the forecasts delivered on the IcySea application (between 3.5 and 3.8 millions of individual datasets depending on the lead time), only 80% of these datasets have been taken into account for training the models used for evaluating the performances of the algorithms. The remaining 20% of the datasets were used to evaluate the performances of the statistical models. The selection of the datasets used for training and testing the models is a random process (according to the forecast start date), and has been repeated 50 times in order to test the influence of this selection on the performances of the calibrated forecasts.

The performances of the calibrated forecasts for the area shown in Figure 3.5.2(a) are presented in Figure 3.5.3. Overall, there is a larger improvement for the speed than for the direction of sea-ice drift. Nevertheless, the calibrated forecasts outperform the TOPAZ4 forecasts for all the lead times for the direction and the speed of sea-ice drift (lower mean absolute errors). Furthermore, most of the calibrated forecasts have lower errors than the TOPAZ4 forecasts (Figure 3.5.3(c,d)). On average, about 53% of the forecasts are improved by the calibration for the drift direction and 57% for the drift speed. The calibrated forecasts delivered on the IcySea application are therefore more accurate than the forecasts produced by the TOPAZ4 system.

3.5.2. Conclusion

Maritime traffic around the Svalbard archipelago has shown changing patterns during the last decade, with expanding seasons and operational areas (Stocker et al. 2020). In this challenging operational environment, sea-ice remains one of the major sources of uncertainty for navigating in this area, and there is a need for sea-ice information that can be easily downloaded and visualised by end-users. The IcySea application addresses this need by delivering satellite images and sea-ice drift forecasts in a user-friendly interface. The calibration method developed for the IcySea application improves the accuracy of sea-ice drift forecasts compared to the forecasts from TOPAZ4. Furthermore, it is designed to work offline and the data can be downloaded under low-bandwidth connections, which is a common limitation in the Arctic. This application should therefore contribute to improving operational planning and safety in the Svalbard area.

Section 3.6. Developing spatial distribution models for demersal species by the integration of trawl surveys data and relevant ocean variables

Authors: Panzeri D., Bitetto I., Carlucci R., Cipriano G., Cossarini G., D'Andrea L., Masnadi F., Querin S., Reale M., Russo T., Scarcella G., Spedicato M.T., Teruzzi A., Vrgoč N., Zupa W., Libralato S.*

Statement of main outcome: Demersal species play a fundamental role in fisheries, thus understanding their distribution and abundance through bottom trawl surveys is crucial for stock and fisheries management. Oceanographic (e.g. biogeochemical, physical) and fishing covariates might be considered, in addition to spatio-temporal variables (latitute, longitude, depth, year and month), to better explain trawl survey data. Here, we analyse biomass indices (kg/km^2) for European hake, common sole, mantis shrimp, red mullet and common cuttlefish from scientific trawl surveys carried out in the Adriatic Sea and the Western Ionian Sea. We used three different Generalised Additive Model (GAM) approaches (Gaussian, Tweedie and Delta) to fit and predict species biomass distribution. In order to evaluate trade-offs in using different covariates, we compared the results obtained from GAM approaches based only on spatiotemporal variables and GAMs including also oceanographic and fishing effort covariates.

The Delta-GAM approach performed better for European hake, mantis shrimp and common cuttlefish, while GAMs based on Gaussian and Tweedie were performing better for the red mullet and common sole, respectively. The results highlighted that adding specific oceanographic and effort covariates to spatiotemporal variables improved the performances of spatial distribution models especially for European hake, mantis shrimp and red mullet. Significant additional explanatory variables were bottom temperature, bottom dissolved oxygen, salinity, particulate organic carbon, and fishing effort for European hake; the same variables and pH for mantis shrimp; chlorophyll-a, pH, sea surface temperature, bottom dissolved oxygen, nitrate and effort for the red mullet; phosphate and salinity for common sole; bottom temperature, bottom dissolved oxygen, and phosphate for the common cuttlefish.

The findings highlight that more accurate estimates of spatial distribution of demersal species biomass from trawl survey data can generally be obtained by integrating oceanographic variables and effort in GAMs approaches with potential impacts on stock assessment and essential fish habitats identification.

Products used

Ref. No.	Product name and type	Documentation
3.6.1	MEDSEA_REANALYSIS_PHYS_006_004	PUM: http://marine.copernicus.eu/documents/PUM/CMEMS-MED-PUM-

(*Continued*)

Continued.

Ref. No.	Product name and type	Documentation
		006-004.pdf QUID: http://marine.copernicus.eu/documents/QUID/CMEMS-MED-QUID-006-004.pdf
3.6.2	MEDSEA_REANALYSIS_BIO_006_008	PUM: http://marine.copernicus.eu/documents/PUM/CMEMS-MED-PUM-006-008.pdf QUID: http://marine.copernicus.eu/documents/QUID/CMEMS-MED-QUID-006-008.pdf
3.6.3	MEDITS: Mediterranean International Trawl Survey; bottom trawl survey up to 800 m depth. From 1994 to 2018 on average 326 sampling sites (hauls) per year were conducted in the Adriatic and Northern Ionian Sea.	Bertrand et al. (2002), Spedicato, Massutí, et al. (2019), MEDITS-Handbook (2017)
3.6.4	SOLEMON: Sole Monitoring, modified beam trawl surveys conducted in the Northern Adriatic Sea up to a depth of 100 m; on average 70 sampling (hauls) per year from 2005 to 2018.	Scarcella et al. (2011), SoleMon Handbook 2019 (http://dcf-italia.cnr.it/assets/lineeguida/lin1/2019/SOLEMON-Handbook_2019_Ver_4.pdf)

3.6.1. 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 Mediterranean and Black Sea, these species constitute approximately 20% of the total landed weight (more than 230,000 tons/year) and 50% of the total landed value (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 large areas of the Mediterranean Sea 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, i.e. fishery-independent data, that 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, Zupa, et al. 2019), and the development of ecosystem models (e.g. Grüss et al. 2018; Moullec et al. 2019). Sampling protocols of multiannual surveys are usually standardised 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.

In order to test the potential benefits on using oceanographic and effort variables in addition to spatio-temporal covariates (latitute, longitude, depth, year and month) to improve species distribution models based on trawl survey data, Generalised Additive Models (GAMs) were chosen for their wide application and suitability with trawl survey data (Grüss et al. 2014; Lauria et al. 2017; Tserpes et al. 2019). GAMs allow to predict species abundance and biomass over the domain (Maunder and Punt 2004; Rubec et al. 2016; Potts and Rose 2018) and provide estimates useful for tuning stock assessment models (Cao et al. 2017; Orio et al. 2017). Furthermore, GAMs are deemed appropriate for mapping species distribution that is useful in ecosystem models (Fulton et al. 2011; Grüss et al. 2014), or for identifying Essential Fish Habitats (e.g. Colloca et al. 2015; 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 (Bastardie et al. 2014; Russo et al. 2018). 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.

In this work, therefore, we propose an integrated approach useful to fisheries management by combining trawl survey data, oceanographic variables and fishing effort estimates. 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 GAM 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.

3.6.2. Material and methods

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 as defined by the FAO-GFCM (General Fisheries Commission for the Mediterranean Sea). We used MEDITS (Mediterranean International Trawl Survey; Spedicato, Massutí, et al. 2019) data from 1994 to 2018 that consists on average 326 sampling sites (bathymetrical range 10–800 m) per year in the three GSAs (Product Ref. 3.6.3) and SOLEMON (Sole Monitoring; Scarcella et al. 2011; Grati et al. 2013) from 2005 to 2018, that consists on average 70 sampling sites per year in GSA 17 (bathymetrical range 10–100 m) (Product Ref. 3.6.4). Indices of demersal species biomass (kg/km^2) were retrieved from the MEDITS dataset for European hake (*Merluccius merluccius*) and red mullet (*Mullus barbatus*) and from the SOLEMON dataset for common sole (*Solea solea*), mantis shrimp (*Squilla mantis*) and common cuttlefish (*Sepia officinalis*).

For each species, GAMs were applied to fit biomass indices by sampling site, set as a response variable, while spatiotemporal variables, oceanographic variables and fishing effort were tested as covariates. Among the spatiotemporal variables we used geographic coordinates (latitude, longitude expressed in UTM coordinates), depth (m), month and year of the observations. Among all the variables available from the 3D monthly CMEMS Mediterranean reanalysis fields (Product Ref. 3.6.1 and 3.6.2) relevant oceanographic variables were considered on the basis of 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 (°C) and dissolved oxygen (mmol/m^3) at the sea bottom, water column averages of nitrate and phosphate concentration (mmol/m^3), chlorophyll-a (mg/m^3), particulate organic carbon (mg/m^3), pH and salinity. These variables were derived from the CMEMS dataset that covers the period 1999–2018, has a spatial horizontal resolution of 1/16° and 72 unevenly vertical levels (Simoncelli et al. 2019; Teruzzi et al. 2019). Furthermore, commercial trawling effort expressed as trawling time (in hours) per year at spatial resolution of 1/16° was estimated from VMS/AIS data for the period 2008–2018 (Russo et al. 2014) and was tested as explanatory variable on the basis of the evidence that fishing effort is a good track of species density. Although different time frames were initially adopted (depeding on the available explanatory variables), here we report the analysis performed on the time frame 2008–2018 that allowed the complete overlap between trawl survey, CMEMS and effort datasets. 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 the spatiotemporal variables, i.e. year, month, depth, latitude, longitude, to be included as explanatory variables. Furthermore, the VIF analysis by species allowed to include additional explanatory variables without collinearity extracted from CMEMS reanalysis and fishing effort: the VIF results showed to be species-specific. Thus the complete model for European hake included the spatiotemporal variables and the bottom temperature, bottom dissolved oxygen, nitrate concentration, salinity, bottom particulate organic carbon, and fishing effort. For the red mullet the following explanatory variables were retained after VIF analysis in the most complete model: month, latitude, longitude, year, depth, pH, chlorophyll-a, sea surface temperature, bottom dissolved oxygen, nitrate, salinity and effort. For the common cuttlefish, the complete set of variables after VIF included month, latitude, longitude, year, depth, bottom temperature, bottom dissolved oxygen, nitrate, phosphate and effort. For common sole the complete set of variables included month, latitude, year, depth, average phosphate, bottom temperature, bottom dissolved oxygen, salinity, average phosphate, pH and effort. For mantis shrimp the set of variables are month, latitude, year, depth, bottom temperature, bottom dissolved oxygen, salinity, particulate organic carbon, pH and fishing effort (more details in Supplemetary Material).

Different GAM distribution families were applied in order 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 biomass data log-transformed for all species, except common cuttlefish, for which better results were obtained by using square root transformation. GAMs were also applied using Tweedie probability distributions with lognormal link on untransformed biomass 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 biomass for presence-only data (Grüss et al. 2014; Lauria et al. 2017). A grid of regular points with the same resolution of the selected CMEMS product (1/16°) and covering the study area was created to predict species biomass distribution by the selected models (Tserpes et al. 2019; Spedicato, Zupa, et al. 2019). For Delta-GAM the final spatial distribution of species biomass as kg/km^2 is obtained by multiplication of Gaussian and Binomial models' predictions to the grid of the model's domain (Grüss et al. 2014; Lauria et al. 2017).

For each species and all GAMs distribution families (Delta, Gaussian and Tweedie), a back-stepwise approach was used. This started from the most complete integrated approach, given by the spatiotemporal variables (geographical coordinates, depth, year, month) combined with all the most meaningful additional biogeochemical, physical and fishing effort variables identified by VIF analysis (model 0). Then the back-stepwise approach consisted in decreasing the number of explanatory variables by successively removing those with lower F statistics till to obtain the model with spatiotemporal variables only. Thus, the back-stepwise approach resulted in a set of models having different explanatory variables to obtain the response variable ($R = \log$ kg/km^2 or presence/absence) (see Supplementary Material). Each model was subjected to a calibration-validation process, thus it was fitted on a training dataset made by randomly choosing 70% of the data (calibration) and testing it on the remaining 30% of records (validation). The training and testing were repeated using 50 runs on datasets randomly selected and without replacement. The best model was selected on the basis of measures of model's performance evaluated through explained deviance (%ED) and prediction errors (AIC, Akaike Information Criterion) on the training datasets as well as correlation coefficient (R^2) of the model predictions on the testing dataset.

For each model with decreasing number of explanatory variables (model 0, model 1, model 2, etc.), the mean of each measure of model's performance (%ED, AIC, R2) was calculated from the 50 runs and compared using the Tukey's test (Tukey 1949). This comparison allows to assess the improvement of performances when different sets of additional variables were used in the models. The best model was chosen based on AIC, but other measures of performance were reported for showing their general consistency.

The chosen model for each species is used to obtain maps of the biomass distribution (kg/km^2) on the most relevant month (July and November, for MEDITS and SOLEMON species, respectively). The maps allowed identifying areas of high biomass density (hot-spots) in the GSAs 17, 18 and 19. Furthermore, a set of spatial indicators (Woillez et al. 2009) permitted to compare models' performances in describing the spatial distribution of demersal species when including or not additional explanatory variables. The set of indicators are the Spreading area (SA), i.e. a measure of the area occupied by the population weighted by the biomass; the latitude of the centroid or centre of gravity of data (CGY), which represents the mean geographic location of the population; the longitude of the centroid (CGX); the distance (D) between the centroid estimated on observations and the centroid estimated on predictions (Woillez et al. 2009; Rufino et al. 2018). Distribution statistics (first and third quartile, median) and performance indicators (mean absolute error MAE and R^2) were also estimated. Comparing such indicators calculated on raw trawl survey data, on models based only on spatiotemporal variables and on the chosen best models using the complete set of significant variables, allow to quantify the improvement of adopting the integrated approach, i.e. embedding biogeochemical, physical and fishing effort, in species distribution models.

3.6.3. Results and discussion

For European hake, mantis shrimp and common cuttlefish the Delta-GAM models were performing better while for the red mullet and common sole the best results were obtained using the Gaussian model and Tweedie, respectively (details are reported in Supplementary material). Figure 3.6.1 shows measures of performance (%ED, AIC, R2) resulting from the back-stepwise approach applied to the most

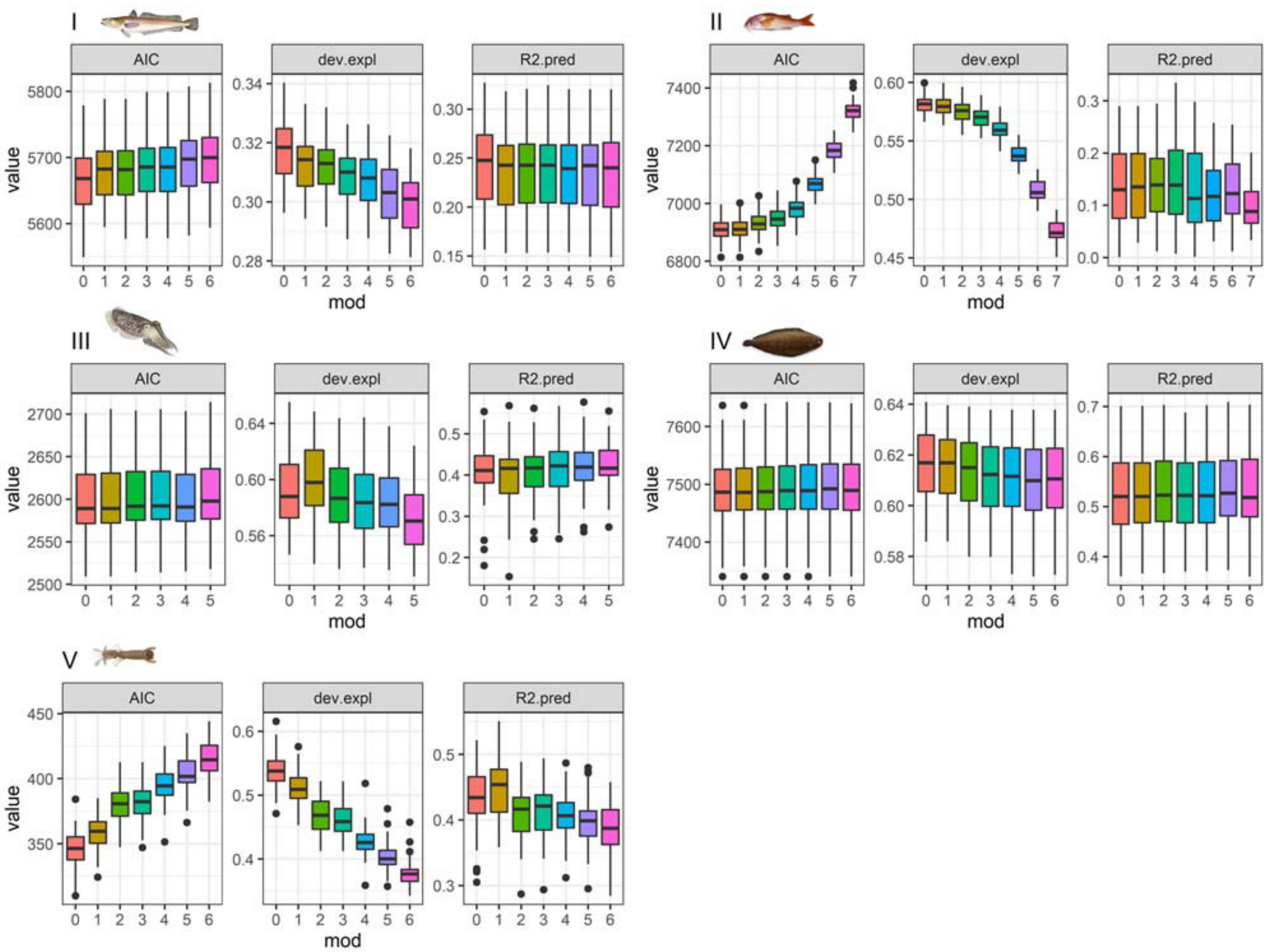

Figure 3.6.1. Performances of the best GAMs in describing the distribution of demersal species for models using a decreasing number of explanatory variables. The best model was Delta-GAM for European hake, common cuttlefish and mantis shrimp (shown the Delta-Gaussian in panels I, III and V, respectively), Gaussian for red mullet (panel II) and Tweedie for common sole (panel IV). For all species the starting model represents the one (model 0) including all the covariates resulting from VIF analysis and including spatiotemporal variables, environmental CMEMS variables and fishing effort (Product Ref. 3.6.1 and 3.6.2, respectively). Successively one variable at each step is removed to reach the minimal model (model 6 for European hake, common sole and mantis shrimp; model 7 for red mullet; model 5 for common cuttlefish) with spatiotemporal variables only. Box-plots synthesise results of the 50 runs of the training/testing procedure in terms of Akaike Information Criterion (AIC), explained deviance (dev-expl) on the 70% training dataset and correlation coefficient (R^2) for the remaining testing dataset.

appropriate family of GAM models for each species (only Delta-Gaussian is reported in the figure for European hake, mantis shrimp and common cuttlefish; the full Delta-GAM results for these species are reported in Supplementary material). Results for the 50 trials of training/testing demonstrate the model improvements when using CMEMS and effort variables in GAMs (Tukey's tests are reported in Supplementary material).

For European hake, the average AIC for Delta-Gaussian increased from 5600 for the model including the complete set of variables (model 0, panel I) to 5700 for the minimal model with spatiotemporal variables only (model 6, panel I). Coherently, the average %ED decreased from 0.32–0.29, and R^2 decreased from 0.24–0.23 from model with complete set of variables to model with spatiotemporal variables (Figure 3.6.1, panel I). For red mullet AIC increased from 6950 to 7340, %ED decreased from 0.57–0.47 and R2 decreased from 0.12–0.09 from the complete to the minimal model (Figure 3.6.1, panel II). For mantis shrimp AIC increased from 350 to 420, %ED decreased from 0.55–0.37, and R2 decreased from 0.44–0.38 from the complete to the minimal model (Figure 3.6.1, panel V). For common cuttlefish and common sole (panels III and IV) the differences in AIC and R2 are less marked when moving from the complete model (0) to the model with spatiotemporal variables (model 5 and 6) but yet the improvement is appreciable in terms of %ED. For all species analysed,

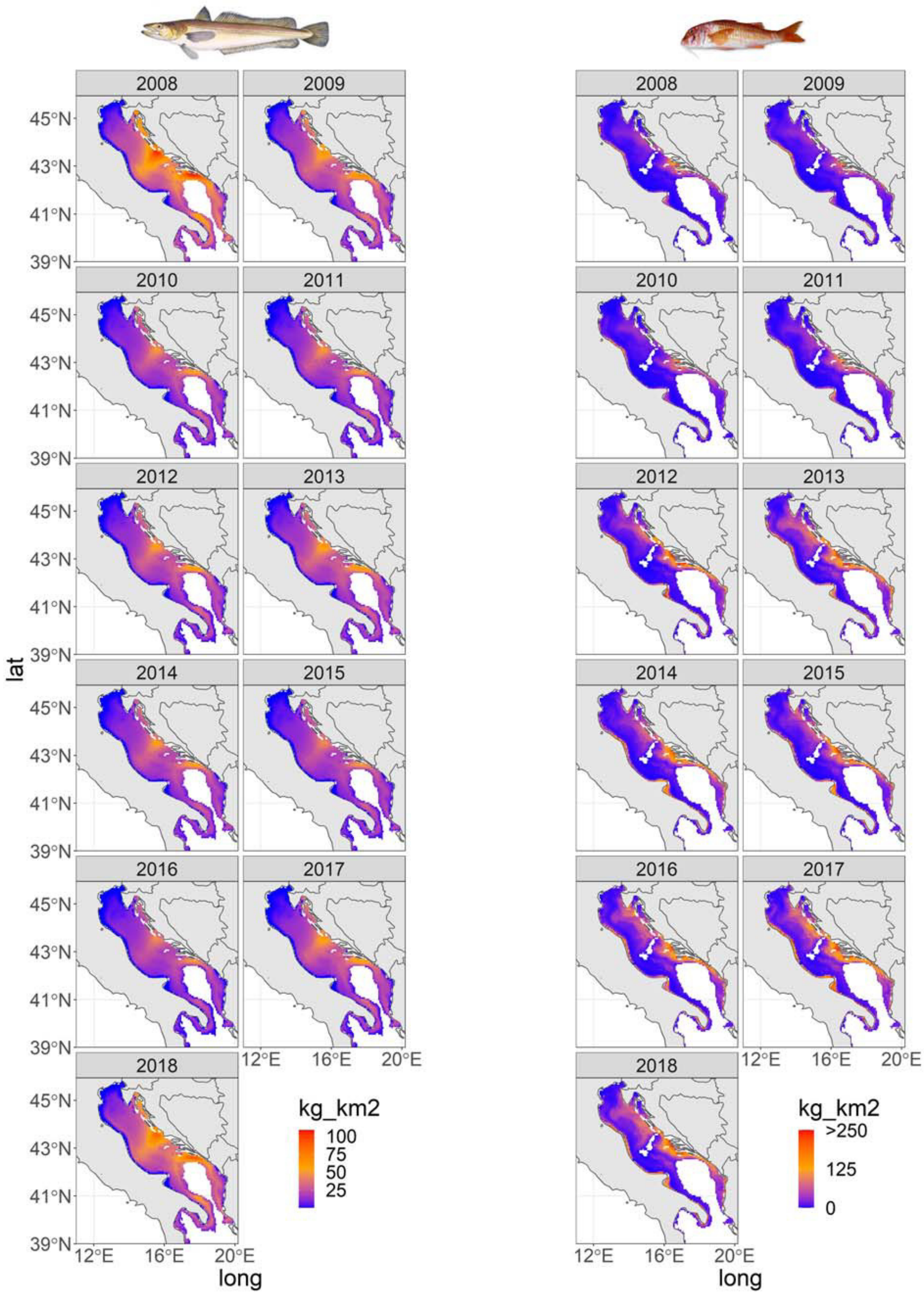

Figure 3.6.2. Yearly maps of estimated biomass (kg/km^2) of European hake (left) and red mullet (right) in the Adriatic and Western Ionian Sea (GSA 17-18-19) obtained with the best GAM model applied on MEDITS trawl survey data for years 2008–2018 (Product Ref. 3.6.3) and with all the additional environmental and effort variables (model 0).

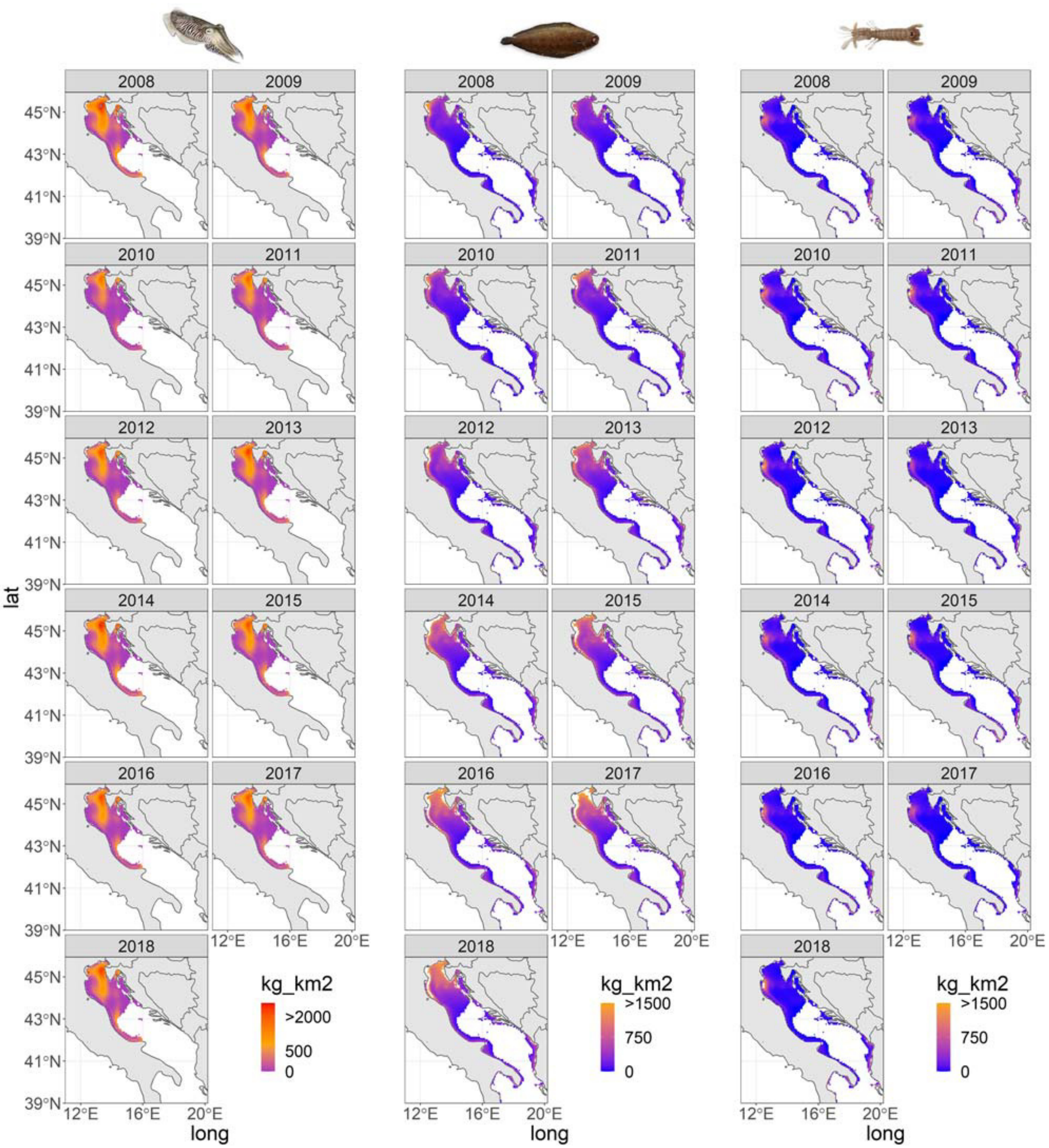

Figure 3.6.3. Yearly maps of estimated biomass (kg/km^2) of common cuttlefish (left), common sole (centre) and mantis shrimp (right) in the Adriatic Sea (GSA 17-18) obtained with the best GAM model applied on SOLEMON trawl survey data for years 2008–2018 (Product Ref. 3.6.4) and with all the additional environmental and effort variables (model 0).

the training/testing approach highlighted that best performances in terms of capabilities to represent trawl survey biomass data (ED% and R2) and performance indicators such as AIC were obtained when the integrated approach was used, i.e. when the spatial model for species distribution included biogeochemical, physical (Product ref. 3.6.1) and fishing effort (Product ref. 3.6.2) as additional explanatory variables (model 0).

For each demersal species the best model has specific significant covariates in addition to spatiotemporal variables. Bottom temperature, bottom dissolved oxygen, salinity, particulate organic carbon, and fishing effort resulted significant variables for European hake. The same variables and pH resulted sigificant for mantis shrimp. Chlorophyll-a, pH, sea surface temperature, bottom dissolved oxygen, nitrate and effort were significant for the red mullet. Bottom temperature, bottom dissolved oxygen, and phosphate for the common cuttlefish. Average phosphate and salinity were significant for common sole (more details in the Supplementary materials).

Figure 3.6.2 shows distribution maps for the years 2008–2018 as obtained by the best complete model for European hake and red mullet based on MEDITS trawl survey data. For European hake (Figure 3.6.2, left panel) the maps highlight higher biomass in 2008 and 2018, hot spots of biomass (as high as 100 kg/km^2) in the central-eastern part of the Adriatic Sea in recent years (particularly in 2018), low biomass of this species, especially in the northern part of the basin, and a prevalence of a north–south gradient. For the red mullet (Figure 3.6.2, right panel) results show that high biomass (up to 200 kg/km^2, particularly in years 2017/2018) is associated to coastal strip in the western part of the basin, while in the eastern part biomass is more widely distributed with a prevalence of south–north gradients. The application of the best complete GAM model for common cuttlefish, common sole and mantis shrimp based on SOLEMON trawl surveys result in distribution maps reported in Figure 3.6.3. The hot spot for common cuttlefish is consistently identified in the North-East Adriatic, in front of Istra peninsula, with highest biomass (peaks of 2000 kg/km$^{2)}$ especially in 2008 and 2014 (Figure 3.6.3, left panels). Common sole is showing higher densities along the North-western coast of the Adriatic, but high biomass are obtained also in the central part of the Northern Adriatic in recent years (2016–2018; central panels). The mantis shrimp resulted to be mainly distributed along the North-western coast in the area interested by the Po

Table 3.6.1. Comparison among indicators calculated on observations, i.e. the original trawl survey data (Product Ref. 3.6.3, 3.6.4), on the results of the GAM model with spatiotemporal variables and on results of the best GAM model including additional oceanographic variables (Product Ref. 3.6.1, 3.6.2) and effort (model 0). Distribution indicators (first and third quartile, median), performance indicators (MAE, R^2) and spatial indicators such as Spreading area (SA), latitudinal centroid (CGY) longitudinal centroid (CGX) and distance (D) of the centroid of model to that of data are reported for the five demersal species analysed. The column 'improvement' reports the improvement on the indicator value when using model with environmental variables with respect to indicator calculated on results of the model without additional variables (observations-model0/observations-model 6 or 7).

	Red mullet (*Mullus barbatus*), GSA 17, 18, 19, 2008–2018			
	Observations	Model 7	Model 0	Improvement
1st.Qu	0	3.11	**2.38**	23%
Median	1.95	8.48	**7.91**	18%
3rd.Qu	23.55	20.36	**23.88**	19%
R2	–	0.08	**0.15**	
MAE	–	39.97	**36.89**	
SA	701.6	1142.5	**1094**	11%
CGX	15.52	15.98	**15.73**	55%
CGY	42.59	**42.52**	42.8	–
D	0	38.68	**29.53**	24%

	Common cuttlefish (*Sepia officinalis*), GSA 17, 2008–2018			
	Observations	Model 6	Model 0	Improvement
1st.Qu	0	9.2	**6.13**	33%
Median	133.48	116.22	**125.45**	53%
3rd.Qu	558.44	**474.37**	465.51	−10%
R2	–	0.56	**0.61**	
MAE	–	240.96	**227.83**	5.44%
SA	235.8	273.21	**272.51**	1.87%
CGX	13.3	**13.29**	12.28	
CGY	44.66	**44.74**	44.75	−12.50%
D		**9.43**	10.2	−8.16%

	European hake (*Merluccius merluccius*), GSA 17, 18, 19, 2008–2018			
	Observations	Model 6	Model 0	improvement
1st.Qu	2.99	3.75	**3.67**	11%
Median	15.75	13.68	**14.82**	55%
3rd.Qu	34.68	25.96	**26.33**	4%
R2	–	0.32	0.32	
MAE	–	16.13	**15.94**	
SA	1552.02	2272.55	**2263.34**	1.30%
CGX	16.18	16.18	16.16	–
CGY	42.09	**42.18**	42.19	−11%
D	0	**10.89**	11.83	−8.63%

	Mantis shrimp (*Squilla mantis*), GSA 17, 2008–2018			
	Observations	Model 6	Model 0	improvement
1st.Qu	0	2.65	**1.22**	54%
Median	36.91	**25.23**	24.5	−6%
3rd.Qu	326.95	90.19	**115.48**	10%
R2	–	0.3	**0.46**	
MAE	–	202.69	**169.76**	
SA	192.95	352.34	**267.37**	53%
CGX	13.01	13.15	**13.01**	100%
CGY	44.26	44.15	**44.23**	72%
D		16.99	**2.93**	82%

	Common sole (*Solea solea*), GSA 17, 2008–2018			
	Observations	Model 6	Model 0	improvement
1st.Qu	127.07	**103.43**	92.21	−47%
Median	439.64	296.35	**302.75**	4.46%
3rd.Qu	1155.34	717.12	**728.82**	2.66%
R2	–	0.35	**0.44**	
MAE	–	538.92	**491.55**	
SA	311.66	447.79	**403.05**	32%
CGX	12.98	13.01	**12.97**	133%
CGY	44.57	44.6	**44.6**	0
D		16.99	**2.83**	83%

river plume with biomasses as high as 1500 kg/km^2 especially in the years 2011, 2012, 2018 (Figure 3.6.3, right panels).

The spatial and temporal distributions shown are coherent with previous results (Sartor et al. 2017). For example, results from Sion et al. (2019) on European hake show for 2011 and 2013 higher biomass values in the eastern-central Adriatic sea, while in 2015 a general lower biomass of this species was estimated, with similar outcomes to the ones we found in this paper (Figure 3.6.2). Tserpes et al. (2019) also highlights a biomass increasing trend for red mullet after 2008, which is in line with the recent stock assessment outcomes (GFCM 2019; STECF 2019). Similarly, Figure 3.6.2 highlights that this biomass increase corresponds to a spreading of the population in the study area.

The set of indicators for evaluating performances of the complete (model 0) or spatiotemporal (model 6 or 7) models contrasted with observations show that the integrated approach embedding biogeochemical, physical and fishing effort variables has improved performances (Table 3.6.1). In particular, indicators in Table 3.6.1 suggest that models' distribution statistics (quartiles and median) are closer to observed data when the integrated approach is used (i.e. the model 0). Exceptions are the first quartile for common sole, the third quartile for common cuttlefish, and the median for mantis shrimp. It is worth to note the relevant improvement of median values for hake and cuttlefish (+55% and +53%, respectively) when the spatial model of species distribution includes additional biogeochemical, physical and fishing effort data (Table 3.6.1). MAE and R^2 showed that consistency of model to the data improves for all species (except R^2 for European hake) when additional variables are included (Table 3.6.1). The spatial indicators used to evaluate the modelling results in terms of variations of the area occupied by the populations and their mean geolocation (e.g. Woillez et al. 2009) show improvements for red mullet, common sole, and mantis shrimp when the models include additional biogeochemical, physical and effort variables. For all these species the centroids of spatial distribution and the spreading area of the best model (model 0) are closer to those estimated on the observed data than to models with no additional explanatory variables (model 6 or 7; Table 3.6.1). For European hake and common cuttlefish, the spreading area improved when additional explanatory variables are included, but not the centroid position. This result and some low improvements of model 0 with respect to the model with spatiotemporal variables only is possibly related to complex influences of other environmental factors such as seabed type and habitats on the spatial distribution of species (in particular for European hake and common cuttlefish). Overall, the approach quantified the relevance of biogeochemical and physical variables derived from CMEMS and fishing effort from VMS/AIS in improving the spatial distribution of demersal species based on trawl survey data. Results highlight species-specific improvements that should be considered also in relation to the use of spatial distribution model (Brodie et al. 2020).

Key objectives of the Common Fisheries Policy (EU 2013) are the achievement of MSY in the short term and the implementation of an ecosystem approach to fisheries management which is often based on fishery indepedent data. Thus we consider that the integrated approach proposed here represents an important step for incorporating anthropogenic (fishing effort) and other environmental stressors (biogeochemical and physical variables) into the advice for fisheries management.

The improved models including enviromental and effort variables, in fact, can be used for a year by year evaluations of species distribution, for explaining and understanding species displacement. This is of paramount importance for a spatially based management of the resources that relies upon the identification of best fishing grounds, spawning or nursery areas, and generally aiming at defining fisheries managed areas (Lauria et al. 2017). The improved accuracy of species distributions based on enviromental and effort variables as obtained inthis study can potentially support co-management initiatives involving fisheries organisations and other stakeholders (e.g. those carried out by the Mediterranean Advisory Council, MEDAC). In particular, sharing such outcomes with the bottom trawl industry could lead to an increase in the awareness of the sector and consequently to the reduction of the alarming footprint of the fisheries in the Adriatic and Western Ionian Seas (Amoroso et al. 2018).

Furthermore, it is largely acknowledged that most of the presently used stock assessment models are too simplistic since they often consider species populations without integrating the role of key environmental drivers, which is a challenging but crucial frontier in the time of global changes. Taking into consideration environmental factors is also pivotal for the MSY objective, as climate change impacts on the fish community would require moving below fishing mortality at FMSY to ensure sustainable exploitation of marine stocks (Travers-Trolet et al., 2020). An optimised approach for the analysis of trawl survey data is relevant for the stock assessments and advices carried out by Scientific Advisory Committee of the General Fisheries

Commission for the Mediterranean Sea (SAC-GFCM) and the Scientific, Technical and Economic Committee for Fisheries (STECF) of the European Commission, as fisheries-independent data are essential for fishery management. The prediction of biomass indices on the whole domain over time with the integrated models proposed here takes into account the influence of relevant oceanographic variables and could be appropriately used for tuning stock assessment models such as, for example, surplus production models that need the catch time series and the survey abundance aggregated indices.

Since most analytical stock assessment models use survey indices by age or length as tuning indices, a further step for future insights is represented by the modelling in similar way also demographic indices, as length and/or age. Moreover, modelling of demographic indices can be useful also for progressing on the geolocation of sensitive life stages of the species, thus addressing further key questions of spatial fishery management.

These spatial distribution models for demersal species were developed for the best extension of trawl survey data to the whole study area from 2008 to 2018. The approach developed here highlights the relevance of integrating oceanographic variables in the analysis of trawl survey data before their use as inputs in stock assessment (Cao et al. 2017) and ecosystem modelling (see for example, Melaku Canu et al. 2010; Grüss et al. 2014; Grüss et al. 2018). This approach sets the basis for providing projections of the potential effects on species distribution and biomass of future environmental changes.

Applying the identified best GAMs models for making future predictions of species distribution is facilitated by the availability of oceanographic variables under future scenarios of climatic changes and appears strongly conditioned to assumptions on the future distribution of fishing effort that are also dependent from policies and regulations. Therefore, using the models developed here for making future scenarios might be considered with caution, needing further specific investigations of model validity to changed conditions. Yet the models can still provide a first order approximation of potential large scale effects, such as displacements of biomass centre of gravity and spreading area due, for example, to climate change. Although the relative distribution pattern might be well predicted by the model, many factors, such as recruitment success and species interactions for example are not included, thus efforts should be addressed in the future for testing additional modelling approaches and for improving the accuracy of these species distribution models.

In conclusion, the present study aims at investigating the influence of environmental variables on the biomass distribution of the most important commercial fishery species in the Adriatic and Western Ionian basin by modelling the data obtained from trawl surveys using different GAM approaches. GAMs are commonly used because they have the advantage of accounting for spatial and temporal autocorrelation of the data. The approach used here robustly demonstrates in which cases oceanographic variables extracted from CMEMS products and effort from VMS/AIS, result in improving species distribution models. Although there is still room for improvements, the work presented here is a remarkable starting point for better understanding species-environment relationships and for understanding the benefits of integrating the CMEMS variables into the modelling of fishery independent data for predicting the species distribution in the Adriatic and Ionian basins.

Acknowledgements

The MEDITS surveys have been carried out with the financial support of the European Commission until 2001 and subsequently within the Data Collection Framework. The SoleMon surveys have been carried out with the financial support of Italian Ministry of Agriculture (MIPAAF) and National Research Council (CNR) until 2016 and subsequently within the Data Collection Framework. The European Commission, the Member States of Italy and Croatia and the FAO AdriaMed Project are thankfully acknowledged. This work results from activities of the project FAIRSEA (Fisheries in the Adriatic Region – a Shared Ecosystem Approach) funded by the 2014–2020 Interreg V-A Italy – Croatia CBC Programme [Standard project ID 10046951].

Section 3.7. A benthic hypoxia index (BH-index) for assessing the Good Environmental Status of the Black Sea's north-western shelf waters

Authors: Arthur Capet, Luc Vandenbulcke, Catherine Meulders, Marilaure Grégoire

Product Table:

Ref. No.	Product name & type	Documentation
3.7.1	BLKSEA_REANALYSIS_BIO_007_005 Black Sea biogeochemistry hindcast	PUM: http://marine.copernicus.eu/documents/PUM/CMEMS-BS-PUM-007-005.pdf QUID: http://marine.copernicus.eu/documents/QUID/

(*Continued*)