unsuitable cells at seafloor locations shallower than 2000 m (solid line) and shallower than 1000 m (dotted line).

2.4.4 Bathymetric distribution

The estimated distributions of CWCs show a clear bathymetric zonation (Figure 2.8b). Environmental values at the locations of high-confidence suitable cells are shown in Figure 2.9. Regional values of nitrate, phosphate and silicate decrease down to about 800 m depth and then they become mostly stationary. At a similar depth there is an oxygen minimum zone in the study area. Temperature and saturation state of aragonite and calcite decrease with depth, but the decrease is gradual for the carbonate compounds and a bit inhomogeneous for the temperature (Figure 2.9). The bathymetric zonation of the modelled CWCs seemed correlated to these regional patterns. In particular, the taxa of group 2 (*N. bellissima*, *N. versluysi*, Coralliidae spp., and *P. johnsoni*) were tightly associated to areas presenting the minimum oxygen saturation and to nutrient concentrations matching the transition from decreasing to stationary background values (Figure 2.8I Figure 2.9c). Species of group 1 (*E. dabneyi*, *P. josephinae* and *D.* aff. *meteor*) only occurred above the oxygen minimum zone and were associated to a very narrow temperature range (11-12 ºC) (Figure 2.8e Figure 2.9a). Species of group 1.5 (*V. flagellum*, *C. verticillata* and *Acanthogorgia* spp.) mostly occurred above and within the oxygen minimum zone. They appeared to tolerate a wide range of nutrient concentration but their suitable cells almost never occupy areas corresponding to the steady state of nutrient concentration below the oxygen minimum (Figure 2.8c Figure 2.9b). They also seemed to tolerate different oxygen saturation levels and wider temperature ranges than the species of group 1 (Figure 2.8e Figure 2.9b). Species of group 2.5 (*D. pertusum*, *M. oculata* and *L.* cf. *expansa*) mostly occurred below and within the zone of minimum oxygen and were tightly associated to the stationary state of nutrient concentrations (Figure 2.8c Figure 2.9d).

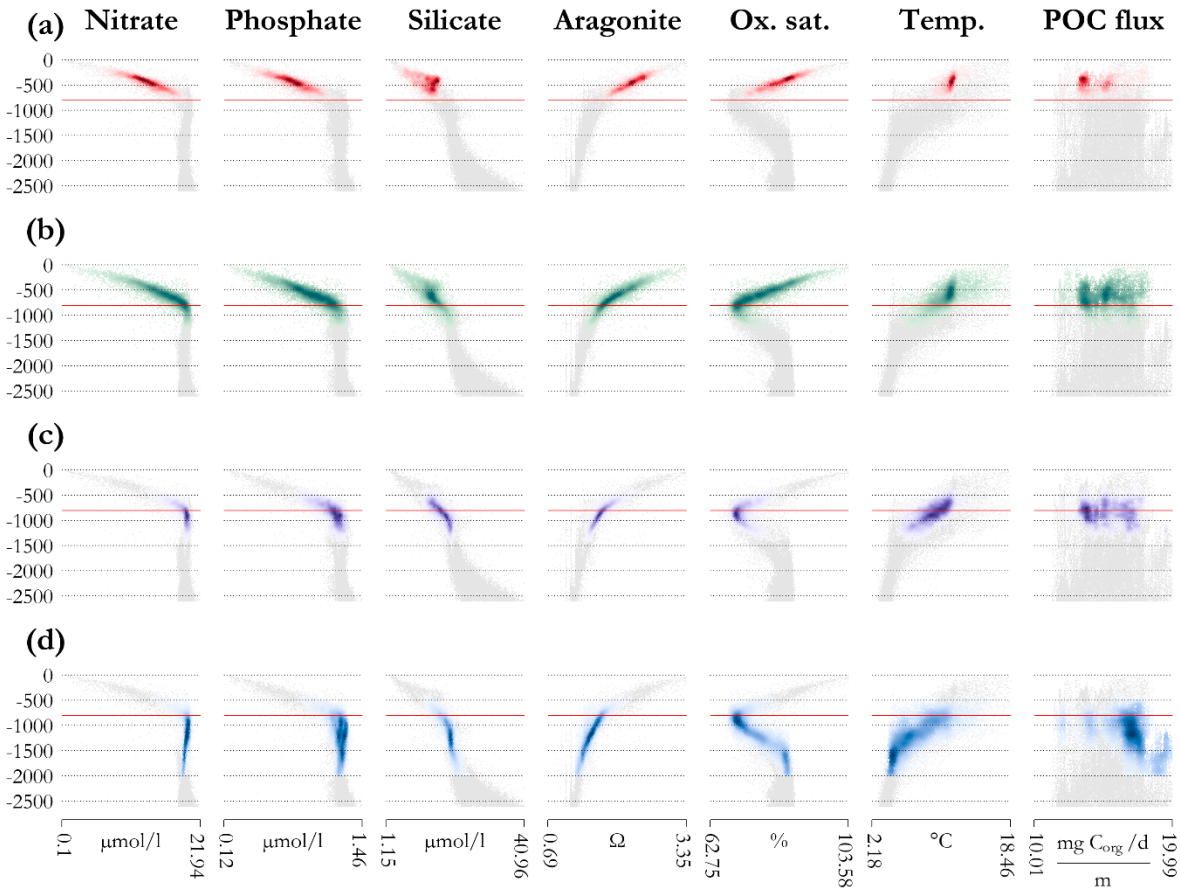

Figure 2.9 – Environmental values associated to high-confidence suitable cells plotted against depth. Plots of aragonite and calcite had a very similar profile, thus only aragonite is shown. The red line marks the minimum concentration of oxygen (~ 4.28 ml l−1 at 805 m). High-confidence suitable cells are combined into four groups. (a) Group 1 in red: *Errina dabneyi*, *Paracalyptrophora josephinae* and *Dentomuricea* aff. *meteor*. (b) Group 1.5 in green: *Viminella flagellum*, *Callogorgia verticillata* and *Acanthogorgia* spp. (c) Group 2 in purple: *Narella bellissima*, *Narella versluysi*, Coralliidae spp., and *Paragorgia johnsoni*. (d) Group 2.5 in blue: *Desmophyllum pertusum*, *Madrepora oculata* and *Leiopathes* cf. *expansa*. The darker the color, the denser are the suitable cells within a specific plot region. In grey are the regional background values. Seafloor areas below 2000 m depth were not modelled. Abbreviations: oxygen saturation (Ox. Sat.), temperature (Temp.), particulate organic carbon (POC).

2.4.5 Niche overlap

The pair-wise niche overlap and the results of niche equivalence and niche similarity tests are shown in Figure 2.10. Considering the taxa of group 1.5, the niches (*i*) *V. flagellum* and *C. verticillata* and (*ii*) *V. flagellum* and *Acanthogorgia* spp. were not more different from each other than expected by chance given the regional environmental conditions (i.e. they were equivalent niches). Considering the taxa of group 1, the niches of *E. dabneyi* and *D*. aff. *meteor* were equivalent. In addition, the niche of *P. josephinae* and the niche of *D*. aff. *meteor* were equivalent to the niche of *C. verticillata*. The niches of the taxa of

group 1 and 1.5 showed Schoener's D overlap values between 0.4 and 0.6 (with 1 being the maximum). *E. dabneyi* and *Acanthogorgia* spp. showed, on average, the lowest overlap values niche with the remaining species of groups 1 and 1.5.

The taxa of group 2 (*N. bellissima*, *N. versluysi*, Coralliidae spp., and *P. johnsoni*) were the ones showing the highest degree of niche overlap (Figure 2.10). They were significantly more similar among themselves than expected by chance, yet their niches were not equivalent. The taxa of group 2 showed a very low niche overlap with species of group 1 and low overlap with species of group 1.5. Between these groups, most of the similarity test were not significant (the niches of pairs of species were not more similar than expected by chance given the niche space they occupy and the regional environmental conditions).

D. pertusum and *M. oculata* occupied an equivalent niche space and showed a very low niche overlap with species of groups 1 and 1.5 and a low overlap with species of group 2. *L.* cf. *expansa* presented a higher degree of overlap with species of group 2 than the remaining species of group 2.5 (i.e. *D. pertusum* and *M. oculata*). However most of the similarity tests between *L.* cf. *expansa* and the taxa of group 2 were not significant when keeping the occurrence values of *L.* cf. *expansa* while shifting the position of a species of group 2 (Figure 2.10).

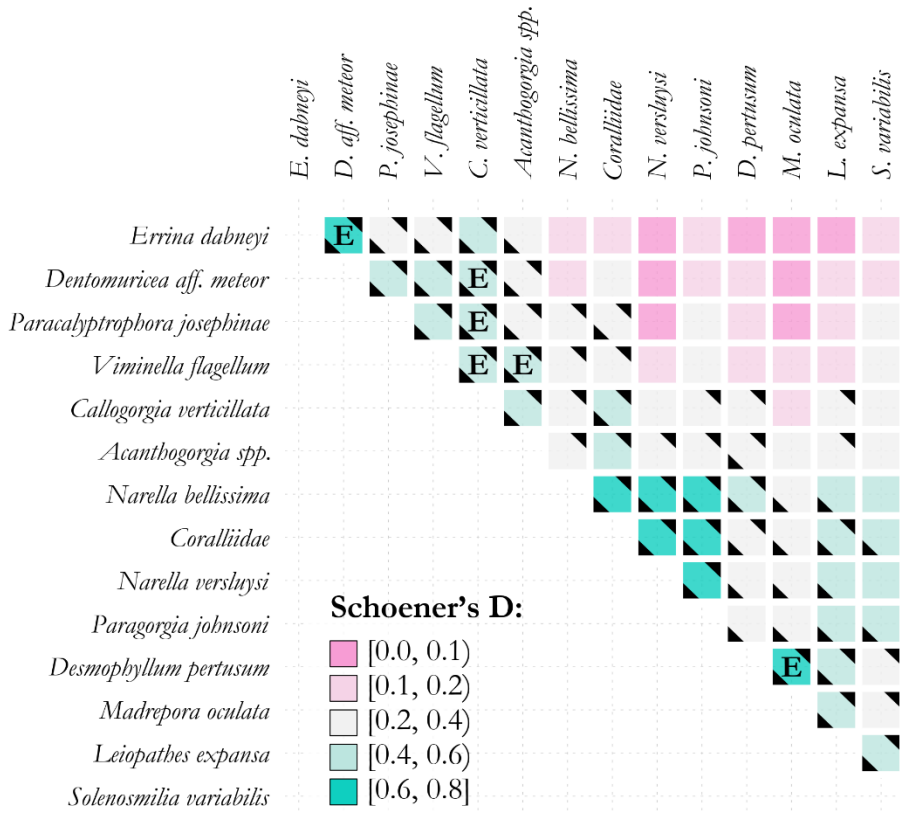

Figure 2.10 – Niche overlap computed based on the predictors 'seawater chemistry', particulate organic carbon flux, temperature, oxygen saturation and current speed. Only high-confidence suitable cells were considered for the computation of the niches. The letter 'E' identifies pair of taxa whose niches were not significantly different from each other (equivalent niches) ($\alpha > 0.05$). The black triangles identify pair of species whose niches were significantly ($P < 0.05$) more similar than expected by chance. Triangles in the bottom-left corners identify significant similarity tests of the row species when random shifting the spatial distribution of the column species. Triangles in the upper-right corners identify significant similarity tests of the column species when random shifting the spatial distribution of the row species.

2.5 Discussion

2.5.1 Performance, caveats and plausibility

Drawing general conclusion based on presence-only models is always a challenge as these models are affected by important limitations (Bowden et al., 2021). In our case, the lack of high resolution layers derived by an oceanographic model integrating tidal effects over topographic structures represented a major drawback. In fact, while communities at depths below the photosynthesis zone rely on food that sinks through the water column, the strength of this pelagic–benthic coupling depends on fluctuating

seabed currents (Jansen et al., 2018; Juva et al., 2020; Mohn et al., 2014). The layer of near-bottom current computed by VIKING20 (Böning et al., 2016) did not considered tidal effects and, hence, showed a limited utility as a predictor variable. The bathymetric grid we used is a mosaic of multibeam and non-multibeam data (EMODnet, 2018). The lack of fine-scale information in non-multibeam areas (e.g. Kurchatov Fracture Zone) may led to erroneous predictions. The coarse native resolution of the environmental layers (other than depth-derived layers) further limited the capacity of HSMs to discriminate between suitable and unsuitable areas. These limitations seemed more evident when assessing the latitudinal and longitudinal distribution of model predictions. On the contrary, HSMs seemed able to capture the major bathymetric patterns existing in the region. Community composition and environmental conditions show a greater change over depth than over latitude (McClain and Rex, 2015; Quattrini et al., 2017; Victorero et al., 2018). Likely, this enhanced variability was better captured by the variables available and, as a consequence, by our models.

According to traditional performance indicators (AUC and TSS), GAM and Maxent models had good or fair performance. However, discriminatory scores alone have a limited utility in evaluating presence-only models without the use of independent datasets, thus the ecological plausibility of model predictions and other measures of model uncertainty should be taken into account (Araújo et al., 2019; Warren et al., 2020). Uncertainties arising from alternative model structures and from biases in the response variables were addressed employing a bootstrap and a fuzzy matching approach. Models tended to better agree in areas where suitability predictions were less affected by the resampling of the input data ($F_{kHC} > F_k$). Until new information become available, these 'high-confidence' zones are the safest to use to make inferences about CWCs distributions. Part of the noise in model predictions appeared to be caused by bycatch records. In fact, areas where bycatch records were the principal source of information (e.g. Princess Alice-Açor) were associated to habitat suitability predictions presenting low or medium confidence. As new data become available, we suggest that future models should only use video records.

Octocoral and *Errina dabneyi* models seemed ecologically sound and in line with the distributional patterns that are emerging form regional surveys (e.g. Dominguez-Carrió et al., 2021; Morato et al., 2021a, 2018b). Model predictions of Scleractinia species and of *Leiopathes* cf. *expansa* were more problematic. The models seemed to capture the vertical zonation but failed to capture the geographic distribution of these species. In particular, the wide geographic extension of their estimated habitat appears unrealistic when compared to field observations. As discussed earlier, our models lack environmental predictors able to correctly link surface productivity, current regimes and food availability at the seafloor. We used POC

fluxes as a proxy, however the pelagic–benthic coupling is way more complex than the simple sinking of organic matter (Jansen et al., 2018). The absence of such predictors probably reduces the predictive capacity of these models. The supply of organic matter mediated by flow dynamics is a determinant factor for the existence of deep-sea stony corals (Fink et al., 2015; Hebbeln et al., 2020; Juva et al., 2020; Raddatz et al., 2020). Similar evidence is also emerging for black coral species of the genus *Leiopathes* (Lim et al., 2020). In general, flow dynamics and nutrient supply can be expected to be important for all CWCs, including Octocorallia and Stylasteridae species (Parrish and Oliver, 2020). Based on the previous considerations, it is likely that the proposed models picture reasonably well the bathymetric distribution of the modelled species. On the contrary, they allow only for cautious extrapolations about the latitudinal and longitudinal patterns existing in the region.

2.5.2 Geographic gradients

Our models suggest that almost all major geomorphic features and island host habitat suitable for the considered CWC taxa. Most of the suitable habitat occurs on the Mid-Atlantic Ridge (MAR), on the Central and on the Southeastern Seamounts and on island slopes (Figure 2.3a). However, the combined habitat of all the modelled species covers a tiny 11% of the study area. Setting this estimate in the context of the North East Central Atlantic, where large abyssal plains dominate the seascape (Figure 2.3b), the proportion of suitable habitats would shrink to a fraction of the total seabed. The fact that the Azores and the Mid-Atlantic Ridge are likely the only regions over a vast area able to sustain viable populations of the considered species suggests that these regions could be important in maintaining metapopulation dynamics (Addamo et al., 2021; Boavida et al., 2019) with implications for international conservation strategies (Badgley et al., 2017; Combes et al., 2021; Dunn et al., 2018). Indeed, Mironov et al. (2006) suggested that the Azores could be a centre of redistribution of marine fauna, i.e. an area where the faunas of different origins are mixing.

In general, CWCs appear to be absent from flat areas and from areas not associated to topographical features. An unsurprising result as topographical features have long been known to be important drivers of CWC distributions (Roberts et al., 2009). In our models, this response was driven by slope and benthic position index (BPI5 and BPI20) with corals avoiding slopes below 5º and near zero BPI values. At the scale considered in the present analysis, these predictors act as a proxy for more direct explanatory variables. Over 1 km^2 (the resolution of the model grid), measures of steepness and relative elevation act as a proxy of the likelihood of having rocky substrates and enhanced tidal and current regimes, essential factors for

the existence of CWC species (Baco et al., 2020; Juva et al., 2020; Mohn et al., 2014; Morato et al., 2021b). In addition, they provide an indication of seafloor complexity, another important driver of benthic species distributions (Henry et al., 2013; Levin et al., 2010). Direct information on surficial sediment types and habitat complexity would likely improve model predictions if finer scales were to be considered, however detailed seafloor characteristics averaged over coarse resolutions generally have low explanatory power (Anderson et al., 2016; Bennecke and Metaxas, 2017; Georgian et al., 2019). It is still important to consider that even if slope and BPI combined provide a valid indication of the availability of hard-substrate in a grid cell, preliminary video analyses in the region suggest that it is unlikely to have continuous rocky patches covering an entire area of 1 km^2. Thus, our models tend to overestimate the extension of suitable habitat.

Geographic gradients, other than those driven by the presence of topographical features and the avoidance of flat areas, appeared weak in the region. In line with previous studies (Braga-Henriques et al., 2013), the modelled taxa seem to exist on both sides of the MAR, the only exception being *Errina dabneyi* mostly limited to the central portion of the study area. The concentration of suitable areas east of the MAR and at latitudes delimited by the northernmost (Corvo) and the southernmost (Santa Maria) islands is probably owed to the abundance of shallow elevations at these latitudes. Wide latitudinal and longitudinal ranges of CWCs have been reported in previous studies covering areas of similar sizes (Auscavitch et al., 2020; Barbosa et al., 2020; Georgian et al., 2019) and represent a general characteristic of deep-sea species (McClain and Rex, 2015).

Different primary production regimes exist in the region with more productive waters located north and less productive waters located south of the archipelago (Amorim et al., 2017; Caldeira and Reis, 2017). The export of particulate organic carbon (POC) from the surface is one of the major factor shaping sub-surface benthic communities (Woolley et al., 2016), therefore, these different regimes could influence the faunal composition of different topographic features. The role of POC flux was not totally clear in our models. It was a shared important variable (i.e. important both for GAM and Maxent models) for *Desmophyllum pertusum*, *Madrepora oculata* and *Leiopathes* cf. *expansa* and it seems to drive their preference toward high productivity regions in the northern part of the MAR. However, habitat suitable for these species was found also in the southern part of the study area. A second possibility is that enhanced carbon fluxes allow these species to inhabit deeper waters in the northern part of the study area resulting in most of their habitat being located there. This pattern emerges in Figure 2.9d. The control of food availability over tolerance ranges and coral fitness have been shown in previous studies (Flögel et

al., 2014; Hanz et al., 2019; Hebbeln et al., 2020). For the remaining GAMs carbon fluxes had low importance, but they still were an important predictor for most of the Maxent models.

In general, all the considered taxa had patches of suitable habitat in the central portion of the study area. On the contrary, the Southern Seamounts and the slopes around Santa Maria, located in less productive waters, seem suitable only for few of the modelled species. This suggests that to the latitudinal gradient in productivity could correspond a gradual latitudinal turnover of species. A moderate rather than a sharp change in the composition of benthic assemblages over latitudinal gradients appears to be a common pattern in the deep sea (McClain and Rex, 2015). Caldeira and Reis (2017) found that the east group of islands and the Southern Seamounts are strongly influenced by the Azores Current and by its westward-propagating eddies, while the remaining portion of the study area is under the influence of the North Azores Current. Silva et al. (2013) found a north-south gradient in the composition of phytoplanktonic species. Future studies taking into account population densities rather than simple presence and new explanatory variables better framing the pelagic–benthic coupling and the hydrodynamic existing in the region will be needed to confirm the existence of a latitudinal turnover of deep-sea benthic species in the region.

2.5.3 Bathymetric gradient

Even acknowledging the limitations of HSMs, our results show some interesting patterns that are worth taking into account in future works. In particular, one of the least important variable explaining the predictions of our models, oxygen saturation, is one of the most important variable for the interpretation of their outputs. The modelled taxa seemed to form distinct groups. One was strictly limited to the upper part of the water column at seafloor locations above the regional minimum concentration of dissolved oxygen (group 1). Temperature seems to be the most important abiotic factor controlling their distribution. A second group of corals results strikingly associated to areas presenting the minimum concentrations of oxygen (group 2). A third group, overlapping with both group 1 and group 2, seems to tolerate most of the abiotic conditions present above and within areas associated to the minimum values of oxygen saturation (group 1.5). Finally, a third group of species (group 2.5) appears to be specular to the species of group 1.5 with suitable habitat present at the oxygen minimum and below (group 2.5). Rephrasing Carney (2005) and his study on the zonation of deep biota on continental margins, these groups seems to represent species truly restricted to upper waters, species extending down from upper waters, species extending up from deep waters and species truly restricted to intermediate waters. In

fact, as discussed in the next paragraphs, the turnover of species around the oxygen minimum can be associated to the presence of stratified water masses.

The reason why oxygen saturation has a poor explanatory value in the present study probably relates to the way its values change along the bathymetric gradient. Correlative models associate the presence of a species to particular values of a predictor variable. The fact that oxygen saturation and dissolved oxygen concentrations assume similar values above and below the oxygen minimum, at very different depths, is a confounding factor for correlative models. Thus, indirect predictors not affected by similar problems could emerge as better explanatory variables. In our case, 'seawater chemistry' was by far the most important predictor in all our models. The grouping of water nutrients with variables measuring seawater carbonate chemistry is unconventional. In fact, these variables have different physiological effects on CWCs. The saturation states of aragonite and calcite are critical parameters for the biomineralization of skeletal structures (Conci et al., 2021; Tambutté et al., 2011) and their importance as predictor variables was evident in studies modelling the distribution of coral species at global scales (Davies and Guinotte, 2011; Yesson et al., 2012). However, the carbonate saturation horizons in the study area occur at greater depths than those modelled (Figure A.6e). Therefore, while the saturation states of aragonite and calcite can still play a role in determining the suitability of particular locations and the vertical turnover of species, it is unlikely that they represent the only factors explaining the observed depth zonation. Similarly, water nutrients influence the physiology of CWCs (Goldberg, 2018), however as far as I know, there is no study identifying the concentration of water nutrient as a key explanatory variable of CWC distributions. Instead, nutrient and carbonate chemistry variables were grouped to act as a proxy variable reflecting the depth zonation of the regional water masses.

The variable 'seawater chemistry' seems to act as a proxy for the vertical stratification of the regional water masses. It also mimics oxygen concentrations above and below the regional oxygen minimum. When considering its individual components in relation to depth, it is possible to see that the concentration of nutrients increases from shallow waters down to a certain depth and then it becomes mostly stationary (Figure 2.9). The depth at which the nutrient concentrations change from declining to stationary corresponds to the depth of the regional minimum concentration of dissolved oxygen (ca. 4.3 ml l^{-1}) and to the depth where the core of intermediate waters occurs (ca. 800 m) (Figure A.1). Model predictions are structured around these patterns (Figure 2.9).

A water mass can be considered as an environmental envelop characterized by a particular combination of oceanographic parameters (Emery, 2001; Liu and Tanhua, 2021) that influences carbonate equilibria and food supply (Azetsu-Scott et al., 2010; Emery, 2001; Fontela et al., 2020; González-Dávila et al., 2011). The effects of water masses on the depth zonation of deep biota is supported by several studies and it is likely caused by multiple processes such as promotion or hampering of dispersal, environmental filtering, spatial segregation, speciation and physical phenomena altering the food supply (e.g. internal waves) (Carney, 2005; Puerta et al., 2020; Quattrini et al., 2017; Radice et al., 2016; Roberts et al., 2021; Victorero et al., 2018). These processes likely produce biodiversity and biogeographic patterns, even at local scales of several hundred meters (Carney, 2005; Puerta et al., 2020). In the Azores, Puerta et al. (2022) found that the regional stratification of the water column was useful to explain the vertical distribution of deep-sea megabenthic assemblages, while Menezes et al. (2006) found that water-mass boundaries and oxygen content influenced the zonation of demersal fish assemblages. In line with these studies, our models suggest that regional water masses play a critical role in the zonation of deep biota and, in particular, of habitat-forming CWCs. In particular, one group of species was strongly associated to intermediate waters (group 2) and a second group of species was strongly associated with upper waters (group 1). Species of groups 1.5 and 2.5 seemed to be limited to upper and intermediate or to deep and intermediate waters, respectively. This suggest that water masses can act as strong or as soft environmental filters depending on the species. In this perspective, intermediate waters and their boundaries represent interesting areas where to explore patterns of α and β diversity of CWC assemblages. In fact, intermediate waters seem to host a distinctive group of species but, at the same time, they seem to act as a transition zone of shallow and deep coral species pretty much like the transition zones described by Carney (2005). While recent analyses provide a preliminary support to this idea (Arantes et al., 2009; Auscavitch et al., 2020; Puerta et al., 2022), the investigation of this hypothesis represents an interesting topic for future studies.

2.5.4 Niche space

The projection of model predictions into the niche space served to reduce eventual geographic biases in habitat availability (Brown and Carnaval, 2019) and to better evaluate the observed distributional patterns. It was not intended to fully characterize the niches of the different species. In fact, as discussed earlier, the absence of important environmental variables and the reduced sampling effort, especially in deep areas, hampers this task. *Narella bellissima*, *Narella versluysi*, Coralliidae spp. and *Paragorgia johnsoni*, the group of taxa associated with intermediate waters (group 2), showed the highest degree of niche overlap (Figure 8). Considering that the only variables used to compute the niche space were

variables describing water properties, this result suggests a strong association of these species with intermediate waters. Species of group 1.5 (*Viminella flagellum*, *Callogorgia verticillata* and *Acanthogorgia* spp.) presented equivalent niches, i.e. statistically not different based on the environmental condition considered (Brown and Carnaval, 2019). This suggests that most of the latitudinal and longitudinal changes observed in their estimated distribution are owed to terrain variables (terrain variables were not used to compute the niche space) and that important predictor variables are likely missing. In line with the previous paragraphs, we argue that the most important variable missing is one that could serve as a proxy for food supply and fluctuating seabed currents. Equivalent niches were also found between species of group 1 and group 1.5. Therefore, with respect to the environmental variability considered, it seems that CWC taxa associated to upper water masses present similar tolerance ranges. Within the group of species inhabiting the deepest portion of the modelled area (group 2.5), *Desmophyllum pertusum* and *Madrepora oculata* showed equivalent niches, while *Leiopathes* cf. *expansa* presented a similar degree of niche overlap both with species of group 2.5 and with species of group 2. When exploring niche differences instead of niche similarities, species restricted to the upper part of the water column (group 1) showed very little overlap with species of intermediate and deep waters. The species characterizing intermediate waters (group 2) showed some overlap both with deep (group 2.5) and with shallow species of group 1.5. Overall, the results obtained in the niche space confirm the idea that the regional water masses act as strong or soft environmental filters depending on the species and that certain species characteristics of deep or of upper water masses appear able to coexist in intermediate waters.

2.5.5 Management

The modelled corals represent indicator taxa of vulnerable marine ecosystems (FAO, 2009) and have the characteristics of benthic foundation organisms (Crotty et al., 2019; Ellison, 2019). Thus, they are of great importance in the region both for ecological research and for spatial management. Acting as foundation species these taxa can be expected to mediate a crucial suite of biotic interactions essential to structure deep-sea benthic communities (Kikvidze and Callaway, 2009). Forming vulnerable marine ecosystems, they are, by definition, greatly affected by human induced disturbance (Clark and Tittensor, 2010; Ragnarsson et al., 2017). In the region CWCs face several threats. Climate change is expected to be particularly severe at the latitudes of the Azores (Levin et al., 2020; Morato et al., 2020c). The incredible amount of lost fishing gears that were observed in recent surveys suggest an extremely high fishing pressure (unpublished data from Dominguez-Carrió et al., 2021) even at remote locations (Morato et al.,

2021b). Finally, deep-sea mining will likely start in the region with important consequences for benthic organisms (Lopes et al., 2019; Ribeiro et al., 2020; Van Dover et al., 2017).

Spatial protection for these species is advised based on the threats they face and in the light of their small estimated habitat (ca. 11% of the modelled area). Because of the vertical zonation of the modelled CWCs, areas where they aggregate generally correspond to extremely steep sites that intersect the different water masses existing in the region (e.g. Gigante Seamount Complex). These are sites were environmental conditions present a rapid change over short distances allowing faunal assemblages of distinct depth zones to mix (McClain and Rex, 2015), therefore, they should represent priority targets for conservation. Our models suggest that the southeastern portion of the study area could harbor distinctive faunal assemblages. If spatial closures were to be implemented in the area following the representativity criteria (Stevens, 2002), networks of marine protected areas should cover all the water masses existing in the region, the interface of different water masses, as well as the two ecological regions that appear to be defined by the Azores Current (Caldeira and Reis, 2017).

Special concern deserves *Errina dabneyi*. This species seems to have an extremely limited range and seems to be endemic to the Azores (Braga-Henriques et al., 2011; Zibrowius and Cairns, 1992). Active conservation actions in its favor should be a priority (Costello and Chaudhary, 2017). However, all the considered species were observed in dense coral gardens and we believe that they should be viewed as important conservation target.

2.6 Conclusions

Our results suggest that in the Azores habitat-forming CWCs are organized in two ecological regions defined by the Azores Current. The depth zonation of CWCs is shaped by regional water masses. Over the considered bathymetric range (0 - 2000 m), CWCs form four groups of species: species restricted to upper water masses, species extending down from upper water masses, species restricted to intermediate waters and species extending up from deep water masses. As all the modelled CWCs represent foundation species, these patterns can be expected to remain valid for all associated species. The geographic and bathymetric patterns identified have important implications for the spatial management of these important cold-water coral species. The Azores is one of the few regions off the Atlantic margins that presents large topographic features intersecting water masses located in the upper portion of the water column. It represents an oasis of suitable habitat for the modelled CWC taxa and likely represent a key

area for maintaining interconnected populations of CWCs with important implications for international conservation initiatives.

Four take home messages emerged from a careful assessment of our models. These messages can help in the design and in the interpretation of future models targeting deep-sea species. One, models can have a different capacity to correctly detect distributional patterns along geographic and bathymetric gradients. Two, the lack of readily available variables acting as a meaningful proxy of the pelagic–benthic coupling hinders the detection of important geographic patterns. Three, unconventional variables mimicking known regional structures or processes can enhance model predictions (in our case 'seawater chemistry' as a proxy for water masses). Four, variables with low model importance can be of great importance in the interpretation of model results especially if they are correlated with other important variables (in our case 'oxygen saturation').

Chapter 3
Isopycnal layers as meaningful marine bioregions in the Azores (NE Atlantic)

3.1 Abstract

This study represents one of the first statistical bioregionalization of deep-sea benthic communities in the Azores (NE Atlantic). It focuses on habitat-forming filter feeder species (Phyla Cnidaria and Porifera) known to represent a critical component of hard-substrate communities in the region. Species archetype models were developed to simultaneously analyze and combine biological and environmental data into meaningful bioregions following a one-stage classification approach. The depth-wise succession of the regional water masses and their stratification into different isopycnal (vertical) layers appears to explain much of the changes in species composition occurring with depth. The boundaries of the different layers seem to act as transitional ecoclines rather than as sharp biogeographic breaks. Our results suggest that the different isopycnal layers can be used as meaningful bioregions, setting an interesting background for ecological research, ecosystem-based management and spatial monitoring. In the region, the main changes in species composition occur at 500, 800 and 1600 m generally confirming the biogeographic boundaries identified in previous studies. Temperature exerts a strong control on the distribution of assemblages restricted to the upper layer. Intermediate water masses appear to be characterized by a distinctive fauna. Food availability seems to exert a strong control on the distribution of assemblages located in the deepest part of the study area. The latter result suggests that the regional productivity gradient in surface waters affects deep-sea assemblages. Many of the described patterns are of general interest as they seem to be caused by environmentally driven processes that most likely affect the functioning and the structure of several components of deep-sea ecosystem in the Azores as well as in other regions.

3.2 Introduction

The distribution of species is a dynamic process influenced by several interacting drivers operating at different scales (D'Amen et al., 2017; Hortal et al., 2015; Myers and Giller, 1988). Understanding how these drivers operate and how biodiversity changes over space and time is a central theme in ecology and biogeography and it is essential for nature management. This task is quite challenging in the context of the deep sea, the most remote and the largest ecosystem on Earth (Folkersen et al., 2018; Glover et al., 2018; Morato et al., 2013; Smith et al., 2020). Species-environment relationships tend to be easier to measure at regional scales and are often the only proxy readily available to estimate the distribution of species. Consequently, habitat suitability models are frequently used to investigate how environmental drivers shape individual species' distributions. However, alternative methods that also combine biological and environmental data appear to be more useful when the focus is on assemblages and communities and, ultimately, on the bioregionalization (biogeographic classification) of a geographic space (Hill et al., 2020; Hui et al., 2013; Woolley et al., 2020). Bioregionalization is the process of identifying bioregions which are defined by Woolley et al. (2020) as "a model of the true distribution of multiple species that share a similar ecological and abiotic preference and sometimes an evolutionary history". A bioregion covers a particular geographic area mapping simultaneously biological and physical phenomena. It defines a meaningful spatial unit that can be used as a basis for research, ecosystem-based management and spatial monitoring (Howell et al., 2020; Rice et al., 2011; Whittaker et al., 2005).

Among the bioregionalization approaches, quantitative biologically-driven approaches are the most promising (Kreft and Jetz, 2010; Warton et al., 2015; Woolley et al., 2020). In particular, one-stage statistical methods that simultaneously model biological and environmental data and cluster groups of species into bioregions present notable advantages (i.e. formal definition of bioregions, result interpretability and characterisation of uncertainty) (Hill et al., 2020; Woolley et al., 2020). For instance, Species Archetype Models (SAMs) identify groups of species showing similar responses to the environment (species archetypes) in a one-stage approach (Dunstan et al., 2011). SEMs have been proposed as a valid tool to address bioregionalization problems (Hill et al., 2020; Hui et al., 2013) and have been used several times to investigate biodiversity patterns in marine ecosystems (Foster et al., 2015; Jansen et al., 2020; Leaper et al., 2014; Murillo et al., 2018; O'Hara et al., 2020; Woolley et al., 2013). Here, SAMs are applied to deep-sea benthic species inhabiting one of the most isolated archipelago in the North Atlantic, the Azores (Morato et al., 2020a; Santos et al., 1995).

The scale addressed by the present study is located in-between ecology and biogeography, in the sphere of ecological biogeography (Jenkins and Ricklefs, 2011; Myers and Giller, 1988). The main objective is to understand how and if the composition of regional deep-sea benthic assemblages changes along environmental gradients. The focus will be on large habitat-forming filter feeders associated with rocky substrates (Phyla Cnidaria and Porifera), an important component of the regional biota whose study has been one of the main targets of recent scientific campaigns (e.g. Morato et al., 2021a, 2020b, 2019, 2018b; Morato and Taranto, 2019; Orejas et al., 2017). Of particular interest would be to recognize persistent environmental structures able to explain the observed distribution of such organisms and, therefore, able to act as a meaningful proxy of marine bioregions. If a clear relationship between hard substrate assemblages and environmental structures were to exist, it could be assumed that a similar relationship would remain valid for other types of regional benthic assemblages. Although this assumption would be of speculative nature, it would still be useful to direct the collection of new data, formulate scientific questions and inform spatial management until a more comprehensive understanding of the forces shaping deep-sea communities is achieved.

Two major persistent environmental structures exist in the region. The first is a latitudinal gradient in productivity with more productive waters located north-west and less productive waters located south-east of the archipelago (Amorim et al., 2017; Caldeira and Reis, 2017). Considering that carbon export from the surface is one of the major factors shaping sub-surface benthic communities (e.g. Jansen et al., 2018; Juva et al., 2020; Woolley et al., 2016), it can be hypothesized that distinctive communities are associated with the two primary production regimes. The second persistent environmental structure is the vertical stratification of water masses and of all correlated environmental variables (Bashmachnikov et al., 2015; Liu and Tanhua, 2021). Water masses have long been recognized as one of the factors contributing to the vertical zonation of deep biota (Carney, 2005; Levin and Sibuet, 2012; Puerta et al., 2020) and have already been used to define biogeographic regions (Chimienti et al., 2019; Salazar et al., 2016; Watling et al., 2013). This far, few studies attempted a bioregionalization of deep-sea communities in the Azores region. In general, there is evidence that the depth-wise stratification of water masses influences the bathymetric zonation of the deep biota (Chapter 2, Braga-Henriques et al., 2013; Menezes et al., 2006; Puerta et al., 2022). The different production regimes appear to influence the distribution of phytoplankton assemblages (Silva et al., 2013), however they did not seem to have an effect on the composition of fish and coral assemblages at latitudes between 36 and 41º N (Braga-Henriques et al.,

2013; Menezes et al., 2006). Considering slightly wider scales (33-42º N), cold-water corals seem to respond differently to the production regimes existing north and south of the region (Chapter 2).

Besides environmental gradients, global biogeographic classifications suggest that a north-south biogeographical gradient could exist in the region. Following a traditional biogeographic method, in which current patterns of endemicity or cosmopolitanism of species are used to infer evolutionary processes (Myers and Giller, 1988; Sanmartín, 2012), Costello et al. (2017) identified several marine realms of endemicity using a two-dimensional approach (i.e. coastal, deep-sea, pelagic and benthic species were pooled together and depth was not considered in their analysis). Interestingly, the southernmost portion of the Azores region appears to be located at the boundary of different realms, the Offshore North and South Atlantic Realms. Following an alternative approach, Watling et al. (2013) classified the deep ocean floor using oceanographic proxies. They assumed that changes in key chemical, physical and environmental parameters, as well as water-mass connectivity, drive the turnover and distribution of marine species and communities. Again, the Azores region appears to be located at the boundary of different provinces but this time the boundary is located north of the archipelago. Even after a recent revision of the same classification based on the global distribution of anthozoans (Phylum Cnidaria) (Watling and Lapointe, 2022), the Azores still sit close to the boundary of two provinces: Boreal Northwest Atlantic and Northeast Atlantic. These classifications suggest that the Azores could represent a biogeographical transition zone and, thus, a place of high phylogenetic diversity where different evolutionary lineages coexist (Ferro and Morrone, 2014). Indeed, Mironov et al. (2006) suggested that the Azores could be a centre of redistribution of marine fauna, i.e. an area where the faunas of different origins are mixing. In general, the Azores are regarded as a hotspot of cold-water coral diversity (Clark and Tittensor, 2010; Sampaio et al., 2019a). Therefore, the composition of benthic assemblages located north and south of the archipelago could differ not only because of different environmental regimes, but also because of distinct evolutionary processes.

The present work aims to: (i) produce species archetype models that correlate benthic fauna occurrence data and environmental gradients; (ii) assess whether the predicted distribution of the different archetypes supports the use of environmental proxies (i.e. primary production or water mass provinces) as meaningful bioregions in the Azores. Based on the data currently available, the study area is limited to bathyal depths (200 2000 m).

3.3 Materials and Methods

3.3.1 Study area

Located on the Mid-Atlantic ridge, the Azores region presents a complex geological setting characterized by volcanic islands, seamounts and crests (Mitchell et al., 2018; Peran et al., 2016) (Figure 3.1). In general, these topographic features present large fractions of exposed bedrock and complex interactions with steady or variable flows, a setting that provides favorable conditions for thriving benthic communities dominated by filter feeders (Morato et al., 2021b; Rogers, 2018). As a result, the Azores host a great diversity of deep-sea corals and sponges that form important and vulnerable communities (Braga-Henriques et al., 2013; Carreiro-Silva et al., 2013; Morato et al., 2020c; Sampaio et al., 2019a; Somoza et al., 2020; Xavier et al., 2021).

Ocean circulation patterns in the Azores are complex (Bashmachnikov et al., 2004; Caldeira and Reis, 2017; Santos et al., 1995). Large-scale circulation is dominated by eastward-flowing branches of the Gulf Stream: The North Atlantic Current in the north and the Azores Current in the south. Eddies systems associated with these currents play an important role in the region (Caldeira and Reis, 2017; Sala et al., 2016). These flows interact with the regional topography and with local winds altering surface temperatures and primary production (Amorim et al., 2017; Bashmachnikov et al., 2004; Caldeira and Reis, 2017). Warmer and less productive waters are located south-east, colder and more productive waters are located north-west of the archipelago. These differences suggest that the Azores is a transitional region between subtropical and temperate waters.

Based on the work of Liu and Tanhua (2021), it is possible to identify the principal water masses around the Azores and approximate their vertical extent (Figure 3.2). In the Atlantic the different water masses are distributed in four main isopycnal (vertical) layers: upper, (ii) intermediate, (iii) deep and overflow and (iv) bottom layer. In the Azores, the influence of upper water masses goes down to about 1100 m (Figure 3.2a-c). The Eastern North Atlantic Central Water (ENACW) has the widest vertical extent with its core at about 500 m depth. The Western North Atlantic Central Water (WNACW) and the Western South Atlantic Central Water (WSACW) are present within the upper 500 m of the water column. The intermediate layer is characterized by the Mediterranean Water (MW) and the Antarctic Intermediate Water (AAIW) (Figure 3.2d-e). Their core is at about 800 m depth. A third intermediate water mass, the Subarctic Intermediate Water (SAIW) has been reported in the northwestern portion of the study area (Bashmachnikov et al., 2015), however, given the lack of sampling stations in this area (Figure 3.2f) it was almost absent in the

data provided by Liu and Tanhua (2021). In the deep and overflow layer, down to 2000 m, the Upper North Atlantic Deep Water (uNADW) is the predominant water mass with its core located at about 1600 m depth (Figure 3.2g). It occupies a wide depth range in the north (500-2000+ m), while in the south it was observed only at depths below 1000 m. The Lower North Atlantic Deep Water (lNADW) and the Northeast Atlantic Bottom Water (NEABW) characterize the deepest portion of the considered bathyal range (200-2000 m) and extend well below 2000 m (Figure 3.2h-i).

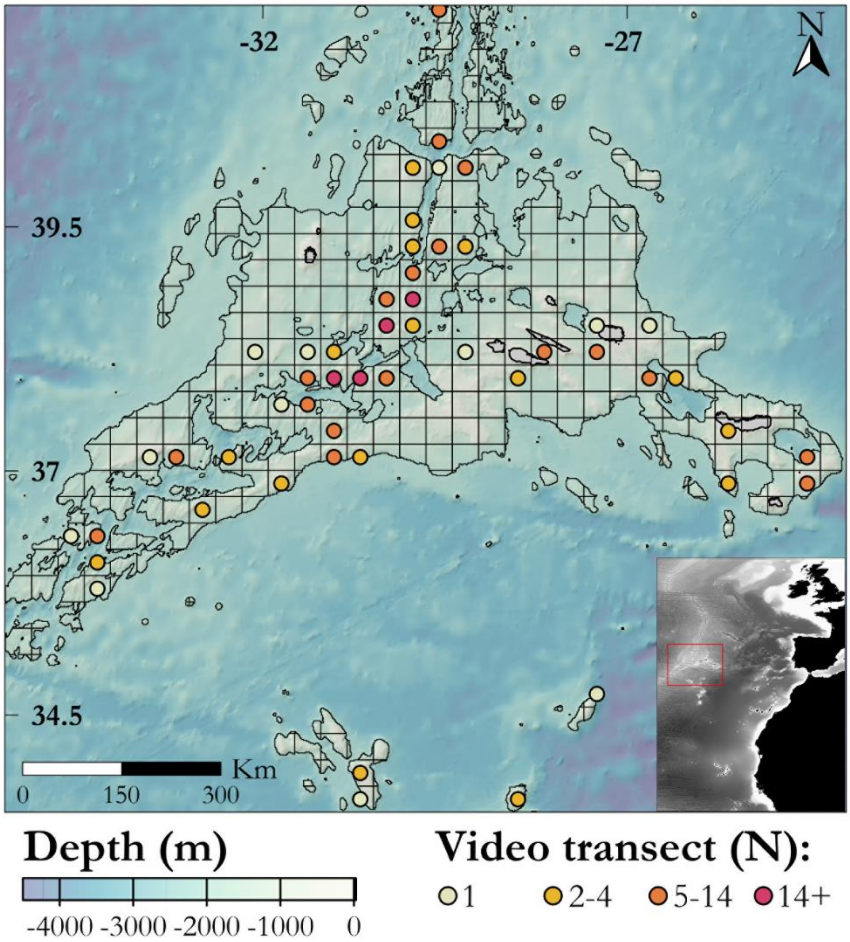

Figure 3.1 – Azores study area (ca. 33.5 to 42º N, 24 to 35º W). Deep-sea benthic video stations (video transects) considered in the present study aggregated on 30x30 km grid for better visualization (note that the model grid is smaller and has a resolution of about 1 km^2). Only stations located at depths shallower than 2000 m are considered. The contour lines identify all the areas shallower than 2000 m.

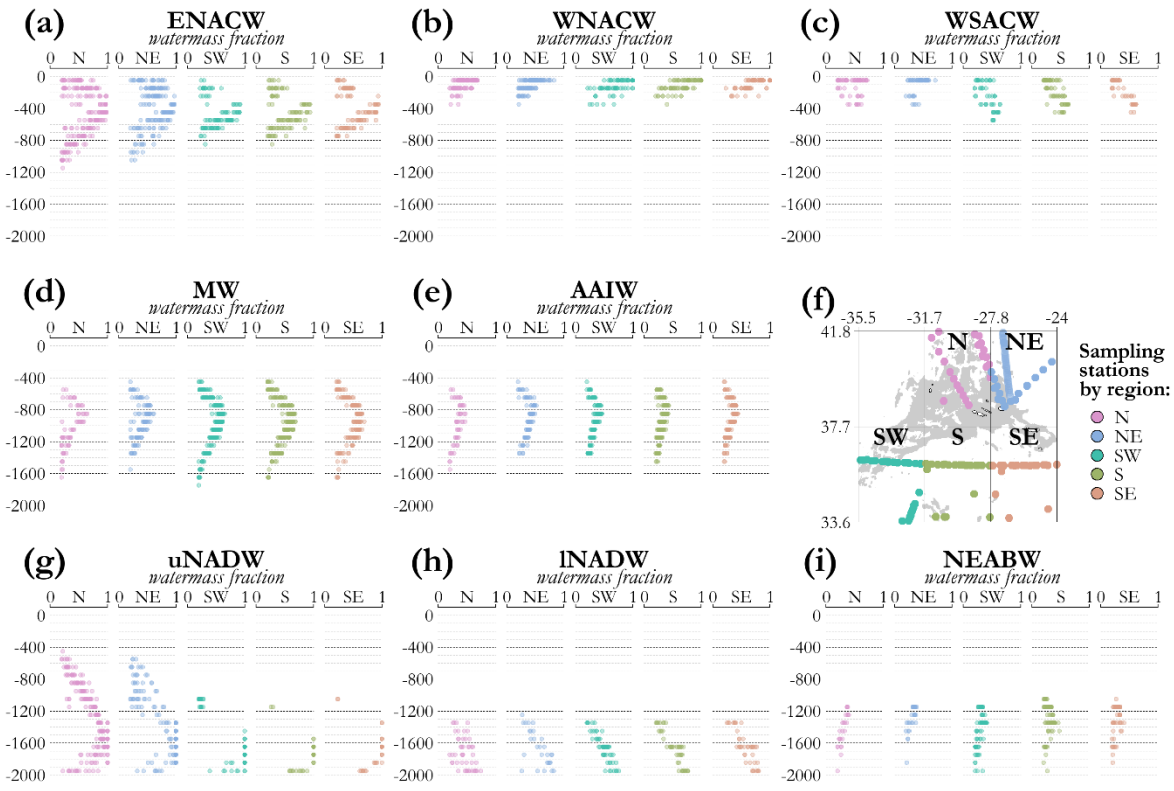

Figure 3.2 – Principal water masses of the Azores based on Liu and Tanhua (2021). Upper layer: (a) Eastern North Atlantic Central Water (ENACW), (b) Western North Atlantic Central Water (WNACW), (c) Western South Atlantic Central Water (WSACW). Intermediate layer: (d) Mediterranean Water (MW), (e) Antarctic Intermediate Water (AAIW). Deep and overflow layer: (g) Upper North Atlantic Deep Water (uNADW), (h) Lower North Atlantic Deep Water (lNADW), (i) Northeast Atlantic Bottom Water (NEABW). Every sampling station is divided in 100 m bins over which water mass fractions are averaged. Only fractions greater than 0.2 are showed on the plots. (f) Sampling stations available in the study area; the grey background identifies seafloor depths shallower than 2000 m.

3.3.2 Biological data

Biological data were derived from the preliminary annotation of videos acquired from deep-sea cameras deployed using different platforms and systems (Appendix B, Table B.1 and Figure 3.1). Most of the recent imaging surveys in the region were directed toward areas of complex morphology expected to present a high percentage of hard substrates (Dominguez-Carrió et al., 2021; Morato et al., 2021a, 2020b, 2019, 2018b; Morato and Taranto, 2019; Orejas et al., 2017). Therefore, this study focuses on habitat-forming species of deep-sea corals and sponges (Phyla Cnidaria and Porifera) representative of regional hard-substrate communities.

In order to reduce biases arising by different image annotators and sampling gears (de Mendonça and Metaxas, 2021; Durden et al., 2016; Sheehan et al., 2016), additional criteria were used to define the list of considered species. Small organisms were excluded because they tend to be hard to identify (Durden et al., 2016), a condition worsened by the use of different camera systems which can detect a small specimen in some cases (e.g. a remotely operated vehicle slowing down and getting closer to a coral patch) but not in some others (e.g. a drift-cam system on a windy day drifting at high speed). The main exception was *Leptopsammia Formosa*, characterized by a bright yellow color easy to identify in most conditions. In order to reduce as much as possible the problem of misidentification (Beisiegel et al., 2017; Howell et al., 2014), preference was given to species with physical specimens deposited in the local reference collection (Sampaio et al., 2019b). In general, there is good confidence about the taxonomic identity of the selected cold-water corals. The taxonomic identification of sponge species is still ongoing and the names used here should be considered only as putative names, at least for the non-lithistid sponges (Xavier et al., 2021). Only species with a peculiar and familiar morphotype, relatively easy to identify in the video annotation process, were considered in the analysis. This led to the exclusion of rare species and of set of species known to present very similar morphologies (e.g. corals of the genus *Acanthogorgia* and of the family Coralliidae).

Biological data were aggregated on a model grid with a resolution of 30-arc second (∼1 km^2, projection UTM 26N). A total of 1612 unique observations, 26 species (Table 3.1) and 281 locations of the model grid at depths between 200 and 2000 m (Table B.2) were used for the species archetype models (SEMs). Only species that occurred at 10 or more locations were retained. At each of these 281 locations, a species was marked as present (when observed in one or more video transects) or absent (when not observed in any video transect).

Table 3.1 – List of species considered by the species archetype models. Number of records (N) after aggregation on the model grid. Species archetype (SA) group, all membership probabilities (τ) were equal to 1.

ID	Phylum	Class	Subclass	Order	Species	N	SA
1	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Acanella arbuscula*	75	G.4
2	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Callogorgia verticillata*	65	G.2
3	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Dentomuricea* aff. *meteor*	40	G.1
4	Cnidaria	Anthozoa	Hexacorallia	Scleractinia	*Desmophyllum pertusum*	30	G.4
5	Cnidaria	Anthozoa	Hexacorallia	Antipatharia	*Elatopathes abietina*	44	G.1
6	Cnidaria	Anthozoa	Hexacorallia	Antipatharia	*Leiopathes* cf. *expansa*	52	G.4
7	Cnidaria	Anthozoa	Hexacorallia	Antipatharia	*Leiopathes glaberrima*	16	G.1
8	Cnidaria	Anthozoa	Hexacorallia	Scleractinia	*Leptopsammia formosa*	82	G.3
9	Cnidaria	Anthozoa	Hexacorallia	Scleractinia	*Madrepora oculata*	58	G.4
10	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Narella bellissima*	42	G.3
11	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Narella versluysi*	69	G.3
12	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Paracalyptrophora josephinae*	18	G.1
13	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Paragorgia johnsoni*	56	G.3
14	Cnidaria	Anthozoa	Octocorallia	Alcyonacea	*Viminella flagellum*	139	G.2
15	Porifera	Demospongiae	Heteroscleromorpha	Tetractinellida	cf. *Characella pachastrelloides*	114	G.2
16	Porifera	Demospongiae	Heteroscleromorpha	Tetractinellida	cf. *Craniella longipilis*	61	G.2
17	Porifera	Demospongiae	Heteroscleromorpha	Haplosclerida	cf. *Haliclona magna*	61	G.3
18	Porifera	Demospongiae	Heteroscleromorpha	Tetractinellida	*Leiodermatium* cf. *pfeifferae*	38	G.1
19	Porifera	Demospongiae	Heteroscleromorpha	Tetractinellida	*Macandrewia* cf. *azorica*	80	G.2
20	Porifera	Demospongiae	Heteroscleromorpha	Tetractinellida	*Neophrissospongia nolitangere*	34	G.1
21	Porifera	Demospongiae	Heteroscleromorpha	Haplosclerida	cf. *Petrosia crassa*	118	G.2
22	Porifera	Demospongiae	Heteroscleromorpha	Bubarida	cf. *Phakellia robusta*	39	G.4
23	Porifera	Demospongiae	Heteroscleromorpha	Bubarida	cf. *Phakellia ventilabrum*	21	G.4
24	Porifera	Hexactinellida	Amphidiscophora	Amphidiscosida	*Pheronema carpenteri*	72	G.4
25	Porifera	Demospongiae	Heteroscleromorpha	Tetractinellida	cf. *Poecillastra compressa*	120	G.3
26	Porifera	Hexactinellida	Hexasterophora	Lyssacinosida	cf. *Regadrella phoenix*	68	G.4

3.3.3 Environmental data

A total a 16 variables were considered to select and develop the SAMs (Figure B.1). Depth values were obtained from an updated bathymetry grid for the Azores that integrates the Emodnet Digital Batimetry (EMODnet Bathymetry Consortium, 2018) with all the multibeam echosounder data available in the region (Luis Rodrigues, unpublished data). The bathymetry grid has a resolution of 30-arc second (~1 km^2). In

general, a large set of depth-related variables are considered to develop deep-sea species models (e.g. Anderson et al., 2016; Georgian et al., 2019; Lecours et al., 2016). In this case only three variables of this kind were considered. The rationale was to avoid generating overcomplicated models. Slope, coarse and fine scale bathymetric position index (BPI) were the terrain variables that contributed the most to models developed for regional cold-water corals (Chapter 2) and, therefore, they were the terrain variables considered here. They were computed using the Benthic Terrain Modeller toolbox (Walbridge et al., 2018). For each cell in a surface, BPI is evaluated by finding the difference between the elevation value of the cell and the mean elevation of all cells in an annulus (Walbridge et al., 2018). It provides an indication of the relative position of a cell (e.g. crest, flat, valley) (Lecours et al., 2016) and it is a variable commonly used to model deep-sea species (e.g. Anderson et al., 2016; Georgian et al., 2019). The fine scale BPI was computed with a radius of 5 km (BPI5), the coarse scale with a radius of 20 km (BPI20). In both cases the inner radius was set to 1 km. Slope is generally used as proxy for the exposure of hard-substrate and for the local acceleration of currents (Lecours et al., 2016; Niedzielski et al., 2013).

The other candidate variables included seafloor values of temperature (ºC), salinity (PSU), dissolved oxygen (mL L^{-1}), oxygen saturation (%), oxygen utilization (mL L^{-1}) and nutrient concentrations (µmol/l) (phosphate, silicate and nitrate) computed by Amorim et al. (2017); near seafloor aragonite (Ω_{ar}) and calcite (Ω_{ca}) saturation computed by Morato et al. (2020c); particulate organic carbon (POC; g m^{-2} d^{-1}) flux estimated using the net primary productivity (NPP) computed by Amorim et al. (2017) and a productivity export model (Lutz et al., 2007). The seafloor variables computed by Amorim et al. (2017) were extracted from the World Ocean Atlas 2013 (Levitus et al., 2015) and seafloor conditions were interpolated using climatological means to create 30-arc second (~1 km^2) grids. The variables computed by Morato et al. (2020c) had a native resolution of 0.5° and were rescaled to a 3 x 3 km grid cell using universal kriging and depth as a covariate. The NPP layer had an original resolution of 9 km.

All predictor variables were projected and rescaled on a 30-arc second grid (~1 km^2, UTM zone 26N) using bilinear interpolation. This resolution was selected to provide a balance between biological and environmental data. The use of coarser resolutions would imply extracting species absences and presences on areas much wider that those actually sampled with detrimental effects on model performance. The use of a finer grid would imply an excessive rescaling of the environmental variables.

Highly correlated variables (Pearson's $r \geq 0.7$) were removed in order to produce interpretable models and reduce the effect of collinearity (Dormann et al., 2013). Most of the environmental variables were

highly correlated with depth (Figure B.1). Since depth and depth-related variables had the highest resolution, their choice was favored over other variables. BPI20 and BPI5 were also correlated. BPI20 was selected because, based on a previous study (Chapter 2), it seemed more appropriate to model the regional fauna at the scale considered here. Since all nutrient variables were highly correlated, it was decided to combine them using the principal component analysis (PCA) to reduce the high collinearity (Dormann et al., 2013). The first principal component (PC1 Nu), explaining 96%, of the variance was used as a predictor variable. Nutrient concentrations and PC1 Nu presented a direct correlation (Figure B.2). The same process was repeated for the highly correlated variables aragonite and calcite saturation. However, even after performing a PCA, the resulting first principal component was highly correlated with depth and with PC1 Nu. Therefore, it was discarded. The final set of variables used to develop the SAMs included six variables: depth, slope, BPI20, POC flux, bottom oxygen saturation and PC1 Nu (Figure B.3).

3.3.4 Analysis

All analyses were performed in the R environment (version 4.2.1) (R Core Team, 2022) using the R packages 'EcoCluster' (version 1.0.0) (Valle et al., 2022) and 'SpeciesMix' (version 0.3.4) (Dunstan et al., 2016). At the time of writing, 'SpeciesMix' was archived because of incompatibilities with the most recent updates of R. After installing the archived version, one internal function presented an error ('SpeciesMix.bernoulli'). The bit of code producing the error was fixed and the package was used without any further problem [old code: if(class(fmM.out$covar) != "try-error"); new code: if(class(fmM.out$covar)[1] != "try-error")].

Species archetype models (SAMs) were used to identify groups of species that respond in a similar fashion to the environment (Dunstan et al., 2013, 2011; Hui et al., 2013). This approach can cluster and predict species occurrence from presence/absence data using finite mixture models. Finite mixture models posit that underlying the total set of observations there is a mixture of subgroups each described by a probability distribution. Based on this assumption, the observed data can be explained combining a finite mixture of regression models (McLachlan and Peel, 2000). In the SAM framework, each subgroup is represented as a logistic generalized linear model (GLM). Species grouped together are described by the same GLM, share a statistically indistinguishable environmental response and, in this sense, represent an 'archetype' (Dunstan et al., 2011; Leaper et al., 2014). Species are assigned to one or more archetype groups in a soft-manner (probabilistic) during model fitting (Dunstan et al., 2011; Hui et al., 2013). Once

fitted, SAMs can be used to make predictions on the occurrence probability of the different archetypes across the study area.

Ultimately, SAMs represent a clustering method. In general, the 'true' number of latent groups in a particular dataset is unknown, therefore, a major challenge of clustering is to identify the optimal number of groups (referred to as archetypes in the case of SAMs) (Dunstan et al., 2011; McLachlan and Peel, 2000; Valle et al., 2022). The standard approach in SAMs and in other clustering methods is to systematically vary the number of groups and to determine the optimal number of groups based on a performance metric like the Bayesian information criterion (BIC) (Dunstan et al., 2013, 2011; Schwarz, 1978). An additional metric, suggested by Dunstan et al. (2011) in order to prevent fitting too many species archetypes, is the minimum mean membership probability [$\min(\pi_i)$]. This probability should never be smaller than the a priori probability of there being at least one species in a group. An alternative approach, directly reducing the chances of overestimating the optimal number of archetypes, has been suggested by Valle et al. (2022). These authors proposed the use of a particular type of sparsity-inducing prior (Ishwaran and James, 2001; Manrique-Vallier, 2016; Sethuraman, 1994), the truncated stick-breaking (TSB) prior. An intuitive way to interpret this approach is to think about it as a sequential process in which every time an additional cluster k is considered, the probability that a sampling unit is assigned to cluster k (θ) depends on the probability that it is not assigned to any other cluster (1, 2, …, k – 1). For large values of k, the probability that a sampling unit is assigned to cluster k drops to values close to zero (Manrique-Vallier, 2016; Valle et al., 2022). Clusters accounting for a small portion of the observations can be ignored. The use of the TSB prior favors solutions with fewer clusters and has the advantage of determining the optimal number of clusters without fitting multiple models. The only requirement is to define a maximum number of groups to truncate the stick-breaking prior. The general advice is to set the maximum number of groups so that most of the posterior membership probability is concentrated on the initial components (i.e. a large number of cluster solutions should be associated with low membership probabilities) (Manrique-Vallier, 2016; Valle et al., 2022). Both the BIC and the TSB prior approach are considered in the choice of the optimal number of species archetypes.

Following the methods used in previous studies (Dunstan et al., 2011; Leaper et al., 2014; Murillo et al., 2018; O'Hara et al., 2020), the choice of the optimal number of archetypes is performed fitting a model with all covariates and quadratic terms (1st and 2nd degree orthogonal polynomials). SAMs were fitted assuming a Bernoulli distribution with a logit link function (Dunstan et al., 2011). The TBS approach is implemented in a single step using the function 'gibbs.SAM' (R package 'EcoCluster'). The BIC approach is

implemented using the function 'SpeciesMix' (R package 'SpeciesMix'). In the latter case, the number of archetypes (G) is altered sequentially. For each number G, models are fitted multiple times with different initializations (50) and the best model (lowest BIC value) is selected. This is a standard computational approach for mixture models used to reduce the chances that the estimation of model parameters during the log-likelihood maximization results from local rather than global maxima (Dunstan et al., 2013, 2011; McLachlan and Peel, 2000). The BIC values of SAMs computed with different values of G are compared based on $\Delta BIC(G) = BIC(G) - BIC(G + 1)$ versus G. The first increase in $\Delta BIC(G)$, which is represented by the first negative value, indicates the optimal number of archetypes. Models are also evaluated considering that $\min(\pi_i)$ should be greater than 1/S (S is the number of species).

After selecting the optimal number of archetypes, the best set of covariates is determined following a stepwise approach to minimize the value of BIC while holding the number of groups constant (e.g. Dunstan et al., 2011; Leaper et al., 2014; Murillo et al., 2018; O'Hara et al., 2020). Again, for each set of covariates, models are fitted multiple times with different initializations (50). Models fitted with different sets of covariates are compared and the model with the lowest BIC value is selected. Parameter relative standard error (RSE) – the standard error (SE) of the absolute value of the parameter multiplied by 100 and divided by the estimated coefficient – is used to evaluate the strength of the relationships between each archetype and covariate. Using the SAM with the optimal number of groups and the optimal set of covariates, prediction across the study area are achieved using the 'predict.archetype' function. Finally, partial response curves are computed for all covariates (Dunstan et al., 2013).

3.3.5 Dummy absences

After running a complete analysis with the available data, it appeared evident that the complete lack of sampling over large flat areas biased model predictions. This was particularly true for archetypes occurring at depths below of 1500 m, where most of the plains without any biological sampling occur (Figure 3.1). All the selected species are sessile invertebrates mostly living on hard substrates. Thus, they are unlikely to colonize flat areas away from major relieves where the chances of having exposed bedrock are limited. In an attempt to improve model predictions, five cells of the model grid located within large plains with flat values (slope < 1º; -100 < BPI20 < 100) were considered as dummy sampling stations where all the considered taxa marked as absent. Using these dummy absences together with the available data SAMs predictions appeared to be more realistic. At the same time, the use of dummy absences did not change the optimal number of archetype groups, species memberships or the general shape of the response to

model covariates. The only notable change was that without the use of dummy absences the best model included the variable PC1 Nu, using dummy absences it included bottom oxygen saturation. These variables present a similar rate of change over depth and they do not display latitudinal trends (Figure B.3). Based on these considerations, it was decided to use the dummy absences. They are listed and plotted Appendix B.2 (Table B.3 and Figure B.4).

3.4 Results

3.4.1 Model selection

A maximum of 16 archetype groups should be expected if the combination of the eight water masses (Figure 3.2) and of the two production regimes existing in the region were to produce distinct bioregions. Therefore, BIC values were computed fitting the species archetype models (SAMs) varying the number of species archetypes (G) from 1 to 16 and TBS values were computed setting the expected maximum number of groups to 16. The two traditional evaluation metrics of SAMs suggest that the optimal value of G is five. In fact, while the first increase in ΔBIC(G) occurred between 6 and 7 archetypes (Figure 3.3a), the value of min(π_i) was above the a priori archetype membership probability (1/S) only selecting five groups (Figure 3.3b). Instead, the TBS approach suggests the existence of four archetype groups that, on average, account for 98.4% of all species (Figure 3.3c). Increasing G from four to five, three archetype groups remained practically unchanged. The species of the fourth group were divided into two distinct archetypes that showed similar environmental responses and a great overlap in their predicted distributions suggesting an overfitting of the data. Based on these results, four archetypes were considered to be the best solution for the available environmental and biological data. Performing the stepwise variable selection with G = 4, the model with the lowest BIC (6135) included five covariates (both linear and quadratic terms): depth, bathymetric position index (BPI20), particulate organic carbon (POC) flux, bottom oxygen saturation and slope (Table B.4). The 26 taxa in the SAMs were assigned to the four archetypes as follows: archetype 1 (G.1) 23%, archetype 2 (G.2) 23%, archetype 3 (G.3) 23% and archetype 4 (G.4) 31%. All species had a membership probability (τ) equal to one suggesting that the archetype groups effectively capture the main environmental responses of the considered taxa (Foster et al., 2015). Species archetype compositions can be seen in Table 3.1.

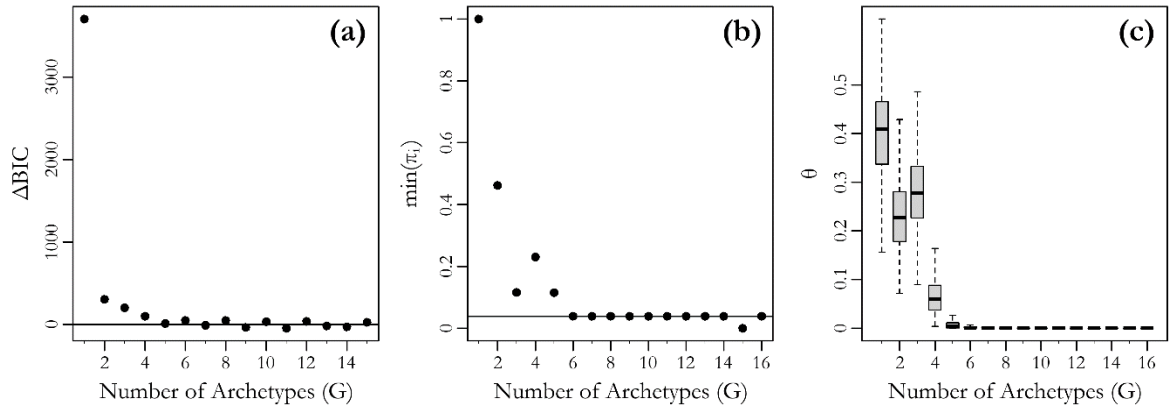

Figure 3.3 – Evaluation metrics of species archetype models (SAMs) with different numbers of archetype groups (G). Plots are (a) ΔBIC(G) vs. G, (b) min(π_i) vs. G and (c) θ vs. G. The solid horizontal line in panel (a) is the ΔBIC(G) = 0, and in panel (b) is the min(π) = 1/S = 1/26. BIC is the Bayesian information criterion, S is the number of species, min(π_i) is the archetype with the minimum value of mean membership probability, θ is the posterior probability of each archetype group computed using the truncated stick-breaking (TSB) prior.

3.4.2 Environmental response

The response of species archetypes to environmental variables is shown in Table 3.2, where model coefficients, standard errors (SE) and relative standard errors (RSE) are summarized. Figure 3.4, presents the partial effects for all covariates in the model. Depth appeared to be the single most important variable in defining the distribution of species archetypes. In fact, among the models fitted using a single covariate, depth models showed the lowest BIC value (Table B.4) and, in general, RSE of depth coefficient were low. Archetypes one and two showed a strong positive response to decreasing depths, but archetype two had no response to depth[2]. Archetype three responded more weakly to depth and was associated with intermediate depth values. Archetype four showed a strong response to increasing depth and was associated with the deepest portion of the modelled region. All archetype groups showed a preference for positive bathymetric position index (BPI20) values, suggesting that all the considered species are associated with regional relieves. Particulate organic carbon (POC) was the most important variable for archetype four which showed a strong preference for high carbon export values. Other groups showed a weaker response to POC and were generally associated with lower POC values then archetype four. Archetype three showed a strong negative response to bottom oxygen saturation. Archetype two preferred slightly more oxygenated waters than group three. Archetypes one and four were only weakly

correlated to bottom oxygen saturation. With the exception of group four, all archetypes avoided very low and very high slope values. Archetypes four showed a mild negative response to slope.

Table 3.2 – Model coefficients, standard errors and relative standard errors. Abbreviations: archetype (arch.), proportion of species in each archetype (π), bathymetric position index with a 20 km radius (BPI20), particulate organic carbon (POC) flux, bottom oxygen saturation (B_Ox), linear term ([…].1), quadratic term ([…].2).

Arch.	π	Intercept	Depth.1	Depth.2	BPI20.1	BPI20.2	POC.1	POC.2	B_Ox.1	B_Ox.2	Slope.1	Slope.2
		Coefficient values										
G.1	0.23	-3.2	30.73	6.65	-2.47	-7.39	-5.09	-1.4	0.49	-3.93	4.19	-3.2
G.2	0.23	-1.29	26.15	-0.08	5.44	-10.11	-3.32	-4.08	-3.55	-3.77	3.11	-2.4
G.3	0.23	-1.39	-3.14	-4.04	0.93	-6.54	2.86	-2.46	-11.48	-2.89	-6.15	-5.07
G.4	0.31	-2.29	-23.85	-12.69	4.77	-3.68	8.45	2.98	-3.57	-0.62	-1.72	-1.3
		Stand error of coefficient values										
G.1	0.23	0.26	8.21	5.81	3.09	2.8	1.84	2.25	2.35	1.57	1.86	1.86
G.2	0.23	0.11	3.61	3.55	1.97	1.78	1.24	1.48	1.67	1.07	1.11	1.12
G.3	0.23	0.07	1.67	1.6	1.51	1.5	1.14	1.43	1.92	1.52	1.24	1.43
G.4	0.31	0.1	1.84	1.51	1.46	1.25	1.11	1.32	1.84	1.54	1.12	1.28
		Relative stand error of coefficient values										
G.1	0.23	8.12	26.72	87.37	125.1	37.89	36.15	160.71	479.59	39.95	44.39	58.12
G.2	0.23	8.53	13.8	4437.5	36.21	17.61	37.35	36.27	47.04	28.38	35.69	46.67
G.3	0.23	5.04	53.18	39.6	162.37	22.94	39.86	58.13	16.72	52.6	20.16	28.21
G.4	0.31	4.37	7.71	11.9	30.61	33.97	13.14	44.3	51.54	248.39	65.12	98.46

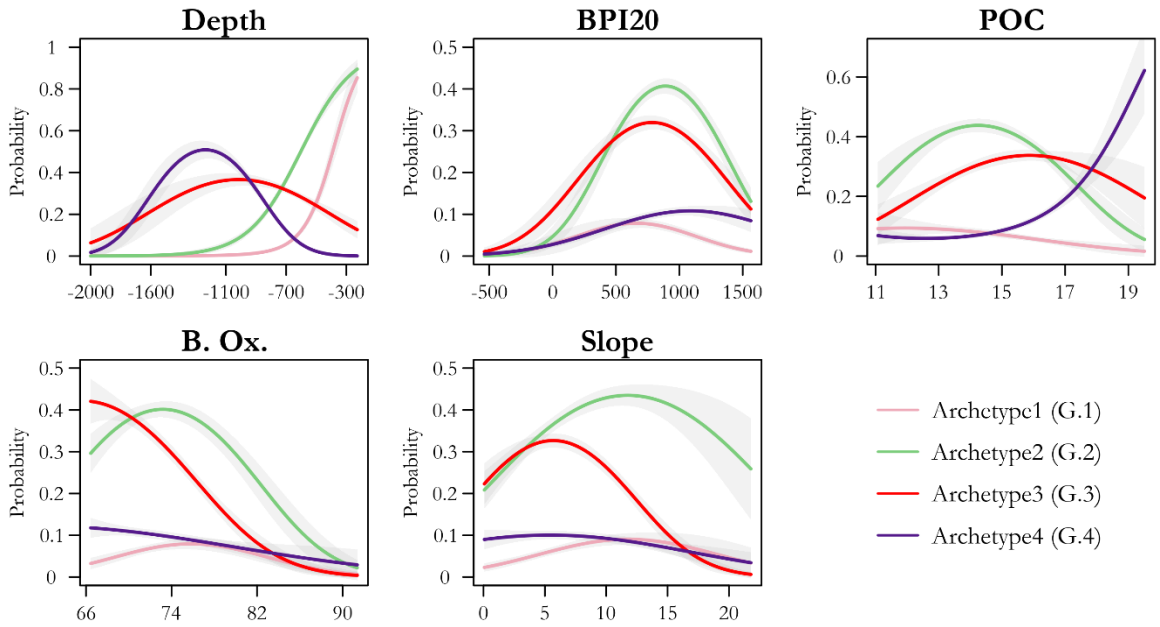

Figure 3.4 – Plot of partial effects for all covariates in the model. Each panel represents the effect of varying each covariate while holding all other covariates at the regional average values. Shaded areas represent the standard errors. Abbreviations and units of measurement: depth (m), bathymetric position index with a 20 km radius (BPI20), particulate organic carbon (POC) flux (mg C_{org} m^{-2} d^{-1}), bottom oxygen saturation (B. Ox.; %) and slope (º).

3.4.3 Benthic invertebrate archetype distributions

There were a number of clear patterns in the spatial prediction of the species archetypes (Figure 3.5, Figure 3.6 Figure 3.7). In general, they showed a marked bathymetric zonation (Figure 3.7a), however they also showed a considerable spatial overlap (Figure 3.7b). Archetype one has the most limited distribution and it is associated with the shallowest relieves and with island slopes (Figure 3.5a). Archetype two has an extensive spatial overlap with archetype 1, but it is also found in deeper waters. In general, it prefers island slopes and the summits of seamounts and crests (Figure 3.5c). Archetype three is located slightly deeper than archetype two. It has a wider distribution than the other two archetypes, it is strongly associated with areas of complex topography, it does not show any latitudinal pattern and it spatially overlaps with archetypes two and four (Figure 3.6a). Archetype four is associated with the deepest relieves and slopes of the modelled area (Figure 3.6c). It presents a wide distribution throughout the study area, however, the highest probabilities of occurrence of this archetype are associated with high latitudes. In general, standard errors were greatest for species archetype one (Figure 3.5b). The high standard errors for this group could result from the fact that, overall, its species summed the smallest number of observations (Table 3.1).

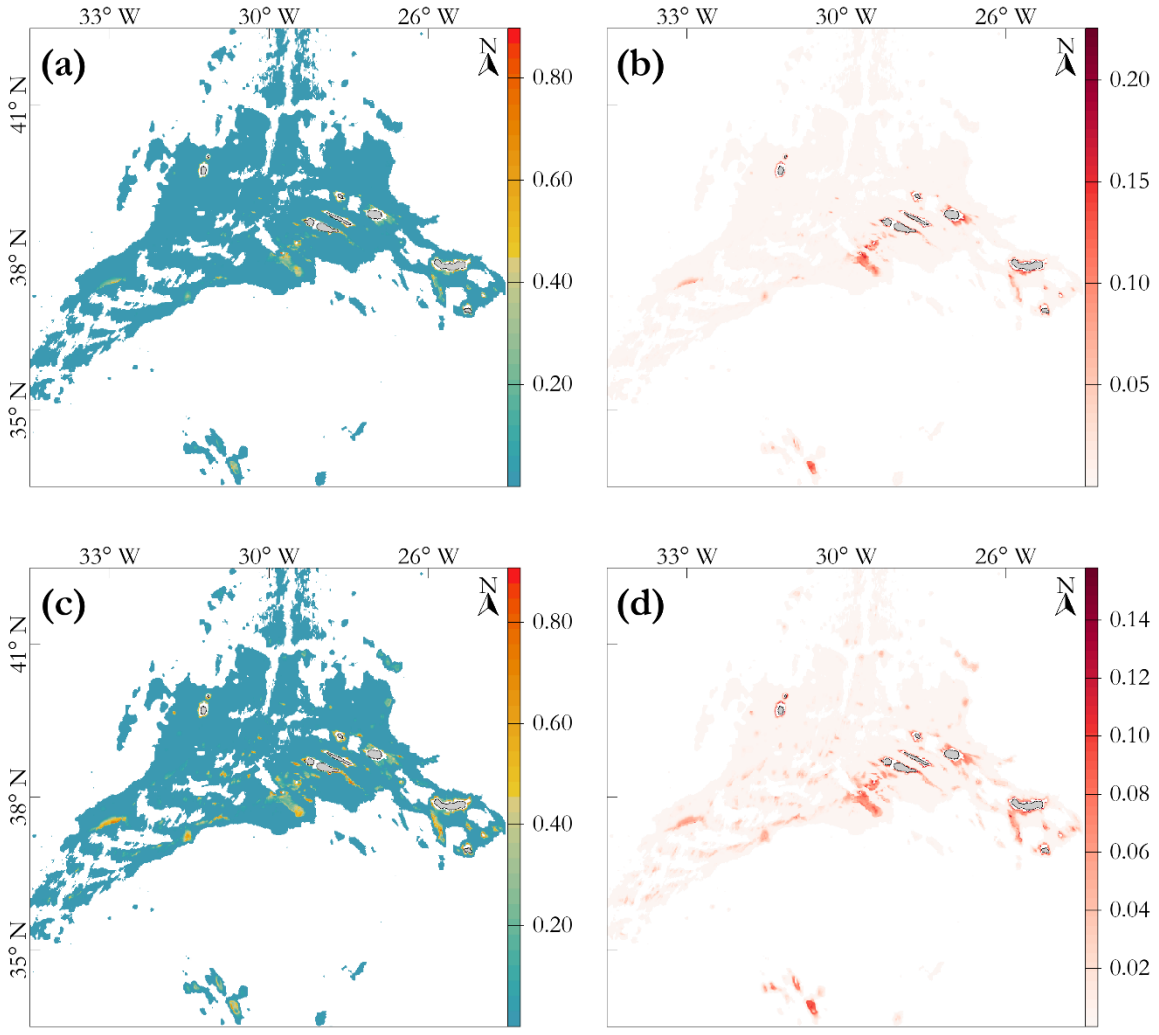

Figure 3.5 – The probability of presence across the study area of species archetypes (a) G.1 and (c) G.2. Standard errors of model predictions of species archetypes (b) G.1 and (d) G.2.

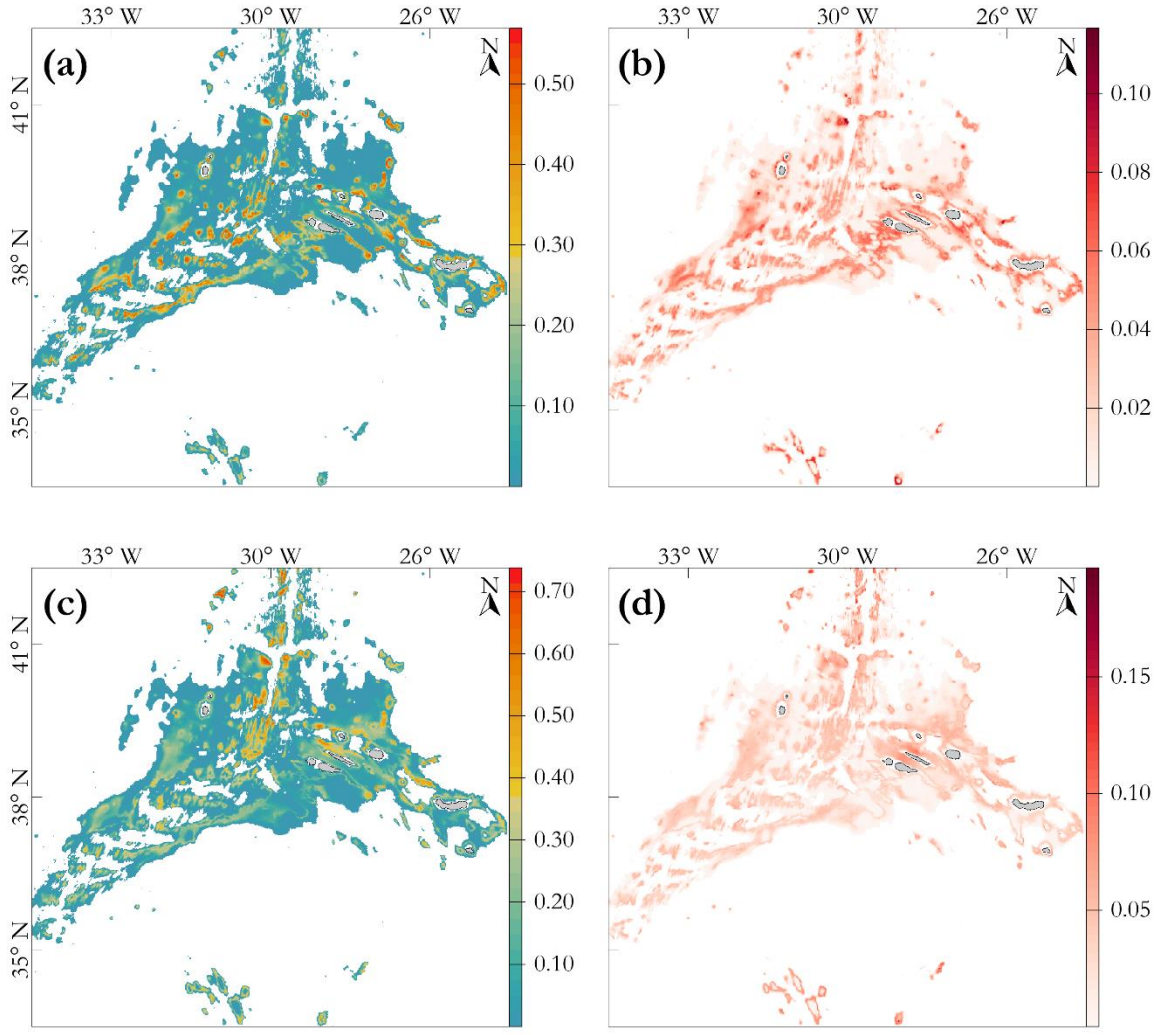

Figure 3.6 – The probability of presence across the study area of species archetypes (a) G.3 and (c) G.4. Standard errors of model predictions of species archetypes (b) G.3 and (d) G.4.

Interesting patterns can be observed projecting the estimated probabilities of presence over two-dimensional plots having depth on one axis and another environmental variable on the other axis (Figure 3.8). In particular, the depth zonation and the partial overlap of the different archetypes become evident. Archetype one is distributed at depths between 200 (the shallowest limit of the study area) and 500 m, archetype two between 200 and 800 m, archetype three between 500 and 1600 m and archetype four between 800 and 2000 m. In general, the depth-wise stratification of the regional water masses matches quite neatly with the vertical succession of the different archetypes (Figure 8). Areas of probable occurrence of archetype one are restricted to a very narrow thermal range. Archetype two shares a similar depth range with archetype one, but it seems able to occupy most of the seafloor locations bathed by surface waters. Archetype three is strongly associated with areas characterized by the lowest oxygenation. These areas correspond to the core depth of intermediate waters. The bathymetric distribution of archetype four is mostly delimited by the core of the deepest limit of intermediate waters (800 – 1600 m). It shows a strong preference for areas presenting high POC flux values. In these areas its depth range extends down to a depth of about 2000 m (the deepest limit of the modelled area).

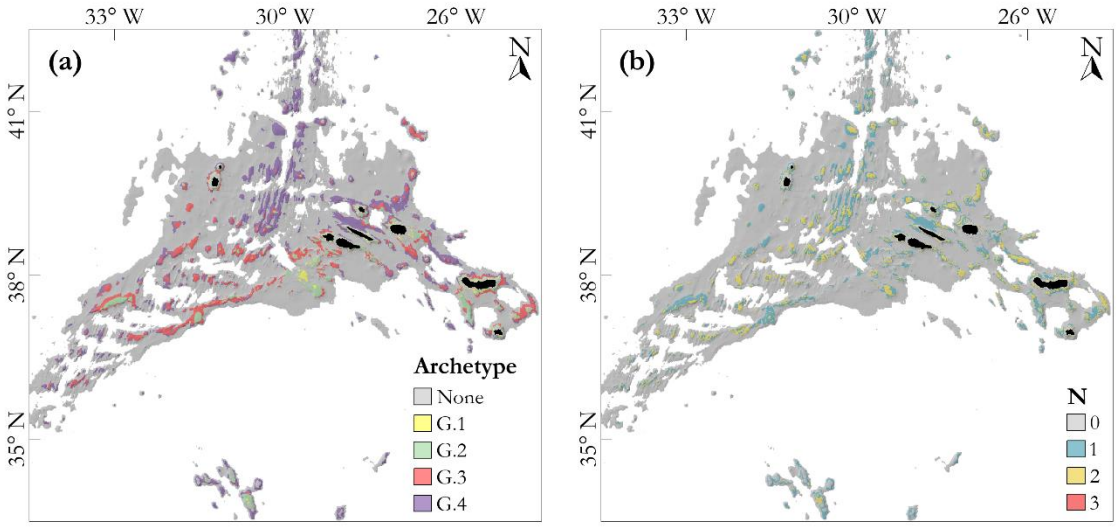

Figure 3.7 – Spatial distribution of species archetypes. For each cell of the model grid, the archetype with the highest probability of presence (a) and the total number of archetypes (b) are shown. In panel (a), if all archetypes in a cell have probabilities smaller than 0.25, then no archetype is shown. In panel (b), only archetypes with probabilities greater than 0.25 are counted.

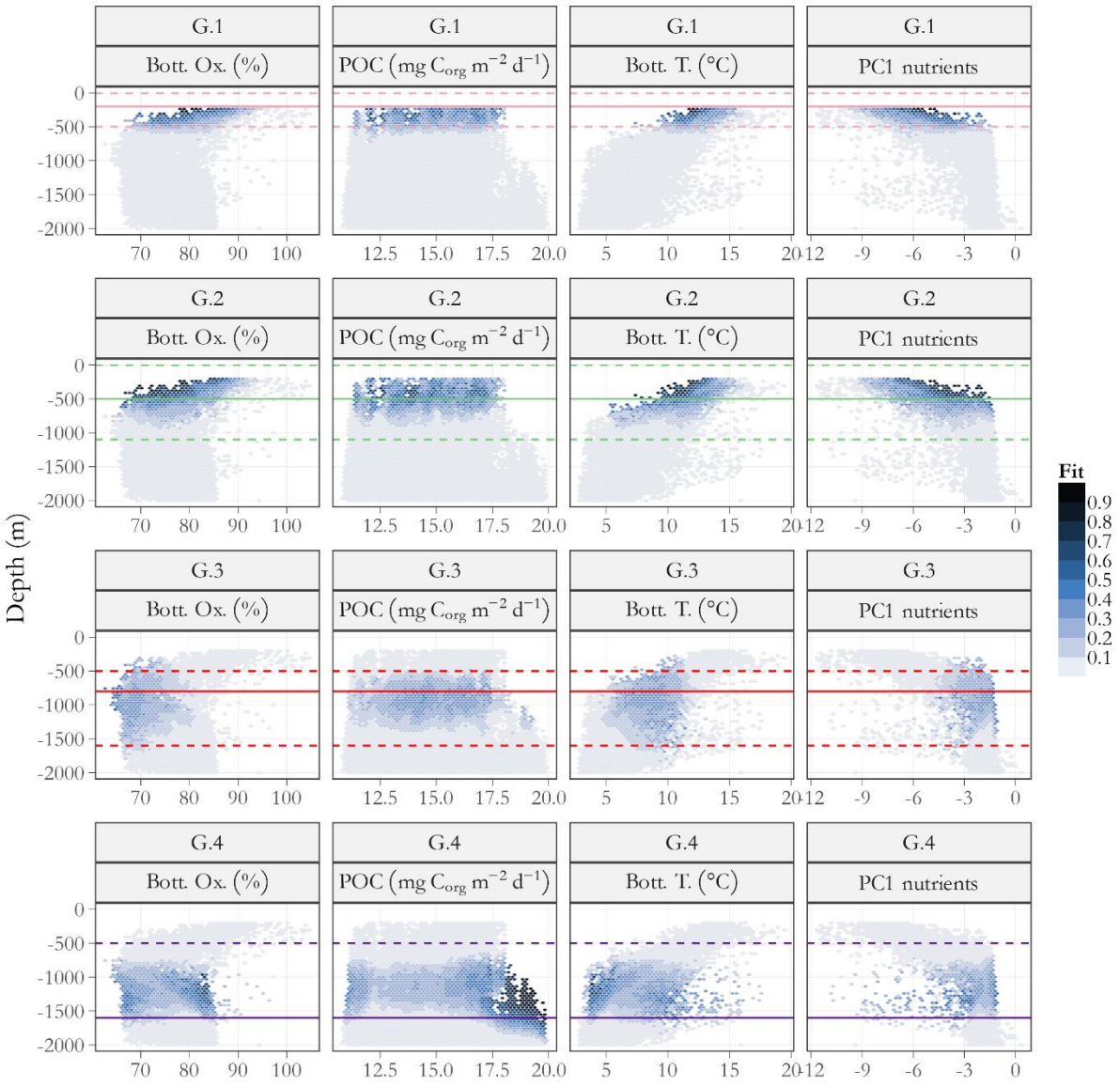

Figure 3.8 – Predicted probability of presence (fit) averaged over 50 two-dimensional bins for different environmental variables against depth. Note that bottom oxygen saturation (Bott. Ox.) and particulate organic carbon (POC) flux were used as model covariates, bottom temperature (Bott. T) and the principal component of sweater nutrients (PC1 nutrients) were not. The horizontal lines show the vertical stratification of the regional water masses. The solid lines identify the core of water masses, the dashed lines delimit the water masses. Each color is associated with one or more water masses: (pink) Western North Atlantic Central Water and Western South Atlantic Central Water; (green) Eastern North Atlantic Central Water; (red) Mediterranean Water and Antarctic Intermediate Water; (purple) Upper North Atlantic Deep Water. The depth range considered is between 200 and 2000 m.

3.5. Discussion

One of the major shortfall of biodiversity knowledge steams from an incomplete description of the geographical distribution of species (Hortal et al., 2015). Overcoming this shortfall is essential for conserving nature and its services and it is a required first step to tackle more complex ecological processes (e.g. dispersal, speciation, disturbance, biotic interactions, etc.) in remote and poorly studied regions such as the deep sea (Hortal et al., 2015; McClain and Hardy, 2010; Whittaker et al., 2005). This study undertakes one of the first statistical bioregionalization of deep-sea benthic assemblages in the Azores. By combining environmental and biological data it provides a simplified model of the distribution of multiple species that, hopefully, will be of help in framing research questions, sampling designs and conservation strategies. Species archetype models (SAMs) are a particularly appealing method that classifies species based on their environmental response and, in this way, simplifies complex multi-species analysis (Dunstan et al., 2011). They implicitly assume that, at certain scales, the observed distribution of a pool of species is mainly driven by the environment and that species within particular geographic regions have a statistically indistinguishable environmental response. These assumptions can be framed within the concepts of environmental filtering and regional species pool (Cornell and Harrison, 2014; Kraft et al., 2015). Environmental filtering can be defined as the effect of environmental conditions selecting those species capable of survival and persistence at a given locale (Kraft et al., 2015). From this perspective, SAMs represent a way to evaluate how plausible is to have environmental filters that spatially sort the regional pool of species into discrete biological regions. Given the complexity of community assembly processes, studies performed exclusively with observational presence/absence data cannot provide conclusive evidence about the role of the environment in structuring regional communities (Cadotte and Tucker, 2017; Thompson et al., 2020; Vellend, 2010). However, acknowledging this limitation, they still provide an indication about the role of the environment and represent a valuable first step to describe regional biodiversity patterns (Woolley et al., 2020). In addition, the effects of the environmental filtering are generally expected to be of relevance at the regional scale addressed here and are particularly evident in marine systems (Levin et al., 2001; Soininen, 2014; Viana and Chase, 2019).

Because of the nature of most of the biological data available, it was decided to focus exclusively on large benthic filter feeders found on hard substrates at depths between 200 and 2000 m (bathyal zone). Although a bioregionalization exercise based on a limited number of taxonomic groups may not be easily generalized (Costello et al., 2017), the use of few and carefully selected taxa presents its advantages. Limiting the analysis to hard substrate communities served to address the regional sampling bias that did

not allow for a proper description of soft substrate communities. Recent surveys (e.g. Dominguez-Carrió et al., 2021; Morato et al., 2021a) confirm that most of the selected species represent a dominant component of coral and sponge aggregations throughout the study region. Cold-water corals and deep-sea sponges tend to be long-lived and highly vulnerable to physical disturbance and climate change (e.g. Carreiro-Silva et al., 2013; Morato et al., 2021b, 2020c; Vieira et al., 2020), thus the species known to aggregate and enhance the structural complexity of deep-water habitats are generally treated as a priority target for conservation (FAO, 2009; OSPAR, 2010). Additionally, the sessile species considered here are likely to play a major role in building and maintaining deep-sea communities as they seem to possess some of the key characteristics of foundation species (Crotty et al., 2019; Ellison, 2019): (i) they represent a dominant element of hard substrate communities in the region reaching high local densities (e.g. Morato et al., 2021b; Tempera et al., 2012a); (ii) they likely mediate several non-trophic interactions especially via habitat provision and nutrient cycling (Beazley et al., 2013; Buhl-Mortensen et al., 2010, 2017; Maldonado et al., 2015; Orejas and Jiménez, 2017). Considering their importance for deep-sea communities, their sessile nature and their long lifespan, these species seem to be good candidates to guide a preliminary bioregionalization using biological and environmental data. Their sessile nature simplifies the identification of environmental structures of biological relevance. In fact, wherever they exist, we can reasonably infer that suitable conditions exist since the moment of their settlement up to the moment of their death. Given their long lifespan, if different pools of deep-sea corals and sponges are actually sorted by different environmental conditions, then we can assume that a long standing and biologically relevant environmental heterogeneity exists in the region. Such heterogeneity can be expected to influence, directly or indirectly, all the species interacting with the foundation species. Other components of the ecosystem are also likely to be influenced by the same heterogeneity. Based on these hypotheses, it is possible formulate new research questions and sampling designs addressing, for example, species abundance distributions, functional traits or intermittent interactions among species and relevant environmental structures (e.g. feeding behavior, diel migrations, interactions during particular life stages, etc.).

The SAMs revealed a number of interesting spatial patterns. As reported in previous studies (Chapter 1, Braga-Henriques et al., 2013; Menezes et al., 2006; Tempera et al., 2012a, 2012b), the strongest distributional patterns in the region occur along the bathymetric gradient confirming, once again, the depth-wise zonation of deep biota (e.g. Carney, 2005; McClain and Rex, 2015). Liu and Tanhua (2021) characterized and mapped the major water masses in the Atlantic Ocean showing that they are distributed

in four main isopycnal (vertical) layers separated by surfaces of neutral density: (i) upper, (ii) intermediate, (iii) deep and overflow and (iv) bottom layer. In the Azores, several overlapping water masses exist within each layer (Figure 3.2). Consequently, it is not possible to detect a relationship between a specific archetype and a specific water mass with the data and the methodology used here. However, our results show a clear correlation between the vertical zonation of the regional benthic fauna and the depth-wise succession of the isopycnal layers. The boundaries of the different layers seem to act as transitional ecoclines rather than as sharp biogeographic breaks. In fact, according to model predictions, the distributions of the different archetypes present several overlapping regions along the depth gradient.

Both archetype one and two were found within the upper layer with archetype one presenting a limited distribution and archetype two being more widespread (Figure 3.5). Temperature was not directly included as a model covariate, but it was considered in the interpretation of the results because of its high correlation with depth ($r > 0.9$). When plotting the predicted probabilities of occurrence of archetype one and two against depth and temperature, archetype one seems to be under a strong thermal control as all suitable locations occur within a narrow range of seafloor temperatures (9 – 15 ºC) (Figure 3.8). Temperature is a major driver of biodiversity because of its physiological importance and it is especially relevant at shallow depths where it likely represents the dominant form of energy (Carney, 2005; McClain and Rex, 2015; Woolley et al., 2016). Therefore, a strong thermal control over certain species inhabiting the uppermost part of the water column seems a reasonable model output. Furthermore, a thermal control over archetype one is supported by the pioneering work of Rakka et al. (2021) on the early life history of deep-sea octocorals. They showed that even small changes in temperature affect embryo development and larval behavior of *Dentomuricea* aff. *meteor* (archetype one). Increasing the incubation temperature from 13 to 15 ºC, the embryo developed faster and the larvae presented a higher swimming velocity and a longer pelagic phase. Species archetype one is mostly found within the upper 500 m of the water column, at depths where there is the largest influence of the water masses found in the upper layer (Liu and Tanhua, 2021; Puerta et al., 2020). Both at global and regional scales, this depth has been considered as an important boundary between coastal and deep-sea habitats (Carney, 2005; Costello et al., 2017; Glover et al., 2022). This study confirms the importance of such boundary and, rephrasing Carney (2005), suggests that it could represent the bathymetric limit of species truly restricted to the upper water masses. Note that previous studies on the distribution of demersal fishes and cold-water corals in the Azores position this bathymetric boundary slightly deeper (600 m), however they confirm the existence of a transition zone at a similar depth (Braga-Henriques et al., 2013; Menezes et al., 2006).

Archetype two has a wider thermal range than archetype one and a slightly preference for less oxygenated waters but, in general, it seems to tolerate most of the conditions existing in the topmost portion of the water column. Based on model predictions, archetype two can potentially co-occur at almost all locations suitable for archetype one and at certain locations with archetype three (Figure 3.5 and Figure 3.6). It was predicted to occur at depths between 200 and 800 m. At 800 m is located the core of the intermediate layer in the region and an oxygen minimum (ca. 4.3 mL L^{-1}) (Amorim et al., 2017; Bashmachnikov et al., 2015; Liu and Tanhua, 2021). As confirmed by previous studies (Braga-Henriques et al., 2013; Menezes et al., 2006), this depth seems to be another important biogeographic boundary in the region. Watling and Lapointe (2022) found a faunal transition at 1000 m in the North East Atlantic which is the same depth where, on average, is situated the core of the intermediate layer (Liu and Tanhua, 2021; Puerta et al., 2020). In general, both regional and basin-scale patterns support the idea that intermediate water masses represent an important biogeographic boundary for benthic assemblages in this part of the Atlantic (Puerta et al., 2022, 2020).

Setting aside the vertical zonation for a moment, it is worth to spend a few more lines about the spatial patterns that appear to be in place at depths between 200 and 500 m. As discussed earlier, there is a group of species strongly controlled by temperature and strictly located within the upper layer (200 – 500 m). Warm and 'shallow' habitats tend to have the highest level of species richness (Costello and Chaudhary, 2017; Menegotto and Rangel, 2018; Saeedi et al., 2019). Within a similar depth range, there is a second group of species not showing any significant response to temperature or POC fluxes that apparently can colonize most of the rocky seafloor locations. Based on these results, we can set forward an interesting hypothesis to be addressed in future studies. Nestedness could be more important than spatial turnover in explaining the level of among-site variation (β-diversity) within the upper layer (Baselga, 2010; Ulrich and Almeida-Neto, 2012). Areas suitable for archetype one could represent areas where species restricted to the upper layer and species in-between the upper and the intermediate layer can coexist. These areas would present the highest number of species. Moving away from these locations the species restricted to the upper layer would be 'lost' and a nestedness pattern would arise. Preliminary observations confirm that species belonging to archetype one and two can coexist in very dense and particularly rich benthic aggregations, e.g. *Dentomuricea* aff. *meteor* (archetype 1) and *Viminella flagellum* (archetype 2) (Morato et al., 2020b, 2019). Environmentally driven nestedness has been shown in previous studies (e.g. Menegotto et al., 2019; Stuart et al., 2017).

Archetype three seems to occur only at locations bathed by intermediate water masses at depths between 500 and 1600 m (Figure 3.2 and Figure 3.8). Archetype three showed a strong negative response to bottom-water oxygen saturation, suggesting that it is associated with the Antarctic Intermediate Water and with the Mediterranean Water. These are the most prominent intermediate water masses at the considered latitudes and longitudes and are both characterized by low oxygen concentrations (Bashmachnikov et al., 2015). As suggested by Menezes et al. (2006), the regional oxygen minimum could exert a direct control over demersal fish assemblages. However, a direct physiological role of bottom-water oxygen concentrations on macrobenthic invertebrates was mostly ascribed to conditions of hypoxia or anoxia (less than 2 mL L^{-1}) (Levin, 2003; Puerta et al., 2020; Rogers, 2000), while the water column throughout the Azores is oxygen-rich with oxygen levels always above 4.3 mL L^{-1} (Amorim et al., 2017). The recent publication of Puerta et al. (2022) confirms that the Mediterranean Water can locally influence the structure of deep-sea megabenthic communities and exert control over connectivity and biogeographic patterns at large spatio-temporal scales. These finding support the idea that species archetype three is composed of species truly restricted to intermediate water masses. In the North East Atlantic, Watling and Lapointe (2022) using distributional data of deep-sea coral (Class Anthozoa) found a fauna transition at around 1500 m which is very close to the deepest limit of the intermediate layer in the region (1600 m).

Archetype four occurs in the deepest portion of the considered bathyal range, mostly at depths between the core and the lowest limit of the intermediate layer (800 – 1600 m). Based on model predictions, archetype four can potentially co-occur with archetype three. Preliminary video analysis confirms that species of these archetypes can coexist in dense assemblages, e.g. *Narella versluysi* (archetype 3) and *Pheronema carpenteri* (archetype 4) (Morato et al., 2020b). The sampling effort below 1500 m depth was very scarce and only three stations were available for the present analysis. Therefore, all results concerning the deepest portion of the study area should be considered cautiously as the maximum depth limit of this archetype and some of its environmental responses could change as new data become available. Nevertheless, we believe that sufficient data was available to correctly identify its minimum depth limit. Archetype four was the archetype showing the strongest response to POC flux. Because of this response, its predicted probabilities of occurrence reached the highest values at high latitudes, in accordance with the productivity gradient that exists in the region (Amorim et al., 2017; Caldeira and Reis, 2017) (Figure 3.6 and Figure 3.8). Interestingly, archetype four seems able to colonize areas deeper than 1600 m only when high carbon export values are present. In general, a high dependence on food supply

is a common pattern in the deep sea and, as depth increases, the chemical energy (export productivity) seems to be the main driver of deep-sea diversity (McClain and Rex, 2015; Woolley et al., 2016). A dependence on export productivity for many of the species included in archetype four has been confirmed by several studies (Fink et al., 2015; Hanz et al., 2021; Hebbeln et al., 2020; Howell et al., 2016; Juva et al., 2020; Lim et al., 2020; Raddatz et al., 2020). Interestingly, Portilho-Ramos et al. (2022) showed that food supply controlled by export production and turbulent hydrodynamics at the seabed exerted the strongest impact on the vitality of *Desmophyllum pertusum* (archetype four) over the last 20,000 years. The high importance of POC flux for the archetype inhabiting the deepest portion of considered depth range seems ecologically sound. Further sampling will be required to confirm our findings.

3.6 Conclusions

This study represents one of the first statistical bioregionalization of deep-sea benthic communities in the Azores. It focuses on habitat-forming filter feeder species (Phyla Cnidaria and Porifera) known to represent a critical component of hard substrate communities in the region. Many of the described patterns are of general interest as they seem to be caused by environmentally driven processes that most likely affect the functioning and the structure of several components of deep-sea ecosystem in the Azores as well as in other regions. The depth-wise succession of the regional water masses and their stratification into different isopycnal layers appears to explain much of the changes in species composition occurring with depth. The ecological processes driving these patterns remain unclear. A water mass can be considered as an environmental envelop characterized by a particular combination of oceanographic parameters (Emery, 2001; Liu and Tanhua, 2021). They could exert a direct physiological control, influence connectivity or mediate food supply via internal waves or density-driven currents (Carney, 2005; Hanz et al., 2021; Lapointe et al., 2020; Puerta et al., 2020; Quattrini et al., 2017; Radice et al., 2016; Roberts et al., 2021). Regardless of the main ecological and evolutionary processes in place, our results suggest that the different isopycnal layers can be used as meaningful bioregions, setting an interesting background for ecological research, ecosystem-based management and spatial monitoring. In the region, the main changes in species composition occur at 500, 800 and 1600 m generally confirming the biogeographic boundaries identified in previous studies. Temperature is the main environmental factor controlling the distribution of assemblages truly restricted to the upper layer. A strong negative response to bottom-oxygen concentrations identifies assemblages truly restricted to the intermediate layer. The isopycnal stratification appear to act as a soft ecological boundary for other species. Food export seems the main controlling factor for species associated with the deep layer (or with the lowest portion of the

intermediate layer), supporting the hypotheses that the regional productivity gradient in surface waters affects deep-sea assemblages. Increasing the study area further north and south will help clarify if there is a fauna transition owed to differences in primary productivity and the relative weight of this ecological gradient in the biogeographic boundaries that were described close to the Azores (Costello et al., 2017; Watling and Lapointe, 2022). Substrate complexity and, in particular, the spatial distribution of patches of hard substrates is likely to be an important distributional driver missing from this analysis. However, we believe that its importance is more evident at finer spatial scales than those considered here. The greatest drawbacks of this work were the paucity of sampling below 1500 m and the lack of a covariate measuring the strength of steady and variable flows at the seafloor. There is growing evidence that the distribution of deep-sea filter feeders is strongly controlled by vertical as well as lateral food supply mediated by hydrodynamic processes (e.g. Hanz et al., 2021; Jansen et al., 2018; Lim et al., 2020; Mohn et al., 2014). The development of a high resolution regional oceanographic model and new biological sampling in deep areas will largely benefit future ecological studies in the region. Species abundance distribution, functional trait, genetic, taxonomic and life history data should also be a priority if we want to understand how selection, drift, speciation, and dispersal (Vellend, 2010) interact at different scales and influence the community assembly process in the region.

Chapter 4
Translate users' perceptions of landscapes and seascapes into computer representation: The R package *scapesClassification*

4.1 Abstract

The R package *scapesClassification* was designed to translate users' perceptions of landscapes and seascapes into computer representation. Particular points of views are encoded in a transparent and explicit manner, can be reproduced on different machines and, ultimately, can be considered as task- or domain-oriented ontologies. The proposed classification algorithms allow to encode mental conceptualization of spaces taking into account salient geographic features and expected topological relationships between classes. In doing so they mimic the way people schematize in their mind spatial configurations and shorten the distance between mental and computer representations. The proposed functions can be used for a range of purposes including: exploratory analyses, pattern retrieval, formalization of expert knowledge into ontology-like scripts, evaluation of automated classification methods, labelling of natural processes and integration of spatial analyses. It has a unique way of evaluating heuristic rules and it is able to bridge raster and object representations of geographic spaces. Ultimately, it represents an additional resource to perform spatial analysis within the R-environment.

4.2 Introduction

The ever-increasing stream of geospatial data is restructuring the way geographic knowledge discovery is addressed (Mennis and Guo, 2009; Reichstein et al., 2019). Particular interest is given to computer based methods able to detect hidden patterns in spatial data and in automated or semi-automated classification processes able to synthetize large quantities of data. In general, it has been advocated that geosciences should be switching from a model-driven to a data-driven approach (Guo et al., 2014; Miller and Goodchild, 2015). Although data-driven science offers exciting opportunities for geosciences (and for the analysis of spatial data), it still poses some challenges such as (i) interpretability of results, (ii) inconsistent or implausible model predictions when assessed in the light of scientific laws or domain knowledge and (iii) conceptual difficulty in labelling natural features and processes – labels that are required for model training and cross-validation (Reichstein et al., 2019). The integration of expert knowledge in data driven models has been proposed to soften these limitations (Arvor et al., 2019; Karpatne et al., 2017). In addition, experts of several scientific domains are not familiar with automated or semi-automated analysis based on machine and deep learning (Farley et al., 2018). Yet, their knowledge remains valuable and when taken into account in a transparent and rigorous way offers a wide range of opportunities for improving research and decision-making processes (Drescher et al., 2013).

One of the ways to capture expert knowledge and use it in the workflow of spatial analysis is to translate human perceptual knowledge into computer representations. As explained in Arvor et al. (2019), the formalization of spatial expert knowledge deals with several challenges. The same geographic concept can be defined in multiple ways and, generally, each definition represents a different point of view. For example, a forest can be defined based on functional, administrative or both functional and administrative attributes producing similar, yet different, representations. A second challenge relates with the duality of geographic concepts. For instance, a mountain can be conceptualized from its real-world physical properties (e.g. a natural elevation rising abruptly from the surrounding level) and from surrogate measures (e.g. topographic position index greater than a certain value). When surrogate measures are used, the definition of classification thresholds tends to be context- and user-specific. Finally, the semantics used to describe geographic concepts are interpreted differently by people having different cultural, social and professional backgrounds (i.e. semantic gaps).

Studies in the area of geospatial semantics try to overcome the complexities inherent to spatial concepts using domain-level and task-oriented ontologies (Fonseca et al., 2002; Kokla and Guilbert, 2020;

Schuurman and Leszczynski, 2006). Ontologies are formal and explicit specifications of shared concepts in a machine-readable way that are used to (i) encode and share expert knowledge and to (ii) bridge human and computer representations (Sun et al., 2019). A practical application of spatial ontologies is to make expert views of geographic concepts explicit, transparent and sharable and to facilitate their integration into computer-based analysis (Arvor et al., 2019; Guilbert and Moulin, 2017). Guilbert and Moulin (2017) point out that people identify geographic entities through salient features that they recognize. For instance, a mountain is commonly perceived as a region of space having a peak and an island as a region of space characterized by a land-sea interface. They suggest that the notion of salient feature could help in the formal definition of spatial ontologies and in the development of general purpose classification software. Classifications can be implemented as multi-step processes where (i) salient features are used to identify a particular type of geographic object and (ii) additional descriptors at locations in the neighborhood or in continuity with the salient feature can be used to define the spatial extent of that object.

Here, we present the open-source R package *scapesClassification*, designed to translate user mental models of seascapes and landscapes into computer representations (classifications). It allows users to encode their spatial perception of a geographic space using spatial data, salient geographic features and the spatial relationships existing among different classes or objects. Users can define flexible and multi-step classification schemes by combining different functions. All classification steps are saved as reproducible outputs (i.e. R scripts) that can be regarded as sharable domain- or task-oriented ontologies. This package deals with raster data and its development was specifically directed toward the classification of digital terrain models (DTMs) for ecological applications. DTMs are surface elevation data providing digital representations of land surface topography (Lecours et al., 2016; Pike et al., 2009; Sofia, 2020). Geomorphological mapping is central to geomorphometry, the science of quantitative land-surface analysis (Lecours et al., 2016; Pike et al., 2009; Sofia, 2020). Geomorphometry plays a central role in landscape and seascape ecology where it helps studying how spatial patterns and spatial heterogeneity influence ecological processes at multiple scales (Pittman et al., 2021; Swanborn et al., 2022; Turner, 2005). In the context of ecological sciences, the topographic setting plays a major role in altering environmental conditions such as climate or oceanic surface and near-bottom currents influencing in this way energy availability and spatial distributions of organisms and communities (e.g. Durden et al., 2020; Irl et al., 2015; Jackson, 2010). Geomorphological maps (classifications) have a wide spectrum of end-uses and are performed using a huge variety of techniques and algorithms (e.g. Jasiewicz and Stepinski, 2013;

Lecours et al., 2016; Pike et al., 2009; Seijmonsbergen et al., 2011; Sofia, 2020; Walbridge et al., 2018). This huge spectrum of classification methodologies is linked to different mental models of space generally representing different domain views useful to address particular aspects of natural phenomena (Freksa and Barkowsky, 1995; Kokla and Guilbert, 2020; Schuurman, 2006; Sofia, 2020). This package does not pretend to substitute well-established geomorphometry methodologies. On the contrary, it can be used to integrate them placing the focus on the end-user and on its specific needs in addressing domain- or task-oriented problems. The suit of functions and algorithms it provides can be used to classify and segment a geographic space using n-dimensional raster layers which can contain elevation and topography-derived layers. At the same time, classifications obtained with different methodologies can be integrated at any point during the classification process as categorical layers and used to direct classification steps.

The software scapesClassification is based on few key concepts:

- Landscapes and seascapes tend to have prominent features easy to identify. These features can be considered as *anchor locations*, locations around which a classification process can start and evolve.
- A classification process can take into account the spatial relationships that are expected to exist among different classes or objects, i.e., where a segment of space or an object is expected to be in relation to other segments and objects.
- Based on such relationships, it is possible to estimate where a certain class is expected to exist and to perform *focal evaluations* of classification rules: rules are only evaluated at suitable locations, thus, limiting possible misclassification cases.
- A classification process is seen as multi-step: as new portions of a raster are classified they can be used to define new focal areas over which classification rules are evaluated.

The software is illustrated through a study case. Seafloor features around the central part of the Azores archipelago (Figure 4.1) are classified from the point of view of a benthic ecologist. Results are evaluated and integrated by an unsupervised classification method.

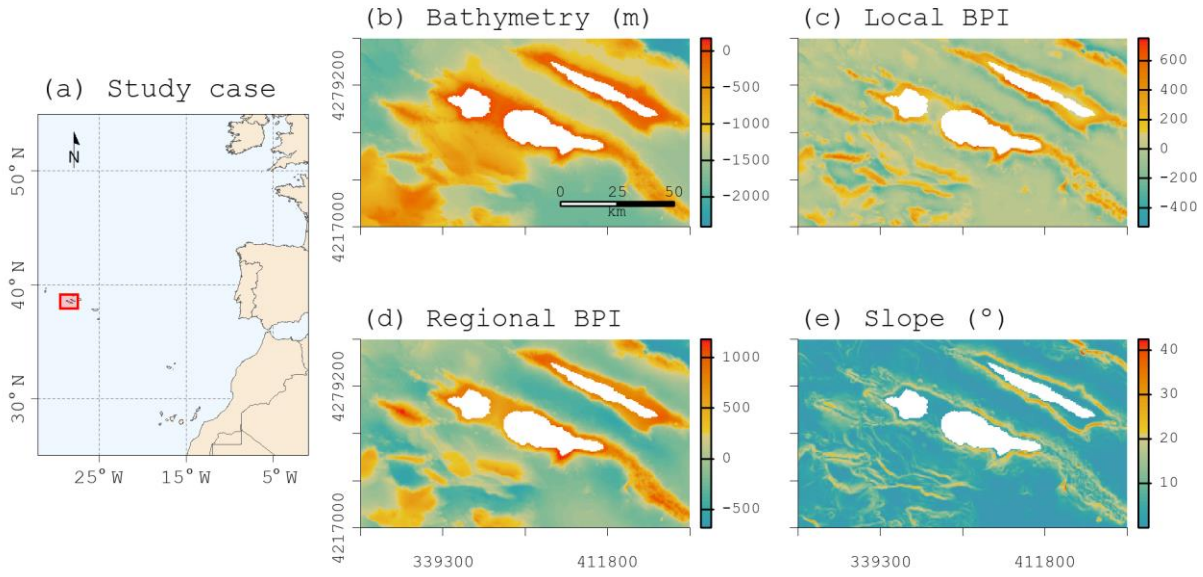

Figure 4.1 – Study case. Central group ('Triangle') of the Azores Archipelago composed of São Jorge (NE), Faial (SW) and Pico (SE) Islands (a). Raster data used to identify seafloor structures (b-e). Panels (b-e) are projected to UTM 26N.

4.3 Materials and Methods

4.3.1 The R environment

R is a multiplatform and free software environment for statistical computing and graphics (R Core Team, 2022). It is widely used for spatial analysis (Bivand, 2021) and it is well integrated in the workflow of ecological and environmental studies (Huang and Yang, 2022; Lai et al., 2019). Packages are the fundamental units of reproducible R code (Wickham, 2015).

4.3.2 Software availability

The latest *scapesClassification* release is available on the Comprehensive R Archive Network (CRAN) (https://cran.r-project.org/package=scapesClassification). The latest software development is available on github (https://github.com/ghTaranto/scapesClassification). Details about its implementation are included as function documentation and on the software page (https://ghtaranto.github.io/scapesClassification/index.html). The code written for the study case and provided in Appendix C can be used as a tutorial guide (all input data are provided as part of the package). Finally, *scapesClassification*

depends on the R package *terra* (Hijmans, 2022) for reading, handling and plotting raster data. It does not have any other external dependency.

4.3.3 (scapes)Classification

The following sections present the general architecture and the logic behind *scapesClassification*. Classification are performed by assigning a numeric identifier to raster cells respecting user-defined conditions. The same numeric identifier can be assigned to all raster cells respecting user-defined conditions. In alternative, a different numeric identifier can be assigned to each discrete group of cells respecting user-defined conditions. In the latter case, each group of cells can be treated as a raster object characterized by identity, spatial embedding and attributes. In the context of *scapesClassification*, discrete groups of cells refer to two or more groups of raster cells respecting a set of conditions separated by raster cells *not* respecting the same set of conditions. Classifications can be implemented in a single-step or, more commonly, in a multi-step process. In a multi-step process, a class is firstly identified based on some salient feature or by its spatial proximity to another class that was previously identified. Next, every location mapped to a class can serve as an anchor location, a location around which classification rules are used to refine the class boundaries. Each step can be implemented using a classification or a raster object function. All the functions of *scapesClassification* are listed and briefly illustrated in Table 4.1. See the package documentation for more information (https://ghtaranto.github.io/scapesClassification/index.html).

4.3.3.1 Raster data

A raster is a spatial data model that divides space in regular shaped and sized pieces (Peuquet, 1984) referred to as cells or pixels. Each cell stores one or more values describing quantitative or qualitative attributes of a geographic space. In the package *terra*, cells are numbered from the upper left cell to the upper right cell and then continuing on the left side of the next row, and so on until the last cell at the lower-right side of the raster (https://rspatial.org/). Functions in *scapesClassification* use this numbering to treat raster data as numeric vectors and data frames.

The functions of *scapesClassification* (Table 4.1) segment raster space into distinct classes and identify and manipulate raster objects within a class. They can occasionally consider spatial vector data to highlight particular locations on a raster. Groups of raster cells are mapped to a class or to an object by a unique numeric identifier. User-defined classification rules take into account data stored as single or multi-layer raster.

Table 4.1 – List of scapesClassification functions. Additional details and visual examples are available in the reference documentation (accessible using the R functions *'?'* or *'help'*) and on the documentation page (https://ghtaranto.github.io/scapesClassification/reference/index.html). Cell refers to raster cell. Function type: format input function (FI); helper function (H); classification function (anchor cells) (C_a); classification function (user-conditions) (C_{uc}); classification function (seed cells) (C_s); classification function (neighbors only) (C_{nbs}); raster object functions (RO). Rule types: absolute test cell condition (ATC); absolute neighborhood condition (ANC); relative focal cell condition (RFC); relative neighborhood condition (RNC).

Function name	Function type	Summary	Rule types	Focal evaluation
attTbl()	FI	Convert raster data into an attribute table	N/A	N/A
ngbList()	FI	Computes the neighborhoods of all cells included in the raster table	N/A	N/A
conditions()	H	Check for spelling and syntax errors in a condition string. Its documentation provide an overview of all rule types	N/A	N/A
cv.2.rast()	H	Convert a class vector to a raster	N/A	N/A
anchor.svo()	C_a	Returns a vector of raster cell numbers extracted at the locations of spatial vector objects	N/A	N/A
anchor.cell()	C_a	Converts a vector of cell numbers into a class vector	N/A	N/A
peak.cell()	C_a	Identify local maxima or minima on a raster surface	N/A	N/A
cond.4.all()	C_a, C_{uc}	Evaluate conditions for all unclassified cells and classify them if conditions are true	ATC	NO
anchor.seed()	C_a, C_s, RO	Returns a vector of cell numbers at the locations of seed cells and growth buffers	ATC, RFC	YES
cond.4.nofn()	C_{uc}, RO	Evaluate conditions for cells neighboring specific classes and classify them if conditions are true. With the argument '*hgrowth = TRUE*' modifies RO	ATC, ANC, RFC, RNC	YES
cond.reclass()	C_{uc}	Evaluate conditions for cells of a class and reclassify them if conditions are true	ATC, ANC	YES
reclass.nbs()	C_{nbs}	Evaluate if members of two classes are contiguous and, if they are, one of them is reclassified	N/A	N/A
classify.all()	C_{nbs}	Classify all cells in that have not yet been classified assigning them to the same class of the nearest classified cell	N/A	N/A
obj.border()	RO	Identify the borders of raster objects	N/A	N/A
obj.nbs()	RO	Identify the shared borders of neighboring raster objects	N/A	N/A
rel.pi()	RO	Compute the relative position index of raster objects	N/A	N/A
pi.sgm()	RO	Segment raster objects based on position index values	N/A	N/A
pi.add()	RO	Add new raster objects based on position index values	N/A	N/A

4.3.3.2 Format inputs

In an initial step, the format input function '*attTbl*' (Table 4.1) convert the raster data into a data frame (referred to as attribute table) and compute the neighborhood of every cell included in the attribute table. Format input functions ignore raster cells having missing values. Attribute tables come with a column named '*Cell*'. This column stores raster cell numbers and associates each row of the attribute table to a cell of the raster. The remaining columns store the data contained in the raster layers. The list of neighborhoods of each raster cells is computed using the function '*ngbList*' and stored in memory to reduce the computational time of classification algorithms. In fact, as explained in the next sections, classification functions use cell neighborhoods to determine at what locations a classification rule has to be evaluated. The neighborhood of a raster cell with coordinates (x, y) counts 8 neighbors with coordinates (x±1, y), (x, y±1) and (x±1, y±1). Cells on the edge of the raster and cells adjacent to cells with missing values have less than 8 neighbors.

4.3.3.3 Class vector

Classification and object functions (Table 4.1) return seascapes or landscapes classifications in the form of a 'class vector'. A class vector stores classification steps by mapping groups of raster cells to numeric classes (or to a missing value if a cell is unclassified). The n^{th} element of a class vector corresponds to the raster cell stored in the n^{th} row of the attribute table. Class vectors serve as both function input and output. Functions receiving a class vector as an input only evaluate classification rules for unclassified cells; previously classified cells retain their classes. This allows for a sequential update of a class vector and, thus, for multi-step classifications. In addition, class vectors can be added as new columns of the attribute table and used to focus particular classification steps within or outside specific areas. The helper function *cv.2.rast* converts a class vectors into a raster (Table 4.1).

4.3.3.4 Rule evaluation

One of the key features of *scapesClassification* is the focal evaluation of classification rules. The term focal refers to the fact that, in general, classification rules are evaluated only in the neighborhoods of cells of a particular set of classes (focal cells). Cells testing positive to the classification rules are classified as focal cells and the same set of rules is automatically applied to new neighboring unclassified cells. The process is repeated as long as new cells are classified as focal cells. In this way, cells that respect the same rule(s) and that are connected the same focal cells are joined into the same class. This process is referred to as class continuity in the software documentation. The documentation of the helper function '*conditions*' (Table 4.1) provides a general overview of the type of rules that can be used and on how they are

evaluated. A more exhaustive explanation is provided on the software documentation page (https://ghtaranto.github.io/scapesClassification/articles/ghp/scapesClassification_01_2_RulesEvalTypes.html)

4.3.3.5 Classification rules

Classifications are based on heuristic rules. Classification rules evaluate to either true or false and determine what raster cells have to be classified. They are passed to *scapesClassification* functions as a single character string and consist of any valid combination of variables, arithmetic (+, -, *, /, ^, %%, %/%), relational (>, <, >=, <=, ==, !=) and logic operators (&, |), and base R functions (e.g., abs(variable_name)). All variables included as columns of the attribute table can be accessed by name. For instance, if there is a column named "slope", it is possible to build a classification rule such as "slope > 5".

Classification rules can be absolute or relative. Absolute conditions compare raster cells and threshold values. The simplest case is when the value of a test cell is compared against a threshold value and the cell is classified if the condition is true (e.g., "slope > 5"). This evaluation is referred to as *absolute test cell condition* in the software documentation. Absolute test cell conditions can be evaluated by the functions '*anchor.seed*', '*cond.4.all*', '*cond.4.nofn*' and '*cond.reclass'* (Table 4.1). A second type of absolute condition (*absolute neighborhood condition*) compares the value of a test cell and of the cells in its neighborhood against a threshold value (nine evaluations). The test cell is classified if the number of positive evaluations is equal or greater than the ones specified by the argument *peval* (positive evaluations). This type of condition is flagged in a condition string by curly brackets (e.g., "slope {} > 5"). Absolute neighborhood condition can be evaluated by the functions '*cond.4.nofn*' and '*cond.reclass*' (Table 4.1).

Relative conditions compare the value of a cell against the values of adjacent cells. In focal evaluations, classification rules are always tested in the neighborhood of a focal cell. A *relative focal cell condition* compares test and focal cell values. This type of condition is flagged by square brackets (e.g., "slope [] > slope"). Square brackets identify the focal cell. Relative focal cell conditions can be evaluated by the functions '*anchor.seed*' and '*cond.4.nofn*' (Table 4.1). A second type of relative condition (*relative neighborhood condition*) compares the value of the test cell against the values of the cells in its neighborhood (eight evaluations). The test cell is classified if the number of positive evaluations is equal or greater than the ones specified by the function argument *peval* (positive evaluations). This type of condition is flagged by curly brackets (e.g., "slope {} > slope"). Curly brackets identify the test cell neighborhood. Relative neighborhood conditions can be evaluated by the functions '*cond.4.nofn*' and '*cond.reclass*' (Table 4.1).

Finally, using the argument *min.bord* – available for the function '*cond.4.nofn*' – it is possible to specify that an unclassified cell is mapped to a class only if the classification rule is true and if it shares with that class a border of two or more cells. In this way, classes are less likely to present irregular boundaries.

4.3.3.6 Classification functions

Classification functions are used to map a raster cell to a specific class (Table 4.1). Anchor cell functions can be used to initialize the classification process by identifying salient features. The seed cell function (*anchor.seed*) can be used to identify a set of cells respecting a particular condition (seed cells) and to evaluate additional sets of conditions to compute growth or isolation buffers around seed cells. This function can also be used as a raster object function by assigning a unique numeric identifier to each discrete group of seed cells. User condition functions evaluate classifications rules at all locations on a raster grid ('*cond.4.all*'), in the neighborhood of specific classes ('*cond.4.nofn*') or to reclassify particular groups of cells ('*cond.reclass*'). The function '*cond.4.nofn*' can also be used as a raster object function setting the argument homogenous growth ('*hgrowth'*) as true. Neighbor only functions performs classification based on the neighborhoods of raster cells. They do not evaluate condition strings.

4.3.3.7 Raster objects

In the context of *scapesClassification*, a raster object is defined as a discrete patch of cells that share some common properties and that have the same numeric identifier (ID). Similar objects can be grouped into the same class. For example, the discrete elevations of an area constitute individual relieves; the set of all relieves constitutes the relief class. Raster object functions are presented in Table 4.1.

The function '*anchor.seed*' identifies raster objects in two steps: (i) it finds cells having specific attributes (e.g., cells surrounding a peak with a position index above a certain threshold); (ii) it assigns a unique identifier to each discrete group of cells (e.g., each discrete patch composed of a peak and of the cells surrounding the peak receives a different ID).

The function '*rel.pi*' computes the relative position index (rPI) of a variable within a raster object. Position index values are standardized or normalized and are computed using the formulas:

- Standardized rPI = (x - mean(x)) / sd(x)
- Normalized rPI = (x - min(x)) / (max(x) - min(x))

Mean, sd, max and min refer to the mean, the standard deviation, the maximum and the minimum of the variable x within the raster object where the index is computed. Other position index functions can use the rPI to segment existing objects ('*pi.sgm*') or to identify new ones ('*pi.add*').

The function '*cond.4.nofn*' implements an algorithm for the homogeneous growth of raster objects (argument *hgrowth = TRUE*). As explained in the rule evaluation section, the evaluation of classification rules is an iterative process that happens in the neighborhood of focal cells. The homogeneous growth algorithm considers, at turn, every raster object as a focal object. Cells of the 1st object are considered as focal cells. Cells in their neighborhood respecting the classification rules are considered as part of the object and regarded as focal cells in the next iteration. Then, the algorithm considers the 2nd object, then the 3rd and so on until the last one. At this point the iteration is completed and a new iteration starts as long as new cells were classified in the previous iteration.

The border function '*obj.border*' identifies the borders of raster objects. The function '*obj.nbs*' identifies the shared borders of neighboring raster objects.

4.3.4 Case study

The R package *scapesClassification* is presented through a study case. Seafloor features around the central part of the Azores archipelago (Figure 4.1a) are classified from the point of view of a benthic ecologist. Island shelf (ISU) and relief (RU) units are identified. Relief units are divided into discrete raster objects. All classification steps can be reproduced using the R scrips provided in Appendix C. In addition, every script has the code to generate an interactive map that facilitates the visualization of the classification results.

The packages *rassta* (Fuentes et al., 2021) was used to complement and compare the user-defined classification with an unsupervised classification method. The packages *data.table* (Dowle and Srinivasan, 2021), *mapview* (Appelhans et al., 2021) and *leaflet* (Cheng et al., 2021) were used to produce interactive maps of the classification results.

4.3.4.1 Benthic ecology perspective

The case study follows the perspective of a benthic ecologist. The classification was motivated by a hypothetical research question: besides being influenced by depth, the distribution of benthic communities changes with elevated features. Thus, elevated features with different shape, area and depth profile could host different communities.

Elevated features are mapped to two different classes the island shelf unit (ISU) and the relief unit (RU). The relief unit is segmented into distinct raster objects that could later be used to help in the design of ecology and spatial analysis, and management strategies.

4.3.4.2 Descriptors

Geomorphic features are identified using 4 raster layers that includes bathymetry and bathymetric derivatives (see Lundblad et al., 2006; Walbridge et al., 2018):

- Bathymetry: depth values (m);
- Local BPI: bathymetric position index (BPI) computed with an outer radius of 10px (ca. 5 km);
- Regional BPI: bathymetric position index (BPI) computed with an outer radius of 40px (ca. 20 km);
- Slope: slope values (º).

The raster layers are plotted in Figure 4.1b-e and are available as part of the package data. They have a resolution of 500m and are projected to UTM zone 26N.

4.3.4.3 Computer representations

The next paragraphs describe the user mental model that guided the classification of the different seafloor geomorphic structures. The actual classification steps are presented as reproducible scripts in Appendix C.

Island shelf unit (ISU). In a simplified representation, the ISU is composed of two main elements: (i) shelves (i.e., relatively flat areas surrounding islands) and (ii) slopes (i.e., areas that connect island shelves to the seafloor). Shelves are adjacent to an island by definition. This property can be considered as a salient feature of the ISU and can be used to identify an initial set of raster cells (*anchor.cells*) that are adjacent to land cells. Interconnected cells adjacent to anchor cells with a gentle slope (slope ≤ 5) are part of the shelf (*flat cells*). Finally, unclassified cells surrounded by shelf cells are also part of the shelf (*hole cells*). In this model a shelf is represented by anchor, flat and hole cells.

Slopes can be considered as areas that connect the shelf to the seafloor. Seafloor and bottom slope areas in proximity of an elevated features generally have negative or near-zero BPI values (Walbridge et al., 2018). Based on the previous properties, it is possible to define a *slope cell* as a cell that has positive BPI values *and* that is contiguous to a shelf or to a slope cell.

Peak cell (PC). Peak cells are used to identify RUs – each relief unit evolves around one or more peak cells. Peak cells are identified using the function *peak.cell*. The function uses a functional definition of local maxima: a raster cell constitutes a local maximum if its value is larger than the values of all the cells in its

neighborhood. Peaks on small relieves (low BPI values) and peaks within the ISU were excluded because they were not considered relevant for this benthic ecology perspective classification.

Relief unit (RU). In a simplified representation, the RU can be conceptualized as set of discrete features that rise above the seafloor and that have one or more local peaks. Each discrete feature can be mapped to a unique identifier and be treated as an individual raster object. Using a simple set of rules, a cell can be considered to be part of an elevated feature if it is connected to a peak by cells of the same class and if it has a positive BPI value. These simple rules, implemented with the function *anchor.seed*, were adequate to identify the relief unit as a whole, but produced some large and some small relief objects. Position index functions (Table 4.1) were used to segment very large objects, remove very small objects and add objects having prominent ridge-like structures high local BPI values.

Position index functions (e.g. '*pi.sgm*', '*pi.add*') can be used to remove cells below a certain relative positon index from a raster object. In this way, a single raster object can be segmented into two or more groups of cells. If the groups have a sufficient size (e.g. >40 cells), then each group become an independent object. After the segmentation, the removed cells can be reassigned to the closest raster object using the homogeneous growth algorithm implemented by the function '*cond.4.nofn*'.

4.3.4.4 Self-organizing map

The self-organizing map (SOM) is an artificial neural network in which neurons are typically arranged as a two-dimensional lattice (Kohonen, 1998). It facilitates the analysis of n-dimensional datasets and can be used to reduce the dimensionality of a dataset to facilitate the discovery of hidden patterns in data by enhancing data visualization. It can be coupled with a cluster analysis (CA) to identify groups of observations sharing a similar pattern (Kaufman and Rousseeuw, 1990). Here, SOM and CA were used for two simple tasks and their use was illustrative rather than exhaustive. SOM and CA were performed using the R package rassta (Fuentes et al., 2021) and the code of all the analysis is shared in Appendix C. The descriptors of Figure 4.1b-e were used as function inputs.

The first task was to visually investigate if the user-defined classification captured the underlying heterogeneity existing in the spatial data. In particular, if it could discriminate accurately between elevated and non-elevated features. In this type of evaluation only two classification outcomes were considered (elevated, non-elevated), therefore, the maximum number of possible clusters was set to two. As this was considered to be a preliminary visual analysis, SOM and CA functions were left with default parameters. The SOM and CA algorithms were initialized 100 times. Every time, the resulting cluster

presenting the shallowest average depth was labelled as the 'elevated feature cluster'. The number of times a cell is included within the 'elevated feature cluster' was considered as a data-driven measure of how likely it is that it actually belongs to an elevated feature.

The second task was to assess if the spatial heterogeneity existing within the relief unit could be captured by meaningful clusters. Descriptor values were clipped to the extent of the relief unit, and the previous analysis was repeated. In this case no maximum number of cluster was expected. Therefore, the optimal number of clusters was estimated using the partitioning around medoids method (Kaufman and Rousseeuw, 1990) and the gap statistic (Tibshirani et al., 2001) as implemented in rassta. Again, this was considered to be an illustrative analysis, thus, SOM and CA functions were left with default parameters.

4.4 Results

The results of each classification script were saved and provided as part of the package data with the file names *ISU_Cells.RDS*, *PKS_Cells.RDS* and *RU_obj.RDS*. In addition, all classification steps and analysis can be reproduced with the scripts provided in Appendix C.

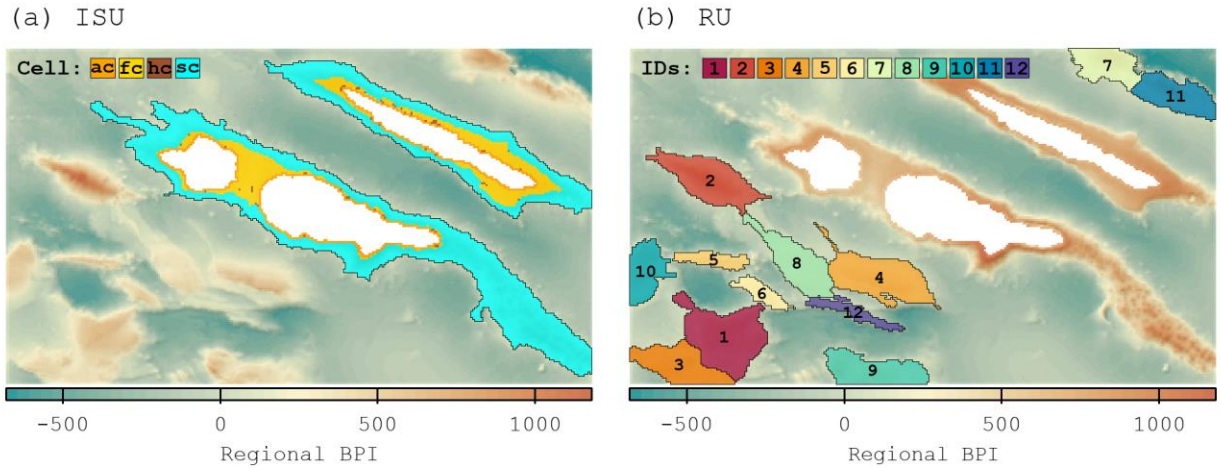

Figure 4.2 – User-defined classification of the island shelf unit (ISU) and of the relief unit (RU). The ISU (a) is composed of anchor (ac), flat (fc), hole (hc) and slope cells (sc). The RU (b) is composed of 12 raster objects, each identifying a distinct elevated feature. Raster objects identifiers (IDs) are ordered from the shallowest to the deepest.

4.4.1 (scapes)Classification

As expected, both the island shelf unit (ISU) and the relief unit (RU) overlaid areas characterized by positive regional BPI values (Figure 4.2). The ISU was composed of two discrete areas, one around São Jorge and one around Pico and Faial Islands (Figure 4.2a). The shelves around each island mostly counted 'flat cells'. 'Anchor cells' followed the boundaries of the island landmasses and defined the salient feature used to initialize the ISU classification. A small percentage of the ISU was composed of 'hole cells' (cells neither classified as anchor or flat cells but that were surrounded by cells of these classes). 'Slope cells' comprised the largest portion of the ISU mainly because of the shelf extending southeast of Pico Island. Probably this area would be better described by an additional class targeting ridges connected to an island shelf. However, this was out of the scope of the present classification.

The RU was composed of twelve discrete features each considered as a distinct raster object (Figure 4.2b). The classification of raster objects consisted of four steps (Figure 4.3). In the first step, the function '*anchor.seed*' identified areas (i) outside the ISU, (ii) connected to a peak and (iii) with positive regional or local BPI values (Figure 4.3a). In the second step, the function '*pi.sgm*' segmented the raster objects by removing cells with negative relative position index (rPI). In addition, it removed objects with less than 40 cells (Figure 4.3b). In the third step, the function '*pi.add*' identified one prominent features (>40 cells) unconnected to existing raster objects that had high local BPI values (local BPI > 100) (Figure 4.3c). Finally, the function '*cond.4.nofn*' assigned raster cells having high local or regional BPI values (BPI > 100) to the closest raster object using the algorithm homogeneous growth (argument *hgrowth* = TRUE) (Figure 4.3c). The final object IDs were sorted based on depth values (the feature having the shallowest minimum depth has the ID number 1).

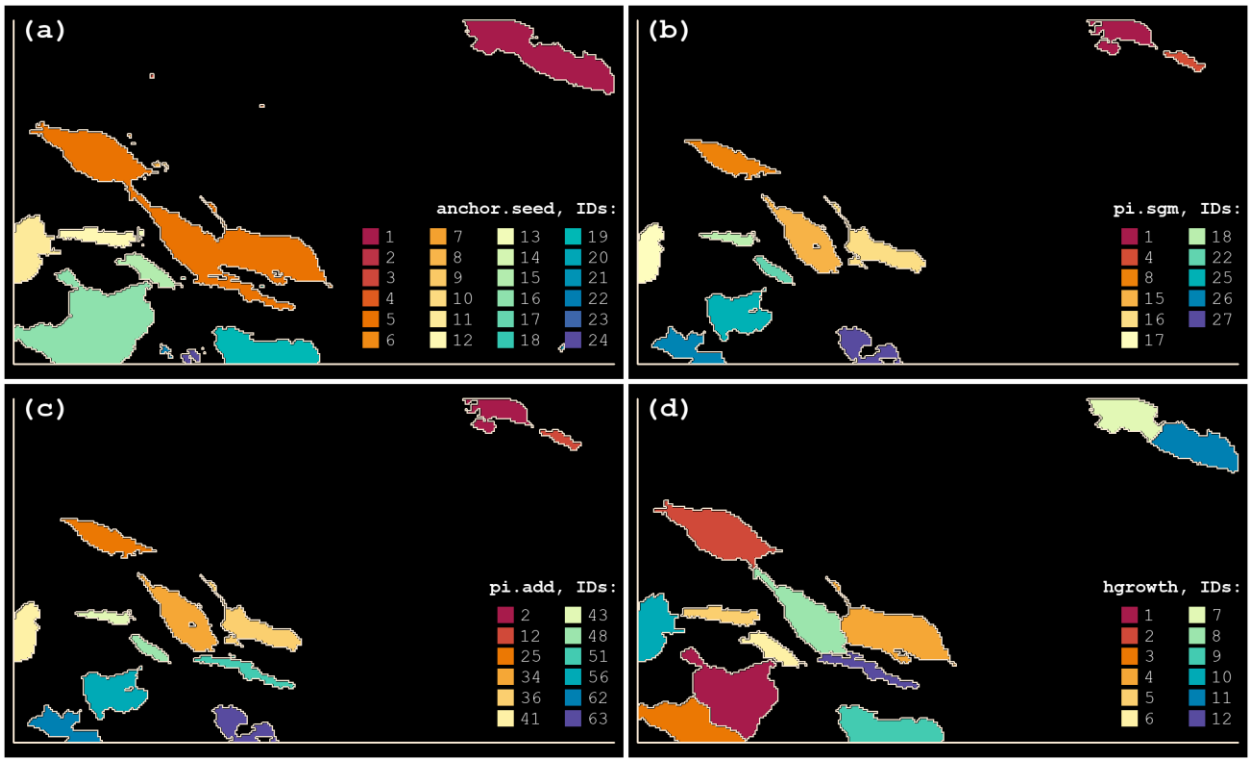

Figure 4.3 – Raster object segmentation of relief units (RU) using the functions '*anchor.seed*' (a), '*pi.sgm*' (b), '*pi.add*' (c) and '*cond.4.nofn*' (d).

4.4.2 Self-organizing map and cluster analysis

Raster cells outside the ISU and the RU were generally included in the non-elevated feature cluster (grey background, Figure 4.4). The few cells outside ISU and RU that were included in the elevated-feature cluster corresponded to objects that were removed along the classification process because of their small size (Figure 4.3). Cells within the RU and the ISU were assigned to the elevated-feature cluster with a certain consistency. Regions characterized by high regional or local BPI or by steep slopes showed the highest consistency. The group of flat cells with high BPI values was the second group of cells most consistently assigned to the elevated-feature cluster. Results for cells characterized by low BPI and slope values were more variable. This last group of cells likely correspond to a transition zone connecting rough and flat areas. Likely, it can be considered as a fuzzy boundary around elevated features.

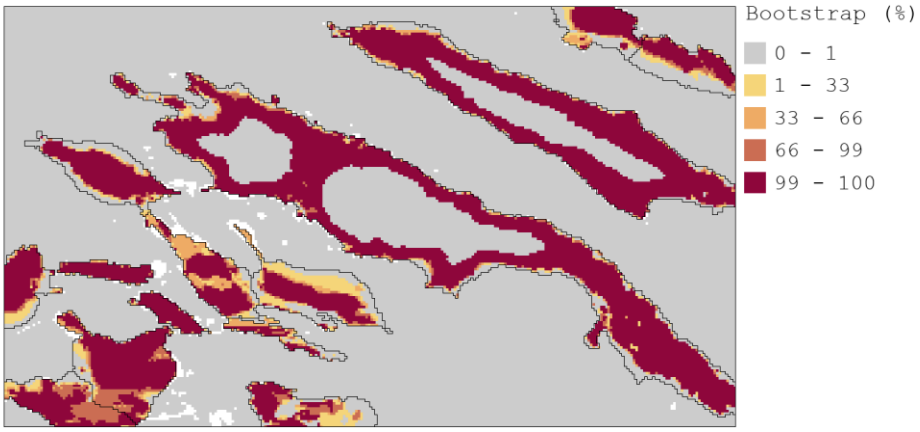

Figure 4.4 – Number of times a cell was assigned to the 'elevated feature cluster' when performing an unsupervised classification based on the self-organizing map technique coupled with cluster analysis. Polygon lines identify the borders of the island shelf unit (ISU) and of the relief unit (RU). White cells are cells included in the 'elevated feature cluster' that were neither in the ISU or in the RU.

In fact, the results of the SOM-CA performed within the RU (Figure 4.5) shows that cluster 4 mostly overlaps with this transition zones. It is characterized by high depth values and by low BPI and slope values. Cluster 1 includes relatively shallow and steep cells with high BPI values likely corresponding to shallow ridges. Cluster 2 has high slope and local BPI values, but low regional BPI values if compared to cluster 1. It likely corresponds to deep ridges and slopes. Finally, cluster 3 includes relatively shallow and flat cells characterized by high regional/low local BPI values. It likely corresponds to flat areas located near the summits of the elevated features.

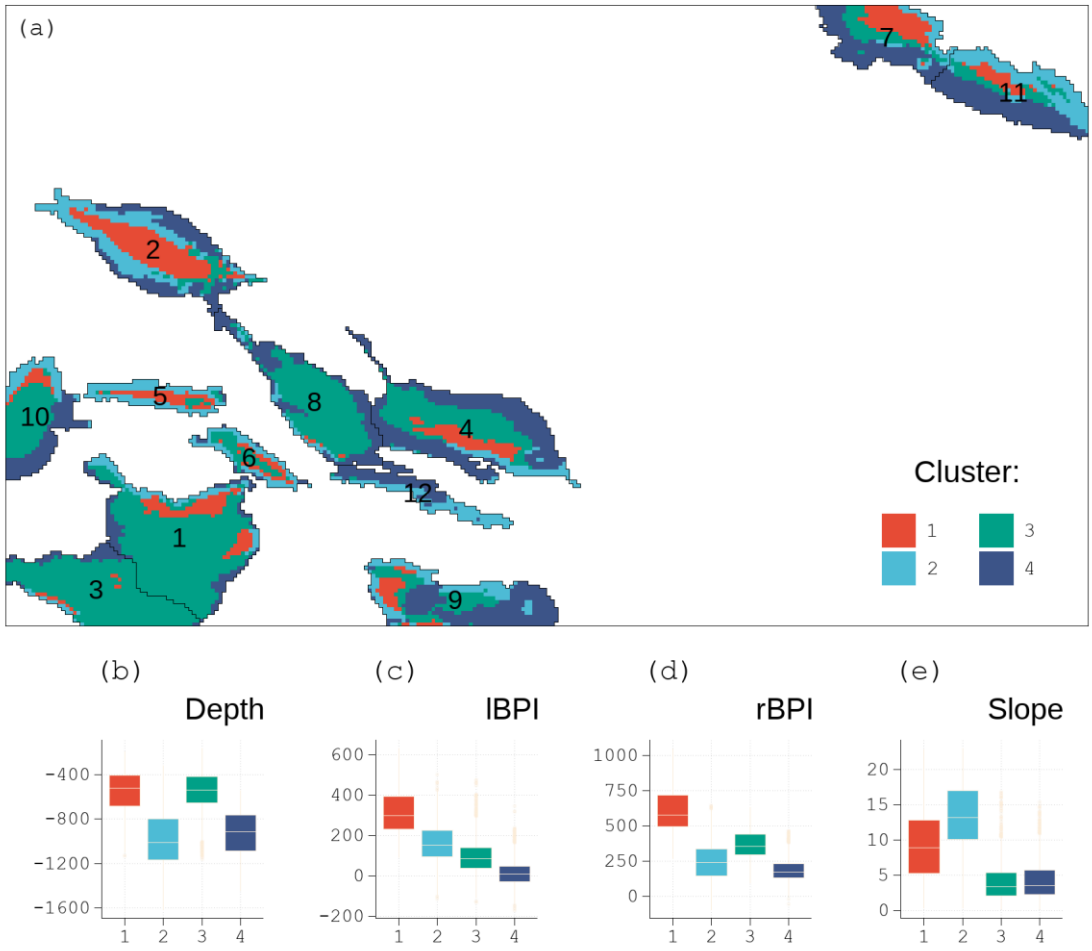

Figure 4.5 – Unsupervised classification of raster the cells included within the relief units. Four clusters were identified (a), each having a distinct terrain signature (b-e). lBPI is local bathymetric position index and rBPI is regional bathymetric position index.

4.5 Discussion

The translation of mental models into computer representation and the retrieval of latent knowledge from spatial data has been a central theme in GIS science (Freksa and Barkowsky, 1995; Kokla and Guilbert, 2020; Schuurman, 2006). Traditional approaches have focused mostly on the definition of general formalizations of perceptual knowledge, however, the importance of task and domain-oriented representations of space has been highlighted in literature (Janowicz et al., 2012; Sun et al., 2019). Different mental models of space generally represent different world views useful to address particular aspects of natural phenomena. The diversity of points of views challenges the knowledge transfer from

one discipline to another, yet it is required to solve particular domain problems (Schuurman and Leszczynski, 2006). In fact, a representation may be suitable for a purpose but unsuitable for another.

The suit of functions available within the *scapesClassification* package provides a series of algorithms for the recovery of particular patterns form n-dimensional raster surfaces. Classifications are based on heuristic rules. While this was a traditional approach in the GIS community (Bishr, 1998), the way they are evaluated is unconventional. The idea of focal evaluations, with rules evaluated only at specific locations, allows users to encode into their classification the topological relationships they expect between and within classes. A particular class can be initially defined based on its salient feature following an innate cognitive process people use to schematize in their mind spatial configurations (Zelinsky-Wibbelt, 1990). In a second step, the class boundaries can be redefined by joining into the same class all the interconnected cells that respect a particular set of rules. Once a particular class is fully defined, it can be used to guide the identification of other classes expected to be adjacent to it. The use of spatial configurations to model geographic spaces represents a powerful method to translate user perceptions into computer representations (Chen et al., 2015). The proposed functions allow for an added flexibility in the encoding of users' perceptions as they can evaluate relative rules (rules that compare the values of adjacent cells) and, therefore, take into account expected trends (i.e., a cell is classified as long as its value is decreasing, increasing or stationary when compared to the value of the focal cell). Another feature of *scapesClassification* is the possibility to segment a raster space into discrete features. The use of discrete spatial units allows for an alternative conceptualization of a study area moving the focus from the field (raster) to the entity (object) (Bishop et al., 2012; Cova and Goodchild, 2002).

The retrieval of recognizable patterns from spatial data can help in the formulation and testing of research and management questions (look for recognizable principle, Andrienko and Andrienko, 2006). In fact, expert judgment is commonly used to develop and evaluate projects at the stages of hypothesis generation, sample design, model development, and interpretation of results (Fazey et al., 2005; Martin et al., 2012) and in the decision-making process (Drescher et al., 2013). In the case study, a deep sea benthic ecologist intends to study the distribution of the seafloor fauna associated with elevated features evaluating if features with different latitude, shape, area and depth profile could host different communities. He tacitly takes into account two other elements to design his classification. First, he knows that the environmental layers he will use to model species distributions have a coarse resolution. Therefore, clipping these data to regions having similar depth profiles could help him detecting hidden latitudinal trends or feature-specific trends (divide and group principle, Andrienko and Andrienko, 2006).

Second, he knows that the local government is planning to implement a new system of marine protected areas. Based on his previous experiences, he believes that the definition of discrete units (relief unit objects), easy to visualize, understand and include in management plans will improve his capacity to synthetize and present his and previous works (e.g., Combes et al., 2021) to policy-makers. While similar applications are available to classify a raster space based on heuristic rules such as the R package terra (Hijmans, 2022) or the Benthic Terrain Modeler (BTM) tool (ArcGIS 10.x) (Walbridge et al., 2018), these applications do not allow for an easy segmentation of the raster space into discrete units. On the contrary other applications are particularly suited to segment a geographic space into discrete spatial units such as software implementing Object-Based Image Analysis (OBIA) approaches (e.g. Arvor et al., 2013; Dekavalla and Argialas, 2017). However, in this latter case, while a user's domain knowledge is used to classify the segmentation output, he/she has little control over the segmentation step implemented with the OBIA approach (Arvor et al., 2019). The R package *scapesClassification* does not aim at substituting these approaches. In fact, any type of raster layer can be included as an input to *scapesClassification* functions. In this way it is possible to integrate (*scapes*)classifications with other classification approaches, building classification rules that take into account, for example, the results of an OBIA segmentation to modify the individual objects based on users' needs encoded in the form of heuristic rules.

Empowering a user with a software that facilitates the retrieval of expected patterns from a spatial dataset represents a way to externalize his tacit knowledge (Fazey et al., 2005). Tacit knowledge is acquired by experience and its externalization is an important pre-requisite for data sharing and knowledge transfer (Chergui et al., 2020; Hampton et al., 2015; Martin et al., 2012). Based on his previous experiences, a particular user could be familiar with a dataset and, thus, expect in advance some pattern type. Having a tool to classify a geographical space based on his previous experience, he is likely to improve his capacity to tailor question oriented analysis. In addition, when using a command program like R to generate user-defined classifications, all classification steps can be encoded into a reproducible script causing several immediate benefits. During the process of encoding and documenting classification steps, the user is forced to reflect about his mental representation and because of this reflection he is likely to recall and organize more effectively his knowledge (Fazey et al., 2005). Script-based classifications can be publicly evaluated and, thus, facilitate open science (Hampton et al., 2015). In the case study, the classification of seafloor structures from the perspective of a benthic ecologist could be considered as an artifact that do not reflect the real variability of the data. However, the classification was explicit and it could be tested (Figure 4.4). It seemed to approximate well the variability existing in the data and it is likely to represent

'*bona fidae*' boundaries (*sensu* Smith and Varzi, 2000). Finally, the sharable classification script can be considered as a task or as domain ontology which embed the view a user has of a space and translate it into a computer representation (Janowicz, 2012). Sharable and explicit ontologies represent a machine mediated communication between humans (Arvor et al., 2019; Janowicz, 2012) that favors the true reproducibility of scientific research (Hampton et al., 2015; Lowndes et al., 2017). In addition, it represent one of the mechanism to label environmental phenomena and include them into automated analysis (Sun et al., 2019).

The R programming language is nowadays wieldy used for spatial analysis and data visualization (Bivand, 2022, 2021; Boettiger et al., 2015; Kelleher and Braswell, 2021). A package like *scapesClassification* developed within the R environment facilitates the inclusion of users' points of views into analytical workflows in an explicit and transparent manner. Considering the study case as an example, in a first step the user defined his objective-oriented classification of seafloor structures. In a second step he used the R package *rassta* to evaluate if meaningful cluster could be identified within the relief unit. The results give the user a general understanding of the relative position of hard-bottom (clusters 1 and 2) and of soft-bottom (clusters 3 and 4) areas, a key information for biodiversity studies on deep-sea benthos. The use of a scripted language to integrate different types of analysis facilitates the communication of a research workflow and of the ideas that sustain it, important aspects to grant the (re-)usability of scientific results (Chen et al., 2019).

4.6 Conclusions

The R package *scapesClassification* was designed to translate users' perceptions of landscapes and seascapes into computer representation. Particular points of view are encoded in a transparent and explicit manner, can be reproduced on different machines and, ultimately, can be considered as task- or domain-oriented ontologies. The proposed functions can be used for a range of purposes including: exploratory analyses, pattern retrieval, formalization of expert knowledge into ontology-like scripts, evaluation of automated classification methods, labelling of natural processes and integration of spatial analyses. It has a unique way of evaluating heuristic rules and it is able to bridge raster and object representations of geographic spaces. Ultimately, it represents an additional resource to perform spatial analysis within the R-environment.

Chapter 5
Distributional patterns and environmental drivers of hard-substrate benthic species from the deep sea of the Azores (NE Atlantic): a synthesis

5.1 Abstract

The deep seabed beyond continental shelf depths (≈200 m) is one of the largest biome on Earth covering an area of over 400 million km^2 and it is currently threatened by multiple stressors such as climate change, resource extraction and pollution. These threats can be particularly severe for deep-sea corals and sponges (Phyla Cnidaria and Porifera), hard-substrate ecosystem engineer species generally regarded as key components of vulnerable marine ecosystems (VMEs). This study presents a synthesis of the distributional patterns and of the environmental drivers of deep-sea hard-substrate communities in the Azores (NE Atlantic) with a special focus on ecosystem engineer species of the Phyla Cnidaria and Porifera. In a first part, it assesses the regional gradient in primary production investigating if the existing literature supports the existence of two oceanic provinces in the Azores. In a second part, a new compilation of bathymetric data is used to identify the location of all the major geomorphic features (e.g. seamounts, ridges, slopes) present in the region and to estimate the total fraction of hard substrate on such features. Geomorphic features are extracted from the topographic surface as discrete spatial entities each with identity, spatial embedding and attributes. The representation of geomorphic features as discrete entities provides an object-oriented perspective that can be of help in communicating scientific results, designing local scale studies based on seascape ecology principles and in synthetizing current knowledge. Finally, all the major environmental drivers shaping hard-substrate benthic communities' distributions are combined to produce in an ecological zonation of the seafloor of the Azores. Four major environmental drivers of deep-sea species distributions are recognized: (*i*) a latitudinal gradient in primary production strongly influenced by the Azores Current-Azores Front (AzC-AzF) system; (*ii*) the depth-wise succession of the regional water masses and their stratification into different isopycnal (vertical) layers; (*iii*) the spatial distribution of prominent geomorphic features such as seamounts ridges and island slopes; (*iv*) the availability of hard substrate for attachment.

5.2 Introduction

The deep seabed beyond continental shelf depths (≈200 m) is one of the largest biome on Earth covering an area of over 400 million km^2 (Danovaro et al., 2014). Its remoteness and immensity has both fascinated and puzzled many ecologists whose attempts to describe natural processes and patterns have been frustrated by a chronic scarcity of data (Howell et al., 2020; McClain and Hardy, 2010; Ramirez-Llodra et al., 2010). Over the last decade, technological advances in seafloor mapping and biological sampling have provided a better understanding of the importance of deep-sea biodiversity and of the ecological services it provides (Danovaro et al., 2017, 2014; Thurber et al., 2014). There is now a great awareness that the deep sea is under multiple stressors including climate change, resource extraction and pollution (Halpern et al., 2019; Levin et al., 2019). These threats can be particularly severe for deep-sea corals and sponges (Phyla Cnidaria and Porifera), hard-substrate ecosystem engineer species generally regarded as key components of vulnerable marine ecosystems (VMEs) (FAO, 2009; Hogg et al., 2010; OSPAR, 2010; Ragnarsson et al., 2017).

Multiple international organizations have proposed a number of policies to reverse the cycle of decline in ocean health. The United Nations General Assembly proclaimed the Decade of Ocean Science for Sustainable Development (2021-2030) (United Nations, 2018). The European Commission has adopted a number of policies to grant a sustainable use of nature space and resources. These policies include the Maritime Spatial Planning Directive (EU Directive 2014/89/EU) and the Marine Strategy Framework Directive (EU Directive 2008/56/EC). In general, these policies call on the scientific community to conduct research that advances understanding of marine ecosystems facilitating in this way the adoption of area-based/ecosystem-based management approaches (Howell et al., 2020; Kirkfeldt et al., 2022). Although the exploration of the deep sea now proceeds at a fast pace (e.g. Dominguez-Carrió et al., 2021; Swanborn et al., 2022; Wölfl et al., 2019), the data about deep-sea communities still cover a limited portion of the seafloor constraining adequate ecosystem-based management approaches (Folkersen et al., 2019; Howell et al., 2020; Townsend et al., 2018). One of the strategy to partially overcome this issue and to promote and direct new collections of samples is the use of ecological maps. Maps represent a powerful tool to synthetize and communicate information about a geographic space (Coetzee et al., 2021). In the context of marine ecology, they are widely used to capture the environmental diversity and the distribution of species and communities (Buhl-Mortensen et al., 2015; Hogg et al., 2016; Howell, 2010; Ismail et al., 2015; Longhurst, 2007; Sayre et al., 2017; Sonnewald et al., 2020; Watling et al., 2013; Woolley et al., 2020). In general, ecological maps can be defined as a visual model (simplification) of complex ecological

phenomena. They rely heavily on physical, spatial, or biological surrogates to describe biodiversity assuming that the biophysical patterns and processes observed for certain taxonomic groups or within particular environmental structures have a general relevance for ecological systems (Brown et al., 2011). The use of continuous coverage data layers is essential as they provide a mean to interpret patchy in situ observations, correlate the presence or the intensity of ecological phenomena with certain conditions and project such interpretations over large geographic areas (Brown et al., 2011).

The present work has the objective to visually synthetize and map what is currently known about the distributional patterns of hard-substrate benthic communities found in the deep sea of the Azores (NE Atlantic). The Archipelago of the Azores consists of nine volcanic islands located close to the Mid-Atlantic Ridge at the junction of the Eurasian, African, and North American plates (Miranda et al., 2018; Santos et al., 1995). This region is characterized by a huge area of deep seafloor scattered across a complex array of plains, ridges, seamounts and slopes (Mitchell et al., 2018; Peran et al., 2016; Tempera et al., 2012a). This work focuses on large deep-sea habitat-forming suspension-feeders (Phyla Cnidaria and Porifera) characteristic of vulnerable marine ecosystems (VMEs) and builds on the many scientific surveys that have explored the deep sea of the Azores in recent years (e.g. Dominguez-Carrió et al., 2021; Morato et al., 2021a, 2020b, 2019, 2018b; Morato and Taranto, 2019; Orejas et al., 2017). These surveys set the bases for a comprehensive characterization of the spatial distribution of deep-sea species (e.g. Morato et al., 2021b) and provided valuable insights into the environmental drivers that shape hard-substrate communities (Chapters 2 and 3). Although the analysis of the collected data is still ongoing, several interesting patterns are emerging. The first and more obvious is the association of habitat-forming corals and sponges with seamounts, ridges and island slopes. These key topographical features present large fractions of exposed bedrock and complex interactions with steady or variable flows, a setting that provides favorable conditions for thriving benthic communities dominated by suspension feeders (Carvalho et al., 2020; Morato et al., 2021b; Rogers, 2018). The second pattern is that the depth-wise stratification of the regional water masses explains much of the changes in species composition occurring with depth (Chapters 2 and 3). The third pattern is somehow less clear than the previous two and additional research is required in order to confirm it, but it appears that the distribution of benthic deep-sea species can be influenced by the regional latitudinal gradient in primary production (Chapters 2 and 3).

The proposed work is composed of two parts. The first part addresses the latitudinal gradient in primary production investigating if previous publications support the idea that there are two distinct ecological

provinces in the Azores. The second part uses a new compilation of bathymetric data gridded at 500 m resolution (Rodrigues et al., unpublished data) to identify the location of all the raised features present in the Azores (e.g. seamounts, ridges, slopes). In doing so, it is possible to estimate the habitat available for hard-substrate communities and allocate it across the different depth zones that capture the vertical zonation of the regional deep biota. The elevated features are identified and segmented into discrete units using the newly developed software *scapesClassification* (Taranto, 2022) (Chapter 4).

5.3 Materials and Methods

5.3.1 Ecological provinces

The existence of different ecological provinces in the Azores is investigated with a literature review. In a first part, articles providing information about *in situ* measurements of planktonic communities and production regimes are assessed. In a second part, articles providing global classification schemes for oceanic waters are investigated and set in the context of the Azores.

5.3.2 Geomorphological classification and input data

This study uses a new compilation of bathymetric data gridded at 500 m resolution (Rodrigues et al., unpublished data) (Figure 5.1) to perform a geomorphic characterization of the deep seafloor in the Azores. The updated bathymetry grid consists of a mosaic of data having different quality and native resolution that integrates publicly available datasets (EMODnet Bathymetry Consortium, 2018; GEBCO Compilation Group, 2022) with recently collected high-resolution multibeam data (e.g. Morato et al., 2018b). It represents the best bathymetry data currently available for the region. Principles of geomorphology and geomorphometry are commonly used to classify the seafloor based on bathymetry data (Brown et al., 2011; Lecours et al., 2016; Swanborn et al., 2022). In fact, shape, complexity and depth of geomorphic features have a strong influence on the distribution of deep-sea communities both at local and regional scales (e.g. Corrêa et al., 2022; Durden et al., 2020; Pierdomenico et al., 2016; Wienberg et al., 2013). In a preliminary phase, geomorphological classifications were performed adopting two traditional approaches: Benthic Terrain Modeler (BTM) tool (ArcGIS 10.x) (Walbridge et al., 2018) and geomorphons (GRASS GIS) (Jasiewicz and Stepinski, 2013). The BTM is a GIS classification approach, which uses a combination of derived terrain attributes and a set of user defined rules to segment the seafloor into geomorphic classes. The geomorphon approach uses a multi-scale pattern recognition approach to classification and mapping of landforms. In short, the class of a particular raster cell is automatically

determined based on its relative elevation along the eight principal compass directions (see Jasiewicz and Stepinski, 2013).

The outputs of such classifications did not seem fit for the purpose of the present manuscript. Especially for the BTM, the classifications appeared to be overly complicated for a synthesis work when traditional rulesets were applied (e.g. Jerosch et al., 2016). In addition, for both methods, it was difficult to determine the biological relevance of all the outputted classes which is a crucial step in classification works of this kind (Howell, 2010). Therefore, it was decided to opt for a different approach producing a more parsimonious output based on simple and ecologically meaningful rules. Recognizing the importance of raised geomorphic features (e.g. seamounts, ridges and island slopes) for the existence of habitat-forming benthic organisms, one of the classification steps was to extract such features from the bathymetric grid. This step was performed in the R environment (version 4.2.1) (R Core Team, 2022) using the R package *scapesClassification* (Taranto, 2022) (Chapter 4). Bathymetry and four terrain indices derived from bathymetry were used in this step: seafloor slope, broad-scale and fine-scale BPI (Bathymetric Position Index) and the standard deviation of depth values computed in a 9 X 9 moving window (Figure 5.1). Slope values are derived for each cell as the maximum rate of change from the cell to its neighbor (Lundblad et al., 2006). Slope values provide a proxy measure about the stability of sediments and the local acceleration of currents (Lecours et al., 2016). BPI values are computed as the elevation difference between a focal cell and the mean elevation of the surrounding cells within an inner and an outer radius (Lundblad et al., 2006). Radii of different sizes are used to detect fine or broad benthic features. In general, a combination of broad-scale and fine-scale BPI is used in geomorphological classifications (Lundblad et al., 2006; Walbridge et al., 2018). Simplifying, positive BPI values identify ridges, negative values identify valleys and near-zero values identify flat areas (Lecours et al., 2016; Lundblad et al., 2006). However, near-zero or negative BPI values do not necessary mark the boundary of a raised feature. In regions of steep slopes and, thus, of large bathymetric variation, upslope BPI values are positive, while downslope BPI values tend to become negative moving toward the seafloor (Lundblad et al., 2006). In order to overcome this issue, the bathymetric variation in a 9 X 9 neighborhood (measured as the standard deviation of depth values) was considered. Seafloor slope, broad-scale and fine-scale BPI were computed with the BTM in ArcGIS 10.5. Broad-scale BPI was computed with an inner and an outer radius of about 1 and 20 km, respectively. Fine-scale BPI was computed with an inner and an outer radius of about 0.5 and 5 km, respectively. These radii were selected considering that the study area spanned over a distance of more than 1000 km both in the latitudinal and in the longitudinal directions. The standard deviation of depth values was computed using the raster calculator of the R package 'terra' (Hijmans, 2022).

The R package *scapesClassification* implements raster classifications in a multi-step process (Chapter 4 and GitHub documentation). It allows to classify and segment a raster surface based both on user-defined heuristic rules and on the topological relationships expected to exist within and between classes and objects. The functions and algorithms it offers were designed for the recovery of particular patterns form n-dimensional raster surfaces following the innate cognitive process people use to schematize in their mind spatial configurations (Chapter 4). In the context of the present study, we can imagine that island shelf (IS) areas are located next to an island landmass and are characterized by a gentle slope down to the shelf break. Similarly, a raised geomorphic feature (RF) can be conceptualized by firstly identifying its salient feature (the presence of one or more summit areas), and then by identifying all the slopes and plateaus that connect the summit area to the seafloor. In general, the classification functions of *scapesClassification* perform focal evaluations (i.e. heuristic rules are only evaluated at suitable locations) and take into account class continuity (i.e. the classification process continues as long as a set of rules is respected by contiguous cells) (Chapter 4 and GitHub documentation). For instance, the existence of a flat area, *per se*, does not qualify an IS cell. However, a flat area next to an island landmass does qualify an IS cell. Unclassified cells adjacent to IS cells that present gentle slopes can also be considered as IS (based on class contiguity). Another feature of *scapesClassification* is that it allows to identify distinct raster objects (i.e. groups of raster cells identified by a unique ID) based on the discontinuity of certain properties or using segmentation algorithms. Each raster object can then be handled as an independent feature and can be modified using heuristic rules (see the section 'Homogeneous growth' in the documentation of the function 'cond.4.nofn'). The software *scapesClassification* was used to identify RFs and to segment them into discrete geomorphic units (raster objects) each with identity, spatial embedding, and attributes. The use of discrete units allows for an alternative conceptualization of the study area moving the focus from the field (raster) to the entity (object) (Bishop et al., 2012; Cova and Goodchild, 2002).

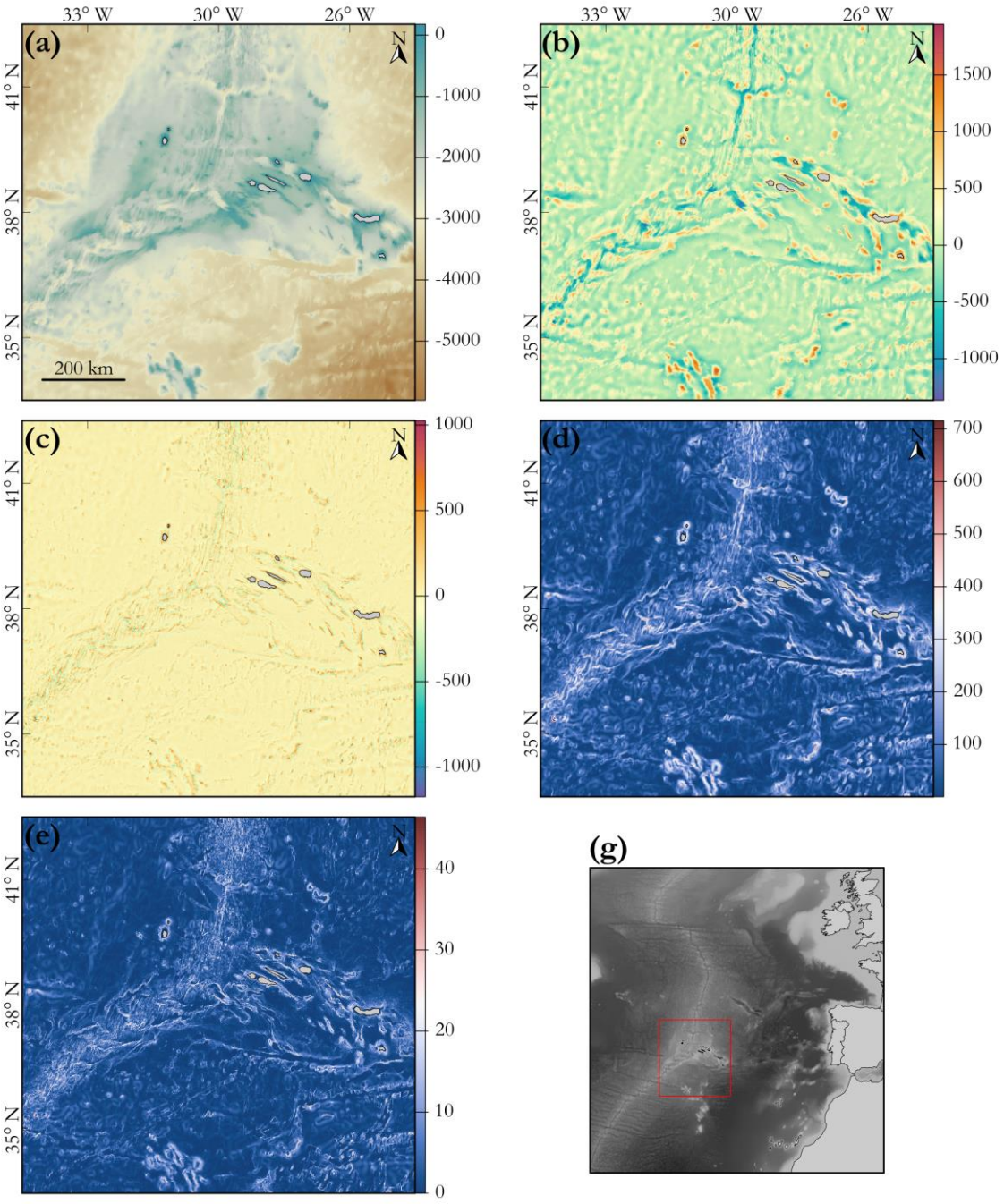

Figure 5.1 – Study area and input layers. Bathymetry (a), coarse-scale BPI (b), fine-scale BPI (c), standard deviation of bathymetric values (d), slope (e), location of the study area in the NE Atlantic (g). BPI is bathymetric position index.

5.3.3 Ecological zonation of the seafloor of the Azores

The information derived from the geomorphological classification and from the literature review are combined to produce an ecological zonation of the seafloor of the Azores mirroring the distributional patterns of regional hard-substrate communities. Besides geomorphology and productivity, two additional patterns are considered. The first is the regional depth zonation of benthic suspension feeders (Braga-Henriques et al., 2013; Tempera et al., 2012b). In particular, the distribution of deep-sea corals and sponges is influenced by the depth-wise stratification of different water masses with the main changes in species composition apparently occurring at depths of 500, 800 and 1600 m (Chapter 3). Watling and Lapointe (2022), using distributional data of anthozoans (Phylum Cnidaria), identified in the North Atlantic an additional fauna transition at 2200 m. These four depth zones are used to classify each raised feature into a different depth class. Finally, the likelihood of having patches of hard substrate is considered in the ecological zonation. Unfortunately, the lack of a continuous layer of backscatter data does not allow for a direct estimate of surficial sediment types (Brown et al., 2011). Instead, it is assumed that steep areas and/or areas on local ridges are the most likely to present patches of hard substrate. Based on the modelled distributions of benthic suspension feeders it appears that, at scales and resolutions similar to the ones considered here, these species avoid areas presenting slope values below five degrees (Chapters 2 and 3). Therefore, raster cells with slope values below five degrees are considered as 'flat' and unsuitable for deep-sea corals and sponges growing on rocky substrates. The only exception to this rule are local ridges, i.e. areas rising abruptly from the surrounding seafloor that are more likely to present rocky substrates at the summit. This type of ridges can be identified by positive fine-scale BPI values (fine-scale BPI > 100) (e.g. Jerosch et al., 2016; Lundblad et al., 2006). The use of slope and of proxy measures of bathymetric variation to identify patches of hard substrate represents a form of tacit knowledge (Fazey et al., 2005). In fact, these proxy measures are implicitly used in the planning of video surveys targeting hard substrate communities and in general they tend to be quite successful (i.e. hard substrate patches are indeed found) (e.g. Morato et al., 2021a).

5.4 Results

5.4.1 Ecological provinces

In the Azores, the oceanographic circulation is dominated by eastward flowing branches of the Gulf Stream: Gulf Stream and Gulf Stream Bifurcations in the north and the Azores Current (AzC) in the south (Caldeira and Reis, 2017; Frazão et al., 2022). The AzC seems to be particularly important in shaping

planktonic and microbial communities and in delimiting different biogeochemical provinces (Baltar and Arístegui, 2017; Frazão et al., 2022; Fründt and Waniek, 2012; Riou et al., 2016). It is part of the North Atlantic Subtropical Gyre, originates near the Grand Banks (40°N, 45°W) and flows south-eastward until it crosses the MAR at approximately 34°N, 37°W (Comas-Rodríguez et al., 2011; Frazão et al., 2022). On average, its core is located between 33°N and 36°N east of Mid Atlantic Ridge (Frazão et al., 2022). The AzC is associated with a thermohaline front – the Azores Front (AzF) – and with two westward countercurrents. The northern countercurrent is known as the Azores Countercurrent (AzCC) and is located at about 35.25°–36.25°N (Comas-Rodríguez et al., 2011; Frazão et al., 2022). The Azores Current-Azores Front (AzC-AzF) system separates the warmer and saltier 18 ºC Western North Atlantic Water originating from the Sargasso Sea to the south from the colder Eastern North Atlantic Water to the north (Frazão et al., 2022; Fründt and Waniek, 2012). This system is extremely dynamic and the position of the AzF shows strong seasonal and interannual variability (Frazão et al., 2022; Fründt et al., 2013; Fründt and Waniek, 2012). Comas-Rodríguez et al. (2011) showed that the AzC and the AzCC could be particular relevant for deep sea organisms as they both reach intermediate water depths (down to ∼2000 m). Interestingly, the AzC-AzF system showed a progressive northward migration starting in the mid-1970s and continuing until 2005 possibly causing an expansion of the oligotrophic areas of the main gyres owed to global warming (Frazão et al., 2022; Fründt and Waniek, 2012; Yang et al., 2020). Several video observations show extensive seafloor areas covered by dead stony corals (Order Scleractinia) at different locations south of the Kurchatov Fracture Zone (~40 ºN) (unpublished data). These species are strongly influenced by food supply controlled by export production and turbulent hydrodynamics at the seabed (Portilho-Ramos et al., 2022). In the future, it would be interesting to assess if the high mortality of scleractinian corals and the decrease in primary production at surface are correlated.

The AzCC appears to be of particular importance as it locally enhances subsurface productivity (Caldeira and Reis, 2017; Fründt and Waniek, 2012; Riou et al., 2016; Schiebel et al., 2011). It is associated with westward propagating eddies and with oceanic filaments that play an important role in sustaining high productivity at depths between 45 and 200 m (Caldeira and Reis, 2017; Riou et al., 2016; Silva et al., 2013). Schiebel et al. (2011) assessed the production and dispersion of coccolithophores in a N–S transect from 47°N to 33°N at 20° W. Coccolithophores were most abundant at the northernmost (45°N–47°N) and southernmost (33°N–37°N) sites while they were least abundant in the middle part of the transect (39°N–43°N). Across all sites *Emiliania huxleyi* was the most common species. However, the coccolithophore assemblages showed significant differences at the extremes of the transect. At the northern end, the second most abundant species was *Gephyrocapsa muellerae* a species that inhabits rather cool waters

under well mixed conditions (Schiebel et al., 2011). At the southern end of the transect, along with *E. huxleyi*, the most abundant species were *Gephyrocapsa ericsonii* and *Florisphaera profunda*. These species are indicative of Subtropical Mode Waters (Schiebel et al., 2011). Interestingly, Schiebel et al. (2011) noticed that *F. profunda* appears to occur at the intersection of decreasing light availability and increasing NO_3 concentration. Riou et al. (2016) found that in the region of the AzCC, where the densities of *F. profunda* are the highest, the N_2-fixation is particularly intense. Importantly, most of the productivity of coccolithophore observed by Schiebel et al. (2011) occurs in subsurface waters down to depths of 150 m. These authors noticed that this apparent bimodal pattern in productivity would have been hard to detect by satellite measurements as the coccolithophore assemblages were dispersed over expanded surface water layers with no significant signal at surface.

In general, the AzC-AzF system seems to delimit distinct oceanic provinces. Its influence extends below surface waters and it appears to have a strong impact on deep ocean particle flux (Fründt and Waniek, 2012). This system has an important effect on the regional latitudinal gradient in productivity. North of the Azores are located nutrient enriched waters sustaining high level of productivity. Moving southward there is a general decrease in primary production. However, at the AzF in the area of the AzCC, primary production seems to locally increase before decreasing again moving further south. Baltar and Arístegui (2017) directly suggested that the AzC-AzF system determines two different ecological regions as it causes sharp differences in the abundance of autotrophic and heterotrophic microbial assemblages with an abrupt decline in the abundance of picophytoplankton south of the front. Changes in microbial assemblages and enzymatic activities extends down to the bathypelagic zone (Baltar and Arístegui, 2017).

The most known bioregionalization of the ocean is probably represented by the biogeochemical provinces of Longhurst (2007). This author assigned all the Azores region to the Northeast Atlantic subtropical gyral. However, Reygondeau et al. (2013) showed that the Longhurst biogeochemical provinces follow a strong seasonal and interannual dynamic. In particular, during summer months the Azores seems to be located at the interface of different provinces and this boundary roughly follows the path of the AzC (Reygondeau et al., 2013). Sonnewald et al. (2020) identified distinct ecological provinces based on a complex unsupervised learning method. Their classification considered the outputs of a numerical model defining planktonic assemblages (Dutkiewicz et al., 2015) and nutrient flux data. The final output was a nested set of eco-provinces. Again, at all nested levels, the Azores were located at the boundaries of distinct provinces and the boundary is approximatively at the latitudes of the AzC-AzF systems. Bock et al. (2022) produced a biogeographical classification of the Global Ocean using Biogeochemical Argo Floats. These

devices sample the water column down to a depth of 1000 m and were used by the authors to identify distinct phytoplankton biomes. Their study suggests that the southern portion of the Azores could be characterized by a seasonal deep chlorophyll maximum biome. This biome is characterized by strong seasonality and shows a bimodal distribution centered at 36.01° ± 3.20°N and 34.98° ± 8.10°S and is associated with 18 ºC Subtropical Mode Waters. At 36º N is approximatively located the AzCC. Finally, Sayre et al. (2017) produced a three-dimensional classification of the ocean based on an aspatial clustering of all the variables from the 2013 World Ocean Atlas. They divided the ocean into 37 ecological marine units (EMUs). When considering surface waters, their classification placed the Azores within a single province. However, moving at a depth of 100 m their classification seemed to follow the path of the AzC and divided the Azores region into two distinct provinces. It was decided to use this classification as an approximation of the different biogeochemical provinces existing in the Azores.

5.4.2 Geomorphic units

The geomorphic units of interests – raised features (RFs) – are identified in multiple phases, each characterized by several steps. All the classification steps are provided in Appendix D.1 in the form of an R script.

In the first phase island shelf (IS) cells are identified considering that these cells are adjacent to island landmasses and that they are relatively flat. Based on the modelled distributions of benthic suspension feeders it appears that, at scales and resolutions similar to the ones considered here, these species avoid areas presenting slope values below five degrees (Chapters 2 and 3). Therefore, the flat areas of the IS are determined using a slope threshold of five degrees (Figure 5.2a). All areas classified as IS are excluded from subsequent classifications.

In the second phase, peak cells are identified using the function ‘peak.cell’. This function identifies local maxima, i.e., raster cells whose elevation value is larger than the values of all the cells in its 3 X 3 neighborhood. Peak cells identified within the IS, at locations adjacent to the IS or on small relieves (fine-scale BPI < 100 or broad-scale BPI < 100) are removed.

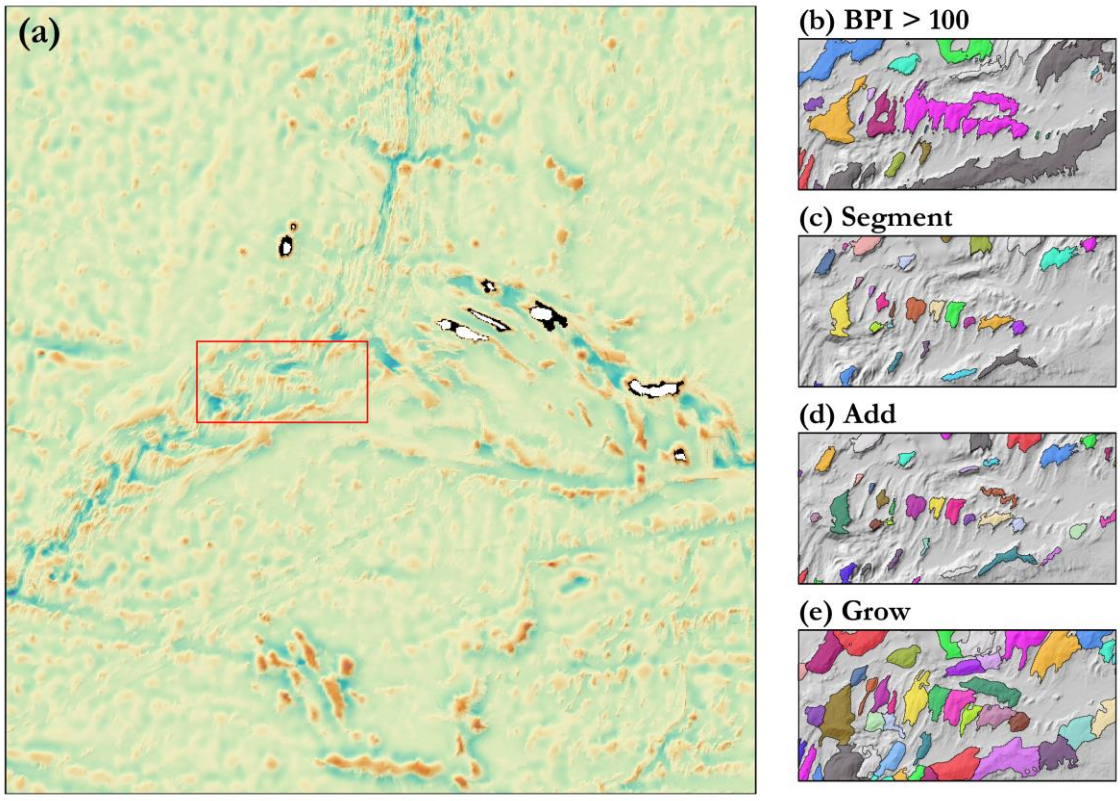

Figure 5.2 – Identification of raised features (RFs). Panel (a) show broad-scale BPI values in the background, island shelf (IS) cells in black and the zoom in area in red. Fist classification step (b): RFs are identified as discrete groups of cells connected to at least one peak cell having broad-scale BPI values > 100. Second classification step (c): RFs are segmented at negative values of relative position index (RPI). Values of RPI are computed independently for each RF using the formula RPI = [D – mean(D)] / sd(D), D is depth, mean(D) is the average depth of the RF, sd(D) is the standard deviation of depth of the RF. Third classification step (d): new discrete groups of cells having broad-scale BPI > 300 or > 400 are classified as RF. Fourth classification step (e): raster cells having broad-scale BPI > 100 or fine-scale BPI > 100 or standard deviation of depth values > 150 are aggregated to the closest RF. BPI is bathymetric position index. Different colors in panels (b-e) identify different RFs. Note that small features (less than 50 raster cells) were excluded from the analysis.

The identification of discrete RFs proceeds in several steps. In the first step groups of cells connected to at least one peak cell (class continuity) with broad-scale BPI > 100 are classified as RFs using the function 'anchor.seed'. Each discrete group of cells receives a different ID and, in this way, can be treated as an independent raster object (Figure 5.2b). This classification step identifies some very large RFs. Therefore, the initial set of RFs is modified in three subsequent steps. First, they are segmented using the functions 'rel.pi' and 'pi.sgm'. The function 'rel.pi' computes the relative position index (RPI) of each cell within a RF based on RPI = [D – mean(D)] / sd(D) [D is the depth of a focal cell, mean(D) is the average depth of the

RF and sd(D) is the standard deviation of depth values within RF]. The function 'pi.sgm' is used to (i) remove from RFs raster cells presenting a negative RPI and (ii) to reassign a different ID to every discrete group of cells counting at least 50 cells (Figure 5.2c). Then, the function 'pi.add' is used twice, first to add features counting at least 50 cells with broad-scale BPI > 300 and then to add features having at least 50 cells with broad-scale BPI > 400. The RFs added in the last steps are the features removed during the segmentation step. However, setting higher BPI threshold values it is possible to detect more groups of unconnected cells each receiving a different ID (Figure 5.2d). Note that in this context connected refers to contiguous cells sharing similar properties (e.g. BPI above a certain threshold); unconnected refers to two or more groups of contiguous cells sharing a particular property (e.g. BPI above a certain threshold) separated by cells not respecting a particular property (e.g. BPI below a certain threshold). Finally, the function 'cond.4.nofn' is used to manipulate RFs. Setting the argument 'hgrowth = TRUE', the function iteratively selects a RF, evaluates a set of conditions for raster cells adjacent to the selected RF (i.e. standard deviation of depth > 150 OR broad-scale BPI > 100 OR fine-scale BPI > 100) and aggregates all cells respecting the conditions to the RF. Then, the function selects a new RF and the process is repeated until all features are evaluated. The process restarts as long as at least one cell is aggregated to a RF. By evaluating sets of conditions only at locations adjacent to a RF in an iterative process, each RF grows homogeneously ('hgrowth'). At the end of the evaluation, all raster cells respecting the set of conditions are aggregated to the closest RF (Figure 5.2e). All the mentioned functions are part of the R Package *scapesClassification* (Taranto, 2022) (Chapter 4). The final set of RFs is shown in Figure 5.3a.

Once identified RFs can be treated as independent objects (Figure 5.3b-d). This can be useful to compare particular spatial patterns within and among distinct features. For example, the three features in Figure 5.3b-d have different areas and shapes. The availability of hard-substrate was approximated using slope and BPI values. Even using this simple approximation, hard-substrate areas seem to have a very different configuration within the different features. Monte Alto is dominated by a flat plateau and most of the areas suitable for sessile hard-substrate suspension feeders are located on the eastern side of the seamount (Figure 5.3b). Gigante 127 is dominated by ridge-like structures (Figure 5.3c). Atlantis presents a flat plateau but it hosts wider patches of hard substrate than Monte Alto (Figure 5.3d).

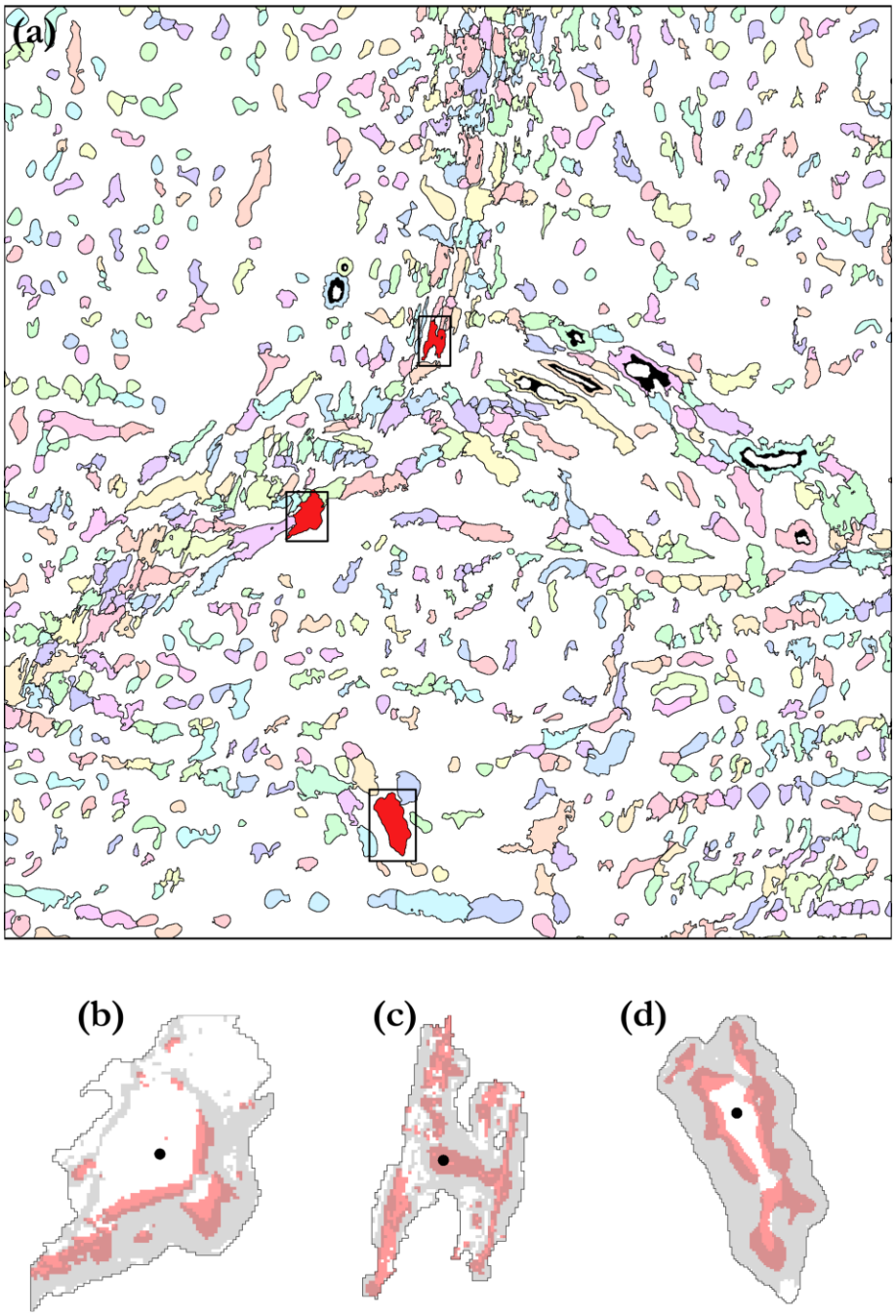

Figure 5.3 – Raised features (RFs) of the Azores. Each RF is identified by a unique ID and can be treated as an independent raster object (a). Panels (b-d) highlight three distinct features: Monte Alto (b), Gigante 127 (c) and Atlantis (d). These features are shown in red in panel (a). In panel (a) colors are randomly assigned to individual RFs. In panels (b-d) grey shades identify slopes (> 5º), red shades identify local ridges (fine-scale BPI > 100) and black dots identify the summit of the RFs.

Table 5.1 – Statistics of three raised features (RFs): Monte Alto (MA), Gigante 127 (G127), Atlantis (AT). MinD is minimum depth (m). Area is the total area of the feature (km^2). Perim. is the perimeter of the feature (km). SA is average roughness of a surface. SDR is surface area ratio. RDG is the area covered by local ridges (fine-scale BPI > 100) (km^2). SL is the area covered by slopes (> 5º) (km^2). The remaining columns report the estimated availability of hard substrate (RDG + SL) in km^2 within distinct depth zones: deeper than 2400 m (DZ 2400L), 2400 – 1600 m (DZ 2400), 1600 – 800 m (DZ 1600), 800 – 500 m (DZ 800) and shallower than 500 m (DZ 500). Numbers in parenthesis are percentages of the total area.

RF	MinD	Area	Perim.	SA	SDR	RDG	SL	DZ 2400L	DZ 2400	DZ 1600	DZ 800	DZ 500
MA	385	1068	212	368.9	0.07	176 (16.5)	367 (34.4)	0 (0)	156 (14.7)	297 (27.9)	87 (8.2)	1 (0.1)
G127	228	597	231	173.5	0.20	220 (37)	262 (43.9)	0 (0)	6 (1)	423 (70.9)	40 (6.7)	13 (2.3)
AT	266	1581	235	623.6	0.07	507 (32.1)	879 (55.6)	143 (9.1)	482 (30.5)	507 (32.1)	131 (8.3)	121 (7.7)

Furthermore, each feature can be defined using metrics characteristics of seascape and landscape ecology. Traditional patch metrics can be used to capture the spatial configuration of discrete patches of hard substrate (Frazier and Kedron, 2017). Surface metrics represent an alternative set of metrics used to address spatial configurations on continuous surfaces (McGarigal et al., 2009). The package R package *geodiv* implements surface metrics in the R environment (Smith et al., 2021). As an example, Table 5.1 presents some of the statistics of three different RFs (Figure 5.3b-d). It includes the surface metrics average roughness (SA) and surface area ratio (SDR) ('sa' and 'sdr' functions, R package *geodiv*) and presents the availability of hard substrate within different depth zones. Average roughness is the average absolute deviation of the surface heights from the mean (McGarigal et al., 2009). In this context it provides a measure of the bathymetric span of a RF. Surface area ratio is the ratio between the surface area to the area of the flat plane with the same x-y dimensions (McGarigal et al., 2009). In this context it can be intended as a measure of the complexity of a RF's surface. Gigante 127 is the smallest seamount covering an area of 597 km^2, has the smallest bathymetric span (SA ~173) and the greatest surface complexity (SDR = 0.2). Atlantis is the largest seamount (1581 km^2) and has the largest bathymetric span (SA ~623). Compared to the other two features, Monte Alto has intermediate values (area ~1068 km^2; SA ~368). Surface complexity is low for both Monte Alto and Atlantis (SDR = 0.07). The estimated percentage of hard substrates is ~51% for Monte Alto, ~81% for Gigante 127 and ~88 for Atlantis. Note that Atlantis is located in an area where bathymetry data have low quality and, as a consequence, all measurement in this area are more likely to bear errors. Gigante 127 and Monte Alto have a very small percentage of hard substrate at depths shallower than 500 m and none below 2400. For all three RFs, most of the hard substrate is

located between 1600 and 800 m depth. Probably this would be the most appropriate depth range to investigate how faunal assemblages change across these three RFs.

5.4.3 Ecological Zonation

Figure 5.4 presents a synthesis of the expected distributional patterns of key habitat forming species (Phyla Cnidaria and Porifera) in the region of the Azores. Table 5.2 presents an estimate of the habitat available for these species. A total area of more than 1.5 million km^2 was considered in the present study. Of this area, the 11% (188,335 km^2) can potentially host hard substrate communities. This value provides an estimate of the availability of hard substrate on the major geomorphic relieves of the region. Hard-substrate locations are assigned to two ecological provinces: 6.8 % of the total area north of the Azores Front-Azores Current system (AzF-AzC), 4.4 % of the total area south of the AzF-AzC system. At depths shallower than 500 m there is a very limited amount of area suitable for hard-substrate communities south of the AzF-AzC (3,890 km^2, 0.2 % of the total area). These areas are associated with the Great Meteor Seamount Chain (Tucholke and Smoot, 1990). Depth zones below 500 m seem to host a similar percentage of hard substrate patches north and south of the AzF-AzC system.

Table 5.2 – Estimated availability of hard substrate on major geomorphic features allocated across distinct depth zones and ecological provinces. TA is total area of the study region. HS is the area expected to host hard substrate. The area expected to host hard substrate is divided into five depth zones and two provinces. The depth zones are: deeper than 2400 m (DZ 2400L), 2400 – 1600 m (DZ 2400), 1600 – 800 m (DZ 1600), 800 – 500 m (DZ 800) and shallower than 500 m (DZ 500). The two provinces are: areas located north of the Azores Front-Azores Current system (North AzF-AzC), areas located south of the Azores Front-Azores Current system (South AzF-AzC). All percentage values are relative to the total area of the study region (TA).

	TA		HS		DZ 2400L		DZ 2400		DZ 1600		DZ 800		DZ 500	
	km²	*%*	*km²*	*%*	*km²*	*%*	*km²*	*%*	*km²*	*%*	*km²*	*%*	*km²*	*%*
All	1,669,749	-	188,335	11.3	67,417	4.0	22,033	1.3	51,534	3.1	23,477	1.4	23,875	1.4
North AzF-AzC	989,844	59.3	112,838	6.8	34,180	2.1	12,559	0.8	33,979	2.0	13,826	0.8	18,295	1.1
South AzF-AzC	675,302	40.7	73,808	4.4	33,238	2.0	9,474	0.6	17,556	1.1	9,650	0.6	3,890	0.2

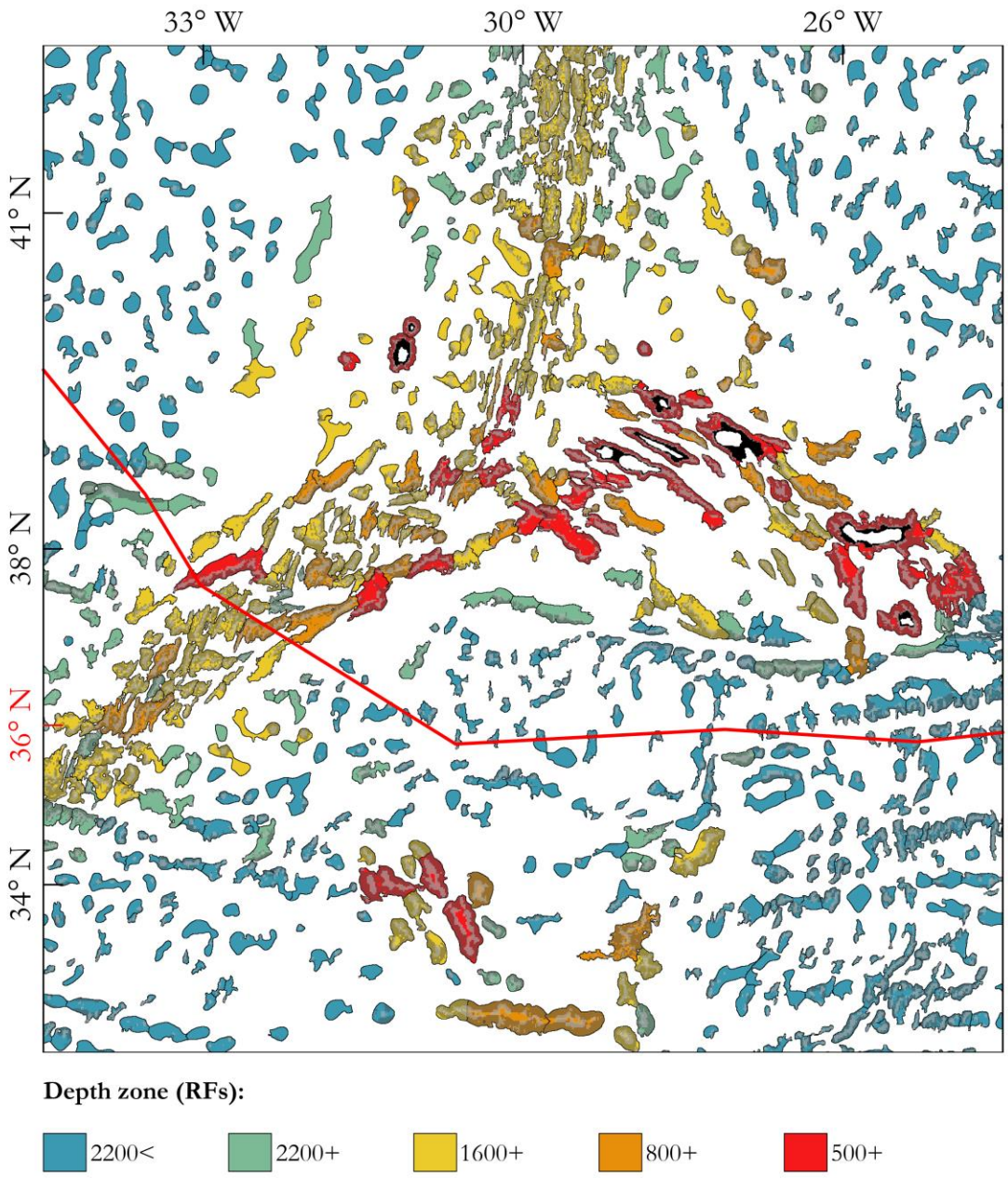

Figure 5.4 – Estimated distributional patterns of deep-sea benthic communities on hard substrates. All regional raised features (RFs) (e.g. seamounts, ridges, island slopes) are highlighted. Island shelf are showed in black. White shades are local ridges. Dark shades are slopes (> 5º). Each RF is assigned to a different depth zone: deeper than 2000 m (2200<), 2200 – 1600 m (2200+), 1600 – 800 m (1600+), 800 – 500 m (800+) and shallower than 500 m (500+). Each RF presents at least 2 km2 within the depth zone it is assigned to. The red line shows the boundaries of two ecological marine units (EMU11 in the north and EMU21 in the south) at a depth of 100 m (adapted from Sayre et al., 2017). This boundary approximates the path of the Azores Current-Azores Front system. The latitude marked in red (36 ºN) is the approximate latitude of the Azores Countercurrent west of the Mid-Atlantic Ridge.

5.5. Discussion

This study presents a synthesis of the distributional patterns and of the environmental drivers of deep-sea hard substrate benthic communities in the Azores (NE Atlantic) with a special focus on ecosystem engineers species of the Phyla Cnidaria and Porifera. It provides an estimate of the availability of hard-substrate habitats associated with prominent geomorphic features, identifies two distinct ecological regions likely to influence benthic deep-sea assemblages and provides and an ecological zonation of the Azores acknowledging the importance of the depth-wise stratification of the regional water masses.

The AzC-AzF system defines two distinct ecological provinces. South of the front are located 18 ºC Subtropical Mode Waters, while north of the front are the more productive temperate Eastern North Atlantic Water (Frazão et al., 2022; Fründt and Waniek, 2012). Immediately north of the front, in the region of the Azores Countercurrent (AzCC) (~36º N), dynamic processes such as oceanic filaments or eddies seem able to raise the nutricline to the euphotic zone and to mix different planktonic assemblages enhancing carbon sequestration and deep ocean particle flux (Caldeira and Reis, 2017; Fründt and Waniek, 2012; Riou et al., 2016; Schiebel et al., 2011; Silva et al., 2013). This complex system is highly dynamic with strong seasonal and interannual variability and its influence extends well below surface waters (Baltar and Arístegui, 2017; Comas-Rodríguez et al., 2011; Frazão et al., 2022; Fründt and Waniek, 2012; Reygondeau et al., 2013). Previous studies in the Azores either did not support the existence of a latitudinal turnover of species (Braga-Henriques et al., 2013; Menezes et al., 2006) or showed only preliminary evidence about the importance of the north-south gradient in productivity for benthic engineer species (Chapter 1 and 2). Latitudinal changes in benthic bathyal assemblages should appear more evident considering wider latitudinal ranges extending further north and south of the front. In a northern section of the Mid-Atlantic Ridge between the Azores and Iceland (42º – 53º N), Gebruk et al. (2010) suggested that the observed turnover in benthic bathyal fauna was caused by the Sub-Polar Front – located at ~52º N (Søiland et al., 2008) – and by differences in food supply to the seafloor. A later study in the same region shared similar conclusions (Alt et al., 2019). Particularly interesting is the article of Yearsley et al. (2020). These authors studied the connectivity patterns of fragmented hydrothermal vent habitats along the Mid-Atlantic Ridge. They used biophysical models built on Argo floats data and, remarkably, they identified persistent barriers to connectivity at about 36°N and 37°N. Although the authors did not acknowledge the existence of the AzC-AzF system at these latitudes, their results strongly suggest that the frontal system can act as a dispersal barrier for deep-sea species. Overall, several studies confirm that the AzC-AzF system is a permanent oceanic feature causing abrupt changes in oceanic conditions. Much of these changes occur

in subsurface and deep waters and are often overlooked in analysis based on remote sensing data. Waters north and south of the system can be considered as distinct ecological provinces which are likely to influence and structure deep-sea communities. Future studies based on quantitative analysis and abundance data will be required to clarify the role of the AzC-AzF system for benthic deep-sea assemblages. Placing a clear boundary for a highly dynamic system such as the AzC-AzF seems somehow arbitrary. The three-dimensional unsupervised mapping of Sayre et al. (2017) at a depth of 100 m seemed to capture quite well the average location of the AzC-AzF system and, therefore, was used to place a boundary in Figure 5.4. This boundary should be intended as reminder that complex oceanographic processes take place approximately at these locations. These processes should be taken into account if we want to better understand the distributional patterns of the deep biota in the Azores as well as the latitudinal turnover in benthic bathyal fauna occurring along the Mid-Atlantic Ridge.

Liu and Tanhua (2021) characterized and mapped the major water masses in the Atlantic Ocean showing that they are distributed in four main isopycnal (vertical) layers separated by surfaces of neutral density: (i) upper, (ii) intermediate, (iii) deep and overflow and (iv) bottom layer. In the Azores, there is a clear correlation between the vertical zonation of the regional benthic and demersal fauna and the depth-wise succession of the isopycnal layers (Chapter 2; Chapter 3; Braga-Henriques et al., 2013; Menezes et al., 2006; Santos et al., 2021), suggesting that the different isopycnal layers can be used as meaningful bioregions (Chapter 3). In general, the influence of water masses on the biodiversity and biogeography of deep-sea benthic ecosystems is widely recognized (e.g. Carney, 2005; Lapointe et al., 2020; Puerta et al., 2020; Roberts et al., 2021; Watling et al., 2013). At depths between 200 and 2000 m, the main changes in the composition of habitat-forming suspension feeders (Phyla Cnidaria and Porifera) occur at 500, 800 and 1600 m (Chapter 3). Watling and Lapointe (2022) identified in the North Atlantic an additional fauna transition at 2200 m for species of the Phylum Cnidaria. This depth zonation was integrated in Figure 5.4. Only few geomorphic features intersect the upper layer (0-500 m) and only a fraction of these features is likely to presents surfaces characterized by hard substrate (~23,000 km^2, 1.4% of the area considered in the present study) (Table 5.2). Therefore, areas of hard-substrate on shallow geomorphic features are likely to host regionally unique assemblages of cold-water corals and sponges, species representing key components of vulnerable marine ecosystems (VMEs) (FAO, 2009; Hogg et al., 2010). Considering the presence of unique and vulnerable benthic assemblages and the strong exposure to regional fisheries (Santos et al., 2021), shallow geomorphic features should be regarded as priority areas for conservation. In addition, these features are particularly indicated to investigate if and how fishery induced disturbance alter the density and composition of benthic habitat-forming species. Intermediate waters in the region

have their core at 800 m and extend down to an approximate depth of 1600 m (Bashmachnikov et al., 2015; Liu and Tanhua, 2021). Along the segment of Mid-Atlantic Ridge included in this study, geomorphic features within this depth range are the most common (Figure 5.4). Therefore, studies specifically targeting this depth zone could be particularly important to effectively disentangle bathymetric, latitudinal and local gradients of diversity. Similarly, studies addressing faunal assemblages occurring on geomorphic features below intermediate waters could be particularly useful in exploring compositional changes east and west of the Mid-Atlantic Ridge (Figure 5.4). So far, these very deep areas are the least explored in the region (Chapter 2 and 3).

The quantification of hard substrate availability and the geomorphological segmentation of regional raised features (i.e. seamounts, ridges and island slopes) was based on a simple approach and it has not the pretention to substitute more robust geomorphological analysis (Bishop et al., 2012; Lecours et al., 2016). The total estimate of hard substrate area is likely affected by the presence of bathymetric artifacts and by the simple methodology used. However, this estimate is based on the best data currently available and represents a valid approximation of the total habitat available for hard-substrate communities. The allocation of hard-substrate areas across depth zones and marine provinces provides a simplified model of the major environmental drivers acting on hard substrate communities (Figure 5.4). Based on this model, it is possible to frame new research designs, conservation strategies and sampling programs. Additionally, an approximate measure of the total habitat available for hard-substrate communities can be considered as a preliminary step required to quantify regional conservation targets (e.g. 30% of hard substrate areas) and a way to move forward national and international conservation goals (e.g. Jetz et al., 2022; Nicholson et al., 2019).

This analysis has the value of being simple and easily interpretable, it can be integrated by nested and more advanced classifications and it provides a different perspective that can be of use in evaluating local and regional scale distributional drivers. The fundamental scope of landscape ecology and of its sister science, seascape ecology, is to determine how spatial and temporal structures influence ecological processes and ecosystem services (Lepczyk et al., 2021; Pittman et al., 2021; Swanborn et al., 2022; Turner, 2005). They tend to represent nature with distinct pattern-oriented constructs such as patches, patch mosaics, and spatial gradients (Pittman et al., 2021). The segmentation of geomorphic features into discrete units – each with identity, spatial embedding, and attributes – can facilitate the implementation of seascape analysis in the deep-sea. The topography of the Azores is very complex (Figure 5.1). When represented as a continuous surface it is very hard to discern the effect of local structures (e.g.

composition and distribution of substrate patches) on ecological processes (e.g. community assembly). Transforming this surface into discrete spatial entities – each likely to host complex assemblages of benthic engineer species – can provide a more suitable framework to address pattern–process relationships at multiple scales. For example, Figure 5.3b-d and Table 5.1 can be used to compare different seamounts and evaluate how the local spatial configuration of hard-substrate patches interacts with regional gradients in productivity and how this interaction influences benthic communities. Surely, the example provided is quite simple and relies on poor estimates of surficial sediment types. Nevertheless, even this simple classification suggests that the total amount of habitat available for the settlement of hard-substrate species changes drastically both between seamounts and between depth zones and this differences should be taken into account in evaluating spatial patterns of species distribution and abundance. If integrated with backscatter and fine-scale data, for example in areas where these data are available, this type of analysis can produce interesting results (e.g. Williams et al., 2010; Zarnetske et al., 2019). In future greater efforts will be required to acquire multibeam bathymetry data in deep areas. Most of the geomorphic features at depths below 2000 m away from the Mid-Atlantic Ridge could be bathymetric artifacts caused by the lack of high quality multibeam data (Lecours et al., 2017). Excluding major ridges and prominent geomorphic features, the seafloor is generally assumed as homogeneous and sedimented. However, recent studies suggest that even abyssal plains could be rockier than expected (Riehl et al., 2020).

The proposed segmentation of raised feature into discrete objects should not be considered as alternative to the classic raster representation of topographic surfaces, but rather as a mean to integrate different conceptual perspectives (Cova and Goodchild, 2002). As discussed, an object perspective can aid in the implementation of patch-based analysis, but it can also be of use in synthetizing and communicating scientific results. In fact, humans clearly perceive the geographic space as populated by discrete objects, to which they give names and ascribe behaviors (Goodchild et al., 2007). By segmenting the seafloor surface, complex conservation outputs (e.g. Combes et al., 2021) could be better shared with policy-makers facilitating in this way the implementation of area-based conservation strategies. In addition, such segmentation could provide a way to organize and archive biological and environmental data. For instance, the interactive maps available on the GitHub page of the R Package '*scapesClassification*' (link) could be populated with information about number of video transects stratified by depth, fishing intensity, taxonomic diversity backscatter data availability, etc. Such a map could also be useful for public outreach if the explored features were to be associated with deep-sea footage. In fact, the use of video data and interactive tool seems to be particularly effective for public engagement and education

(Hoeberechts et al., 2015). Efforts aimed at exposing the public to the deep sea could increase awareness and support marine conservation and are particularly important in a region such as the Azores where the seascape is dominated by deep waters.

5.6 Conclusions

This study presents a synthesis of the distributional patterns and of the environmental drivers of deep-sea hard substrate benthic communities in the Azores (NE Atlantic) with a special focus on ecosystem engineers species of the Phyla Cnidaria and Porifera. Understanding how these important species distribute is of pivotal importance for effective area-based management and for the implementation of national and international policies aiming to reverse the cycle of decline in ocean health (e.g. Combes et al., 2021; Galparsoro et al., 2015; Gonçalves, 2021; Grehan et al., 2017; Kazanidis et al., 2020; Rice et al., 2012; Rueda et al., 2016; Tunnicliffe et al., 2020). Four major distributional drivers can be recognized in the region: (i) a latitudinal gradient in primary productivity strongly influenced by the Azores Current-Azores Front (AzC-AzF) system; (ii) the depth-wise succession of the regional water masses and their stratification into different isopycnal (vertical) layers; (iii) the spatial distribution of prominent geomorphic features such as seamounts ridges and island slopes; (iv) the availability of hard substrate for attachment. The results provide an estimate of the availability of hard-substrate habitats associated with prominent geomorphic features (i.e. seamounts, ridges and island slopes), identify two distinct ecological regions likely to influence benthic deep-sea assemblages and synthetize the distributional patterns in an ecological zonation of the Azores acknowledging the importance of the depth-wise stratification of the regional water masses. The total fraction of rocky substrates on prominent geomorphic features available for the settlement of sessile suspension feeders was estimated to cover an area of about 188,000 km^2 (~11% of the study area). Geomorphic features are extracted from an updated bathymetric grid of the Azores and treated as discrete spatial entities. This approach adopts an object-oriented perspective about the major geomorphic features existing in the region and, in doing so, it provides a more suitable framework to address pattern–process relationships following seascape and landscape ecology strategies. Overall, several studies confirm that the AzC-AzF system is a permanent oceanic feature causing abrupt changes in oceanic conditions. Much of these changes occur in subsurface and deep waters and are often overlooked in analysis based on remote sensing data. The conditions imposed by the AzC-AzF system are likely to influence distribution and abundance of benthic organisms in the Azores. The different frontal systems occurring along the Mid-Atlantic Ridge could be particularly important in explaining latitudinal changes in species composition even at the seafloor. Previous works seem to confirm the importance of

the Sub-Polar Front (~52° N). Future studies will have to clarify the biogeographic role of the AzC-AzF system (~36º N).

Chapter 6
General Discussion

The Azores are strategically located to investigate and explore the deep sea. Thanks to regional, national and international research programmes the exploration of the deep sea in the region gained new momentum (e.g. Dominguez-Carrió et al., 2021; Morato et al., 2021a, 2020b, 2019, 2018b; Morato and Taranto, 2019; Orejas et al., 2017). We now have a much better understanding of the different benthic habitats existing in the region and the many collected specimens will allow for a proper taxonomic description of the regional deep sea fauna. However, over the first years of my thesis work, I felt that there was very limited information to actually provide a useful evaluation of the environmental status of deep sea biodiversity in the region. I believe it will be possible to provide a scientifically sound evaluation of the environmental status of deep sea biodiversity in the Azores once all the collected data will be processed. Therefore, much of the efforts of this thesis were placed in describing the distributional patterns of species representing an important component of regional benthic communities, habitat-forming suspension feeders of the Phyla Cnidaria and Porifera. Understanding how these important species distribute is of pivotal importance for effective area-based management and for the implementation of national and international policies aiming to reverse the cycle of decline in ocean health (e.g. Combes et al., 2021; Galparsoro et al., 2015; Gonçalves, 2021; Grehan et al., 2017; Kazanidis et al., 2020; Rice et al., 2012; Rueda et al., 2016; Tunnicliffe et al., 2020). In addition, many of the described patterns are of general interest as they seem to be caused by environmentally driven processes that most likely affect the functioning and the structure of several components of deep-sea ecosystem in the Azores.

Chapter two described the distributional patterns of habitat-forming cold-water corals using habitat suitability models. It suggested that CWCs are organized in two ecological regions defined by the Azores Current and that the depth zonation of CWCs is shaped by regional water masses. Over the considered bathymetric range (0 - 2000 m), CWCs form four groups of species: species restricted to upper water masses, species extending down from upper water masses, species restricted to intermediate waters and species extending up from deep water masses. As all the modelled CWCs represent foundation species, these patterns can be expected to remain valid for all associated species. Because of the results of Chapter two it was decided to further investigate how hard substrate communities change along bathymetric and

latitudinal gradients and to evaluate if the results of Chapter two remained valid using additional biological data and a different modelling approach. Chapter three provides one of the first statistical bioregionalization of deep-sea benthic communities in the Azores. It used species archetype models and generally confirmed the results of Chapter two. The depth-wise succession of the regional water masses and their stratification into different isopycnal layers appears to explain much of the changes in species composition occurring with depth. The ecological processes driving these patterns remain unclear. A water mass can be considered as an environmental envelop characterized by a particular combination of oceanographic parameters. They could exert a direct physiological control, influence connectivity or mediate food supply via internal waves or density-driven currents. Regardless of the main ecological and evolutionary processes in place, the results suggest that the different isopycnal layers can be used as meaningful bioregions, setting an interesting background for ecological research, ecosystem-based management and spatial monitoring. In the region, the main changes in species composition occur at 500, 800 and 1600 m generally confirming the biogeographic boundaries identified in previous studies. Temperature is the main environmental factor controlling the distribution of assemblages truly restricted to the upper layer. A strong negative response to bottom-oxygen concentrations identifies assemblages truly restricted to the intermediate layer. The isopycnal stratification appear to act as a soft ecological boundary for other species. Food export seems the main controlling factor for species associated with the deep layer (or with the lowest portion of the intermediate layer), supporting the hypotheses that the regional productivity gradient in surface waters affects deep-sea assemblages.

In Chapter four I developed a new software (scapesClassification) to be used for the ecological zonation of Chapter five. While a number of methodologies are available to classify seascapes, I felt the need of using a knowledge-oriented approach that could capture and integrate the distributional drivers identified in Chapters two and three with an additional distributional driver of hard substrate communities, i.e. the availability of hard substrate on major topographical features. Traditional geomorphological classification methods generally rely on data that were not available at the time of writing (e.g. continuous backscatter layer) and tend to segment a seascape in many classes (or units) whose ecological meaning was hard to evaluate. A simple knowledge—oriented classification seemed to be more suitable to synthetize the distributional patterns of hard substrate communities in an ecological map.

Chapter five presents a synthesis of the distributional patterns and of the environmental drivers of deep-sea hard substrate benthic communities in the Azores (NE Atlantic). Four major distributional drivers can be recognized in the region: (i) a latitudinal gradient in primary productivity strongly influenced by the Azores Current-Azores Front (AzC-AzF) system; (ii) the depth-wise succession of the regional water masses and their stratification into different isopycnal (vertical) layers; (iii) the spatial distribution of prominent geomorphic features such as seamounts ridges and island slopes; (iv) the availability of hard substrate for attachment. The results provide an estimate of the availability of hard substrate habitats associated with prominent geomorphic features (i.e. seamounts, ridges and island slopes), identify two distinct ecological regions likely to influence benthic deep-sea assemblages and synthetize the distributional patterns in an ecological zonation of the Azores acknowledging the importance of the depth-wise stratification of the regional water masses. The total fraction of rocky substrates on prominent geomorphic features available for the settlement of sessile suspension feeders was estimated to cover an area of about 188,000 km^2 (~11% of the study area). Geomorphic features are extracted from an updated bathymetric grid of the Azores and treated as discrete spatial entities. This approach adopts an object-oriented perspective about the major geomorphic features existing in the region and, in doing so, it provides a more suitable framework to address pattern–process relationships following seascape and landscape ecology strategies. Overall, several studies confirm that the AzC-AzF system is a permanent oceanic feature causing abrupt changes in oceanic conditions. Much of these changes occur in subsurface and deep waters and are often overlooked in analysis based on remote sensing data. The conditions imposed by the AzC-AzF system are likely to influence distribution and abundance of benthic organisms in the Azores. The different frontal systems occurring along the Mid-Atlantic Ridge could be particularly important in explaining latitudinal changes in species composition even at the seafloor. Previous works seem to confirm the importance of the Sub-Polar Front (~52° N). Future studies will have to clarify the biogeographic role of the AzC-AzF system (~36º N).

References

Addamo, A.M., Miller, K.J., Häussermann, V., Taviani, M., Machordom, A., 2021. Global-scale genetic structure of a cosmopolitan cold-water coral species. Aquat. Conserv. Mar. Freshw. Ecosyst. 31, 1–14. https://doi.org/10.1002/aqc.3421

Afoullouss, S., Sanchez, A.R., Jennings, L.K., Kee, Y., Allcock, A.L., Thomas, O.P., 2022. Unveiling the Chemical Diversity of the Deep-Sea Sponge Characella pachastrelloides. Mar. Drugs 20, 52. https://doi.org/10.3390/md20010052

Aguirre-Gutiérrez, J., Serna-Chavez, H.M., Villalobos-Arambula, A.R., Pérez de la Rosa, J.A., Raes, N., 2015. Similar but not equivalent: ecological niche comparison across closely–related Mexican white pines. Divers. Distrib. 21, 245–257. https://doi.org/10.1111/ddi.12268

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Alt, C.H.S., Kremenetskaia (Rogacheva), A., Gebruk, A. V., Gooday, A.J., Jones, D.O.B., 2019. Bathyal benthic megafauna from the Mid-Atlantic Ridge in the region of the Charlie-Gibbs fracture zone based on remotely operated vehicle observations. Deep Sea Res. Part I Oceanogr. Res. Pap. 145, 1–12. https://doi.org/10.1016/j.dsr.2018.12.006

Amorim, P., Perán, A.D., Pham, C.K., Juliano, M., Cardigos, F., Tempera, F., Morato, T., 2017. Overview of the Ocean Climatology and Its Variability in the Azores Region of the North Atlantic Including Environmental Characteristics at the Seabed. Front. Mar. Sci. 4, 1–16. https://doi.org/10.3389/fmars.2017.00056

Anderson, O.F., Guinotte, J.M., Rowden, A.A., Tracey, D.M., Mackay, K.A., Clark, M.R., 2016. Habitat suitability models for predicting the occurrence of vulnerable marine ecosystems in the seas around New Zealand. Deep Sea Res. Part I Oceanogr. Res. Pap. 115, 265–292. https://doi.org/10.1016/j.dsr.2016.07.006

Andrienko, N., Andrienko, G., 2006. Exploratory Analysis of Spatial and Temporal Data. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-31190-4

Appelhans, T., Detsch, F., Reudenbach, C., Woellauer, S., 2021. mapview: Interactive Viewing of Spatial Data in R. R package version 2.10.0. https://CRAN.R-project.org/package=mapview.

Arantes, R., Castro, C., Pires, D., Seoane, J., 2009. Depth and water mass zonation and species associations of cold-water octocoral and stony coral communities in the southwestern Atlantic. Mar. Ecol. Prog. Ser. 397, 71–79. https://doi.org/10.3354/meps08230

Araújo, M.B., Anderson, R.P., Márcia Barbosa, A., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R.B., Zimmermann, N.E., Rahbek, C., 2019. Standards for distribution models in biodiversity assessments. Sci. Adv. 5, eaat4858. https://doi.org/10.1126/sciadv.aat4858

Armstrong, C.W., Aanesen, M., Rensburg, T.M., Sandorf, E.D., 2019. Willingness to pay to protect cold water corals. Conserv. Biol. 0, cobi.13380. https://doi.org/10.1111/cobi.13380

Arvor, D., Belgiu, M., Falomir, Z., Mougenot, I., Durieux, L., 2019. Ontologies to interpret remote sensing images: Why do we need them? GIScience Remote Sens. 56, 911–939. https://doi.org/10.1080/15481603.2019.1587890

Arvor, D., Durieux, L., Andrés, S., Laporte, M.-A., 2013. Advances in Geographic Object-Based Image Analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. ISPRS J. Photogramm. Remote Sens. 82, 125–137. https://doi.org/10.1016/j.isprsjprs.2013.05.003

Auscavitch, S.R., Deere, M.C., Keller, A.G., Rotjan, R.D., Shank, T.M., Cordes, E.E., 2020. Oceanographic Drivers of Deep-Sea Coral Species Distribution and Community Assembly on Seamounts, Islands, Atolls, and Reefs Within the Phoenix Islands Protected Area. Front. Mar. Sci. 7, 1–15. https://doi.org/10.3389/fmars.2020.00042

Azetsu-Scott, K., Clarke, A., Falkner, K., Hamilton, J., Jones, E.P., Lee, C., Petrie, B., Prinsenberg, S., Starr, M., Yeats, P., 2010. Calcium carbonate saturation states in the waters of the Canadian Arctic Archipelago and the Labrador Sea. J. Geophys. Res. 115, C11021. https://doi.org/10.1029/2009JC005917

Baco, A.R., Morgan, N.B., Roark, E.B., 2020. Observations of vulnerable marine ecosystems and significant adverse impacts on high seas seamounts of the northwestern Hawaiian Ridge and Emperor Seamount Chain. Mar. Policy 115, 103834. https://doi.org/10.1016/j.marpol.2020.103834

Badgley, C., Smiley, T.M., Terry, R., Davis, E.B., DeSantis, L.R.G., Fox, D.L., Hopkins, S.S.B., Jezkova, T., Matocq, M.D., Matzke, N., McGuire, J.L., Mulch, A., Riddle, B.R., Roth, V.L., Samuels, J.X., Strömberg, C.A.E., Yanites, B.J., 2017. Biodiversity and Topographic Complexity: Modern and Geohistorical Perspectives. Trends Ecol. Evol. 32, 211–226. https://doi.org/10.1016/j.tree.2016.12.010

Baltar, F., Arístegui, J., 2017. Fronts at the Surface Ocean Can Shape Distinct Regions of Microbial Activity and Community Assemblages Down to the Bathypelagic Zone: The Azores Front as a Case Study. Front. Mar. Sci. 4, 1–13. https://doi.org/10.3389/fmars.2017.00252

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol. Evol. 3, 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Barbosa, R.V., Davies, A.J., Sumida, P.Y.G., 2020. Habitat suitability and environmental niche comparison of cold-water coral species along the Brazilian continental margin. Deep Sea Res. Part I Oceanogr. Res. Pap. 155, 103147. https://doi.org/10.1016/j.dsr.2019.103147

Bart, M.C., Hudspith, M., Rapp, H.T., Verdonschot, P.F.M., de Goeij, J.M., 2021. A Deep-Sea Sponge Loop? Sponges Transfer Dissolved and Particulate Organic Carbon and Nitrogen to Associated Fauna. Front. Mar. Sci. 8, 1–12. https://doi.org/10.3389/fmars.2021.604879

Bartoń, K., 2022. MuMIn: Multi-Model Inference. R package version 1.46.0.

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. Glob. Ecol. Biogeogr. 19, 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Bashmachnikov, I., Lafon, V.M., Martins, A.M., 2004. SST stationary anomalies in the Azores region, in: Bostater, Jr., C.R., Santoleri, R. (Eds.), Remote Sensing of the Ocean and Sea Ice 2004. p. 148. https://doi.org/10.1117/12.565596

Bashmachnikov, I., Mohn, C., Pelegrí, J.L., Martins, A., Jose, F., Machín, F., White, M., 2009. Interaction of Mediterranean water eddies with Sedlo and Seine Seamounts, Subtropical Northeast Atlantic. Deep Sea Res. Part II Top. Stud. Oceanogr. 56, 2593–2605. https://doi.org/10.1016/j.dsr2.2008.12.036

Bashmachnikov, I., Nascimento, Neves, F., Menezes, T., Koldunov, N. V., 2015. Distribution of intermediate water masses in the subtropical northeast Atlantic. Ocean Sci. 11, 803–827. https://doi.org/10.5194/os-11-803-2015

Beazley, L.I., Kenchington, E.L., Murillo, F.J., Sacau, M. del M., 2013. Deep-sea sponge grounds enhance diversity and

abundance of epibenthic megafauna in the Northwest Atlantic. ICES J. Mar. Sci. 70, 1471–1490. https://doi.org/10.1093/icesjms/fst124

Beisiegel, K., Darr, A., Gogina, M., Zettler, M.L., 2017. Benefits and shortcomings of non-destructive benthic imagery for monitoring hard-bottom habitats. Mar. Pollut. Bull. 121, 5–15. https://doi.org/10.1016/j.marpolbul.2017.04.009

Bennecke, S., Metaxas, A., 2017. Is substrate composition a suitable predictor for deep-water coral occurrence on fine scales? Deep Sea Res. Part I Oceanogr. Res. Pap. 124, 55–65. https://doi.org/10.1016/j.dsr.2017.04.011

Bishop, M.P., James, L.A., Shroder, J.F., Walsh, S.J., 2012. Geospatial technologies and digital geomorphological mapping: Concepts, issues and research. Geomorphology 137, 5–26. https://doi.org/10.1016/j.geomorph.2011.06.027

Bishr, Y., 1998. Overcoming the semantic and other barriers to GIS interoperability. Int. J. Geogr. Inf. Sci. 12, 299–314. https://doi.org/10.1080/136588198241806

Bivand, R., 2022. R Packages for Analyzing Spatial Data: A Comparative Case Study with Areal Data. Geogr. Anal. 1–31. https://doi.org/10.1111/gean.12319

Bivand, R.S., 2021. Progress in the R ecosystem for representing and handling spatial data. J. Geogr. Syst. 23, 515–546. https://doi.org/10.1007/s10109-020-00336-0

Boavida, J., Becheler, R., Choquet, M., Frank, N., Taviani, M., Bourillet, J., Meistertzheim, A., Grehan, A., Savini, A., Arnaud-Haond, S., 2019. Out of the Mediterranean? Post-glacial colonization pathways varied among cold-water coral species. J. Biogeogr. 46, 915–931. https://doi.org/10.1111/jbi.13570

Bock, N., Cornec, M., Claustre, H., Duhamel, S., 2022. Biogeographical Classification of the Global Ocean From BGC-Argo Floats. Global Biogeochem. Cycles 36. https://doi.org/10.1029/2021GB007233

Boettiger, C., Chamberlain, S., Hart, E., Ram, K., 2015. Building Software, Building Community: Lessons from the rOpenSci Project. J. Open Res. Softw. 3. https://doi.org/10.5334/jors.bu

Böning, C.W., Behrens, E., Biastoch, A., Getzlaff, K., Bamber, J.L., 2016. Emerging impact of Greenland meltwater on deepwater formation in the North Atlantic Ocean. Nat. Geosci. 9, 523–527. https://doi.org/10.1038/ngeo2740

Bowden, D.A., Anderson, O.F., Rowden, A.A., Stephenson, F., Clark, M.R., 2021. Assessing Habitat Suitability Models for the Deep Sea: Is Our Ability to Predict the Distributions of Seafloor Fauna Improving? Front. Mar. Sci. 8. https://doi.org/10.3389/fmars.2021.632389

Braga-Henriques, A., Carreiro-Silva, M., Porteiro, F.M., de Matos, V., Sampaio, Í., Ocaña, O., Ávila, S.P., 2011. The association between a deep-sea gastropod Pedicularia sicula (Caenogastropoda: Pediculariidae) and its coral host Errina dabneyi (Hydrozoa: Stylasteridae) in the Azores. ICES J. Mar. Sci. 68, 399–407. https://doi.org/10.1093/icesjms/fsq066

Braga-Henriques, A., Porteiro, F.M., Ribeiro, P.A., de Matos, V., Sampaio, Í., Ocaña, O., Santos, R.S., 2013. Diversity, distribution and spatial structure of the cold-water coral fauna of the Azores (NE Atlantic). Biogeosciences 10, 4009–4036. https://doi.org/10.5194/bg-10-4009-2013

Bridges, A.E.H., Barnes, D.K.A., Bell, J.B., Ross, R.E., Howell, K.L., 2022. Depth and latitudinal gradients of diversity in seamount benthic communities. J. Biogeogr. 1–12. https://doi.org/10.1111/jbi.14355

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H., Guisan, A., 2012. Measuring ecological niche overlap from

occurrence and spatial environmental data. Glob. Ecol. Biogeogr. 21, 481–497. https://doi.org/10.1111/j.1466-8238.2011.00698.x

Brown, C.J., Smith, S.J., Lawton, P., Anderson, J.T., 2011. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuar. Coast. Shelf Sci. 92, 502–520. https://doi.org/10.1016/j.ecss.2011.02.007

Brown, J.L., Carnaval, A.C., 2019. A tale of two niches: methods, concepts, and evolution. Front. Biogeogr. 11. https://doi.org/10.21425/F5FBG44158

Buhl-Mortensen, L., Buhl-Mortensen, P., Dolan, M.J.F., Gonzalez-Mirelis, G., 2015. Habitat mapping as a tool for conservation and sustainable use of marine resources: Some perspectives from the MAREANO Programme, Norway. J. Sea Res. 100, 46–61. https://doi.org/10.1016/j.seares.2014.10.014

Buhl-Mortensen, L., Vanreusel, A., Gooday, A.J., Levin, L.A., Priede, I.G., Buhl-Mortensen, P., Gheerardyn, H., King, N.J., Raes, M., 2010. Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. Mar. Ecol. 31, 21–50. https://doi.org/10.1111/j.1439-0485.2010.00359.x

Buhl-Mortensen, P., Buhl-Mortensen, L., Purser, A., 2017. Trophic Ecology and Habitat Provision in Cold-Water Coral Ecosystems, in: Rossi, S., Bramanti, L., Gori, A., Orejas, C. (Eds.), Marine Animal Forests. Springer International Publishing, Cham, pp. 1–42. https://doi.org/10.1007/978-3-319-17001-5_17-2

Cadotte, M.W., Tucker, C.M., 2017. Should Environmental Filtering be Abandoned? Trends Ecol. Evol. 32, 429–437. https://doi.org/10.1016/j.tree.2017.03.004

Cairns, S.D., 2007. Deep-water corals: An overview with special reference to diversity and distribution of deep-water scleractinian corals. Bull. Mar. Sci. 81, 311–322.

Caldeira, R.M.A., Reis, J.C., 2017. The Azores Confluence Zone. Front. Mar. Sci. 4, 1–14. https://doi.org/10.3389/fmars.2017.00037

Carney, R., 2005. Zonation of Deep Biota on Continental Margins, in: Oceanography and Marine Biology. pp. 211–278. https://doi.org/10.1201/9781420037449.ch6

Carreiro-Silva, M., Andrews, A., Braga-Henriques, A., de Matos, V., Porteiro, F., Santos, R., 2013. Variability in growth rates of long-lived black coral Leiopathes sp. from the Azores. Mar. Ecol. Prog. Ser. 473, 189–199. https://doi.org/10.3354/meps10052

Carvalho, F.C., Cárdenas, P., Ríos, P., Cristobo, J., Rapp, H.T., Xavier, J.R., 2020. Rock sponges (lithistid Demospongiae) of the Northeast Atlantic seamounts, with description of ten new species. PeerJ 8, e8703. https://doi.org/10.7717/peerj.8703

Cathalot, C., Van Oevelen, D., Cox, T.J.S., Kutti, T., Lavaleye, M., Duineveld, G., Meysman, F.J.R., 2015. Cold-water coral reefs and adjacent sponge grounds: hotspots of benthic respiration and organic carbon cycling in the deep sea. Front. Mar. Sci. 2, 1–12. https://doi.org/10.3389/fmars.2015.00037

Chen, J., Cohn, A.G., Liu, D., Wang, S., Ouyang, J., Yu, Q., 2015. A survey of qualitative spatial representations. Knowl. Eng. Rev. 30, 106–136. https://doi.org/10.1017/S0269888913000350

Chen, X., Dallmeier-Tiessen, S., Dasler, R., Feger, S., Fokianos, P., Gonzalez, J.B., Hirvonsalo, H., Kousidis, D., Lavasa, A., Mele, S., Rodriguez, D.R., Šimko, T., Smith, T., Trisovic, A., Trzcinska, A., Tsanaktsidis, I., Zimmermann, M., Cranmer, K., Heinrich, L., Watts, G., Hildreth, M., Lloret Iglesias, L., Lassila-Perini, K., Neubert, S., 2019. Open is not enough. Nat. Phys. 15, 113–119. https://doi.org/10.1038/s41567-018-0342-2

Cheng, J., Karambelkar, B., Xie, Y., 2021. leaflet: Create Interactive Web Maps with the JavaScript "Leaflet" Library. R package version 2.0.4.1. https://CRAN.R-project.org/package=leaflet.

Chergui, W., Zidat, S., Marir, F., 2020. An approach to the acquisition of tacit knowledge based on an ontological model. J. King Saud Univ. - Comput. Inf. Sci. 32, 818–828. https://doi.org/10.1016/j.jksuci.2018.09.012

Chimienti, G., Bo, M., Taviani, M., Mastrototaro, F., 2019. Occurrence and Biogeography of Mediterranean Cold-Water Corals, in: Orejas, C., Jiménez, C. (Eds.), Mediterranean Cold-Water Corals: Past, Present and Future. Coral Reefs of the World, vol 9. Springer, Cham., pp. 213–243. https://doi.org/10.1007/978-3-319-91608-8_19

Clark, M.R., Tittensor, D.P., 2010. An index to assess the risk to stony corals from bottom trawling on seamounts. Mar. Ecol. 31, 200–211. https://doi.org/10.1111/j.1439-0485.2010.00392.x

Coetzee, S., Carow, S., Snyman, L., 2021. How are maps used in research? An exploratory review of PhD dissertations. Adv. Cartogr. GIScience ICA 3, 1–8. https://doi.org/10.5194/ica-adv-3-3-2021

Comas-Rodríguez, I., Hernández-Guerra, A., Fraile-Nuez, E., Martínez-Marrero, A., Benítez-Barrios, V.M., Pérez-Hernández, M.D., Vélez-Belchí, P., 2011. The Azores Current System from a meridional section at 24.5°W. J. Geophys. Res. 116, C09021. https://doi.org/10.1029/2011JC007129

Combes, M., Vaz, S., Grehan, A., Morato, T., Arnaud-Haond, S., Dominguez-Carrió, C., Fox, A., González-Irusta, J.M., Johnson, D., Callery, O., Davies, A., Fauconnet, L., Kenchington, E., Orejas, C., Roberts, J.M., Taranto, G., Menot, L., 2021. Systematic Conservation Planning at an Ocean Basin Scale: Identifying a Viable Network of Deep-Sea Protected Areas in the North Atlantic and the Mediterranean. Front. Mar. Sci. 8, 1–27. https://doi.org/10.3389/fmars.2021.611358

Conci, N., Vargas, S., Wörheide, G., 2021. The Biology and Evolution of Calcite and Aragonite Mineralization in Octocorallia. Front. Ecol. Evol. 9, 1–19. https://doi.org/10.3389/fevo.2021.623774

Cornell, H. V., Harrison, S.P., 2014. What Are Species Pools and When Are They Important? Annu. Rev. Ecol. Evol. Syst. 45, 45–67. https://doi.org/10.1146/annurev-ecolsys-120213-091759

Corrêa, P.V.F., Jovane, L., Murton, B.J., Sumida, P.Y.G., 2022. Benthic megafauna habitats, community structure and environmental drivers at Rio Grande Rise (SW Atlantic). Deep Sea Res. Part I Oceanogr. Res. Pap. 186, 103811. https://doi.org/10.1016/j.dsr.2022.103811

Costello, M.J., Chaudhary, C., 2017. Marine Biodiversity, Biogeography, Deep-Sea Gradients, and Conservation. Curr. Biol. 27, R511–R527. https://doi.org/10.1016/j.cub.2017.04.060

Costello, M.J., Tsai, P., Wong, P.S., Cheung, A.K.L., Basher, Z., Chaudhary, C., 2017. Marine biogeographic realms and species endemicity. Nat. Commun. 8, 1057. https://doi.org/10.1038/s41467-017-01121-2

Cova, T.J., Goodchild, M.F., 2002. Extending geographical representation to include fields of spatial objects. Int. J. Geogr. Inf. Sci. 16, 509–532. https://doi.org/10.1080/13658810210137040

Crotty, S.M., Altieri, A.H., Bruno, J.F., Fischman, H., Bertness, M.D., 2019. The Foundation for Building the Conservation Capacity of Community Ecology. Front. Mar. Sci. 6, 1–8. https://doi.org/10.3389/fmars.2019.00238

Cyr, F., Haren, H., Mienis, F., Duineveld, G., Bourgault, D., 2016. On the influence of cold-water coral mound size on flow hydrodynamics, and vice versa. Geophys. Res. Lett. 43, 775–783. https://doi.org/10.1002/2015GL067038

D'Amen, M., Rahbek, C., Zimmermann, N.E., Guisan, A., 2017. Spatial predictions at the community level: from current approaches to future frameworks. Biol. Rev. 92, 169–187. https://doi.org/10.1111/brv.12222

Danovaro, R., Corinaldesi, C., Dell'Anno, A., Snelgrove, P.V.R., 2017. The deep-sea under global change. Curr. Biol. 27, R461–R465. https://doi.org/10.1016/j.cub.2017.02.046

Danovaro, R., Fanelli, E., Canals, M., Ciuffardi, T., Fabri, M.-C., Taviani, M., Argyrou, M., Azzurro, E., Bianchelli, S., Cantafaro, A., Carugati, L., Corinaldesi, C., de Haan, W.P., Dell'Anno, A., Evans, J., Foglini, F., Galil, B., Gianni, M., Goren, M., Greco, S., Grimalt, J., Güell-Bujons, Q., Jadaud, A., Knittweis, L., Lopez, J.L., Sanchez-Vidal, A., Schembri, P.J., Snelgrove, P., Vaz, S., Angeletti, L., Barsanti, M., Borg, J.A., Bosso, M., Brind'Amour, A., Castellan, G., Conte, F., Delbono, I., Galgani, F., Morgana, G., Prato, S., Schirone, A., Soldevila, E., 2020. Towards a marine strategy for the deep Mediterranean Sea: Analysis of current ecological status. Mar. Policy 112, 103781. https://doi.org/10.1016/j.marpol.2019.103781

Danovaro, R., Snelgrove, P.V.R., Tyler, P., 2014. Challenging the paradigms of deep-sea ecology. Trends Ecol. Evol. 29, 465–475. https://doi.org/10.1016/j.tree.2014.06.002

Davies, A.J., Guinotte, J.M., 2011. Global Habitat Suitability for Framework-Forming Cold-Water Corals. PLoS One 6, e18483. https://doi.org/10.1371/journal.pone.0018483

de Mendonça, S.N., Metaxas, A., 2021. Comparing the Performance of a Remotely Operated Vehicle, a Drop Camera, and a Trawl in Capturing Deep-Sea Epifaunal Abundance and Diversity. Front. Mar. Sci. 8, 1–17. https://doi.org/10.3389/fmars.2021.631354

Dekavalla, M., Argialas, D., 2017. Object-based classification of global undersea topography and geomorphological features from the SRTM30_PLUS data. Geomorphology 288, 66–82. https://doi.org/10.1016/j.geomorph.2017.03.026

Dominguez-Carrió, C., Fontes, J., Morato, T., 2021. A cost-effective video system for a rapid appraisal of deep-sea benthic habitats: The Azor drift-cam. Methods Ecol. Evol. 12, 1379–1388. https://doi.org/10.1111/2041-210X.13617

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop.). 36, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dowle, M., Srinivasan, A., 2021. data.table: Extension of `data.frame`. R package version 1.14.2. https://CRAN.R-project.org/package=data.table.

Drake, J.M., Randin, C., Guisan, A., 2006. Modelling ecological niches with support vector machines. J. Appl. Ecol. 43, 424–432. https://doi.org/10.1111/j.1365-2664.2006.01141.x

Drescher, M., Perera, A.H., Johnson, C.J., Buse, L.J., Drew, C.A., Burgman, M.A., 2013. Toward rigorous use of expert knowledge in ecological research. Ecosphere 4, 1–26. https://doi.org/10.1890/ES12-00415.1

Dunn, D.C., Van Dover, C.L., Etter, R.J., Smith, C.R., Levin, L.A., Morato, T., Colaço, A., Dale, A.C., Gebruk, A. V., Gjerde, K.M., Halpin, P.N., Howell, K.L., Johnson, D., Perez, J.A.A., Ribeiro, M.C., Stuckas, H., Weaver, P., 2018. A strategy for the conservation of biodiversity on mid-ocean ridges from deep-sea mining. Sci. Adv. 4, 1–16. https://doi.org/10.1126/sciadv.aar4313

Dunstan, P.K., Foster, S.D., Darnell, R., 2016. SpeciesMix: Fit Mixtures of Archetype Species. R package version 0.3.4.

Dunstan, P.K., Foster, S.D., Darnell, R., 2011. Model based grouping of species across environmental gradients. Ecol. Modell. 222, 955–963. https://doi.org/10.1016/j.ecolmodel.2010.11.030

Dunstan, P.K., Foster, S.D., Hui, F.K.C., Warton, D.I., 2013. Finite Mixture of Regression Modeling for High-

Dimensional Count and Biomass Data in Ecology. J. Agric. Biol. Environ. Stat. 18, 357–375. https://doi.org/10.1007/s13253-013-0146-x

Durden, J., Bett, B., Schoening, T., Morris, K., Nattkemper, T., Ruhl, H., 2016. Comparison of image annotation data generated by multiple investigators for benthic ecology. Mar. Ecol. Prog. Ser. 552, 61–70. https://doi.org/10.3354/meps11775

Durden, J.M., Bett, B.J., Ruhl, H.A., 2020. Subtle variation in abyssal terrain induces significant change in benthic megafaunal abundance, diversity, and community structure. Prog. Oceanogr. 186, 102395. https://doi.org/10.1016/j.pocean.2020.102395

Dutkiewicz, S., Hickman, A.E., Jahn, O., Gregg, W.W., Mouw, C.B., Follows, M.J., 2015. Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. Biogeosciences 12, 4447–4481. https://doi.org/10.5194/bg-12-4447-2015

Elith, J., Ferrier, S., Huettmann, F., Leathwick, J., 2005. The evaluation strip: A new and robust method for plotting predicted responses from species distribution models. Ecol. Modell. 186, 280–289. https://doi.org/10.1016/j.ecolmodel.2004.12.007

Elith, J., Leathwick, J.R., 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Ellison, A.M., 2019. Foundation Species, Non-trophic Interactions, and the Value of Being Common. iScience 13, 254–268. https://doi.org/10.1016/j.isci.2019.02.020

Emery, W.J., 2001. Water Types And Water Masses, in: Encyclopedia of Ocean Sciences. Elsevier, pp. 3179–3187. https://doi.org/10.1006/rwos.2001.0108

EMODnet Bathymetry Consortium, 2018. EMODnet Digital Bathymetry (DTM 2018). https://doi.org/10.12770/18ff0d48-b203-4a65-94a9-5fd8b0ec35f6

F. Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography (Cop.). 30, 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

FAO, 2009. International guidelines for the management of deep-sea fisheries in the high seas. Rome, FAO.

Farley, S.S., Dawson, A., Goring, S.J., Williams, J.W., 2018. Situating ecology as a big-data science: Current advances, challenges, and solutions. Bioscience 68, 563–576. https://doi.org/10.1093/biosci/biy068

Fazey, I., Fazey, J.A., Fazey, D.M.A., 2005. Learning More Effectively from Experience. Ecol. Soc. 10, art4. https://doi.org/10.5751/ES-01384-100204

Ferro, I., Morrone, J.J., 2014. Biogeographical transition zones: a search for conceptual synthesis. Biol. J. Linn. Soc. 113, 1–12. https://doi.org/10.1111/bij.12333

Fink, H.G., Wienberg, C., De Pol-Holz, R., Hebbeln, D., 2015. Spatio-temporal distribution patterns of Mediterranean cold-water corals (Lophelia pertusa and Madrepora oculata) during the past 14,000 years. Deep. Res. Part I Oceanogr. Res. Pap. 103, 37–48. https://doi.org/10.1016/j.dsr.2015.05.006

Flögel, S., Dullo, W.C., Pfannkuche, O., Kiriakoulakis, K., Rüggeberg, A., 2014. Geochemical and physical constraints for the occurrence of living cold-water corals. Deep. Res. Part II Top. Stud. Oceanogr. 99, 19–26. https://doi.org/10.1016/j.dsr2.2013.06.006

Folkersen, M.V., Fleming, C.M., Hasan, S., 2019. Depths of uncertainty for deep-sea policy and legislation. Glob. Environ. Chang. 54, 1–5. https://doi.org/10.1016/j.gloenvcha.2018.11.002

Folkersen, M.V., Fleming, C.M., Hasan, S., 2018. The economic value of the deep sea: A systematic review and meta-analysis. Mar. Policy 94, 71–80. https://doi.org/10.1016/j.marpol.2018.05.003

Fonseca, F.T., Egenhofer, M.J., Agouris, P., Cmara, G., 2002. Using ontologies for integrated geographic information systems. Trans. GIS 6, 231–257. https://doi.org/10.1111/1467-9671.00109

Fontela, M., Pérez, F.F., Carracedo, L.I., Padín, X.A., Velo, A., García-Ibañez, M.I., Lherminier, P., 2020. The Northeast Atlantic is running out of excess carbonate in the horizon of cold-water corals communities. Sci. Rep. 10, 14714. https://doi.org/10.1038/s41598-020-71793-2

Foster, S.D., Dunstan, P.K., Althaus, F., Williams, A., 2015. The cumulative effect of trawl fishing on a multispecies fish assemblage in south-eastern Australia. J. Appl. Ecol. 52, 129–139. https://doi.org/10.1111/1365-2664.12353

Frazão, H.C., Prien, R.D., Schulz-Bull, D.E., Seidov, D., Waniek, J.J., 2022. The Forgotten Azores Current: A Long-Term Perspective. Front. Mar. Sci. 9, 1–14. https://doi.org/10.3389/fmars.2022.842251

Frazier, A.E., Kedron, P., 2017. Landscape Metrics: Past Progress and Future Directions. Curr. Landsc. Ecol. Reports 2, 63–72. https://doi.org/10.1007/s40823-017-0026-0

Freksa, C., Barkowsky, T., 1995. On the Relations between Spatial Concepts and Geographic Objects, in: Burrough, P., Frank, A. (Eds.), Geographic Objects with Indeterminate Boundaries. Taylor & Francis, London.

Fründt, B., Müller, T.J., Schulz-Bull, D.E., Waniek, J.J., 2013. Long-term changes in the thermocline of the subtropical Northeast Atlantic (33°N, 22°W). Prog. Oceanogr. 116, 246–260. https://doi.org/10.1016/j.pocean.2013.07.004

Fründt, B., Waniek, J., 2012. Impact of the azores front propagation on deep ocean particle flux. Open Geosci. 4, 531–544. https://doi.org/10.2478/s13533-012-0102-2

Fuentes, B.A., Dorantes, M.J., Tipton, J.R., 2021. rassta: Raster-based Spatial Stratification Algorithms. EarthArXiv, 2021. https://doi.org/10.31223/X50S57 1–16. https://doi.org/10.31223/X50S57

Galparsoro, I., Rodríguez, J.G., Menchaca, I., Quincoces, I., Garmendia, J.M., Borja, Á., 2015. Benthic habitat mapping on the Basque continental shelf (SE Bay of Biscay) and its application to the European Marine Strategy Framework Directive. J. Sea Res. 100, 70–76. https://doi.org/10.1016/j.seares.2014.09.013

GEBCO Compilation Group, 2022. GEBCO_2022 Grid.

Gebruk, A. V., Budaeva, N.E., King, N.J., 2010. Bathyal benthic fauna of the Mid-Atlantic Ridge between the Azores and the Reykjanes Ridge. J. Mar. Biol. Assoc. United Kingdom 90, 1–14. https://doi.org/10.1017/S0025315409991111

Georgian, S.E., Anderson, O.F., Rowden, A.A., 2019. Ensemble habitat suitability modeling of vulnerable marine ecosystem indicator taxa to inform deep-sea fisheries management in the South Pacific Ocean. Fish. Res. 211, 256–274. https://doi.org/10.1016/j.fishres.2018.11.020

Georgian, S.E., Deleo, D., Durkin, A., Gomez, C.E., Kurman, M., Lunden, J.J., Cordes, E.E., 2016. Oceanographic patterns and carbonate chemistry in the vicinity of cold-water coral reefs in the Gulf of Mexico: Implications for resilience in a changing ocean. Limnol. Oceanogr. 61, 648–665. https://doi.org/10.1002/lno.10242

Glover, A.G., Higgs, N.D., Horton, T., 2022. World Register of Deep-Sea species (WoRDSS). Accessed at

https://www.marinespecies.org/deepsea on 2022-05-19. https://doi.org/10.14284/352

Glover, A.G., Wiklund, H., Chen, C., Dahlgren, T.G., 2018. Managing a sustainable deep-sea 'blue economy' requires knowledge of what actually lives there. Elife 7, 1–7. https://doi.org/10.7554/eLife.41319

Goldberg, W.M., 2018. Coral Food, Feeding, Nutrition, and Secretion: A Review, in: Kloc, M., Kubiak, J.Z. (Eds.), Marine Organisms as Model Systems in Biology and Medicine. Results and Problems in Cell Differentiation. vol 65. Springer, Cham, pp. 377–421. https://doi.org/10.1007/978-3-319-92486-1_18

Gomes-Pereira, J.N., Carmo, V., Catarino, D., Jakobsen, J., Alvarez, H., Aguilar, R., Hart, J., Giacomello, E., Menezes, G., Stefanni, S., Colaço, A., Morato, T., Santos, R.S., Tempera, F., Porteiro, F., 2017. Cold-water corals and large hydrozoans provide essential fish habitat for Lappanella fasciata and Benthocometes robustus. Deep. Res. Part II Top. Stud. Oceanogr. 145, 33–48. https://doi.org/10.1016/j.dsr2.2017.09.015

Gonçalves, G., 2021. Fishing Impact on Deep-Sea Cold-Water Coral Communities of the Azores: Contribution to Descriptor 6 of the Marine Strategy Framework Directive.

González-Dávila, M., Santana-Casiano, J.M., Fine, R.A., Happell, J., Delille, B., Speich, S., 2011. Carbonate system in the water masses of the Southeast Atlantic sector of the Southern Ocean during February and March 2008. Biogeosciences 8, 1401–1413. https://doi.org/10.5194/bg-8-1401-2011

Goodchild, M.F., Yuan, M., Cova, T.J., 2007. Towards a general theory of geographic representation in GIS. Int. J. Geogr. Inf. Sci. 21, 239–260. https://doi.org/10.1080/13658810600965271

Grehan, A.J., Arnaud-Haond, S., D'Onghia, G., Savini, A., Yesson, C., 2017. Towards ecosystem based management and monitoring of the deep Mediterranean, North-East Atlantic and Beyond. Deep Sea Res. Part II Top. Stud. Oceanogr. 145, 1–7. https://doi.org/10.1016/j.dsr2.2017.09.014

Guilbert, E., Moulin, B., 2017. Towards a common framework for the identification of landforms on terrain models. ISPRS Int. J. Geo-Information 6. https://doi.org/10.3390/ijgi6010012

Guo, H., Wang, L., Chen, F., Liang, D., 2014. Scientific big data and Digital Earth. Chinese Sci. Bull. 59, 5066–5073. https://doi.org/10.1007/s11434-014-0645-3

Hagen-Zanker, A., 2009. An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation. Int. J. Geogr. Inf. Sci. 23, 61–73. https://doi.org/10.1080/13658810802570317

Hagen, A., 2003. Fuzzy set approach to assessing similarity of categorical maps. Int. J. Geogr. Inf. Sci. 17, 235–249. https://doi.org/10.1080/13658810210157822

Halpern, B.S., Frazier, M., Afflerbach, J., Lowndes, J.S., Micheli, F., O'Hara, C., Scarborough, C., Selkoe, K.A., 2019. Recent pace of change in human impact on the world's ocean. Sci. Rep. 9, 11609. https://doi.org/10.1038/s41598-019-47201-9

Hampton, S.E., Anderson, S.S., Bagby, S.C., Gries, C., Han, X., Hart, E.M., Jones, M.B., Lenhardt, W.C., MacDonald, A., Michener, W.K., Mudge, J., Pourmokhtarian, A., Schildhauer, M.P., Woo, K.H., Zimmerman, N., 2015. The Tao of open science for ecology. Ecosphere 6, art120. https://doi.org/10.1890/ES14-00402.1

Hanz, U., Roberts, E.M., Duineveld, G., Davies, A., van Haren, H., Rapp, H.T., Reichart, G., Mienis, F., 2021. Long-term Observations Reveal Environmental Conditions and Food Supply Mechanisms at an Arctic Deep-Sea Sponge Ground. J. Geophys. Res. Ocean. 126. https://doi.org/10.1029/2020JC016776

Hanz, U., Wienberg, C., Hebbeln, D., Duineveld, G., Lavaleye, M., Juva, K., Dullo, W.-C., Freiwald, A., Tamborrino, L., Reichart, G.-J., Flögel, S., Mienis, F., 2019. Environmental factors influencing benthic communities in the

oxygen minimum zones on the Angolan and Namibian margins. Biogeosciences 16, 4337–4356. https://doi.org/10.5194/bg-16-4337-2019

Hastie, T., Tibshirani, R., 1990. Generalized Additive Models. Chapman & Hall/CRC, New York/Boca Raton.

Hebbeln, D., Wienberg, C., Dullo, W.-C., Freiwald, A., Mienis, F., Orejas, C., Titschack, J., 2020. Cold-water coral reefs thriving under hypoxia. Coral Reefs. https://doi.org/10.1007/s00338-020-01934-6

Henderson, M.J., Huff, D.D., Yoklavich, M.M., 2020. Deep-Sea Coral and Sponge Taxa Increase Demersal Fish Diversity and the Probability of Fish Presence. Front. Mar. Sci. 7, 1–19. https://doi.org/10.3389/fmars.2020.593844

Henry, L.-A., Moreno Navas, J., Roberts, J.M., 2013. Multi-scale interactions between local hydrography, seabed topography, and community assembly on cold-water coral reefs. Biogeosciences 10, 2737–2746. https://doi.org/10.5194/bg-10-2737-2013

Hertzog, L.R., Besnard, A., Jay-Robert, P., 2014. Field validation shows bias-corrected pseudo-absence selection is the best method for predictive species-distribution modelling. Divers. Distrib. 20, 1403–1413. https://doi.org/10.1111/ddi.12249

Hijmans, R., Phillips, Leathwick, J., Elith, J., 2022. dismo: Species Distribution Modeling. R package version 1.3-5.

Hijmans, R.J., 2022. terra: Spatial Data Analysis. R package version 1.5-21. https://CRAN.R-project.org/package=terra.

Hill, N., Woolley, S.N.C., Foster, S., Dunstan, P.K., McKinlay, J., Ovaskainen, O., Johnson, C., 2020. Determining marine bioregions: A comparison of quantitative approaches. Methods Ecol. Evol. 11, 1258–1272. https://doi.org/10.1111/2041-210X.13447

Hoeberechts, M., Owens, D., Riddell, D.J., Robertson, A.D., 2015. The Power of Seeing: Experiences using video as a deep-sea engagement and education tool, in: OCEANS 2015 - MTS/IEEE Washington. IEEE, pp. 1–9. https://doi.org/10.23919/OCEANS.2015.7404592

Hogg, M.M., Tendal, O.S., Conway, K.W., Pomponi, S.A., Van Soest, R.W.M., Gutt, J., Krautter, M., Roberts, J.M., 2010. Deep-sea Sponge Grounds: Reservoirs of Biodiversity, UNEP-WCMC Biodiversity Series No. 32. UNEP-WCMC, Cambridge, UK.

Hogg, O.T., Huvenne, V.A.I., Griffiths, H.J., Dorschel, B., Linse, K., 2016. Landscape mapping at sub-Antarctic South Georgia provides a protocol for underpinning large-scale marine protected areas. Sci. Rep. 6, 33163. https://doi.org/10.1038/srep33163

Hortal, J., De Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M., Ladle, R.J., 2015. Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. Annu. Rev. Ecol. Evol. Syst. 46, 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400

Howell, K.-L., Piechaud, N., Downie, A.-L., Kenny, A., 2016. The distribution of deep-sea sponge aggregations in the North Atlantic and implications for their effective spatial management. Deep Sea Res. Part I Oceanogr. Res. Pap. 115, 309–320. https://doi.org/10.1016/j.dsr.2016.07.005

Howell, K.L., 2010. A benthic classification system to aid in the implementation of marine protected area networks in the deep/high seas of the NE Atlantic. Biol. Conserv. 143, 1041–1056. https://doi.org/10.1016/j.biocon.2010.02.001

Howell, K.L., Bullimore, R.D., Foster, N.L., 2014. Quality assurance in the identification of deep-sea taxa from video

and image analysis: response to Henry and Roberts. ICES J. Mar. Sci. 71, 899–906. https://doi.org/10.1093/icesjms/fsu052

Howell, K.L., Hilário, A., Allcock, A.L., Bailey, D.M., Baker, M., Clark, M.R., Colaço, A., Copley, J., Cordes, E.E., Danovaro, R., Dissanayake, A., Escobar, E., Esquete, P., Gallagher, A.J., Gates, A.R., Gaudron, S.M., German, C.R., Gjerde, K.M., Higgs, N.D., Le Bris, N., Levin, L.A., Manea, E., McClain, C., Menot, L., Mestre, N.C., Metaxas, A., Milligan, R.J., Muthumbi, A.W.N., Narayanaswamy, B.E., Ramalho, S.P., Ramirez-Llodra, E., Robson, L.M., Rogers, A.D., Sellanes, J., Sigwart, J.D., Sink, K., Snelgrove, P.V.R., Stefanoudis, P. V., Sumida, P.Y., Taylor, M.L., Thurber, A.R., Vieira, R.P., Watanabe, H.K., Woodall, L.C., Xavier, J.R., 2020. A Blueprint for an Inclusive, Global Deep-Sea Ocean Decade Field Program. Front. Mar. Sci. 7, 1–25. https://doi.org/10.3389/fmars.2020.584861

Huang, T.-Y., Yang, L., 2022. An empirical exploration of the vibrant R ecosystem. arXiv.

Hui, F.K.C., Warton, D.I., Foster, S.D., Dunstan, P.K., 2013. To mix or not to mix: comparing the predictive performance of mixture models vs. separate species distribution models. Ecology 94, 1913–1919. https://doi.org/10.1890/12-1322.1

Irl, S.D.H., Harter, D.E. V., Steinbauer, M.J., Gallego Puyol, D., Fernández-Palacios, J.M., Jentsch, A., Beierkuhnlein, C., 2015. Climate vs. topography - spatial patterns of plant species diversity and endemism on a high-elevation island. J. Ecol. 103, 1621–1633. https://doi.org/10.1111/1365-2745.12463

Ishwaran, H., James, L.F., 2001. Gibbs Sampling Methods for Stick-Breaking Priors. J. Am. Stat. Assoc. 96, 161–173. https://doi.org/10.1198/016214501750332758

Ismail, K., Huvenne, V.A.I., Masson, D.G., 2015. Objective automated classification technique for marine landscape mapping in submarine canyons. Mar. Geol. 362, 17–32. https://doi.org/10.1016/j.margeo.2015.01.006

Jackson, A.C., 2010. Effects of topography on the environment. J. Mar. Biol. Assoc. United Kingdom 90, 169–192. https://doi.org/10.1017/S0025315409991123

Janowicz, K., 2012. Observation-Driven Geo-Ontology Engineering. Trans. GIS 16, 351–374. https://doi.org/10.1111/j.1467-9671.2012.01342.x

Janowicz, K., Scheider, S., Pehle, T., Hart, G., 2012. Geospatial semantics and linked spatiotemporal data – Past, present, and future. Semant. Web 3, 321–332. https://doi.org/10.3233/SW-2012-0077

Jansen, J., Dunstan, P.K., Hill, N.A., Koubbi, P., Melbourne-Thomas, J., Causse, R., Johnson, C.R., 2020. Integrated assessment of the spatial distribution and structural dynamics of deep benthic marine communities. Ecol. Appl. 30, 1–16. https://doi.org/10.1002/eap.2065

Jansen, J., Hill, N.A., Dunstan, P.K., McKinlay, J., Sumner, M.D., Post, A.L., Eléaume, M.P., Armand, L.K., Warnock, J.P., Galton-Fenzi, B.K., Johnson, C.R., 2018. Abundance and richness of key Antarctic seafloor fauna correlates with modelled food availability. Nat. Ecol. Evol. 2, 71–80. https://doi.org/10.1038/s41559-017-0392-3

Jasiewicz, J., Stepinski, T.F., 2013. Geomorphons — a pattern recognition approach to classification and mapping of landforms. Geomorphology 182, 147–156. https://doi.org/10.1016/j.geomorph.2012.11.005

Jenkins, D.G., Ricklefs, R.E., 2011. Biogeography and ecology: two views of one world. Philos. Trans. R. Soc. B Biol. Sci. 366, 2331–2335. https://doi.org/10.1098/rstb.2011.0064

Jerosch, K., Kuhn, G., Krajnik, I., Scharf, F.K., Dorschel, B., 2016. A geomorphological seabed classification for the Weddell Sea, Antarctica. Mar. Geophys. Res. 37, 127–141. https://doi.org/10.1007/s11001-015-9256-x

Jetz, W., McGowan, J., Rinnan, D.S., Possingham, H.P., Visconti, P., O'Donnell, B., Londoño-Murcia, M.C., 2022.

Include biodiversity representation indicators in area-based conservation targets. Nat. Ecol. Evol. 6, 123–126. https://doi.org/10.1038/s41559-021-01620-y

Juva, K., Flögel, S., Karstensen, J., Linke, P., Dullo, W.-C., 2020. Tidal Dynamics Control on Cold-Water Coral Growth: A High-Resolution Multivariable Study on Eastern Atlantic Cold-Water Coral Sites. Front. Mar. Sci. 7, 1–23. https://doi.org/10.3389/fmars.2020.00132

Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V., 2017. Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. IEEE Trans. Knowl. Data Eng. 29, 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168

Kaufman, L., Rousseeuw, P.J. (Eds.), 1990. Finding Groups in Data, Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA. https://doi.org/10.1002/9780470316801

Kazanidis, G., Orejas, C., Borja, A., Kenchington, E., Henry, L.-A., Callery, O., Carreiro-Silva, M., Egilsdottir, H., Giacomello, E., Grehan, A., Menot, L., Morato, T., Ragnarsson, S.Á., Rueda, J.L., Stirling, D., Stratmann, T., van Oevelen, D., Palialexis, A., Johnson, D., Roberts, J.M., 2020. Assessing the environmental status of selected North Atlantic deep-sea ecosystems. Ecol. Indic. 119, 106624. https://doi.org/10.1016/j.ecolind.2020.106624

Kelleher, C., Braswell, A., 2021. Introductory overview: Recommendations for approaching scientific visualization with large environmental datasets. Environ. Model. Softw. 143, 105113. https://doi.org/10.1016/j.envsoft.2021.105113

Kikvidze, Z., Callaway, R.M., 2009. Ecological Facilitation May Drive Major Evolutionary Transitions. Bioscience 59, 399–404. https://doi.org/10.1525/bio.2009.59.5.7

Kirkfeldt, T.S., van Tatenhove, J.P.M., Calado, H.M.G.P., 2022. The Way Forward on Ecosystem-Based Marine Spatial Planning in the EU. Coast. Manag. 50, 29–44. https://doi.org/10.1080/08920753.2022.2006879

Klompmaker, A.A., Jakobsen, S.L., Lauridsen, B.W., 2016. Evolution of body size, vision, and biodiversity of coral-associated organisms: evidence from fossil crustaceans in cold-water coral and tropical coral ecosystems. BMC Evol. Biol. 16, 132. https://doi.org/10.1186/s12862-016-0694-0

Kohonen, T., 1998. The self-organizing map. Neurocomputing 21, 1–6. https://doi.org/10.1016/S0925-2312(98)00030-7

Kokla, M., Guilbert, E., 2020. A Review of Geospatial Semantic Information Modeling and Elicitation Approaches. ISPRS Int. J. Geo-Information 9, 146. https://doi.org/10.3390/ijgi9030146

Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S., Levine, J.M., 2015. Community assembly, coexistence and the environmental filtering metaphor. Funct. Ecol. 592–599. https://doi.org/10.1111/1365-2435.12345

Kreft, H., Jetz, W., 2010. A framework for delineating biogeographical regions based on species distributions. J. Biogeogr. 37, 2029–2053. https://doi.org/10.1111/j.1365-2699.2010.02375.x

Lai, J., Lortie, C.J., Muenchen, R.A., Yang, J., Ma, K., 2019. Evaluating the popularity of R in ecology. Ecosphere 10, 1–7. https://doi.org/10.1002/ecs2.2567

Landis, J.R., Koch, G.G., 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33, 159. https://doi.org/10.2307/2529310

Lapointe, A.E., Watling, L., France, S.C., Auster, P.J., 2020. Megabenthic assemblages in the lower bathyal (700–3000 m) on the New England and Corner Rise Seamounts, Northwest Atlantic. Deep Sea Res. Part I Oceanogr. Res. Pap. 165, 103366. https://doi.org/10.1016/j.dsr.2020.103366

Leaper, R., Dunstan, P.K., Foster, S.D., Barrett, N.S., Edgar, G.J., 2014. Do communities exist? Complex patterns of overlapping marine species distributions. Ecology 95, 2016–2025. https://doi.org/10.1890/13-0789.1

Lecours, V., Devillers, R., Edinger, E.N., Brown, C.J., Lucieer, V.L., 2017. Influence of artefacts in marine digital terrain models on habitat maps and species distribution models: a multiscale assessment. Remote Sens. Ecol. Conserv. 3, 232–246. https://doi.org/10.1002/rse2.49

Lecours, V., Dolan, M.F.J., Micallef, A., Lucieer, V.L., 2016. A review of marine geomorphometry, the quantitative study of the seafloor. Hydrol. Earth Syst. Sci. 20, 3207–3244. https://doi.org/10.5194/hess-20-3207-2016

Lepczyk, C.A., Wedding, L.M., Asner, G.P., Pittman, S.J., Goulden, T., Linderman, M.A., Gang, J., Wright, R., 2021. Advancing Landscape and Seascape Ecology from a 2D to a 3D Science. Bioscience 71, 596–608. https://doi.org/10.1093/biosci/biab001

Levin, L.A., 2003. OXYGEN MINIMUM ZONE BENTHOS: ADAPTATION AND COMMUNITY RESPONSE TO HYPOXIA, in: Oceanography and Marine Biology, An Annual Review, Volume 41. CRC Press, pp. 9–9. https://doi.org/10.1201/9780203180570-3

Levin, L.A., Bett, B.J., Gates, A.R., Heimbach, P., Howe, B.M., Janssen, F., McCurdy, A., Ruhl, H.A., Snelgrove, P., Stocks, K.I., Bailey, D., Baumann-Pickering, S., Beaverson, C., Benfield, M.C., Booth, D.J., Carreiro-Silva, M., Colaço, A., Eblé, M.C., Fowler, A.M., Gjerde, K.M., Jones, D.O.B., Katsumata, K., Kelley, D., Le Bris, N., Leonardi, A.P., Lejzerowicz, F., Macreadie, P.I., McLean, D., Meitz, F., Morato, T., Netburn, A., Pawlowski, J., Smith, C.R., Sun, S., Uchida, H., Vardaro, M.F., Venkatesan, R., Weller, R.A., 2019. Global Observing Needs in the Deep Ocean. Front. Mar. Sci. 6, 1–32. https://doi.org/10.3389/fmars.2019.00241

Levin, L.A., Etter, R.J., Rex, M.A., Gooday, A.J., Smith, C.R., Pineda, J., Stuart, C.T., Hessler, R.R., Pawson, D., 2001. Environmental Influences on Regional Deep-Sea Species Diversity. Annu. Rev. Ecol. Syst. 32, 51–93. https://doi.org/10.1146/annurev.ecolsys.32.081501.114002

Levin, L.A., Sibuet, M., 2012. Understanding Continental Margin Biodiversity: A New Imperative. Ann. Rev. Mar. Sci. 4, 79–112. https://doi.org/10.1146/annurev-marine-120709-142714

Levin, L.A., Sibuet, M., Gooday, A.J., Smith, C.R., Vanreusel, A., 2010. The roles of habitat heterogeneity in generating and maintaining biodiversity on continental margins: an introduction. Mar. Ecol. 31, 1–5. https://doi.org/10.1111/j.1439-0485.2009.00358.x

Levin, L.A., Wei, C., Dunn, D.C., Amon, D.J., Ashford, O.S., Cheung, W.W.L., Colaço, A., Dominguez-Carrió, C., Escobar, E.G., Harden-Davies, H.R., Drazen, J.C., Ismail, K., Jones, D.O.B., Johnson, D.E., Le, J.T., Lejzerowicz, F., Mitarai, S., Morato, T., Mulsow, S., Snelgrove, P.V.R., Sweetman, A.K., Yasuhara, M., 2020. Climate change considerations are fundamental to management of deep-sea resource extraction. Glob. Chang. Biol. 26, 4664–4678. https://doi.org/10.1111/gcb.15223

Levitus, S., Boyer, T.P., Garcia, H.E., Locarnini, R.A., Zweng, M.M., Mishonov, A. V, Reagan, J.R., Antonov, J.I., Baranova, O.K., Biddle, M., Hamilton, M., Johnson, D.R., Paver, C.R., Seidov, D., 2015. World Ocean Atlas 2013.

Lim, A., Wheeler, A.J., Price, D.M., O'Reilly, L., Harris, K., Conti, L., 2020. Influence of benthic currents on cold-water coral habitats: a combined benthic monitoring and 3D photogrammetric investigation. Sci. Rep. 10, 19433. https://doi.org/10.1038/s41598-020-76446-y

Linley, T.D., Lavaleye, M., Maiorano, P., Bergman, M., Capezzuto, F., Cousins, N.J., D'Onghia, G., Duineveld, G., Shields, M.A., Sion, L., Tursi, A., Priede, I.G., 2017. Effects of cold-water corals on fish diversity and density (European continental margin: Arctic, NE Atlantic and Mediterranean Sea): Data from three baited lander systems. Deep Sea Res. Part II Top. Stud. Oceanogr. 145, 8–21. https://doi.org/10.1016/j.dsr2.2015.12.003

Liu, C., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography (Cop.). 28, 385–393. https://doi.org/10.1111/j.0906-7590.2005.03957.x

Liu, M., Tanhua, T., 2021. Water masses in the Atlantic Ocean: characteristics and distributions. Ocean Sci. 17, 463–486. https://doi.org/10.5194/os-17-463-2021

Longhurst, A., 2007. Ecological Geography of the Sea. Academic Press, London.

Lopes, C.L., Bastos, L., Caetano, M., Martins, I., Santos, M.M., Iglesias, I., 2019. Development of physical modelling tools in support of risk scenarios: A new framework focused on deep-sea mining. Sci. Total Environ. 650, 2294–2306. https://doi.org/10.1016/j.scitotenv.2018.09.351

Lowndes, J.S.S., Best, B.D., Scarborough, C., Afflerbach, J.C., Frazier, M.R., O'Hara, C.C., Jiang, N., Halpern, B.S., 2017. Our path to better science in less time using open data science tools. Nat. Ecol. Evol. 1, 0160. https://doi.org/10.1038/s41559-017-0160

Lundblad, E.R., Wright, D.J., Miller, J., Larkin, E.M., Rinehart, R., Naar, D.F., Donahue, B.T., Anderson, S.M., Battista, T., 2006. A Benthic Terrain Classification Scheme for American Samoa. Mar. Geod. 29, 89–111. https://doi.org/10.1080/01490410600738021

Lutz, M.J., Caldeira, K., Dunbar, R.B., Behrenfeld, M.J., 2007. Seasonal rhythms of net primary production and particulate organic carbon flux to depth describe the efficiency of biological pump in the global ocean. J. Geophys. Res. 112, C10011. https://doi.org/10.1029/2006JC003706

Maldonado, M., Aguilar, R., Bannister, R.J., Bell, J.J., Conway, K.W., Dayton, P.K., Díaz, C., Gutt, J., Kelly, M., Kenchington, E.L.R., Leys, S.P., Pomponi, S.A., Rapp, H.T., Rützler, K., Tendal, O.S., Vacelet, J., Young, C.M., 2015. Sponge Grounds as Key Marine Habitats: A Synthetic Review of Types, Structure, Functional Roles, and Conservation Concerns, in: Marine Animal Forests. Springer International Publishing, Cham, pp. 1–39. https://doi.org/10.1007/978-3-319-17001-5_24-1

Manrique-Vallier, D., 2016. Bayesian population size estimation using Dirichlet process mixtures. Biometrics 72, 1246–1254. https://doi.org/10.1111/biom.12502

Martin, T.G., BURGMAN, M.A., FIDLER, F., KUHNERT, P.M., LOW-CHOY, S., MCBRIDE, M., MENGERSEN, K., 2012. Eliciting Expert Knowledge in Conservation Science. Conserv. Biol. 26, 29–38. https://doi.org/10.1111/j.1523-1739.2011.01806.x

McClain, C.R., Hardy, S.M., 2010. The dynamics of biogeographic ranges in the deep sea. Proc. R. Soc. B Biol. Sci. 277, 3533–3546. https://doi.org/10.1098/rspb.2010.1057

McClain, C.R., Rex, M.A., 2015. Toward a Conceptual Understanding of β-Diversity in the Deep-Sea Benthos. Annu. Rev. Ecol. Evol. Syst. 46, 623–642. https://doi.org/10.1146/annurev-ecolsys-120213-091640

McGarigal, K., Tagil, S., Cushman, S.A., 2009. Surface metrics: an alternative to patch metrics for the quantification of landscape structure. Landsc. Ecol. 24, 433–450. https://doi.org/10.1007/s10980-009-9327-y

McLachlan, G., Peel, D., 2000. Finite mixture models. Wiley, New York, NY.

Menegotto, A., Dambros, C.S., Netto, S.A., 2019. The scale-dependent effect of environmental filters on species turnover and nestedness in an estuarine benthic community. Ecology 100. https://doi.org/10.1002/ecy.2721

Menegotto, A., Rangel, T.F., 2018. Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness. Nat. Commun. 9. https://doi.org/10.1038/s41467-018-07217-7

Menezes, G., Sigler, M., Silva, H., Pinho, M., 2006. Structure and zonation of demersal fish assemblages off the

Azores Archipelago (mid-Atlantic). Mar. Ecol. Prog. Ser. 324, 241–260. https://doi.org/10.3354/meps324241

Mennis, J., Guo, D., 2009. Spatial data mining and geographic knowledge discovery—An introduction. Comput. Environ. Urban Syst. 33, 403–408. https://doi.org/10.1016/j.compenvurbsys.2009.11.001

Middelburg, J.J., Mueller, C.E., Veuger, B., Larsson, A.I., Form, A.U., Oevelen, D. van, 2016. Discovery of symbiotic nitrogen fixation and chemoautotrophy in cold-water corals. Sci. Rep. 5, 17962. https://doi.org/10.1038/srep17962

Mienis, F., Bouma, T., Witbaard, R., van Oevelen, D., Duineveld, G., 2019. Experimental assessment of the effects of coldwater coral patches on water flow. Mar. Ecol. Prog. Ser. 609, 101–117. https://doi.org/10.3354/meps12815

Miller, H.J., Goodchild, M.F., 2015. Data-driven geography. GeoJournal 80, 449–461. https://doi.org/10.1007/s10708-014-9602-6

Miranda, J.M., Luis, J.F., Lourenço, N., 2018. The Tectonic Evolution of the Azores Based on Magnetic Data, in: Active Volcanoes of the World. pp. 89–100. https://doi.org/10.1007/978-3-642-32226-6_6

Mironov, A.N., Krylova, E.M., 2006. Origin of the fauna of the Meteor Seamounts, north-eastern Atlantic, in: Mironov, Alexander N., Southward, A.J., Gebruk, A. V. (Eds.), Biogeography of the North Atlantic Seamounts. KMK Scientific Press, pp. 22–57.

Mitchell, N.C., Stretch, R., Tempera, F., Ligi, M., 2018. Volcanism in the Azores: A marine geophysical perspective. Act. Volcanoes World 101–126. https://doi.org/10.1007/978-3-642-32226-6_7

Mohn, C., Rengstorf, A., White, M., Duineveld, G., Mienis, F., Soetaert, K., Grehan, A., 2014. Linking benthic hydrodynamics and cold-water coral occurrences: A high-resolution model study at three cold-water coral provinces in the NE Atlantic. Prog. Oceanogr. 122, 92–104. https://doi.org/10.1016/j.pocean.2013.12.003

Morato, T., Afonso, P., Menezes, G.M., Santos, R.S., Silva, M.A., 2020a. Editorial: The Azores Marine Ecosystem: An Open Window Into North Atlantic Open Ocean and Deep-Sea Environments. Front. Mar. Sci. 7, 1–3. https://doi.org/10.3389/fmars.2020.601798

Morato, T., Carreiro-Silva, M., Dominguez-Carrió, C., Taranto, G.H., Rodrigues, L., Brito, J., Gomes, S., 2018a. MapGES / ATLAS Project: August 2018 Cruise on board of R/V Arquipélago.

Morato, T., Carreiro-Silva, M., Taranto, G.H., Dominguez-Carrió, C., Ramos, M., Ríos, N., Fauconnet, L., Ocaña Vicente, O., Calado, A., Afonso, A., Ramos, B., Souto, M., Bettencourt, R., 2018b. Blue Azores Program Expedition 2018 On Board The Nrp Gago Coutinho. https://doi.org/10.5281/zenodo.3416898

Morato, T., Dominguez-Carrió, C., Evans, S., Taranto, G.H., Laura, T., De Sousa, L., Mohn, C., Carreiro-Silva, M., 2021a. iMAR: Integrated assessment of the distribution of vulnerable ecosystem along the Mid-Atlantic Ridge in the Azores region.

Morato, T., Dominguez-Carrió, C., Gomes, S., Taranto, G.H., Ramos, M., Fauconnet, L., Rodrigues, L., Carreiro-Silva, M., 2020b. MapGES_2020 Cruise Report: Exploration of Azores deep-sea habitats, summer 2020.

Morato, T., Dominguez-Carrió, C., Gomes, S., Taranto, G.H., Blasco-Ferre, J., Ramos, M., Fauconnet, L., Zarate, C., Carreiro-Silva, M., 2019. MapGES 2019: Summer 2019 cruise on board of N/I Arquipélago.

Morato, T., Dominguez-Carrió, C., Mohn, C., Ocaña Vicente, O., Ramos, M., Rodrigues, L., Sampaio, Í., Taranto, G.H., Fauconnet, L., Tojeira, I., Gonçalves, E.J., Carreiro-Silva, M., 2021b. Dense cold-water coral garden of Paragorgia johnsoni suggests the importance of the Mid-Atlantic Ridge for deep-sea biodiversity. Ecol. Evol.

11, 16426–16433. https://doi.org/10.1002/ece3.8319

Morato, T., González-Irusta, J., Dominguez-Carrió, C., Wei, C., Davies, A., Sweetman, A.K., Taranto, G.H., Beazley, L., García-Alegre, A., Grehan, A., Laffargue, P., Murillo, F.J., Sacau, M., Vaz, S., Kenchington, E., Arnaud-Haond, S., Callery, O., Chimienti, G., Cordes, E., Egilsdottir, H., Freiwald, A., Gasbarro, R., Gutiérrez-Zárate, C., Gianni, M., Gilkinson, K., Wareham Hayes, V.E., Hebbeln, D., Hedges, K., Henry, L., Johnson, D., Koen-Alonso, M., Lirette, C., Mastrototaro, F., Menot, L., Molodtsova, T., Durán Muñoz, P., Orejas, C., Pennino, M.G., Puerta, P., Ragnarsson, S.Á., Ramiro-Sánchez, B., Rice, J., Rivera, J., Roberts, J.M., Ross, S.W., Rueda, J.L., Sampaio, Í., Snelgrove, P., Stirling, D., Treble, M.A., Urra, J., Vad, J., Oevelen, D., Watling, L., Walkusz, W., Wienberg, C., Woillez, M., Levin, L.A., Carreiro-Silva, M., 2020c. Climate-induced changes in the suitable habitat of cold-water corals and commercially important deep-sea fishes in the North Atlantic. Glob. Chang. Biol. gcb.14996. https://doi.org/10.1111/gcb.14996

Morato, T., Kvile, K.Ø., Taranto, G.H., Tempera, F., Narayanaswamy, B.E., Hebbeln, D., Menezes, G.M., Wienberg, C., Santos, R.S., Pitcher, T.J., 2013. Seamount physiography and biology in the north-east Atlantic and Mediterranean Sea. Biogeosciences 10, 3039–3054. https://doi.org/10.5194/bg-10-3039-2013

Morato, T., Taranto, G.H., 2019. Cruise report for 64PE456 onboard of R/V Pelagia Terceira Island 2019 - Hopper tow-cam video footage.

Murillo, F., Kenchington, E., Tompkins, G., Beazley, L., Baker, E., Knudby, A., Walkusz, W., 2018. Sponge assemblages and predicted archetypes in the eastern Canadian Arctic. Mar. Ecol. Prog. Ser. 597, 115–135. https://doi.org/10.3354/meps12589

Myers, A.A., Giller, P.S., 1988. Process, pattern and scale in biogeography, in: Analytical Biogeography. Springer Netherlands, Dordrecht, pp. 3–12. https://doi.org/10.1007/978-94-009-1199-4_1

Nicholson, E., Fulton, E.A., Brooks, T.M., Blanchard, R., Leadley, P., Metzger, J.P., Mokany, K., Stevenson, S., Wintle, B.A., Woolley, S.N.C., Barnes, M., Watson, J.E.M., Ferrier, S., 2019. Scenarios and Models to Support Global Conservation Targets. Trends Ecol. Evol. 34, 57–68. https://doi.org/10.1016/j.tree.2018.10.006

Niedzielski, T., Høines, Å., Shields, M.A., Linley, T.D., Priede, I.G., 2013. A multi-scale investigation into seafloor topography of the northern Mid-Atlantic Ridge based on geographic information system analysis. Deep Sea Res. Part II Top. Stud. Oceanogr. 98, 231–243. https://doi.org/10.1016/j.dsr2.2013.10.006

O'Hara, T.D., Williams, A., Woolley, S.N.C., Nau, A.W., Bax, N.J., 2020. Deep-sea temperate-tropical faunal transition across uniform environmental gradients. Deep Sea Res. Part I Oceanogr. Res. Pap. 161, 103283. https://doi.org/10.1016/j.dsr.2020.103283

Orejas, C., Addamo, A., Alvarez, M., Aparicio, A., Alcoverro, D., Arnaud-Haond, S., Bilan, M., Boavida, J., Cainzos, V., Calderon, R., Cambeiro, P., Castano, M., Fox, A., Gallardo, M., Gori, A., Guitierrez, C., Henry, L.-A., Hermida, M., Jimenez, J.A., Lopez-Jurado, J.L., Lozano, P., Mateo-Ramirez, A., Mateu, G., Matoso, J.L., Mendez, C., Morillas, A., Movilla, J., Olariaga, A., Paredes, M., Pelayo, V., Pineiro, S., Rakka, M., Ramirez, T., Ramos, M., Reis, J., Rivera, J., Romero, A., Rueda, J.L., Salvador, T., Sampaio, I., Sanchez, H., Santiago, R., Serrano, A., Taranto, G., Urra, J., Velez-Belchi, P., Viladrich, N., Zein, M., 2017. Cruise Summary Report - MEDWAVES survey (MEDiterranean out flow WAter and Vulnerable EcosystemS).

Orejas, C., Jiménez, C., 2017. The Builders of the Oceans – Part I: Coral Architecture from the Tropics to the Poles, from the Shallow to the Deep, in: Rossi, S., Bramanti, L., Gori, A., Orejas, C. (Eds.), Marine Animal Forests. Springer International Publishing, Cham, pp. 627–655. https://doi.org/10.1007/978-3-319-21012-4_10

OSPAR, 2010. Background Document for Coral Gardens, Biodiversity Series,.

Palma, C., Lillebø, A.I., Borges, C., Souto, M., Pereira, E., Duarte, A.C., Abreu, M.P. de, 2012. Water column

characterisation on the Azores platform and at the sea mounts south of the archipelago. Mar. Pollut. Bull. 64, 1884–1894. https://doi.org/10.1016/j.marpolbul.2012.06.015

Paradis, E., Schliep, K., 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Parrish, F.A., Oliver, T.A., 2020. Comparative Observations of Current Flow, Tidal Spectra, and Scattering Strength in and Around Hawaiian Deep-Sea Coral Patches. Front. Mar. Sci. 7, 1–16. https://doi.org/10.3389/fmars.2020.00310

Peran, A.D., Pham, C.K., Amorim, P., Cardigos, F., Tempera, F., Morato, T., 2016. Seafloor Characteristics in the Azores Region (North Atlantic). Front. Mar. Sci. 3, 2014–2017. https://doi.org/10.3389/fmars.2016.00204

Peuquet, D.J., 1984. A CONCEPTUAL FRAMEWORK AND COMPARISON OF SPATIAL DATA MODELS. Cartogr. Int. J. Geogr. Inf. Geovisualization 21, 66–113. https://doi.org/10.3138/D794-N214-221R-23R5

Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E., Blair, M.E., 2017. Opening the black box: an open-source release of Maxent. Ecography (Cop.). 40, 887–893. https://doi.org/10.1111/ecog.03049

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography (Cop.). 31, 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

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. Ecol. Appl. 19, 181–197. https://doi.org/10.1890/07-2153.1

Pierdomenico, M., Martorelli, E., Dominguez-Carrió, C., Gili, J.M., Chiocci, F.L., 2016. Seafloor characterization and benthic megafaunal distribution of an active submarine canyon and surrounding sectors: The case of Gioia Canyon (Southern Tyrrhenian Sea). J. Mar. Syst. 157, 101–117. https://doi.org/10.1016/j.jmarsys.2016.01.005

Pike, R.J., Evans, I.S., Hengl, T., 2009. Geomorphometry: A Brief Guide, in: Hengl, T., Reuter, H.I. (Eds.), Geomorphometry Concepts, Software, Applications. Developments in Soil Science, vol. 33, Elsevier, pp. 3–30. https://doi.org/10.1016/S0166-2481(08)00001-9

Pittman, S., Yates, K., Bouchet, P., Alvarez-Berastegui, D., Andréfouët, S., Bell, S., Berkström, C., Boström, C., Brown, C., Connolly, R., Devillers, R., Eggleston, D., Gilby, B., Gullström, M., Halpern, B., Hidalgo, M., Holstein, D., Hovel, K., Huettmann, F., Jackson, E., James, W., Kellner, J., Kot, C., Lecours, V., Lepczyk, C., Nagelkerken, I., Nelson, J., Olds, A., Santos, R., Scales, K., Schneider, D., Schilling, H., Simenstad, C., Suthers, I., Treml, E., Wedding, L., Yates, P., Young, M., 2021. Seascape ecology: identifying research priorities for an emerging ocean sustainability science. Mar. Ecol. Prog. Ser. 663, 1–29. https://doi.org/10.3354/meps13661

Portilho-Ramos, R. da C., Titschack, J., Wienberg, C., Siccha Rojas, M.G., Yokoyama, Y., Hebbeln, D., 2022. Major environmental drivers determining life and death of cold-water corals through time. PLOS Biol. 20, e3001628. https://doi.org/10.1371/journal.pbio.3001628

Puerta, P., Johnson, C., Carreiro-Silva, M., Henry, L.-A., Kenchington, E., Morato, T., Kazanidis, G., Rueda, J.L., Urra, J., Ross, S., Wei, C.-L., González-Irusta, J.M., Arnaud-Haond, S., Orejas, C., 2020. Influence of Water Masses on the Biodiversity and Biogeography of Deep-Sea Benthic Ecosystems in the North Atlantic. Front. Mar. Sci. 7, 1–25. https://doi.org/10.3389/fmars.2020.00239

Puerta, P., Mosquera-Giménez, Á., Reñones, O., Domínguez-Carrió, C., Rueda, J.L., Urra, J., Carreiro-Silva, M., Blasco-Ferre, J., Santana, Y., Gutiérrez-Zárate, C., Vélez-Belchí, P., Rivera, J., Morato, T., Orejas, C., 2022. Variability of deep-sea megabenthic assemblages along the western pathway of the Mediterranean outflow water. Deep Sea Res. Part I Oceanogr. Res. Pap. 185, 103791. https://doi.org/10.1016/j.dsr.2022.103791

Quattrini, A.M., Gómez, C.E., Cordes, E.E., 2017. Environmental filtering and neutral processes shape octocoral community assembly in the deep sea. Oecologia 183, 221–236. https://doi.org/10.1007/s00442-016-3765-4

R Core Team, 2022. R: A language and environment for statistical computing.

Raddatz, J., Titschack, J., Frank, N., Freiwald, A., Conforti, A., Osborne, A., Skornitzke, S., Stiller, W., Rüggeberg, A., Voigt, S., Albuquerque, A.L.S., Vertino, A., Schröder-Ritzrau, A., Bahr, A., 2020. Solenosmilia variabilis-bearing cold-water coral mounds off Brazil. Coral Reefs, Coral Reefs of the World 39, 69–83. https://doi.org/10.1007/s00338-019-01882-w

Radice, V.Z., Quattrini, A.M., Wareham, V.E., Edinger, E.N., Cordes, E.E., 2016. Vertical water mass structure in the North Atlantic influences the bathymetric distribution of species in the deep-sea coral genus Paramuricea. Deep Sea Res. Part I Oceanogr. Res. Pap. 116, 253–263. https://doi.org/10.1016/j.dsr.2016.08.014

Ragnarsson, S.Á., Burgos, J.M., Kutti, T., Beld, I. Van Den, Egilsdóttir, H., Arnaud-haond, S., Grehan, A., 2017. The impact of anthropogenic activity on cold-water corals, Marine Animal Forests. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-17001-5

Rakka, M., Godinho, A., Orejas, C., Carreiro-Silva, M., 2021. Embryo and larval biology of the deep-sea octocoral Dentomuricea aff. meteor under different temperature regimes. PeerJ 9, e11604. https://doi.org/10.7717/peerj.11604

Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C.R., Levin, L.A., Martinez Arbizu, P., Menot, L., Buhl-Mortensen, P., Narayanaswamy, B.E., Smith, C.R., Tittensor, D.P., Tyler, P.A., Vanreusel, A., Vecchione, M., 2010. Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. Biogeosciences 7, 2851–2899. https://doi.org/10.5194/bg-7-2851-2010

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., Maury, O., 2013. Dynamic biogeochemical provinces in the global ocean. Global Biogeochem. Cycles 27, 1046–1058. https://doi.org/10.1002/gbc.20089

Ribeiro, M.C., Ferreira, R., Pereira, E., Soares, J., 2020. Scientific, technical and legal challenges of deep sea mining. A vision for Portugal – Conference report. Mar. Policy 114, 103338. https://doi.org/10.1016/j.marpol.2018.11.001

Rice, J., Arvanitidis, C., Borja, A., Frid, C., Hiddink, J.G., Krause, J., Lorance, P., Ragnarsson, S.Á., Sköld, M., Trabucco, B., Enserink, L., Norkko, A., 2012. Indicators for Sea-floor Integrity under the European Marine Strategy Framework Directive. Ecol. Indic. 12, 174–184. https://doi.org/10.1016/j.ecolind.2011.03.021

Rice, J., Gjerde, K.M., Ardron, J., Arico, S., Cresswell, I., Escobar, E., Grant, S., Vierros, M., 2011. Policy relevance of biogeographic classification for conservation and management of marine biodiversity beyond national jurisdiction, and the GOODS biogeographic classification. Ocean Coast. Manag. 54, 110–122. https://doi.org/10.1016/j.ocecoaman.2010.10.010

Riehl, T., Wölfl, A.-C., Augustin, N., Devey, C.W., Brandt, A., 2020. Discovery of widely available abyssal rock patches reveals overlooked habitat type and prompts rethinking deep-sea biodiversity. Proc. Natl. Acad. Sci. 117, 15450–15459. https://doi.org/10.1073/pnas.1920706117

Riou, V., Fonseca-Batista, D., Roukaerts, A., Biegala, I.C., Prakya, S.R., Magalhães Loureiro, C., Santos, M., Muniz-Piniella, A.E., Schmiing, M., Elskens, M., Brion, N., Martins, M.A., Dehairs, F., 2016. Importance of N2-Fixation on the Productivity at the North-Western Azores Current/Front System, and the Abundance of Diazotrophic

Unicellular Cyanobacteria. PLoS One 11, e0150827. https://doi.org/10.1371/journal.pone.0150827

Rix, L., De Goeij, J.M., Mueller, C.E., Struck, U., Middelburg, J.J., Van Duyl, F.C., Al-Horani, F.A., Wild, C., Naumann, M.S., Van Oevelen, D., 2016. Coral mucus fuels the sponge loop in warm-and cold-water coral reef ecosystems. Sci. Rep. 6, 1–11. https://doi.org/10.1038/srep18715

Roberts, E., Bowers, D., Meyer, H., Samuelsen, A., Rapp, H., Cárdenas, P., 2021. Water masses constrain the distribution of deep-sea sponges in the North Atlantic Ocean and Nordic Seas. Mar. Ecol. Prog. Ser. 659, 75–96. https://doi.org/10.3354/meps13570

Roberts, J.M., Cairns, S.D., 2014. Cold-water corals in a changing ocean. Curr. Opin. Environ. Sustain. 7, 118–126. https://doi.org/10.1016/j.cosust.2014.01.004

Roberts, J.M., Wheeler, A., Freiwald, A., Cairns, S.D., 2009. Cold-Water Corals. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511581588

Rogers, A.D., 2018. The Biology of Seamounts: 25 Years on, in: Advances in Marine Biology. pp. 137–224. https://doi.org/10.1016/bs.amb.2018.06.001

Rogers, A.D., 2000. The role of the oceanic oxygen minima in generating biodiversity in the deep sea. Deep Sea Res. Part II Top. Stud. Oceanogr. 47, 119–148. https://doi.org/10.1016/S0967-0645(99)00107-1

Rovelli, L., Attard, K., Bryant, L., Flögel, S., Stahl, H., Roberts, J.M., Linke, P., Glud, R., 2015. Benthic O2 uptake of two cold-water coral communities estimated with the non-invasive eddy correlation technique. Mar. Ecol. Prog. Ser. 525, 97–104. https://doi.org/10.3354/meps11211

Rovelli, L., Carreiro-Silva, M., Attard, K., Rakka, M., Dominguez-Carrió, C., Bilan, M., Blackbird, S., Morato, T., Wolff, G., Glud, R., 2022. Benthic O2 uptake by coral gardens at the Condor seamount (Azores). Mar. Ecol. Prog. Ser. 688, 19–31. https://doi.org/10.3354/meps14021

Rueda, J.L., González-García, E., Krutzky, C., López-Rodriguez, F.J., Bruque, G., López-González, N., Palomino, D., Sánchez, R.F., Vázquez, J.T., Fernández-Salas, L.M., Díaz-del-Río, V., 2016. From chemosynthesis-based communities to cold-water corals: Vulnerable deep-sea habitats of the Gulf of Cádiz. Mar. Biodivers. 46, 473–482. https://doi.org/10.1007/s12526-015-0366-0

Saeedi, H., Costello, M.J., Warren, D., Brandt, A., 2019. Latitudinal and bathymetrical species richness patterns in the NW Pacific and adjacent Arctic Ocean. Sci. Rep. 9, 9303. https://doi.org/10.1038/s41598-019-45813-9

Sala, I., Harrison, C.S., Caldeira, R.M.A., 2016. The role of the Azores Archipelago in capturing and retaining incoming particles. J. Mar. Syst. 154, 146–156. https://doi.org/10.1016/j.jmarsys.2015.10.001

Salazar, G., Cornejo-Castillo, F.M., Benítez-Barrios, V., Fraile-Nuez, E., Álvarez-Salgado, X.A., Duarte, C.M., Gasol, J.M., Acinas, S.G., 2016. Global diversity and biogeography of deep-sea pelagic prokaryotes. ISME J. 10, 596–608. https://doi.org/10.1038/ismej.2015.137

Sampaio, Í., Braga-Henriques, A., Pham, C., Ocaña, O., de Matos, V., Morato, T., Porteiro, F.M., 2012. Cold-water corals landed by bottom longline fisheries in the Azores (north-eastern Atlantic). J. Mar. Biol. Assoc. United Kingdom 92, 1547–1555. https://doi.org/10.1017/S0025315412000045

Sampaio, Í., Freiwald, A., Porteiro, F.M., Menezes, G., Carreiro-Silva, M., 2019a. Census of octocorallia (cnidaria: Anthozoa) of the azores (NE atlantic) with a nomenclature update. Zootaxa 4550, 451–498. https://doi.org/10.11646/zootaxa.4550.4.1

Sampaio, Í., Morato, T., Porteiro, F., Gutiérrez-Zárate, C., Taranto, G., Pham, C., Gonçalves, J., Carreiro-Silva, M.,

2019b. The Value of a Deep-Sea Collection of the Azores (NE Atlantic Ocean): Marine invertebrate biodiversity in an era of global environmental change. Biodivers. Inf. Sci. Stand. 3. https://doi.org/10.3897/biss.3.37209

Sanmartín, I., 2012. Historical Biogeography: Evolution in Time and Space. Evol. Educ. Outreach 5, 555–568. https://doi.org/10.1007/s12052-012-0421-2

Santos, R., Medeiros-Leal, W., Novoa-Pabon, A., Silva, H., Pinho, M., 2021. Demersal fish assemblages on seamounts exploited by fishing in the Azores (NE Atlantic). J. Appl. Ichthyol. 37, 198–215. https://doi.org/10.1111/jai.14165

Santos, R.S., Hawkins, S., Monteiro, L.R., Alves, M., Isidro, E.J., 1995. Marine research, resources and conservation in the Azores. Aquat. Conserv. Mar. Freshw. Ecosyst. 5, 311–354. https://doi.org/10.1002/aqc.3270050406

Sayre, R.G., Wright, D.J., Breyer, S.P., Butler, K.A., Van Graafeiland, K., Costello, M.J., Harris, P.T., Goodin, K.L., Guinotte, J.M., Basher, Z., Kavanaugh, M.T., Halpin, P.N., Monaco, M.E., Cressie, N.A., Aniello, P., Frye, C.E., Stephens, D., 2017. A three-dimensional mapping of the ocean based on environmental data. Oceanography 30, 90–103. https://doi.org/10.5670/oceanog.2017.116

Schiebel, R., Brupbacher, U., Schmidtko, S., Nausch, G., Waniek, J.J., Thierstein, H.-R., 2011. Spring coccolithophore production and dispersion in the temperate eastern North Atlantic Ocean. J. Geophys. Res. 116, C08030. https://doi.org/10.1029/2010JC006841

Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C., 2001. Estimating the Support of a High-Dimensional Distribution. Neural Comput. 13, 1443–1471. https://doi.org/10.1162/089976601750264965

Schuurman, N., 2006. Formalization matters: Critical GIS and ontology research. Ann. Assoc. Am. Geogr. 96, 726–739. https://doi.org/10.1111/j.1467-8306.2006.00513.x

Schuurman, N., Leszczynski, A., 2006. Ontology-Based Metadata. Trans. GIS 10, 709–726. https://doi.org/10.1111/j.1467-9671.2006.01024.x

Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Stat. 6, 461–464. https://doi.org/10.1214/aos/1176344136

Seijmonsbergen, A.C., Hengl, T., Anders, N.S., 2011. Semi-Automated Identification and Extraction of Geomorphological Features Using Digital Elevation Data, in: Developments in Earth Surface Processes. Elsevier B.V., pp. 297–335. https://doi.org/10.1016/B978-0-444-53446-0.00010-0

Sethuraman, J., 1994. A constructive definition of Dirichlet priors. Stat. Sin.

Sheehan, E. V., Vaz, S., Pettifer, E., Foster, N.L., Nancollas, S.J., Cousens, S., Holmes, L., Facq, J., Germain, G., Attrill, M.J., 2016. An experimental comparison of three towed underwater video systems using species metrics, benthic impact and performance. Methods Ecol. Evol. 7, 843–852. https://doi.org/10.1111/2041-210X.12540

Silva, A., Brotas, V., Valente, A., Sá, C., Diniz, T., Patarra, R.F., Álvaro, N.V., Neto, A.I., 2013. Coccolithophore species as indicators of surface oceanographic conditions in the vicinity of Azores islands. Estuar. Coast. Shelf Sci. 118, 50–59. https://doi.org/10.1016/j.ecss.2012.12.010

Smith, A.C., Dahlin, K.M., Record, S., Costanza, J.K., Wilson, A.M., Zarnetske, P.L., 2021. The <scp>geodiv r</scp> package: Tools for calculating gradient surface metrics. Methods Ecol. Evol. 12, 2094–2100. https://doi.org/10.1111/2041-210X.13677

Smith, B., Varzi, A.C., 2000. Fiat and Bona Fide Boundaries. Philos. Phenomenol. Res. 60, 401. https://doi.org/10.2307/2653492

Smith, C.R., Tunnicliffe, V., Colaço, A., Drazen, J.C., Gollner, S., Levin, L.A., Mestre, N.C., Metaxas, A., Molodtsova, T.N., Morato, T., Sweetman, A.K., Washburn, T., Amon, D.J., 2020. Deep-Sea Misconceptions Cause Underestimation of Seabed-Mining Impacts. Trends Ecol. Evol. 35, 853–857. https://doi.org/10.1016/j.tree.2020.07.002

Soetaert, K., Mohn, C., Rengstorf, A., Grehan, A., van Oevelen, D., 2016. Ecosystem engineering creates a direct nutritional link between 600-m deep cold-water coral mounds and surface productivity. Sci. Rep. 6, 35057. https://doi.org/10.1038/srep35057

Sofia, G., 2020. Combining geomorphometry, feature extraction techniques and Earth-surface processes research: The way forward. Geomorphology 355, 107055. https://doi.org/10.1016/j.geomorph.2020.107055

Søiland, H., Budgell, W.P., Knutsen, Ø., 2008. The physical oceanographic conditions along the Mid-Atlantic Ridge north of the Azores in June–July 2004. Deep Sea Res. Part II Top. Stud. Oceanogr. 55, 29–44. https://doi.org/10.1016/j.dsr2.2007.09.015

Soininen, J., 2014. A quantitative analysis of species sorting across organisms and ecosystems. Ecology 95, 3284–3292. https://doi.org/10.1890/13-2228.1

Somoza, L., Medialdea, T., González, F.J., Calado, A., Afonso, A., Albuquerque, M., Asensio-Ramos, M., Bettencourt, R., Blasco, I., Candón, J.A., Carreiro-Silva, M., Cid, C., De Ignacio, C., López-Pamo, E., Machancoses, S., Ramos, B., Ribeiro, L.P., Rincón-Tomás, B., Santofimia, E., Souto, M., Tojeira, I., Viegas, C., Madureira, P., 2020. Multidisciplinary Scientific Cruise to the Northern Mid-Atlantic Ridge and Azores Archipelago. Front. Mar. Sci. 7, 1–9. https://doi.org/10.3389/fmars.2020.568035

Sonnewald, M., Sonnewald, M., Dutkiewicz, S., Hill, C., Forget, G., 2020. Elucidating ecological complexity: Unsupervised learning determines global marine eco-provinces. Sci. Adv. 6, 1–12. https://doi.org/10.1126/sciadv.aay4740

Stevens, T., 2002. Rigor and Representativeness in Marine Protected Area Design. Coast. Manag. 30, 237–248. https://doi.org/10.1080/08920750290042183

Stuart, C.T., Brault, S., Rowe, G.T., Wei, C., Wagstaff, M., McClain, C.R., Rex, M.A., 2017. Nestedness and species replacement along bathymetric gradients in the deep sea reflect productivity: a test with polychaete assemblages in the oligotrophic north-west Gulf of Mexico. J. Biogeogr. 44, 548–555. https://doi.org/10.1111/jbi.12810

Sun, K., Zhu, Y., Pan, P., Hou, Z., Wang, D., Li, W., Song, J., 2019. Geospatial data ontology: the semantic foundation of geospatial data integration and sharing. Big Earth Data 3, 269–296. https://doi.org/10.1080/20964471.2019.1661662

Swanborn, D.J.B., Huvenne, V.A.I., Pittman, S.J., Woodall, L.C., 2022. Bringing seascape ecology to the deep seabed: A review and framework for its application. Limnol. Oceanogr. 67, 66–88. https://doi.org/10.1002/lno.11976

Sweetman, A.K., Thurber, A.R., Smith, C.R., Levin, L.A., Mora, C., Wei, C.-L., Gooday, A.J., Jones, D.O.B., Rex, M., Yasuhara, M., Ingels, J., Ruhl, H.A., Frieder, C.A., Danovaro, R., Würzberg, L., Baco, A.R., Grupe, B.M., Pasulka, A., Meyer, K.S., Dunlop, K.M., Henry, L.-A., Roberts, J.M., 2017. Major impacts of climate change on deep-sea benthic ecosystems. Elem Sci Anth 5, 4. https://doi.org/10.1525/elementa.203

Symonds, M.R.E., Moussalli, A., 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behav. Ecol. Sociobiol. 65, 13–21. https://doi.org/10.1007/s00265-010-1037-6

Tambutté, S., Holcomb, M., Ferrier-Pagès, C., Reynaud, S., Tambutté, É., Zoccola, D., Allemand, D., 2011. Coral

biomineralization: From the gene to the environment. J. Exp. Mar. Bio. Ecol. 408, 58–78. https://doi.org/10.1016/j.jembe.2011.07.026

Taranto, G.H., 2022. scapesClassification: User-Defined Classification of Raster Surfaces. R package version 1.0.0.

Tempera, F., Giacomello, E., Mitchell, N.C., Campos, A.S., Braga Henriques, A., Bashmachnikov, I., Martins, A., Mendonça, A., Morato, T., Colaço, A., Porteiro, F.M., Catarino, D., Gonçalves, J., Pinho, M.R., Isidro, E.J., Santos, R.S., Menezes, G., 2012a. Mapping Condor Seamount Seafloor Environment and Associated Biological Assemblages (Azores, NE Atlantic), in: Seafloor Geomorphology as Benthic Habitat. Elsevier, pp. 807–818. https://doi.org/10.1016/B978-0-12-385140-6.00059-1

Tempera, F., Pereira, J., Braga-Henriques, A., Porteiro, F.M., Morato, T., de Matos, V., Souto, M., Guillaumont, B., Santos, R.S., 2012b. Cataloguing deep-sea biological facies of the Azores. Rev. Investig. Mar. 19, 36–38.

Thompson, P.L., Guzman, L.M., De Meester, L., Horváth, Z., Ptacnik, R., Vanschoenwinkel, B., Viana, D.S., Chase, J.M., 2020. A process-based metacommunity framework linking local and regional scale community ecology. Ecol. Lett. 23, 1314–1329. https://doi.org/10.1111/ele.13568

Thuiller, W., Georges, D., Gueguen, M., Engler, R., Breiner, F., Lafourcade, B., 2022. biomod2: Ensemble Platform for Species Distribution Modeling. R package version 4.0.

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD - a platform for ensemble forecasting of species distributions. Ecography (Cop.). 32, 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x

Thurber, A.R., Sweetman, A.K., Narayanaswamy, B.E., Jones, D.O.B., Ingels, J., Hansman, R.L., 2014. Ecosystem function and services provided by the deep sea. Biogeosciences 11, 3941–3963. https://doi.org/10.5194/bg-11-3941-2014

Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic. J. R. Stat. Soc. Ser. B (Statistical Methodol. 63, 411–423. https://doi.org/10.1111/1467-9868.00293

Townsend, M., Davies, K., Hanley, N., Hewitt, J.E., Lundquist, C.J., Lohrer, A.M., 2018. The Challenge of Implementing the Marine Ecosystem Service Concept. Front. Mar. Sci. 5, 1–13. https://doi.org/10.3389/fmars.2018.00359

Tucholke, B.E., Smoot, N.C., 1990. Evidence for age and evolution of Corner Seamounts and Great Meteor Seamount Chain from multibeam bathymetry. J. Geophys. Res. 95, 17555. https://doi.org/10.1029/JB095iB11p17555

Tunnicliffe, V., Metaxas, A., Le, J., Ramirez-Llodra, E., Levin, L.A., 2020. Strategic Environmental Goals and Objectives: Setting the basis for environmental regulation of deep seabed mining. Mar. Policy 114, 103347. https://doi.org/10.1016/j.marpol.2018.11.010

Turner, M.G., 2005. Landscape Ecology: What Is the State of the Science? Annu. Rev. Ecol. Evol. Syst. 36, 319–344. https://doi.org/10.1146/annurev.ecolsys.36.102003.152614

Ulrich, W., Almeida-Neto, M., 2012. On the meanings of nestedness: back to the basics. Ecography (Cop.). 35, 865–871. https://doi.org/10.1111/j.1600-0587.2012.07671.x

United Nations, 2018. Revised Roadmap for the UN Decade of Ocean Science for Sustainable Development.

Valavi, R., Elith, J., Lahoz-Monfort, J.J., Guillera-Arroita, G., 2019. blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Methods Ecol. Evol. 10, 225–232. https://doi.org/10.1111/2041-210X.13107

Valle, D., Jameel, Y., Betancourt, B., Azeria, E.T., Attias, N., Cullen, J., 2022. Automatic selection of the number of clusters using Bayesian clustering and sparsity-inducing priors. Ecol. Appl. 32, 1–19.

https://doi.org/10.1002/eap.2524

Van Dover, C.L., Ardron, J.A., Escobar, E., Gianni, M., Gjerde, K.M., Jaeckel, A., Jones, D.O.B., Levin, L.A., Niner, H.J., Pendleton, L., Smith, C.R., Thiele, T., Turner, P.J., Watling, L., Weaver, P.P.E., 2017. Biodiversity loss from deep-sea mining. Nat. Geosci. 10, 464–465. https://doi.org/10.1038/ngeo2983

van Proosdij, A.S.J., Sosef, M.S.M., Wieringa, J.J., Raes, N., 2016. Minimum required number of specimen records to develop accurate species distribution models. Ecography (Cop.). 39, 542–552. https://doi.org/10.1111/ecog.01509

Vellend, M., 2010. Conceptual Synthesis in Community Ecology. Q. Rev. Biol. 85, 183–206. https://doi.org/10.1086/652373

Viana, D.S., Chase, J.M., 2019. Spatial scale modulates the inference of metacommunity assembly processes. Ecology 100, e02576. https://doi.org/10.1002/ecy.2576

Victorero, L., Robert, K., Robinson, L.F., Taylor, M.L., Huvenne, V.A.I., 2018. Species replacement dominates megabenthos beta diversity in a remote seamount setting. Sci. Rep. 8, 4152. https://doi.org/10.1038/s41598-018-22296-8

Vieira, R.P., Bett, B.J., Jones, D.O.B., Durden, J.M., Morris, K.J., Cunha, M.R., Trueman, C.N., Ruhl, H.A., 2020. Deep-sea sponge aggregations (Pheronema carpenteri) in the Porcupine Seabight (NE Atlantic) potentially degraded by demersal fishing. Prog. Oceanogr. 183, 102189. https://doi.org/10.1016/j.pocean.2019.102189

Walbridge, S., Slocum, N., Pobuda, M., Wright, D., 2018. Unified Geomorphological Analysis Workflows with Benthic Terrain Modeler. Geosciences 8, 94. https://doi.org/10.3390/geosciences8030094

Warren, D.L., Glor, R.E., Turelli, M., 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution (N. Y). 62, 2868–2883. https://doi.org/10.1111/j.1558-5646.2008.00482.x

Warren, D.L., Matzke, N.J., Iglesias, T.L., 2020. Evaluating presence-only species distribution models with discrimination accuracy is uninformative for many applications. J. Biogeogr. 47, 167–180. https://doi.org/10.1111/jbi.13705

Warton, D.I., Foster, S.D., De'ath, G., Stoklosa, J., Dunstan, P.K., 2015. Model-based thinking for community ecology. Plant Ecol. 216, 669–682. https://doi.org/10.1007/s11258-014-0366-3

Watling, L., Guinotte, J., Clark, M.R., Smith, C.R., 2013. A proposed biogeography of the deep ocean floor. Prog. Oceanogr. 111, 91–112. https://doi.org/10.1016/j.pocean.2012.11.003

Watling, L., Lapointe, A., 2022. Global biogeography of the lower bathyal (700–3000 m) as determined from the distributions of cnidarian anthozoans. Deep Sea Res. Part I Oceanogr. Res. Pap. 181, 103703. https://doi.org/10.1016/j.dsr.2022.103703

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M., Willis, K.J., 2005. Conservation Biogeography: assessment and prospect. Divers. Distrib. 11, 3–23. https://doi.org/10.1111/j.1366-9516.2005.00143.x

Wickham, H., 2015. R Packages: Organize, Test, Document, and Share Your Code. O'Reilly Media, Inc. , Boston, Massachusetts, USA.

Wienberg, C., Wintersteller, P., Beuck, L., Hebbeln, D., 2013. Coral Patch seamount (NE Atlantic) – a sedimentological and megafaunal reconnaissance based on video and hydroacoustic surveys. Biogeosciences 10, 3421–3443. https://doi.org/10.5194/bg-10-3421-2013

Williams, A., Althaus, F., Dunstan, P.K., Poore, G.C.B., Bax, N.J., Kloser, R.J., McEnnulty, F.R., 2010. Scales of habitat heterogeneity and megabenthos biodiversity on an extensive Australian continental margin (100-1100 m depths). Mar. Ecol. 31, 222–236. https://doi.org/10.1111/j.1439-0485.2009.00355.x

Wölfl, A.-C., Snaith, H., Amirebrahimi, S., Devey, C.W., Dorschel, B., Ferrini, V., Huvenne, V.A.I., Jakobsson, M., Jencks, J., Johnston, G., Lamarche, G., Mayer, L., Millar, D., Pedersen, T.H., Picard, K., Reitz, A., Schmitt, T., Visbeck, M., Weatherall, P., Wigley, R., 2019. Seafloor Mapping – The Challenge of a Truly Global Ocean Bathymetry. Front. Mar. Sci. 6, 1–16. https://doi.org/10.3389/fmars.2019.00283

Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Woolley, S.N.C., Foster, S.D., Bax, N.J., Currie, J.C., Dunn, D.C., Hansen, C., Hill, N., O'Hara, T.D., Ovaskainen, O., Sayre, R., Vanhatalo, J.P., Dunstan, P.K., 2020. Bioregions in Marine Environments: Combining Biological and Environmental Data for Management and Scientific Understanding. Bioscience 70, 48–59. https://doi.org/10.1093/biosci/biz133

Woolley, S.N.C., McCallum, A.W., Wilson, R., O'Hara, T.D., Dunstan, P.K., 2013. Fathom out: biogeographical subdivision across the Western Australian continental margin - a multispecies modelling approach. Divers. Distrib. 19, 1506–1517. https://doi.org/10.1111/ddi.12119

Woolley, S.N.C., Tittensor, D.P., Dunstan, P.K., Guillera-Arroita, G., Lahoz-Monfort, J.J., Wintle, B.A., Worm, B., O'Hara, T.D., 2016. Deep-sea diversity patterns are shaped by energy availability. Nature 533, 393–396. https://doi.org/10.1038/nature17937

Xavier, J.R., Rees, D.J., Pereira, R., Colaço, A., Pham, C.K., Carvalho, F.C., 2021. Diversity, Distribution and Phylogenetic Relationships of Deep-Sea Lithistids (Porifera, Heteroscleromorpha) of the Azores Archipelago. Front. Mar. Sci. 8, 1–20. https://doi.org/10.3389/fmars.2021.600087

Yang, H., Lohmann, G., Krebs-Kanzow, U., Ionita, M., Shi, X., Sidorenko, D., Gong, X., Chen, X., Gowan, E.J., 2020. Poleward Shift of the Major Ocean Gyres Detected in a Warming Climate. Geophys. Res. Lett. 47. https://doi.org/10.1029/2019GL085868

Yearsley, J.M., Salmanidou, D.M., Carlsson, J., Burns, D., Van Dover, C.L., 2020. Biophysical models of persistent connectivity and barriers on the northern Mid-Atlantic Ridge. Deep Sea Res. Part II Top. Stud. Oceanogr. 180, 104819. https://doi.org/10.1016/j.dsr2.2020.104819

Yesson, C., Taylor, M.L., Tittensor, D.P., Davies, A.J., Guinotte, J., Baco, A.R., Black, J., Hall-Spencer, J.M., Rogers, A.D., 2012. Global habitat suitability of cold-water octocorals. J. Biogeogr. 39, 1278–1292. https://doi.org/10.1111/j.1365-2699.2011.02681.x

Zarnetske, P.L., Read, Q.D., Record, S., Gaddis, K.D., Pau, S., Hobi, M.L., Malone, S.L., Costanza, J., M. Dahlin, K., Latimer, A.M., Wilson, A.M., Grady, J.M., Ollinger, S. V., Finley, A.O., 2019. Towards connecting biodiversity and geodiversity across scales with satellite remote sensing. Glob. Ecol. Biogeogr. 28, 548–556. https://doi.org/10.1111/geb.12887

Zelinsky-Wibbelt, C., 1990. The semantic representation of spatial configurations, in: Proceedings of the 13th Conference on Computational Linguistics. Association for Computational Linguistics, Morristown, NJ, USA, pp. 299–304. https://doi.org/10.3115/991146.991199

Zibrowius, H., Cairns, S.D., 1992. Revision of the northeast Atlantic and Mediterranean Stylasteridae (Cnidaria: Hydrozoa).

List of Figures

List of Tables

Appendix A
Supplementary data Chapter 2

A.1 Species selection and input data

A.1.1 Regional water masses

The water masses of the Atlantic Ocean have been recently described by Liu and Tanhua (2021) using conservative (temperature and salinity) and non-conservative (oxygen, silicate, phosphate and nitrate) measures. Based on their results, it is possible to identify the principal water masses occurring in the Azores (Appendix A.1.1, Figure A.1) and approximate the limits of the different water layers. The influence of upper water masses goes down to about 1200m (Figure A.1a-c). The Eastern North Atlantic Central Water (ENACW) has the widest vertical extent with its core at about 400 m depth. The Western North Atlantic Central Water (WNACW) and the Western South Atlantic Central Water (WSACW) are present within the upper 400 m of the water column. The intermediate layer is characterized by the Mediterranean Water (MW) and the Antarctic Intermediate Water (AAIW) (Figure A.1d-e). Their core is around 800 m depth. A third intermediate water mass, the Subarctic Intermediate Water (SAIW) has been reported in the northwestern portion of the study area (Bashmachnikov et al., 2015), however, given the lack of sampling stations in this area (Figure A.1f) it was almost absent in the data provided by Liu and Tanhua (2021). In the deep and overflow layer, down to 2000 m (Figure A.1g-i), the Upper North Atlantic Deep Water (uNADW) is the predominant water mass with its core located at about 1600 m depth. It occupies a wide depth range in the north (400-2000+ m), while in the south it only appears at depths below 1000 m. The Lower North Atlantic Deep Water (lNADW) and the Northeast Atlantic Bottom Water (NEABW) characterize the deepest portion of the considered depth range (0-2000m) and extend below 2000 m, where they become the most abundant water masses.

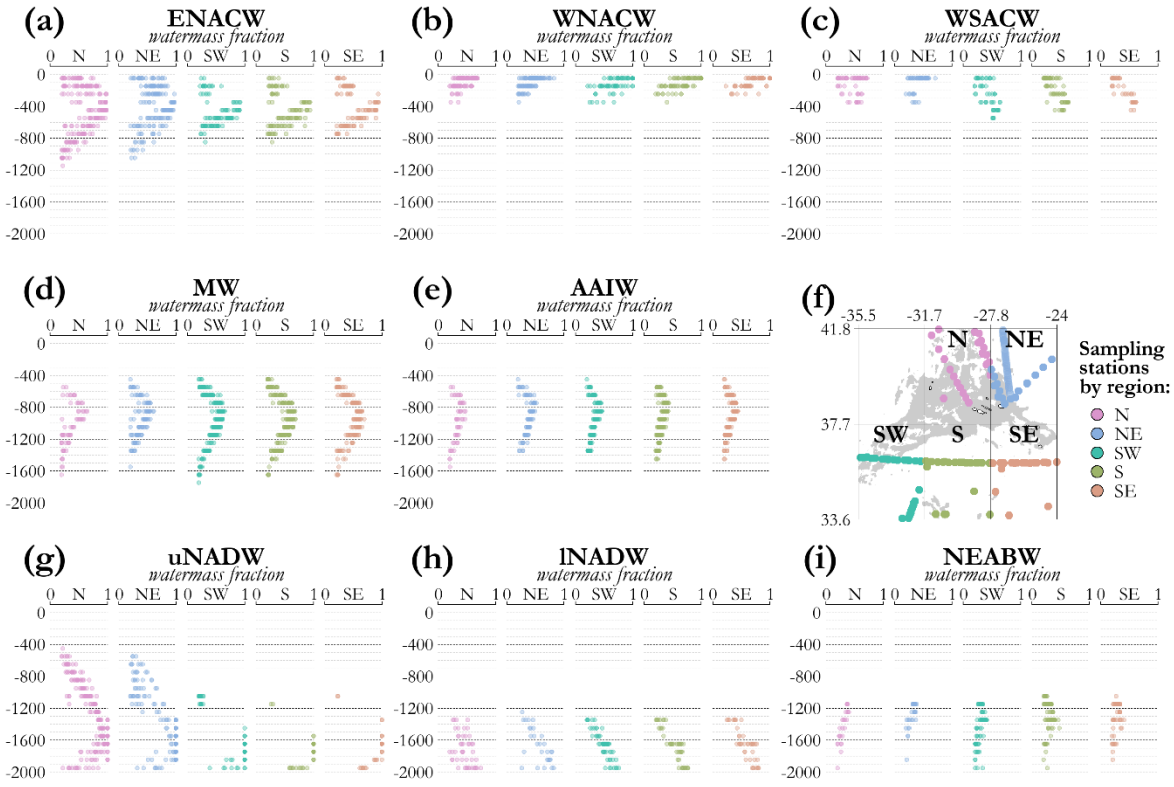

Figure A.1 – Principal water masses of the Azores based on Liu and Tanhua (2021). Upper layer: (a) Eastern North Atlantic Central Water (ENACW), (b) Western North Atlantic Central Water (WNACW), (c) Western South Atlantic Central Water (WSACW). Intermediate layer: (d) Mediterranean Water (MW), (e) Antarctic Intermediate Water (AAIW). Deep and overflow layer: (g) Upper North Atlantic Deep Water (uNADW), (h) Lower North Atlantic Deep Water (lNADW), (i) Northeast Atlantic Bottom Water (NEABW). Every sampling station is divided in 100 m bins over which water mass fractions are averaged. Only fractions greater than 0.2 are showed on the plots. (f) Sampling stations available in the study area; the grey background identifies seafloor depths shallower than 2000 m.

A.1.2 Species selection and biological data

A list of CWCs particularly vulnerable to human activities and with the characteristics of foundation species in the Azores was defined considering three traits: relative abundance, structural complexity and relative vulnerability to bottom longline and handline fishing, the main benthic extractive activities in the region. Traits could only be scored qualitatively adopting an expert-driven process, were relative and referred only to colonies formed in the Azores. Underwater videos and collected specimens informed the expert driven classification of CWC with respect to the three traits. The general criteria adopted to classify CWC are reported in Table A.1.

Presence records were aggregated at higher taxonomic levels when species identifications were uncertain. This was the case for the genus *Acanthogorgia* (mostly *A. armata* and *A. hirsuta* in the region; Braga-Henriques et al., 2013) and the family Corallidae (*Pleurocorallium johnsoni*, *Hemicorallium niobe* and *Hemicorallium tricolor*; Sampaio et al., 2019).

The modelled corals are shown in Figure A.2 and include: *Leiopathes* cf. *expansa*, *Callogorgia verticillata*, *Narella bellissima*, *Narella versluysi*, *Paracalyptrophora josephinae*, *Viminella flagellum*, *Acanthogorgia* spp., *Dentomuricea* aff. *meteor*, *Desmophyllum pertusum*, *Madrepora oculata*, *Solenosmilia variabilis*, Coralliidae, *Paragorgia johnsoni* and *Errina dabneyi*. From the initial pool of species considered, *Candidella imbricata*, c.f. *Eguchipsammia* sp. and *Leptosammia* sp. were excluded because of insufficient presence points. Most of the presence records (76%) were extracted form underwater images collected during the cruises reported in Table A.2. Additional presence records (24%) were extracted from longline and handline bycatch data (Figure A.3 and Figure A.4).

Table A.1 – Qualitative criteria adopted to define relative abundance, structural complexity and relative vulnerability of CWCs in the Azores.

Trait	Low/Small	Medium	High/Large
Relative abundance	When present the CWC species is generally not very abundant (e.g. *Callogorgia verticillata*)	When present the CWC species is known to occasionally be the most abundant of the coral assemblage (e.g. *Dentomuricea aff. meteor*)	When present the CWC species tends to be the most abundant of the coral assemblage (e.g. *Viminella flagellum*);
Structural complexity	Whip corals (e.g., *Narella versluysi*)	Branched colonies of medium size (e.g., *Narella bellissima*)	Branched colonies of large size (e.g. *Paragorgia johnsoni*)
Relative vulnerability	Colonies persist and dominate coral assemblages in heavily fished areas and do not show serious damages owed to longlines (e.g., *Viminella flagellum* and *Acanthogorgia* spp.)	Not applicable	Branched colonies with rigid supporting axes (e.g. Coralliidae); extensive damages to CWC colonies from longlines (e.g., *Paracalyptrophora josephinae*); very vulnerable life histories (e.g. *Leiopathes* spp.)

Table A.2 – Video surveys that provided the data used in the present study. Remotely operated vehicle (ROV).

Cruise name	Vessel	Year	System	Platform
Meteor Cruise M 58/3	RV Meteor	2003	ROV	QUEST
EMEPC/LUSO/G3	NRP Gago Coutinho	2008	ROV	LUSO
Treasure 2014	RV Pelagia	2014	Tow-cam	Hopper
Biometore	NI Arquipélago	2015	ROV	LUSO
MEDWAVES	RV Sarmiento de Gamboa	2016	ROV	Liropus
Treasure 2016	RV Pelagia	2016	Tow-cam	Hopper
MapGES 2018	NI Arquipélago	2018	Drift-cam	Azor
Blue Azores 2018	NRP Gago Coutinho	2018	ROV	LUSO
Nico 2018	RV Pelagia	2018	Tow-cam	Hopper
Greenpeace Pole-to-Pole	MV L'Esperanza	2019	ROV	SEAEYE COUGAR-XT
MapGES 2019	NI Arquipélago	2019	Drift-cam	Azor
Rainbow 2019	RV Pelagia	2019	Tow-cam	Hopper
iMAR 2021/Eurofleets+	RV Pelagia	2021	Tow-cam	Hopper

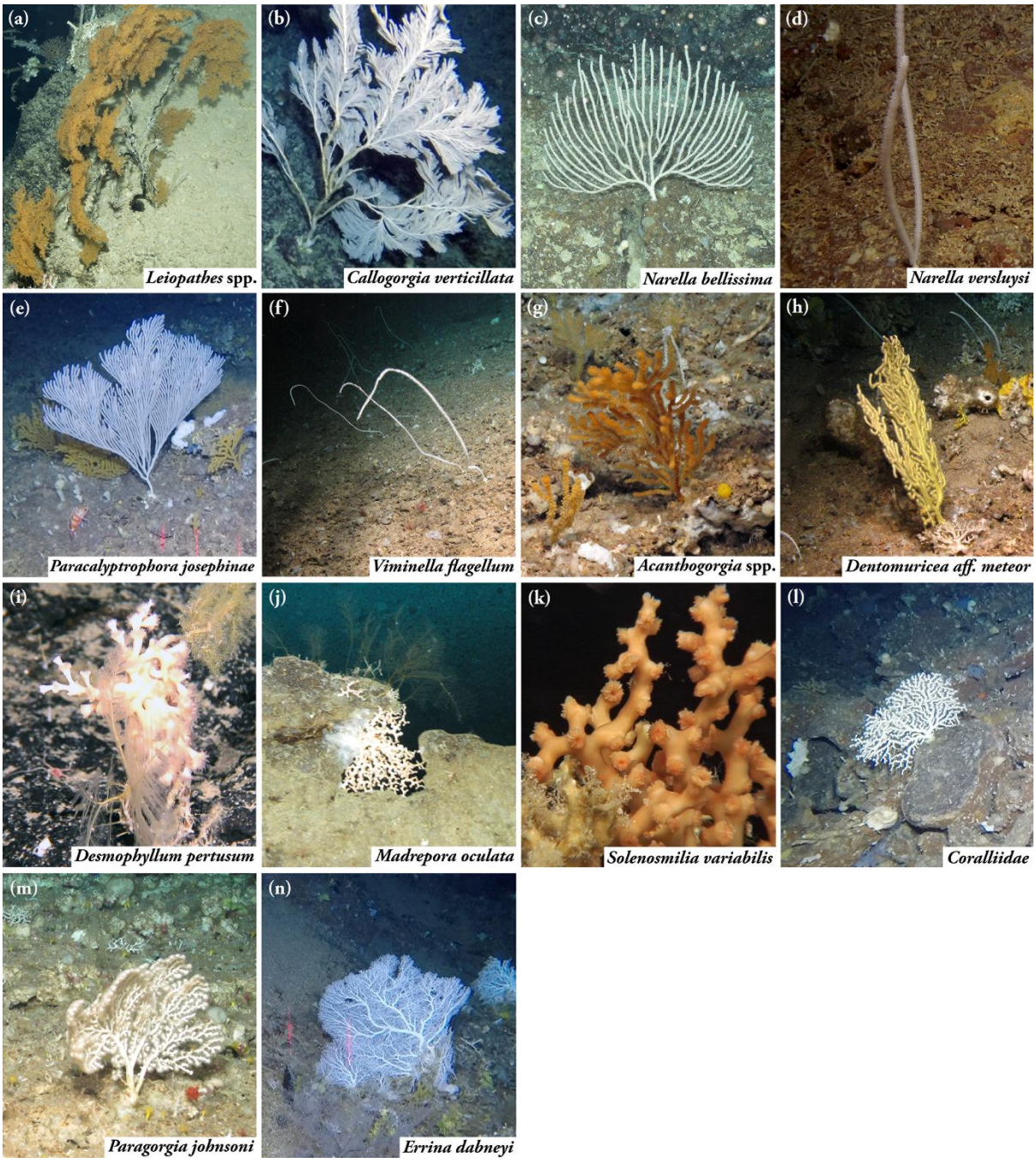

Figure A.2 – Cold water coral taxa selected to develop habitat suitability models. These species were selected considering their vulnerability to human activities and their potential as foundation species. © MEDWAVES, ATLAS project (*Leiopathes* cf. *expansa*, *N. bellissima*, *N. versluysi*, *D. pertusum* and *M. oculata*); © IMAR/OKEANOS-UAz, Drift camera (*C. verticillata*, *P. josephinae*, Coralliidae, *E. dabneyi*); © ROV Luso / EMEPC / 2018 Oceano Azul Expedition, organized by Oceano Azul Foundation & partners (*V. flagellum*, *Acanthogorgia* spp. and *D.* aff. *meteor*); © EMEPC/Luso/Açores/2013 (*S. variabilis*).

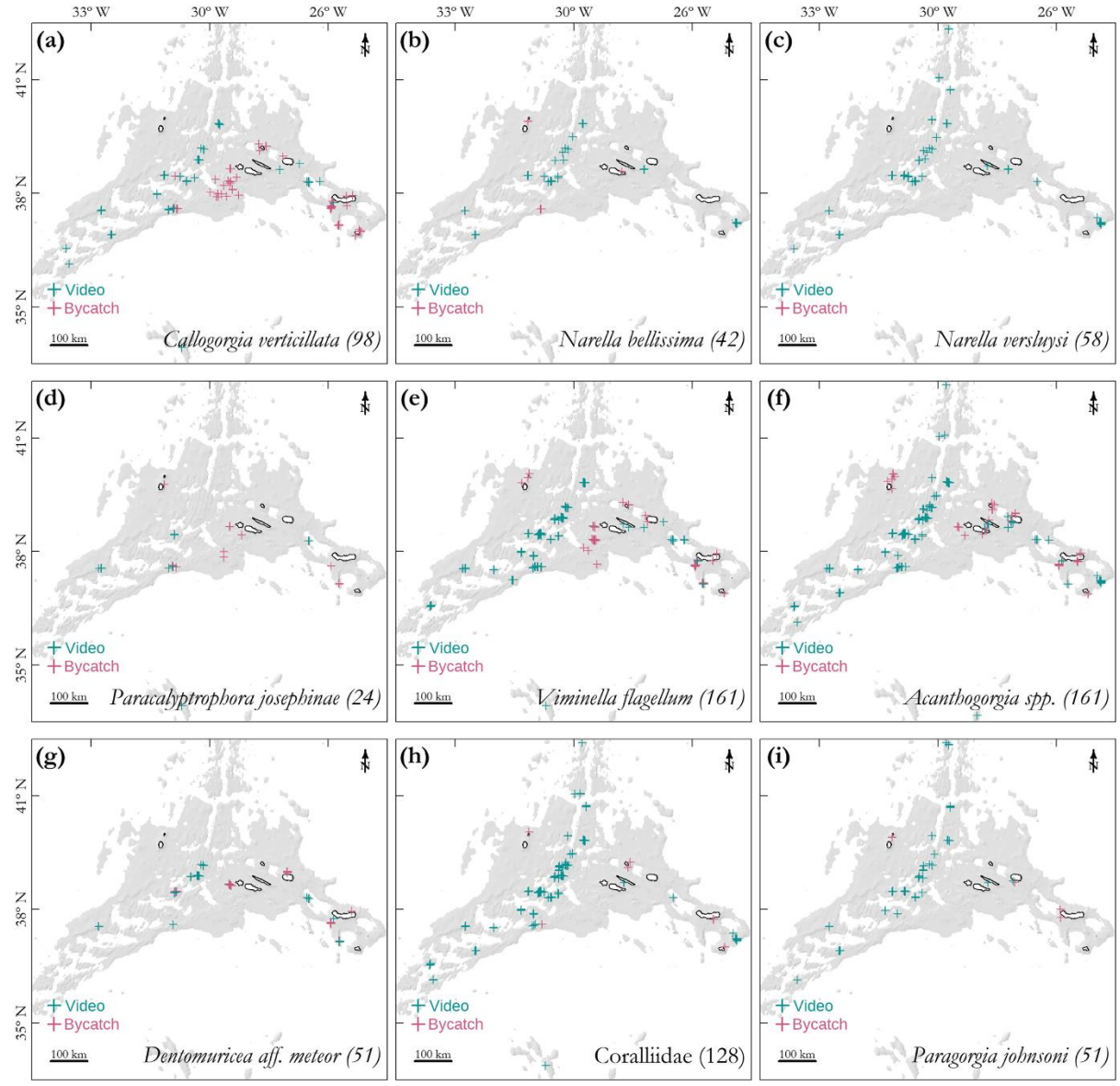

Figure A.3 – Presence records of the selected octocoral (Alcyonacea) taxa. Suborders: (a-e) Calcaxonia; (f-g) Holaxonia; (h-i) Scleraxonia. In parenthesis the number of records used to fit the habitat suitability models.

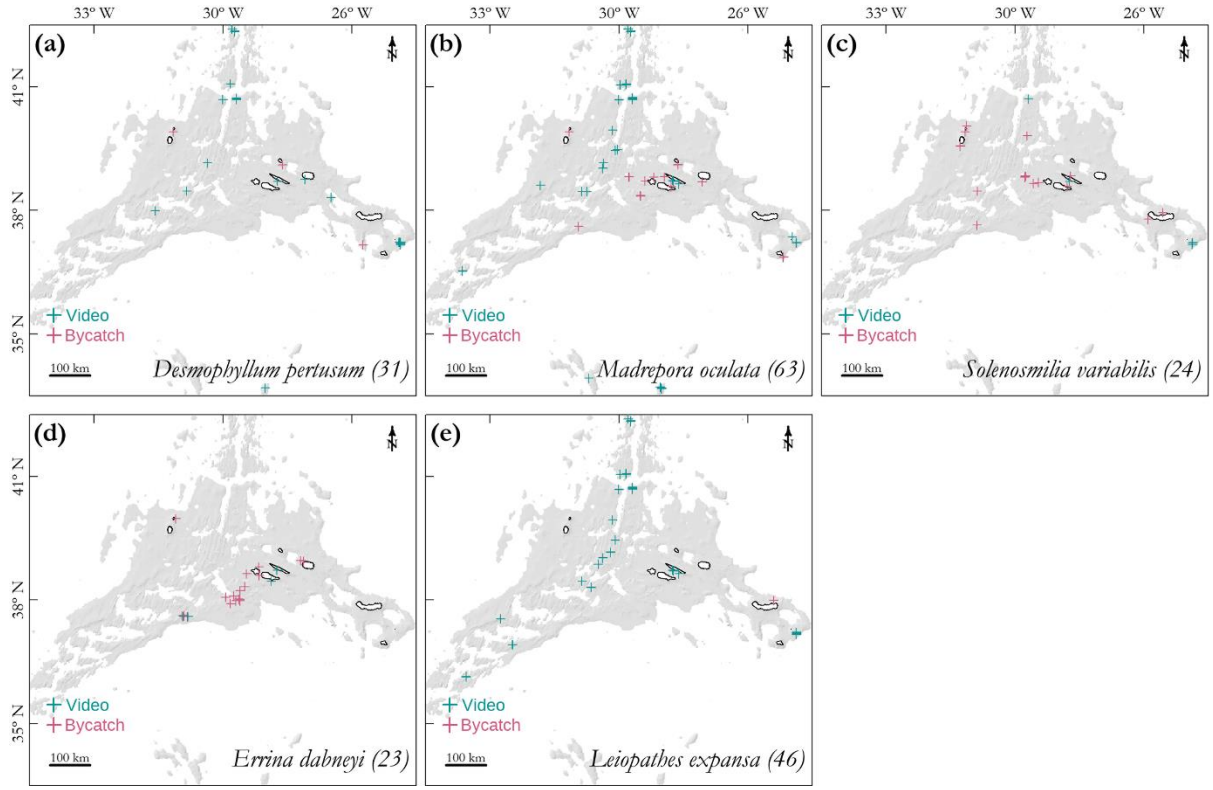

Figure A.4 – Presence records of the selected Scleractinia (a-c), Anthoathecata (Stylasteridae) (d) and Antipatharia (e) species. In parenthesis the of number records used to fit the habitat suitability models.

A.1.3 Environmental data

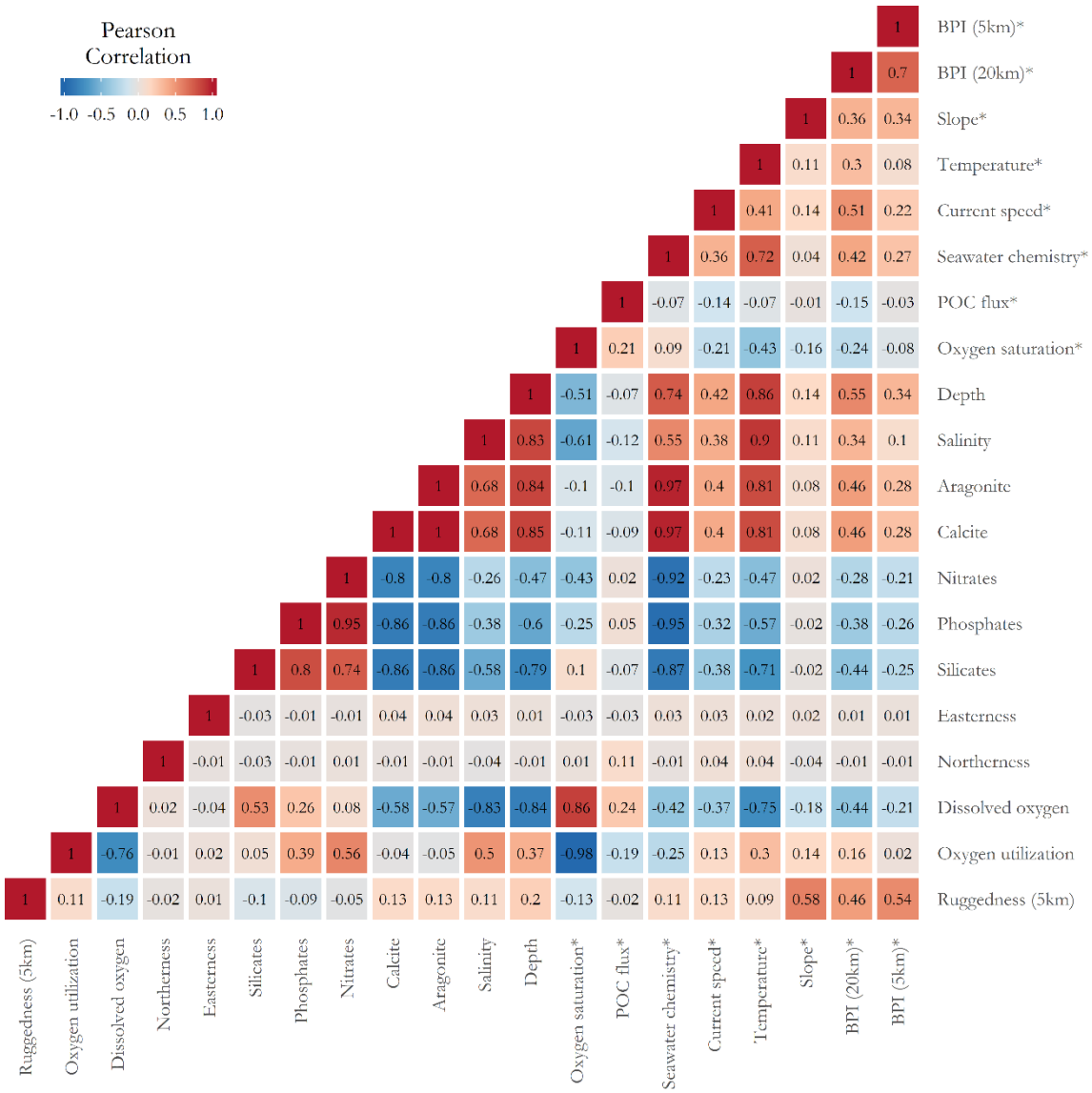

Figure A.5 – Pearson's correlation coefficient among all the considered predictor variables. (*) Final set of predictor variables used in the modelling exercise.

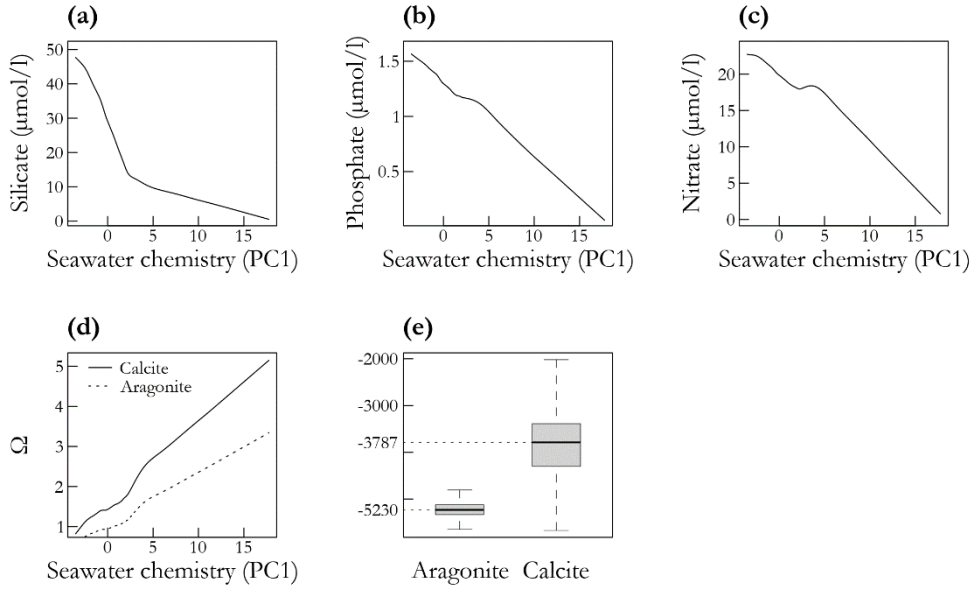

Figure A.6 – Principal component analysis (PCA) of aragonite, calcite, nitrate, phosphate and silicate. (a-d) Locally weighted scatterplot smoothing (LOWESS) of the unscaled values of silicate, phosphate, nitrate, calcite and aragonite against the first principal component ('seawater chemistry'). (e) Depths corresponding to aragonite and calcite undersaturation zones ($\Omega < 1$). Seawater chemistry, concentration and saturation values refers to seafloor locations at depth shallower than 2000 m.

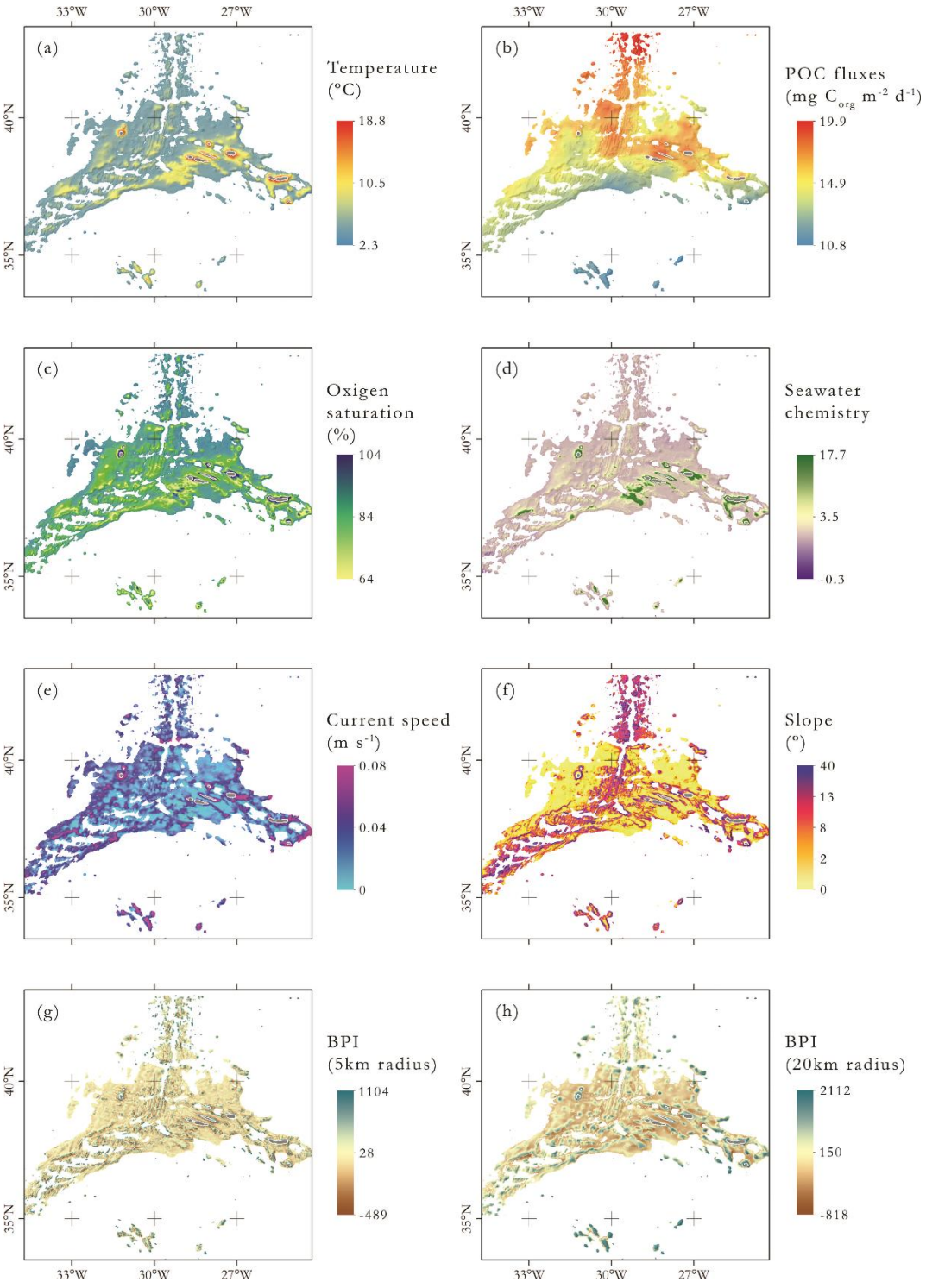

Figure A.7 – Environmental layers used in the habitat suitability models. (b) POC: Particulate organic carbon export. (d) Seawater chemistry: first principal component of the principal component analysis combining nitrate, phosphate, silicate, aragonite and calcite concentrations. (g-h) BPI: Bathymetric position index.

A.2 Model settings

Table A.3 – Similarity matrix. Values of 1 mean that two categories are identical and values of zero that they are totally different. US: unsuitable cells; LC: low-confidence suitable cells; MC: medium-confidence suitable cells; HC: high-confidence suitable cells.

	US	LC	MC	HC
US	-	0	0	0
LC	-	1	0.6	0.3
MC	-	-	1	0.6
HC	-	-	-	1

A.3 Model predictions

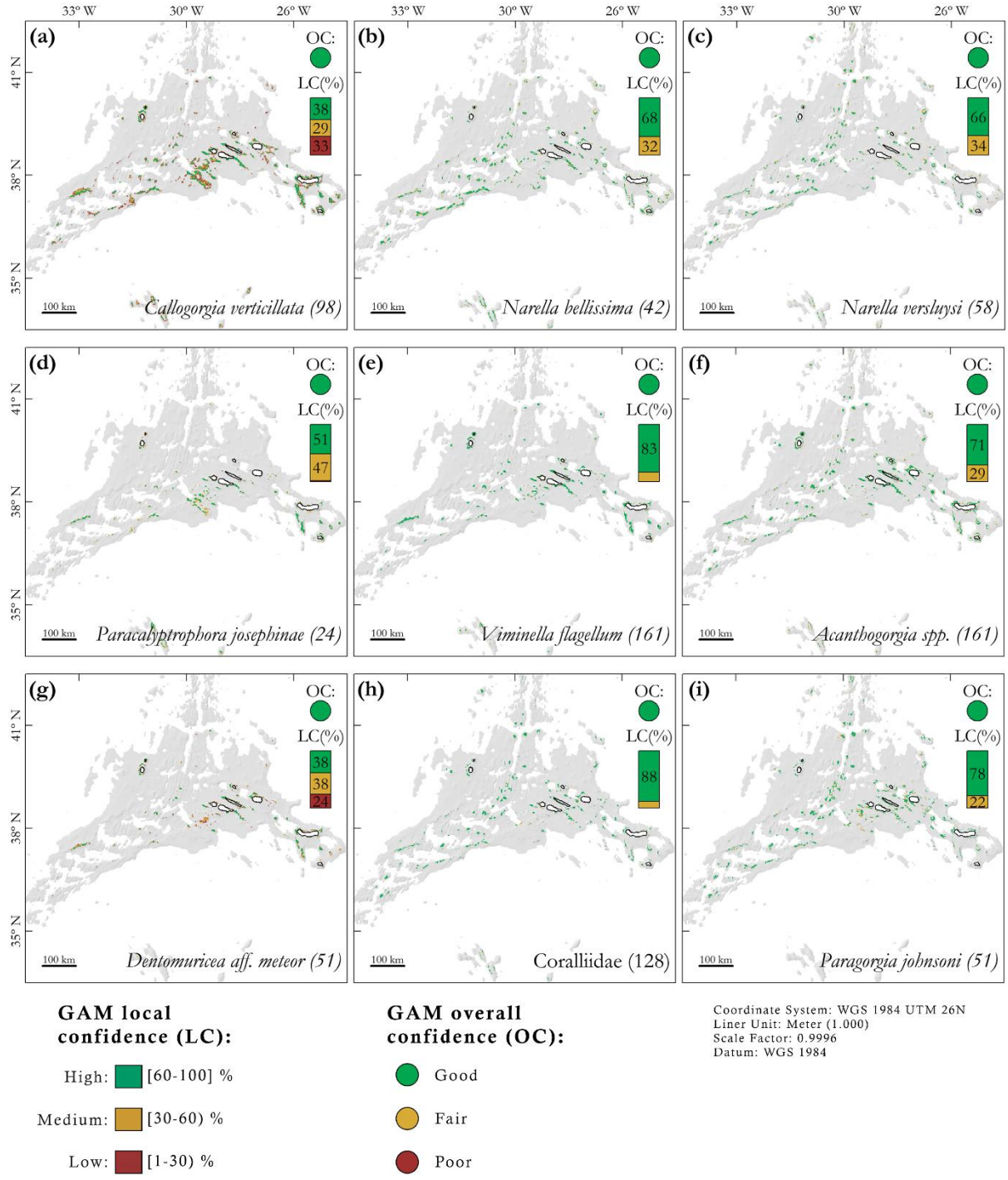

Figure A.8 – Habitat suitability maps produced by GAMs for Octocorallia (Alcyonacea). Local confidence (LC): the number of times a cell is classified as suitable when models are fitted with resampled input data. Overall confidence (OC): determined by the lowest score among AUC_{GAM} and TSS_{GAM} – area under the curve (AUC); true skill statistics (TSS). In parenthesis the number of records used to fit the models.

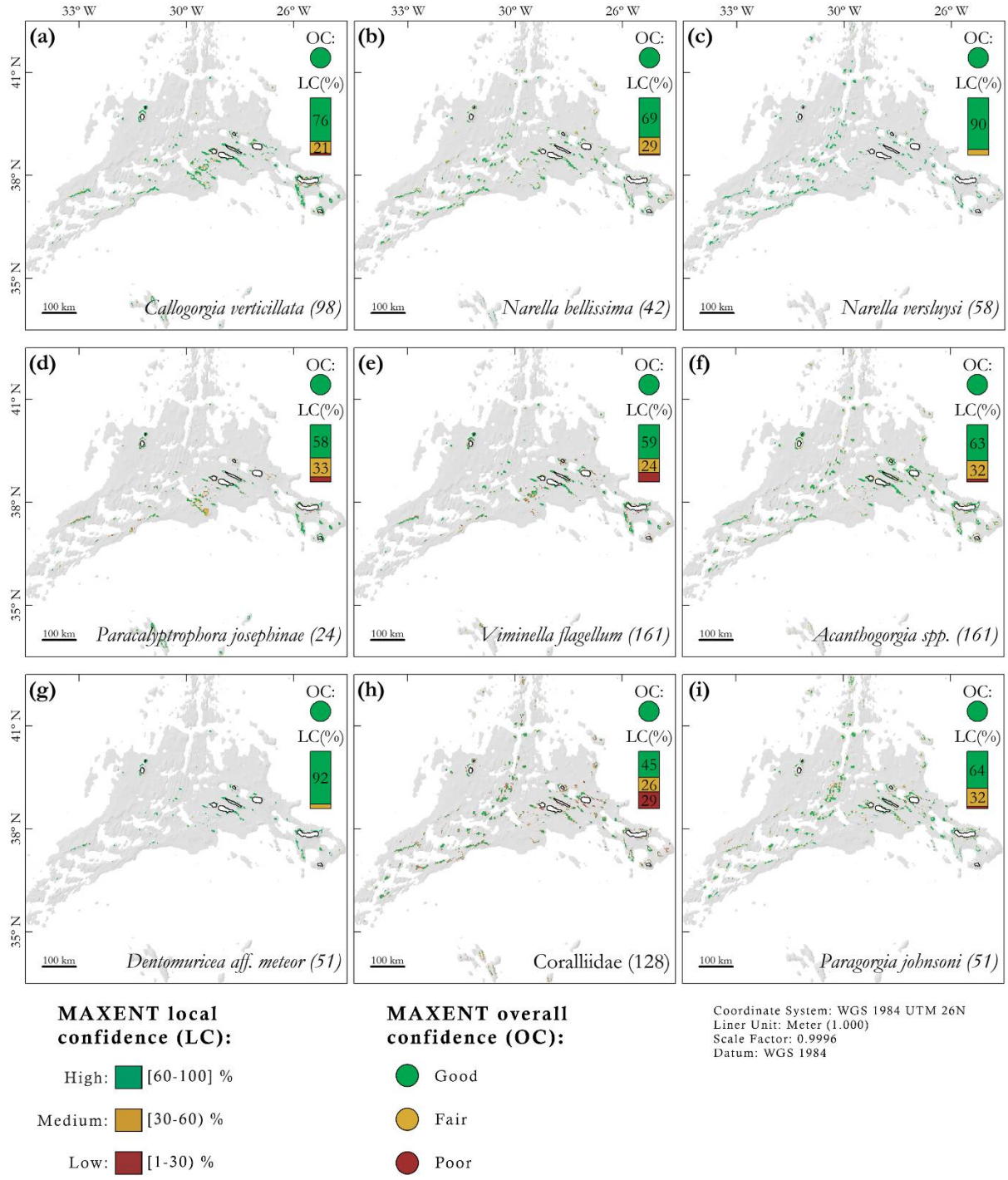

Figure A.9 – Habitat suitability maps produced by Maxent models for Octocorallia (Alcyonacea). Local confidence (LC): the number of times a cell is classified as suitable when models are fitted with resampled input data. Overall confidence (OC): determined by the lowest score among AUC_{Maxent} and TSS_{Maxent} – area under the curve (AUC); true skill statistics (TSS). In parenthesis the number of records used to fit the models.

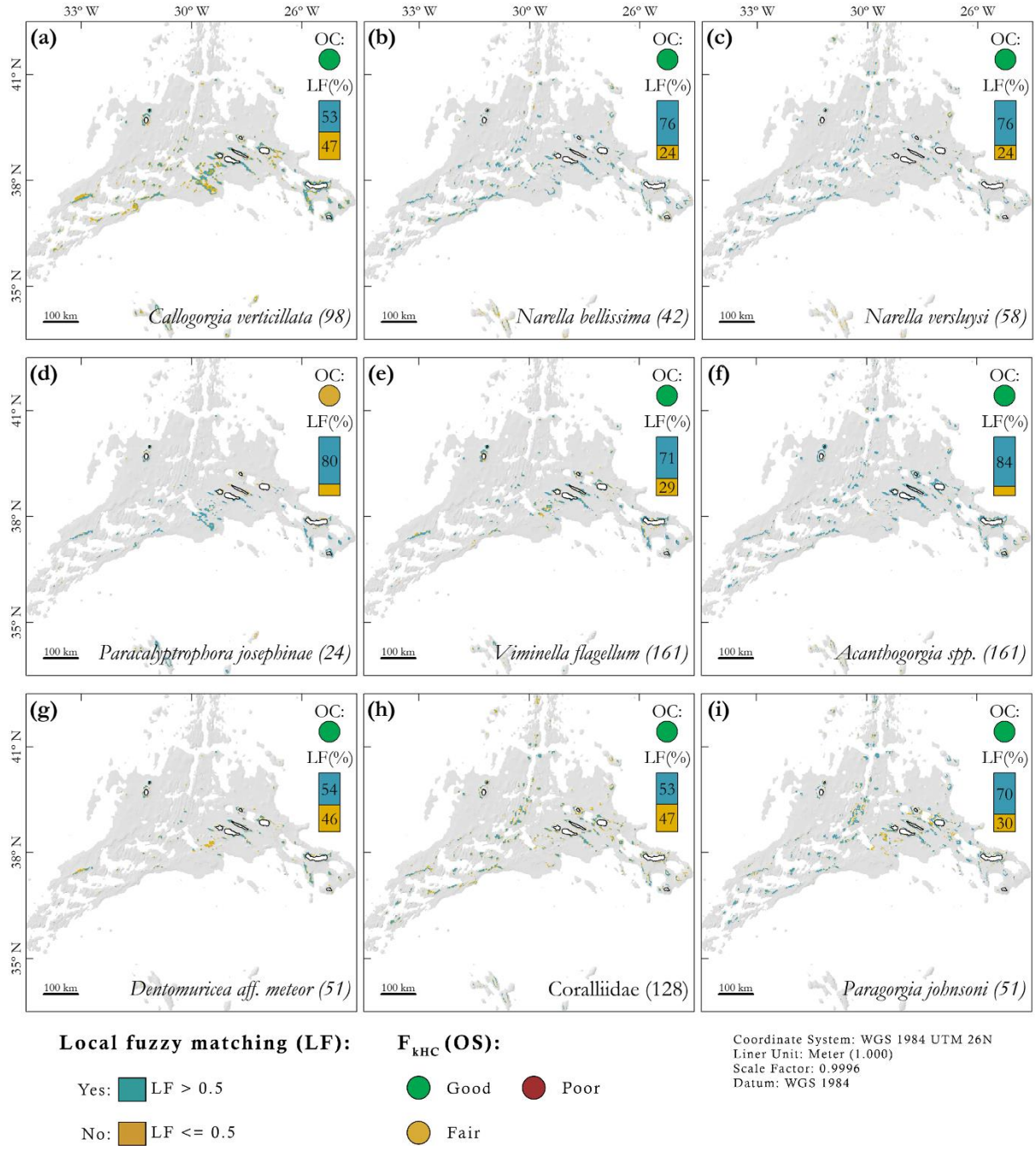

Figure A.10 – Fuzzy similarity of the habitat suitability maps produced by GAM and Maxent models for Octocorallia (Alcyonacea). Local fuzzy matching (LF): computes the fuzzy matching of GAM and Maxent suitable cells using two membership functions (i) category similarity and (ii) distance decay. Fuzzy matching values greater than 0.5 means that the two cells are more similar than different. Overall similarity (OS): determined by the improved fuzzy kappa score for high quality cells (F_{kHC}). The score measures the degree of overlap of GAM and Maxent high confidence cells.

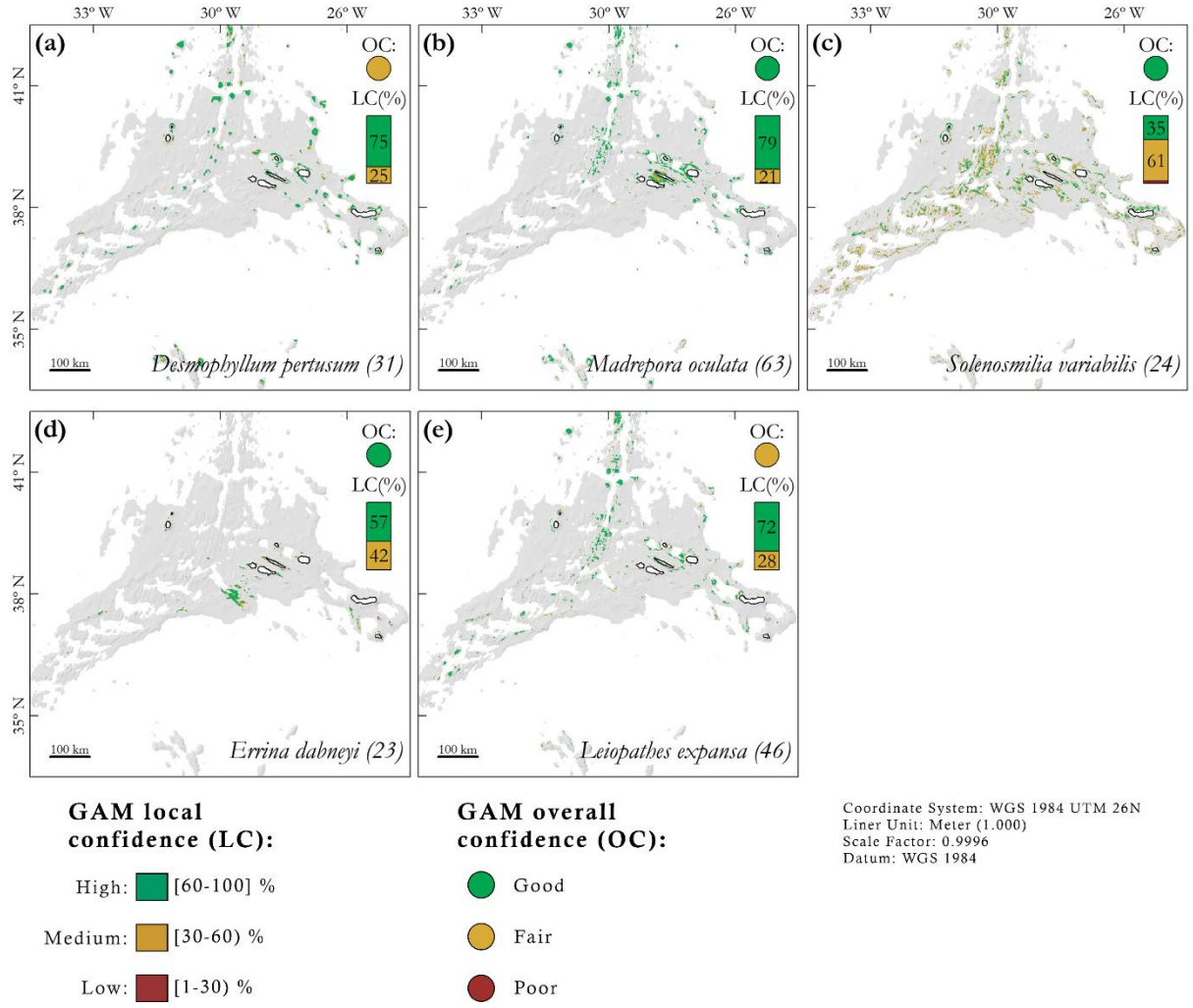

Figure A.11 – Habitat suitability maps produced by GAMs for Scleractinia (a-c), Anthoathecata (c) and Antipatharia (e). Local confidence (LC): the number of times a cell is classified as suitable when models are fitted with resampled input data. Overall confidence (OC): determined by the lowest score among AUC_{GAM} and TSS_{GAM} – area under the curve (AUC); true skill statistics (TSS). In parenthesis the number of records used to fit the models.

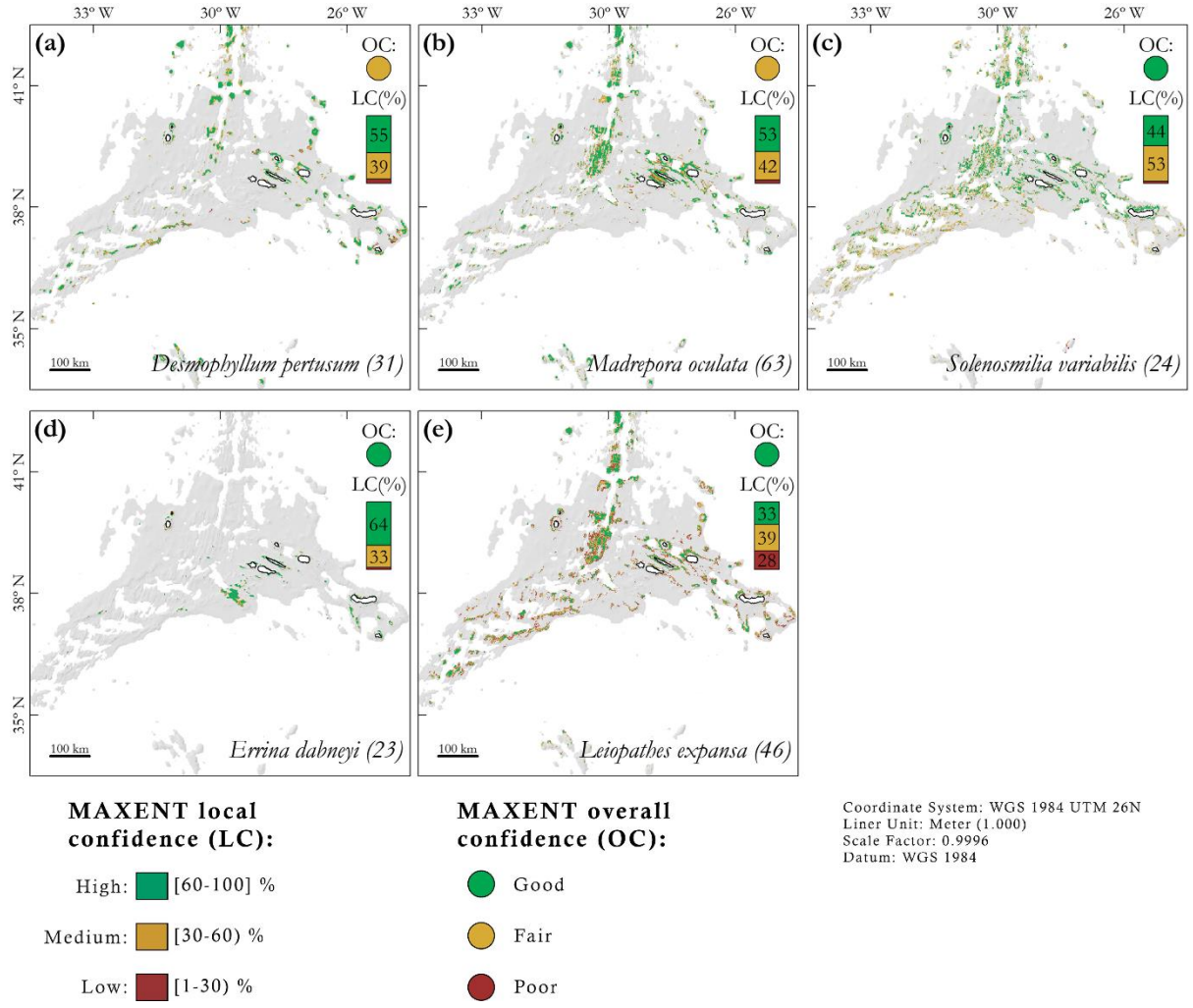

Figure A.12 – Habitat suitability maps produced by Maxent models for Scleractinia (a-c), Anthoathecata (c) and Antipatharia (e). Local confidence (LC): the number of times a cell is classified as suitable when models are fitted with resampled input data. Overall confidence (OC): determined by the lowest score among AUC_{Maxent} and TSS_{Maxent} – area under the curve (AUC); true skill statistics (TSS). In parenthesis the number of records used to fit the models.

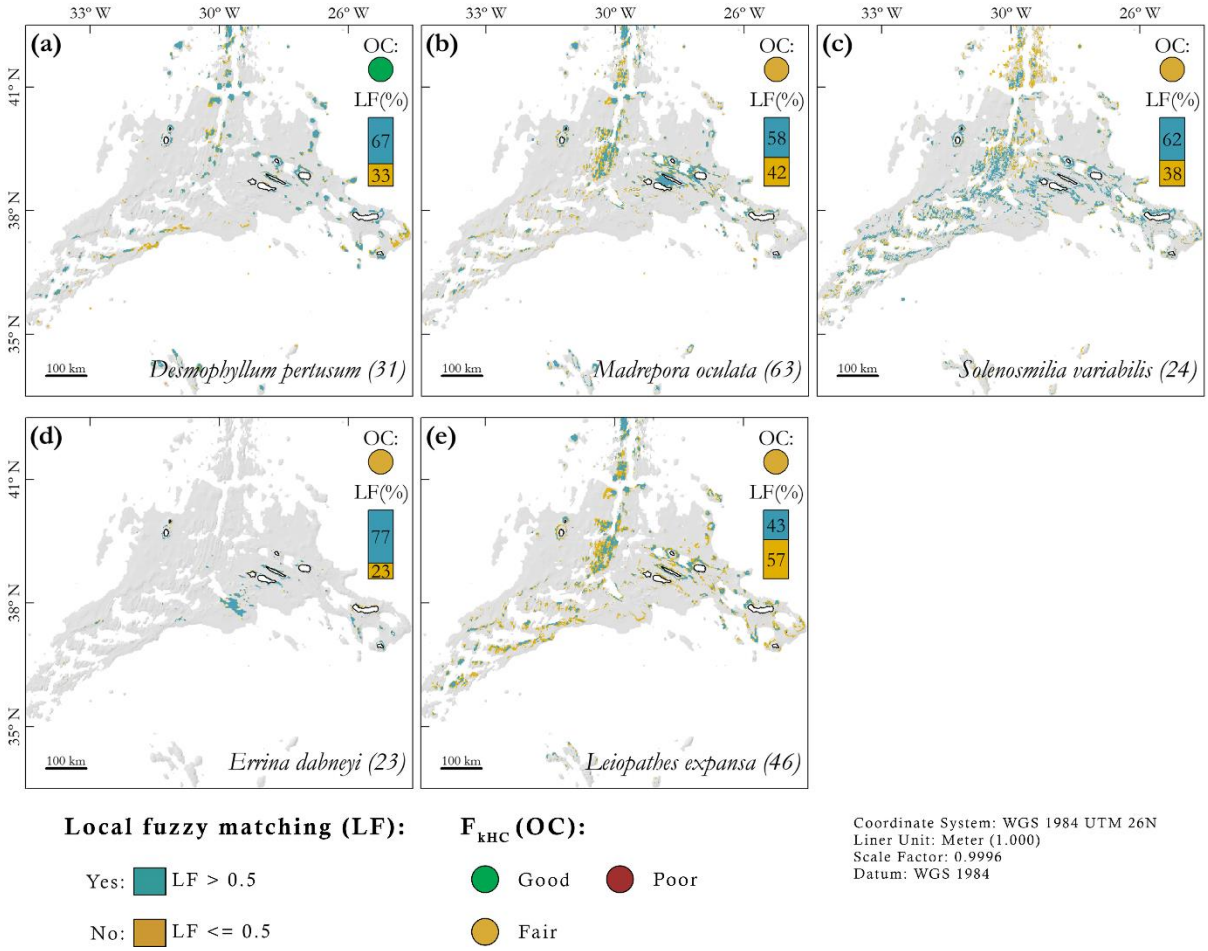

Figure A.13 – Fuzzy similarity of the habitat suitability maps produced by GAM and Maxent models for Scleractinia (a-c), Anthoathecata (c) and Antipatharia (e). Local fuzzy matching (LF): computes the fuzzy matching of GAM and Maxent suitable cells using two membership functions (i) category similarity and (ii) distance decay. Fuzzy matching values greater than 0.5 means that the two cells are more similar than different. Overall similarity (OS): determined by the improved fuzzy kappa score for high quality cells (F_{kHC}). The score measures the degree of overlap of GAM and Maxent high confidence cells.

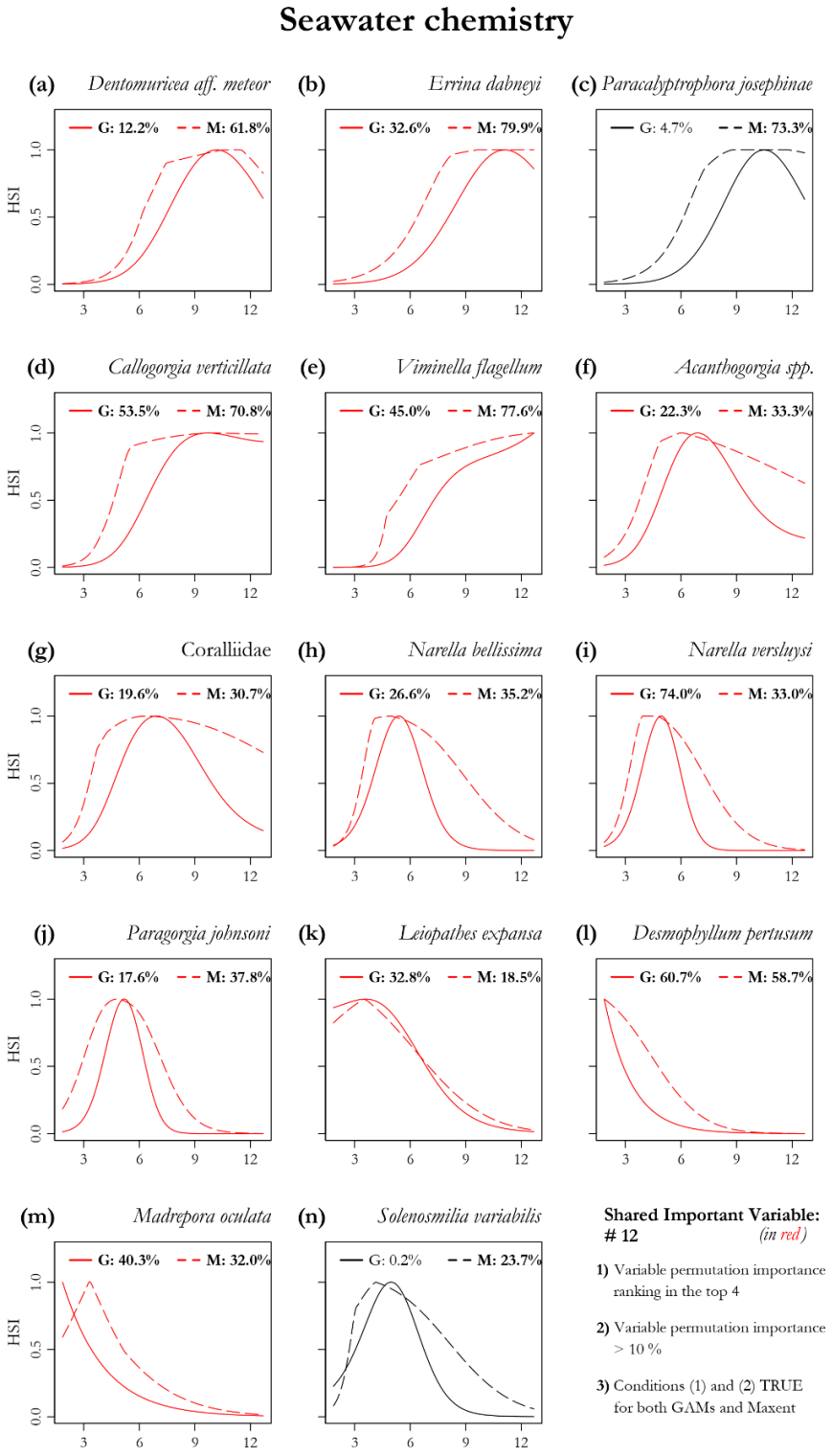

Figure A.14 – Response curves for GAM (solid lines) and Maxent (dashed lines) models, predictor *seawater chemistry*. Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

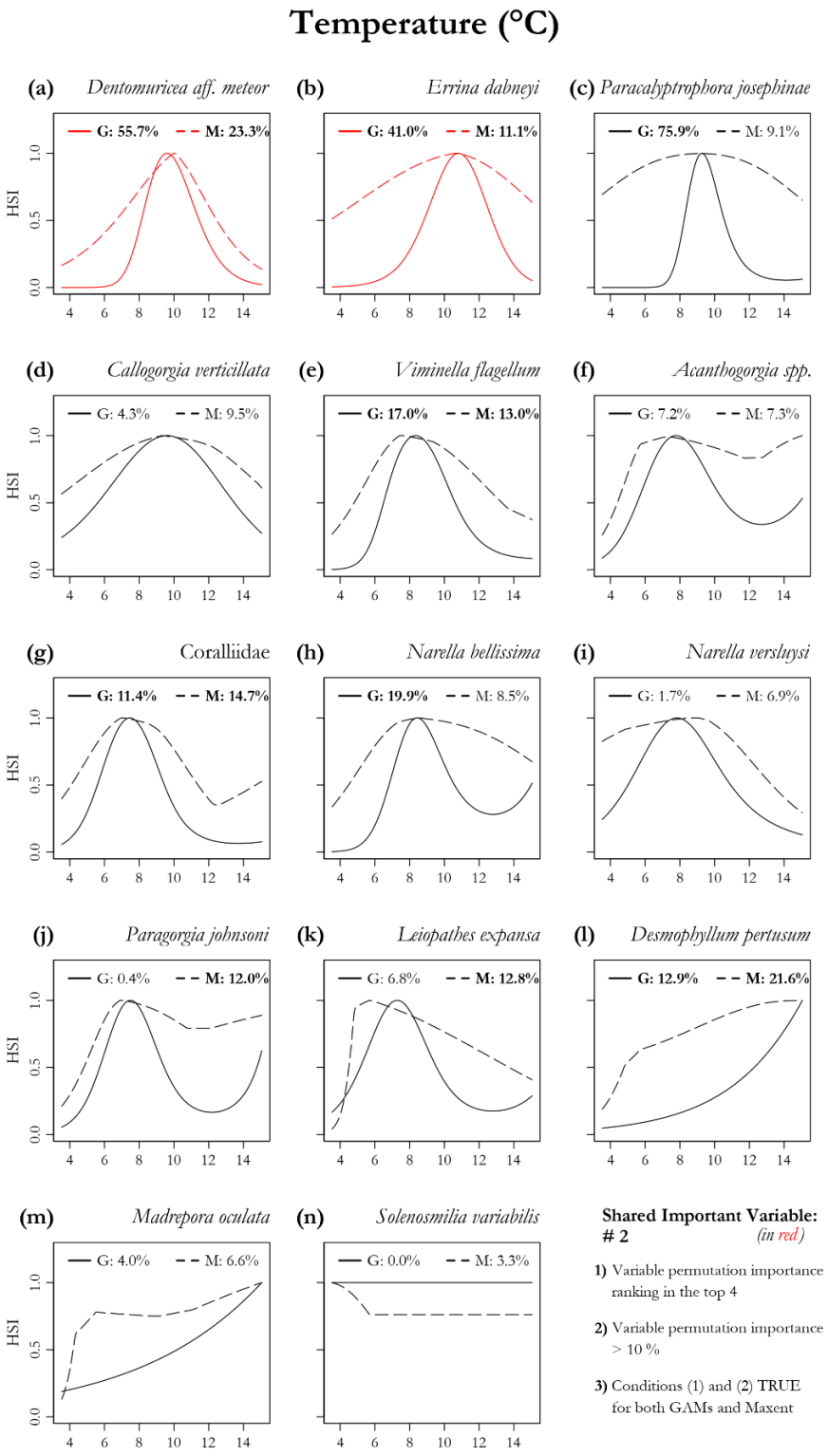

Figure A.15 – Response curves for GAM (solid lines) and Maxent (dashed lines) models, predictor *seafloor temperature*. Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

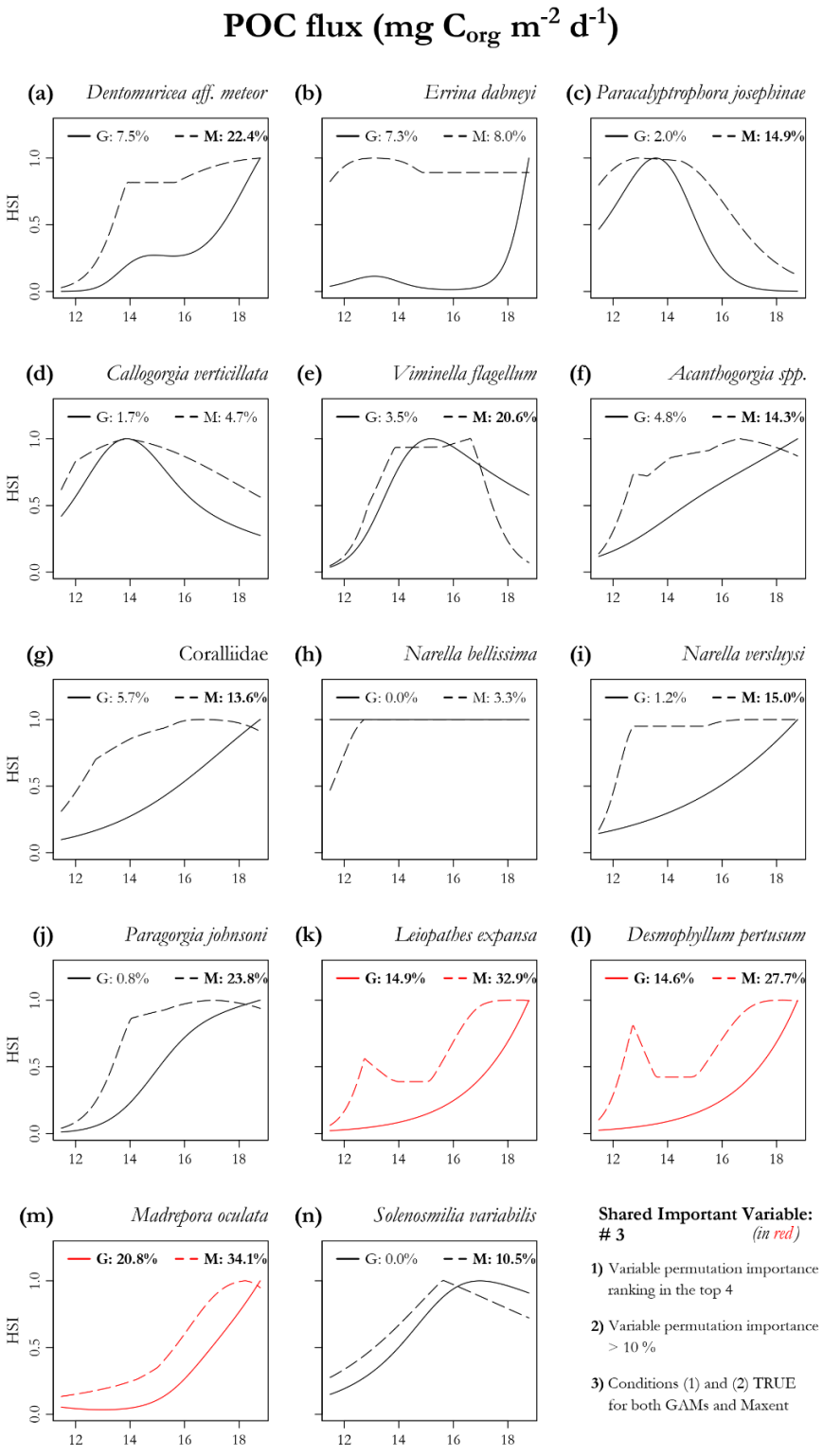

Figure A.16 – Response curves for GAM (solid lines) and Maxent (dashed lines) models, predictor *POC flux* (Particulate Organic Carbon flux). Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

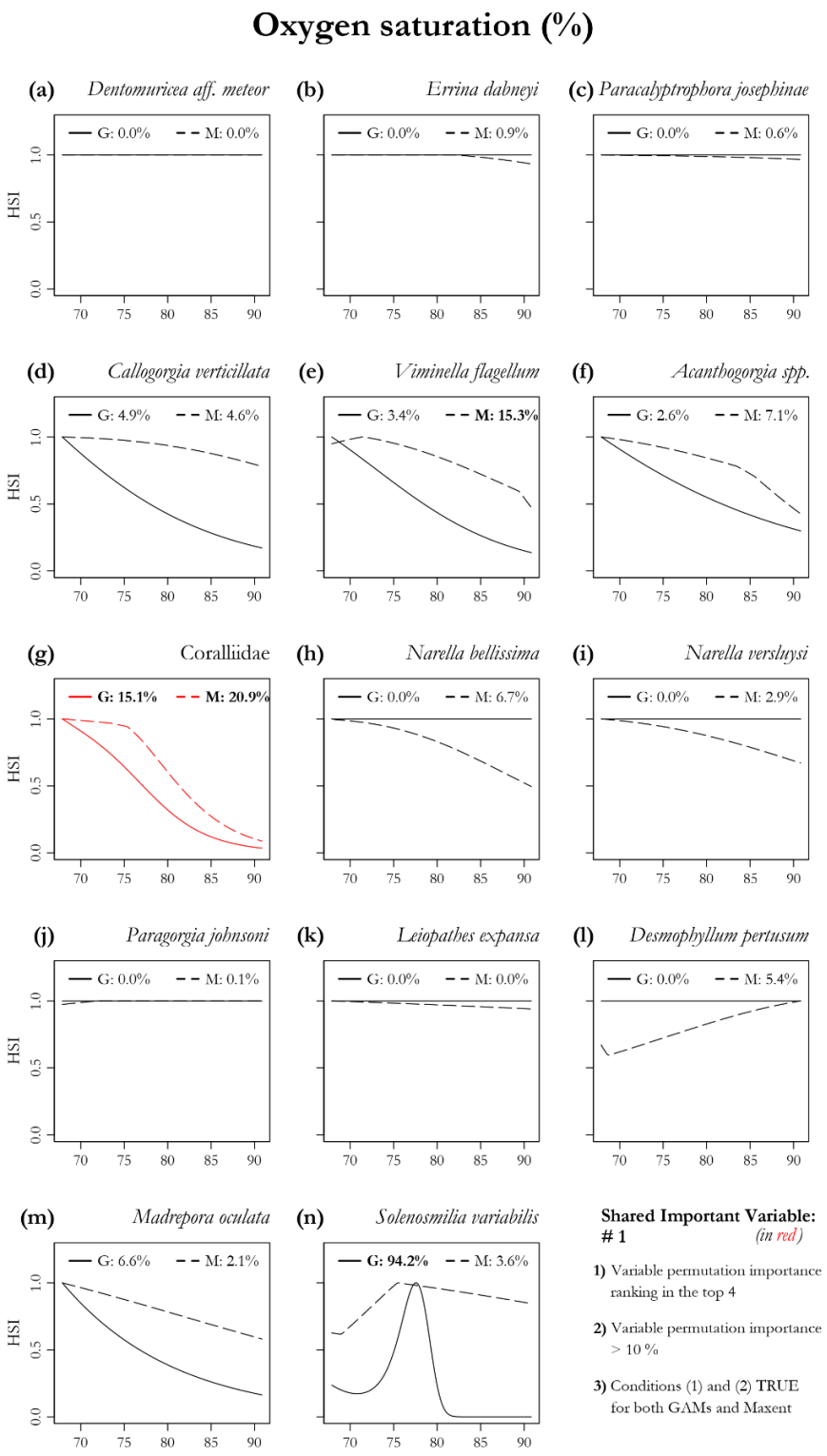

Figure A.17 – Response curves for GAM (solid lines) and Maxent (dashed lines) models, predictor *oxygen saturation*. Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

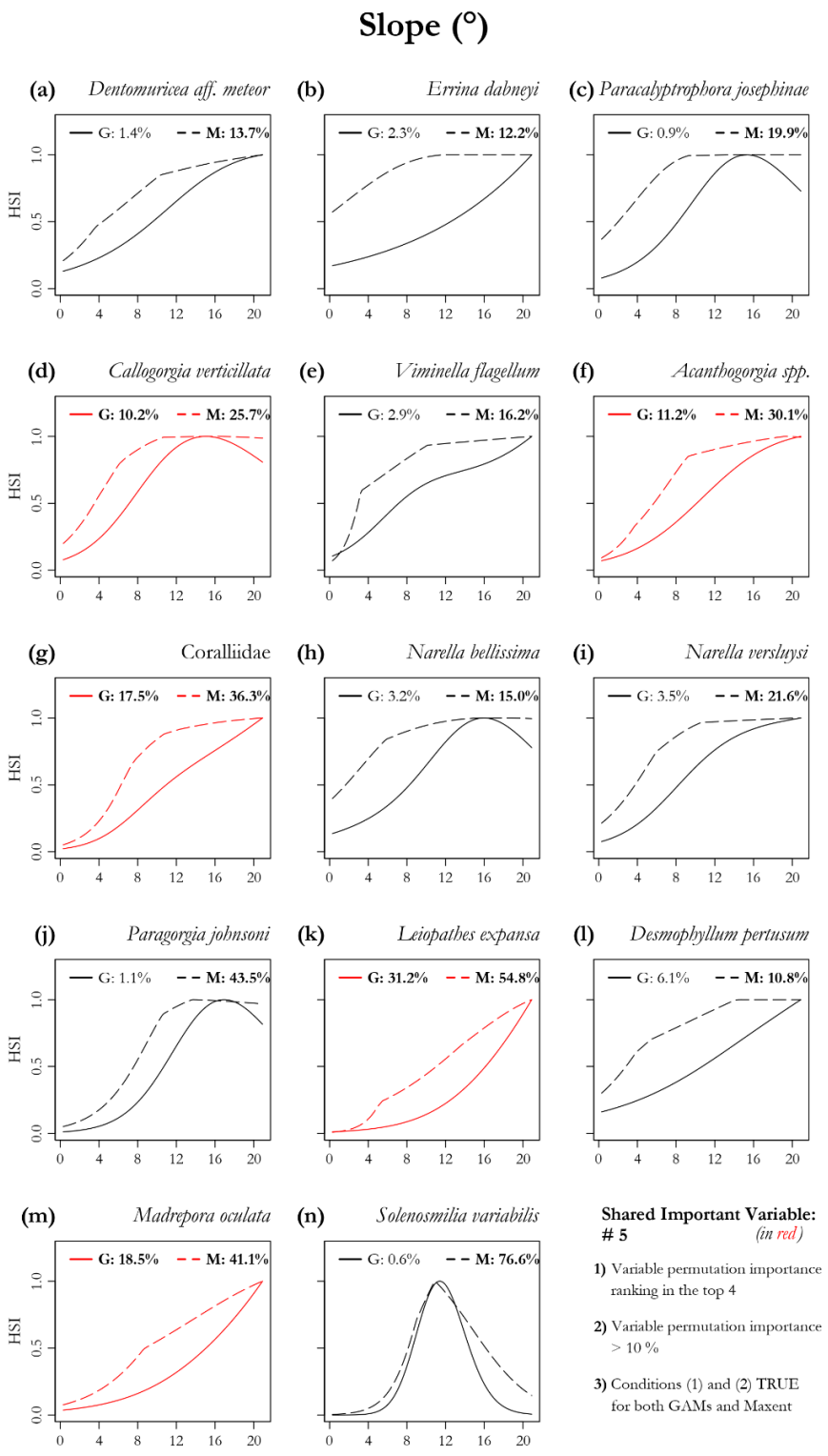

Figure A.18 – Response curves for GAM (solid lines) and Maxent (dashed lines) models, predictor *slope*. Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

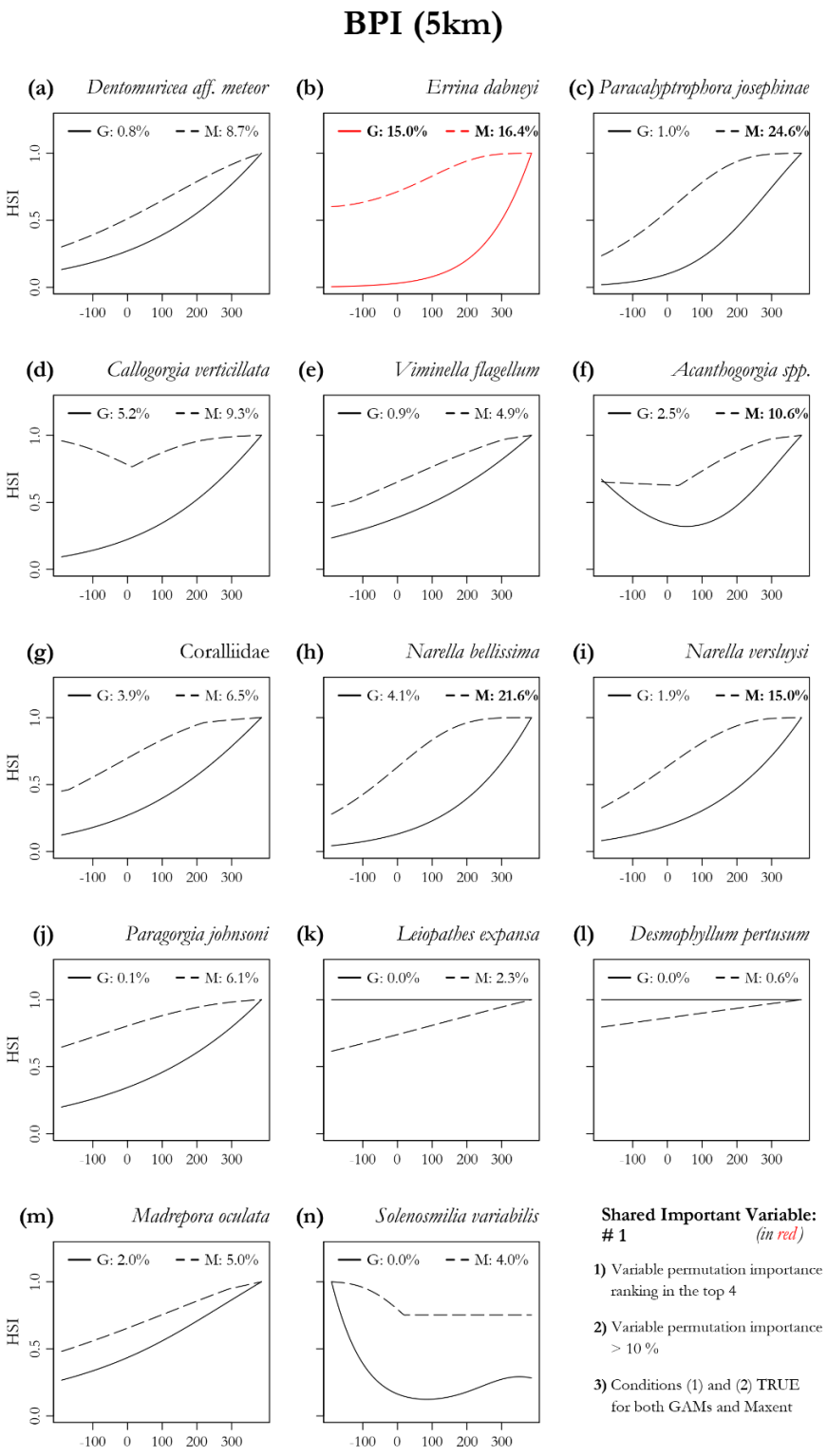

Figure A.19 – Response curves for GAM (solid lines) and Maxent (dashed lines) models, predictor *BPI 5 km* (Bathymetric Position Index computed on a 5 km radius). Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

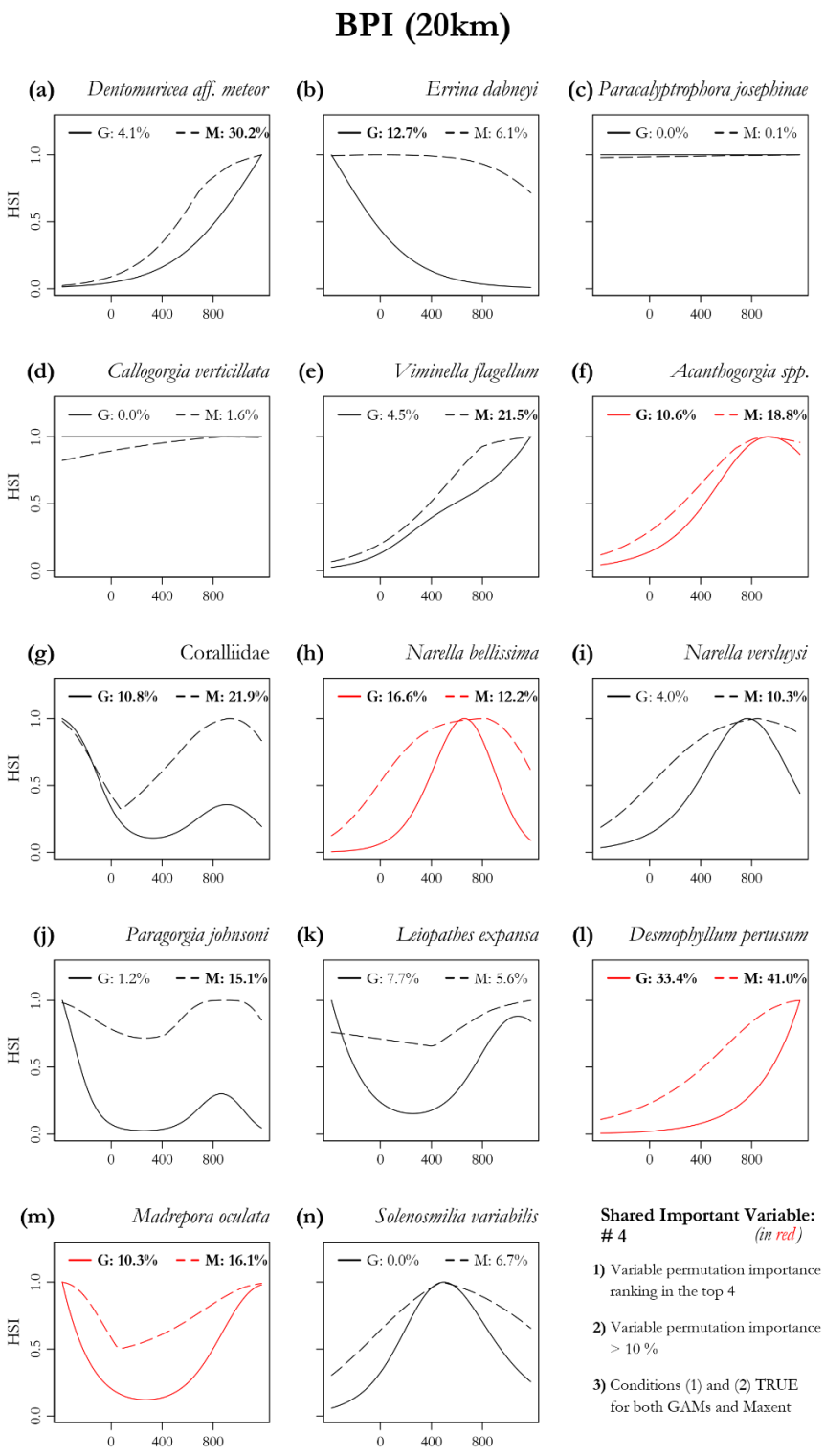

Figure A.20 – Response curves for GAM (solid lines) and Maxent (dashed lines) models, predictor *BPI 20 km* (Bathymetric Position Index computed on a 20 km radius). Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

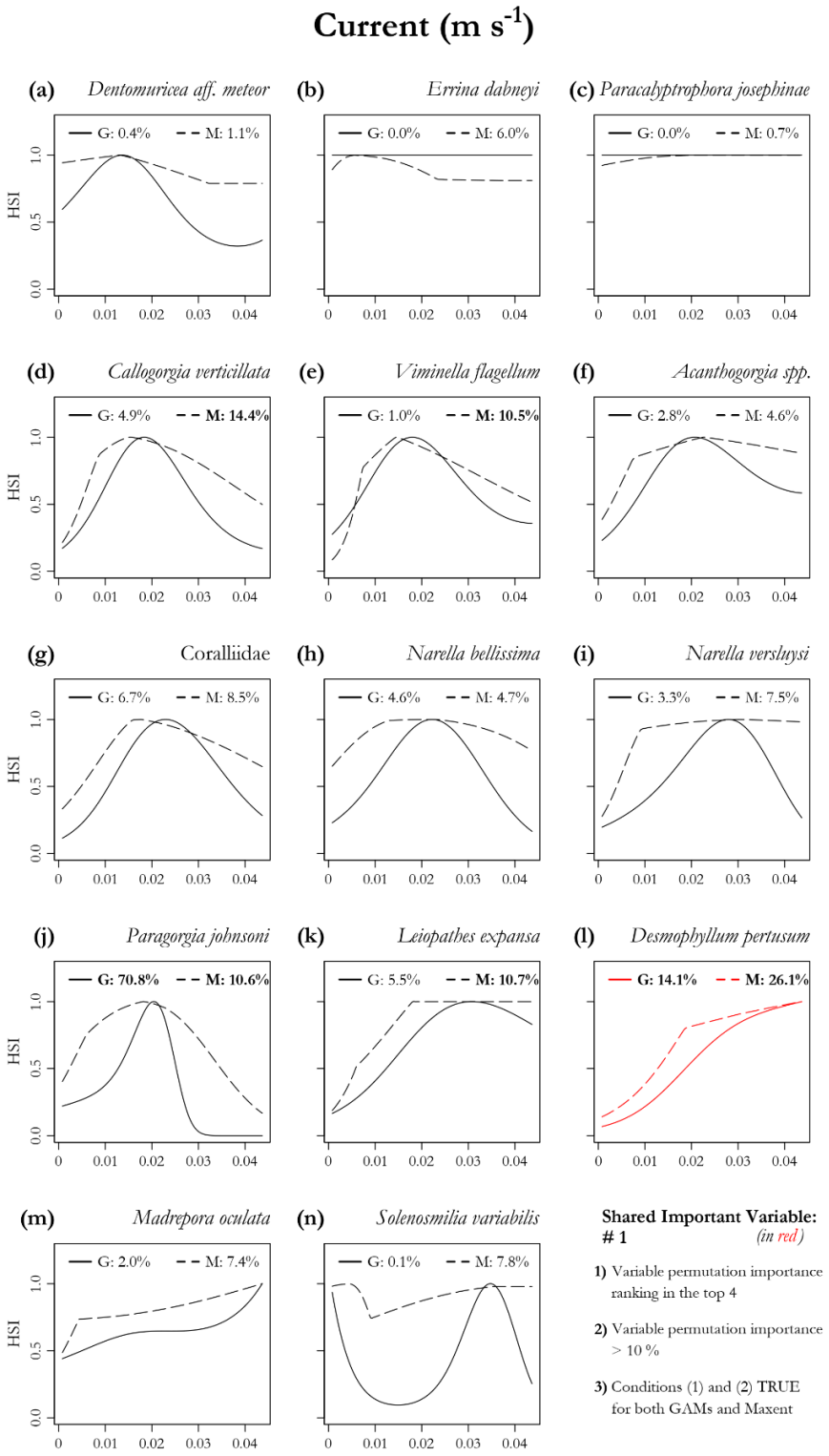

Figure A.21 – Response curves for GAM (solid lines) and Maxent (dashed lines) models regarding the predictor *bottom current speed*. Percentage values represent the permutation variable importance (in bold values > 10 %). Shared important variables are colored in red. Species are ordered according to the mean depth of their estimated suitable habitat. HSI: Habitat Suitability Index.

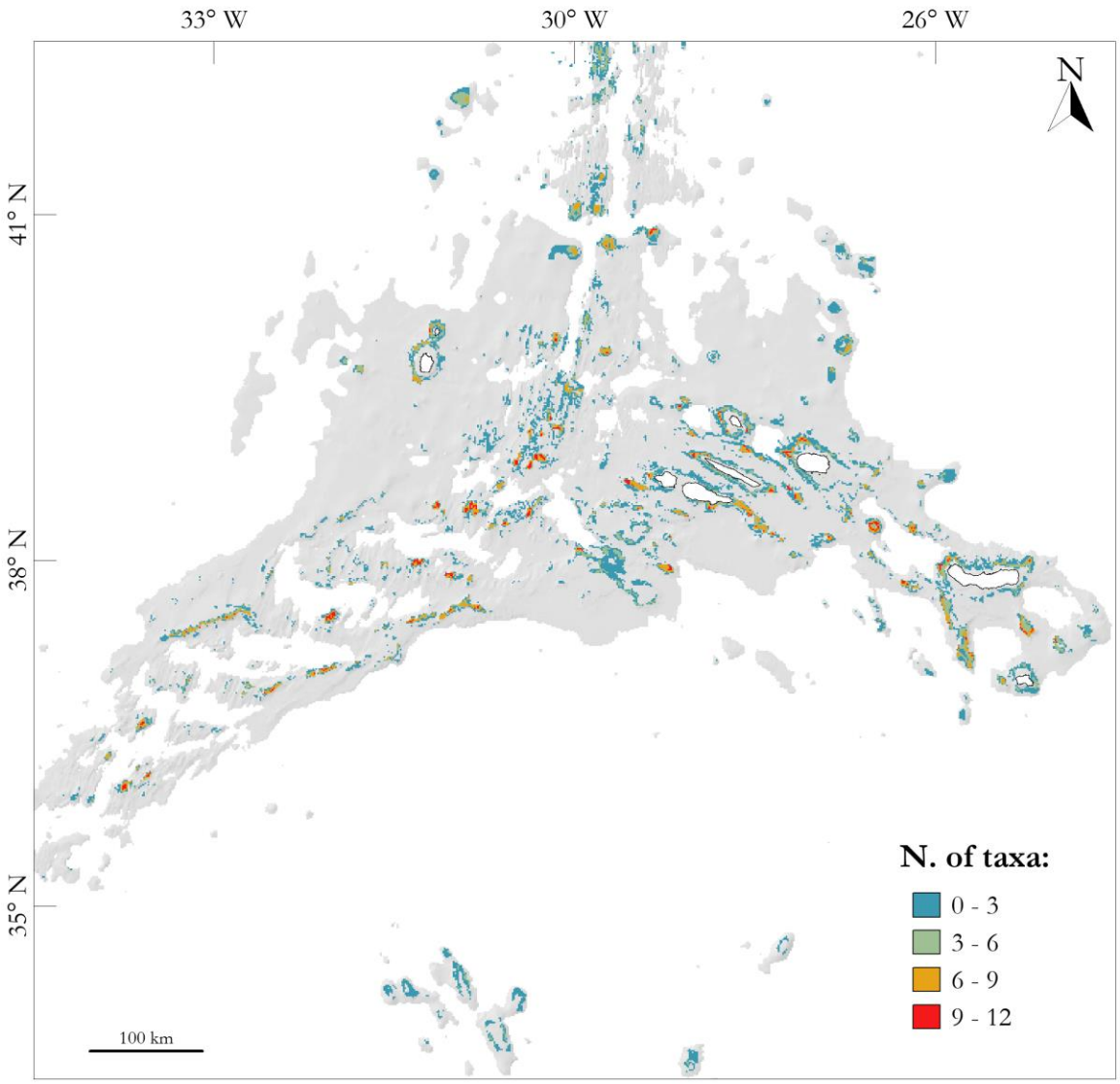

Figure A.22 – Overlapping habitat suitability predictions.

Appendix B
Supplementary data Chapter 3

B.1 Input data

Table B.1 – Video surveys that provided the data used in the present study. Remotely operated vehicle (ROV).

Cruise name	Vessel	Year	System	Platform
Meteor Cruise M 58/3	RV Meteor	2003	ROV	QUEST
EMEPC/LUSO/G3	NRP Gago Coutinho	2008	ROV	LUSO
Treasure 2014	RV Pelagia	2014	Tow-cam	Hopper
Biometore	NI Arquipélago	2015	ROV	LUSO
MEDWAVES	RV Sarmiento de Gamboa	2016	ROV	Liropus
Treasure 2016	RV Pelagia	2016	Tow-cam	Hopper
MapGES 2018	NI Arquipélago	2018	Drift-cam	Azor
Blue Azores 2018	NRP Gago Coutinho	2018	ROV	LUSO
Nico 2018	RV Pelagia	2018	Tow-cam	Hopper
Greenpeace Pole-to-Pole	MV L'Esperanza	2019	ROV	SEAEYE COUGAR-XT
MapGES 2019	NI Arquipélago	2019	Drift-cam	Azor
Rainbow 2019	RV Pelagia	2019	Tow-cam	Hopper
iMAR 2021/Eurofleets+	RV Pelagia	2021	Tow-cam	Hopper

Table B.2 – Sampling stations. Each station has the following characteristics: (i) it is represented by the central coordinates of a cell of the model grid (~1 km^2); (ii) it can contain one or more video transects; (iii) it has at least one presence record. Species IDs are relative to Table 3.1.

Station	x	y	Depth	Cruise	Species IDs
1	-29.19	42.46	-1079	iMAR 2021/Eurofleets+	1; 9; 24
2	-29.12	42.46	-1310	iMAR 2021/Eurofleets+	1; 6; 22; 24; 26
3	-29.2	42.45	-1039	iMAR 2021/Eurofleets+	1; 4; 9; 22; 24; 25
4	-29.19	42.45	-1040	iMAR 2021/Eurofleets+	1; 4; 6; 9; 22; 24; 25; 26
5	-29.12	42.45	-1021	iMAR 2021/Eurofleets+	1; 4; 6; 9; 17; 22; 24; 25; 26
6	-29.2	42.44	-855	iMAR 2021/Eurofleets+	1; 4; 6; 9; 22; 24; 25; 26
7	-29.2	42.43	-825	iMAR 2021/Eurofleets+	1; 4; 6; 9; 11; 24; 25; 26
8	-29.51	41.98	-1161	iMAR 2021/Eurofleets+	1; 9; 22; 24
9	-29.51	41.97	-1074	iMAR 2021/Eurofleets+	1; 4; 6; 8; 9; 11; 13; 17; 22; 23; 24; 25; 26
10	-29.51	41.96	-1209	iMAR 2021/Eurofleets+	1; 6; 8; 9; 10; 11; 25
11	-29.51	41.95	-1178	iMAR 2021/Eurofleets+	8
12	-29.43	41.93	-1061	iMAR 2021/Eurofleets+	1; 4; 6; 9; 13; 22; 24; 25; 26
13	-29.43	41.92	-958	iMAR 2021/Eurofleets+	1; 6; 9; 11; 13; 24; 25; 26
14	-29.43	41.91	-967	iMAR 2021/Eurofleets+	1; 4; 6; 9; 11; 24; 26
15	-29.52	40.75	-1119	iMAR 2021/Eurofleets+	1; 6; 9; 24; 25
16	-29.5	40.75	-1020	iMAR 2021/Eurofleets+	1; 4; 6; 9; 17; 24; 25; 26
17	-29.69	40.74	-812	iMAR 2021/Eurofleets+	8; 11; 21; 23; 24; 25
18	-29.52	40.74	-1017	iMAR 2021/Eurofleets+	1; 6; 9; 15; 17; 24; 25; 26
19	-29.69	40.73	-844	iMAR 2021/Eurofleets+	1; 6; 9; 11; 23; 24; 25; 26
20	-29.68	40.73	-1105	iMAR 2021/Eurofleets+	1
21	-29.52	40.73	-974	iMAR 2021/Eurofleets+	1; 6; 9; 15; 22; 24; 25; 26
22	-29.68	40.72	-1264	iMAR 2021/Eurofleets+	1; 9; 22; 25
23	-29.35	40.46	-673	iMAR 2021/Eurofleets+	2; 9; 10; 11; 13; 14; 24; 25; 26
24	-29.36	40.45	-867	iMAR 2021/Eurofleets+	14; 16; 17; 19; 21; 25
25	-29.35	40.45	-732	iMAR 2021/Eurofleets+	21
26	-29.32	40.44	-813	iMAR 2021/Eurofleets+	1; 6; 8; 9; 11; 13; 15; 17; 22; 23; 24; 25; 26
27	-29.32	40.43	-973	iMAR 2021/Eurofleets+	1; 4; 6; 9; 11; 13; 22; 23; 24; 25; 26
28	-29.32	40.42	-1049	iMAR 2021/Eurofleets+	1; 4; 6; 9; 13; 15; 20; 22; 24; 25; 26
29	-29.32	40.41	-1019	iMAR 2021/Eurofleets+	1; 4; 6; 9; 11; 13; 15; 16; 17; 22; 24; 25; 26
30	-29.72	40.4	-1844	iMAR 2021/Eurofleets+	1; 4; 6; 8; 9; 17; 24; 25; 26
31	-29.32	40.4	-1080	iMAR 2021/Eurofleets+	1; 4; 6; 9; 13; 24; 25; 26
32	-29.72	40.39	-1800	iMAR 2021/Eurofleets+	6; 8; 9; 17; 21; 25
33	-29.88	39.74	-1171	iMAR 2021/Eurofleets+	1; 6; 9; 24; 25; 26
34	-29.86	39.72	-1229	iMAR 2021/Eurofleets+	1; 9; 15; 22; 24; 25; 26
35	-29.86	39.71	-1190	iMAR 2021/Eurofleets+	1; 4; 6; 9; 11; 13; 15; 22; 23; 24; 25; 26
36	-29.87	39.7	-874	iMAR 2021/Eurofleets+	1; 6; 8; 9; 11; 13; 15; 17; 19; 20; 22; 23; 24; 25; 26

Station	x	y	Depth	Cruise	Species IDs
37	-29.86	39.7	-1219	iMAR 2021/Eurofleets+	1; 6; 9; 11; 23; 24; 25; 26
38	-29.87	39.69	-807	iMAR 2021/Eurofleets+	1; 8; 11; 23; 24; 25
39	-29.44	39.63	-916	MapGES 2019	2; 8; 10; 11; 13; 14; 25
40	-29.41	39.62	-725	MapGES 2019	2; 8; 10; 11; 25; 26
41	-29.4	39.62	-787	MapGES 2019	2; 8; 10; 11
42	-29.38	39.62	-901	MapGES 2019	2; 8; 10; 15
43	-29.34	39.62	-874	MapGES 2019	2; 8; 11; 13
44	-29.38	39.61	-739	MapGES 2019	2; 8; 10; 11; 13; 25
45	-29.37	39.61	-643	MapGES 2019	8; 14; 15; 21
46	-29.38	39.6	-756	MapGES 2019	8; 13; 14; 15; 16; 19; 21; 25
47	-29.37	39.6	-584	MapGES 2019	8; 14; 16; 17; 21; 25
48	-29.36	39.6	-600	MapGES 2019	8; 13; 14; 19
49	-29.34	39.6	-588	MapGES 2019	14
50	-29.33	39.6	-627	MapGES 2019	8
51	-29.32	39.6	-769	MapGES 2019	8; 13; 15; 17; 25
52	-29.36	39.59	-691	MapGES 2019	2; 14
53	-29.34	39.59	-599	MapGES 2019	8; 14
54	-29.32	39.59	-620	MapGES 2019	8; 13; 14; 15; 21; 25
55	-29.7	39.3	-1055	iMAR 2021/Eurofleets+	1; 10; 11; 22; 23; 24; 25; 26
56	-29.71	39.29	-1029	iMAR 2021/Eurofleets+	24
57	-29.7	39.29	-991	iMAR 2021/Eurofleets+	1; 8; 9; 10; 11; 19; 20; 22; 23; 24; 25; 26
58	-29.78	39.27	-1276	iMAR 2021/Eurofleets+	1; 13; 22; 25; 26
59	-29.71	39.28	-917	iMAR 2021/Eurofleets+	1; 8; 9; 11; 17; 22; 23; 24; 25; 26
60	-29.78	39.26	-1253	iMAR 2021/Eurofleets+	1; 4; 6; 9; 13; 22; 24; 25; 26
61	-29.71	39.27	-830	iMAR 2021/Eurofleets+	1; 19; 22; 23; 24; 25; 26
62	-29.78	39.25	-1146	iMAR 2021/Eurofleets+	1; 6; 9; 13; 15; 23; 24; 25; 26
63	-29.91	39.05	-702	MapGES 2019	15; 17; 19; 21; 25; 26
64	-29.91	39.04	-686	MapGES 2019	19; 21; 25
65	-29.92	39.01	-789	Blue Azores 2018	2; 10; 11; 13; 14; 15; 16; 17; 19; 21; 24; 25; 26
66	-29.9	39.01	-752	MapGES 2019	3; 14; 15; 18; 19; 21
67	-29.92	39	-550	Blue Azores 2018	3; 5; 7; 14; 15; 17; 18; 19; 21; 25
68	-29.9	39	-456	MapGES 2019	3; 5; 14; 15; 16; 18; 19; 20; 21
69	-29.9	38.99	-275	MapGES 2019	2; 3; 7; 14; 15; 18; 19; 20; 21
70	-29.88	38.99	-366	MapGES 2019	2; 3; 7; 14; 20; 21
71	-29.92	38.98	-733	Blue Azores 2018	3
72	-29.91	38.98	-509	MapGES 2019	3; 5; 6; 14; 15; 16; 18; 19; 20; 21
73	-29.9	38.98	-404	MapGES 2019	6; 14; 15; 18; 21
74	-29.88	38.98	-367	MapGES 2019	2; 3; 7; 14; 21
75	-29.87	38.98	-305	MapGES 2019	2; 3; 7; 12; 14

Station	x	y	Depth	Cruise	Species IDs
76	-29.86	38.98	-373	Blue Azores 2018	14
77	-29.85	38.98	-351	Blue Azores 2018	2; 3; 5; 14; 15; 18; 21
78	-29.83	38.98	-657	Blue Azores 2018	2; 10; 11; 13; 14; 17; 19; 25; 26
79	-30.11	38.97	-994	iMAR 2021/Eurofleets+	1; 4; 9; 13; 15; 22; 23; 24; 25; 26
80	-29.9	38.97	-680	MapGES 2019	6; 14; 15; 16; 21
81	-29.86	38.97	-613	Blue Azores 2018	3; 5; 14; 15; 17; 18; 21; 24
82	-29.84	38.97	-605	Blue Azores 2018	3; 24
83	-30.11	38.96	-937	iMAR 2021/Eurofleets+	1; 13; 21; 22; 23; 24; 25; 26
84	-30.1	38.95	-826	iMAR 2021/Eurofleets+	1; 8; 10; 11; 17; 19; 25; 26
85	-30.12	38.94	-814	iMAR 2021/Eurofleets+	8; 17; 25
86	-30.1	38.94	-933	iMAR 2021/Eurofleets+	2; 17; 19; 21; 25
87	-30	38.92	-691	MapGES 2019	8; 11; 19; 25
88	-30	38.91	-712	MapGES 2019	8; 10; 11; 17; 19; 24; 25; 26
89	-30.11	38.87	-973	iMAR 2021/Eurofleets+	1; 13; 22; 23; 24; 25; 26
90	-30.11	38.86	-1152	iMAR 2021/Eurofleets+	1; 6; 9; 13; 17; 22; 24; 25; 26
91	-30.11	38.85	-1311	iMAR 2021/Eurofleets+	1; 6; 9; 13; 22; 24; 25; 26
92	-30.04	38.75	-757	Blue Azores 2018	11; 14; 19; 25; 26
93	-30.04	38.74	-468	Blue Azores 2018	3; 5; 7; 14; 17; 19
94	-30.02	38.74	-487	MapGES 2019	14; 19; 21
95	-29.98	38.74	-639	MapGES 2019	14; 21
96	-30.02	38.73	-306	MapGES 2019	2; 3; 7; 14; 15; 18; 19; 21
97	-29.99	38.73	-264	MapGES 2019	2; 3; 14; 20; 21
98	-30.22	38.71	-688	Blue Azores 2018	8; 11; 13; 14; 15; 16; 21; 25
99	-30.02	38.72	-455	Blue Azores 2018	5; 14; 15; 16; 19; 21; 25
100	-30	38.72	-389	Blue Azores 2018; MapGES 2019	3; 5; 14; 15; 19; 21
101	-29.99	38.72	-361	MapGES 2019	5; 14; 15; 16; 19; 21
102	-29.96	38.72	-359	Blue Azores 2018	2; 3; 5; 10; 14; 15; 18; 19; 20; 21
103	-29.95	38.72	-625	Blue Azores 2018	15; 17; 21
104	-30.24	38.7	-619	Blue Azores 2018; MapGES 2019	3; 5; 6; 8; 11; 14; 15; 16; 21
105	-30.22	38.7	-682	Blue Azores 2018	3; 5; 8; 10; 13; 14; 15; 16; 21; 25
106	-29.99	38.71	-667	MapGES 2019	14; 19
107	-30.25	38.69	-647	MapGES 2019	3; 14; 15; 19; 21
108	-30.24	38.69	-503	MapGES 2019	3; 5; 14; 15; 16; 17; 18; 19; 20; 21
109	-30.25	38.68	-605	MapGES 2019	5; 8; 14; 15; 16; 21
110	-30.24	38.68	-724	MapGES 2019	5; 14; 15; 16; 21
111	-30.09	38.69	-616	MapGES 2019	25
112	-30.25	38.67	-591	MapGES 2019	5; 8; 13; 14; 15; 21; 25
113	-30.11	38.68	-757	MapGES 2019	14; 15; 16; 19; 21

Station	x	y	Depth	Cruise	Species IDs
114	-30.11	38.67	-542	MapGES 2019	8; 13; 14; 16; 21
115	-27.35	38.67	-626	Nico 2018	14
116	-26.86	38.67	-422	EMEPC/LUSO/G3	2; 7; 8; 14
117	-27.35	38.66	-894	Nico 2018	1; 8
118	-27.31	38.64	-793	Nico 2018	4; 8; 13
119	-28.12	38.61	-908	EMEPC/LUSO/G3	6; 8; 9
120	-28.1	38.61	-505	EMEPC/LUSO/G3	6; 8; 9; 11; 14
121	-28.09	38.59	-1125	EMEPC/LUSO/G3	1; 4; 6; 8; 9; 13
122	-28.08	38.59	-1072	EMEPC/LUSO/G3	8
123	-29.05	38.55	-222	Greenpeace Pole-to-Pole	3; 14
124	-27.45	38.55	-1094	Nico 2018	1
125	-27.44	38.55	-1278	Nico 2018	1
126	-27.93	38.53	-649	EMEPC/LUSO/G3	1; 6; 8; 9; 14
127	-27.47	38.54	-788	Nico 2018	10; 11
128	-27.45	38.54	-976	Nico 2018	1; 8; 10; 11
129	-27.47	38.53	-744	Nico 2018	1; 2; 8; 10; 11; 14
130	-27.48	38.52	-1200	Nico 2018	1
131	-27.47	38.52	-1063	Nico 2018	1; 8
132	-31.87	38.41	-1121	Meteor Cruise M 58/3	9
133	-31.02	38.33	-632	MapGES 2019	2; 15; 16; 21; 25
134	-30.69	38.34	-581	MapGES 2019	17; 25
135	-30.67	38.34	-667	MapGES 2019	5; 13; 14; 15; 16; 18; 19; 21; 25
136	-30.66	38.34	-905	iMAR 2021/Eurofleets+	1; 11; 13; 16; 17; 25; 26
137	-31.05	38.32	-687	MapGES 2019	2; 10; 11
138	-31.03	38.32	-633	MapGES 2019	2; 5; 8; 10; 11; 13; 14; 15; 17; 20; 21; 25
139	-31.02	38.32	-591	MapGES 2019	5; 8; 13; 14; 15; 16; 21; 25
140	-30.69	38.33	-539	MapGES 2019	5; 8; 14; 15; 16; 21; 25; 26
141	-30.67	38.33	-843	iMAR 2021/Eurofleets+	1; 22; 25; 26
142	-30.63	38.33	-627	MapGES 2019	5; 13; 14; 15; 16; 17; 19; 21; 25
143	-31.06	38.31	-760	MapGES 2019	10; 11; 15
144	-31.05	38.31	-654	MapGES 2019	2; 5; 8; 10; 11; 13; 14; 15; 17; 21; 25; 26
145	-30.69	38.32	-645	iMAR 2021/Eurofleets+	11; 25; 26
146	-30.67	38.32	-799	iMAR 2021/Eurofleets+	22; 23; 25; 26
147	-30.66	38.32	-615	MapGES 2019	4; 5; 13; 14; 15; 16; 17; 21; 22; 24; 25
148	-30.72	38.31	-455	MapGES 2019	5; 14; 15; 16; 17; 18; 19; 21
149	-30.71	38.31	-423	MapGES 2019	3; 12; 14; 15; 19; 21
150	-30.68	38.31	-691	iMAR 2021/Eurofleets+	6; 9; 11; 17; 18; 22; 23; 25; 26
151	-30.67	38.31	-718	MapGES 2019	17; 22; 25
152	-30.66	38.31	-546	MapGES 2019	3; 5; 14; 15; 16; 19; 21

Station	x	y	Depth	Cruise	Species IDs
153	-30.54	38.31	-850	iMAR 2021/Eurofleets+	1; 9; 11; 15; 24; 25; 26
154	-28.28	38.36	-814	Blue Azores 2018	24
155	-30.75	38.3	-737	MapGES 2019	24; 26
156	-30.71	38.3	-518	MapGES 2019	14
157	-30.7	38.3	-714	MapGES 2019	14; 15; 16
158	-30.65	38.3	-415	MapGES 2019	5; 14; 15; 18; 19; 21
159	-30.56	38.3	-955	iMAR 2021/Eurofleets+	15; 24; 25
160	-30.54	38.3	-749	iMAR 2021/Eurofleets+	2; 10; 11; 15; 17; 18; 19; 20; 21; 22; 23; 24; 25; 26
161	-30.12	38.32	-458	MapGES 2019	10; 11; 13; 14; 15; 17; 19; 21; 24; 25; 26
162	-28.27	38.35	-880	Blue Azores 2018	24
163	-30.74	38.29	-501	MapGES 2019	5; 14; 15; 16; 19; 21; 25
164	-30.72	38.29	-493	MapGES 2019	2; 3; 7; 12; 14; 15; 16; 18; 19; 20; 21; 25
165	-30.63	38.29	-440	MapGES 2019	5; 8; 14; 25
166	-30.55	38.29	-795	iMAR 2021/Eurofleets+	11; 24
167	-30.09	38.31	-473	MapGES 2019	2; 3; 12; 14; 15; 16; 18; 19; 21
168	-30.1	38.29	-527	MapGES 2019	14; 15; 16; 19; 21
169	-30.09	38.29	-578	MapGES 2019	2; 14; 15; 16; 17; 19; 21; 25
170	-30.11	38.28	-756	MapGES 2019	10; 15; 21; 25; 26
171	-30.09	38.28	-661	MapGES 2019	2; 14
172	-30.33	38.21	-702	MapGES 2019	2; 8; 10; 11; 13; 17; 21; 25; 26
173	-30.34	38.2	-641	MapGES 2019	2; 8; 10; 11; 14; 17; 19; 21; 25; 26
174	-30.33	38.2	-589	MapGES 2019	8; 10; 11; 13; 17; 21; 25
175	-30.42	38.19	-830	MapGES 2019	6; 17; 24; 25
176	-30.34	38.19	-571	MapGES 2019	2; 5; 8; 14; 15; 16; 21
177	-30.33	38.19	-611	MapGES 2019	2; 14
178	-30.32	38.19	-652	MapGES 2019	8; 10; 11; 13; 15; 17; 18; 19; 21; 22; 24; 25
179	-30.3	38.19	-780	MapGES 2019	10; 11; 25
180	-26.62	38.24	-406	MapGES 2018	2; 3; 14
181	-26.59	38.24	-373	MapGES 2018	2; 14
182	-26.57	38.24	-586	EMEPC/LUSO/G3	14
183	-26.56	38.24	-740	EMEPC/LUSO/G3	2; 4; 8; 11; 14
184	-26.22	38.23	-502	MapGES 2018	2; 14
185	-30.42	38.18	-738	MapGES 2019	6; 8; 10; 11; 15; 21; 25
186	-26.22	38.22	-624	MapGES 2018	14
187	-26.21	38.22	-659	MapGES 2018	14
188	-26.57	38.21	-407	MapGES 2018	2; 3; 12; 14
189	-26.56	38.21	-516	MapGES 2018	2; 3; 12; 14
190	-31.23	37.87	-722	MapGES 2019	5; 14; 15; 16; 21; 25
191	-31.22	37.87	-674	MapGES 2019	2; 14; 15; 16; 20; 21

Station	x	y	Depth	Cruise	Species IDs
192	-31.52	37.85	-960	EMEPC/LUSO/G3	4
193	-31.23	37.86	-534	MapGES 2019	2; 14; 15; 16; 19; 20
194	-31.21	37.86	-567	MapGES 2019	14; 21
195	-31.23	37.85	-618	MapGES 2019	2; 14; 25
196	-31.24	37.84	-827	MapGES 2019	8; 13; 14; 15; 16; 19; 21; 25
197	-31.23	37.84	-641	MapGES 2019	2; 8; 13; 14; 15; 21; 25
198	-30.85	37.78	-779	MapGES 2019	8; 13; 14; 15; 25
199	-30.84	37.78	-687	MapGES 2019	8; 14; 15; 16; 17; 18; 21; 25
200	-30.85	37.77	-515	MapGES 2019	8; 13; 14; 15; 16; 19; 21; 25
201	-30.84	37.77	-583	MapGES 2019	14; 15; 16; 21; 25
202	-30.84	37.76	-650	MapGES 2019	14; 15; 16; 21; 25
203	-25.81	37.72	-432	EMEPC/LUSO/G3	2; 3; 7; 14
204	-25.81	37.71	-600	EMEPC/LUSO/G3	14
205	-30.79	37.53	-555	MapGES 2019	14; 15; 19; 21
206	-30.77	37.53	-419	MapGES 2019	14; 15; 19; 21
207	-30.76	37.53	-458	MapGES 2019	3; 7; 12; 14; 15; 18; 19; 20; 21
208	-30.7	37.53	-438	MapGES 2019	14; 15; 18; 19; 20; 21
209	-30.77	37.52	-412	MapGES 2019	2; 3; 7; 9; 12; 14; 15; 18; 19; 20; 21
210	-30.76	37.52	-301	MapGES 2019	2; 3; 7; 12; 14; 15; 18; 20; 21
211	-30.72	37.52	-413	MapGES 2019	2; 3; 7; 12; 14; 15; 17; 18; 19; 20; 21; 25
212	-30.7	37.52	-592	MapGES 2019	2; 14; 15; 16; 18; 19; 20; 21
213	-30.59	37.52	-495	MapGES 2019	14; 15; 16; 18; 19; 20; 21
214	-30.58	37.52	-472	MapGES 2019	2; 14; 15; 16; 18; 19; 20; 21
215	-30.72	37.51	-616	MapGES 2019	15; 16; 19; 20; 21
216	-30.58	37.51	-593	MapGES 2019	14; 15; 19
217	-30.85	37.49	-532	MapGES 2019	2; 5; 14; 15; 16; 17; 18; 19; 20; 21
218	-30.81	37.49	-733	MapGES 2019	15; 16; 17; 19; 21; 25
219	-32.9	37.4	-523	Rainbow 2019	2; 5; 12; 14; 15; 16; 18; 19; 21
220	-30.86	37.48	-598	MapGES 2019	2; 5; 15; 16; 19; 20; 21; 25
221	-30.85	37.48	-474	MapGES 2019	2; 5; 7; 12; 14; 15; 16; 18; 19; 20; 21
222	-30.84	37.48	-632	MapGES 2019	2; 12; 14; 15; 19; 20; 21
223	-32.97	37.38	-509	Rainbow 2019	3; 12; 14; 15; 16; 17; 18; 19; 20; 21
224	-32.91	37.39	-585	Rainbow 2019	2; 5; 8; 12; 13; 14; 15; 21
225	-32.9	37.39	-528	Rainbow 2019	2; 5; 12; 14; 15; 18; 19; 21
226	-32.99	37.37	-469	Rainbow 2019	3; 14; 19; 21
227	-32.97	37.37	-639	Rainbow 2019	3; 12; 14; 15; 19; 20; 21
228	-32.92	37.38	-836	Rainbow 2019	6; 10; 11; 17; 18; 21; 24; 25; 26
229	-32.91	37.38	-666	Rainbow 2019	2; 5; 8; 10; 11; 13; 14; 15; 16; 17; 18; 24; 25
230	-32.03	37.41	-739	MapGES 2019	14; 16

Station	x	y	Depth	Cruise	Species IDs
231	-32.04	37.4	-551	MapGES 2019	5; 14; 15; 16; 21; 25
232	-32.03	37.4	-571	MapGES 2019	5; 14
233	-32.04	37.39	-562	MapGES 2019	5; 14; 15; 21
234	-24.77	37.35	-1394	MedWaves	1
235	-24.74	37.34	-1422	MedWaves	1
236	-24.57	37.34	-971	MedWaves	9
237	-24.77	37.34	-1085	MedWaves	1
238	-24.74	37.33	-1274	MedWaves	1; 8; 9; 11
239	-31.38	37.18	-582	MapGES 2019	15; 21
240	-31.47	37.17	-525	MapGES 2019	14; 15; 17; 20; 21
241	-31.46	37.17	-465	MapGES 2019	14; 15; 20; 21
242	-31.43	37.17	-396	MapGES 2019	14; 15; 16; 18; 19; 20; 21
243	-24.67	37.21	-607	MedWaves	4; 8; 10; 11
244	-24.65	37.21	-645	MedWaves	1; 4; 8; 10; 11
245	-24.62	37.21	-824	MedWaves	4; 8; 10; 11
246	-24.63	37.2	-912	MedWaves	1; 4; 6; 8; 9; 10; 11
247	-24.62	37.2	-952	MedWaves	1; 8; 9; 10; 11
248	-24.63	37.19	-1185	MedWaves	1; 4; 6; 8; 9; 11
249	-24.63	37.18	-1212	MedWaves	1; 6; 10; 11
250	-24.63	37.17	-1261	MedWaves	1; 4; 6; 7; 8; 11; 24
251	-24.63	37.16	-1333	MedWaves	1; 6; 11
252	-25.64	37.17	-336	MapGES 2018	3; 14
253	-24.63	37.15	-1500	MedWaves	1; 4; 6
254	-25.65	37.15	-361	MapGES 2018	3; 14
255	-25.64	37.15	-468	MapGES 2018	3; 14
256	-32.55	36.82	-827	Nico 2018	2; 6; 8; 10; 11; 13; 15; 17; 19; 21; 22; 24; 25; 26
257	-32.54	36.82	-605	Nico 2018	2; 6; 8; 13; 21; 25
258	-32.56	36.81	-678	Nico 2018	2; 10; 11; 13; 15; 17; 21; 22; 24; 25
259	-32.55	36.81	-732	Nico 2018	8; 10; 11; 13; 16; 17; 19; 21; 22; 24; 25; 26
260	-33.87	36.45	-1026	Treasure 2014	1; 17; 24
261	-33.87	36.44	-818	Treasure 2014	1; 8; 17; 18; 19; 21; 24; 25; 26
262	-33.86	36.44	-608	Treasure 2016	8; 14; 16; 21
263	-33.87	36.43	-538	Treasure 2014	5; 14; 15; 17; 18; 19; 20; 21
264	-33.88	36.42	-651	Treasure 2016	5; 14; 15; 16; 19
265	-33.87	36.42	-586	Treasure 2016	5; 14; 15; 19
266	-33.89	36.41	-880	Treasure 2016	14
267	-33.88	36.41	-613	Treasure 2016	2; 5; 8; 14; 15; 16; 17; 19; 21; 26
268	-33.9	36.4	-1123	Treasure 2016	24
269	-33.89	36.4	-774	Treasure 2016	1; 2; 8; 9; 11; 14; 15; 16; 17; 19; 21; 24; 25; 26

Station	x	y	Depth	Cruise	Species IDs
270	-33.89	36.39	-1096	Treasure 2016	8; 9; 17; 21; 22; 24; 25; 26
271	-33.77	36.04	-1334	Treasure 2016	24; 26
272	-33.76	36.04	-907	Treasure 2016	6; 24; 26
273	-33.76	36.03	-804	Treasure 2016	2; 6; 13; 17; 19; 21; 23; 24; 25; 26
274	-27.45	34.92	-1159	Biometore	1
275	-30.31	34.15	-366	Biometore	2; 8; 9; 12; 14
276	-30.32	34.1	-721	Biometore	1
277	-28.38	33.99	-1101	Biometore	9
278	-28.36	33.97	-837	Biometore	4; 9
279	-28.35	33.96	-781	Biometore	9
280	-30.15	33.91	-991	Biometore	1
281	-28.35	33.95	-835	Biometore	9

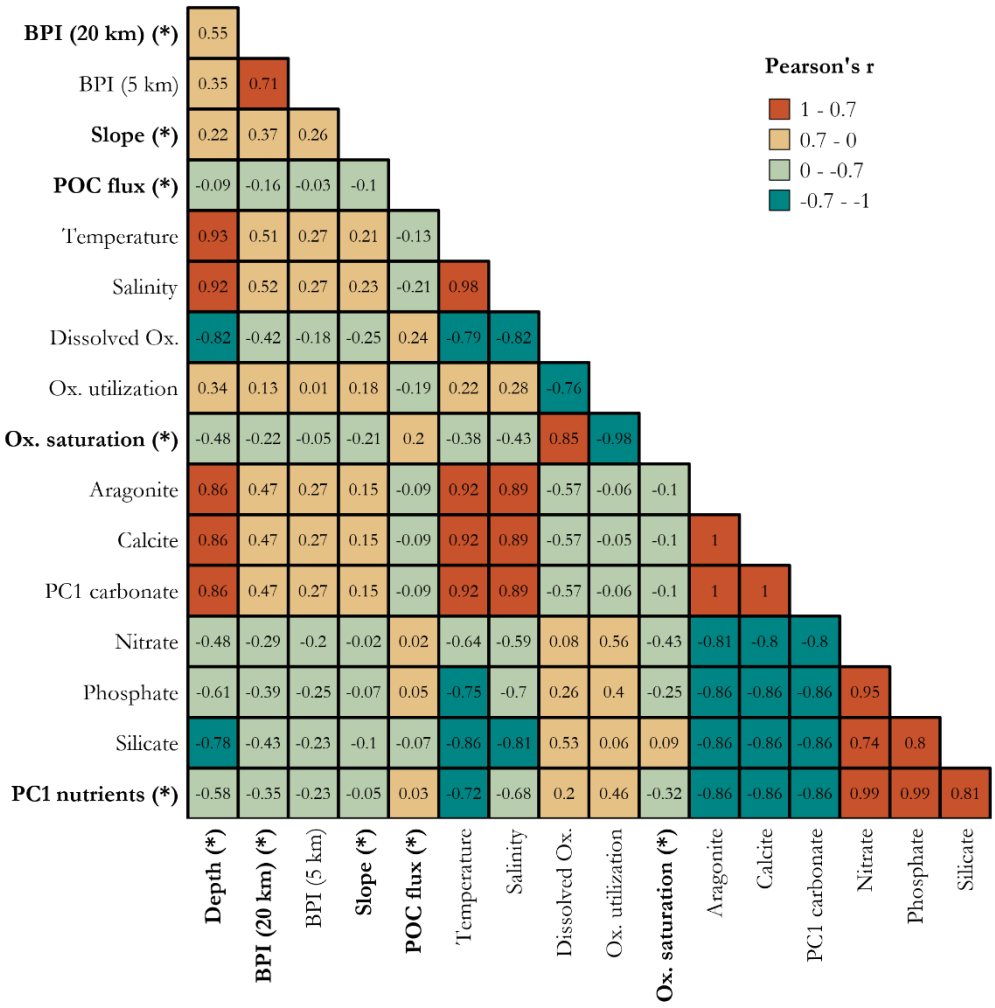

Figure B.1 – Pair-wise Pearson's correlation coefficient (r) among all candidate variables. (*) Variables used to select and develop the species archetype models. Abbreviations and units of measurement: depth (m), bathymetric position index (BPI), slope (º), particulate organic carbon (POC) flux (mg C_{org} m^{-2} d^{-1}), temperature (ºC), dissolved oxygen (dissolved ox.; mL L^{-1}), oxygen utilization (ox. utilization; mL L^{-1}), oxygen saturation (ox. saturation; %), aragonite (Ω_{ar}), calcite (Ω_{ca}), 1st principal component of aragonite and calcite (PC1 carbonate), nitrate (µmol/l), phosphate (µmol/l), silicate (µmol/l) and 1st principal component of water nutrients (PC1 nutrients).

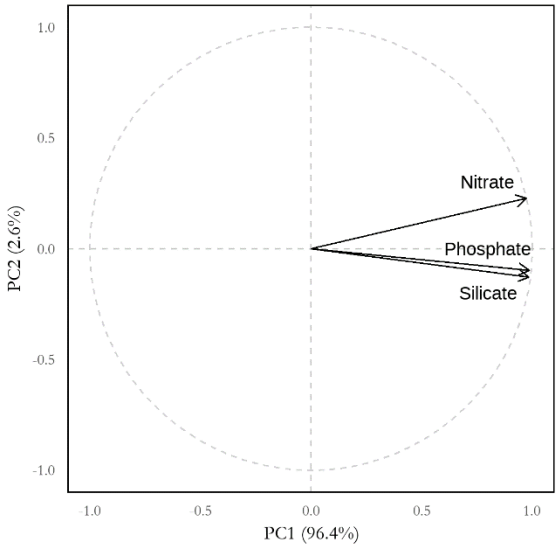

Figure B.2 – Circle of correlation after the principal component analysis of nutrient concentrations (nitrate, phosphate and silicate) with contributions of variables to PC1 and PC2 axes.

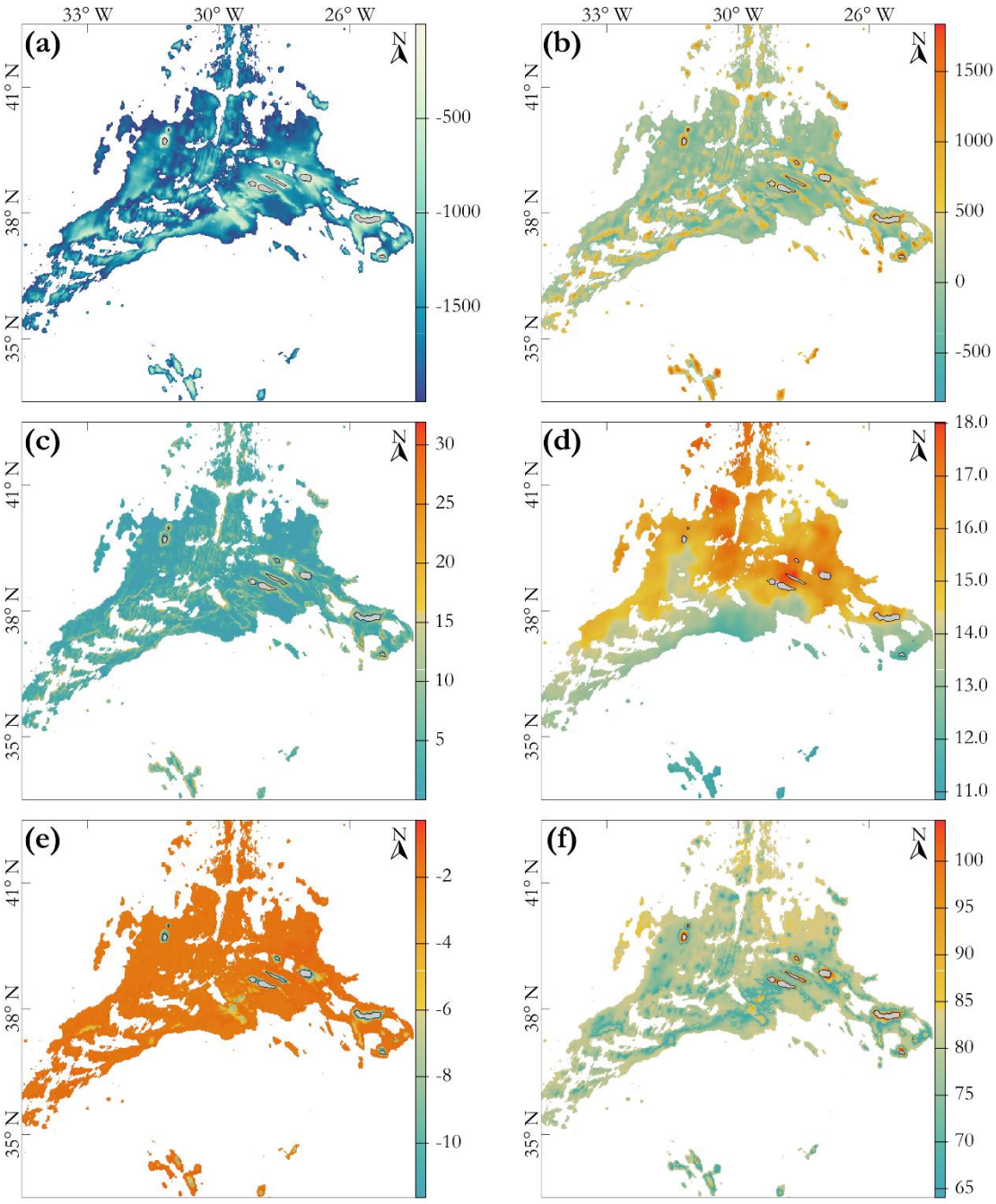

Figure B.3 – Environmental layers used to select and develop the species archetype model. (a) depth (m); (b) bathymetric position index with a 20 km radius (BPI20); (c) slope (º); (d) particulate organic carbon (POC) flux (mg C_{org} m^{-2} d^{-1}); (e) first principal component of the principal component analysis of water nutrients (nitrate, phosphate and silicate) (PC1 nutrients); bottom oxygen saturation (%). The plots display only locations shallower than 2000 m.

B.2 Dummy absences

Table B.3 – Dummy absences used to develop the species archetype models. Abbreviations and units of measurement: depth (m); bathymetric position index with a 20 km radius (BPI20); slope (º); particulate organic carbon (POC) flux (mg C_{org} m^{-2} d^{-1}); first principal component of the principal component analysis of water nutrients (nitrate, phosphate and silicate) (PC1 Nu); bottom oxygen saturation (BOX; %).

ID	x	y	Depth	POC	Slope	BPI20	PC1 Nu	BOX
1	-31.53	39.22	-1470.56	14.73	0.5	25.28	-2.09	78.61
2	-27.5	39.53	-1746.4	15.76	0.76	-35.05	-1.87	83.36
3	-30.15	40.2	-1777.13	17.39	0.19	-88.41	-1.97	82.22
4	-28.03	38.05	-1832.16	13.01	0.28	-38.89	-1.94	82.79
5	-29.74	37.54	-1473.6	12.05	0.37	56.16	-1.94	72.41

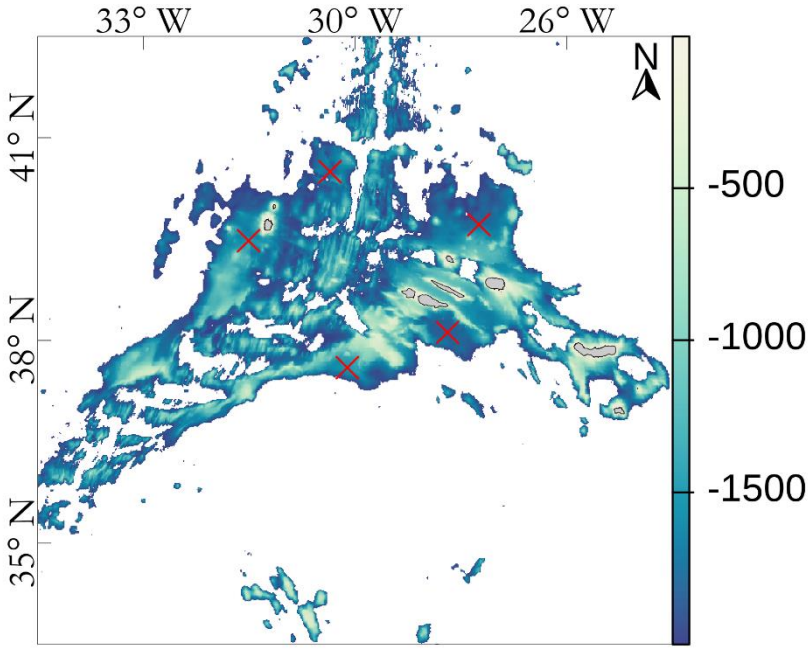

Figure B.4 – Locations of the dummy absences. In the background depth values (m).

B.3 Model selection

Table B.4 – Forward selection of covariate terms with four archetype groups. The best model is highlighted in bold. Abbreviations: bathymetric position index with a 20 km radius (BPI20), particulate organic carbon (POC) flux, first principal component of water nutrients (PC1_Nu), bottom oxygen saturation (B_Ox), linear term ([…].1), quadratic term ([…].2).

Order	AIC	BIC	Formula
1	6665.434	6679.273	~1 + Depth.1
2	7265.145	7278.984	~1 + POC.1
3	7345.18	7359.019	~1 + Slope.1
4	7268.388	7282.227	~1 + BPI20.1
5	6694.281	6708.12	~1 + PC1_Nu.1
6	7194.32	7208.159	~1 + B_Ox.1
7	6378.038	6396.91	~1 + Depth.1 + Depth.2
8	7111.653	7130.524	~1 + POC.1 + POC.2
9	7300.102	7318.973	~1 + Slope.1 + Slope.2
10	7100.954	7119.825	~1 + BPI20.1 + BPI20.2
11	6566.099	6584.97	~1 + PC1_Nu.1 + PC1_Nu.2
12	7179.786	7198.658	~1 + B_Ox.1 + B_Ox.2
13	6317.565	6341.469	~1 + Depth.1 + Depth.2 + POC.1
14	6352.609	6376.513	~1 + Depth.1 + Depth.2 + Slope.1
15	6359.739	6383.643	~1 + Depth.1 + Depth.2 + BPI20.1
16	6359.812	6383.716	~1 + Depth.1 + Depth.2 + PC1_Nu.1
17	6338.678	6362.581	~1 + Depth.1 + Depth.2 + B_Ox.1
18	6279.355	6308.291	~1 + Depth.1 + Depth.2 + POC.1 + POC.2
19	6315.514	6344.45	~1 + Depth.1 + Depth.2 + Slope.1 + Slope.2
20	6263.635	6292.571	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2
21	6317.991	6346.927	~1 + Depth.1 + Depth.2 + PC1_Nu.1 + PC1_Nu.2
22	6317.752	6346.688	~1 + Depth.1 + Depth.2 + B_Ox.1 + B_Ox.2
23	6194.54	6228.509	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1

Order	AIC	BIC	Formula
24	6237.729	6271.698	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + Slope.1
25	6240.929	6274.897	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + PC1_Nu.1
26	6225.765	6259.733	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + B_Ox.1
27	6174.48	6213.481	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2
28	6220.147	6259.148	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + Slope.1 + Slope.2
29	6203.672	6242.673	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + PC1_Nu.1 + PC1_Nu.2
30	6206.785	6245.786	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + B_Ox.1 + B_Ox.2
31	6144.278	6188.312	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + Slope.1
32	6151.993	6196.026	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + PC1_Nu.1
33	6140.086	6184.12	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1
34	6130.81	6179.876	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + Slope.1 + Slope.2
35	6127.688	6176.753	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + PC1_Nu.1 + PC1_Nu.2
36	6123.506	6172.572	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1 + B_Ox.2
37	6091.94	6146.038	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1 + B_Ox.2 + Slope.1
38	6119.491	6173.589	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1 + B_Ox.2 + PC1_Nu.1
39	6076.733	**6135.864**	**~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1 + B_Ox.2 + Slope.1 + Slope.2**
40	6119.011	6178.142	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1 + B_Ox.2 + PC1_Nu.1 + PC1_Nu.2
41	6074.108	6138.271	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1 + B_Ox.2 + Slope.1 + Slope.2 + PC1_Nu.1
42	6072.866	6142.061	~1 + Depth.1 + Depth.2 + BPI20.1 + BPI20.2 + POC.1 + POC.2 + B_Ox.1 + B_Ox.2 + Slope.1 + Slope.2 + PC1_Nu.1 + PC1_Nu.2

Appendix C
Supplementary data Chapter 4

C.1 R code to identify the Island Shelf Unit (ISU)

```
##########################################################################################
##############

# User-defined classification of raster surfaces: The R package scapesClassification


# ISU script

# Author: Gerald Hechter Taranto


# A detailed explanation of all classification steps is available on the documentation page

# https://ghtaranto.github.io/scapesClassification/index.html



# Results are plotted only if PLOT is TRUE

PLOT <- T

##########################################################################################


#### 1. INSTALL REQUIRED PACKAGES ####

# install.packages("scapesClassification")

# install.packages("terra")


# Required for interactive maps

# install.packages("raster")

# install.packages("mapview")
```

```
# install.packages("leaflet")

#### 2. LOAD LIBRARIES ####
library(scapesClassification)
library(terra)



#### 3. PLOTS & INTERACTIVE MAPS
library(mapview)
library(leaflet)
mapviewOptions(basemaps = c("Esri.OceanBasemap"), na.color = "grey88",
               homebutton = FALSE, query.position = "bottomright", legend.opacity = 0.9,
               fgb = FALSE)

# palette for background raster
palRYB <- c("#2892c7", "#fafa64", "#e81014")
palYGB <- c("#ffff80" ,"#3dba65", "#0d1178")
palBYR <- c("#4575b5", "#ffffbf", "#d62f27")

#### 3. LOAD DATA ####
grd <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T)
grd <- grd[grepl("\\.grd", grd)]
grd <- grd[!grepl("hillshade", grd)]
rstack <- rast(grd)

if(PLOT){
  plot(rstack, mar = c(3, 1.6, 1.6, 4), cex.main= 1.1,
       main = c("(a) Bathymetry", "(b) Local BPI", "(c) Regional BPI", "(d) Slope"))
```

```
  readline(prompt = "Plot input data. Press [enter] to continue.")

}

##############################################################################################

#### 4. FORMAT INPUTS ####

##############################################################################################

# COMPUTE ATTRIBUTE TABLE

atbl <- attTbl(rstack, var_names = c("bathymetry", "local_bpi", "regional_bpi", "slope"))

# COMPUTE NEIGHBORHOOD LIST

nbs <- ngbList(rstack, rNumb = TRUE, attTbl = atbl) # neighbors are identified by row numbers (see
?ngbList)

##############################################################################################

#### 5. ISLAND SHELF UNIT (ISU)

##############################################################################################

# LOAD SHAPE FILE

shp <- system.file("extdata", "Azores.shp", package = "scapesClassification")

# EXTRACT LAND POSITIONS

anchorLAND <-

  anchor.svo(r = rstack, dsn = shp, only_NAs = TRUE, fill_NAs = TRUE)

# IDENTIFY ANCHOR CELLS

anchorCELL <-
```

```
 anchor.cell(attTbl = atbl, r = rstack, anchor = anchorLAND, class = 1,
          class2cell = FALSE, # class not attributed to land cells
          class2nbs  = TRUE)  # class attributed to cell adjacent to land cells


# FLAT CELLS
flatCELLS <-
 cond.4.nofn(atbl, nbs, rNumb = TRUE, classVector = anchorCELL, # update previous class vector
          class = 2,         # flat cell class
          nbs_of = c(1, 2),   # focal cells
          cond = "slope <= 5",# condition (slope refers to the column `slope` of `atbl`)
          min.bord = 0.2)     # a test cell is classified only if cond=TRUE AND
# if at least 20% of its neighbors belong to a shelf cell class
# HOLE CELLS
holeCELLS <-
 cond.4.nofn(atbl, nbs, rNumb = TRUE, classVector = flatCELLS, # update previous class vector
          class = 3,         # hole cell class
          nbs_of = c(1, 2, 3),# focal cells
          cond = "TRUE",      # conditions are always true
          min.bord = 0.6)     # a test cell is classified only if 60% of its neighbors
# belong to a shelf cell class
# SLOPE CELLS
slopeCELLS <-
 cond.4.nofn(atbl, nbs, rNumb = TRUE, classVector = holeCELLS, # update previous class vector
          class = 4,          # slope cell class
          nbs_of = c(1,2,3,4), # focal cells
          cond = "regional_bpi > 100 | local_bpi{} > 100", # conditions
          peval = 0.4)    # minimum number of positive evaluations (neighborhood condition, ?conditions)
```

```
#############################################################################################
# 5.1 ISU Interactive map
#############################################################################################

if(PLOT){

 # Note that all raster classifications are converted into polygons (as.polygons()) before plotting
 # for a better visualization. Note also that mapview do not support the terra package yet, therefore
 # all objects of class terra are converted into objects of class raster before plotting
 # (as(xx,"Spatial") OR raster::raster(XX)).

 # Land shape file
 island <- terra::vect(shp)

 # Class vector slopeCELLS contains all classes as it is the last one to be updated
 unique(slopeCELLS)

 # class = 1, Anchor cell
 # class = 2, Flat cell
 # class = 3, Hole cell
 # class = 4, Slope cell

 # Class vector 2 raster
 ISU  <- cv.2.rast(rstack, slopeCELLS)
 LAND <- rstack[[1]]; LAND[]  <- NA; LAND[anchorLAND] <- 1

 # Raster 2 polygon
 ISU  <- as.polygons(ISU, dissolve = TRUE)
```

```
  LAND <- as.polygons(LAND, dissolve = TRUE)
  LAND$dummy <- 1 #trick map label


  m <-
    mapview(as(ISU,"Spatial"), alpha.regions = 0.8, col.regions =
c("orange","#FFD700","#A0522D","cyan"),
          layer.name = "ISU", popup = NULL, label = "ISU") +
    mapview(as(LAND,"Spatial"), alpha.regions = 1,
          col.regions = "black", layer.name = "Land cell", popup = NULL, label = "LAND") +
    mapview(raster::raster(rstack[["regional_bpi_exm"]]), layer.name = "Regional BPI", col.regions =
palBYR) +
    mapview(raster::raster(rstack[["local_bpi_exm"]]), layer.name = "Local BPI", col.regions = palBYR) +
    mapview(raster::raster(rstack[["slope_exm"]]), layer.name = "Slope (deg)", col.regions = palYGB)


  m@map %>%
    hideGroup(c("Local BPI", "Slope (deg)")) %>%
    addLayersControl(overlayGroups = c("Land cell", "ISU", "Regional BPI", "Local BPI", "Slope (deg)"),
                  position = "topleft",
                  options = layersControlOptions(collapsed = TRUE))%>%
    addMiniMap(tiles = "Esri.OceanBasemap",
             position = "bottomleft",
             toggleDisplay = TRUE,
             zoomLevelOffset = -5)


}
```

C.2 R code to identify the Peak Cells (PKS)

```
# User-defined classification of raster surfaces: The R package scapesClassification

# PKS script
# Author: Gerald Hechter Taranto

# A detailed explanation of all classification steps is available on the documentation page
# https://ghtaranto.github.io/scapesClassification/index.html


# Results are plotted only if PLOT is TRUE
PLOT <- T
##############################################################################################


#### 1. INSTALL REQUIRED PACKAGES ####
# install.packages("scapesClassification")
# install.packages("terra")

# Required for interactive maps
# install.packages("raster")
# install.packages("mapview")
# install.packages("leaflet")

#### 2. LOAD LIBRARIES ####
library(scapesClassification)
```

```
library(terra)



#### 3. PLOTS & INTERACTIVE MAPS
library(mapview)
library(leaflet)
mapviewOptions(basemaps = c("Esri.OceanBasemap"), na.color = "grey88",
               homebutton = FALSE, query.position = "bottomright", legend.opacity = 0.9,
               fgb = FALSE)

# palette for background raster
palRYB <- c("#2892c7", "#fafa64", "#e81014")
palYGB <- c("#ffff80" ,"#3dba65", "#0d1178")
palBYR <- c("#4575b5", "#ffffbf", "#d62f27")

# hillShade
hs <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T)
hs <- rast(hs[grepl("hillshade\\.grd", hs)])

#### 3. LOAD DATA ####
grd <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T)
grd <- grd[grepl("\\.grd", grd)]
grd <- grd[!grepl("hillshade", grd)]
rstack <- rast(grd)

if(PLOT){
  plot(rstack, mar = c(3, 1.6, 1.6, 4), cex.main= 1.1, col = hcl.colors(50, palette =  "Zissou 1"),
       main = c("(a) Bathymetry", "(b) Local BPI", "(c) Regional BPI", "(d) Slope"))
```

```
  readline(prompt = "Plot input data. Press [enter] to continue.")
}

# Load ISU classification
ISU <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T,
                  pattern = "ISU_Cells\\.RDS")

################################################################################################
#### 4. FORMAT INPUTS ####
################################################################################################

# COMPUTE ATTRIBUTE TABLE
atbl <- attTbl(rstack, var_names = c("bathymetry", "local_bpi", "regional_bpi", "slope"))

# Use ISU classification as an attribute table filter
atbl$ISU <- readRDS(ISU)
atbl$ISU[!is.na(atbl$ISU)] <- 1 # Class 1 identifies all ISU cells

head(atbl)

# COMPUTE NEIGHBORHOOD LIST
nbs <- ngbList(rstack, rNumb = TRUE, attTbl = atbl) # neighbors are identified by row numbers (see
?ngbList)

head(nbs)

################################################################################################
#### 5. PEAK CELLs (PC)
```

```
##############################################################################################

# PEAK CELL
atbl$pc <- peak.cell(atbl, nbs, rNumb = TRUE,
                     p_col = "bathymetry", # column on which local maxima are searched
                     p_edge = FALSE)       # local maxima or minima are not searched on edge cells

# FILTER PEAK CELL

# Remove:
# (i)  peak cell within the ISU           -> !is.na(ISU)
# (ii) peak cell not on prominent features -> BPI < 100

cond <- "regional_bpi < 100 | local_bpi < 100 | !is.na(ISU)"

# Filter
atbl$pc <- cond.reclass(atbl, nbs, rNumb = TRUE, classVector = atbl$pc, # filter class vector atbl$pc
                        cond = cond,  # filter condition
                        class = 1,    # peak cells meeting conditions...
                        reclass = NA) # are reclassified as NA-cells

##############################################################################################
# 5.1 PKS Interactive map
##########################################################################
##############################################################################################

if(PLOT){

  # Note that all raster classifications are converted into polygons (as.polygons()) before plotting
```

```
# for a better visualization. Note also that mapview do not support the terra package yet, therefore
# all objects of class terra are converted into objects of class raster before plotting
# (as(xx,"Spatial") OR raster::raster(XX)).

# Load land cells
lc <- readRDS(list.files(system.file("extdata", package = "scapesClassification"),
                         full.names = T, pattern = "Land_Cells.RDS"))
LAND <- rstack[[1]]; LAND[] <- NA; LAND[lc] <- 1

# Class vector 2 raster
PC <- cv.2.rast(rstack, atbl$pc)
ISU<- cv.2.rast(rstack, atbl$ISU)

# Raster 2 polygon
PC   <- as.polygons(PC,  dissolve = TRUE)
ISU  <- as.polygons(ISU, dissolve = TRUE)
LAND <- as.polygons(LAND,dissolve = TRUE)

PC$dummy   <- 1 #trick map label
ISU$dummy <- 1 #trick map label
LAND$dummy   <- 1 #trick map label

m <-
  mapview(raster::raster(hs), col.regions = c("gray0", "gray50", "grey100"), alpha.regions=1,
          legend = FALSE, label = FALSE) +
  mapview(as(PC,"Spatial"), alpha.regions = 0.8,
          col.regions = "black", layer.name = "Peak cell", popup = NULL, label = "Peak cell") +
  mapview(as(ISU,"Spatial"), alpha.regions = 1, col.regions = c("gray"),
```

```
        layer.name = "ISU", popup = NULL, label = "ISU") +
   mapview(as(LAND,"Spatial"), alpha.regions = 1,
        col.regions = "black", layer.name = "Land cell", popup = NULL, legend = FALSE) +
   mapview(raster::raster(rstack[["regional_bpi_exm"]]), layer.name = "Regional BPI", col.regions =
palRYB) +
   mapview(raster::raster(rstack[["bathymetry_exm"]]), layer.name = "Depth (m)", col.regions = palRYB)


  m@map %>%
   hideGroup(c("Depth (m)")) %>%
   addLayersControl(overlayGroups = c("Peak cell", "ISU", "Regional BPI", "Depth (m)"),
                    position = "topleft",
                    options = layersControlOptions(collapsed = TRUE))
}
```

C.3 R code to identify the Relief Units (RUs)

```
# User-defined classification of raster surfaces: The R package scapesClassification

# RU script
# Author: Gerald Hechter Taranto

# A detailed explanation of all classification steps is available on the documentation page
# https://ghtaranto.github.io/scapesClassification/index.html


# Results are plotted only if PLOT is TRUE
PLOT <- T
################################################################################################


#### 1. INSTALL REQUIRED PACKAGES ####
# install.packages("scapesClassification")
# install.packages("terra")

# Required for interactive maps & plots
# install.packages("raster")
# install.packages("mapview")
# install.packages("leaflet")
# install.packages("leafpop")
# install.packages("data.table")
```

```
#### 2. LOAD LIBRARIES ####
library(scapesClassification)
library(terra)



#### 3. PLOTS & INTERACTIVE MAPS
library(data.table)
library(mapview)
library(leaflet)
library(leafpop)

mapviewOptions(basemaps = c("Esri.OceanBasemap"), na.color = "grey88",
               homebutton = FALSE, query.position = "bottomright", legend.opacity = 0.9,
               fgb = FALSE)

# palette for background raster
palRYB <- c("#2892c7", "#fafa64", "#e81014")
palYGB <- c("#ffff80" ,"#3dba65", "#0d1178")
palBYR <- c("#4575b5", "#ffffbf", "#d62f27")

#### 3. LOAD DATA ####
grd <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T)
grd <- grd[grepl("\\.grd", grd)]
grd <- grd[!grepl("hillshade", grd)]
rstack <- rast(grd)

if(PLOT){
  plot(rstack, mar = c(3, 1.6, 1.6, 4), cex.main= 1.1, col = hcl.colors(50, palette =  "Zissou 1"),
```

```
    main = c("(a) Bathymetry", "(b) Local BPI", "(c) Regional BPI", "(d) Slope"))

 readline(prompt = "Plot input data. Press [enter] to continue.")
}

# Load ISU and PKS classification
ISU <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T,
            pattern = "ISU_Cells\\.RDS")

PKS <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T,
            pattern = "PKS_Cells\\.RDS")

##############################################################################################
#### 4. FORMAT INPUTS ####
##############################################################################################

# COMPUTE ATTRIBUTE TABLE
atbl <- attTbl(rstack, var_names = c("bathymetry", "local_bpi", "regional_bpi", "slope"))

# Use ISU and PKS classification as attribute table filters
atbl$ISU <- readRDS(ISU)
atbl$ISU[!is.na(atbl$ISU)] <- 1 # Class 1 identifies all ISU cells

atbl$PKS <- readRDS(PKS)

head(atbl)

# COMPUTE NEIGHBORHOOD LIST
```

```
nbs <- ngbList(rstack, rNumb = TRUE, attTbl = atbl) # neighbors are identified by row numbers (see
?ngbList)


head(nbs)


##############################################################################################

#### 5. RELEF UNIT (RU)

##############################################################################################


# 1. ANCHOR.SEED

# Tasks:

# Identify raster cells on prominent features in connection with peak cells

# that are not within island shelf units (ISU).

atbl$RO <-

  anchor.seed(atbl, nbs, rNumb = TRUE, silent = TRUE,


          class = NULL, # A new ID is assigned to every

                  # discrete group of cells...


          cond.seed   = "PKS == 1", # ...in continuity with peak cells...


          cond.growth = "regional_bpi >= 100 | local_bpi >=100", # ...that are on prominent features.


          cond.filter = "is.na(ISU)") # Only consider non-ISU cells

                          # on prominent features.


rRO1 <- cv.2.rast(rstack, atbl$RO) # store raster for plotting

##############################################################################################
```

```
# 2. REL.PI
# Tasks:
# Compute the standardized relative position index (`rPI`) of each raster object.
atbl$rPI <- rel.pi(atbl, RO="RO", el="bathymetry")
##############################################################################################

# 3. PI.SGM
# Tasks:
# (i) Segment large raster objects that have negative rPI or negative local bpi;
# (ii) Remove raster objects with less than 40 cells.
atbl$RO <-
  pi.sgm(atbl, nbs, rNumb = TRUE, RO = "RO", mainPI = "rPI", secPI = "local_bpi",
         cut.mPI = 0, cut.sPI = 0, min.N = 40)

rRO2 <- cv.2.rast(rstack, atbl$RO) # store raster for plotting
##############################################################################################

# 4. PI.ADD
# Tasks:
# Add new raster objects composed of cells with local_bpi>100;
# An object is added only if it is disjoint from other objects,
# if it has more than 40 cells and if it is not within the ISU.
atbl$RO <- pi.add(atbl, nbs, rNumb = TRUE, RO = "RO", mainPI = "local_bpi", add.mPI = 100,
                  min.N = 40, cond.filter = "is.na(ISU)")

rRO3 <- cv.2.rast(rstack, atbl$RO) # store raster for plotting
##############################################################################################
```

```
# 4. COND.4.NOFN (HGROWTH)
# Task:
# Homogeneous growth of raster objects;
# At turns, each raster object adds contiguous cells having
# regional or local BPI values > 100
IDs <- unique(atbl$RO)[!is.na(unique(atbl$RO))]

atbl$RO <- cond.4.nofn(atbl, nbs, rNumb = TRUE, classVector = atbl$RO,
                       nbs_of = IDs, class = NULL,
                       cond = "regional_bpi > 100 | local_bpi > 100",
                       hgrowth = TRUE)



# Sort id based on depth
atbl$RO <- anchor.seed(atbl, nbs, rNumb = TRUE, silent = TRUE, class = NULL,
        cond.filter = "!is.na(RO)", # cell within RU class
        cond.seed   = "!is.na(RO)", # cell within RU class
        cond.growth = "RO[] == RO", # cell having the same class as the focal cell
        sort.col    = "bathymetry", # order seed evaluation by bathymetry
        sort.seed   = "max")       # cells with the highest bathymetric values are evaluated first


rRO4 <- cv.2.rast(rstack, atbl$RO) # store raster for plotting

##############################################################################
# 5.1 RU Static plots
##############################################################################
```

```
if(PLOT){

  # Classification steps
  plot(c(rRO1,rRO2,rRO3,rRO4), type = "classes", mar = c(3, 1.6, 1.6, 4), cex.main= 1.1,
       col = hcl.colors(50, palette =  "Spectral"),
       main = c("(a) rRO1, anchor.seed", "(b) rRO2, pi.sgm", "(c) rRO3, pi.add",
                "(d) rRO4, hgrowth"))

  readline(prompt = "Classification steps. Press [enter] to continue.")

}

if(PLOT){

  # Position index
  atbl$rPI <- rel.pi(atbl, RO="RO", el="bathymetry", type = "n")

  # Shallowest peak for each RO
  library(data.table)
  atbl <- as.data.table(atbl)
  atbl[!is.na(RO), mbat := bathymetry == max(bathymetry), by = "RO"]
  pc <- atbl[which(mbat), Cell]
  atbl$mbat <- NULL

  # Plot
  r_rPI <- cv.2.rast(rstack, atbl$rPI) # relative PI
```

```
 terra::plot(r_rPI, type="interval", breaks=seq(0,1,0.1),
          plg = list(title="rPI", title.adj=0.1),
          col = hcl.colors(50, palette =  "Terrain 2", rev=T),
          main = "Normalized relative PI")

 terra::points(xyFromCell(rstack, pc), pch=20, col="red", cex=1.3)
 legend(ext(r_rPI)[2], ext(r_rPI)[4]-25000, legend=c("Peak"), bty = "n", xpd=TRUE,
       col=c("red"), pch=20, cex=0.9, pt.cex = 1.5)

 readline(prompt = "Normalized relative PI. Press [enter] to continue.")
}

##############################################################################################
# 5.1 RU Interactive map
##############################################################################################

if(PLOT){

 # Note that all raster classifications are converted into polygons (as.polygons()) before plotting
 # for a better visualization. Note also that mapview do not support the terra package yet, therefore
 # all objects of class terra are converted into objects of class raster before plotting
 # (as(xx,"Spatial")).

 # CONVERT ATTRIBUTE TABLE TO ATTRIBUTE 'DATA' TABLE
 atbl <- as.data.table(atbl)

 # COMPUTE STATISTICS
 RO_stat <- atbl[!is.na(RO),
```

```
          .(N = .N,                    # NUMBER OF CELLS
            minD = round(max(bathymetry), 2), # MIN DEPTH
            avgD = round(mean(bathymetry), 2),# AVERAGE DEPTH
            maxD = round(min(bathymetry), 2), # MAX DEPTH
            sdD = round(sd(bathymetry), 2)),  # STANDARD DEVIATION

          by = "RO"]

# RENAME AND RANK RO IDs BY MIN. DEPTH
new_IDs <- RO_stat[, data.table::frankv(minD, order = -1L)]
atbl[!is.na(RO), RO := new_IDs[match(RO, RO_stat$RO)] ]
RO_stat[, RO := new_IDs][order(RO)]

RO_stat

# Class vector 2 raster
rRO <- cv.2.rast(rstack, atbl$RO)
ISU <- cv.2.rast(rstack, atbl$ISU)

# Raster 2 polygon
rRO <- as.polygons(rRO, dissolve = TRUE)
ISU <- as.polygons(ISU, dissolve = TRUE)

ISU$dummy <- 1 #trick map label

# PEAKS
pcV <- vect(xyFromCell(rstack, pc), crs = crs(rRO))
```

```
# BUILD POP UP TABLE
names(rRO) <- "ID"

# Index vector IDs into RO_stat
ind <- match(rRO$ID, RO_stat$RO)

# Add columns
rRO$N    <- RO_stat$N[ind]
rRO$minD <- RO_stat$minD[ind]
rRO$avgD <- RO_stat$avgD[ind]
rRO$maxD <- RO_stat$maxD[ind]
rRO$sdD  <- RO_stat$sdD[ind]

m <-
  mapview(as(rRO,"Spatial"), layer.name = "RO", zcol = "ID", alpha.regions = 0.8, legend = TRUE,
          popup = popupTable(as(rRO,"Spatial"), feature.id = FALSE, row.numbers = FALSE),
          col.regions = hcl.colors(length(rRO$ID), palette =  "Spectral")) +
  mapview(as(pcV,"Spatial"), layer.name = "Peak", alpha.regions = 0.9, col.regions = "grey", legend=
TRUE,
          popup = FALSE) +
  mapview(raster::raster(rstack[["bathymetry_exm"]]), layer.name = "Depth (m)", col.regions = palRYB)

m@map %>%
  hideGroup(c("Depth (m)")) %>%
  addLayersControl(overlayGroups = c("RO", "Peak", "Depth (m)"),
                   position = "topleft",
                   options = layersControlOptions(collapsed = TRUE))%>%
  addMiniMap(tiles = "Esri.OceanBasemap",
```

```
        position = "bottomleft",
        toggleDisplay = TRUE,
        zoomLevelOffset = -5)
}
```

C.4 R code to compute the Self-Organizing Map (SOM)

```
# User-defined classification of raster surfaces: The R package scapesClassification

# SOM script
# Author: Gerald Hechter Taranto
# Journal:

# A detailed explanation of all classification steps is available on the documentation page
# https://ghtaranto.github.io/scapesClassification/index.html


# Results are plotted only if PLOT is TRUE
PLOT <- T
################################################################################################


#### 1. INSTALL REQUIRED PACKAGES ####
# install.packages("scapesClassification")
# install.packages("terra")

# Required for interactive maps & plots
# install.packages("raster")
# install.packages("mapview")
# install.packages("leaflet")
# install.packages("leafpop")
# install.packages("data.table")
```

```
#### 2. LOAD LIBRARIES ####
library(scapesClassification)
library(terra)



#### 3. PLOTS & INTERACTIVE MAPS
library(data.table)
library(mapview)
library(leaflet)
library(leafpop)

mapviewOptions(basemaps = c("Esri.OceanBasemap"), na.color = "grey88",
               homebutton = FALSE, query.position = "bottomright", legend.opacity = 0.9,
               fgb = FALSE)

# palette for background raster
palRYB <- c("#2892c7", "#fafa64", "#e81014")
palYGB <- c("#ffff80" ,"#3dba65", "#0d1178")
palBYR <- c("#4575b5", "#ffffbf", "#d62f27")

#### 3. LOAD DATA ####
grd <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T)
grd <- grd[grepl("\\.grd", grd)]
grd <- grd[!grepl("hillshade", grd)]
rstack <- rast(grd)

if(PLOT){
```

```
  plot(rstack, mar = c(3, 1.6, 1.6, 4), cex.main= 1.1, col = hcl.colors(50, palette =  "Zissou 1"),
       main = c("(a) Bathymetry", "(b) Local BPI", "(c) Regional BPI", "(d) Slope"))

  readline(prompt = "Plot input data. Press [enter] to continue.")
}

# Load ISU, PKS and RU classification
ISU <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T,
                  pattern = "ISU_Cells\\.RDS")

PKS <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T,
                  pattern = "PKS_Cells\\.RDS")

RO <- list.files(system.file("extdata", package = "scapesClassification"), full.names = T,
                 pattern = "RU_obj\\.RDS")


################################################################################
#### 4. FORMAT INPUTS ####
################################################################################

# COMPUTE ATTRIBUTE TABLE
atbl <- attTbl(rstack, var_names = c("bathymetry", "local_bpi", "regional_bpi", "slope"))

# Use ISU and PKS classification as attribute table filters
atbl$ISU <- readRDS(ISU)
atbl$ISU[!is.na(atbl$ISU)] <- 1 # Class 1 identifies all ISU cells
```

```
atbl$PKS <- readRDS(PKS)

atbl$RO  <- readRDS(RO)


head(atbl)


# COMPUTE NEIGHBORHOOD LIST

nbs <- ngbList(rstack, rNumb = TRUE, attTbl = atbl) # neighbors are identified by row numbers (see
?ngbList)


##############################################################################################

#### 5. SELF ORGANIZING MAP ####

##############################################################################################


# install.packages("rassta")

library(rassta)


########################

#### 5.1 RO vs. SOM####

########################


# TIME CONSUMING!


# SET SEED FOR REPRODUCIBILITY

# set.seed(380)

#

# # SCALE VALUES

# r <- terra::scale(rstack)

#

# # VARIABILITY ANALYSIS
```

```
# dt <- data.table(bat = values(rstack[[1]], mat=F))
# N <- rep(0, length = ncell(r))
#
# itr <- 1
# maxITR <- 100
#
# while(itr <= maxITR){
#
#   print(itr)
#
#   # SOM WITH A MAXIMUM OF 2 CLUSTERS
#   r.som <- som_gap(r, K.max = 2)
#
#   # RASTERIZATION of SOM
#   r.sompam0 <- som_pam(ref.rast = r[[1]], kohsom = r.som$SOM, k = r.som$Kopt)
#
#   # plot(r.sompam0$sompam.rast[[2]], main = "Elevaton Units",
#   #      type = "classes", col = hcl.colors(r.som$Kopt, "Mint", rev = TRUE),
#   #      mar = c(1.5, 2, 1.5, 2.5))
#
#   dt[, clu := values(r.sompam0$sompam.rast[[2]], mat=F)]
#   cl <- dt[, mean(bat), by=clu]
#   cl <- cl$clu[which.max(cl$V1)]
#
#   N[which(dt$clu == cl)] <- N[which(dt$clu == cl)] + 1
#
#   itr <- itr + 1
#
```

```
# }
#
# rt <- r[[1]]; rt[] <- NA; rt[] <- N
# plot(rt)

###############################################
#### 5.2 DATA HETEREOGENEITY WITHIN RO ####
###############################################

# CROP RASTER DATA TO THE RO REGION
rRO <- cv.2.rast(rstack, (atbl$RO))
r   <- terra::mask(rstack, rRO)

# SCALE VALUES
r <- terra::scale(r)

# SET SEED FOR REPRODUCIBILITY
set.seed(381)

# SOM WITH A MAXIMUM OF 9 CLUSTERS
r.som <- som_gap(r, K.max = 9)

# RASTERIZATION of SOM
r.sompam1 <- som_pam(ref.rast = r[[1]], kohsom = r.som$SOM, k = r.som$Kopt)

plot(r.sompam1$sompam.rast[[2]], main = "Nested Features",
    type = "classes", col = hcl.colors(r.som$Kopt, "Terrain 2"),
    mar = c(1.5, 2, 1.5, 2.5))
```

Appendix D
Supplementary data Chapter 5

D.1 R code to identify the geomorphic units of the Azores

```
# LOAD PACKAGES
library(scapesClassification)
library(terra)
library(data.table)
library(scapesClassification)

# 1. LOAD INPUT DATA AND CREATE FOLDERS

# 1.1 RASTERS
rasters <- character(6L)

rasters[1] <- "./InputLayers/bathymetry_Luis.tif"      # bathymetry
rasters[2] <- "./InputLayers/Slope_Luis.tif"          # slope
rasters[3] <- "./InputLayers/geomorph_az_3_50_fl_2.tif"# gmorph
rasters[4] <- "./InputLayers/bpi_2_100_Luis.tif"      # global bpi
rasters[5] <- "./InputLayers/bpi_2_40_Luis.tif"       # regional bpi
rasters[6] <- "./InputLayers/bpi_1_10_Luis.tif"       # local bpi

rstack <- rast(rasters)
names(rstack) <- c("bathymetry", "slope", "gmorph", "global_bpi", "regional_bpi", "local_bpi")
```

```
# 1.2 ATTRIBUTE TABLE
dt <- attTbl(rstack, var_names = NULL)
dt <- as.data.table(dt)

# 1.3 8-NIGHBOR LIST
# nbs <- ngbList(rstack, rNumb = TRUE, attTbl = dt)
# saveRDS(nbs, file.path(out_dir, "nbs.RDS"))
nbs <- readRDS(file.path(out_dir, "nbs.RDS"))

# 1.4 CREATE FOLDERS TO SAVE OUTPUTS
if(!dir.exists("./GMU_Class/01_GMU_IS")){dir.create("./GMU_Class/01_GMU_IS")}
if(!dir.exists("./GMU_Class/01_GMU_IS/STEPS")){dir.create("./GMU_Class/01_GMU_IS/STEPS")}



################################################################################
# 2. ISLAND SHELF ####
################################################################################

# Rule-set 1
# A) Extract anchor points from the Azores Islands shapefile
anchor <- anchor.svo (r = rstack[[1]],
                      dsn = "E:/Documents/GIS_LAYERS/Gerald_Taranto/General/Layers/Azores.shp",
                      only_NAs = TRUE,
                      fill_NAs = TRUE,
                      plot = TRUE,
                      saveRDS = "./GMU_Class/land_pixel_number.RDS",
                      writeRaster = "./GMU_Class/01_GMU_IS/STEPS/IS_01_anchor.tif",
```

```
                overWrite = TRUE)


# B) Cells adjacent to landmasses are island shelf cells
classVector1 <- anchor.cell(dt, rstack, anchor,
                        class = 1,
                        classVector = NULL,
                        class2cell = F,
                        class2nbs = T,
                        overwrite_class = F,
                        plot = TRUE,
                        writeRaster = NULL,
                        overWrite = FALSE)


cv.2.rast(rstack[[1]], classVector1, index = dt$Cell, writeRaster =
"./GMU_Class/01_GMU_IS/STEPS/IS_02_adj.tif", overWrite = T)


# Rule-set 2
# A)
# Bathymetric data around Terceira Island are of poor quality and the shelf
# break does not show sharp discontinuities in slope values. Set maximum
# number of iteration and minimum border in the functions below to overcome this
# problem.
classVector2 <- cond.4.nofn(dt, nbs, rNumb  = T,
                        classVector = classVector1,
                        nbs_of = 1,
                        cond = "slope <= 5",
                        class = 1,
                        max.iter = 5)
```

```
# B) Minimum border and max 30 iteration - avoid seafloor around Terceira to be
# cataloged as shelf
classVector3 <- cond.4.nofn(dt, nbs, rNumb = T,
                    classVector = classVector2,
                    nbs_of = 1,
                    cond = "slope <= 5",
                    class = 1,
                    min.bord = 0.2,
                    max.iter = 30)

cv.2.rast(rstack[[1]], classVector3, dt$Cell, writeRaster =
"./GMU_Class/01_GMU_IS/STEPS/IS_03_slopes.tif", overWrite = T)

# Rule-set 3
# Local_bpi increase as we approximate to the shelf break; use this property to
# set shelf area boundaries
classVector4 <- cond.4.nofn(dt, nbs, rNumb = T,
                    classVector = classVector3,
                    nbs_of = 1,
                    cond = "slope <= 5 & local_bpi > local_bpi[]",
                    class = 1)

cv.2.rast(rstack[[1]], classVector4, dt$Cell, writeRaster =
"./GMU_Class/01_GMU_IS/STEPS/IS_04_bpi.tif", overWrite = T)

# Rule-set 4
# Fill up classification holes (min.border = 60%)
classVector5 <- cond.4.nofn(dt, nbs, rNumb = T,
                    classVector = classVector4,
```

```
                nbs_of = 1,
                cond = "TRUE",
                class = 1,
                min.bord = 0.6)

cv.2.rast(rstack[[1]], classVector5, dt$Cell, writeRaster = "./GMU_Class/01_GMU_IS/IS_01_shelf.tif",
overWrite = T)
saveRDS(classVector5, "./GMU_Class/01_GMU_IS.RDS")

#############################################################################
# 3. PEAKS (PKS) ####
#############################################################################

# Rule-set 1
# Find all local maxima
dt$PKS <- peak.cell(dt, nbs, rNumb = TRUE,
            p_col = "bathymetry", # column on which local maxima are searched
            p_edge = FALSE)      # local maxima or minima are not searched on edge cells


# Rule-set 2
# Local maxima only on relieves and not within island shelves
dt$PKS[dt$global_bpi < 100 | dt$local_bpi < 100 | !is.na(dt$SHELF)] <- NA

# Rule-set 3
# Remove local maxima adjacent to island shelves
dt$PKS <- cond.reclass(dt, nbs, rNumb = TRUE,
                classVector = dt$PKS,
                class = 1,
```

```
                reclass = NA,
                cond = "!is.na(SHELF{})", # RECLASS IF ANY NEIGHBOR IS A SHELF CELL
                peval = 0.01)

cv.2.rast(rstack[[1]], dt$PKS, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_PKS/01_PKS_ALL.tif", overWrite = T)
saveRDS(dt$PKS, "./GMU_Class/02_GMU_PKS.RDS")

#############################################################################
# 4. RAISED FEATURES (RO)  ####
#############################################################################

# 1.1 Raster objects (Regional BPI > 100)
dt$RO <- anchor.seed(dt, nbs, rNumb = TRUE, silent = FALSE,

                class = NULL, # A new ID is assigned to every
                # discrete group of cells...

                cond.seed   = "PKS == 1", # ...in continuity with peak cells...

                cond.growth = "regional_bpi >= 100", # ...that are on prominent features.

                sort.col = "bathymetry", sort.seed = "max", # shallow relieves are identified first

                cond.filter = "regional_bpi >= 100 & is.na(SHELF)") # Only consider non-ISU cells


cv.2.rast(rstack[[1]], dt$RO, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_RO/STEPS/01_RO_ALL.tif", overWrite = T)
```

```
# 1.2 Segment raster objects (Relative BPI < 0 and regional BPI < 200)
dt$rPI <- rel.pi(dt, RO="RO", el="bathymetry", type = "s")
dt$RO_1 <- pi.sgm(dt, nbs, rNumb = TRUE,
                  RO = "RO",
                  mainPI = "rPI",
                  secPI = "regional_bpi",
                  cut.mPI = 0,
                  cut.sPI = 200,
                  min.N = 40)

cv.2.rast(rstack[[1]], dt$RO_1, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_RO/STEPS/02_RO_SGM.tif", overWrite = T)

# 1.3 Segment large raster objects (N > 3000) (Relative BPI < 0 and regional BPI < 200)
dt$rPI <- NULL
dt[!is.na(RO_1), N := .N , by = "RO_1"]
dt[N > 3000, RO_2 := RO_1]

dt$rPI <- rel.pi(dt, RO="RO_2", el="bathymetry", type = "s")

dt$RO_2 <- pi.sgm(dt, nbs, rNumb = TRUE,
                  RO = "RO_2",
                  mainPI = "rPI",
                  cut.mPI = 0,
                  min.N = 100)

dt$RO_2b <- dt$RO_1
dt[N > 2000, RO_2b := RO_2]
```

```
# Integrate new RO into the main RO class vector
dt$RO_2c <- anchor.seed(dt, nbs, rNumb = TRUE, silent = FALSE,

                class = NULL, # A new ID is assigned to every
                # discrete group of cells...

                cond.seed   = "!is.na(RO_2b)", # ...in continuity with peak cells...

                cond.growth = "RO_2b == RO_2b", # ...that are on prominent features.

                sort.col = "bathymetry", sort.seed = "max") # shallow relieves are identified first


cv.2.rast(rstack[[1]], dt$RO_2c, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_RO/STEPS/03_RO_SGM.tif", overWrite = T)

# 1.4 Add new RO with cells having regional BPI > 300 (non connected with existing raster objects)
dt$RO_3 <- pi.add(dt, nbs, rNumb = TRUE,
                RO = "RO_2c",
                mainPI = "regional_bpi",
                add.mPI = 300,
                min.N = 50,
                cond.filter = "is.na(SHELF)")

cv.2.rast(rstack[[1]], dt$RO_3, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_RO/STEPS/04_RO_ADD.tif", overWrite = T)

# 1.5 Add new RO with cells having regional BPI > 400 (non connected with existing raster objects)
```

```
dt$RO_4 <- pi.add(dt, nbs, rNumb = TRUE,
                  RO = "RO_3",
                  mainPI = "regional_bpi",
                  add.mPI = 400,
                  min.N = 50,
                  cond.filter = "is.na(SHELF)")


cv.2.rast(rstack[[1]], dt$RO_4, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_RO/STEPS/05_RO_ADD.tif", overWrite = T)


# 1.6 Homogeneous growth RO adding cells with decreasing bathymetry and regional_bpi > 100
IDs <- unique(dt$RO_4)[!is.na(unique(dt$RO_4))]
dt$RO_5 <- cond.4.nofn(dt, nbs, rNumb = TRUE, classVector = dt$RO_4,
                       nbs_of = IDs, class = NULL,
                       cond = "bathymetry < bathymetry[] & regional_bpi > 100 & is.na(SHELF)",
                       hgrowth = TRUE)


cv.2.rast(rstack[[1]], dt$RO_5, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_RO/STEPS/06_RO_GROW.tif", overWrite = T)


# 1.7 Homogeneous growth RO adding cells with regional_bpi > 100
IDs <- unique(dt$RO_5)[!is.na(unique(dt$RO_5))]
dt$RO_6 <- cond.4.nofn(dt, nbs, rNumb = TRUE, classVector = dt$RO_5,
                       nbs_of = IDs, class = NULL,
                       cond = "regional_bpi > 100 & is.na(SHELF)",
                       hgrowth = TRUE)


cv.2.rast(rstack[[1]], dt$RO_6, index = dt$Cell, writeRaster =
"./GMU_Class/02_GMU_RO/STEPS/07_RO_GROWb.tif", overWrite = T)
```

```
cv.2.rast(rstack[[1]], dt$RO_6, index = dt$Cell, writeRaster = "./GMU_Class/02_GMU_RO/RO.tif",
overWrite = T)

saveRDS(dt$RO_6, "./GMU_Class/03_GMU_RO.RDS")


##############################################################################

# 5. ROs (deep slopes & crests) ####

##############################################################################


# Deep slopes identified as areas of high bathymetry variation


# Lower slopes adjacent to RO with a lower bathymetry

IDs <- unique(dt$RO)[!is.na(unique(dt$RO))]

dt$ROds <- cond.4.nofn(dt, nbs, rNumb = TRUE, classVector = dt$RO,

                nbs_of = IDs, class = NULL,

                min.bord = 0.2,

                directional = TRUE,

                cond = "bathymetry <= bathymetry{} & bat_sd  > 150 & is.na(SHELF)",

                hgrowth = TRUE, peval = 0.2)


# Lower slopes adjacent to RO with a lower bathymetry

IDs <- unique(dt$ROds)[!is.na(unique(dt$ROds))]

dt$ROds <- cond.4.nofn(dt, nbs, rNumb = TRUE, classVector = dt$ROds,

                nbs_of = IDs, class = NULL,

                min.bord = 0.2,

                directional = TRUE,

                cond = "local_bpi > 100 & is.na(SHELF)",

                hgrowth = TRUE)
```

```
# Fill holes (unclassified cells surrounded by classified cells)
IDs <- unique(dt$ROds)[!is.na(unique(dt$ROds))]
dt$ROds <- cond.4.nofn(dt, nbs, rNumb = TRUE, classVector = dt$ROds,
                       nbs_of = IDs, class = NULL,
                       min.bord = 0.6,
                       cond = "TRUE",
                       hgrowth = TRUE)

cv.2.rast(rstack[[1]], dt$ROds, dt$Cell, writeRaster = "./GMU_Class/04_GMU_ROds/ROds.tif", overWrite
= T)
saveRDS(dt$ROds, "./GMU_Class/04_GMU_ROds.RDS")
```

UNIVERSIDADE DOS AÇORES

Faculdade de Ciências e Tecnologia

Rua Professor Doutor Frederico Machado, N.º 4

9901-862 Horta

Açores, Portugal

2022

TD

Mapping deep-sea biodiversity and good environmental status in the Azores: assisting with the implementation of EU Marine Strategy Framework Directive

Gerald Hechter Taranto