 doi:10.20944/preprints202008.0424.v1

A Fuzzy Inference System for Seagrass Distribution Modeling in the Mediterranean Sea

Dimitra Papaki, Nikolaos Kokkos and Georgios Sylaios *

Democritus University of Thrace, Department of Environmental Engineering, Laboratory of Ecological Engineering and Technology, Xanthi, Greece; dimitr595papaki@gmail.com ; nikolaoskokkos@gmail.com

* Correspondence: gsylaios@env.duth.gr; Tel.: (+30-25410-79398)

Abstract: A Mamdani-type fuzzy-logic model has been developed to link Mediterranean seagrass abundance to the prevailing environmental conditions. Big Databases, as UNEP-WCMC (seagrass abundance), CMEMS and EMODnet (oceanographic/environmental) and human-impact parameters were utilized for this expert system. Model structure and input parameters were tested according to their capacity to accurately predict seagrass families at specific locations. The optimum FIS comprised of four input variables: water depth, sea surface temperature and nitrates and bottom chlorophyll-a concentration, exhibiting fair accuracy (76%). Results illustrated that *Posidoniaceae* prefers cool (16-18°C) and low chlorophyll-a presence (< 0.2 mg/m^3); *Zosteraceae* favors cool (16-18°C) and mesotrophic waters (Chl-a > 0.2 mg/m^3), but also slightly warmer (18-19.5 °C) with lower Chl-a levels (< 0.2 mg/m^3); *Cymodoceaceae* lives from warm, oligotrophic (19.5-21.0°C and Chl-a < 0.3 mg/m^3) to moderately warm mesotrophic sites (18-21.3°C and 0.3 – 0.4 mg/m^3 Chl-a). Finally, *Hydrocharitaceae* thrives in warm Mediterranaean waters (21-23°C) of low chlorophyll-a content (< 0.25 mg/m^3). Climate change scenarios showed that *Posidoniaceae* and *Zosteraceae* tolerate bathymetric changes, *Posidoniaceae* and *Zosteraceae* are mostly affected by sea temperature rise, while *Hydrocharitaceae* exhibits tolerance in higher sea temperature rise. This FIS could be used by national and regional policy-makers and public authorities.

Keywords: seagrass; fuzzy inference system; modeling; species abundance; Mediterranean Sea

1. Introduction

Seagrasses, the only submerged marine plants with an underground root and rhizome system forming beds and meadows, play a key role in the ecosystem services of the global coastal zone in relation to nutrients biogeochemical cycling, carbon sequestration, sediments stabilization, fish sheltering and food-web structure [1]. Although the global species diversity of seagrasses is relatively low (< 60 species), understanding their distribution is particularly important for ecologists developing bioregional models, covering all oceans and climatic zones together with the respective species assemblages [2]. In parallel, most aquatic ecosystem health assessment studies rely on seagrass species richness and distribution, since these serve as valuable bio-indicators reflecting prompt environmental changes, especially pollutants release and eutrophication events. *Halophila minor* and *Halophila ovalis* act as bio-indicators for trace metals pollution and sediments accumulation [3], *Zostera marina* acts as eutrophication indicator [4], while the genus *Cystoseira* act as heavy metal bioaccumulative and tolerant bioindicator of pollution [5], rapid coastal development and human intervention [6]. Finally, *Posidonia oceanica* meadows is directly linked to the degree of human impact.

Based on the above, it evident that seagrass species face significant challenges due to their high vulnerability related to a variety of anthropogenic disturbances concentrated along the coastal zone [2], and their sensitivity in environmental changes driven by climate change [7], ultimately leading to their global decline [8]. Focusing in the Mediterranean bioregion, there are present nine seagrass species; however, in western Mediterranean Sea, 4 of 5 species present appear declining in population, implying that priority conservation policies are needed. To unveil the complex interrelation between oceanographic, environmental, morphodynamic and human impact conditions on one hand and the seagrass species distribution on the other, modern data-driven models have been developed and implemented, following Machine Learning and Artificial Neural Networks techniques [9-10]. Such models are capable to explore the seagrass presence/absence dynamics, and link seagrass genes and species to the main environmental drivers determining their distribution, thus providing the hidden preferences affecting their abundance.

In the present work, fuzzy logic modeling will be employed to explore the nonlinear dynamics among the environment-ecosystem-human gradient and their impact on seagrass presence and distribution in the Mediterranean Sea. Such model will further be utilized to assess the response of the main Mediterranean seagrass species on the influence of climate change (gradual water temperature increase, sea level rise) and eutrophication. Such FIS model could be used by policy-makers and public authorities responsible for the Marine Strategy Framework Directive implementation.

The work was conducted in the framework of the Horizon 2020 ODYSSEA Project. One of the aims of this project is develop the appropriate algorithmic machinery, never previously experimented, aggregating data from diverse databases, attempting at establishing hidden relations among parameters, identifying repeated patterns, behaviors and trends.

2. Materials and Methods

2.1. The Seagrass Dataset

The dataset includes 1,771 locations in which the major seagrass families (*Cymodoceaceae*, *Zosteraceae*, *Posidoniaceae*, *Hydrocharitaceae*, *Ruppiaceae*) were observed, which were further divided at species level, covering the whole Mediterranean Sea. This initial dataset on seagrass species distribution was provided by the UNEP-WCMC (United Nations Environment Programme - World Conservation Monitoring Centre) [11]. The dataset illustrates the global distribution of seagrass species and has the form of a geo-referenced shapefile. Data were filtered and only instances located in the Mediterranean Sea were kept (Table 1, [12]). A limited number of problematic points was identified, based mostly on the depth zone distribution and the distance to coast. Furthermore, *Ruppiaceae* data were excluded from present analysis due to its limited occurrence (<2.0%).

Table 1. Seagrass genes presence in UNEP-WCMC dataset (last column presents seagrasses at species level).

Seagrass Gene	Instances	Percentage	Seagrass Species
Cymodoceaceae	1,337	75.49%	*C. nodosa*
Zosteraceae	187	10.56%	*Z. noltii, Z. marina*
Posidoniaceae	125	7.07%	*P. oceanica*
Hydrocharitaceae	94	5.30%	*H. stipulacea*
Ruppiaceae	28	1.58%	*R maritima, R. cirrhosa*

2.2. The Environmental Drivers Dataset

The distribution of seagrasses in the Mediterranean Sea is assumed to vary in relation to a series of physical, chemical, biological, seabed- and human-related parameters. For each geolocation in which seagrass species is reported, hydrographic data (water temperature, salinity, currents, waves) and water quality data (nutrients, dissolved oxygen, chlorophyll-a, net primary production rates) at the surface and bottom of the water column were retrieved from the Copernicus Marine Environmental Service (CMEMS) data products. These are gridded, mean-monthly oceanographic data, covering the whole Mediterranean Sea during the period 1987 to 2015 (Table 2). The only exception was the wave dataset, that was reported on hourly basis and subsequently was converted into mean-monthly values.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 August 2020 doi:10.20944/preprints202008.0424.v1

Table 2. Hydrographic and water quality parameters retrieved from CMEMS and periods these data cover. S, B represent Surface/Bottom values.

Parameter [units]	CMEMS Data Product	Period
Significant Wave Height [m] [S]	MEDSEA_HINDCAST_WAV_006_012	2006 – 2015
Water Velocity [m/s] [S, B]	MEDSEA_REANALYSIS_PHYS_006_004	1987 – 2015
Water Temperature [degC] [S, B]	MEDSEA_REANALYSIS_PHYS_006_004	1987 – 2015
Salinity [psu] [S, B]	MEDSEA_REANALYSIS_PHYS_006_004	1987 – 2015
Chlorophyll-a [mg/m^3] [S, B]	MEDSEA_REANALYSIS_BIO_006_008	1999 – 2015
Nitrate [mmol/m^3] [S, B]	MEDSEA_REANALYSIS_BIO_006_008	1999 – 2015
Phosphate [mmol/m^3] [S, B]	MEDSEA_REANALYSIS_BIO_006_008	1999 – 2015
Dissolved Oxygen [mmol/m^3] [S, B]	MEDSEA_REANALYSIS_BIO_006_008	1999 – 2015
Net Primary Production Rate [mol/m^3/s] [S, B]	MEDSEA_REANALYSIS_BIO_006_008	1999 – 2015

Other parameters potentially affecting the distribution of seagrass species, as the water depth and the substrate conditions (varying from mud to rock) were provided by the European Marine Observation and Data Network (EMODnet).

Finally, human impact parameters influencing the Mediterranean seagrass distribution, as the distance of each seagrass observation point to the closest point of human influence (port, coastal city, river mouth, distance to coast) was computed, using the haversine distance method.

The final dataset consisted of 573 data points, including the four main marine seagrass families of the Mediterranean Sea, namely *Zosteraceae*, *Hydrocharitacheae*, *Cymodoceaceae* and *Posidoniaceae* together with the mean-monthly values of a series of physical, water quality, biological, seabed and human impact parameters that act as drivers to explain marine flora distribution and response to environmental changes. This dataset was randomly divided into two parts: the 90% of data was defined as training dataset while the remaining 10% acted as validation dataset.

2.3. The Fuzzy Inference System

A Fuzzy Inference System (FIS) is an engine that applies reasoning to compute fuzzy outputs, and involves a knowledge base which defines rules and membership functions (MFs). The system is built using a set of ‘if–then’ rules having the general form “If x is A then z is C”, where A and C are linguistic values defined by fuzzy sets in the universes of discourse X and Z, respectively. The if-part and the then-part are called the antecedent and the consequent of a rule, respectively.

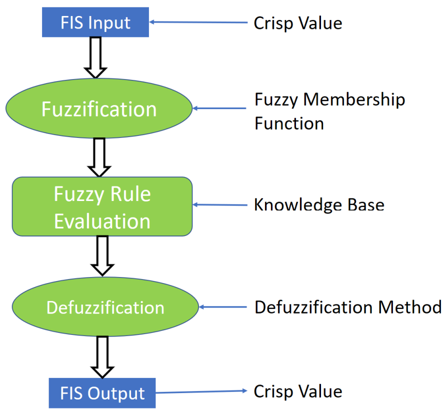

Figure 1. Fuzzy Inference System Procedures.

In this work, a Mamdani-type fuzzy inference system was developed, which comprises of fours steps (Figure 1): fuzzification of input variables, rules construction and evaluation, aggregation of rules output and defuzzification. The first step involves taking the crisp inputs, i.e. each numerical value, and the determination of the degree to which this value belongs to the appropriate fuzzy sets (defined here as Very Low, Low, Medium, High and Very High) through the selected membership function (in this case a trapezoidal membership function). After fuzzification, the second step involves taking the fuzzified inputs and applying them to the antecedents in a series of constructed fuzzy rules. If a given fuzzy rule has multiple antecedents, then the fuzzy operator (in this case AND) is used to obtain a single number that represents the result of the antecedent evaluation. This number (the true value) is then applied to the consequent membership function. The third step involves the aggregation, a process that produces an overall output by considering the membership functions of all rule consequents and combining them into a single fuzzy set. Finally, defuzzification is the process which leads to a final output as a crisp number. The input in the defuzzification process is the aggregate output fuzzy set and the output is a single number. There are several defuzzification methods, but the most popular one is the centroid technique [13, 14], which is followed here.

The herein developed FIS consists of four physical, chemical, biological, seabed- and human-related input parameters, a series of well-designed fuzzy rules, based on the most frequent parameters interrelation in the training set, and receiving one aggregate value representing the seagrass family abundance. Such FIS should be able to respond to the following question: "Given a series of environmental driving conditions prevailing in an area, what is the seagrass family favored by these conditions?".

The trapezoidal membership functions used to transform crisp inputs into membership values (step 1) belonging to the various fuzzy sets of the FIS have the form:

$$trapezoidal(x;a,b,c,d) = \max\left(\min\left(\frac{x-a}{b-a},1,\frac{d-x}{d-c}\right),0\right) \quad (1)$$

where a, b, c and d are the membership function parameters, as shown in Table 3 for indicative environmental parameters included in the FIS (Figures 2 and 3).

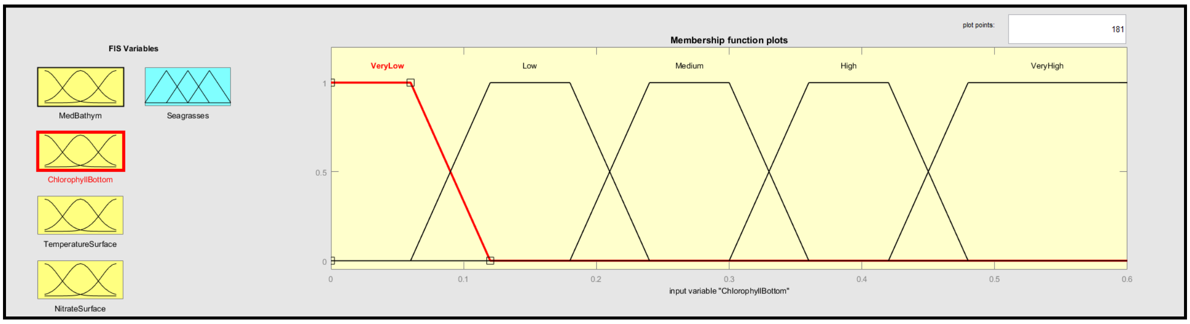

Figure 2. Trapezoidal membership functions for the fuzzy sets of parameter Bottom Chlorophyll-a imported in the FIS.

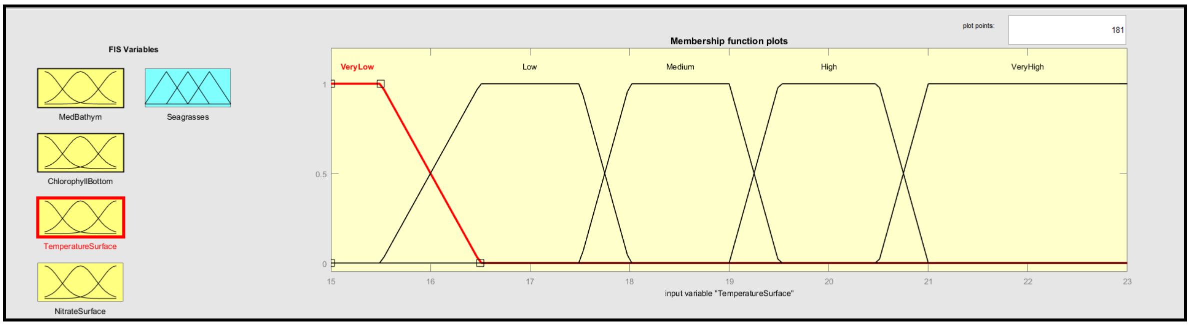

Figure 3. Trapezoidal membership functions for the fuzzy sets of parameter Surface Temperature imported in the FIS.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 August 2020 doi:10.20944/preprints202008.0424.v1

Table 3. Parameters for membership functions used in the fuzzy inference system.

Parameter	Very Low			Low				Medium				High				Very High		
	a=b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c=d
Water Depth [m]	0	10	20	10	20	30	40	30	40	80	100	80	100	120	130	120	130	150
Distance from Cities [km]	0.01	0.26	0.63	0.62	0.63	0.87	1.25	0.87	1.25	1.50	1.80	1.50	1.80	2.10	2.30	2.10	2.30	2.70
Distance from River Mouths [km]	0.12	1.00	2.39	1.00	2.39	3.30	4.66	3.30	4.66	5.57	6.94	5.57	6.94	7.84	8.75	7.84	8.75	10.00
Distance from Ports [km]	0.00	0.12	0.25	0.12	0.25	0.37	0.52	0.37	0.52	0.62	0.75	0.62	0.75	0.87	1.00	0.87	1.00	1.20
Distance from Coast [km]	0.00	0.13	0.34	0.13	0.34	0.48	0.68	0.48	0.68	0.82	1.00	0.82	1.00	1.17	1.31	1.17	1.31	1.50
Surface Temperature [degC]	15.0	15.5	16.5	15.5	16.5	17.5	18.0	17.5	18.0	19.0	19.5	19.0	19.5	20.5	21.0	20.5	21.0	23.0
Bottom Temperature [degC]	11.0	12.0	13.0	12.0	13.0	14.0	15.0	14.0	15.0	17.0	18.0	17.0	18.0	20.0	21.0	20.0	21.0	22.0
Bottom Salinity [psu]	36.0	36.3	36.7	36.3	36.7	37.0	37.4	37.0	37.4	38.1	38.5	38.1	38.5	39.2	39.6	39.2	39.6	40.0
Surface Nitrates [mmol/m^3]	0.0	1.5	3.5	1.5	3.5	5.0	7.0	5.0	7.0	9.0	11.0	9.0	11.0	13.0	15.0	13.0	15.0	16.0
Bottom Nitrates [mmol/m^3]	0.0	1.4	3.4	1.4	3.4	4.8	6.9	4.8	6.9	8.2	10.3	8.2	10.3	11.7	13.1	11.7	13.1	15.0
Bottom Phosphates [mmol/m^3]	0.0	0.04	0.08	0.04	0.08	0.11	0.15	0.11	0.15	0.19	0.23	0.19	0.23	0.26	0.30	0.26	0.30	0.38
Bottom Dissolved Oxygen [mmol/m^3]	145	156	167	156	167	178	189	178	189	200	211	200	211	222	233	222	233	255
Bottom Chl-a [mg/m^3]	0.00	0.06	0.12	0.06	0.12	0.18	0.24	0.18	0.24	0.30	0.36	0.30	0.36	0.42	0.48	0.42	0.48	0.60

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 August 2020 doi:10.20944/preprints202008.0424.v1

The IF-THEN fuzzy rules constructed according to the most frequent interrelation of input parameters observed in the training set (step 2) have the form: Rule 1: "If the water depth is Low, AND the water temperature at the sea bottom is Low, AND the chlorophyll-a concentration at the sea bottom is Low, AND the nitrates at the sea surface is Very Low THEN then the Seagrass Family favored in these conditions is *Posidoniaceae*". A set of similar fuzzy rules may be developed and inserted in the FIS. When feeding the FIS with more data, i.e., developing and adding more rules, the model provides more accurate estimates of the favored Seagrass Family abundance. In our case, 45 such rules were used in total, in order to associate the most frequently appearing independent antecedents in the training sets with a rule.

The seagrass family is the output parameter of the herein developed FIS, using a triangular membership function, of the form:

$$triangular(x;a,b,c)=\max\left(\min\left(\frac{x-a}{b-a},\frac{c-x}{c-b}\right),0\right) \quad (2)$$

The defuzzification membership functions of the FIS are discrete, as the result should a definite seagrass species. As explained, the final FIS output is a crisp number, thus according to the herein defined schema, if the output ranges in the interval [0, 0.25] then the seagrass family is *Zosteraceae*; in case the FIS output ranges in the interval (0.25, 0.50] then the seagrass family is *Hydrocharitaceae*; in the interval (0.50, 0.75] then it is *Cymodoceaceae* and finally in the range (0.75, 1.00] then it is *Posidoniaceae* (Figure 4).

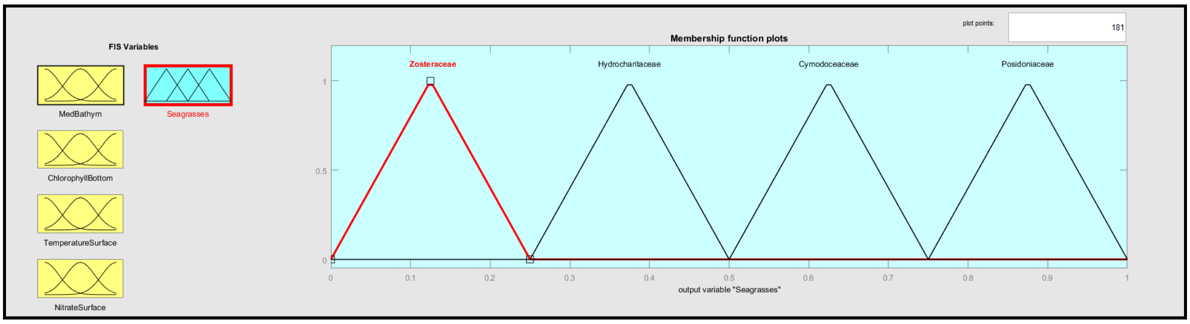

Figure 4. Triangular membership functions for the fuzzy sets of parameter Seagrass Family exported from the FIS.

Several experimental Fuzzy Inference Systems were developed, comprised of different combinations of the four input parameters and keeping constant the one output parameter, i.e., the seagrass family. The FIS was developed using the Fuzzy Logic Toolbox operating under MATLAB 7.0 (The Mathworks, Inc., Natick, MA, USA).

2.4. Evaluation Metrics

The most commonly used measure to evaluate this type of FIS model results is the Classification Accuracy of the system, being the fraction of relevant/correct instances to the whole validation dataset. Accuracy is a good metric for balanced datasets, like the case studied here for seagrass family abundance, and is defined as:

$$Accuracy=\frac{MS}{MS+MF} \quad (3)$$

where MS represents the number of model successes, i.e., the number of true (correct) assessments and MF the number of failures (wrong) assessments produced by the examined FIS.

3. Results

3.1. Data Analysis

Seagrass geolocation data were imported and mapped using a QGIS (QGIS 3.14 Pi). The spatial distribution of *Cymodoceaceae* in the Mediterranean Sea is illustrated in Figure 5.

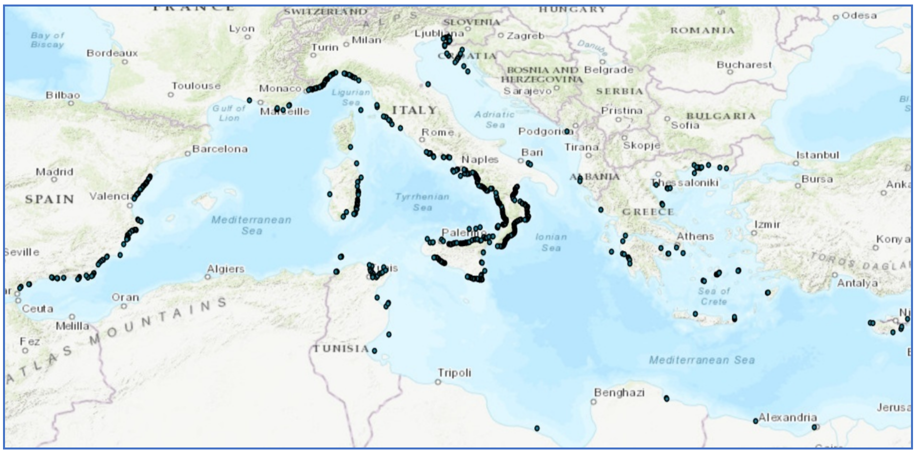

Figure 5. Spatial distribution of the seagrass family *Cymodoceaceae* in the Mediterranean Sea (UNEP-WCMC database).

The *Cymodoceaceae* records are distributed by 35% in the Tyrrhenian Sea, 29% in the Alboran Sea and 20% in the Balearic Sea. The *Posidoniaceae* data are located by 52% in the Alboran Sea, 30% in the Balearic Sea and 8% in the Tyrrhenian Sea.

The statistical parameters of the physicochemical data collected from the CMEMS databases for these points are summarized in Table 4.

Table 4. Summary statistics for the environmental parameters in the locations of *Cymodoceaceae* (CMEMS database).

	Temperature [degC]	Salinity [psu]	Dissolved Oxygen [mmol/m³]	Nitrate [mmol/m³]	Chlorophyll-a [mg/m³]
Mean	17.62	38.10	230.38	0.52	0.12
Stand Deviation	1.93	0.62	4.67	1.52	0.09
Min	13.46	31.89	214.34	0.02	0.04
Max	23.15	40.18	253.57	22.81	1.10
Q1	15.98	38.00	227.51	0.15	0.15
Q3	19.25	38.34	232.88	0.39	0.39
Median	17.62	38.07	230.44	0.23	0.11

In Figure 6 the salinity variability in the sites favored by the various Mediterranean families is represented by a series of boxplots. *Hydrocharitaceae*, and most specifically the seagrass species *Halophilla*, appears favoring the most saline environments, within the narrow yearly-averaged salinity range from 39 to 39.5 psu. *Posidoniaceae* and *Cymodoceaceae* appear favoring lower, almost similar salinity levels, ranging between 37.3 and 39.3 psu. In parallel Figure 6 indicates that *Posidoniaceae*, and mostly *P. oceanica* favors stable in salinity marine environments, while *Cymodoceaceae* and mostly *C. nodosa* may grow in variable salinity levels.

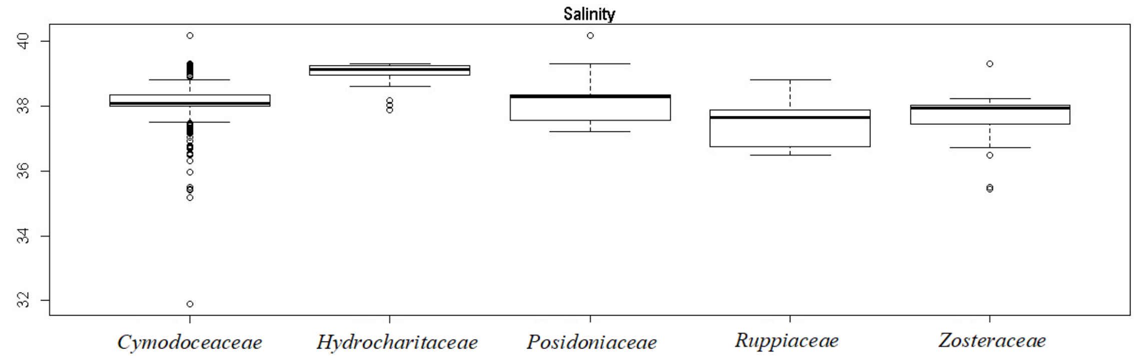

Figure 6. Boxplots of water salinity [psu] at the locations of seagrass families' abundance in the Mediterranean Sea.

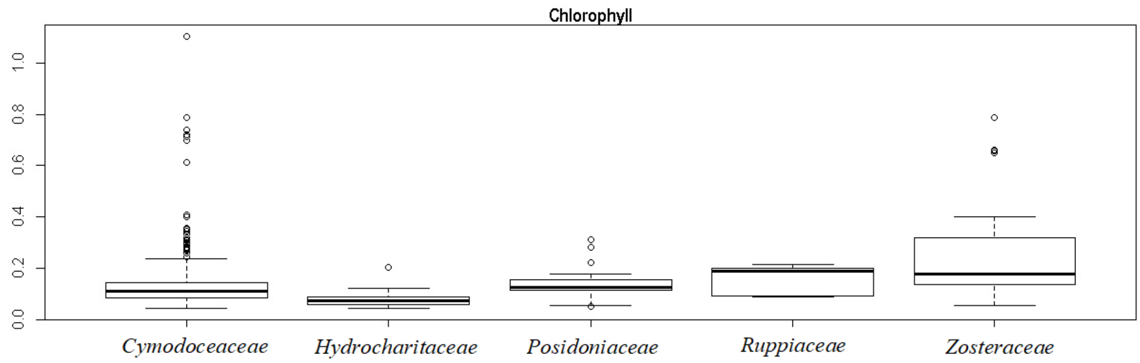

Figure 7. Boxplots of chlorophyll-a concentration [mg/m³] at the locations of seagrass families' abundance in the Mediterranean Sea.

For chlorophyll-α, the mean value for each seagrass family is shown in Figure 7. Each family prefers specific and different ranges of chlorophyll-α concentration. For *Hydrocharitaceae* and *Posidoniaceae*, chlorophyll-α ranges in a narrow specific range of values, while for the other families this is more spread, explaining their tolerance in eutrophication.

3.2. Experimental Tests Analysis

In the initial experimental test, four input parameters were imported in the developed FIS: the distance from major cities, the concentration of phosphates at sea surface, the seabed salinity and the benthic concentration of chlorophyll-a. Some basic fuzzy rules of this initial test are shown in Table 5.

Table 5. Indicative fuzzy rules in initial FIS test.

No	Distance from Cities	Phosphates [S]	Salinity [B]	Chl-a [B]	Seagrass Family
1	M	VL	M	L	*Zosteraceae*
2	L	VL	M	M	*Zosteraceae*
3	VL	H	M	VL	*Zosteraceae*
4	VL	VL	M	H	*Zosteraceae*
5	VL	VL	M	M	*Zosteraceae*
6	L	L	H	VL	*Hydrocharitaceae*
7	M	L	H	VL	*Hydrocharitaceae*
8	VL	L	H	VL	*Hydrocharitaceae*
9	L	VL	M	L	*Cymodoceaceae*
10	VL	VL	M	L	*Cymodoceaceae*
11	L	L	M	L	*Cymodoceaceae*
12	L	VL	H	L	*Posidoniaceae*
13	M	M	H	VL	*Posidoniaceae*

This experimental test showed limited precision (52%), however, after rules improvements and excluding *Posidoniaceae* from the model, due to the limited records in the training dataset, model precision reached 63% in the testing/validation dataset.

In another test, the developed FIS was constructed from parameters as the distance of each geolocation with seagrass from the coastline, the bottom water temperature and salinity and the benthic chlorophyll-a concentration. Despite improvements in overlapping fuzzy rules, this scenario reached precision of only 61% in the validation procedure. Overall, 22 experimental FIS of similar nature were developed and tested for their precision in assessing the seagrass family favoring specific environmental conditions.

The optimum FIS was found to be comprised of variables as the water depth, sea surface temperature, nitrates concentration at sea surface and chlorophyll-a concentration at sea bottom. This FIS exhibited seagrass family assessment accuracy of the order of 72%, and after fuzzy rules manipulation and adjustment, the FIS reached accuracy of 76%. However, accuracy was not evenly distributed over the Mediterranean. Higher accuracy (82%) was found in the eastern Mediterranean Sea, especially along the coastal zones of the Adriatic, Ionian and Aegean Seas, while decreased to 68% in the western basin, mostly in Alboran, Tyrrhenian Seas and the coastline of Northern Africa. Some indicative fuzzy rules for this test are shown in Table 6.

Table 6. Indicative fuzzy rules in the optimum FIS.

No	Water Depth	Chl-a [B]	Temperature [S]	Nitrates [S]	Seagrass Family
1	VL	M	L	VL	*Zosteraceae*
2	VL	H	L	L	*Zosteraceae*
3	VL	L	M	VL	*Zosteraceae*
4	VL	M	L	L	*Zosteraceae*
5	H	VL	M	VL	*Hydrocharitaceae*
6	H	VL	VH	VL	*Hydrocharitaceae*
7	L	VL	VH	VL	*Hydrocharitaceae*
8	M	VL	H	VL	*Hydrocharitaceae*
9	L	L	H	VL	*Cymodoceaceae*
10	VL	L	H	VL	*Cymodoceaceae*
11	L	M	H	VL	*Cymodoceaceae*
12	L	VL	H	VL	*Cymodoceaceae*
13	L	L	L	VL	*Posidoniaceae*
14	VL	L	L	VL	*Posidoniaceae*
15	M	M	M	VL	*Posidoniaceae*
16	M	H	H	VL	*Posidoniaceae*

3.3. Favorable Environment per Seagrass Family

Merging the seagrass database of seagrass occurrence over the Med with the CMEMS and EMODnet data, we were able to exploit the most favorable conditions per seagrass family. Figure 8 illustrates these favorable conditions, produced by the best FIS model, for all four Mediterranean seagrass families examined. The three-dimensional surface displays the limits in surface water temperature and chlorophyll-a at sea bottom levels, that each seagrass family favors the most, at relatively shallow and moderately nutrient-rich waters. *Posidoniaceae*, governed mostly by *P. oceanica*, prefers the cooler waters at the range of 16-18°C and low chlorophyll-a presence (< 0.2 mg/m^3). *Zosteraceae* seabeds are favored at cool (16-18°C) and more mesotrophic waters (Chl-a > 0.2 mg/m^3), but also at slightly warmer waters (18-19.5 $^{\circ}$C) with lower Chl-a levels (< 0.2 mg/m^3). Strong tolerance in broad range of environmental conditions is also seen by *Cymodoceaceae*, ranging from warm, oligotrophic waters (19.5-21.0°C and Chl-a < 0.3 mg/m^3) to moderately warm mesotrophic areas (18-21.3°C and

doi:10.20944/preprints202008.0424.v1

0.3 – 0.4 mg/m³ Chl-a). Finally, *Hydrocharitaceae* seem to thrive in the warm Mediterranaean waters (mean annual temperature 21-23°C) of low chlorophyll-a concentration (< 0.25 mg/m³).

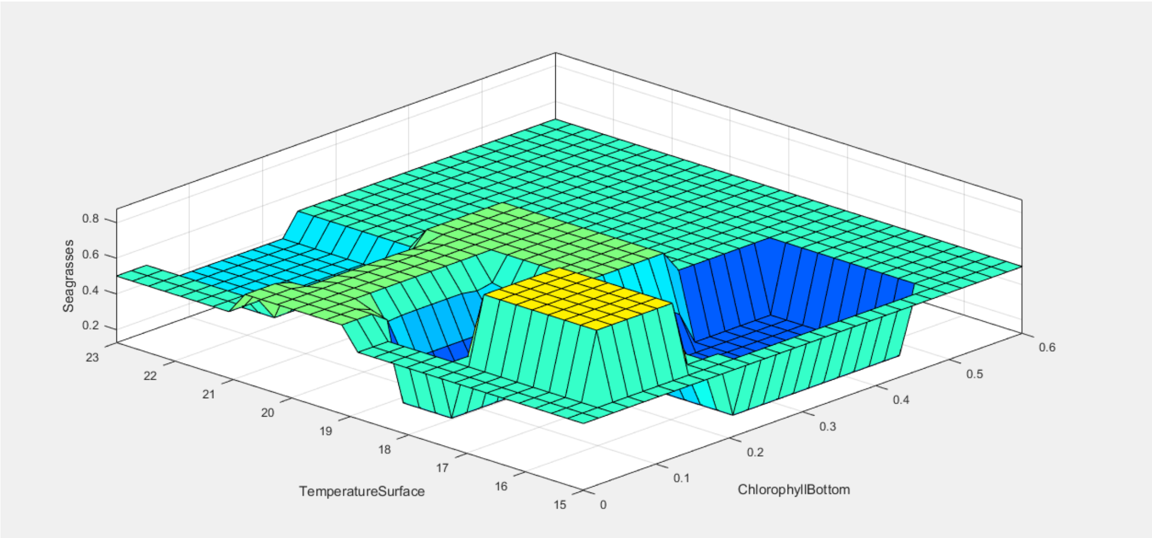

Figure 8. Three-dimensional surface produced by the best FIS, showing the levels of preference of various seagrass families. (Yellow: *Posidoniaceae*; Dark Blue: *Zosteraceae*; Light Green: *Cymodoceaceae*; Light Blue: *Hydrocharitaceae*).

3.4. FIS Sensitivity Analysis and Climate Change Scenarios

As the optimum FIS was found, one could perform sensitivity tests to examine the response of the various seagrass families in the change of external, environmental conditions. These tests could in fact act as scenarios simulating the impact of climate change and increased human pressure on seagrass communities, thus examining the level of seagrass tolerance and their progressive species replacement. Each parameter was forced to vary locally by -15%, -10%, -5%, +5%, +10% and +15% of its initial value. Then, the optimum FIS could assess and return the most appropriate seagrass family favoring these environmental conditions. The herein developed tests are 'closed', i.e., seagrass families may change among them but their disappearance is not allowed. This is a deficiency of the presently developed model that will be resolved in a future work.

Scenario 1 examined the impacts produced by changing the bathymetry in the locations with seagrass beds. The most sensitive seagrass families were *Hydrocharitaceae* and *Cymodoceaceae*. The reduction in water depth by 10 and 15% led to the reduction in the relative abundance of *Hydrocharitaceae* by 12 and 14%, respectively, being replaced by *Cymodoceaceae*. On the contrary, sea level rise seemed to leave unaffected all seagrass families.

Scenario 2 studied the response of seagrass families on water temperature changes. All four families were sensitive to these variations. *Posidoniaceae* exhibited higher tolerance in the reduction of mean water temperature, while *Hydrocharitaceae* was the family with the higher tolerance in the rise of sea temperature. More specifically, 43% and 57% of the locations with *Posidoniaceae* were found to be replaced by *Cymodoceaceae* when the water temperature rose by 10 and 15%, respectively. In parallel, 35% of the *Cymodoceaceae* sea beds was found to be replaced by *Posidoniaceae* as water temperature was reduced by 15%. Another 18% of these *Cymodoceaceae* meadows was found to change into *Zosteraceae* under these conditions. *Posidoniaceae* replaced 43% of the initial *Zosteraceae* seabeds when local water temperature drops by 15%. On the contrary, as water temperature rises by 5%, the *Zosteraceae* seabeds change into *Hydrocharitaceae* and *Cymodoceaceae* by 43 and 14%, respectively.

Figure 9 illustrates the seagrass meadows at family level in the Mediterranean Sea that remain unchanged for each water temperature change. As explained, *Posidoniaceae* is more tolerant in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 August 2020 doi:10.20944/preprints202008.0424.v1

cooler water temperature environments, while *Hydrocharitaceae* and *Cymodoceaceae* in mild sea temperature rise conditions, allowing only the former to sustain the higher water temperature increase.

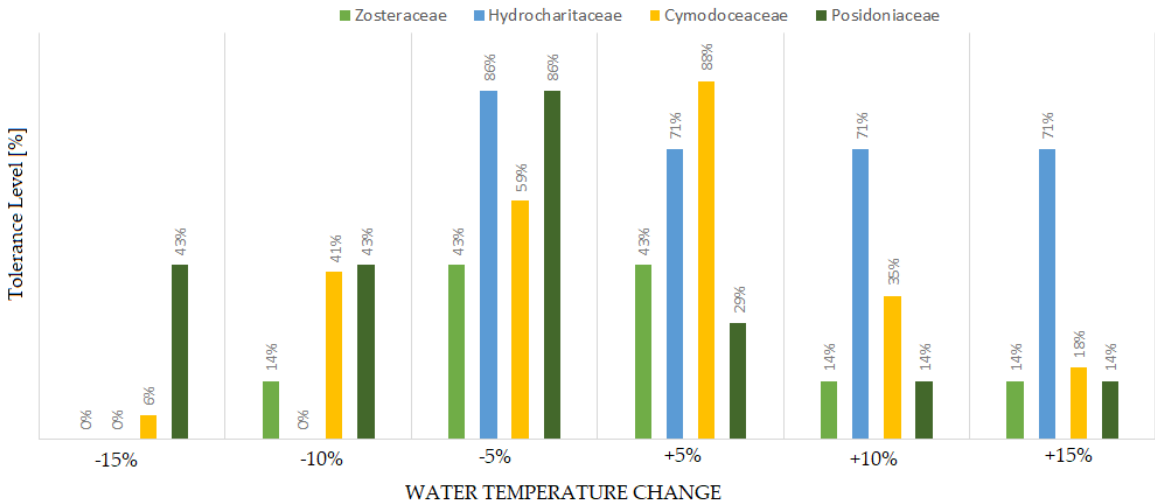

Figure 9. Seagrass tolerance per family on water temperature change from initial values.

The final scenario examined the response of seagrasses on the change of bottom chlorophyll-a levels. *Hydrocharitaceae* remained unaffected by changes in chlorophyll-a levels to any direction. Lower levels in chlorophyll-a led to the replacement of *Zosteraceae* into *Hydrocharitaceae* (29 and 40% in 5 and 10% reduction, respectively) and *Cymodoceaceae* (29% under 15% drop). On the other hand, an increase of 15% in the chlorophyll-a annual levels led to the reduction of 14% in the points with *Posidoniaceae* meadows.

4. Discussion

The exploitation of patterns between seabed habitat types and environmental parameters is particularly important for ecologists, since the abundance, patterns and distributions of seabed biological communities are directly linked to water column dynamics and quality [15]. In parallel, seagrass presence/absence and internal body growth asymmetries provide valuable metrics for marine ecosystems' ecological status evaluation [16]. Although the interrelation between physical, chemical, biological, seabed- and human-related drivers to seagrass species distribution is strongly non-linear, and therefore difficult to be revealed, modern data-driven models and techniques are capable to exploit and explore these hidden patterns. In parallel, the development of systematic, freely-available, diverse databases over the latest decades, and the development of tools for Big Data fusion and aggregation, lead to the better understanding of coastal benthic processes and human impacts.

In this work, a fuzzy logic model was developed and implemented covering the whole Mediterranean basin, with the aim to suggest the main environmental parameters affecting the distribution of seagrass, at family level. Fuzzy logic models of Mamdani type have never been implemented for such complex task. In a similar concept, [17] considered a series of morphodynamic, environmental and human impact variables and employed Machine Learning algorithms to predict the presence–absence of *P. oceanica* seagrass species. However, the dataset used was limited and rather unbalanced, towards absence records, affecting model's reliability. A more comprehensive study was performed by [18] detecting seagrass presence/absence and distinguishing seagrass families in the Mediterranean through supervised learning methods. Analysis in the relative variables' strength, chlorophyll-a and distance to coast appear more relevant

 doi:10.20944/preprints202008.0424.v1

to explain seagrass abundance, while chlorophyll-a, salinity, distance to major cities and nutrients were found as the main drivers for detecting seagrass family.

Several parameters combinations were tested by applying a trial and error approach. Our optimum FIS was found to consist of four main drivers: water depth, chlorophyll-a concentration at sea bottom, surface water temperature and nitrates. Water depth is an important parameter to determine the seagrass families, since it indirectly expresses changes in seabed temperature, pressure and light availability. *Cymodoceaceae*, mostly *C. nodosa*, favors living in a range of bathymetry, from shallow waters to depths of 60 m in sheltered to semi-exposed coasts, while *P. oceanica* is also present up to 50 m depth [19]. Above-ground biomass, leaf biomass and shoot density of both species was found to decline from shallow to higher depths in Spanish Mediterranean Sea [20], in agreement to our present findings.

Water temperature is also a significant driver for seagrass, as it affects its growth rates and reproductive patterns, enzymic and metabolic functions, when ranging within a physiological optimum, while extreme heating may enhance mortality [21]. Through the developed FIS, the limits of sea water temperature preference per family in the Mediterranean were defined: *Posidoniaceae* and *Zosteraceae* favor cooler systems, *Cymodoceaceae* range from moderate to warm areas, while *Hydrocharitaceae* grow in warmer waters. The above preferences determine the species that will be mostly affected in a temperature rising world. Under such conditions, the decline of one species may lead to the recovery and replacement by another [22], following the patterns defined by the FIS. In some examples of disturbance, no recovery was observed, although this case was not examined by the FIS.

Chlorophyll-a levels were important for seagrass family identification. As seen in Figure 8, each family prefers specific and different values of chlorophyll-α. *Posidoniaceae* and *Hydrocharitaceae* favor oligotrophic systems, *Zosteraceae* is abundant in mesotrophic environments, while *Cymodoceaceae* is tolerant to a wide range of chlorophyll-a concentrations. According to [18], chlorophyll-α levels in winter months (mostly in December) is the key parameter determining seagrass presence/absence and family identification.

Nutrients, mostly in the form of nitrates, also represent key environmental driver for seagrass species abundance, since food availability controls seagrass growth, distribution and metabolism [23]. Several nutrient sources, mostly rivers, outflow along the Mediterranean coastline, providing the appropriate nutrient levels to seagrass sustainability. The main difference from study [18] is that the FIS considered as important the impact of nitrates, while the machine learning model the influence of phosphates.

In terms of seagrass resilience to climate change impacts, the developed FIS indicated that: a) *Zosteraceae* is highly tolerant to changes in water depth and nutrient levels, as well as to small changes in chlrophyll-a levels, but extremely sensitive to temperature changes; b) *Hydrocharitaceae* is resistant to increases in water temperature, chlorophyll-a and nutrient levels, but sensitive to bathymetric changes; c) *Cymodoceaceae* appears affected by changes in sea water temperature, but intolerant in changes of the trophic status; and d) *Posidoniaceae* is threatened by the increase in water temperature and resistant in the variability of water depth and nitrate levels.

Based on the above it is evident that tolerance toward disturbances as well as growth and recolonization potentials differ among species and various seagrass species therefore show different temporal and spatial dynamics.

5. Conclusions

This work has developed a simple but novel, self-learning expert-system application, based on Mamdani fuzzy logic and its inference system development, for the prediction of the Mediterranean seagrass habitats occurrence at family level, according to the environmental conditions prevailing in an area. The system was developed utilizing diverse databases, as UNEP-WCMC for seagrass distribution and CMEMS and EMODnet for environmental conditions prevailing at each seagrass site. The optimum model receives input values from four parameters, namely water depth, sea surface temperature, surface nitrates and bottom chlorophyll-a and produces the most favorable seagrass family with fair classification accuracy (~76%).

Since the proposed model (in terms of input variables) follows the "trial and error" approach, it could be further improved through extended testing. The inclusion of species disappearance, as an additional output state, is also another option, advancing the FIS and making the model more realistic. However, expert-knowledge is needed for such advancement. Other tests could include finding the most accurate shape and boundaries of the membership functions and examining the most appropriate defuzzification method.

The present FIS is capable to describe the favorable living conditions per seagrass family over the Mediterranean Sea. *Posidoniaceae*, and mostly *P. oceanica*, prefers cool, oligotrophic waters, *Zosteraceae* seabeds favor a wider temperature range (16-19.5°C) and mesotrophic waters, *Cymodoceaceae*, and mostly *C. nodosa*, appears tolerant to a broad range of living conditions, from warm, oligotrophic to moderately warm mesotrophic areas, while *Hydrocharitaceae* prefer the warmer, oligotrophic parts of the Mediterranaean.

The FIS has the capacity to exploit the impact of environmental change on seabed habitats, as those induced by climate change, illustrating that (a) *Hydrocharitaceae* and *Cymodoceaceae* are the most sensitive families in bathymetric reduction, while *Posidoniaceae* and *Zosteraceae* are insensitive in any water variation; (b) *Cymodoceaceae* is the family with the higher tolerance in mild sea temperature rise (+5%); (c) *Hydrocharitaceae* exhibits tolerance in higher sea temperature rise (10-15%); (d) *Posidoniaceae* and *Zosteraceae* are mostly affected by temperature rise, at any level, and (e) *Posidoniaceae* exhibited higher tolerance in mean water temperature reduction.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: title, Table S1: title, Video S1: title.

Author Contributions: Conceptualization, G.S.; methodology, G.S.; software, G.S.; validation, D.P., N.K., G.S; formal analysis, D.P.; resources, N.K.; data curation, N.K.; writing—original draft preparation, G.S.; writing—review and editing, N.K., G.S.; visualization, D.P.; supervision, G.S.; project administration, G.S.; funding acquisition, G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by European Union's Horizon 2020 Research and Innovation Program (H2020-BG-12-2016-2), grant number No. 727277 – ODYSSEA (Towards an integrated Mediterranean Sea Observing System). The article reflects only authors' view and that the Commission is not responsible for any use that may be made of the information it contains.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Barnes, R.S.K.; Claassens, L. Do beds of subtidal estuarine seagrass constitute a refuge for macrobenthic biodiversity threatened intertidally? *Biodivers Conserv* **2020**, *In Press*, https://doi.org/10.1007/s10531-020-02019-0.
2. Short, F.; Carruthers, T.; Dennison, W.; Waycott, M. Global seagrass distribution and diversity: A bioregional model. *J Exp Mar Biol Ecol* **2007**, *350*, 3-20.
3. Ahmad, F.; Azman, S.; Said, M.I.M.; Lavania-Baloo Tropical seagrass as a bioindicator of metal accumulation. *Sains Malays* **2015**, *44*, 203–210.
4. Lee, K.-S.; Short, F.T.; Burdick, D.M. Development of a nutrient pollution indicator using the seagrass, *Zostera marina*, along nutrient gradients in three New England estuaries. *Aquat Bot* **2004**, *78*, 197–216.
5. Ferrat, L.; Pergent-Martini, C. ; Roméo, M. Assessment of the use of biomarkers in aquatic plants for the evaluation of environmental quality: application to seagrasses. *Aquat Toxic* **2003**, *65*, 187–204.

6. Orfanidis, S.; Papathanasiou, V.; Gounaris, S.; Theodosiou, Th. Size distribution approaches for monitoring and conservation of coastal *Cymodocea* habitats. *Aquat Conserv* **2009**, *20*, 177–188.
7. Short, F.T.; Neckles, H.A. The effects of global climate change on seagrasses. *Aquat Bot* **1999**, *3-4*, 169-196.
8. Short, F.; Polidoro, B.; Livingstone, S.R. et al. Extinction risk assessment of the world's seagrass species. *Biol Conserv* **2011**, *144*, 1961-1971.
9. Giannoulaki, M.; Belluscio, A.; Colloca, F.; Fraschetti, S.; Scardi, M.; Smith, C.; Panayotidis, P.; Valavanis, V.; Spedicato, M. Mediterranean sensitive habitats (Mediseh). Final project report. **2013**, DG MARE Specific Contract SI2, 600741, 557.
10. Effrosynidis, D. ; Arampatzis, A. ; Sylaios, G. Seagrass detection in the mediterranean: A supervised learning approach. *Ecol Inform* **2018**, *48*, 158-170.
11. Weatherdon, L.; Fletcher, R.; Jones, M. et al. Manual of Marine and Coastal Datasets of Biodiversity Importance. **2015**, UNEP World Conservation Monitoring Centre, Cambridge, UK.
12. Effrosynidis, D. ; Arampatzis, A.; Sylaios, G. Seagrass and hydrographic data for the Mediterranean Sea. *Data Brief* **2019**, *25*, 104286.
13. Ocampo-Duque, W.; Ferré-Huguet, N. ; Domingo, J.L. ; Schuhmacher, M. Assessing water quality in rivers with fuzzy inference systems: a case study. *Environ Int* **2006**, *32*, 733–742.
14. Kotti, I.P.; Sylaios, G.K.; Tsihrintzis, V.A. Fuzzy logic models for BOD removal prediction in free-water surface constructed wetlands. *Ecol Eng* **2013**, *51*, 66-74.
15. Boudouresque, C.F.; Bernard, G. ; Pergent, G. ; Shili, A. ; Verlaque, M. Regression of Mediterranean seagrasses caused by natural processes and anthropogenic disturbances and stress: a critical review. *Bot Mar* **2009**, *52*, 395–418.
16. Orfanidis, S.; Papathanasiou, V.; Mittas, N.; Theodosiou, Th.; Ramfos, A.; Tsiolia, S.; Kosmidou, M.; Kafas, A.; Mystikou, A.; Papadimitriou, A. Further improvement, validation, and application of CymoSkew biotic index for the ecological status assessment of the Greek coastal and transitional waters. *Ecol Indic* **2020**, *118*, 106727.
17. Giannoulaki, M.; Belluscio, A.; Colloca, F.; Fraschetti, S.; Scardi, M.; Smith, C.; Panayotidis, P.; Valavanis, V.; Spedicato, M. Mediterranean sensitive habitats (Mediseh). 2013, Final Project Report. DG MARE Specific Contract SI2 600 (741), 557.
18. Effrosynidis, D.; Arampatzis, A. ; Sylaios, G. Seagrass detection in the mediterranean: A supervised learning approach. *Ecol Inform* **2018**, *48*, 158-170.
19. Barbera, C.; Tuya, F.; Boyra, A.; Sanchez-Jerez, P. ; Blanch, I. ; Haroun, R.J. Spatial variation in the structural parameters of *Cymodocea nodosa* seagrass meadows in the Canary Islands: A multi-scaled approach. *Bot Mar* **2005**, *48*, 122-126.
20. Olesen, B.; Enríquez, S.; Duarte, C.M.; Sand-Jensen, K. Depth-acclimation of photosynthesis, morphology and demography of *Posidonia oceanica* and *Cymodocea nodosa* in the Spanish Mediterranean Sea. *Mar Ecol Progr Ser* **2002**, *236*, 89-97.
21. Collier, C.J.; Waycott, M. Temperature extremes reduce seagrass growth and induce mortality. *Mar Poll Bull* **2014**, *83*, 483-490.
22. Peterson, B.J.; Rose, C.D.; Rutten, L.M.; Fourqurean, J.W. Disturbance and recovery following catastrophic grazing: Studies of a successional chronosequence in a seagrass bed. *Oikos*, **2002**, *97*, 361–370.
23. Brix, H.; Lyngby, J. Uptake and translocation of phosphorus in eelgrass (*Zostera marina*). *Mar Biol* **1985**, *90*, 111–116.