

FAIRSEA (ID 10046951)

“Fisheries in the Adriatic Region - a Shared Ecosystem Approach”

D.4.8.1 First analysis of layers of information

Work Package:	WP4 - Implementation of a shared and integrated platform Activity 4.8 Development of an integrated platform for an Ecosystem Approach to Fisheries
Type of Document	Analysis of the outputs of HYDRO, BGC, FSTAT, BSTAT, and EFFORT for defining a region of overlap, possible areas of management, comparison with areas of competence of regional authorities is performed.
Use	Internal
Responsible PP	LP-OGS and PP3-CNR
Authors	Libralato S., Cossarini G. [LP-OGS] Panzeri D. [LP-OGS & University of Trieste] Masnadi F., Scarcella G. [PP3-CNR] Spedicato M.T., Zupa W., Bitetto I. [PP5-COISPA] Vrgoč N., Isajlović I. [PP1-IOF] Carlucci R., Cipriano G., Manfredi C., Piccinetti C. [PP6-CONISMA] Krstulović Šifner S. [PP10-UNIST] External service contribution: Russo T., D’Andrea L. [Università Tor Vergata]
Version and date	V.02, 30/06/2021

Deliverable 4.8.1

First analysis of layers of information

FAIRSEA – Fisheries in the Adriatic Region – a shared Ecosystem Approach

FAIRSEA is financed by Interreg V-A IT-HR CBC Programme (Priority Axis 1 – Blue innovation)

Start date: 01 January 2019

End date: 28 February 2021 (extended to: 31 August 2021)

Acronyms used	4
Executive Summary.....	5
The Project FAIRSEA.....	6
1. INTRODUCTION.....	7
2. Objectives of the First analysis of layers of information	9
3. Layers of information uploaded in the Integrated platform	10
3.1 The FAIRSEA Integrated Platform structure and software	10
3.2 Modules details.....	10
4. MATERIALS.....	14
4.1 Study area	14
4.2 Inputs from BSTAT (Activity 4.3)	16
4.4 Inputs from HYDRO and BGC modules (Activities 4.1 and 4.2)	16
4.5 Inputs from EFFORT module (Activity 4.5).....	17
5. Integrating layers using Species Distribution Models (SDM).....	17
5.1 Stepwise approach, training and testing of SDM	18
5.1 Stepwise approach, training and testing of SDM	19
5.2 Predicted errors	24
5.3 Identification of hot spots on results of SDM.....	27
6. Using results for management insights: matching hot spots	29
6.1 Capabilities and Limitations.....	31
7 Final remarks.....	32
7.1 Notes.....	33
8. REFERENCES	34

Acronyms used

AIS	Automatic Informative System
CFP	Common Fisheries Policy
EAF	Ecosystem Approach to Fisheries
EAFM	Ecosystem Approach to Fisheries Management
FAIRSEA	Fisheries in the Adriatic Region – a Shared Ecosystem Approach
GAM	Generalized Additive Models
GIS	Geographic Information System
GSA	FAO Geographical Sub Areas
GVC	Generalized Cross-Validation
IT	Italy
LOA	Length OverAll
LP	Lead Partner
MEDITS	International Bottom Trawl Survey in the Mediterranean
OGS	Istituto Nazionale di Oceanografia e di Geofisica Sperimentale
PP	Project Partner
RCP	Representative Concentration Pathway
SOLEMON	Rapido trawl surveys in the Northern Adriatic Sea
VMS	Vessel Monitoring System
WP	Work packages

Executive Summary

This activity presents the application of analyses to the layers of information collected in the Integrated platform of FAIRSEA for the detection of hot spots of aggregation of target species as a preliminary basis for ecological identification of fisheries management areas.

A set of different Spatial Distribution Models (SDM; Gaussian GAM, Tweedie GAM, Delta-GAM, Random Forest and Gradient Boosting Machine) was applied to detect aggregations of juveniles and adults of 6 demersal species in the Adriatic-Ionian area (GSA17, 18 and 19). The SDM ensemble was applied to abundance indices (N/km²) of European hake (*Merluccius merluccius*), red mullet (*Mullus barbatus*) and Norway lobster (*Nephrops norvegicus*), from MEDITS trawl surveys in the three GSAs, and abundance indices of common sole (*Solea solea*), mantis shrimp (*Squilla mantis*) and common cuttlefish (*Sepia officinalis*) from SOLEMON beam trawl survey in GSA17. The ensemble of SDMs was implemented using spatio-temporal variables (e.g., latitude, longitude, depth) of trawl survey data from module BSTAT (Act. 4.4) and including additional environmental covariates such as data obtained from modules HYDRO (Act. 4.1), BGC (Act. 4.2) and fishing effort patterns estimated from VMS data (module EFFORT, Act. 4.5).

The Getis Index was applied to the results of SDMs and used to identify the hot-spots for juveniles and adults of each species in the period 2008-2018 in the Adriatic-Ionian basin. The ensemble of SDMs allowed identifying also areas of overlapping hot-spots among species that, when occurring, were also located. Results point out that environmental covariates improve representation of the past data, and permit to project possible effects of future climate change scenarios. Results allow to preliminary identify areas for possible management measures that will be tested in other activities of the Interreg IT-HR Project FAIRSEA.

The Project FAIRSEA

The FAIRSEA is a European Territory Cooperation project financed under the priority 1 “Blue innovation”, Specific Objective 1.1 “Enhance the framework conditions for innovation in the relevant sectors of the blue economy within the cooperation area” of the INTERREG V-A Italy–Croatia Programme 2014-2020. The project focuses on the fisheries sector, key driver for the blue growth of the Adriatic communities, towards a sustainable co-management of resources and marine ecosystem protection. The transboundary nature of marine resources requires a cross-border cooperation and a shared “Vision” to properly tackle and address the different socio-economic and environmental challenges related to fisheries activities management. In this context, FAIRSEA Project aims at enhancing transnational capacity and cooperation in order to promote the sharing of knowledge and good practices between regional and transnational key actors in the sector of sustainable fisheries management in the Adriatic Sea as well as to implement innovative approaches adopting an ecosystem approach to fisheries (EAF).

Coordinated by the OGS of Trieste (IT), the project involves a consortium of 12 strategic and operational partners from Italy and Croatia that will make to best use of their complementary expertise to address and support the application of the EAF ensuring a strong and interactive engagement of institutional, technical and socio-economic stakeholder in project activities.

The main result of the FAIRSEA Project will be the development of an integrated platform for a quantitative ecosystem approach to fisheries that goes across territorial boundaries and across several disciplines. The platform will integrate biological/ecological processes (i.e. considering water mass circulation, physical-chemical properties, plankton productivity, dynamics of resources including their interactions) and fisheries bio-economic dynamics (including fisheries displacement). This high technological and innovative platform will be used as a planning tool to implement demonstrative testing of applicable fisheries policies both at local (subareas) and Adriatic scales. It will provide scientific basis for formulating and evaluating the shared management advice in the local and international participatory processes, involving management authorities, experts and

stakeholders. The Project will also provide an answer to the need of reference points, best practices and guidelines for the optimisation between ecological and socio-economical sustainability of fisheries in the Adriatic Sea.

1. INTRODUCTION

Marine fish and invertebrates that live and feed close to the marine seabed, i.e., the demersal species, play a fundamental role in fisheries in the Adriatic and Ionian Sea (FAO, 2018). In order to ensure the sustainability of exploitation, a set of fisheries management measures and restrictions are adopted also considering scientific information on the status of resources. Clearly, management actions are particularly relevant and impacting in areas where demersal resources play a central role in local fishing communities and economies, such as the Adriatic and Ionian seas. Therefore, it is of paramount importance to increase accuracy of scientific information used to inform management.

Scientific bottom trawl surveys provide quantification of abundance and biomass (hereafter termed indices) by species as fully described in module BSTAT (Activity 4.3 BSTAT of the WP4). These fishery-independent data are used for manifold purposes related to management: stock assessment (e.g., Cotter et al., 2009), evaluation of spatio-temporal distribution of demersal resources (e.g., Carlucci et al., 2009), estimates of population and community densities (e.g., Spedicato et al., 2019b), and the development of ecosystem models (e.g., Gruss et al., 2018; Moullec et al., 2019). Sampling protocols of multiannual surveys are usually standardized for sampling design, gear geometry, sampling season, sampling locations to allow comparability of the trawl survey data across space and time. However, unavoidable small deviances (e.g., sampling period) or changes (e.g., vessel) during sampling may affect the abundance and biomass indices obtained from trawl surveys.

The analysis would like to test the potential benefits on integrating oceanographic variables from module HYDRO (Act. 4.1), BGC (Act. 4.2) and effort variables from module EFFORT (Act. 4.5) in addition to spatiotemporal covariates (latitude,

longitude, depth, year and month) to improve species distribution models based on trawl survey data.

Generalized Additive Models (GAMs), and machine learning models (MLMs), Random forest (RF, Breiman L., 2001) and Gradient Boosting machine, (GBM, Schapire R.E., 2003) were chosen for their wide application and suitability with trawl survey data (Gruss et al, 2014, Lauria et al. 2017; Tserpes et al., 2019). GAMs and MLMs allow to predict species abundance and biomass over the domain (Maunder and Punt, 2004; Rubec et al., 2016; Potts and Rose, 2018, Luan J. et al., 2018) and provide estimates useful for tuning stock assessment models (Orio et al., 2017; Cao et al., 2017). Furthermore, GAMs and MLMs are deemed appropriate for mapping species distribution that is useful in ecosystem models (Fulton et al., 2011; Gruss et al., 2014, Luan et al., 2018), or for identifying Essential Fish Habitats (Druon et al., 2015).

In addition to monitoring deviances, environmental changes and anthropogenic stressors may cause life-history responses, and their impacts on survey estimates are difficult to disentangle. Satellite data are successfully used to provide environmental variables (e.g. sea surface temperature; sea surface chlorophyll concentration) to be included in models to describe the spatial distribution of some pelagic species (Giannoulaki et al., 2008; Schismenou et al., 2017). However, these variables might be insufficient to model the distribution of demersal species, which may require additional oceanographic variables close to seabed such as those provided by the Copernicus Marine Environment Monitoring Service (CMEMS). The relative high number and the quality of the CMEMS products, as well as their high temporal coverage and spatial resolution, provide biogeochemical and physical oceanographic variables that can be useful to improve the analysis of abundance and biomass indices derived from trawl surveys (e.g., Sion et al., 2019; Tserpes et al., 2019).

In addition, the displacement of fishing fleets derived from satellite-based tracking devices, such as Vessel Monitoring System (VMS) and/or Automatic Identification System (AIS), is a valuable source of information on the distribution and spatial aggregation of marine resources (Russo et al., 2018; Bastardie et al., 2014). The yearly

distribution of fisheries, in fact, represents a good track of the distribution of the targeted resource rather than a measure of the direct impact on it (which is a much longer term effect). Thus increasing accuracy of distribution of the species might be gained embedding fishing effort among the explanatory variables.

2. Objectives of the First analysis of layers of information

Analysis of the outputs of HYDRO, BGC, FSTAT, BSTAT, and EFFORT for defining a region of overlap, possible areas of management, comparison with areas of competence of regional authorities is performed.

The information resulting from modules of the WP4 and embedded in the Integrated platform were used in an interdisciplinary and integrative quantitative analysis with the objectives to provide supportive science-based evidences useful for fisheries management.

The integrated approach useful to fisheries management presented here is obtained by combining trawl survey data from BSTAT module, relevant physical and biogeochemical oceanographic variables from modules HYDRO and BGC, respectively. Moreover, fishing effort estimates were considered as a result of the module EFFORT of the integrated platform. Biomass indices of demersal fish from scientific trawl surveys carried out in the Adriatic Sea and in the Western Ionian Sea (Adriatic-Ionian macro-region, EUSAIR, 2014) are analysed with a set of approaches using as explanatory variables the relevant biogeochemical and physical variables from CMEMS products and the distribution of fishing effort from VMS/AIS data. The objective of the study is to contrast models with spatiotemporal variables only and with different sets of additional explanatory variables in order to explore the improvement on estimates of demersal species distribution when environmental variables and effort are included into species distribution models.

3. Layers of information uploaded in the Integrated platform

3.1 The FAIRSEA Integrated Platform structure and software

The software development of the IP is being entrusted to the external software development company INKODE (<https://inkode.it/>).

FAIRSEA IP is a web-GIS application based on open source software. All services are deployed by [Docker](#) containers, main services are:

- Backend: based REST API developed in [Python](#) with on [Django](#), [Django Rest Framework](#) and [GeoDjango](#)
- Frontend: a Single Page Application based on [AngularJS](#) with [Angular Material](#) framework
- Database: [PostgreSQL](#) with [PostGIS](#)
- GIS software: [Geoserver](#)
- Charts and dashboards: [Plotly](#) and [Grafana](#)

Other used libraries and services: [GDAL](#), [scipy](#), [Shapely](#), [netCDF4](#), [Pandas](#), [MapProxy](#), [Pillow](#)

Alpha/testing version of the IP running at <http://fairsea.caspar.inkode.it:8887/>.
View only testing credentials: username “viewer”, password “fairsea2020”.

3.2 Modules details

In Table 1 each individual module and sub-module of the Integrated Platform will be briefly described. All layers uploaded in the platform can be visualized and downloaded in the form of NetCDF, shapefile or zipped folders according to the module of interest.

Modules (Sub-modules) name	Modules description	Characteristics within Integrated platform
<p>4.1 HYDRO – Hydrodynamic circulation and connectivity</p> <ul style="list-style-type: none"> - <i>Connectivity and main circulation patterns in the Adriatic Sea</i> - <i>Scenario analysis of future circulation</i> 	<p>Results on the space-time distribution of temperature, salinity and current velocity at different vertical layers for the past 20 years and in the 21st century under emission scenarios RCP4.5 and RCP8.5</p>	<p>The physical variables can be mapped through different layers that can be activated or not thanks to an interactive legend. Through a drop-down menu the user can choose to view different vertical (depth) and temporal (seasons) levels (Fig. 1).</p>
<p>4.2 BGC – Biogeochemical processes and dynamics</p> <ul style="list-style-type: none"> - <i>Production patterns in the Adriatic Sea</i> - <i>Future scenarios of production patterns</i> 	<p>Results on the space-time distribution of nitrogen, chlorophyll, primary production, plankton biomass and oxygen indicators for the past 20 years and in the 21st century under emission scenarios RCP4.5 and RCP8.5</p>	<p>The biogeochemical variables can be mapped through different layers that can be activated or not thanks to an interactive legend. Through a drop-down menu the user can choose to view different vertical (depth) and temporal (seasons) levels.</p>
<p>4.3 BSTAT – Spatial distribution of marine resources</p> <ul style="list-style-type: none"> - <i>GSA17</i> - <i>GSA18</i> - <i>GSA19</i> 	<p>Results of the application of spatio-temporal analyses, through BioIndex and BioStand routine (https://www.coispa.it), to fishery independent survey (MEDITS, SOLEMON) data of GSA17, GSA18 and GSA19 for the time series 1994-2018.</p>	<p>Plots and data table from BSTAT are divided in <i>Bioindex</i> and <i>BioStand</i> folder. Moreover, spatial distribution of interesting species in the GSA17 can be mapped through different layers.</p>

Modules (Sub-modules) name	Modules description	Characteristics within Integrated platform
4.4 FSTAT – Fisheries production and capacity	Information from time series of catches (i.e. landings and discards), fleet consistency and economic data (e.g. costs, revenues) by fleet segment.	Within the platform the outputs can be visualized through an interactive dashboard. Through a filter menu the user can choose to view different years, GSAs, species, fleet segment levels (Fig. 2).
4.5 EFFORT – Effort distribution and fleet displacement	Results of assessment of the fishing effort distribution for both fishing vessels equipped with tracking devices (i.e. VMS or AIS) and non-equipped vessels (typically the fishing vessel smaller than 12 m)	Within the platform the outputs by gears and LOA can be mapped through different layers that can be activated or not thanks to an interactive legend. Through a drop-down menu the user can choose to view different years.
Summary Module	Interaction workspace between module layers.	The workspace gives the user the possibility of simple calculations on different layers on a regional/county basis (average, sum, min and max value) (Fig. 3).

Table 1 Description of modules in the integrated platform

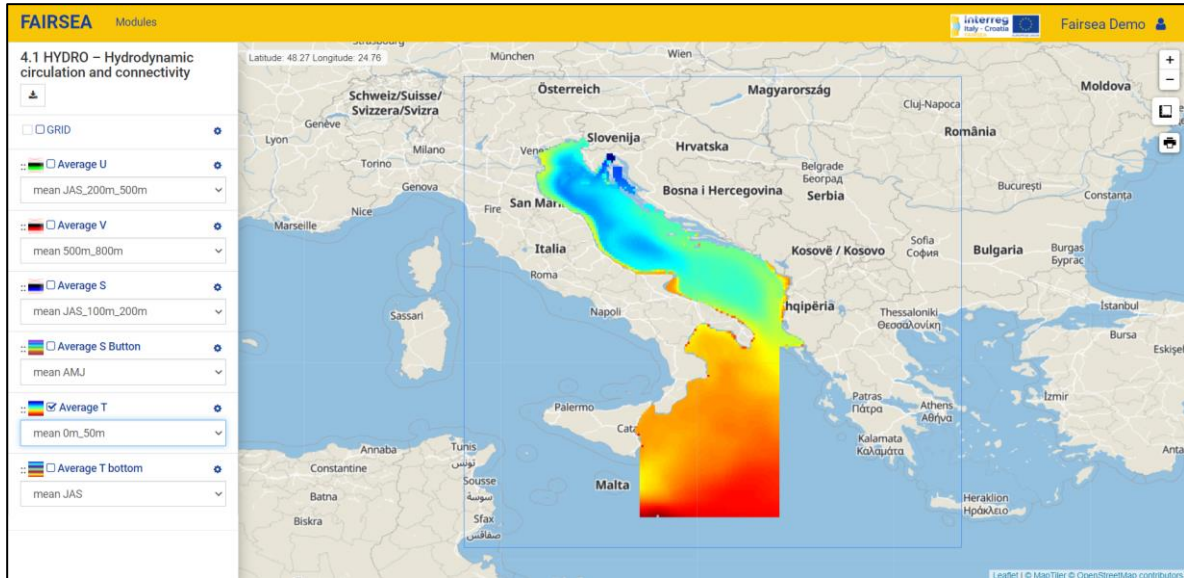


Fig. 1 - HYDRO output layout for “Connectivity and main circulation patterns in the Adriatic Sea” within FAIRSEA IP. This example displays the annual average of temperature values (T) in the project study area at 0-50 m depth (0m-50m).

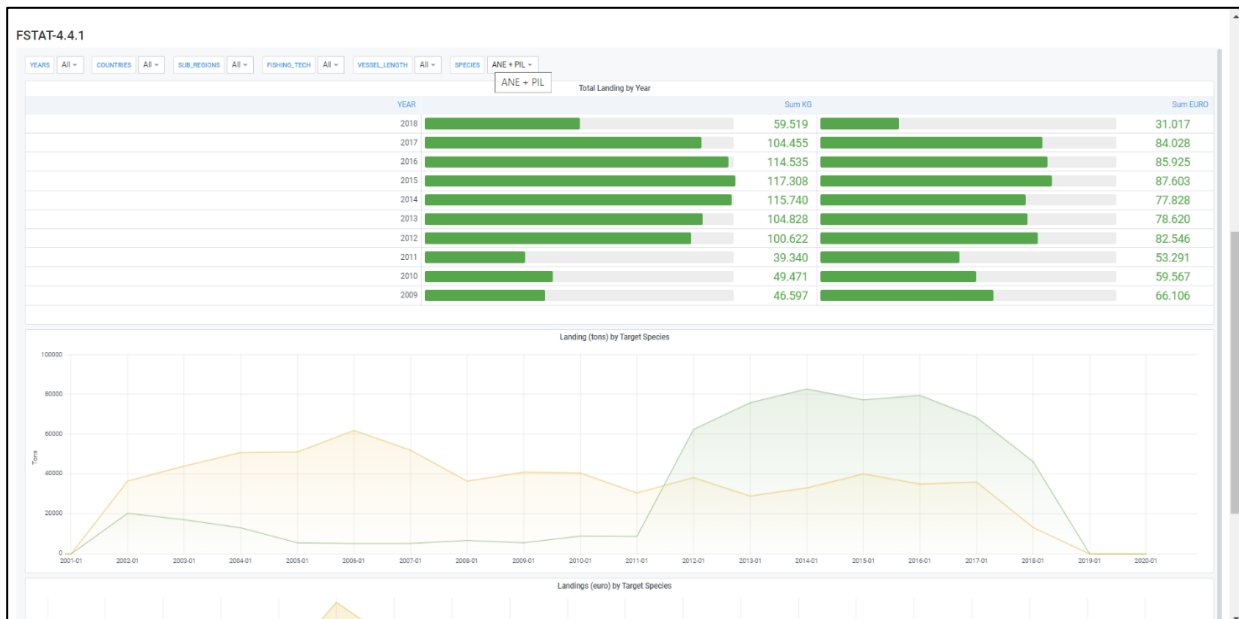


Fig. 2 - Overview of FSTAT interactive dashboard output within FAIRSEA IP. This example displays the landings in tons and Euro of anchovies and sardines (ANE and PIL).



Eledone moschata (ELEDMOS)	Bottom Otter Trawl (OTB) with 18-24
Metrics on N_Km2	Metrics on X2018
Grid filter: Marche	Grid filter: Marche
Average: 4.004161743697478	Average: 825.5096899224806
Min: 0	Min: 0
Sum: 952.9904949999997	Sum: 141987.66666666666
Max: 66.91217	Max: 8332.66666666667

Fig. 3 - Interaction workspace between module layers in the IP. This example displays the spatial distribution and summary calculation on *Eledone moschata* from MEDITS survey and the fishing effort of bottom trawler in the Marche region waters.

4. MATERIALS

4.1 Study area

The study area is the Adriatic Sea and a part of the Ionian Sea, corresponding to the Geographic Sub Areas (GSA) 17, 18 and 19, as defined by the FAO-GFCM (General Fisheries Commission for the Mediterranean Sea). We used the bathymetry range from 10 m to 800 m, depth range of MEDITS survey, and from 10 m to 100 m for

SOLEMON survey, and the latitude and longitude range for the mapping of the species is included between 12'.00" E and 20'.00" E for longitude and 39.5'.00" N and 46'.00" N for latitude (Fig. 4).

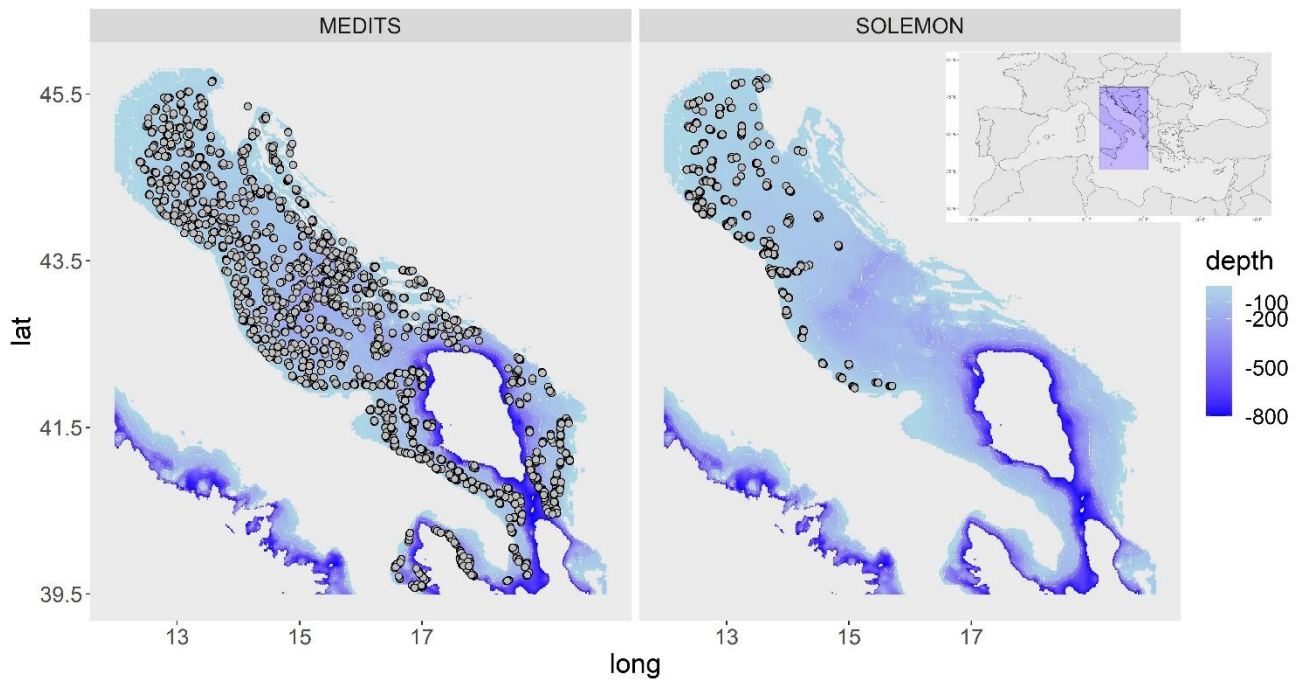


Fig. 4: Haul position (from 2008-2018) for the two surveys in Adriatic-Ionian Sea, MEDITS left panel and SOLEMON right panel.

The central part of the Adriatic basin is characterized by a meso-Adriatic depression, called Pomo pit, reached 260 m, involved in some restricted fisheries recommendation. The southern part of the basin (GSA 18) is characterized by a steep continental slope to deepest bathymetry at 1250 m and its circulation is characterized by a cyclonic gyre. The dynamics of the Southern Adriatic and Ionian Sea are linked through the Bimodal Oscillating System (BiOS) mechanism (Civitaresse, et al. 2010).

4.2 Inputs from BSTAT (Activity 4.3)

We used data from the bottom trawl surveys conducted in the Adriatic Sea and North Western Ionian Sea, i.e., in the geographical sub-areas (GSAs) 17, 18, and 19. MEDITS (Mediterranean International Trawl Survey; Spedicato et al., 2019a) data that consists on average 326 sampling sites per year in the three GSAs and SOLEMON (Sole Monitoring; Scarcella G., 2011, Grati et al., 2013), that consists on average 70 sampling sites per year in GSA 17 (see D4.3.1 “Spatio-temporal distribution of marine species”). We considered only the data from 2008 to 2018 (instead of the entire time series 1994-2018 MEDITS and 2005-2018 SOLEMON) due to the employment of the effort data (see following section).

Indices of demersal species density (n/km^2 , n =number of individual) were retrieved from the MEDITS dataset for European hake (*Merluccius merluccius*, bony fish), red mullet (*Mullus barbatus*, bony fish) and Norway lobster (*Neprhos norvegicus*, crustacean) and from the SOLEMON dataset for common sole (*Solea solea*, bony fish), mantis shrimp (*Squilla mantis*, crustacean) and common cuttlefish (*Sepia officinalis*, mollusc). We divided the species in adult and juvenile by using a specific threshold in length based on biological information.

For each species, we fit indices of individuals km^{-2} by sampling site, set as a response variable in a set of species distribution models (SDM) developed on the basis of different approaches.

4.4 Inputs from HYDRO and BGC modules (Activities 4.1 and 4.2)

Among oceanographic variables available from modules HYDRO and BGC relevant oceanographic variables for the period 1999-2018 were considered based on known ecological importance for chosen demersal species (Carlucci et al., 2018; Bitetto et al., 2019) as well as proxies for productivity and favourable environments. The relevant oceanographic variables considered were the water temperature ($^{\circ}\text{C}$) and dissolved oxygen (mmol m^{-3}) at the sea bottom and surface, water column averages of nitrate and phosphate concentration (mmol/m^3), chlorophyll-a (mg m^{-3}), particulate organic carbon (mg m^{-3}), pH at the sea bottom, and surface salinity.

4.5 Inputs from EFFORT module (Activity 4.5)

We used effort data expressed as trawling time (kw hours⁻¹) per year at the spatial resolution of 1/16° taken from the module EFFORT (D4.5.1 Fishing effort map distribution). These data are estimated from VMS data for the period 2008-2018 (see also Russo et al., 2014a). Total effort data from single bottom otter trawl (OTB) and beam trawls (TBB) applied to species on the basis of the target of fisheries. The effort is based on data from VMS for fleets belonging to both Italy and Croatia (D4.5.1 Fishing effort map distribution).

5. Integrating layers using Species Distribution Models (SDM)

A few different approaches were used to implement SDMs: Generalized Additive Models (GAM), Random forest model (RF) and gradient boosted machine (GBM). In the case of GAM, we applied different distribution families to demonstrate the potential benefits of using additional variables disregarding the model structure. GAMs were developed using Gaussian probability distributions with identity link on trawl survey density data log-transformed for all species. GAMs were also applied using Tweedie probability distributions with a lognormal link on untransformed density indices. Furthermore, the Delta-GAM approach was implemented in two steps: i) a binomial occurrence model was used to fit presence/absence data (binomial family error distribution logit link function), ii) a Gaussian distribution model with identity link function on transformed density for presence-only data (Gruss et al, 2014, Lauria et al. 2017). For Delta-GAM the final spatial distribution of species biomass as n/km² is obtained by multiplication of Gaussian and Binomial models' predictions to the grid of the model's domain (Gruss et al., 2014, Lauria et al. 2017). About random forest model, we applied 5000 trees on data log-transformed for all species, with a minimum number of variables for each split equal to 1/3 of the explained variable (R package randomForest, Breiman L., 2018). Finally, for the GBM model, we applied 10000 trees with shrinkage of 0.01 (gbm package R, <https://github.com/gbm-developers/gbm>).

A grid of regular points with the same resolution of the selected CMEMS product ($1/16^\circ$) and covering the study area was created to predict species density distribution by the selected models (Spedicato et al., 2019b; Tserpes et al., 2019).

In the SDMs, the explanatory variables tested were the geoposition variables (latitude and longitude of the haul survey expressed in decimal degrees) as well as additional oceanographic and fishing effort variables. Among the geoposition variables we used geographic coordinates, depth (m), and year of the observations.

The explanatory variables were preliminarily selected using the VIF approach (Variance Inflation Factor; Sheather, 2009) with a threshold of $VIF < 5$ to avoid collinearity (see also Orio et al., 2017; Sion et al., 2019).

The results of the VIF analysis identified for all the species, year, depth, latitude, longitude, to be included as explanatory variables and constituted the minimal spatiotemporal model (ST). Furthermore, the VIF analysis by species allowed to include additional explanatory variables without collinearity extracted from HYDRO and BGC reanalysis and fishing effort: the VIF results showed to be species-specific.

5.1 Stepwise approach, training and testing of SDM

For each species and all approaches (Delta, Gaussian, Tweedie, RF, and GBM), a forward-stepwise approach was used. This started from the simplest model given by the minimal spatiotemporal explanatory variables (model A, with latitude, longitude, depth, and year) to combine it with all the most meaningful additional biogeochemical, physical, and fishing effort variables (from model B to I), identified by VIF analysis. Then the forward-stepwise approach consisted of increasing the number of explanatory variables by successively adding those with high F statistics till the model with full explanatory variables was obtained. Thus, the forward-stepwise approach resulted in a set of models having different explanatory variables to obtain the response variable ($R = \log(\text{ind km}^{-2})$ or presence/absence). Each model was subjected to a calibration-validation process, thus it was fitted on a spatial training dataset made by randomly choosing 70% of the data used for training the model (calibration) and assessing the fitting model on the remaining 30% of records (testing). The training and testing were repeated using 5 spatially fold (fig. 5) on

datasets randomly selected and without replacement, using the BlockCV package (Valavi et al., 2019). The best model was selected based on measures of the model’s performance such as explained deviance (%ED) and prediction errors (AIC, Akaike Information Criterion) of the training datasets; correlation coefficient (R^2) and mean absolute error (MAE) of the model predictions on the testing dataset.

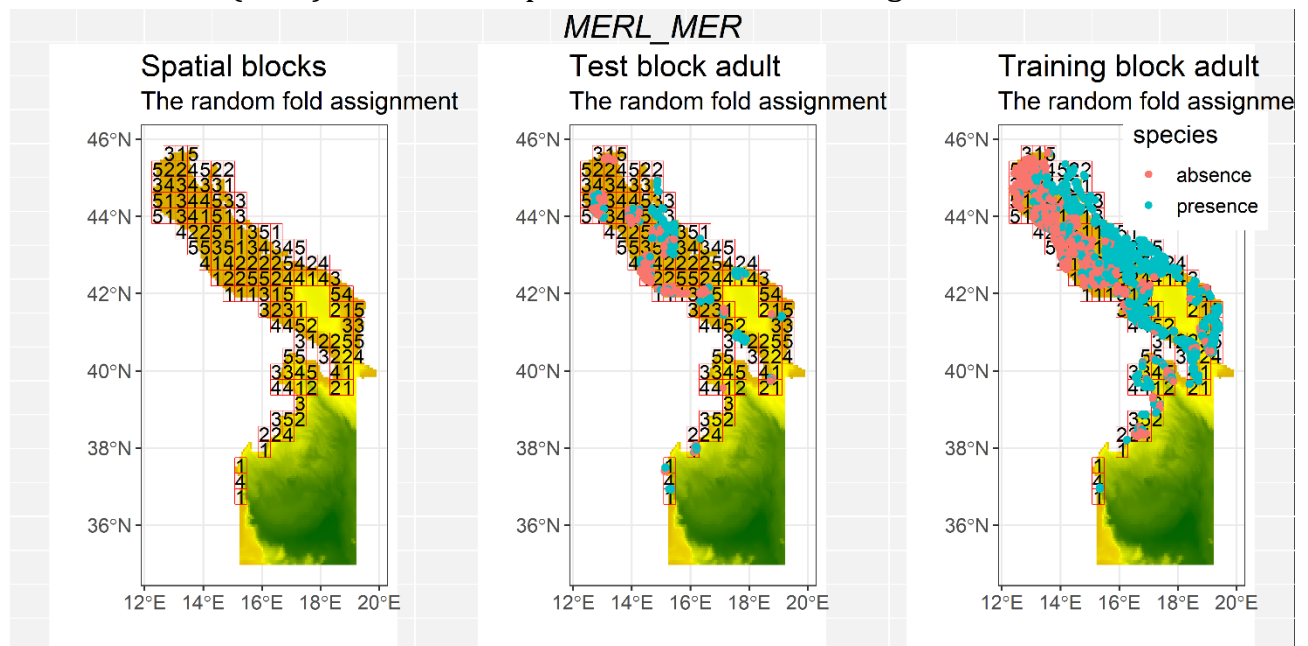


Fig. 5. Spatial identification of blocks used for testing and training the models. Example of Hake (MERL_MER MEDITS code, *Merluccius merluccius*) and spatial blocks used for training and testing the model (from package BlockCV R). Left panel: subdivision of the study area in blocks for 5 fold. Central panel: test blocks for fold 1, Right panel: training plot for the remaining fold: 2,3,4,5.

5.1 Stepwise approach, training and testing of SDM

Furthermore, the validation process includes an application of the best model for each approach and species on the sampled data (MEDITS and SOLEMON) and evaluation of MAE and R^2 (Tables 2,3,4,5). Indicators of performances for the models applied to European hake, red mullet and Norway lobster using MEDITS trawl survey data are reported for adults and juveniles in Tables 2 and 3 respectively. Indicators of performances for the models applied on data from SOLEMON trawl survey data are reported for adults and juveniles in Tables 4 and 5 respectively.

metrics	model	ST	ST+Env	ST	ST+Env	
			Hake (adults)		Red mullet (adults)	
AIC	Delta	2495,45/5795,46	2304,69/5721,45	2258,71/5785,93	2117,566/5721	
	Gaus	15207.98	14824.69	17077.54	16643.26	
	TW	32836.43	32453.31	32553.71	32170.56	
	RF	X	X	X	X	
	GBM	X	X	X	X	
Dev. Expl.	Delta	54.00%	57.00%	53.60%	55.75	
	Gaus	60.60%	64.60%	59.90%	64.60%	
	TW	50.30%	55.80%	61.30%	67.00%	
	RF	81.00%	82.7%	81.35%	82.91%	
	GBM	61.00%	66.55%	62.00%	68.00%	
MAE	Delta	117.53	112.5	601.19	585.67	
	Gaus	163.84	158.06	860.68	793.72	
	TW	145.14	135.53	872.23	785.56	
	RF	86.43	59.88	509.41	371.17	
	GBM	153.16	142.91	849.46	732.92	
			Norway lobster (adults)			
AIC	Delta	2196,49/1857,65	2093,53/1831,79			
	Gaus	12762.13	12560			
	TW	10695.5	10614.76			
	RF	X	X			
	GBM	X	X			
Dev. Expl.	Delta	33.10%	36.80%			
	Gaus	36,9%%	40.80%			
	TW	56.20%	60.40%			
	RF	57.26%	59.00%			
	GBM	43.00%	47.00%			
MAE	Delta	17.58	17.21			
	Gaus	17.46	17.19			
	TW	20.08	19.34			
	RF	13.4	11.06			
	GBM	16.84	16.64			

Table 2. Training results for different models and evaluated using different indicators for adults of MEDITS species, i.e., European hake, Norway lobster and red mullet.

metrics	model	ST	ST+Env	ST	ST+Env
		Hake (juveniles)		Red mullet (juveniles)	
AIC	Delta	1351,755/5315,19	1197,23/5248,15	728,35/1903,25	649,83/1869,85
	Gaus	13215.79	12968.01	10241.99	9565.45
	TW	27137.62	27019.39	10455.56	10169.61
	RF	X	X	X	X
	GBM	X	X	X	X
Dev. Expl.	Delta	57.00%	60.80%	62.00%	65.95%
	Gaus	65.60%	68.30%	56.70%	67.40%
	TW	63.20%	65.00%	79.20%	86.70%
	RF	84.60%	86.71%	83.06%	86.16%
	GBM	67.00%	72.00%	60.00%	70.00%
MAE	Delta	255.32	251.68	808.23	780.57
	Gaus	304.57	292.36	842.89	825.81
	TW	296.67	288.09	1388.89	920.35
	RF	206.73	148.68	724.85	623.08
	GBM	284.39	276.37	844.92	822.69
		Norway lobster (juveniles)			
AIC	Delta	830,36/577,82	794,69/569,95		
	Gaus	7930.14	7898.12		
	TW	5306.69	5283.87		
	RF				
	GBM				
Dev. Expl.	Delta	38.90%	42.45%		
	Gaus	28.40%	29.10%		
	TW	69.00%	72.40%		
	RF	54.86%	56.15%		
	GBM	40.87%	44.62%		
MAE	Delta	4.15	4.03		
	Gaus	4.82	4.82		
	TW	4.69	4.48		
	RF	3.87	3.45		
	GBM	16.84	16.64		

Table 3. Training results for different models and evaluated using different indicators for juveniles of MEDITS species, i.e., European hake, Norway lobster and red mullet.

metrics	model	ST	ST+Env	ST	ST+Env
		Sole (adults)		Mantis shrimp (adults)	
AIC	Delta	271,16/1678,51	270,35/1729,55	322,89/1102,94	366,59/1164,44
	Gaus	2621.06	2654.59	2677.54	2706.25
	TW	9760.86	9764.04	5921.28	5945.32
	RF	X	X	X	X
	GBM	X	X	X	X
Dev. Expl.	Delta	60.95%	56.40%	58.60%	47.75%
	Gaus	65,7%%	63.80%	66.60%	66.00%
	TW	61.20%	60.50%	64.30%	63.20%
	RF	70.60%	71.55%	72.93%	71.08%
	GBM	74.00%	77%%	75.00%	64.00%
MAE	Delta	289.76	300.2	258.4	297.26
	Gaus	322.89	324.41	301.3	309.66
	TW	305.91	306.42	276.6	279.94
	RF	185.69	154.21	197.23	157.61
	GBM	279.7	286.51	275.23	335.52
		Cuttlefish (adults)			
AIC	Delta	451,37/1078,10	458,17/1086,47		
	Gaus	2583.97	2632.71		
	TW	6040.15	6074.89		
	RF	X	X		
	GBM	X	X		
Dev. Expl.	Delta	51.80%	49.75%		
	Gaus	68.90%	66.70%		
	TW	62.00%	58.80%		
	RF	66.72%	66.63%		
	GBM	76.00%	64.00%		
MAE	Delta	86.55	92.98		
	Gaus	98.38	108.43		
	TW	91.82	98.16		
	RF	61.51	59.59		
	GBM	87.49	100.85		

Table 4. Training results for different models and evaluated using different indicators for adults of SOLEMON species, i.e., common sole, common cuttlefish and mantis shrimp.

metrics	model	ST	ST+Env	ST	ST+Env
			Common sole (juveniles)	Mantis shrimp (juveniles)	
AIC	Delta	101,73/884,01	84,87/924,05	367,27/1246,39	373,04/1266,24
	Gaus	1284.04	1301.69	2767.26	2802.75
	TW	4394.93	4390.81	6350.31	6327.68
	RF	X	X	X	X
	GBM	X	X	X	X
Dev. Expl.	Delta	76,05%%	72.15%	58.50%	56.20%
	Gaus	81,8%%	79.30%	66,7%%	65.40%
	TW	72,1%%	72.10%	64,1%%	66.40%
	RF	78%%	79.60%	73.20%	64.70%
	GBM	87.00%	90.00%	75.00%	61.00%
MAE	Delta	531.4	591.64	312.35	330.03
	Gaus	580.29	571.53	336.09	358.6
	TW	570.6	550.43	328.36	309.6
	RF	435.25	346.01	230.86	228.9
	GBM	465.59	430.3	319.53	398.43
			Cuttlefish (juveniles)		
AIC	Delta	466,35/961,62	454,58/993,49		
	Gaus	2484.36	2539.68		
	TW	5556.38	5599.83		
	RF	X	X		
	GBM	X	X		
Dev. Expl.	Delta	56.95%	56.75%		
	Gaus	67.40%	64.40%		
	TW	66.30%	64.00%		
	RF	66.20%	65.22		
	GBM	73.00%	78.00%		
MAE	Delta	110.01	115.52		
	Gaus	116.05	126.66		
	TW	113.63	123.97		
	RF	81.62	78.14		
	GBM	110.11	105.16		

Table 5. Training results for different models and evaluated using different indicators for juveniles of SOLEMON species, i.e., common sole, common cuttlefish and mantis shrimp.

5.2 Predicted errors

Using the SDM ensemble the predicted values were calculated on a grid of regular points (with the same resolution of the selected CMEMS product 1/16°) covering the study area are compared with the data surveys. The model error was calculated as:

$$\text{Err}_i = \left(\frac{|y_i - x_i|}{x_i} \right) * 100 \quad \text{Eq. (1)}$$

where y_i is the model predicted value on the grid and x_i is the survey data value, for each haul (i) of the survey. The spatial distribution of the model error is mapped by triangular irregular network (TIN) interpolation (REF) algorithm in R. This method is useful to predict values for unmeasured location and TIN utilizes the points to constitute many non-overlapping triangles that cover the entire region according to a set of rules. The ground surface is described approximately with triangles (Advance Remote Sensing, 2020). Furthermore, a mean percentage error for each year, for ensemble model and each species, was calculated as:

$$\frac{\sum_{j=1}^n \left(\left(\frac{|y_i - x_i|}{x_i} \right) * 100 \right)}{n_j} \quad \text{Eq. (2)}$$

Results as presented in Fig. 6 for model ensemble and for adults and juveniles of each species highlight the generally better performances of SDM models with spatiotemporal and including additional environmental and effort variables, i.e., the integrated model better perform in describing the data available.

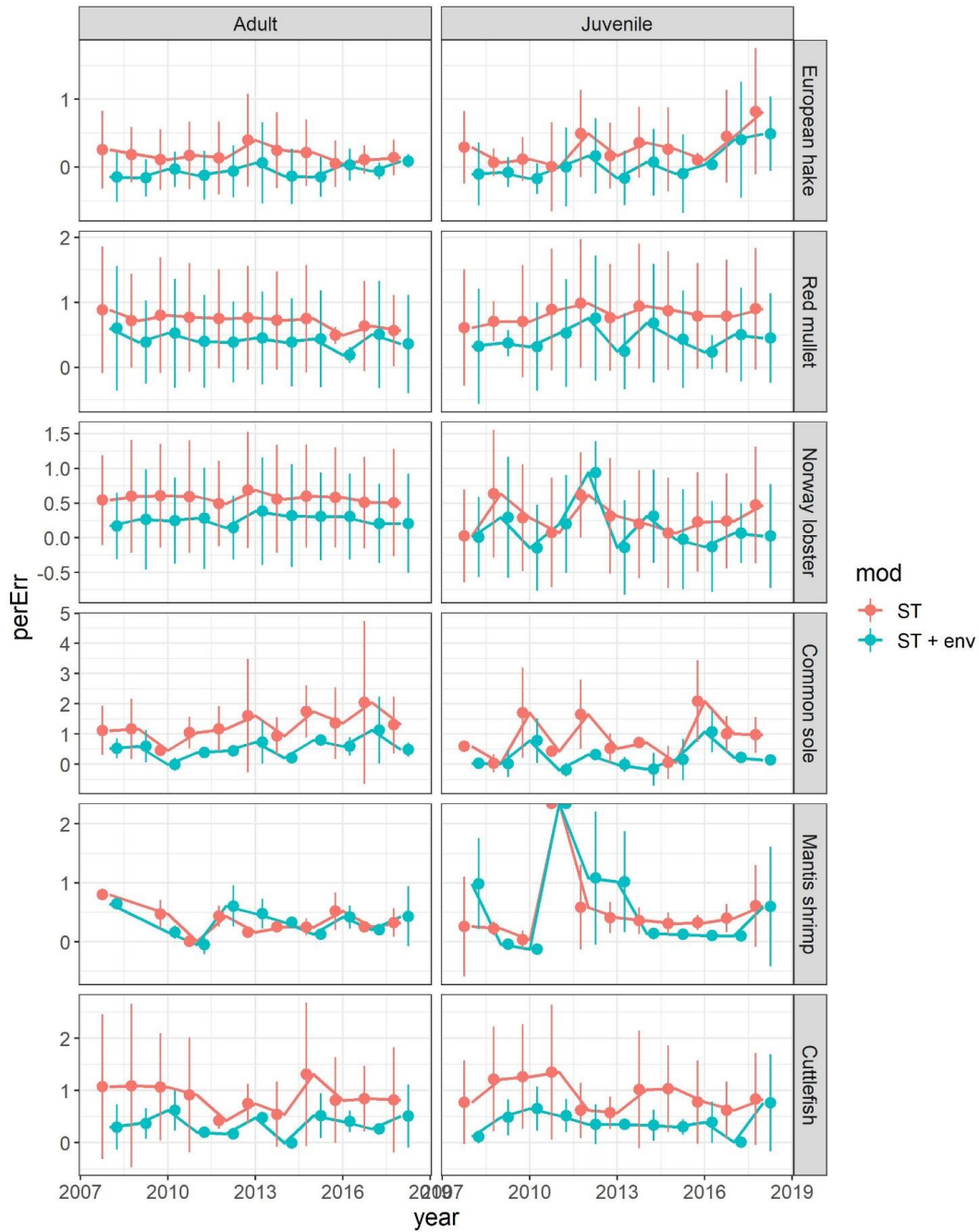


Fig. 6. Annual error (average and SD) for weighted ensemble model for the 6 species. Legend: ST, spatiotemporal, ST + Env, spatiotemporal + environmental variables and effort.

Species	Adult		Juvenile	
	ST	ST+Env	ST	ST+Env
Hake	0.59	0.65	0.54	0.58
Red mullet	0.14	0.22	0.32	0.4
N. lobster	0.43	0.48	0.32	0.36
Cuttlefish	0.73	0.75	0.74	0.74
Mantis shrimp	0.68	0.69	0.69	0.71
Sole	0.70	0.67	0.60	0.65

Table 6. Pearson correlation values between survey data and model prediction, for each species and size, for ST and ST+Env model.

The average error of SDM ensemble with and without environmental and effort variables, as well as the correlation coefficient as represented in Table 6 for all species highlight a generally much better performance of integrated models in describing trawl survey data.

5.3 Identification of hot spots on results of SDM

The best model for each species and each approach is used to obtain maps of the density distribution (ind km⁻²) for each year that allow identifying areas of high biomass density (hot-spots) in the GSAs 17, 18, and 19. To identify a hot spot area we used Getis index (Getis A. & Ord J.K, 1992) that calculate areas of aggregation by comparing statistically the average of a group of cells with average of the model domain abundances. The number of cells used was equal to 8 (after several tests).

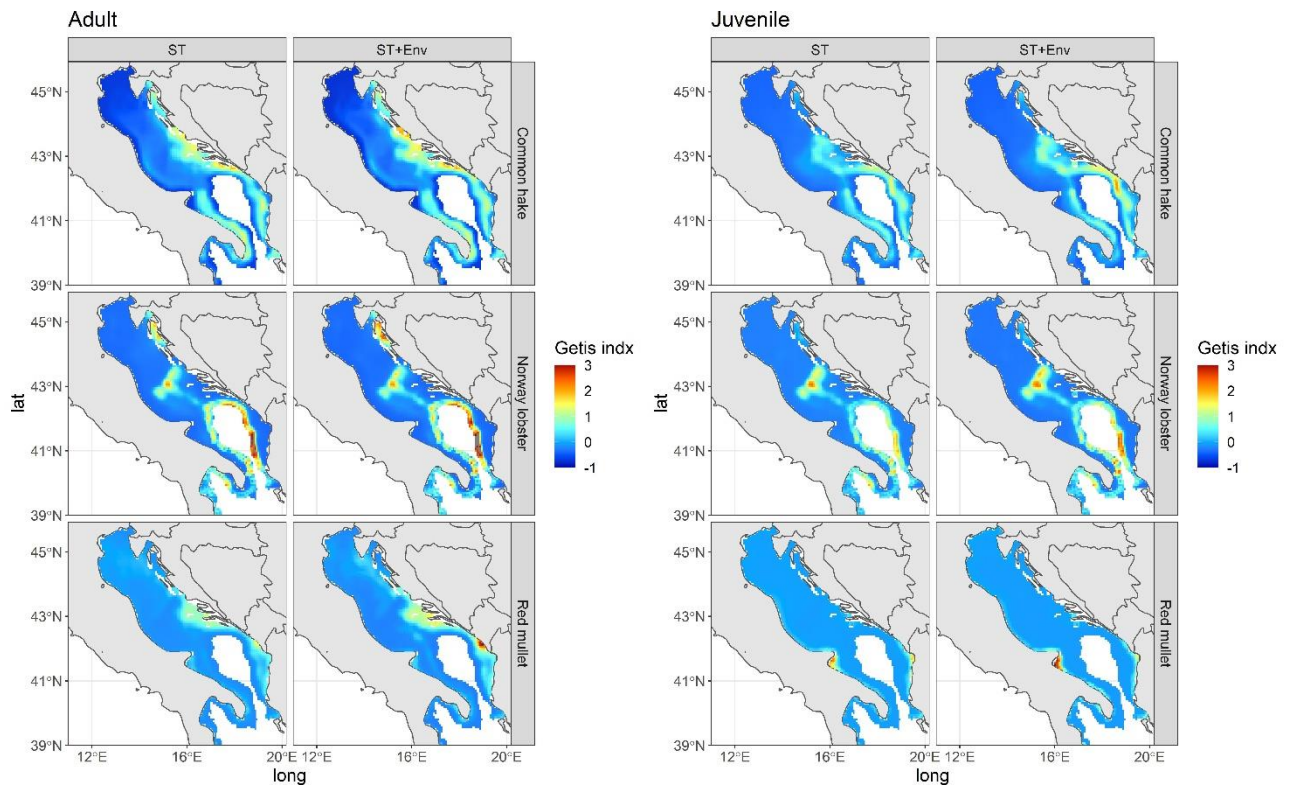


Fig. 7: Results of the Getis index applied at the weighted ensemble model for European hake, Norway lobster and Red Mullet based on data from MEDITS survey. Each panel shows spatial-temporal model (ST) and spatial-temporal + environmental model (ST+Env) for adult (left) and juvenile (right).

In order to account of the different performances among models in representing the data, the weighted average of the 5 approaches was used as ensemble weighted mean. The weights used were the R^2 obtained in the validation process. The weighted averages of the Getis index were obtained both using the spatiotemporal variables only, and including also relevant oceanographic and effort variables (ST and ST+Env, respectively). The resulting hot spots for both adults and juveniles are reported in Figure 7 (European hake, Norway lobster and red mullet) and Figure 8 (common sole, common cuttlefish and mantis shrimp).

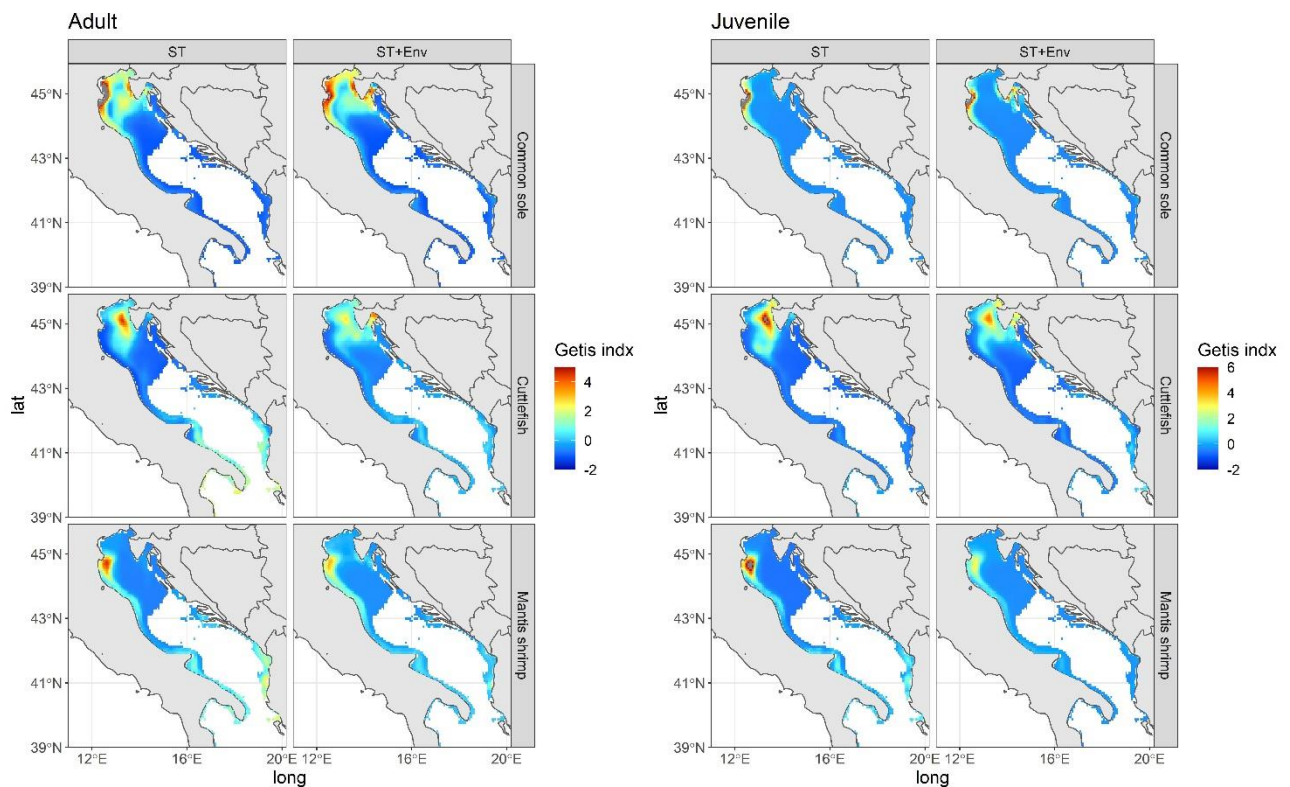


Fig. 8: Results of the Getis index applied at the weighted ensemble model for common sole, common cuttlefish and mantis shrimp derived from SOLEMON survey data by species. Each panel shows spatial-temporal model (ST) and spatial-temporal + environmental model (ST+Env) for adult (left) and juvenile (right).

Figures 7 and 8 show that although the identification of hot spots seems robust to the inclusion or not of the environmental and effort variables, the most integrated model is not only better (Figure 6) but it is also highlighting more granularity and gradients in the species distribution. Some differences in the identification of hot spots occur when including or not environmental variables (from HYDRO e BGC) and effort (from module EFFORT) that might be relevant for management purposes.

Notably hot spots are clearly specific for each species and are coherent with available information and previous scientific achievements.

6. Using results for management insights: matching hot spots

Hot spots by species, located as presented in figures 7 and 8, represent potential areas for spatial fisheries management. However, in the context of mixed fisheries as that of the Adriatic, might be useful to restrict and focus the potential areas of intervention to areas favourable for aggregation of more than one species. Matching the hot spots obtained from the ensemble of models, therefore, might represent a very useful approach to identify most relevant ecological areas for management.

A preliminary analysis on the hot spots identified by Getis index in order to identify areas where 2 or more species aggregates. This was done both for juveniles and adults.

Getis values for the threshold of 75 percentiles were filtered for each species and the matching grid cells for two or more species were identified, with or without environmental variable (ST+env and ST respectively). This approach applied for both adults and juveniles resulted in Figures 9 and 10, respectively.

Different colours among overlapping areas help to disentangle the overlapping species. Red area includes shared hot spots for European hake and Norway lobster. Violet area for European hake and red mullet, and light green area for red mullet and Norway lobster. These areas are more common in the southern part of the basin. Similarly, in yellow was represented the shared hot spot area between cuttlefish and common sole, light grey area between common sole and mantis shrimp, and dark grey area between cuttlefish and mantis shrimp. The hot spots areas for these species concentrate in the northern part of the basin.

It appears that there are areas where three species overlap their hot spots, and these are visible for adults in the area close to the Pomo Pit.

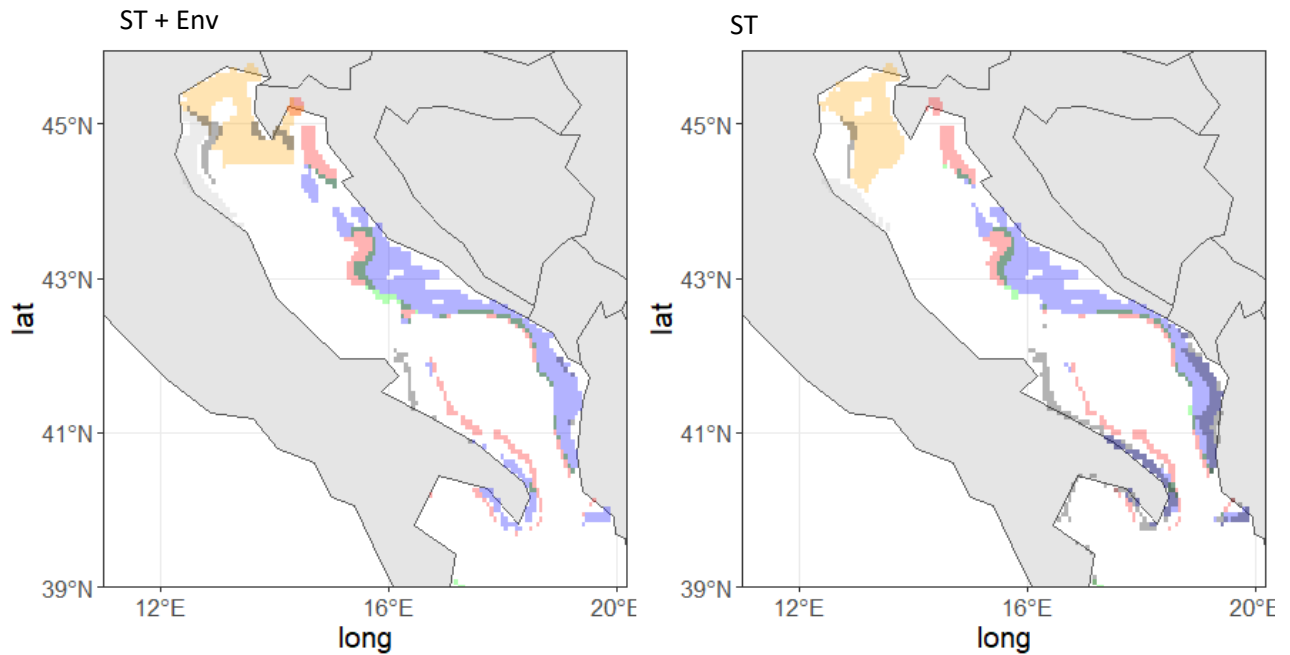


Fig. 9: Representation of matching area (from Getis analysis, $\geq 75\%$) among hot spots for adults of the analysed species. The approach is applied for ensemble of both model ST and ST+ env. The colours are: red area includes shared grid cells for European Hake and Norway lobster. Violet area for European hake and red mullet, and light green area for red mullet and Norway lobster. About SOLEMON species we have: yellow is a shared area between cuttlefish and common sole, light grey area between common sole and mantis shrimp, and finally, dark grey area between cuttlefish and mantis shrimp. **Hot spots matching for adults of more than 2 species.**

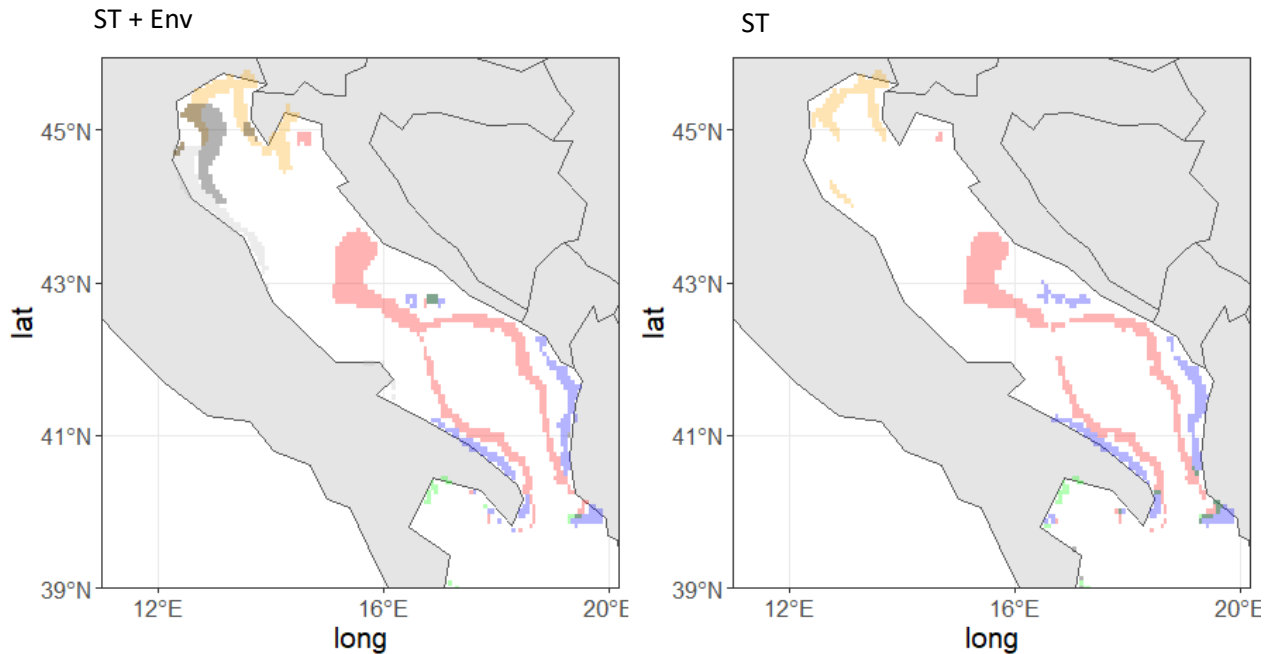


Fig. 10: Representation of matching area (from Getis analysis, $\geq 75\%$) among hot spots for juveniles of the analysed species. The approach is applied for ensemble of both model ST and ST+ env. The colours are: red area includes shared grid cells for European Hake and Norway lobster. Violet area for European hake and red mullet, and light green area for red mullet and Norway lobster. About SOLEMON species we have: yellow is a shared area between cuttlefish and common sole, light grey area between common sole and mantis shrimp, and finally, dark grey area between cuttlefish and mantis shrimp. **Hot spots matching for juveniles of more than 2 species.**

6.1 Capabilities and Limitations

The approach is statistically sound and the application of ensemble of different models permit to consider the capabilities of each model in fitting the data, possibly resulting in a best overarching ensemble model.

One important limitation of the approach relies on the definition of threshold between juveniles and adults. This threshold is based on the average length at first maturity but might be opportune to consider other options also based on local biological and fisheries knowledge.

Furthermore, the significance of results strongly relies on the trawl survey data accuracy in representing the species abundance and distribution. More than the quality of trawl survey is the representation of one seasonal sampling moment that should worry: clearly the MEDITS is better representative of summer periods and SOLEMON of the November periods. This should be kept in mind when considering the results for juveniles and adults reported here as the match between sampling season and relevance of the season for biological cycle of each species should be considered.

7 Final remarks

A procedure was set for integrating at best the data from environmental variables from modules 4.1 HYDRO and 4.2 BGC, effort data from module 4.5 EFFORT and trawl survey independent fishery data available in 4.3 BSTAT. Such approach is statistically robust, obtained with a protocol for training and testing repetitively the several Species Distribution models. The resulting ensemble proved to be better in representing fishery independent data when embedding also environmental and effort variables.

The approach is at the base of the identification of hot spots of aggregation for juveniles and adults of 6 important demersal target species: European hake (*Merluccius merluccius*, bony fish), red mullet (*Mullus barbatus*, bony fish) and Norway lobster (*Nepros norvegicus*, crustacean), common sole (*Solea solea*, bony fish), mantis shrimp (*Squilla mantis*, crustacean) and common cuttlefish (*Sepia officinalis*, mollusc).

The available layers of information integrated resulted in hot spots that are also overlapping. Overlapping areas among hot spots might be considered preferred areas for management as they can support the rebuilding and/or protect critical stages that might be identified in fisheries stock assessment. The located hot spots areas identified here are preliminary and anyway are of ECOLOGICAL significance. This does not necessarily means that these are areas to be managed. Further approaches

can be applied to simulate fisheries management in these areas for testing their appropriateness in terms of social, economic and ecological fisheries sustainability.

7.1 Notes

The work presented here is also part of the PhD of Diego Panzeri which is a PhD student of the XXXV cycle at Università degli Studi di Trieste, currently developing a PhD thesis in synergy with FAIRSEA activities.

8. REFERENCES

- Bastardie, F., Rasmus Nielsen, J., and Miethe, T. (2014). DISPLACE: a dynamic, individual-based model for spatial fishing planning and effort displacement—Integrating underlying fish population models. *Can. J. Fish. Aquat. Sci.* 71, 366–386. doi: 10.1139/cjfas-2013-0126.
- Boldrin A, Langone L, Misericocchi S, Turchetto M, Acri F (2005) Po River plume on the Adriatic continental shelf: dispersion and sedimentation of dissolved and suspended matter during different river discharge rates. *Mar Geol* 222–223:135–158
- Breiman L., 2001. Random forest. *March. Learn.* 45,5-32.
- Cao, J., Thorson, J. T., Richards, R. A., & Chen, Y., (2017). Spatiotemporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(11), 1781-1793.
- Carlucci, R., G. Lembo, P. Maiorano, F. Capezzuto, C. A. Marano, L. Sion, M. T. Spedicato, N. Ungaro, A. Tursi, and G. D’Onghia. (2009). Nursery areas of red mullet (*Mullus barbatus*), hake (*Merluccius merluccius*), and deepwater rose shrimp (*Parapenaeus longirostris*) in the Eastern-central Mediterranean Sea. *Estuarine Coastal and Shelf Science* 83:529–538.
- Civitaresse, G., Gačić, M., Lipizer, M., and Eusebi Borzelli, G. L. (2010): On the impact of the Bimodal Oscillating System (BIOS) on the biogeochemistry and biology of the Adriatic and Ionian Seas (Eastern Mediterranean), *Biogeosciences*, 7, 3987–3997, <https://doi.org/10.5194/bg-7-3987-2010>.
- Cotter J., Petitgas P., Abella A., Apostolaki P., Mesnil B., Politou C-Y., Rivoirard J., Rochet M.-J., Spedicato M.T., Trenkel V.M., Woillez M. (2009). Towards an ecosystem approach to fisheries management (EAFM) when trawl surveys provide the main source of information. *Aquat. Living Resour.* 22: 243-254. <https://doi.org/10.1051/alr/2009025>.

Druon, J. N., Fiorentino, F., Murenu, M., Knittweis, L., Colloca, F., Osio, C., ... & Sbrana, M. (2015). Modelling of European hake nurseries in the Mediterranean Sea: an ecological niche approach. *Progress in oceanography*, 130, 188-204.

Fulton, E. A., Link, J. S., Kaplan, I. C., Savina-Rolland, M., Johnson, P., Ainsworth, C., ... & Smith, D. C. (2011). Lessons in modelling and management of marine ecosystems: the Atlantis experience. *Fish and fisheries*, 12(2), 171-188.

Getis, A. and J.K. Ord. 1992. The Analysis of Spatial Association by Use of Distance Statistics in *Geographical Analysis* 24(3).

Giannoulaki, M.; Valavanis, V.D.; Palialexis, A.; Tsagarakis, K.; Machias, A.; Somarakis, S.; Papaconstantinou, C. (2008). Modelling the presence of anchovy *Engraulis encrasicolus* in the Aegean Sea during early summer, based on satellite environmental data. *Hydrobiologia* 612: 225-240.

Grüss, Arnaud, Michael Drexler, and Cameron H. Ainsworth, (2014). Using delta generalized additive models to produce distribution maps for spatially explicit ecosystem models. *Fisheries Research* 159: 11-24.

Grüss, A., Chagaris, D. D., Babcock, E. A., & Tarnecki, J. H. (2018). Assisting ecosystem-based fisheries management efforts using a comprehensive survey database, a large environmental database, and generalized additive models. *Marine and Coastal Fisheries*, 10(1), 40-70

Lauria, V., Garofalo G., Fiorentino F., Massi D., Milisenda G., Piraino S., Russo T., & Gristina M., (2017). Species distribution models of two critically endangered deep-sea octocorals reveal fishing impacts on vulnerable marine ecosystems in central Mediterranean Sea. *Scientific reports* 7.1: 1-14.

Luan J, Zhang C, Xu B, Xue Y, Ren Y (2018) Modelling the spatial distribution of three Portunidae crabs in Haizhou Bay, China. *PLOS ONE* 13(11): e0207457. <https://doi.org/10.1371/journal.pone.0207457>.

Maunder, M. N., & Punt, A. E., (2004). Standardizing catch and effort data: a review of recent approaches. *Fisheries research*, 70(2-3), 141-159.

M. Marini, B.H. Jones, A. Campanelli, F. Grilli, C.M. Lee (2008). Seasonal variability and Po River plume influence on biochemical properties along western Adriatic coast. *J. Geophys. Res.*, 113. Article C05S90, 10.1029/2007JC004370.

Moullec F., L. Veleza, P. Verley, N. Barrier, C. Ulses, P. Carbonara, A. Esteban, C. Follesa, M. Gristina, A. Jadaud, A. Ligas, E. López Díaz, P. Maiorano, P. Peristeraki, M. T. Spedicato, I. Thasitis, M. Valls, F. Guilhaumon, Y.-J. Shin. (2019). Capturing the big picture of Mediterranean marine biodiversity with an end to end model of climate and fishing impacts. *Progress in Oceanography*, 178:102179.

Orio, A., Florin, A.-B., Bergström, U., Šics, I., Baranova, T., Casini, M. 2017. Modelling indices of abundance and size-based indicators of cod and flounder stocks in the Baltic Sea using newly standardized trawl survey data. *ICES Journal of Marine Science*, 74(5), 1322-1333, doi:10.1093/icesjms/fsx005.

Potts, S. E., & Rose, K. A. (2018). Evaluation of GLM and GAM for estimating population indices from fishery independent surveys. *Fisheries Research*, 208, 167-178.

Pranovi, F., Monti, M. A., Caccin, A., Colla, S., & Zucchetta, M. (2016). Recreational fishing on the West coast of the Northern Adriatic Sea (Western Mediterranean) and its possible ecological implications. *Regional Studies in Marine Science*, 3, 273-278.

Punzo, E., Gomiero, A., Tassetti, A.N. et al. Environmental Impact of Offshore Gas Activities on the Benthic Environment: A Case study (2017). *Environmental Management* 60, 340–356. <https://doi.org/10.1007/s00267-017-0886-4>

Rubec, P. J., Kiltie, R., Leone, E., Flamm, R. O., McEachron, L., & Santi, C. (2016). Using delta-generalized additive models to predict spatial distributions and population abundance of juvenile pink shrimp in Tampa Bay, Florida. *Marine and coastal fisheries*, 8(1), 232-243.

Russo, T., Morello, E. B., Parisi, A., Scarcella, G., Angelini, S., Labanchi, L., et al. (2018). A model combining landings and VMS data to estimate landings by fishing ground and harbor. *Fish. Res.* 199, 218–230. doi: 10.1016/j.fishres.2017.11.002.

Schapire, R.E. (2003). The Boosting Approach to Machine Learning: An Overview. In *Nonlinear Estimation and Classification*; Springer: New York, NY, USA, 2003; pp. 149–171.

Schismenou, E., Tsoukali, S., Giannoulaki, M., & Somarakis, S. (2017). Modelling small pelagic fish potential spawning habitats: eggs vs spawners and in situ vs satellite data. *Hydrobiologia*, 788(1), 17-32.

Sion L., W. Zupa, C. Calculli, G. Garofalo, M. Hidalgo, A. Jadaud, E. Lefkaditou, A. Ligas, P. Peristeraki, I. Bitetto, F. Capezzuto, R. Carlucci, A. Esteban, C. Follesa, B. Guijarro, Z. Ikica, I. Isajlovic, G. Lembo, C. Manfredi, J. L. Pérez, C. Porcu, I. Thasitis, G. Tserpes, P. Carbonara. (2019). Spatial distribution pattern of European hake, *Merluccius merluccius* (Pisces: Merlucciidae), in the Mediterranean Sea. *Scientia Marina*, 83S1: 21-32.

Spedicato M.T., Zupa W., Carbonara P., Fiorentino F., Follesa M.C., Galgani F., García-Ruiz C., Jadaud A., Ioakeimidis C., Lazarakis G., Lembo G., Mandic M., Maiorano P., Sartini M., Serena F., Cau A., Esteban A., Isajlovic I., Micallef R., Thasitis I. (2019b). Spatial distribution of marine macro-litter on the seafloor in the northern Mediterranean Sea: the MEDITS initiative. *Sci. Mar.* 83S1: 257-270. <https://doi.org/10.3989/scimar.04987.14A>.

Teruzzi, A., Bolzon, G., Cossarini, G., Lazzari, P., Salon, S., Crise, A., & Solidoro, C. (2019). Mediterranean Sea Biogeochemical Reanalysis (CMEMS MED-Biogeochemistry) [Data set]. Copernicus Monitoring Environment Marine Service (CMEMS). https://doi.org/10.25423/MEDSEA_REANALYSIS_BIO_006_008.

Tserpes G., Massuti E., Fiorentino F., Facchini M.T., Viva C., Jadaud A., Joksimovic A., Pesci P., Piccinetti C., Sion L., Thasitis I., Vrgoc N. (2019). Distribution and spatio-

temporal biomass trends of red mullets across the Mediterranean. *Sci. Mar.* 83S1: 43-55. <https://doi.org/10.3989/scimar.04888.21A>.

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). block CV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10(2), 225-232.