Chapter 7

Modelling the past, present and future distribution of invasive seaweeds in Europe

Samuel Bosch[1,2], Eduardo Gomez Giron[1], Brezo Martínez[3] and Olivier De Clerck[1]

[1]*Research Group Phycology, Biology Department, Ghent University, Krijgslaan 281/S8, 9000 Ghent, Belgium*
[2]*Flanders Marine Institute (VLIZ), InnovOcean site, Wandelaarskaai 7, 8400 Ostend, Belgium*
[3]*Biology and Geology Department, Rey Juan Carlos University, Tulipán sn.,Móstoles 28933, Spain*

Manuscript in preparation.

Abstract

Invasive species can cause significant problems at ecosystem, economic and social levels. Assessing the potential geographic range of such species in invaded regions is therefore increasingly promoted for proactive ecological management. Unfortunately, because invasive species are by definition not at equilibrium within recipient environments, there is considerable uncertainty on how to model their distributions. In this study we evaluated the performance of species distribution modelling, trained with native and/or non-European distribution records, as a tool for predicting the spread of invasive seaweeds at various stages of the invasion process. We estimated the level of niche expansion observed under analog and non-analog conditions and assessed which areas in Europe are expected to be disproportionally impacted by migrations of introduced seaweeds due to climate change. Our results indicate that due to considerable niche expansion in non-analog conditions including only native records is generally not sufficient to predict the range of invasive species. Including distribution records from non-European invaded regions on the other hand significantly increases the predictive power of the models and reduces the measured niche expansion in analog and non-analog conditions considerably. The European change and turnover maps combined with an assessment of the uncertainty therein predict an increased habitat suitability in northern Europe (northern UK, Scandinavia, Iceland), while southern European are likely to become less suitable. In addition to the overall picture, uncertainty in the estimates is apparent for specific regions but correlates only moderately to changes in habitat suitability.

Introduction

Invasive species rank as one of the greatest threats to marine coastal biodiversity (McGeoch et al., 2010; Seebens et al., 2013). The European shores sadly stand out as a hotspot for introduced species (Molnar et al., 2008) and seaweeds represent one of the largest groups of marine aliens, accounting for 10 to 30% of all marine introduced species in Europe (Schaffelke et al., 2006; Williams & Smith, 2007; Zenetos et al., 2012; Katsanevakis et al., 2013). Several seaweeds, such as kelp or fucoids in the Atlantic Ocean and the canopy-forming *Cystoseira* species in the Mediterranean Sea, are true ecosystem engineers or foundation species (Jones et al., 1994; Mineur et al., 2015). Consequently, changes in seaweed communities can provoke cascading effects influencing the entire ecosystem, including for example changes in abundances of herbivores and understorey coralline algae (Monteiro et al., 2009; Harley et al., 2012; Verges et al., 2014; Wernberg et al., 2016).

The ecological importance of seaweeds combined with high introduction rates of non-native species, highlights the need for methods able to accurately predict the future distribution of invasive species, preferably at the early stages of the invasion process. Species distribution modelling (SDM) links species occurrences with the environmental characteristics, and has the potential to predict distributions in a geographically explicit framework, including extrapolation in space and time. SDM can be used to identify areas with suitable habitat, assess whether introductions are likely to be successful, anticipate arrival points, and predict the extent of potential spread following an introduction. However, arrival points also heavily depend on the introduction vector (e.g. shipping, aquaculture) and the level of human activity related to these introduction vectors (Reiss et al., 2015). SDM can thus, supplemented with information on introduction vectors, help us inform decisions about preventive and control actions.

The predictive power of SDM, however, is very much dependent on the assumption that species are at equilibrium with their environment, which implies that distribution records reflect stable relationships with environment. The very nature of invasive alien species, which are possibly still in the process of range expansion in the introduced range, means that this assumption is not met for these organisms (Elith et al., 2010). Furthermore, the biotic interactions in the native and introduced environment may differ leading to changes in the geographical and environmental range (DeWalt et al., 2004; Mitchell et al., 2006). Therefore SDM of invasive or range-shifting species is particularly challenging, and requires the development of advanced modelling techniques potentially integrating mechanistic and correlative approaches (Kearney & Porter, 2009). Mechanistic approaches may model species

distributions by modelling the body temperature based on functional traits of organisms instead of using the air or sea surface temperature as indicators of environmental stress (Kearney et al., 2010; Helmuth et al., 2011) However, data availability for mechanistic models is limited while distribution data to build correlative models is more widely available (Elith et al., 2010). Improving transferability of correlative models to other time/space datasets has been accomplished by reducing overfitting and sample selection bias . Such models can be obtained by using different model choices in the background selection (Barbet-Massin et al., 2012; Martínez et al., 2015), restricting model complexity (regularization and number of variables) (Wenger & Olden, 2012), eliminating sample selection bias (Verbruggen et al., 2013; Radosavljevic & Anderson, 2014) or applying ensemble models (Hijmans & Graham, 2006; Araújo & New, 2007).

In order to accurately predict the introduced geographic range, the environmental niche of the native and introduced populations of the species should be similar (Guisan et al., 2014). While, Wasof et al. (2015) have shown that environmental niches are generally conserved between separated populations of alpine plants, for introduced seaweeds this has not been shown. We, furthermore, distinguish niche expansion or niche conservatism in analog and non-analog conditions (Guisan et al., 2014). Analog conditions are environmental conditions occurring both in the native and invaded range, while non-analog conditions are only occurring in one of the ranges. Although calculating niche change metrics in non-analog climates provides little insight in the evolution of the niche of a species, the change in niche metrics in non-analog conditions is still highly relevant for predicting the distribution of the species in invaded ranges (Petitpierre et al., 2012; Webber et al., 2012; Guisan et al., 2014).

In this study, we aim to improve predictions of invasive seaweeds in Europe and to map areas that will be disproportionally affected by changes in invasive seaweed distributions due to climate change. To this end, we first analyse the predictive performance of SDM towards the identification of suitable habitats in Europe at different stages of the invasion history based on a case study with five invasive seaweeds. Next, we calculated niche expansion and relate it to the performance of SDM. Finally, using an expanded dataset of 15 commonly recorded and widespread non-native seaweeds, we identified geographic risk areas in Europe by comparing current and future climate species distribution models.

Methods

Records collection

In order to explore the modelling of invasive seaweeds and their environmental niche in Europe we collected species records for five invasive species, for which we have ample distribution records in the introduced and native range: *Codium fragile* subsp*. fragile, Dictyota cyanoloma, Grateloupia turuturu, Sargassum muticum* and *Undaria pinnatifida*. Distribution records were classified as native or invasive, by region (Asia, Europe, America, Africa and Australia) and by year whenever possible. With respect to *C. fragile* subsp. *fragile* we decided to include species records for all subspecies as the identification of subspecies of *C. fragile* is notoriously difficult based on morphological criteria and the invasive subspecies is found in the entire range of the species (Brodie et al., 2007b; Provan et al., 2008; McDonald et al., 2015). Moreover, DNA barcodes and morphometric data indicate that *C. fragile* may actually consist of two species, the invasive subspecies *fragile* and a second species grouping all remaining subspecies (Verbruggen et al., 2016). Distribution records were collected from different data portals including the Macroalgal Herbarium Portal (macroalgae.org), Global Biodiversity Information Facility (gbif.org), Australia's Virtual Herbarium (avh.chah.org.au), Natural History Museum London (nhm.ac.uk), Muséum National d'Histoire Naturelle (mnhn.fr) with records updated until March 2016. For the last part of our study we want to uncover areas in Europe that will be affected by displacements of introduced seaweeds due to climate change. Therefore, we collected occurrences for an additional set of ten seaweeds: *Asparagopsis armata*, *Bonnemaisonia hamifera*, *Colpomenia peregrina*, *Dasya sessilis*, *Dasysiphonia japonica*, *Gracilaria vermiculophylla*, *Grateloupia subpectinata*, *Lomentaria hakodatensis*, *Polysiphonia harveyi* and *Polysiphonia morrowii*. Together with the five species from the first part they form a set of 15 representative and widely introduced seaweeds in Europe.

The quality of the distribution records was checked by geographic visualization and verification of mismatches between the location where the records were found and the coordinates recorded (Marcelino & Verbruggen, 2015). Duplicate records were eliminated as well as records located in the same grid cell of the environmental data in the same year. Records within the boundaries of the landmask were moved to the nearest ocean grid cell if located within 1,000 meters from an ocean grid cell. Records further than 1,000 meters from an ocean cell were deleted.

Environmental data

The Bio-ORACLE dataset was used as a source for environmental predictor variables. It consists of global rasters with a spatial resolution of 5 arcmin (Tyberghein et al., 2012) and it is primarily designed for global-scale niche modelling of shallow water marine organisms (Marcelino & Verbruggen, 2015). The environmental layers were retrieved using the *sdmpredictors* R package (Bosch et al., 2016).

Predictor selection is a major concern when building transferable SDMs. Many studies have addressed the consequences of variable selection (Rödder & Lötters, 2009; Verbruggen et al., 2013; Barbet-Massin & Jetz, 2014). The problem underlying this issue is the absence of causal links between predictor and response variables which may constrain the predictive power of the model (Austin, 2002; Martínez et al., 2015). In this study, the selection of variables was made a priori, taking general knowledge on the physiology and ecology of seaweeds into account (Lüning, 1990; Hurd et al., 2014). In addition variables with high correlation were not selected.

Four variables were selected a priori as potentially influencing seaweed distributions (Table 1). Sea surface temperature is suspected to be the main variable driving the distribution of seaweeds. It can affect the performance of growth, photosynthesis, reproduction and survival (Breeman, 1988; Lüning, 1990; Eggert, 2012). We used two temperature measures: maximum sea surface temperature and sea surface temperature range. Two more variables were added: mean photosynthetically active radiation and sea surface salinity. Seaweeds are photosynthetic organisms and therefore the quantity of light can affect their growth and limit habitat suitability. Salinity can influence osmotic dynamics limiting nutrient absorption and affect membrane integrity (Hurd et al., 2014), thus influencing growth, fitness and survival of seaweeds and therefore limit suitable habitats (Martins et al., 1999), as for example is the case in the Baltic Sea (Nyström Sandman et al., 2013).

Table 1. Overview of the ranges (minimum, median and maximum) of the environmental data used for modelling invasive seaweeds both for global and coastal data with the values for Europe between brackets. The different layers are maximum and range of sea surface temperature (SST), mean photosynthetically active radiation (PAR) and sea surface salinity.

	Minimum		Median		Maximum	
Layer	Global	Coastal	Global	Coastal	Global	Coastal
SST (max)	-1.5 (0.7)	-1.5 (0.6)	25.3 (19.2)	20.2 (19.9)	35.9 (32.7)	37.6 (34.4)
SST (range)	0.1 (2.5)	0.0 (1.3)	4.1 (7.0)	5.3 (12.2)	29.6 (25.1)	31.2 (29.4)
PAR (mean)	0.5 (23.7)	0.5 (8.2)	39.6 (31.3)	34.4 (32.6)	52.3 (47.1)	66.9 (55.0)
Salinity	0.0 (2.1)	0.0 (1.8)	34.7 (35.5)	33.4 (33.9)	40.7 (40.6)	41.5 (41.5)

Distribution modelling

Species distributions were modelled using four different algorithms: surface range envelope, which is equivalent to bioclim (SRE, Busby, 1991), generalized linear model (GLM), maximum entropy (MaxEnt, Phillips et al., 2004) and random forests (Breiman, 2001). For MaxEnt and GLM, we explored the complexity of the models fitted by building models with linear and quadratic features. The complexity of SRE cannot be controlled and for random forests different settings were not explored. Additionally, an ensemble model (Araújo & New, 2007) was built by averaging the results of the most transferable models. Distributions were modelled using the R (R Core Team, 2016) packages *biomod2* (Georges & Thuiller, 2013; Thuiller et al., 2016) and *dismo* (Hijmans et al., 2016).

Sample selection bias is one of the main problems impacting the transferability of models, leading to an overrepresentation of conditions in places where collecting effort is higher and thereby inflating model performance indices (Hijmans, 2012). In order to reduce sample selection bias, and therefore also environmental bias, a spatial occurrence thinning method was used (Veloz, 2009). Presence records were eliminated with the R package *spThin* (Aiello-Lammens et al., 2015) for two different thinning distances, 30 and 100 kilometres, and the results of these were compared for the different model algorithms and complexities.

Presence-only methods use species occurrence records and background points, which are selected randomly in the study area. As suggested by Phillips & Dudík (2008) we used 10000 background points in our study, which is adequate to cover the whole area. Two different sets of training background points were generated, one with points restricted to all pixels adjacent to land (coastal) and one with global background points.

In order to evaluate the different modelling options an evaluation dataset is needed (Arlot & Celisse, 2010). As no independently sampled evaluation data was available, three different ways to obtain the evaluation dataset from our dataset were explored by either splitting 'randomly', 'temporally' or 'spatially' (Roberts et al., 2016). The random splitting method consists of randomly splitting the dataset into training and testing. In the temporal approach, we used records from the earlier years to build the model and more recent occurrence records to evaluate them. Testing absence points were selected using pairwise distances such that the distance between the test occurrences and pseudo-absence points is the same as the distance between training and test occurrences (Hijmans, 2012). Finally, the spatial approach consists of dividing datasets based on geography. European

occurrence records and coastal pseudo-absence points are used to evaluate the model built with records outside of Europe. The European region was determined as the area with longitude between -34° and 65° and latitude between 29° and 73°.

Three different metrics were used to evaluate model performance: area under the receiver operating curve (AUC) (Hanley & McNeil, 1982), Cohen's kappa and the H-measure. Although AUC values have been criticized in the context of species distribution modelling (Lobo et al., 2008), its use was motivated because it is objective, threshold independent and insensitive to imbalanced datasets (Hand, 2009). Cohen´s kappa measures the agreement between predictions of the model and observations but corrects for agreement expected by chance. Kappa is sensitive to imbalanced datasets. Therefore, it has been corrected by creating a multitude of kappa values calculated from random balanced subsamples and taking the first quartile as the final kappa value. The H-measure (Hand, 2009), is similar to AUC but has as additional property that it is independent of the distribution of the empirical scores.

Starting from the model choices resulting in the most transferable and robust models, SDMs were created for all five species for different timeframes in their invasive history. Models were fitted with an increasing number of records starting with all records from the last year prior to introduction in Europe (T1). We assessed the ability of these and successive models (T2, T3 and T4), which cumulatively included more invasive records, to predict the European distribution.

This invasive history was analysed for two scenarios: a restricted and a global one. The restricted scenario consisted of a first model (T1), built with all native records and subsequent models with invasive records from Europe cumulatively added according to the invasive history (Table 2). The global scenario, on the other hand, consisted of models fitted with all available native and invasive records at the specific timeframe. Therefore, T1 models included native and invasive records from all non-European areas known at T1.

As continuous model projections are sometimes difficult to interpret, the creation of binary presence/absence maps can be a useful tool for risk assessment. We used as threshold the value that maximizes the sum of sensitivity and specificity (maxSSS) to create binary suitability maps because it is one of the best performing methods to create threshold maps when absences are not reliable (Liu et al., 2013).

Niche shifts

In order to measure the overlap in the realized niche of the species between the native and invaded range, three different indices have been calculated: niche expansion, niche stability and niche unfilling (Guisan et al., 2014). Niche expansion measures conditions in niche space not occupied in the native range. On the other hand, niche stability measures the conditions shared between both distributions. The niche stability is comparable to the niche overlap as assessed through Schoener's D or Hellinger's I. Lastly, niche unfilling measures the conditions occupied in the native range but not in the invasive range (Guisan et al., 2014). Ordination techniques, more specifically PCA-env, have been show to measure the niche overlap between two distributions better than SDM methods (Broennimann et al., 2012). The PCA-env method compares kernel smoothed species occurrence densities in an ordinated environmental space, which allows for direct comparisons of species-environment relationships in environmental space. Similar to the distribution modelling, niche measures were calculated for the restricted and global scenario and by either only taking into account analog conditions or also including non-analog conditions (Guisan et al., 2012; Webber et al., 2012). The study area used to distinguish between analog and non-analog conditions was defined as all ecoregions (*sensu* Spalding et al., 2007) wherein occurrences are located. All indices were transformed to percentages of the entire niche. These analyses were performed using the R package *ecospat* (Broennimann et al., 2016).

Risk areas

In order to identify areas at risk in Europe we modelled current and future climate distributions with the same predictors used to build transferable models. Additionally, we fitted models for two extra predictor sets by substituting maximum sea surface temperature with the mean or minimum sea surface temperature, since according to Synes and Osborn (2011) it is rarely clear which temperature variable is most applicable.

Ensemble models for these three sets of predictors were created by averaging the output of SDMs built with coastal background and all occurrence records using generalized linear models (GLM) with quadratic features, MaxEnt with quadratic features, random forests (RF) and surface range envelope model (SRE). These current climate models were subsequently projected to the three IPCC climate scenarios B1 (550 ppm stabilization), A1B (720 ppm stabilization) and A2 (>800 ppm) for the year 2100 (Jueterbock et al., 2013). The maximum sum of sensitivity and specificity (maxSSS) was used as a threshold for converting the current and future climate SDMs to binary maps. These binary maps for the 15 species were summed to

get a map of the number of species predicted in the current and future climate. The change maps were obtained by subtracting the current and future climate count maps for the three predictor sets, and subsequently calculating the mean and standard deviation of these. The mean anomaly map then reflects the change in number of species in the different areas, and the standard deviation map indicates the uncertainty of the results. Additionally, maps of the mean and standard deviation of the species turnover were calculated by counting, based on the binary maps, for each raster cell the number of species that either are predicted in the current climate and are not predicted in the future climate or vice versa.

Results

Distribution modelling

The most transferable species distribution models were obtained by creating an ensemble based on MaxEnt and GLM with quadratic features, random forests and SRE, with models fitted using coastal background and species specific spatial thinning settings. Further information on the modelling choices is provided in the Supporting information.

Based on these modelling choices we present the model performance when cumulatively more records are included in the training set representing different time points during the history of the introduction process. We repeated this process for two different setups, the first one using all non-European records (global scenario) and the second one using only native records (restricted scenario). Fig. 1 represents the performance of the model for the two different scenarios for all species, while the previously introduced Table 2 gives an overview of the number of records available for each timeframe.

For the restricted scenario, when only native records are used to build the model, AUC values are generally lower (left most values in Fig. 1A). The highest AUC is measured for *D. cyanoloma* (AUC = 0.872 with 11 records from the native range). While models for *U. pinnatifida* (AUC = 0.618) and *S. muticum* (AUC = 0.607) perform similarly with 77 and 71 records in the native range, respectively. Both *G. turuturu* (AUC = 0.536 with 43 records) and *C. fragile* (AUC = 0.485 with 44 records) have the lowest AUC. The AUC increases when occurrence records from Europe are included in the model (second, third and fourth value for each line in Fig. 1A and B), and this increase in AUC is larger at the early compared to the later phases of the introduction process.

Table 2 Data used to build the models in the global and restricted scenarios. The number of records included in the models (T1, T2, T3 and T4) is determined by the cumulative sum of the records (Timeframe). The number of records for the restricted scenario is determined by the sum of the Native and European records and for the global scenario by the sum of the Native, European and Non-European records. All records, including records after the last timeframe or without a year indication, were only used for calculating the niche metrics.

Species	Timeframe	Native	European	Non-European
Codium fragile	Before 1845 = T_1	44	0	9
	Before 1940 = T_2		41	49
	Before 1965 = T_3		169	164
	Before 1990 = T_4		471	350
	All		965	917
Dictyota cyanoloma	Before 1935 = T_1	11	0	0
	Before 2008 = T_2		5	0
	Before 2010 = T_3		15	2
	All		41	2
Grateloupia turuturu	Before 1969 = T_1	43	0	1
	Before 1985 = T_2		13	1
	Before 2000 = T_3		44	3
	All		170	36
Undaria pinnatifida	Before 1971 = T_1	77	0	0
	Before 1990 = T_2		7	1
	Before 2000 = T_3		50	4
	All		165	51
Sargassum muticum	Before 1970 = T_1	71	0	93
	Before 1975 = T_2		12	112
	Before 1985 = T_3		143	159
	Before 2000 = T_4		412	179
	All		1447	345

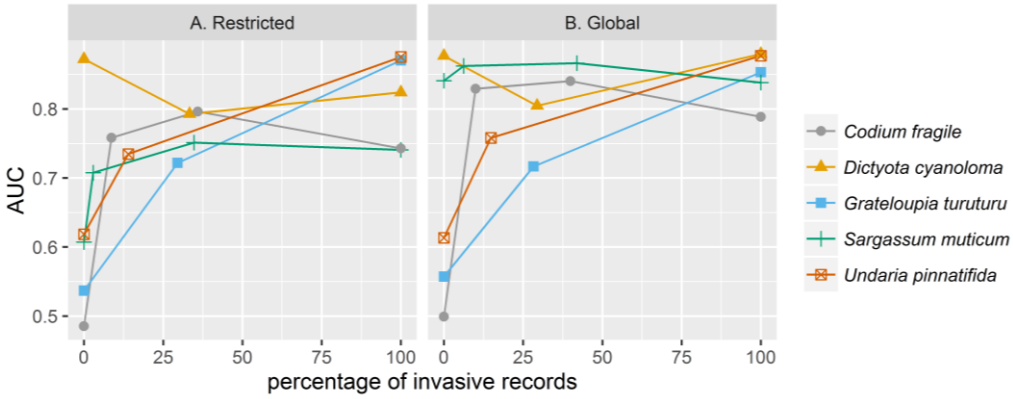

Figure 1 Evolution of the area under the curve (AUC) values at different points along the invasive history. The left figure (A) shows the AUC values for the restricted scenario, in which at T1 only native records are used for model fitting. Subsequent time points use both native and European records. For the right figure (B) all occurrence records (European and non-European) known at the specific year are used for modelling. The x-axis represents the percentage of invasive records used to build the model with the total number of records included in the last model as 100%.

Model performances of the global scenario are quite similar to those of the restricted scenario with the exceptions for *C. fragile* and *S. muticum* (Fig. 1B). Both the models for *S. muticum* and *C. fragile* have markedly higher AUC values when, next to native records, invasive records from other parts of the world known before the introduction in Europe (T1) are used to create the SDM. The other evaluation metrics (H-measure and kappa) show similar trends (Fig. S2 in Supporting information).

Fig. 2 shows maps of the model predictions of *S. muticum* for the timeframes T1 and T4 for the restricted and global scenarios. While both T1 models generally predict low habitat suitability, the threshold map of the global scenario overall reflects the present European invaded area well (Fig. 2B). However, the model failed to predict parts of the French and Catalonian coasts in the Mediterranean Sea. Model predictions from the restricted scenario (Fig. 2A and C) tend to overpredict the Mediterranean and Baltic sea and underpredict Portugal and the North of the British Isles. Maps of the other species are available in Figs. S3 to S6 in Supporting information.

Niche shifts

The niche analysis was performed in both the restricted and global scenario and niche indices were measured with or without taking into account non-analog conditions. Generally very low niche unfilling was measured with the highest value being 3% for *G. turuturu*. From Fig. 3, which reports the niche expansion, we notice that except for *D. cyanoloma* there is virtually no niche expansion between the invasive records in Europe and the non-European records, regardless of whether non-analog conditions are included in the niche expansion. When the niche of the native records is compared with the niche of the European records (restricted scenario), we see that in analog conditions there is 20 % niche expansion for *C. fragile* and about 10 % niche expansion for *S. muticum*. Niche expansion is highest when non-analog conditions are also taken into account with almost 50% for *C. fragile*, around 20% for *G. turuturu* and *S. muticum* and 10 % for *U. pinnatifida*. *D. cyanoloma*, which is not introduced in regions outside Europe, has less than 10 % niche expansion between native and invaded range when non-analog conditions are taken into account.

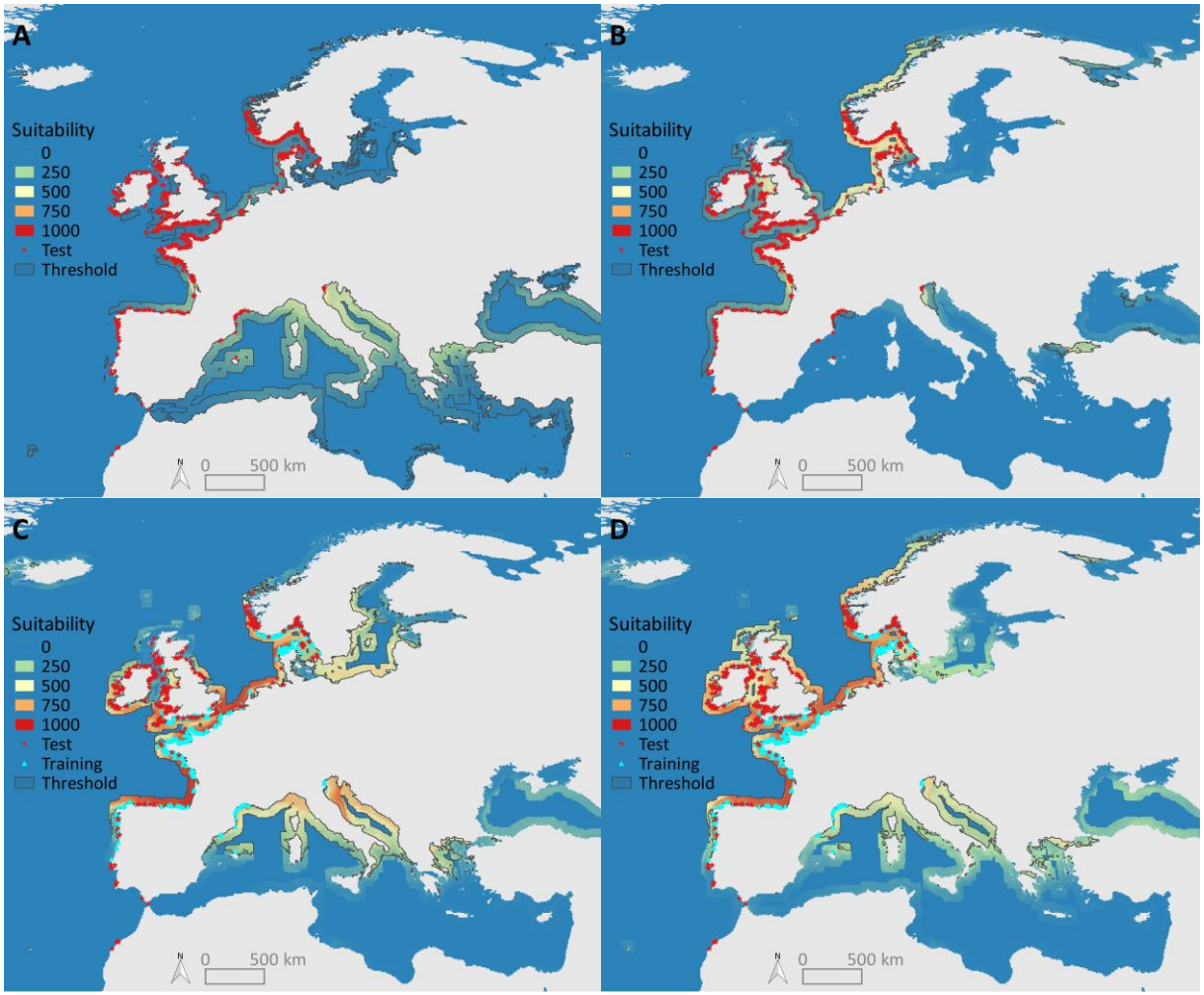

Figure 2. European predictions of suitable areas for *Sargassum muticum*. Red stars are locations used as test occurrences and the training records are in cyan triangles. The records for fitting the models are: A) only native records, B) native records and all invasive records known before the introduction of *S. muticum* in Europe, C) native records and all European records known in the year 2000 (T4) and D) native records and all invasive records known in 2000. A) and C) represent models from the restricted scenario while B) and D) are models from the global scenario. For further information about the number of records included in each model we refer to Table 2. Maps of the other species are available in Supporting information (Figs S3-S5).

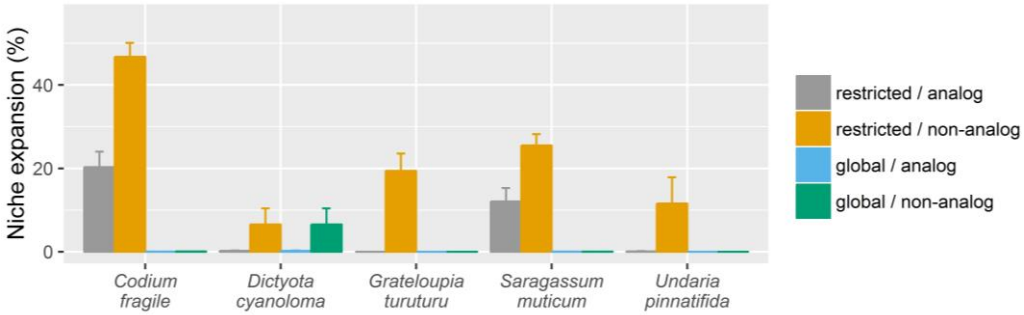

Figure 3 Niche expansion between the native and European occurrence records for four different setups. The setups have differences in the records used for comparing niches and in what is considered as niche expansion. We compared the niche expansion in Europe for the native records (restricted) and the native and non-European introduced distribution records (global). For the analog scenario niche expansion is only measured in environmental space that is available in both native and invaded area. In the non-analog scenario all niche expansion is reported. The error bars represent the standard error of using either 30 or 100 km spatial thinning.

Risk areas

Regarding the assessment of areas at risk in Europe, we see that the largest increase in number of introduced species is predicted in the northern areas of Europe, more specifically along the coasts of Iceland, Denmark and Norway by 2100 for the IPCC scenario B1 (Fig. 4A). Smaller increases in the number of introduced species are predicted for the United Kingdom, the Netherlands and Belgium. Areas with the biggest decreases, effectively becoming less suitable for the modelled list of species, are mostly located in the Mediterranean region. Additionally, some smaller spots in the Atlantic show a decrease in the number of introduced species predicted. The standard deviation map (Fig. 4B) clearly shows that some areas with larger gain also have a higher uncertainty and that the northern regions have higher standard deviations. The Pearson correlation between the absolute value of the mean and standard deviation of the change maps is only 0.55 (Table 4), which indicates a low to moderate correlation. The turnover maps for the same IPCC scenario B1 (Fig. 4C) reveal a high turnover for the regions with big increases and decreases as identified by the change maps and a number of additional areas with changes in species composition. This is most notably the case for the southern coasts of Great-Britain and Ireland. The Pearson correlation between the mean and standard deviation of the turnover maps is 0.66 (Table 4).

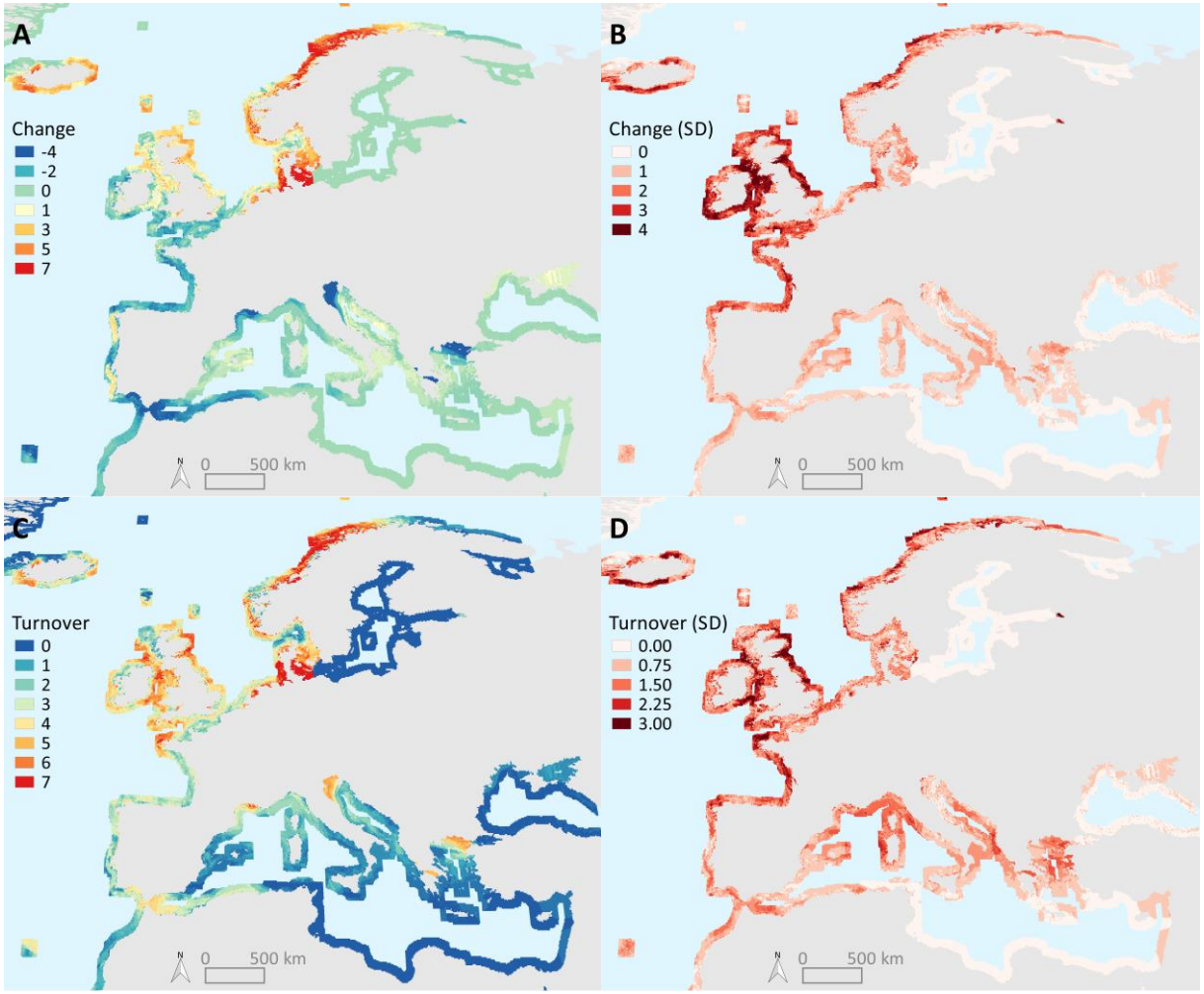

Figure 4. The change in number and turnover of introduced seaweeds predicted by 2100 under the IPCC climate change scenario B1. The top left map (A) shows the mean difference in number of introduced species between the current climate and climate change scenario B1 from SDMs build for 15 species. The top right map (B) indicates the standard deviation in model predictions when using the mean, minimum or maximum sea surface temperature as one of the four predictors for building the distribution models for each species. The bottom left (C) and right (D) maps indicate the mean and standard deviation of the turnover of introduced seaweeds. For the results based on two other IPCC scenarios, A1B (Fig. S7) and A2 (Fig. S8) we refer to Supporting information.

For the two other IPCC scenarios, A1B (Fig. S7) and A2 (Fig. S8) in Supporting information, we observe the same trends. This is confirmed by the Pearson correlation between the change and turnover maps of the different climate change scenarios which are all highly or very highly correlated, with the smallest correlation for the mean maps being 0.88 and for the standard deviation maps 0.82 (Table 4).

Table 4 Pearson correlation between the mean and standard deviation (SD) of the change (left) and turnover (right) maps for the different climate change scenarios (B1, A1B and A2). The left numbers are the correlations between the change maps, the right numbers represent the correlations between the turnover maps.

		B1		**A1B**		**A2**	
		Mean	SD	Mean	SD	Mean	SD
B1	Mean	1 / 1					
	SD	0.55 / 0.67	1 / 1				
A1B	Mean	0.90 /0.93	-	1 / 1			
	SD	-	0.85 / 0.82	0.53 / 0.66	1 / 1		
A2	Mean	0.88 / 0.91	-	0.95 / 0.97	-	1/ 1	
	SD	-	0.82 / 0.79	-	0.88 / 0.84	0.50 / 0.69	1 / 1

Discussion

Distribution modelling

Modelling the distribution of invasive species requires extrapolation to locations where the species have not previously been recorded. Therefore, general models with high transferability are needed (Randin et al., 2006). Several studies have researched methods to split occurrences into evaluation and training sets (Arlot & Celisse, 2010; Hijmans, 2012; Radosavljevic & Anderson, 2014; Roberts et al., 2016). The results of this study (Table S1 and Fig. S1) show that the random splitting method inflates the values of the evaluation metrics due to two different causes. Firstly, the environmental space distribution of the training and test occurrence records is the same. Hijmans (2012) already stated that closer testing and training points lead to artefacts in model evaluation and inflation of evaluation metrics. Secondly, background points and occurrences are more distant from each other in the environmental space leading to higher performance metrics. On the other hand, the temporal splitting approach shows low performance because background points and occurrences are distributed in the same environmental space rendering the differentiation between background points and presences difficult. With the spatial splitting approach, background points are well discriminated from occurrences but the occurrences used for testing the models and those used to build the model have a different distribution. In this context, we conclude that the spatial approach is the

most realistic cross-validation method for model selection of invasive species distribution models. This approach corroborates results by Radosavljevic & Anderson (2013), who demonstrated the power of geographic approaches to split the data to improve transferability and, therefore, to model invasive species.

Models with overly complex response curves limit the transferability of SDMs due to overfitting (Wenger & Olden, 2012; Merow et al., 2013; Verbruggen et al., 2013; Duque-Lazo et al., 2016). However, our results showed that including only very simple features results in models with a low performance as compared to using quadratic features, indicating that using only linear features results in underfitting (Hastie et al., 2009; Merow et al., 2014; Moreno-Amat et al., 2015). This can potentially be explained by the inability of the models to capture the relationship of predictors like maximum sea surface temperature (SST (max)) with the species distribution as an organism's response to temperature behaves like a quadratic curve and not a linear curve. With respect to algorithms we used GLM, MaxEnt, RF and SRE. GLM and MaxEnt generally performed well when models were tested with an independent European dataset (Table S3). However, a general trend does not exist which could be due to the variability in model performance of algorithms for different species (Elith et al., 2006; Araújo & New, 2007).In this context ensemble models prove to be a good solution to capture differences between model algorithms in a single transferable SDM (Table S3).

Background selection has a big impact on model transferability. The model performance of the coastal background is consistently higher (Table S3). The motivation for this approach was the impossibility of seaweeds, being coastal organisms, to survive in deep oceanic areas (Lüning, 1990; Marcelino & Verbruggen, 2015) and the usage of a similar approach in previous studies (Pauly et al., 2011; Martínez et al., 2015). Masking out training background data from the middle of the ocean improved model transferability when evaluated with testing records and coastal background points from Europe. Disregarding the absence of suitable substrate, the open ocean could hold environmental conditions suitable for the species. But when they are included in the background data they have to be classified as absences as species are not able to live there. By removing open ocean background data from the training set the number of false absences is thus significantly decreased, resulting in better and more transferable SDMs. Interestingly, differences between occurrence thinning parameters were species dependent. The lack of a common trend results from the idiosyncratic nature of the invasive process and recording history of the individual species. We followed the recommendations made in the literature (Phillips et al., 2009; Anderson & Raza,

2010; Barnes et al., 2014) and selected the wider thinning distances for those species with a small difference in performance for the two thinning distances since sample selection bias is one of the main drivers constraining transferability.

Threshold maps of the predictions of the European invasion allow us to visualize areas for which presences and absences were incorrectly predicted (Fig. 2). Species distributions were not predicted accurately when models were built without any invasive records from Europe. However, prediction improves rapidly when only 10 per cent of the European records were included to build the model. The addition of distribution records from other regions (global scenario case) improves the prediction and model performance mainly in *C. fragile* and *S. muticum* as these are the species with more records from other invaded regions. Especially, for *S. muticum* this resulted in a marked improvement of the T1 and later models. This implies that the inclusion of, the mostly Californian, invasive records known before the invasion in Europe add essential information about the environmental niche of *S. muticum* for modelling the European distribution. The fact that models perform generally better when including more records and when records from other invaded areas are included supports the idea that the whole environmental niche may not be recorded in the native range. Our results confirm that correlative models which aim to predict biological invasions should use all available records in order to capture the environmental niche better (Broennimann & Guisan, 2008; Verbruggen et al., 2013). But, in order to further improve the performance of species distribution models, next to extensive sampling of the native area, additional factors such as eco-physiological data and biotic interactions may need to be included.

For the other species the performance of T1 models is really low even if they are built with a relatively high number of records. For example the T1 model of *U. pinnatifida* was built with 77 native records and was barely able to predict the records in Europe. *D. cyanoloma* is a special case due to the few number of records available. The low number of records in *D. cyanoloma* models could explain the high AUC values and, therefore, it is probably heavily affected by the stochasticity of the known invasion process. Other factors contributing to the low initial performance include: differences in distributions in the environmental space between native and European occurrences, oversampling of specific native areas, lack of knowledge of the native distribution and lower competition in the invaded area.

Niche shifts

When using only analog climatic conditions, as suggested by Petitpierre et al. (2012), *C. fragile* and *S. muticum* are the species with the highest niche expansion in the

restricted scenario with both more than 10%, considered by Strubbe et al. (2013) to be a significant amount of niche expansion. The inclusion of distribution records from other invaded areas reduced the niche expansion to nearly zero. These results are very similar to previous studies of non-native plants and birds, with respectively 7 out of 50 and 8 out of 28 species displaying more than 10% niche expansion in analog conditions (Petitpierre et al., 2012; Strubbe et al., 2013). In contrast to non-native plants and birds, where niche unfilling was more prevalent than niche expansion, no niche unfilling was measured. This might potentially indicate a lack of sampling in the native range.

However, we agree with Webber et al. (2012) that studies aiming to forecast biological invasions should include non-analog conditions, as those studies are based on extrapolation in analog but also non-analog conditions. The difference of niche expansion with or without inclusion of non-analog conditions in the restricted scenario is 20% for *C. fragile* and *G. turuturu* and more than 10% for *S. muticum* and *U. pinnatifida* which could significantly constrain the prediction of introduced species. However the inclusion of records from invaded areas outside of Europe eliminated all significant niche expansion. This might explain why including other invaded records resulted in an improvement of the distribution models.

Risk areas

The increase in number of introduced species in the more northern areas of Europe is in accordance with Jueterbock et al. (2013) who predicted a northward shift for three North Atlantic seaweeds. But, the predicted risk areas are influenced by the fact that we only took into account previously known invasive seaweeds in Europe, for instance the predicted decrease in introduced species in the Mediterranean Sea was to be expected given that the rising temperatures in the Mediterranean will render it unsuitable for several of the modelled species. This doesn't necessarily imply that new species will not be introduced as the increased temperature might make it suitable for other, predominantly subtropical to tropical species that have not yet been reported in Europe. The big increase in predicted suitability for invasive seaweeds in the Northern Atlantic by 2100 might be tempered by the fact that the introduction and distribution of species depends on more factors than only the environmental suitability. Although temperature can be considered to be the main factor restricting the distribution of seaweeds (Breeman, 1988; Lüning, 1990; Eggert, 2012), it is also limited by other abiotic (bathymetry, light, substrate) and biotic factors (competition and grazing) at smaller geographic scales (Marcelino & Verbruggen, 2015).

Another inherent factor of uncertainty in the results is the fact that we used models to predict the distributions of species in the current and future climate. One of the important factors contributing to this uncertainty is the selection of predictor variables (Synes & Osborne, 2011). By calculating the mean and standard deviation of models built with the minimum, mean and maximum temperature we tried to mitigate and visualize this uncertainty. By employing spatial thinning we reduced sampling bias and thus reduced overfitting, which in turn improves the transferability of the models in space and time (Boria et al., 2014). A limiting factor of distribution modelling of introduced seaweeds for future climate predictions is the availability of future climate predictions of other abiotic factors, like pH and phosphor, that have been shown to be important predictors of seaweed distributions (Verbruggen et al., 2013).

The turnover maps show that the species composition in certain places will be altered, even when there is a limited increase in the total number of introduced species. If one or a few 'leverage species' become suitable or unsuitable this may result in sweeping community-level changes (Harley et al., 2006).

Conclusion

Distribution modelling of invasive seaweeds is a challenging task. In this study we showed that using coastal background, spatial thinning and an ensemble of models with quadratic features results in transferable distribution models. However, predicting the invasion through time can yield poorly performing models when the known distribution records don't reflect the environmental niche of the species. Change and turnover maps combined with an assessment of the uncertainty therein are valuable tools. They allow for a cost-effective monitoring of coastlines, as not all European coastlines will be evenly impacted by climate change.

Acknowledgements

The research was carried with financial support from the ERANET INVASIVES project (EU FP7 SEAS-ERA/INVASIVES SD/ER/010) and financial support provided by VLIZ as part of the Flemish contribution to the LifeWatch ESFRI. The authors would like to thank the INVASIVES consortium members for contributing data and providing suggestions; Sara Martinez and Zoë De Corte for collecting data and exploratory work.

Supporting information

In this supporting information we first aim to build transferable SDMs by selecting the best modelling choices. In order to be able to select transferable models the cross-validation (CV) splitting approaches have to be compared. Of the three cross-validation data splitting approaches (CV) the random splitting approach yielded the highest values across the evaluation metrics used: AUC, kappa and H-measure (Table S1). On the other hand, the temporal splitting approach resulted in the lowest evaluation metrics for all species, with some of the models becoming indistinguishable from random models. The random CV resulted in occurrence training and test sets with very similar densities in environmental space, that are dissimilar from the background test points (Fig. S1). The spatial CV has occurrence training and test records that are both dissimilar from each other and from the background points. Lastly, with the temporal splitting approach both the occurrence training and test sets and the background test points are very similar. From this point onward options were only evaluated using the spatial CV.

None of the two thinning distances explored (30 versus 100 km) performed better for all species (Table S2). Therefore the thinning procedure was kept species-specific. The thinning distances used were 30 km for *G. turuturu* and *S. muticum* and 100 km for the other species.

Table S3 shows that models built with coastal background have higher AUC values. They are the most transferable for all features, species and algorithms. As the Surface Range Envelope algorithm only uses occurrence data to build the model, the resulting AUC values are the same for both types of background.

Including quadratic features results in a higher transferability of the models for both MaxEnt and GLM (Table S3). Regarding the different algorithms tested, although generalized linear models consistently have a high AUC (Table S3A), other coastal background algorithms sometimes perform better than GLM depending on the species. In addition, MaxEnt coastal models with quadratic features perform somewhat similarly to coastal GLMs with quadratic features. The ensemble model built using all the algorithms, with the coastal background and quadratic features (for MaxEnt and GLM), generally performs well. The results for the other evaluations metrics show the same trends as those for AUC described here (Table S3B and C).

Table S1. Model performance for different algorithms and cross-validation (CV) data splitting approaches for all species. Values in red indicate high values while values in white indicate low values. A coastal background was used for all models and MaxEnt and GLM algorithms were performed with quadratic features. Thinning distances used for the different species were 100, 100, 30, 30 and 100 km, respectively. The three evaluation metrics used are AUC (A), kappa (B) and the H-measure (C).

A. AUC

CV	Algorithm	*D. cyanoloma*	*C. fragile*	*G. turuturu*	*S. muticum*	*U. pinnatifida*
Spatial	GLM	0.89	0.769	0.86	0.866	0.899
	MaxEnt	0.764	0.765	0.682	0.862	0.755
	RF	0.713	0.63	0.641	0.652	0.764
	SRE	0.727	0.796	0.743	0.613	0.813
Year	GLM	0.576	0.598	0.798	0.593	0.607
	MaxEnt	0.502	0.593	0.663	0.578	0.556
	RF	0.46	0.568	0.749	0.558	0.529
	SRE	0.567	0.588	0.743	0.621	0.626
Random	GLM	0.964	0.908	0.938	0.939	0.948
	MaxEnt	0.949	0.91	0.937	0.94	0.935
	RF	0.915	0.908	0.949	0.957	0.931
	SRE	0.893	0.835	0.805	0.848	0.86

B. Kappa

CV	Algorithm	*D. cyanoloma*	*C. fragile*	*G. turuturu*	*S. muticum*	*U. pinnatifida*
Spatial	GLM	0.636	0.345	0.657	0.667	0.702
	MaxEnt	0.424	0.345	0.357	0.638	0.471
	RF	0.182	-0.024	-0.086	0.039	-0.038
	SER	0.455	0.595	0.486	0.226	0.615
Year	GLM	0.067	0.088	0.378	0.151	0.058
	MaxEnt	0.067	0.076	0.324	0.196	0.084
	RF	-0.067	0.069	0.135	0.063	0.027
	SER	0.133	0.176	0.486	0.242	0.254
Random	GLM	0.857	0.675	0.78	0.725	0.824
	MaxEnt	0.786	0.675	0.805	0.725	0.765
	RF	0	0.147	0.146	0.469	0.059
	SER	0.786	0.669	0.61	0.689	0.721

C. H-measure

CV	Algorithm	*D. cyanoloma*	*C. fragile*	*G. turuturu*	*S. muticum*	*U. pinnatifida*
Europe	GLM	0.539	0.301	0.492	0.509	0.608
	MaxEnt	0.354	0.299	0.185	0.506	0.296
	RF	0.255	0.1	0.132	0.094	0.28
	SER	0.391	0.394	0.381	0.093	0.454
Year	GLM	0.155	0.084	0.39	0.097	0.108
	MaxEnt	0.128	0.079	0.232	0.105	0.109
	RF	0.12	0.036	0.288	0.032	0.04
	SER	0.024	0.056	0.281	0.092	0.111
Random	GLM	0.853	0.542	0.756	0.648	0.752
	MaxEnt	0.884	0.542	0.732	0.648	0.705
	RF	0.783	0.529	0.736	0.735	0.694
	SER	0.768	0.449	0.544	0.531	0.61

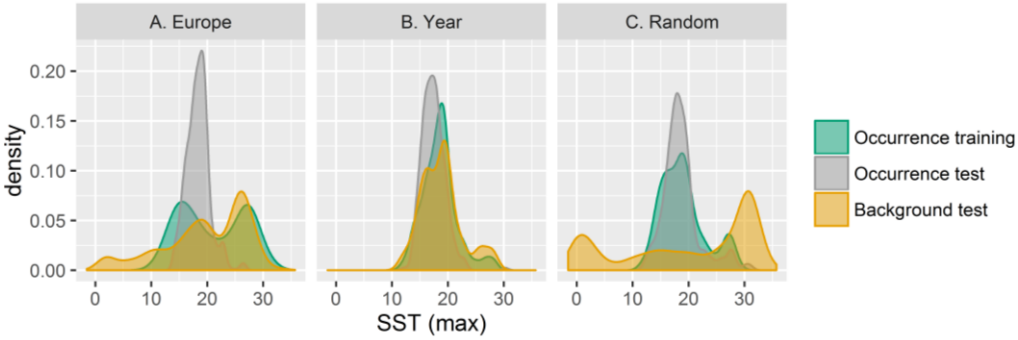

Figure S1. Distribution of occurrences (training and test) and background test for the maximum sea surface temperature for *Sargassum muticum* for the three different splitting approaches: Europe (A), year (B) and random (C).

Table S2. Model performance for the two thinning distances, 30km and 100km for the three different metrics: AUC (A), kappa (B) and H-measure (C). GLM and MaxEnt models were built with quadratic features. The best thinning distance is marked in green.

A. AUC

Algorithm	*D. cyanoloma*		*C. fragile*		*G. turuturu*		*S. muticum*		*U. pinnatifida*	
	30 km	100 km	30 km	100 km	30 km	100 km	30 km	100 km	30 km	100 km
GLM	0.881	0.895	0.766	0.769	0.86	0.665	0.866	0.688	0.907	0.899
MaxEnt	0.728	0.764	0.82	0.765	0.682	0.631	0.862	0.683	0.779	0.755
RF	0.705	0.713	0.725	0.63	0.641	0.59	0.652	0.566	0.817	0.764
SRE	0.727	0.727	0.805	0.796	0.743	0.75	0.613	0.593	0.827	0.813

B. Kappa

Algorithm	*D. cyanoloma*		*C. fragile*		*G. turuturu*		*S. muticum*		*U. pinnatifida*	
	30 km	100 km	30 km	100 km	30 km	100 km	30 km	100 km	30 km	100 km
GLM	0.636	0.636	0.278	0.345	0.657	0.443	0.667	0.211	0.692	0.702
MaxEnt	0.394	0.424	0.487	0.345	0.357	0.314	0.638	0.166	0.51	0.471
RF	0.212	0.182	0.216	-0.024	-0.086	-0.014	0.039	-0.005	0.038	-0.038
SRE	0.455	0.455	0.612	0.595	0.486	0.5	0.226	0.184	0.663	0.615

C. H-measure

Algorithm	*D. cyanoloma*		*C. fragile*		*G. turuturu*		*S. muticum*		*U. pinnatifida*	
	30 km	100 km	30 km	100 km	30 km	100 km	30 km	100 km	30 km	100 km
GLM	0.534	0.539	0.261	0.301	0.492	0.206	0.509	0.169	0.599	0.608
MaxEnt	0.312	0.354	0.395	0.299	0.185	0.158	0.506	0.168	0.324	0.296
RF	0.25	0.255	0.195	0.1	0.132	0.083	0.094	0.036	0.383	0.28
SRE	0.389	0.391	0.418	0.394	0.381	0.399	0.093	0.076	0.512	0.454

Table S3 Overview of the effect of the different modelling choices on the model performance. Columns are divided by species and background type (coastal and global background), rows represent modelling algorithms and feature types (linear or quadratic), and the values are the performance metrics area under the curve (A), kappa (B) and H-measure (C). A higher value (red) indicates a high transferability, while low values (white) indicate a poor performance. Thinning distances used for the different species were 100, 100, 30, 30 and 100 km, respectively. Ensemble models for the different species were built with the options selected with an asterisk (*).

A. AUC

	D. cyanoloma		*C. fragile*		*G. turuturu*		*S. muticum*		*U. pinnatifida*	
Algorithm	Coastal	Global	Coastal	Global	Coastal	Global	Coastal	Global	Coastal	Global
GLM Q	0.895	0.768	0.769	0.666	0.86	0.383	0.866	0.771	0.899	0.827
GLM L	0.651	0.647	0.49	0.351	0.529	0.329	0.505	0.492	0.56	0.508
MaxEnt Q	0.764	0.662	0.765	0.536	0.682	0.379	0.862	0.762	0.755	0.529
MaxEnt L	0.661	0.632	0.485	0.359	0.527	0.386	0.51	0.484	0.564	0.487
RF	0.713	0.655	0.63	0.498	0.641	0.46	0.652	0.392	0.764	0.745
SRE	0.727	0.727	0.796	0.797	0.743	0.743	0.613	0.614	0.813	0.808
Ensemble	0.894		0.831		0.802		0.843		0.885	

B. Kappa

	D. cyanoloma		*C. fragile*		*G. turuturu*		*S. muticum*		*U. pinnatifida*	
Algorithm	Coastal	Global	Coastal	Global	Coastal	Global	Coastal	Global	Coastal	Global
GLM Q	0.636	0.515	0.345	0.297	0.657	-0.071	0.667	0.286	0.702	0.538
GLM L	0.455	0.485	0.259	-0.012	0.157	-0.229	0.099	0.077	0.327	0.115
MaxEnt Q	0.424	0.394	0.345	0.136	0.357	-0.043	0.638	0.358	0.471	0.135
MaxEnt L	0.455	0.394	0.253	0.003	0.171	-0.057	0.102	0.047	0.317	0.106
RF	0.182	0.061	-0.024	0	-0.086	-0.143	0.039	-0.031	-0.038	0.317
SRE	0.455	0.455	0.595	0.59	0.486	0.486	0.226	0.226	0.615	0.625
Ensemble	0.545		0.463		0.4		0.461		0.519	

C. H-measure

	D. cyanoloma		*C. fragile*		*G. turuturu*		*S. muticum*		*U. pinnatifida*	
Algorithm	Coastal	Global	Coastal	Global	Coastal	Global	Coastal	Global	Coastal	Global
GLM Q	0.539	0.367	0.301	0.151	0.492	0.068	0.509	0.34	0.608	0.366
GLM L	0.256	0.239	0.065	0.133	0.091	0.083	0.033	0.074	0.119	0.043
MaxEnt Q	0.354	0.251	0.299	0.056	0.185	0.057	0.506	0.339	0.296	0.052
MaxEnt L	0.293	0.25	0.068	0.133	0.096	0.065	0.034	0.074	0.123	0.039
RF	0.255	0.17	0.1	0.01	0.132	0.027	0.094	0.067	0.28	0.257
SRE	0.391	0.391	0.394	0.394	0.381	0.381	0.093	0.093	0.454	0.454
Ensemble	0.562		0.451		0.414		0.478		0.526	

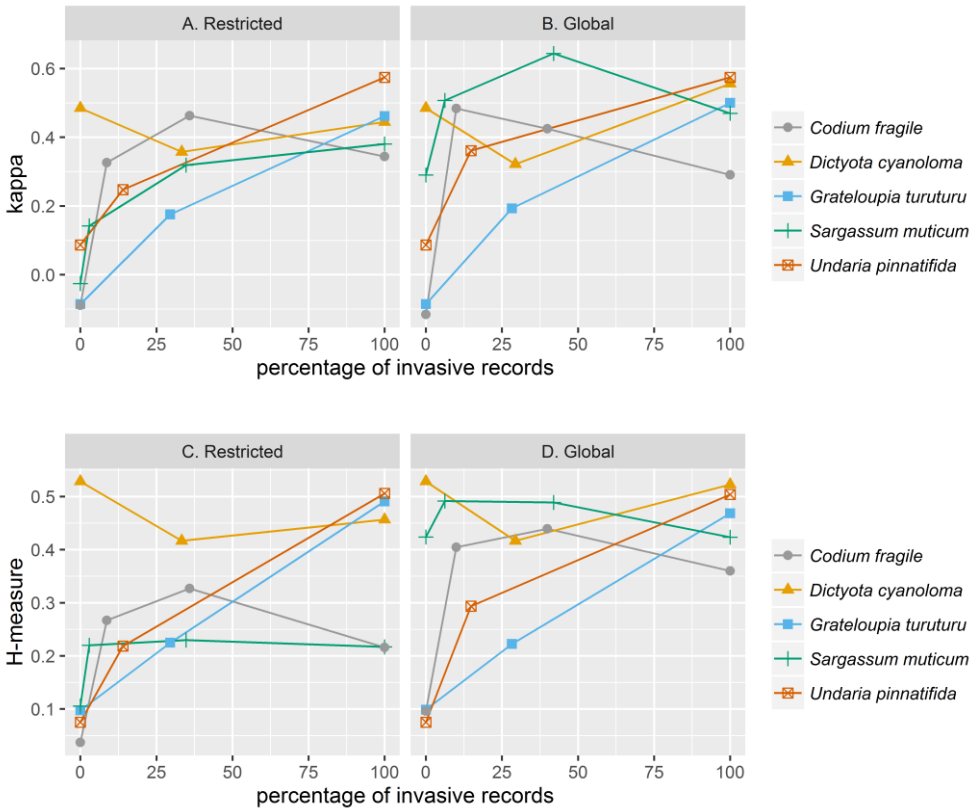

Figure S2. Evolution of kappa (A, B) and H-measure (C, D) at different points along the invasive history. The left figure (A, C) shows the values for the restricted scenario, in which at T1 only native records are used for model fitting. Subsequent time points use both native and European records. For the right figure (B, D) all occurrence records known at the specific years are used for modelling. The x-axis represents the percentage of invasive records included to build the model with the total number of records included in the last model as 100%.

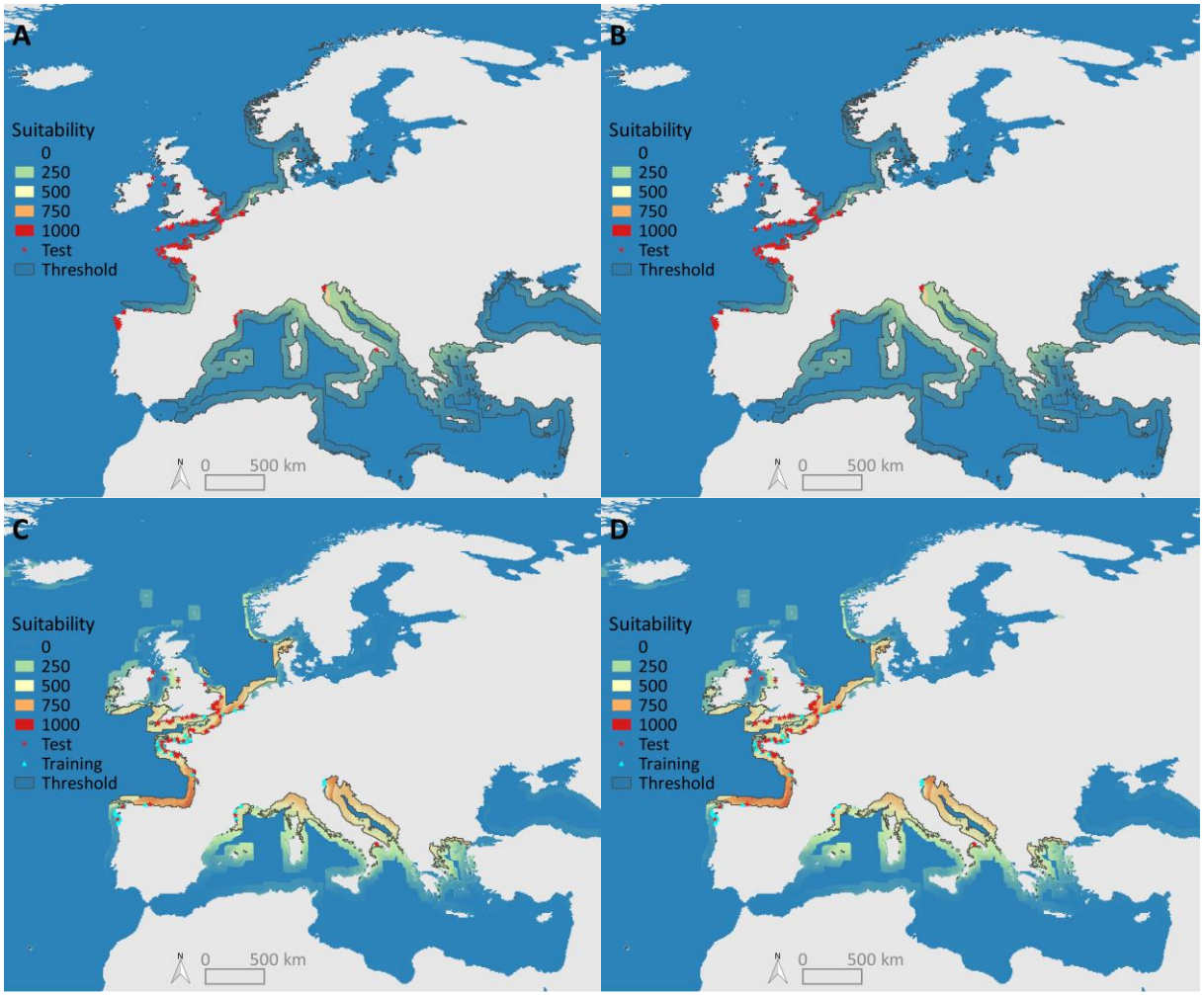

Figure S3. European predictions of suitable areas for *Grateloupia turuturu*. Red stars are locations used as test occurrences and the training records are in cyan triangles. The records for fitting the models are: A) only native records, B) native records and all invasive records known before the introduction of *G. turuturu* in Europe, C) native records and all European records known in the year 2000 (T3) and D) native records and all invasive records known in 2000. A) and C) represent models from the restricted scenario while B) and D) are models from the global scenario. For further information about the number of records included in each model we refer to Table 2.

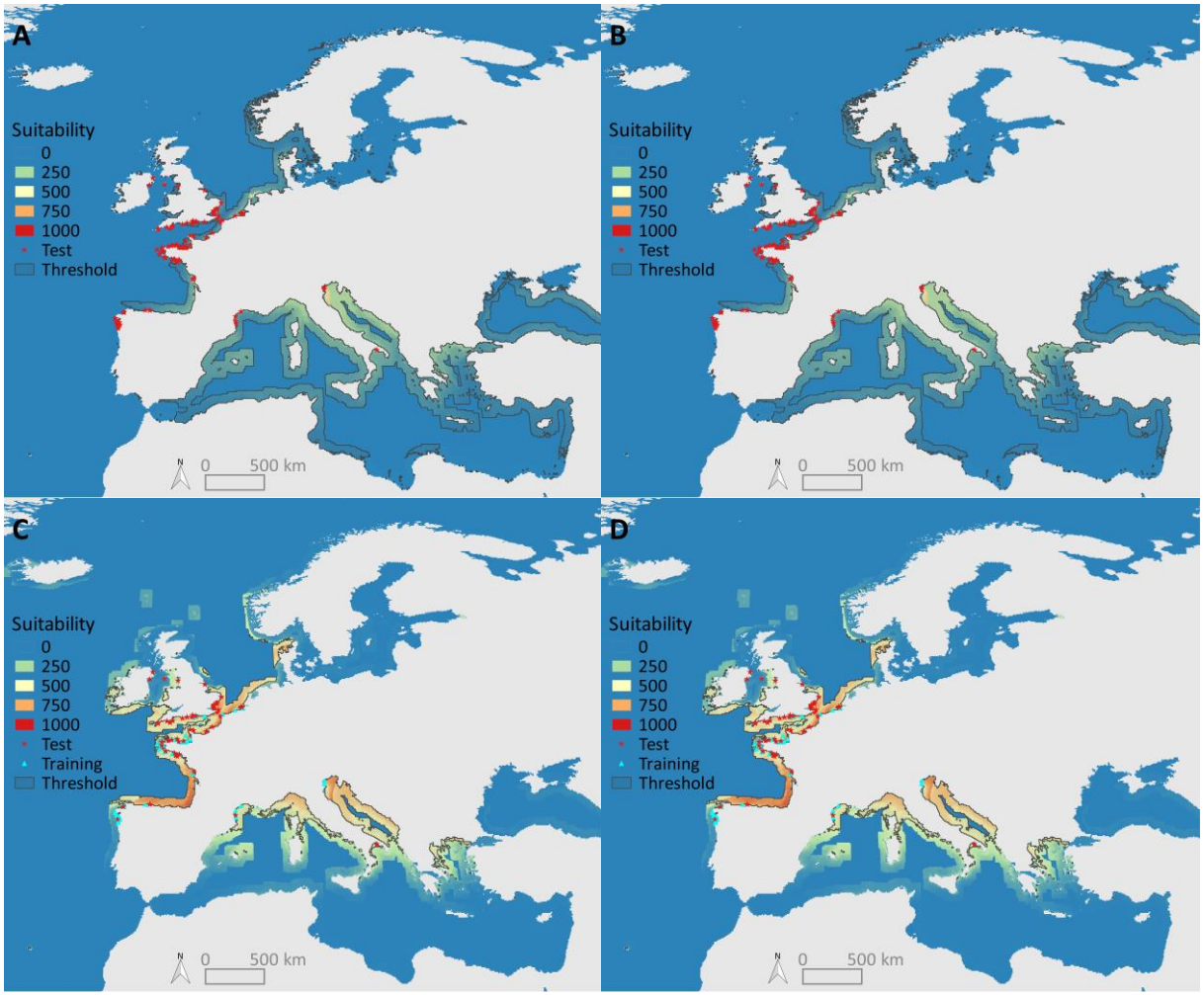

Figure S4. European predictions of suitable areas for *Undaria pinnatifida*. Red stars are locations used as test occurrences and the training records are in cyan triangles. The records for fitting the models are: A) only native records, B) native records and all invasive records known before the introduction of *U. pinnatifida* in Europe, C) native records and all European records known in the year 2000 (T3) and D) native records and all invasive records known in 2000. A) and C) represent models from the restricted scenario while B) and D) are models from the global scenario. For further information about the number of records included in each model we refer to Table 2.

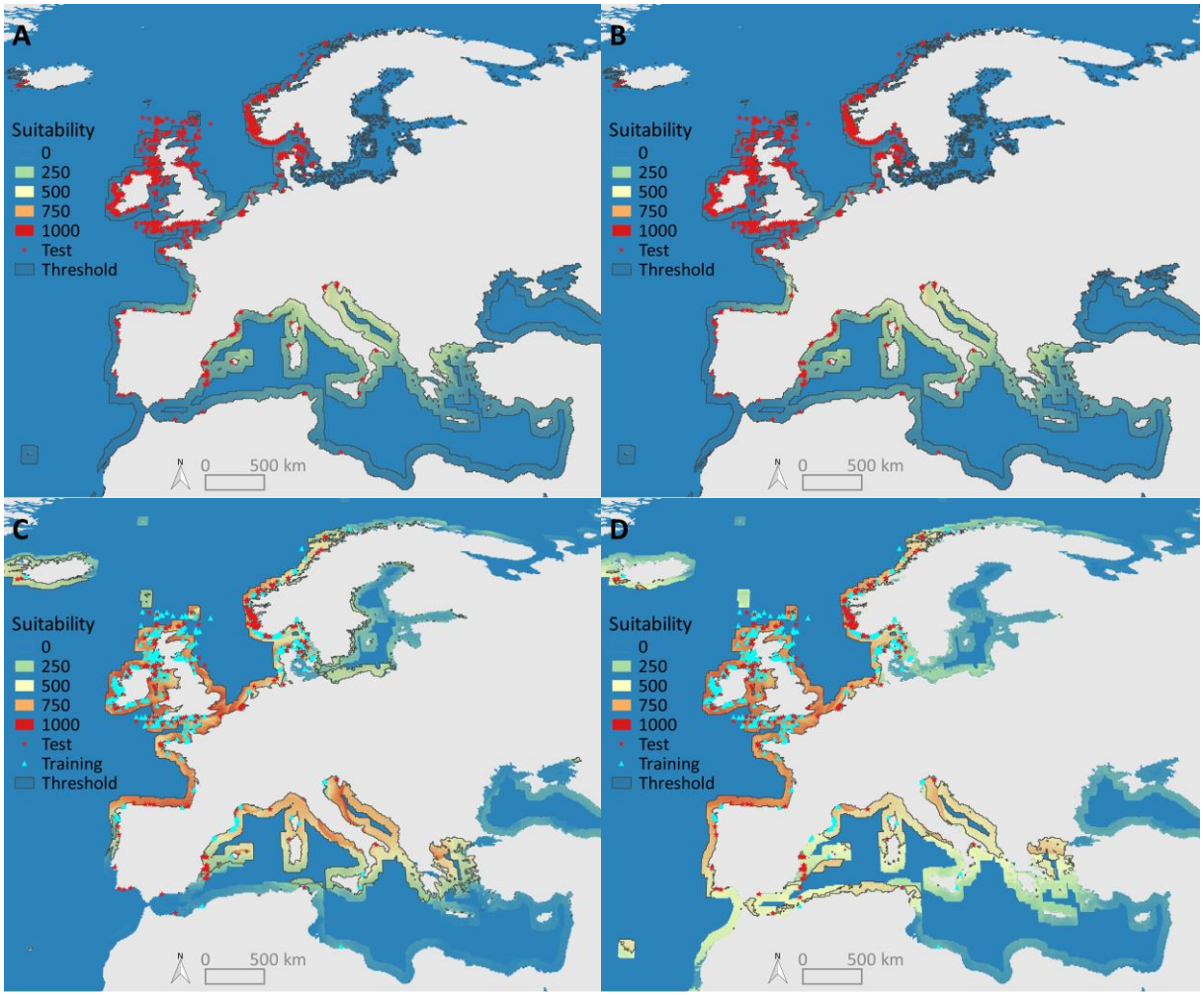

Figure S5. European predictions of suitable areas for *Codium fragile* subsp. *fragile*. Red stars are locations used as test occurrences and the training records are in cyan triangles. The records for fitting the models are: A) only native records, B) native records and all invasive records known before the introduction of *C. fragile* in Europe, C) native records and all European records known in the year 1990 (T4) and D) native records and all invasive records known in 1990. A) and C) represent models from the restricted scenario while B) and D) are models from the global scenario. For further information about the number of records included in each model we refer to Table 2.

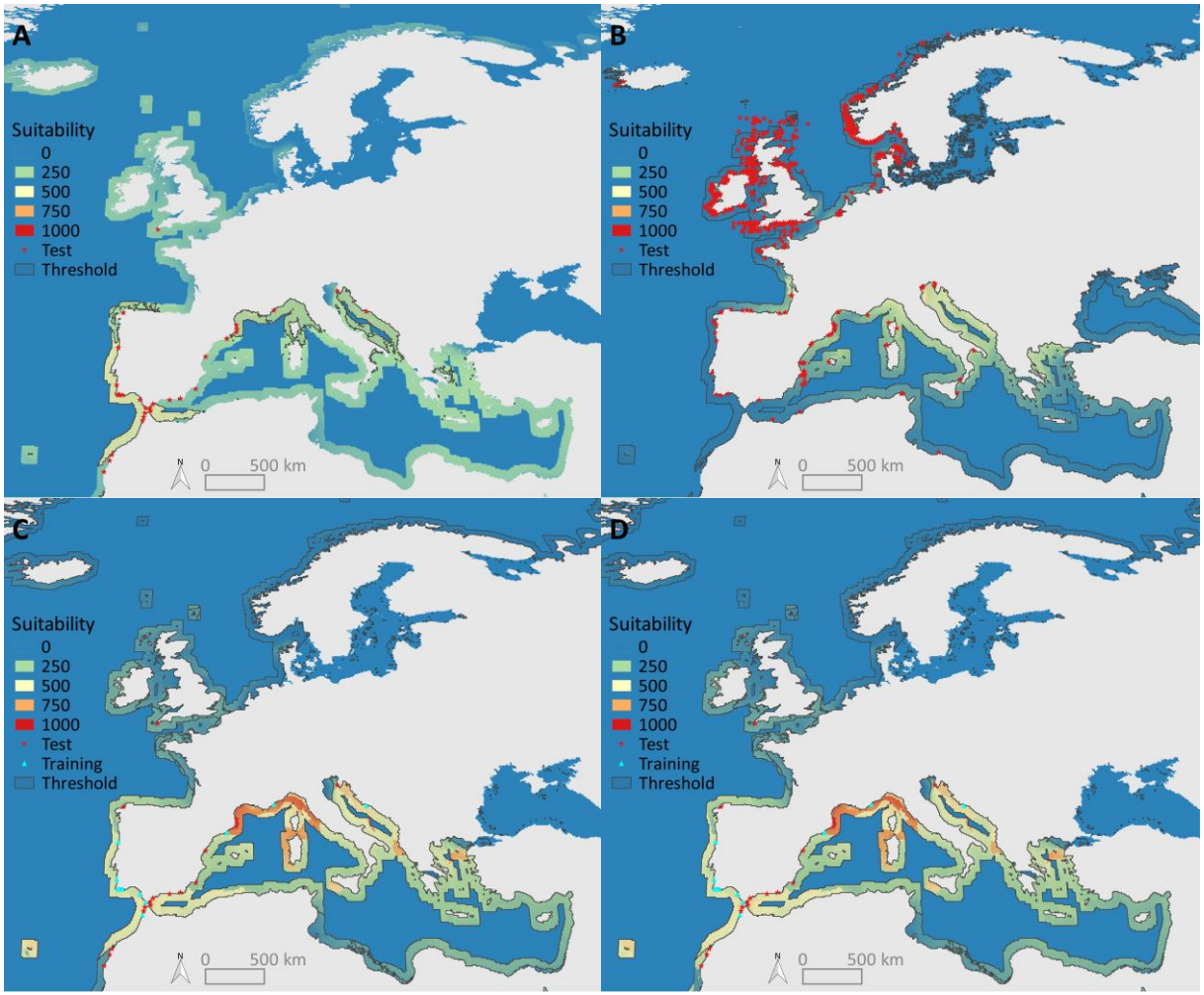

Figure S6. European predictions of suitable areas for *Dictyota cyanoloma*. Red stars are locations used as test occurrences and the training records are in cyan triangles. The records for fitting the models are: A) only native records, B) native records and all invasive records known before the introduction of *D. cyanoloma* in Europe, C) native records and all European records known in the year 2010 (T3) and D) native records and all invasive records known in 2010. A) and C) represent models from the restricted scenario while B) and D) are models from the global scenario. For further information about the number of records included in each model we refer to Table 2.

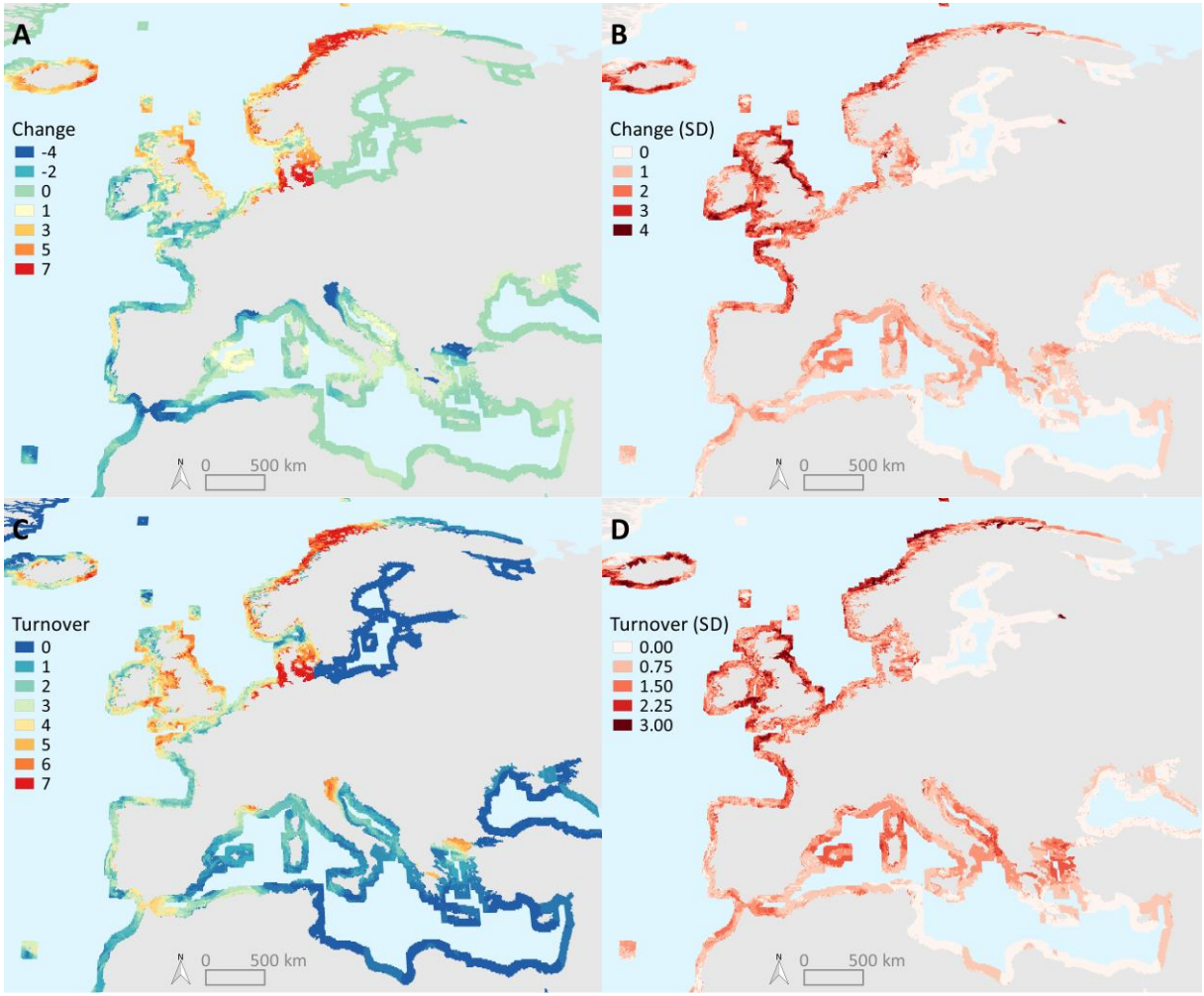

Figure S7. The change in number and turnover of introduced seaweeds predicted by 2100 under the IPCC climate change scenario A1B. The top left map (A) shows the mean difference in number of introduced species between the current climate and climate change scenario A1B from SDMs build for 15 species. The top right map (B) indicates the standard deviation in model predictions when using the mean, minimum or maximum sea surface temperature as one of the four predictors for building the distribution models for each species. The bottom left (C) and right (D) maps indicate the mean and standard deviation of the turnover of introduced seaweeds.

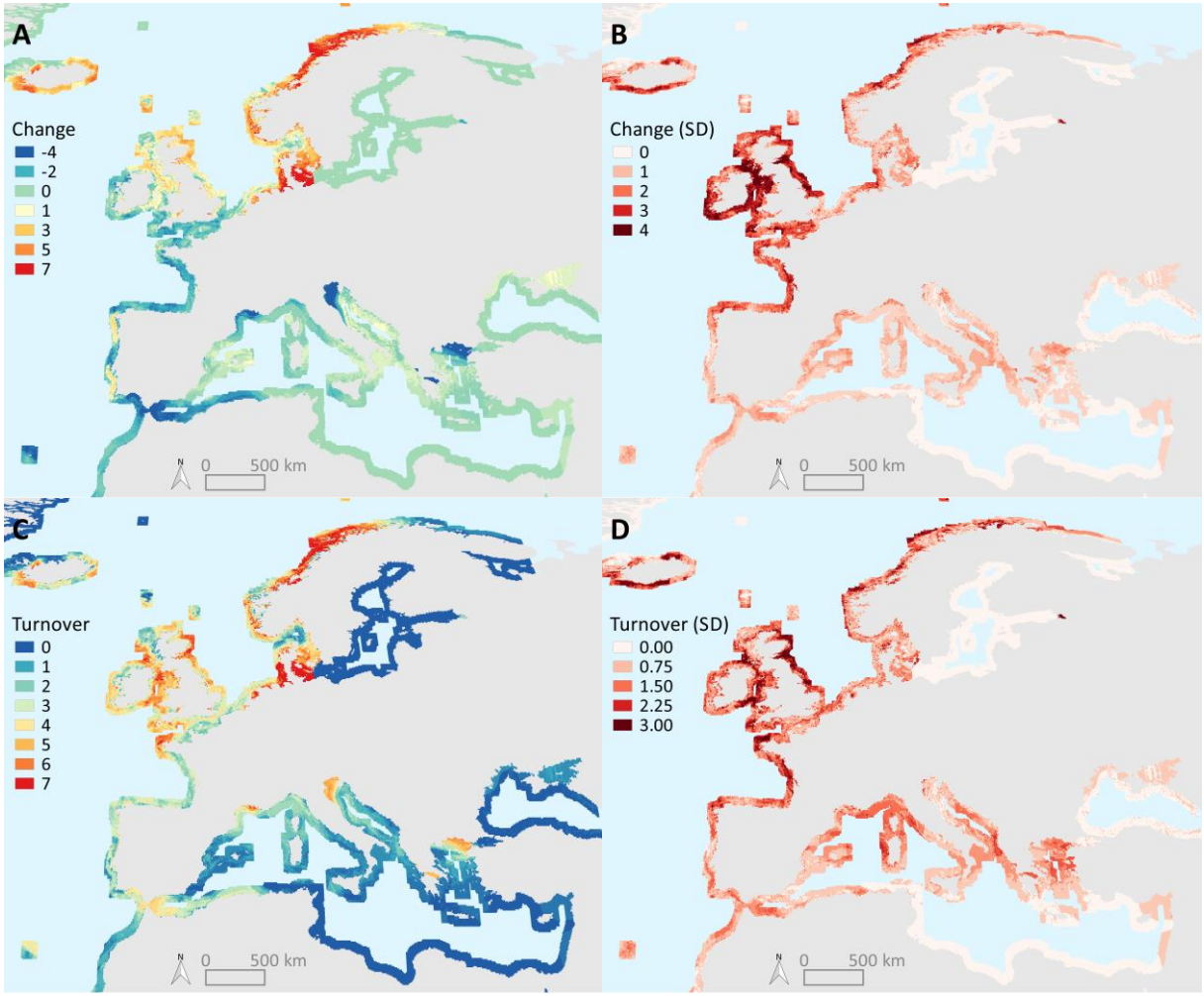

Figure S8. The change in number and turnover of introduced seaweeds predicted by 2100 under the IPCC climate change scenario A2. The top left map (A) shows the mean difference in number of introduced species between the current climate and climate change scenario A2 from SDMs build for 15 species. The top right map (B) indicates the standard deviation in model predictions when using the mean, minimum or maximum sea surface temperature as one of the four predictors for building the distribution models for each species. The bottom left (C) and right (D) maps indicate the mean and standard deviation of the turnover of introduced seaweeds.