

See discussions, stats, and author profiles for this publication at: <http://www.researchgate.net/publication/275954823>

Large-Scale Prediction of Seagrass Distribution Integrating Landscape Metrics and Environmental Factors: The Case of *Cymodocea nodosa* (Mediterranean–Atlantic)

ARTICLE *in* ESTUARIES AND COASTS · MAY 2015

Impact Factor: 2.54 · DOI: 10.1007/s12237-015-9966-y

READS

193

4 AUTHORS:



Rosa María Chefaoui

Universidade do Algarve

9 PUBLICATIONS 289 CITATIONS

SEE PROFILE



Jorge Assis

Universidade do Algarve

23 PUBLICATIONS 99 CITATIONS

SEE PROFILE



Carlos M. Duarte

University of Western Australia

678 PUBLICATIONS 27,607 CITATIONS

SEE PROFILE



Ester A Serrao

Universidade do Algarve

201 PUBLICATIONS 3,449 CITATIONS

SEE PROFILE

*Large-Scale Prediction of Seagrass
Distribution Integrating Landscape Metrics
and Environmental Factors: The Case of
Cymodocea nodosa (Mediterranean–
Atlantic)*

**Rosa M. Chefaoui, Jorge Assis, Carlos
M. Duarte & Ester A. Serrão**

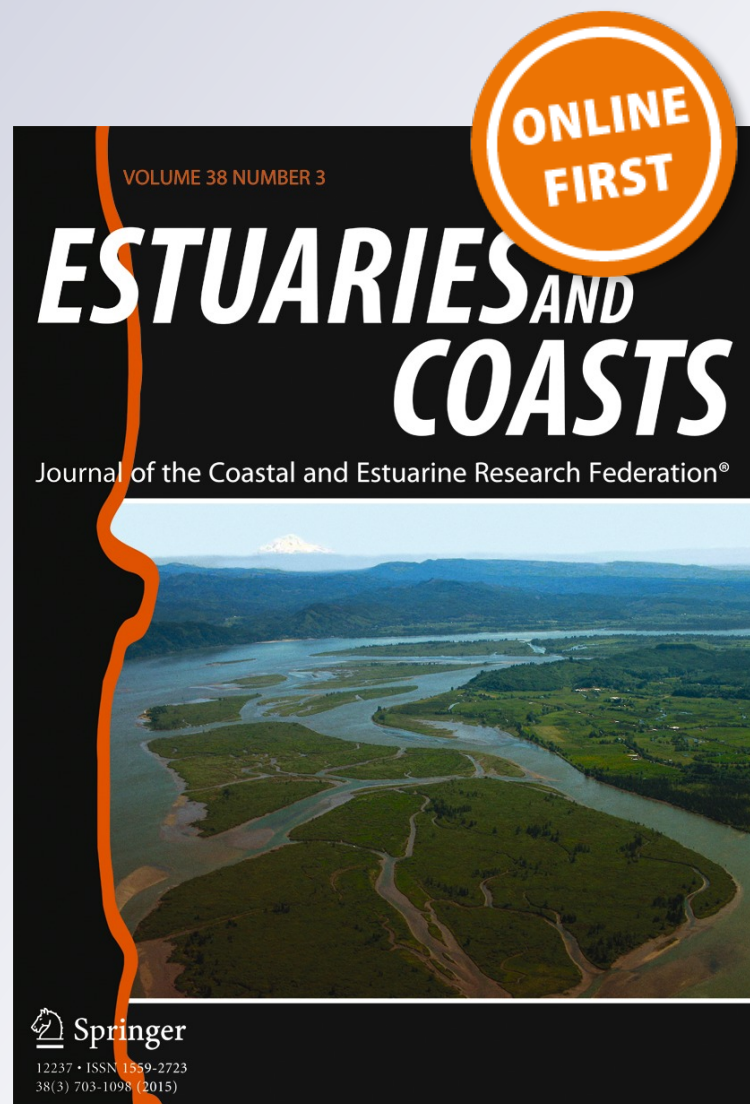
Estuaries and Coasts

Journal of the Coastal and Estuarine
Research Federation

ISSN 1559-2723

Estuaries and Coasts

DOI 10.1007/s12237-015-9966-y



Your article is protected by copyright and all rights are held exclusively by Coastal and Estuarine Research Federation. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Large-Scale Prediction of Seagrass Distribution Integrating Landscape Metrics and Environmental Factors: The Case of *Cymodocea nodosa* (Mediterranean–Atlantic)

Rosa M. Chefaoui¹ · Jorge Assis¹ · Carlos M. Duarte² · Ester A. Serrão¹

Received: 3 November 2014 / Revised: 19 March 2015 / Accepted: 19 March 2015
 © Coastal and Estuarine Research Federation 2015

Abstract Understanding the factors that affect seagrass meadows encompassing their entire range of distribution is challenging yet important for their conservation. Here, we predict the realized and potential distribution for the species *Cymodocea nodosa* modelling its environmental niche in the Mediterranean and adjacent Atlantic coastlines. We use a combination of environmental variables and landscape metrics to perform a suite of predictive algorithms which enables examination of the niche and find suitable habitats for the species. The most relevant environmental variables defining the distribution of *C. nodosa* were sea surface temperature (SST) and salinity. We found suitable habitats at SST from 5.8 °C to 26.4 °C and salinity ranging from 17.5 to 39.3. Optimal values of mean winter wave height ranged between 1.2 and 1.5 m, while waves higher than 2.5 m seemed to limit the presence of the species. The influence of nutrients and pH,

despite having weight on the models, was not so clear in terms of ranges that confine the distribution of the species. Landscape metrics able to capture variation in the coastline enhanced significantly the accuracy of the models, despite the limitations caused by the scale of the study. We found potential suitable areas not occupied by the seagrass mainly in coastal regions of North Africa and the Adriatic coast of Italy. The present study describes the realized and potential distribution of a seagrass species, providing the first global model of the factors that can be shaping the environmental niche of *C. nodosa* throughout its range. We identified the variables constraining its distribution as well as thresholds delineating its environmental niche. Landscape metrics showed promising prospects for the prediction of coastal species dependent on the shape of the coast. By contrasting predictive approaches, we defined the variables affecting the distributional areas that seem unsuitable for *C. nodosa* as well as those suitable habitats not occupied by the species. These findings are encouraging for its use in future studies on climate-related marine range shifts and meadow restoration projects of these fragile ecosystems.

Communicated by Charles Simenstad

Electronic supplementary material The online version of this article (doi:10.1007/s12237-015-9966-y) contains supplementary material, which is available to authorized users.

✉ Rosa M. Chefaoui
 rosa.chef@gmail.com; rmdiaz@ualg.pt

Jorge Assis
 jorgemfa@gmail.com

Carlos M. Duarte
 carlos.duarte@kaust.edu.sa

Ester A. Serrão
 eserrao@ualg.pt

¹ Centro de Ciências do Mar (CCMAR), Universidade do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

² Red Sea Research Center (RSRC) Building 2, Level 3, Room 3219, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia

Keywords Coastal morphology · *Cymodocea nodosa* · Environmental niche · Landscape metrics · Seagrass · Species distribution modelling

Introduction

Seagrass meadows rank amongst the most valuable ecosystems on Earth (Costanza et al. 1998), yet they are declining worldwide due to coastal habitat alterations, the impacts of climate change, and ecological degradation (see, e.g., Orth et al. 2006; Duarte et al. 2008; Duarte 2011; Waycott et al. 2009). Whereas this is a global phenomenon, most efforts at understanding variables explaining seagrass performance and

loss have focused at local scales. A focus on local scales may, for instance, overlook global redistribution patterns, such as biogeographic range shifts, which have been documented for most other marine taxa, but not yet for seagrasses (Poloczanska et al. 2013). Indeed, local declines could, in some cases, be related to large-scale biogeographical range shifts rather than local factors. Hence, there is a need to move beyond local scales towards understanding factors determining global patterns of seagrass distribution.

An interesting model species to study what variables explain its global distribution is the seagrass *Cymodocea nodosa* (Ucria) Ascherson, because it is the main marine subtidal seagrass that can be found extending across the gradient of environmental conditions which change sharply between the Mediterranean and adjacent Atlantic coastal regions. Although in the Mediterranean another seagrass species, the endemic *Posidonia oceanica*, is found in greater abundance, *C. nodosa* is present throughout the Mediterranean Sea and is the dominant keystone species structuring subtidal seagrass ecosystems along NW Africa and SW Europe (occurring from Senegal northwards to SW Iberia; see, e.g., Alberto et al. 2008; Cunha and Araújo 2009; Green and Short 2003). There, populations have also been reported to be affected by biotic interactions (Tuya et al. 2013), or population bottlenecks (Cunha et al. 2013). *C. nodosa* develops ecologically important meadows providing a favorable habitat for fishes and invertebrates (e.g., Espino et al. 2011; Gartner et al. 2013; Guidetti and Bussotti 2000; Tuya et al. 2014a), or stabilizing sediments (Cabaço et al. 2010). Due to its importance, *C. nodosa* is listed for protection by the OSPAR (<http://www.ospar.org/>) and Bern (<http://www.coe.int/>) conventions.

Another aspect that renders *C. nodosa* a suitable model species for large-scale distribution modeling is the availability of a wealth of literature on its physiology and ecology. Some examples include research on its photosynthetic and nutrient status (e.g., Enriquez et al. 2004), metal bioaccumulation (Malea et al. 2013), photoacclimation capacity (Olivé et al. 2013), associated assemblages (e.g. Tuya et al. 2014b; Verdiell-Cubedo et al. 2007), changes in the structure and growth dynamics of the meadows (Cancemi et al. 2002), and the genetic structure of the populations (Alberto et al. 2008), among other questions. However, most studies, except the assessment of genetic structure, have only been carried out at local spatial scales, mostly along the European coastline, due to the difficulties of accomplishing such experiments at wider extents. So, there is still insufficient knowledge on *C. nodosa* encompassing its entire distributional range.

Niche modelling algorithms allow the characterization of conditions suitable for the species and to predict its distributions under different conditions. Possible uses of species distribution models (SDMs) in ecology and conservation biology are numerous: invasion risk assessment, future and past projections with climate change, discovery of new populations,

reserve selection, and niche evolution among others (see, e.g., Guisan and Thuiller 2005; Araújo and Peterson 2012). The combined use of geographic information systems (GIS) with niche and distribution modelling has proven to be an effective methodology for analyzing global patterns and ecological requirements of marine species (see, e.g., Bryan and Metaxas 2007; Tittensor et al. 2010). However, so far, such modelling approaches have not been applied to seagrasses at broad spatial scales, with the exception of a recent study of *Zostera noltii* (Valle et al. 2014).

Here, we examine the factors affecting the distribution of *C. nodosa* over its entire geographic range and estimate its potential and realized distribution using several SDM techniques. It has been previously described (e.g., Zaniewski et al. 2002; Chefaoui and Lobo 2008; Mateo et al. 2010) that, according to the type of absences used, different predictions can be obtained from a potential (wider) to a realized (closer to the known niche) distribution gradient. We estimated these two distant niche hypotheses, although we recognize that estimates can rarely describe the real potential and realized niches due to the lack of evidence from different sources (see Jiménez-Valverde et al. 2008). With the purpose of representing these niche hypotheses, we used: (i) presence-absence algorithms to estimate an approximation to the realized niche of the species using random selection of pseudo-absences and (ii) a presence-only method to estimate its potential distribution according to suitable habitats. The estimates of the niche so obtained are considered as two different hypotheses for the distribution of *C. nodosa*, and the main factors characterizing these niches are discussed and contrasted against data from physiological experiments reported in the literature. Besides environmental variables, *C. nodosa* seems to be influenced by coastal morphology and hydrodynamics, as it occurs preferentially in bays sheltered from waves (Cunha and Araújo 2009) or lagoons behind *Posidonia* reefs along its distribution. In an attempt to capture this influence of coastlines on the distribution of *C. nodosa*, we integrate certain landscape metrics found to be suitable indicators for coastal features (Chefaoui 2014) as predictors in our models. Although this previous research showed best performance at small scales, we attempt to apply the same methodology at a wider scale, where the loss of resolution of the coastline affects the precision of the results. In particular, we hypothesize that landscape metrics are able to enhance the performance of SDMs of coastal species even at coarse spatial extents and resolutions, hence delivering on the aims of this study to: (i) explore the niche of *C. nodosa* across its entire geographic distribution under two different predictive hypotheses; (ii) test the ability of landscape metrics to improve the prediction of species such as *C. nodosa*, whose habitat is linked to the morphology of the coasts; and (iii) to identify suitable sites where prospective sampling or restoration actions could be developed.

Methods

Species Data and Study Area

C. nodosa presence records were compiled mainly from literature, using also internet databases: Algaebase (Guiry and Guiry 2014) and the Global Biodiversity Information Facility (GBIF; <http://www.gbif.org/>). All presences were reviewed to discard or correct those having referencing errors. After georeferencing species data to the same resolution as predictor layers ($0.083^{\circ} \times 0.083^{\circ}$ latitude–longitude ~ 9.2 km), 202 unique presence cells remained. The entire species' geographic range was considered, which includes the Mediterranean Sea, North-East Atlantic coasts, and the Black Sea. The Red Sea was excluded from the study area as this species, unlike others of its genus, has neither been found there previously (see Alberto et al. 2008) nor have there been any reports of recent colonization through the Suez Canal. From this geographic extent, we selected those cells ranging from 0 to 35 m depth, the known habitat for *C. nodosa* (Zarranz et al. 2010).

Environmental and Landscape Metrics Predictors

A number of environmental variables were first selected according to their possible association with *C. nodosa*, based on the known species biology (Duarte 1995; Marbà et al. 1996; Pérez et al. 1991; Fernández-Torquemada and Sánchez-Lizaso 2006). Maximum, minimum, and range values at the sites where the species was reported to be present were calculated for diffuse attenuation coefficient (Kd), sea surface temperature (SST), and wave height from satellite data. Kd has a direct relation with the traditional Secchi disk (Chen et al. 2007) and was derived from SeaWiFS (Sea Wide Field Sensor) satellite radiance (Gohin et al. 2005) and MODIS (Moderate Resolution Imaging Spectroradiometer; Huot et al. 2005). SST was obtained from the OSTIA system (Stark et al. 2007), which combines infrared and microwave satellite data with in situ measurements. Significant wave height was derived from AVISO altimeter data (Schaeffer et al. 2012). A set of environmental variables were obtained from Bio-ORACLE (Tyberghein et al. 2012): dissolved oxygen, nitrate, phosphate, pH, photosynthetically available radiation (PAR) and salinity (see Table 1 for abbreviations). Besides, bottom slope and bottom aspect were calculated from bathymetric data derived from the General Bathymetric Chart of the Oceans (GEBCO 2010).

To assess if coastal shape plays a role in determining the presence of the species, a total of nine landscape metrics were chosen attending to previous successful efforts at predicting Iberian Peninsula's coastal features (Chefaoui 2014): edge density (ED), mean perimeter–area ratio (PARA_MN), percentage of landscape (PLAND), and mean shape index (SHAPE_MN). In addition, we

Table 1 Abbreviations employed in the text

Abbreviation	Definition
SDM	Species distribution model
ENFA	Environmental niche factor analysis
FDA	Flexible discriminant analysis
GAM	Generalized additive model
GBM	Generalized boosting model
GLM	Generalized linear model
MARS	Multiple adaptive regression splines
MD	Mahalanobis distance
RF	Random forest
ENV set	Set of environmental variables
Kd	Diffuse attenuation coefficient
PAR	Photosynthetically available radiation
SST	Sea surface temperature
MET set	Set of landscape metrics
ECON_MN	Mean edge contrast index distribution
FRAC_AM	Area-weighted mean fractal dimension index
PARA_MN	Mean perimeter–area ratio
PLAND	Percentage of landscape
SHAPE_MN	Mean shape index
TECI	Total edge contrast index

also considered the metrics that Cushman et al (2008) found to be universal, strong, and consistent measuring edge contrast (contrast between a patch and its neighborhood): edge contrast index (ECON_MN and ECON_AM) and total edge contrast index (TECI); and patch shape complexity: fractal dimension index (FRAC_AM and FRAC_MN). All landscape metrics were computed at class level in FRAGSTATS (v4; McGarigal et al. 2012), a Spatial Pattern Analysis Program for Categorical Maps, using a moving-window square of 27×27 pixels, a size found to be adequate for the scale of the study (Chefaoui 2014). Coastline was rasterized from high-resolution world vector shoreline (GSHHG v. 2.2.2), extracted using GEODAS desktop software (GEODAS-NG v. 1.1.1.1) at a wider extent than the study area to avoid loss of data in the extremes during metric computations (see Chefaoui 2014). A Pearson correlation test was used to check highly correlated predictors ($r \geq |0.80|$, $p < 0.001$). As the discarding of variables may be an arbitrary procedure, we performed a preliminary ENFA analysis (Ecological Niche Factor Analysis; Hirzel et al. 2002) to choose between correlated variables. ENFA is a presence-only method which identifies how is each variable able to relate the presences with the factors defining the main environmental gradients in the region considered (see, e.g., Hirzel et al. 2002; Calenge 2006). ENFA was fit using “adehabitat” (Calenge 2006)

in R. The variables selected were grouped into three sets: (i) environmental (ENV), (ii) landscape metrics (MET), and (iii) environmental+landscape metrics (ENV+MET) (see Table 2).

Modelling Approach

Six presence–absence modelling techniques were used to predict the more restricted distribution (closer to the realized) of *C. nodosa* and to assess the importance of the three sets of predictors: generalized linear model (GLM), generalized additive model (GAM), generalized boosting model (GBM), flexible discriminant analysis (FDA), multiple adaptive regression splines (MARS), and randomForest (RF). All models were implemented using the “biomod2” package (Thuiller et al. 2013) in R (R Core Team 2013). Parameters were configured directly in biomod2 for all techniques with exception of GLMs, which were in first instance performed in “MASS” package (Venables and Ripley 2002) to manually allow interactions among landscape metric predictors. As reliable absence data was not available, we created pseudo-absences. We used ten times more pseudo-absences than presences (2020 cells), setting a prevalence of 0.1, this proportion has been used before with good results (e.g., Chefaoui and Lobo 2008). Though some studies recommend selecting a higher number of pseudo-absences (e.g., Barbet-Massin et al. 2012), our study is hindered by the low number of cells of the study area (12,330 cells). Pseudo-absences were extracted at random to offer more realistic validation results in terms of comparison of techniques and sets of variables (Chefaoui and Lobo 2008). To increase the representativeness of these pseudo-absences, we used three different sets of pseudo-absences for each model.

We split the data into two sets, 70 % for calibration and 30 % for model validation. In order to validate, the models we performed a repeated random sub-sampling validation repeating ten times the process (calibration and evaluation). Model performance was measured using the area under the receiver operating characteristic (ROC) curve (AUC), ROC-derived sensitivity (presences correctly predicted) and specificity (absences correctly predicted) (Fielding and Bell 1997), and the true skill statistic (TSS; Allouche et al. 2006). The threshold used to transform predicted probability into binary values was that which optimized ROC and TSS scores after testing a set of thresholds (Thuiller et al. 2013). Student's independent *t* tests were used to compare validation results among the best models with each set of variables, using a Welch's procedure (Welch 1947) to adjust for possible unequal variances.

To test the performance obtained with each type of predictor, we ran models using the three sets of variables (ENV, MET, ENV+MET). For each set of variables, a total of 180

models were computed, 10 runs×3 pseudo-absence sets×6 modelling techniques.

To reduce the variability in the models due to the algorithm or pseudo-absence set used, an ensemble approach (Araújo and New 2007) was accomplished. Ensemble modelling was performed using all the repetitions and sets of pseudo-absences of the algorithm with best accuracy scores, excluding those models with TSS values <0.8 or AUC <0.95. We performed two types of ensembles: “mean of probabilities”, estimating the mean probabilities over the selected models; and “committee averaging”, calculating the average of binary predictions. Subsequently, we selected the one which achieved the best evaluation scores, giving priority to obtaining a high sensitivity.

The potential distribution of *C. nodosa* was estimated by Mahalanobis distance (MD), a presence-only method, and compared with the projection of the best ensemble of presence–absence models mentioned above. MD algorithm generates the elliptic envelope for a species using just presences (Clark et al. 1993). This analysis is closely related to ENFA but is superior to obtain an environmental suitability map (Calenge et al. 2008). We scaled the variables before the computation to allow them to have the same variance. MD was fit using “adehabitat” (Calenge 2006) in R.

Variables Testing and Response Plots

To estimate the importance of each variable, we used two measures: the scores derived from ENFA and the ones obtained from the best presence–absence models. We used the scores obtained from the two uncorrelated axes computed by ENFA: the marginality (direction of maximum difference between the species niche and the available conditions in the study area) and the specialization factor (the ratio of ecological variance of the species in relation to the mean habitat) (Hirzel et al. 2002).

The importance of the variables was also estimated for the best presence–absence algorithm, the one used for the ensemble, performing three permutations. It was calculated with the correlation between the full model and the model without one variable iteratively, using a procedure similar to “randomForest” (Liaw and Wiener 2002). The subtraction (1 minus correlation) gives a value ranging from 0 (lowest importance) to 1 (highest importance) (see Thuiller et al. 2013). Subsequently, a mean of importance scores per variable was calculated for each run and pseudo-absence set. Comparison between ENFA and GAM scores was made using the Pearson correlation coefficient. In addition, we estimated the thresholds of the most important variables found to constrain the reported distribution of the seagrass by means of the response plots derived from the ensemble. Two-dimensional (2D) and three-dimensional (3D) response plots were obtained according to Elith et al. (2005), creating strings of varying values over the

Table 2 Variables used to obtain species distribution models for *C. nodosa*

Predictors (units)	Min.-max. (study area)	Mean values (study area)	Min.-max. (presences)	Mean values (presences)	Min.-max. (suitable GAMs)	Mean values (suitable GAMs)	Min.-max. (suitable MD)	Mean values (suitable MD)
Environmental (ENV)								
Minimum summer Kd (m ⁻¹)	0.017–3.594	0.177	0.021–1.118	0.087	0.021–2.406	0.099	0.017–0.347	0.047
Minimum winter Kd (m ⁻¹)	0.018–4.034	0.180	0.026–1.134	0.105	0.023–1.843	0.117	0.019–0.400	0.071
Mean summer SST (°C)	13.57–31.25	21.25	17.57–25.90	21.10	14.57–26.36	21.07	18.38–24.63	21.33
Minimum winter SST (°C)	1.95–26.25	13.32	5.89–20.01	14.030	5.84–20.68	13.54	6.931–18.58	13.682
Mean winter wave height (m)	0.643–4.312	1.33	0.806–2.443	1.377	0.771–2.891	1.326	0.888–2.206	1.289
Nitrate (μmol N L ⁻¹)	0.003–12.49	1.469	0.005–7.316	1.052	0.003–10.54	1.124	0.293–3.450	0.775
Phosphate (μmol P L ⁻¹)	0.018–0.877	0.211	0.021–0.871	0.151	0.018–0.871	0.16	0.022–0.378	0.120
pH (unitless)	7.671–8.633	8.245	8.042–8.540	8.238	7.671–8.543	8.238	8.065–8.423	8.226
PAR max. (mol quanta m ⁻² d ⁻¹)	36.16–64.61	56.83	47.71–64.29	58.23	36.16–64.29	58.08	53.39–63.26	58.66
Salinity (unitless)	1.63–39.98	33.07	17.50–39.32	35.96	17.50–39.32	35.95	24.40–39.32	37.50
Bottom slope (degrees)	0–18.65	1.64	0.024–15.08	2.867	0.024–18.65	2.59	0.027–12.145	2.069
Bottom aspect (degrees)	15.99–339.2	185.18	38.11–307.2	170.23	34.7–312.8	177.8	28.89–295.75	173.32
Landscape metrics (MET)								
ECON_MN (%)	50.00–100.0	93.98	70.03–100.0	94.46	69.72–100.0	94.23	77.43–100.0	95.12
FRAC_AM (unitless)	1.00–1.14	1.102	1.035–1.136	1.095	1.021–1.139	1.098	1.041–1.134	1.100
PARA_MN (unitless)	0.829–4.000	2.079	0.844–3.093	2.078	0.842–3.402	2.135	1.026–3.333	2.155
PLAND (%)	0.14–43.75	12.79	4.115–40.19	10.143	2.743–40.74	10.854	3.704–28.121	9.802
SHAPE_MN (unitless)	1.00–5.16	2.59	1.167–4.737	2.440	1.083–4.750	2.435	1.268–4.526	2.430
TECI (%)	50.00–100.0	95.77	89.61–100.0	96.59	87.76–100.0	96.51	89.58–100.0	96.75

Values for suitable area are referred to the ensemble model of generalized additive models (GAMs) and to the Mahalanobis distance (MD) model, with probability of occurrence transformed into binary data using the same threshold as for validation

SST sea surface temperature, Kd diffuse attenuation coefficient, PAR photosynthetically available radiation, ECON_MN mean edge contrast index distribution, FRAC_AM area-weighted mean fractal dimension index, PARA_MN mean perimeter–area ratio, PLAND percentage of landscape, SHAPE_MN mean shape index, TECI total edge contrast index

range of target variables while holding the rest constant at their mean value. All analyses were conducted in R.

Results

Presence–Absence Modelling and Assessment of Sets of Variables

There were high correlations among the initial set of variables, so we selected the most ecologically representative for the species in conformity with the preliminary ENFA scores. As all of the highly correlated different measures of SST also showed similar ENFA scores, we chose those not correlated and able to capture the differences between summer and winter SST. After correlation and ENFA analysis, a total of 18 variables (12 environmental and 6 landscape metrics) were selected (Tables 2 and 3) and used to create the same three sets of predictors for all the models.

GAM performed better than the rest of techniques using ENV and ENV+MET sets of variables (Table 4). ENV+MET set performed significantly better than ENV (AUC, $t=11.013$, $p=7.87e^{-16}$, 95 % CI, 0.023 to 0.034; TSS, $t=9.553$, $p=1.82e^{-13}$, 95% CI, 0.063 to 0.097; sensitivity, $t=3.007$, $p=0.003$, 95 % CI, 0.822 to 4.095; specificity, $t=8.521$, $p=8.31e^{-12}$, 95 % CI, 4.259 to 6.874). As can be seen from the data in Table 4, the MET set showed the worst validation scores, and no agreement was found among validation measures of the different techniques, thus it was not considered for subsequent analyses.

According to these results, ensemble modelling was performed with GAM models using all the variables (ENV+MET set). Between the two ensemble algorithms used, the committee averaging identified better the presences and also obtained better specificity and TSS scores (AUC=0.934; TSS=0.806; sensitivity=92.079; specificity=88.489) than the ensemble based on the mean of probabilities (AUC=0.945;

TSS=0.778; sensitivity=91.584; specificity=86.609), so it was selected to map the probabilities of occurrence for the distribution of *C. nodosa* (Fig. 1).

Estimation of Variable Importance

Marginality scores for environmental variables of ENV+MET and ENV sets of predictors were highly correlated ($r=1$, $p=2.2e^{-16}$). The ENFA marginality scores (Online resource 1) are in general all moderately above zero, indicating that *C. nodosa* occurs predominantly in waters with higher values of each of the variables than the overall mean conditions of the study area here selected. Seventy-two percent of the variables obtained similar scores of marginality, ranging from 0.23 to 0.29, showing no differences between environmental variables and landscape metrics. The specialization scores (Online resource 1) showed a more restricted range of the species, with respect to the study area, on the landscape metric total edge contrast index (TECI) and the environmental variables pH, PAR-max, and SST.

Measures of the importance of the variables in GAMs were estimated from the sets of predictors which performed better: ENV+MET and ENV. Environmental variables performed similarly in ENV+MET and ENV sets, summing the scores of 5.75 and 5.18, respectively (Fig. 2). Unlike ENFA, the relative importance of environmental predictors influencing the ENV+MET model was threefold higher (73.71 %) than that of landscape metrics (26.29 %). Winter SST and phosphate obtained the higher scores of variable importance in both sets of predictors. Among the landscape metrics, the mean shape index (SHAPE_MN) was the most influential predictor.

ENFA scores of coefficients on the marginality factor were not correlated with GAM variable importance scores (Fig. 3). Thus, some variables were found to be important just for one of the models, e.g., phosphate. The variables which showed scores above the first quartile of both models were: salinity, SST of summer, SST of winter, pH, and mean winter wave

Table 3 Landscape metrics used at class level to obtain the species distribution models of *C. nodosa* (based on McGarigal et al. 2012)

Metric	Description
Mean edge contrast index distribution (ECON_MN)	The average edge contrast along the patch perimeter (unit: %)
Area-weighted mean fractal dimension index (FRAC_AM)	Shape complexity across a range of spatial scales. Two times the logarithm of class perimeter divided by the logarithm of class area (unitless)
Mean perimeter-area ratio (PARA_MN)	Ratio of the patch perimeter to area, it measures shape complexity (unitless)
Percentage of landscape (PLAND)	The proportion of the landscape occupied by the class (unit: %)
Mean shape index (SHAPE_MN)	Measure of the complexity of the class shape compared with a standard shape (square) of the same size (unitless)
Total edge contrast index (TECI)	The percentage of edge contrast in relation to the maximum possible (unit: %)

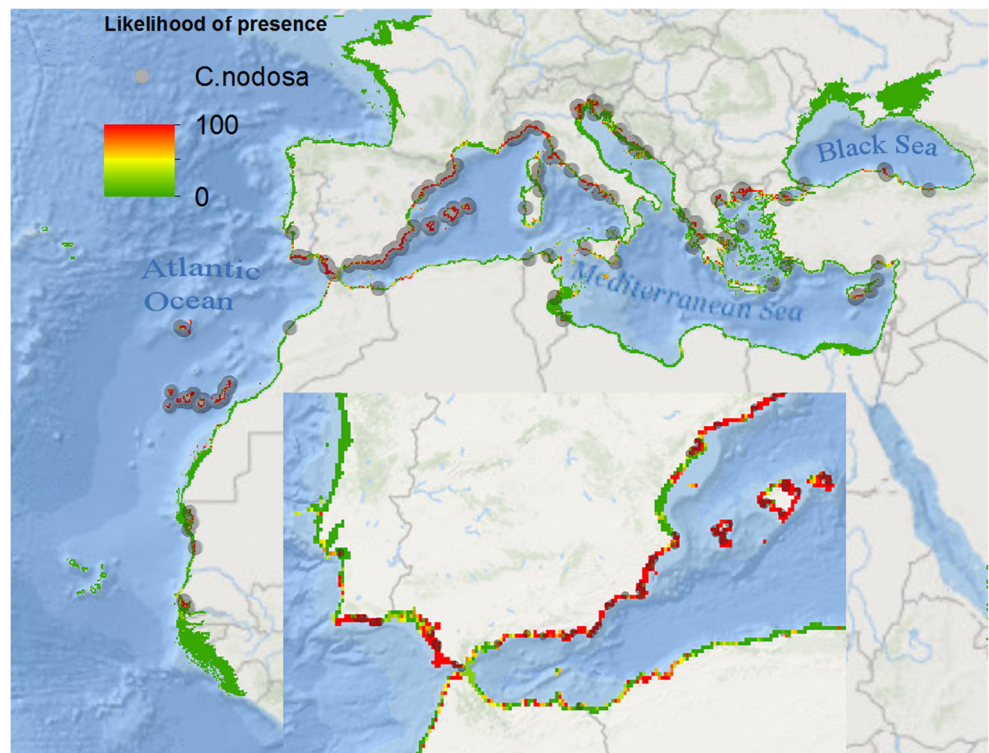
Table 4 Mean validation scores of the 30 models obtained using each technique and the three sets of predictors

Model	AUC mean (\pm SD)	TSS mean (\pm SD)	Sensitivity (\pm SD)	Specificity (\pm SD)
All variables (ENV+MET)				
GLM	0.839 (\pm 0.008)	0.547 (\pm 0.008)	76.939 (\pm 4.377)	78.630 (\pm 4.167)
GBM	0.874 (\pm 0.013)	0.587 (\pm 0.025)	79.398 (\pm 4.417)	80.501 (\pm 3.458)
GAM	0.957 (\pm 0.006)	0.805 (\pm 0.021)	91.858 (\pm 1.678)	89.450 (\pm 1.237)
FDA	0.836 (\pm 0.016)	0.524 (\pm 0.028)	76.885 (\pm 5.719)	76.650 (\pm 5.395)
MARS	0.843 (\pm 0.019)	0.573 (\pm 0.032)	78.525 (\pm 3.821)	79.532 (\pm 4.589)
RF	0.891 (\pm 0.011)	0.624 (\pm 0.018)	80.819 (\pm 2.645)	82.596 (\pm 2.398)
Environmental (ENV)				
GLM	0.829 (\pm 0.009)	0.527 (\pm 0.022)	77.268 (\pm 4.529)	76.259 (\pm 3.896)
GBM	0.871 (\pm 0.013)	0.579 (\pm 0.026)	80.437 (\pm 4.699)	78.883 (\pm 3.823)
GAM	0.928 (\pm 0.007)	0.725 (\pm 0.019)	89.399 (\pm 2.292)	83.883 (\pm 1.399)
FDA	0.832 (\pm 0.020)	0.511 (\pm 0.032)	78.524 (\pm 5.519)	74.411 (\pm 4.899)
MARS	0.846 (\pm 0.013)	0.576 (\pm 0.033)	80.710 (\pm 2.914)	77.953 (\pm 2.601)
RF	0.889 (\pm 0.010)	0.626 (\pm 0.024)	81.257 (\pm 3.991)	82.046 (\pm 3.704)
Landscape metrics (MET)				
GLM	0.672 (\pm 0.013)	0.281 (\pm 0.025)	64.863 (\pm 7.923)	65.429 (\pm 6.598)
GBM	0.731 (\pm 0.015)	0.395 (\pm 0.036)	59.344 (\pm 4.588)	81.160 (\pm 4.194)
GAM	0.714 (\pm 0.011)	0.323 (\pm 0.020)	68.852 (\pm 9.541)	64.774 (\pm 8.614)
FDA	0.714 (\pm 0.014)	0.361 (\pm 0.020)	66.885 (\pm 4.529)	70.401 (\pm 4.146)
MARS	0.710 (\pm 0.015)	0.357 (\pm 0.030)	65.737 (\pm 3.932)	70.968 (\pm 3.265)
RF	0.744 (\pm 0.018)	0.395 (\pm 0.037)	63.442 (\pm 7.089)	76.991 (\pm 7.370)

The best models are highlighted in bold

SD standard deviation, GLM generalized linear model, GBM generalized boosting model, GAM generalized additive model, FDA flexible discriminant analysis, MARS multiple adaptive regression splines, RF randomForest

Fig. 1 Distribution predicted by committee averaging ensemble of generalized additive models (GAMs) for *C. nodosa*. Documented presence points of the species are shown as grey circles



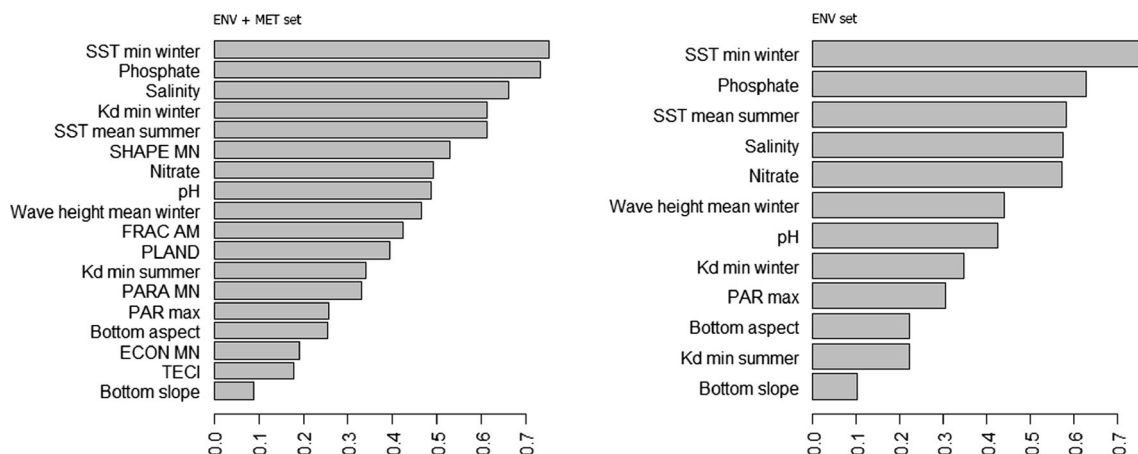


Fig. 2 Scores of the importance of the variables obtained from the best generalized additive models (GAMs) using environmental and landscape metrics (ENV+MET set) and just environmental variables (ENV set). (*ECON_MN* mean edge contrast index distribution; *FRAC_AM* area-

weighted mean fractal dimension index; *PARA_MN* mean perimeter–area ratio; *PLAND* percentage of landscape; *SHAPE_MN* mean shape index; *TECI* total edge contrast index; *Kd* diffuse attenuation coefficient; *PAR* Photosynthetically available radiation; *SST* sea surface temperature)

height. Though with less relevance, three landscape metrics were also found in this interval: area-weighted mean fractal dimension index (*FRAC_AM*), mean shape index

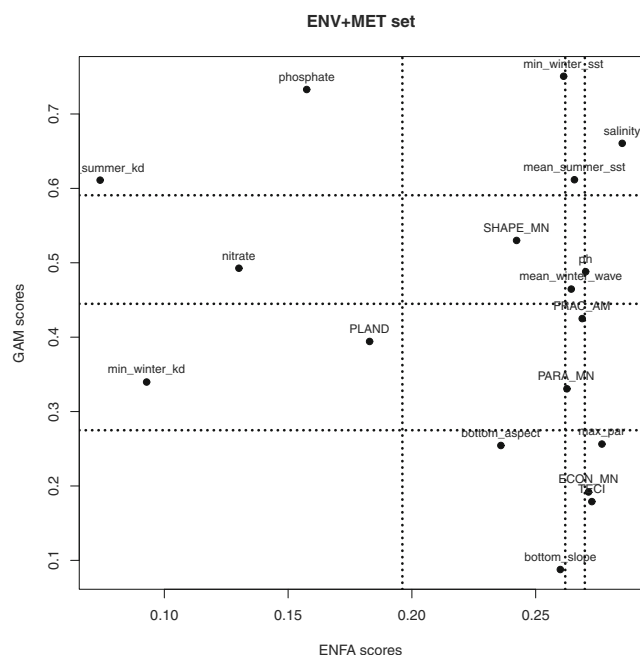


Fig. 3 Relation between the importance found for the variables of environmental and landscape metrics (ENV+MET set) in generalized additive models (GAMs) and environmental niche factor analysis (ENFA). The variables that determine the known distribution are located in the upper zone, as we go down in the right zone of the figure we find those providing information on the potential distribution. ENFA scores refer to the coefficients of the marginality factor of *C. nodosa* on each predictor. Dotted lines indicate the quartiles of the bivariate distribution. (*ECON_MN* mean edge contrast index distribution; *FRAC_AM* area-weighted mean fractal dimension index; *PARA_MN* mean perimeter–area ratio; *PLAND* percentage of landscape; *SHAPE_MN* mean shape index; *TECI* total edge contrast index; *Kd* diffuse attenuation coefficient; *PAR* photosynthetically available radiation; *SST* sea surface temperature)

(*SHAPE_MN*), and mean perimeter–area ratio (*PARA_MN*). Figure 4 provides the correlations between the environmental variables identified as relevant for ensemble of GAMs and ENFA model. In this figure, some environmentally and spatially distant localities in the Black Sea and Mauritania have been identified (see “Discussion”). 2D and 3D response plots (Figs. 5 and 6) allowed detection of main thresholds for the most relevant variables and their pairwise interactions.

Estimation of Potential Habitats

The MD projection of *C. nodosa* (Fig. 7) was less accurate predicting the presences than the ensemble committee averaging of GAMs (Fig. 8); by contrast, MD predicted better the presences than the ensemble discarded (mean of probabilities). To enable a comparison, we applied the same threshold to MD than that applied to GAM ensemble; suitable area predicted was higher in MD model (accounting for 17.41 % of the study area) than in GAM ensemble (14.72 %); MD provided a greater estimate of potentially suitable areas. From the suitable area, 29.9 % was congruent for both models. In contrast, both models coincidentally found 76.71 % of the study area as unsuitable. The range of values of the environmental variables found in suitable habitats predicted by MD is narrower than ensemble model prediction (Table 2).

Discussion

This study has been able to demonstrate for the first time that large-scale predictive models provide a fine delimitation of the variables constraining the distribution of a seagrass species throughout its range. Though the set of environmental

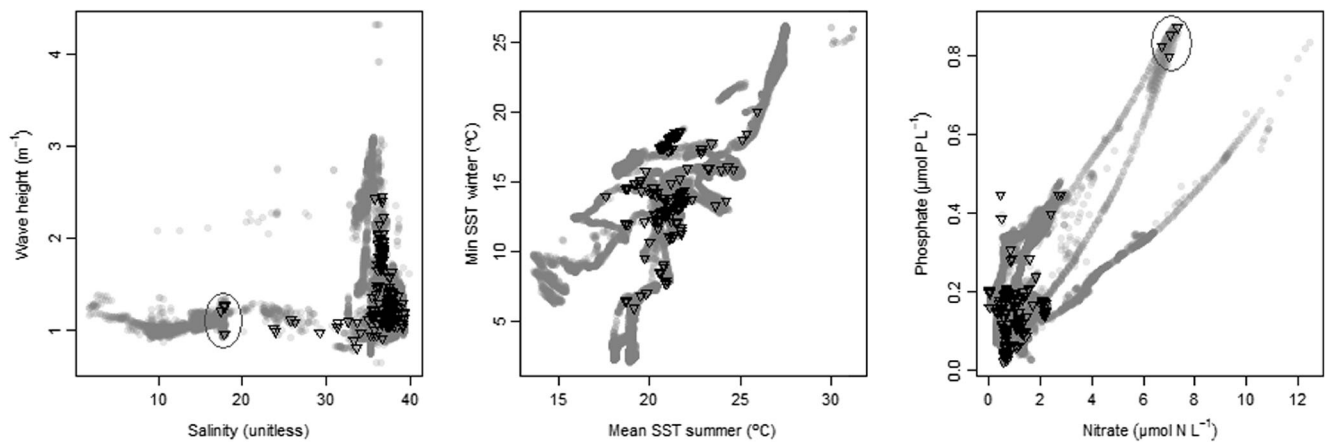


Fig. 4 Distribution of values of environmental variables found among the most important for ensemble of generalized additive models (GAMs) and environmental niche factor analysis (ENFA). Grey dots represent all

the cells of the study area. Triangles correspond to presences of *C. nodosa*. Ellipse in the left includes Black Sea records, while ellipse in the right includes Mauritanian records (SST: sea surface temperature)

variables achieved the best performance, the inclusion of landscape metrics improved the different accuracy measures of GAMs ensemble from 2.7 % to 9.9 % (Table 4).

A comparison between the two variable importance measures (Fig. 3) showed that some variables were found relevant

in both procedures: SST and salinity. SST seems to be an important factor delimiting *C. nodosa* distribution: the species tends to be absent from the coldest and warmest waters of the study area (Table 2) (Figs. 4, 5, and 6). This result is similar for winter and summer SST, and it is comparable for all values

Fig. 5 2D response plots for the most relevant variables found in the ensemble “committee averaging” of generalized additive models (GAMs). It should be noted that interaction terms of GAMs are not taken into account in these plots, thus the individual responses obtained here are different from those presented in Table 1. Horizontal dotted line shows the cut-off used for binary transformation of the ensemble. Breakpoints at which changes in the values occur in relation to the likelihood of presence are marked by open circles. Top boxplots display the distribution of occurrence records over the range of each variable in the study area (SST: sea surface temperature)

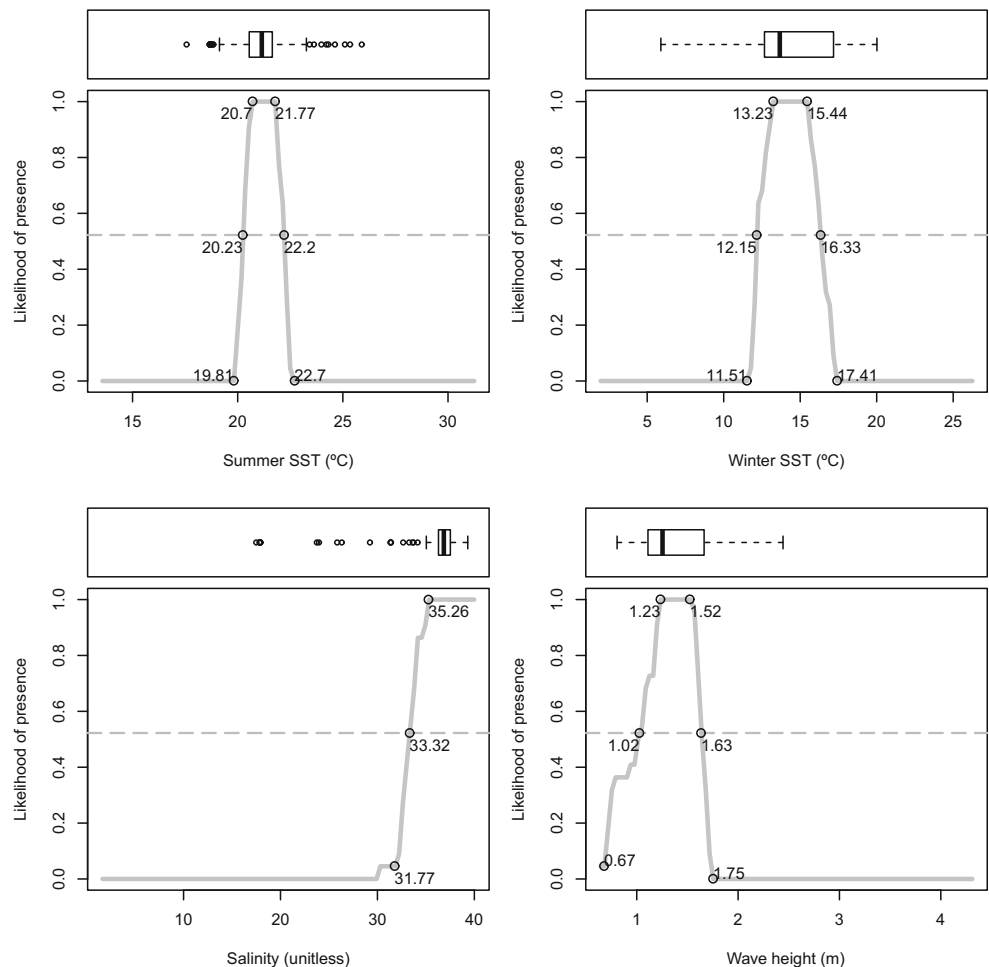
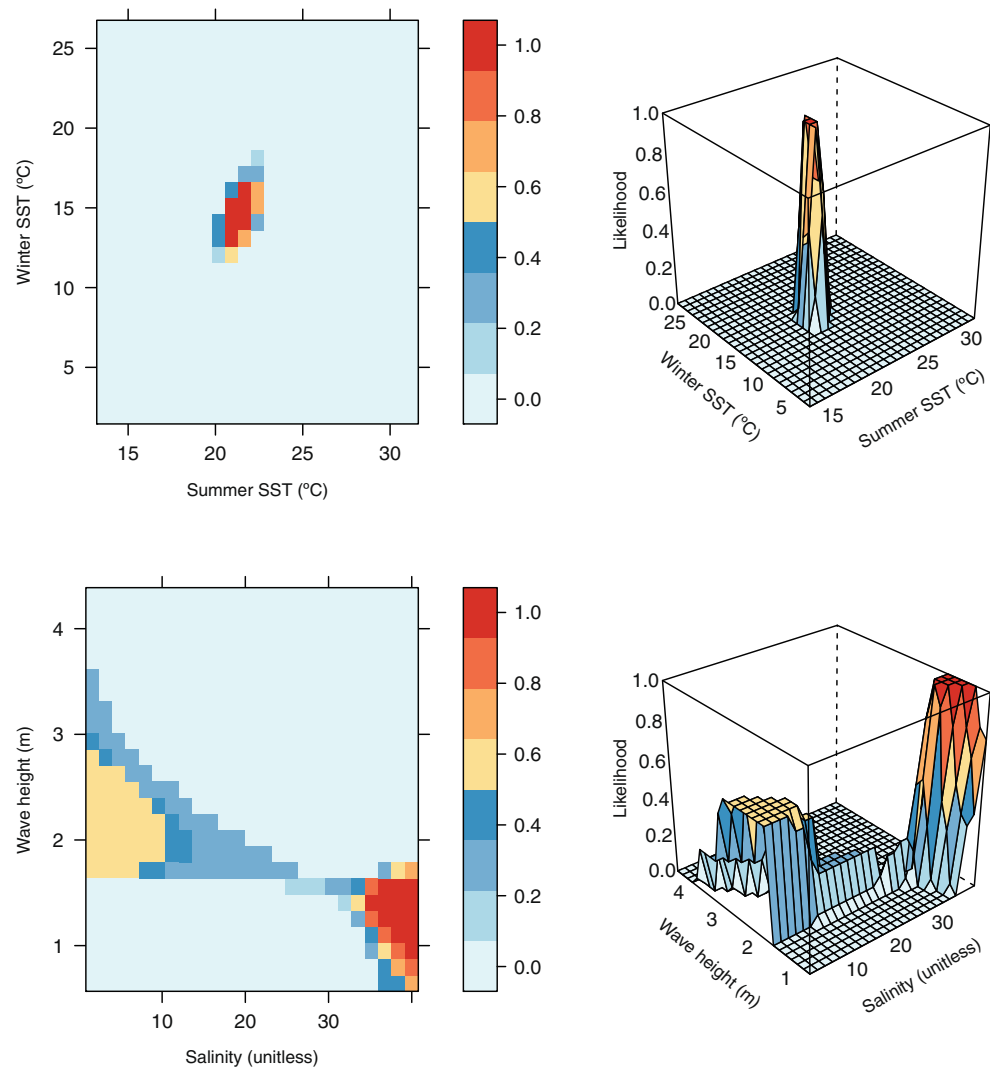


Fig. 6 3D bivariate response plots showing the interaction between the most relevant variables found in the ensemble “committee averaging” of generalized additive models (GAMs) in the environmental space of the study area. (SST: sea surface temperature)

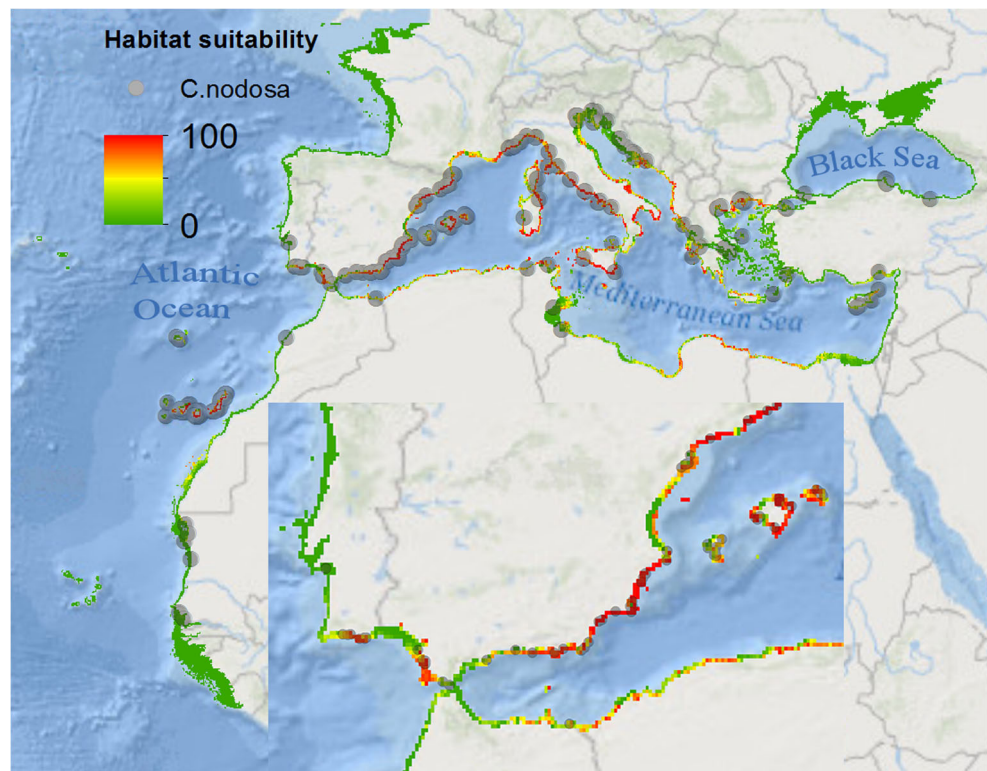


used (min., max., or mean), due to the high correlation between them. Suitable areas found for the species seem to be constrained to temperatures ranging from 5.8 °C to 20.7 °C in winter and above 14.57 °C and below 26.4 °C in summer (Table 2). The optimal temperature response for the highest likelihood of occurrence of *C. nodosa* ranges from 20.7 °C to 21.77 °C in summer and 13.2 °C to 15.44 °C in winter, though slightly wider values are found just for suitable areas (see Figs. 5 and 6). Previous research has indicated that temperature fluctuations have a greater impact on *C. nodosa* than on the rest of the seagrass species occurring in the NW Mediterranean (Marbà et al. 1996). Although little information is available on the range of extreme temperatures limiting *C. nodosa* populations at biogeographical scales, our results are consistent with the report of an upper threshold of 29 °C to 30 °C beyond which *C. nodosa* declines (Olsen et al. 2012). It is therefore likely that SST is defining both the northern and southern range limits of *C. nodosa*'s distribution. In contrast, our optimal temperature values are not as high as 24.5 °C, the

average of optimal growth reported by a review based on five studies for this species (Lee et al. 2007).

Salinity has also been identified as an important predictor for the species, suitable habitats ranging between 17.5 and 39.3 (Table 2), and optimal values ranging from 35.26 to 39.9 (Figs. 5 and 6). In the study area, low salinity waters (ranging from 1.6 to 18) were found in estuarine environments in the Black Sea, where some presence records exist (Fig. 4) as the eastern limit of distribution for *C. nodosa*. Laboratory studies have also shown significant effects of salinity on the growth and survival of *C. nodosa*, reaching lethal levels at salinities below 17 and above 50 (Fernández-Torquemada and Sánchez-Lizaso 2006) or 54 (Pagès et al. 2010). Considering that the maximum salinity in our study area is 39.98 (Table 2), hypersalinity may not be restricting the occurrence of *C. nodosa*, while low salinity might be responsible for the lack of suitable habitats across much of the Black Sea (Fig. 1 and 7). Nonetheless, it is important to bear in mind that an increase of salinity in punctual areas, caused for example

Fig. 7 Potential distribution predicted by Mahalanobis distance algorithm for *C. nodosa*



by desalination plants or small lagoons with restricted circulation, could be damaging for *Cymodocea* meadows.

Phosphate, and nitrate to a lesser degree, had considerable influence on the GAM predictions. Accordingly, the finding of abundant *C. nodosa* on the Mauritanian coast, where phosphate concentration is above $0.7 \mu\text{mol P L}^{-1}$, and nitrate $> 6 \mu\text{mol N L}^{-1}$, confirms that the species is also able to tolerate

relatively high nutrient concentrations, well above the average values for the study area (see Fig. 4). A possible explanation for these high scores of importance of the variables in GAMs might be the rarity of waters with high concentrations of phosphate and nitrate among the cells of the study area, which limits also the abundance of observations in waters with these conditions. These results are, however, also consistent with experiments demonstrating nutrient, particularly P, limitation in situ for *C. nodosa* (Pérez et al. 1991, 1994), and observations concluding that colonization, but both seeds and clonal dispersal, is also nutrient-limited (Duarte and Sand-Jensen 1996). However, the limitation on this factor is difficult to estimate at large scales as seagrasses are able to assimilate nutrients from both column water and sediments, nutrients in sediments being a better source as they do not produce adverse effects on growth associated with the proliferation of epiphytes in nutrient rich waters (Duarte 1995; Lee et al. 2007). Thus, our data on water column dissolved nutrients may not reflect the full effects of nutrient limitation, as they do not include sediment sources (e.g., Pérez and Romero 1992).

Other factors identified as relevant just by MD are: photosynthetically available radiation (PAR), pH, and mean winter wave height. We found optimal values of mean winter wave height between 1.2 and 1.5 m (Figs. 5 and 6) and no occurrences above 2.5 m. The importance of waves was lower than expected through field experience maybe because this resolution (10 km) does not capture the variation of wave height at the scale that meadows require. PAR and pH do not seem to be

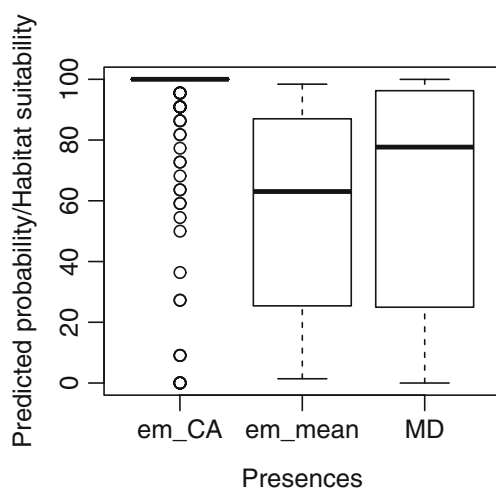


Fig. 8 Comparison among the predicted probabilities of the ensembles of generalized additive models (GAMs) and the habitat suitability of Mahalanobis distance, for the 202 presences of *C. nodosa* (MD: Mahalanobis distance; *em_CA*: ensemble “committee averaging” of GAMs; *em_mean*: ensemble “mean of probabilities” of GAMs)

major factors in delimiting the distribution of the species in the study area according to available data (Table 2). Moreover, the validity of pH to delineate the occurrence of *C. nodosa* is questionable, as seagrass are able, through their photosynthetic activity, to elevate pH (Duarte et al. 2013; Hendriks et al. 2014). Thus, the use of pH in predicting the occurrence of *C. nodosa* may be circular and especially problematic if in situ pH measurements close to the meadows are used, although this effect is attenuated using coarser data resolution as here.

The differences found in the importance of the variables when using different approaches can be explained by the dissimilar computational procedures used by ENFA and that used in the presence-absence models. While ENFA marginality factor scores provide information about how much each variable accounts for the differentiation between the species niche and the global area, the importance of the variables for GAMs ensemble measures the influence of each predictor in a model able to differentiate presences from absences. Due to the lack of real absences, the pseudo-absences used here could be contingent or methodological rather than environmental (see Lobo et al. 2010), thus the importance obtained with the last procedure must be interpreted as that retrieved from the known distribution of the species. Hence, using the information provided by the two procedures can be a useful practice to discriminate the importance of the variables between the known and the potential distribution and also when it is difficult to choose a threshold to discard previously the lowest scored variables in ENFA. Our results (Fig. 3) show a gradation in the contribution of different variables to determining the known distribution of *C. nodosa* as well as those that provide information on its potential niche.

Landscape Metrics as Predictors

The current study found that the incorporation of landscape metrics as predictors in SDMs enhanced, albeit modestly, the predictive power of the models. Although when assessed separately, environmental variables were superior predictors to landscape metrics, the set of predictors consisting of both environmental variables and landscape metrics was the most effective to predict *C. nodosa* distribution. According to the difference between the values of landscape metrics found in presence and suitable cells and their range of variation in the study area (Table 2), the species seems to occur preferentially on coastlines with high edge contrast and not too irregular or complex shapes. Here, the edge contrast reflects the magnitude of difference between the coastline and the rest of the landscape and is described by TECI and ECON_MN. In combination with measures related to the complexity of the shape as SHAPE_MN, FRAC_AM and PARA_MN, the models increase their ability to discern between convoluted shapes and straight lines. The coastal features defined by these

metrics could be the ones that contribute to better dissipate the wave energy and allow the establishment of seagrass meadows.

The finding of the predictive power of landscape metrics has important implications for future studies to predict the distribution of species dependent on coastal morphology. Although the resolution used in this study was not the most applicable for the inclusion of these metrics (Chefaoui 2014), the models obtained here are very accurate. As behavior of metrics is scale-dependent (e.g., Turner et al. 1989; McGarigal et al. 2012), these ranges could vary in studies using other resolutions and moving-window sizes. The extent (overall size of the study area) used here is greater than those previously used, thus these results are encouraging, showing that these landscape metrics may be effective predictors across a wide range of scales. In addition, using finer scales would also allow us to distinguish better among different coastal morphologies such as capes, bays, and beaches, because a smaller resolution would reduce the stair-stepping pattern of perimeter line segments (see McGarigal et al. 2012).

Potential Distribution of *C. nodosa*

Data comparing the prediction results obtained for the presences (Fig. 8) showed that the ensemble of GAMs selected (committee averaging) is more accurate than MD. MD, the profile method, provided a wider potential distribution as absences are not used to calibrate the model, while the GAMs obtained with random pseudo-absences tended to obtain restricted predictions, closed to the occurrences used (Figs. 1 and 7). These results are consistent with those of other studies (e.g., Zaniewski et al. 2002; Chefaoui and Lobo 2008; Mateo et al. 2010), and these two hypotheses on the distribution of the species provide a basis for comparison.

There are two important observations from the obtained models. First, there are large spans of coastline within the distributional areas that are unsuitable to support *C. nodosa* meadows. These include estuarine areas in the Black Sea, where low salinity appears to limit the occurrence of *C. nodosa*, much of the northern African coast and the Eastern Mediterranean, where maximum summer temperatures above the 29 °C to 30 °C threshold may limit *C. nodosa*, and areas in the Gulf of Lyons, Northern Portugal, NW Spain, W France, the Italian coast of the Adriatic Sea, and the Atlantic African coast, where very low water temperatures in winter or during upwelling seasons, in the case of the Atlantic African Coast, may be limiting (Fig. 1). The second important finding is that *C. nodosa* has not been reported to be occupying all available suitable habitats, according to MD predictions (Fig. 7). Most of the potential suitable areas where the seagrass is not recorded are located in coastal regions of North Africa and the Adriatic coast of Italy. North African countries such as Algeria, Libya, and Egypt are under-surveyed regions due to

limited capacity and investment in research and development possibly aggravated now by politician instability. Thus, the lack of confirmed absence of the species in this area prevents us from verifying whether it is in a state of non-equilibrium with the environment. On the contrary, some regions predicted as non-suitable by both models seem quite reliable. In Cape Verde Islands, where no records were found, a very low probability of presence was predicted. The absence of *C. nodosa* from these islands has been recently confirmed (P. Wirtz, personal communication). The Adriatic coast of Italy, another region without records of the species, has been predicted as suitable to some extent by MD while the ensemble of GAMs predicted very low probabilities. In this case, the absences and GAMs' predictions seem also quite reliable as the marine ecosystem in this area receives significant research and observational effort.

Though the northern limit of the species on the coast of Portugal is well delimited by recent surveys (Cunha et al. 2013), the southern limit on the west coast of Africa is not yet clearly defined due to lack of survey effort in the area (Cunha and Araújo 2009). It has been hypothesized that the species may be present further south beyond the coast of Senegal (Cunha and Araújo 2009); however, both models identified these regions as unsuitable, probably due to the higher winter and summer SST (other possible explanations, as seasonal upwelling cold water filaments, are not well captured by the data used).

The information derived from the models developed here can be useful to find potential suitable habitats for restoration of meadows. Restoration projects (e.g., Biomares LIFE project, LIFE06 NAT/P/000192) need a great certainty about the suitability of the places where the seagrass are transplanted to be successful, as they need to be sheltered from storms and offer appropriate environmental conditions. The importance of habitat selection in seagrass restoration has been emphasized by several authors, and GIS-based habitat suitability maps have already been successfully employed (see Van Katwijk et al (2009) for a review). Indeed, seagrass restoration projects often suffer from relatively low success, possibly because some of these efforts might target unsuitable areas. Some potential suitable habitats detected in the North of Morocco and Algeria by both models could be considered as areas where *C. nodosa*, if not already present, could survive with higher reliability.

Moreover, the models developed here can also be used to predict, when combined with downscaled climatic models, the future distribution of suitable sites, identifying areas where *C. nodosa* is presently absent but may grow in the future. Indeed, recent analyses have shown the poleward migration of distributional limits of marine species by, on average, 30.6 km decade⁻¹ (Poloczanska et al. 2013). This is a rough average value derived from a synthesis of studies, none of which address seagrass species. However, the model produced here allow for

more precise predictions as the models suggest that the poleward spread of the range of *C. nodosa* should track the 26.4 °C isotherm for the trailing edge and the 5.84 °C isotherm for the leading edge, the temperature thresholds for the occurrence of *C. nodosa* derived from this analysis. These findings enhance our understanding of the factors that can be shaping *C. nodosa*'s environmental niche throughout its range of distribution, as well as the location of suitable areas that could not have been derived from local or experimental studies. We emphasize the relevance of SST, which seems to be defining the northern and southern limit of the species. The importance for the conservation of these ecosystems makes it convenient for future studies of climate-related marine range shifts, along the Atlantic part of its distribution, which is subjected to a latitudinal thermal gradient.

Acknowledgments We thank the three anonymous referees for their helpful comments. We are also grateful to Cymon Cox for his guidance to run R on the CCMAR GYRA cluster and to Damien Georges and Wilfried Thuiller for their help with biomod2. RC was supported by the postdoctoral fellowship SFRH/BPD/85040/2012 from the Fundação para a Ciência e a Tecnologia (FCT, Portugal). JA was supported by the postdoctoral fellowship CCMAR/BPD/0045/2013 from FCT. We acknowledge FCT project EXTANT (EXCL/AAG-GLO/0661/2012).

References

- Alberto, F., S. Massa, P. Manent, E. Diaz-Almela, S. Arnaud-Haond, C.M. Duarte, and E.A. Serrão. 2008. Genetic differentiation and secondary contact zone in the seagrass *Cymodocea nodosa* across the Mediterranean–Atlantic transition region. *Journal of Biogeography* 35: 1279–1294.
- Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43: 1223–1232.
- Araújo, M.B., and M. New. 2007. Ensemble forecasting of species distributions. *Trends in Ecology & Evolution* 22: 42–47.
- Araújo, M.B., and A.T. Peterson. 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology* 93: 1527–1539.
- Barbet-Massin, M., F. Jiguet, C.H. Albert, and W. Thuiller. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* 3: 327–338.
- Bryan, T.L., and A. Metaxas. 2007. Predicting suitable habitat for deep-water gorgonian corals on the Atlantic and Pacific continental margins of North America. *Marine Ecology Progress Series* 330: 113–126.
- Cabaço, S., Ó. Ferreira, and R. Santos. 2010. Population dynamics of the seagrass *Cymodocea nodosa* in Ria Formosa lagoon following inlet artificial relocation. *Estuarine, Coastal and Shelf Science* 87: 510–516.
- Calenge, C. 2006. The package adehabitat for the R software: a tool for the analysis of space and habitat use by animals. *Ecological Modelling* 197: 516–519.
- Calenge, C., G. Darmon, M. Basille, A. Loison, and J.M. Jullien. 2008. The factorial decomposition of the Mahalanobis distances in habitat selection studies. *Ecology* 89: 555–566.

- Cancemi, G., M.C. Buia, and L. Mazzella. 2002. Structure and growth dynamics of *Cymodocea nodosa* meadows. *Scientia Marina* 66: 365–373.
- Chefaoui, R.M. 2014. Landscape metrics as indicators of coastal morphology: a multi-scale approach. *Ecological Indicators* 45: 139–147.
- Chefaoui, R.M., and J.M. Lobo. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling* 210: 478–486.
- Chen, Z., F.E. Muller-Karger, and C. Hu. 2007. Remote sensing of water clarity in Tampa Bay. *Remote Sensing of Environment* 109: 249–259.
- Clark, J.D., J.E. Dunn, and K.G. Smith. 1993. A multivariate model of female black bear habitat use for a geographic information system. *The Journal of Wildlife Management* 57: 519–526.
- Costanza, R., R. d'Arge, R. De Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R.V. O'Neill, J. Paruelo, R.G. Raskin, P. Sutton, and M. van den Belt. 1998. The value of the world's ecosystem services and natural capital. *Ecological Economics* 25: 3–15.
- Cunha, A.H., and A. Araújo. 2009. New distribution limits of seagrass beds in West Africa. *Journal of Biogeography* 36: 1621–1622.
- Cunha, A.H., J.F. Assis, and E.A. Serrão. 2013. Seagrasses in Portugal: a most endangered marine habitat. *Aquatic Botany* 104: 193–203.
- Cushman, S.A., K. McGarigal, and M.C. Neel. 2008. Parsimony in landscape metrics: strength, universality, and consistency. *Ecological Indicators* 8: 691–703.
- Duarte, C.M. 1995. Submerged aquatic vegetation in relation to different nutrient regimes. *Ophelia* 41: 87–112.
- Duarte, C. M. 2011. Seagrass meadows. <http://www.eoearth.org/view/article/155952>.
- Duarte, C.M., and K. Sand-Jensen. 1996. Nutrient constraints on establishment from seed and on vegetative expansion of the Mediterranean seagrass *Cymodocea nodosa*. *Aquatic Botany* 54: 279–286.
- Duarte, C. M., J. Borum, F. Short, and D. Walker 2008. Seagrass ecosystems: Their global status and prospects. In *Aquatic Ecosystems*, ed N. Polunin, 281–294. Cambridge University Press.
- Duarte, C.M., I.E. Hendriks, T.S. Moore, Y.S. Olsen, A. Steckbauer, L. Ramajo, J. Carstensen, J.A. Trotter, and M. McCulloch. 2013. Is ocean acidification an open-ocean syndrome? Understanding anthropogenic impacts on seawater pH. *Estuaries and Coasts* 36: 221–236.
- Elith, J., S. Ferrier, F. Huettmann, and J. Leathwick. 2005. The evaluation strip: A new and robust method for plotting predicted responses from species distribution models. *Ecological Modelling* 186: 280–289.
- Enriquez, S., N. Marbà, J. Cebrián, and C.M. Duarte. 2004. Annual variation in leaf photosynthesis and leaf nutrient content of four Mediterranean seagrasses. *Botanica Marina* 47: 295–306.
- Espino, F., F. Tuya, A. Brito, and R.J. Haroun. 2011. Ichthyofauna associated with *Cymodocea nodosa* meadows in the Canarian Archipelago (central eastern Atlantic): community structure and nursery role. *Ciencias Marinas* 37: 157–174.
- Fernández-Torquemada, Y., and J.L. Sánchez-Lizaso. 2006. Effects of salinity on growth and survival of *Cymodocea nodosa* Ucria Ascherson and *Zostera noltii* Hornemann. *Biologia Marina Mediterranea* 13: 46–47.
- Fielding, A.H., and J.F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49.
- Gartner, A., F. Tuya, P.S. Lavery, and K. McMahon. 2013. Habitat preferences of macroinvertebrate fauna among seagrasses with varying structural forms. *Journal of Experimental Marine Biology and Ecology* 439: 143–151.
- GEBCO General Bathymetric Chart of the Oceans 2010. Br. Oceanogr. Data Cent., Liverpool, U. K. http://www.gebco.net/data_and_products/gridded_bathymetry_data/.
- GEODAS-NG Software version 1.1.1.1. National Geophysical Data Center, Boulder, Colorado. http://www.ngdc.noaa.gov/mgg/gdas/gx_announce.html Accessed Aug 2013.
- Gohin, F., S. Loyer, M. Lunven, C. Labry, J.M. Froidefond, D. Delmas, M. Huret, and A. Herbland. 2005. Satellite-derived parameters for biological modelling in coastal waters: illustration over the eastern continental shelf of the Bay of Biscay. *Remote Sensing of Environment* 95: 29–46.
- Green, E. P. and F. T. Short (eds) 2003. World atlas of seagrasses. Univ of California Press.
- GSHHG. A global self-consistent, hierarchical, high-resolution geography database version 2.2.2. National Geophysical Data Center, Boulder, Colorado. <http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html>. Accessed Aug 2013.
- Guidetti, P., and S. Bussotti. 2000. Fish fauna of a mixed meadow composed by the seagrasses *Cymodocea nodosa* and *Zostera noltii* in the Western Mediterranean. *Oceanologica Acta* 23: 759–770.
- Guiry, M. D., and G. M. Guiry 2014. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. <http://www.algaebase.org>. Accessed 13 Feb 2014.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993–1009.
- Hendriks, I.E., Y.S. Olsen, L. Ramajo, L. Basso, A. Steckbauer, T. Moore, J. Howard, and C.M. Duarte. 2014. Photosynthetic activity buffers ocean acidification in seagrass meadows. *Biogeosciences* 11: 333–346.
- Hirzel, A.H., J. Hausser, D. Chessel, and N. Perrin. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology* 7: 2027–2036.
- Huot, Y., C.A. Brown, and J.J. Cullen. 2005. New algorithms for MODIS sun-induced chlorophyll fluorescence and a comparison with present data products. *Limnology and Oceanography: Methods* 3: 108–130.
- Jiménez-Valverde, A., J.M. Lobo, and J. Hortal. 2008. Not as good as they seem: The importance of concepts in species distribution modelling. *Diversity and Distributions* 14: 885–890.
- Lee, K.S., S.R. Park, and Y.K. Kim. 2007. Effects of irradiance, temperature, and nutrients on growth dynamics of seagrasses: a review. *Journal of Experimental Marine Biology and Ecology* 350: 144–175.
- Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. *R News* 2: 18–22.
- Lobo, J.M., A. Jiménez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. *Ecography* 33: 103–114.
- Malea, P., T. Kevrekidis, and M. Potouroglou. 2013. Seasonal variation of trace metal Mn, Zn, Cu, Pb, Co, Cd concentrations in compartments of the seagrass *Cymodocea nodosa*. *Botanica Marina* 56: 169–184.
- Marbà, N., J. Cebrián, S. Enriquez, and C.M. Duarte. 1996. Growth patterns of Western Mediterranean seagrasses: species-specific responses to seasonal forcing. *Marine Ecology Progress Series* 133: 203–215.
- Mateo, R.G., T.B. Croat, A.M. Felicísimo, and J. Muñoz. 2010. Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Diversity and Distributions* 16: 84–94.
- McGarigal K., Cushman S. A., and Ene E. 2012. FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps. University of Massachusetts, Amherst. <http://www.umass.edu/landeco/research/fragstats/fragstats.html>. Accessed Aug 2013.

- Olivé, I., J.J. Vergara, and J.L. Pérez-Lloréns. 2013. Photosynthetic and morphological photoacclimation of the seagrass *Cymodocea nodosa* to season, depth and leaf position. *Marine Biology* 160: 285–297.
- Olsen, Y.S., M. Sánchez-Camacho, N. Marbà, and C.M. Duarte. 2012. Mediterranean seagrass growth and demography responses to experimental warming. *Estuaries and Coasts* 35: 1205–1213.
- Orth, R.J., T.J. Carruthers, W.C. Dennison, C.M. Duarte, J.W. Fourqurean, K.L. Heck, A.R. Hughes, G.A. Kendrick, W.J. Kenworthy, S. Olyarnik, F.T. Short, M. Waycott, and S.L. Williams. 2006. A global crisis for seagrass ecosystems. *Bioscience* 56: 987–996.
- Pages, J.F., M. Pérez, and J. Romero. 2010. Sensitivity of the seagrass *Cymodocea nodosa* to hypersaline conditions: a microcosm approach. *Journal of Experimental Marine Biology and Ecology* 386: 34–38.
- Pérez, M., and J. Romero. 1992. Photosynthetic response to light and temperature of the seagrass *Cymodocea nodosa* and the prediction of its seasonality. *Aquatic Botany* 43: 51–62.
- Pérez, M., J. Romero, C.M. Duarte, and K. Sand-Jensen. 1991. Phosphorus limitation of *Cymodocea nodosa* growth. *Marine Biology* 109: 129–133.
- Pérez, M., C.M. Duarte, J. Romero, K. Sand-Jensen, and T. Alcoverro. 1994. Growth plasticity in *Cymodocea nodosa* stands: the importance of nutrient supply. *Aquatic Botany* 47: 249–264.
- Poloczanska, E.S., C.J. Brown, W.J. Sydeman, W. Kiessling, D.S. Schoeman, P.J. Moore, K. Brander, J.F. Bruno, L. Buckley, M.T. Burrows, C.M. Duarte, B.S. Halpern, J. Holding, C.V. Kappel, M.I. O'Connor, J.M. Pandolfi, C. Parmesan, F. Schwing, S.A. Thompson, and A.J. Richardson. 2013. Global imprint of climate change on marine life. *Nature Climate Change* 3: 919–925.
- R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/> Accessed Mar 2013.
- Schaeffer, P., Y. Faugere, J.F. Legeais, A. Ollivier, T. Guinle, and N. Picot. 2012. The CNES_CLS11 global mean sea surface computed from 16 years of satellite altimeter data. *Marine Geodesy* 35: 3–19.
- Stark, J.D., C.J. Donlon, M.J. Martin and M.E. McCulloch 2007. OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system. In: OCEANS 2007-Europe, IEEE, 1–4.
- Thuiller, W., D. Georges, and R. Engler 2013. biomod2: Ensemble platform for species distribution modeling. R package version 3.1-25.
- Tittensor, D.P., C. Mora, W. Jetz, H.K. Lotze, D. Ricard, E.V. Berghe, and B. Worm. 2010. Global patterns and predictors of marine biodiversity across taxa. *Nature* 466: 1098–1101.
- Turner, M.G., R.V. O'Neill, R.H. Gardner, and B.T. Milne. 1989. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology* 3: 153–162.
- Tuya, F., H. Hernandez-Zerpa, F. Espino, and R. Haroun. 2013. Drastic decadal decline of the seagrass *Cymodocea nodosa* at Gran Canaria eastern Atlantic: interactions with the green algae *Caulerpa prolifera*. *Aquatic Botany* 105: 1–6.
- Tuya, F., R. Haroun, and F. Espino. 2014a. Economic assessment of ecosystem services: monetary value of seagrass meadows for coastal fisheries. *Ocean & Coastal Management* 96: 181–187.
- Tuya, F., L. Png-Gonzalez, R. Riera, R. Haroun, and F. Espino. 2014b. Ecological structure and function differs between habitats dominated by seagrasses and green seaweeds. *Marine Environmental Research* 98: 1–13.
- Tyberghein, L., H. Verbruggen, K. Pauly, C. Troupin, F. Mineur, and O. De Clerck. 2012. Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* 21: 272–281.
- Valle, M., G. Chust, A. del Campo, M.S. Wisz, S.M. Olsen, J.M. Garmendia, and Á. Borja. 2014. Projecting future distribution of the seagrass *Zostera noltii* under global warming and sea level rise. *Biological Conservation* 170: 74–85.
- Van Katwijk, M.M., A.R. Bos, V.N. De Jonge, L.S.A.M. Hanssen, D.C.R. Hermus, and D.J. De Jong. 2009. Guidelines for seagrass restoration: importance of habitat selection and donor population, spreading of risks, and ecosystem engineering effects. *Marine Pollution Bulletin* 58: 179–188.
- Venables, W.N., and B.D. Ripley. 2002. *Modern applied statistics with S*. New York: Springer.
- Verdiell-Cubedo, D., F.J. Oliva-Paterna, and M. Torralva-Forero. 2007. Fish assemblages associated with *Cymodocea nodosa* and *Caulerpa prolifera* meadows in the shallow areas of the Mar Menor coastal lagoon. *Limnetica* 26: 341–350.
- Waycott, M., C.M. Duarte, T.J. Carruthers, R.J. Orth, W.C. Dennison, S. Olyarnik, A. Calladine, J.W. Fourqurean, K.L. Heck Jr., A.R. Hughes, G.A. Kendrick, W.J. Kenworthy, F.T. Short, and S.L. Williams. 2009. Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences* 106: 12377–12381.
- Welch, B.L. 1947. The generalization of student's' problem when several different population variances are involved. *Biometrika* 34: 28–35.
- Zaniewski, A.E., A. Lehmann, and J.M. Overton. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling* 157: 261–280.
- Zarranz, M.E., N. González-Henríquez, P. García-Jiménez, and R.R. Robaina. 2010. Restoration of *Cymodocea nodosa* seagrass meadows through seed propagation: germination in vitro, seedling culture and field transplants. *Botanica Marina* 53: 173–181.