HELLENIC REPUBLIC
National and Kapodistrian University of Athens
EST. 1837

hcmr
ΕΛΚΕΘΕ

Modelling the distribution of the mesopelagic fish *Maurolicus muelleri* in the Greek seas

Alexandros Alamanellis-Zisimopoulos, 22010

Department of Geology and Geoinvironment, School of Science,

National and Kapodistrian University of Athens,

Msc "Oceanography and Management of the Marine Environment"

Examination Committee:

i) Dr. Konstantinos Tsagkarakis

ii) Dr. Dionysios Raitsos-Exarchopoulos

iii) Dr. Marianna Giannoulaki

Athens, 2023

Contents

Abstract

Mueller's pearlside (*Maurolicus muelleri*) is a mesopelagic fish with a wide distribution in the Pacific and Atlantic Oceans, including the Mediterranean Sea. The potential habitat of the species in the Greek seas was identified using ensemble modelling derived from eight different Species Distribution Models (SDMs). Acoustic data collected in the Ionian and Aegean Seas between 2018 and 2020 were used along with biological sampling performed with a pelagic trawl for determining the species presence and its characteristic echo-traces in the study areas. *M. muelleri* echo-traces appear as large aggregations of smaller, densely-populated shoals, that contribute to a cloud-shaped formation. Presence/absence observations, paired with environmental and bathymetry data, were modeled using the Biomod2 software package. Out of 10 different Species Distribution Modelling approaches applied, those eight with the highest evaluation metrics were used for ensemble modelling, producing annual probability maps of *M. muelleri* presence in the Greek seas for the period 2018-2020. Bottom depth showed the highest variable importance in all models, followed by the surface chlorophyll concentration (mg/m^3) and sea current geostrophic velocity (m/s). Probability of presence was high in the zone between 100 m and 250 m during every year, presenting widespread suitable habitats in the Greek seas. Hotspots were identified in the North Aegean Sea and near the shelf breaks throughout the Greek Seas. The results contribute to the knowledge of mesopelagic fish distribution and their potential interaction with charismatic species and assemblages of commercial importance.

Περίληψη

Το είδος *Maurolicus muelleri* είναι μεσοπελαγικό ψάρι με ευρεία εξάπλωση στον Ειρηνικό και τον Ατλαντικό Ωκεανό, συμπεριλαμβανομένης και της Μεσογείου. Τα πιθανά ενδιαιτήματα του είδους στις ελληνικές θάλασσες προσδιορίστηκαν με τη χρήση ενός συνδυαστικού μοντέλου που προέκυψε από οκτώ επιμέρους στατιστικά μοντέλα κατανομής ειδών (Species Distribution Models, SDMs). Χρησιμοποιήθηκαν ακουστικά δεδομένα που συλλέχθηκαν από το Ιόνιο και το Αιγαίο Πέλαγος από το 2018 έως το 2020, καθώς και βιολογικές δειγματοληψίες που πραγματοποιήθηκαν με πελαγική τράτα για τον προσδιορισμό της παρουσίας του είδους και των χαρακτηριστικών ηχογραμμάτων του στις περιοχές μελέτης. Τα ηχογράμματα του *M. muelleri* εμφανίζονται ως μεγάλες συγκεντρώσεις μικρότερων, πυκνών κοπαδιών, που συμβάλλουν σε ένα σχηματισμό που προσομοιάζει σε νέφος. Τα δεδομένα παρουσίας-απουσίας, σε συνδυασμό με περιβαλλοντικά δεδομένα και δεδομένα βαθυμετρίας, μοντελοποιήθηκαν χρησιμοποιώντας το πακέτο λογισμικού Biomod2. Από τα 10 διαφορετικά SDMs που εφαρμόστηκαν, τα 8 με τους υψηλότερους δείκτες αξιολόγησης χρησιμοποιήθηκαν για τη δημιουργία του συνδυαστικού μοντέλου, παράγοντας χάρτες που δείχνουν την πιθανότητα παρουσίας του *M. muelleri* στις ελληνικές θάλασσες για κάθε έτος από το 2018 έως το 2020. Το βάθος του πυθμένα παρουσίασε την υψηλότερη σημαντικότητα σε όλα τα μοντέλα, με τη συγκέντρωση χλωροφύλλης (mg/m^3) στην επιφάνεια της θάλασσας και την γεωστροφική ταχύτητα ρευμάτων (m/s) να ακολουθούν. Η πιθανότητα παρουσίας ήταν υψηλή στη ζώνη μεταξύ 100m και 250m, με ισοδύναμα αποτελέσματα μεταξύ των τριών ετών, παρουσιάζοντας μεγάλο εύρος κατάλληλων ενδιαιτημάτων στις ελληνικές θάλασσες, με hotspots που εντοπίστηκαν στο Βόρειο Αιγαίο και κοντά στην ηπειρωτική υφαλοκρηπίδα των Ελληνικών θαλασσών. Η μελέτη συμβάλλει στη γνώση για την κατανομή των μεσοπελαγικών ψαριών στις ελληνικές θάλασσες αλλά και την πιθανή αλληλεπίδρασή τους με εμπορικής σημασίας είδη.

1. Introduction

The mesopelagic zone, also known as the twilight zone, is defined as the zone that stands between the euphotic, epipelagic zone and the bathypelagic zone. In terms of bathymetry the mesopelagic zone is usually defined as the area between 200 m and 1000 m depth [1]. Only a small fraction of sun light reaches the mesopelagic zone [2], thus rendering the conditions insufficient for photosynthesis and primary production. Despite the limited presence of light in this large area that covers 20% of the global ocean mass, adaptations of the organisms that occupy the mesopelagic zone, make the light sufficient for predation [3].

Mesopelagic fish constitute the most abundant group of vertebrates in the marine environment [4] and are present in every ocean [5]. They occupy the twilight zone during the day while many species perform Diel Vertical Migrations (DVM) during the night to feed upon the zooplankton of the epipelagic zone. Thus, they play a key role in the energy influx within the mesopelagic ecosystem. Mesopelagic fish assemblages form distinct layers in the mesopelagic zone, called the deep scattering layers (DSL), that can be identified by acoustic means. The global biomass of mesopelagic species had been largely underestimated in the past, with older studies suggesting that it is ~1,000 million tons [5]; in contrast, a more recent study has estimated the global biomass to be one order of magnitude higher (between 11,000 and 15,000 million tons) [6]. Therefore, their important role in the mesopelagic ecosystem underscores the significance of understanding their assemblages and distribution. Underwater acoustics are commonly employed as a tool for assessing mesopelagic fish abundance and distribution.

Underwater acoustics, such as echosounder devices transmit sound waves vertically through the water column. By examining the backscatter of the transmitted sound, information can be acquired about the distance from the target (e.g. sea bottom and fish schools) and its density (Horne, 2000). Different fish species, due to their respective anatomy, size and behavior inside the water column, may form schools that have different density and/or echo trace characteristics. By combining acoustic data alongside supplementary biological sampling, echo traces can be attributed to specific species or groups of species.

Since broad-scale sampling is extremely time-consuming and expensive especially under a regular monitoring scheme, there is a need to develop techniques to indirectly quantify the distribution of marine organisms. Species distribution models (SDMs) are algorithms that use presence/absence or abundance data to relate them with environmental data, in order to predict and visualize the spatial distribution of the species inside the study area (Elith *et al*., 2006; Giannoulaki *et al*., 2014). SDMs are useful tools that contribute to understanding the dynamics between environmental parameters and habitat suitability which is essential for ecosystem management, fisheries and prediction of changes in future

species habitats due to climate change. SDMs are widely used in studies for marine species but the single SDM approach poses a risk of uncertainty. Increasing number of studies have shown that using ensemble SDMs helps assessing the factor of uncertainty and produce improved predictions, considering the variability amongst results produced with different algorithms (Grenouillet *et al*. 2011; Ramirez-Reyes *et al*., 2021).

The mesopelagic zone in the Mediterranean Sea is occupied by fewer fish species compared to the adjacent Atlantic Ocean, with the most abundant families being Myctophidae, Sternoptychidae, Phosichthyidae, Gonostomatidae, Stomiidae and Paralepididae [11]. Mueller's pearlside *Maurolicus muelleri* is a small mesopelagic fish belonging to the Sternoptychidae family with world-wide distribution [12]. It is usually found in the shallower part of the mesopelagic zone and is even considered to define its upper limit (Kaartvedt *et al*. 2019), while it is known to perform DVM to feed [13]. It is not currently a target for commercial fishing although there have been attempts and an ongoing interest for its exploitation [14,15]. Studies have shown that it is part of the diet of different commercial and charismatic species, such as *Thunnus thynnus* [16], *Eythynnus alleteratus* [17], gadoids [18] and marine mammals [19], suggesting that it a species of great ecological interest. Still, little is known about the preferred habitats of mesopelagic fishes. In the Greek seas *M. muelleri*, is one of the most prevalent species, usually occupying areas close to the seafloor, near the continental shelf (Kapelonis *et al*., 2023). With the scientific interest around those species rapidly growing, new knowledge about their habits and distribution is essential. Thus, obtaining information on their spatial distribution and the link with different environmental parameters, provides knowledge on the dynamics of the mesopelagic ecosystem, on potential climate change effects, and probable implications deriving from their exploitation.

The objectives of the current study were: *(i)* to model the distribution of *M. muelleri*, using modern techniques, in particular ensemble species distribution modeling (SDM), *(ii)* to investigate the importance of certain environmental parameters in the prediction of the suitable habitat for *M. muelleri*, and *(iii)* to project suitable habitats into maps to identify areas with high probability of presence. To this end, species presence and absence from acoustic data were combined with a set of environmental parameters, through 10 different SDM algorithms. An ensemble algorithm was created and the probability of the species' presence was projected in maps of the Greek Seas, providing valuable knowledge for understanding the ecology of the species and the mesopelagic zone in general.

2. Materials & Methods

2.1. Study Areas

The study area covered several parts of the Greek seas, including open seas as well as enclosed gulfs. The Aegean Sea is located in the eastern part of the Mediterranean (Fig.1), separating the Greek and Turkish mainland and is the area where the majority of the Greek fisheries operate. It is characterized by high environmental variability; the Northern part of the Aegean Sea is considered one of the most productive areas of the Eastern Mediterranean due to nutrient rich water inflows, from big rivers and the Black Sea [21–24], while the central and southern parts are quite oligotrophic. The Cretan Sea is located at the Southernmost part of the Aegean, north of the island of Crete, (Fig.1), and is the deepest basin of the Aegean, reaching ~2500 m depth [25]. It is considered one of the most oligotrophic areas in the Greek seas and the biggest part of its primary productivity is attributed to upwelling from gyres and vertical mixing of water layers [26]. On the west Aegean Sea, between Evia Island and the mainland, is located the Gulf of Evia (Fig.1), a generally shallow gulf with a deep basin in its Northern part where maximum depth is ~440 m, that is considered oxygen deficient from ~350m and below (HCMR unpublished data). The Saronic Gulf, with a maximum depth of ~450 m, is also part of the west Aegean Sea, located south of the Attica peninsula and the Greek capital of Athens. Due to its proximity to the metropolitan area of Athens, it is considered heavily disturbed from anthropogenic pollution, mostly attributed to urban and industrial waste [27–29].

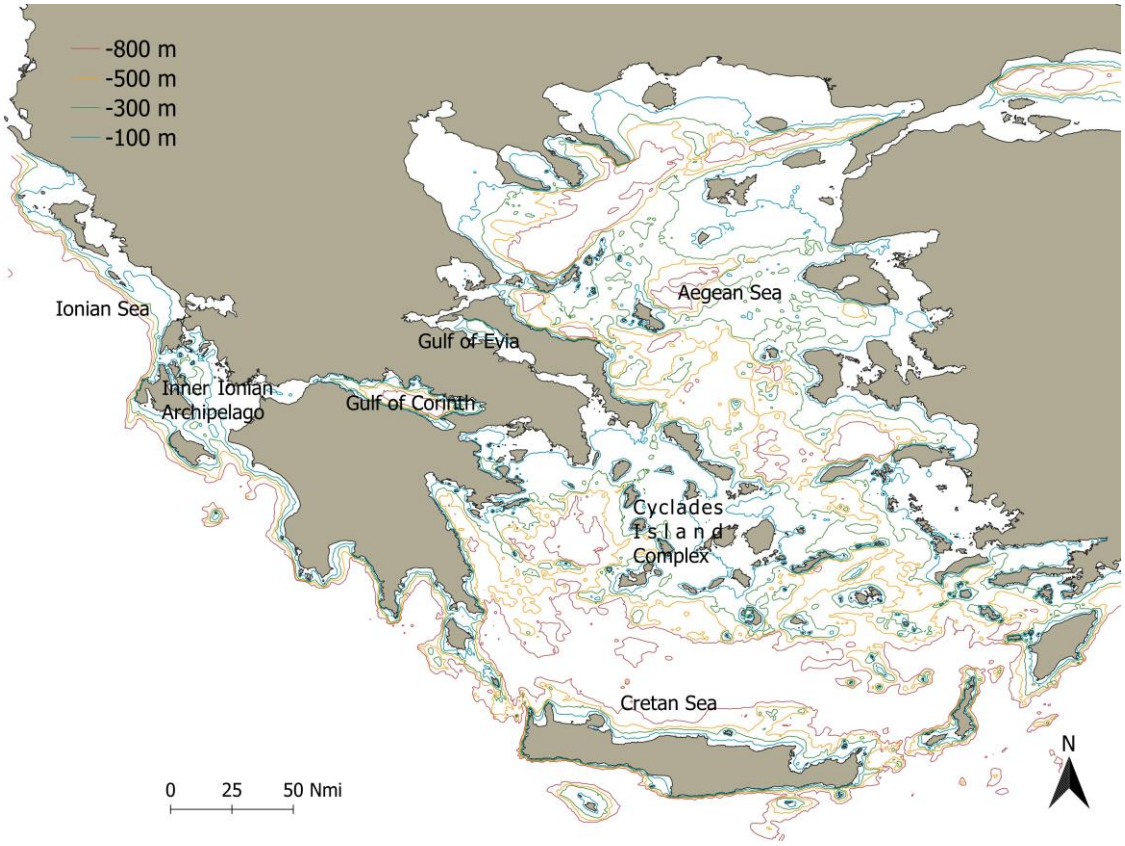


Fig. 1. Contour map of the Greek Seas and study areas

The Ionian Sea, located between the Greek and Italian mainland (Fig.1), is the deepest sea in the Mediterranean with a maximum depth of 5121m [30]. The Eastern, Greek part of the Ionian is considered oligotrophic in comparison with the Aegean Sea, with no influence from big river inflows [31] and a relatively narrow continental shelf. The Hellenic Trench, a long deep trough starting from the central Ionian Sea which heads south of Crete, has been identified as an important marine mammal area as it hosts several cetaceans that inhabit deep waters [32–34]. Part of the Ionian is the Gulf of Corinth (Fig.1), a deep gulf with a maximum depth of 935m and a steep slope, that is connected to the Ionian through the much shallower Patraikos Gulf (Kapelonis *et al*., 2023). It is considered of great ecological importance as it is a habitat for a large number of marine mammals, especially dolphins [35,36].

2.2. Acoustic and Biological data

2.2.1. Sampling

The presence of *M. muelleri* was explored with acoustic surveys held in the Aegean and Ionian seas (Eastern Mediterranean Sea, Greece) (Leonori *et al*., 2021; Kapelonis *et al*., 2023) on board R/V Philia, in the framework of monitoring and research projects targeting small pelagic and/or mesopelagic fish (Leonori *et al*., 2021; Kapelonis *et al*., 2023), during the time period 2018-2020. Acoustic data were collected using a hull-mounted Simrad EK80 split beam echosounder, traveling at a speed of ~8 kn. Out of the four split-beam transducers (38, 120, 200 and 333 kHz), only the 38 kHz frequency was capable of operating throughout the depth range of all the surveys, and thus, used for the analysis. The echosounder system was calibrated, before each survey, with the standard target method (Demer *et al*., 2015). Echo sounding took place during daytime, both in predefined transects and opportunistically during transit, while the total distance covered during the sampling was approximately 2500 nm (Fig. 2).

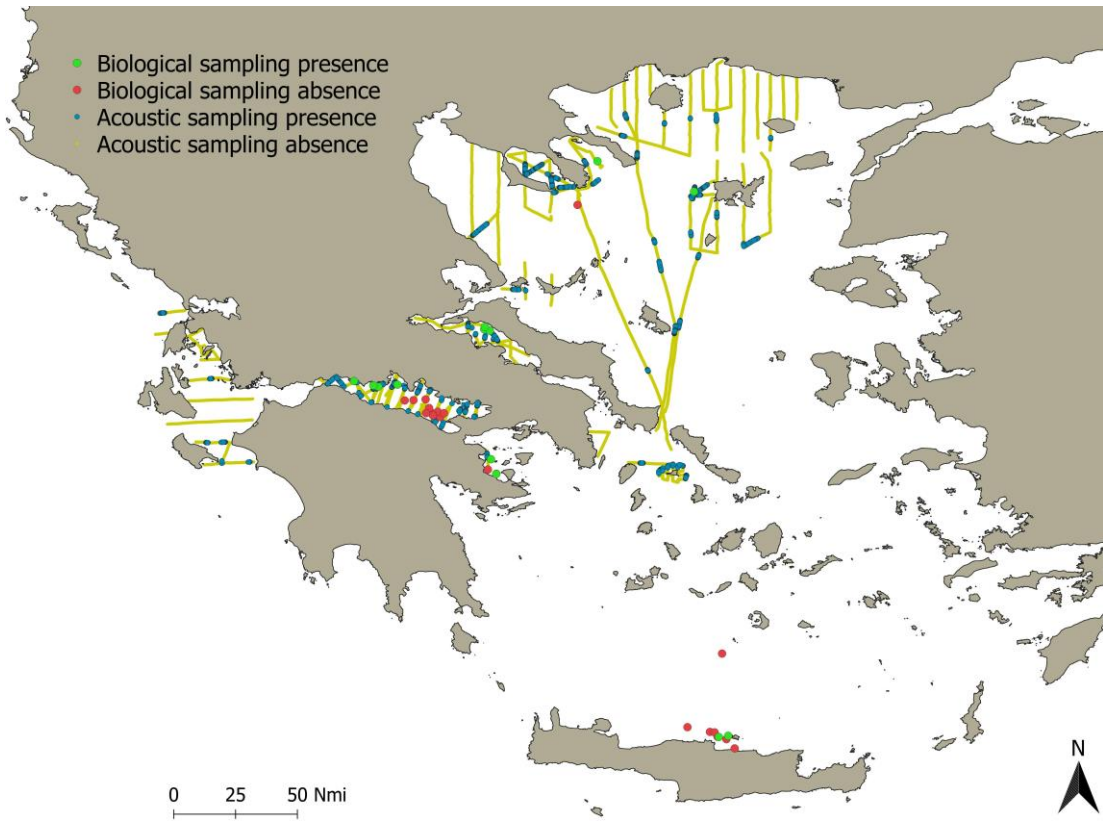


Fig. 2. Map of the Greek seas, presenting the locations of observed *M. muelleri* characteristic echo-type presence (blue) and absence (yellow) during the acoustic sampling surveys and the locations of observed *M.* presence (green) and absence (red) during the biological sampling, in the time period between years 2018 and 2020.

Biological sampling for echo-trace identification took place in parallel with the acoustic surveys in selected parts of the study areas (Fig. 2) using a pelagic trawl of 7 m vertical opening, 12 m horizontal opening and a 16 mm (stretched) mesh cod-end, typically towed at 3-4 knots.

2.2.2. Analysis

Biological sampling and acoustic data were used alongside former knowledge (Kapelonis *et al*, 2023), to produce an expert-based protocol for recognizing the characteristic echo-traces of *M. muelleri*.

The entirety of the acoustic data was grouped by transect and through the Echoview software, a calibration procedure took place. The software settings were tuned for each transect, based on the environmental conditions of the area (temperature, salinity) recorded by CTD casts. The bottom line was automatically set by the Software's bottom detection feature and manual corrections took place where necessary. Different echo-trace regions were identified and outlined manually in polygons. The echo traces that resembled the characteristics attributed to *M. muelleri* were manually labeled. The acoustic tracks were divided into a number of intervals based on an elementary distance sampling unit (EDSU) with length of 200 m. Using a threshold of -70 dB, the Nautical Area Scattering Coefficient (NASC-s_A) [m^2 nmi $^{-2}$] was calculated for each interval, which is the acoustic backscatter (s_v) integration of the *M. muelleri* region samples for the specified interval scaled to 1 nmi^2. Species presence was assigned to each interval with a non-zero *M. muelleri* NASC, and absence was assigned to all other intervals.

2.3. Environmental and Bathymetry data

In order to associate *M. muelleri* presence with environmental parameters, a list of variables was explored and their values were retrieved for each interval sampled. Bathymetry data were directly obtained from the acoustic integration data by means of the average bottom depth per interval and a set of 13 environmental variables were acquired from respective databases (Table 1). For satellite-derived parameters, annual mean environmental values for the time period of 2018-2020 were used (instead of monthly or daily ones) because the presence data were collected in different seasons, and no extended seasonal horizontal migrations have been reported for *M. muelleri* in the study area (Kapelonis *et al*., 2023). Variables related with chlorophyl, temperature and organic matter such as Surface Chlorophyl-a Concentration (CHL, mg/m^3), Photosynthetic Active Radiation (PAR,

mol·m^{-2}·s^{-1}), Sea Surface Temperature (SST, °C), Euphotic Depth (ZEU, m) and Particulate Organic Carbon (POC, mg/m^3) were acquired from Ocean Color (oceancolor.gsfc.nasa.gov). Variables related with hydrology, i.e. Eddy Kinetic Energy (EKE, cm²/s²), Sea Level Anomaly (SLA, m), meridional component of the absolute geostrophic velocity current (VADT, m/s) and zonal component of the absolute geostrophic velocity (UADT, m/s) were acquired from AVISO+ (www.aviso.altimetry.fr), and Salinity (SAL, psu) and Sea Bottom Temperature (SBT, °C) were acquired from Copernicus Marine Data Service (marine.copernicus.eu). Distance to Coast (DTC, m) was calculated from Euclidian distance through GIS. Sea Bottom Slope (SLO, degrees from North), was acquired from GEBCO (www.gebco.net), and so was Sea Bottom Depth (DEP, m) data for the needs of SDM projection (see below). More information on each variable, i.e. transformations, spatial resolution and data source(s), is listed in Table 1.

2.4. Data preprocessing

All variables were tested for collinearity in order to omit those that were highly correlated, to improve the interpretability of the models (including diagnostic/evaluation metrics) and to reduce the model error for some SDMs. Four of the variables (CHL, EKE, POC, ZEU) presented high Pearson's Correlation Coefficient values (PCC>0.7; Fig. 3), thus only Sea Surface Chlorophyl Concentration (CHL) was kept and the rest were not used for the SDM training to avoid model overfitting [39]. Model fitting has higher success probability when the data used follow the Gaussian distribution, thus transformations were applied to environmental data where necessary (Table 1, Sup. Fig. 1) and all the data were standardized by subtracting the mean and scaling to unit variance (i.e. z-transform).

Table. 1. List of environmental variables tested along with data sources, available resolutions and applied transformations.

Variable	Tranformations	Source	Sensor/Product	Resolution
Depth	cubic root, z-transform	simrad ek80, gebco.net	GEBCO_2023	450 m
CHL	reversed, z-transform	oceancolor.gsfc.nasa.gov	MODIS-A SMI	4 km
Slope	log, z-transform	gebco.net	GEBCO_2023	450 m
UADT	z-transform	www.aviso.altimetry.fr	SSALTO/DUACS (merged product)	0.125°
VADT	z-transform	www.aviso.altimetry.fr	SSALTO/DUACS (merged product)	0.125°
PAR	z-transform	oceancolor.gsfc.nasa.gov	MODIS-A SMI	4 km
SLA	z-transform	www.aviso.altimetry.fr	SSALTO/DUACS (merged product)	0.125°
SST	z-transform	oceancolor.gsfc.nasa.gov	MODIS-A SMI	4 km
SBT	z-transform	marine.copernicus.eu	MEDSEA_MULTIYEAR_PHY_006_004_E3R1	4 km
ZEU	z-transform	oceancolor.gsfc.nasa.gov	MODIS-A SMI	4 km
POC	reversed, z-transform	oceancolor.gsfc.nasa.gov	MODIS-A SMI	4 km
EKE	reversed, z-transform	www.aviso.altimetry.fr	SSALTO/DUACS (merged product)	10 km
SAL	square root, z-transform	marine.copernicus.eu	MEDSEA_MULTIYEAR_PHY_006_004_E3R1	0.333°
DTC	cubic root, z-transform	Euclidian distance	GIS calculated	400m

The presence/absence ratio of the hydroacoustic data for *M. muelleri* was ~9%. It has been shown that the number of absences used in SDMs has great influence in their performance [40]. The preferred presence/absence ratio appears to be different for each type of SDM, with regression techniques such as generalized additive model (GAM) and generalized linear model (GLM), having higher accuracy when more absence data are used, while machine learning techniques such as gradient boosting machine (GBM), random forest (RF) and multiple adaptive regression splines (MARS), have higher accuracy when the presence/absence ratio is equal to 1 [40]. The magnitude of the influence that presence/absence ratio has on SDM accuracy also varies amongst different techniques, with little influence on regression techniques and high influence on machine learning techniques [40]. Also, the 1:1 presence/absence ratio, is shown to provide better SDM predictions, when the absence data are taken from a large area [41]. Thus, the presence/absence data that were used in the present study, were balanced using an

algorithm that removes absence values randomly from the dataset until the number of presence and absence data is equal.

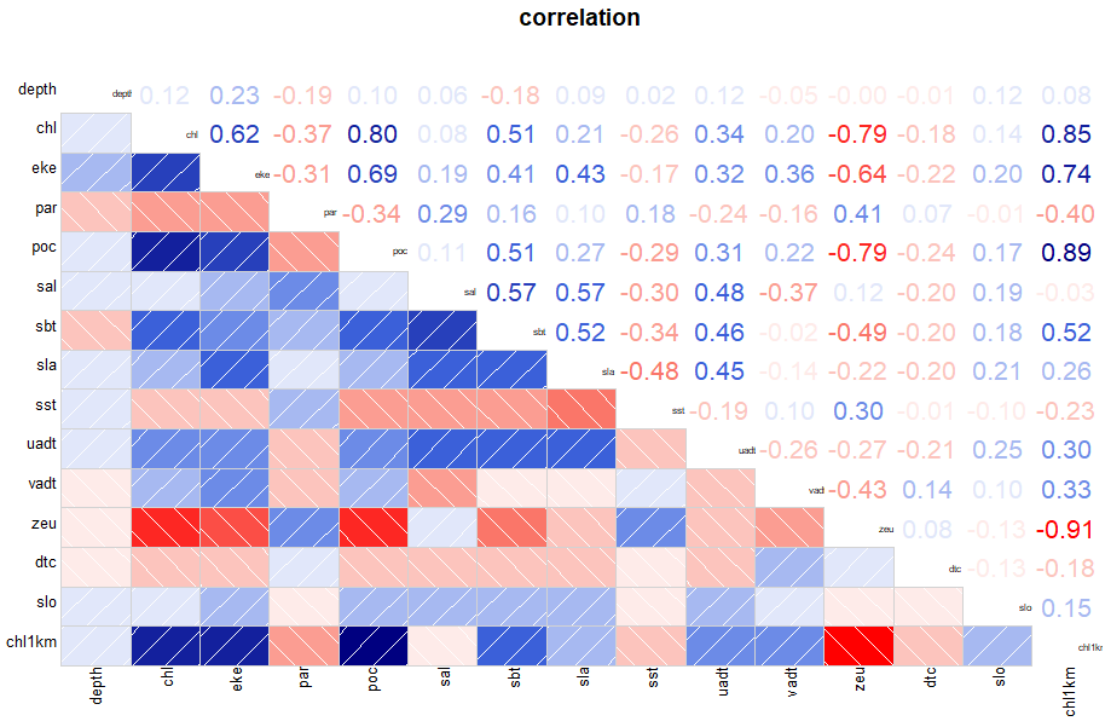


Fig. 3. Pearson's Correlation Coefficient plot of the environmental variables tested

2.5. Modelling

The Biomod2 R software package [42,43], was used to fit multiple SDMs and produce an ensemble model. Biomod2 combines any number of the following 10 SDM algorithms: classification tree analysis (CTA), generalized additive model (GAM), gradient boosting machine (GBM), generalized linear model (GLM), surface range envelop (SRE), artificial neural network (ANN), random forest (RF), multiple adaptive regression splines (MARS), flexible discriminant analysis (FDA), and maximum entropy (Maxent) to form an ensemble SDM algorithm that predicts the probability of the targeted species' presence inside the study area.

The cross-validation strategy that was used for individual SDM performance evaluation, was the "random" Biomod2 strategy. The dataset was split in two groups, all the models used 80% of the data for calibration, then the remaining 20% of the data were used for calculating two different evaluation metrics: the true skill statistic (TSS) and the receiver operating characteristic (ROC) area under the curve (AUC). Both of them were calculated from specificity and sensitivity values for each individual SDM, using the true presence

ratio and false presence ratio deriving from the two groups of the split dataset. TSS values > 0.6 and AUC ≥ 0.8 are considered good to excellent [44].

Different combinations of tuning settings were tested for each SDM by running it three times for each set of settings and then comparing their mean evaluation metrics, TSS and ROC. The default Biomod2 settings for species distribution modelling showed the highest performance in every model, thus they were used for the final model runs.

An additional round of SDM runs took place and Variable importance indexes were calculated for each individual SDM by Biomod2; environmental variables that showed really low importance indexes (<0.1) in all SDMs were removed one at a time and the SDMs ran again, with variables that continued to show low importance indexes being removed. Bottom depth, Sea Surface Chlorophyl Concentration, Sea Bottom Slope, Photosynthetic Active Radiation (PAR) and Absolute Dynamic Topography (VADT and UADT), was the final set of variables that were used for model training. Each of the variables that were removed, was individually placed back in with the final set for an extra test run. None of those variables showed higher variable importance on those test runs, thus the final variable set remained the same.

The SDMs that met the evaluation metric thresholds, in particular TSS>0.7 and ROC>0.8, were used to create an ensemble model algorithm. The ensemble model was created using a weighted mean of the SDMs that passed the evaluation metric thresholds; the individual SDMs were weighted based on their TSS scores. Response curve diagrams were created for each environmental variable, to identify how the range of the variable values influences the probability of *M. muelleri* presence.

2.6. Projection

The final ensemble algorithm was used to predict the probability of *M. muelleri* presence in the entirety of the Greek seas. Using Biomod2 projection function, and three different environmental parameter datasets for the years 2018, 2019 and 2020, three maps were created that presented the probability of presence (0-1) of the species for each year. The maps were rendered using the Qgis software (www.qgis.org/en/site) and were processed to provide a clear image of the probability of presence gradients.

3. Results

3.1. Acoustic sampling

Biological sampling showed that echotraces associated with *M. muelleri* presence were largely monospecific as only few or none individuals of other species were caught. The shoals of *M. muelleri* inside the study areas showed a distinguished behavior compared to

other mesopelagic fish: their echotraces appeared as sparse, extended clusters of smaller densely populated schools that occupied areas often larger than 1 nmi in length and usually ranging from depths of 100 m to 250 m (Fig. 4). Echo-traces of juvenile fish aggregations appeared in shallower zones than larger individuals, as indicated by biological sampling-data gathered inside the gulf of Corinth. There was a limited number of pelagic hauls that *M. muelleri* was present along with other species; in such cases the echotraces were of different form, i.e., dense layers at short distance from the seabed. Since such layer-type formations (contrast to the abovementioned shoals) usually didn't contain *M. muelleri*, it was assumed that in these few instances, the species was caught at the boundary of its distribution, when the two echo-types overlapped. Therefore, such acoustic layers were not taken into account for modelling, probably ignoring a small portion of the species distribution. Out of 24,045 intervals analyzed, *M. muelleri* was documented in 1,922, with ~90% of the documented presence occurring in the depth zone of 100-300 m (Fig. 5).

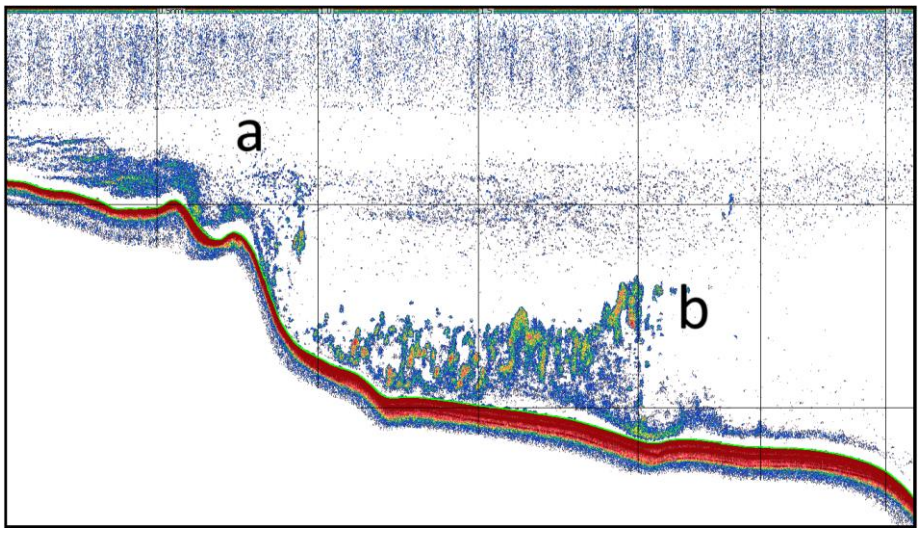


Fig. 4. Echogram showing shoals of juvenile (a) and adult (b) *M. muelleri*

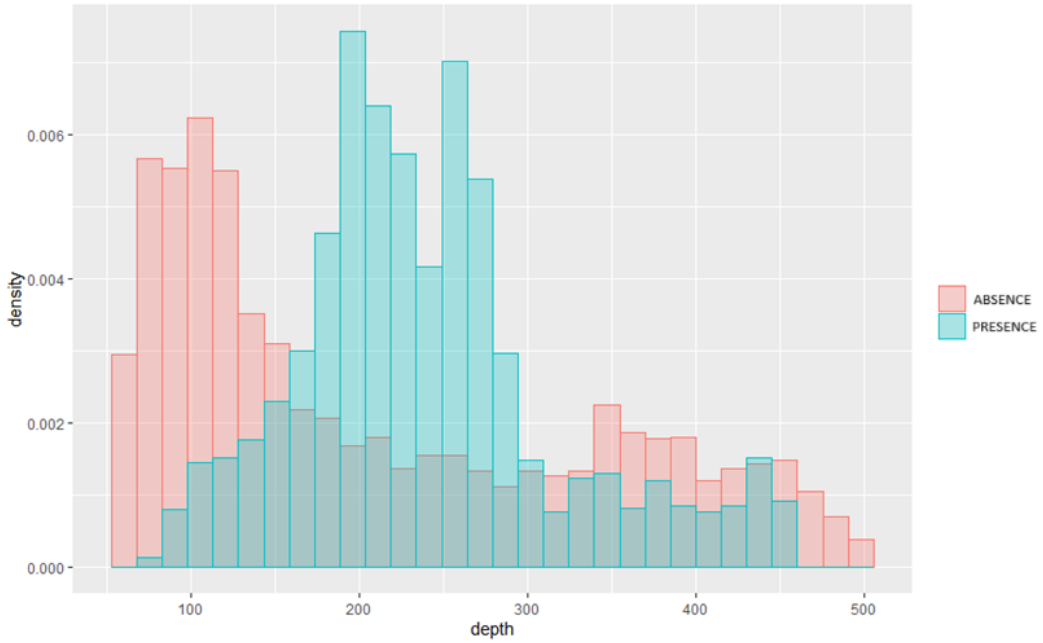


Fig. 5. Histogram demonstrating the density of presence (blue) and absence (red) observations for *M. muelleri* in relation to bottom depth (m), during the acoustic sampling surveys that took place between years 2018 and 2020.

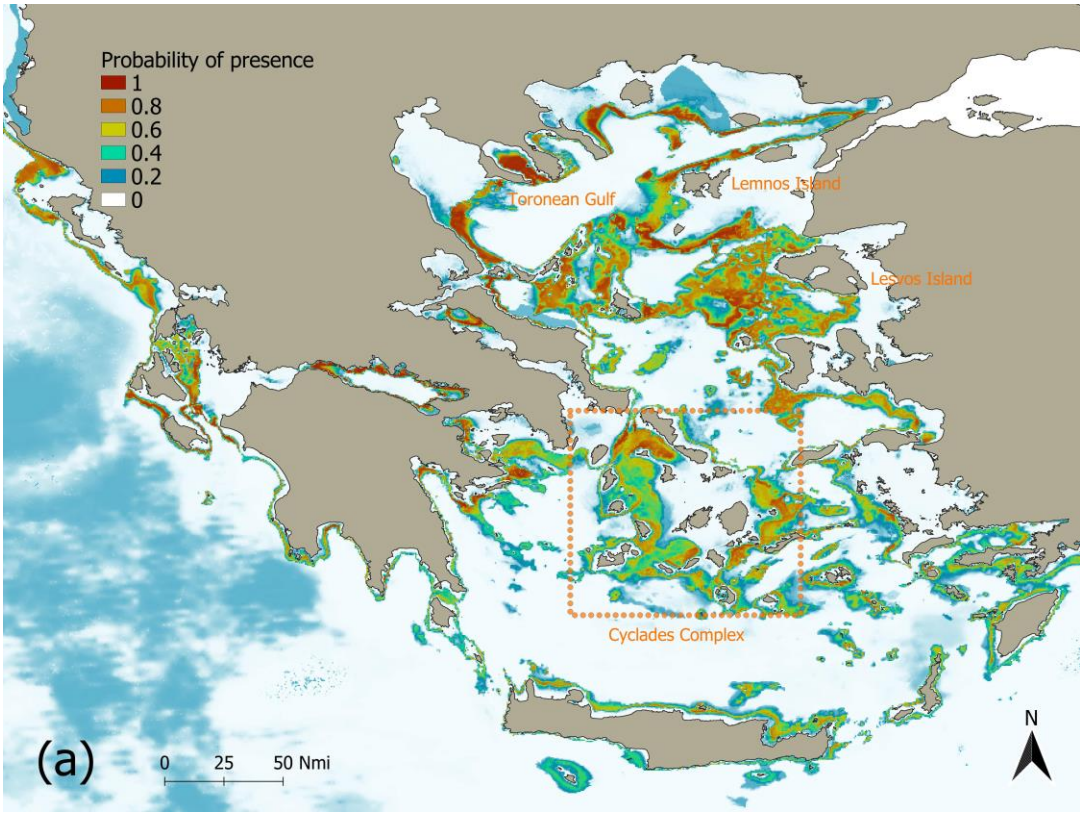
3.2. Modelling

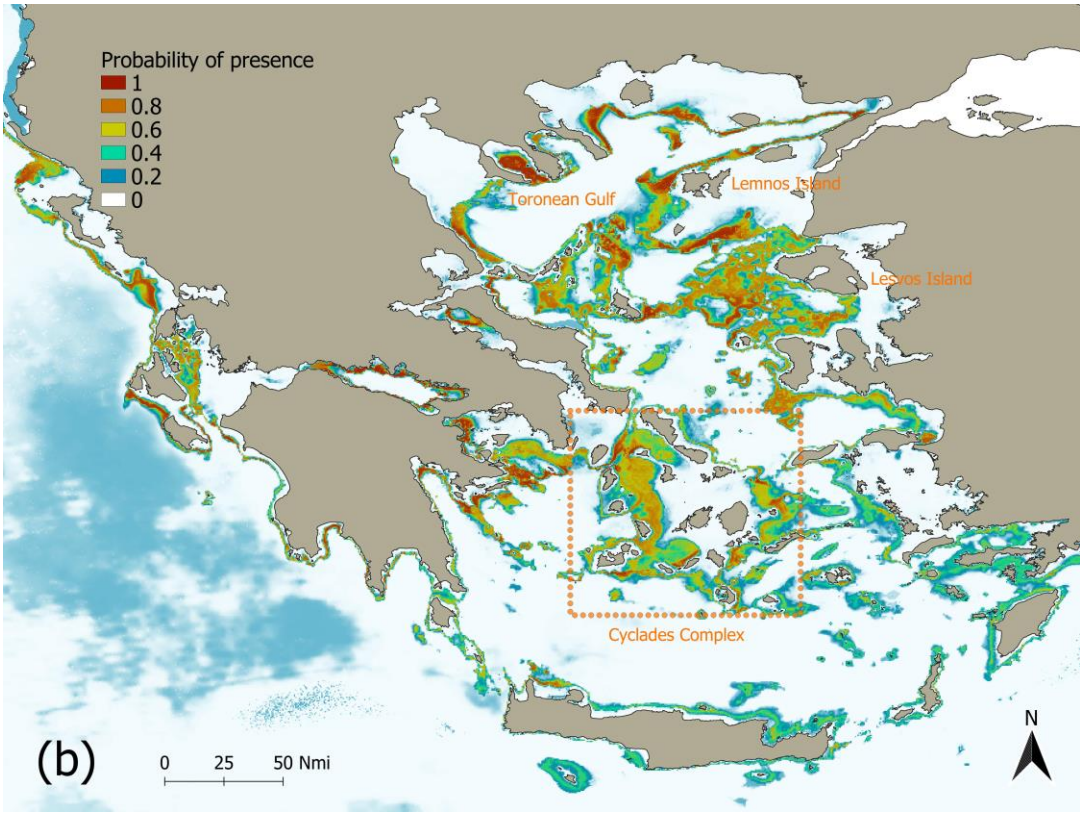
Biomod2 individual model runs were 80% successful, with eight out of ten modeling methods that the software package supports, showing higher evaluation metric scores than the acceptable predetermined threshold (TSS>0.7, ROC>0.8) (Table 2). Sea bottom depth was the highest contributing variable in all modeling methods with mean variable importance index of (0.82). CHL concentration was second with a mean variable importance index of (0.12), followed by UADT (0.11), bottom slope (0.06), PAR (0.06) and VADT (0.05), the last three environmental parameters had significantly lesser impact in model training. The ensemble modelling merged algorithm presented great fitting to the data with evaluation metric scores well above thresholds (TSS=0.88, ROC=0.99). According to the response curve diagrams produced by Biomod2, the highest probability of presence was associated with depths ranging from 200 m to 240 m, chlorophyl concentration ranging from 0.4 mg/m^3 to 1.5 mg/m^3 and UADT (west-east direction, sea current speed) ranging from -0.04 m/s to -0.02 m/s (Sup. Fig. 2).

Table. 2. Evaluation metric scores (True skill statistic, Receiver operating characteristic) and environmental variable importance indexes for each Species Distribution Model that was tested.

	Evaluation metric score		Variable importance					
SDMs	**TSS**	**ROC**	**Depth**	**CHL**	**Slope**	**UADT**	**VAD**	**PAR**
RF	0.952	0.997	0.546	0.206	0.093	0.206	0.065	0.061
CTA	0.863	0.950	0.799	0.213	0.122	0.270	0.091	0.157
GBM	0.856	0.971	0.904	0.093	0.025	0.027	0.001	0.035
MAXENT	0.814	0.958	0.904	0.048	0.061	0.062	0.010	0.057
MARS	0.809	0.948	0.864	0.160	0.014	0.078	0.000	0.060
FDA	0.777	0.943	0.947	0.130	0.008	0.025	0.000	0.050
ANN	0.777	0.927	0.977	0.083	0.024	0.068	0.030	0.059
GLM	0.773	0.927	0.959	0.066	0.011	0.030	0.029	0.021
GAM	0.646	0.879	0.718	0.110	0.013	0.278	0.111	0.030
SRE	0.473	0.737	0.628	0.067	0.197	0.106	0.110	0.054

Three different maps were projected using the Biomod2 ensemble model, presenting the probability of *M. muelleri* presence in the Greek seas from 2018 to 2020, one for each year. The three maps presented almost identical patterns due to bathymetry being the key variable in model training, which remains unchanged through the years. Probability of presence was extremely high in the North and North-East Aegean Sea (Fig. 6), with key factor being the extended areas that range from 150 m to 250 m depth, forming cyclical patterns at the boundaries of deep basins (e.g. the North Aegean Trough as well as south of Lemnos and west of Lesvos islands). In addition, a second extended area with somewhat lower probability of presence was identified in the Aegean Sea, covering similar depth ranges around the Cyclades Island complex and the Argo-Saronic Gulf (Fig. 6). The probability of presence was also much higher in the northern parts of the Greek seas decreasing progressively in the southern areas, following the downwards trend of Chlorophyl-a sea surface concentration from the more productive North to the South Aegean Sea. *M. muelleri* suitable habitats were also identified by the model in enclosed Gulfs presenting smoother bottom slopes, like the Toronean Gulf and the northern part of the Gulf of Evia, which provide extended areas with suitable depth ranges in combination with other conditions. In the Ionian Sea, the areas that presented high probability of presence are narrow and in close distance to the shore (Fig. 6) due to the steeper slope and the narrow continental shelf. Highest probability has been identified inside the Gulf of Corinth and in the Inner Ionian Sea archipelago (Fig. 6).


Probability of presence
1
0.8
0.6
0.4
0.2
0
Toronean Gulf
Lemnos Island
Lesvos Island
Cyclades Complex
N
(a)
0 25 50 Nmi


Probability of presence
1
0.8
0.6
0.4
0.2
0
Toronean Gulf
Lemnos Island
Lesvos Island
Cyclades Complex
N
(b)
0 25 50 Nmi

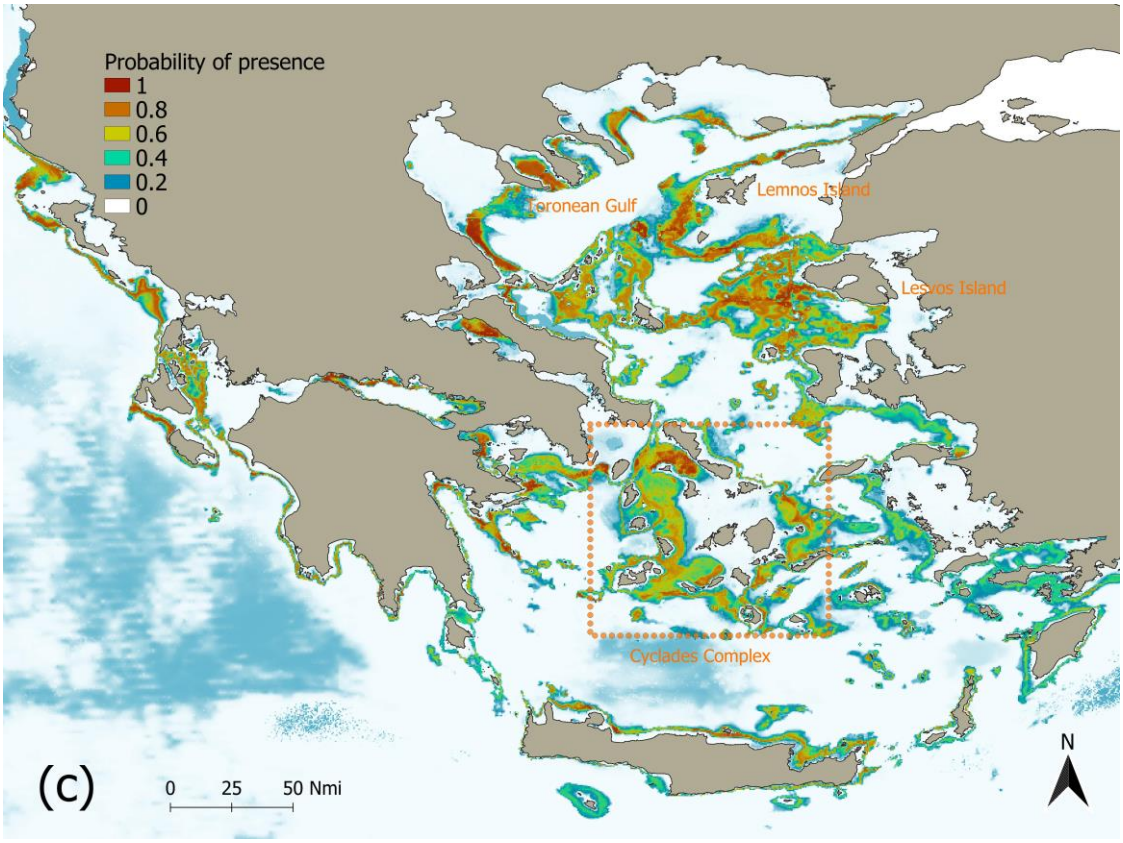


Fig. 6. Mueller's pearlside probability of presence maps for years 2018 (a), 2019 (b) and 2020 (c)

4. Discussion

Species distribution models are powerful tools for understanding the patterns that dictate habitat suitability for marine species. Nonetheless, studies that are based on single-algorithm SDM projections, introduce the risk of uncertainty when choosing a single modeling technique. Herein, the Biomod2 ensemble SDM was successfully applied to model the distribution of *M. muelleri* in the Greek Seas. This approach minimizes the risk of uncertainty and high variability of results through the procedure of weighing each individual SDM based on its performance and then combining all to produce a final algorithm.

Little research on mesopelagic fish has taken place in the Mediterranean Sea, due to the lack of commercial use for this group of fish and the demanding sampling required (e.g. owing to the deep-sea environment and offshore surveys needed). Recent studies have shown that mesopelagic fish -and *M. muelleri* in particular- are major zooplankton consumers being at the same time possible prey for commercial fish such as tunas and charismatic species such as dolphins [18,17,16,19]. These, combined with the possibility of them

being used as fisheries resource, has increased the scientific interest around these groups of fishes. *M. muelleri*, being one of the most abundant mesopelagic fish in the Mediterranean and the Greek seas (Kapelonis *et al*., 2023), is a great candidate for contributing to the knowledge of the dynamics of mesopelagic fish assemblages. In addition, the distinct acoustic typology and the monospecific character of the *M. muelleri* aggregations allow for modelling the species distribution with high degree of certainty. The Greek seas, characterized by high variability in topography and environmental parameters, with examples as the deep Ionian Sea and the topographically complex Aegean Sea archipelagos, the highly productive Northern Aegean and the oligotrophic South Aegean, as well as the existence of deep Gulfs with varying characteristics (e.g. with steeper or smoother slopes), serve as a great study area for SDM applications.

Overall SDM performance in the current study was high, with eight out of ten SDMs being above the evaluation metric thresholds (TSS>0.7, ROC>0.8), even with the relatively strict thresholds that were set [45,46]. The large number of SDMs with high performance used to compose the final ensemble algorithm, likely contributes to the exceptional performance of the latter (TSS=0.88, ROC=0.99). Bottom depth is known to affect mesopelagic fish distribution [47,48]. From the results of the current work, bottom depth appeared as the key predictor for habitat suitability with the highest mean variable importance (0.83). The importance of bottom depth as a predictor can be explained by visualizing the monospecific aggregations of the species. They occupy a bathymetrically niche area (~150-250 m), that is located shallower in the water column than the typical DSLs and often in stronger association with the seabed. *M. muelleri* schools in the Eastern Mediterranean don't seem to overlap extensively with other mesopelagic fish aggregations; this behavior can potentially be attributed to a strategy towards reducing interspecific competition by avoiding the competition with other mesopelagic fishes in the same area. In addition, juvenile fish occupying areas closer to the surface than larger individuals, is a behavior that is potentially associated with the avoidance of intraspecific competition. *M. muelleri* shoals appear to have similar behavior in different parts of the world, they are found in the same depth range (~150-200 m) during daytime in the North Atlantic Ocean [49], with juveniles found shallower than larger individuals [49]. They, occupy areas near the continental shelf, in the North (Gjøsæter *et al*., 1981) and South Atlantic ocean [51]. Furthermore, *M. muelleri* is the most abundant fish species of the upper mesopelagic zone in the Norwegian Fjords (Gjøsæter *et al*., 1981). As a species that occupies the part of the mesopelagic zone that provides the highest amount of light possible [13,1], it's position inside the water column is affected by the light [13]. Moreover, *M. muelleri* shoals usually appear to be in close distance to the bottom and the continental shelf break [52]. The combination of the latter two observations can potentially explain the bathymetrically niche area that *M. muelleri* occupies.

CHL and UADT mean variable importance (0.12 and 0.11 respectively) seem low compared to depth, however due to the SDM-weighting nature of the ensemble algorithm and significantly higher importance (>0.2) in SDMs that presented high TSS scores (CTA: 0.86, RF: 0.95), those variables had a larger contribution to the projection than initially expected. CHL concentration being a proxy for productivity, has been shown to be one of the most important variables for species habitat suitability and biomass prediction of small pelagics [53] and mesopelagics (Irigoien *et al*., 2014). UADT, a variable associated with sea current movement on the North-South axis was highlighted as an important predictor for Atlantic mackerel habitat suitability in the Mediterranean Sea (Giannoulaki *et al*., 2017). The relatively low variable importance of UADT and CHL, is possibly associated with the usage of presence/absence data in the SDM runs instead of abundance data. It is probable that a modelling approach that uses abundance data as dependent variable would highlight more the effect of variables related to circulation, productivity and/or food availability because such variables and their gradient are expected to affect more population density; removing the factor of abundance and the habitat's capacity to support fish populations, potentially limits the impact of CHL and UADT on individual SDM runs. Slope was initially expected to have higher contribution to the ensemble SDM algorithm formation, due to *M. muelleri* echotypes appearing in multiple locations close to the shelf break (Fig. 4), where bottom slope gets its higher values. Those expectations were not met, possibly because the shelf break in the study area is often located between depths of 150-250 m that are ideal for *M. muelleri* presence according to the ensemble SDM algorithm.

The annual projected maps (2018-2020) presented little variance in the probability of presence as the dynamic environmental parameters had lower importance in contrast with the non-dynamic (Depth). Little variability was identified only in areas with high probability of presence, with values varying from 0.8 to 1 between years 2018 and 2020 (Fig. 6). Sea surface temperature (SST) and sea bottom temperature (SBT), were amongst the variables that had minimal impact in all the models and were excluded. CHL concentration, is known to be influenced by changes in sea temperature [54,55], mostly due to vertical mixing within the water column. In addition, climate-related changes in precipitation and river runoffs can affect primary productivity, with contrasting patterns forecasted in the Eastern Mediterranean (e.g. increasing: Macias *et al*., 2015; Richon *et al*., 2019). Thus, climate change can possibly affect the habitat of *M. muelleri* due to changes in CHL concentration.

The areas with high probability of presence for *M. muelleri* , projected by the current study, overlap to a great extent with fishing grounds for commercial species in the Greek Seas [58]. *M. muelleri* potential distribution in the North and Central Aegean Sea presented extremely high resemblance to the fishing grounds for the Blackmouth catshark (*Galeus melastomus*) [58]. Fishing grounds for other commercial species (*Trachurus trachurus, Lophius budegassa, Merluccius merluccius, Eledone cirrhosa, Illex coindenti*) also overlapped to a

great extent with *M. muelleri* spatial distribution [58], mostly in the North and Central Aegean Sea, the Saronic Gulf and the Inner Ionian Archipelago. The habitats of the Common dolphin (*Delphinus delphis*) and Bottlenose dolphin (*Tursiops truncates*) during the early summer period [59], also cover a great part of *M. muelleri* spatial distribution in the North-East Aegean, Inner Ionian Archipelago and the Cyclades Island Complex. There is a high probability that several of those commercial and charismatic species have strong interactions and possibly feed on *M. muelleri* individuals, setting *M. muelleri* an important part of the marine food webs in the Greek and the Eastern Mediterranean Sea ecosystems.

The current study provides valuable knowledge on *M. muelleri* spatial distribution, behavior, dynamics between the environment and the species' potential habitat and important information on the ecology of the mesopelagic zone in the Greek seas. Nevertheless, more aspects of *M. muelleri* and mesopelagic zone dynamics should be explored. Abundance data could be used in habitat modeling to explore the potential of other environmental parameters that can get overlooked when using presence/absence data. Data from potential mesopelagic-boundary community [60] DSLs that seem to contain *M. muelleri* mixed with other mesopelagic species should be better studied and included in SDMs alongside with monospecific aggregation data, to provide additional insight on the species distribution. Future projections of environmental data can be used in SDM to provide information on the magnitude of climate change's impact on *M. muelleri* habitat extent. Finally, the current approach could be applied to additional echo-traces from the twilight zone (e.g. DSLs at different depth zones) to increase our knowledge on other mesopelagic fishes and the ecology of the mesopelagic zone in general.

Acknowledgements

This endeavor would not have been possible without the invaluable assistance and guidance of my supervisor Dr. Konstantinos Tsagkarakis, I will forever be deeply indebted to him. I would like to extend my deepest appreciation to the other two members of my examination committee, Dr. Dionysios Raitsos-Exarchopoulos and Dr. Marianna Giannoulaki for generously providing their knowledge and feedback throughout the long process of conducting this study. Also, I am extremely grateful to Zacharias Kapelonis and Vasilis Valavanis for sharing their highly valuable expertise in numerous occasions. Lastly, I would like to mention that this study was supported by the Hellenic Centre for Marine Research (H.C.M.R.)

References

1. Kaartvedt, S., Langbehn, T. J. & Aksnes, D. L. Enlightening the ocean's twilight zone. *ICES J. Mar. Sci.* 76, 803–812 (2019).

2. Behrenfeld, M. J. Abandoning Sverdrup's Critical Depth Hypothesis on phytoplankton blooms. *Ecology* 91, 977–989 (2010).

3. de Busserolles, F. & Marshall, N. J. Seeing in the deep-sea: visual adaptations in lanternfishes. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160070 (2017).

4. Catul, V., Gauns, M. & Karuppasamy, P. K. A review on mesopelagic fishes belonging to family Myctophidae. *Rev. Fish Biol. Fish.* 21, 339–354 (2011).

5. Gjøsaeter, J., Kawaguchi, K. & Nations, F. and A. O. of the U. *A Review of the World Resources of Mesopelagic Fish*. (Food & Agriculture Org., 1980).

6. Irigoien, X. *et al.* Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nat. Commun.* 5, 3271 (2014).

7. John K. Horne. Acoustic approaches to remote species identification: a review. *Fish Ocean.* (2000).

8. Elith, J. *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151 (2006).

9. Grenouillet, G. Ensemble modelling of species distribution: the effects of geographical and environmental ranges.

10. Ramirez-Reyes, C. *et al.* Embracing Ensemble Species Distribution Models to Inform At-Risk Species Status Assessments. *J. Fish Wildl. Manag.* 12, 98–111 (2021).

11. Olivar, M. P. *et al.* Vertical distribution, diversity and assemblages of mesopelagic fishes in the western Mediterranean. *Deep Sea Res. Part Oceanogr. Res. Pap.* 62, 53–69 (2012).

12. Rees, D. J., Poulsen, J. Y., Sutton, T. T., Costa, P. A. S. & Landaeta, M. F. Global phylogeography suggests extensive eucosmopolitanism in Mesopelagic Fishes (Maurolicus: Sternoptychidae). *Sci. Rep.* 10, 20544 (2020).

13. Staby, A. & Aksnes, D. L. Follow the light—diurnal and seasonal variations in vertical distribution of the mesopelagic fish Maurolicus muelleri. *Mar. Ecol. Prog. Ser.* 422, 265–273 (2011).

14. Prellezo, R. Exploring the economic viability of a mesopelagic fishery in the Bay of Biscay. *ICES J. Mar. Sci.* 76, 771–779 (2019).

15. Vastenhoud, B. M. J., Bastardie, F., Andersen, K. H., Speirs, D. C. & Nielsen, J. R. Economic viability of a large vessel mesopelagic fishery under ecological uncertainty. *Front. Mar. Sci.* 10, (2023).

16. Olafsdottir, D., MacKenzie, B. R., Chosson-P, V. & Ingimundardottir, T. Dietary Evidence of Mesopelagic and Pelagic Foraging by Atlantic Bluefin Tuna (Thunnus thynnus L.) during Autumn Migrations to the Iceland Basin. *Front. Mar. Sci.* 3, (2016).

17. Falautano, M., Castriota, L., Finoia, M. G. & Andaloro, F. Feeding ecology of little tunny Euthynnus alletteratus in the central Mediterranean Sea. *J. Mar. Biol. Assoc. U. K.* 87, 999–1005 (2007).

18. Giske, J. *et al.* Vertical distribution and trophic interactions of zooplankton and fish in Masfjorden, Norway. *Sarsia* (1990).

19. Giménez, J. *et al.* Feeding ecology of Mediterranean common dolphins: The importance of mesopelagic fish in the diet of an endangered subpopulation. *Mar. Mammal Sci.* 34, 136–154 (2018).

20. Kapelonis, Z. Seasonal patterns in the mesopelagic fish community and associated deep scattering layers of an enclosed deep basin.

21. Lykousis, V. *et al.* Major outputs of the recent multidisciplinary biogeochemical researches undertaken in the Aegean Sea. *J. Mar. Syst.* 33–34, 313–334 (2002).

22. Karageorgis, A. P., Nikolaidis, N. P., Karamanos, H. & Skoulikidis, N. Water and sediment quality assessment of the Axios River and its coastal environment. *Cont. Shelf Res.* 23, 1929–1944 (2003).

23. Tsiaras, K. P. *et al.* Inter-annual productivity variability in the North Aegean Sea: Influence of thermohaline circulation during the Eastern Mediterranean Transient. *J. Mar. Syst.* 96–97, 72–81 (2012).

24. Tsiaras, K. P., Petihakis, G., Kourafalou, V. H. & Triantafyllou, G. Impact of the river nutrient load variability on the North Aegean ecosystem functioning over the last decades. *J. Sea Res.* 86, 97–109 (2014).

25. Georgopoulos, D. *et al.* Hydrology and circulation in the Southern Cretan Sea during the CINCS experiment (May 1994–September 1995). *Prog. Oceanogr.* 46, 89–112 (2000).

26. Psarra, S., Tselepides, A. & Ignatiades, L. Primary productivity in the oligotrophic Cretan Sea (NE Mediterranean): seasonal and interannual variability. *Prog. Oceanogr.* 46, 187–204 (2000).

27. Galanopoulou, S., Vgenopoulos, A. & Conispoliatis, N. Anthropogenic Heavy Metal Pollution in the Surficial Sediments of the Keratsini Harbor, Saronikos Gulf, Greece. *Water. Air. Soil Pollut.* 202, 121–130 (2009).

28. Tsangaris, C. *et al.* Impact of dredged urban river sediment on a Saronikos Gulf dumping site (Eastern Mediterranean): sediment toxicity, contaminant levels, and biomarkers in caged mussels. *Environ. Sci. Pollut. Res.* 21, 6146–6161 (2014).

29. Papantoniou, G., Zervoudaki, S., Assimakopoulou, G., Stoumboudi, M. Th. & Tsagarakis, K. Ecosystem-level responses to multiple stressors using a time-dynamic food-web model: The case of a re-oligotrophicated coastal embayment (Saronikos Gulf, E Mediterranean). *Sci. Total Environ.* 903, 165882 (2023).

30. Emig, C. C. & Geistdoerfer, P. The Mediterranean deep-sea fauna: historical evolution, bathymetric variations and geographical changes. *Carnets Géologie Noteb. Geol.* (2004) doi:10.4267/2042/3230.

31. Pitta, P., Karakassis, I., Tsapakis, M. & Zivanovic, S. Natural vs. mariculture induced variability in nutrients and plankton in the eastern Mediterranean. *Hydrobiologia* 391, 179–192 (1998).

32. Gkikopoulou, K. C. Distribution and abundance estimation of sperm whales (Physeter macrocephalus) along the Hellenic Trench in eastern Mediterranean. (University of St Andrews, 2013).

33. Frantzis, A., Alexiadou, P. & Gkikopoulou, K. C. Sperm whale occurrence, site fidelity and population structure along the Hellenic Trench (Greece, Mediterranean Sea). *Aquat. Conserv. Mar. Freshw. Ecosyst.* 24, 83–102 (2014).

34. Thompson, K. F., Webber, T., Karantzas, L., Gordon, J. & Frantzis, A. Summer and winter surveys of deep waters of the Hellenic Trench, Greece, provide insights into the spatial and temporal distribution of odontocetes. *Endanger. Species Res.* 52, 163–176 (2023).

35. Bearzi, G., Bonizzoni, S., Agazzi, S., Gonzalvo, J. & Currey, R. J. C. Striped dolphins and short-beaked common dolphins in the Gulf of Corinth, Greece: Abundance estimates from dorsal fin photographs. *Mar. Mammal Sci.* 27, E165–E184 (2011).

36. Bonizzoni, S. *et al.* Modelling dolphin distribution within an Important Marine Mammal Area in Greece to support spatial management planning. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 29, 1665–1680 (2019).

37. Leonori, I. *et al.* The history of hydroacoustic surveys on small pelagic fishes in the European Mediterranean Sea. (2021) doi:10.12681/mms.26001.

38. Demer, D. A.; Berger, L.; Bernasconi, M.; Bethke, E.; Boswell, K.; Chu, D.; Domokos, R. Et Al. Calibration of acoustic instruments. (2015) doi:10.17895/ICES.PUB.5494.

39. Breiner, F. T., Guisan, A., Bergamini, A. & Nobis, M. P. Overcoming limitations of modelling rare species by using ensembles of small models. *Methods Ecol. Evol.* 6, 1210–1218 (2015).

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

41. Hong, H., Miao, Y., Liu, J. & Zhu, A.-X. Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping. *CATENA* 176, 45–64 (2019).

42. Thuiller, W. BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.* 9, 1353–1362 (2003).

43. Thuiller, W., Georges, D. & Engler, R. biomod2: Ensemble platform for species distribution modelling. 2, (2014).

44. Komac, B., Esteban, P., Trapero, L. & Caritg, R. Modelization of the Current and Future Habitat Suitability of Rhododendron ferrugineum Using Potential Snow Accumulation. *PLOS ONE* 11, e0147324 (2016).

45. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43, 1223–1232 (2006).

46. Beaumont, L. J. *et al.* Which species distribution models are more (or less) likely to project broad-scale, climate-induced shifts in species ranges? *Ecol. Model.* 342, 135–146 (2016).

47. Ross, S. W., Quattrini, A. M., Roa-Varón, A. Y. & McClain, J. P. Species composition and distributions of mesopelagic fishes over the slope of the north-central Gulf of Mexico. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 57, 1926–1956 (2010).δ

48. Sutton, T. T. *et al.* A global biogeographic classification of the mesopelagic zone. *Deep Sea Res. Part Oceanogr. Res. Pap.* 126, 85–102 (2017).

49. Staby, A., Røstad, A. & Kaartvedt, S. Long-term acoustical observations of the mesopelagic fish Maurolicus muelleri reveal novel and varied vertical migration patterns. *Mar. Ecol. Prog. Ser.* 441, 241–255 (2011).

50. Gbjgseter, J. A. K. LIFE HISTORY AND ECOLOGY OF MAUROLICUS MUEL- LERI (GONOSTOMATIDAE) IN NORWEGIAN WATERS.

51. Armstrong, M. J. & Prosch, R. M. Abundance and distribution of the mesopelagic fish Maurolicus muelleri in the southern Benguela system. *South Afr. J. Mar. Sci.* 10, 13–28 (1991).

52. Kapelonis, Z. *et al.* Seasonal patterns in the mesopelagic fish community and associated deep scattering layers of an enclosed deep basin. *Sci. Rep.* 13, 17890 (2023).

53. Giannoulaki, M. *et al.* Characterizing the potential habitat of European anchovy *Engraulis encrasicolus* in the Mediterranean Sea, at different life stages: *Habitat of anchovy in the Mediterranean*. *Fish. Oceanogr.* 22, 69–89 (2013).

54. Raitsos, D. E., Pradhan, Y., Brewin, R. J. W., Stenchikov, G. & Hoteit, I. Remote Sensing the Phytoplankton Seasonal Succession of the Red Sea. *PLOS ONE* 8, e64909 (2013).

55. Dunstan, P. K. *et al.* Global patterns of change and variation in sea surface temperature and chlorophyll a. *Sci. Rep.* 8, 14624 (2018).

56. Macias, D. M., Garcia-Gorriz, E. & Stips, A. Productivity changes in the Mediterranean Sea for the twenty-first century in response to changes in the regional atmospheric forcing. *Front. Mar. Sci.* 2, (2015).

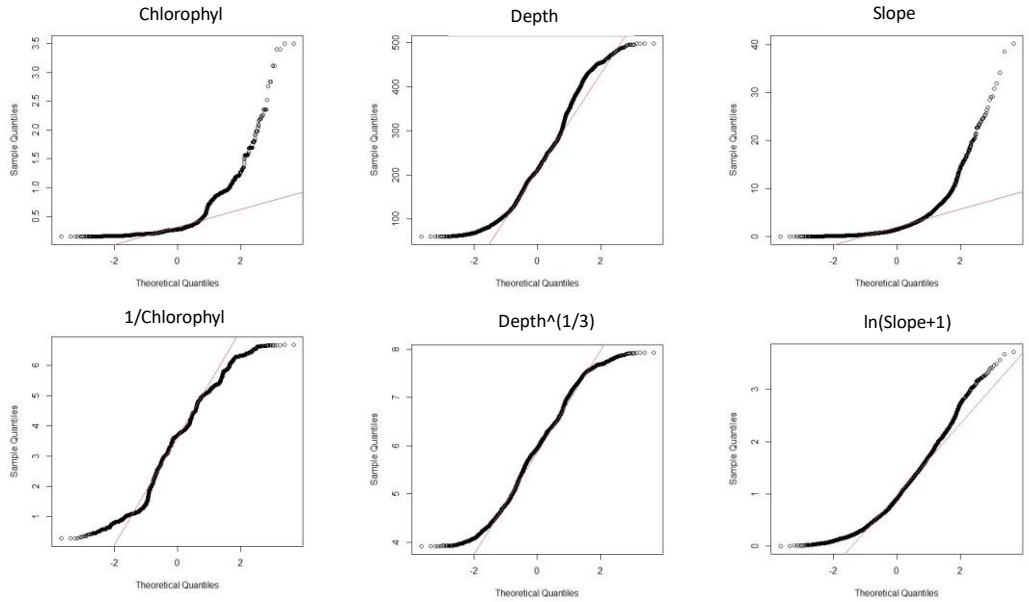
57. Richon, C. *et al.* Biogeochemical response of the Mediterranean Sea to the transient SRES-A2 climate change scenario. *Biogeosciences* 16, 135–165 (2019).

58. Maina, I. *et al.* A methodological approach to identify fishing grounds: A case study on Greek trawlers. *Fish. Res.* 183, 326–339 (2016).

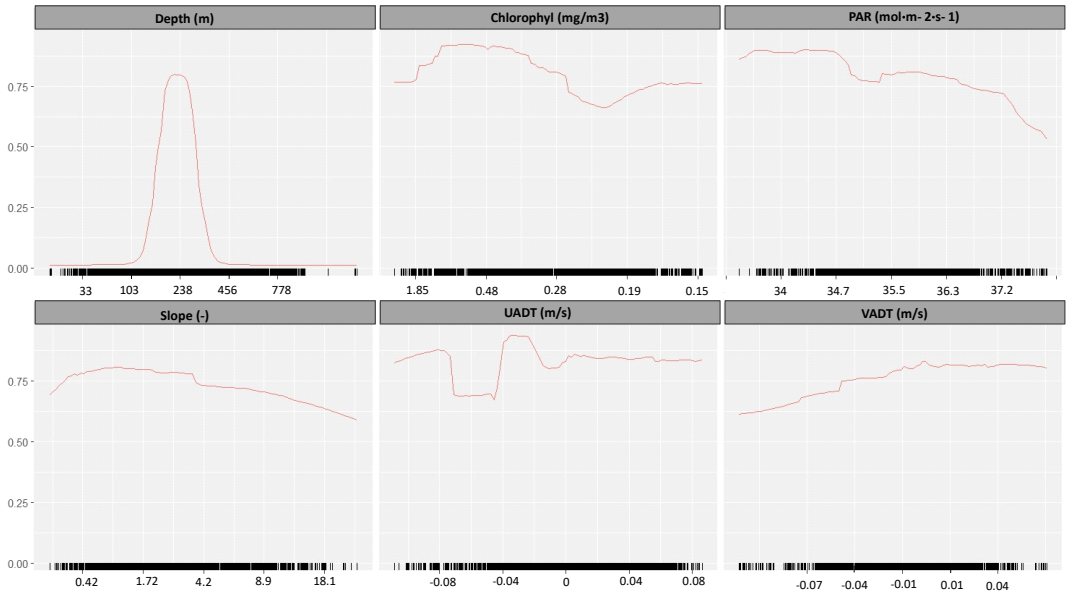
59. Giannoulaki, M. *et al.* Linking small pelagic fish and cetacean distribution to model suitable habitat for coastal dolphin species, Delphinus delphis and Tursiops truncatus, in the Greek Seas (Eastern Mediterranean). *Aquat. Conserv. Mar. Freshw. Ecosyst.* 27, 436–451 (2017).

60. Reid, S. B., Hirota, J., Young, R. E. & Hallacher, L. E. Mesopelagic-boundary community in Hawaii: Micronekton at the interface between neritic and oceanic ecosystems. *Mar. Biol.* 109, 427–440 (1991).

Supplementary material



Sup. Fig 1. Qq-norm plots, comparing the data distribution of the three environmental parameters that got transformed with the normal distribution (red line). Up: Before tranformations. Down: After tranformations



Sup. Fig 2. Response curve plots for the environmental parameters used in Biomod2 ensemble algorithm