

UNIVERSIDADE DO ALGARVE
Faculdade de Ciências e Tecnologia

Modeling the suitable habitat of a structural red gorgonian species in the Mediterranean/east Atlantic regions predicts a wider distribution than previously known

Inga Isabel Esteves Barata Ferreira da Silva

Dissertation

Marine Biology Master

Supervisors

Prof. Dra. Ester Serrão

Universidade do Algarve, Faro, Portugal

Dr. Jorge Assis

CCMAR, Faro, Portugal

UNIVERSIDADE DO ALGARVE
Faculdade de Ciências e Tecnologia

Modeling the suitable habitat of a structural red gorgonian species in the Mediterranean/east Atlantic regions predicts a wider distribution than previously known

Inga Isabel Esteves Barata Ferreira da Silva

Dissertation

Marine Biology Master

Supervisors

Prof. Dra. Ester Serrão

Universidade do Algarve, Faro, Portugal

Dr. Jorge Assis

CCMAR, Faro, Portugal

Mestrado: Biologia Marinha

Título: Modeling the suitable habitat of a structural red gorgonian species in the Mediterranean/east Atlantic regions predicts a wider distribution than previously known

Autoria: Inga Isabel Esteves Barata Ferreira da Silva

Ano: 2015

Orientação científica do trabalho: Prof^ª. Dra. Ester Serrão & Dr. Jorge Assis

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

«Copyright»

A Universidade do Algarve tem o direito, perpétuo e sem limites geográficos, de arquivar e publicitar este trabalho através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, de o divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

ACKNOWLEDGEMENTS

Em primeiro lugar, quero agradecer aos meus orientadores, Prof^a Dra. Ester Serrão e Dr. Jorge Assis, a oportunidade de realizar este estudo, foi completamente de encontro ao que pretendia fazer quando decidi tirar este mestrado. Obrigada por todo o conhecimento transmitido ao longo deste tempo.

Um muito obrigado à Joana Boavida, por todo o trabalho que desenvolveu comigo, pela disponibilidade e prontidão em ajudar, e por todo o conhecimento partilhado. Aprendi muito!

Quero agradecer à minha família, Mãe, Pai, Lari, Catarina, Madalena e Margarida, pelo apoio incondicional, pelas visitas, pela força e paciência no decorrer deste trabalho.

Aos meus queridos amigos, porque a vida sem eles não era a mesma coisa, obrigada pela força, apoio e boa energia!

Muito obrigada à família do Sul, por terem estado sempre presentes, nos momentos bons e menos bons, por toda a ajuda, amizade e carinho que me deram, e por acreditarem em mim! Foram um pilar muito importante, merci merci!

Um obrigada muito especial à Tamára e ao Filipe, por me acompanharem no decorrer deste trabalho, pela amizade, conselhos, ajuda, por estarem sempre por perto...obrigada por tudo!

A todos....obrigada, muito obrigada!

RESUMO:

O clima tem um papel fundamental na forma como molda a distribuição global dos organismos, sendo por isso importante conhecer e compreender os limites de tolerância dos mesmos a variações ambientais. Devido à dificuldade de amostragem em ecossistemas profundos, existe ainda uma grande lacuna no que diz respeito aos factores que controlam a distribuição das espécies que neles habitam, nomeadamente as comunidades bentónicas. O Mediterrâneo é um ecossistema que apresenta características muito particulares, abrigando uma vasta diversidade de organismos essencialmente associados às comunidades coralígenas. A *Paramuricea clavata* é uma das espécies estruturantes mais importantes destas comunidades, tendo um papel muito importante no funcionamento e manutenção do ecossistema. Nos últimos anos, o aumento da temperatura observado no Mediterrâneo tem provocado eventos de mortalidade massiva junto das comunidades bentónicas, levando à perda de populações desta gorgónia. A distribuição conhecida desta espécie era até hoje restrita ao Mediterrâneo, principalmente na zona Norte. Contudo, novas populações foram encontradas no Atlântico adjacente, junto a costa sul e oeste de Portugal e no mar Alboran, indicando uma distribuição mais ampla. Boosted Regression Tress (BRT) é um método de modelação recentemente aplicada na área da ecologia, que permite relacionar dados ambientais com dados de presença, possibilitando a descrição de padrões de distribuição para as espécies. Este método usa algoritmos que aprendem a relação existente entre os dados tipos de dados, permitindo a descrição do seu padrão. BRT resulta da junção de dois algoritmos, árvores de regressão e agregação, o que permite que o modelo final seja construído tendo em conta as interações existentes entre os dados. A flexibilidade do BRT permite contornar os principais problemas associados à modelação de dados ecológicos tais como: 1) capacidade de lidar com diferentes tipos de dados; 2) minimização da perda de função que ocorre devido à tendência de amostragem (sampling bias), existência de locais na área de estudo com maior numero de amostras que outros; 3) é dos métodos menos afetados pela resolução dos dados; e 4) reduz o sobre-ajuste do modelo aos dados da amostra (overfitting), por considerar todas as possíveis combinações destes. Por ser uma técnica que tem revelado boa capacidade de previsão da distribuição, nós utilizamos o BRT para descrever o habitat adequado para a *P. clavata*. Deste modo, determinámos quais as variáveis mais relevantes para

descrição da distribuição desta espécie e quais os seus limites de tolerância. O método mostrou-se eficaz, apresentado um excelente poder preditivo (AUC: 0.98; TSS: 0.96; sensibilidade: 0.98, especificidade: 0.98). O habitat adequado para esta gorgónia é caracterizado por uma topografia de fundo complexa, elevada produtividade e hidrodinamismo de moderado a forte. Os nossos resultados revelaram a topografia do fundo, representada pelo declive, como sendo a variável mais importante para a caracterização do habitat desta espécie, apresentando uma contribuição de 21.5% e um intervalo de tolerância sem limite máximo, $> 0.3^\circ$. Esta variável é um bom indicador da dinâmica que ocorre no fundo, sendo que elevações topográficas estão normalmente associadas a alterações no padrão das correntes, e a zonas de maior produtividade. A Temperatura tem também um papel muito importante (contribuição de 19%), contudo, o intervalo de tolerância é bastante restrito (12.3 a 26.5°C), limitando assim a sua área de distribuição. Através deste modelo, verificámos também que, apesar dos nossos dados de presença ocorrerem entre os 15 e 100 m de profundidade, esta consegue atingir os 200 m. Este facto é interessante, visto que a partir de uma determinada profundidade as perturbações ocorridas à superfície não se fazem sentir com tanta intensidade. Neste sentido, a existência de populações profundas pode ser um indicador de zonas de refúgio. A distribuição da espécie ocorre numa estreita linha da costa, ao longo do Mediterrâneo e Atlântico adjacente, e também em ilhas e outras em elevações topográfica do fundo, como montes e bancos submarinos. Para além dos novos locais de presença ocorridos no Atlântico e mar Alboran, o nosso modelo prevê uma boa probabilidade de ocorrência para a região sul do Mediterrâneo, nomeadamente ao longo da costa de Marrocos, Argélia e costa noroeste da Tunísia. Para o Atlântico, a área entre Agadir, Marrocos, e a costa sudoeste da Península Ibérica revela também a possibilidade de ocorrência desta gorgónia. Tendo em conta que se tratam de regiões pouco estudadas, os nossos dados sugerem estas áreas como potenciais alvos para estudos futuros de monitorização e conservação destes habitats. Com as alterações climáticas previstas para o Mediterrâneo, há uma forte possibilidade de ocorrer perda de grande parte das populações da espécie, especialmente das que habitam zonas menos profundas.

Palavras-chave: Limites de tolerância das espécies, espécies estruturantes de ecossistema, Modelação de Nicho Ecológico (ENM), habitat adequado, *Paramuricea clavata*

ABSTRACT:

Understand how the distribution of deep-sea species is shaped by environment is a challenge for ecologists, given the difficulty of sampling at these locations. *Paramuricea clavata* is a long-lived species that inhabits the Mediterranean coralligenous assemblages. Its known distribution was so far described as being restricted to Mediterranean Sea, however, new populations were found in the neighboring Atlantic Ocean, suggesting a wider distribution. BRT is a modeling technique that uses algorithms to describe the relation between environmental and presence data, in order to assess species distribution. This method has been applied to ecological studies, showing good accuracy in predicting suitable habitats, since it copes well with the major problems associated with niche modeling. We use BRT to describe the suitable habitat of *Paramuricea clavata* by determining the variables that limit its distribution, and the species' tolerance range for these variables. Our results accurately revealed that the suitable habitat is mainly characterized by complex bottom topography (represented by slope) and temperature, presenting a wide range for the first variable ($> 0.3^\circ$) and a well-defined range for the second (12.3 to 26.5°C). In addition to the previous known distribution, new potential sites of occurrence were revealed in the Mediterranean Sea, along the coast of Morocco, Algeria, Tunisia and Strait of Sicily, and on the Atlantic coasts between Morocco and southwest Iberia. With this study we were able to describe, for the first time, the ecological niche and potential global distribution of this important structuring species. Moreover, our results contribute to the achievement of future studies of monitoring and conservation of benthic communities in poorly-sampled areas. With the predicted climate change for Mediterranean, the occurrence of mortality events in benthic communities will continue, especially to those found in shallower areas, putting *P. clavata* at risk.

Keywords: Species range limits, ecosystem structuring species, Ecological Niche modeling (ENM), suitable habitat, *Paramuricea clavata*

INDEX:

ACKNOWLEDGEMENTS	I
RESUMO.....	II
ABSTRACT	V
INDEX	VI
FIGURES LIST	VII
TABLES LIST	VIII
ABREVIATONS LIST.....	IX
1 INTRODUCTION.....	1
1.1 SPECIES DISTRIBUTION	1
1.2 ECOLOGICAL NICHE MODELING.....	3
1.3 FOCAL SPECIES: <i>PARAMURICEA CLAVATA</i>	7
2 MATERIAL AND METHODS.....	11
2.1 STUDY LOCATION.....	11
2.2 DISTRIBUTION AND ENVIRONMENTAL DATA	12
2.3 PREPARING PSEUDO-ABSENCES	13
2.4 MODELING PROCEDURE	14
3 RESULTS.....	17
3.1 MODEL PERFORMANCE	17
3.2 THE NICHE OF <i>PARAMURICEA CLAVATA</i>	19
3.3 PREDICTIVE SUITABLE HABITAT	21
4 DISCUSSION.....	23
4.1 MODEL PERFORMANCE:	23
4.2 CHARACTERIZATION OF <i>PARAMURICEA CLAVATA</i> ECOLOGICAL NICHE.....	25
4.3 PREDICTIVE SUITABLE HABITAT	29
4.4 PREDICTION OF RANGE-SHIFTS	31
5 CONCLUSION.....	35
6 BIBLIOGRAPHY	37
7 APPENDIX	46

FIGURES LIST

Figure 1 Focal species, <i>Paramuricea clavata</i> .	8
Figure 2 Study area and location of the records points of <i>Paramuricea clavata</i> used in the model. Coordinates of the occurrence points and respective references are available in Appendix 1.	11
Figure 3 Partial dependence plots of the fitted values of the final predictor variables used for modeling. X-axis represents the range of each predictor and y-axis describes the influence of the predictor on the occurrence of <i>Paramuricea clavata</i> . The vertical dotted line corresponds to the tipping point. Temperatures were modeled as monotonic variables.	16
Figure 4 Change in predictive deviance obtained in model reduction and confidence interval. Red dashed line represents the actual number of variables removed. The initial model contained eleven predictors, of which four were removed: maximum nitrate, minimum nitrate, maximum phosphate and minimum silicate.....	17
Figure 5 Relative contribution (%) of each predictor for the production of the suitable habitat map of <i>Paramuricea clavata</i> , after model reduction.	18
Figure 6 Response curves of fitted values to each predictor used for modeling. The range of each predictor is represented in the x-axis. The occurrence of <i>Paramuricea clavata</i> is represented in the y-axis, were 1 = presence and 0 = absence.	19
Figure 7 Predicted habitat suitability for <i>Paramuricea clavata</i> . Numbers refer to some locations mentioned in the text: 1 - Galicia (Spain), 2 - West Portugal, including Berlengas islands, Cape Espichel and Sines, 3 - Gorringe Bank (seamounts), 4 - Algarve Basin, including submerged caves and Tavira (south Portugal), 5 - Agadir (west Morocco), 6 - Alboran Sea, including islands and seamounts, 7 - Ligurian Sea and Pelagos Sanctuary, 8 - Tyrrhenian Sea and Vercelli seamount, 9 - Adriatic Sea, 10 - Ionian Sea, 11 - Aegean Sea, 12 - Levantine Sea with Cyprus to the left and the Syrian-Lebanon border to the right. Some numbers were placed near the site they are referring to, not to overlap suitable habitat areas.....	22
Figure 8 Intersection of suitable habitat of <i>Paramuricea clavata</i> with mortality events in the Mediterranean Sea ().	33

TABLES LIST

- Table 1** Data type and sources of the environmental variables used for modeling the distribution of *Paramuricea clavata*..... 13
- Table 2** Environmental predictors chosen to characterized the ecological limits of *Paramuricea clavata* and corresponding tipping points (TP) obtained from the model (TP.Mod.) and from a literature review (TP.Rev.). TP are levels beyond which the species is unlikely to occur. N.a. indicates data is not available. Occurrence range describes the predictor variation found within our presence data. Raster range describes the predictor variation along the study area..... 20

ABREVIATONS LIST

AUC	Area under the receiver operating characteristic curve
BODC	British Oceanographic Data Centre
BRT	Boosted Regression Trees
CV	Cross-validation
ENFA	Ecological niche factor analysis
ENM	Ecological Niche Modeling
ESM	Environmental Suitable Map
GAM	Generalized Additive Model
GBIF	Global Biodiversity Information Facility
GIS	Geographic Information System
GLM	Generalized Linear Model
HSM	Habitat Suitability Map
<i>lr</i>	Learning rate
NOAA	National Oceanic and Atmospheric Administration
NODC	National Oceanographic Data Center
<i>nt</i>	Number of trees
OBIS	Ocean Biogeographic Information System
POC	Particulate Organic Carbon
PSS	Practical Salinity Scale
<i>tc</i>	Tree complexity
TSS	True Skill Statistics

1 INTRODUCTION

1.1 Species distribution

Understanding how organisms and their environments are related and interact with each other is one of the central issues of ecology science. The environment plays an important role in the spatial distribution of organisms, therefore, some organisms are found in several regions of the world while others are confined to certain areas (Fenchel & Finlay, 2004). For a habitat to be suitable, it must have to favor biological processes like reproduction, growth and survival, and so each species has specific ecological niche requirements and distribution pattern (Sexton, *et al.*, 2009). Although most of species can be found in more than one region, their distribution is determined by physiological tolerance to environmental gradients and biotic interactions. These biotic and abiotic factors determined the range limits of species (Sexton, *et al.*, 2009).

Species range limits are defined by environmental tipping points, boundaries that characterize the tolerance range of a species to a particular environmental variable (Botero *et al.*, 2015; Sexton *et al.*, 2009). When these boundaries are exceeded, the habitat became unsuitable, occurring shifts and contractions on species ranges (Botero *et al.*, 2015). In nature, all environmental variables have a range of variation to which the species survive, and they characterized its ecological niche. Although independent, these variables are correlated, since closer habitats are more similar than distant ones, not meaning that all the adjacent areas are suitable for a certain species. Identification of species' tipping points is important for the understanding of population dynamics, allowing a better evaluation of the actual state of organisms, especially when facing climate change (Botero *et al.*, 2015; Sexton *et al.*, 2009). Within a species, the geographic distribution of individuals may be different, with some populations living in the center of their range, while others live at the margins. Populations occurring in the marginal range of their distribution are known as edge populations (Sexton *et al.*, 2009). According to the latitude limit they occupy, these populations can be define as leading edge, if they occur in the upper latitudinal distribution limit, or rear edge, if they occur in the lower latitudinal limit. While leading edge populations represents the front of

colonization, rear edge represents the older populations that have persisted in the same region despite environmental oscillation, and so they present higher genetic diversity (Hampe & Petit, 2005). Species living at the rear edge are normally confined to smaller habitats surrounded by unsuitable environmental conditions, so they present a distinctive genetic pool diversity that reflects their evolutionary history (Hampe & Petit, 2005). A slightly change in environmental conditions at the range edge can lead to transposition of species' tolerance limits, putting these populations at risk of extinction (Botero *et al.*, 2015). Because they already live closer to their limits, these organisms are good biological indicators of environmental change (Sexton *et al.*, 2009; Woolbright *et al.*, 2014), so studying them may give us a highlight of how organisms can behave when exposed to extreme environmental conditions (Sexton *et al.*, 2009).

In the last years, we have been experience climate oscillation. While some organisms are able to adapt this variation, others are shifting their distribution (Parmesan, 2006). One of the most affected regions with increasing temperature has been the Mediterranean Sea (Rivetti *et al.*, 2014). It is characterized by unique climatic features, hosting a wide biodiversity despite its small area (Coll *et al.*, 2010). This region has been affected by climate variations, being described several mass mortality events in benthic communities, especially in coralligenous assemblages (Bally & Garrabou, 2007; Cerrano *et al.*, 2000; Cigliano & Gambi, 2007; Coma *et al.*, 2009; Crisci *et al.*, 2011; Cupido *et al.*, 2008; Gambi *et al.*, 2010; Garrabou *et al.*, 2009; Huete-Stauffer *et al.*, 2011; Linares *et al.*, 2005; Martin *et al.*, 2002; Rivetti *et al.*, 2014). These assemblages are composed of many structuring species, organisms that directly influence the equilibrium of their ecosystem, by making changes in the physical environment that allow the maintenance of other organisms (Crain & Bertness, 2006; Jones *et al.*, 1994). The most important engineers are those who limit the habitat requirements to other species, being key species in the community structure (Crain & Bertness, 2006). Variations in environment conditions have impact not only on structuring species but in all the associated community, since they have the ability to control and maintain some ecosystem properties (Crain & Bertness, 2006). A better understanding of how these organisms live and how their habitats are composed is of major importance. Hereupon, many studies have been conducted to determine the habitat of these organisms through habitat suitability maps (HSM), produced from ecological modeling (Bryan & Metaxas, 2007; Davies & Guinotte, 2011; Davies *et al.*,

2008; Giusti *et al.*, 2014; Guinotte & Davies, 2014; Leverette & Metaxas, 2005; Vierod *et al.*, 2014).

1.2 Ecological niche modeling

Ecological Niche Modeling (ENM) is a strategy used for estimating the potential distribution of species, based on the spatial variation of the environmental variables that promote their settlement. A statistical algorithm is applied to relate environmental data (predictors) with known occurrence data to classify regions with resembling environments (Guisan & Thuiller, 2005; Guisan & Zimmernann, 2000; Kearney & Porter, 2009; Pearson *et al.*, 2007). With ENM it is possible to predict species' distribution in unknown locations or in different time scales, by extrapolating in space and time. Besides the environmental niche, this modeling technique has also been widely used to address population ecology issues, like, species' invasion (*e.g.* Peterson & Vieglais, 2001), predictions of past distribution based on biogeographic patterns (*e.g.* Neiva *et al.*, 2014) and also, assessment of the impact of climate change in future distribution (*e.g.* Assis *et al.*, 2014).

Modeling techniques had a rapid growth in the last years due to technological advances and its applications had expanded to other areas, including ecology. The incorporation of Geographical Information Systems (GIS) in ecological studies, made possible to manipulate environmental and distribution data in a specific geographical space (Guisan & Thuiller, 2005). Several environmental variables have been measured over the years, making it possible to generate useful databases for modeling. (*e.g.*, BODC, NOAA – NODC). Beside environmental data, species known occurrences datasets have also been created, by assigning to each occurrence point its geographic coordinates (*e.g.*, GBIF, OBIS). Each variable is presented as a layer, and the study area is a result of the combination of multiple layers. The ability to use both data modeled in the same format is what makes possible to extrapolate in space and time (Elith & Leathwick, 2009; Guisan & Thuiller, 2005).

The estimation of species distribution using environmental information can be obtained from two types of models that use environmental information for the estimation of species distribution: 1) mechanistic model, the potential extent of species' niche is evaluated considering the physiological traits of organisms in response to environmental conditions (Kearney & Porter, 2009); and 2) correlative models, describe the abiotic interactions between the organism and its local environment, requiring only information about environmental variables and species occurrences (Pearson *et al.*, 2007). These two approaches provide different kinds of information, so before the development of the model, the study purpose must be settled. Correlatives are the most commonly and easier models to use, and also, the most appropriated model for the assessment and prediction of species' spatial distribution (Kearney *et al.*, 2010; Pearson *et al.*, 2007).

The assumptions behind ENM methods theory should be taken into consideration before models' applications. First assumption relies on the choice of predictors, since the environmental variables chosen to describe species' range limits determined the description level of the predictive model (Elith & Leathwick, 2009; Guisan & Thuiller, 2005). Predictors selection must be carefully performed, since they have to directly or indirectly evidence the most important types of influences on species, such as: 1) regulators, factors that regulate the physiological adaptation of species to environmental conditions; 2) disturbances, any facts which has an impact on the functioning of the environmental system; and 3) resources, all the elements that provide energy for the persistence and survival of species (Guisan & Thuiller, 2005). When relevant variables are not considered, overprediction can occur with the species being described in places where, in fact, does not occur (Pearson *et al.*, 2007). Second assumption states that the species is in a "pseudo-equilibrium with its environment" (Elith & Leathwick, 2009; Guisan & Thuiller, 2005). These assumptions have to be fulfilled taking into consideration the model's projections in time and space (Guisan & Thuiller, 2005).

Ecological niche is an important concept to highlight when studying the distribution of species. There are two different approaches of this concept operating at different spatial scales, fundamental and realized niche (Guisan & Thuiller, 2005). Fundamental niche describes the abiotic interactions between the physical environment and resources of a potential area occupied by an organism, the suitable habitat (Hutchinson, 1957; Pearson *et al.*, 2007). Because only part of the suitable habitat is

really occupied by organisms, another definition assigned is realized niche, which defines the areas where the species actually occurs in nature including biotic interactions and limiting factors (Hutchinson, 1957; Pearson *et al.*, 2007). For the purpose of our work, we consider the fundamental niche concept, since it operates on a global scale, and because our data only includes environmental information.

The basic process of modeling is to use occurrence data inserted in a geographic space, extrapolated to an environmental space and project them back to the geographic space. The choice of the modeling procedure should be made according with the purpose of the study. There are three different categories of modeling methods, based on the type of data used: presence-only, presence-absence and presence-pseudo-absence. Since our modeling approach is based on the fundamental niche, we are interested in unsuitable conditions as well, so absence data is useful for model construction (Guisan & Thuiller, 2005). Therefore, for a better prediction of the suitable habitat we choose to use presence-pseudo-absence data. After deciding what type of data to be used, it is necessary to proceed to its treatment in order to optimize the performance of the model. In this step, both distribution and environmental data were transferred to the geographical space of the study area, with the appropriate spatial scale, resolution and extent. Spatial scale is directly related to the extent of the study area and resolution. Resolution defines the properties of data, so, it is determined according to the resolution available for both distribution and environmental data. The ideal is to use the same resolution for both data, and the finer resolution the better quality of model predictions (Elith & Leathwick, 2009; Guisan & Thuiller, 2005). Spatial scale is also dependent on the previous knowledge of the motion ability of the target species. When modelling highly mobile species, each cell of the geographic space must contain all the possible environments the species can inhabit, increasing the complexity of the model. Sessile organisms or with restricted motion ability are more easily to model because they occupied a narrow range of environments, and so, absences are more accurately obtained (Guisan & Thuiller, 2005).

For the application of model algorithm, two different categories of modeling methods are available based on the type of data used, presence-only and presence-absence. The main difference between these methods relies on the outcome distribution, where presence-only predictions described only sites where the species occur (*e.g.*

Ecological Niche Factor Analysis (ENFA), Bioclim and Domain), and presence-absence described both suitable and unsuitable conditions for occurrence (*e.g.* Generalized linear models (GLM), generalized additive models (GAM) and classification and regression tree analyses) (Hirzel and Le Lay, 2008; Lobo *et al.*, 2010; Tsoar *et al.*, 2007). We choose to use presence-pseudo-absence data for a better prediction of the suitable habitat, and the model algorithm explored is the machine learning method, Boosted Regression Trees (BRT).

BRT is a modelling technique that combines probable models in order to optimize the performance of a single one (Elith *et al.*, 2006; Leathwick *et al.*, 2006). It is an advanced form of regression developed from Machine Learning approach, a technique known by its artificial intelligence (Friedman *et al.*, 2000; Schapire, 2001). These algorithms relate known distribution and environmental data in order to learn the relationship between them, and then extrapolated to unknown locations. BRT method results from the merge of two algorithms, regression trees and boosting. In regression trees, also known as decision trees, the geographical space of the study area is divided in rectangles identifying regions with the most homogenous response, and then a constant is fitted to each region. Since this is a stagewise process, the input response of one variable is dependent on values of inputs higher in the tree, and so, the interactions between predictors are automatically modeled (Elith *et al.*, 2008). For boosting, the decision trees previously modeled are gradually combined to minimize the loss of function, which is used to measure the error margin associated with models prediction, uncorrected predictions like false presences and false absences. The lower this value the better performance of the model (Elith *et al.*, 2008). First decision tree added in this step represents the regression that better reduces the loss of function. The second decision tree is fitted to the residuals of the first one, and so on. In the end, the final model is the result of a linear combination of all the trees (Elith *et al.*, 2008). This method is characterized by its affordable way to visualize the modeled information. Usually, BRT is performed several times to exclude the highest probability of error and to ensure that all the tested parameters are properly examined (Elith *et al.*, 2008).

Regression models are considered good predictors for modeling ecological data because they manage to quantify and illustrate the relationship between predictors and the response (Elith *et al.*, 2008; De'ath, 2007; Moisen *et al.*, 2006; Recknagel, 2001).

To measure the predictive performance of the models, a validation technique is required (Kohavi, 1995; Pearson *et al.*, 2007). Cross-validation (CV) is an evaluation process of predictive models accuracy that divides the distribution data in two different sets, one for training the models and other for testing them. This procedure is repeated several times and, for each time, it trains and tests the model with a different sorted data, constructing a more accurate final model (Kohavi, 1995). For binary methods, the accuracy of the final model should be measure. True Skill Statistics (TSS) is a method used to access model accuracy though a confusion matrix that relates the presence records with variables in order to obtain true presence, false presence, true absence and false absence. From this matrix we get sensitivity, the proportion of true presences well predicted (omission errors), and specificity, proportion of true absences well predicted (commission errors), which are independent from each other and from prevalence, proportion of locations with presence records. After these settings, the potential distribution of species is described by the probability of its occurrence in each cell along the study area.

1.3 Focal species: *Paramuricea clavata*

The model species chosen for this study was *Paramuricea clavata* (Risso, 1826), currently known as red gorgonian (**Figure**). This is a long-lived organism that normally dwells in vertical rocky walls, but can occur in horizontal substrate, both exposed to high hydrodynamics (Gori *et al.*, 2011 spatial; Linares *et al.*, 2008a). Associated with benthic coralligenous communities, *P. clavata* is essentially found in the Western Mediterranean Sea, and also in Adriatic, Aegean Sea and in the neighboring Atlantic Ocean (Appendix 1). They exhibit a bathymetric range that goes from 5 to 200m (Mokhtar-Jamäi *et al.*, 2011), and present a latitudinal gradient in the depth limits, being the western populations shallower than the eastern ones (Linares *et al.*, 2008a).

P. clavata is a gonochoric surface brooder, meaning that each individual is unisexual and instead of remaining into the water column, the fertilized eggs stay attached to the mother colony (Coma *et al.*, 1995; Linares *et al.*, 2008b). The sex ratio



Figure 2 Focal species, *Paramuricea clavata* .

within these organisms is 1:1, although, the gonadal development is different between sexes (Coma *et al.*, 1995; Linares *et al.*, 2008b). The development of female gonads takes between 13 to 18 months to reach maturity, while male gonads only need 6 to 7 months to reach the same degree of ripeness (Coma *et al.*, 1995). To avoid the loss of gametes in the reproductive season the spawning event is synchronized, being stimulated by the increased of seawater temperature and the lunar cycle, full and new moon. There are two different spawning events, with the first occurring in late May, beginning of June, and the second with an interval of 8 to 10 days (Coma *et al.*, 1995; Linares *et al.*, 2008b). After hatching, the larvae remain suspended in the water column for a short period of time (only a few minutes) and quickly settle near the mother colony (Coma *et al.*, 1995). These organisms exhibit negative phototaxis, which means they go in the opposite direction of light looking for a place to settle, probably to elude competition with algae (Linares *et al.*, 2008b). Although they present a mobile early life phase and are exposed to high currents speed, the feature describe above favors the sinking of the larva, thus reducing the potential dispersion of the species (Coma *et al.*, 1995; Linares *et al.*, 2008b).

As sessile organisms, gorgonians assemblages depend on the physical environment to reach food resources, so the red gorgonian is considered a benthic suspension feeder (Ribes *et al.*, 1999). This species shares its habitat with other suspension feeders that have different feeding strategies and different food target items, reducing competition between them (Gili & Coma, 1998). Moreover, suspension feeders usually have a diverse diet that allows them to explore different food sources, according to their availability in a water column (Coma *et al.*, 1994; Ribes *et al.*, 1999). *P. clavata* filters the surrounding waters and feed on suspension organisms ranging from above 3µm (nanoeukaryotes) to 700µm (zooplankton), and also on detrital POC (particulate organic carbon) (Coma *et al.*, 1994; Gili & Coma, 1998; Ribes *et al.*, 1999). While detrital POC and zooplankton represent an important food resource, live POC has low contribution probably due to the limited capture capacity of this species (Coma *et al.*, 1994). They are considered an ecological success, because they practically don't waste energy in food collection (Ribes *et al.*, 1999), storing it to invest in growth and reproduction (Coma *et al.*, 1998).

The interaction between the substratum and water column made by these organisms, links the benthic and pelagic systems. Besides this connection, suspension feeders regulate abiotic and biotic process of their habitat, making them suitable for many other species (Dame *et al.*, 2001; Gili & Coma, 1998). The ability of *P. clavata* to interfere with its habitat properties, considered it as ecosystem engineer species (Ballesteros, 2006; Cupido *et al.*, 2008; Mistri & Ceccherelli, 1994). It is known that this gorgonian is sensitive to high temperatures and several reports showed an immediate response resulting in mass mortality events along the northern west Mediterranean Sea (Bally & Garrabou, 2007; Cerrano *et al.*, 2000; Coma *et al.*, 2009; Crisci *et al.*, 2011; Cupido *et al.*, 2008; Gambi *et al.*, 2010; Garrabou *et al.*, 2009; Huete-Stauffer *et al.*, 2011; Linares *et al.*, 2005, Martin *et al.*, 2002; Rivetti *et al.*, 2014; Vezzulli *et al.*, 2010). If the increasing temperature continues, the future of these gorgonian and the associated community may be compromised. Recent knowledge of its occurrence in the Atlantic Ocean, along the Portuguese coast, aroused interest for their actual distribution, since this gorgonian was considered to be endemic of the Mediterranean Sea (Linares *et al.*, 2008b; Mokhtar-Jamaï *et al.*, 2011). This coupled with the warning situation in

which this species is found and the improvement of modeling techniques, challenges us to model its ecological niche.

The goal of our study was to estimate the suitable habitat of a structuring species, *P. clavata*, in its global distribution. Specifically, we aimed to 1) determine the relative importance of several environmental variables in order to identify which best describe the fundamental niche of the species, and express this in terms of probability in a habitat suitability map; 2) predict the total distribution to uncover if there are other suitable habitat areas beyond the known occurrence sites reported in literature; and 3) determine the tipping points of the species at its global distribution.

2 MATERIAL AND METHODS:

2.1 Study location

The study area included the entire Mediterranean Sea and the adjacent Atlantic Ocean, from 24°71N and 36°E to 51°85N and 18°W, corresponding to the total distributional range of the red gorgonian, *Paramuricea clavata* (**Figure 2**). This area spans about 4000 km horizontally and encompasses the entire littoral coastlines, islands, underwater features such as submarine mountains and continental shelf slopes. Heterogeneous shores characterize the geography of the area; prominent capes and cliffs demarcate narrow continental shelves with steep slopes and vertical rocky walls, interspaced with flat shores, protected embayments, extended sandy or muddy areas, presenting distinct climate and ocean exposures. Populations of *P. clavata* are found growing over rocky substrates from 15 to about 200 m depth (Mokhtar-Jamaï *et al.*,2011; this study).



Figure 2 Study area and location of the records points of *Paramuricea clavata* used in the model. Coordinates of the occurrence points and respective references are available in Appendix 1.

2.2 Distribution and environmental data

We compiled a database with distributional records (presence data from 1980 to 2014) of *P. clavata* from the literature (Figure ; Appendix 1) and online species databases (GBIF; OBIS). This was completed by a field survey because recent studies (Cúrdia, 2012; Rodrigues, 2008) and divers' observations (<http://www.deepreefs.com>) indicated that the species could be present in the Atlantic Ocean, along the south and west coasts of Portugal. To discover the distribution along the Portuguese coastline, SCUBA dives performed by DeepReefs were made from Tavira (south Portugal) to the Berlengas islands (west Portugal). Additional observation records from Greece were obtained by expert knowledge (see Appendix 1). All coordinates were confirmed in QGIS 2.4 Chugiak, and points that occurred on land or at improbable depths were eliminated. The dataset included 164 occurrences gridded to 0.05 degrees' resolution (about 5 km). Only one occurrence per cell was kept for modeling. Finally, to reduce possible spatial autocorrelation caused by more thoroughly sampled areas, data were visually assessed and some occurrence points were eliminated, keeping a relative balance between geographical regions (Barbet-Massin *et al.*, 2012). The final dataset used for modeling included 50 occurrences of *P. clavata*.

Eleven environmental variables were selected based on the known ecology and physiological tolerances of *P. clavata* (Table 1). Three-dimensional profiles of salinity, ocean temperature and current velocity were obtained from the Global Observed Ocean Physics Temperature Salinity Reprocessing (Guinehut *et al.*, 2012), nutrients (phosphate, nitrate and silicate) were obtained from the World Ocean Database 2013 (Levitus *et al.*, 2013), and slope was calculated using the General Bathymetric Chart of the Oceans (GBCO; BODC, 2013). The environmental data were gridded to the same resolution of the distribution data (0.05 degrees) using trilinear interpolation (location and depth, as deduced from the GBCO), with a vertical coverage from 15 to 200 m. Environmental predictors were tested for correlation (pairwise Pearson *et al.* correlation), because when correlated variables are used for modeling it is not possible to identify which one has better predictive power (Dormann, 2011.), influencing the model performance (Farber & Kadmon, 2003; Fourcade *et al.*, 2014; Guinotte & Davies, 2014). No predictors were eliminated prior to modeling since only nitrate and

phosphate minima presented a correlation higher than 0.8 (Appendix 2), but because nutrient limitation can influence the distribution and abundance of suspension feeders (*e.g.*, Schoener, 1974), both variables were kept.

Table 1 Data type and sources of the environmental variables used for modeling the distribution of *Paramuricea clavata*.

Environmental Predictor	Native resolution of grid	Units	Source	Type of data
Currents (minimum)	0.25°	m/s	Guinehut <i>et al.</i> , 2012	Raster
Nitrates (maximum)	1°	μmol/L	Levitus <i>et al.</i> , 2013	Raster
Nitrates (minimum)	1°	μmol/L	Levitus <i>et al.</i> , 2013	Raster
Phosphates (maximum)	1°	μmol/L	Levitus <i>et al.</i> , 2013	Raster
Phosphates (minimum)	1°	μmol/L	Levitus <i>et al.</i> , 2013	Raster
Salinity (minimum)	0.25°	PSS	Guinehut <i>et al.</i> , 2012	Raster
Silicates (maximum)	1°	μmol/L	Levitus <i>et al.</i> , 2013	Raster
Silicates (minimum)	1°	μmol/L	Levitus <i>et al.</i> , 2013	Raster
Slope	0.08°	Degrees	BODC, 2013	Raster
Temperature (maximum)	0.25°	°C	Guinehut <i>et al.</i> , 2012	Raster
Temperatures (minimum)	0.25°	°C	Guinehut <i>et al.</i> , 2012	Raster

2.3 Preparing pseudo-absences

Boosted Regression Trees uses binary occurrence data (presence and absence) but accurate absence data were unavailable for *P. clavata*, so pseudo-absences were generated through a Mahalanobis Distance function (Farber *et al.*, 2003). The function uses the presences and normalized environmental predictors (Calenge *et al.*, 2008) to determine an Environmental Suitability Map (ESM), specifying for each cell how distant (in probability of occurrence) their local habitat is from the niche optimum. Pseudo-absences were randomly selected from the ESM cells that presented a probability ≤ 0.2 . This threshold favors sensitivity (correctly predicted presences), over specificity (correctly predicted absences), a desirable criterion in conservation to avoid rare/important species being wrongly classified as absent (Jiménez-Valverde & Lobo, 2007). Lower thresholds, *e.g.* 0.1 (Chefaoui & Lobo, 2007), may lead to over-prediction of the niche by selecting pseudo-absences from a very narrow set of conditions in those map cells. To produce a model with better accuracy and reduce over-prediction, we followed the recommendations of Chefaoui & Lobo (2007) for choosing pseudo-

absences in areas distant from presence data, and of Barbet-Massin *et al.* (2012) for randomly selecting the same proportion of absence as presence data. After this data treatment, both distribution and environmental data are ready for the application of the model algorithm.

2.4 Modeling procedure

For prediction of the distribution of *P. clavata* a BRT model was developed. The performance of this model was optimized with four parameters: bag fraction, corresponding to the fraction of the training data randomly selected for the next tree to be fitted; learning rate (*lr*), responsible for shrinking the contribution of each tree added to the model; number of trees (*nt*), which corresponds to the number of iterations occurred in each model; and tree complexity (*tc*), which regulates the number of nodes in a tree and, hence, the level of interaction between predictors (De'ath, 2007; Elith *et al.*, 2008).

The best BRT model was built in two stages, one using all predictors (full model, see in Appendix 3) and the other selecting only the predictors with greater relevance to our species (reduced model, see in Appendix 3), to obtain the lowest model deviance. In the first stage, full model, the best combination of values for all model parameters were determined through a 10-fold cross-validation procedure over *lr* values of 0.01, 0.005, 0.001 and 0.0005, *tc* from 1 to 11, corresponding to the number of environmental predictors, and a bag fraction of 0.5 (Elith *et al.*, 2008). For each model, the optimal combination of parameters was determined. Each model run included a 10-fold cross validation comprising nine unique data subsets used for model training, and a unique omitted subset used for model testing (Elith *et al.*, 2008). Model deviance accounts for the loss of function in fitted models, so lower deviance values correspond to a better model performance. The lower value of deviance obtained with the full model indicated a better model performance with the removal of the variables with lower contribution (Elith *et al.*, 2008). For model reduction (reduced model), four predictors were removed to examine the change in model deviance. The best combination of values for all model parameters was again determined with a 10-fold cross-validation procedure, using the

same values for *lr* and bag fraction, but a *tc* from 1 to 7. The combination of parameters that corresponded to the lowest cross-validation deviance was used for the fit of the entire *P. clavata* training dataset to the optimal BRT model. After determined the best parameters and obtained the optimal predictive model, the HSM was produced.

The final accuracy of the predictive models was verified through True Skill Statistics (TSS), by comparing the predicted distributions with the presence data. The measure of accuracy of this method is threshold-dependent since it is applied to binary data (presence-absence) (Allouche *et al.*, 2006). Although a specific threshold could be established, we prefer to use all possible values (100), and the one with best TSS was chosen. The relative influence of each predictor used in the final BRT model was measured based on the number of times a predictor is selected for splitting the tree, weighted by the squared improvement to the model it produced, averaged over all trees (Elith *et al.*, 2008). In order to visualize the effect of the predictors in the final model we produce partial dependence plots. This function produces a graphical representation of the interactions of each predictor when combined with the others in the final model (Elith *et al.*, 2008).

Environmental tipping points were determined (minimum and maximum levels with occurrences) for the variables that best described the niche of *P. clavata* by plotting the probability of occurrence versus the range of each predictor (**Figure 3**). The tipping point value was established according to the threshold that maximizes the model accuracy.

Niche modeling and area calculations were performed in R version 3.1.1 (R Foundation for Statistical Computing, 2014) using the packages raster, rgdal, adehabitat, SDMTools, sp, dismo, parallel, doParallel and gbm. For a more detailed of the procedure, see the R script available in Appendix 3.

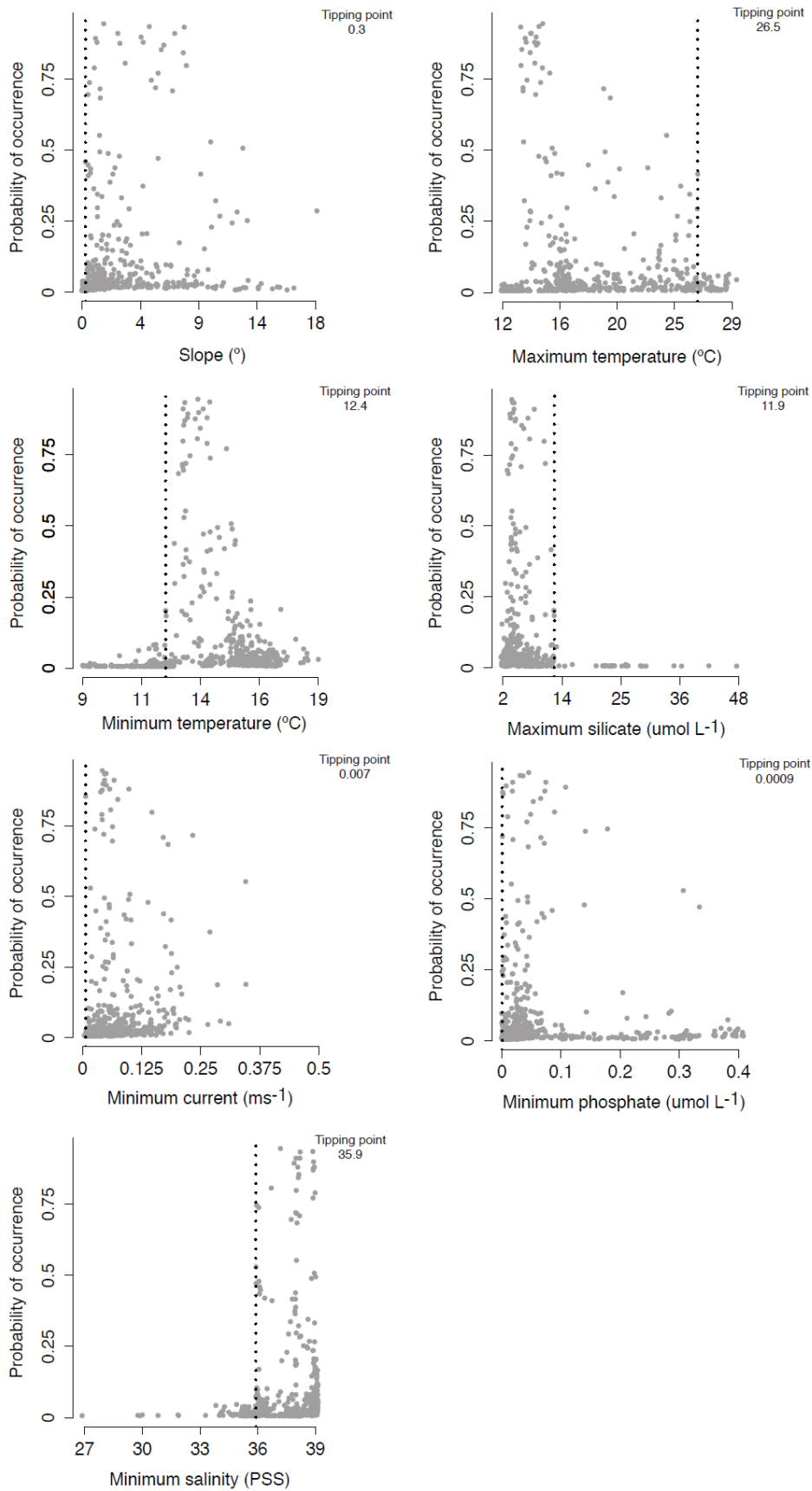


Figure 3 Partial dependence plots of the fitted values of the final predictor variables used for modeling. X-axis represents the range of each predictor and y-axis describes the influence of the predictor on the occurrence of *Paramuricea clavata*. The vertical dotted line corresponds to the tipping point. Temperatures were modeled as monotonic variables.

3 RESULTS

3.1 Model performance

In total, the depth of the 50 records used for modeling varied between 15 and 100 m. The eleven environmental variables considered in the initial model were reduced to seven (**Figure 4**), which gave excellent prediction ability to the model (AUC: 0.98, TSS: 0.96, specificity: 0.98, sensitivity: 0.98, deviance explained: 0.60) and allowed an accurate description of the species distribution. The global distribution of *P. clavata* was best explained by slope (21.8%), maximum and minimum ocean temperature (17.8% and 17.3%, respectively) (**Figure 5**). BRT showed that the suitable habitat of *P. clavata* is also controlled by silicate (13.4%), current velocity (11.2%), phosphate (10.3%) and salinity (8.3%).

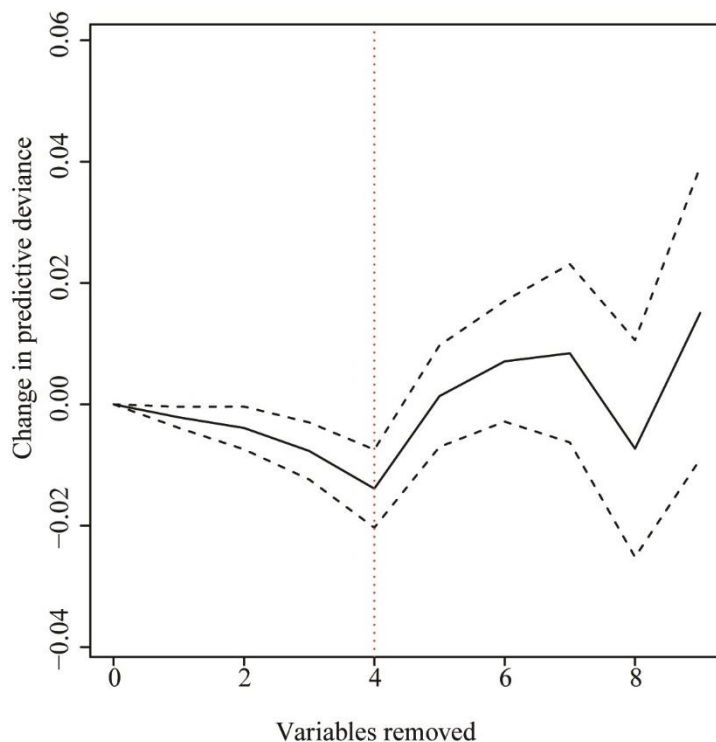


Figure 4 Change in predictive deviance obtained in model reduction and confidence interval. Red dashed line represents the actual number of variables removed. The initial model contained eleven predictors, of which four were removed: maximum nitrate, minimum nitrate, maximum phosphate and minimum silicate.

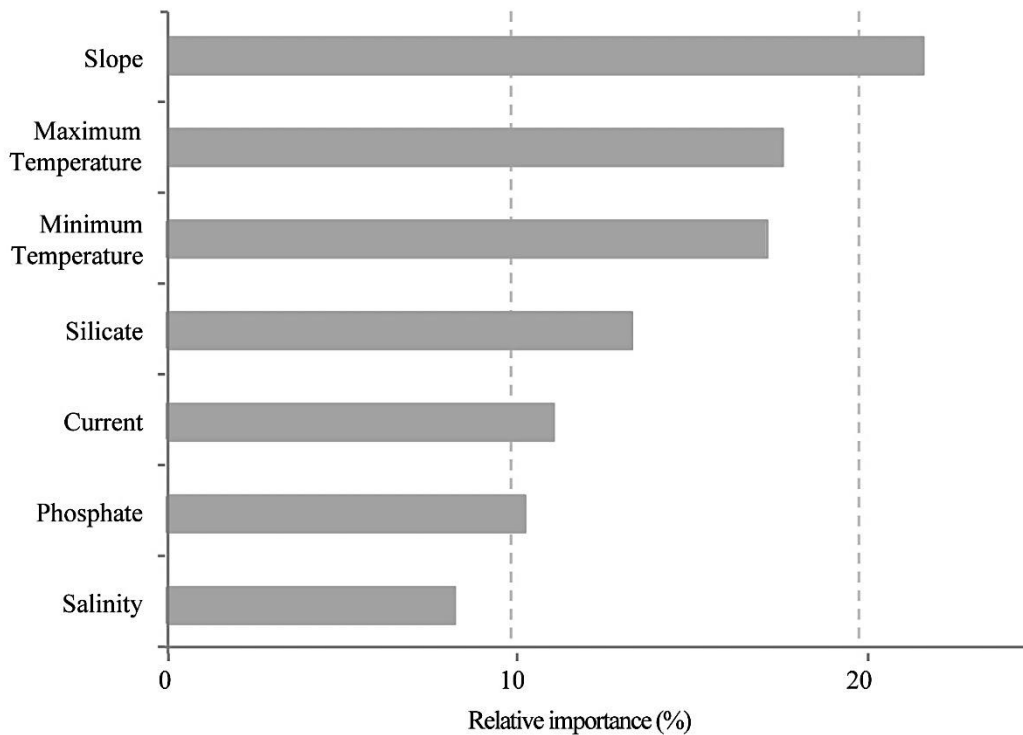


Figure 5 Relative contribution (%) of each predictor for the production of the suitable habitat map of *Paramuricea clavata*, after model reduction.

Partial dependence plots were produced to describe the response curve of each predictor in the final model, taking into consideration the interaction between all predictors (**Figure 6**). Analyzing all response curves, we clearly see a negative effect of the slope above 10°, meaning that below 200 m of depth, the habitat is unsuitable for *P. clavata*. Silicate, currents and phosphate didn't present different response curves when interacting with the set of predictors. Since the variation of predictors is small for currents and phosphate, we must be careful when interpreting the data. With respect to the salinity curve, it presents a rapid change in the response when exceeds the tipping point, although, some individual may occur in areas below 36 (PSS). This variable is clearly dependent on the other predictors, considering that it shows little variation over the study area. When interacting with other predictors, the effect of minimum temperature is maintained, with the habitat being suitable above 12°C. This indicates that the response of this predictor is independent of the others. The same is not observed for maximum temperature, since the critical value is lower than the obtained for the tipping point.

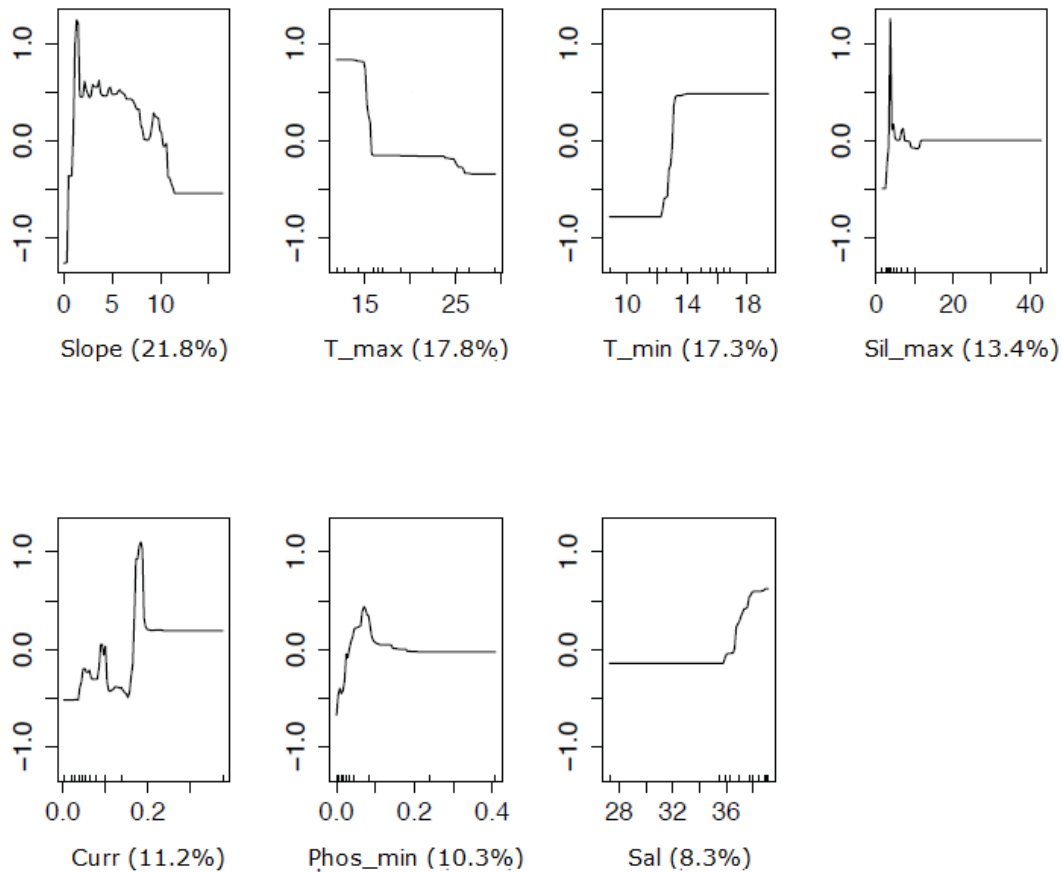


Figure 6 Response curves of fitted values to each predictor used for modeling. The range of each predictor is represented in the x-axis. The occurrence of *Paramuricea clavata* is represented in the y-axis, where 1 = presence, 0 = absence.

3.2 The niche of *Paramuricea clavata*

Our model allowed the estimation of tolerance levels for the environmental variables that best predict the distribution of *P. clavata* (summarized in Table 2) and a comparison with published limits. At the sites where this coral occurs (Appendix 1), the range of optimal conditions indicated by the model was mostly concordant with values obtained from the literature. Together, these values define the species-specific tipping points, here defined as the points at which the build-up of small changes in the environmental variable will culminate on a significant effect for the species, for example a small increase in the maximum temperature tolerated by *P. clavata* (26.5 °C) will strongly decrease its probability of occurrence (partial dependence: **Figure 3 and 6**).

Table 2 Environmental predictors chosen to characterized the ecological limits of *Paramuricea clavata* and corresponding tipping points (TP) obtained from the model (TP.Mod.) and from a literature review (TP.Rev.). TP are levels beyond which the species is unlikely to occur. N.a. indicates data is not available. Occurrence range describes the predictor variation found within our presence data. Raster range describes the predictor variation along the study area.

Environmental predictor	Units	Raster range	Occurrences range	TP.Rev	TP.Mod	References
Slope	Degrees	0.0005 - 34.63	0.26 - 12.40	Steep slope	<0.26	Garrabou <i>et al.</i> , 2002
Max. Temperature	°C	11.9 - 29.4	13.4 - 26.5	>25	>26.5	Garrabou <i>et al.</i> , 2002; Kipson <i>et al.</i> , 2014; Bally & Garrabou, 2007
Min. Temperature	°C	8.8 - 19.5	12.4 - 15.9	12	<12.4	Garrabou <i>et al.</i> , 2002; Gori <i>et al.</i> , 2013
Max. Silicate	µmol/L	1.42 - 51.77	2.48 - 11.92	N.a.	>11.9	N.a.
Current	m/s	0.0004 - 0.42	0.007 - 0.34	0.104 ^a	<0.007	Ribes <i>et al.</i> , 1999
Min. Phosphate	µmol/L	0 - 0.43	0.001 - 0.33	Almost undetectable	<0.001	Ballesteros, 2006
Salinity	PSS	26.9 - 39.1	35.9 - 39.1	37 - 38	<35.9	Ballesteros, 2006

^a This was the only water flow value obtained in the literature and does not correspond to a true tipping point but rather to the local average current speed found in the cited study.

The ecological limits found by our model indicated that *P. clavata* tolerates a wide range of a few environmental variables (*e.g.* slope and current velocity) but that it requires a well-defined range for other environmental predictors (*e.g.* ocean temperatures). Specifically for slope, no maximum was detected indicating that the species might occur in substrates with slope steeper than 18°, the maximum slope value in the presence model cells. The species can grow in vertical walls and overhangs.

The model also indicated that this species habitat occurs in areas with maximum silicate concentration below 12 µM, some current velocity (>0.7 cm.s⁻¹), that it is not excluded from oligotrophic areas which present an almost absence of phosphate, and, naturally, is dependent on marine waters. Although the species tolerates oligotrophic areas and low current speed, the fitted functions of nutrients and currents (**Figure 6**) show that it likely thrives better in areas with stronger water flow (current at about 20 cm.s⁻¹) and levels of phosphate minimum near 0.1 µM.

3.3 Predictive suitable habitat

The niche modeling results showed that *P. clavata* has a restricted habitat along a widespread distribution. The construction of a habitat suitability map allowed identifying probable areas for *P. clavata* extending almost continually throughout the coastlines of the west Mediterranean Sea spreading eastwards until the Aegean Sea, and in the neighboring Atlantic Ocean from North Morocco to west Portugal, including offshore seamounts (**Figure 7**). The majority of highly suitable habitat was predicted to occur along the Catalan and Balearic coasts, along the Ligurian coastlines, including Corsica and Sardinia, along the west and south coasts of Italy, extending into Sicily and into the Ionian and Aegean Seas. Other significant regions of suitable habitat for *P. clavata* are found beyond already known locations, in the Alboran Sea, especially in the Moroccan and Algerian coastlines and in offshore seamounts and islands (*e.g.* Alboran Island), continuing along Tunisia and extending along the shallow sea bottom into Sicily. Although with a much lower probability and area, suitable habitat for *P. clavata* was also found along the Libyan and Egypt coasts, and a few lower probability cells in Israel, Lebanon, Syria, north Cyprus and southeastern Turkey. Major discontinuities in the species niche were found to occur in the southeastern Mediterranean, from the eastern Greek islands and Turkey to the Levantine Sea and in the northern Adriatic Sea. Additionally, the model identified the Atlantic distributional limits for *P. clavata* as located south of Agadir on the west coast of Morocco and north of the Berlengas islands in Portugal, although a discontinuity occurs between southwestern Morocco and the Strait of Gibraltar. The global estimated suitable habitat area for *P. clavata* is 239 626 km².

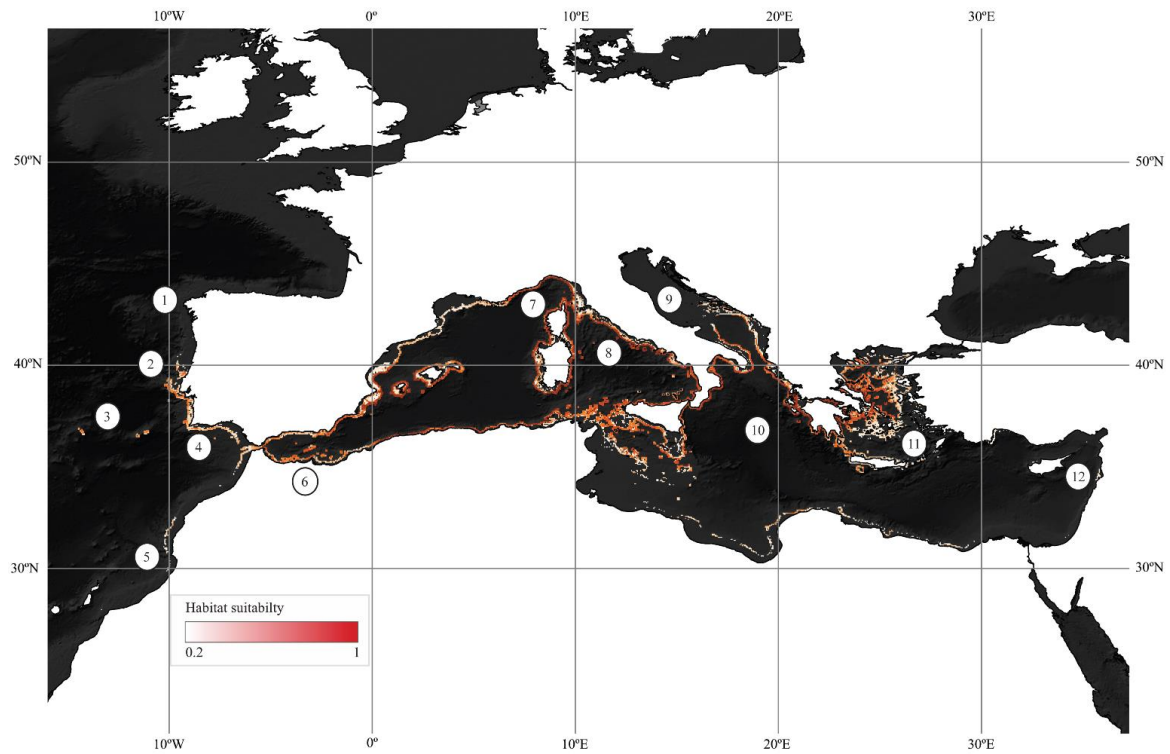


Figure 7 Predicted habitat suitability for *Paramuricea clavata*. Numbers refer to some locations mentioned in the text: 1 - Galicia (Spain), 2 - West Portugal, including Berlengas islands, Cape Espichel and Sines, 3 - Gorringe Bank (seamounts), 4 - Algarve Basin, including submerged caves and Tavira (south Portugal), 5 - Agadir (west Morocco), 6 - Alboran Sea, including islands and seamounts, 7 - Ligurian Sea and Pelagos Sanctuary, 8 - Tyrrhenian Sea and Vercelli seamount, 9 - Adriatic Sea, 10 - Ionian Sea, 11 - Aegean Sea, 12 - Levantine Sea with Cyprus to the left and the Syrian-Lebanon border to the right. Some numbers were placed near the site they are referring to, not to overlap suitable habitat areas.

4 DISCUSSION:

4.1 Model performance:

BRT was the model technique chosen for the prediction of the suitable habitat of *P. clavata* and it proved to have good predictive power by presenting high accuracy values. This machine learning approach allow the recognition of patterns and it is flexible when dealing with different types of predictors, which improves the accuracy of the models (Elith *et al.*, 2008). Due to the characteristics it presents, this method has been used not only to predict the current distribution of species (*e.g.*, Bryan & Metaxas, 2007; Franklin *et al.*, 2013), but also to infer how the distribution patterns varied in the past (*e.g.*, Neiva *et al.*, 2014) and how will they vary in the future (as a consequence of climate change) (*e.g.*, Assis *et al.*, 2014). Although ENM is an important tool to access the distribution pattern of the species, it presents some limitations that may condition the validation of the suitable habitat generated, such as sampling bias, loss of function, data resolution and overfitting.

Our model was built based on the fundamental niche of the species, which means that suitable habitat obtained only reflects the environmental conditions favorable for the species presence. Biotic factors, such as competition and dispersion, are also important variables in shaping the distribution of species, because they can prevent their occurrence in habitat considered favorable (Elith & Leathwick, 2009; Pearson *et al.*, 2007). Thus, it is possible that some of the potential areas predicted for the occurrence of *P. clavata* are not actually occupied, because the species has not reach yet those places, or because the presence of other species leads to competition for food, which makes the habitat less efficient for the ecology of the species. So, models based on environmental variables do not include biological barriers of colonization, creating uncertainty about the model's predictive power (Araújo & Guisan, 2006; Mokhtar-Jamäi *et al.*, 2011). However, we believe that our model has achieved good distribution limits of the species range, given that the obtained tipping points are in concordance with published data (**Table 2**), and they could explain the high value obtained for sensitivity.

Other concerns related to the good performance of the ENM are sampling bias, the choice of pseudo-absences, data resolution and overfitting (Elith *et al.*, 2008; Pearson *et al.*, 2007). The spatial distribution of both presence and pseudo-absences points influence the performance of the model. With respect to presence points, since we intended to predict the global distribution of *P. clavata*, some regions of our study area were more sampled than others, meaning that important locations of occurrence of the species may not be included for modelling. To minimize sampling bias, we didn't include all the presence points in the model, and, BRT was programmed to reduce bias through its stagewise process, sequential modeling of the trees based on the residuals of the previous tree (Elith *et al.*, 2008; De'ath, 2007). For the generation of pseudo-absences, BRT overcomes the bias through a random selection of pseudo-absences based on a relation between the geographical space and the environmental variables. Regarding the appropriated number of pseudo-absences, it is recommended for BRT to use the same proportion of the presence records (Barbet-Massin *et al.*, 2012), as we did.

The environmental dataset chosen for the characterization of the fundamental niche of *P. clavata* doesn't represent a full description of the abiotic requirements of the species. A limitation for variables selection is the availability of data and its resolution for modeling. For marine environments, the description of the sea floor is difficult to obtain, and so there are few available datasets with fine resolution for the characterization of deep ocean habitats (Bryan & Metaxas, 2007; Davies *et al.*, 2008; Tittensor *et al.*, 2009; Rengstorf *et al.*, 2012; Vierod *et al.*, 2014). Data resolution (cells grain size) is very important for the description of a physical space because that's what defines the data variability (Vierod *et al.*, 2014). Finer resolution allows to capture with more detail the interactions between environmental variables and the occurrence, which coarse resolution does not detect, directly influencing the projection of suitable habitat (Guisan & Thuiller, 2005; Vierod *et al.*, 2014; Rengstorf *et al.*, 2012). Because the choice of grain size for environmental variables affects predictions, coarse data introduces uncertainty in the model (Guisan *et al.*, 2007; Rengstorf *et al.*, 2012; Vierod *et al.*, 2014). It's important that both distribution and environmental data to have the same resolution, especially when modeling benthic habitat where a small environmental variation can translate a different habitat, leading to overestimation of the suitable habitat (Guisan & Thuiller, 2005; Rengstorf *et al.*, 2012). Although, given that the resolution available for variables is not the same, mixing them will lead to the

loss of information. The major limitation for modelling the deep-sea is the lack of data with good resolution, which causes uncertainty in model projections. Since benthic terrain characteristics are a good representation of the bottom dynamics, good resolution of these data allows the detection of small elevations and bottom features that locally alter the distribution of benthos (Vierod *et al.*, 2014). Hard substrate is an important variable for the distribution of corals, including *P. clavata*. Since it presents high variability in small scale, this variable is difficult to measure and therefore there is insufficient data to use it as a predictor. To address this situation, terrain variables like slope and bathymetry can be used as a proxy of substrate type, low sediment deposition and can also give an insight about the most relevant position of the current flow. In this way they are good indicators of hydrodynamic and productive areas for benthic suspension feeders (Bryan & Metaxas, 2007; Dolan *et al.* 2008; Guinan *et al.* 2009; Rengstorf *et al.*, 2012; Vierod *et al.*,2014; Wilson *et al.* 2007). Using a coarse resolution of this predictors leads to the loss of bottom detail features like steep slopes, seamounts and other small elevations, underestimating the importance of these predictors. While data resolution is a problem in many modeling techniques, BRT is considered to be one of the less affected by cells grain size (Guisan *et al.*, 2007).

Overfitting is another problem associated with SDM since this also decreases the predictive power of the models. While most of the models try to avoid this, BRT reduces overfitting through CV, by randomly splitting the distribution data in k equal parts/sets for training and testing the model. In this way, all the possible data combinations are used for model validation, which means that overfitted data is also considered (Elith *et al.*, 2008).

4.2 Characterization of *Paramuricea clavata* ecological niche

The use of ENM for the description of the suitable habitat of species has become an efficient tool to determine where species occur. Several studies have been conducted to determine the suitable habitat of deep-water corals, however, the vast majority was made at a local scale (*e.g.* Bryan & Metaxas, 2007; Davies *et al.*, 2008; Dolan *et al.*, 2008; Giusti *et al.*, 2014; Leverette & Metaxas, 2005;). In this sense, this study was the

first not only in determining the environmental envelope that describes the suitable habitat of a very important structuring species of the Mediterranean coralligenous assemblages, *P. clavata*, but also in estimating its potential global distribution. We also identified the variables that best explain this distribution and determined the species tolerance range for these variables. Moreover, the compilation of known data with new occurrence points provided a unique opportunity to determine the potential suitable habitat for red gorgonian.

The suitable habitat of *P. clavata* predicted by our model was mainly characterized by productive and hydrodynamic marine environments associated with a complex topography, occurring in a narrow coastline, around islands and seamounts, as well as other rocky elevations. Although sea bottom topography is characterized by several variables, in this model it was represented by the slope. This terrain variable was considered the most relevant predictor for the geographic distribution of *P. clavata*, presenting a wide range of tolerance with no maximum value attributed (according to Ballesteros, 2006; Gori *et al.*, 2011). Coralligenous assemblages generally inhabit in sloping areas associated with rocky substrate, because these features allowed the change of speed and direction of currents (Leverette & Metaxas, 2005; Yesson *et al.*, 2012). The restricted distribution of the red gorgonian is largely explained by the complexity and relief of the bottom, having high preference for vertical slopes (Gori *et al.*, 2011). Depending on the variation of other conditions, like current speed and primary production, these assemblages may also occur in more horizontal locations, where the relief of the bottom may be less pronounced (Gori *et al.*, 2011). Studies determining the suitable habitat for other octocoral species have also used slope as an environmental predictor for quantification of the bottom topography, presenting itself as a good proxy of the complexity of the seabed. (Bryan & Metaxas, 2007; Davies *et al.*, 2008; Davies & Guinotte, 2011; Dolan *et al.*, 2008; Giusti *et al.*, 2014; Guinan *et al.*, 2009; Leverette & Metaxas, 2005; Tong *et al.*, 2012; Yesson *et al.*, 2012).

Besides slope, temperature was also considered very important in shaping the distribution of *P. clavata*. This variable is considered very important for the regulation of species' biological processes, and so, the species tolerance range presented for this predictor was well defined, 12.4 - 26.5 °C. In the case of *P. clavata*, temperature has a key role both in terms of food availability and reproductive success. With respect to the

reproductive success, this variable is extremely important since spawning events are synchronized by a temperature increase occurring between May and June (Coma *et al.*, 1995; Linares *et al.*, 2008b). An abnormal variation in temperature can prevent the meeting of the gametes, since their development presents different time periods (Gori *et al.*, 2007). For long-lived organisms like gorgonians, where reproductive success is very low, temperature oscillation may be crucial for their maintenance. Several mortality events were reported for populations of *P. clavata* as result of an irregular increase in seawater temperature (Cerrano *et al.*, 2000; Coma *et al.*, 2009; Crisci *et al.*, 2011; Cupido *et al.*, 2008; Gambi *et al.*, 2010; Garrabou *et al.*, 2009; Huete-Stauffer *et al.*, 2011; Linares *et al.*, 2005; Martin *et al.*, 2002; Rivetti *et al.*, 2014; Vezzulli *et al.*, 2010), highlighting the importance of the maximum temperature for the survival of the species. Food availability is greater during winter because periods of heavy rain stir up the sediment and debris accumulated at the bottom, increasing the amount of dissolved and suspended matter. In summer, these flood events are reduced, leading to stratification of the water column, and thus, decreasing the levels of dissolved and suspended matter (Gasith & Resh, 1999). Since suspension feeders are properly adapted to the Mediterranean climate, they can survive periods of food scarcity in summer, by decreasing the metabolic activity (Coma *et al.*, 1994; Ribes *et al.*, 1999). To address this situation, they store energy during periods of greatest abundance of food, winter (Gori *et al.*, 2013; Rossi *et al.*, 2005). This explains the lower contribution of nutrients for the distribution, indicating that the hydrodynamics of the dwelling place is crucial to the survival of the species. Analyzing the data of the two temperatures along the study area, we noticed that they exhibit different gradients (see Appendix 4). The variation of the minimum temperature presented a latitudinal and longitudinal gradient, with minimum temperature increasing from north to south and from west to east. For maximum temperature, a vertical gradient was observed, with shallow areas presenting higher temperatures than deeper ones, and a longitudinal, from west to east. Comparing the two temperature gradients, we observed that this variable is highest in the south eastern Mediterranean, and in shallower areas. As we head towards the northwest of the Mediterranean Sea and Atlantic Ocean, or to greater depths, temperature decreases. This temperature gradient coupled with the strong association to slope, may explain the restricted suitable habitat for the red gorgonian.

Benthic communities are normally associated to areas exposed to moderate-high hydrodynamics (Gori *et al.*, 2011). Currents are very important for feeding *P. clavata* because when they flow against vertical rocky walls, the direction and intensity of the flow speed change traversing the wall towards the top with higher velocity (Leichter & Witman, 1997). These currents transport microorganisms, influence the diffusion rate of nutrients and promote the re-suspension of the particulate matter, increasing the available food for this gorgonian and thus promoting its growth (Coma *et al.*, 1998; Mistri, 1994). We used minimum current velocity as a representation of hydrodynamics, and the tipping point obtained was lower than expected (0.007ms^{-1}). The coarse resolution data used for this predictor could explain the low value since, small local changes in direction and speed of currents caused by topographical elevations may not be taking into account, and because each cell represents the mean value of the region. Besides the feeding importance, currents are probably the main way of dispersion for this species, since the larvae presented phototaxis, and so, quickly settle down near the mother colony in order to avoid light (Linares *et al.*, 2008).

As suspension feeder, *P. clavata* feed essentially on microorganisms and suspended organic matter found in the water column. Prey capture and ingestion of detrital POC exhibit seasonal variation, with primary production peaks (and consequent zooplankton peak) occurring in spring and autumn (Coma *et al.*, 1994), and the abundance and ingestion rates of POC being higher during winter and spring (Ribes *et al.*, 1999). Productivity presented lower contribution for the persistence of red gorgonian, and was quantified in our model by minimum phosphate and maximum silicate concentrations. Phosphate is considered as one of the most important limiting factors for primary production in the Mediterranean Sea (Ballesteros, 2006). When available, this nutrient is quickly consumed by phytoplankton, and so, lower concentrations indicate the presence of primary producers (Davies *et al.*, 2008; Montagna *et al.*, 2006). High concentration levels of phosphate in the water could lead to an exponential growth of benthic communities, causing eutrophication and, consequently, habitat destruction (Davies *et al.*, 2008). With concern to silicates, the amount of dissolved silica in the water could also be considered as a proxy of productivity, with high levels of dissolved silicate indicating lower levels of primary producers (Bonilla & Piñón, 2002; Davies *et al.*, 2008). In the presence of diatoms, dissolved silica concentrations decrease because these microorganisms use this element

for the formation of their frustule (Martin-Jézéquel *et al.*, 2000). Besides productivity indicators, diatoms include the diet of the red gorgonian (Ballesteros, 2006; Coma *et al.*, 1994; Ribes *et al.*, 1999). Known as an oligotrophic system, the Mediterranean Sea exhibits low concentrations of nutrients (Coll *et al.*, 2010). Concerning phosphates concentrations, they were low throughout the study area, even in the adjacent Atlantic Ocean. In contrast, silicate concentrations presented some variations, being higher in the northern Adriatic and along the coast of Tunisia, where the concentrations of phosphates are very small (see Appendix 5). This correlation allows a better evaluation of the importance of productivity in habitat description of *P. clavata*.

Salinity (minimum) was considered the predictor with less relevance for the characterization of the suitable habitat of the red gorgonian, presenting a tipping point of 35.9 PSS. This result is explained by the gradient of the salinity data observed along the study area, which presents a small variation (see Appendix 6). There are few studies evaluating the importance of salinity for the distribution of octocorals (Berkelmans *et al.*, 2012; Williamson *et al.*, 2011), however, this predictor has been used to classify the type of environment in habitat modeling (Guinan *et al.*, 2009; Martin *et al.*, 2014; Rengstorf *et al.*, 2012; Yesson *et al.*, 2012).

4.3 Predictive suitable habitat

The predictive suitable habitat obtained for *P. clavata* described a wider distribution area than documented. Analyzing the HSM shown in figure X, we observed that the suitable habitat distribution along the Mediterranean Sea exhibited a latitudinal and longitudinal gradient, with the highest probability of occurrence of *P. clavata* in the Western Mediterranean. This gradient was observed in other studies (Gori *et al.*, 2011; Linares *et al.*, 2008a), indicating a good prediction by our model of the distribution pattern of this species. For the Atlantic Ocean distribution, only a latitudinal gradient was visible from Agadir (Morocco) to Berlengas islands. As was mentioned in the previous section, temperature plays a crucial role in shaping the distribution of the red gorgonian. When comparing temperature gradients (see Appendix 4) with the HSM a possible pattern was observed. Minimum temperature appeared to have greater

influence in the longitudinal distribution of *P. clavata* because both gradients conversely coincide, with the probability of occurrence increasing as the temperature decreases over the study area. Although depth has not been used for modeling, the upper depth limit described by Linares *et al.* (2008a) for distribution of this gorgonian also described a longitudinal gradient, with species occurring in deeper zones in the Eastern Mediterranean. For this pattern, maximum temperature has a greater influence on the vertical distribution, since higher temperatures occur at lower depths. In a general way, most of the suitable habitats occur with a minimum temperature between 11-14°C and a maximum temperature of approximately 16°C (see Appendix 4). In this sense, we can state that the two temperatures have regulate in the distribution of suitable habitats for *P. clavata*.

For the Mediterranean Sea, our model shows new potential suitable habitats in the western side, especially in the Alboran Sea and continuing along the coasts of Algeria, Tunis until Sicily channel (**Figure 7**). The dynamics presented by these regions is directly influence by the exchange of the water masses between the North Atlantic Ocean and the Mediterranean Sea (Milot, 1999). For the Alboran Sea, our study has provided new occurrence points in the Alboran Island and near the Strait of Gibraltar, in Ceuta, which supports the probable presence of the red gorgonian along this sea area. The lack of scientific information concerning the biological composition of the Algerian coast, makes difficult to prove the presence of our species. However, given the high suitability shown along this coast, we search for more information around this region and found two online dive videos showing populations of *P. clavata* occurring near the coast of Oran and Habibas Island. Hereupon, we contacted the diver who published the videos and he confirmed the presence of *P. clavata*, also sending photos for visual confirmation of the species. This validation reinforces the accuracy of our model. With respect to the Strait of Sicily, this area was considered to be good for the proliferation of benthic communities, because elevated topographic features, like seamounts and banks, enable the settlement of several species, by providing shelter and food for their survival (UNEP/MAP, 2015). We didn't find information to confirm our results, although, a study performed in the Tyrrhenian Sea have shown the presence of *P. clavata* in the Vercelli Seamount (Bo *et al.*, 2011). This site was considered in our suitable habitat map as an area of high probability of occurrence, which confirms our predictions. While the western sub-basin proved to be highly suitable, the eastern side showed a different

pattern, with the northeast being more favorable than the southeast. Besides the lower probability of occurrence along the southeastern side, Martin *et al.* (2014) recently published a study concerning the spatial distribution of coralligenous habitats in the Mediterranean Sea, and they also obtained suitable habitats near the coast of Israel, Lebanon, Syria, and in the north of Cyprus Island.

The main contribution of our work for the distribution of *P. clavata* was the new occurrence points found along the south and west coasts of Portugal and in the Gorringe Ridge. Berlengas islands were considered the northern distribution limit, since north of this region the species is absent and it is replaced by the sister-species *P. macrospina* from the Galicia region northwards (Altuna, 1994; 2010). Our model also described suitable areas near the coast of Agadir, Morocco, but there are no studies proving the existence of the red gorgonian. Although, the geographic proximity and climatic influence of Mediterranean Sea make possible the occurrence of this species.

The northern and central Adriatic Sea and the Libyan-Tunisian Gulf were considered as two unsuitable habitats for the red gorgonian, and these areas presented high values for maximum silica, which indicates low levels of productivity. This result, in agreement with other studies, suggest a negative correlation between the amount of dissolved silica and the occurrence of corals, although without a plausible explanation for this relationship (Bonilla & Piñón, 2002; Davies *et al.*, 2008).

4.4 Prediction of range-shifts

One of the main problems we face today are climate fluctuations and the impact they have on the distribution of species. Climate change has leading to the occurrence of abnormal temperature increased periods, causing rapid responses by the species, particularly at the distribution level (Walther *et al.*, 2002). Mediterranean Sea has been one of the most affected regions with climate change, and since it's an ecosystem with very particular characteristics, small changes in its climate have direct influence on the species it hosts (Lejeusne *et al.*, 2009; Rivetti *et al.*, 2014). Several studies have been conducted in the Mediterranean Sea with the purpose of assessing the impact of these

climate oscillations in benthic communities (Cerrano *et al.*, 2000; Cigliano & Gambi, 2007; Crisci *et al.*, 2011; Cupido *et al.*, 2008; Gambi *et al.*, 2010; Garrabou *et al.*, 2009; Huete-Stauffer *et al.*, 2011; Linares *et al.*, 2005; Martin *et al.*, 2002; Rivetti *et al.*, 2014; Vezzulli *et al.*, 2010). Mortality events observed in these studies occurred in the North-western Mediterranean: Balearic, Tyrrhenian and Ligurian Sea. With X affected regions, Ligurian sea presented the highest number of events occurred. One of the species that suffered the greatest population loss was *P. clavata*. The main causes leading to the death of this gorgonian were the increase of seawater temperature in summer, and disease contraction from opportunistic pathogenic microorganisms (Cerrano *et al.*, 2000; Crisci *et al.*, 2011; Cupido *et al.*, 2008; Gambi *et al.*, 2010; Garrabou *et al.*, 2009; Huete-Stauffer *et al.*, 2011; Linares *et al.*, 2005; Martin *et al.*, 2002; Vezzulli *et al.*, 2010). As it was observed in this study, *P. clavata* presents a restricted temperature tolerance range, being sensitive to small changes in seawater temperature. During summer periods, where temperature reaches its highest values, suspension feeders have the ability to decrease their metabolic rate due to food scarcity (Coma *et al.*, 1994; Ribes *et al.*, 1999). In this condition they are more vulnerable to external variations, which may cause physiological stress in irregular warmer periods, allowing infection by pathogenic microorganisms (Cerrano *et al.*, 2000; Gambi *et al.*, 2010; Garrabou *et al.*, 2009; Huete-Stauffer *et al.*, 2011; Martin *et al.*, 2002; Vezzulli *et al.*, 2010). Temperature oscillations have also direct impact on the reproductive cycle of *P. clavata*, since thermal stress affects the development of gametes, reducing the viability of embryos and larvae. Because this species is a slow growing organism, it depends on the reproductive success for the maintenance of their populations (Kipson *et al.*, 2012). Comparing the HSM obtained for the red gorgonian with the areas affected by mortality events (**Figure 8**), we clearly see a match between mortality events and locations where the probability of the species occurrence was higher. This situation jeopardizes the persistence of our species, since much of their suitable habitat is already compromised. Likewise, the associated communities are also threatened.

Due to the impact that climate change has so far in the Mediterranean Sea, several studies have been developed to determine the main changes predicted to occur in the future for this vulnerable ecosystem (*e.g.* Albouy *et al.*, 2013; Giorgi & Lionello, 2008; Philippart *et al.*, 2011; Sánchez *et al.*, 2004). In a general way, a continuous increase in

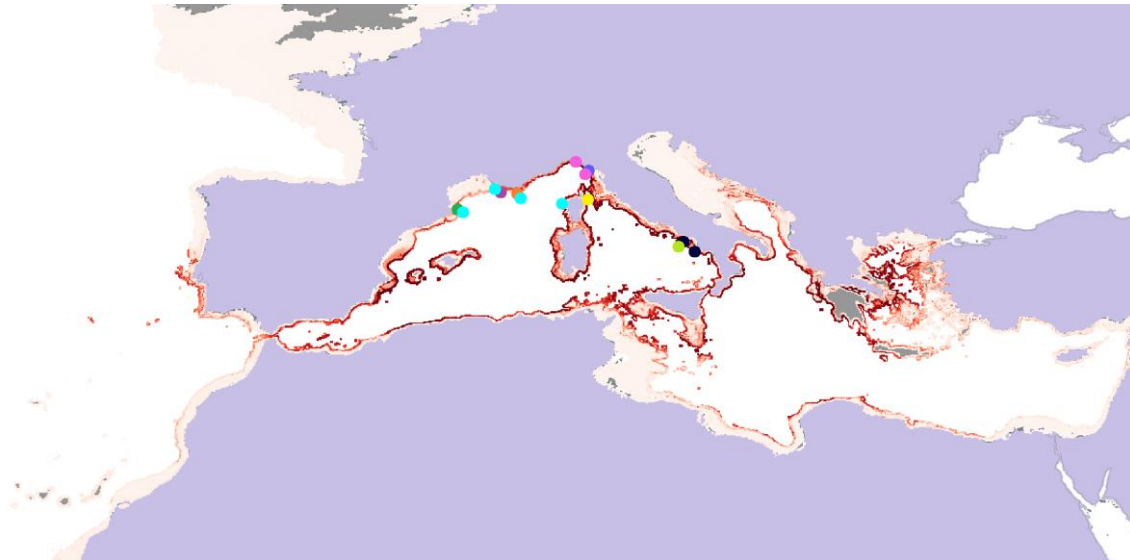


Figure 8 Intersection of suitable habitat of *Paramuricea clavata* with mortality events in the Mediterranean Sea: Cerrano *et al.*, 2000 ● (pink) ; Cigliano & Gambi, 2007 ● (green) ; Coma *et al.*, 2009 ● (dark green) ; Crisci *et al.*, 2011 ● (cyan) ; Cupido *et al.*, 2008 ● (blue) ; Gambi *et al.*, 2010 ● (black) ; Huete-Stauffer *et al.*, 2011 ● (yellow) ; Linares *et al.*, 2005 ● (orange) ; Martin *et al.*, 2002 ● (purple)

mean seawater temperature is expected. In a study conducted by Giorgi & Lionello (2008), besides temperature increase, a marked decrease in precipitation is observed, with increasing of dry periods (in agreement with Sánchez *et al.*, 2004). An interannual variability is also predicted, especially in summer, which would lead to an increase of mortality events. Philippart *et al.*, (2011) predicted that these mortality events will be a result of both extremely high and low temperatures. Moreover, these authors have also reported species distribution shifts, increase of epidemic events and the occupation of niches by invasive species. Predicted warming for Mediterranean waters alters the habitat availability for species, with tropical species being the most favored as the area of their suitable habitat increases. The invasion of tropical species in the Mediterranean puts at risk its biodiversity, as these invasive species will interfere with the normal habitat function (Raitsos *et al.*, 2010).

Albouy *et al.*, (2013) described the impact of future climate change in Mediterranean fishes, and concluded that along the coastal area there was a significant loss in species richness. The main causes for this loss were: shifts in species distribution, reduction in species range size and loss of suitable habitats. In this study, the authors also determined the expected increase (in °C) of sea surface temperatures

(mean, minimum and maximum) for two periods of future times, 2040-59 and 2080-99. A global warming is expected for the entire Mediterranean Sea, with an increase of about 1.4°C in 2040-59 and 2.8°C in 2080-99. Aegean Sea showed higher temperature rise, being expected a variation of 1.6-2°C for 2040 and 2.8-3.8°C for 2080. Comparing these temperature data with our HSM, we observed that one of the places where *P. clavata* is most likely to occur, in the entire Mediterranean Sea and especially in the Eastern basin, coincides with the most affected area with temperature increase, the Aegean Sea. Given this hypothesis, there is a strong probability of these populations not persisting in the future. The upper depth limit of *P. clavata* in the Eastern Mediterranean is lower than the west, indicating that some populations may persist in deeper habitats during to climate change. Regions that provide environmental conditions for the persistence of the species act as refugia areas. (Assis *et al.*, 2015; Bongaerts *et al.*, 2010). Deep refugia are of great importance for the persistence of the species, not only to serve as shelter to face less favorable environmental conditions, but because they allow the survival of populations that present a unique and high genetic variability, rear edge populations (Assis *et al.*, 2015_deep; Hampe & Petit, 2005).

Considering the predicted rise in seawater temperature and taking into account the temperature tolerance range of the species, it is important to understand the impact climate change will have on the distribution of *P. clavata*. One of the main applications of ENM is the extrapolation of data in time and space, allowing the estimation of species' distribution in the past and for the future (Elith & Leathwick, 2009). An example of the use of these models to predict the future distribution based on the increase temperature was carried out with *Fucus vesiculosus*. Currently, the distribution of this species occurs along the north coast of the Atlantic Ocean. When extrapolated to future, the distribution of this species suffers a major contraction especially on its lower limit, affecting rear edge populations (Assis *et al.*, 2014). Given this scenario, there is a strong possibility that the most ancient lineages, as well as, the genetic variability presented by this species to be lost in the coming years. The same situation was reported for *Saccorhiza polyschide*. This species also occurs along the north Atlantic coasts, and has some populations inhabiting in the Western Mediterranean (Assis *et al.*, 2015). Observing the predicted distributions for the future all the Mediterranean populations were lost. As verified for *F. vesiculosus*, once again the populations that suffer greater loss were the ones at the rear edge. However, in the case of *S. polyschide* it is possible

that some populations may occur at greater depths, which means, in deep refugia areas, safeguarding the genetic diversity of this species.

Just as these species are at risk of losing their older populations, *P. clavata* is in the same situation. The loss of rear edge populations has a major implication for the maintenance of the species, since they exhibit an old and distinct genetic variability. Although they don't have direct influence in the recolonization of lost areas, these population retains the genes that allows them to adapt to new conditions. Therefore, identification of rear edge population is very important for understanding populations dynamics.

5 CONCLUSION:

BRT was used in this study to predict, with excellence, the suitable habitat *P. clavata* and determine the environmental tipping points of the seven selected variables (slope, temperature, currents, silicates, phosphates and salinity) considered the most relevant for the description of its niche. We choose this model due to its complexity, recognized as artificial intelligence, to understand and interpret environmental standards, minimizing the main associated problems.

Detection of new points of occurrence in the Atlantic Ocean and Alboran Sea, areas until now unknown, provided a unique opportunity to investigate the global distribution of this species and to test the predictive performance of the BRT. In addition to the Atlantic and Alboran Sea, our study pointed to the occurrence of populations of *P. clavata* along the coast of Algeria and Seamount Vercelli, which were later confirmed. This validation leads us to conclude that the suitable habitat was well modeled, serving as a good indicator for conducting studies in Tunisia coast and along the Atlantic and Mediterranean coasts of Morocco, regions presenting high probability of occurrence of red gorgonian

Massive mortality events occurred in the Mediterranean Sea in the last years, alert to the fragile situation in which these habitats are. With the prediction of a continuous rise in temperature, especially for the Mediterranean, the future of *P. clavata* is compromised given that it presents low tolerance to temperature variations. In addition

to climatic change, this gorgonian is also threatened by human activity, (*e.g.* diving, destructive fishing and pollution), affecting not only local populations but the species in its global range (Ballesteros, 2006; Mokhtar-Jamaï *et al.*, 2011). The loss of unique traits of edge and deep populations compromise the ability to respond to environmental changes. Therefore, the results presented in this study may serve as a basis for future studies, like predicting past and future distribution in order to understand the ecological niche of this species, and develop for a better management approaches for its preservation.

6 BIBLIOGRAPHY

Albouy, C., Guilhaumon, F., Leprieur, F., Lasram, F. B. R., Somot, S., Aznar, R., Velez, L., Le Loc'h, F. & Mouillot, D. (2013). Projected climate change and the changing biogeography of coastal Mediterranean fishes. *Journal of Biogeography*, 40(3), 534-547.

Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). In *Journal of Applied Ecology*, 43, (pp. 1223-1232).

Altuna Prados, Á. (1994). Observaciones biogeográficas sobre los cnidarios bentónicos de la costa Vasca. In *KOBIE (Serie Ciencias Naturales), Bilbao*.

Altuna Prados, Á. (2010). Listado de los cnidarios (Cnidaria) bentónicos del Golfo de Vizcaya y zonas próximas (Atlántico NE). In *Donostia-San Sebastián*. Retrieved from <http://www.faunaiberica.es/faunaib/altuna7.pdf>

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. In *Journal of Biogeography*, 33, (pp. 1677–1688).

Assis, J. F., Coelho, N. C., Lamy, T., Valero, F. A. & Serrão, E. A. (2015). Deep reefs are climatic refugia for genetic diversity of marine forests. In *Journal of Biogeography*, (pp. 1-12).

Assis, J. F., Serrão, E. A., Claro, B., Perrin, C., & Pearson *et al.*, G. A. (2014). Climate-driven range shifts explain the distribution of extant gene pools and predict future loss of unique lineages in a marine brown alga. In *Molecular Ecology*, 23, (pp. 2797–810).

Bongaerts, P., Ridgway, T., Sampayo, E. M., & Hoegh-Guldberg, O. (2010). Assessing the ‘deep reef refugia’ hypothesis: focus on Caribbean reefs. In *Coral Reefs*, 29(2), 309-327.

Ballesteros, E. (2006). Mediterranean coralligenous assemblages: A synthesis of present knowledge. In *Oceanography and Marine Biology: An Annual Review*, 44, (pp. 123-195).

Bally, M. & Garrabou, J. (2007). Thermodependent bacterial pathogens and mass mortalities in temperate benthic communities: a new case of emerging disease linked to climate change. In *Global Change Biology*, 13, (pp. 2078-2088).

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

Berkelmans, R., Jones, A. M. & Schaffelke, B. (2012). Salinity thresholds of *Acropora* spp. on the Great Barrier Reef. In *Coral Reefs*, 31, (pp. 1103-1110)

Bo, M., Bertolino, M., Borghini, M., Castellano, M., Covazzi Harriague, A., Di Camillo, C. G., Gasparini, G., Misic, C., Povero, P., Pusceddu, A., Schroeder, K. & Bavestrello, G. (2011). Characteristics of the mesophotic megabenthic assemblages of the vercelli seamount (north tyrrhenian sea). In *PloS One*, 6, e16357.

BODC (2013) General Bathymetric Chart of the Oceans, Liverpool, U.K.

Bonilla, H. R., & Piñón, G. C. (2002). Influence of temperature and nutrients on species richness of deep water corals from the western coast of the Americas. In *Hydrobiologia*, 471(1-3), (pp. 35-41).

- Botero, C. A., Weissing, F. J., Wright, J. & Rubenstein, D. R. (2015). Evolutionary tipping points in the capacity to adapt to environmental change. In *PNAS*, *112*, (pp. 184-189).
- Bryan, T. L., & Metaxas, A. (2007). Predicting suitable habitat for deep-water gorgonian corals on the Atlantic and Pacific Continental Margins of North America. In *Marine Ecology Progress Series*, *330*, (pp.113–126).
- Calenge, C., Darmon, G., Basille, M., Loison, A., & Jullien, J.M. (2008). The factorial decomposition of the Mahalanobis distances in habitat selection studies. In *Ecology*, *89*, (pp.555–66).
- Cerrano, C., Bavestrello, G., Bianchi, C. N., Cattaneo-Vietti, R., Bava, S., Morganti, C., Morri, C., Picco, P., Sara, G., Schiaparelli, S., Siccardi, A. & Sponga, F. (2000). A catastrophic mass-mortality episode of gorgonians and other organisms in the Ligurian Sea (North-western Mediterranean), summer 1999. In *Ecology Letters*, *3*, (pp. 284-293).
- Chefaoui, R. M. & Lobo, J. M. (2007). Assessing the Conservation Status of an Iberian Moth Using Pseudo- Absences. In *The Journal of Wildlife Management*, *71*, (pp. 2507-2516).
- Cigliano, M. & Gambi, M. C. (2007). The long hot summer: a further mortality event of gorgonians along the Phlaegrean islands (Tyrrhenian Sea). *Biologia Marina Mediterranea*, *14*, (pp. 292-293).
- Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Lasram, F. B. R., Aguzzi, J., Ballesteros, E., Bianchi, C. N., Corbera, J., Dailianis, T., Danovaro, R., Estrada, M., Frogli, C., Galil, B. S., Gasol, J. M., Gertwagen, R., Gil, J., Guilhaumon, F., Kesner-Reyes, K. Kitsos, M. S., Koukouras, A., Lampadariou, N., Laxamana, E., de la Cuadra, C. M L. F., Lotze, H. K., Martin, D., Mouillot, D., Oro, D., Raicevich, S., Rius-Barile, J., Saiz-Salinas, J. I., Vicente, C. S., Somot, S., Templado, J., Turon, X., Vafidis, D. and Villanueva, R. (2010). The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats. In *PLoS ONE*, *5*, e.11842.
- Coma, R., Gili, J. M., Zabala, M. & Riera, T. (1994). Feeding and prey capture cycles in the aposymbiotic gorgonian *Paramuricea clavata*. In *Marine Ecology Progress Series*, *115*, (pp. 257-270).
- Coma, R., Ribes, M., Serrano, E., Jiménez, E., Salat, J. & Pascual, J. (2009). Global warming-enhanced stratification and mass mortality events in the Mediterranean. In *Proceedings of the National Academy of Sciences*, *106*, (pp. 6176-6181).
- Coma, R., Ribes, M., Zabala, M. & Gili, J. (1995). Reproduction and cycle of gonadal development in the Mediterranean gorgonian *Paramuricea clavata*. In *Marine Ecology Progress Series*, *117*, (pp. 173-183).
- Coma, R., Ribes, M., Zabala, M. & Gili, J. (1998). Growth in a Modular Colonial Marine Invertebrate. In *Estuarine, Coastal and Shelf Science*, *47*, (pp. 459-470).
- Coppo, S., Diviacco, G. & Tunesi, L. (2009). Environmental and conservation relevance of the Punta Manara coralligenous beds (Eastern Ligurian Sea). In *Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions*. Tabarka, Tunisia (pp. 76-82).
- Hipótese de exemplo para falar do futuro - Couce, E., Ridgwell, A., & Hendy, E. J. (2013). Future habitat suitability for coral reef ecosystems under global warming and ocean acidification. In *Global Change Biology*, *19*, (pp. 3592–3606).

- Crain, C. M. & Bertness, M. D. (2006). Ecosystem Engineering across Environmental Gradients : Implications for Conservation and Management. In *BioScience*, 56, (pp. 211-218).
- Crisci, C., Bensoussan, N., Romano, J. C., & Garrabou, J. (2011). Temperature anomalies and mortality events in marine communities: insights on factors behind differential mortality impacts in the NW Mediterranean. In *PloS one*, 6, e23814.
- Cupido, R., Cocito, S., Manno, V., Ferrando, S., Peirano, A., Iannelli, M., Bramanti, L. & Santangelo, G. (2012). Sexual structure of a highly reproductive, recovering gorgonian population: quantifying reproductive output. In *Marine Ecology Progress Series*, 469, (pp. 25-36).
- Cupido, R., Cocito, S., Sgorbini, S., Bordone, A., & Santangelo, G. (2008). Response of a gorgonian (*Paramuricea clavata*) population to mortality events: recovery or loss?. In *Aquatic Conservation: Marine and Freshwater Ecosystems*, 18, (pp. 984-992).
- Cúrdia, J. L. da S. (2012). Gorgonians of the South of Portugal: Biology, ecology and conservation. Ph.D. Thesis, University of Aveiro.
- Dame, R. F., Bushek, D., & Prins, T. C. (2001). Benthic suspension feeders as determinants of ecosystem structure and function in shallow coastal waters. In *Ecological comparisons of sedimentary shores*. Springer Berlin Heidelberg (pp. 11-37).
- Davies, A. J. & Guinotte, J. M. (2011). Global habitat suitability for framework-forming cold-water corals. In *PLoS ONE*, 6, e18483.
- Davies, A. J., Wisshak, M., Orr, J. C., & Murray Roberts, J. (2008). Predicting suitable habitat for the cold-water coral *Lophelia pertusa* (Scleractinia). In *Deep Sea Research Part I: Oceanographic Research Papers*, 55, (pp. 1048–1062).
- De'ath, G. (2007). Boosted trees for ecological modeling and prediction. *Ecology*, 88, 243–251.
- Dolan, M. F., Grehan, A. J., Guinan, J. C., & Brown, C. (2008). Modelling the local distribution of cold-water corals in relation to bathymetric variables: Adding spatial context to deep-sea video data. In *Deep Sea Research Part I: Oceanographic Research Papers*, 55, (pp. 1564-1579).
- Dormann, C. (2011). Modelling species' distributions. In *Modelling complex ecological dynamics: an introduction into ecological modelling*, (pp. 179-196).
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. In *Annual Review of Ecology, Evolution, and Systematics*, 40, (pp.677-697).
- Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC., O. J., Peterson, A. T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira, R., Schapire, R. E., Soberon, J., Williams, S., Wisz, M. S. & Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. In *Ecography*, 29, (pp. 129–151).
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. In *The Journal of Animal Ecology*, 77, (pp. 802–13).
- Farber, O., & Kadmon, R. (2003). Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. In *Ecological Modelling*, 160, (pp. 115–130).

- Fenchel, T. O. M., & Finlay, B. J. (2004). The ubiquity of small species: patterns of local and global diversity. In *Bioscience*, *54*, (pp. 777-784).
- Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. In *PLoS ONE*, *9*, e97122.
- Franklin, E. C., Jokiel, P. L., & Donahue, M. J. (2013). Predictive modeling of coral distribution and abundance in the Hawaiian Islands. In *Marine Ecology Progress Series*, *481*, (pp. 121–132).
- Friedman, J.H., Hastie, T. & Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting. In *Annals of Statistics*, *28*, (pp. 337–407).
- Gambi, M. C., Barbieri, F., Signorelli, S. & Saggiomo, V. (2010). Mortality events along the Campania coast (Tyrrhenian Sea) in summers 2008 and 2009 and relation to thermal conditions. In *Biologia Marina Mediterranea*, *17*, (pp. 126-127).
- Garrabou, J., Ballesteros, E., & Zabala, M. (2002). Structure and Dynamics of North-western Mediterranean Rocky Benthic Communities along a Depth Gradient. In *Estuarine, Coastal and Shelf Science*, *55*, (pp. 493–508).
- Garrabou, J., Coma, R., Bensoussan, N., Bally, M., Chevaldonné, P., Cigliano, M., Diaz, D., Harmelin, J. G., Gambi, M. C., Kersting, D. K., Ledoux, J. B., Lejeusne, C., Linares, C., Marschal, C., Pérez, T., Ribes, M., Romano, J. C., Serrano, E., Teixido, N., Torrents, O., Zabala, M., Zuberer, F. & Cerrano, C. (2009). Mass mortality in Northwestern Mediterranean rocky benthic communities: effects of the 2003 heat wave. In *Global Change Biology*, *15*, (pp. 1090-1103).
- Gasith, A. & Resh, V. H. (1999). Streams in Mediterranean climate regions: abiotic influences and biotic responses to predictable seasonal events. In *Annual review of ecology and systematics*, *30*, (pp. 51-81).
- Gili, J. M., & Coma, R. (1998). Benthic suspension feeders: their paramount role in littoral marine food webs. In *Trends in ecology & evolution*, *13*, (pp. 316-321).
- Giorgi, F., & Lionello, P. (2008). Climate change projections for the Mediterranean region. *Global and Planetary Change*, *63*, 90-104.
- Giusti, M., Canese, S., Angiolillo, M., Bo, M., Salvati, E., Cardinali, M. & Greco, S. (2009). Three-dimensional distribution of *Gerardia savaglia* in relation to depth, orientation and slope of the substrata in the south Tyrrhenian Sea. In Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia (pp. 96-100).
- Giusti, M., Innocenti, C., & Canese, S. (2014). Predicting suitable habitat for the gold coral *Savalia savaglia* (Bertoloni, 1819) (Cnidaria, Zoantharia) in the South Tyrrhenian Sea. In *Continental Shelf Research*, *81*, (pp. 19–28).
- Gori, A., Linares, C., Rossi, S., Coma, R., & Gili, J. M. (2007). Spatial variability in reproductive cycle of the gorgonians *Paramuricea clavata* and *Eunicella singularis* (Anthozoa, Octocorallia) in the Western Mediterranean Sea. In *Marine Biology*, *151*, (pp. 1571-1584).
- Gori, A., Linares, C., Viladrich, N., Clavero, A., Orejas, C., Fiorillo, I., Ambroso, S., Gili, J. M. & Rossi, S. (2013). Effects of food availability on the sexual reproduction and biochemical composition of

the Mediterranean gorgonian *Paramuricea clavata*. In *Journal of Experimental Marine Biology and Ecology*, 444, (pp. 38-45).

Gori, A., Rossi, S., Berganzo, E., Pretus, J. L., Dale, M. R., & Gili, J. M. (2011). Spatial distribution patterns of the gorgonians *Eunicella singularis*, *Paramuricea clavata*, and *Leptogorgia sarmentosa* (Cape of Creus, Northwestern Mediterranean Sea). In *Marine biology*, 158, (pp. 143-158).

Guinan, J., Brown, C., Dolan, M. F., & Grehan, A. J. (2009). Ecological niche modelling of the distribution of cold-water coral habitat using underwater remote sensing data. In *Ecological Informatics*, 4, (pp. 83-92).

Guinehut, S., Dhomps, A.L., Larnicol, G. & Le Traon, P.Y. (2012) High Resolution 3-D temperature and salinity fields derived from in situ and satellite observations. In *Ocean Science Discussions*, 9, (pp. 1313–1347).

Guinotte, J. M., & Davies, A. J. (2014). Predicted deep-sea coral habitat suitability for the U.S. West coast. *PloS One*, 9, e93918. Hirzel, A. H., & Le Lay, G. (2008). Habitat suitability modelling and niche theory. In *Journal of Applied Ecology*, 45, (pp. 1372–1381).

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

Guisan, A. & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. In *Ecological modelling*, 135, (pp. 147-186).

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

Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: the rear edge matters. In *Ecology letters*, 8, (pp. 461-467).

Hirzel, A. H., & Le Lay, G. (2008). Habitat suitability modelling and niche theory. In *Journal of Applied Ecology*, 45, (pp. 1372-1381).

Huete-Stauffer, C., Vielmini, I., Palma, M., Navone, A., Panzalis, P., Vezzulli, L., Misic, C. & Cerrano, C. (2011). *Paramuricea clavata* (Anthozoa, Octocorallia) loss in the Marine Protected Area of Tavolara (Sardinia, Italy) due to a mass mortality event. In *Marine Ecology*, 23, (pp. 107-116).

Hutchinson, G. E. (1957). The multivariate niche. In *Cold Spring Harbor Symposium on Quantitative Biology*, 22, (pp. 415-421).

Jiménez-Valverde, A., & Lobo, J. M. (2007). Threshold criteria for conversion of probability of species presence to either-or presence-absence. In *Acta Oecologica*, 31, (pp. 361–369).

Jones *et al.*, 1994 – Jones, C. G., Lawton, J. H., & Shachak, M. (1994). Organisms as ecosystem engineers. In *Oikos*, 69, (pp. 373–386).

Kearney, M. R., Wintle, B. A. & Porter, W. P. (2010). Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. In *Conservation Letters*, 3, (pp. 203-213).

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. In *Ecology letters*, 12, (pp. 334-350).

Kipson, S., Linares, C., Čížmek, H., Cebrián, E., Ballesteros, E., Bakran- Petricioli, T., & Garrabou, J. (2014). Population structure and conservation status of the red gorgonian

Kipson, S., Linares, C., Teixidó, N., Bakran-Petricioli, T. & Garrabou, J. (2012). Effects of thermal stress on early developmental stages of a gorgonian coral. In *Marine Ecology Progress Series*, 470, (pp. 69-78).

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *IJCAI*, 14, (pp. 1137-1145).

Leathwick, J. R., Elith, J., Francis, M. P., Hastie, T. & Taylor, P. (2006). Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. In *Marine Ecology Progress Series*, 321, (pp. 267–281).

Leichter, J. J., & Witman, J. D. (1997). Water flow over subtidal rock walls: relation to distributions and growth rates of sessile suspension feeders in the Gulf of Maine Water flow and growth rates. In *Journal of Experimental Marine Biology and Ecology*, 209, (pp. 293-307).

Lejeusne, C., Chevaldonné, P., Pergent-Martini, C., Boudouresque, C. F., & Perez, T. (2009). Climate change effects on a miniature ocean: the highly diverse, highly impacted Mediterranean Sea. In *Trends in ecology & evolution*, 25(4), (pp. 250-260).

Leverette, T. L., & Metaxas, A. (2005). Predicting habitat for two species of deep-water coral on the Canadian Atlantic continental shelf and slope. In A. Freiwald & J. M. Roberts (Eds.), *Cold-water Corals and Ecosystems*, (pp. 467–479).

Levitus, S., Antonov, J.I., Baranova, O.K., Boyer, T.P., Coleman, C.L., Garcia, H.E., Grodsky, A.I., Johnson, D.R., Locarnini, R. a, Mishonov, A.V., Reagan, J.R., Sazama, C.L., Seidov, D., Smolyar, I., Yarosh, E.S. & Zweng, M.M. (2013) The World Ocean Database. In *Data Science Journal*, 12, WDS229–WDS234.

Linares, C., Coma, R., Diaz, D., Zabala, M., Hereu, B. & Dantart, L. (2005). Immediate and delayed effects of a mass mortality event on gorgonian population dynamics and benthic community structure in the NW Mediterranean Sea. In *Marine Ecology Progress Series*, 305, (pp. 127-137).

Linares, C., Coma, R., Garrabou, J., Bianchimani, O., Drap, P., Serrano, E. & Zabala, M. (2009). Contribution to the conservation of coralligenous communities through studies on population ecology of mediterranean gorgonians. In Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia (pp. 107-112).

Linares, C., Coma, R., Garrabou, J., Díaz, D. & Zabala, M. (2008a). Size distribution, density and disturbance in two Mediterranean gorgonians: *Paramuricea clavata* and *Eunicella singularis*. In *Journal of Applied Ecology*, 45, (pp. 688-699).

Linares, C., Coma, R., Mariani, S., Díaz, D., Hereu, B., & Zabala, M. (2008b). Early life history of the Mediterranean gorgonian *Paramuricea clavata*: implications for population dynamics. In *Invertebrate Biology*, 127, (pp. 1-11).

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

Martin, C. S., Giannoulaki, M., De Leo, F., Scardi, M., Salomidi, M., Knitweiss, L., Pace, M. L., Garofalo, G., Gristina, M., Ballesteros, E., Bavestrello, G., Belluscio, A., Cebrian, E., Gerakaris, V., Pergent, G., Pergent-Martini, C., Schembri, P. J., Terribile, K., Rizzo, L., Ben Souissi, J., Bonacorsi, M., Guarnieri, G., Krzelj, M., Macic, V., Punzo, E., Valavanis, V. & Fraschetti, S. (2014). Coralligenous and

maërl habitats: predictive modelling to identify their spatial distributions across the Mediterranean Sea. In *Scientific Reports*, 4, (pp. 1–8).

Martin, Y., Bonnefont, J. L., & Chancerelle, L. (2002). Gorgonians mass mortality during the 1999 late summer in French Mediterranean coastal waters: the bacterial hypothesis. In *Water research*, 36, (pp.779-782).

Martin-Jézéquel, V., Hildebrand, M., & Brzezinski, M. A. (2000). Silicon metabolism in diatoms: implications for growth. In *Journal of Phycology*, 36(5), (pp. 821-840).

Millot, C. (1999). Circulation in the western Mediterranean Sea. In *Journal of Marine Systems*, 20, (pp. 423-442).

Mistri, M. (1994). Ecological observations on a population of the Mediterranean gorgonian *Paramuricea clavata* (Risso, 1826). In *Bolletino di zoologia*, 61(02), (pp. 163-166).

Mistri, M., & Ceccherelli, V. U. (1994). Growth and secondary production of the Mediterranean gorgonian *Paramuricea clavata*. In *Marine Ecology-Progress Series*, 103, (pp. 291-291).

Moisen, G.G., Freeman, E.A., Blackard, J.A., Frescino, T.S., Zimmermann, N.E. & Edwards, T.C. (2006). Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. In *Ecological Modelling*, 199, (pp. 176–187).

Mokhtar-Jamaï, K., Pascual, M., Ledoux, J.B., Coma, R., Féral, J.P. & Garrabou J. (2011). From global to local genetic structuring in the red gorgonian *Paramuricea clavata*: the interplay between oceanographic conditions and limited larval dispersal. In *Molecular Ecology*, 20, (pp. 3291-3305).

Montagna, P., McCulloch, M., Taviani, M., Mazzoli, C., & Vendrell, B. (2006). Phosphorus in cold-water corals as a proxy for seawater nutrient chemistry. In *Science*, 312, (pp. 1788-1791).

Neiva, J., Assis, J., Fernandes, F., Pearson *et al.*, G. A. & Serrão, E. A. (2014). Species distribution models and mitochondrial DNA phylogeography suggest an extensive biogeographical shift in the high-intertidal seaweed *Pelvetia canaliculata*. In *Journal of Biogeography*, 41(6), (pp. 1137-1148).

OBIS (2013). Global biodiversity indices from the Ocean Biogeographic Information System. Intergovernmental Oceanographic Commission of UNESCO. Web. <http://www.iobis.org> (consulted on 2014/03/10)

OCEANA (2014). The seamounts of the Gorringer Bank.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. In *Annual Review of Ecology, Evolution and Systematics*, 37, (pp. 643-669).

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

Peterson, A. T., & Vieglais, D. A. (2001). Predicting Species Invasions Using Ecological Niche Modeling: New Approaches from Bioinformatics Attack a Pressing Problem A new approach to ecological niche modeling, based on new tools drawn from biodiversity informatics, is applied to the challenge of predicting potential species' invasions. In *BioScience*, 51, (pp. 363-371).

Philippart, C. J., Anadón, R., Danovaro, R., Dippner, J. W., Drinkwater, K. F., Hawkins, S. J., Oguz, T., O'Sullivan, G. & Reid, P. C. (2011). Impacts of climate change on European marine

ecosystems: observations, expectations and indicators. In *Journal of Experimental Marine Biology and Ecology*, 400(1), (pp. 52-69).

Raitsos, D. E., Beaugrand, G., Georgopoulos, D., Zenetos, A., Pancucci-Papadopoulou, A. M., Theocharis, A., & Papathanassiou, E. (2010). Global climate change amplifies the entry of tropical species into the Eastern Mediterranean Sea. In *Limnology and Oceanography*, 55(4), (pp. 1478-1484).

Recknagel, F. (2001). Applications of machine learning to ecological modelling. In *Ecological Modelling*, 146, (pp. 303-310).

Rengstorf, A. M., Grehan, A., Yesson, C., & Brown, C. (2012). Towards High-Resolution Habitat Suitability Modeling of Vulnerable Marine Ecosystems in the Deep-Sea: Resolving Terrain Attribute Dependencies. In *Marine Geodesy*, 35, (pp. 343–361).

Ribes, M., Coma, R., & Gili, J.-M. (1999). Heterogeneous feeding in benthic suspension feeders: the natural diet and grazing rate of the temperate gorgonian *Paramuricea clavata* (Cnidaria: Octocorallia) over a year cycle. In *Marine Ecology Progress Series*, 183, (pp. 125–137).

Rivetti, I., Frascchetti, S., Lionello, P., Zambianchi, E., & Boero, F. (2014). Global Warming and Mass Mortalities of Benthic Invertebrates in the Mediterranean Sea. In *PloS one*, 9(12), (pp. 1-22).

Rodrigues SC. Dados ecológicos de gorgónias (Octocorallia: Alcyonacea) – contributo para a conservação e gestão de actividades subaquáticas no Parque Marinho Professor Luiz Saldanha (Portugal). M.Sc. Thesis, University of Lisbon. 2008. Available: http://www.cpas.pt/ecologia/publicacoes/RODRIGUES_2008_Dados_ecol_Gorg_Gestao_Conserv_PML_S.pdf

Rossi, S., Gili, J. M., Coma, R., Linares, C., Gori, A., & Vert, N. (2005). Temporal variation in protein, carbohydrate, and lipid concentrations in *Paramuricea clavata* (Anthozoa, Octocorallia): evidence for summer–autumn feeding constraints. In *Marine Biology*, 149, (pp. 643-651).

Rouanet, E., Lelong, P., Lecalard, C. & Martin, Y. (2009). Estimation of coralligenous assemblages' ecological state around Embiez archipelago. In Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia (pp. 239-241).

Salomidi, M., Smith, C., Katsanevakis, S., Panayotidis, P., & Papathanassiou, V. (2009). Some observations on the structure and distribution of gorgonian assemblages in the eastern Mediterranean Sea. In Proceedings of the 1st Mediterranean symposium on the conservation of the coralligenous and other calcareous bio-concretions. Tabarka, Tunisia (pp. 242-245).

Sánchez, E., Gallardo, C., Gaertner, M. A., Arribas, A., & Castro, M. (2004). Future climate extreme events in the Mediterranean simulated by a regional climate model: a first approach. In *Global and Planetary Change*, 44(1), (pp. 163-180).

Schapire, R. (2001) The boosting approach to machine learning – an overview. MSRI Workshop on Nonlinear Estimation and Classification, 2002 (eds D.D. Denison, M. H. Hansen, C. Holmes, B. Mallick & B. Yu). Springer, New York. Elith, J., Leathwick J. R. and Hastie, T. (2008). "A working guide to boosted regression trees." *Journal of Animal Ecology* 77(4): 802-813.

Schoener, T. W. (1974). Resource Partitioning in Ecological Communities. In *Science (New York, N.Y.)*, 185, (pp. 27-39).

- Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. In *Annual Review of Ecology, Evolution, and Systematics*, 40, (pp. 415-436).
- Tittensor, D. P., Baco, A. R., Brewin, P. E., Clark, M. R., Consalvey, M., Hall-Spencer, J., Rowden, A. A., Schlacher, T., Stocks, K. I. & Rogers, A. D. (2009). Predicting global habitat suitability for stony corals on seamounts. In *Journal of Biogeography*, 36, (pp. 1111–1128).
- Tong, R., Purser, A., Guinan, J., & Unnithan, V. (2012). Modeling the habitat suitability for deep-water gorgonian corals based on terrain variables. In *Ecological Informatics*, 13, (pp. 123–132).
- Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., & Kadmon, R. (2007). A comparative evaluation of presence-only methods for modelling species distribution. In *Diversity and distributions*, 13, (pp. 397-405).
- UNEP/MAP, United Nations Environment Programme - Mediterranean Action Plan (2015). Sicily Channel/Tunisian Plateau: Topography, circulation and their effects on biological component.
- Vezzulli, L., Previati, M., Pruzzo, C., Marchese, A., Bourne, D. G., Cerrano, C. & the VibrioSea Consortium (2010). Vibrio infections triggering mass mortality events in a warming Mediterranean Sea. In *Environmental microbiology*, 12, 2007-2019.
- Vierod, A. D. T., Guinotte, J. M., & Davies, A. J. (2014). Predicting the distribution of vulnerable marine ecosystems in the deep sea using presence-background models. In *Deep Sea Research Part II: Topical Studies in Oceanography*, 99, (pp. 6–18).
- Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J., Fromentin, J. M., Hoegh-Guldberg, O. & Bairlein, F. (2002). Ecological responses to recent climate change. In *Nature*, 416(6879), (pp. 389-395).
- Williamson, E. A., Strychar, K. B., Withers, K., & Sterba-Boatwright, B. (2011). Effects of salinity and sedimentation on the Gorgonian Coral, *Leptogorgia virgulata* (Lamarck 1815). In *Journal of Experimental Marine Biology and Ecology*, 409, (pp. 331-338).
- Wilson, M. F., O'Connell, B., Brown, C., Guinan, J. C., & Grehan, A. J. (2007). Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. In *Marine Geodesy*, 30, (pp. 3-35).
- Woolbright, S. A., Whitham, T. G., Gehring, C. A., Allan, G. J. & Bailey, J. K. (2014). Climate relicts and their associated communities as natural ecology and evolution laboratories. In *Trends in Ecology & Evolution*, 29, (pp. 406-416).
- Yesson, C., Taylor, M. L., Tittensor, D. P., Davies, A. J., Guinotte, J. M., Baco, A., Black, J., Hall-Spencer, J. M. & Rogers, A. D. (2012). Global habitat suitability of cold-water octocorals. In *Journal of Biogeography*, 39, (pp. 1278–1292).

7 APPENDIX:

APPENDIX 1

Table A.1.1 Location of record points and source. Depth is provided when available. LON – longitude; LAT – latitude. Reference ID provided in Table A.2.

REFERENCE ID	LON	LAT	LOCATION	DEPTH (m)	INCLUDED IN MODEL (1=YES)
Atlantic					
2	-9,546	39,472	Berlengas, Portugal	15	1
2	-9,253	38,419	Cape Espichel, Portugal	39	1
2	-9,252	38,419	Cape Espichel, Portugal	45	0
2	-9,253	38,408	Cape Espichel, Portugal	24	0
2	-9,226	38,407	Cape Espichel, Portugal	44	0
2	-8,926	37,898	Sines, Portugal	50	1
2	-8,919	36,972	Sagres, Portugal	59	0
2	-8,920	36,971	Sagres, Portugal	59	0
2	-8,922	36,967	Sagres, Portugal	60	1
2	-8,704	37,004	Lagos, Portugal	61	1
2	-8,492	37,063	Portimao, Portugal	37	1
2	-8,505	36,999	Portimao, Portugal	50	0
2	-8,599	36,996	Portimao, Portugal	60	0
2	-8,500	36,992	Portimao, Portugal	49	0
2	-8,489	36,991	Portimao, Portugal	48	0
2	-8,528	36,988	Portimao, Portugal	55	0
2	-7,635	37,057	Tavira, Portugal	30	1
16	-11,596	36,368	Gorringe, Portugal	40	1
Western Mediterranean					
5	-5,313	35,752	Marinasmir, Morocco	40	0
5	-5,248	35,642	Martyl, Morocco	35	0
5	-5,300	35,918	Monte Hacho, Ceuta, Spain	26	1
20	-3,218	35,831	Alboran, Spain	100	1
20	-3,049	35,893	Alboran, Spain	46	0
20	-2,982	35,959	Alboran, Spain	50	1
20	-0,653	37,652	Cabo de Palos, Spain	-	1
20	-0,029	37,604	Alboran, Spain	100	1
6	0,667	39,917	Columbretes Archipelago, Spain	-	1
5	0,670	39,892	Bajo del Carallot, Columbretes Island, Spain	40	0
18	1,460	41,210	Barcelona, Spain	-	1
5	1,670	38,970	Bledes (Na Gorra), Ibiza Island, Spain	37	1
18	2,700	39,520	Maiorca, Spain	-	1
5	2,960	39,124	L'Imperial, Cabrera Island, , Spain	40	1

4	3,217	42,200	Medes Islands (MPA), Spain	28	0
5	3,219	42,233	Punta Falconera, Cap de Creus, Spain	25	1
18	3,220	42,040	Catalan sea, Spain	-	1
5	5,239	43,331	Tombant Moulon Est, Cote Bleue, France	20	0
5	5,338	43,207	Pharillons, Marseille, France	20-40	1
5	5,390	43,173	Riou Sud, Marseille, France	20-40	0
5	5,391	43,187	Grotte Peres, Marseille, France	10-20	0
5	5,393	43,173	Imperiales de Terre, Marseille, France	8	0
5	5,396	43,179	Petit Congloue, Marseille, France	10-20	0
5	5,402	43,176	Grand Congloue, Marseille, France	7	0
5	5,452	43,201	Morgiou, Marseille, France	30	0
5	5,499	43,198	Castelvieil, Marseille, France	10	0
			Plate aux Meros, Embiez Archipelago,		
9	5,750	43,080	France	-	0
9	5,751	43,074	Merveilleuse, Embiez Archipelago, France	-	0
9	5,766	43,073	Athena, Embiez Archipelago, France	-	0
			Pierre à Christian, Embiez Archipelago,		
9	5,774	43,071	France	-	1
5	6,363	43,019	Montremian, Port-Cros, France	20	0
5	6,397	42,989	Gabinieri, Port-Cros, France	22	0
14	6,398	42,987	Gabinieri, Port Cros National Park, France	-	1
5	6,407	42,995	Pointe du Vaisseau, Port-Cros, France	20	0
18	6,630	43,270	Saint-Tropez, France	-	1
5	8,537	42,373	Garganellu, North Corsica, France	20	0
5	8,546	42,380	Palazzu, North Corsica, France	28	0
5	8,550	42,380	Palazzinu, North Corsica, France	25	0
5	8,551	42,350	Baja Casju, North Corsica, France	25	0
5	8,554	42,332	Punta Muchillina, North Corsica, France	20	1
18	8,560	43,350	Corsica, France	-	0
1	8,625	41,718	Corsica, France	-	1
3	8,649	41,745	Corsica, France	-	0
17	9,260	41,350	Lavezzi islands, Corsica, France	-	1
18	8,910	44,390	Liguria, Italy	-	1
5	9,196	44,306	Altare, Portofino, Italy	25	0
11	9,405	44,251	Punta Manara, Italy	56	1
12	9,870	44,069	Gulf of La Spezia, Italy	21	0
15	9,510	44,013	Tinetto Island, Italy	21	0
15	9,505	44,012	Tinetto Shoal, Italy	21	1
18	9,710	40,900	Sardinia, Italy	-	1
18	14,290	40,830	Tyrrhenian sea, Italy	-	1
5	13,894	40,692	Sant'Angelo, Ischia Island, Italy	32	1
10	16,118	38,749	Calabria, Italy	75	0
10	16,124	38,745	Calabria, Italy	75	1
10	15,829	38,363	Calabria, Italy	50	0

10	15,823	38,328	Calabria, Italy	50	0
10	15,819	38,321	Calabria, Italy	50	1
13	15,657	38,267	Punta Faro, Messina, Sicily, Italy	55	1
Eastern Mediterranean					
19	14,837	44,887	Prvic, Croatia	23-40	0
19	14,493	44,725	Cutin Mali, Croatia	25-38	0
19	14,734	44,704	Lun, Croatia	25-40	0
5	15,111	44,018	Fulija Island, Dugi Otok, Croatia	40	1
19	15,220	43,836	Obrucan, Croatia	25-55	0
19	15,255	43,804	Balun, Croatia	27-60	0
19	15,266	43,800	Mana, Croatia	27-60	0
19	15,919	43,513	Mulo, Croatia	35-50+	0
19	15,942	43,511	Smokvica, Croatia	30-50+	0
19	15,969	43,493	Planka, Croatia	32-50+	0
very					
21	24,831	34,753	South Crete, Greece	deep	1
7	23,573	35,657	Gramvousa island, Greece	95	1
21	26,489	37,576	Fourni Island, Greece	-	1
7	24,613	38,106	S. Evia, Greece	61	1
21	22,505	38,212	Korinthiakos Gulf, Greece	-	1
21	23,597	38,463	Evia Island, Greece	-	1
7	23,898	38,696	E. Evia, Greece	63	1
3	26,551	38,949	Lesbos Island, Greece	-	0
1	26,590	38,960	Lesbos Island, Greece	-	0
8	26,534	38,969	Lesvos Island, Greece	-	1
1	26,510	39,050	Lesbos Island, Greece	-	0
7	24,108	39,075	N. Sporades, Greece	50	1
1	23,850	39,130	Alonissos Island, Greece	-	0
3	23,950	39,147	Peristera Island, Greece	-	0
1	23,982	39,193	Peristera Island, Greece	-	0
8	26,443	39,306	Lesvos Island, Greece	-	1
8	26,147	39,337	Lesvos Island, Greece	-	0
21	24,135	39,345	Alonissos Island, Northern Sporades, Greece	-	0
7	26,437	39,358	N. Lesvos, Greece	48	0
7	24,235	39,383	N. Sporades, Greece	60	1
1	24,840	39,520	Agios Efstratios, Greece	-	0
3	24,848	39,548	Agios Efstratios, Greece	-	1
1	24,300	40,110	Agio Oros, Greece	-	0
1	24,290	40,120	Agio Oros, Greece	-	0
3	24,252	40,147	Agio Oros, Greece	-	0
3	24,241	40,150	Agio Oros, Greece	-	1
5	26,577	39,370	Ayvalik, Ezerbey Sigiligi, Turkey	24	0

Table A.1.2 References ID of the records used to model the niche of *Paramuricea clavata*.

	REFERENCE ID (dataset)
1	GBIF Data Portal, www.gbif.net . 2014-03-10. Senckenberg: Collection Cnidaria SMF, 207950808.
2	This study.
3	OBIS (2013). Global biodiversity indices from the Ocean Biogeographic Information System. Intergovernmental Oceanographic Commission of UNESCO. Web. http://www.iobis.org (consulted on 2014/03/10)
4	Linares, C., Coma, R., Mariani, S., Díaz, D., Hereu, B. & Zabala, M. (2008). Early life history of the Mediterranean gorgonian <i>Paramuricea clavata</i> : implications for population dynamics. In <i>Invertebrate Biology</i> , 127, (pp.1-11).
5	Mokhtar-Jamäi, K., Pascual, M., Ledoux, J.B., Coma, R., Féral, J.P. & Garrabou J. (2011). From global to local genetic structuring in the red gorgonian <i>Paramuricea clavata</i> : the interplay between oceanographic conditions and limited larval dispersal. In <i>Molecular Ecology</i> , 20, (pp. 3291-3305).
6	Agell, G., Rius, M., & Pascual, M. (2009). Isolation and characterization of eight polymorphic microsatellite loci for the Mediterranean gorgonian <i>Paramuricea clavata</i> . In <i>Conservation genetics</i> , 10(6), (pp. 2025-2027).
7	Salomidi, M., Smith, C., Katsanevakis, S., Panayotidis, P., & Papathanassiou, V. (2009). Some observations on the structure and distribution of gorgonian assemblages in the eastern Mediterranean Sea. In <i>Proceedings of the 1st Mediterranean symposium on the conservation of the coralligenous and other calcareous bio-concretions. Tabarka, Tunisia</i> (pp. 242-245).
8	Gerovasileiou, V., Sini, M. I., Poursanidis, D., & Koutsoubas, D. (2009). Contribution to the knowledge of Coralligenous communities in the NE Aegean Sea. In <i>Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia</i> (pp. 14-16).
9	Rouanet, E., Lelong, P., Lecalard, C. & Martin, Y. (2009). Estimation of coralligenous assemblages' ecological state around Embiez archipelago. In <i>Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia</i> (pp. 239-241).
10	Angiolillo, M., Canese, S., Salvati, E., Giusti, M., Cardinali, M., Bo, M. & Greco, S. (2009). Presence of corallium rubrum on coralligenous assemblages below 50m along Calabrian coast (South Italy). In <i>Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia</i> (pp. 47-52).
11	Coppo, S., Diviacco, G. & Tunesi, L. (2009). Environmental and conservation relevance of the Punta Manara coralligenous beds (Eastern Ligurian Sea). In <i>Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia</i> (pp. 76-82).
12	Cupido, R., Cocito, S., Peirano, A. and Santangelo, G. (2009). Mortality and Resilience: A ten-year monitoring of gorgonian population trends in the eastern Ligurian Sea. In <i>Proceedings of</i>

- the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia* (pp. 83-88).
- 13 Giusti, M., Canese, S., Angiolillo, M., Bo, M., Salvati, E., Cardinali, M. & Greco, S. (2009). Three-dimensional distribution of *Gerardia savaglia* in relation to depth, orientation and slope of the substrata in the south Tyrrhenian Sea. In *Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia* (pp. 96-100).
 - 14 Linares, C., Coma, R., Garrabou, J., Bianchimani, O., Drap, P., Serrano, E. & Zabala, M. (2009). Contribution to the conservation of coralligenous communities through studies on population ecology of mediterranean gorgonians. In *Proceedings of the 1st Mediterranean Symposium on the Coralligenous Conservation and other calcareous bio-concretions. Tabarka, Tunisia* (pp. 107-112).
 - 15 Cupido, R., Cocito, S., Manno, V., Ferrando, S., Peirano, A., Iannelli, M., Bramanti L. & Santangelo G. (2012). Sexual structure of a highly reproductive, recovering gorgonian population: quantifying reproductive output. In *Marine Ecology Progress Series, 469*, (pp. 25-36).
 - 16 OCEANA (2014). The seamounts of the Gorringe Bank.
 - 17 Mistri, M. (1994). Ecological observations on a population of the Mediterranean gorgonian *Paramuricea clavata* (Risso, 1826). In *Bolletino di zoologia, 61(02)*, (pp. 163-166).
 - 18 Rivetti, I., Frascetti, S., Lionello, P., Zambianchi, E., & Boero, F. (2014). Global Warming and Mass Mortalities of Benthic Invertebrates in the Mediterranean Sea. In *PloS one, 9(12)*, (pp. 1-22).
 - 19 Kipson, S., Linares, C., Čížmek, H., Cebrián, E., Ballesteros, E., Bakran- Petricioli, T., & Garrabou, J. (2014). Population structure and conservation status of the red gorgonian *Paramuricea clavata* (Risso, 1826) in the Eastern Adriatic Sea. In *Marine Ecology*, (pp. 1-12).
 - 20 Branco, M. (2014). Title. B.Sc. Thesis, University of Algarve, Portugal. Pp.
 - 21 Vasilis Gerovasileiou and Maria Sini personal communication; Hellenic Center for Marine Research, Institute of Marine Biology, Biotechnology and Aquaculture, Greece
-

APPENDIX 2.

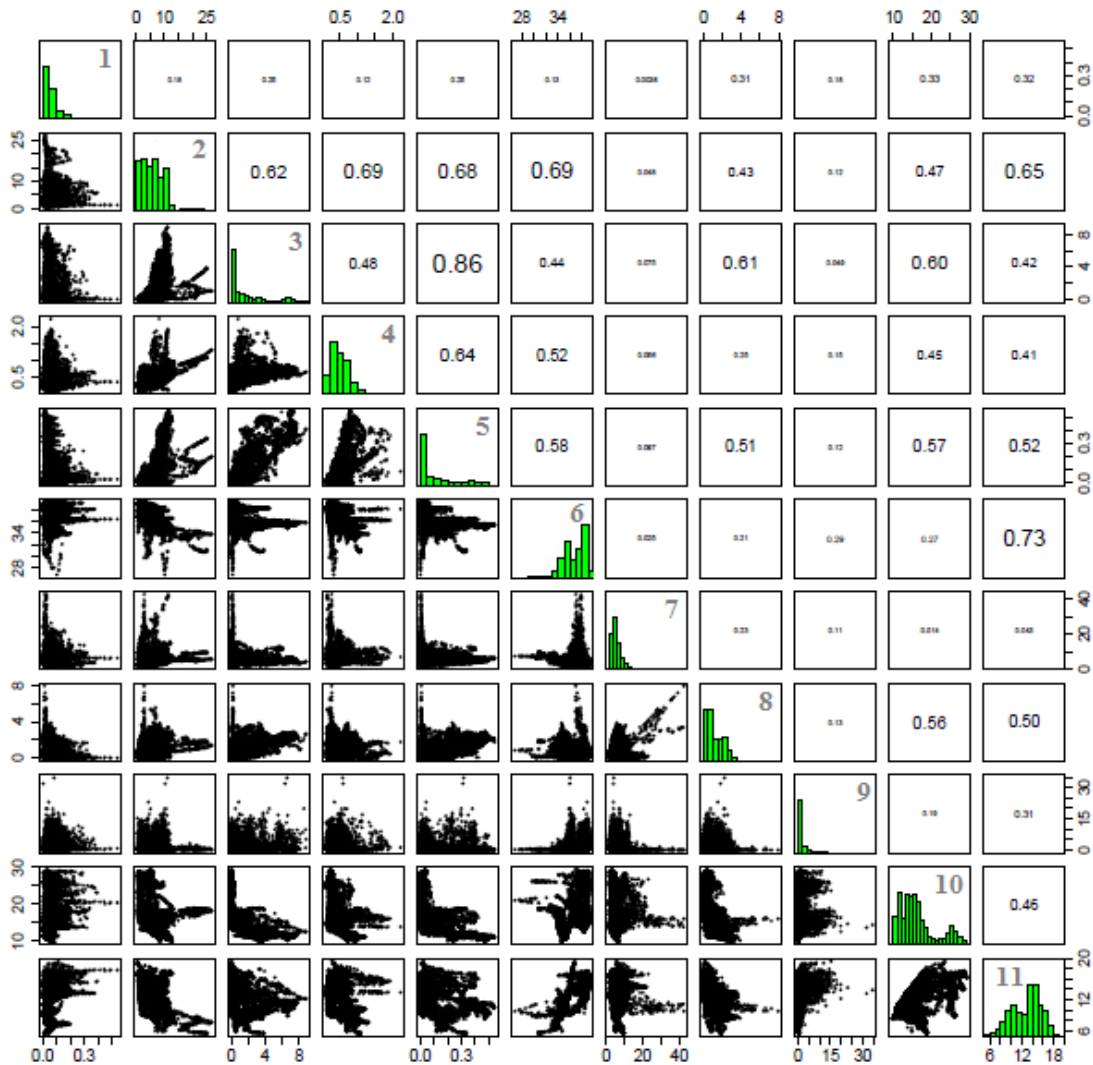


Figura A.2.1 Correlation matrix of the environmental predictors chosen to describe the suitable habitat of *Paramuricea clavata*. From top to bottom: currents minimum (currents_min)(1), nitrate maximum (2) and minimum (3) (nitrate_max/min), phosphate maximum (4) and minimum (5) (phosphate_max/min), salinity minimum (6) (salinity_min), silicate maximum (7) and minimum (8) (silicate_max/min), slope (9) and temperature maximum (10) and minimum (11) (temperature_max/min). This matrix represents the pairwise Pearson *et al.* correlation coefficient between all predictors, $r \leq 0.4$ indicates a weak correlation between variables and $r \geq 0.8$ corresponds to a strong correlation.

APPENDIX 3.

```
##Script developed for the prediction of the suitable habitat of Paramuricea clavata, 2015
## Silva 2015
## -----
##
## DEPENDENCES

library(raster)
library(rgdal)
library(adehabitat)
library(SDMTools)
library(sp)
library(dismo)
library(parallel)
library(doParallel)
library(gbm)

setwd("C:/Users/Inga/Desktop/Thesis/Final Version/Data treatment/Model_Proced_Out2015")

## -----
## Pairs Plot
## -----

rasters <- list.files("Data/Environment", pattern="tif", full.names = TRUE)#caminho para os rasters
rasters <- stack(rasters)
names(rasters) <- c("curr",
  "nit_max",
  "nit_min",
  "phos_max",
  "phos_min",
  "sal",
  "sil_max",
  "sil_min",
  "slope",
  "t_max",
  "t_min")#atribuir nomes aos rasters

pairs(rasters)#matriz de correlacao das variaveis

## -----
## Managing occurrences and producing Pseudo-absences
## -----

model.cells <- raster("Data/Model cells Present.tif")#abrir
model.cells.r <- as.data.frame(model.cells, xy=TRUE, centroids=TRUE)
model.cells.r <- model.cells.r[!is.na(model.cells.r[,3]),]

domain <- "crop.north"

occurrences <- read.table("Data/Occurrences/pontos_paramuricea_14-jan-2015_croacia.txt", sep="\t",
header = TRUE)
occurrences$IncludedInModel[is.na(occurrences$IncludedInModel)] <- 0
occurrences <- cbind(occurrences$LON[occurrences$IncludedInModel ==
1],occurrences$LAT[occurrences$IncludedInModel == 1])
```

```

plot(occurrences)

## -----
##Preparing distribution data

occurrences.final <- data.frame()

for (i in 1:nrow(occurrences)) {
  near.cells <- as.data.frame( sort( spDistsN1( as.matrix(model.cells.r[,1:2]),
      as.matrix(occurrences[i,1:2]),longlat=TRUE),
      decreasing = FALSE,index.return = TRUE))
  occurrences.final <- rbind(occurrences.final,c(model.cells.r[near.cells[1,2],1],
      model.cells.r[near.cells[1,2],2],1))
}

occurrences.r <- rasterize(occurrences.final[,1:2], model.cells)
occurrences.r.points <- as.data.frame(occurrences.r, , xy=TRUE, centroids=TRUE)
occurrences.r.points <- occurrences.r.points[ which(!is.na(occurrences.r.points[,3])) ,1:2 ]

plot(occurrences.r.points)

# -----
##Preparing environmental data

raster.stack <- rasters
names <- names(raster.stack)

if(domain == "crop.north") {
  xmin <- -31.5 ; xmax <- 36.5 ; ymin <- 27.5 ; ymax <- 46
  resolution <- 0.1
  lons <- seq(from=xmin, to=xmax, by=resolution)
  lats <- seq(from=ymin, to=ymax, by=resolution)
  shape <- raster(ncol=length(lons)-1,nrow=length(lats)-1)
  extent(shape) <- extent(xmin,xmax,ymin,ymax)
  raster.stack <- crop(raster.stack,shape)
}

# -----
# transform rasters kasc to asc format

for (i in 1:length(names(raster.stack))) { assign( names(subset(raster.stack,i)) ,
      asc.from.raster(subset(raster.stack,i)) ) }

predictors <- as.kasc(list( #nit_max=nit_max,
  #nit_min=nit_min,
  #phos_max=phos_max,
  #phos_min=phos_min,
  #sil_max=sil_max,
  #sil_min=sil_min,
  curr=curr,
  sal=sal,
  slope=slope,
  t_max=t_max,
  t_min=t_min ))

# -----

```

Mahalanobis distance

```
df <- kasc2df(predictors)
pc <- dudi.pca(df$tab, scannf=FALSE)
tab <- pc$tab
ka <- df2kasc(tab, df$index, predictors)
hsm <- mahasuhab(ka, occurrences.r.points, type = "probability")

predicted <- raster.from.asc(hsm)
plot(predicted)

points(occurrences.r.points)
projection(predicted) <- CRS("+proj=longlat +datum=WGS84")
writeRaster(predicted,filename=paste("Results/Mahalanob_final_Reduced_",
                                   "out2015",sep=""), format="GTiff",overwrite=T)
```

#Obtaining pseudo.absences

```
predicted <- as.data.frame(predicted, xy=TRUE, centroids=TRUE)
pseudo_absences <- predicted[which(predicted[,3] <= 0.2),]
colnames(pseudo_absences) <-c("Lon", "Lat", "PA")
pseudo_absences[,3] <- 0

occurrences.r.points.f <- cbind(occurrences.r.points,rep(1,nrow(occurrences.r.points)))
colnames(occurrences.r.points.f) <-c("Lon", "Lat", "PA")
final.data <- rbind(occurrences.r.points.f,pseudo_absences)

if (domain == "crop.north") { final.data <- final.data[final.data[,2] <= ymax,]
}

plot(final.data[which(final.data[,3] == 0),1:2], col="black")
points(final.data[which(final.data[,3] == 1),1:2], col="red")

write.table(final.data, file = paste("Data/Occurrences/presence_absences_out2015.txt",sep=""),
           sep = ",", row.names = FALSE, col.names = TRUE, na = "NA", dec = ".")
```

Produce Full Model with Best TC and LR

```
n.predictors <- length(names(raster.stack))

final.data.occurrences <- read.table("Data/Occurrences/presence_absences_out2015.txt",
                                   header = TRUE, sep = ",", dec = ".")

zeros <- which(final.data.occurrences[,3] == 0)
ones <- which(final.data.occurrences[,3] == 1)

zeros <- sample(zeros, length(ones) * 10, replace=FALSE)
final.data.occurrences <- final.data.occurrences[c(ones,zeros),]

plot(final.data.occurrences[which(final.data.occurrences[,3] == 0),1:2])
points(final.data.occurrences[which(final.data.occurrences[,3] == 1),1:2],col="red")
```

Produce dataset for modeling

```
dataset.brt <- data.frame( PA = final.data.occurrences[,3] ,  
                          extract( raster.stack, final.data.occurrences[,1:2], method="simple"))
```

```
# -----
```

Best fit model

```
predictions.brt <- function(set.number){
```

```
  assign( paste("learning.rate",set.number,sep=""), parameters.comb[set.number,1] )  
  assign( paste("tree.complex",set.number,sep=""), parameters.comb[set.number,2] )  
  assign(paste("model",set.number,sep=""),
```

```
    gbm.step( data=dataset.brt,  
              gbm.x = variables,  
              gbm.y = 1,  
              family = "bernoulli",  
              tree.complexity = get(paste("tree.complex",set.number,sep="")),  
              learning.rate = get(paste("learning.rate",set.number,sep="")),  
              bag.fraction = 0.5,  
              var.monotone = monotonic.response,  
              verbose=FALSE)
```

```
)
```

```
vector.results <- data.frame( as.numeric(get(paste("model",set.number,sep=""))$gbm.call$learning.rate),  
                             as.numeric(get(paste("model",set.number,sep=""))$gbm.call$tree.complexity),  
                             as.numeric(get(paste("model",set.number,sep=""))$cv.statistics$deviance.mean) )
```

```
  return(as.data.frame(vector.results))  
}
```

```
# -----
```

1a - Full Model

```
variables <- 2:(n.predictors+1)
```

```
monotonic.response <- c(0,0,0,0,0,+1,0,0,0,-1,+1) # UNUSED c(0,-1,1,-1,1,1,-1,1,0,-1,1)
```

```
# -----
```

```
lr <- c(0.01 , 0.005 , 0.001 , 0.0005)
```

```
tc <- 1:length(variables)
```

```
parameters.comb <- expand.grid(lr=lr,tc=tc)
```

```
sets <- 1:nrow(parameters.comb)
```

```
cl <- makeCluster(2) ; registerDoParallel(cl)
```

```
parameters <- foreach(k=sets, .verbose=F, .packages=c("dismo","rgdal","raster","SDMTTools")) %dopar%  
{ predictions.brt(k) }
```

```
stopCluster(cl)
```

```
parameters <- do.call(rbind.data.frame, parameters)
```

```
best.paramet <- parameters[which.min(parameters[,3]),1:2]
```

```
best.model <- gbm.step( data=dataset.brt,  
                       gbm.x = variables,  
                       gbm.y = 1,  
                       family = "bernoulli",
```

```

        tree.complexity = best.paramet[,2],
        learning.rate = best.paramet[,1],
        bag.fraction = 0.5,
        var.monotone = monotonic.response,
        verbose=FALSE)

system("say Best model")

model.simp <- gbm.simplify(best.model, n.drops = 10)

# -----
##Second round -> Reduced model

# -----
# Best fit model

predictions.brt <- function(set.number){

  assign( paste("learning.rate",set.number,sep=""), parameters.comb[set.number,1] )
  assign( paste("tree.complex",set.number,sep=""), parameters.comb[set.number,2] )
  assign(paste("model",set.number,sep=""),

    gbm.step( data=dataset.brt,
              gbm.x = variables,
              gbm.y = 1,
              family = "bernoulli",
              tree.complexity = get(paste("tree.complex",set.number,sep="")),
              learning.rate = get(paste("learning.rate",set.number,sep="")),
              bag.fraction = 0.5,
              var.monotone = monotonic.response,
              verbose=FALSE)
    )

  vector.results <- data.frame( as.numeric(get(paste("model",set.number,sep=""))$gbm.call$learning.rate),
                              as.numeric(get(paste("model",set.number,sep=""))$gbm.call$tree.complexity),
                              as.numeric(get(paste("model",set.number,sep=""))$cv.statistics$deviance.mean) )

  return(as.data.frame(vector.results))

}

# -----
# 2a - Reduced Model
how.many.to.drop <- 4
variables <- model.simp$pred.list[[how.many.to.drop]]
monotonic.response <- monotonic.response[variables-1]

# Final variables reduced model
colnames(dataset.brt)[variables]

# -----

lr <- c(0.01 , 0.005 , 0.001 , 0.0005)
tc <- 1:length(variables)
parameters.comb <- expand.grid(lr=lr,tc=tc)
sets <- 1:nrow(parameters.comb)

```

```

cl <- makeCluster(2) ; registerDoParallel(cl)
parameters <- foreach(k=sets, .verbose=F, .packages=c("dismo", "rgdal", "raster", "SDMTools")) %dopar%
{ predictions.brt(k) }
stopCluster(cl)

parameters <- do.call(rbind.data.frame, parameters)
best.paramet <- parameters[which.min(parameters[,3]),1:2]

best.model <- gbm.step( data=dataset.brt,
  gbm.x = variables,
  gbm.y = 1,
  family = "bernoulli",
  tree.complexity = best.paramet[,2],
  learning.rate = best.paramet[,1],
  bag.fraction = 0.5,
  var.monotone = monotonic.response,
  verbose=FALSE)

system("say Best model")

# -----
# Sumario de estadísticas (Deviance explained, etc.)

summary(best.model)
best.model$cv.statistics
1 - best.model$cv.statistics$deviance.mean

#dev.off()
gbm.plot(best.model,smooth=FALSE,show.contrib=TRUE, y.label="Marginal effect on gs")
gbm.perspec(best.model,x=7,y=6)

# -----
## Plot Predictive Map

predictive.map <- predict(rasters,best.model,n.trees=best.model$gbm.call$best.trees,
  type="response")
plot(predictive.map)
writeRaster(predictive.map,filename="Results/final_nov2015.map",format="GTiff",overwrite=T)

# -----
# TSS

prediction.brt.point <- extract(predictive.map, final.data.occurrences[,1:2], method="bilinear")
Accur <- accuracy(final.data.occurrences[,3],prediction.brt.point,threshold=100)
Accur[which.max(Accur$specificity + Accur$sensitivity),]
Accur[which.max(Accur$specificity + Accur$sensitivity),4] + Accur[which.max(Accur$specificity +
Accur$sensitivity),5] - 1

# -----
# Determine the Areas

area.cells = area(predictive.map)
prediction.area <- calc(stack(area.cells,predictive.map),function (x) x[[1]] * x[[2]])
sum(values(prediction.area),na.rm=T)

```

```

# -----
# Partial Dependence Plots

prediction.brt.point <- extract(predictive.map, final.data.occurrences[,1:2],
                               method="bilinear")
Accur <- accuracy(final.data.occurrences[,3],prediction.brt.point,threshold=100)
Threshold <- Accur[which.max(Accur$specificity + Accur$sensitivity),1]

names(raster.stack)
raster.number <- 8
probability <- extract(predictive.map , final.data.occurrences[,1:2])
values.env.var <- extract(subset(raster.stack,raster.number), final.data.occurrences[,1:2])
#dev.off()

plot(values.env.var,probability, xlab=names(raster.stack)[raster.number],
      ylab="Probability of occurrence",pch=20, col="#a19e9f", axes = FALSE,
      # xlim for raster 5: phophate_min is xlim=c(0,0.4)
      # xlim for raster 1: curr is xlim=c(0,0.5)
      xlim=c(0,0.5)
      # if can plot the automatic x-axis, run following xlim
      #xlim=c(round(min(values.env.var,na.rm=T),digits=3),
      # round(max(values.env.var,na.rm=T),digits=3)))
)
axis(1, lwd = 1, at=seq(0,
                       0.5,
                       length.out = 5),
     lab=(seq(0,
              0.5,
              length.out = 5)))
# automatic x-axis. Do not run if x labels are incorrectly plotted (ie, use code above)
axis(1, lwd = 1, at=seq(round(min(values.env.var,na.rm=T)),
                       round(max(values.env.var,na.rm=T)),
                       length.out = 5),
     lab=round(seq(round(min(values.env.var,na.rm=T)),
                  round(max(values.env.var,na.rm=T)),
                  length.out = 5)))
axis(2, lwd = 1, at=seq(0,1,length.out = 5), lab=seq(0,1,length.out = 5))

# plot this code just for variables named "min"
abline(v=min(values.env.var[which(probability > Threshold)]),lwd = 2,lty=3)
min(values.env.var[which(probability > Threshold)], na.rm=T)

# plot this code just for variables named "max"
abline(v=max(values.env.var[which(probability > Threshold)]),lwd = 2,lty=3)
max(values.env.var[which(probability > Threshold)], na.rm=T)

# -----

```

APPENDIX 4.

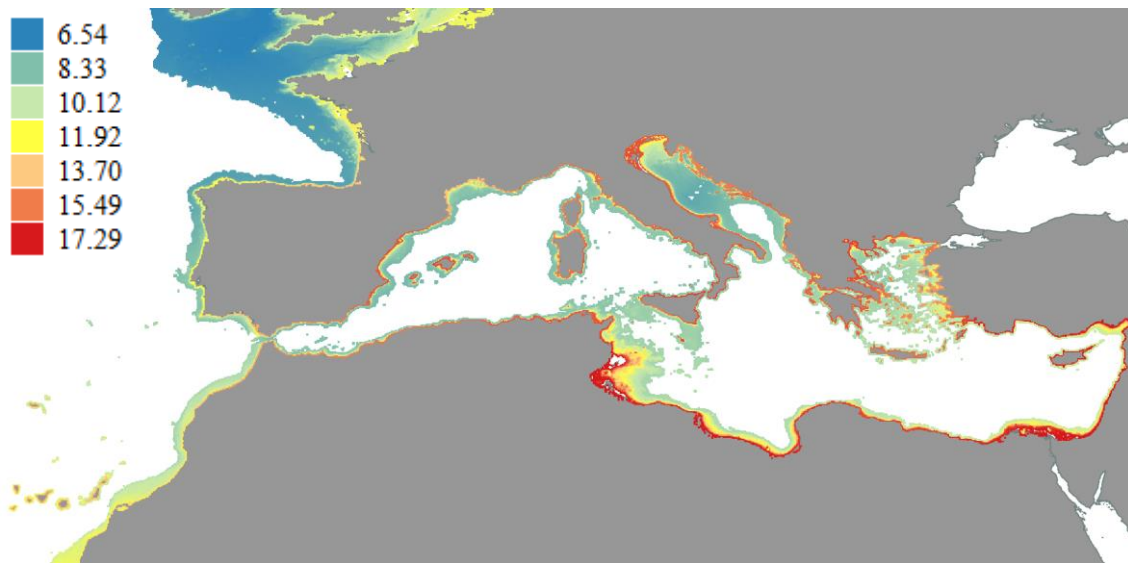


Figure A.4.1 Data of the minimum temperature gradient (6.54 to 17.29°C) along the study area, obtained from Guinehut *et al.*, 2012

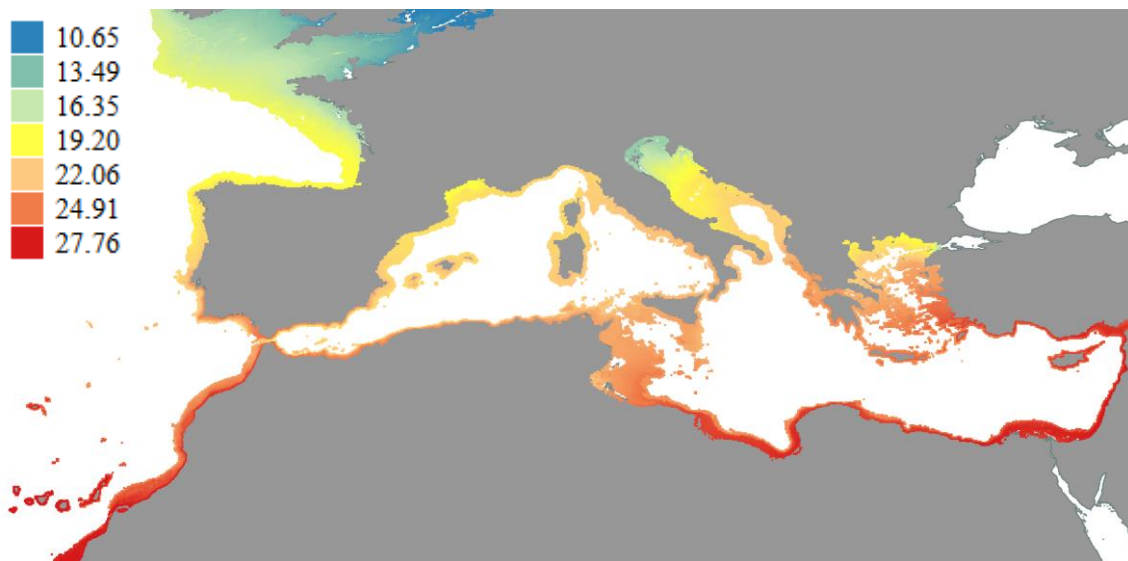


Figure A.4.2 Data of the maximum temperature gradient (10.65 to 27.76 °C) along the study area, obtained from Guinehut *et al.*, 2012

APPENDIX 5.

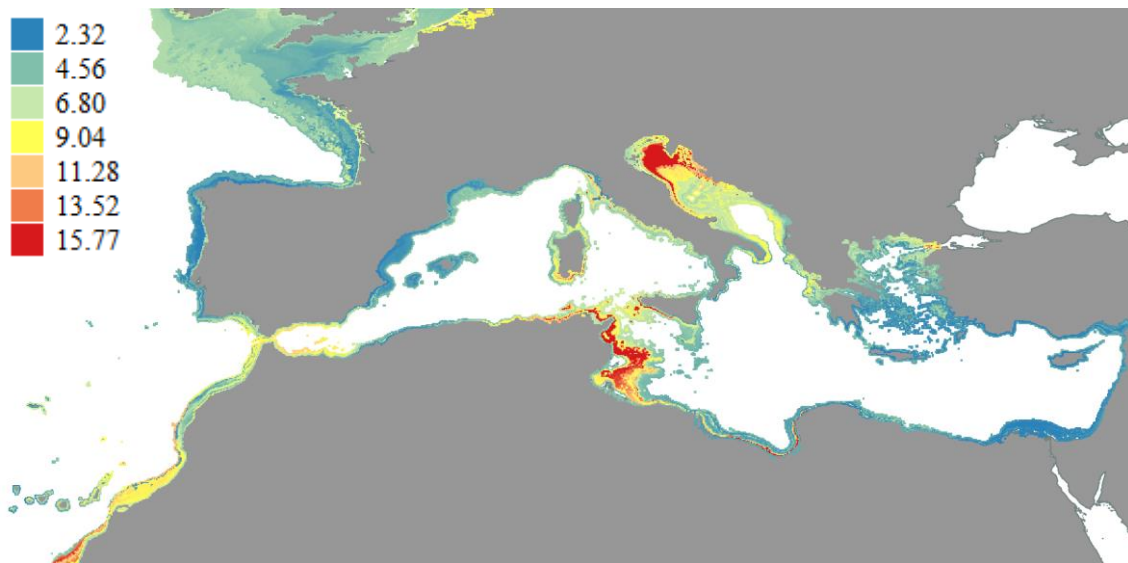


Figure A.5.1 Data of the minimum phosphate gradient (0.00 to 0.48 $\mu\text{mol/L}$) along the study area, obtained from [Levitus *et al.*, 2013](#)

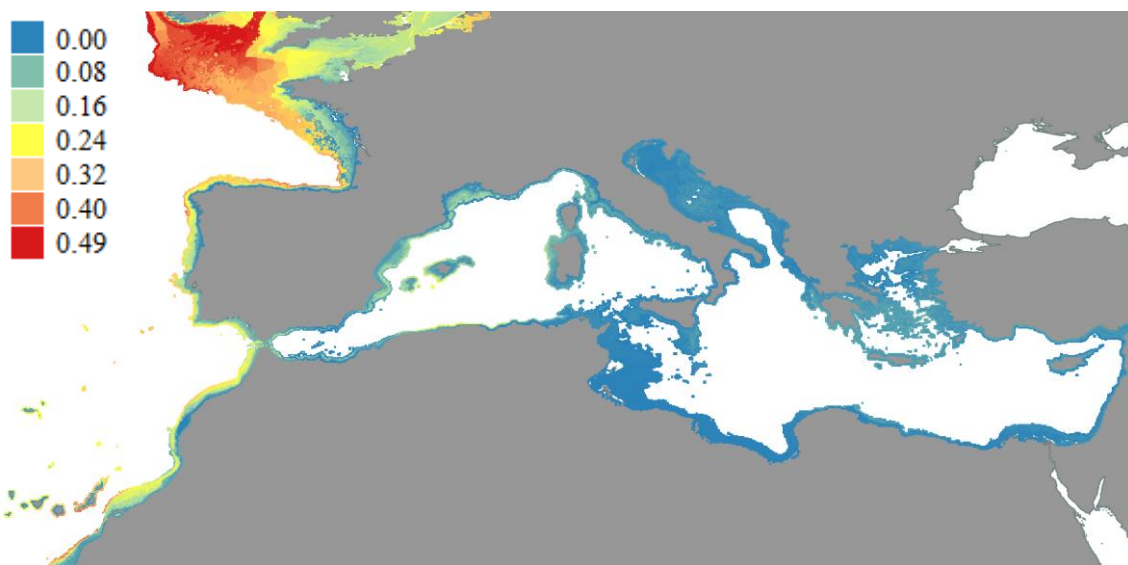


Figure A.5.2 Data of the maximum silicate gradient (2.32 to 15.76 $\mu\text{mol/L}$) along the study area, obtained from [Levitus *et al.*, 2013](#)

APPENDIX 6.

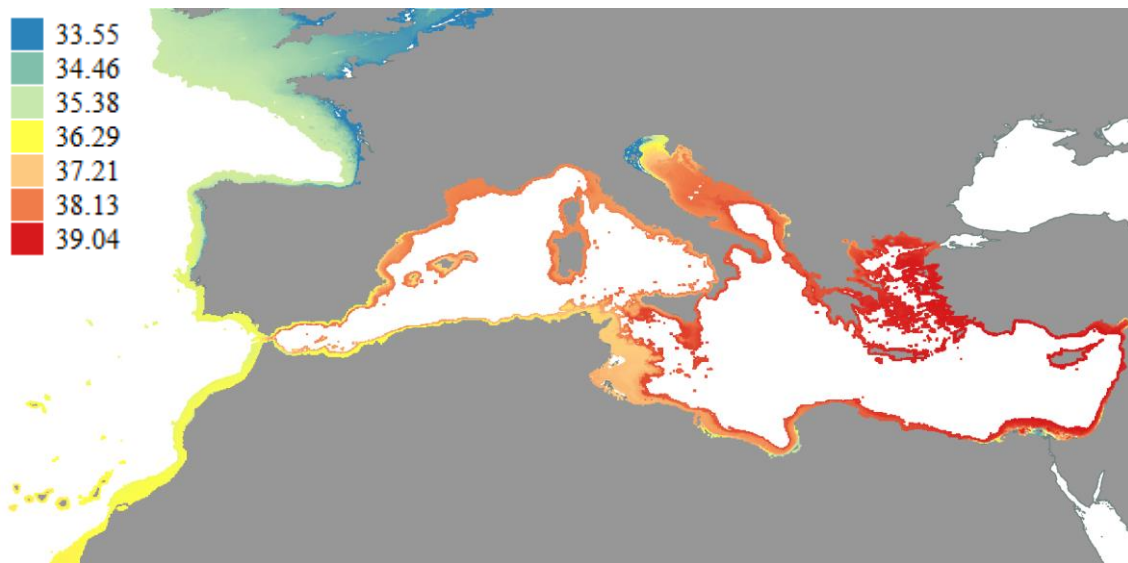


Figure A.6 Data of the salinity gradient (33.55 to 39.05 PSS) along the study area, obtained from Guinehut *et al.*, 2012