

HABITAT QUALITY ESTIMATION FOR TURBOT ON THE ROMANIAN CONTINENTAL SHELF

Sorin Constantin^{1,2*}, George Țiganov³, Ioan-Daniel Șerban¹, Răzvan Mateescu³,
Marius Budileanu¹

¹*Terrasigna, 3 Logofat Luca Stroici Street, 020581 Bucharest, Romania*

²*Research Institute of the University of Bucharest,
90-92 Panduri Street, Bucharest, Romania*

³*National Institute for Marine Research and Development "Grigore Antipa"
300 Mamaia Blvd., 900581 Constanta, Romania*

**Corresponding author: sorin.constantin@terrasigna.com*

ABSTRACT

Turbot is one of the most important species for the Romanian fisheries, in terms of ecological and economical value. Its distribution can be indirectly characterized through the analysis of the habitat quality, given by various environmental factors. In this study, a regional MaxEnt calibrated model for turbot habitat assessment is proposed. It was built using in-situ occurrence data and oceanographic variables available through the Copernicus Marine Service. The obtained model has relatively good prediction capabilities and can be used for habitat suitability index (HSI) mapping. The differences in terms of spatial and temporal patterns for the HSI were shown using monthly products for a time span of ten years. Good and stable habitat conditions, from an inter-annual point of view, were identified and discussed.

Keywords: turbot, Species Distribution Modelling, Habitat Suitability Index, Romanian continental shelf

INTRODUCTION

The turbot is a demersal fish species. The accepted scientific name is *Scophthalmus maeoticus*. However, its taxonomic status remains debated (Cardinale *et al.*, 2021). Some studies and authorities consider it a distinct species (Turan *et al.*, 2019), found exclusively in the Black Sea, while others classify it as a subspecies or synonym of *Scophthalmus maximus* (e.g. Bailly and Chanet, 2010), which has a relatively large distribution in the European waters, from the Northeast Atlantic, throughout the Mediterranean and along the European coasts to the Arctic Circle. In the Black Sea it populates the continental shelf area, especially in the north-western part of the basin. Adults live on sandy or mixed bottoms and are commonly found in brackish waters. It feeds mainly on various bottom-living fishes, and, sometimes, on larger crustaceans and bivalves (Froese and Pauly, 2025). The marine habitat for turbot has been described by previous studies. Nita *et al* (2011) provide an in-depth description of the general habitat preferences for the species in the Black Sea, by making references to different behaviors, such as spawning or wintering. Their observations are based on localized in-situ environmental data and overall biological knowledge regarding the species.

Species Distribution Modelling (SDM) is a general term used to describe processes that try to link known occurrences of a species to environmental predictors in order to estimate its potential geographic range. By combining species presence

records (and, when available, absence) with spatially explicit variables, such as temperature, salinity and other oceanographic variables, SDM identifies ecological niches where the species is most likely to be found and persist. Common algorithms include machine - learning methods (e.g. MaxEnt) and statistical approaches (e.g. generalized linear or additive models) (Anderson *et al.*, 2003). The resulting models generate maps of habitat suitability, which can be used to assess current distributions, forecast range shifts under climate change scenarios, or inform conservation planning and resource management. For the following exercise, the Maximum Entropy Modeling (MaxEnt) approach was used.

The main objective of the current study was to calibrate a Species Distribution Model, using MaxEnt, for adult individuals of turbot that can be applied in the Romanian Black Sea area and to make brief observations on the spatial and temporal dynamics of the habitat quality, as given by the aforementioned model. Better knowledge regarding the turbot's potential distribution at different time periods can be used to take proper actions for an improved management of fisheries and also help avoid by-catches (such as cetaceans – Radu and Anton, 2014). Other studies have used MaxEnt to model the habitat quality of turbot in the Black Sea. For example, Zlateva *et al* (2023), using a similar approach to the one presented hereafter, in our study, found that the species' habitat preferences in the Bulgarian waters are strongly influenced by temperature and dissolved oxygen.

MATERIALS AND METHODS

MaxEnt is a machine-learning algorithm designed for presence-only species distribution modeling. It estimates the probability distribution of a population across a landscape by finding the distribution of maximum entropy (i.e., closest to uniform) subject to environmental constraints derived from known occurrence points (Elith *et al.*, 2011). By comparing environmental values at these locations to background data, MaxEnt identifies the combination of predictor variables that best distinguishes suitable from unsuitable habitats (Phillips *et al.*, 2006). The model optimizes feature weights through regularization to prevent overfitting and outputs a continuous habitat suitability map – the Habitat Suitability Index (HSI).

In order to train a MaxEnt model, two main input datasets are required: in-situ catch data showing occurrence records for a particular species (turbot, in our case) and associated environmental predictors.

In-situ catch data were collected by the Romanian National Institute for Marine Research and Development during the annual field campaigns. The scientific fishing activity consists of approximately 1-hour long profiles, using a demersal trawl. For the present study, a total of 685 such profiles were available (Fig. 1), between 2014 and 2022 (for each year), for 141 days. They overlap 26 unique months, as depicted in Table 1. As it can be observed, these are heterogeneously distributed along a calendar year, with in-situ measurements available only during late spring – beginning of summer (May and June) and for the cold season, between October and December. The average length of the profiles is 4.2 km, with a median value of 4.1 km. These were collected in

areas with a water depth ranging from 7 m to 71 m. For each profile, the total weight and the number of fish were recorded. Within the current study, these were used only as presence and absence information, for model training. These locally collected occurrences are extremely valuable, especially when considering the scarcity of such openly available datasets for the Black Sea area. While other regions are rich in such information, our study area lacks almost completely this sort of knowledge. For example, when consulting the Ocean Biodiversity Information System (OBIS, 2025), one of the best-known global databases that provides open-access marine biodiversity data, only 8 turbot occurrences were found for the entire Black Sea (out of almost 25000 available worldwide). Search was performed using both "*Scophthalmus maximus*" and "*Scophthalmus maeoticus*" (considered a subspecies of the first one) as keywords. Thus, publicly available datasets, while extremely useful in global applications, are less suitable for specific regional endeavors.

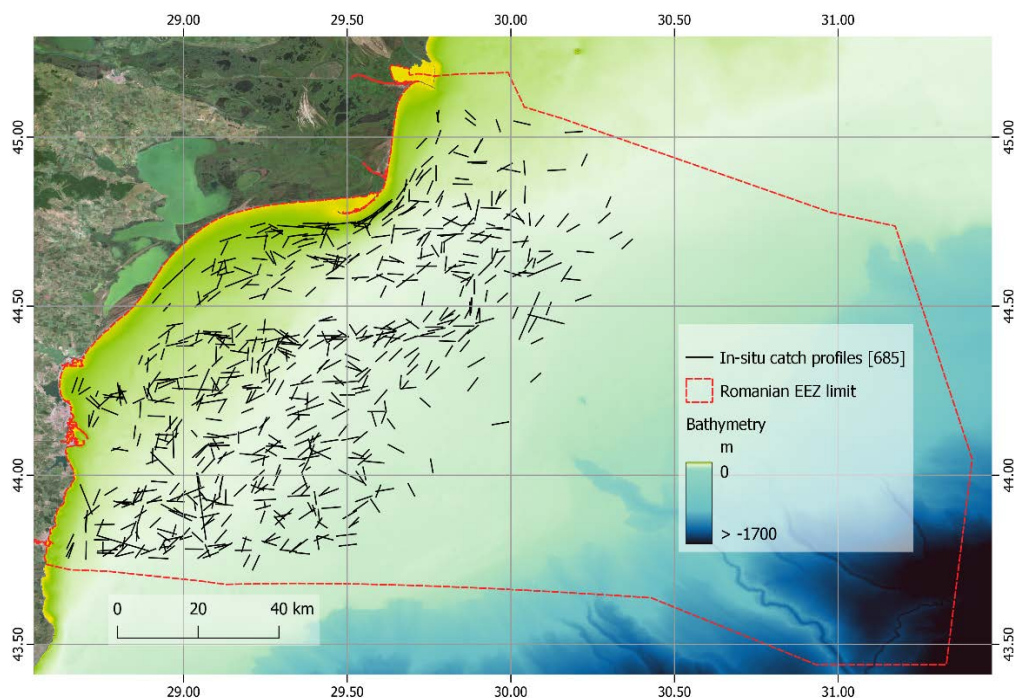


Fig. 1. Area of Interest, depicted by the distribution of available in-situ catch data as fishing profiles

The environmental parameters (predictors) used for the current study can be grouped into two categories, depending on the depth of the water column that they refer to: surface and bottom variables. The surface indicators make reference to the upper layers, at -1.5 m depth (thus could be considered as pseudo-surface), while the bottom ones characterize the environment at the sea floor. The first category comprises Sea Surface Temperature (SST, °C), Sea Surface Salinity (SSS, psu), Net Primary Production (NPP, mg m⁻³ day⁻¹), Sea Surface Height (SSH, m) and currents velocity (Cvel, m s⁻¹).

Table 1. Temporal distribution of the fishing profiles containing in-situ catch data

Month	Years	Number of profiles
May	2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022	279
June	2019, 2022	45
October	2017, 2018, 2021, 2022	113
November	2014, 2015, 2016, 2017, 2018, 2019, 2020, 2022	168
December	2015, 2016, 2019, 2020	80
		Sum = 685

The second category contains dissolved oxygen concentration on the shelf bottom (O2b, mmol m⁻³) and Sea Bottom Temperature (SBT, °C). All variables were acquired from the Copernicus Marine Service (CMEMS) and are part of the following products: Black Sea Physics Reanalysis - BLKSEA_MULTIYEAR_PHY_007_004 (Lima *et al*, 2020) and Black Sea Biogeochemistry Reanalysis - BLKSEA_MULTIYEAR_BGC_007_005 (Grégoire *et al*, 2020). These are the result of numerical modeling approaches and are delivered with multiple temporal granularities (from daily to climatology datasets) at 0.025° spatial resolution.

Since substrate characteristics are known to be of particular importance for the turbot distribution, the first model definitions included such information, as well. The EUSeaMap23 dataset from EMODnet was used, which comprises five different substrate types for our area of interest: sand, muddy sand, mixed sediment, sandy mud and fine mud. However, the results showed no discrimination capabilities for this particular dataset, which had an insignificant contribution to the overall characterization of the habitat. Thus, the substrate type was removed from the following model versions.

For model training, daily observations were employed. The final model was subsequently applied to monthly products to analyze the spatial and temporal patterns of habitat quality for turbot.

Another environmental predictor is represented by the water depth. Bathymetric data was obtained from GEBCO (General Bathymetric Chart of the Oceans) at approximately 500 m spatial resolution (GEBCO Bathymetric Compilation Group, 2024). A regional subset corresponding to our study area was extracted from the global dataset.

The first step was to create spatial and temporal match-ups between occurrence data and the environmental predictors. Out of the total 685 in-situ profiles, 487 had records of turbot (occurrences). These were first selected. Each in-situ profile (as a geospatial vector – line) is then correlated with the predictors (which come as raster files) from that particular day. Values are then extracted along the lines/profiles. Thus, depending on the length of a particular line, several pixels/cells (all that are intersected) values will be extracted and stored as relevant environmental predictors for turbot occurrences. Apart from the 487 profiles that were used as presence data, the remaining 198, with no recorded catch of turbot, served as locations for the extraction of background (pseudo-absences) information. The same procedure as for occurrences

was used for pseudo-absences extractions, as well. Since these are insufficient, in terms of total number (Merow *et al.*, 2013), the background data were supplemented by selecting 400 random days (out of a total of 796) from the months when catches are available. For each day, 25 random points were generated for background extractions, which resulted in 10000 points. A probability distribution mask was used to generate these random points, which assures that these locations are selected with a higher probability in areas where occurrences were recorded, while still using the other regions, but with less frequency. Finally, pixel values were extracted for each environmental predictor, for all random locations.

In the end, a MaxEnt model was trained using the above-mentioned selection of occurrences and background associated environmental predictors. For model evaluation purposes, 90% of the dataset was used for training, while 10% was kept for validation.

RESULTS AND DISCUSSION

The MaxEnt model, trained on 1458 presence records (161 kept for testing) and 11619 background points, achieved reasonable performance. Gain measures improvement over random guessing. Training gains (0.335 regularized, 0.450 unregularized) and a test gain of 0.402 show the model reliably predicts on new data. AUC (Area Under the Curve) measures a species distribution model's ability to distinguish presence from absence/background. Values range from 0.5 (random) to 1 (perfect discrimination). For our model, the training AUC was 0.81, with test AUC 0.80 (± 0.016), reflecting acceptable discrimination between presence and background and minimal overfitting. In other words, the AUC value indicates that the model correctly ranks true presences above background about 80% of the time, and the very small training–test AUC gap (0.01) together with a low test standard error (± 0.016) imply that this discriminatory ability is consistent and unlikely due to overfitting. Omission rates were low across common thresholds.

Variable contributions analysis reveal that sea surface temperature (SST) had the highest influence (20.8% contribution), followed by bathymetry (18.5%) and current velocity magnitude (Cvel; 17.0%). The others are at 11.5% (SSH), 9.7% (SSS), 9.2% (NPP), 8.6% (SBT) and 4.7% (O2b). Jackknife tests confirm SST's primacy: using SST alone yielded the highest gain, and omitting it caused the largest drop. Bathymetry similarly stood out on test data, indicating unique environmental information in these predictors. The utilization of pseudo-surface environmental parameters is related to one major assumption of this study, which implies that variables will have an impact on the bottom conditions, as well, to a certain extent. This is partially sustained by the results, which suggests that these pseudo-surface information have a higher relative influence on the predictions. However, the uncertainties associated to each dataset are not accounted for. Thus, it is not reasonable to argue that bottom parameters are less influential, in general, but rather only when these particular datasets are considered.

The regularization multiplier was set to 0.5, balancing model complexity and generalization. Overall, with test AUC ≈ 0.80 , low omission rates, and modest training-

test metric gaps, the model effectively mitigated overfitting and yields a fairly robust representation of habitat suitability.

As previously mentioned, the inclusion of substrate type information into the model was assessed. The very low contribution of this particular variable to the model discrimination capabilities can be linked to the subtle differences between bottom sediments found in our particular area of interest and in our in-situ occurrence data, considering the general preferences of turbot for soft substrates. Most the the studied region overlaps such types (sand 1.10%, muddy sand 2.83%, mixed sediment 42.27%, sandy mud 29.65% and fine mud 23.93%). Only 0.22% is covered by “rock or other hard substrata” and “coarse substrate” categories, in limited areas, very close to the shoreline. None of our in-situ occurrences data overlap them. For the other classes, the ratio between the number of recorded in-situ occurrences and in-situ absences is relatively similar. This shows that, given our in-situ dataset, there is no clear preference for one type of soft substrate over another. The Maxent model was then applied to monthly predictor products for the 2013-2022 period. This resulted in monthly HSI maps, highlighting the quality of the habitat for turbot and, indirectly, the probability of the species to be found in specific areas. Since the occurrences used for the definition of the methodology are representative only for five months (May to June and October to December), the results shown hereafter will make reference only to these particular periods of the year. While the extrapolation of the model to other months, especially the ones close to the training period (e.g. July), could be considered feasible, we do not recommend such practice, unless the outputs can be validated with independent new data.

The monthly HSI products were masked using the bathymetry data, according to the maximum depth coming from the occurrences information and in agreement with the envelope proposed by the AquaMaps initiative, which is -70 m (Kesner-Reyes *et al.*, 2020). To highlight how the model estimates beyond this depth, a slightly higher threshold was selected for masking out deep regions: -80 m.

Maps in Fig. 2 show the multi-annual monthly averages for the estimated turbot HSI, computed for a ten-years period (2013-2022). The index values range is between 0 and 1. Low numbers (orange shades) mark poor environmental conditions for turbot, while green to blue colors depict good and very good habitat suitability. Areas with water depth below -70 m are always predicted as unsuitable conditions. Close to the coastline and in front of the Danube mouths, the model also predicts a less favorable habitat.

The best conditions are usually located in the middle of the study region, in areas with water depths between approximately -40 m and -70 m. The highest values for the index can be observed in May (median HSI of 0.60), October (median HSI of 0.57) and November (median HSI of 0.58), while the lowest ones were estimated for June (median HSI of 0.36) and December (median HSI of 0.43). During May, the optimum conditions are located to the south of the Romanian littoral. These favorable areas migrate north during October and especially November.

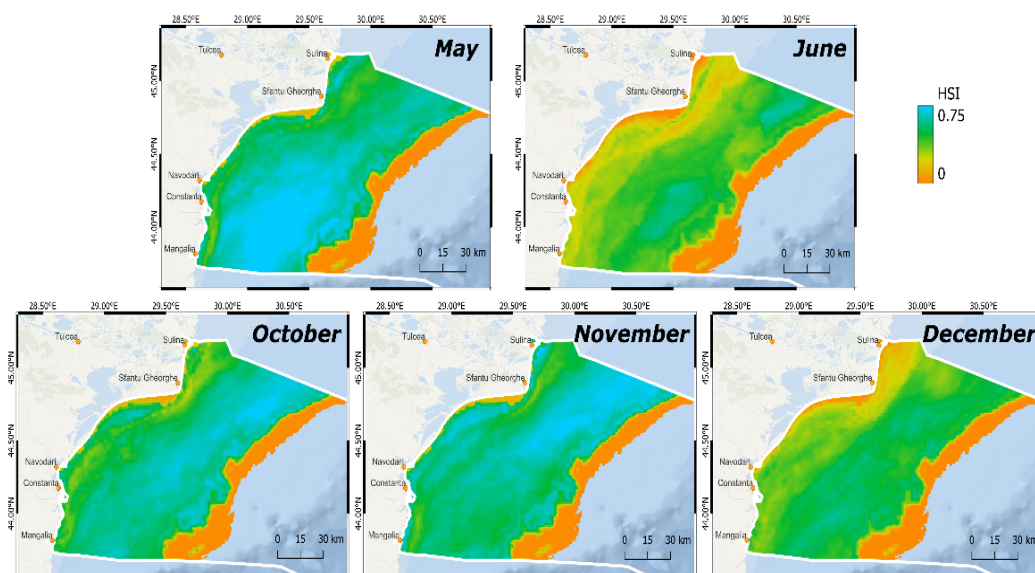


Fig. 2. Multi-annual (2013-2022) monthly averages for the turbot Habitat Suitability Index (HSI), as predicted by the regional calibrated model

Out of the total investigated region, between the 0 m and -80 m isobaths, the surface characterized by above average HSI (> 0.50) occupies an area between 64% and 75% for the best months (May, October and November) and between 10% and 29% for the less favorable period (June and December).

In order to complement the information provided by the multi-annual averages of the HSI, the standard deviation, for each pixel, was also computed (Fig. 3).

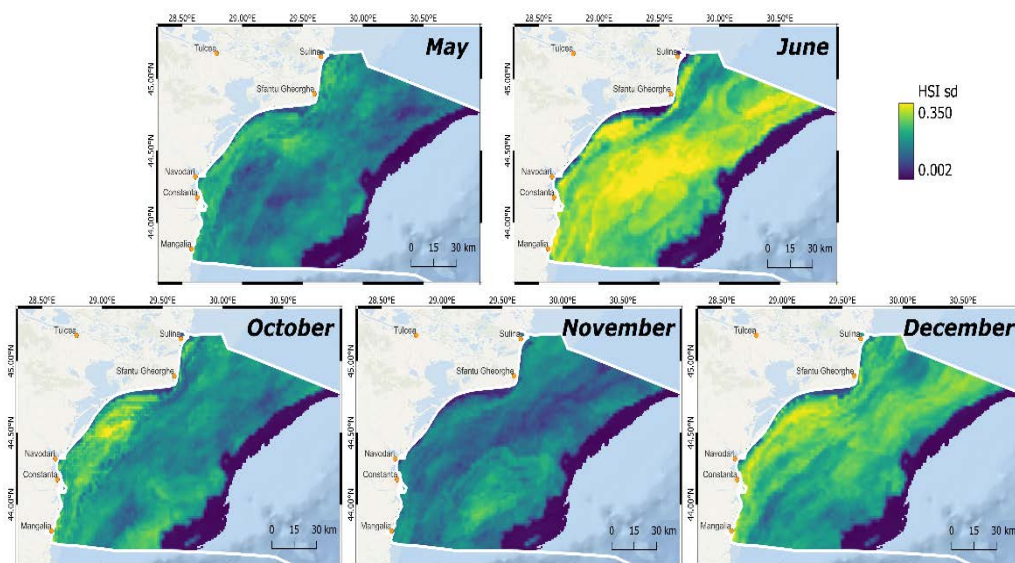


Fig. 3. Multi-annual (2013-2022) monthly standard deviation for the turbot Habitat Suitability Index (HSI), as predicted by the regional calibrated model

These results show the spatial patterns of the inter-annual dynamics of the HSI. Low values overlap areas where the habitat quality remains relatively constant from one year to another, while high values depict regions where the environmental conditions change significantly from one year to another.

Overall, the coastal zones are the most dynamic from this point of view, together with the ones in front of the Danube mouths. The month of June shows the highest inter-annual variability for the analyzed period, followed by December.

By combining the two results shown above, the HSI and its standard deviation, it is possible to showcase various functional areas. This has been done by normalizing both the multi-annual HSI (HSI_N) and the HSI standard deviation (HSI_{sd_N}) and then computing the ratio between them (HSI_N / HSI_{sd_N}). The breakpoints for the classes shown in Fig. 4 were computed by taking the 25th, 50th and 75th percentiles of all the values for this ratio, from all five analyzed months. Therefore, Class 1 can be defined as areas where the habitat quality is both high and stable from one year to another. These could be considered as the most important regions for the turbot's environmental conditions. The last one, Class 4, represents the opposite: low HSI values, combined with high inter-annual changes. Thus, the lowest quality in terms of habitat conditions for turbot, at high temporal scales. Classes 2 and 3 show intermediate conditions.

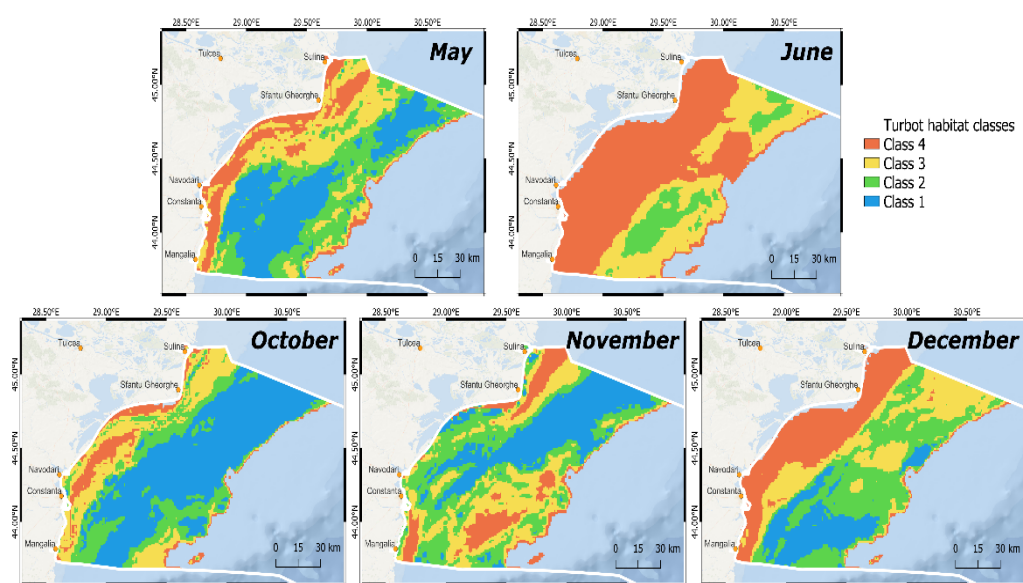


Fig. 4. Multi-annual (2013-2022) monthly classes of the turbot habitat quality, computed based on the average HSI and its standard deviation

The spatial extent of Class 1 was also determined, since it can be considered that, in these regions, there is the highest probability to regularly encounter favorable habitat conditions for the turbot, each year. Therefore, it is important to know how many square kilometers can be characterized by such conditions, each month, given the regional calibrated model proposed in this study. The largest extension of Class 1 can

be observed in October ($\approx 7780 \text{ km}^2$), followed by May ($\approx 4870 \text{ km}^2$), November ($\approx 4490 \text{ km}^2$) and December ($\approx 2480 \text{ km}^2$). For the month of June, Class 1 is not present.

CONCLUSIONS

The current study represents a first attempt to define and calibrate a regional ecological model that can be used to estimate the habitat conditions for turbot. The quality of the results shows a high potential to use such approaches for fisheries management activities. However, more in-situ data would be required in order to improve the methodology towards a fully operational product. The proposed model was built taking into consideration the occurrences and potential distribution of adult individuals. Thus, it does not take into account the specificities related to the reproduction migration or any juvenile behavior. Also, the model may be biased for environmental conditions that fall at the temporal edges of the observed occurrences. If we recall that temperature is the most important predictor, then June and December might represent the less common conditions given the available in-situ data: too warm or too cold. Such limitations of the model can only be accounted for and mitigated if more occurrence information becomes available in the future, which can be used to characterize a broader range of environmental conditions. The inclusion and evaluation of a more accurate substrate type dataset might also improve a similar model.

While the examples shown in this paper rely on monthly datasets, the model can be applied to any level of temporal granularity, such as daily observations, which can provide detailed insights regarding intra-monthly changes of the habitat quality. From a climatological point of view, we analyzed a relatively short period of time, between 2013 and 2022. However, applying such a methodology to longer time series (e.g. 30 years) could be used to derive information about general trends. Areas with degrading habitats could be highlighted, as well as regions where the environmental conditions are improving. The latter ones might be used for the designation of marine protected areas for the purpose of protecting the turbot by offering a refuge from fishing activities.

The results shown in this study provide an initial perspective upon the quality of the environmental habitat for turbot, at multi-annual monthly level, based on a regional calibrated ecological model. The quantity and quality of the in-situ data used to calibrate this model are one of the key elements of the work performed. Further activities can be dedicated to improving the in-situ database, by adding other time periods to the list, and to enhance such modelling approaches.

Acknowledgement. This research received funding from the European Commission's Horizon 2020 Research and Innovation Program under GA no. 101037643—ILIAD ("Integrated Digital Framework for Comprehensive Maritime Data and Information Services"). Partial activities were supported by the Romanian Ministry of Research, Innovation and Digitalization (grant no. 760054-JUST4MPA), in the PNRR-III-C9-2022-I8 call.

REFERENCES

- Anderson, R. P., Lew, D., Peterson, A. T. (2003). Evaluating predictive models of species' distributions: Criteria for selecting optimal models. *Ecological Modelling*, 162, <http://biodi.sdsc.edu/>
- Bailly, N., Chanet, B. (2010). *Scophthalmus Rafinesque, 1810: the valid generic name for the turbot, S. maximus (Linnaeus, 1758) [Pleuronectiformes: Scophthalmidae]. Cybium*, 34(3): 257-261. <https://doi.org/10.26028/CYBIUM/2010-343-003>
- Cardinale, M., Chanet, B., Martínez Portela, P., Munroe, T.A., Nimmegeers, S., Shlyakhov, V., Turan, C., Vansteenbrugge, L. (2021). *Scophthalmus maximus*. The IUCN Red List of Threatened Species 2021: e.T198731A144939322. <https://dx.doi.org/10.2305/IUCN.UK.2021-2.RLTS.T198731A144939322.en>
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1):43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Froese, R., Pauly, D. (Ed). (2025). *FishBase*. World Wide Web electronic publication. <https://www.fishbase.org> (accessed 04/2025)
- GEBCO Bathymetric Compilation Group. (2024). The GEBCO_2024 Grid – a continuous terrain model of the global oceans and land. *NERC EDS British Oceanographic Data Centre NOC*. <https://doi.org/10.5285/1c44ce99-0a0d-5f4f-e063-7086abc0ea0f>
- Grégoire, M., Vandenbulcke, L., Capet, A. (2020). Black Sea biogeochemical reanalysis (CMEMS BS-Biogeochemistry) (Version 1) [Data set]. *Copernicus Monitoring Environment Marine Service (CMEMS)*. https://doi.org/10.25423/CMCC/BLKSEA_REANALYSIS_BIO_007_005_BAMHBI
- Kesner-Reyes, K., Garilao, C., Kaschner, K., Barile, J., Froese, R. (2020). AquaMaps: Algorithm and data sources for marine organisms. In R. Froese, D. Pauly (Ed.), *FishBase*. <https://www.fishbase.org> (version 10/2019)
- Lima, L., Aydogdu, A., Escudier, R., Masina, S., Ciliberti, S. A., Azevedo, D., Peneva, E. L., Causio, S., Cipollone, A., Clementi, E., Cretí, S., Stefanizzi, L., Lecci, R., Palermo, F., Coppini, G., Pinardi, N., Palazov, A. (2020). Black Sea physical reanalysis (CMEMS BLK-Physics, E3R1 system) (Version 1) [Data set]. *Copernicus Monitoring Environment Marine Service (CMEMS)*. https://doi.org/10.25423/CMCC/BLKSEA_MULTIYEAR_PHY_007_004
- Merow, C., Smith, M. J., Silander, J. A. (2013). A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36(10):1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>
- Niță, V., Diaconescu, S., Zaharia, T., Maximov, V., Nicolae, C., Micu, D. (2011). The characterization of the main habitat types populated by the Black Sea turbot in its different stages of development. *AACL Bioflux*, 4(5): 552-570.

- OBIS, (2025). *Ocean Biodiversity Information System*. Intergovernmental Oceanographic Commission of UNESCO. <https://obis.org>
- Phillips, S. B., Anderson, R. P., Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **6**(2–3):231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Radu, G., Anton, E. (2014). Impact of turbot fishery on cetaceans in the Romanian Black Sea area. *Sci. mar.*, 78(S1):103–109. <https://doi.org/10.3989/scimar.04029.27A>
- Turan, C., Ivanova, P., Gürlek, M., Yağlioğlu, D., Ergüden, D., Karan, S., Doğdu, S.A., Uyan, A., Öztürk, B., Nikolov, V. and Raykov, V. (2019). Phylogenetic relationships of turbot species (Scophthalmidae) inferred from the mitochondrial COIII gene and morphological characters. *NESciences* 4: 28-41
- Zlateva, I., Ivanova, P., Dzhembekova, N., Doncheva, V., Popov, I., Slabakova, V., Raev, Y., Raykov, V., Dimitrov, D. (2023). Spatial distribution and genetic diversity of turbot (*Scophthalmus maximus*, Linnaeus, 1758) in Bulgarian Black Sea waters relative to fishing pressure and their abiotic environment. *J.Mar.Sci.Eng.* 11:1982. <https://doi.org/10.3390/jmse11101982>