UNIVERSIDAD DE CANTABRIA

ESCUELA DE DOCTORADO DE LA UNIVERSIDAD DE CANTABRIA
DOCTORADO EN INGENIERÍA DE COSTAS, HIDROBIOLOGÍA Y GESTIÓN DE SISTEMAS ACUÁTICOS (IH2O)

TESIS DOCTORAL

EVALUACIÓN DE LOS EFECTOS DEL CAMBIO CLIMÁTICO SOBRE LOS ECOSISTEMAS LITORALES

PhD THESIS

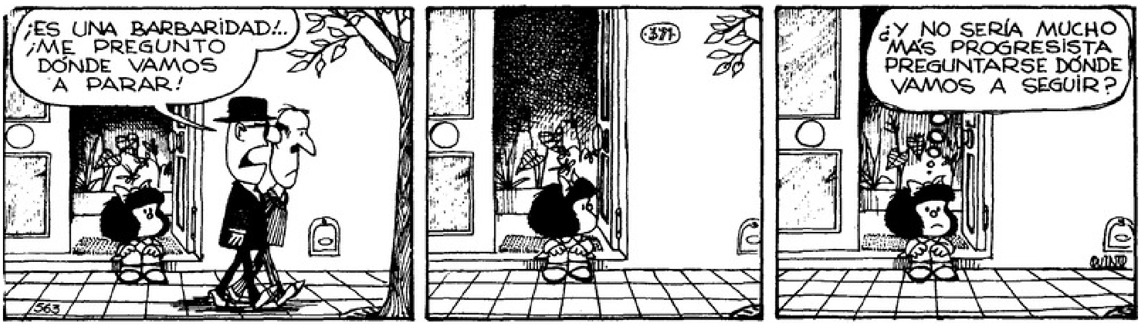
ASSESSMENT OF THE EFFECTS OF CLIMATE CHANGE ON LITTORAL ECOSYSTEMS

Presentada por: **CAMINO FERNÁNDEZ DE LA HOZ**

Dirigida por: **Prof./Dra ARACELI PUENTE TRUEBA**
Prof./Dr JOSÉ ANTONIO JUANES DE LA PEÑA

Santander, mayo de 2019

A la memoria de mi abuela


¡ES UNA BARBARIDAD!.. ¡ME PREGUNTO DÓNDE VAMOS A PARAR!
563
371
¿Y NO SERÍA MUCHO MÁS PROGRESISTA PREGUNTARSE DÓNDE VAMOS A SEGUIR?
QUINO

AGRADECIMIENTOS

En primer lugar, me gustaría agradecer a Araceli y Jose su confianza al darme la oportunidad de formar parte del IH y compartir sus conocimientos y entusiasmo por la investigación conmigo.

Esta tesis es el resultado de la interacción de varios grupos, lo que me ha dado la oportunidad de trabajar y aprender de perfiles muy diversos. A todos ellos les agradezco su tiempo y paciencia. Hago extensible el agradecimiento a todos los compañeros de IHCantabria con lo que he interactuado en estos años, por la buena disposición que he encontrado siempre. De forma muy especial quiero agradecer a Ecolit toda la ayuda prestada y el buen ambiente que generan en el día a día. Y principalmente le quiero agradecer a Elvira todo el tiempo dedicado a esta tesis y su generosidad para compartir su conocimiento y echarme una mano siempre que se lo he pedido.

I would like to thank also to the Phycology Research Group at Ghent University for hosting me during my stay, especially to Olivier and Soria for all the knowledge shared during field campaigns and laboratory work. Bedankt.

Esta tesis no habría sido posible sin el apoyo económico de la Fundación Instituto de Hidráulica Ambiental de la Universidad de Cantabria y el Ministerio de Economía y Competitividad (BES-2016-076434).

Y por supuesto, mil gracias a mi familia y amigos, a los que estaban y a los que han llegado, por los buenos ratos y el apoyo sin reservas. Y de manera muy especial estoy agradecida a mis padres, mi hermana y Javi. Imposible haber llegado hasta aquí y haber disfrutado tanto del proceso sin vuestro amor y confianza.

A todos, gracias de corazón.

Camino

Contents

Acronym list

AT	Air temperature
AUC	Area Under the receiver-operating Curve
bathy	Bathymetry
BRT	Boosted regression trees
CART	Classification and Regression Trees
CCR	Correct classification rate
CMIP5	Coupled model intercomparison project
Days	Number of days over a defined threshold
f	Fall
FDA	Flexible discriminant analysis
GAMs	Generalized additive models
GCMs	General circulation models
GLMs	Generalized linear models
Hs	Significant wave height
IPCC	Intergovernmental Panel on Climate Change
IQR	Interquartile range
Kd	Light attenuation coefficient
MARS	Multivariate adaptative regression splines
max	Maximum
Maxent	Maximum entropy
MDA	Mixture discriminant analysis
MESS	Multivariate environmental similarity surface
min	Minimum
MSE	Mean squared errors
Nit	Nitrate
OCLE	Open access database on Climate change effects on Littoral Ecosystems
P	Percentile
P1	Period 1

P1b	Period 1 backward
P2	Period 2
P2f	Period 2 forward
RCP	Representative concentration pathway
RF	Random forest
rg	Range
ROC	Receiver-operating curve
Rss	Radiation
s	Summer
Sal	Salinity
SDM	Species distribution modelling
SLR	Sea level rise
SOM	Self-organizing maps
sp	Spring
SST	Sea Surface Temperature
std	Standard deviation
sub	Substrate
SVM	Support vector machine
Tm	Bottom shear stress
TR	Tidal range
TSS	True skill statistics
Urms	Bottom orbital speed
Uw	Currents
VIF	Variable inflation factor
w	Winter
Ws	Wind speed

Resumen

Resumen

Capítulo I. Introducción y contexto de la investigación

Existen evidencias de los efectos del cambio climático en el medio marino, y con ello en las especies que alberga y los servicios ecosistémicos que proveen (Costanza *et al.*, 1997). Entre las principales amenazas del cambio climático en los sistemas marinos y costeros se encuentran el aumento del nivel del mar, el calentamiento y la acidificación del agua (IPCC, 2014).

En los productores primarios todos estos cambios interactúan produciendo efectos sinérgicos que repercuten en todo el ecosistema, dado que sustentan la estructura del mismo. Los organismos fotosintéticos, como las algas, tienen un papel clave en los ciclos de los nutrientes, la captura y transferencia de la energía y la generación de hábitats para la alimentación, puesta y refugio de numerosos organismos. De hecho, en Europa ya se están observando alteraciones como la sustitución de especies de aguas frías por especies tolerantes a temperaturas más altas (Fernández, 2016) o la disminución de biomasa y abundancia de determinadas comunidades bentónicas (Piñeiro-Corbeira *et al.*, 2016; Borja *et al.*, 2018).

Por lo tanto, resulta crucial llevara cabo un análisis del riesgo asociado al cambio climático en estas zonas donde conviven tantos elementos sociales, económicos y ambientales. Entre las metodologías de análisis de riesgos existentes, la propuesta por el IPCC (IPCC, 2014) es una de las más utilizadas, aunque su aplicación para el estudio de los ecosistemas marinos requiere una adaptación previa de los conceptos que la integran:

- Amenazas

  Se entienden las amenazas como la ocurrencia potencial de un evento que pueda causar daño o pérdida a un ecosistema. Para su correcta definición, resulta esencial establecer qué variables y parámetros determinan la distribución de la especie objeto de estudio.

- Exposición

  La exposición se relaciona con la presencia de especies que puedan verse afectadas negativamente por las amenazas.

- Vulnerabilidad

  La predisposición de las especies para ser negativamente afectadas define su vulnerabilidad. Este concepto abarca su sensibilidad o susceptibilidad, así como su capacidad para adaptarse y sobrellevar los cambios.

- Riesgo

  El riesgo se deriva de la interacción de las amenazas, la exposición y la vulnerabilidad en el medio natural.

Existen una gran cantidad de métodos para evaluar los efectos del cambio climático sobre las especies, entre los que destaca el modelado de distribución de especies (SDM). Estos constituyen una herramienta aplicada con éxito en el análisis de los efectos del cambio climático sobre los sistemas naturales (Assis et al., 2017a; Buonomo et al., 2018; Chefaoui et al., 2018; Jonsson et al., 2018). A continuación, se presenta una revisión del estado del conocimiento de cada uno de los conceptos presentados previamente.

La adquisición de información ambiental, ligada a la fisiología de la especie, a la resolución apropiada es un paso crucial en el modelado de especies. Esta ha de ser obtenida de las mejores fuentes disponibles, que incluyen observaciones directas, a través de las medidas *in situ*, lo que conlleva una falta de homogeneidad espacial y temporal; observaciones indirectas de satélites que permite solucionar las limitaciones de las medidas *in situ* al aumentar la frecuencia y el área de los datos (Álvarez-Martínez *et al.*, 2018); y el modelado numérico o reanálisis que complementa a las anteriores, al permitir reconstruir series históricas largas (Perez *et al.*, 2017). Entre las bases de datos ambientales existentes en el medio marino, Bio-Oracle es sin duda la más utilizada, ya que provee de información de cobertura global de las variables temperatura del agua, salinidad, nutrientes, clorofila, hielo, velocidad de las corrientes, fitoplancton, producción primaria, hierro y luz. Sin embargo, a pesar de los importantes avances que se han producido en los últimos años en la disponibilidad de datos, existen algunas cuestiones que necesitan un mayor desarrollo.

La disponibilidad de predictores adecuados, desde un punto de vista ecológico, espacial y temporalmente homogéneos, tanto para condiciones históricas como en escenarios de cambio climático, es todavía una necesidad en el medio marino.

Una de las limitaciones para trabajar en grandes áreas es la falta de información biológica obtenida de forma homogénea y con una resolución espacial y temporal adecuada. Al igual que en el caso de las amenazas, existen diferentes tipologías de fuentes de datos. Por un lado, las medidas directas procedentes de muestreos en campo. Idealmente, estas campañas deberían estar diseñadas sin sesgos y de manera que provean tanto presencias como ausencias (Franklin, 2009). Sin embargo, este tipo de observaciones directas son inviables en grandes áreas, por lo que en los últimos años se han generado bases de datos que permiten recopilar información de diferentes fuentes, como Global Biodiversity Information (GBIF, 2013) o Ocean Biogeographic Information Systems (OBIS, 2015). A pesar del importante avance que suponen, su cobertura espacial y temporal es limitada.

A pesar de los avances realizados en la aplicación de los SDM, para obtener resultados rigurosos es necesario solventar aspectos como i) la selección de predictores relacionados con la ecología de las especies, ii) la aplicación de algoritmos que aseguren la transferencia temporal de los resultados cuando se modelan los efectos del cambio climático y iii) la evaluación objetiva de las salidas generadas por el modelo.

La construcción de los SDM supone la asunción de algunos aspectos que es necesario tener en cuenta a la hora de interpretar los resultados (Bedia *et al.*, 2013): que todos los predictores relevantes para la distribución de las especies objeto de estudio están incluidos (Guisan *et al.*, 2017); que las observaciones biológicas son representativas de su distribución biogeográfica; que la especie está en equilibrio con el medio en el espacio y en el tiempo (Araújo & Pearson, 2005), sin estar limitada por su capacidad de dispersión (Pulliam, 2000).

La evaluación de los resultados de los modelos es crucial para determinar su capacidad para representar la realidad y la incertidumbre asociada a la extrapolación (Barry & Elith, 2006; Liu *et al.*, 2009). Esta evaluación debe llevarse a cabo tanto desde el punto de vista de la capacidad del modelo para predecir datos independientes (testing) como del realismo ecológico de los resultados.

Para concluir, la evaluación rigurosa de los efectos del cambio climático en las especies marinas necesita de la definición de predictores ecológicamente ligados a la distribución de la especie y obtenidos de fuentes validadas a las resoluciones espaciales y temporales adecuadas. Así mismo, en la selección de los algoritmos se debe asegurar su capacidad de transferencia, así como la evaluación de los resultados en el espacio geográfico y ambiental para caracterizar adecuadamente la incertidumbre del modelo.

Con base en todo lo anterior, el **objetivo general** de esta tesis es la evaluación del riesgo de las macroalgas en Europa debido al cambio climático, para lo cual es necesario el desarrollo de los siguientes objetivos específicos:

- Creación de una base de datos de variables relacionadas con la distribución de especies, con el objetivo de permitir el estudio de la variabilidad de las condiciones meteo-oceanográficas en el periodo histórico y futuro en Europa (Capítulo II).

- Selección de las especies representativas de las macroalgas, tanto intermareales como submareales, en Europa (Capítulo III).
- Desarrollo de una metodología para la selección de algoritmos que aseguren la transferencia temporal para la evaluación de los efectos del cambio climático mediante el uso de modelos de distribución de especies (Capítulo IV).
- Determinación de los cambios en los rangos de distribución de las macroalgas en Europa debido al cambio climático e identificación de las variables y parámetros más relevantes para las especies representativas (Capítulo V).

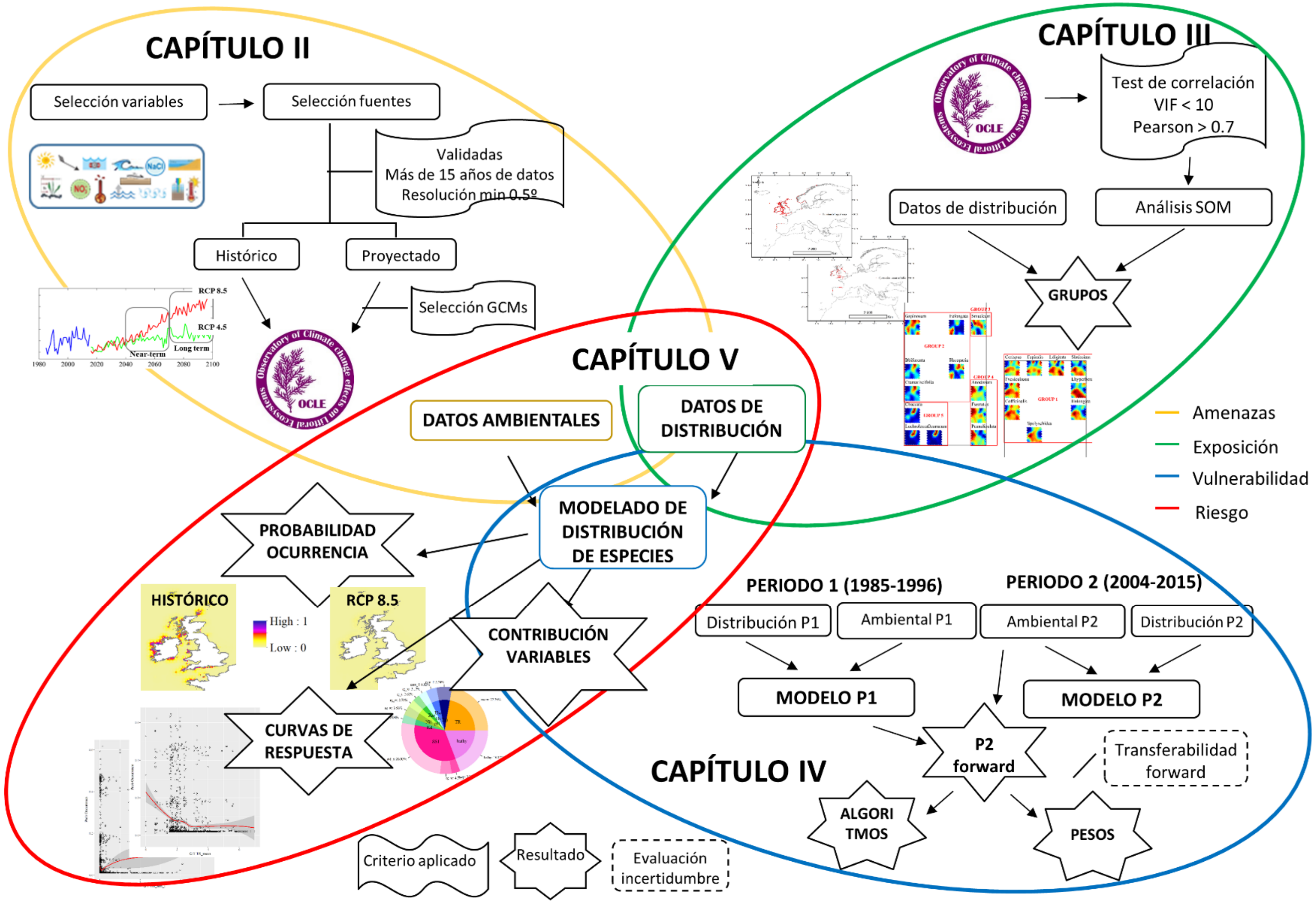


Figura 1. Resumen gráfico de la tesis.

Capítulo II. OCLE: una base de datos de libre acceso de los efectos del cambio climático en los ecosistemas litorales y oceánicos de Europa

El estudio de las distribuciones históricas y futuras de las especies marinas se ve limitado habitualmente por la falta de datos ambientales relevantes para las especies (IPCC, 2014), lo que resulta más evidente cuando estos abarcan grandes áreas (Robinson *et al.*, 2017).

Las bases de datos existentes en el medio marino (Ready *et al.*, 2010; Tyberghein *et al.*, 2012; Sbrocco & Barber, 2013; Assis *et al.*, 2017b; Vega *et al.*, 2017) han supuesto un avance muy importante para el desarrollo de estudios globales. Sin embargo, existen todavía una serie de carencias, como la ausencia de variables hidrodinámicas, su falta de homogeneidad temporal, la definición de predictores con sentido ecológico o la proyección con los mejores GCMs.

Con el objetivo de solventar estas necesidades se ha generado la base de datos OCLE, para la evaluación histórica y futura de las comunidades marinas en Europa.

La selección de las variables a incluir se ha orientado hacia los productores primarios, dado su papel crucial como base de todo el ecosistema. Para asegurar la homogeneidad temporal y espacial de los datos, aquellas variables cuya distribución era heterogénea fueron descartadas. Se han incluido variables generales de aspectos meteo-oceanográficos como temperatura (Fralick *et al.*, 1990; Valle *et al.*, 2014), luz (Best *et al.*, 2001; Larkum *et al.*, 2006; Riis *et al.*, 2012), salinidad (Touchette, 2007; Nejrup & Pedersen, 2008) y nutrientes (Hughes *et al.*, 2004; Martínez *et al.*, 2012a). Estas variables se complementaron con información relativa a las condiciones estresantes que limitan a organismos intermareales y submareales, como la desecación, caracterizada por el rango de marea (Pearson *et al.*, 2009), la velocidad del viento (Lipkin *et al.*, 1993), la altura de ola significante (Ondiviela *et al.*, 2014; Jensen & Denny, 2015; Jones *et al.*, 2015; Quintano *et al.*, 2015) y el nivel del mar (Short & Neckles, 1999). Dado el impacto que el arranque produce en las comunidades submareales, éste se ha incluido en OCLE, caracterizado por la velocidad en el fondo (Young *et al.*, 2015), la velocidad de las corrientes (Fonseca & Fisher, 1986; Infantes *et al.*, 2011) y la tensión tangencial en el fondo (Pace *et al.*, 2017).

Para cada una de estas variables, se han calculado parámetros que caracterizan, no solo sus valores medios sino también las condiciones más extremas (máxima, mínima, media, desviación estándar, rango y percentiles 10, 25, 50, 75 y 90), tanto anualmente, como estacionalmente y para todo el periodo (1985-2015). Para la temperatura del aire y del mar y la tensión tangencial en el fondo se han incorporado los umbrales existentes en la literatura para la supervivencia de los organismos bentónicos (número de días consecutivos por encima del percentil 90 de la temperatura del aire y del mar (Torresan *et al.*, 2016) y que se sobrepasa el valor de 2.2 Nt/m^2 (Vousdoukas *et al.*, 2012)). Las variables disponibles se proyectaron considerando los RCPs 4.5 y 8.5 para el medio (2040-2069) y el largo plazo (2070-2099) y se calcularon los mismos parámetros que en el periodo histórico (Figura).

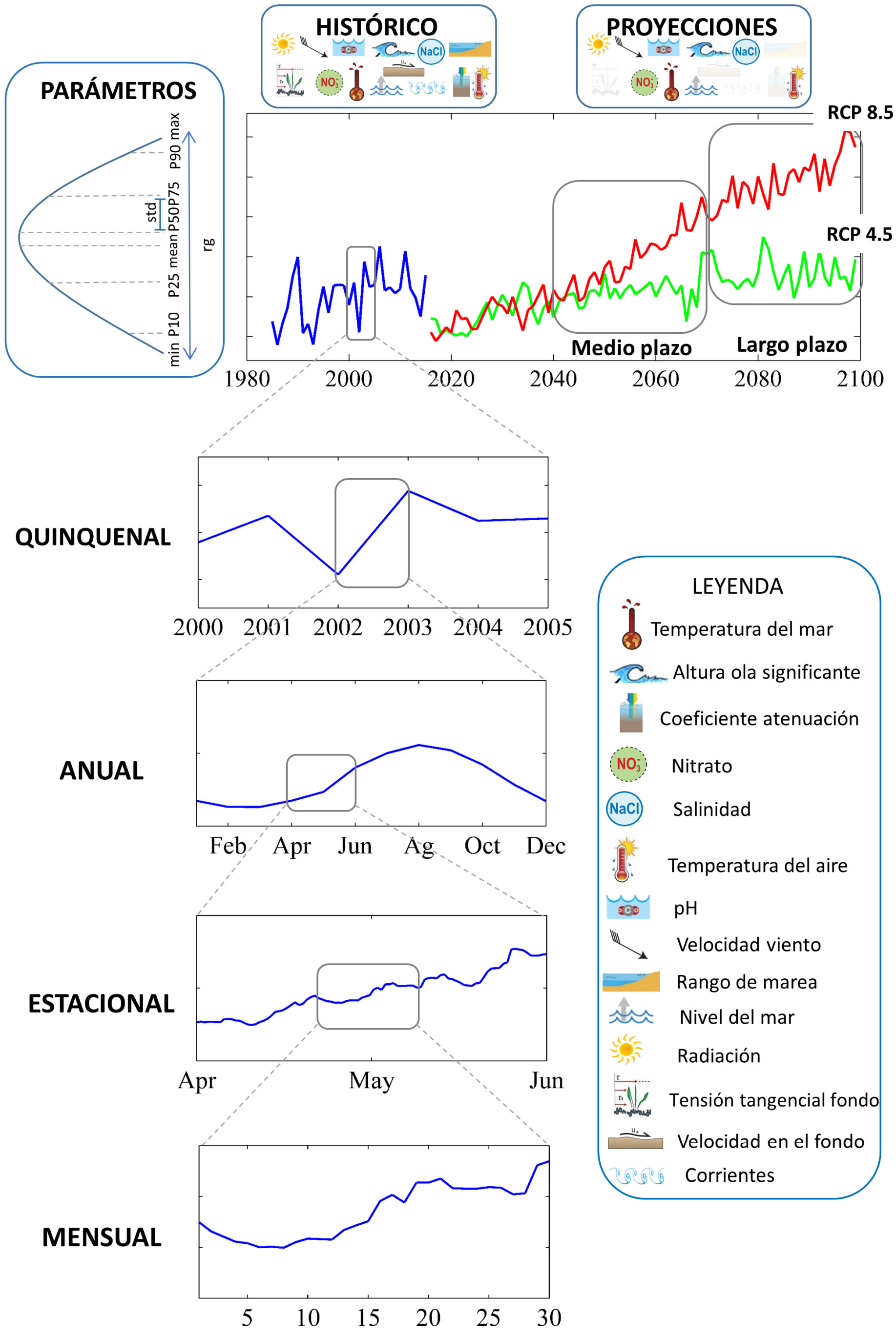


Figura 2. Parámetros calculados y su resolución temporal. Símbolos obtenidos de Integration and Application Network (ian.umces.edu/symbols).

Los datos históricos se recopilaron de satélite (Schuckmann *et al.*, 2016), reanálisis (Stark *et al.*, 2007; Saha *et al.*, 2010; Donlon *et al.*, 2012; Reguero *et al.*, 2012; Cid *et al.*, 2014; Perez *et al.*, 2017) y medidas *in situ* (Weatherall *et al.*, 2015). Para asegurar su homogeneidad espacial y temporal, solo las fuentes con series temporales superiores a 15 años y una resolución menor de 0.5º fueron utilizadas. Además, los datos finales se compararon con estudios existentes (Rhein *et al.*, 2013; EEA, 2009; Collins *et al.*, 2013).

Para los datos proyectados se utilizó el CMIP5 (Taylor *et al.*, 2012). No obstante, dada la cantidad de GCMs desarrollados por diferentes instituciones, se llevó a cabo una evaluación para determinar los más adecuados para la zona de estudio. Se calculó el MSE entre la serie de datos histórica disponible y el periodo coincidente de los GCMs (1985-2005) para cada una de las regiones definidas en la Directiva Marco de la Estrategia Marina (European Commission, 2008) para evitar sesgos debidos a procesos locales. Se definieron los outliers como aquellos GCMs con más del 20% de sus valores fuera del rango $MSE_{mean} \pm MSE_{std}$ (Chai & Draxler, 2014) (Tabla 1).

Tabla 1. Variables seleccionadas (unidades entre paréntesis), periodos, fuentes y metodologías aplicadas para su obtención. Para las proyecciones se especifican los GCMs utilizados.

VARIABLE	Periodo	Método	Fuente
Temperatura superficial del agua (ºC)	01/01/1985 – 31/12/2015	Reanálisis	OSTIA (NASA)
	01/01/2010 – 31/12/2099	Proyecciones CNRM-CM5, GFDL-ESM2G, IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Altura de ola significante (m)	01/01/1985 – 31/12/2015	Reanálisis	GOW (IH Cantabria)
	01/01/2010 – 31/12/2099	Proyecciones GFDL-ESM2G, GFDL-ESM2M, IPSL-CM5A-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Batimetría (m)	-	Medidas de satélite e *in situ*	GEBCO 2014 (BODC).
Coeficiente de atenuación de la luz (m^{-1})	25/01/1998–27/12/2015	Medidas de satélite	Copernicus Marine System (ESA)
Sustrato	-	Reanálisis y medidas *in situ*	EMODNET EUSeaMap
Nitrato (mol/m^3)	16/01/1998–16/12/2014	Reanálisis	Copernicus Marine System (ESA)
	15/01/2010–15/12/2099	Proyecciones IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Salinidad (psu)	01/01/1985 – 31/12/2015	Reanálisis	CFSR
	15/01/2010–15/12/2099	Proyecciones IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR	CMIP5

Temperatura del aire (ºC)	01/01/1985 – 31/12/2015	Reanálisis	CFSR
	01/01/2010 – 31/12/2099	Proyecciones CNRM-CM5, GFDL-ESM2G, GFDL-ESM2M, IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
pH	13/01/1985 – 8/11/2005	Reanálisis	CMIP5
	16/01/2010– 16/12/2099	Proyecciones IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR	CMIP5
Velocidad del viento (m/s)	01/01/1985 – 31/12/2015	Reanálisis	CFSR
	01/01/2010 – 31/12/2099	Proyecciones CNRM-CM5, GFDL-ESM2G, GFDL-ESM2M, IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-MR	CMIP5
Rango de marea (m)	01/01/1985 – 31/12/2013	Reanálisis	GOST (IH Cantabria)
Ascenso del nivel del mar (m)	-	-	-
	01/01/2010 – 31/12/2099	Proyecciones (IPCC 2014)	(Slangen et al., 2014)
Radiación (W/m^2)	01/01/1985 – 31/12/2005	Proyecciones GFDL-ESM2G, IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
	01/01/2010 – 31/12/2099	Proyecciones GFDL-ESM2G, IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Tensión tangencial en el fondo (N/m^2)	01/01/1985 – 31/12/2013	Reanálisis	Desarrollo propio
Velocidad en el fondo (m/s)	01/01/1985 – 31/12/2013	Reanálisis	GOW (IH Cantabria)
Corrientes (m/s)	01/01/1985 – 31/12/2013	Reanálisis	GOST (IH Cantabria)

OSTIA, Operational Sea surface Temperature and sea-Ice concentration Analysis; NASA, National Aeronautics and Space Administration; CMIP5, World Climate Research Programme; GOW, Global Ocean Wave; GEBCO, General Bathymetric Chart of the Oceans; BODC, British Oceanographic Data Centre; ESA, European Space Agency; CFSR, NCEP Climate Forecast System Reanalysis; GOST, Global Ocean Surges Tides. Todos los datos tienen una resolución especial de 0.1º y 0.5º.

Para cada una de las variables de la Tabla 1 se recopiló la información diaria, cuando fue posible, o mensual.

En el caso de la variable de preparación propia, la tensión tangencial en el fondo, su cálculo se basó en la información horaria disponible en GOW (Perez *et al.*, 2017) y GOST (Cid *et al.*, 2014), aplicando la formulación de Soulsby (Soulsby, 1997). La rugosidad en el fondo se derivó del tipo de sustrato, clasificada en cinco tipologías (fango, arena, grava, mixto y roca), de acuerdo con Soulsby (1983).

Todos los análisis se llevaron a cabo con Climate Data Operators (CDO 1.7), NetCDF Operators (NCO 4.4.5), Matlab 8.1 y ArcGis 10.1.

OCLE permite el acceso a información homogénea y rigurosa de 16 variables relevantes en la distribución de especies y 12 parámetros derivados, tanto para el estudio del periodo histórico como en escenarios de cambio climático.

La información generada está disponible para su consulta y descarga gratuita en http://ocle.ihcantabria.com/. Dado el volumen de información generada, en la página web se ha agregado quinquenalmente la información histórica y su descarga se puede filtrar espacialmente, por coordenadas geográficas o regiones de la Directiva Marco de la Estrategia Marina. La descarga se produce en formato .csv, lo que permite su utilización en modelos de distribución de especies o en sistemas de información geográfica.

El desarrollo de OCLE constituye una herramienta de gran utilidad para la investigación en el medio marino, ya que provee la más completa base de datos hasta la fecha de predictores, desagregados temporalmente, tanto para condiciones históricas como en escenarios de cambio climático en Europa.

Capítulo III. Selección de especies de algas clave en Europa

Estudiar un gran número de especies exige un esfuerzo que se puede economizar mediante el uso de técnicas de clasificación como SOM (Kohonen, 1998). Aunque su aplicación con información ambiental está ampliamente extendida (Peterson *et al.*, 2011; Galván *et al.*, 2016; Guisan *et al.*, 2017), su uso con datos biológicos es escasa, a pesar de haber demostrado buenos resultados (Ramos *et al.*, 2017).

Por lo tanto, el objetivo de este capítulo es agrupar las especies características de Europa de acuerdo con sus requerimientos ambientales para definir aquellas que se pueden considerar clave.

Las variables ambientales fueron seleccionadas de acuerdo con su relación con la ecología de las algas (Martínez *et al.*, 2012b; Young *et al.*, 2015; Cefalì *et al.*, 2016a). La información se recolectó para el periodo 1985-2015 de la base de datos OCLE (de la Hoz *et al.*, 2018b) para cada punto de la malla. Se consideraron todas las variables y parámetros disponibles (Figura 2), lo que generó un total de 642 predictores.

Para evitar multicolinelidad entre los predictores se aplicó un test de correlación con el umbral de Pearson < 0.7 y VIF < 10, mediante el paquete de R *usdm* (Naimi *et al.*, 2014). Finalmente se consideraron 32 predictores, a los que se añadieron las diferentes categorías del sustrato (fango, arena, gruesos, mixto y roca), dando un total de 36 predictores. Para cada uno de estos predictores se extrajeron los patrones ambientales con el análisis SOM (Figura 3)

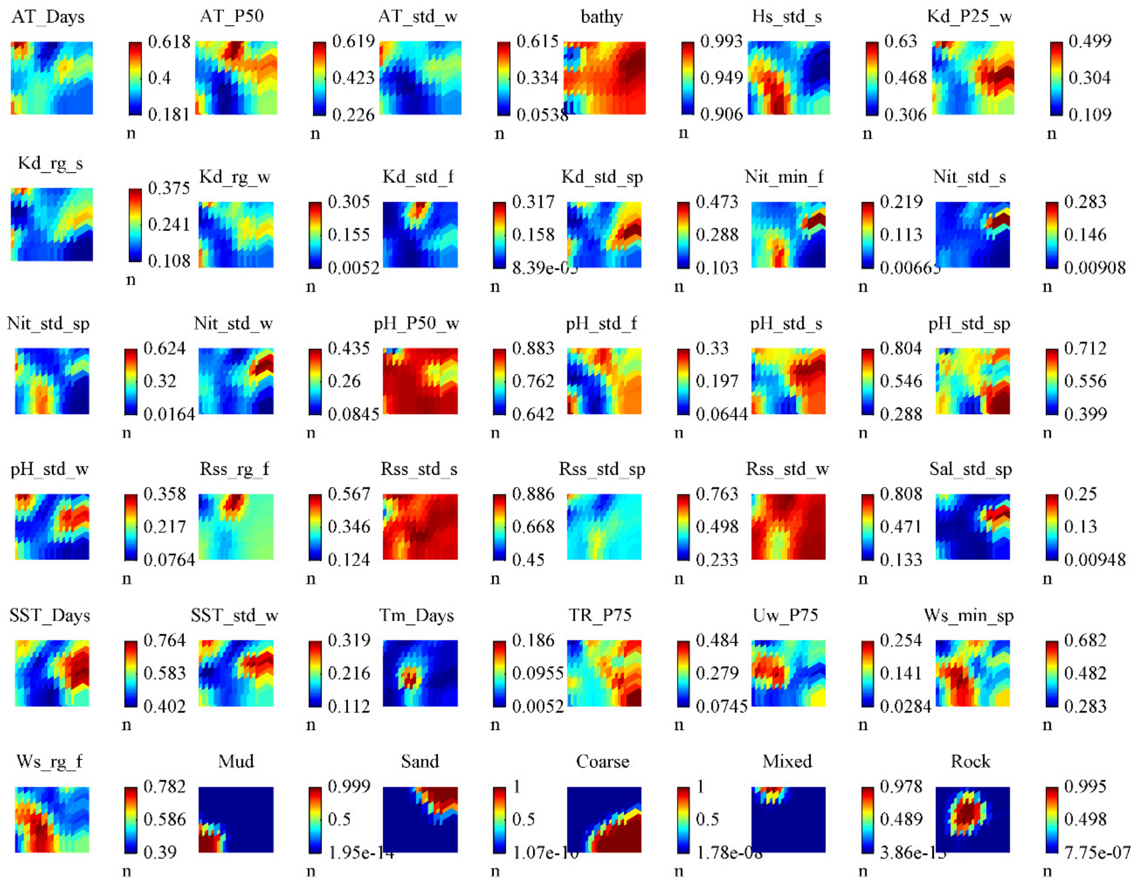


Figura 3. Visualización de los predictores en escala de colores en la SOM.

Por lo tanto, se seleccionaron 21 especies de macroalgas, tanto intermareales como submareales, de las diferentes regiones biogeográficas de Europa, incluyendo tanto especies nativas como la invasora *Sargassum muticum*. De cada una de estas especies, se extrajo la información relativa a las presencias para el mismo periodo que la información ambiental, es decir, de 1985 a 2015. Todas las observaciones fueron comprobadas para eliminar errores de referenciación y, para evitar sobreajustes, solo una presencia se utilizó para cada punto de la malla (Peterson *et al.*, 2011).

Las similitudes en los requerimientos ambientales de las especies se evaluaron proyectando sus presencias sobre la SOM previamente obtenida para los predictores (Figura), utilizando la herramienta de Matlab 8.1 component planes (Vesanto, 1999). Las especies más similares en cuanto a su respuesta ecológica aparecen más próximas, por lo que se definieron cinco grupos aplicando el criterio de experto (Figura 4).

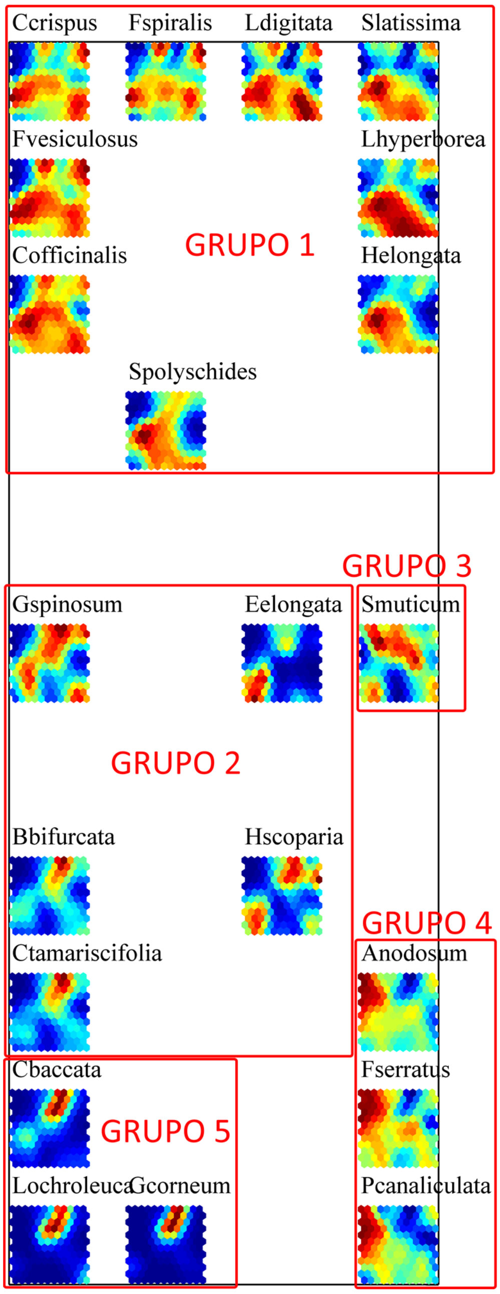


Figura 4. Component planes de las especies consideradas sobre la SOM obtenida en la Figura .

Las especies incluidas en el Grupo 1 se caracterizan por vivir en aguas frías y zonas expuestas (Norton, 1977; Sjotun & Fredriksen, 1995; Fernandez, 2011; Takolander *et al.*, 2017), por lo que son más frecuentes en las zonas en las que el número de días por encima del umbral de SST es más bajo y sin embargo para Tm los valores son mayores.

El Grupo 2 incluye especies que se sitúan desde el intermareal medio hasta el submareal y cuyos requerimientos ambientales se relacionan con condiciones propias del Mediterráneo o el Sur del Atlántico, como valores estables de SST o valores bajos de nitrato (Fredriksen & Rueness, 1989; Rueness & Fredriksen, 1989; Celis-Pla *et al.*, 2014).

El patrón que presenta *Sargassum muticum* es claramente diferente al de las demás especies, lo que justifica su aislamiento en el Grupo 3. Esto puede ser consecuencia de su carácter invasor en Europa (Stæhr *et al.*, 2000; Sánchez *et al.*, 2005), por lo que probablemente no está en equilibrio con el medio de esta zona (Araújo & Pearson, 2005).

Las especies que conforman el Grupo 4 se encuentran en zonas más altas de la costa (Anadón, 1983; Little & Kitching, 1996), por lo que su distribución está probablemente más influenciada por factores climáticos relacionados con los tiempos de emersión, como la AT o Ws (Duarte Mora, 2015).

El Grupo 5 incluye especies principalmente submareales y afines a aguas cálidas que viven en zonas rocosas y preferentemente protegidas (Salinas & Granado, 1984; Juanes & Gutiérrez, 1992; Smale *et al.*, 2015).

Para cada uno de los grupos se seleccionó una especie característica para cumplir con el criterio de tener al menos 5-10 presencias por cada predictor (Guisan & Zimmermann, 2000; Franklin, 2009; Araújo & Peterson, 2012) y poder ser utilizada en la generación de los SDM. Si ninguna de las especies cumplía con este criterio, se eligió aquella más próxima al mismo. Las especies seleccionadas fueron las siguientes:

- *Saccorhiza polyschides*
- *Gelidium spinosum*
- *Sargassum muticum*
- *Pelvetia canaliculata*
- *Cystoseira baccata*

Capítulo IV. Transferencia temporal de los modelos de distribución de especies marinas

Comprender la distribución espacial de las especies marinas resulta esencial para su conservación y gestión. Para ello el uso de los SDM es una herramienta muy útil y en auge en los últimos años que ha llevado al desarrollo de numerosos algoritmos. Esto conlleva una alta incertidumbre asociada a su elección, dado que cada algoritmo presenta una serie de fortalezas y debilidades asociadas (cf. Franklin, 2009; Peterson *et al.*, 2011; Guisan *et al.*, 2017). Es evidente que su elección es muy dependiente del tipo de pregunta que se pretende resolver, pero, así como en el medio terrestre se han desarrollado trabajos para tratar de guiar esta selección (Segurado & Araujo, 2004; Aguirre-Gutiérrez *et al.*, 2013; Raina & Rao, 2014; Beaumont *et al.*, 2016), el medio marino ha sido mucho menos estudiado (Robinson *et al.*, 2011). Con el objetivo de disminuir esta incertidumbre, el promediado de las salidas de los diferentes algoritmos (ensemble) ha demostrado mejorar los resultados (Araujo & New, 2007). Sin embargo, no existe consenso sobre cómo ponderar los algoritmos al promediarlos, por lo que se hace necesario definir un procedimiento objetivo para hacerlo (Peterson *et al.*, 2011).

Otro aspecto importante en los estudios de cambio climático es la capacidad de transferencia de los resultados a escenarios futuros, aspecto que resulta crucial en el modelado de distribución de especies y que por lo tanto debe ser evaluado.

En base a las necesidades expuestas, el objetivo de este capítulo es desarrollar una metodología, ampliamente aplicable, para seleccionar los algoritmos y su ponderación en el medio marino que asegure la transferencia temporal de los resultados en estudios de cambio climático.

La metodología propuesta se divide en tres pasos principales (Figura 5):

1) Selección de los datos y control de calidad

Se seleccionaron los predictores incluidos en OCLE en el periodo 1985-2015, dada su relación con la ecología de las cinco algas seleccionadas en el Capítulo III (Martínez *et al.*, 2012b; Young *et al.*, 2015; Cefalì *et al.*, 2016b).

Se definieron dos periodos independientes en los que evaluar la transferencia temporal de los algoritmos: 1985-1996 (P1) y 2004-2015 (P2). Se estableció un periodo de siete años entre ellos para asegurar su independencia, considerando una vida media para las algas de 3-5 años (Dixon, 1966; Chapman & Chapman, 1980; Lüning *et al.*, 1990; Little & Kitching, 1996).

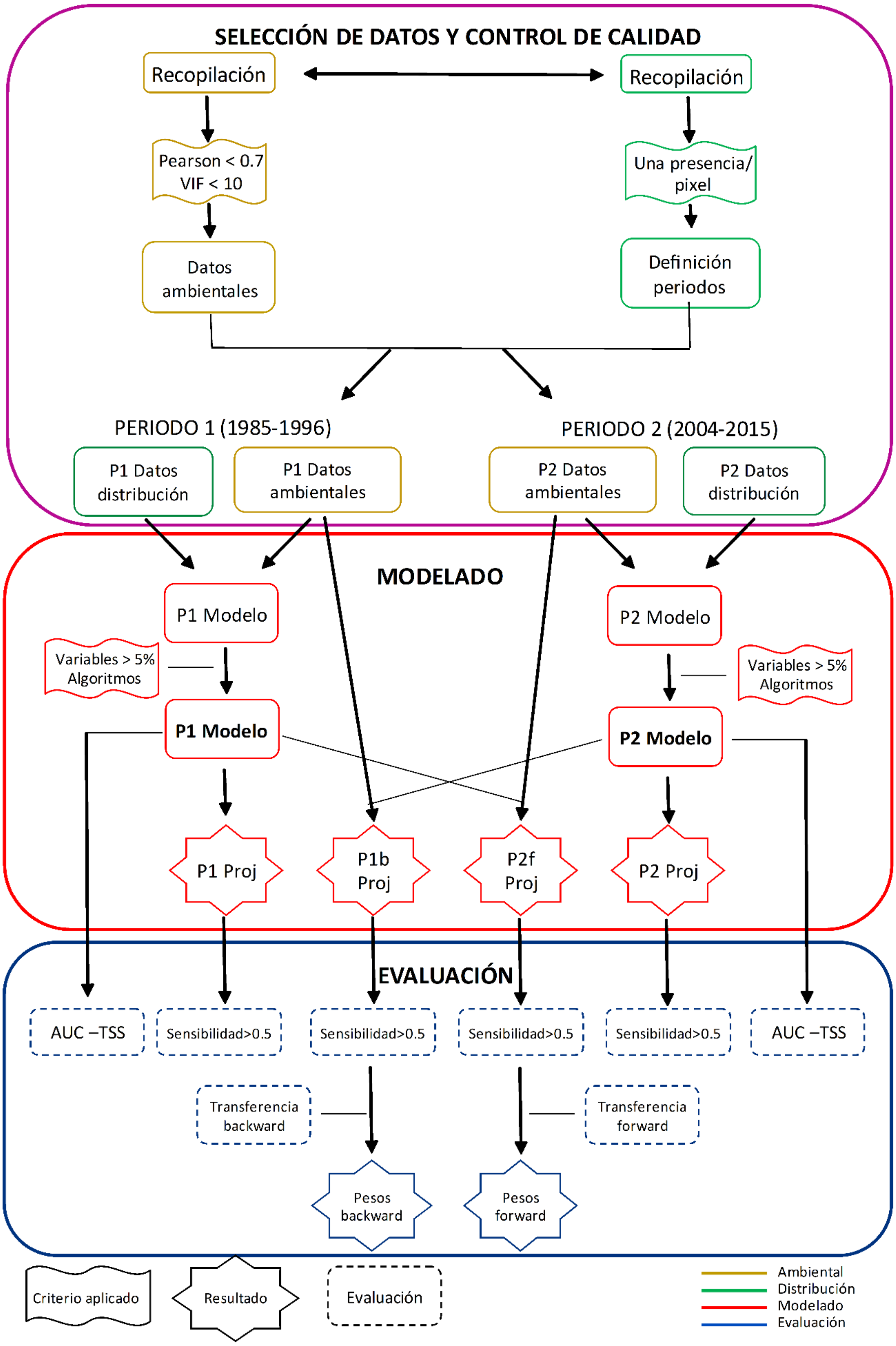


Figura 5. Diagrama de la metodología propuesta.

2) Modelado

Tanto la información biológica como la física se separó en los dos periodos definidos para cada uno de los grupos representados por una especie. Se utilizó un 70% de las presencias de cada grupo para entrenar el modelo y el 30% restante se utilizó para testar los resultados. Se aplicaron nueve algoritmos de modelado para reflejar las diferentes aproximaciones existentes: análisis de regresión (GLM, GAM, MARS y MDA), técnicas de clasificación (CART, SVM, FDA) y de aprendizaje (RF, Maxent). Dado que solo se disponía de presencias y algunos de los modelos necesitan de ausencias, se generaron 1000 pseudoausencias de forma aleatoria en el espacio geográfico, de forma que no hubiera solapamiento con las presencias (Phillips & Dudík, 2008; Guisan *et al.*, 2017; Iturbide *et al.*, 2018). Para cada uno de los grupos se evaluó la colinealidad de los predictores, con los mismos criterios establecidos para la agrupación de las especies (coeficiente de Pearson < 0.7 y VIF < 10).

Los modelos se ejecutaron de forma independiente para cada uno de los periodos con el paquete *sdm* (Naimi & Araujo, 2016) y se proyectó al otro periodo. Por lo tanto, se obtuvieron proyecciones forward (P2f: del P1 al P2) y backward (P1b: del P2 al P1).

Se calculó la sensibilidad de cada algoritmo en cada proyección como el porcentaje de presencias correctamente predichas (por encima del 50% de la máxima probabilidad) (Thuiller *et al.*, 2005).

Se ejecutaron los modelos nuevamente con los algoritmos con sensibilidad por encima del umbral y con los predictores con una contribución superior al 5%, con los mismos puntos de entrenamiento y testado que anteriormente (Tabla 2).

Tabla 2. Algoritmos y predictores utilizados para construir los modelos finales de cada grupo.

	GRUPO 1	**GRUPO 2**	**GRUPO 3**	**GRUPO 4**	**GRUPO 5**
	RF, CART	GLM, GAM, RF, MAXENT, MDA	RF, MAXENT	GLM, RF, MAXENT, MDA	RF, MAXENT
PREDICTORES P1	SST_std_s TR_rg_f bathy SST_rg_sp Ws_min_sp Uw_min_sp AT_Days Rss_rg_f Tm_Days Rss_rg_s Ws_std_f Hs_max_f Rss_rg_sp sub Ws_min_w SST_rg_w Sal_rg_f	SST_std_s SST_max_w TR_mean bathy SST_rg_sp Ws_min_w Uw_P75 Sal_rg_f sub Ws_min_sp Rss_rg_sp Rss_rg_w Hs_max_f SST_Days AT_Days SST_std_w Tm_max_f Rss_std	Sal_rg_f Urms_mean TR_rg_f sub Ws_min_s bathy SST_rg_w SST_Days Ws_max_s Hs_std_f Tm_max_f	bathy SST_max_sp SST_P50_f sub Sal_rg_w Rss_std TR_rg_f Ws_min_s SST_std_f Hs_P50 Uw_P75 Uw_max_w AT_rg_w Rss_mean_s Sal_max	Urms_max TR_max SST_rg_w bathy Tm_mean SST_rg_sp SST_Days sub AT_Days Tm_Days
PREDICTORES P2	SST_rg_s TR_mean bathy Sal_max Uw_min_f Ws_min_w Tm_Days Rss_rg_w Tm_min_s Rss_rg_sp AT_rg_s sub	Rss_rg_w Sal_rg_f Ws_min_sp bathy TR_max sub Tm_mean_s AT_rg_s Ws_min_w Urms_max Tm_max_f SST_Days Rss_rg_s Hs_std_s	TR_mean SST_max_sp Ws_rg_s Ws_min_sp bathy Urms_min_w Sal_rg_f AT_P10_s Rss_rg_f SST_Days	bathy sub SST_P75_sp Ws_min_s Uw_P75 AT_max_s AT_rg_w Sal_rg_w Rss_std TR_max Rss_max_s SST_rg	TR_max Rss_rg_sp AT_Days AT_rg_s Sal_max Urms_max bathy AT_mean Tm_max_f Hs_max Sal_rg_f sub AT_max

3) Evaluación

Se aplicaron diferentes métricas para evaluar tanto la capacidad predictiva de los modelos, como su transferencia y los pesos de cada algoritmo. Para todos los grupos los valores de AUC y TSS, ejecutados con el paquete *pROC* (Robin *et al.*, 2011), se encontraron por encima de 0.8 (Tabla 3), lo que pude considerarse buena capacidad predictiva (Swets, 1988; Araujo *et al.*, 2005; Engler *et al.*, 2011; Gallien *et al.*, 2012)

Tabla 3. Media y desviación estándar de AUC and TSS pesados para cada grupo en el periodo P2f.

	GRUPO 1	**GRUPO 2**	**GRUPO 3**	**GRUPO 4**	**GRUPO 5**
AUC	0.87±0.25	0.92±0.07	0.92±0.01	0.97±0.07	0.91±0.01
TSS	0.80±0.39	0.90±0.02	0.92±0.07	0.93±0.05	0.96±0

Una vez ejecutados los modelos finales, la sensibilidad en P2f fue calculada para obtener los pesos de cada algoritmo para cada grupo (Tabla 4).

Tabla 4. Pesos asignados a cada algoritmo en cada grupo.

GRUPO	**MODELO**	**PESO**
1	RF	70.91
	CART	29.09
2	GLM	26.04
	GAM	21.88
	RF	21.88
	MAXENT	22.92
	MDA	7.29
3	RF	44.83
	MAXENT	55.17
4	GLM	7
	RF	25.5
	MAXENT	40
	MDA	27.5
5	RF	85.72
	MAXENT	14.29

La transferencia temporal de los modelos se evaluó con las ecuaciones de Tuanmu (Tuanmu *et al.*, 2011), con valores por encima de 0.8 para todos los grupos (Tabla 5), lo que se considera una capacidad de transferencia buena.

Tabla 5. Transferencia forward, backward y total para cada grupo.

TRANSFERENCIA	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4	GRUPO 5
Forward	1	0.88	0.86	0.98	0.84
Backward	0.98	0.92	0.96	0.92	0.86
Total	0.97	0.87	0.83	0.90	0.83

De acuerdo con las métricas de evaluación consideradas y la inspección visual de los resultados, se puede considerar que los modelos desarrollados tienen una buena capacidad predictiva y son ecológicamente realistas.

En cuanto a la selección de algoritmos, de acuerdo con otros autores (Elith *et al.*, 2006; Franklin, 2009; Thibaud *et al.*, 2014) los algoritmos de aprendizaje generan modelos con mayor sensibilidad que las técnicas basadas en regresiones, probablemente debido a que gestionan de manera más eficiente las relaciones complejas entre las especies y su medio (Thibaud *et al.*, 2014).

Los resultados de la transferencia de los modelos muestran la especificidad por especies, como ya habían señalado otros autores (Randin *et al.*, 2006). A pesar de los buenos resultados obtenidos en la evaluación de la transferencia, es necesario tener en cuenta que la extrapolación a condiciones ambientales nuevas implica una alta incertidumbre (Sequeira *et al.*, 2018), por lo que es necesario evaluar la similaridad entre los dos medios.

En lo que a las relaciones entre las especies y su medio se refiere, los resultados sustentan la relevancia de las variables hidrodinámicas en la distribución de las algas, como se ha señalado por otros autores (Ramos *et al.*, 2014; de la Hoz *et al.*, 2018a; Jonsson *et al.*, 2018). Además, ha quedado patente el importante rol de las condiciones extremas en la supervivencia de las especies, lo que apoya el uso de predictores relacionados con las condiciones estresantes del medio (Galván *et al.*, 2016; Torresan *et al.*, 2016; Assis *et al.*, 2017a; Vranken *et al.*, 2018), a pesar de que su uso en el modelado de especies no está generalizado.

La metodología desarrollada ha demostrado dar buenos resultados con las especies de algas consideradas; sin embargo, se considera que se puede aplicar a otras zonas geográficas o con otras especies de manera igualmente efectiva y eficiente. Además, los pesos definidos para el promedio de los algoritmos permiten acotar la incertidumbre asociada a la selección de algoritmos, mejorando la rigurosidad de los resultados en estudios de cambio climático.

Capítulo V. Efectos del cambio climático en la distribución de las algas en Europa

En este capítulo se utilizaron SDM para evaluar los efectos del cambio climático en Europa sobre las cinco especies definidas como representativas en el Capítulo III. Para ello se ha aplicado la metodología desarrollada en el Capítulo IV, incluyendo la proyección de las variables en escenarios de cambio climático (Figura 6).

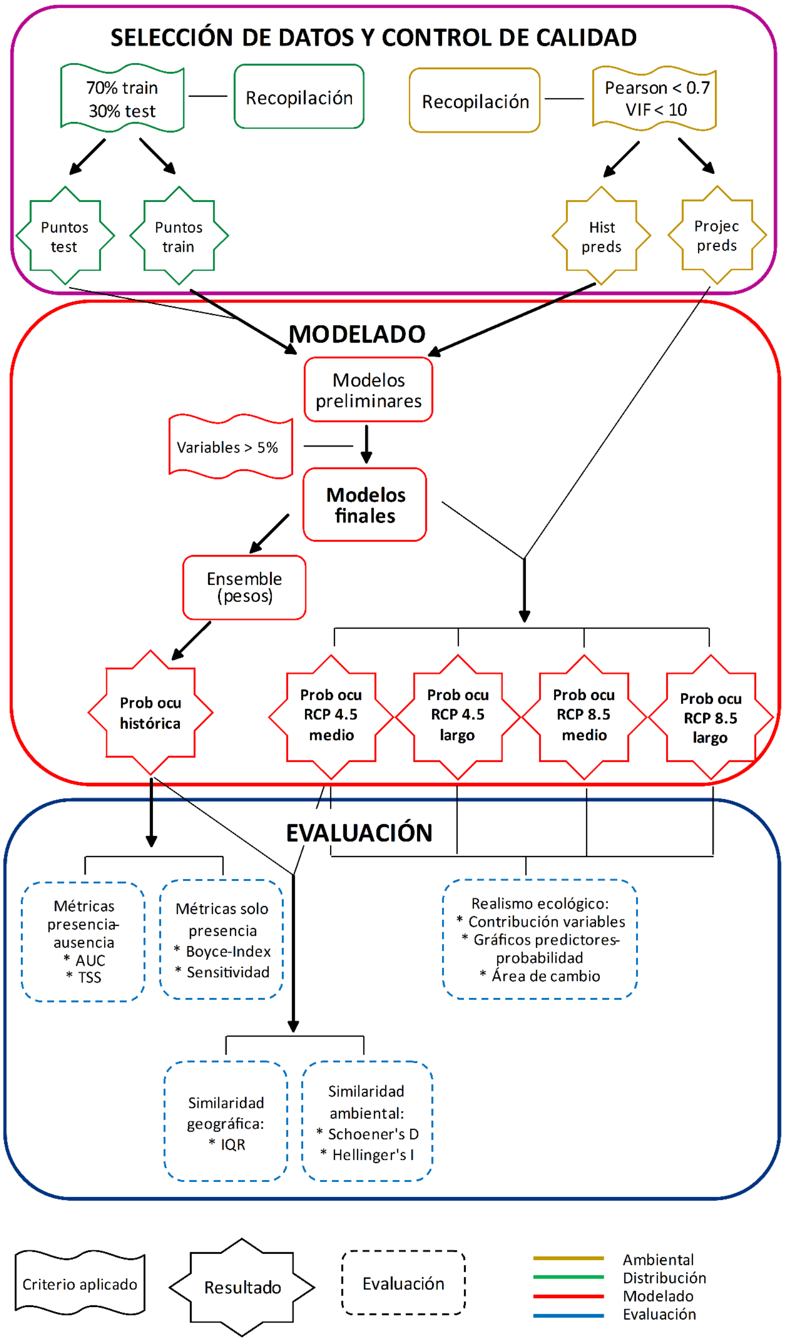


Figura 6. Diagrama de la metodología propuesta.

Las presencias de las especies se recopilaron en el periodo 1985-2015 y se filtraron para considerar solamente una por punto de la malla. El 70% se utilizaron para entrenar el modelo y el 30% restante para validarlo. La falta de ausencias se suplió con la generación de 1000 pseudoausencias aleatorias, sin solapar con las presencias (Barbet-Massin *et al.*, 2012; Guisan *et al.*, 2017).Por otro lado, se obtuvieron los predictores de OCLE (de la Hoz *et al.*, 2018b) para el mismo periodo, cuya correlación se evaluó con los mismos criterios que en los Capítulos III y IV: coeficiente de Pearson < 0.7 y VIF < 10, con el paquete de R *usdm* (Naimi *et al.*, 2014).Para cada grupo se ejecutaron aquellos algoritmos que demostraron ser más adecuados para asegurar la transferencia temporal de los resultados en el Capítulo IV, promediados de acuerdo con los pesos obtenidos (Tabla 4).

La capacidad predictiva de los modelos se evaluó tanto en el espacio ambiental como geográfico. En primer lugar, se utilizaron métricas para comparar las probabilidades de las presencias predichas con las ausencias observadas, mediante el AUC y el TSS. Dado que los datos biológicos solo disponían de presencias, se aplicaron también índices basados en ellas, el índice de Boyce, el CCR y la sensibilidad. De forma complementaria se evaluó la superposición de nichos con la D de Schoener (Schoener, 1989) y la I modificada de Hellinger (Warren *et al.*, 2008), con el paquete de R *ecospat* (Broennimann *et al.*, 2016). Por otro lado, la similaridad geográfica se evaluó mediante el IQR (Acuna & Rodriguez, 2004) para detectar aquellas áreas donde el medio no es análogo, y por lo tanto las proyecciones han de interpretarse con cautela. Se consideraron outliers las zonas con valores inferiores a Q1-1.5(IQR) y mayores de Q3+1.5(IQR). Para cada una de las especies consideradas se obtuvo la contribución de cada predictor promediando los resultados de los algoritmos considerados, de acuerdo con sus pesos. Finalmente, los modelos finales de cada especie se aplicaron por separado para cada predictor para obtener sus umbrales fisiológicos.

Como resultado más importante, cabe destacar que la bondad de los modelos finales y su capacidad de discriminación puede considerarse muy buena de acuerdo con las métricas aplicadas, con valores de AUC y TSS por encima de 0.93 y 0.84, respectivamente (Tabla 6). El índice de Boyce indica predicciones muy buenas, con valores por encima de 0.91 en todos los casos, y con sensibilidades buenas para todas las especies, con más del 60% de las presencias correctamente predichas, excepto en el caso de *P. canaliculata* en el que el alto valor de CCR (0.74) ha limitado estos valores. Además, la pequeña diferencia entre los valores de entrenamiento y testado de AUC (AUC_{diff}) indica que no se ha producido un sobreajuste de los modelos.

Tabla 6. Métricas de evaluación para cada especie.

	S. polyschides	***G. spinosum***	***S. muticum***	***P. canaliculata***	***C. baccata***
AUC	0.93	0.98	0.99	0.97	0.98
TSS	0.84	0.93	0.93	0.86	0.94
ÍNDICE BOYCE	0.99	0.94	0.99	0.99	0.91
CCR	0.43	0.35	0.11	0.74	0.31
SENSIBILIDAD	73.30	61.67	100	11.44	69.23
AUC_{diff}	0.039	0.007	0.004	0.0002	0.019

La comparación de los nichos ambientales mostró que las condiciones en el RCP 8.5 eran más diferentes de las históricas que las de RCP 4.5 y por el largo plazo más que por el medio plazo (Tabla).

Tabla 7. Valores de solapamiento de nichos

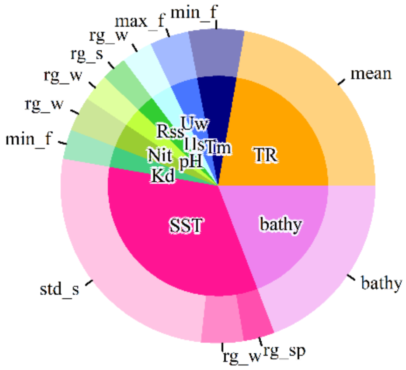
	RCP 45m	**RCP 45l**	**RCP 85m**	**RCP 85l**
Schoener's D	0.45	0.40	0.39	0.29
Hellinger’s I	0.71	0.63	0.61	0.56

En cuanto a la contribución de los predictores a los modelos, destaca la importancia de las variables hidrodinámicas y los parámetros indicativos de variabilidad y condiciones extremas (Figura 7).

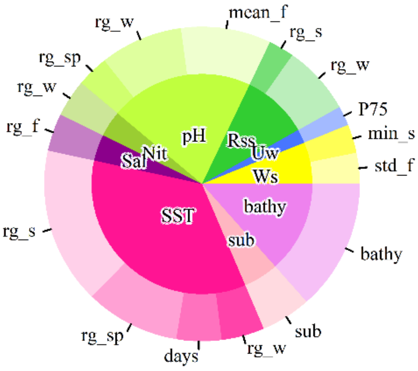
A la vista de las diferencias en la probabilidad de presencia de las especies, destaca la disminución predicha para *S. polyschides* y *G. spinosum* en su área de distribución (Figura 8 y Figura 9). En el caso de la invasora *S. muticum*, se observa una disminución importante de su probabilidad, prácticamente desapareciendo en la Península Ibérica; sin embargo, su probabilidad aumenta en las áreas más al norte (Figura 10). La probabilidad de presencia de *P. canaliculata* en la costa oeste de Gran Bretaña podría disminuir ligeramente, mientras que en las zonas próximas a la Bretaña y el Canal de la Mancha aumentan (Figura 11). Las diferencias en la probabilidad de presencia predicha para *C. baccata* se incrementa en los RCPs considerados en el entorno de Gran Bretaña y el Golfo de Vizcaya (Figura 12).

Las predicciones se pueden considerar rigurosas, puesto que, de acuerdo con el criterio establecido, solo algunos puntos aislados constituyen outliers.

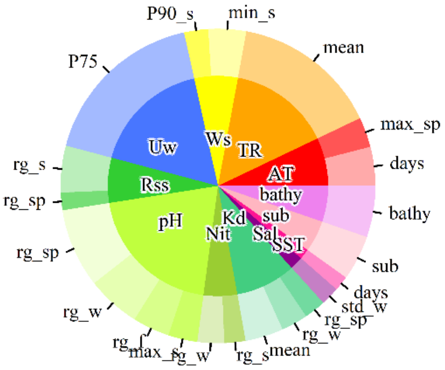
Saccorhiza polyschides



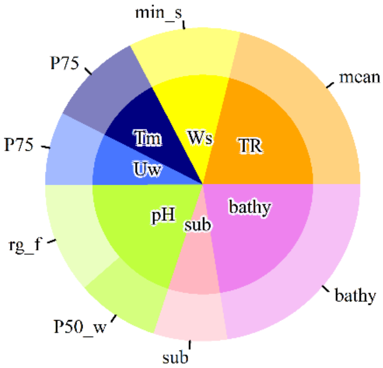
Gelidium spinosum



Sargassum muticum



Pelvetia canaliculata



Cystoseira baccata

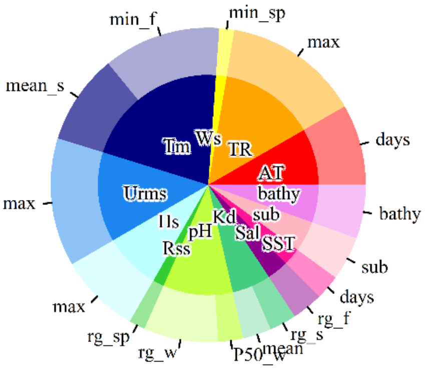


Figura 7. Contribución de cada variable a los modelos finales. Los círculos internos indican las variables y las subdivisiones exteriores sus parámetros.

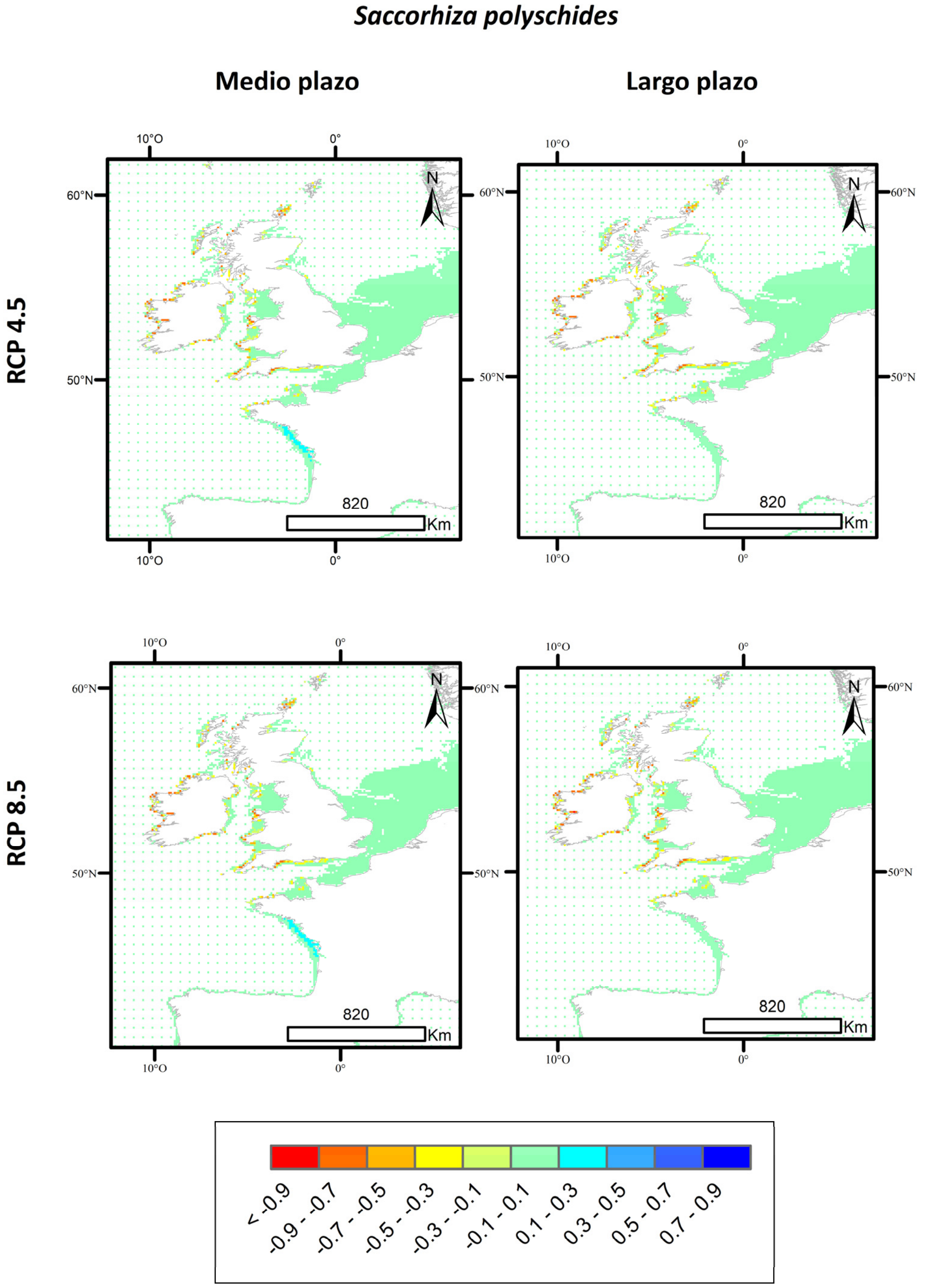


Figura 8. Diferencias en la probabilidad de presencia predicha para los escenarios RCP 4.5 y RCP 8.5 en el medio y largo plazo con el periodo histórico para *Saccorhiza polyschides*. Valores positivos (en azul) indica ganancias en los RCPs; mientras que los valores negativos (en rojo) representa pérdidas. Los mapas se centran en las zonas con valores diferentes de cero para facilitar su visualización.

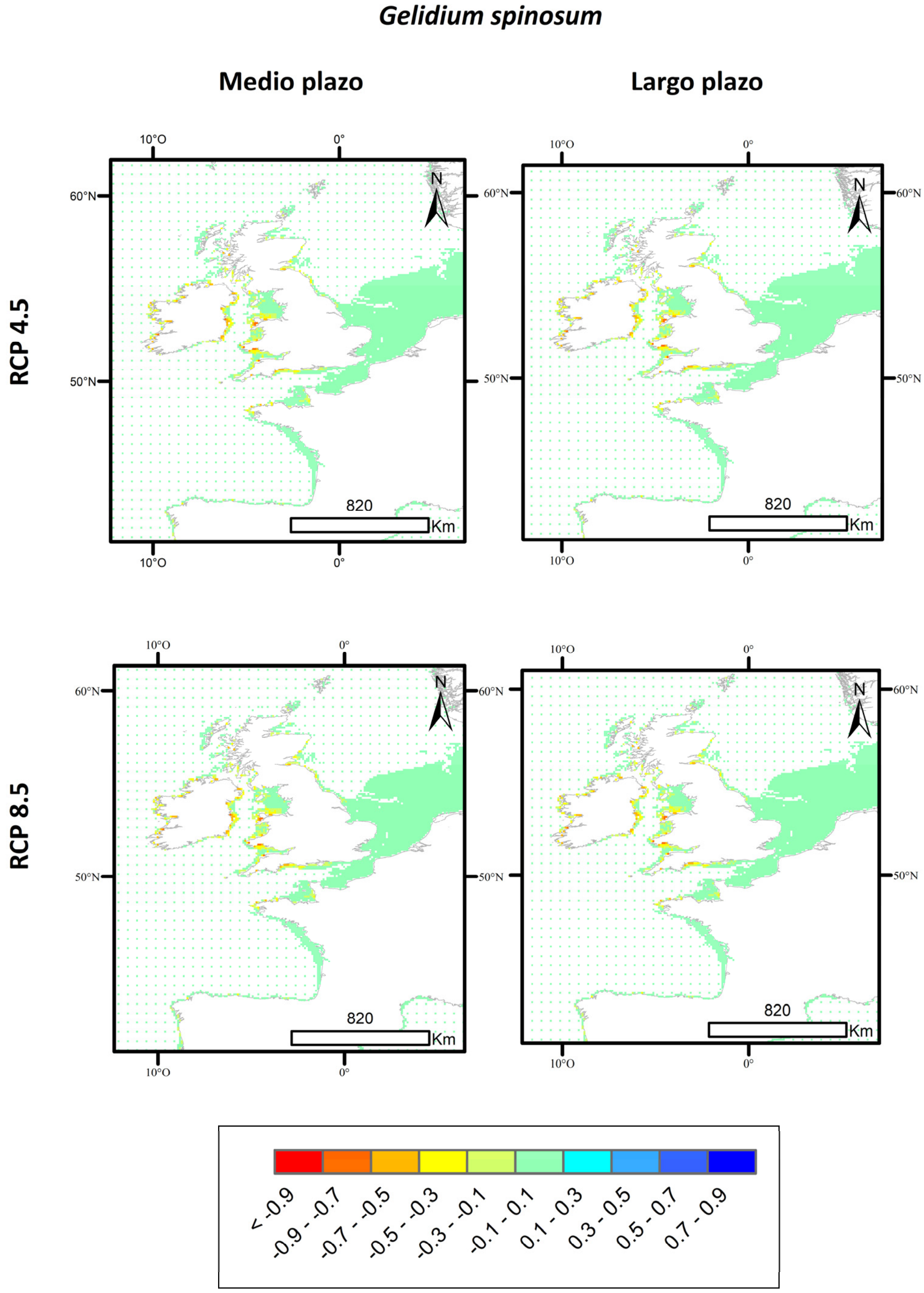


Figura 9. Diferencias en la probabilidad de presencia predicha para los escenarios RCP 4.5 y RCP 8.5 en el medio y largo plazo con el periodo histórico para *Gelidium spinosum*. Valores positivos (en azul) indica ganancias en los RCPs; mientras que los valores negativos (en rojo) representa pérdidas. Los mapas se centran en las zonas con valores diferentes de cero para facilitar su visualización.

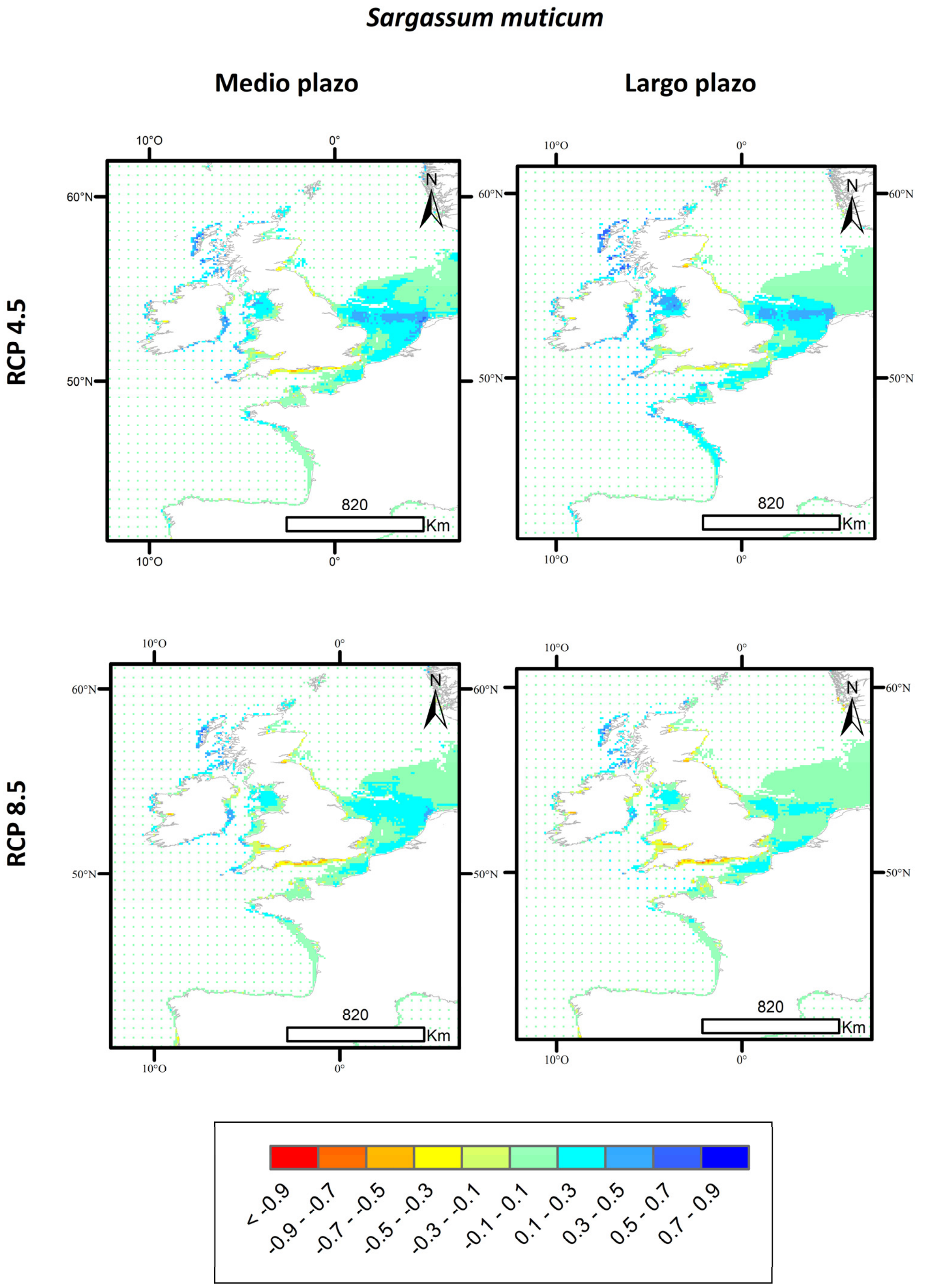
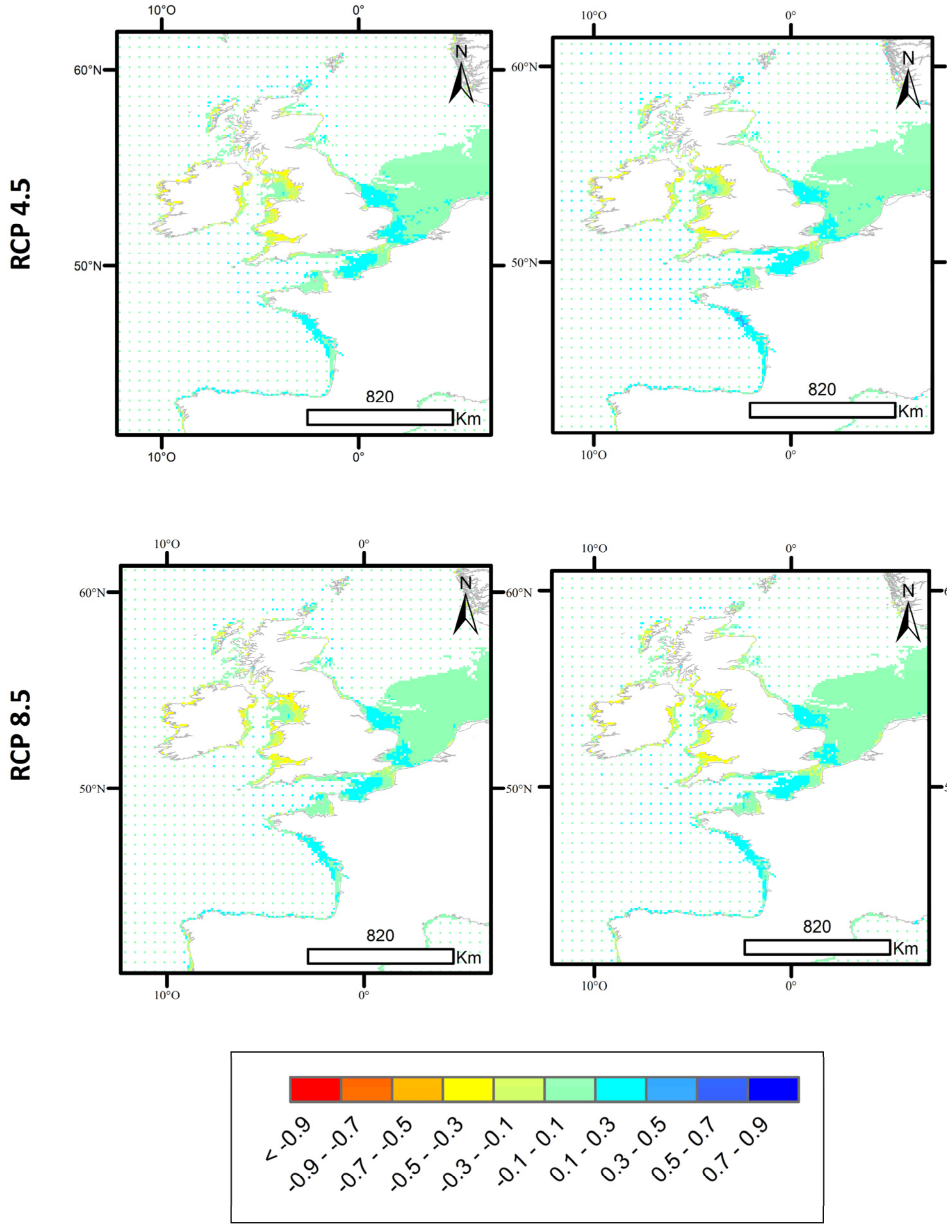


Figura 10. Diferencias en la probabilidad de presencia predicha para los escenarios RCP 4.5 y RCP 8.5 en el medio y largo plazo con el periodo histórico para *Sargassum muticum*. Valores positivos (en azul) indica ganancias en los RCPs; mientras que los valores negativos (en rojo) representa pérdidas. Los mapas se centran en las zonas con valores diferentes de cero para facilitar su visualización.



Pelvetia canaliculata

Medio plazo **Largo plazo**

RCP 4.5

RCP 8.5

Figura 11. Diferencias en la probabilidad de presencia predicha para los escenarios RCP 4.5 y RCP 8.5 en el medio y largo plazo con el periodo histórico para *Pelvetia canaliculata*. Valores positivos (en azul) indica ganancias en los RCPs; mientras que los valores negativos (en rojo) representa pérdidas. Los mapas se centran en las zonas con valores diferentes de cero para facilitar su visualización.

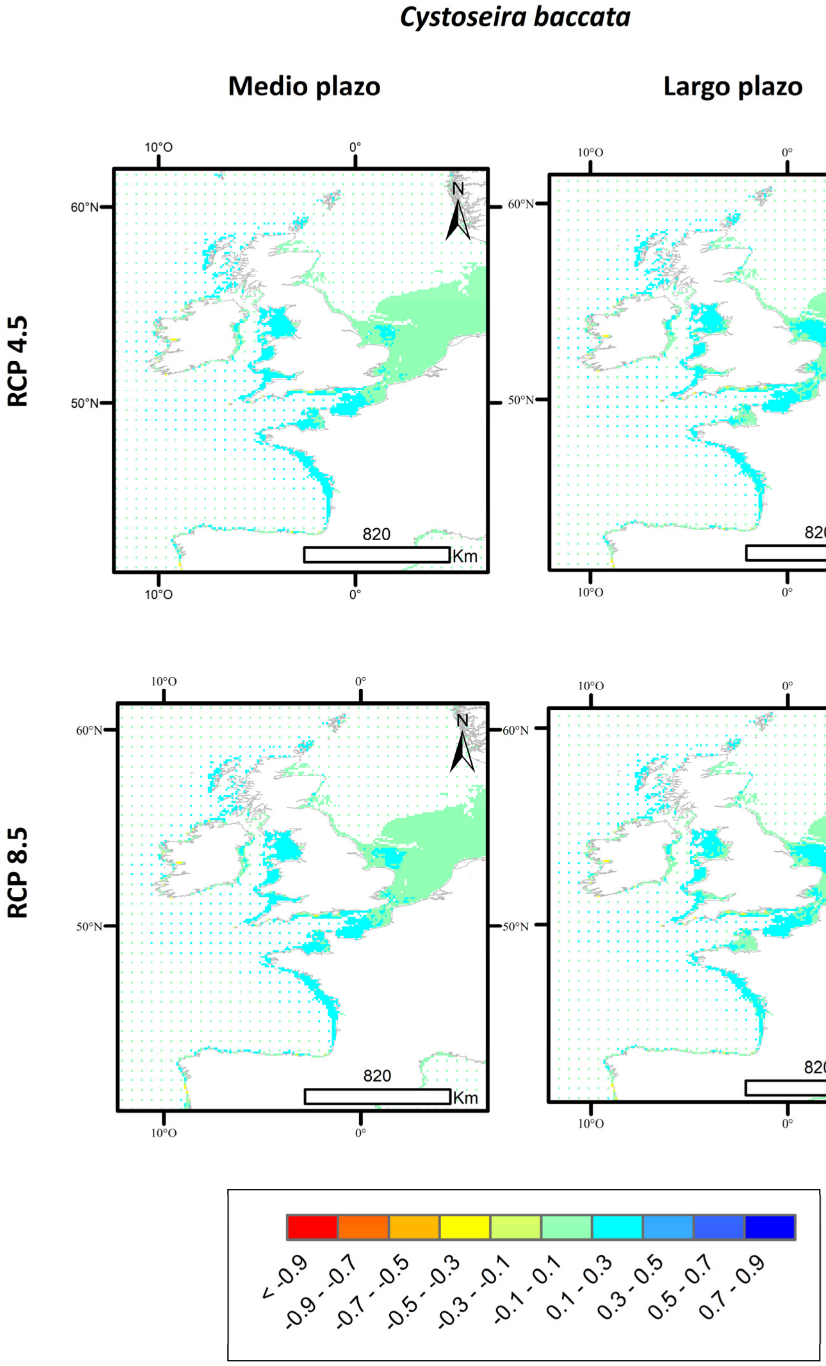


Figura 12. Diferencias en la probabilidad de presencia predicha para los escenarios RCP 4.5 y RCP 8.5 en el medio y largo plazo con el periodo histórico para *Cystoseira baccata*. Valores positivos (en azul) indica ganancias en los RCPs; mientras que los valores negativos (en rojo) representa pérdidas. Los mapas se centran en las zonas con valores diferentes de cero para facilitar su visualización.

Los resultados obtenidos respaldan los modelos generados, dados los valores de rigurosidad de las métricas utilizadas para las especies consideradas. Una de las contribuciones más destacables de este trabajo es el uso combinado de un amplio número de predictores ambientales (de la Hoz *et al.*, 2018b), especialmente hidrodinámicos, con la selección de algoritmos transferibles en el tiempo (Regos *et al.*, 2019).

La importancia de los predictores hidrodinámicos en la distribución de las macroalgas, a pesar de haber sido establecido por diversos autores (Ramos *et al.*, 2014; de la Hoz *et al.*, 2018a; Jonsson *et al.*, 2018) y respaldado por los resultados expuestos, no se incluyen habitualmente en los SDM. Además, se ha confirmado la relevancia de variables ampliamente utilizadas, como la SST (Bosch *et al.*, 2017; Martínez *et al.*, 2018a; Piñeiro-Corbeira *et al.*, 2018). Otro aspecto destacable en lo que a los predictores se refiere es el papel decisivo de las condiciones extremas en la distribución de las macroalgas, lo que apoya la importancia de considerar los factores estresantes (Galván *et al.*, 2016; Assis *et al.*, 2017a; Vranken *et al.*, 2018).

En el caso de *S. polyschides*, los resultados refuerzan estudios previos que establecen la importancia de la SST en su distribución (Norton, 1977; Fernandez, 2011), especialmente de los valores máximos en la estación de crecimiento. Además, su preferencia por zonas con rangos de temperatura pequeños (Lüning *et al.*, 1990) ensalza la importancia de la variabilidad en su distribución. *S. polyschides* vive en el submareal (Assis *et al.*, 2013), pero puede extenderse hasta zonas más altas, lo que explica la importancia de TR como un factor limitante. El modelo ha seleccionado también la disponibilidad de nitrato como un factor crítico para su crecimiento. La importancia de Tm mínima en otoño refleja su preferencia por zonas de baja energía, a pesar de ser capaz de tolerar corrientes fuertes (Norton, 1978). En cuanto a su distribución geográfica, en el periodo histórico aparece confinada en las costas de Reino Unido y la Bretaña, en línea con el conocimiento actual (Burrows *et al.*, 2018). Sin embargo, el modelo no ha identificado las presencias del norte y oeste de la Península Ibérica señaladas por otros autores (Fernandez, 2011; Assis *et al.*, 2013). Esto puede deberse a la escasez de datos en esta área durante el entrenamiento, por lo que se podría haber excluido del modelo alguna combinación ambiental que se da en esta zona. En cuanto a las proyecciones, éstas detectan su desaparición, especialmente en la zona sur, en línea con otros autores (Araújo *et al.*, 2016; Assis *et al.*, 2018). Sin embargo, autores como Assis *et al.* (2018) predicen una migración hacia el norte de la especie en el escenario RCP 8.5; mientras que el modelo desarrollado predice su ausencia también en esas zonas, aunque no se descarta una posible migración en profundidad (Assis *et al.*, 2016) porque, aunque con menor probabilidad, permanecen algunas presencias en Gran Bretaña y la Bretaña francesa. Este cambio drástico en la presencia de la especie está en línea con las observaciones de los últimos años en el norte de España, donde en 30 años se ha constatado que las poblaciones presentaban generaciones más cortas y una fracción menor era reproductiva (Fernandez, 2011).

Aunque la relación de *G. spinosum* con la SST ya ha sido estudiada (Rueness & Fredriksen, 1989), su variabilidad no se suele considerar a pesar de su importancia, como ha quedado patente en este trabajo. Además, los resultados mostraron la relevancia de los valores extremos, como el parámetro del número de días que se sobrepasa el percentil 90 del periodo histórico, cuyos valores altos disminuyen la probabilidad de ocurrencia de *G. spinosum*, de acuerdo con las curvas de respuesta. Este efecto letal de la SST ha sido ya referenciado (Fralick *et al.*, 1990), aunque su valor acumulativo no se había evaluado hasta ahora. La luz constituye un factor muy importante para *G. spinosum* (Fralick *et al.*, 1990; Rico, 1992), especialmente su rango en invierno y verano. Su distribución histórica presentaba mayores probabilidades en las comunidades del norte (Irlanda y oeste de Gran Bretaña) que en las del sur (Bretaña y Península Ibérica). Las predicciones de los modelos sugerían una extinción de las especies en los RCPs considerados, incluso en el medio plazo, en línea con las observaciones de pérdida de su biomasa entre 1985 y 2002 (Sánchez *et al.*, 2005).

El predictor más importante de *S. muticum* fue el percentil 75 de Uw, en línea con su alta tolerancia a las corrientes, incluso en condiciones extremas (Josefsson & Jansson, 2011). Dado que se tiene que enfrentar a la desecación, variables como TR, Ws o AT se han incluido en el modelo. Es destacable también la importancia del parámetro extremo de número de días consecutivos que se sobrepasa el percentil 90 del periodo histórico, tanto en la temperatura del aire como del agua (Belsher, 1989). El modelo ha detectado también su preferencia por zonas someras y sustratos rocosos (Thomsen *et al.*, 2006). La reproducción de *S. muticum* se da en primavera y verano, razón por la que la variabilidad en la disponibilidad de luz en estas estaciones se ha incluido en el modelo. Como ocurre con la mayor parte de las especies invasoras, probablemente no esté en equilibrio con el medio en Europa (Araújo & Pearson, 2005), lo que explica que sean muchas las variables que determinan su distribución. Dado que los SDMs asumen que este equilibrio existe y que el modelo está entrenado con datos de la zona invadida exclusivamente, los resultados en este caso se deben interpretar con cautela. Con esto en mente, los resultados reflejan el conocimiento actual sobre la distribución de *S. muticum*, lo que sugiere que el modelo funciona adecuadamente. La distribución histórica ha detectado todas las zonas invadidas de Europa, excepto la costa mediterránea de Francia (Critchley *et al.*, 1990). El motivo podría ser que no se dispuso de puntos en el Mediterráneo para entrenar el modelo, lo que puede haber provocado errores de omisión. En los escenarios de cambio climático, los modelos predicen la extensión de su área de distribución hacia el norte y un incremento en su probabilidad de ocurrencia. Esta tendencia continuaría con la línea ya detectada por otros autores (Karlsson & Loo, 1999; Sánchez *et al.*, 2005; Engelen *et al.*, 2015; Yesson *et al.*, 2015).

P. canaliculata tolera un amplio rango de condiciones (Lüning *et al.*, 1990), por lo que la velocidad del viento y la batimetría tienen un peso importante en el modelo, así como el rango de TR. El modelo final incorpora variables tanto terrestres como marinas, en línea con otros autores (Neiva *et al.*, 2014). Las proyecciones muestran un descenso en su probabilidad de ocurrencia, aunque se mantiene estable en algunos emplazamientos. Esto puede deberse a la alta resiliencia que presentan las especies que viven en la zona más alta del intermareal donde tienen que lidiar con las condiciones más desfavorables (Harker *et al.*, 1999; Skene, 2004). Además, su probabilidad se incrementa en algunas zonas, de acuerdo con la migración hacia el norte reportada por otros autores (Lima *et al.*, 2007; Lamela-Silvarrey *et al.*, 2012; Neiva *et al.*, 2014; Piñeiro-Corbeira *et al.*, 2016).

Dado que *C. baccata* es capaz de vivir en un amplio rango de profundidad, exposición y sedimentación (García-Fernández & Bárbara, 2016), su distribución está determinada por variables terrestres (TR, AT, Ws) y marinas (SST, Urms, Tm, Hs). La mayoría de los parámetros seleccionados en el modelo de *C. baccata* están ligados a condiciones extremas, lo que pone de relieve su importancia. Las curvas de respuesta de Urms, Tm y Hs apoyan su preferencia por zonas protegidas, a pesar de ser capaz de tolerar condiciones de semi-exposición (Díez *et al.*, 2012; Méndez-Sandín & Fernández, 2016). El modelo captura la distribución sur de la especie en el Atlántico (García-Fernández & Bárbara, 2016), cuyos límites se expanden en el RCP 8.5. Estos resultados parecen continuar la tendencia de extensión del periodo de crecimiento y el aumento de biomasa detectados en el norte de España entre 1997 y 2007 (Méndez-Sandín & Fernández, 2016). A pesar de la coherencia de los resultados con los estudios existentes, éstos han de interpretarse con cautela, puesto que el modelo se entrenó con un número pequeño de presencias.

En general se puede considerar que los resultados obtenidos son coherentes con los estudios predictivos existentes. La tendencia principal es la desaparición de los límites sur de distribución de las especies (Jüterbock, 2013; Assis *et al.*, 2018; Wilson *et al.*, 2019). Sin embargo, los modelos desarrollados predicen un escenario más dramático que otros autores que contemplan la migración hacia el norte de las especies, mientras que los resultados aquí expuestos predicen su desaparición. Esto puede deberse a que en este estudio se han incorporado variables hidrodinámicas que no se suelen incluir en el modelado, así como predictores extremos y de variabilidad con una resolución temporal más detallada que habitualmente. A esto se une el filtrado de algoritmos en base a su capacidad de transferencia temporal, lo que disminuye la incertidumbre en escenarios de cambio climático.

No obstante, los estudios de modelado están sujetos a varias fuentes de incertidumbre. Por un lado, la falta o escasez de observaciones de las especies en algunas áreas, especialmente importante en el caso de la Península Ibérica por ser el límite sur de distribución de muchas macroalgas (Ramos *et al.*, 2016b). Esto puede provocar un sesgo, por lo que la interpretación de los resultados ha de hacerse con precaución. Además, en el contexto de cambio climático los cambios en la distribución de las especies son el resultado de la interacción de muchos factores abióticos y bióticos y la capacidad de dispersión y adaptación de la especie (Soberón & Peterson, 2005). Teniendo en cuenta estos aspectos, las métricas de evaluación y la coherencia ecológica de los resultados obtenidos apoya el uso de este estudio como herramienta para la determinación del riesgo de las macroalgas debido al cambio climático.

Conclusiones

De acuerdo con el objetivo general de esta tesis se han evaluado los efectos del cambio climático en las macroalgas de Europa. Para conseguirlo se han desarrollado una serie de objetivos específicos, con cada uno de los elementos que integran la evaluación de riesgos relacionados con el cambio climático (IPCC, 2014).

Las amenazas (objetivo específico 1) se han generado y puesto a disposición de otros investigadores en la web OCLE. Para ello se ha homogeneizado espacial y temporalmente información procedente de satélite, datos *in situ* y reanálisis de 16 variables y 10 parámetros relacionados con la ecología de las macroalgas, para el periodo histórico (1985-2015) y los escenarios de cambio climático RCP 4.5 y RCP 8.5 en el medio (2040-2069) y largo plazo (2070-2099).

Por otro lado, las observaciones de 21 especies de macroalgas características de Europa se han recolectado y filtrado en el periodo 1985-2015. Estas se agruparon de acuerdo con sus requerimientos ambientales con la técnica SOM en cinco grupos y se seleccionó una especie representativa de cada uno de los grupos (objetivo específico 2).

Puesto que el objetivo general de la tesis es evaluar los efectos del cambio climático, fue necesario establecer una metodología para guiar la selección de algoritmos de modelado y sus pesos para asegurar la transferencia temporal de los resultados (objetivo específico 3).

Utilizando los datos de entrada de los objetivos específicos 1 y 2, la metodología del objetivo 3 se aplicó en toda Europa en los dos escenarios de cambio climático seleccionados, cumpliendo así con el objetivo específico 4. Las conclusiones principales que se pueden extraer del análisis efectuado es que las especies más amenazadas son las incluidas en los grupos representados por *S. polyschides* y *G. spinosum*; mientras que los correspondientes a *S. muticum* y *C. baccata* expandirían su rango de distribución. La probabilidad de ocurrencia de *P. canaliculata* podría disminuir, pero las predicciones parecen indicar una buena capacidad para gestionar los cambios. Además, ha quedado patente la importancia de considerar predictores hidrodinámicos en el modelado de especies, como la tensión tangencial de fondo, las corrientes o la altura de ola significante.

De forma más específica, a continuación se exponen las conclusiones obtenidas en cada capítulo.

OCLE: una base de datos de libre acceso de los efectos del cambio climático en los ecosistemas litorales y oceánicos de Europa

- La base de datos OCLE constituye un avance importante en la disponibilidad de información ambiental relacionada con los ecosistemas marinos, que facilita la caracterización de las amenazas en el contexto de los estudios sobre los efectos del cambio climático en los ecosistemas litorales.
- OCLE contribuye al conocimiento existente con la inclusión de variables hidrodinámicas relevantes para la distribución de especies marinas, como la altura de ola significante o la tensión tangencial en el fondo.
- La variedad de parámetros incluidos en OCLE a una alta resolución temporal permite definir predictores con sentido ecológico.
- El formato de salida (.csv) facilita su uso en diversas tipologías de estudios, con diferentes objetivos y escalas, como el modelado de especies o la clasificación ecológica de grandes áreas.

Selección de especies de algas clave en Europa

- El procedimiento basado en la técnica SOM facilita la selección objetiva de especies representativas en función de sus requerimientos ambientales, aspecto que permite la optimización y desarrollo de estudios predictivos a gran escala.
- Las especies *Saccorhiza polyschides*, *Gelidium spinosum*, *Sargassum muticum*, *Pelvetia canaliculata* y *Cystoseira baccata* se pueden considerar representativas, en términos de sus requerimientos ambientales, de 21 especies intermareales y submareales de Europa.
- La vinculación de dichas especies representativas con una serie de zonas donde se registran, con mayor probabilidad, las condiciones de mayor habitabilidad para cada grupo de especies representa un paso muy importante en la caracterización de la exposición.

Transferencia temporal de los modelos de distribución de especies marinos: el papel de la selección de algoritmos

- En general, los algoritmos de aprendizaje muestran una mayor sensibilidad que los basados en regresiones cuando se evalúan temporalmente.
- La transferencia temporal de todas las especies es buena, de acuerdo con las métricas de evaluación utilizadas.
- Las variables hidrodinámicas (altura de ola significante, tensión tangencial en el fondo, corrientes, etc.) son relevantes en la distribución de especies desde el intermareal medio hasta el submareal.
- Los predictores relacionados con las condiciones extremas (máximos, días que se sobrepasa un determinado umbral) resultaron decisivos para todas las especies consideradas.
- La metodología desarrollada constituye una herramienta efectiva para los estudios de cambio climático en áreas marinas, ya que establece una guía para la selección de algoritmos y pesos que optimicen la obtención de modelos con una adecuada transferencia temporal.
- Los resultados son robustos y satisfactorios para ser aplicados de forma generalizada en estudios marinos con diferentes objetivos y áreas geográficas, permitiendo la comparación entre predicciones y promoviendo el uso de modelos transferibles, especialmente en trabajos de cambio climático.

Efectos del cambio climático en la distribución de las algas en Europa

- Los modelos desarrollados han demostrado una buena bondad de ajuste y capacidad de discriminación.
- Aunque se requiere un mayor desarrollo, los resultados de esta tesis identifican las áreas y especies en riesgo de entre las cinco consideradas, con diferentes requerimientos ecológicos (desde la zona alta del intermareal al submareal) y patrones de distribución (nativas e invasoras), por lo que se podrían utilizar como representantes de otras con requerimientos similares.
- Las especies más amenazadas serían *Sacchoriza polyschides* y *Gelidium spinosum*, mientras que *Sargassum muticum* y *Cystoseira baccata* podrían expandir su rango de distribución. *Pelvetia canaliculata* podría disminuir su probabilidad de ocurrencia, pero probablemente sea capaz de asumir los cambios.

Summary

Summary

The general objective of this thesis is to assess the effects of climate change on macroalgae distribution in Europe. To reach this goal it was necessary the definition of the different elements of risk: hazards, exposure and vulnerability. Bearing this in mind, it was tried to provide an assessment of key macroalgae in European seas, through the development of tools and methodologies useful for marine researchers.

First, an ecologically-driven database of present and future drivers for marine life in Europe, the Open access database on Climate change effects on Littoral and oceanic Ecosystems (OCLE), was developed to fill the existing gaps, such as the absence of hydrodynamic variables, the lack of homogeneity in time intervals, the ecological reliability of parameters and the use of the best information available to project variables in climate change scenarios.

Data from numerical models, satellite and *in situ* measurements were compiled for all European seas at two spatial resolutions (0.1º up to 50 m depth and 0.5º for oceanic waters). Data were gathered for homogeneous and long time series (1985-2015 and 2015-2099) for 16 variables (sea surface temperature, significant wave height, bathymetry, light attenuation coefficient, substrate, nitrate, salinity, air temperature, pH, wind speed, tidal range, sea level, radiation, bottom shear stress, bottom orbital speed and currents). A quality control was applied to ensure the accuracy of data and the selection of the best general circulation models for two climate change scenarios (RCP 4.5 and RCP 8.5). For each variable, 10 statistical parameters were calculated (maximum, minimum, mean, standard deviation, range and percentiles 10, 25, 50, 75 and 90). Additionally, OCLE provides supplementary indicators relevant to marine ecosystems, such as persistence of air and sea surface temperature or the bottom shear stress over a harmful threshold for benthic communities. These indicators were calculated over different time scales: full period, year, five years, seasons and months. Besides, the high temporal resolution of raw data allows obtaining other parameters or compound indicators, which opens a broad field of potential applications and research in the European seas, such as the novelty of the bottom shear stress.

OCLE complements existing databases in the marine area by providing met-ocean information not included in other databases, such as the bottom shear stress or the significant wave height. Besides, the information is temporally disaggregated to allow the definition of precise parameters, which is a relevant aspect to link them to species distribution. Data are freely available at http://ocle.ihcantabria.com.

Once the predictors were defined, it was necessary the selection of species to work with. A set of 21 characteristics species for European ecosystems were initially considered. In

order to optimise the analysis of such information it was necessary the identification of key species. For this purpose, distribution of 21 taxa in the intertidal and subtidal area in Europe per year and season was compiled from global databases, field campaigns, own published data and collection of an important amount of valuable unpublished data. At present, macroalgae data is available along north and northwest Iberian Peninsula in OCLE website (http://ocle.ihcantabria.com). Then, species were grouped according to their abiotic requirements. A complete set of parameters was collected from 1985 to 2015 from the OCLE database for all European seas, after multicollinearity assessment (Pearson<0.7 and VIF<10). To extract environmental data patterns a SOM analysis was conducted in the final predictors. Species presences were projected over the SOM to identify groups of species with similar ecological responses, according to their requirements, representative species for each group were selected. As a result, five groups were defined characterized by the species *Saccorhiza polyschides*, *Gelidium spinosum*, *Sargassum muticum*, *Pelvetia canaliculata* and *Cystoseira baccata*.

To assess the risk due to climate change, SDMs were selected as an appropriate tool as they have been demonstrated to deal with climate change scenarios properly. However, this implies extrapolation to other time, thus the uncertainty is high. To reduce it, a step-wise methodology to select the most transferable algorithms in time in the marine environment was developed:

1) Collection of predictors related to the five representative seaweeds ecology and their records. Environmental and distribution data were divided into two independent periods to evaluate temporal transferability.

2) A model for each species was built in each period with nine algorithms (GLM, GAM, MARS, MDA, CART, SVM, FDA, RF, MAXENT) and projected into the other period. Predictor contributions to the final models were obtained.

3) Assessment of predictive performance for each model, using the area under the receiver operating characteristic curve and true skill statistics metric for both models' accuracy and temporal transferability capabilities. All values were over 0.8 for all groups. In turn, the geographical pattern of all models showed to be ecologically coherent. The algorithms and their weights that better fit to generate transferable models in time in the marine environment were retained for each species.

The objective framework developed demonstrates a high predictive power and ecological realism to generate temporally transferable models in the marine environment. Results were satisfactory and robust for being broadly applied in marine research with different objectives and geographical areas, allowing a comparative

framework between species predictions and promoting the use of transferable models, especially in climate change studies.

Finally, the effects of climate change on five seaweeds species (*Saccorhiza polyschides*, *Gelidium spinosum*, *Sargassum muticum*, *Pelvetia canaliculata* and *Cystoseira baccata*) were assessed.

First, records for the representative species previously selected were spatially filtered and predictors for historical and the RCPs 4.5 and 8.5 for the mid-term (2040-2069) and the long term (2070-2099) were evaluated to remove correlated ones with the same criteria as previously (Pearson<0.7 and VIF<10) for each species. According to the above exposed methodology, algorithms for each species were applied to generate models that were assessed by comparison of probabilities and observations (area under the curve, true skill statistics, Boyce index, sensitivity, correct classification rate), niches overlap (Schoener's D, Hellinger's I), geographical similarity (interquartile range) and ecological realism.

Models built demonstrated a very good goodness-of-fit and discrimination performance, therefore it can be inferred that the most endangered species would be *Sacchoriza polyschides* and *Gelidium spinosum*, meanwhile *Sargassum muticum* and *Cystoseira baccata* would be predicted to expand their range. *Pelvetia canaliculata* might decrease its probability of occurrence but it would be able to cope with changes.

According to the general objective of this thesis, results should be considered as a first step in the assessment of the magnitude of modelled range changes in seaweeds in Europe with different ecological requirements (from upper littoral to subtidal) and distributional patterns (native and invasive). The novelty of this work is the supply to marine researchers and managers with a complete pull of species physiological thresholds by the inclusion of hydrodynamic variables and extreme parameters, whose importance in species distributions has been demonstrated. Therefore, it helps to fill the gap in knowledge between seaweeds ecology and environmental drivers on the one-hand and between science and managers on the other, by paying particular attention to building robust models with objective, reproducible, globally applicable and efficient methodology.

Chapter I

Introduction and background to the research

Chapter I. Introduction and background to the research

1. Motivation for the research

1.1. Climate change in marine systems

Marine environments are complex systems, governed by the energy flow from the base of food webs, sustained by primary producers, through to top predators and pathogens, and then back again through decomposition and detrital pathways. This equilibrium maintains the functioning of the ecosystem and the ecosystem services derived. According to Costanza *et al.* (1997) ecosystem services are the benefits that human populations derive, directly or indirectly, from ecosystem functions. In coastal and ocean systems, examples of the wealth of free natural benefits that society depends upon include the provision of natural resources (food, fuel, biochemical resources); nutrient recycling; the regulation of the global climate (oxygen production, atmospheric carbon dioxide removal); protection from extreme weather and climate events; and aesthetic, cultural and supporting services (IPCC, 2014).

Although the Earth has experienced climate changes in the past, the current rate of environmental modification is much faster, taking place within a few generations of a species. In fact, the resilience of marine ecosystems to adjust to climate change impacts is likely to be reduced by both the range of environmental factors and their rate of change (IPCC, 2014). In marine systems, the main threats are sea level rise, sea water warming and acidification. These changes all pose risks for marine life and may affect the oceans' ability to perform the wide range of functions that are vitally important for environmental and human health. They also occur in an environment subjected to other human pressures that affect ocean conditions, such as overfishing, pollution or nutrient runoff. The wide range of factors that affect ocean conditions and the complex ways these factors interact make it difficult to isolate the role that any one factor plays in the context of climate change and to identify with precision the combined effects of these multiple drivers.

Sea level rise

Coastal environments are particularly sensitive to sea level rise because they will increasingly experience adverse effects such as submergence, coastal flooding and coastal erosion, and are perceived as the most important risk to human systems. Under RCP 8.5 an increase in mean sea level of 0.71±0.28 m is predicted by 2100, although the sea level rise may be higher when considering regional variations and local factors the sea level rise can be higher, reaching up to 30 % higher along the North Atlantic Ocean and the Antarctic Circumpolar Current, and up to 20% higher in the subtropical and equatorial regions (Slangen *et al.*, 2014). This has serious implications for coastal cities, deltas, and low-lying states, which will also face the increase in the most intense tropical cyclones. Changes in storms and associated storm surges may further contribute to changes in sea level extremes. In marine areas, rising sea level causes changes in parameters such as available light, salinity and temperature, and therefore its impact is related mostly to an organism's capacity to keep up with the vertical rise of the sea (IPCC, 2014). Sea level rise and changes in storm wave and surge characteristics lead to enhanced dune erosion, allowing waves to attack dunes at a higher level, and thus a decrease in safety levels (de Winter & Ruessink, 2017). Ecosystems at the land sea border, such as saltmarshes and mangroves, are also been affected by sea level rise. A dieback in mangroves due to subsidence has been observed (Lovelock *et al.*, 2015; Albert *et al.*, 2017); meanwhile seagrasses and algae growing at their depth limit will be sensitive as increasing water depth reduces light penetration or hydrodynamic conditions. This could result in complete loss at the deeper edge, and biomass and growth are likely to reduce for the remainder of the meadow (Waycott *et al.*, 2007; Albert *et al.*, 2017).

Ocean acidification

Ocean acidification is another factor of concern. Anthropogenic CO_2 enters oceans and chemical reactions turn some of it to carbonic acid, which acidifies the water. The global pH of coastal waters has been predicted to decrease from about 8.16 in the year 1850 to 7.83 in 2100 (Lerman *et al.*, 2011), showing considerable spatial variability, with higher increases in acidity in areas where eutrophication or coastal upwelling occurs (IPCC, 2014). Marine species that are dependent on calcium carbonate ($CaCO_3$), such as shellfish, sea stars and corals, may find it difficult to build their shells and skeletons under ocean acidification (IPCC, 2014). Reported effects in corals caused by ocean acidification include decreased productivity and bleaching in corals and crustose coralline algae (Anthony *et al.*, 2008; Albright *et al.*, 2010); decreased recruitment and settlement (Albright *et al.*, 2010; Ateweberhan *et al.*, 2013); reduced primary polyp growth (Anlauf *et al.*, 2011); and effects on early life history (Albright *et al.*, 2010).

Sea surface warming

Sea surface warming projection varies considerably between emission scenarios, ranging from about 1°C (RCP 2.6) to more than 3°C in RCP 8.5, with an average rate of 0.18 ± 0.16°C per decade. Mixing and advection processes will gradually transfer the additional heat to deeper levels of about 2000 m at the end of the 21st century. In addition to the projected warming of the sea surface, more frequent heat waves are expected, which will provoke the decline of some marine ecosystems, such as temperate seagrass and kelp forests. This will facilitate the establishment of invasive species, resulting in spatial shifts of native species and a global redistribution, with implications for biodiversity and the maintenance of ecosystem services.

The vulnerability of most organisms to warming is set by their physiology, which defines their temperature ranges and, therefore, warmer temperatures have direct effects on species adjusted to specific, and sometimes narrow, temperature ranges, and are fatal to those already adapted to living at the limit of their thermal sensitivity. Warmer temperatures consequently cause changes in abundance through local extinctions and latitudinal expansions or shifts.

Ocean warming is considered the main driver in intertidal species shift, reaching up to 50 km per decade over the past 30 years in the North Pacific and North Atlantic, much faster than most recorded shifts of terrestrial species (Helmuth *et al.*, 2006). This effect is mainly evident within species living in their geographic limits (Lima *et al.*, 2007; Wethey & Woodin, 2008; Nicastro *et al.*, 2013) or confined areas (Marbà *et al.*, 2014). Seawater warming also induces stratification, which limits the exchange of gases between water layers and enhances oxygen consumption by heterotrophic organisms. Consequently, the oxygen depletes, causing a community shift toward lower species richness and hypoxia-tolerant specialists, usually dominated by microbes (IPCC, 2014).

Interactions of temperature and ocean acidification enhance sensitivity to temperature extremes in organisms such as corals, coralline algae, molluscs, crustaceans and fishes. Light and individual nutrients can also interact with temperature and acidification in primary producers. The integrated and synergistic effects of these multiple stressors on marine ecosystems must thus be considered altogether, not as independent issues (Doney, 2010), taking into account the increase in the frequency of extreme events (Smale & Wernberg, 2013), ecosystems are at increasing risk of being negatively affected by the combined consequences of these drivers because during the next decades.

1.2. Impacts on macroalgae communities

Seaweeds are key structural species characterising intertidal and subtidal communities along the coast of Europe. They play an important role in nutrient cycling, energy capture and transfer, and coastal defence (Krause-Jensen & Duarte, 2014). Kelps also provide extensive substrata for colonising organisms, ameliorate conditions for understorey assemblages, and provide three-dimensional habitat structure for a vast array of marine plants and animals, including a number of commercially important species (Teagle *et al.*, 2017). They are also the basis of the food web and impacts on these habitats will have cascade effects in entire ecosystems (Mineur *et al.*, 2015).

Changes have already been reported in Europe, such as the disappearance and replacement of cold-temperate canopy species by warm-temperate species (Fernández, 2016), their biomass and abundance decline (Borja *et al.*, (2018); Piñeiro-Corbeira *et al.* (2016)) and the disappearance of their southern limits (Casado de Amezúa *et al.*, 2015; Araújo *et al.*, 2016). The presence of invasive species has also been frequently reported in Europe in recent years (Karlsson & Loo, 1999; Willette & Ambrose, 2009; Steen *et al.*, 2017) adding to the stress experienced by species due to climate change.

An accurate definition of the risk of these valuable communities is essential so as to guide adaptive measurements.

1.3. Risk assessment

Given the growing challenges expected in coastal areas under climate change and the complex dynamics of the coastal socio-ecological systems, as well as the multiple management goals, the competing preferences of stakeholders and the social conflicts involved, there is an imperative need to accurately assess these risks. The IPCC has thus proposed an approach including the core Hazard-Exposure-Vulnerability-Risk concepts in an assessment of the climate-related impacts. This methodology is mainly oriented toward the impacts of climate change on human systems, and therefore the risk assessment of marine ecosystems needs some specific developments of the main concepts of the methodology (Figure 1.1):

- Hazards

  Hazards are potential events that may cause damage to, and the loss of, ecosystems. The adequate definition of variables and their parameters in the ecology of species or ecosystems is therefore crucial as an object of study that determines their biogeographical distribution.

- Exposure

  Exposure is related to the presence of species or ecosystems, environmental services and resources that could be adversely affected. The identification of target species or habitats and the collection of unbiased data about their distribution and ecology is therefore a key step.

- Vulnerability

  The propensity or predisposition of species or ecosystems to be adversely affected defines their vulnerability. It encompasses a variety of concepts that need to be addressed, including sensitivity or susceptibility to harm and a lack of capacity to cope and adapt.

- Risk

  Once the hazards, exposure and vulnerability are adequately conceptualised, risk is defined as their interaction within natural systems.

It is clear that the study of coastal ecosystems evolution, in light of this changing climate of uncertainty, requires an adaptation of existing methodologies to reach an accurate risk assessment for marine ecosystems.

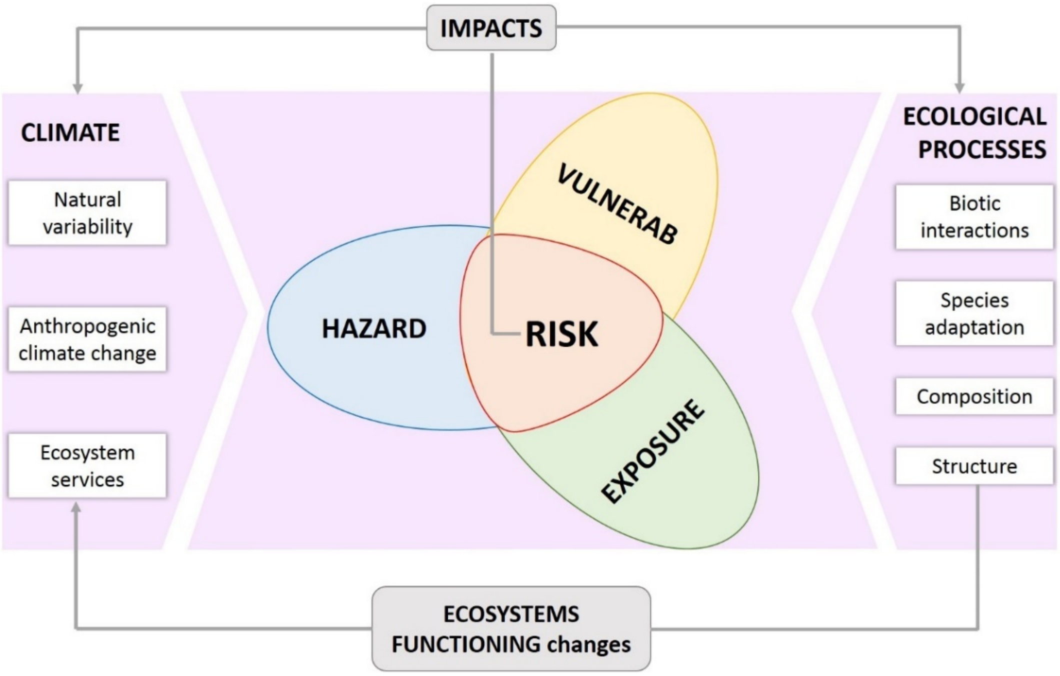


Figure 1.1. Illustration of the core concepts of the IPCC with the thesis objectives integrated. Modified from IPCC (2014)

2. State of knowledge

Several decision support systems have been developed in recent years for vulnerability assessment to encourage climate adaptation planning in coastal and marine areas (Brugère & De Young, 2015; Ellison, 2015; Torresan *et al.*, 2016). They are mainly based on the definition of the sensitivity and adaptive capacity of systems through the characterisation of indicators (Fritzsche *et al.*, 2017) and focused on multicriteria decision analysis implemented by expert judgment (De Lange *et al.*, 2009).

Despite its broad application, the concepts of vulnerability are not clear and homogeneous in the different methodologies (Adger, 2006). The IPCC report (IPCC, 2014) brought together existing terms in its proposed methodology (Figure 1.1), however, as it is oriented toward socio-ecological systems it requires its adaptation to marine ecosystems study. This implies the development of an approach that uses a biophysical threat as the point of departure in order to accurately define the hazard. This requires an in-depth knowledge of the ecology of the species or communities to be assessed and the collection of important variables at the appropriate spatial and temporal resolution. This bias also affects the collection of biological data, the elements required to quantify exposure, which is a key aspect to define what systems are vulnerable to. A suitable definition of these elements, hazards and exposure, constitutes the main source of bias in vulnerability assessment due to the difficulty to accessing to data with the resolution and coverage needed, especially when working in large areas.

Although the use of expert criteria to define vulnerability has had good results in several cases (De Lange *et al.*, 2009; Brandt *et al.*, 2017), its applicability to other researchers and its comparability is limited, therefore a more impartial tool would be useful. The SDMs are an interesting option, which have emerged in recent years as a very useful tool with which to reduce knowledge gaps in the understanding of the link between environmental conditions and species distribution, helping to respond to changes more efficiently. To obtain accurate results, however, it is necessary to select input data (hazards and exposure) accurately, as noted above. An extensive number of algorithms are available, and therefore their selection constitutes another possible source of bias (Segurado & Araujo, 2004). When working with climate change studies algorithms must be able to transfer results to other times, which is not usually assessed but is crucial in order to understand the effects (Araujo *et al.*, 2005).

A review of present knowledge and emerging questions about these issues is summarised in the following sections.

2.1. Hazards

A crucial step in habitat suitability modelling is the acquisition of spatially explicit environmental variables at the right resolution, which are sufficiently accurate to determine a species niche as close to its ecophysiological needs as possible. The more detailed the knowledge of the physiological drivers of species distributions, the more accurate the predictive ability of the model (Robinson *et al.*, 2011). Data must therefore be obtained from the best sources available in the study area, which guarantee temporal and spatial homogeneity.

The available data sources are usually grouped into three categories:

Direct observations

One of the most traditional ways to collect abiotic data is through *in situ* measurements, where the data quality is strongly dependent on the quality of the analytical procedure control. Despite numerous regional and national surveys and observation programs, a homogeneous data bank has not been compiled in a systematic way for the whole of Europe.

Indirect observations

The growing availability and reliability of environmental information through remote sensing techniques in recent years may provide a tool to overcome the limitations of *in situ* measurements. Satellite imagery has the advantage of increasing data collection frequency and expanding areal coverage, offering repeatable, standardised and verifiable information over various time periods when field sampling is not available (Turner *et al.*, 2003; Muller-Karger *et al.*, 2005; de Barbosa Araujo *et al.*, 2015; Álvarez-Martínez *et al.*, 2018).

Modelling

As a complement to these sources, numerical modelling outputs provide spatially homogenous data at high spatial and temporal resolution. These datasets historically reconstruct climate conditions, and provide long-time series, however, they should be validated/calibrated against real observations (Perez *et al.*, 2017).

When working in large marine areas, the most valuable reference is Bio-Oracle (Tyberghein *et al.*, 2012; Assis *et al.*, 2017b), which provides surface and benthic layers for water temperature, salinity, nutrients, chlorophyll, sea ice, current velocity, phytoplankton, primary productivity, iron and light at high resolution and global coverage. Other remarkable databases are MARSPEC (Sbrocco & Barber, 2013), offering variables derived from bathymetry, slope, salinity and sea surface temperature, Aquamaps (Ready *et al.*, 2010), focused on marine animals; and Hexacoral (Fautin & Buddemeier, 2002), which aims to understand spatial and temporal patterns in biogeochemistry and biogeography. Some databases cover both land and sea areas, such as the MERRAclim (Vega *et al.*, 2017), which offers decadal data for 19 derived parameters of air temperature and humidity atmospheric water vapour.

Despite the important contributions of these marine databases, various important questions still require further investigation. The first issue to be considered is the common absence of data on hydrodynamic variables with global coverage (e.g. wave height, current speed or bottom and wind stress), despite its relevance to species distribution (Ramos *et al.*, 2014; Callaghan *et al.*, 2015; Young *et al.*, 2015; Puente *et al.*, 2017; de la Hoz *et al.*, 2018a). Bottom shear stress is very important when studying benthic vegetation, due to its effect on settlement and survival (Sutcliffe *et al.*, 2014; Pace *et al.*, 2017), but this kind of information does not seem to be currently available for large areas.

A second concern involves to the lack of homogeneity in the time intervals used to calculate various parameters and the consequent limitations on long-term multi-criteria retrospective analysis.

The third issue that rises in this analysis applies to the ecological reliability of the selected parameters (Petitpierre *et al.*, 2017). Most databases only provide mean, minimum and maximum values for long periods, although many environmental affecting life cycles and species distributions seem to act on extreme events occurring on shorter time scales (Seabra *et al.*, 2015; Galván *et al.*, 2016), especially in a climate change context (Lima & Wethey, 2012; IPCC, 2013; Vasseur *et al.*, 2014). The formulation of biologically-meaningful parameters using datasets and increasing time resolutions involves two key steps in order to get more realistic results.

The GCMs, which take into account the RCPs introduced in the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2014) provide the best information available, however, one of the main challenges associated with GCMs is model uncertainty, and therefore the assessment of their performance is necessary in order to select those with better ability to reproduce the climate in the study area (Perez *et al.*, 2014). Although there are some works in the North-East Atlantic Ocean region (Perez *et al.*, 2014), no evaluation of all European seas has been conducted. The fifth phase of the CMIP5 (Taylor *et al.*, 2012) constitutes the most current set of coordinated climate model experiments. A common procedure with which to evaluate the ability of GCMs is to compare outputs of model simulations against historical reconstructions (reanalysis) or observations.

Accurate predictors for historical and future climate change conditions that are temporally and spatially homogenous from an ecological point of view, are still a need in the marine environment.

2.2. Exposure

The collection of biological records is usually the most difficult step, but it is crucial for the definition of vulnerability. It is more remarkable when working in large areas, where there is a lack of homogeneous information.

The available data sources in this field are usually grouped into three categories, as for hazards:

Direct observations

Ideally, biological data is from well-designed field campaigns, with accurately located species presences and absences in the area (Franklin, 2009), however, when working in large areas, it is not possible to sample the whole study area, and online global databases are a good option to fill the existing gaps. Two of the most widely used databases for large scale and global distribution of species are the Global Biodiversity Information Facility (GBIF) (GBIF, 2013) and the Ocean Biogeographic Information System (OBIS) (OBIS, 2015). They provide reasonable information about species distribution. Despite the important advance in open access information that they imply, there are some gaps and bias in the collections, such as uncertainty in species identification, lack of design, incomplete or uneven spatial coverage of the true distribution of a species, lack of temporal information and long series, georeferencing errors or spatial autocorrelation in sample locations.

Indirect observations

Various indirect techniques have been used to solve some of the above challenges. An example is the work developed by de la Hoz *et al.* (2018a), available as Annex A, which constitutes the base of this thesis. The basis of the study was that, as biological processes are difficult to directly record over large areas, abiotic characteristics can be used as surrogate parameters. The identification and mapping of large-scale ecological units for which conditions are similar and whose communities may also be similar, was used as an approach.

Typologies were defined by grouping abiotic variables with a combination of clustering techniques: SOM (Kohonen, 1998) and k-means (Hartigan & Wong, 1979). Although they have been used extensively with environmental data (Peterson *et al.*, 2011; Galván *et al.*, 2016; Guisan *et al.*, 2017) and their power to classify distribution data according to ecological requirements has been proven (Ramos *et al.*, 2017), they have not been used cwith a biological grouping in large marine areas, which might be a very useful tool for working with extensive number of species.

Results were consistent with the ecological requirements of the five studied seagrasses (*Posidonia oceanica, Zostera marina, Zostera noltei, Cymodocea nodosa, Halophila stipulacea*), supporting the suitability of the classification and the proposed methodology for mapping large areas, and the use of environmental information as a proxy for species presence.

Nevertheless, the geographical coverage of typologies obtained was too big if the aim of the study is to manage high resolution species distribution.

Modelling

A further step in defining the occurrence of species in geographical and environmental space is the development of SDMs.

SDMs generate hypotheses about mechanisms controlling the spatial patterns of species distributions at geographical scales, based on the relationship between species occurrence data and environmental predictors (Peterson *et al.*, 2011). The theoretical framework for this approach is that of the ecological niche. Grinnell (1917) referred to the niche as the climatic and habitat requirements of a species, that is the environmental factors, expressed geographically. In a contrasting interpretation, Elton (1927) viewed the niche as the functional role of an animal in a community (its local effects), therefore placing emphasis on the impacts. Hutchinson (1957, 1978) provided a conceptualisation of the Grinnellian niche as a hypervolume of environmental states that allows a given species to establish, reproduce and survive in a certain geographical region or biotic community. A species may therefore maintain local populations if its reproductive rate is non-negative across space and through time (Peterson *et al.*, 2011).

SDMs can usually be interpreted as describing a species' Grinnellian niche, which may not define the complete n-dimensional Hutchinson's conceptualisation, but determine an approximate set of conditions in which species can occur (Tingley *et al.*, 2009). This is considered the best alternative for modelling non-equilibrium distributions between species and the environment (Holt & Keitt, 2000), despite being considered static approaches, as they are fed with environmental limiting factors collected recently or past periods, which change very slowly at an evolutionary time-scale (Elith *et al.*, 2010), assuming niche conservatism (Pearman *et al.*, 2008; Wiens *et al.*, 2010). Hutchinson (1957) also introduced the concept of the fundamental niche as comprising those environmental abiotic conditions that allow a species to maintain sustainable population growth, and the realised niche as the part of the fundamental niche that can be occupied when competition or other forms of biotic interactions are included (Peterson *et al.*, 2011).

Although this description only considers abiotic conditions, when building models to explain and predict the distribution of organisms it is necessary to take into account that there are three main conditions to be met in order for a species to occupy a site and maintain populations (Simões & Peterson, 2018). The first condition is a matter of species dispersal capacity, including all factors limiting its distribution, such as barriers to migration, biotic and abiotic dispersal factors, etc. The second refers to the abiotic habitat suitability for the target species, which implies that the combination of abiotic environmental variables at the site is included in the environmental conditions that the species needs to grow and maintain viable populations. The third condition concerns biotic interactions with other organisms, either positive (commensalism, mutualism) or negative (competition, predation), which themselves are dictated by the environment through their effect on all organisms in the local community (Guisan *et al.*, 2017). Therefore, modelling habitat suitability fed with environmental information also depends on dispersal limitation and biotic interactions.

SDMs can be classified according to their theoretical basis, in correlative or mechanistic models. Mechanistic models involve determining the ecological physiology of a species, such as the range of conditions under which a species can live (Peterson *et al.*, 2011). This information can be obtained through laboratory experiments, which constitutes its main drawback, as they are not readily applicable to large numbers of species. A related problem is that most climatic data are not in the form of variables related directly to limiting physiological mechanisms, and therefore laborious measurements and calculations are needed. On the other hand, correlative methods estimate the existing fundamental niche from observations of the presence of a species in relation to environmental variation. Its feasibility is higher as the information required is easily attainable, however its use has been contentious, since it is likely that the occupied distribution of a species already includes reductions due biotic interactions and dispersal limitations (Peterson *et al.*, 2011). As correlative approaches are most broadly applicable to diverse questions regarding the ecology and geography of biodiversity phenomena, most techniques used to model species' distributions rely on them rather than mechanistic models.

Even though the processes driving a species' relationship with an environmental gradient may vary, the correlative approach can still characterise, in a statistical sense, the complex response of a species to these different processes, as reflected in its distribution (Barry & Elith, 2006). As correlative models are the most commonly applied nowadays, they will be the only to be considered.

SDMs usually quantify a species' realised niche by taking all known localities of the species and extracting the attributes of the environmental variables that define the species' niche for each. Once species observations are related to the environmental data a model of the species-environmental relationship is obtained, resulting in a continuous index of habitat suitability or probability of species occurrence (Peterson *et al.*, 2011; Guisan *et al.*, 2017). This implies that environmental and distributional information is needed as input data to build these models. This need can constitute a limitation when models are calibrated with existing large scale databases about species presence–absence (or presumed absence), collected at varying resolution, accuracy and reliability (Peterson *et al.*, 2011). Alternatively, the selection of variables and their parameters that control population dynamics (such as growth, survival or reproduction) must be representative, geographically and temporally, however, these data (i.e. obtained through long periods and large geographical scales) are rarely available for entire populations.

2.3. Vulnerability and risk

To assess the effects on natural systems of climate change, SDMs is one of the most powerful tools as they have been successfully applied (Assis *et al.*, 2017a; Buonomo *et al.*, 2018; Chefaoui *et al.*, 2018; Jonsson *et al.*, 2018) and the information needed is easy to find ready-to-use in global databases. Various issues have to be taken into account when considering extrapolation to other climates, being the transferability one of the most important. This implies the evaluation of the existing algorithms in terms of their ability to be projected to different environments.

2.3.1. Types of algorithms

Nowadays, there are many algorithms with which to build species models. Complete reviews can be found in Peterson *et al.* (2011); Guisan *et al.* (2017); Gobeyn *et al.* (2018). One classification refers to the type of distribution data needed as input, distinguishing between presence-only, presence-absence or the intermediate type, presence-background algorithms. This implies that background points are generated as pseudo-absences. The geographic or environmental extent over which these background points are sampled has also been shown to potentially affect model performance (Barbet-Massin *et al.*, 2012). Approaches using predictions from a preliminary model to stratify the selection of background have also been proposed (Engler *et al.*, 2004). Compared to simple random selection, this approach has the disadvantage of amplifying any bias in the initial species data (Barbet-Massin *et al.*, 2012). Sampling background points randomly is the strategy with the least assumptions and should be used by default if there are no strong arguments in favour of a different, more taxon-specific approach (Barbet-Massin *et al.*, 2012). Constraining the sampling to areas where the species has not been observed to prevent conflicting presence and absence observations is likely to increase initial bias and may result in over-predictions of the species ranges (Wisz & Guisan, 2009; Hanberry *et al.*, 2012).

Presence-only and presence-background approaches are the simplest methods available, usually based on very simple rules and assumptions. They can roughly be separated into different categories:

a) Envelope approaches.

These can be geographic (focused on the geographic distribution of a species or population) or environmental (based on the potential environmental drivers of species distributions). Geographic approaches define the extent of occurrence of a species as the area contained within the shortest continuous geographic boundary (e.g. convex hull); meanwhile environmental defines the ecological niche of a species as the *n*-dimensional bounding box that encloses all the records of the species in the environmental space defined by *n* pre-selected variables (Busby, 1991). In a way, this is similar to Hutchinson's view of the realized niche except that it only considers presence data and does not provide an estimate of habitat suitability (Guisan *et al.*, 2017). The rectilinear envelope is defined in the environmental space by means of the most extreme (minimum and maximum) records of the species along each selected environmental variable. In order to reduce the sensitivity of model predictions to outliers (i.e. sink populations), species records can be sorted along each variable and only the records that lied within a certain percentile range of these environmental gradients can be used for model construction. Its main disadvantage is that it assumes independent rectilinear bounds and that all variables are known, and it will cause over-prediction when not enough variables are included and under-prediction with too many (or even spurious) variables (Barry & Elith, 2006). They also assume the relationship between the presence of a given species and any given variable is binary, and thus a single presence record under an extreme climatic condition at the edge of a species' range, for example, has the same weight as thousands of presences recorded in the core of the range (Barry & Elith, 2006). This highlights the importance of carefully selecting the variables (Austin, 1990).

b) Distance-based approaches.

These are refined alternatives to envelope approaches as they are usually built on the distance between the environmental centroid of the study area and that defined for the species (Guisan & Zimmermann, 2000; Elith & Leathwick, 2009). This approach is meant to overcome some of the limitations previously discussed such as variable selection and variable importance, which can be used to calculate the axes of the environmental space. Various approaches have been proposed, such as those based on principal component analysis (PCA-sp, (Robertson *et al.*, 2001)) or the ecological niche factor analysis (ENFA, (Hirzel *et al.*, 2002).

c) Regression-based approaches.

These approaches are the most commonly used in ecology and particularly in habitat suitability modelling (Guisan *et al.*, 2017). Regression fit a curve through a set of points using some goodness-of-fit criterion. There are numerous variations in how a response variable (e.g. presence-absence, abundance, biomass) is related to a set of pre-selected environmental predictors because the response can be calculated with respect to many variables; it can be linear or non-linear; the models can be purely additive or include interactions between predictor variables (Barry & Elith, 2006). The most well-known examples are GLMs, GAMs and MARS. GLMs relates the linear models with the response variable via a link function, allowing the magnitude of the variance for each measurement to act as a function of its predicted value. GAMs automatically fit response curves as closely as possible to the data given the permitted level of smoothing, and thus do not require a shape to be postulated from a specific parametric function. GAMs are therefore useful when the relationship between variables is expected to be of a more complex form, not easily fitted with standard parametric functions of the predictors or where there is no a priori reason for using a particular shape (Hastie & Tibshirani, 1990). Like GAM, MARS is a more flexible regression technique than GLM, as it also does not require any assumptions to be made about the underlying functional relationship between the species and the environmental variables.

d) Classification and machine-learning approaches.

These rely on the concept of classifying observations into homogeneous groups. Recursive partioning, such as CART, is meant to explain the variation for a single response variable with one or more explanatory variables. The response variable can be either discrete (classification trees) or continuous (regression trees). A decision tree is grown by repeatedly splitting the data, defined at each split (node) by a rule based on a single explanatory variable. At each split the data is partitioned into two mutually exclusive groups. The criteria for segmenting the data are based on either minimising the classification error rate in the case of a classification tree or maximising the inter-class variance in the case of a regression tree. The splitting procedure is then re-applied to each group separately, repeating the same procedure at parallel nodes, thus growing the tree iteratively (Guisan *et al.*, 2017). One advantage of recursive partioning method is that it does not rely on assumptions about the relationship between the explanatory variable and the response variable of interest. It also does not expect the dependent variable to follow any specific distribution, thus it is entirely data-driven. Discriminant analyses are methods used to classify individuals (e.g. sites, samples and populations) into groups (low, moderate or high suitability) based on a set of features (e.g. environmental variables) to describe them. Because it relies on linear combinations of predictor variables, it is not always relevant for modelling species distributions. Various extensions have been proposed in the past, notably the FDA (Hastie *et al.*, 1994), which allows the user to replace the linear combination with non-parametric functions such as MARS.

e) Bagging and boosting techniques.

A classification or regression method is applied to various resampling of the original dataset or through a stage-based framework, respectively, and then combined (ensembled) using different weighting schemes (Guisan *et al.*, 2017). RF and BRT have gained momentum in the ecological literature in recent years (Elith *et al.*, 2008; Peterson & Herkül, 2017; Kotta *et al.*, 2018). RF has been developed to check for overfitting by adding some stochasticity to the process of building the trees, but also at each node of each tree (Breiman, 2001). BRT use a forward stage-wise procedure that iteratively fits simple trees to the training data, while gradually increasing the focus on poorly modelled observations (by fitting residuals to the same predictors again). The general idea is to compute a sequence of very simple trees, where each tree is fitted to the residuals of the set of trees so far developed to improve the predictive ability of the model and the bias and variance of estimates (Friedman, 2002).

f) Maximum entropy.

This states that the probability distribution that best represents the data is that with the greatest entropy. The pixels representing the distribution of species occurrences constitute the sample points and their environmental features are the explanatory variables (Elith *et al.*, 2011).

Considering the variety of existing algorithms, it is clear that selection implies the introduction of bias in the model, which has to be taken into account when interpreting the results.

2.3.2. SDMs assumptions

Some assumptions are made when building species models. Both environmental and distributional data imply some bias that has a great impact on the resulting models (Bedia *et al.*, 2013).

First, it is assumed that all important environmental predictors required to capture the species niche are included in the model at the relevant resolution (Guisan *et al.*, 2017). Due to the lack of knowledge about species physiology and/or the availability of reliable data, however this is rarely achieved (Araújo & Guisan, 2006). This generates an unexplained variance that biases the quantification of the climate niche. Important predictors that are unavailable should thus be identified prior to model fitting, and implications anticipated to ensure successful predictions and avoid drawing spurious conclusions. Whether species observations are appropriate to fit a model can only be determined according to the later use of the model to answer the initial aims of the study. For example, if simple presence-absence is available, one cannot exclude sink populations (i.e. outside the species' fundamental niche, where mortality is greater than fecundity and populations cannot maintain viable populations without constant immigration; see Pulliam (2000)). If the models are calibrated from data that include sink populations, this may seriously mislead further applications, for instance if the predictions are used to guide conservation decision-making (Guisan *et al.*, 2013).

The availability of species records will also determine the algorithm chosen, as explained above, so that the inherent bias linked to this choice.

Finally, one of the most controversial assumptions related to SDMs, is that the species-environmental relation is considered at equilibrium in space or time (Araújo & Pearson, 2005). Observed patterns are assumed to reflect the species' full biotic potential, implying that the species can potentially occur in all environmentally suitable locations and that its distribution has not been constrained locally by factors, such as historic accidents. Whether this is plausible depends on the scale of the model, the dispersal ability of the organisms and the history and biology of the species (Pulliam, 2000). In practice, errors resulting from the equilibrium assumption are most acute when trying to predict distributions of species recently introduced to new locations (Barry & Elith, 2006). Recent findings suggest that niche shifts do occur in previously unrecognised situations, owing to ecological processes changing the realised niche, and/or evolutionary processes altering the fundamental niche. Furthermore, we still cannot predict which species will be affected by niche changes, nor it is clear under which ecological conditions or over what time periods these niche changes will be observed (Pearman *et al.*, 2008). Caution is therefore needed in circumstances in which this assumption does not hold, such as biological invasions or when species are recolonising a territory after major environmental changes. Some authors suggest the assessment of equilibrium of species with their environment and how long it would take to reach a new equilibrium (Araújo & Pearson, 2005), which constitutes an interesting consideration.

Following Guisan *et al.* (2017), some methodological assumptions, which complement the theoretical assumptions, can be considered. The first is the appropriateness of the statistical methods: different types of response variables (semi-quantitative data vs quantitative data) require different types of statistical methods (Guisan & Zimmermann, 2000). Failure to identify the correct method can lead to errors and uncertainty in the predictions. Secondly, estimation of the spatial distribution of errors associated with environmental predictors is essentially used to calculate spatial uncertainty in the model predictions, but it is usually considered that predictors are measured without error, which implies an important bias. Another bias in input data involves species data because unbiased species data is assumed when modelling. Bias is likely to arise when the chosen sampling design lacks a random component or when the data are gathered without employing a designed sampling strategy, leading to partial niche quantification. The independence of species observations is also assumed but sometimes they are spatially autocorrelated. This is difficult to solve, however as there is no simple way of checking whether observations are independent, especially in biological systems where species interact with each other and disperse into neighbouring sites (Naimi *et al.*, 2011).

Although some of these assumptions cannot be removed, other can be minimised, such as the inclusion of a high pull of predictors, whose relationship with species ecology is clear. In any case, these assumptions have to be taken into account when interpreting model results.

2.3.3. SDMs assessment

Assessing the predictive power of a model is of paramount importance, for both theoretical and applied issues. It is a crucial step in any modelling exercise, as it evaluates the capacity of a given model to reflect the “truth”, its inherent uncertainty in the parameter estimations and whether it can be applied under other conditions (Barry & Elith, 2006; Liu *et al.*, 2009). A sound evaluation primarily depends on the intended use of a model and therefore on the aims of the underlying modelling study. The minimum requirement for a model that aims to explain patterns is that it is robust in its ecological assumptions, whereas a model that aims to predict distributions needs to be robust from a predictive capacity perspective (Guisan *et al.*, 2002).

Model strength has to be assessed, both in terms of the application of the model to predict independent data (predictive ability, “testing”) and to the ecological meaningfulness of the underlying model (ecological realism). This implies a certain degree of independence in the test dataset (Hastie *et al.*, 2009).

Generally, measures can be grouped as threshold-dependent and threshold-independent (Liu *et al.*, 2009). The main limitation of first group is that an arbitrary choice of threshold is required in converting probabilistic predictions to binary ones. This is the case for metrics based on a confusion-matrix, such as sensitivity, specificity, positive predictive value, TSS or Cohen’s Kappa, however, some authors have argued that it is better to consider the discrepancy between the model inferences (such as predicted probability of occurrences) and the actual observations (Guisan *et al.*, 2017). This leads to statistics such as the deviance of the model (Hastie *et al.*, 2009) or AUC (Hanley & McNeil, 1982). The most commonly used statistic is the AUC, which is obtained by plotting, for each threshold, the proportion of true positive against the proportion of false positive and by computing the area under the curve thus defined.

Various judgment scales have been proposed to interpret AUC values. Swets (1988) defined AUC values between 0.5 and 0.7 as translating “poor” predictions, values between 0.7 and 0.9 as “useful” prediction and values above 0.9 as good predictions. Araujo *et al.* (2005) proposed a refined scale with AUC>0.9 being “excellent”; 0.8<AUC<0.9 being “good”; 0.7<AUC<0.8 “fair”; 0.6<AUC<0.7 “poor”; 0.5<AUC<0.6 “fail” and AUC<0.5 being “counter-predictions” (similar to negative correlation between observations and predictions). Most models built with background data are also evaluated with standard statistics (e.g. AUC), however, the use of pseudo-absences renders model evaluation even more difficult than model fitting because the probabilities predicted have no absolute value *per se*. Only occurrence data can be used in this case to evaluate calibration. If it is used to assess discrimination, the evaluation can only be partial because it can only properly assess the model’s ability to predict presences.

To deal with this issue, some approaches have been proposed for evaluating presence-only predictions that do not require the selection of a single threshold, such as the Boyce index (Boyce *et al.*, 2002). This method splits the model predictions into regular bins or classes and then assesses the proportion of presences actually found within each bin compared to the proportion of modelling cells in the same bin. A model that adequately predicts the distribution of a given species should predict large numbers of presences in the high prediction bins (i.e. high proportion of presences with high values of habitat suitability) and fewer and fewer presences as one moves toward the lower prediction bins (i.e. toward low habitat suitability for the species). It takes a value between -1 and +1, with a value tending toward -1 indicating predictions no different from those obtained by chance, and values toward -1 indicating counter-predictions (Hirzel *et al.*, 2006).

Some measures of fit ignore the geographical pattern of the predictions, especially those based on confusion matrices, and do not identify where and how the errors occur, either spatially or environmentally. When assessing SDMs predictions therefore have to be ‘realistic’ at each geographical location in the study area or may simply be content with similar levels of realism at all geographical locations sharing some common set of environmental attributes (Barry & Elith, 2006). The question of environmental availability is important because it can affect the way that key assumptions (such as niche conservatism/stability) are evaluated, and also how some models are parameterised. The question of environmental analogy is different, though complementary. It relates to the identification of environments that exist in one area/period but not in the other. In recent years, several tests have been developed to compare different environments, such as the MESS (Elith *et al.*, 2010). Assessing environmental availability and analogy together is important because it can affect the quantification of the realized niche in each area or time period and thus affect the way models are built, compared and projected between areas and time periods (Guisan *et al.*, 2014).

Finally, it is good practice to consider more than one evaluation metric for comparing model predictions with observations, in order to provide information in different aspects of the model's predictive power (Guisan *et al.*, 2017). To obtain a real evaluation of the model, it is essential to use an independent dataset to assess it, both in environmental and geographical space.

2.3.4. SDMs transferability

Model generality is the ability of a model to predict events outside the modelled system via projection or transfer to a different resolution, geographic location, or time period (Araújo *et al.*, 2019). To assess how predictive a model is, it has to be tested against independent data, which is critical when forecasting the effects of climate change on biodiversity and considered the most robust type of evaluation by ecological modellers (Fielding & Bell, 1997; Araujo *et al.*, 2005). The definition of the independence of datasets is a crucial aspect. When running an independent evaluation in the same area and period, one of the potential problems is that the training and test datasets are not spatially independent. In turn, the spatial structure of the environmental predictors can be confidently transferred from one situation to the other in this evaluation.

If no independent set of data is used, the results will not guarantee that the model can be applied to another area (Randin *et al.*, 2006) or time (Araujo *et al.*, 2005) where the spatial co-variation between predictors is different or has changed, such as when attempting to anticipate biological invasions (Thuiller *et al.*, 2005) or the impact of climate change on biodiversity. In such cases, the model's transferability to these distinct conditions implies an extrapolation, which is statistically challenging, as it forces the algorithm to make predictions for novel environmental conditions, which may be erroneous (Elith & Leathwick, 2009).

The first important consideration when projecting models in time or space is therefore to assess whether the same environmental conditions prevail in both areas or periods (training and projection). If independent data for other period or time are available, testing the model's transferability to novel conditions should ideally be tested (Randin *et al.*, 2006; Tuanmu *et al.*, 2011; Petitpierre *et al.*, 2017).

2.3.5. SDMs ensemble

The ensemble technique has been used in recent years to eliminate (or at least limit) model selection bias (Araujo & New, 2007; Marmion *et al.*, 2009). The rationale behind using and ensembling several models is that two or more models may have very similar predictive performance even when they contain different environmental predictors and/or yield vastly different spatial predictions, making it difficult to know which of the equivalent candidate models to use. Furthermore, the “best” model may not necessarily be the best one for predictions in a different area or under new conditions, or some models may be more sensitive (than one or other models) to sampling bias, which might also reduce model transferability (Randin *et al.*, 2006). Ensemble modelling is particularly powerful in these situations, as it maps both the main trend (i.e. mean, median, percentile) and the overall variation (and thus uncertainty) across all models. Various metrics have been proposed for use as criteria for selecting the models to be considered and to weight them, such as the AUC of a ROC plot (Swets, 1988) or the TSS (Allouche *et al.*, 2006).

In ecology, the question of how to best build multi-model ensembles from different approaches is still wide open.

To obtain an accurate assessment of climate change effects over marine species, the definition of predictors from reliable sources and explanatory of the ecology of species and the selection of modelling algorithms is crucial to ensure their transferability. Model evaluation, both in geographical and environmental space, is a key aspect in determining the uncertainty of the final predictions.

2. Objectives of the thesis

The overall objective of this thesis was the risk assessment of macroalgae distribution due to climate change in Europe. The following specific objectives had to be achieved to reach this aim:

- Creation of a database of variables linked with species distribution, to allow the historical and future variability of meteo-oceanographic conditions in Europe to be studied.
- Identification of key species of macroalgae from intertidal and subtidal assemblages threated by climate change in Europe.
- Development of a temporally transferable methodology to assess the effects of climate change over macroalgae, using species distribution modelling.
- Assessment of climate change driven range shifts of representative seaweeds in Europe.

3. Layout of the thesis

The thesis is organised as follows:

First, the motivations for the research, the state of knowledge and the objectives or the thesis are presented in Chapter I. As the study area is common for all chapters; it has been described at the end of the first.

The following chapters (II, III, IV and V) address the objectives of the thesis. Each chapter includes a brief introduction and sections for methodology, results and discussion, in some cases merged, constituting edited and modified versions of different articles published in or submitted to SCI journals.

Finally, general conclusions and future research lines are described in Chapter VI, followed by the references cited in the thesis and the Annexes.

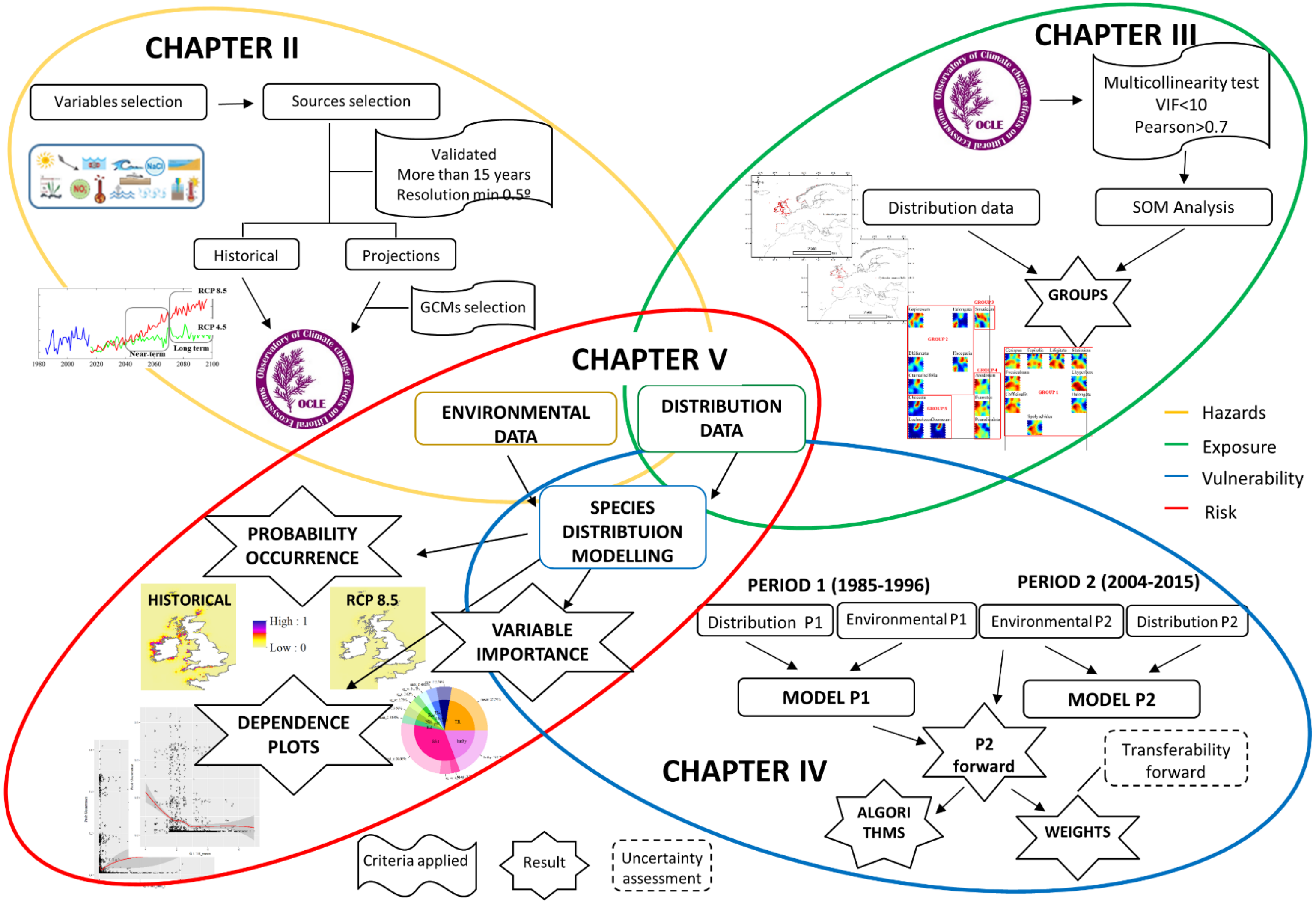


Figure 1.2. Graphical summary of the thesis.

4. Study area

This thesis covers all European seas, which includes a wide range of marine and coastal ecosystems, ranging from the stable environment of the deep ocean to highly dynamic coastal waters. These ecosystems provide a home for up to 48 000 species (EEA, 2015), in enclosed seas as the Mediterranean or the Black Sea and open areas as the Atlantic region. Europe is thus an interesting geographical area in which to assess the effects of climate change over macroalgae.

To characterise European seas a mesh have been defined at two different resolutions: 0.1º for coastal waters (until 50 m depth), as an "equally-stratified" sampling strategy to better characterise the potential habitat of coastal ecosystems (Cefalì *et al.*, 2018); and 0.5º for oceanic waters. Due to the influence of the Red Sea in the Mediterranean, especially for invasive species, it has been included in the mesh. Additionally, the northern coasts of Africa are also part of the mesh to allow models to be trained in areas with different environmental conditions.

The final mesh had 20800 points, or "virtual sensors" (14049 coastal and 6831 offshore) (Figure 1.3).

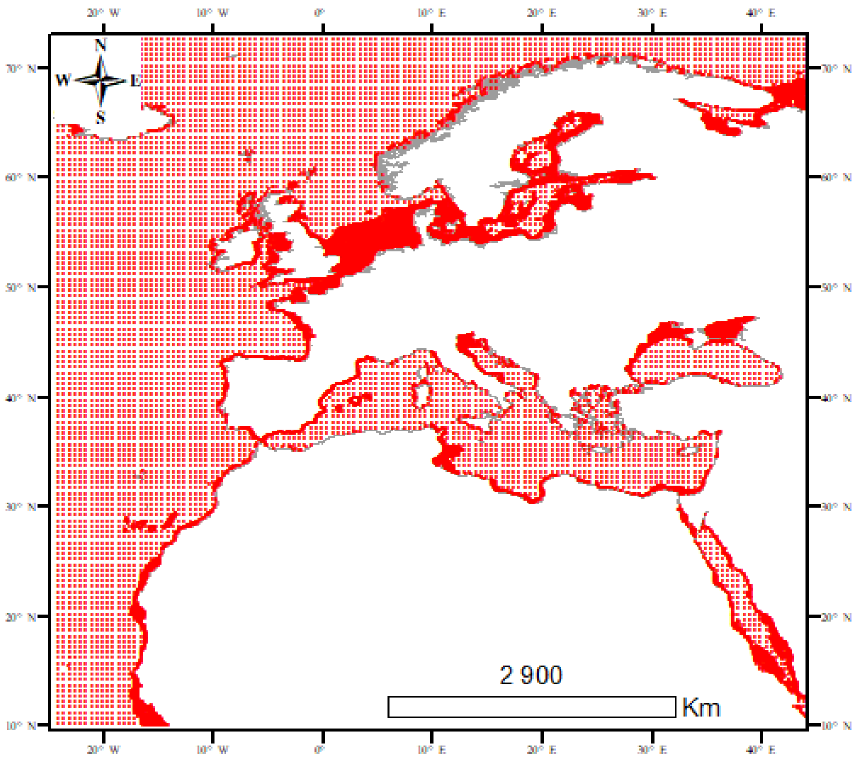


Figure 1.3. Study area showing the grid squares that define the sampling units.

Chapter II

OCLE: a European Open access database on Climate change effects on Littoral and oceanic Ecosystems

Chapter II. OCLE: a European Open access database on Climate change effects on Littoral and oceanic Ecosystems

This chapter is an edited and modified version of the research article published in 2018 in the journal Progress in Oceanography, vol. 168, pp. 222–231, by de la Hoz, C.F.; Ramos, E.; Acevedo, A;, Puente, A.; Losada, I.J. and Juanes, J.A. with the title "Ocle: a European Open Access Database on Climate Change Effects on Littoral and Oceanic Ecosystems".

1. Introduction

Studies on historical and future distribution of marine species are frequently limited by the lack of relevant data on abiotic components (IPCC, 2014), especially when working over large areas (Robinson *et al.*, 2017). Marine existing databases (such as Bio-Oracle (Tyberghein *et al.*, 2012; Assis *et al.*, 2017b), MARSPEC (Sbrocco & Barber, 2013), Aquamaps (Ready *et al.*, 2010) or the MERRAclim (Vega *et al.*, 2017)) contributed remarkably to the advance in global studies. However, it is necessary to address the detected gaps (inclusion of hydrodynamic variables, lack of homogeneity in time intervals, ecological reliability of parameters and projected to climate change scenarios considering the best information available (RCPs)) in order to meet current and future needs for species distribution studies. Homogeneous and complete high-resolution data, integrated at different time scales, ecological-sounded parameters, based on abiotic conditions that determine the ecology of the species of interest have to be included. Additionally, raw data have to be controlled and homogenised to guarantee the quality of the derived products. Concerning temporal periods, different resolutions should be available to allow researchers to define specific parameters for each species. Besides, data should fit to the spatial scale of the work, covering the study area with the necessary detail. Finally, the access to the data have to be free and very intuitive for users, reducing to the maximum the weight and the computing resources used for getting the information.

Trying to comply with these requirements and using the best data available, to our best knowledge, this study presents the open access database on climate change effects on littoral and oceanic ecosystems (OCLE), an ecological-driven database of present and future hazards for marine life in Europe. With the development of this database the first specific objective of this thesis is reached. As a first step the database is oriented toward seagrasses and algae, due to their key role in the food chain of marine ecosystems, contributing to the maintenance of biodiversity and providing ecosystem services (Duarte *et al.*, 2013; Ondiviela *et al.*, 2014; Mazarrasa *et al.*, 2017). However, the aim of OCLE is to provide researchers with open access accurate information for marine studies, not only for coastal studies, but also for oceanic waters.

2. Material and methods

2.1. Variables and parameters

The variables included in OCLE were first selected because of their functional relationship with seagrasses and macroalgae distributions. Those variables with a heterogeneous distribution in space and/or time were discarded. General meteo-oceanographic variables (hereinafter referred to as met-ocean variables) were considered first, including different physical and chemical factors, such as temperature (Fralick *et al.*, 1990; Valle *et al.*, 2014), light (Best *et al.*, 2001; Larkum *et al.*, 2006; Riis *et al.*, 2012), salinity (Touchette, 2007; Nejrup & Pedersen, 2008) or nutrients (Hughes *et al.*, 2004; Martínez *et al.*, 2012b). Those were complemented with other variables related to the stressful conditions that limit intertidal organisms distributions, such as desiccation, a decisive survival factor characterized by the tidal range (Pearson *et al.*, 2009), the wind speed (Lipkin *et al.*, 1993), the significant wave height (Ondiviela *et al.*, 2014; Jensen & Denny, 2015; Jones *et al.*, 2015; Quintano *et al.*, 2015) and sea level (Short & Neckles, 1999), especially under future scenarios. A final group of variables regarding exposure of subtidal species to uprooting conditions was also taken into account. Stress to high energy conditions is characterized by the bottom orbital speed (Young *et al.*, 2015), the currents speed (Fonseca & Fisher, 1986; Infantes *et al.*, 2011) and the, significantly more complex variable, bottom shear stress (Pace *et al.*, 2017).

For each variable, a complete set of parameters was selected in order to reflect in a more holistic perspective the state of the environment, as a proxy of ecological processes. For historical data, the maximum, minimum, mean, standard deviation, range and percentiles 10, 25, 50, 75 and 90 were calculated at each virtual sensor, for seasonal (winter was defined as January–March, spring as April–June, summer as July–September and autumn as October–December (as per Zavatarelli *et al.* (2000)), monthly, yearly, five-yearly and full (1985-2015) periods (Figure 2.1). Besides, according to the more detailed information available and their close relationship to macrophytes distributions, some specific and relevant parameters to detect changes in extreme conditions of sea and air temperatures (i.e. number of consecutive days over the percentile 90 (Torresan *et al.*, 2016) and for the shear stress (i.e. number of days over 2.2 Nt/m^2 (Vousdoukas *et al.*, 2012)) were calculated (Figure 2.1). Furthermore, for future projections, the same group of parameters were calculated on a seasonal, yearly and full period, considering both the mid-term (2040-2069) and the long term (2070-2099) for two RCPs, namely RCP 4.5 and RCP 8.5.

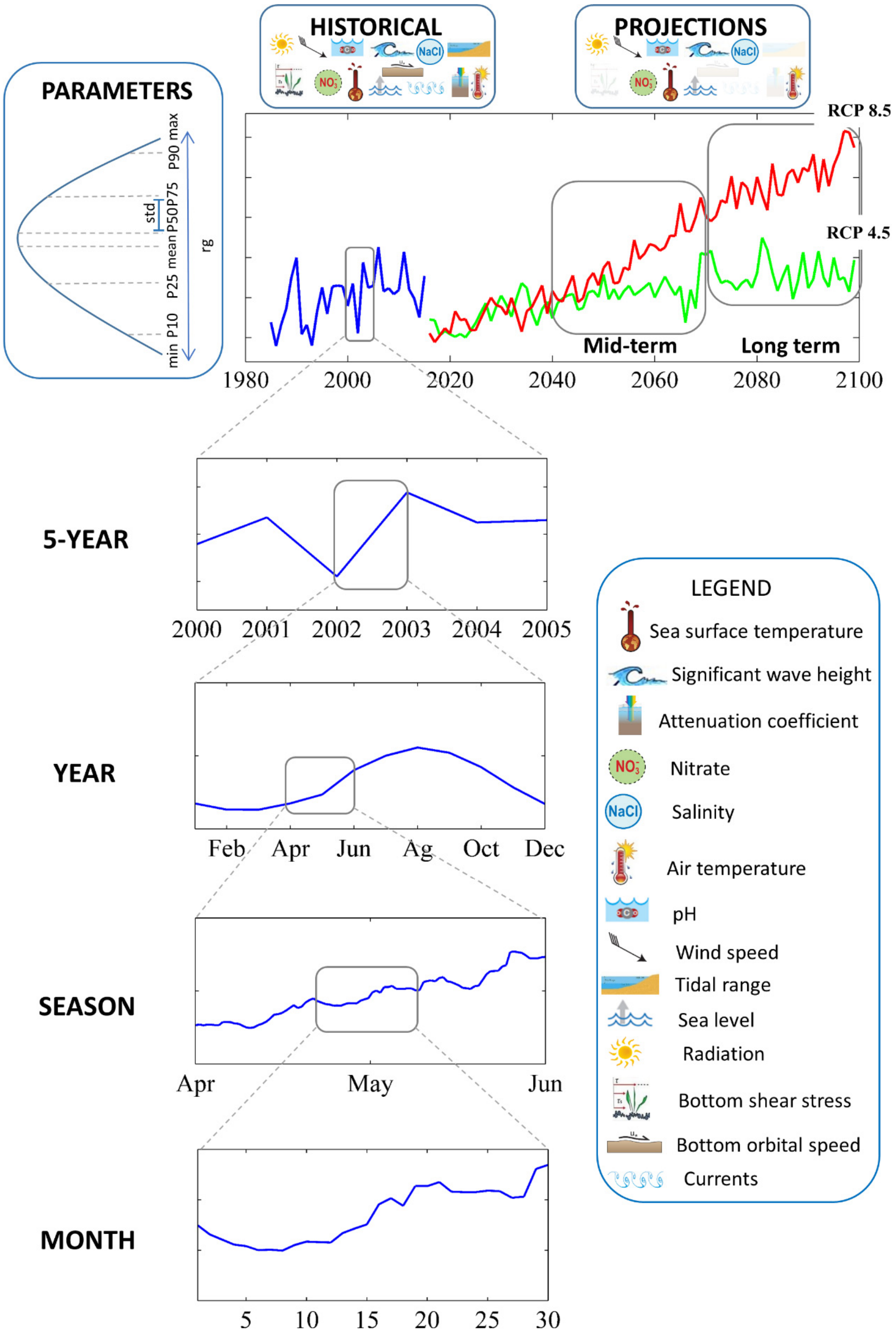


Figure 2.1. Parameters calculated and their temporal resolution. max, maximum; min, minimum; P, percentile; std, standard deviation; rg, range. Symbols for diagrams courtesy of the Integration and Application Network (ian.umces.edu/symbols).

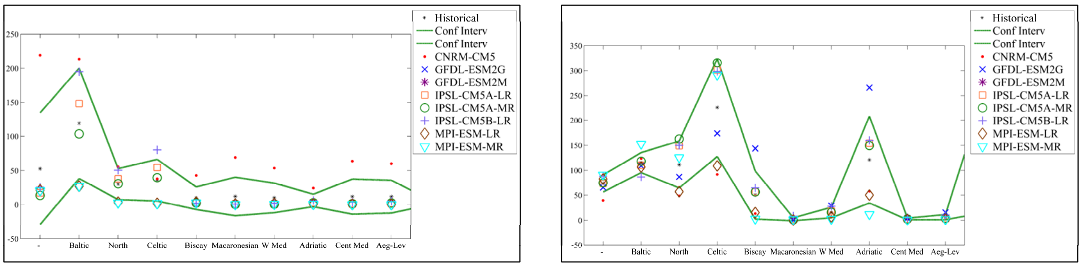
2.2. Data sources and methods

Historical data were compiled from satellite (Schuckmann *et al.*, 2016), reanalysis (Stark *et al.*, 2007; Saha *et al.*, 2010; Donlon *et al.*, 2012; Reguero *et al.*, 2012; Cid *et al.*, 2014; Perez *et al.*, 2017) and *in situ* measurements (Weatherall *et al.*, 2015).

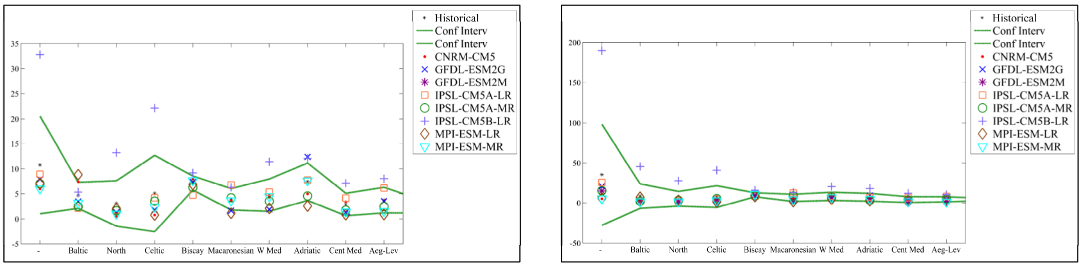
A quality control was established along all steps. First, only validated sources were selected, either with instrumental data (Saha *et al.*, 2010, 2014; Perruche *et al.*, 2015; Garnesson *et al.*, 2016; Schuckmann *et al.*, 2016), remotely sensed information (Donlon *et al.*, 2012) or both of them (Cid *et al.*, 2014; Perez *et al.*, 2017). To get a temporal and spatially homogeneous database, only sources with time series longer than 15 years and a spatial resolution lower than 0.5º were taken into account. Final selected data were compared with existing studies (e.g. Rhein *et al.* (2013); EEA (2009); Collins *et al.* (2013)).

For future projections data from the fifth phase of CMIP5 (Taylor *et al.*, 2012) were used. CMIP5 provides results of a set of coordinated climatic model experiments using GCMs at two times scales, a mid-term (2040-2069) and a long-term period (2070-2099) for different RCPs. This experimental set has been selected for its high skill to represent projections at the North-East Atlantic Region (Perez *et al.*, 2014) and because it is the reference set provided by the IPCC for climate research and impact and risk assessment. Quality assurance and control procedures for projected data were based on MSE between the historical data series selected and those of the GCMs for the reference period (1985-2005). This analysis was carried out for each of the Marine Strategy Framework Directive region (European Commission, 2008) to avoid a bias by local processes. Outliers (GCMs with more than 20% of their values out of the limits within the $MSE_{mean} \pm MSE_{std}$) were discarded (Chai & Draxler, 2014). More detailed information is available in Figure 2.2.

NITRATE **SALINITY**



SEA SURFACE TEMPERATURE **AIR TEMPERATURE**



WIND SPEED **SIGNIFICANT WAVE HEIGHT**

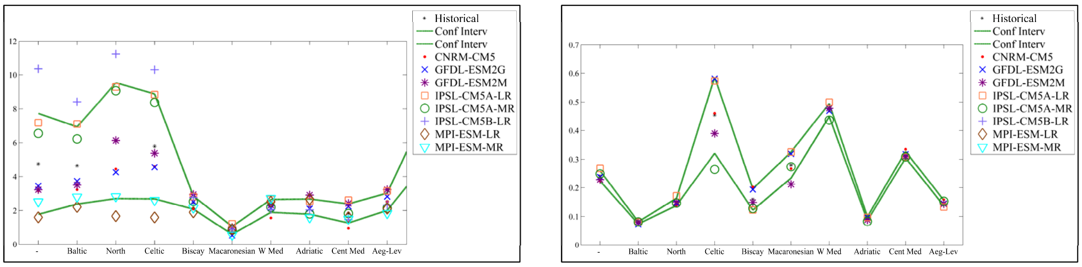


Figure 2.2. Mean squared errors (MSE) values by Marine Strategy Framework Directive region for the Global Circulation Models considered. In the x-axis, from left to right: Baltic, the Baltic Sea; North, the Greater North Sea; Celtic, the Celtic Seas; Biscay, the Bay of Biscay and the Iberian Coast; Macaronesian, the Macaronesian biogeographic region; W Med, the Western Mediterranean Sea; Adriatic, the Adriatic Sea; Cent Med, the Ionian Sea and the Central Mediterranean Sea; Aeg-Lev, the Aegean-Levantine Sea. Information about Global Circulation Models is detailed in Table 2.2.

The final data sources are shown in Table 2.1. Historical and projected periods are shown for each variable and final GCMs selected are named. More detailed information about original data sources and GCMs is available in Table 2.2.

Table 2.1. Variables selected with indication of periods, sources and the applied method to data gathering. For projections, the GCMs used are specified.

VARIABLE	Period	Method	Source
Sea Surface Temperature (SST) (ºC)	01/01/1985 – 31/12/2015	Reanalysis	OSTIA dataset (NASA)
	01/01/2010 – 31/12/2099	Projections CNRM-CM5, GFDL-ESM2G, IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Significant wave height (Hs) (m)	01/01/1985 – 31/12/2015	Reanalysis	GOW (IH Cantabria)
	01/01/2010 – 31/12/2099	Projections GFDL-ESM2G, GFDL-ESM2M, IPSL-CM5A-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Bathymetry (m)	-	Satellite and *in situ* measurements	GEBCO 2014 (BODC).
Light attenuation coefficient (Kd) (m^{-1})	25/01/1998– 27/12/2015	Satellite measurements	Copernicus Marine System (ESA)
Substrate	-	Renalysis and *in situ* measurements	EMODNET EUSeaMap
Nitrate (mol/m^3)	16/01/1998– 16/12/2014	Reanalysis	Copernicus Marine System (ESA)
	15/01/2010– 15/12/2099	Projections IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Salinity (psu)	01/01/1985 – 31/12/2015	Reanalysis	CFSR
	15/01/2010– 15/12/2099	Projections IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR	CMIP5
Air Temperature (ºC)	01/01/1985 – 31/12/2015	Reanalysis	CFSR
	01/01/2010 – 31/12/2099	Projections CNRM-CM5, GFDL-ESM2G, GFDL-ESM2M, IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
pH	13/01/1985 – 8/11/2005	Reanalysis	CMIP5
	16/01/2010– 16/12/2099	Projections IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR	CMIP5
Wind speed (m/s)	**01/01/1985 – 31/12/2015**	**Reanalysis**	**CFSR**
	01/01/2010 – 31/12/2099	Projections CNRM-CM5, GFDL-ESM2G, GFDL-ESM2M, IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-MR	CMIP5

Tidal range (m)	01/01/1985 – 31/12/2013	Reanalysis	GOST (IH Cantabria)
Sea level rise (m)	-	-	-
	01/01/2010 – 31/12/2099	Projections (IPCC 2014)	(Slangen et al., 2014)
Radiation (W/m^2)	01/01/1985 – 31/12/2005	Projections GFDL-ESM2G, IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
	01/01/2010 – 31/12/2099	Projections GFDL-ESM2G, IPSL-CM5A-LR, IPSL-CM5A-MR, IPSL-CM5B-LR, MPI-ESM-LR, MPI-ESM-MR	CMIP5
Bottom shear stress (N/m^2)	01/01/1985 – 31/12/2013	Reanalysis	Own development
Bottom orbital speed (m/s)	01/01/1985 – 31/12/2013	Reanalysis	GOW (IH Cantabria)
Currents (m/s)	01/01/1985 – 31/12/2013	Reanalysis	GOST (IH Cantabria)

OSTIA, Operational Sea surface Temperature and sea-Ice concentration Analysis; NASA, National Aeronautics and Space Administration; CMIP5, World Climate Research Programme; GOW, Global Ocean Wave; GEBCO, General Bathymetric Chart of the Oceans; BODC, British Oceanographic Data Centre; ESA, European Space Agency; CFSR, NCEP Climate Forecast System Reanalysis; GOST, Global Ocean Surges Tides. All databases are at a spatial resolution of 0.1º and 0.5º, according to the defined mesh (Figure 1.3).

For each variable, data available from the original sources closest to the defined virtual sensors were selected. Bottom shear stress calculations were based on hourly waves and currents data, obtained from GOW (Perez *et al.*, 2017) and GOST (Cid *et al.*, 2014) databases (Table 2.1), applying the formulation of Soulsby (Soulsby, 1997). The bed roughness was derived from the substrate type (standardized in five typologies: mud, sand, coarse, mixed and rock), according to Soulsby (1983). This formulation was selected because it has demonstrated good results in other studies (Antunes do Carmo *et al.*, 2003; Tomás *et al.*, 2012; Roulund *et al.*, 2016). All variables were temporally homogenised through the compilation of raw to daily, when possible, or monthly data. For projections, parameters were calculated for each GCM, RCP and period, and averaged with the ensemble method (Arnell *et al.*, 2014). For sea level projections, the IPCC values were considered for the whole period for each RCP (Slangen *et al.*, 2014). Analyses were conducted using Climate Data Operators (CDO 1.7), NetCDF Operators (NCO 4.4.5), Matlab 8.1 and ArcGis 10.1.

Table 2.2. Global Circulation Models information, the developer institution, the atmospheric resolution and links to more specific information.

Model	Institution	Country	Atmospheric resolution (lat x lon, number of layers)
CNRM-CM5	Centre National de Recherches Météorologiques	France	1°L42
GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory	USA	1°L63
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory	USA	1°L50
IPSL-CM5A-LR	Institute Pierre Simon Laplace (IPSL)	France	1.9° x 3.75° (96 x 96 L 39)
IPSL-CM5A-MR	Institute Pierre Simon Laplace (IPSL)	France	1.25° x 2.5° (143 x 144 L 39)
IPSL-CM5B-LR	Institute Pierre Simon Laplace (IPSL)	France	1.9° x 3.75° (96 x 96 L 39)
MPI-ESM-LR	Max Planck Institute for Meteorology (MPI-M)	Germany	1,5°L40
MPI-ESM-MR	Max Planck Institute for Meteorology (MPI-M)	Germany	0,4°L40

3. Results and discussion

3.1. OCLE data

According to the aforementioned detected gaps, OCLE represents a step further in the capacity to characterize marine and coastal systems from an integrated temporal and spatial perspective. OCLE provides homogeneous and open access accurate information of 16 variables and 12 derived parameters for historical and projected periods, which completes the existing databases. Facing requirements for making predictions based on retrospective analysis of species distributions at different scales, this work has considered four fundamental aspects: hydrodynamic characterization, spatial and temporal dimensions, biologically-meaningful parameters and reliable climate change projections.

3.1.1 Hydrodynamic characterization

This is a key aspect because this type of variables are usually omitted in other databases, despite their ecological importance for sessile species (Callaghan *et al.*, 2015; Ramos *et* *al.*, 2016b; Pace *et al.*, 2017; de la Hoz *et al.*, 2018a). Furthermore, their temporal and spatial variability along the European seas constitutes a critical element to be considered for a holistic physical characterization. Consequently, it is essential to include variables that allow a complementary interpretation of the physical resistance of species, such as waves, currents and bottom orbital speed, both included in OCLE (Figure 2.3).

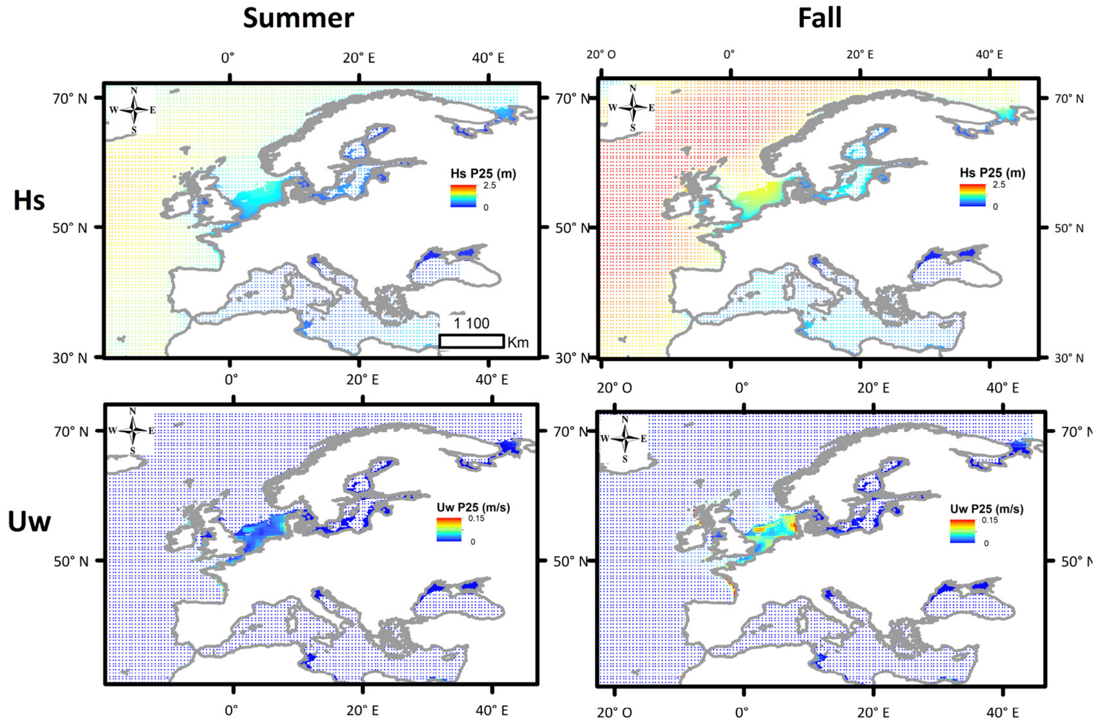


Figure 2.3. Percentile 25 of significant wave height (Hs, upper panel) and currents (Uw, bottom panel), for summer (left) and fall (right) seasons.

One important contribution of OCLE is the development of a derived variable, the bottom shear stress, which allows detecting the areas where the energy of the system is higher and, consequently, the stress for benthic organisms (Pace *et al.*, 2017). Although regional studies of bottom shear stress have been carried out by other authors (Tomás *et al.*, 2012; Dalyander *et al.*, 2013; Rengstorf *et al.*, 2014; Alekseenko *et al.*, 2017), a broad scale characterization along Europe has not been developed so far. This gap has been solved in OCLE providing average and extreme parameters of bottom shear stress. The historical distribution of the winter 90th percentile reflects the spatial differences when considering potential bottom shear stress impacts on marine flora (Figure 2.4), as previously demonstrated in the North Sea, the west coasts of Ireland or the Gulf of Gabès (Schanz & Asmus, 2003; Ben Brahim *et al.*, 2015; Kregting *et al.*, 2016).

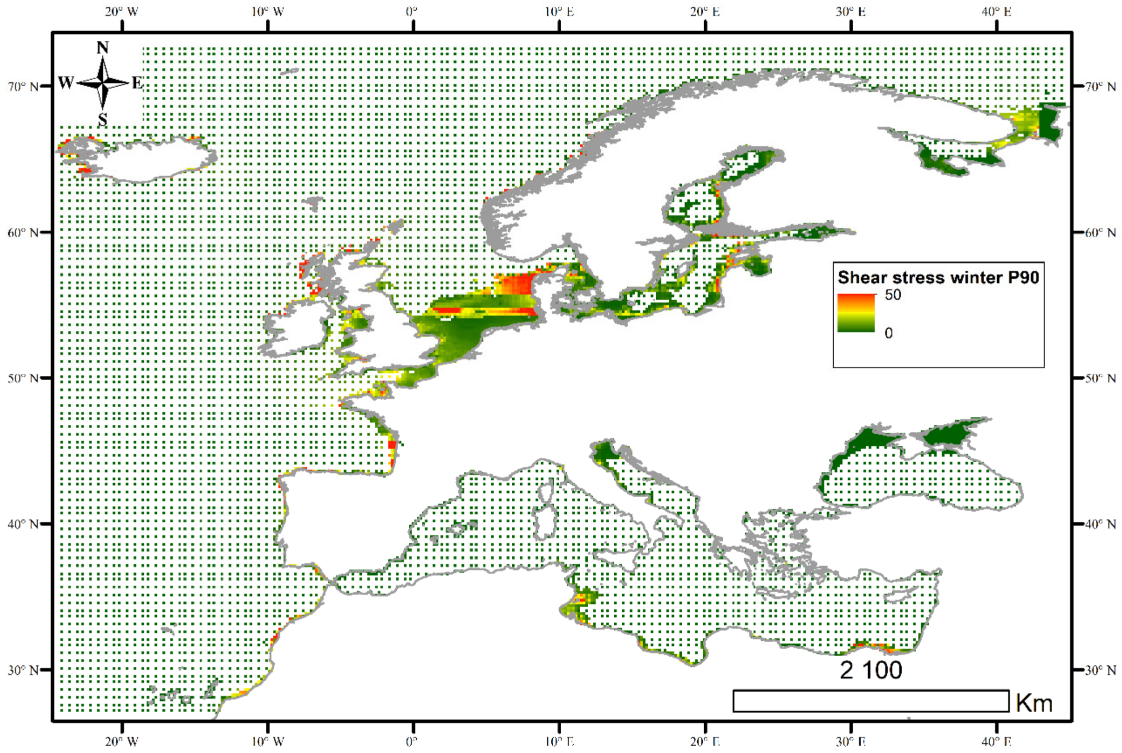


Figure 2.4. Shear stress percentile 90 average values of all winters of the historical period (N/m^2).

3.1.2 Spatial and temporal dimensions

One of the first requirements in the design of an end-users oriented database is the integration of their needs. Most databases available provide a unique value for the whole period considered, which largely limits the type of hypothesis posed concerning prospective and retrospective trend analysis on species distribution (Thurstan *et al.*, 2015). OCLE offers data with a higher resolution from daily or monthly to full period considered (Figure 2.1). This allows detecting not only average environmental conditions, but also extreme conditions, which affect many species responses (Galván *et al.*, 2016).

The interannual variability results crucial because species can respond to yearly episodes, most of them lost when using long-term averaged values. Similarly, intra-annual variability provides a huge potential for testing research questions regarding ecological responses of marine species versus historical and projected patterns in the abiotic environment linked to different natural or anthropogenic scenarios. For instance, seasonal differences in the potential influence of river (Spillman *et al.*, 2007), upwelling zones (Johns *et al.*, 1992; Sarhan *et al.*, 2000; Albérola & Millot, 2003), the entrance of other waters (Omstedt, 1987; Poulos & Collins, 1997) and the inputs in the North Sea due to the industrial activity on nitrate discharges (Burson *et al.*, 2016) may be estimated from Figure 2.5 at the European scale.

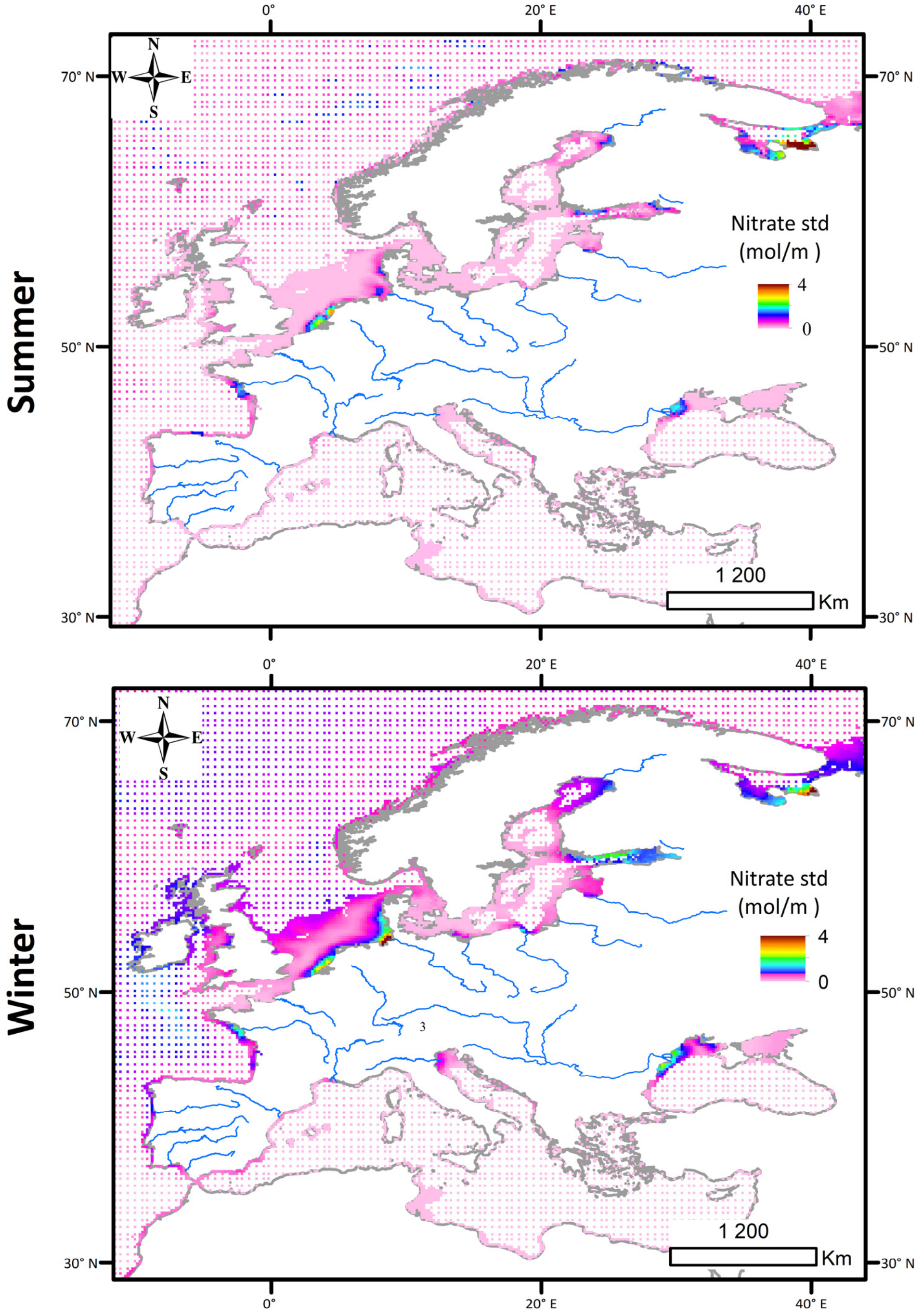


Figure 2.5. Nitrate standard deviation of the year 2000 in summer (upper) and winter (bottom). In blue, main European rivers.

Two spatial resolutions are available in OCLE in order to adapt the analysis scale and computing resources to the end-users needs. This allows addressing from general trend studies, at a spatial resolution of 0.5º (e.g. Figure 2.3) to detailed studies, at 0.1º resolution (Figure 2.6). In general terms, this dual approach will facilitate more precise studies at the coastal zone, where the influence on macrophytes-based communities occur; whereas, a balance between a lower spatial resolution and a much more detailed abiotic information compared to previous works is available for oceanic areas. At the same time, OCLE opens the possibility for very specific studies by downscaling the variables at even higher resolutions (Galván *et al.*, 2016; Tomás *et al.*, 2016).

Finally, an important issue to ensure the potential applications of the database is the temporal and spatial homogeneity among variables. Some existing databases combine variables with diverse periods, due to the difficulty to collect temporal homogeneous data. However, a huge effort has been developed in this work to provide temporally homogeneous information for all variables (within the period 1985-2015), allowing the comparison among them (Table 2.1).

3.1.3 Biologically-meaningful parameters

Abiotic conditions that control the settlement, survival and reproduction of marine species are key factors that may determine, together with biological interactions, the species distribution (Araújo & Guisan, 2006). Currently, studies on specific thresholds for different physical and chemical variables affecting functional processes of macrophytes are mostly addressed in laboratory or *in situ* experiments. Their application to field conditions and under usually more complex environments may limit their transferability (Valiela, 2001). Variables and parameters established in OCLE have been selected to cover some ecological processes mainly related to the macroalgae and seagrasses distributions, for both average and extreme abiotic conditions. The later ones are crucial for the survival of those species which are living in their limits of distribution (Araújo *et al.*, 2016). In spite of the important advantages of OCLE, the involvement of the scientific community in the development of more specific parameters (Bosch *et al.*, 2017) and the definition of other variables related to species distributions at different scales is necessary (Martínez *et al.*, 2012b; Ramos *et al.*, 2016c). For example, according to the threshold for bottom shear stress proposed by Vousdoukas *et al.* (2012), an analysis regarding the distribution of macrophytic communities through time (1985-2014) in response to the estimated increase in the stress in the intertidal zones is carried out along the coasts of Denmark (Figure 2.6).

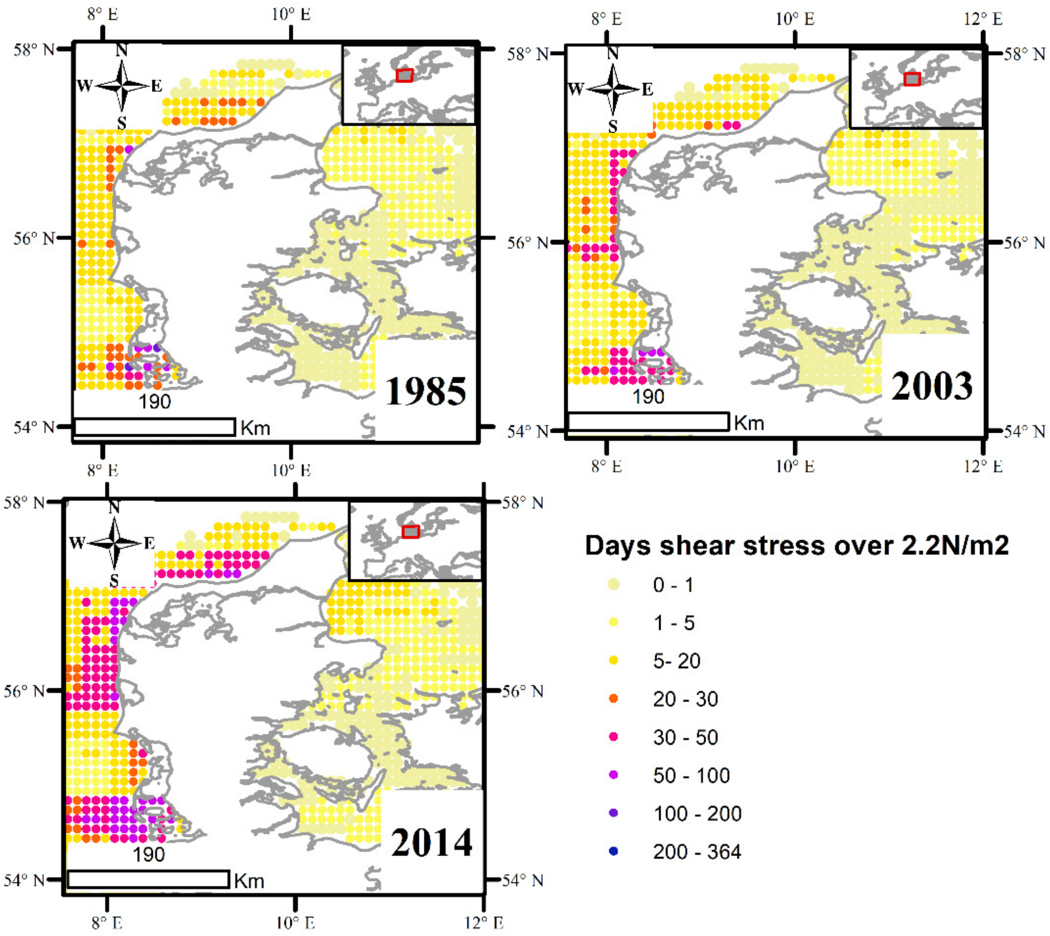


Figure 2.6. Number of consecutive days in each year (1985 (top left), 2003 (top right) and 2014 (bottom left)) with bottom shear stress values over 2.2 N/m^2.

To arise the definition of these thresholds, the availability of a broad set of spatial and temporal information on biological-sounded parameters at the right scale is crucial. This information can be used in SDMs for the statistical determination of limiting factors and the establishment of parameters response curves. This approach may be used as a proxy to understand the present and predict future distributions.

3.1.4 Reliable climate change projections

Concerning projections, OCLE does include, to our knowledge, the best information available (IPCC, 2014). Two future scenarios (RCP 4.5, RCP 8.5) were considered for the mid (2040-2069) and long term (2070-2099), for eight of the variables (Table 2.1). Besides, the bias of the GCMs has been reduced thanks of the ensemble technique that has been applied in the quality control process (Meier *et al.*, 2011; Camus *et al.*, 2017). Results are in agreement with IPCC predictions, as the averaged increment of 2.7ºC of sea surface temperature by 2090 in the RCP 8.5, with higher values in the Mediterranean and enclosed seas than Atlantic (EEA, 2009; Collins *et al.*, 2013) (Figure 2.7).

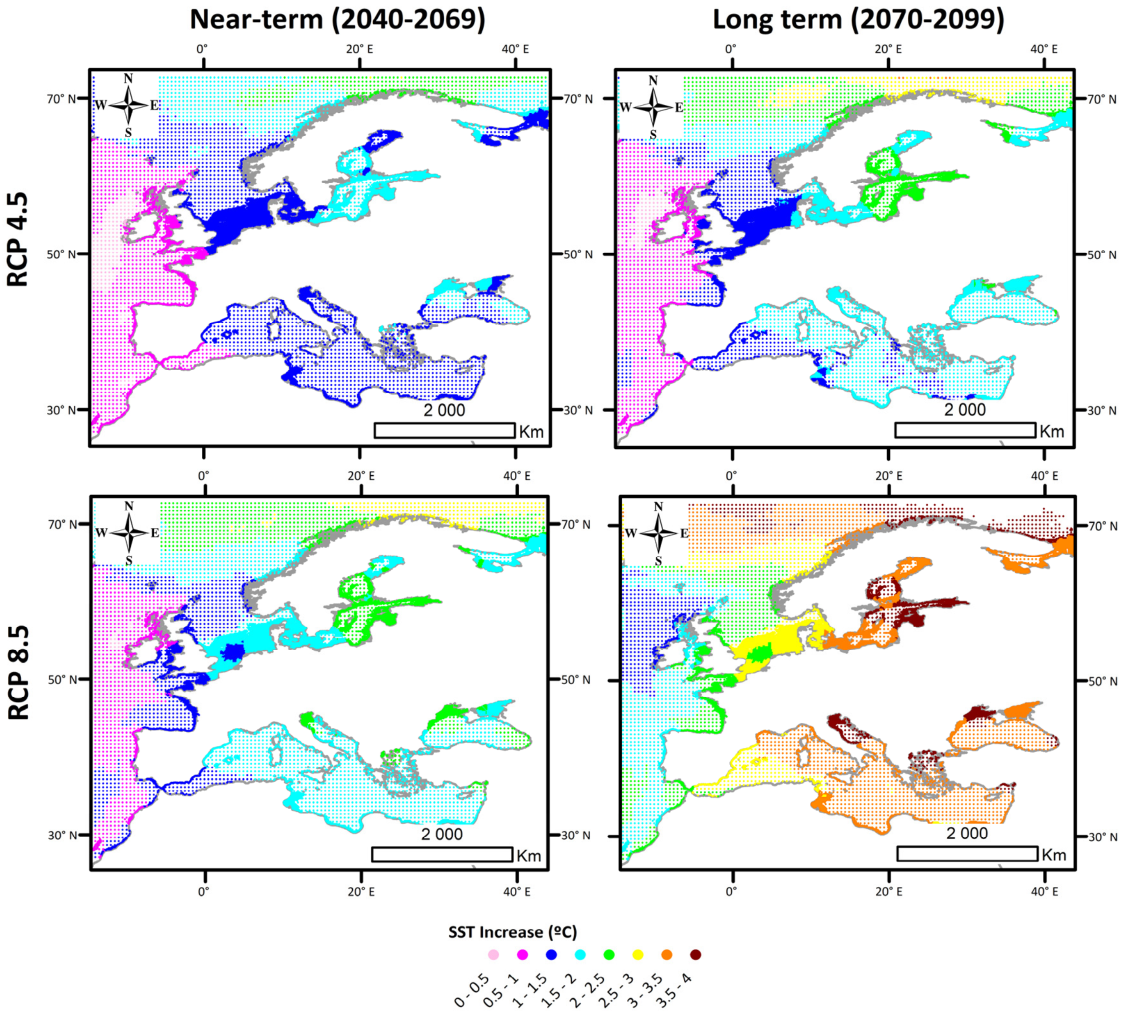
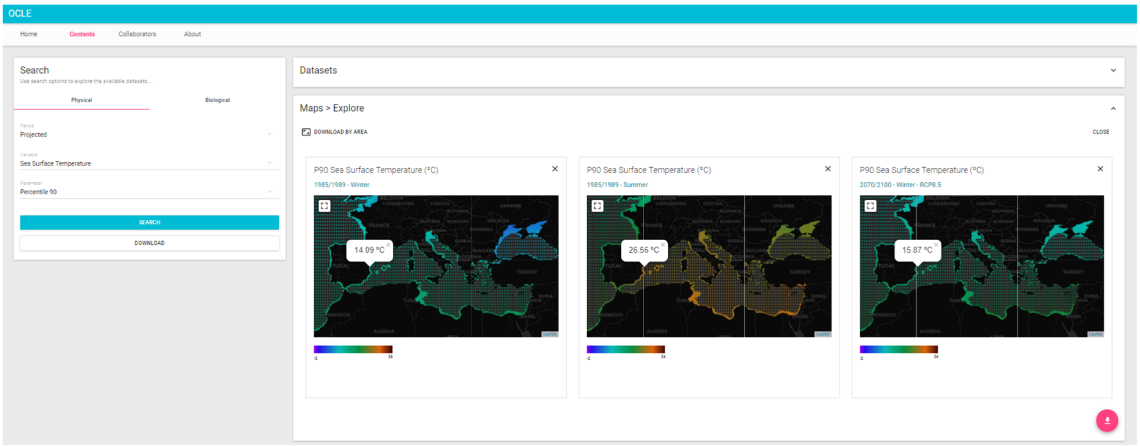


Figure 2.7. SST increase with respect to the reference period for the RCP 4.5 (upper panel) and RCP 8.5 (lower panel) for the mid-term (left) and long term (right).

3.2. OCLE website

OCLE results are available for free and stepwise download at http://ocle.ihcantabria.com, to allow users' choice of the most appropriate data for their research needs, regarding the period (historical or projected), the variables (16 options) and parameters of interest (12 choices). The full historical records (1985-2015) are split into five-yearly datasets.

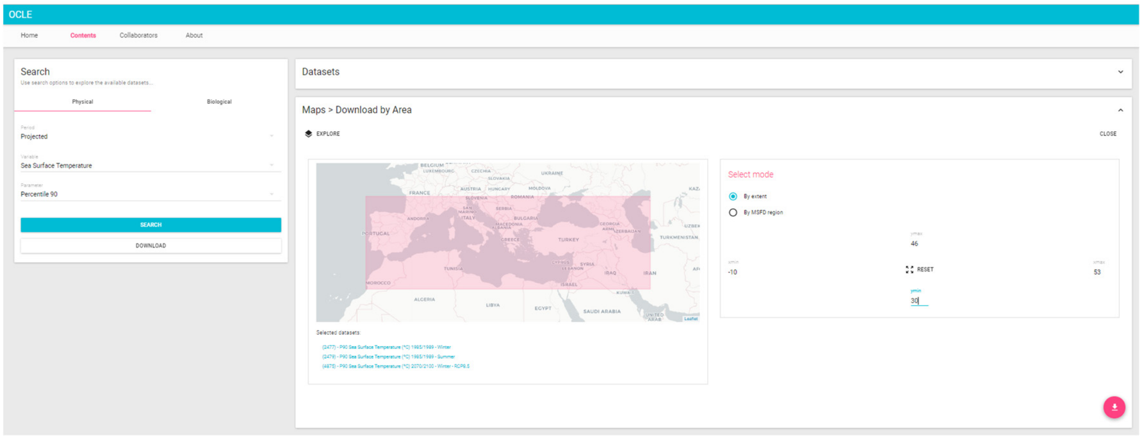
Concerning projections, data for the mid-term (2040-2069) and the long term (2070-2099) are available for each RCP considered. Researchers can select the datasets of interest and explore the data on a map (Figure 2.8) or downloading the information in a .csv format to be used in SDMs or to be visualized in geographic information systems.



Period: Projected
Variable: Sea Surface Temperature
Parameter: Percentile 90

Figure 2.8. OCLE website interface to explore data available.

To avoid computational overcharges, OCLE allows the spatial filtering, by zone selection over the map, coordinates screening, coastal areas (until 50 m depth) or predefined regions according to the European Commission (2008) (Figure 2.9). A user manual is available on the website and as Annex B.



Period: Projected
Variable: Sea Surface Temperature
Parameter: Percentile 90

● By extent
○ By MSFD region

ymax: 46
xmin: -10
xmax: 53
ymin: 30

Selected datasets:
(2477) - P90 Sea Surface Temperature (°C) 1985/1989 - Winter
(2479) - P90 Sea Surface Temperature (°C) 1985/1989 - Summer
(4875) - P90 Sea Surface Temperature (°C) 2070/2100 - Winter - RCP8.5

Figure 2.9. OCLE website interface to download the data by selecting the area of interest.

Additionally, it is possible to access to yearly data and customized parameters on request through the website. This opens up the possibility to a broad field of work generating more specific parameters.

The development of OCLE database constitutes a very useful tool for marine researchers as it provides the most complete set of predictors, temporally disaggregated, for historical and climate change conditions in Europe.

Chapter III

Identification and classification of key seaweeds species in Europe

Chapter III. Identification and classification of key seaweeds species in Europe

This chapter is part of an edited and modified version of the research article submitted to the journal Ecological Indicators, by de la Hoz, C.F., Ramos, E, Puente, A. and Juanes, J.A. with the title "Temporal transferability of marine distribution models: the role of algorithm selection".

1. Introduction

Understanding wide spatial distributions of marine species is essential for their conservation and management. However, in the marine area cartographies are biased because of logistical complications and economic constraints of broad scale subtidal sampling (Bekkby *et al.*, 2009; Belando *et al.*, 2015). These limitations have increased the interest and reliance on SDMs, as a tool for marine resources management (Robinson *et al.*, 2017). SDMs are based on a quantification of species-environment relationships by using numerical tools (Guisan & Zimmermann, 2000; Elith & Leathwick, 2009), and thus they represent a cost-effective way to provide efficient large-scale mapping where field-collected data are limited or currently unavailable (Gorman *et al.*, 2013; Álvarez-Martínez *et al.*, 2018).

Additionally, SDMs allow to quantifying the environmental niche of species and its changes in time and space, therefore they are widely used to assess the impacts of climate change on species distributions (Buonomo *et al.*, 2018; Martínez *et al.*, 2018b).

Modelling many species is time-consuming, as adjusting model settings is required according to their different relationships with environment (Phillips & Dudík, 2008). To deal with this, different classification techniques are available (Pearson, 1901), such as k-means (Hartigan & Wong, 1979) or SOM (Kohonen, 1998). Although they have been used extensively with environmental data (Peterson *et al.*, 2011; Galván *et al.*, 2016; Guisan *et al.*, 2017) and their power to classify distribution data according to ecological requirements have been proven (Ramos *et al.*, 2017), it has not been conducted a generalization in large marine areas, which might be a very useful tool for working with extensive number of species.

Bearing in mind this need, the overall objective of this chapter is to standardize a technique to group species according to their environmental requirements to reach the specific objective 2 of this thesis. As a case study, seaweeds have been selected as their distribution is strongly affected by environmental factors (Lüning *et al.*, 1990), they are ecological engineers of the whole marine system and they provide a wide number of ecosystem services (Liquete *et al.*, 2016).

2. Material and methods

2.1. Environmental data

Environmental variables were selected according to their relationship with seaweed ecology (Martínez *et al.*, 2012a; Young *et al.*, 2015; Cefalì *et al.*, 2016). Information for a complete set of parameters was collected from 1985 to 2015 from the OCLE database (de la Hoz *et al.*, 2018b) for each mesh point, including maximum value, minimum value, mean value, standard deviation, range and percentiles 10th, 25th, 50th, 75th and 90th for each variable (Table 2.1, Chapter II). Additionally, the number of consecutive days that percentile 90th of sea surface and air temperature is exceeded (Torresan *et al.*, 2016) and the shear stress harmful threshold 2.2 Nt/m^2 is surpassed (Vousdoukas *et al.*, 2012) were calculated.

For each point, parameters were calculated seasonally and for the whole period 1985-2015, creating a pull of 642 predictors (including bathymetry and substrate) for the analyses. To avoid multicollinearity, tests were conducted with the criteria of a Pearson correlation coefficient < 0.7 (Green, 1979; Suárez-Seoane *et al.*, 2018; Fabri-Ruiz *et al.*, 2019; Rodríguez *et al.*, 2019) and a VIF < 10 (Chatterjee & Hadi, 2006; Guisan *et al.*, 2017; Mpakairi *et al.*, 2017), using the R package *usdm* (Naimi *et al.*, 2014). VIF estimates the severity of the effect of multicollinearity, by measuring the extent to which variance increases in a regression due to collinearity compared to when uncorrelated variables are used (Montgomery & Peck, 1982). VIF test are especially recommended when numerous variables are added to a regression, as it detects the variables' linear correlation structure. Usually values from five to ten are considered as critical for multi-variable correlation (Guisan *et al.*, 2017). Thus, 32 predictors remained (Table 3.1). The substrate categories (mud, sand, coarse substrate, mixed, rock) were considered as different parameters, therefore a total of 36 predictors were finally included.

Table 3.1. Variables and their parameters remained after collinearity test.

VARIABLE	PARAMETER	VARIABLE	PARAMETER
AT	Days	SST	Days
	P50		std_w
	std_w	Sal	std_sp
Kd	P25_w	pH	P50_w
	rg_s		std_f
	rg_w		std_s
	std_f		std_sp
	std_sp		std_w
Nit	min_f	Hs	std_s
	std_s	Tm	Days
	std_sp	TR	P75
	std_w	Uw	P75
Rss	rg_f	Ws	min_sp
	std_s		rg_f
	std_sp	Bathymetry	-
	std_w	Substrate	-

To extract environmental data patterns a SOM analysis (Kohonen, 1998) was conducted. The optimal map size (number of units) was chosen based on the heuristic formula proposed by Vesanto et al. (2000), $M = 5\sqrt{N}$, where M is the number of map units an N is the number of points in the mesh. As N=843, thus M=145.17 and the SOM size selected was 144. As a result, data were transformed from the high-dimensional space into a bidimensional lattice, allowing an intuitive visualization of predictor patterns (Figure 3.1).

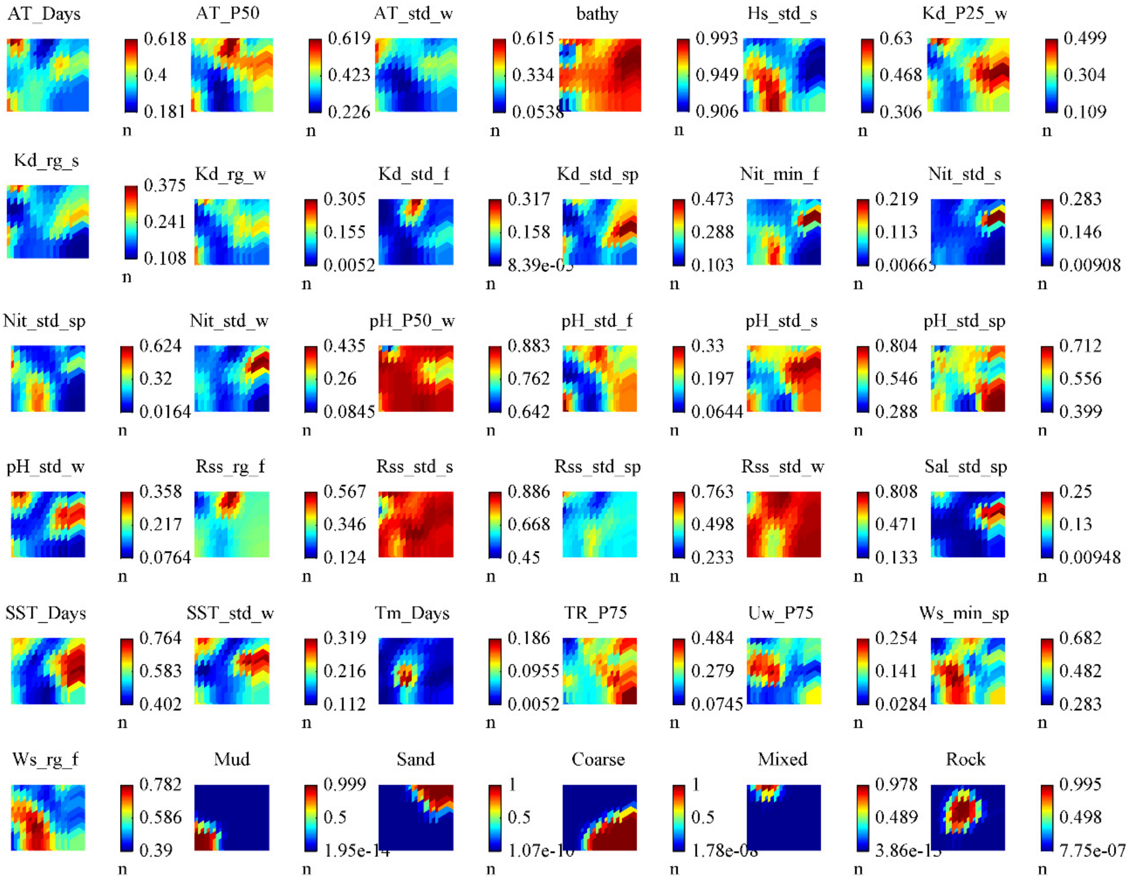


Figure 3.1. Visualization of predictors in a coloured scale on the SOM.

2.2. Distribution data

We selected 21 intertidal and subtidal macroalgae from different biogeographic regions around the European Regional Seas, one of them being the invasive *Sargassum muticum* (Table 3.2).

Table 3.2. List of species considered, with reference to their general area of distribution, preferred tidal level and the number of records.

SPECIES	DISTRIBUTION	TIDAL LEVEL	NUMBER OF RECORDS
Laminaria hyperborea	North Atlantic	Subtidal	507
Laminaria digitata	North Atlantic	Subtidal	473
Laminaria ochroleuca	Atlantic	Subtidal	37
Saccorhiza polyschides	Atlantic and Mediterranean Sea	Subtidal	299
Saccharina latissima	Atlantic	Low intertidal to subtidal	465
Himanthalia elongata	Atlantic	Low intertidal to subtidal	231
Ascophyllum nodosum	North Atlantic	Intertidal	482
Pelvetia canaliculata	Atlantic	Upper intertidal	449
Fucus spiralis	Atlantic and West Mediterranean	Upper intertidal	451
Fucus vesiculosus	Atlantic	Intertidal	409
Bifurcaria bifurcata	South Atlantic	Lower intertidal	130
Halopteris scoparia	Atlantic and Mediterranean Sea	Upper intertidal to subtidal	53
Sargassum muticum	Atlantic	Lower intertidal and upper subtidal	242
Fucus serratus	North Atlantic	Middle intertidal	536
Cystoseira baccata	Atlantic	Lower intertidal	70
Cystoseira tamariscifolia	Atlantic and Mediterranean Sea	Middle intertidal to shallow subtidal	115
Chondrus crispus	Atlantic and Mediterranean Sea	Subtidal	412
Corallina officinalis	Atlantic	Lower intertidal to subtidal	486
Ellisolandia elongata	Atlantic	Lower intertidal to subtidal	33
Gelidium corneum	South Atlantic	Subtidal	29
Gelidium spinosum	Atlantic and Mediterranean Sea	Lower intertidal	170

Biological records with sampling time information were collected from 1985 to 2015 from the sources of Table 3.3.

Table 3.3. List of sources of occurrence data, including the year of the data collection and the citation of the original studies.

SOURCE	**YEARS**
Guinda, Juanes, Puente, & Echavarri-Erasun (2012)	2005
IH Cantabria	2015
Juanes & Gutiérrez (1992) and Juanes, Gutierrez, Puente (1991)	1990, 1991
Guinda, Juanes, Puente, & Revilla (2008)	2006, 2007, 2008
IH Cantabria	2009, 2010, 2015, 2016
IH Cantabria	2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016
Ramos, Puente, & Juanes (2016)	2011
Ramos *et al.*,(2017)	2012
IHCantabria (2008)	1992, 1993, 1994, 1998, 2003, 2004
UC (2006)	2003, 2004, 2005, 2006
Ramos et al. (2014)	2004-2011
OBIS (2015)	1985-2015
GBIF (2013)	1985-2015

All occurrences were double-checked and corrected for referencing errors and misplacement. Species bias was reduced by getting only one record by pixel (Peterson *et al.*, 2011). The minimum distance between pixels were 0.1º, which guaranteed that there was not spatial autocorrelation (Jonsson *et al.*, 2018). Final records considered can be geographically visualized in the Figures 3.2 – 3.5.

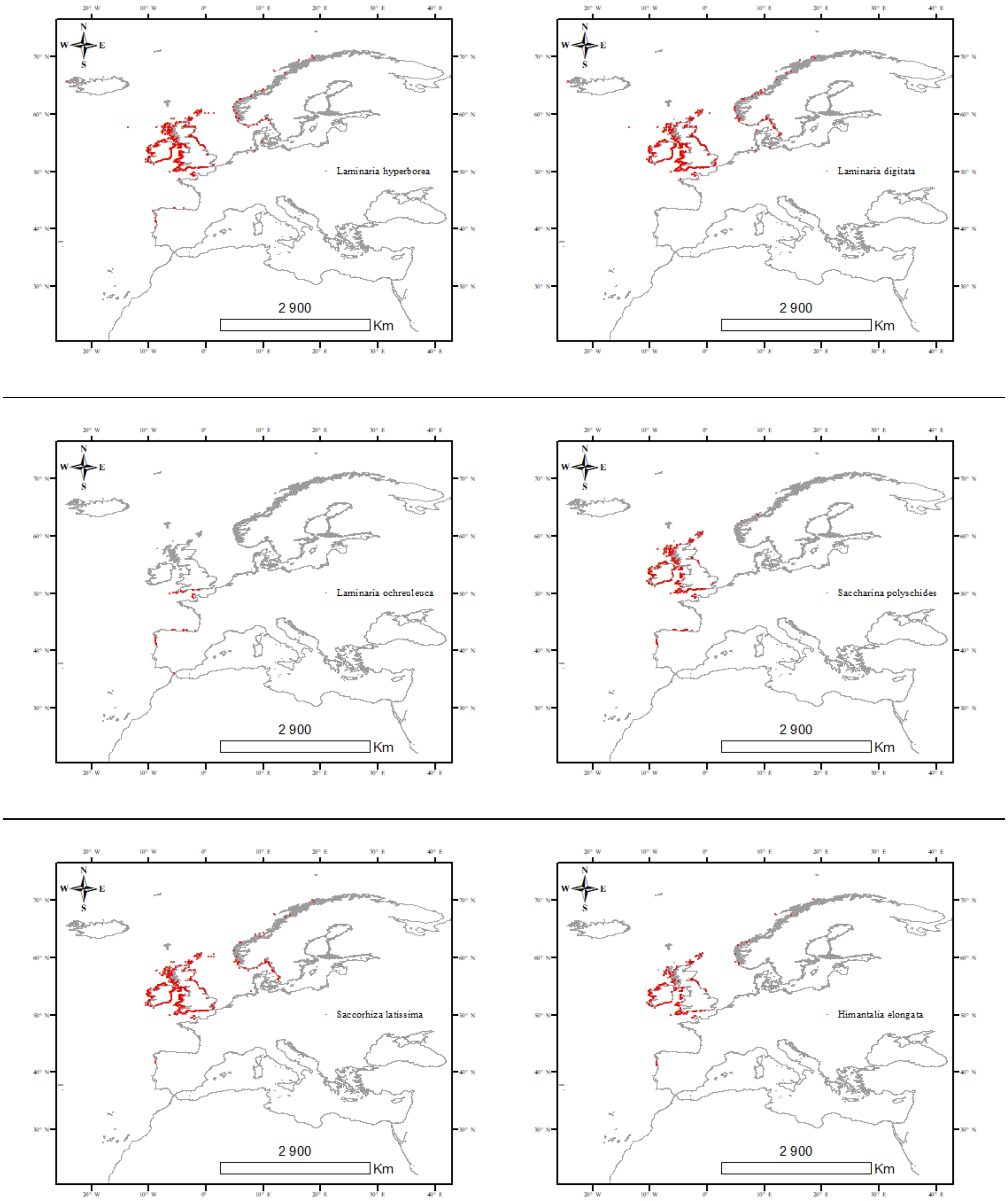


Figure 3.2. Geographic distribution of records for the species considered. From top left: *Laminaria hyperborea, Laminaria digitata, Laminaria ochroleuca, Saccorhiza polyschides, Saccharina latissima, Himantalia elongata*.

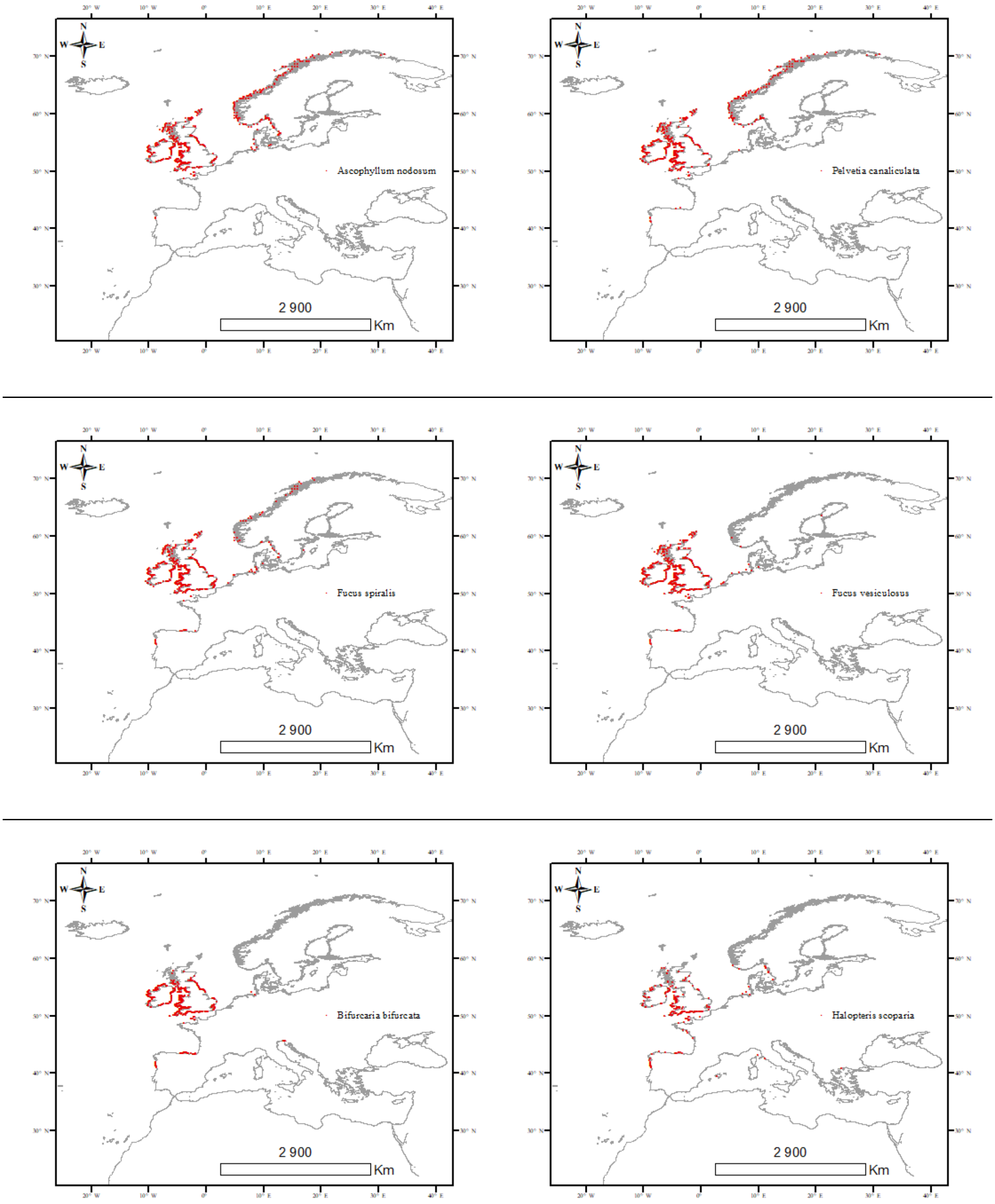


Figure 3.3. Geographic distribution of records for the species considered. From top left: *Ascophyllum nodosum, Pelvetia canaliculata, Fucus spiralis, Fucus vesiculosus, Bifurcaria bifurcata, Halopteris scoparia*.

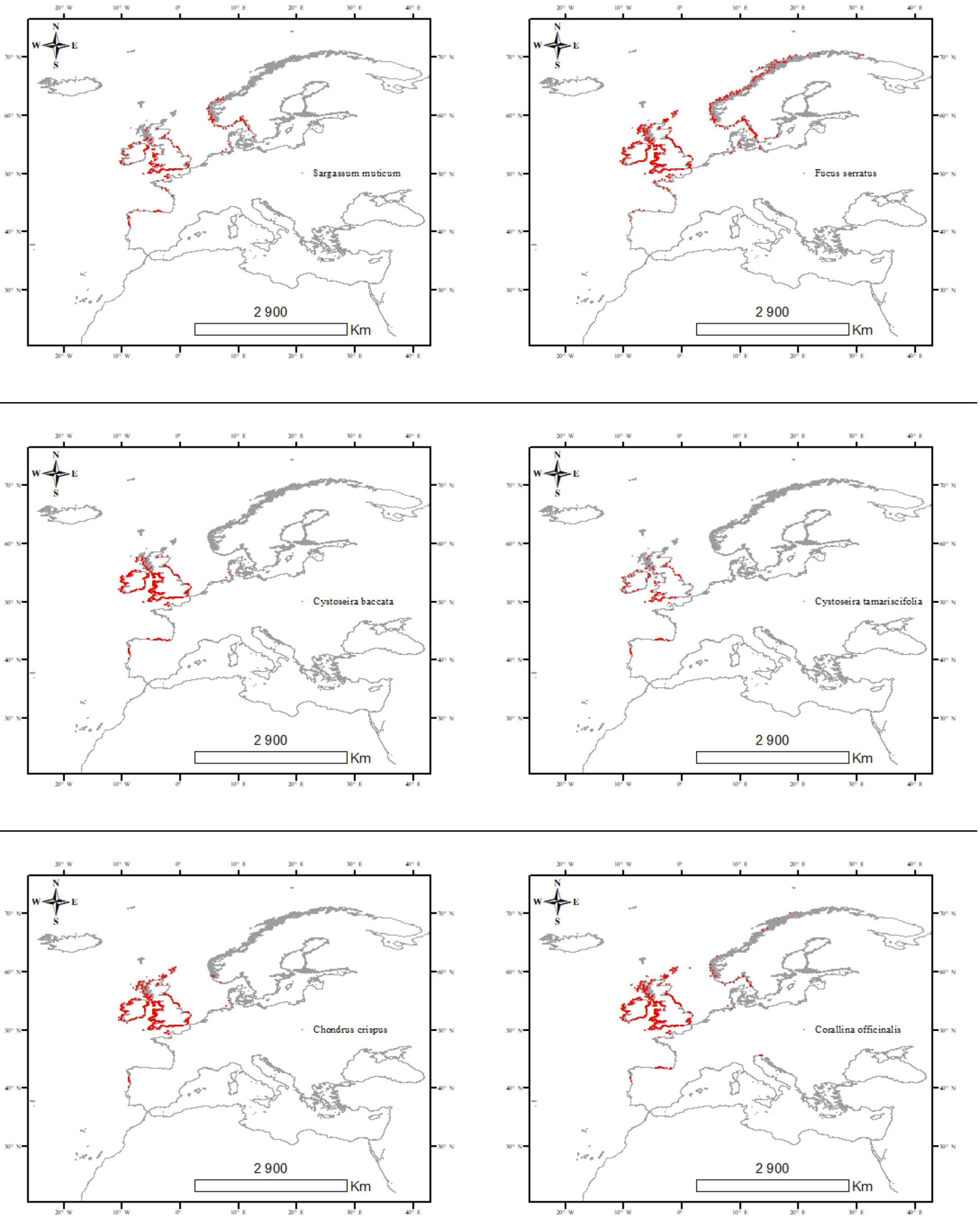


Figure 3.4. Geographic distribution of records for the species considered. From top left: *Sargassum muticum, Fucus serratus, Cystoseira baccata, Cystoseira tamariscifolia*, *Chondrus crispus, Corallina officinalis*.

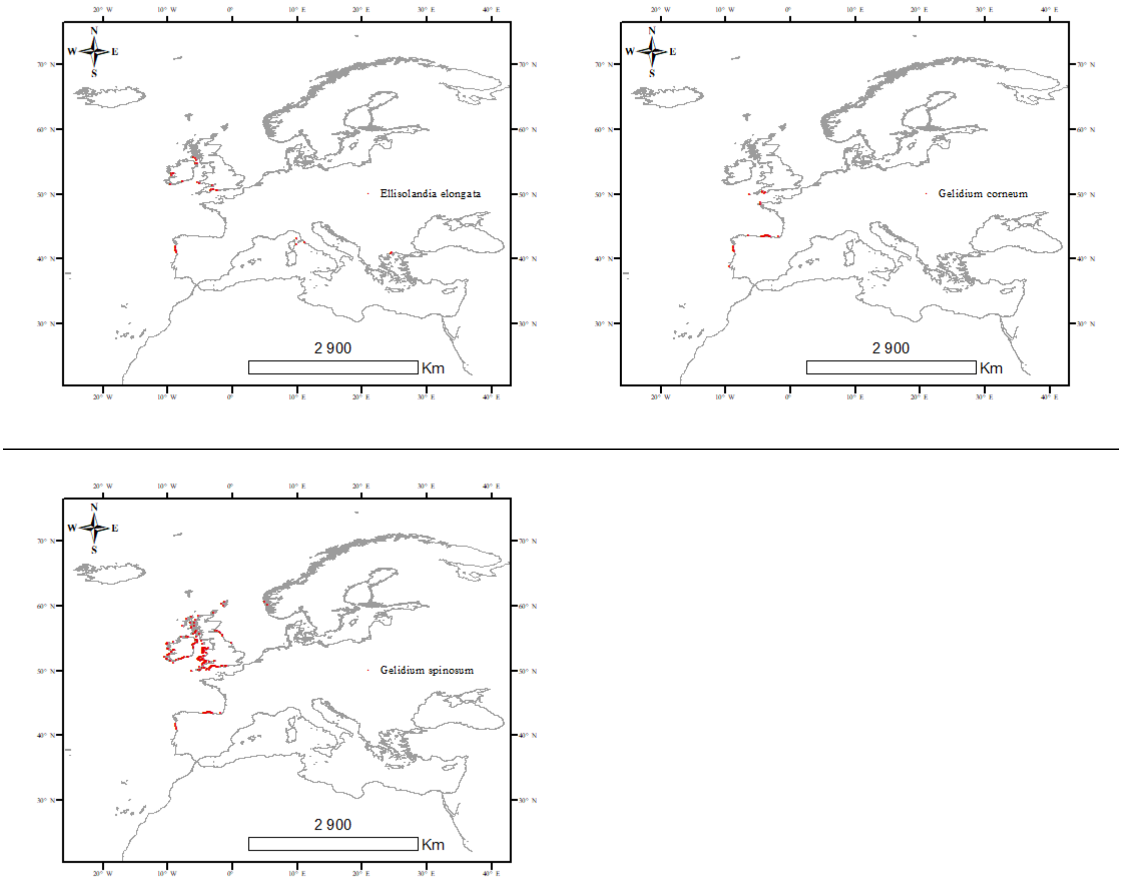


Figure 3.5. Geographic distribution of records for the species considered. From top left: *Ellisolandia elongata, Gelidium corneum, Gelidium spinosum*.

To facilitate interpretation of the results, similarities in the species ecological response were assessed by grouping them according to their environmental requirements. This was undertaken through projecting species presences over the previously obtained SOM (Figure 3.1) using component planes (Vesanto, 1999).

3. Results and discussion

More similar species were represented closer; therefore, five groups were defined by expert criteria visual assessment of their patterns (Figure 3.6).

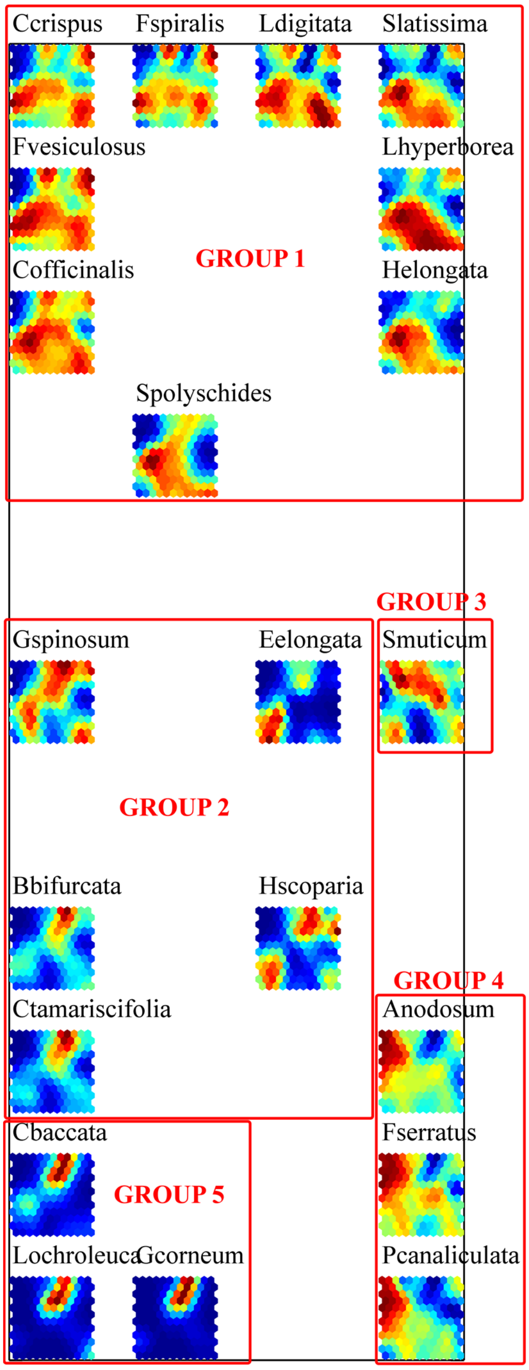


Figure 3.6. Component planes ordering the taxa on the previously trained SOM (Figure 3.1) visualized in a coloured scale.

The first Group was characterized for cold-water species that can live in exposed areas (Norton, 1977; Sjotun & Fredriksen, 1995; Fernandez, 2011; Takolander *et al.*, 2017), therefore species high values are located in the lowest areas of the number of days over SST but they include the highest values of the number of days over Tm threshold.

In the Group 2, all species are located in the mid-intertidal to subtidal and related to Mediterranean or South Atlantic conditions, such as stable SST or low values of nitrate (Fredriksen & Rueness, 1989; Rueness & Fredriksen, 1989; Celis-Pla *et al.*, 2014).

The pattern of *Sargassum muticum* is far from the other species, which justifies its isolation in the Group 3. This is likely due to its invasive condition in Europe (Stæhr *et al.*, 2000; Sánchez *et al.*, 2005), therefore it is probably not in equilibrium with the environment (Araújo & Pearson, 2005).

The species included in the Group 4 are found high on the shore (Anadón, 1983; Little & Kitching, 1996), therefore their distribution is probably more influenced by climatic factors related to emersion times, such as the percentile 50 of AT or the Ws (Duarte Mora, 2015).

The Group 5 comprises mainly subtidal warm species that live in rocky substrate and sheltered areas (Salinas & Granado, 1984; Juanes & Gutiérrez, 1992; Smale *et al.*, 2015).

A representative species of each group was selected to accomplish the recommendations of having 5-10 presences for each predictor (Guisan & Zimmermann, 2000; Franklin, 2009; Araújo & Peterson, 2012), therefore 180-360 records. When no species accomplished these criteria, the one closer to limits was selected. The final species were:

1. *Saccorhiza polyschides*

   This kelp is the main annual forest forming species in the North-East Atlantic (Smale & Wernberg, 2013). It usually occupies rocky substrate up to 35 m deep (Norton, 1970) and its growth starts in spring and maximum size is reached in the summer (Pereira, 2014). It is used commercially as fertiliser and as an alginate source.

2. *Gelidium spinosum*

   G. spinosum can reach 20 m deep (Santelices, 1974) on rocky substrate or pools, especially in exposed coasts between the maximum low tide in neap tides and in spring tides (Rico, 1992). The highest productivity has been reported during autumn and winter (Juanes & Fernández, 1988). Its use to produce agar is an important income for some regions.

3. *Sargassum muticum*

   The brown algae *S. muticum* is often found in shallow waters, in the lower intertidal and upper subtidal zones. Its distribution is strongly depth-dependent, with a maximum distribution at intermediate depths of 2–4 m (Thomsen *et al.*, 2006) and controlled by the availability of hard substrate (Stæhr *et al.*, 2000). Its introduction was most likely due to shipment of Japanese oyster and it constitutes one of the most conspicuous changes in macroalgal vegetation in recent years. *S. muticum* is firmly established as a major colonizer of lower intertidal and shallow subtidal regions from the western coast of Norway to Portugal (Steen, 2004). Due to its invasive behaviour it competes with other macroalgae species for both substrate and light and may cause variations in species abundance or even prevent their recruitment (Stæhr *et al.*, 2000; Sánchez *et al.*, 2005).

4. *Pelvetia canaliculata*

   P. canaliculata is a perennial, hermaphroditic fucoid that forms, together with several other kelp/fucoid seaweeds, an endemic structural assemblage that typifies the cold-temperate shores of the North-East Atlantic (Lüning *et al.*, 1990). It lives in the high intertidal, on sheltered to semi-exposed rocky shores. Its commercial interest is linked to its use as fertiliser and to produce flour for the cattle.

5. *Cystoseira baccata*

   This perennial brown alga is present in the North-East Atlantic, from United Kingdom to the North of Spain (Cabioc’h *et al.*, 1995), preferably in areas with high radiation and low hydrodynamic energy (Salinas & Granado, 1984). It is used commercially as fertiliser and as an alginate source.

The proposed classification is able to select representative species, reducing the modelling effort when working with a large number of species.

Chapter IV

Temporal transferability of marine distribution models: the role of algorithm selection

Chapter IV. Temporal transferability of marine distribution models: the role of algorithm selection

This chapter is part of an edited and modified version of the research article submitted to the journal Ecological Indicators, by de la Hoz, C.F., Ramos, E, Puente, A. and Juanes, J.A. with the title "Temporal transferability of marine distribution models: the role of algorithm selection".

1. Introduction

Understanding wide spatial distributions of marine species is essential for their conservation and management. However, compared to terrestrial environments, marine species distributions are often inadequately characterized because of logistical complications and economic constraints of broad scale subtidal sampling (Bekkby *et al.*, 2009; Belando *et al.*, 2015) which results in lower spatial resolution occurrence records biased towards the coast and economically important areas (Robinson *et al.*, 2011). These limitations have increased the interest and reliance on SDMs, as a tool for marine resources management (Martin *et al.*, 2015; Robinson *et al.*, 2017). SDMs are based on a quantification of species-environment relationships by using numerical tools (Guisan & Zimmermann, 2000; Elith & Leathwick, 2009), and thus they represent a cost-effective way to provide efficient large-scale mapping where field-collected data are limited or currently unavailable (Gorman *et al.*, 2013; Álvarez-Martínez *et al.*, 2018). Among existing SDMs, correlative models are widely used as they are the most broadly applicable to diverse questions regarding the ecology and geography of biodiversity phenomena (Peterson *et al.*, 2011), especially those based on presence-only data due to their more easily parameterization and data availability.

Because of the increase of modelling techniques, modellers are usually confronted with the difficulty of which statistical algorithm to choose, and therefore, their associated uncertainties, which are largely uncharacterised. Each algorithm has its own strengths and weaknesses (for review see Franklin (2009); Peterson et al. (2011); Guisan, Thuiller, & Zimmermann (2017)) and there is no ultimate one to answer every possible question in ecological modelling (Guisan *et al.*, 2017). There is a wide body of literature with many attempts to guide this selection in the terrestrial environment (Segurado & Araujo, 2004; Aguirre-Gutiérrez *et al.*, 2013; Raina & Rao, 2014; Beaumont *et al.*, 2016), but marine ecosystems are much less studied (Robinson et al., 2011). For instance, the inconsistency in algorithm performance has led to the aggregation of the outputs from different algorithms into ensemble models (Araujo & New, 2007). While it is hard to provide general guidance (although see Araújo *et al.* (2019)), ensemble approaches have been suggested to increase the accuracy of models (Marmion *et al.*, 2009; Grenouillet *et al.*, 2011). An important point when applying ensembles is defining the criteria for these consensus approaches. According to Marmion et al. (2009) the weighted consensus method is the most efficient, but it is necessary to define the objective weighting scheme (Peterson *et al.*, 2011).

Another important issue when estimating the effect of climate change impacts over marine species using SDMs deals with the possibility of transferring results obtaining at one particular time through different future scenarios. The assessment of model transferability needs a careful interpretation based upon independent dataset, which is a limitation when not enough data are available to ensure such independence (Yates *et al.*, 2018). Although the spatial transferability of SDMs have been most studied (Randin *et al.*, 2006; Barbosa *et al.*, 2009; Lauria *et al.*, 2015; Sequeira *et al.*, 2016, 2018; Moon *et al.*, 2017), relatively few studies have focused on temporal transferability (Araujo *et al.*, 2005; Tapia *et al.*, 2017) because getting independent historical datasets needs long series and paleoecological are subjected to collector's and taphonomic biases (Yates *et al.*, 2018). Moreover, transferability assessments have not yet been standardised, leading to disparate interpretations that preclude comparisons of relative performance among model transfers (Sequeira *et al.*, 2018).

Bearing in mind the noted gaps, the overall objective of this chapter is to develop a methodology, broadly applicable, for selecting the modelling techniques and their weights that better fit in the marine environment, considering their temporal transferability for climate change studies. This way, the specific objective 3 of this thesis is accomplished. As a case study, seaweeds have been selected as their distribution is strongly affected by environmental factors (Lüning *et al.*, 1990), they are ecological engineers of the whole marine system and they provide a wide number of ecosystem services (Liquete *et al.*, 2016).

2. Methodological approach

The methodology presented is divided into three main steps (Figure 4.1):

1) Data selection and quality control. Predictors related to seaweed ecology and biological records of representative macroalgae (referred as groups hereinafter) were collected from 1985 to 2015. To get temporally independent data, two different periods were defined: 1985-1996 and 2004-2015.

2) Modelling. The best algorithms for each group were selected. Settings were adjusted for each algorithm by including only the most explanatory predictors, whose contributions were obtained for the final models.

3) Assessment. Model performance metrics, temporal transferability and the weight of each algorithm for each model yielded were calculated.

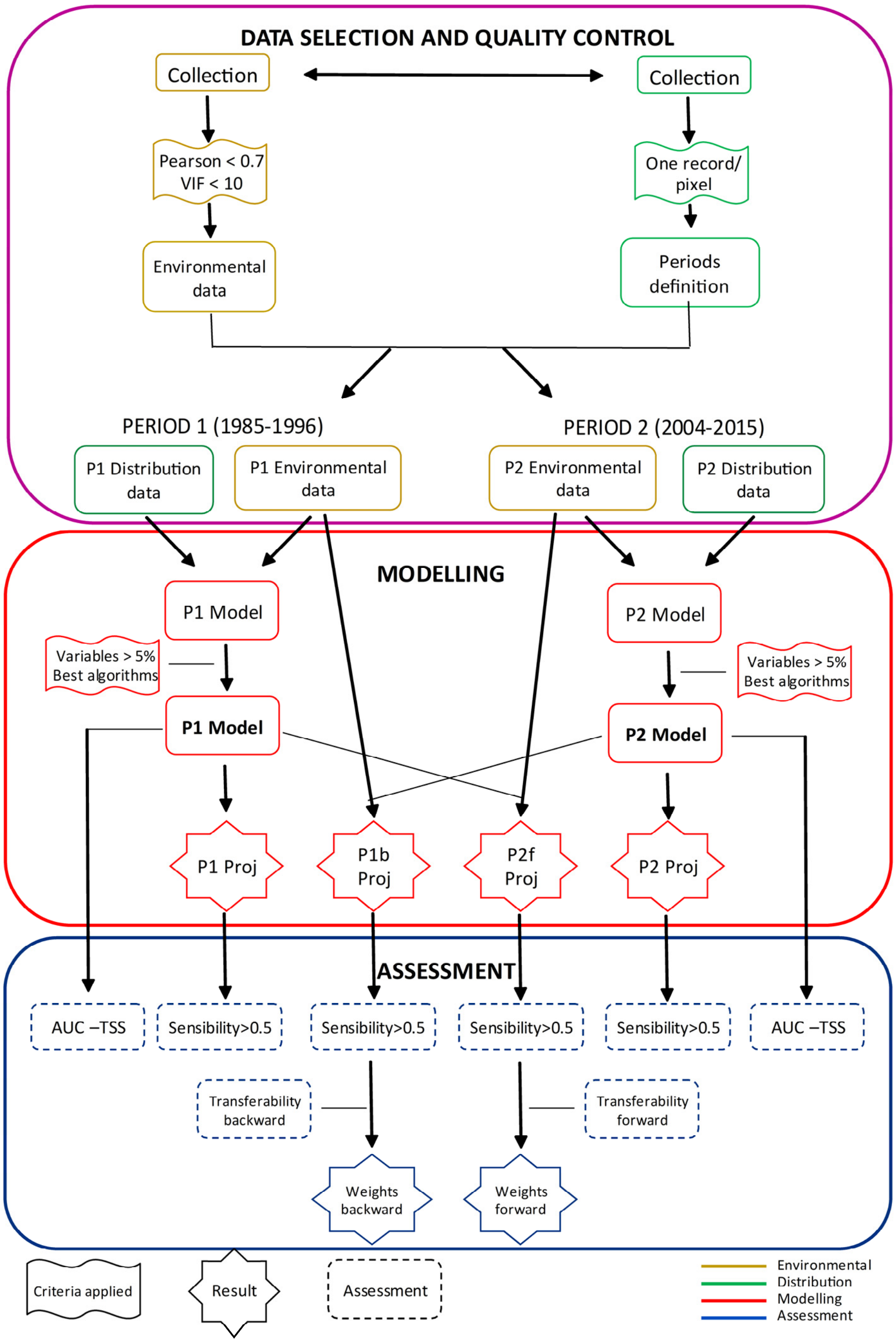


Figure 4.1. Flow diagram of the methodology proposed. Wave rectangles represent the criteria applied, stars the results and dashed rectangle the assessment techniques.

2.1. Data selection and quality control

2.1.1. Environmental data

Environmental variables were selected according to their relationship with seaweed ecology (Martínez *et al.*, 2012a; Young *et al.*, 2015; Cefalì *et al.*, 2016). Information for a complete set of parameters was collected from 1985 to 2015 from the OCLE database (de la Hoz *et al.*, 2018b) for each mesh point, including maximum value, minimum value, mean value, standard deviation, range and percentiles 10th, 25th, 50th, 75th and 90th for each variable (Table 2.1, Chapter II). Additionally, the number of consecutive days that 90th percentile of sea surface and air temperature is exceeded (Torresan *et al.*, 2016) and the shear stress harmful threshold 2.2 Nt/m^2 is surpassed (Vousdoukas *et al.*, 2012) were calculated as a proxy of stressful conditions.

For each point, parameters were calculated seasonally and for the whole period 1985-2015.

2.1.2. Distribution data

According to species grouping results (Chapter III), the five species selected were used as representative of the 21 macroalgae (Table 3.2, Chapter III): *Saccorhiza polyschides*, *Gelidium spinosum*, *Sargassum muticum*, *Pelvetia canaliculata* and *Cystoseira baccata*.

Two different periods were considered to assess models' temporal transferability. To ensure the independence of data between periods, a review of life history of algae was conducted, defining an average life history of 3-5 years for macroalgae (Dixon, 1966; Chapman & Chapman, 1980; Lüning *et al.*, 1990; Little & Kitching, 1996). To get the maximum independence possible between training (P1: 1985-1996) and the validation period (P2: 2004-2015), a gap of 7 years (1997-2003) between them was considered. Occurrence records of Table 3.3 were then classified into P1 or P2, following the criteria defined previously.

2.2. Modelling

2.2.1. Data

A 30% of species data were used for testing and 70% for training (Broennimann *et al.*, 2007). Training points were selected according to expert criteria to represent the ideal conditions of species (Figure 4.2).

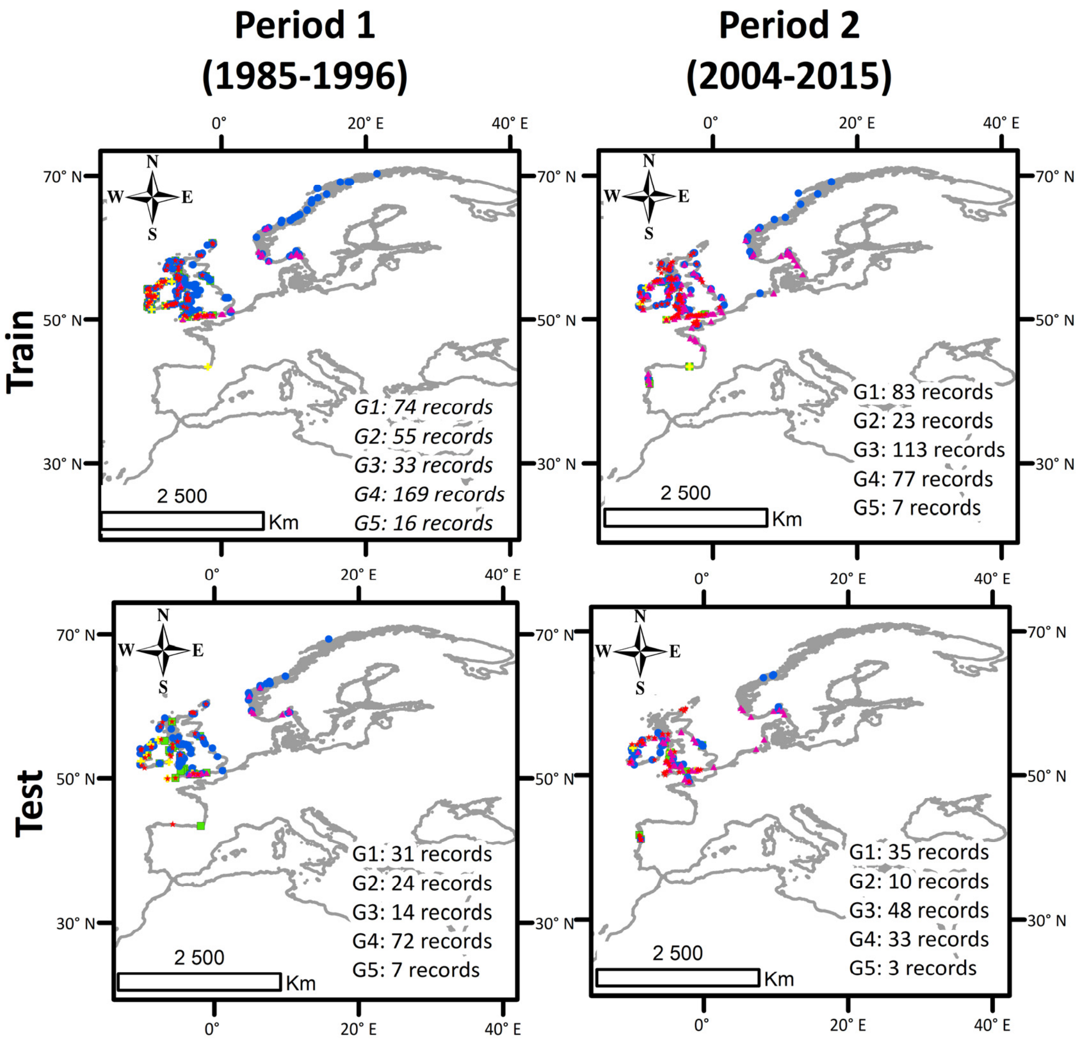


Figure 4.2. Species records used for training and testing in the two periods. Group 1 (*S. polyschides*) is represented with red stars, Group 2 (*G. spinosum*) with green squares, Group 3 (*S. muticum*) with pink triangles, Group 4 (*P. canaliculata*) with blue circles and Group 5 (*C. baccata*) with yellow crosses.

As only presence information was available, 1000 background points were randomly generated in the geographic space, not overlapping presence sites, which has been reported as the method that provide more interpretable results (Phillips & Dudík, 2008; Guisan *et al.*, 2017; Iturbide *et al.*, 2018).

Collinearity was tested in the variables with information for both periods with the criteria of a Pearson correlation coefficient < 0.7 and a VIF < 10, using the R package *usdm* (Naimi *et al.*, 2014). VIF values from five to ten are considered as critical for multi-variable correlation (Guisan *et al.*, 2017). The final predictors for each group and period can be consulted in Table 4.1.

Table 4.1. Parameters not correlated for each species group and period.

	PERIOD 1 (1985-1996)	**PERIOD 2 (2005-2015)**
GROUP 1 *Saccharina polyschides*	1. AT_Days	1. AT_Days
	2. AT_min _fall	2. AT_rg_summer
	3. Bathy	3. AT_std_winter
	4. Hs_max_fall	4. Bathy
	5. Rss_rg_fall	5. Hs_max_fall
	6. Rss_rg_summer	6. Rss_rg_spring
	7. Rss_rg_spring	7. Rss_rg_winter
	8. Sal_rg_fall	8. Sal_max
	9. SST_Days	9. Sal_rg_spring
	10. SST_rg_spring	10. SST_Days
	11. SST_rg_winter	11. SST_rg_summer
	12. SST_rg_summer	12. SST_std_winter
	13. Sub	13. Sub
	14. Tm_days	14. Tm_Days
	15. TR_rg_fall	15. Tm_min_summer
	16. Uw_min_spring	16. TR_mean
	17. Ws_min_spring	17. Uw_min_fall
	18. Ws_min_winter	18. Ws_min_summer
	19. Ws_std_fall	19. Ws_min_spring
		20. Ws_min_winter

GROUP 2 *Gelidium spinosum*	1. AT_Days	1. AT_rg_summer
	2. Bathy	2. Bathy
	3. Hs_max_fall	3. Hs_std_summer
	4. Rss_rg_spring	4. Rss_rg_summer
	5. Rss_rg_winter	5. Rss_rg_spring
	6. Rss_std	6. Rss_rg_winter
	7. Sal_rg_fall	7. Sal_rg_fall
	8. SST_Days	8. SST_Days
	9. SST_max_winter	9. Sub
	10. SST_rg_spring	10. Tm_max_fall
	11. SST_std_summer	11. Tm_mean_summer
	12. SST_std_winter	12. TR_max
	13. Sub	13. Urms_max
	14. Tm_max_fall	14. Ws_min_spring
	15. TR_mean	15. Ws_min_winter
	16. Uw_P75	
	17. Ws_min_spring	
	18. Ws_min_winter	
	19. Ws_std_fall	
GROUP 3 *Sargassum muticum*	1. AT_Days	1. AT_Days
	2. Bathy	2. AT_P10_summer
	3. Hs_std_fall	3. Bathy
	4. Sal_rg_fall	4. Hs_max_fall
	5. SST_Days	5. Rss_rg_fall
	6. SST_rg_winter	6. Rss_rg_summer
	7. Sub	7. Rs_rg_winter
	8. Tm_max_fall	8. Rss_std_spring
	9. TR_rg_fall	9. Sal_rg_fall
	10. Urms_mean	10. SST_Days
	11. Ws_max_summer	11. SST_max_spring
	12. Ws_min_summer	12. Sub
		13. Tm_Days
		14. Tm_max_fall
		15. TR_mean
		16. Urms_min_winter
		17. Ws_min_spring
		18. Ws_min_winter
		19. Ws_rg_summer

GROUP 4 *Pelvetia canaliculata*	1. AT_Days	1. AT_Days
	2. AT_rg_winter	2. AT_max_summer
	3. Bathy	3. AT_rg_winter
	4. Hs_P50	4. Bathy
	5. Rss_mean_summer	5. Hs_max_fall
	6. Rss_rg_spring	6. Rss_max_summer
	7. Rss_std	7. Rss_rg_fall
	8. Sal_max	8. Rss_rg_spring
	9. Sal_rg_winter	9. Rss_rg_winter
	10. SST_Days	10. Rss_std_
	11. SST_max_spring	11. Sal_rg_winter
	12. SST_P50_fall	12. SST_Days
	13. SST_rg_winter	13. SST_P75_spring
	14. SST_std_fall	14. SST_rg
	15. Sub	15. Sub
	16. Tm_Days	16. Tm_Days
	17. TR_rg_fall	17. TR_max
	18. Uw_max_winter	18. Uw_P75
	19. Uw_P75	19. Ws_min_summer
	20. Ws_min_summer	20. Ws_min_winter
	21. Ws_rg_winter	21. Ws_std_fall
GROUP 5 *Cystoseira baccata*	1. AT_Days	1. AT_Days
	2. Bathy	2. AT_max
	3. SST_Days	3. AT_mean
	4. SST_rg_spring	4. AT_rg_summer
	5. SST_rg_winter	5. Bathy
	6. Sub	6. Hs_max
	7. Tm_Days	7. RSS_rg_spring
	8. Tm_mean	8. Sal_max
	9. TR_max	9. Sal_rf_fall
	10. Urms_max	10. SST_Days
	11. Ws_min_summer	11. Sub
		12. Tm_max_fall
		13. TR_max
		14. Urms_max

2.2.2. Model implementation

In this work modelling algorithms or algorithms are defined as the specific sequence of instructions for solving a problem or developing a task; meanwhile model is used as a simplified representation of some aspects of nature for the purpose of research (*sensu* Peterson *et al.* (2011).

Nine correlative modelling algorithms were applied to reflect the different approaches available and run sensitivity analyses of model performance: regression analysis (GLM (McCullagh & Nelder, 1989); GAM (Hastie & Tibshirani, 1990); MARS (Friedman, 1991); MDA (Hastie *et al.*, 1994)), classification (CART (Breiman, 1984); SVM (Vapnik, 1995); FDA (Hastie *et al.*, 1994)) and machine-learning methods (RF (Breiman, 2001); MAXENT(Phillips *et al.*, 2006)). As the goal was to evaluate model performance under commonly adopted settings, algorithms were not tuned and run under default settings (Qiao *et al.*, 2019; Regos *et al.*, 2019).

Models were trained separately for each period, with the corresponding environmental and distribution data. Then, each calibrated model in one period was projected onto the other period, forward in case of model P1 to P2 (P2f) and backward in case of model P2 to P1 (P1b) (Figure 4.1).

Sensitivity was calculated as the percentage of presences correctly predicted (over 50% of the highest probability) (Thuiller *et al.*, 2005) (Figure 4.2).

Besides, a visual inspection of predictions was conducted to identify possible abnormalities, such as unnatural patterns due to algorithms or data-driven effects. According to results, SVM was discarded because all area was predicted as suitable. In the case of group 3, the percentages for GLM, MDA and FDA were very low because all coastal points were predicted as highly suitable. In the group 4, the pattern of GAM was not consistent, thus it was removed. For the group 5 only RF accomplished the criteria and MAXENT sensitivity was good in the P1. Finally, only the algorithms with ecological coherence were considered and retained for further analyses.

As climate change studies are the aim of this thesis, only the results in P2f were considered. An ensemble model was built for each group by averaging the results of the most transferable algorithms according to weights.

The contributions of the different predictors were calculated for the selected models (Table 4.3) and those with a contribution lower than 5% were discarded.

New models were running with the selected algorithms and predictors (Table 4.4) and the same training and testing points.

Table 4.2. Percentage of presences correctly predicted for each group, temporal period and algorithm.

	GROUP 1				GROUP 2				GROUP 3				GROUP 4				GROUP 5			
	P1	P1b	P2	P2f	P1	P1b	P2	P2f	P1	P1b	P2	P2f	P1	P1b	P2	P2f	P1	P1b	P2	P2f
GLM	27.62	37.14	14.41	18.64	22.78	6.33	21.21	75.76	19.15	8.51	6.83	95.03	15.77	98.34	20.00	95.45	17.39	4.35	10.00	10.00
GAM	37.14	6.67	22.03	1.69	29.11	3.80	21.21	57.58	68.09	19.15	17.39	100	44.40	0.83	42.73	97.27	8.70	-	-	10.00
RF	71.43	37.14	37.29	52.54	69.62	15.19	18.18	63.64	76.60	59.57	17.39	12.42	72.61	42.74	52.73	70.00	69.57	21.74	10.00	60.00
MAXENT	79.05	37.14	52.54	10.17	79.75	39.24	42.42	60.61	82.98	72.34	30.43	21.12	80.50	47.72	78.18	83.64	82.61	34.78	10.00	10.00
CART	45.71	3.81	24.58	34.75	34.18	-	-	12.12	-	2.13	100.00	-	14.94	26.97	10.91	2.73	26.09	-	-	10.00
SVM	37.14	12.38	20.34	84.75	20.25	41.77	72.73	51.52	53.19	29.79	15.53	0.00	7.47	53.94	10.91	18.18	91.30	100.00	10.00	100.00
MARS	22.86	13.33	12.71	3.39	39.24	2.53	12.12	6.06	19.15	25.53	4.97	0.62	27.80	0.41	33.64	27.27	13.04	4.35	10.00	10.00
MDA	47.62	32.38	34.75	11.86	34.18	7.59	36.36	30.30	19.15	14.89	14.29	93.79	51.87	48.55	46.36	37.27	47.83	13.04	10.00	10.00
FDA	34.29	22.86	19.49	1.69	15.19	2.53	15.15	9.09	8.51	12.77	1.86	99.38	39.83	33.61	35.45	20.00	39.13	4.35	10.00	10.00

Table 4.3. Contribution of each predictor (in parenthesis in percentage) by groups and periods ordered from higher to lower importance.

	GROUP 1	**GROUP 2**	**GROUP 3**	**GROUP 4**	**GROUP 5**
PREDICTORS PERIOD 1	SST_std_s (64)	SST_std_s (39)	Sal_rg_f (17)	bathy (39)	Urms_max (60)
	TR_rg_f (63)	SST_max_w (29)	Urms_mean (17)	SST_max_sp (32)	TR_max (47)
	bathy (54)	TR_mean (27)	TR_rg_f (15)	SST_P50_f (20)	SST_rg_w (35)
	SST_rg_sp (33)	bathy (24)	sub (15)	sub (18)	bathy (26)
	Ws_min_sp (27)	SST_rg_sp (24)	Ws_min_s (14)	Sal_rg_w (16)	Tm_mean (26)
	Uw_min_sp (18)	Ws_min_w (21)	bathy (13)	Rss_std (15)	SST_rg_sp (25)
	AT_Days (17)	Uw_P75 (21)	SST_rg_w (13)	TR_rg_f (14)	SST_Days (22)
	Rss_rg_f (17)	Sal_rg_f (19)	SST_Days (10)	Ws_min_s (9)	sub (10)
	Tm_Days (15)	sub (18)	Ws_max_s (10)	SST_std_f (8)	AT_Days (7)
	Rss_rg_s (12)	Ws_min_sp (17)	Hs_std_f (10)	Hs_P50 (8)	Tm_Days (6)
	Ws_std_f (11)	Rss_rg_sp (17)	Tm_max_f (8)	Uw_P75 (8)	
	Hs_max_f (11)	Rss_rg_w (16)		Uw_max_w (7)	
	Rss_rg_sp (11)	Hs_max_f (14)		AT_rg_w (6)	
	sub (9)	SST_Days (12)		Rss_mean_s (6)	
	Ws_min_w (8)	AT_Days (10)		Sal_max (5)	
	SST_rg_w (7)	SST_std_w (8)			
	Sal_rg_f (6)	Tm_max_f (8)			
		Rss_std (7)			
PREDICTORS PERIOD 2	SST_rg_s (78)	Rss_rg_w (43)	TR_mean (19)	bathy (32)	TR_max (47)
	TR_mean (67)	Sal_rg_f (40)	SST_max_sp (14)	sub (24)	Rss_rg_sp (39)
	bathy (49)	Ws_min_sp (39)	Ws_rg_s (12)	SST_P75_sp (23)	AT_Days (33)
	Sal_max (23)	bathy (35)	Ws_min_sp (12)	Ws_min_s (18)	AT_rg_s (33)
	Uw_min_f (21)	TR_max (34)	bathy (11)	Uw_P75 (17)	Sal_max (32)
	Ws_min_w (18)	sub (26)	Urms_min_w (10)	AT_max_s (16)	Urms_max (30)
	Tm_Days (17)	Tm_mean_s (26)	Sal_rg_f (10)	AT_rg_w (14)	bathy (30)
	Rss_rg_w (14)	AT_rg_s (23)	AT_P10_s (9)	Sal_rg_w (11)	AT_mean (29)
	Tm_min_s (11)	Ws_min_w (18)	Rss_rg_f (7)	Rss_std (9)	Tm_max_f (26)
	Rss_rg_sp (10)	Urms_max (13)	SST_Days (6)	TR_max (8)	Hs_max (21)
	AT_rg_s (10)	Tm_max_f (12)		Rss_max_s (5)	Sal_rg_f (16)
	sub (5)	SST_Days (10)		SST_rg (5)	sub (14)
		Rss_rg_s (8)			AT_max (12)
		Hs_std_s (7)			

Table 4.4. Filtered models and predictors.

	GROUP 1	**GROUP 2**	**GROUP 3**	**GROUP 4**	**GROUP 5**
	RF, CART	GLM, GAM, RF, MAXENT, MDA	RF, MAXENT	GLM, RF, MAXENT, MDA	RF, MAXENT
PREDICTORS P1	SST_std_s TR_rg_f bathy SST_rg_sp Ws_min_sp Uw_min_sp AT_Days Rss_rg_f Tm_Days Rss_rg_s Ws_std_f Hs_max_f Rss_rg_sp sub Ws_min_w SST_rg_w Sal_rg_f	SST_std_s SST_max_w TR_mean bathy SST_rg_sp Ws_min_w Uw_P75 Sal_rg_f sub Ws_min_sp Rss_rg_sp Rss_rg_w Hs_max_f SST_Days AT_Days SST_std_w Tm_max_f Rss_std	Sal_rg_f Urms_mean TR_rg_f sub Ws_min_s bathy SST_rg_w SST_Days Ws_max_s Hs_std_f Tm_max_f	bathy SST_max_sp SST_P50_f sub Sal_rg_w Rss_std TR_rg_f Ws_min_s SST_std_f Hs_P50 Uw_P75 Uw_max_w AT_rg_w Rss_mean_s Sal_max	Urms_max TR_max SST_rg_w bathy Tm_mean SST_rg_sp SST_Days sub AT_Days Tm_Days
PREDICTORS P2	SST_rg_s TR_mean bathy Sal_max Uw_min_f Ws_min_w Tm_Days Rss_rg_w Tm_min_s Rss_rg_sp AT_rg_s sub	Rss_rg_w Sal_rg_f Ws_min_sp bathy TR_max sub Tm_mean_s AT_rg_s Ws_min_w Urms_max Tm_max_f SST_Days Rss_rg_s Hs_std_s	TR_mean SST_max_sp Ws_rg_s Ws_min_sp bathy Urms_min_w Sal_rg_f AT_P10_s Rss_rg_f SST_Days	bathy sub SST_P75_sp Ws_min_s Uw_P75 AT_max_s AT_rg_w Sal_rg_w Rss_std TR_max Rss_max_s SST_rg	TR_max Rss_rg_sp AT_Days AT_rg_s Sal_max Urms_max bathy AT_mean Tm_max_f Hs_max Sal_rg_f sub AT_max

Variable importance scores revealed differences depending on the groups, although the highest relevance was achieved in general by SST, TR and bathymetry (Figure 4.3 and Table 4.5). This was especially clear in the case of the group 1, for which the summer SST variability, the TR and the bathymetry were within the most important variables for both periods. In the case of the group 2, bathymetry and TR were crucial in both periods but there were some remarkable differences. Meanwhile in P1 SST and Rss parameters had more importance, the P2 was more diverse in the variables represented. For both periods of group 3, salinity range in fall and the minimum Ws were within the predictors that contribute the most. Besides bathymetry and SST days contributed equally for P1 and P2; meanwhile Urms and TR parameters were present in both periods with different importance. In the group 4, bathymetry, substrate, TR and the standard deviation of Rss were important in both periods with the same importance. However, the minimum Ws in summer and the AT range in winter presented different contribution in the different periods. For the group 5 the maximum values of TR and Urms contributed importantly to both periods. Temperature was also important in both cases, but in P1 SST contributed more, meanwhile AT did it in P2. It is also remarkable that the extreme conditions of AT were in both periods, but P1 included also for SST and Tm. In P2 the maximum salinity contributed importantly.

Models were built, evaluated and projected using the *sdm* package (Naimi & Araujo, 2016).

Table 4.5. Contribution of each variable (in brackets in percentage) by groups and periods ordered from higher to lower importance.

	GROUP 1	GROUP 2	GROUP 3	GROUP 4	GROUP 5
PREDICTORS P1	SST_std_s (65) TR_rg_f (50) bathy (45) Uw_min_sp (35) SST_rg_sp (25) Rss_rg_s (19) SST_rg_w (19) Tm_Days (19) Rss_rg_sp (17) Ws_min_w (16) Hs_max_f (11) Ws_min_sp (9) Sal_rg_f (9) Rss_rg_f (8) Ws_std_f (7) Sub (7) AT_Days (6)	SST_std_s (44) bathy (40) SST_max_w (37) SST_rg_sp (28) TR_mean (22) Uw_P75 (17) Ws_min_sp (16) Sal_rg_f (16) SST_Days (14) Rss_std (14) Ws_min_w (13) sub (13) Rss_rg_w (11) Rss_rg_sp (11) Tm_max_f (8) Hs_max_f (7) SST_std_w (7) AT_Days (5)	Sal_rg_f (37) Urms_mean (34) TR_rg_f (31) Ws_min_s (29) sub (26) bathy (25) Ws_max_s (23) SST_rg_w (22) Hs_std_f (18) SST_Days (17) Tm_max_f (12)	SST_P50_f (47) bathy (46) SST_max_sp (37) Sal_rg_w (25) TR_rg_f (22) sub (20) Uw_max_w (18) SST_std_f (17) Rss_std (14) Rss_mean_s (13) Hs_P50(13) AT_rg_w (9) Uw_P75(9) Ws_min_s (9) Sal_max (8)	Urms_max (72) SST_rg_w (42) TR_max (40) Tm_mean (35) SST_Days (27) SST_rg_sp (22) bathy (22) AT_Days (15) Tm_Days (13) sub (13)
PREDICTORS P2	SST_rg_s (78) TR_mean (55) bathy (53) Sal_max (19) AT_rg_s (19) Uw_min_f (16) Tm_min_s (14) Tm_Days (13) Rss_rg_sp (13) Ws_min_w (11) sub (11) Rss_rg_w (5)	Sal_rg_f (50) Rss_rg_w (48) Bathy (37) Ws_min_sp (34) TR_max (31) Tm_mean_s (28) sub (25) AT_rg_s (20) Tm_max_f (16) Urms_max (11) SST_Days (10) Ws_min_w (8) Hs_std_s (6) Rss_rg_s (5)	SST_max_sp (37) Ws_min_sp (28) Sal_rg_f (27) Ws_rg_s(27) TR_mean(27) bathy (27) Rss_rg_f (25) AT_P10_s (23) Urms_min_w (21) SST_days_Days (17)	bathy (38) SST_P75_sp (38) sub (25) Sal_rg_w (24) Uw_P75 (23) Ws_min_s (21) TR_max (19) AT_rg_w (17) AT_max_s (17) Rss_std (10) Rss_max_s (7) SST_rg (5)	TR_max (70) AT_rg_s (59) AT_Days (54) AT_mean (54) Sal_max (53) Urms_max (51) Rss_rg_sp (48) Tm_max_f (46) Hs_max (37) bathy (33) sub (24) AT_max (21) Sal_rg_f (15)

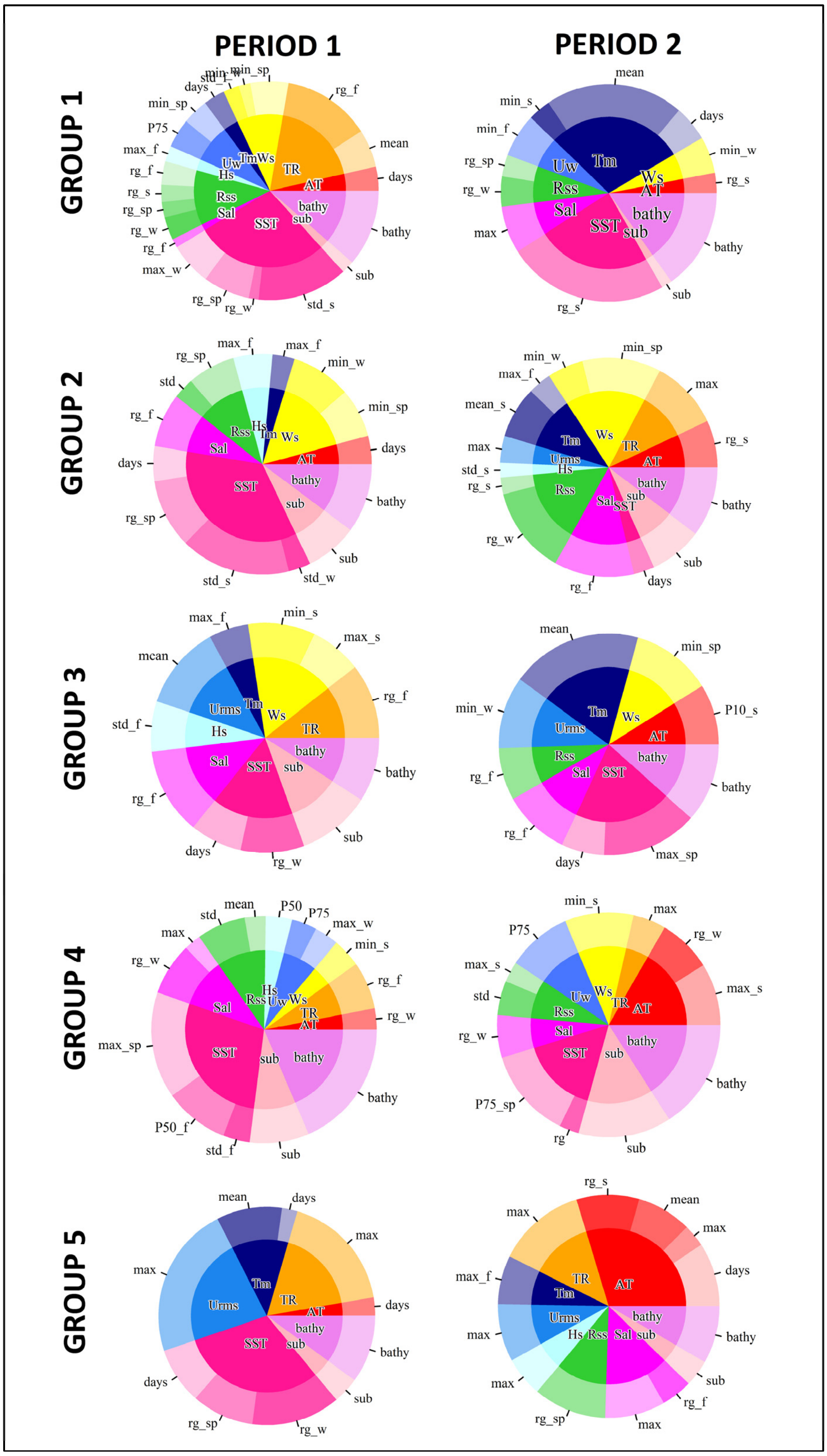


Figure 4.3. Contribution of each variable to the final models. Inner circle colours refer to variables and exterior subdivisions to their parameters. Relative values can be consulted in Table 4.5.

2.3. Assessment

To assess the predictive accuracy and performance of models the AUC (Hanley & McNeil, 1982) and the TSS (Allouche *et al.*, 2006) were employed. AUC varies from 0.5 for a model whose predictions are no better than random, to 1 for a model achieving perfect agreement with the observed data. TSS varies between 0 (random model) and 1 (perfect agreement). All AUC and TSS weighted values were over 0.8 (Table 4.6), which can be considered good (Swets, 1988; Araujo *et al.*, 2005; Engler *et al.*, 2011; Gallien *et al.*, 2012). For both metrics the highest variability was observed in group 1, specially in TSS values. Despite the wide use of AUC and TSS for SDMs assessment, it is necessary to take into account their limitations, previously established by other authors (Allouche *et al.*, 2006; Araújo & Guisan, 2006; Lobo *et al.*, 2008). Nevertheless, as the aim of this study is to be general, it has been considered appropriate to include these metrics.

Table 4.6. AUC and TSS weighted average and standard deviation values for each group in P2f.

	GROUP 1	**GROUP 2**	**GROUP 3**	**GROUP 4**	**GROUP 5**
AUC	0.87±0.25	0.92±0.07	0.92±0.01	0.97±0.07	0.91±0.01
TSS	0.80±0.39	0.90±0.02	0.92±0.07	0.93±0.05	0.96±0

AUC and TSS analysis were performed in R, through *pROC* (Robin *et al.*, 2011) and *sdm* (Naimi & Araujo, 2016) packages respectively. AUC plots are represented in Figure 4.4 – Figure 4.8.

GROUP 1 *Saccorhiza polyschides*

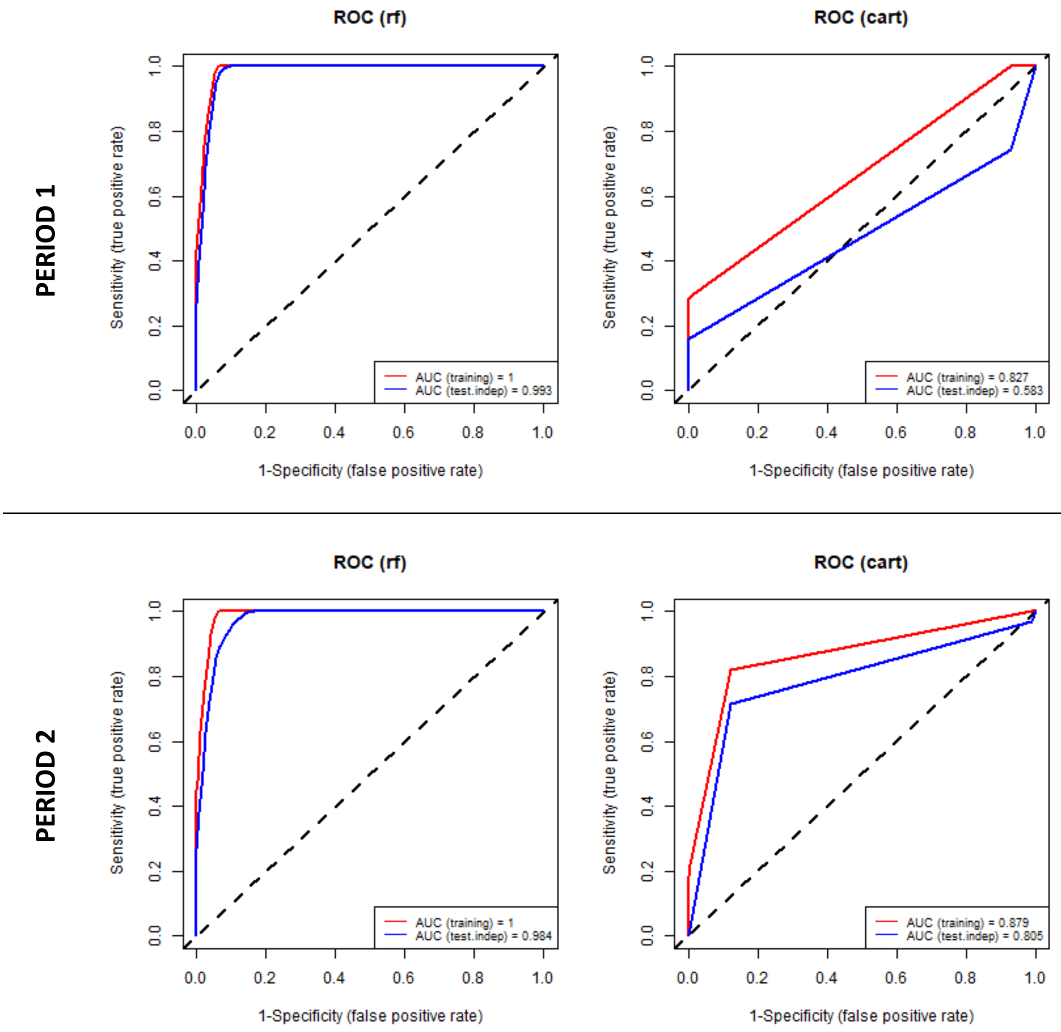


Figure 4.4. AUC ROC plots for Saccorhiza polyschides modelled using the final selected algorithms (in brackets above each graph) for periods 1 and 2. Red line represents training AUC and blue line testing AUC.

GROUP 2 *Gelidium spinosum*

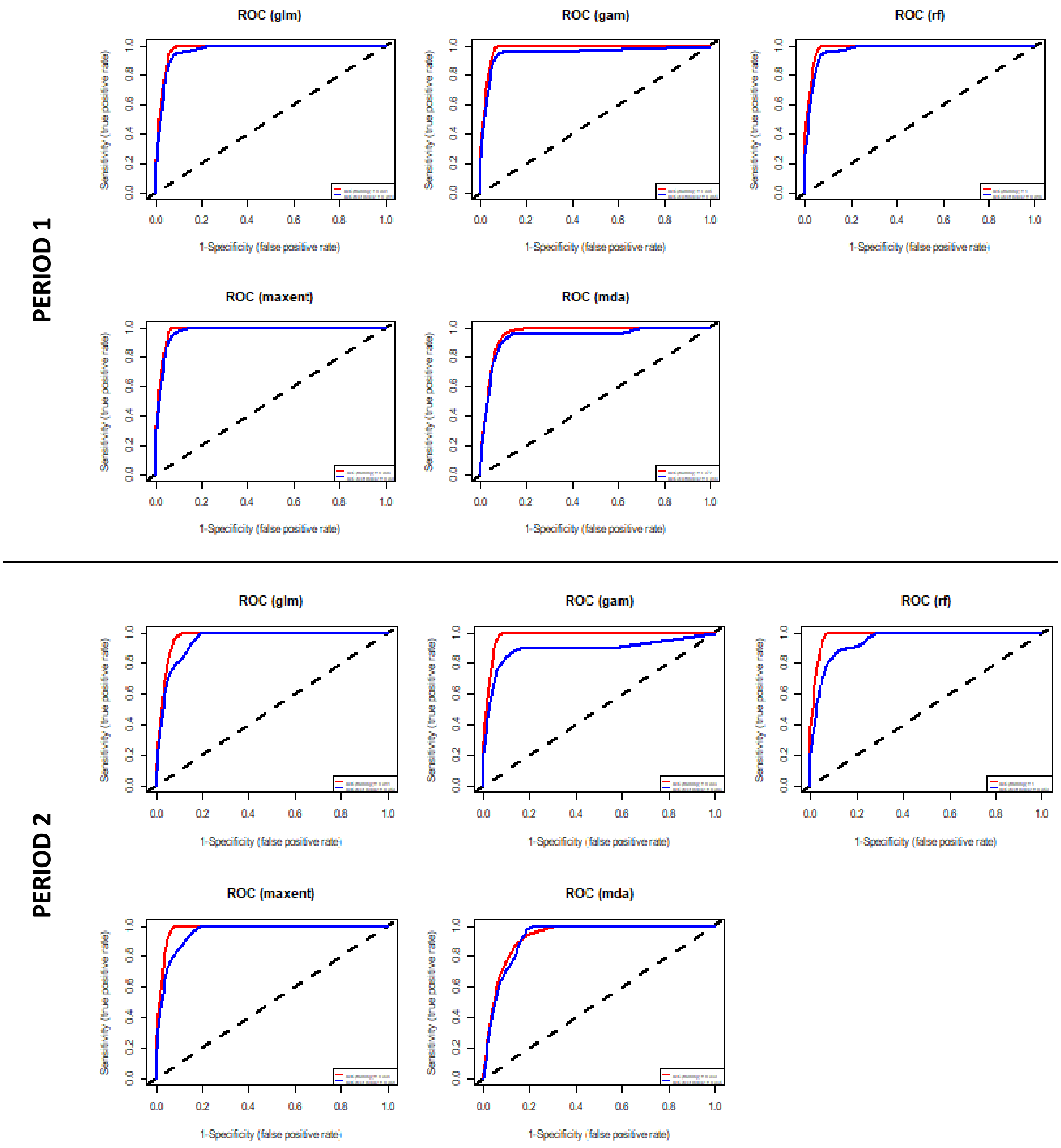


Figure 4.5. AUC ROC plots *Gelidium spinosum* modelled using the final selected algorithms (in brackets above each graph) for periods 1 and 2. Red line represents training AUC and blue line testing AUC.

GROUP 3 *Sargassum muticum*

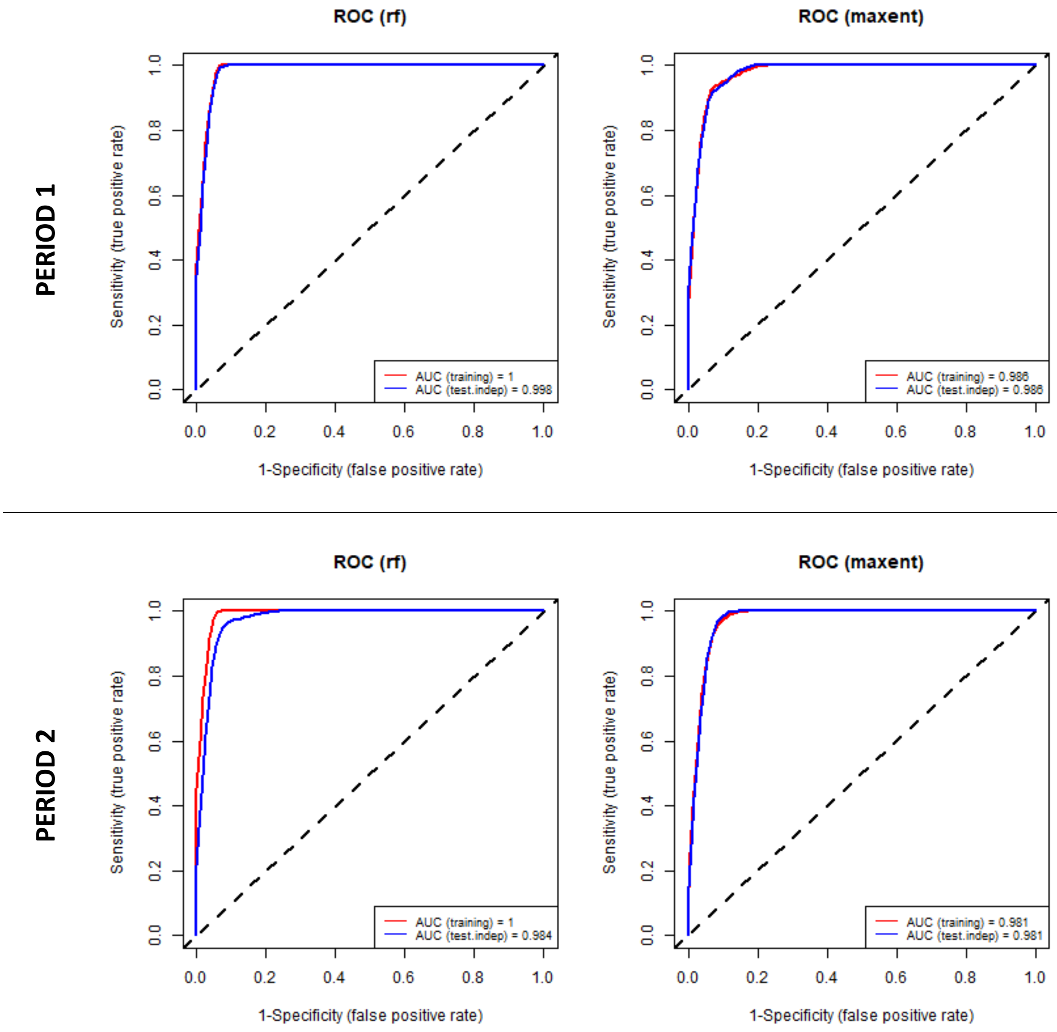


Figure 4.6. AUC ROC plots for Sargassum muticum modelled using the final selected algorithms (in brackets above each graph) for periods 1 and 2. Red line represents training AUC and blue line testing AUC.

GROUP 4 *Pelvetia canaliculata*

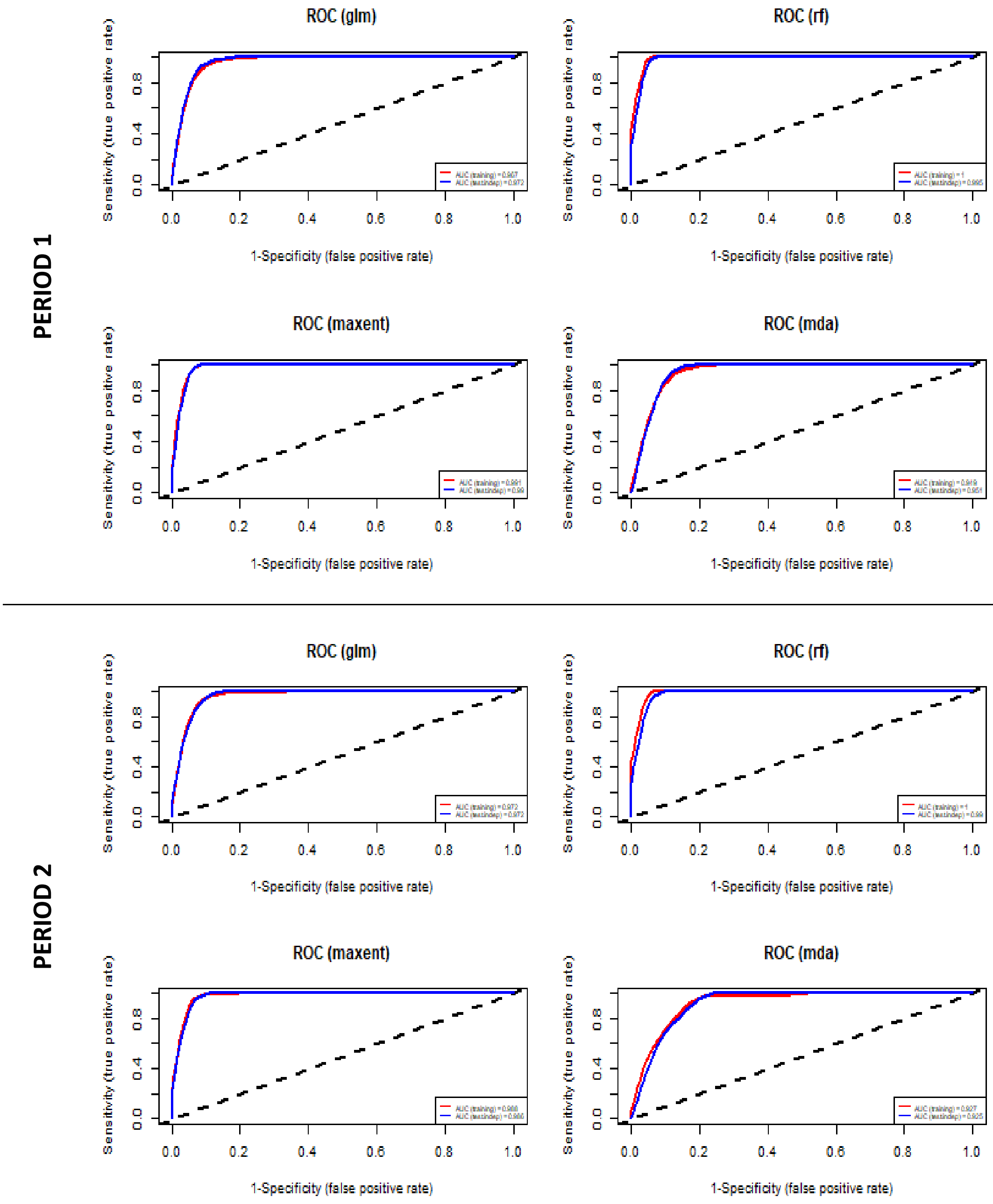


Figure 4.7. AUC ROC plots for Pelvetia canaliculata modelled using the final selected algorithms (in brackets above each graph) for periods 1 and 2. Red line represents training AUC and blue line testing AUC.

GROUP 5 *Cystoseira baccata*

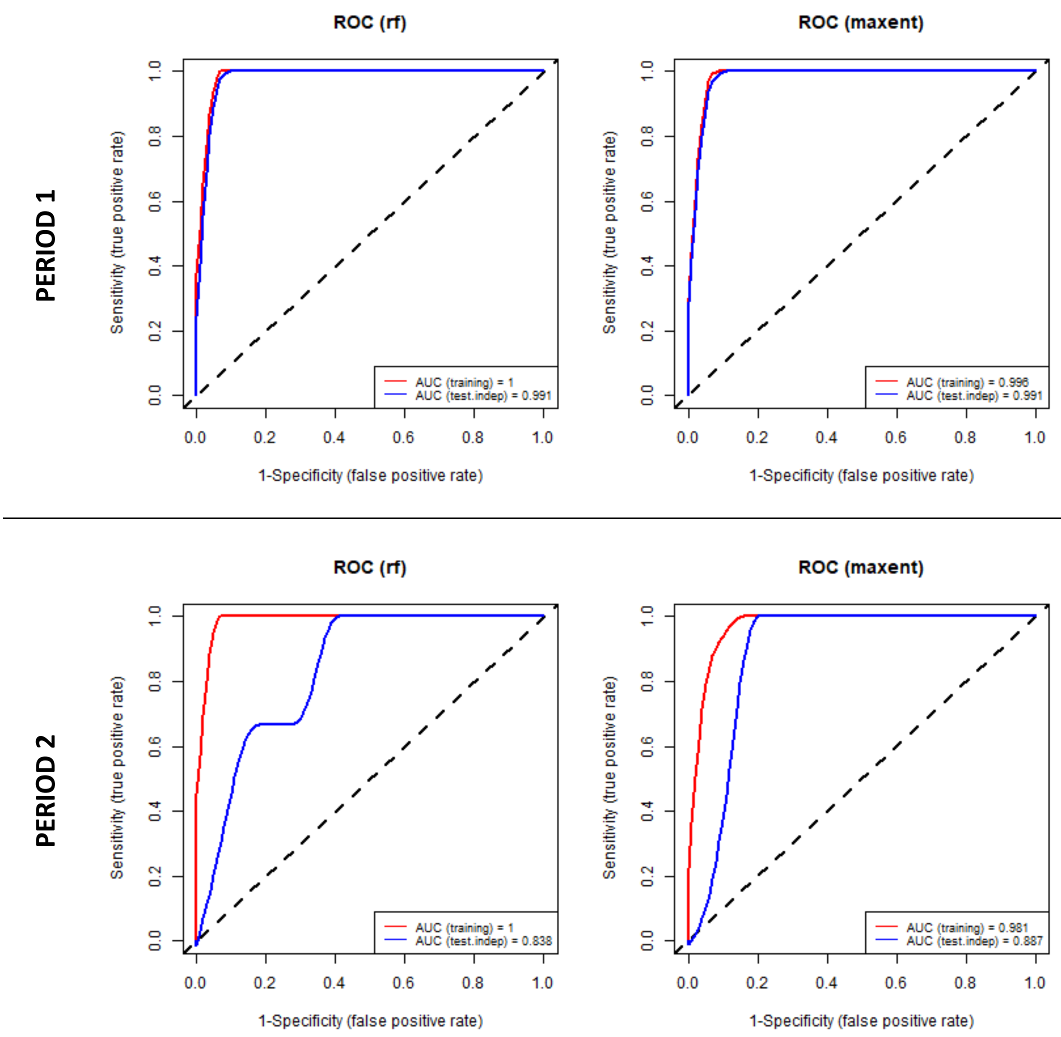


Figure 4.8. AUC ROC plots for *Cystoseira baccata* modelled using the final selected algorithms (in brackets above each graph) for periods 1 and 2. Red line represents training AUC and blue line testing AUC.

After executing the models with the final algorithms, the sensitivity in P2f was calculated as previously to get the weights that each algorithm represents for each group (Table 4.7). According to these weights, the ensemble was calculated.

Table 4.7. Weight assigned for each algorithm at each group.

GROUP	MODEL	WEIGHT
1	RF	70.91
	CART	29.09
2	GLM	26.04
	GAM	21.88
	RF	21.88
	MAXENT	22.92
	MDA	7.29
3	RF	44.83
	MAXENT	55.17
4	GLM	7
	RF	25.5
	MAXENT	40
	MDA	27.5
5	RF	85.72
	MAXENT	14.29

Backward and forward transferability was assessed using Tuanmu et al. (2011) equations, which could be considered good because for all groups the indices were over 0.8 (Table 4.8). Backward transferability was in general better than forward, but in all cases reasonable.

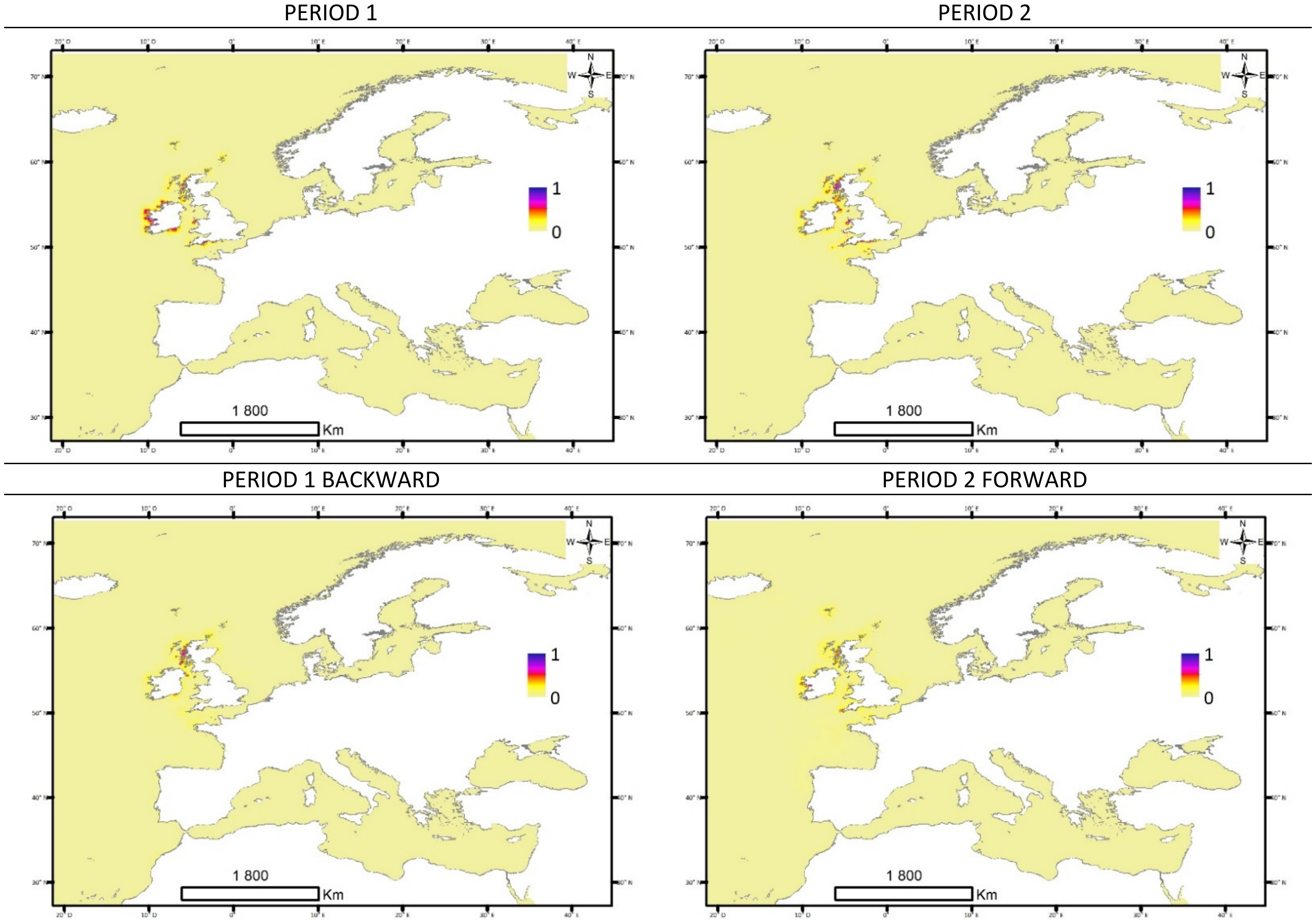
Table 4.8. Single directions and overall temporal transferability indices by group.

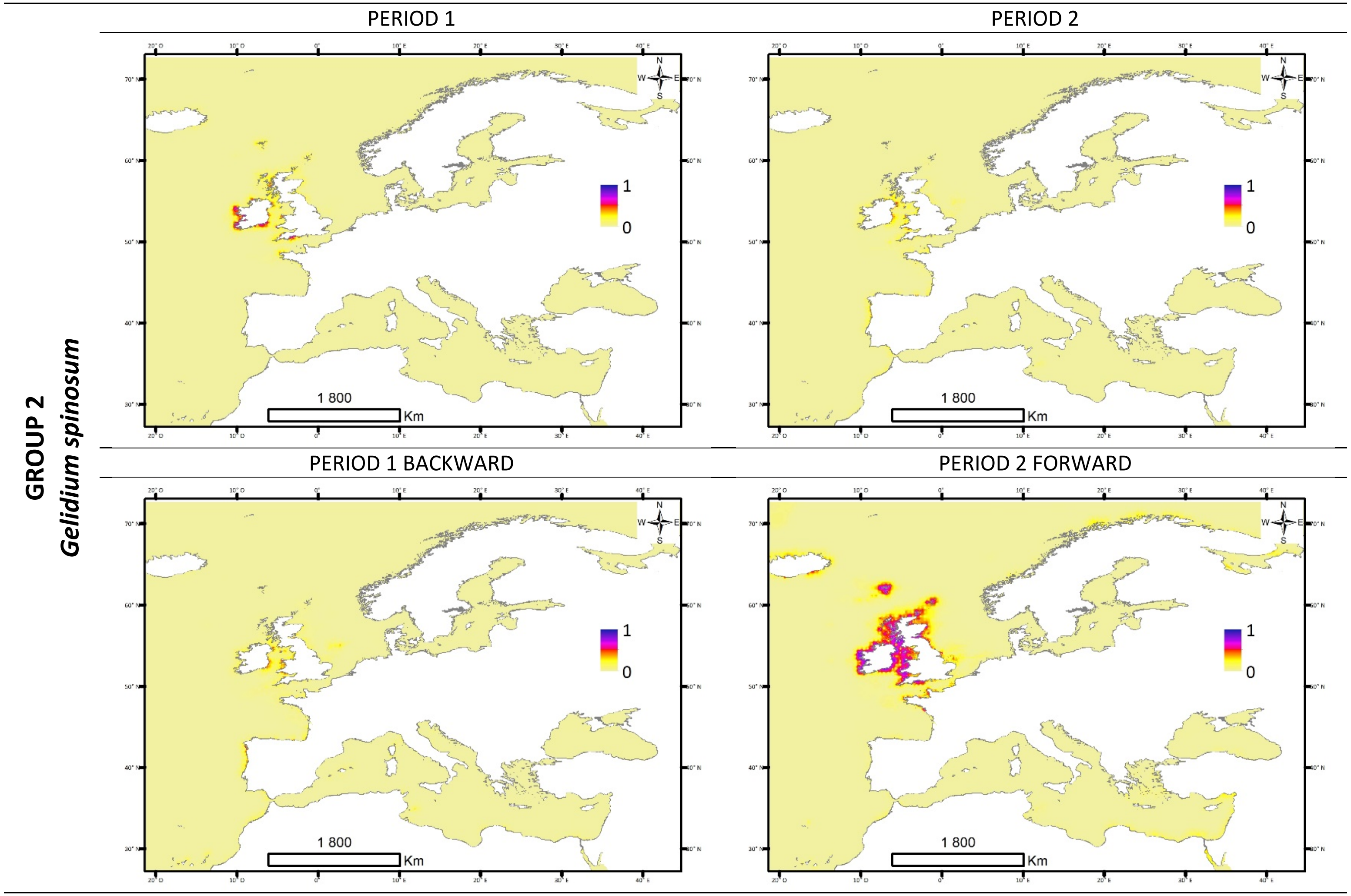
TRANSFERABILITY	GROUP 1	GROUP 2	GROUP 3	GROUP 4	GROUP 5
Forward	1	0.88	0.86	0.98	0.84
Backward	0.98	0.92	0.96	0.92	0.86
Overall	0.97	0.87	0.83	0.90	0.83

Moreover, the overlap of species niches was assessed using Hellinger's I (modified by Warren, Glor, & Turelli, 2008) in the R package *ecospat* (Broennimann et al., 2016). For both periods it was 0.67, indicative of high levels of overlap (Broennimann et al., 2012).

The geographical distribution of probabilities in the different periods projected, P1 and P1b and P2 and P2f patterns, was similar for all groups (Figure 4.9).

GROUP 1
Saccorhiza polyschides

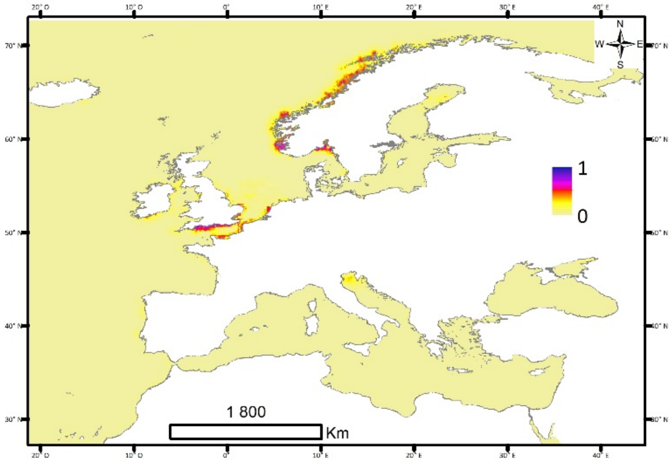

PERIOD 1
PERIOD 2
PERIOD 1 BACKWARD
PERIOD 2 FORWARD


GROUP 2
Gelidium spinosum
PERIOD 1
PERIOD 2
PERIOD 1 BACKWARD
PERIOD 2 FORWARD
N
W
E
S
1
0
1 800
Km

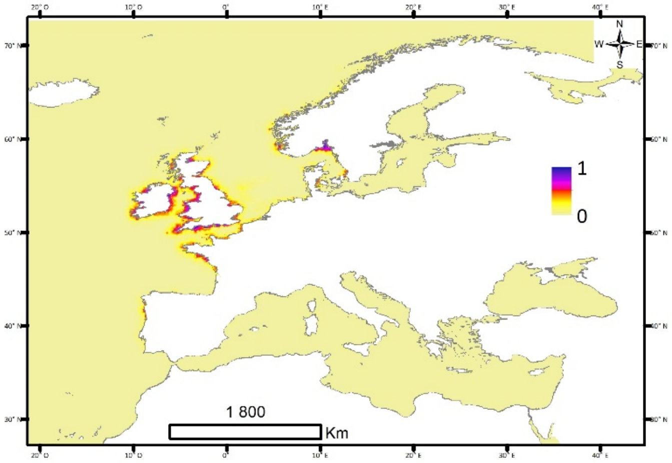
GROUP 3

Sargassum muticum

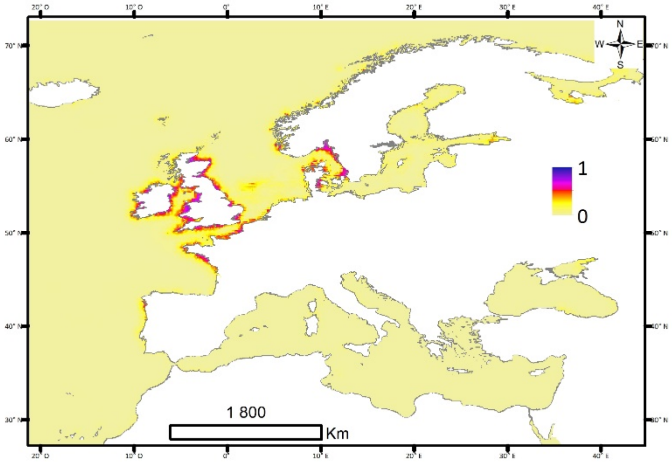
PERIOD 1



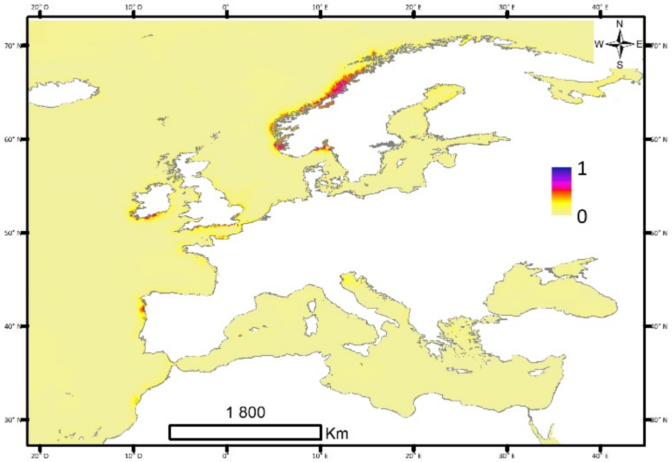
PERIOD 2

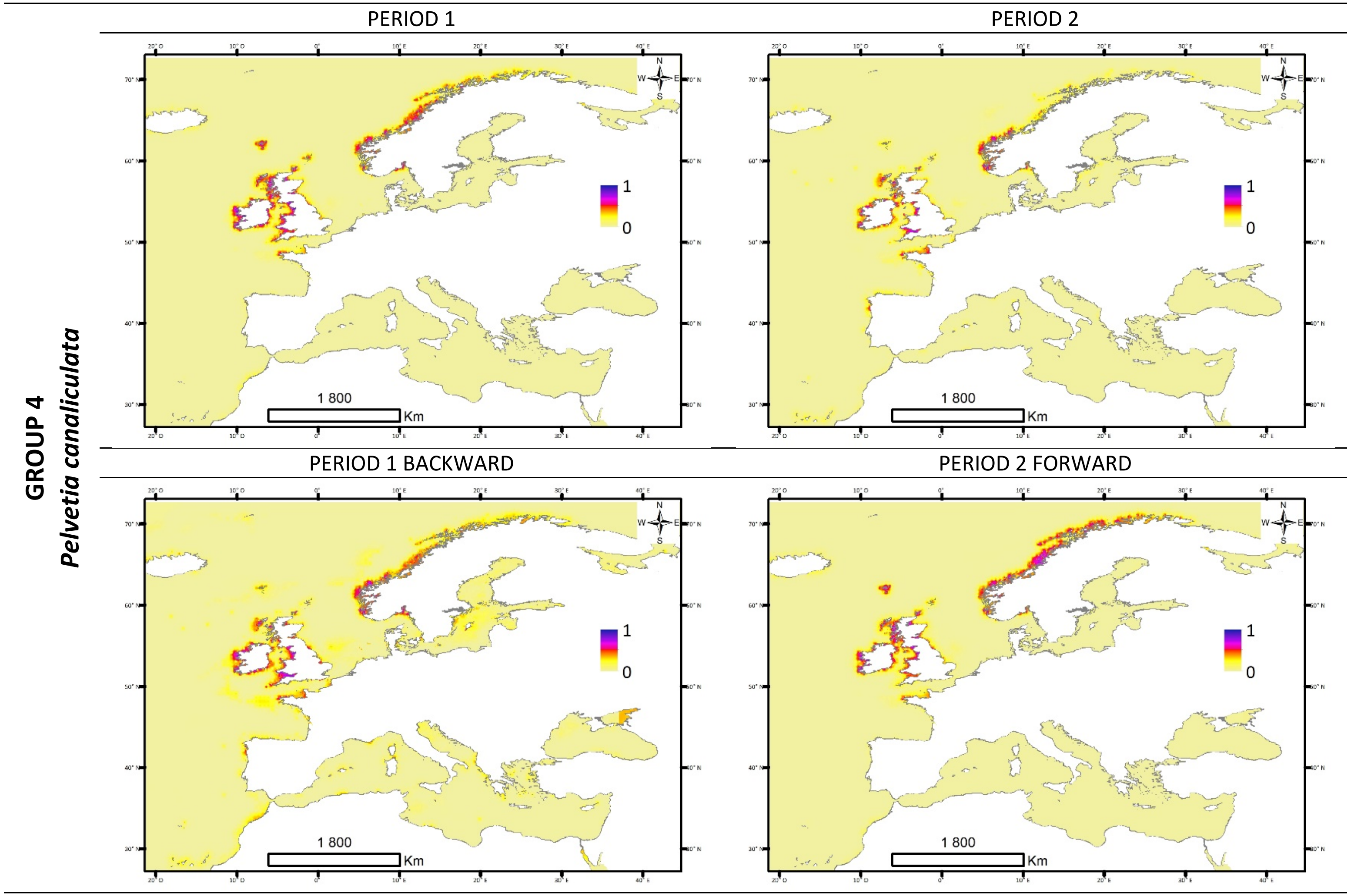


PERIOD 1 BACKWARD



PERIOD 2 FORWARD





GROUP 4
Pelvetia canaliculata

PERIOD 1

PERIOD 2

PERIOD 1 BACKWARD

PERIOD 2 FORWARD

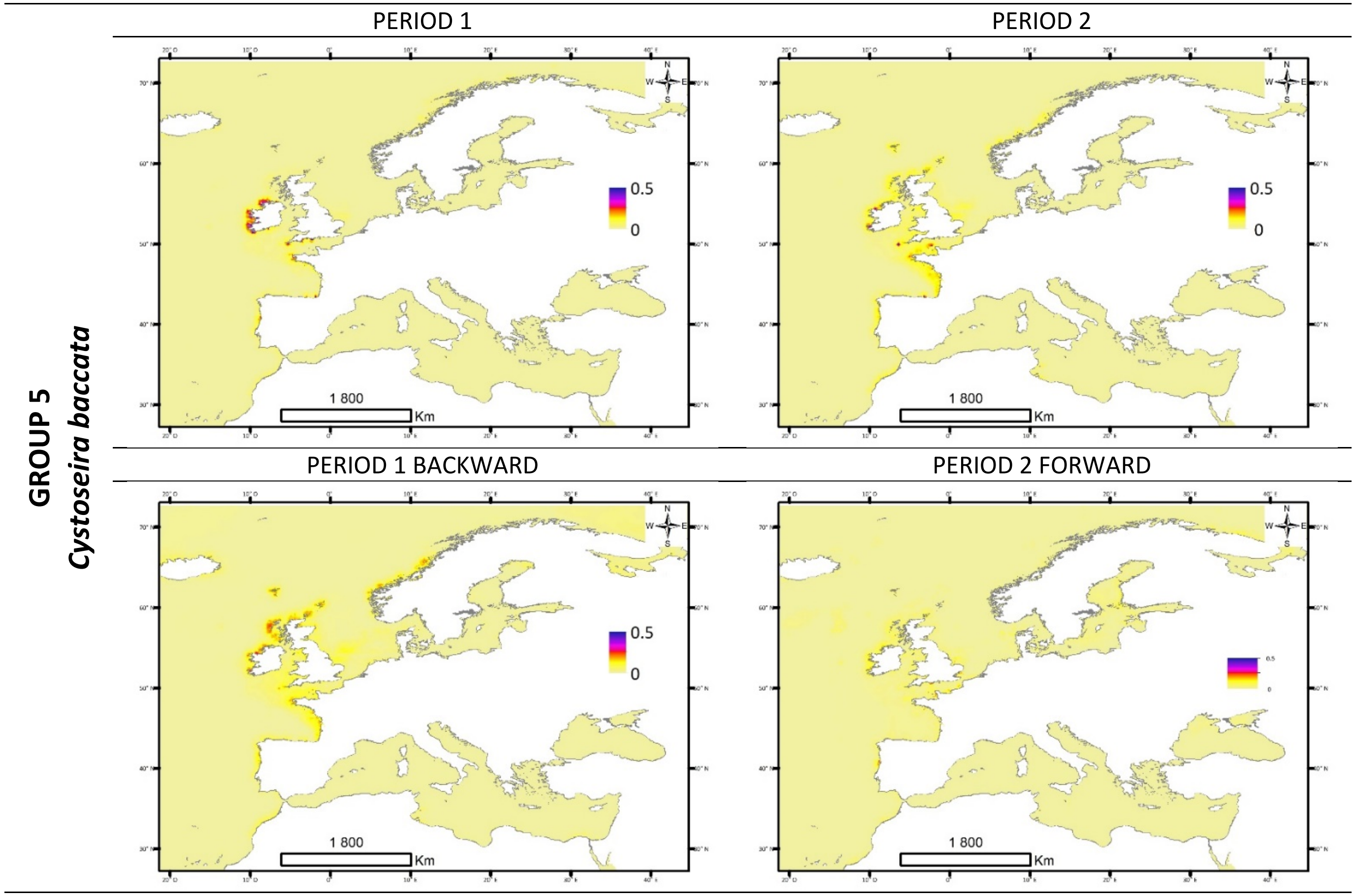


Figure 4.9. Distribution predicted by committee weighted ensemble for the four periods, from left: period 1, period 2, period 1 backward, period 2 forward for the five key species considered: *Saccorhiza polyschides*, *Gelidium spinosum*, *Sargassum muticum*, *Pelvetia canaliculata*, *Cystoseira baccata*. Notice that the bar scale is the same for all species (0 for low values and 1 for high values), but for *Cystoseira baccata* it reaches only 0.5 as maximum value.

3. Discussion

This chapter evaluates algorithms and weights selection in related to model transferability across time in the marine environment. Mathematical and visual results show the good predictive performance and transferability of the models, which outputs are ecologically realistic. The proposed method has the potential to become a time-and cost-effective tool for conservation and management, especially for the seaweeds considered for which settings have already been defined. In addition, the methodology is flexible enough to be applied to other species or geographical areas and, consequently, it has a great potential for climate change studies.

3.1. Modelling

Predictive accuracy varied for each group from one algorithm to another because their specific ability to capture species–environment relationships (Araujo et al., 2005; Randin et al., 2006). In general, in agreement with other authors (Elith *et al.*, 2006; Franklin, 2009; Thibaud *et al.*, 2014), machine learning algorithms produce models with higher sensitivity than regression-based approaches because of their more efficient management of complex species-environment relationships (Thibaud *et al.*, 2014). On the other hand, SVM, MARS and FDA presented low sensitivity. Concerning SVM, it works as an envelope method, using the outlying observations in environmental space to enclose in a larger subsets of environments (Peterson *et al.*, 2011). It is designed to find the hyperplane that maximally separates two target classes (Guo *et al.*, 2005) and the lack of absence data in one-class models results in an over-prediction of distributions (Zaniewski *et al.*, 2002; Engler *et al.*, 2004). Other algorithms that were discarded in this study were FDA and MARS. FDA is an extension of linear discriminant analysis that allows to replace the linear combination, which is not always relevant in SDMs, with non-parametric function such as MARS (Guisan *et al.*, 2017). MARS can fit response curves with levels of complexity similar to GAMs, but their link functions consist of series of connected straight line segments, in contrast to the smooth curves of GAMs (Peterson *et al.*, 2011). As a consequence, MARS projections showed unrealistic distributions, as has been already reported at regional level (Iturbide *et al.*, 2018). When comparing the performance of different algorithms, sample size is a key aspect (Wisz *et al.*, 2008; Thibaud *et al.*, 2014; Breiner *et al.*, 2018). In the group 5, the sample was small, although MAXENT and RF showed to be much less affected by this issue than regression algorithms (Aguirre-Gutiérrez *et al.*, 2013; Thibaud *et al.*, 2014).

Nevertheless, another aspect that it is necessary to consider when comparing modelling algorithms, is that sensibility test to different settings were not outperformed in this work and need further research in comparable conditions to define their impact (Robinson et al., 2011; Breiner, Nobis, Bergamini, & Guisan, 2018).

AUC and TSS showed that for all groups models' accuracy was good (Swets, 1988; Araujo et al., 2005; Engler et al., 2011). For both metrics, the group 1 showed the lowest model performance and the highest variation, due to the low values of CART algorithms. CART low performance ability was captured in its low weight compared to RF, demonstrating the significance of weights.

Several authors have established that ensemble methods outperform individual models (Marmion *et al.*, 2009; Grenouillet *et al.*, 2011; Crossman *et al.*, 2012; Guisan *et al.*, 2017), even for small samples sizes (Breiner *et al.*, 2018), particularly weighted average consensus methods provide the most robust predictions (Marmion *et al.*, 2009). However, the selection of the weights to consider is a key step, still not resolved for general applications. Therefore, the weights assessed in this study can be a valuable contribution to marine SDMs in order to reduce uncertainty and improve accuracy from ensemble modelling under future scenarios.

3.2. Transferability

Working with transferred models is critical, especially in the context of climate change studies (Regos *et al.*, 2019), because good performance in the internal evaluation does not guarantee transferable models (Huang *et al.*, 2016). A good option to solve this problem is using models that have already been successfully transferred to other time and validated with independent data (Sequeira *et al.*, 2018), which are also able to capture the inter-annual variability in each period (Tuanmu *et al.*, 2011; Huang *et al.*, 2016). Both aspects were considered in this study, with 12 years data for each period and a gap of seven years to ensure the independence of time frames (Dixon, 1966; Chapman & Chapman, 1980; Lüning *et al.*, 1990; Little & Kitching, 1996). According to Tuanmu et al. (2011), models have good temporal transferability, with differences between groups, as it is species-specific (Randin *et al.*, 2006). For example, group 5 presented the lowest values, which could be a consequence of the low number of records available, suffering from transferability estimation bias (Huang *et al.*, 2016). Therefore, the results of this group need to be considered carefully.

Another issue to take into account is that temporal transferability involves extrapolation into a new environmental space (Elith & Leathwick, 2009; Schibalski *et al.*, 2014). Good results may be expected in this case, assuming high niches overlapping between two periods (Broennimann *et al.*, 2012; Sequeira *et al.*, 2018). As well as climate analogy, environmental availability assessment need to be done when the range of predictors in the object period extends beyond that encountered during model building to detect the areas where the climate is not analogous and predictions have to be considered carefully

(Sequeira *et al.*, 2018). Nevertheless, in this work, models were trained in a larger area than Europe (Figure 1.3, Chapter I) in order to include a large combination of environmental conditions in both periods (data-rich scenario according to Sequeira et al. (2018)). Therefore, results are useful to better understand the factors affecting transferability (Barbet-Massin *et al.*, 2010)y. In the context of climate change, this is especially relevant as the changes of a predictor's importance over time can modify the interactions between climate and species (Schibalski *et al.*, 2014), and the knowledge about how this will vary across time is still lacking (Zharikov *et al.*, 2009). Each algorithm captures it in a different way (Araujo *et al.*, 2005; Randin *et al.*, 2006; Beaumont *et al.*, 2016), and therefore the model with the highest accuracy in present situations may not necessarily be the best one for predictions in another time (Randin *et al.*, 2006). In these situations, transferability assessment and ensemble modelling are especially powerful (Guisan *et al.*, 2017) and, consequently, the weighted ensemble by their temporal transferability ability proposed in this work draws inference in a probabilistic way to better inform modellers selection decisions.

3.3. Ecological realism

The minimum requirement for a model that aims to explain patterns is that it is robust in its ecological assumptions. In this study, results are ecologically strong in the two main steps of the methodology.

First, this study highlights the relevance of including hydrodynamic variables, whose link with marine species have been reported by different authors (Ramos *et al.*, 2014; de la Hoz *et al.*, 2018a; Jonsson *et al.*, 2018). It was clear the importance of these variables in mid-intertidal and subtidal species, with a high contribution of Tm, Uw and Urms in groups 2, 3 and 5 but not in group 4, characterized by high TR (Figure 4.3). It is also remarkable that all groups exposed to desiccation presented a high importance of parameters derived of AT, Ws and TR, especially their maximum and variability indicators. This highlights their significant role in species survival, although in marine studies they are not usually considered. Furthermore, this work also empathizes the decisive role of extreme conditions, whose parameters were present in all groups, supporting the importance of considering stressful conditions (Galván *et al.*, 2016; Torresan *et al.*, 2016; Assis *et al.*, 2017a; Vranken *et al.*, 2018). Besides, the relevance of widely used variables, such as SST was also confirmed by results (Bosch *et al.*, 2017). It can be said that the methodology presented was successful in predicting the distribution of important habitat, as the patterns of distribution were well reproduced in the projected periods (P1b and P2f) (Figure 4.9).

The methodological approach presented here has the potential to become a time and cost-effective tool for climate change studies in marine areas, as it guides researchers in the algorithms and their weights selection to develop models temporally transferable.

The proposed method has nevertheless sound limitations that require further development, such as those linked to: 1) work with presence-only data (and in the case of group 5 with a small sample size), 2) the assumptions that species are in equilibrium with their environments, and 3) the need to test algorithm settings and records bias; the results presented here show the ecological significance of the methodological proposal, which guarantees the accuracy and transferability of model outputs.

Results are satisfactory and robust for being broadly applied in marine research with different objectives and geographical areas, allowing a comparative framework between species predictions and promoting the use of transferable models, especially in climate change studies.

Chapter V

Climate change induced range shifts in seaweeds distributions in Europe

Chapter V. Climate change induced range shifts in seaweeds distributions in Europe.

This chapter is an edited and modified version of the research article published in 2019 in the journal Marine Environmental Research, vol. 148, pp. 1-11, by de la Hoz, C.F.; Ramos, E; Puente, A. and Juanes, J.A. with the title "Climate change induced range shifts in seaweeds distributions in Europe".

1. Introduction

There are evidences of how climate change is affecting marine communities and consequently, the ecosystem services they provide (Doney *et al.*, 2012; Liquete *et al.*, 2016). Among marine flora, seaweeds dominate intertidal and subtidal environments (Lüning *et al.*, 1990). Their role as engineering species defines the structure and functioning of the benthic assemblages through biotic and non-biotic interactions (Schiel, 2006; Méndez-Sandín & Fernández, 2016). Therefore, a comprehensive understanding of consequences of climate change in their distributional patterns is a core requirement for adopting an ecosystem-based management (Sax *et al.*, 2013), such as resources management, establishment of marine protected areas or measures to face invasions (Tamburello *et al.*, 2014; Duarte *et al.*, 2017). To reach this objective it is required the identification of the variables and parameters that determine species ecology and the definition of their physiological thresholds. When working in large areas, hydrological variables (e.g. wave height, current speed or bottom and wind stress) are usually not considered, although there is evidence of their relevance in species distribution (Ramos *et al.*, 2014; Pace *et al.*, 2017; de la Hoz *et al.*, 2018a). Likewise, the formulation of biologically-meaningful parameters is essential to get realistic models, but abiotic information availability is usually limited to general parameters, not taking into account for example extreme events (althoug see Marshall *et al.,* 2013; Smale & Wernberg, 2013; Cavole *et al.,* 2016; Hughes *et al.,* 2017), whose importance in climate change context will be decisive (Seabra *et al.*, 2015). Concerning species tolerance thresholds, experimentation provides valuable information (Martínez *et al.*, 2012b; Celis-Pla *et al.*, 2017) but extrapolation from laboratory to field conditions can be tricky. One of the most applied tools to deal with this are SDMs, whose applications in recent years have increased significantly because they map efficiently large areas where field-collected data are limited or unavailable (Álvarez-Martínez *et al.*, 2018). They are based on a quantification of species-environment relationships by using numerical tools (Elith & Leathwick, 2009). Nevertheless, their implementation in decision-support systems is sometimes limited due to managers' difficulty to develop them.

The objective of this chapter is to provide an overview of climate change driven range shifts of representative seaweeds in Europe in order to reach the general objective of this thesis. Besides, dependence plots are expected to be generated to improve the knowledge of species physiological thresholds in a cost-effective way.

2. Material and methods

The methodology presented is divided into three main steps (Figure 5.1):

1) Data selection and quality control. Predictors related to seaweed ecology (environmental data) and biological records of five macroalgae (distribution data) were collected from 1985 to 2015.

2) Modelling. The best algorithms for each species were selected and their settings adjusted by including only the most explanatory predictors (with a contribution higher than 5%).

3) Assessment. Model performance metrics were calculated and models assessed in the environmental and geographic space. Besides the ecological realism of results was tested.

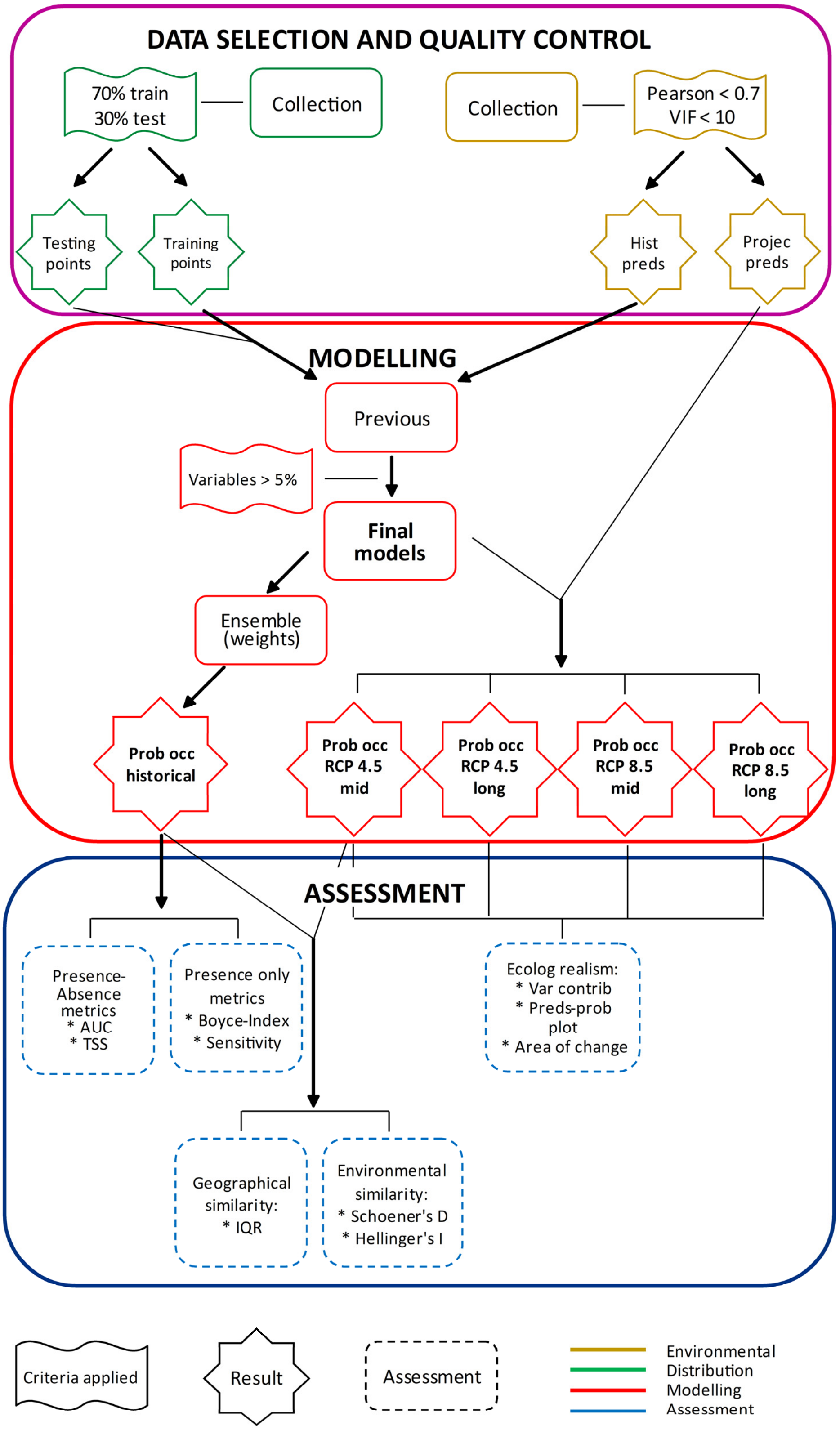


Figure 5.1. Diagram methodology followed in this work. Wave rectangles represent the criteria applied, stars the results and dashed rectangle the assessment techniques.

2.1. Distribution data

Five species were considered in this study: *Saccorhiza polyschides*, *Gelidium spinosum*, *Sargassum muticum*, *Pelvetia canaliculata* and *Cystoseira baccata*. They were selected as representative of 21 seaweeds (Table 3.2, Chapter III) defined by grouping them according to results of Chapter IV.

Biological records were collected from 1985 to 2015 from the sources of Table 3.3 (Chapter III). Only records with geographic information were considered. Besides, all presences were checked by experts to remove possible geographic errors. Occurrences were gridded and only one per cell was considered. Representative training points were selected (70%), using 30% of data for testing (Figure 5.2).

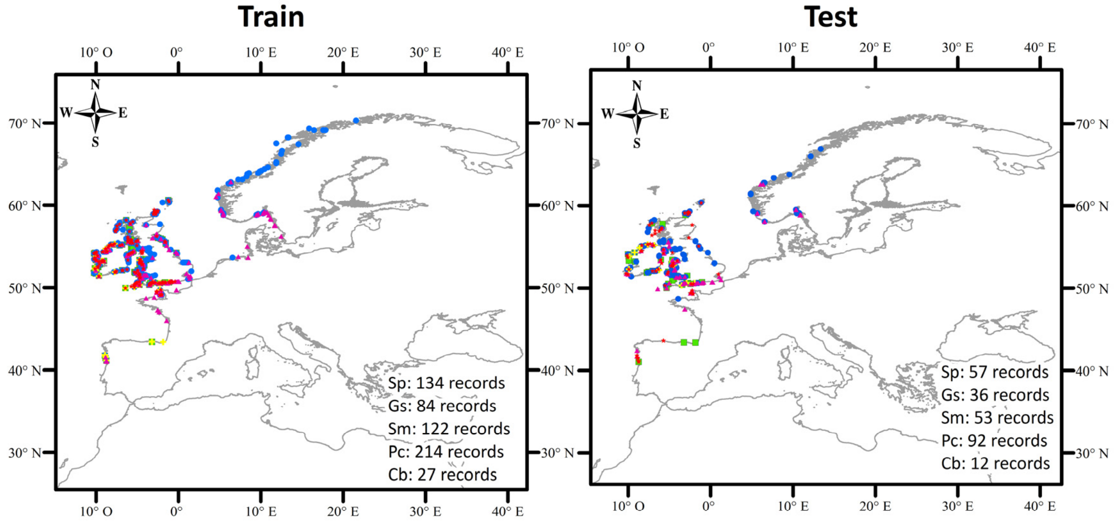


Figure 5.2. Presences records used for training and testing. *S. polyschides* is represented with red stars, *G. spinosum* with green *squares, S. muticum with pink triangles, P. canaliculata with blue circles and C. baccata* with yellow crosses.

Pseudo-absences were randomly generated in the geographic space (1000 points), not overlapping presence sites (Barbet-Massin *et al.*, 2012; Guisan *et al.*, 2017), with the R package *sdm* (Naimi & Araujo, 2016).

2.2. Environmental data

Abiotic variables were collected from OCLE database (de la Hoz *et al.*, 2018b) for each mesh point. For the historical period (1985-2015) 15 variables were considered: sea surface temperature, significant wave height, salinity, air temperature, wind speed, tidal range, bottom shear stress, bottom orbital speed, currents, radiation, light attenuation coefficient, pH, nitrate, bathymetry, substrate (categorised in mud, sand, coarse substrate, mixed sediment and rock). Nine variables were projected to the RCPs 4.5 and

8.5 (IPCC, 2014) for the mid-term (2040-2069) and the long term (2070-2099): sea surface temperature, significant wave height, salinity, air temperature, wind speed, radiation, pH, nitrate, sea level rise. Units and periods of raw data can be consulted in Table 2.1. For each variable, a complete set of parameters was considered: maximum, minimum, mean, standard deviation, range and percentiles 10th, 25th, 50th, 75th and 90th. Additionally, to detect changes in extreme conditions for SST and AT the number of consecutive days that the percentile 90th is exceeded (Torresan *et al.*, 2016) and for Tm the number of days over 2.2 Nt/m^2 (Vousdoukas *et al.*, 2012) were calculated. Furthermore, for projections, the same parameters were calculated seasonally, yearly and for the full period.

Collinearity of parameters was tested for the 642 parameters with the criteria of Pearson correlation coefficient < 0.7 and VIF < 10 in the points with species presence, using the R package *usdm* (Naimi *et al.*, 2014). VIF (Montgomery & Peck, 1982) values from five to ten are considered as critical for multi-variable correlation (Guisan *et al.*, 2017).

2.3. Modelling

Nine correlative algorithms were ran (GLM (McCullagh & Nelder, 1989), GAM (Hastie & Tibshirani, 1990), MARS (Friedman, 1991), MDA (Hastie *et al.*, 1994), CART (Breiman, 1984), SVM (Vapnik, 1995), FDA (Hastie *et al.*, 1994), RF (Breiman, 2001) and MAXENT (Phillips *et al.*, 2006) in two independent periods to test their temporal transferability. For each species, the algorithms with high sensibility when projected and their weights were retained according to results of Chapter IV (Table 4.7).

A preliminary model was built with all non-correlated predictors and those with a contribution higher than 5% were included in the final models (Coro *et al.*, 2015). New models were ran with the selected predictors and the same training and testing points and projected into the historical and climate change conditions. An ensemble (Araujo & New, 2007) was built for each species by averaging the results of the algorithms according to weights (Figure 5.1).

The models were built, evaluated and projected using the *sdm* package (Naimi & Araujo, 2016).

2.4. Assessment

Models were assessed considering different metrics to provide information about their predictive power (Figure 5.1). First, predicted probabilities of presence versus presence-absence observations were calculated, through the area under the AUC (Hanley & McNeil, 1982) and the TSS (Allouche *et al.*, 2006). AUC varies from 0.5 for a model whose predictions are no better than random, to 1 for perfect agreement with the observed data. TSS varies between 0 (random model) and 1 (perfect agreement). Models with AUC> 0.9 were considered "very good", in the range 0.7-0.9 "reasonable predictions", and values <0.7 "poorly accurate" (Swets, 1988). For TSS the predictive power of models were classified as "poor" for values < 0.4, "good" in the range 0.4-0.8 and "excellent" from 0.8-1 (Allouche *et al.*, 2006). Analysis of AUC and TSS were performed in R, using the *pROC* (Robin *et al.*, 2011) and *sdm* (Naimi & Araujo, 2016) packages, respectively. Additionally, model overfitting was calculated as the difference between training and testing AUC, with higher values indicating over-parameterizing (Warren & Seifert, 2010; Jiménez-Alfaro *et al.*, 2018).

Probabilistic predictions were compared to presence-only observations with the Boyce index (Boyce *et al.*, 2002). Besides, as the goal of the study was to detect species and their areas at risk in climate change scenarios, sensitive models were desired, therefore the CCR was used as threshold to define presences.

As different abiotic combinations exist between the historical and projected scenarios, niche overlap was tested on the environmental conditions and on the geographical area. In the first case, the assessment was conducted through Schoener's D (Schoener, 1989) and modified Hellinger's I metrics (Warren *et al.*, 2008) with the PCA-env ordination in the R package *ecospat* (Broennimann *et al.*, 2016). They range from 0 (no overlap) to 1 (identical niches), values higher than 0.6 are considered as high overlap (Broennimann *et al.*, 2012). On the other hand, the geographical similarity was assessed through the IQR (Acuna & Rodriguez, 2004) to detect the areas where the climate is non-analogous, therefore predictions in those areas have to be considered carefully. IQR was calculated in each point of the mesh for historical and projected maps and values lower than Q1-1.5(IQR) and higher than Q3+1.5(IQR) were considered as outliers.

Concerning the ecological realism of models output, different analyses were conducted in R (R Core Team, 2018). For each species the contribution of predictors to model were obtained by averaging the algorithms results according to weights (Table 4.7). Besides, the fitted model of each species was applied separately for each predictor to get the physiological thresholds. Additionally, the relative area of change between periods was calculated after reclassifying the output using CCR as threshold. Additionally, to mathematical assessment, results were also visually evaluated.

3. Results

3.1. Assessment metrics

Final models' goodness-of-fit and discrimination performance could be considered very good according to AUC and TSS, with values over 0.93 and 0.84 respectively (Table 5.1). Presence-only metric Boyce-Index showed close to perfect predictions, with values over 0.91 in all cases. Boyce index plots are represented in Figure S5.1. Sensitivity can also be considered good for all species, with more that 60% of presences well predicted, except in the case of *P. canaliculata* because of the high value of the CCR used as threshold (0.74) (Table 5.1). The difference between training and testing AUC (AUC_{diff}) showed low values for all species, indicating that the risk that models were over-parametrized was small.

Table 5.1. Summary of the assessment metrics by species.

	S. polyschides	***G. spinosum***	***S. muticum***	***P. canaliculata***	***C. baccata***
AUC	0.93	0.98	0.99	0.97	0.98
TSS	0.84	0.93	0.93	0.86	0.94
BOYCE INDEX	0.99	0.94	0.99	0.99	0.91
CCR	0.43	0.35	0.11	0.74	0.31
SENSITIVITY	73.30	61.67	100	11.44	69.23
AUC_{diff}	0.039	0.007	0.004	0.0002	0.019

Niche overlaps comparison showed that abiotic environments were more different for the RCP 8.5 than for RCP 4.5 and for long term than mid-term, showing a clear gradation in both metrics (Table 5.2). According to Figure S5.2, a medium overlap can be considered in general, although it decreased in the RCP 8.5 and the long term (Broennimann *et al.*, 2012).

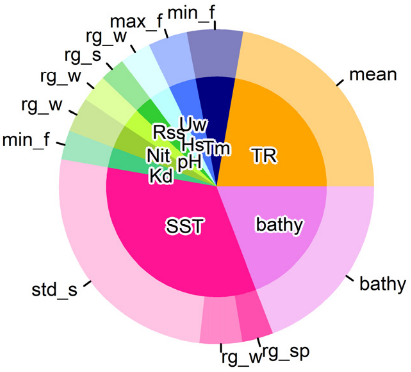
Table 5.2. Overlap values for the RCPs considered.

	RCP 45m	**RCP 45l**	**RCP 85m**	**RCP 85l**
Schoener's D	0.45	0.40	0.39	0.29
Hellinger's I	0.71	0.63	0.61	0.56

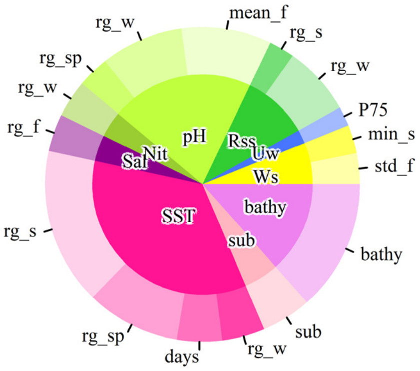
3.2. Ecological realism

Concerning predictors contribution (Figure 5.3), it was remarkable the importance of hydrodynamic variables, such as Tm, Uw, Urms or Hs, present in all models. Besides, SST also contributed in all models except in *P. canaliculata*'s. Desiccation variables, as AT or Ws, were also important for all species, except for *S. polyschides*. Variables such as Kd, pH, bathymetry, substrate or Rss were important for all species with different percentage. Concerning parameters, extreme and variability indicators were the best represented. Numerical value of predictors contribution can be checked in Table S5.1. Note that substrate categories were represented by numbers, being 1 mud, 2 sand, 3 coarse, 4 mixed and 5 rocky.

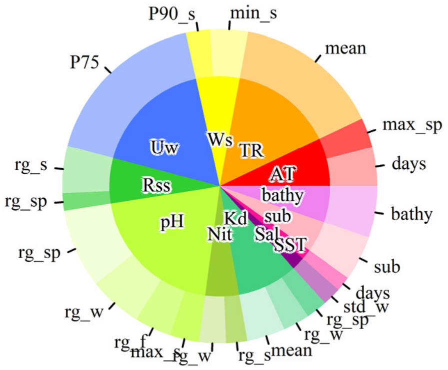
Saccorhiza polyschides



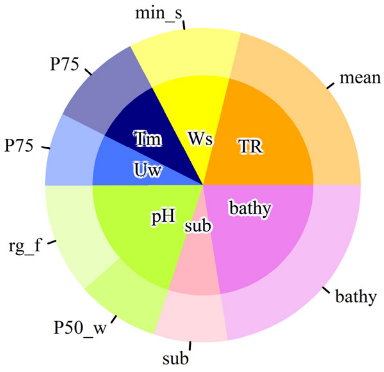
Gelidium spinosum



Sargassum muticum



Pelvetia canaliculata



Cystoseira baccata

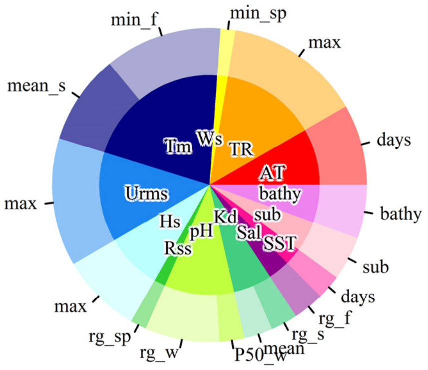


Figure 5.3. Predictors' contribution by species. Inner circles represent the variables and external subdivisions their parameters.

Dependence plots (Figures S5.3 – S5.7) showed species physiological thresholds for each predictor, such as the high probabilities in semi-exposed rocky shores and high intertidal of *P. canaliculata*, sheltered areas in the case of *C. baccata* and hard substrates for *S. muticum*. Besides species' tolerance to predictors is clearly defined by the different probabilities of presence. For example, it is exposed *S. muticum*'s tolerance to high values of currents and temperatures and low variability in nitrate.

According to the predicted geographical distribution of species, Figures 5.4-5.8 showed the difference in the probability of occurrence species between RCPs and historical predictions. In addition, all RCPs and periods were represented in Figures S5.8-S5.12 with different scales to allow a better visualization. Visual assessment showed that for *S. polyschides* and *G. spinosum* a remarkable decrease in habitat suitability was predicted in the whole distribution area, reaching values lower than 0.2 in the RCP 8.5 long term. *S. muticum* southern areas reduce their suitability, until disappear in the Iberian Peninsula; meanwhile probability increase in northern areas. *P. canaliculata* probabilities of occurrence in the west of Britain might decrease slightly, whereas the suitability of some areas around the Brittany and the English Channel increase. The suitable habitat for *C. baccata* was predicted to increase in all RCPs around UK and the Gulf of Biscay. Only some isolated points were detected as outliers in the IQR assessment (Figure S5.13), which usually correspond to low probabilities, therefore results can be considered accurate.

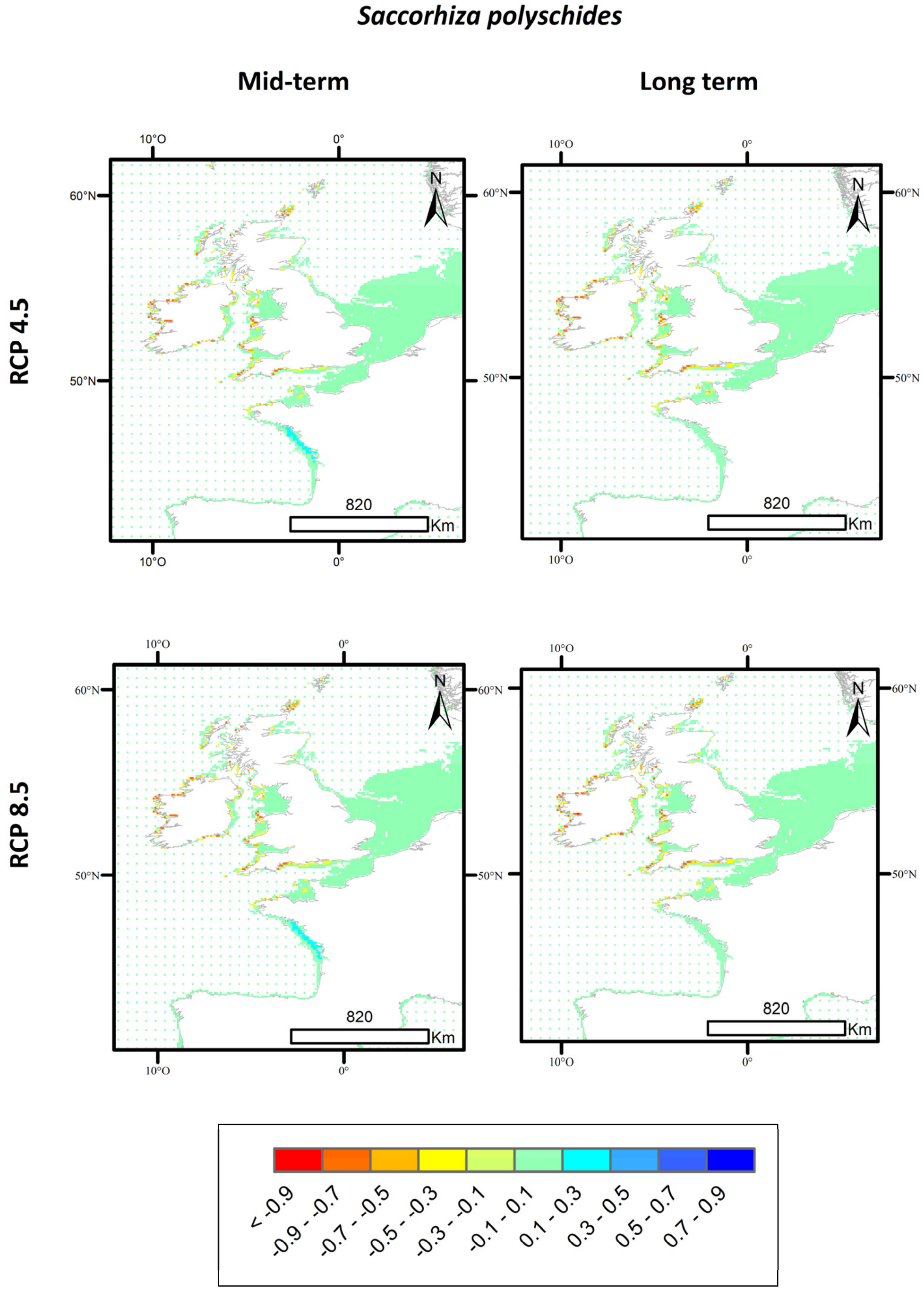


Figure 5.4. Differences in the probability of presence predicted by the weighted ensemble of predictions (RCP 4.5 and 8.5 in the mid and long term) and historical period for Saccorhiza polyschides. Positive values (in blue) indicate gains in the RCPs; meanwhile negative values (in red) represent losses. Note maps have been zoomed to the areas where values were different to zero to facilitate the visualization.

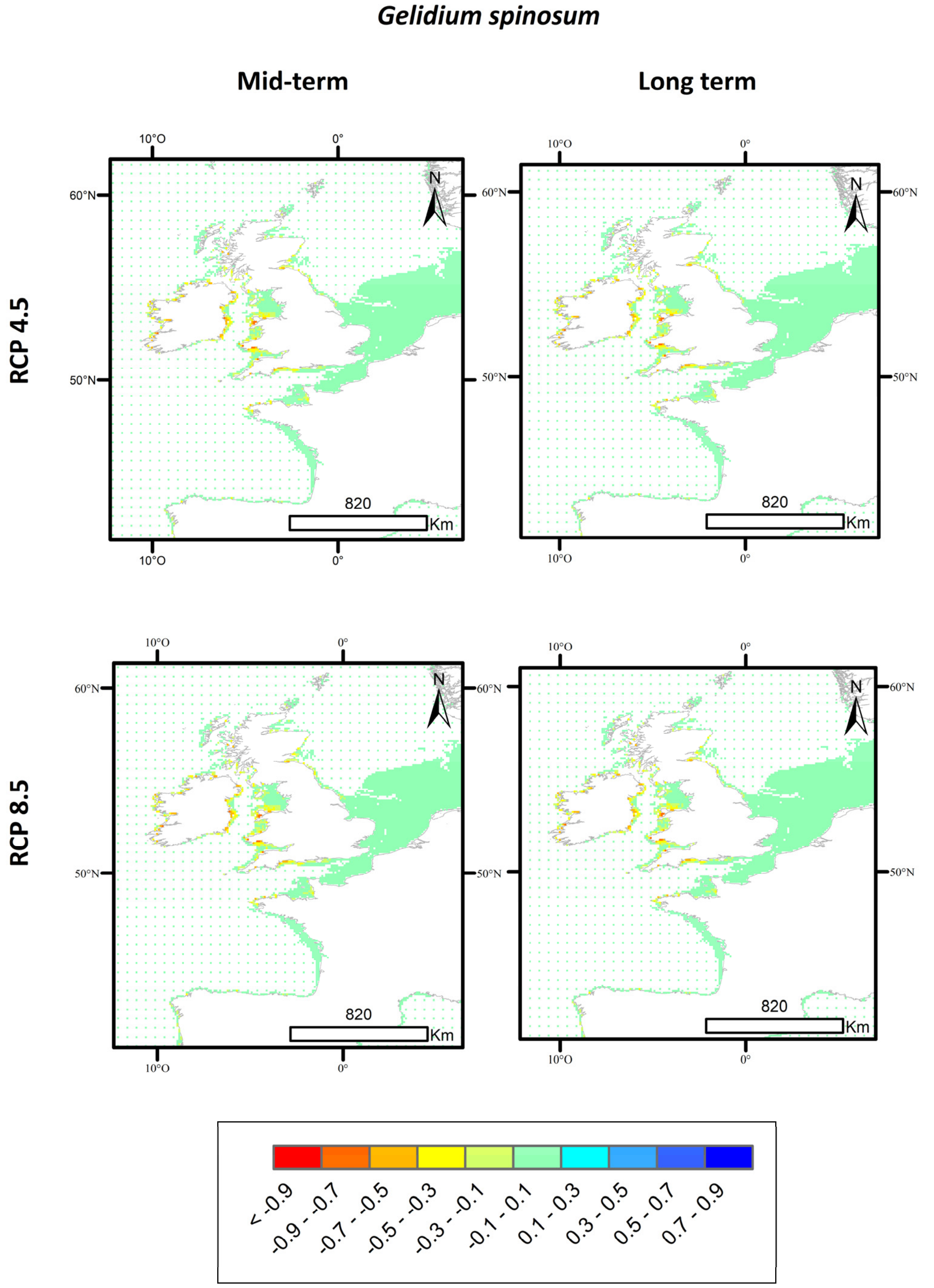


Figure 5.5. Differences in the probability of presence predicted by the weighted ensemble of predictions (RCP 4.5 and 8.5 in the mid and long term) and historical period for Gelidium spinosum. Positive values (in blue) indicate gains in the RCPs; meanwhile negative values (in red) represent losses. Note maps have been zoomed to the areas where values were different to zero to facilitate the visualization.

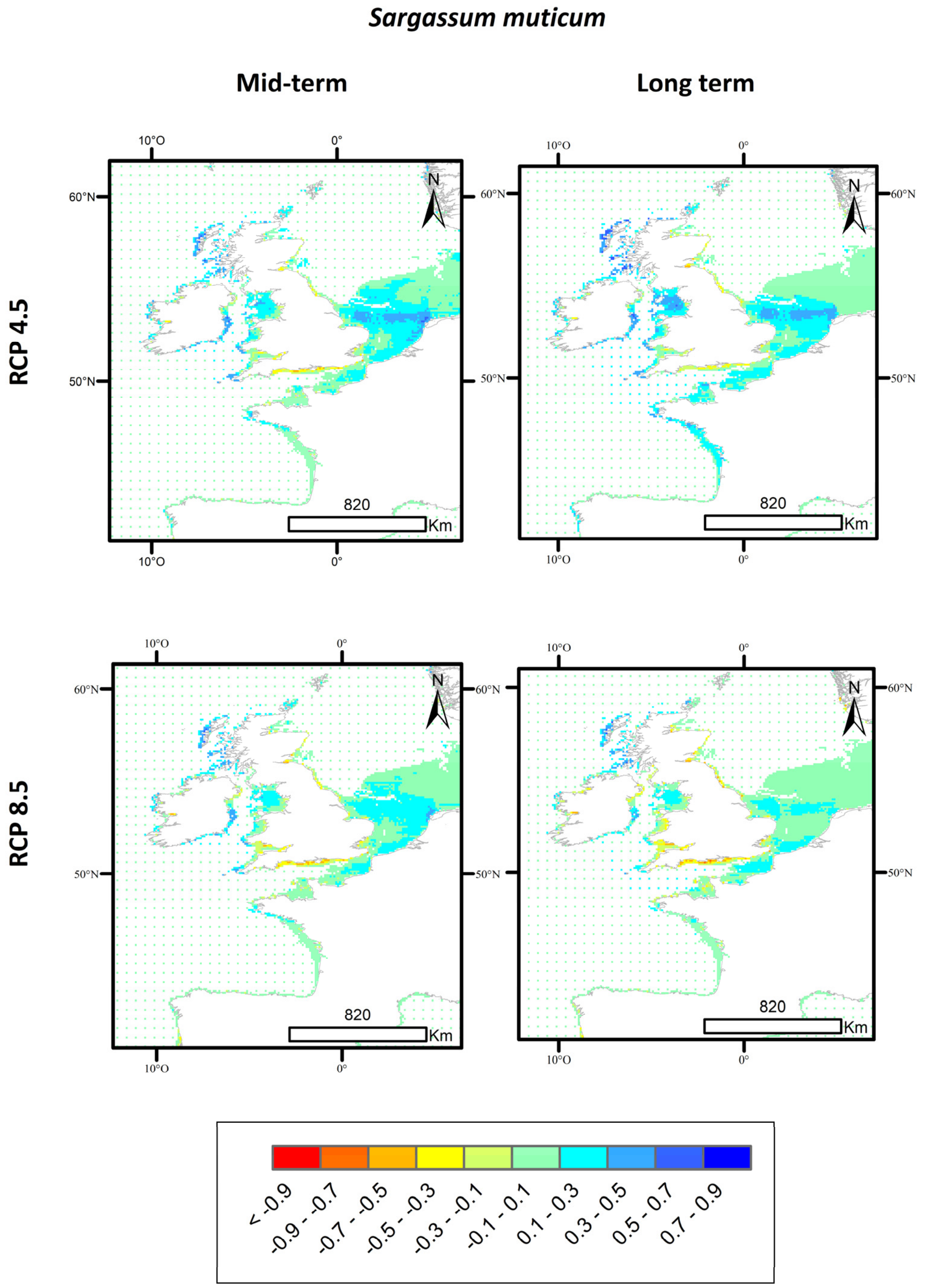


Figure 5.6. Differences in the probability of presence predicted by the weighted ensemble of predictions (RCP 4.5 and 8.5 in the mid and long term) and historical period for *Sargassum muticum*. Positive values (in blue) indicate gains in the RCPs; meanwhile negative values (in red) represent losses. Note maps have been zoomed to the areas where values were different to zero to facilitate the visualization.

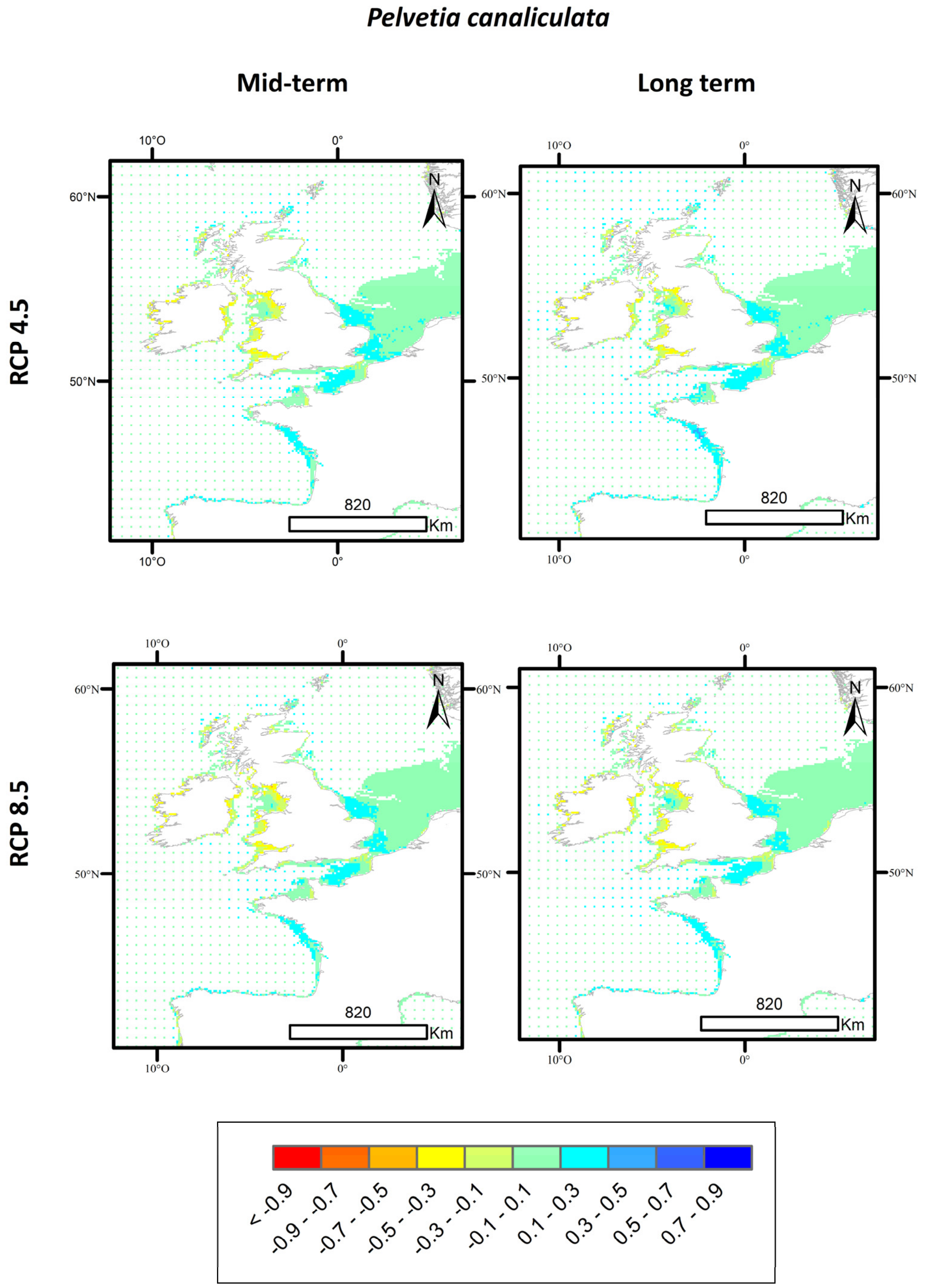


Figure 5.7. Differences in the probability of presence predicted by the weighted ensemble of predictions (RCP 4.5 and 8.5 in the mid and long term) and historical period for Pelvetia canaliculata. Positive values (in blue) indicate gains in the RCPs; meanwhile negative values (in red) represent losses. Note maps have been zoomed to the areas where values were different to zero to facilitate the visualization.

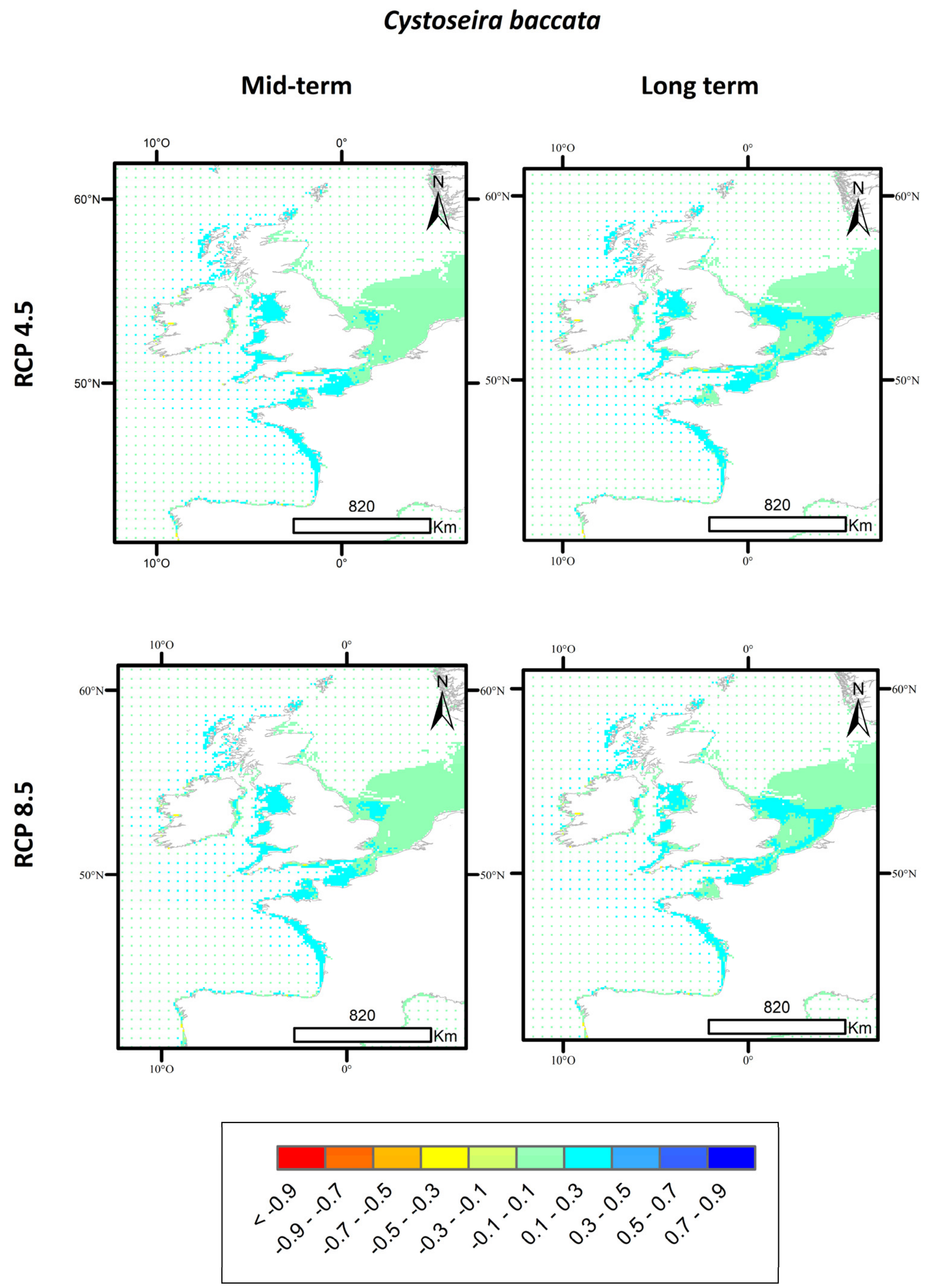


Figure 5.8. Differences in the probability of presence predicted by the weighted ensemble of predictions (RCP 4.5 and 8.5 in the mid and long term) and historical period for Cystoseira baccata. Positive values (in blue) indicate gains in the RCPs; meanwhile negative values (in red) represent losses. Note maps have been zoomed to the areas where values were different to zero to facilitate the visualization.

4. Discussion

Models developed in this work predicted, with good performance, important changes in the habitat suitability of the five seaweeds considered. According to Chapter IV, results can be transferable to 21 species and broadly applied in Europe, becoming a very useful management tool in wide-ranging climate change studies in the marine environment. One of the main contributions of this work is the inclusion of the most complete set of parameters in seaweeds' assessment until now (de la Hoz *et al.*, 2018b), which improves the existing knowledge in inferring the environmental forces shaping their distributions.

Despite not being usually considered in SDMs, this study reinforces previous findings about the importance of hydrodynamic variables in marine flora distribution (Ramos *et al.*, 2014; de la Hoz *et al.*, 2018a; Jonsson *et al.*, 2018). Besides, the relevance of widely used variables, such as SST was also confirmed (Bosch *et al.*, 2017; Martínez *et al.*, 2018b; Piñeiro-Corbeira *et al.*, 2018). Results remark the decisive role of extreme conditions in seaweeds distribution, supporting the importance of considering stressful parameters (Galván *et al.*, 2016; Assis *et al.*, 2017a; Vranken *et al.*, 2018).

Previous studies have already established the importance of SST for *S. polyschides* (Fernandez, 2011), especially focused on their maximum during growing season. Additionally, this work remarks the importance of variability (Werner *et al.*, 2016), reason why it was reported to survive only in areas with narrow temperature range (Lüning *et al.*, 1990). *S. polyschides* lives in the subtidal, but often extends on to the lower shore, which explained the importance of TR as a limiting factor. The model selected also nitrate availability as a critical factor of *S. polyschides* growth. The importance of Tm minimum in fall shows its preference for areas of low energy, although it is able to tolerate strong currents (Norton, 1978). Concerning its geographical distribution, for the historical period it was confined to the coasts of UK and the Brittany, in agreement with current knowledge (Burrows *et al.*, 2018). However, the model did not identify the records in the north and west coast of the Iberian Peninsula reported by other authors (Fernandez, 2011; Assis *et al.*, 2013). This can be the consequence of the scarcity of data in this area during the training process, which could have excluded some environmental combination important for the area. Projections agreed with previous studies that suggest their practically disappearance, especially towards the southern edge (Méndez-Sandín & Fernández, 2016; Assis *et al.*, 2018). Meanwhile other studies, such as Assis *et al.* (2018), detect the northward migration of *S. polyschides* in the RCP 8.5, our results suggest an important disappearance in the following decades, although some migration in depth would happen for the populations situated below the summer thermocline (Assis *et al.*, 2016) because low probabilities remained in UK and Brittany. This quick range shift was in the line of the field observations carried out in the north of

Spain, where after 30 years populations had shorter generations and a smaller fraction becomes reproductive (Fernandez, 2011).

The link of *G. spinosum* with SST has been studied (Rueness & Fredriksen, 1989), however its variability is usually not considered despite its importance according to this work. Besides, results showed the importance of extreme SST, with the inclusion in the model of the Days parameter, whose high values decreased *G. spinosum* probability of occurrence. This lethal effect of high temperatures has already been reported (Fralick *et al.*, 1990), however the consecutive effect has not been considered until now. Light is an important factor to *G. spinosum* (Fralick *et al.*, 1990), especially its range in extreme seasons, winter and summer. Its historical distribution presented higher probabilities of occurrence in the north populations (Ireland and west of UK) than in southern ones (Brittany and Iberian Peninsula). Models predictions suggested an extinction of the species under both RCPs considered, even in the mid-term. This trend toward decline has already been observed in its biomass between 1985 and 2002 (Sánchez *et al.*, 2005). These authors suggest a competition for light, which could be softened in climate change scenarios. However, for other *Gelidium* species high light stress triggers photobiological changes that could involve photoinhibition (Quintano *et al.*, 2019).

The most important parameter for *S. muticum* was Uw_P75, according to the studies that establish a high tolerance of *S. muticum* to currents, even at extreme conditions (Josefsson & Jansson, 2011). As it has to deal with desiccation, variables such as TR, Ws or AT were also included in the model. It was also remarkable the importance of the Days parameter in both temperatures, but coping better with high SST values than with AT ones (Belsher, 1989). Model also captured its preference by shallow waters and rocky substrates (Thomsen *et al.*, 2006). Reproduction occurs in spring and summer, reason why the variability in the light availability during these seasons were included in the model. As a consequence of its invasive condition it is probably not in equilibrium (Araújo & Pearson, 2005) and many variables influence its distribution. As SDMs assume this balance and the model was calibrated with data of the invaded range only, results must be considered carefully. Bearing this in mind, results agreed with existing knowledge, which suggests that model worked well. Historical distribution captured the existing invaded areas, except the Mediterranean coasts of France (Critchley *et al.*, 1990). No occurrence points were available in the Mediterranean for the analysis, which could be the reason of the underprediction in this area. Under climate change, model predicted a range expansion northward and an increase in its probability of occurrence. This trend continues the one already observed since its introduction in Europe (Karlsson & Loo, 1999; Stæhr *et al.*, 2000; Sánchez *et al.*, 2005; Engelen *et al.*, 2015; Yesson *et al.*, 2015).

P. canaliculata tolerates a wide range of exposure conditions (Lüning *et al.*, 1990), reason why wind speed and bathymetry had an important weight in the model and it could bear with a high range of TR. The final model incorporated both terrestrial and marine variables, as models of other authors (Neiva *et al.*, 2014). This fucoid is widely distributed in Europe and projections showed a decrease in its probability of occurrence, but some sites might be stable. This could be the consequence of the high resilience that high intertidal species present (Harker *et al.*, 1999; Skene, 2004). Besides, its probability increases in some areas, in agreement with the northward migration reported for the European Atlantic (Lima *et al.*, 2007; Lamela-Silvarrey *et al.*, 2012; Neiva *et al.*, 2014; Piñeiro-Corbeira *et al.*, 2016).

As *C. baccata* inhabits in a wide range of depth, exposure and sedimentation conditions (García-Fernández & Bárbara, 2016), its distribution was determined by terrestrial (TR, AT, Ws) and marine variables (SST, Urms, Tm, Hs). It is remarkable that most of the parameters selected for the model were the extreme ones, which highlighted its importance. Urms, Tm and Hs dependence plots supported the preference of *C. baccata* for sheltered areas, although it can tolerate semi-exposed conditions (Díez *et al.*, 2003; Méndez-Sandín & Fernández, 2016). The model captured its southern distribution in the Atlantic (García-Fernández & Bárbara, 2016), whose limits increase in the RCP 8.5. Together with *S. muticum*, they were the only species that gained area. These results agreed with the changes detected in the north of Spain between 1977 and 2007, with a lengthened of *C. baccata* period of growth and an increase of its biomass (Méndez-Sandín & Fernández, 2016), suggesting a maintenance of this trend. As the model was constructed with a low number of records, it must therefore be interpreted cautiously, although results were in close agreement with previous studies.

For all species, important changes occur in their range of distribution, which might modify the assemblage physical environment and the structure of the understory species. These shifts in species compositions have already been reported in the north of Spain (Fernandez, 2011; Díez *et al.*, 2012), which opens the gap of how biotic interactions would be modified and how the ecosystem services they provided would change.

Although uncertainties with regards to the modelled species distributional range shifts due to climate change are considerable because they are the product of a multifactor complex interaction of abiotic and biotic conditions, dispersal potential and evolutionary capacity to adapt (Soberón & Peterson, 2005), this study constitutes the baseline for the identification of the species and locations potentially in threat, showing good accuracy. Additionally, the use of a broad pull of predictors provides a complete library of species tolerance thresholds, obtained in a cost-effective way, and very useful for researchers. It is remarkable the great significance of hydrodynamic variables and extreme parameters, which demonstrate the need to include them in species distribution modelling.

In general, it can be said that the trend obtained for kelps and fucoids is in agreement with other predictive models. The main trend is the disappearance of southern communities (Jüterbock, 2013; Assis *et al.*, 2018; Wilson *et al.*, 2019), which has been also detected in this study. However, meanwhile some fucoids are predicted to migrate northward, our model suggests a more dramatic scenario. That could be the consequence of including some hydrodynamic variables that are usually not considered and the methodological approach that selects the algorithms and their weights according to their ability to be transferable in time.

Modelled species distributional range shifts are subjected to a large number of uncertainties, such as the lack of biological records. In the context of this work, the scarcity of data in the Iberian Peninsula can affect the models, due to the importance of this area as distributional limit of algae (Ramos *et al.*, 2016a), therefore, results have to be interpreted with caution. Besides, in climate change scenarios, distributional range shifts are the product of a multifactor complex interaction of abiotic and biotic conditions, dispersal potential and evolutionary capacity to adapt (Soberón & Peterson, 2005). Taking these aspects into account, the good accuracy of results obtained supports that this study constitutes the baseline for identification of the species and locations potentially in threat. It can be considered an initial step in the assessment of a large number of marine species with different requirements (intertidal and subtidal) and distributional patterns (native and invasive), by improving the knowledge in species-predictors relationship.

Chapter VI

Conclusions

Chapter VI. Conclusions

The overall aim of this thesis was the risk assessment of macroalgae distribution due to climate change in Europe. This objective has been reached through specific goals, which correspond to the core concepts of the WGII AR5 l. The methodology is based on four main elements:

Hazards

Hazards (specific objective 1) have been generated, spatial and temporal homogenously, from satellite, reanalysis and *in situ* sources for 16 variables and 10 parameters for each variable in the historical period (1985-2015) and two climate change scenarios (RCP 4.5 and RCP 8.5) for the mid and long term (2070 and 2099, respectively). This information is available for marine researchers to be used in the risk assessment of marine communities to climate change.

Exposure

On the other hand, species records from 1985 to 2015 have been collected and filtered for 21 representative macroalgae species. As the number of species was too large to manage in modelling, they were grouped with the self-organizing technique according to their ecological requirements, so five species were used as representative of the five final groups (specific objective 2).

Vulnerability

As the objective of the methodology was to assess climate change effects, it was necessary to test its temporal transferability. To do this a framework was developed to guide the algorithms selection and their weights in ensemble approaches when modelling species distribution, ensuring their temporal transferability with two independent sets of abiotic and biotic data, previously obtained (specific objective 3). Besides, the most complete pull until now of species physiological thresholds by the inclusion of hydrodynamic variables and extreme parameters, whose importance in species distributions has been demonstrated, has been generated.

Risk

Once the input data (hazards from OCLE and exposure of the five selected species) and the modelling framework were defined, all elements were integrated to obtain the species probability of presence in the historical and climate change scenarios at European level (specific objective 4).

As a result, it can be inferred that the most endangered species would be *Sacchoriza polyschides* and *Gelidium spinosum*, meanwhile *Sargassum muticum* and *Cystoseira baccata* would be predicted to expand their range. *Pelvetia canaliculata* might decrease its probability of occurrence but it would be able to cope with changes. Besides, it is remarkable the importance of hydrodynamic variables, such as bottom shear stress, currents, bottom orbital speed or significant wave height for all species, supporting the importance of including this information when modelling marine species.

In addition to the general conclusions exposed, results allow the extraction of more specific findings for each chapter, detailed below.

Chapter II. OCLE: a European Open access database on Climate change effects on Littoral and oceanic Ecosystems

- OCLE database constitutes a step further in the availability of environmental information linked with marine ecosystems. It allows hazards definition in the context of studies of climate change effects on marine ecosystems.
- OCLE contributes to the existing knowledge by providing outstanding hydrodynamic variables for marine communities, such as the own developed variable bottom shear stress.
- The variety of parameters included in OCLE and their high temporal resolution, allow researchers to define biologically meaningful predictors of species or communities.
- The output format (.csv) can be used in diverse kinds of marine studies, with different purposes and scales, such as species distribution modelling or the physical and ecological classification of large areas.

Chapter III. Identification and classification of key seaweeds species in Europe

- The application of the SOM analysis to group species shows that, in terms of minimum effort and highest accuracy, it is possible to deal with a high number of species, considering the ecological differences between them.
- The species *Saccorhiza polyschides*, *Gelidium spinosum*, *Sargassum muticum*, *Pelvetia canaliculata* and *Cystoseira baccata* can be considered representative, in terms of their abiotic requirements, of 21 subtidal and intertidal species in Europe.
- The establishment of the link of species with the areas with higher probability of presence represents a step further in exposure characterization.

Chapter IV. Temporal transferability of macroalgae distribution models: the role of algorithm selection

- In general, machine learning algorithms show higher sensitivity that regression-based approaches when evaluated in another period.
- Temporal transferability of all species is good, according to evaluation metrics considered.
- Hydrodynamic variables (e.g. significant wave height, shear stress, currents, etc.) have a significant importance in mid-intertidal and subtidal species distribution.
- Predictors related to extreme conditions (maximum, number of days over a threshold) are decisive for all species considered.
- This methodological approach constitutes a time and cost-effective tool for climate change studies in marine areas by providing a starting point to guide researchers in the algorithm and their weights selection to ensure models temporal transferability.
- Results are satisfactory and robust for being broadly applied in marine research with different objectives and geographical areas, allowing a comparative framework between species predictions and promoting the use of transferable models, especially in climate change studies.

Chapter V. Climate change induced range shifts in seaweeds distributions in Europe

- Models built have been demonstrated a very good goodness-of-fit and discrimination performance.
- Although requiring further development, the initial results presented here identified the species and areas at risk for five algae with different ecological requirements (from upper littoral to subtidal) and distributional patterns (native and invasive), therefore it can be used as surrogate for many others with similar ecological requirements.
- The most endangered species would be *Sacchoriza polyschides* and *Gelidium spinosum*, meanwhile *Sargassum muticum* and *Cystoseira baccata* would be predicted to expand their range. *Pelvetia canaliculata* might decrease its probability of occurrence but it would be able to cope with changes.

Results of this thesis should be considered as a first step in the assessment of the magnitude of modelled range changes in seaweeds in Europe. Tools developed help to fill the gap in knowledge between seaweeds ecology and environmental drivers on the one-hand and between science and managers on the other, by paying particular attention to building robust models with objective, reproducible, globally applicable and efficient methodology.

Future research

During this thesis, some aspects have been detected to need further development. Although they have been exposed in the discussion section of each chapter, they have been summarized below:

- Regarding hazards, it is necessary to advance in the definition of ecologically sounded parameters. This challenging task needs the interaction among experts of different disciplines to model with the best knowledge available. Within the abiotic variables' definition, some aspects as the temporal influence of events such as ENSO or NAO needs to be assessed, as this work has demonstrated the importance of extreme events for species distribution.
- Local studies require the collection of data at appropriate detail level by downscaling the existing information at coarse resolution. This approach requires the definition of an accurate procedure, able to capture local characteristics.
- The improvement of biological data quality, by field procedures homogenization and increase in geographical and temporally cover, would allow the inclusion of absences in models, which has been demonstrated to increase the reliability of predictions. This is challenging but different approaches are arising in the last years, such as the application of environmental DNA (Muha *et al.*, 2017).
- Concerning modelling, as this field has evolved quickly in the last years, important advances have been achieved. However, there is a lack of agreement for general studies. Besides, some assumptions made when modelling, such as the equilibrium of species with their environment needs to be assessed, especially for invasive species.
- A further step in modelling is the inclusion of biotic relationships and species movement capacities additional to abiotic conditions. Although the definition of these relationships in climate change scenarios are difficult to predict, some attempts have been done which demonstrate the increase in ecological realism of resulting models (Meier *et al.*, 2010; Jaeschke *et al.*, 2012; Giannini *et al.*, 2013; Atauchi *et al.*, 2018; Simões & Peterson, 2018).
- Ideally, when temporal information is available and the biotic interactions are clearly defined in the ecosystem, the use of mechanistic models arises as the best tool to make predictions.
- The combination of niche-based SDMs, environmental data from the field, and phylogenetic information can help us to better define both the timescales over which shifts of the realized niche occur and which species are most likely to

experience them. Additional empirical data on the life history characteristics, ecological circumstances and evolutionary histories that are associated with particular dynamics of species niches will also improve confidence in the predictions supplied by niche-based SDMs (Pearman *et al.*, 2008).

References

References

A

Acuna E, Rodriguez C (2004) A meta analysis study of outlier detection methods in classification. In: *International IPSI 2004 Conference, Symposium on Challenges in Internet and Interdisciplinary Research*, p. 25. Mayaguez, Puerto Rico.

Adger WN (2006) Vulnerability. *Global Environmental Change*, **16**, 268–281.

Aguirre-Gutiérrez J, Carvalheiro LG, Polce C, van Loon EE, Raes N, Reemer M, Biesmeijer JC (2013) Fit-for-Purpose: Species Distribution Model Performance Depends on Evaluation Criteria - Dutch Hoverflies as a Case Study. *PLoS ONE*, **8**.

Albérola C, Millot C (2003) Circulation in the French mediterranean coastal zone near Marseilles: The influence of wind and the Northern Current. *Continental Shelf Research*, **23**, 587–610.

Albert S, Leon JX, Grinham AR et al. (2017) Environmental Research Letters Related content Interactions between sea-level rise and wave exposure on reef island dynamics in the Solomon Islands Winners and losers as mangrove, coral and seagrass ecosystems respond to sea-level rise in Solomon Islands. *Environ. Res. Lett*, **12**.

Albright R, Mason B, Miller M, Langdon C (2010) Ocean acidification compromises recruitment success of the threatened Caribbean coral *Acropora palmata*. *Proceedings of the National Academy of Sciences*, **107**, 20400–20404.

Alekseenko E, Roux B, Fougere D, Chen PG (2017) The effect of wind induced bottom shear stress and salinity on *Zostera noltii* replanting in a Mediterranean coastal lagoon. *Estuarine, Coastal and Shelf Science*, **187**, 293–305.

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Álvarez-Martínez JM, Jiménez-Alfaro B, Barquín J, Ondiviela B, Recio M, Silió-Calzada A, Juanes JA (2018) Modelling the area of occupancy of habitat types with remote sensing. *Methods in Ecology and Evolution*, **9**, 580–593.

Anadón R (1983) Zonación en la costa asturiana: variación longitudinal de las comunidades de macrófitos en diferentes niveles de marea. *Investigación Pesquera*, **45**, 125–141.

Anlauf H, D'Croz L, O'Dea A (2011) A corrosive concoction: The combined effects of ocean warming and acidification on the early growth of a stony coral are multiplicative. *Journal of Experimental Marine Biology and Ecology*, **397**, 13–20.

Anthony KRN, Kline DI, Diaz-Pulido G, Dove S, Hoegh-Guldberg O (2008) Ocean acidification causes bleaching and productivity loss in coral reef builders. *Proceedings of the National Academy of Sciences*, **105**, 17442–17446.

Antunes do Carmo JS, Temperville A, Seabra-Santos FJ (2003) Bottom friction and time-dependent shear stress for wave-current interaction. *Journal of Hydraulic Research*, **41**, 27–37.

Araujo M, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.

Araujo MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species-climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.

Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.

Araújo MB, Pearson RG (2005) Equilibrium of species' distributions with climate. *Ecography*, **28**, 693–695.

Araújo RM, Assis J, Aguillar R et al. (2016) Status, trends and drivers of kelp forests in Europe: an expert assessment. *Biodiversity and Conservation*, **25**, 1319–1348.

Araújo MB, Anderson RP, Barbosa AM et al. (2019) Standards for distribution models in biodiversity assessments. *Science advances*, **5**, 1–10.

Arnell NW, Brown S, Gosling SN et al. (2014) The impacts of climate change across the globe: A multi-sectoral assessment. *Climatic Change*, **134**, 457–474.

Assis J, Castilho Coelho N, Alberto F, Valero M, Raimondi P, Reed D, Alvares Serrão E (2013) High and Distinct Range-Edge Genetic Diversity despite Local Bottlenecks. *PLoS ONE*, **8**, 1–11.

Assis J, Coelho NC, Lamy T, Valero M, Alberto F, Serrão EÁ (2016) Deep reefs are climatic refugia for genetic diversity of marine forests. *Journal of Biogeography*, **43**, 833–844.

Assis J, Berecibar E, Claro B, Alberto F, Reed D, Raimondi P, Serrão EA (2017a) Major shifts at the range edge of marine forests: The combined effects of climate changes and limited dispersal. *Scientific Reports*, **7**, 1–10.

Assis J, Tyberghein L, Verbruggen H, Serrão E, De Clerck O (2017b) Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, **27**, 277–284.

Assis J, Araújo MB, Serrão EA (2018) Projected climate changes threaten ancient refugia of kelp forests in the North Atlantic. *Global Change Biology*, **24**, e55–e66.

Atauchi PJ, Peterson AT, Flanagan J (2018) Species distribution models for *Peruvian Plantcutter* improve with consideration of biotic interactions. *Journal of Avian Biology*, 1–8.

Ateweberhan M, Feary DA, Keshavmurthy S, Chen A, Schleyer MH, Sheppard CRC (2013) Climate change impacts on coral reefs: Synergies with local effects, possibilities for acclimation, and management implications. *Marine Pollution Bulletin*, **74**, 526–539.

Austin MP (1990) Community theory and competition in vegetation. *Community theory and competition in vegetation.*, 215–238.

B

Barbet-Massin M, Thuiller W, Jiguet F (2010) How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography*, **33**, 878–886.

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Barbosa AM, Real R, Mario Vargas J (2009) Transferability of environmental favourability models in geographic space: The case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling*, **220**, 747–754.

de Barbosa Araujo CC, Atkinson PM, Dearing JA (2015) Remote sensing of ecosystem services: A systematic review. *Ecological Indicators*, **52**, 430–443.

Barry S, Elith J (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology*, **43**, 413–423.

Beaumont LJ, Graham E, Duursma DE et al. (2016) Which species distribution models are more (or less) likely to project broad-scale, climate-induced shifts in species ranges? *Ecological Modelling*, **342**, 135–146.

Bedia J, Herrera S, Gutiérrez JM (2013) Dangers of using global bioclimatic datasets for ecological niche modeling. Limitations for future climate projections. *Global and Planetary Change*, **107**, 1–12.

Bekkby T, Rinde E, Erikstad L, Bakkestuen V (2009) Spatial predictive distribution modelling of the kelp species *Laminaria hyperborea*. *ICES Journal of Marine Science*, **66**, 2106–2115.

Belando MD, García Muñoz R, Ramos Segura A, Franco-Navarro I, Bernardeau-esteller J, García Moreno P, Ruiz JM (2015) Distribution and abundance of *Cymodocea nodosa* meadows and *Pinna nobilis* populations in the Mar Menor coastal lagoon (Murcia, SE of Spain). In: *PeerJ PrePrints*.

Belsher T (1989) *Sargassum muticum (Yendo) Fensholt sur le littoral français. Synthese des études 1983-1989*. Plouzane, France, 103 pp.

Best EP., Buzzelli CP, Bartell SM, Wetzel RL, Boyd WA, Doyle RD, Campbell KR (2001) Modeling submersed macrophyte growth in relation to underwater light climate: Modeling approaches and application potential. *Hydrobiologia*, **444**, 43–70.

Borja A, Chust G, Fontán A, Garmendia JM, Uyarra MC (2018) Long-term decline of the canopy-forming algae *Gelidium corneum*, associated to extreme wave events and reduced sunlight hours, in the southeastern Bay of Biscay. *Estuarine, Coastal and Shelf Science*, **205**, 152–160.

Bosch S, Tyberghein L, Deneudt K, Hernandez F, De Clerck O (2017) In search of relevant predictors for marine species distribution modelling using the MarineSPEED benchmark dataset. *Diversity and Distributions*, **24**, 144–157.

Boyce MS, Vernier PR, Nielsen SE, Schmiegelow FKA (2002) Evaluating resource selection functions. *Ecological Modelling*, **157**, 281–300.

Ben Brahim M, Mabrouk L, Hamza A, Mahfoudhi M, Bouain A, Aleya L (2015) Spatial scale variability in shoot density and epiphytic leaves of *Posidonia oceanica* on Kerkennah Island (Tunisia) in relation to current tide effects. *Marine Ecology*, **36**, 1311–1331.

Brandt LA, Benscoter AM, Harvey R et al. (2017) Comparison of climate envelope models developed using expert-selected variables versus statistical selection. *Ecological Modelling*, **345**, 10–20.

Breiman L (1984) *Classification and Regression Trees*. Routledge, New York, 368 pp.

Breiman L (2001) Random forests. *Machine Learning*, **45**, 5–32.

Breiner FT, Nobis MP, Bergamini A, Guisan A (2018) Optimizing ensembles of small models for predicting the distribution of species with few occurrences. *Methods in Ecology and Evolution*, **9**, 802–808.

Broennimann O, Treier UA, Müller-Schärer H, Thuiller W, Peterson AT, Guisan A (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.

Broennimann O, Fitzpatrick MC, Pearman PB et al. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.

Broennimann O, Di Cola V, Petitpierre B et al. (2016) ecospat: Spatial Ecology Miscellaneous Methods.

Brugère C, De Young C (2015) *Assessing climate change vulnerability in fisheries and aquaculture: Available methodologies and their relevance for the sector, by Cecile Brugère and Cassandra De Young. FAO Fisheries and Aquaculture Technical Paper No. 597. Rome, Italy.* Rome, 98 pp.

Buonomo R, Chefaoui RM, Lacida RB, Engelen AH, Serrão EA, Airoldi L (2018) Predicted extinction of unique genetic diversity in marine forests of *Cystoseira spp*. *Marine Environmental Research*, **138**, 119–128.

Burrows MT, Fox CJ, Moore P et al. (2018) *Wild seaweed harvesting as a diversification* *opportunity for fishermen*. 171 pp.

Burson A, Stomp M, Akil L, Brussaard CPD, Huisman J (2016) Unbalanced reduction of nutrient loads has created an offshore gradient from phosphorus to nitrogen limitation in the North Sea. *Limnology and Oceanography*, **61**, 869–888.

Busby JR (1991) *BIOCLIM - a bioclimate analysis and prediction system*. 64–68 pp.

C

Cabioc’h J, Floc’h JY, Le Toquin A, Boudouresque CF, Meinesz A, Verlaque M (1995) *Guía de las algas de los mares de Europa: Atlántico y Mediterráneo*. Ediciones Omega, S.A., Barcelona, Spain, 249 pp.

Callaghan DP, Leon JX, Saunders MI (2015) Wave modelling as a proxy for seagrass ecological modelling: Comparing fetch and process-based predictions for a bay and reef lagoon. *Estuarine, Coastal and Shelf Science*, **153**, 108–120.

Camus P, Losada ÍJ, Izaguirre C, Espejo A, Menendez M, Pérez J (2017) Statistical wave climate projections for coastal impact assessments. *Earth’s Future*, **5**, 918–933.

Casado de Amezúa M del P, Altamirano M, Bárbara I et al. (2015) *Categorización, según su grado de amenaza, de especies de macrófitos fundadores de bosques y praderas marinas en el atlántico español*. 126 pp.

Cavole L, Demko A, Diner R et al. (2016) Biological impacts of the 2013–2015 warm-water anomaly in the Northeast Pacific: winners, losers, and the future. *Oceanography*, **29**.

Cefalì ME, Cebrian E, Chappuis E, Pinedo S, Terradas M, Mariani S, Ballesteros E (2016) Life on the boundary: Environmental factors as drivers of habitat distribution in the littoral zone. *Estuarine, Coastal and Shelf Science*, **172**, 81–92.

Cefalì ME, Ballesteros E, Lluı J, Chappuis E, Terradas M, Mariani S, Cebrian E (2018) The optimal sampling design for littoral habitats modelling : A case study from the north-western Mediterranean. *PLoS ONE*, **13**, 1–18.

Celis-Pla PSM, Martinez B, Quintano E et al. (2014) Short-term ecophysiological and biochemical responses of *Cystoseira tamariscifolia* and *Ellisolandia elongata* to environmental changes. *Aquatic Biology*, **22**, 227–243.

Celis-Pla PSM, Martinez B, Korbee N, Hall-Spencer JM, Figueroa FL (2017) Ecophysiological responses to elevated CO2 and temperature in *Cystoseira tamariscifolia* (Phaeophyceae). *Climatic Change*, 1–15.

Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, **7**, 1247–1250.

Chapman VJ, Chapman DJ (1980) *Seaweeds and their uses*, Third edit edn (ed Hall C and).

Springer Netherlands, London ; New York, 334 pp.

Chatterjee S, Hadi AS (2006) *Regression Analysis by Example*. John Wiley and Sons, 424 pp.

Chefaoui RM, Duarte CM, Serrão EA (2018) Dramatic loss of seagrass habitat under projected climate change in the Mediterranean Sea. *Global Change Biology*, **24**, 4919–4928.

Cid A, Castanedo S, Abascal AJ, Menéndez M, Medina R (2014) A high resolution hindcast of the meteorological sea level component for Southern Europe: the GOS dataset. *Climate Dynamics*, **43**, 1–18.

Collins M, Knutti R, Arblaster J et al. (2013) Long-term Climate Change: Projections, Commitments and Irreversibility. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM), pp. 1029–1136. Cambridge, United Kingdom and New York, NY, USA.

Coro G, Magliozzi C, Ellenbroek A, Pagano P (2015) Improving data quality to build a robust distribution model for *Architeuthis dux*. *Ecological Modelling*, **305**, 29–39.

Costanza R, D'Arge R, de Groot R et al. (1997) The value of the world's ecosystem services and natural capital. *Nature*, **387**, 253–260.

Critchley AT, Farnham WF, Yoshida T, Norton TA (1990) A bibliography of the invasive alga *Sargassum muticum* (Fucales: Sargassaceae). *Botanica Marina*, **33**, 551–562.

Crossman ND, Bryan BA, Summers DM (2012) Identifying priority areas for reducing species vulnerability to climate change. *Diversity and Distributions*, **18**, 60–72.

D

Dalyander PS, Butman B, Sherwood CR, Signell RP, Wilkin JL (2013) Characterizing wave- and current- induced bottom shear stress: U.S. middle Atlantic continental shelf. *Continental Shelf Research*, **52**, 73–86.

Díez I, Muguerza N, Santolaria A., Ganzedo U, Gorostiaga JM (2012) Seaweed assemblage changes in the eastern Cantabrian Sea and their potential relationship to climate change. *Estuarine, Coastal and Shelf Science*, **99**, 108–120.

Dixon PS (1966) On the form of the thallus in the Florideophyceae. In: *Trends in Plant Morphogenesis* (ed Cutter EG), pp. 45–63. Longmans, Green & co, London.

Doney SC (2010) The Growing Human Footprint on Coastal and Open-Ocean Biogeochemistry. *Science*, **328**, 1512.

Doney SC, Ruckelshaus M, Emmett Duffy J et al. (2012) Climate Change Impacts on Marine Ecosystems. *Annual Review of Marine Science*, **4**, 11–37.

Donlon CJ, Martin M, Stark J, Roberts-Jones J, Fiedler E, Wimmer W (2012) The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sensing of Environment*, **116**, 140–158.

Duarte CM, Losada IJ, Hendriks IE, Mazarrasa I, Marbà N (2013) The role of coastal plant communities for climate change mitigation and adaptation. *Nature Climate Change*, **3**, 961–968.

Duarte CM, Wu J, Xiao X, Bruhn A, Krause-Jensen D (2017) Can Seaweed Farming Play a Role in Climate Change Mitigation and Adaptation? *Frontiers in Marine Science*, **4**.

Duarte Mora L (2015) *Distribución geográfica de macroalgas boreoatlánticas en su límite sur: mecanismos, patrones de cambio e impactos ecológicos*. Universidad Rey Juan Carlos, 213 pp.

E

EEA (2009) *Rising sea temperatures , ice-free Arctic summers and a changing marine food chain*. Copenhagen, Denmark, 8 pp.

EEA (2015) *The European environment — state and outlook 2015: synthesis report*. Copenhagen, Denmark, 212 pp.

Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.

Elith J, Graham CH, Anderson RP et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. M. *ECography*, **29**, 129–151.

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Ellison JC (2015) Vulnerability assessment of mangroves to climate change and sea-level rise impacts. *Wetlands Ecology and Management*, **23**, 115–137.

Elton CS (Charles S (1927) *Animal ecology*. University of Chicago Press, 209 pp.

Engelen AH, Serebryakova A, Ang P et al. (2015) Circumglobal invasion by the brown seaweed *Sargassum muticum*. *Oceanography and Marine Biology: An annual review*, **53**, 81–126.

Engler R, Guisan A, Rechsteiner L (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, **41**, 263–274.

Engler R, Randin CF, Thuiller W et al. (2011) 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology*, **17**, 2330–2341.

European Commission (2008) Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). *Official Journal of the European Communities*, **164**, 19–40.

F

Fabri-Ruiz S, Danis B, David B, Saucède T (2019) Can we generate robust species distribution models at the scale of the Southern Ocean? *Diversity and Distributions*, **25**, 21–37.

Fautin DG, Buddemeier RW (2002) *Biogeoinformatics of hexacorallia (corals, sea anemones, and their allies): interfacing geospatial, taxonomic, and environmental data for a group of marine invertebrates*.

Fernandez C (2011) The retreat of large brown seaweeds on the north coast of Spain: the case of *Saccorhiza polyschides*. *European Journal of Phycology*, **46**, 352–360.

Fernández C (2016) Current status and multidecadal biogeographical changes in rocky intertidal algal assemblages: The northern Spanish coast. *Estuarine, Coastal and Shelf Science*, **171**, 35–40.

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, **24**, 38–49.

Fonseca M, Fisher J (1986) A comparison of canopy friction and sediment movement between four species of seagrass with reference to their ecology and restoration . *Marine Ecology Progress Series*, **29**, 15–22.

Fralick RA, Baldwin HP, Neto AI, Hehre EJ (1990) Physiological responses of *Pterocladia* and *Gelidium* (Gelidiales, Rhodophyta) from the Azores, Portugal. *Hydrobiologia*, **204**, 479–482.

Franklin J (2009) *Mapping species distributions: spatial inference and prediction*, Cambridge edn (eds Usher M, Saunders D, Peet R, Dobson A). 320 pp.

Fredriksen S, Rueness J (1989) Culture Studies of *Gelidium latifolium* (Grev.) Born et Thur. (Rhodophyta) from Norway. Growth and Nitrogen Storage in Response to Varying Photon Flux Density, Temperature and Nitrogen Availability. *Botanica Marina*, **32**, 539–546.

Friedman JH (1991) Multivariate Adaptive Regression Splines. *The Annals of Statistics*, **19**, 1–67.

Friedman JH (2002) Stochastic gradient boosting. *Computational Statistics and Data Analysis*, **38**, 367–378.

Fritzsche K, Schneiderbauer S, Bubeck P, Kienberger S, Buth M, Zebisch M, Kahlenborn W (2017) *The Vulnerability Sourcebook. Concept and guidelines for standardised vulnerability assessments*.

G

Gallien L, Douzet R, Pratte S, Zimmermann NE, Thuiller W (2012) Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography*, **21**, 1126–1136.

Galván C, Puente A, Castanedo S, Juanes JA (2016) Average vs. extreme salinity conditions: Do they equally affect the distribution of macroinvertebrates in estuarine environments? *Limnology and Oceanography*, **61**, 984–1000.

García-Fernández A, Bárbara I (2016) Studies of the *Cystoseira* assemblages in Northern Atlantic Iberia. *Anales del Jardín Botánico de Madrid*, **73**, 1–39.

Garnesson P, Mangin A, Gohin F (2016) *Quatliy information document. Ocean Colour - Global. Optics/Chlorophyll Observation Products*. 1–37 pp.

GBIF (2013) GBIF Backbone Taxonomy, 2013-07-01.

Giannini TC, Chapman DS, Saraiva AM, Alves-dos-Santos I, Biesmeijer JC (2013) Improving species distribution models using biotic interactions: A case study of parasites, pollinators and plants. *Ecography*, **36**, 649–656.

Gobeyn S, Mouton AM, Cord AF, Kaim A, Volk M, Goethals PLM (2018) Evolutionary algorithms for species distribution modelling: A review in the context of machine learning. *Ecological Modelling*, **392**, 179–195.

Gorman D, Bajjouk T, Populus J, Vasquez M, Ehrhold A (2013) Modeling kelp forest distribution and biomass along temperate rocky coastlines. *Marine Biology*, **160**, 309–325.

Green RH (1979) *Sampling design and statistical methods for environmental biologists*. Wiley, New York, 257 pp.

Grenouillet G, Buisson L, Casajus N, Lek S (2011) Ensemble modelling of species distribution: The effects of geographical and environmental ranges. *Ecography*, **34**, 9–17.

Grinnell J (1917) The niche-relationships of the California thrasher. *The Auk*, **34**, 427–433.

Guinda X, Juanes JA, Puente A, Revilla JA (2008) Comparison of two methods for quality assessment of macroalgae assemblages, under different pollution types. *Ecological Indicators*, **8**, 743–753.

Guinda X, Juanes JA, Puente A, Echavarri-Erasun B (2012) Spatial distribution pattern analysis of subtidal macroalgae assemblages by a non-destructive rapid assessment method. *Journal of Sea Research*, **67**, 34–43.

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.

Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157**, 89–100.

Guisan A, Tingley R, Baumgartner JB et al. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424–1435.

Guisan A, Petitpierre B, Broennimann O, Daehler C, Kueffer C (2014) Unifying niche shift studies : Insigts from biological invasions. **29**.

Guisan A, Thuiller W, Zimmermann NE (2017) *Habitat suitability and distribution models*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 462 pp.

Guo Q, Kelly M, Graham CH (2005) Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecological Modelling*, **182**, 75–90.

H

Hanberry BB, Hong SH, Palik BJ (2012) Pseudoabsence Generation Strategies for Species Distribution Models. *PLoS ONE*, **7**, 7–10.

Hanley JA, McNeil BJ (1982) The meaning and use of the area under the receiver operating characteristic (ROC) curve. *Radiology*, **143**, 29–36.

Harker M, Berkaloff C, Lemoine Y (1999) Effects of high light and desiccation on the operation of the xanthophyll cycle in two marine brown algae. *European Journal of Phycology*, 37–41.

Hartigan JA, Wong MA (1979) Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*, **28**, 100.

Hastie TJ, Tibshirani RJ (1990) Generalized Additive Models. *Monographs on Statistics & Applied Probability*, **I**, 352.

Hastie T, Tibshirani R, Buja A (1994) Flexible discriminant analysis by optimal scoring. *Journal of the American Statistical Association*, **89**, 1255–1270.

Hastie T, Tibshirani R, Friedman J (2009) *The elements of statistical learning data mining, inference, and prediction*, 2nd edition edn. Springer, Berlin, 763 pp.

Helmuth B, Mieszkowska N, Moore P, Hawkins SJ (2006) Living on the Edge of Two Changing Worlds: Forecasting the Responses of Rocky Intertidal Ecosystems to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 373–404.

Hirzel AH, Hausser J, Chessel D, Perrin N (2002) Ecological-Niche Factor Analysis: How to compute habitat-suitability maps without absence data? *Ecology*, **83**, 2027–2036.

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, **199**, 142–152.

Holt RD, Keitt TH (2000) Alternative causes for range limits: A metapopulation perspective. *Ecology Letters*, **3**, 41–47.

Huang J, Frimpong EA, Orth DJ (2016) Temporal transferability of stream fish distribution models: Can uncalibrated SDMs predict distribution shifts over time? *Diversity and Distributions*, **22**, 651–662.

Hughes ARR, Bando KJ, Rodriguez LF, Williams SL (2004) Relative effects of grazers and nutrients on seagrasses: a meta-analysis approach. *Marine Ecology Progress Series*, **282**, 87–99.

Hughes TP, Kerry JT, Álvarez-Noriega M et al. (2017) Global warming and recurrent mass bleaching of corals. *Nature*, **543**, 373–377.

Hutchinson GE (1957) Concluding remarks. In: *Cold spring Harbor symposia on quantitative biology*, Vol. 22, pp. 415–427.

Hutchinson GE (1978) *An introduction to population ecology*. Yale University Press, New Haven, 260 pp.

I

IHCantabria (2008) *Desarrollo de una metodología para la evaluación del estado de las aguas superficiales del entorno de afección de saneamientos litorales*.

Infantes E, Orfila A, Bouma TJ, Simarro G, Terrados J (2011) *Posidonia oceanica* and *Cymodocea nodosa* seedling tolerance to wave exposure. *Limnology and Oceanography*, **56**, 2223–2232.

IPCC (2013) *Working Group I: contribution to the IPCC fifth assessment report climate change 2013: the physical science basis*. 103 pp.

IPCC (2014) *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland, 151 pp.

Iturbide M, Bedia J, Gutiérrez JM (2018) Background sampling and transferability of species distribution model ensembles under climate change. *Global and Planetary Change*, **166**, 19–29.

J

Jaeschke A, Bittner T, Jentsch A, Reineking B, Schlumprecht H, Beierkuhnlein C (2012)

Biotic Interactions in the Face of Climate Change: A Comparison of Three Modelling Approaches. *PLoS ONE*, **7**.

Jensen MM, Denny MW (2015) Experimental determination of the hydrodynamic forces responsible for wave impact events. *Journal of Experimental Marine Biology and Ecology*, **469**, 123–130.

Jiménez-Alfaro B, Suárez-Seoane S, Chytrý M et al. (2018) Modelling the distribution and compositional variation of plant communities at the continental scale. *Diversity and Distributions*, **24**, 978–990.

Johns B, Marsaleix P, Estournel C, Véhil R (1992) On the wind-driven coastal upwelling in the Gulf of Lions. *Journal of Marine Systems*, **3**, 309–320.

Jones T, Gardner JP a., Bell JJ (2015) Modelling the effect of wave forces on subtidal macroalgae: A spatial evaluation of predicted disturbance for two habitat-forming species. *Ecological Modelling*, **313**, 149–161.

Jonsson PR, Kotta J, Andersson HC, Herkül K, Virtanen E, Sandman AN, Johannesson K (2018) High climate velocity and population fragmentation may constrain climate-driven range shift of the key habitat former *Fucus vesiculosus*. *Diversity and Distributions*, **24**, 892–905.

Josefsson M, Jansson K (2011) *NOBANIS - Invasive Alien Species. Fact Sheet Sargassum muticum*. 1–10 pp.

Juanes JA, Fernández C (1988) Ciclo anual y producción de *Gelidium latifolium* (Grev.) Thur. et Born. (1876), en la región de cabo Peñas (Asturias, N de España). *Investigacion Pesquera*, **52**, 109–122.

Juanes JA, Gutiérrez LM (1992) Cartografía y evaluación de biomasa de *Gelidium sesquipedale* (Clem.) Born. et Thur en la costa oriental de Cantabria (N de España). *Informes Técnicos de Scientia Marina. Consejo Superior de Investigaciones Científicas.*, **166**, 3–22.

Jüterbock A (2013) *Climate change impact on the seaweed Fucus serratus, a key foundational species on North Atlantic rocky shores*. University of Nordland, 66 pp.

K

Karlsson J, Loo LO (1999) On the distribution and continuous expansion of the Japanese seaweed *Sargassum muticum* in Sweden. *Botanica Marina*, **42**, 285–294.

Kohonen T (1998) The self-organizing map. *Neurocomputing*, **21**, 1–6.

Kotta J, Valdivia N, Kutser T, Toming K, Rätsep M, Orav-Kotta H (2018) Predicting the cover and richness of intertidal macroalgae in remote areas: a case study in the Antarctic Peninsula. *Ecology and Evolution*, 1–9.

Krause-Jensen D, Duarte CM (2014) Expansion of vegetated coastal ecosystems in the future Arctic. **1**, 1–10.

Kregting L, Blight AJ, Elsäßer B, Savidge G (2016) The influence of water motion on the growth rate of the kelp *Laminaria digitata*. *Journal of Experimental Marine Biology and Ecology*, **478**, 86–95.

L

de la Hoz CF, Ramos E, Puente A, Méndez FJ, Menéndez M, Juanes JA, Losada IJ (2018a) Ecological typologies of large areas. An application in the Mediterranean Sea. *Journal of Environmental Management*, **205**, 59–72.

de la Hoz CF, Ramos E, Acevedo A, Puente A, Losada íñigo J, Juanes JA (2018b) Ocle: a European Open Access Database on Climate Change Effects on Littoral and Oceanic Ecosystems. *Progress in Oceanography*, **168**, 222–231.

Lamela-Silvarrey C, Fernández C, Anadón R, Arrontes J (2012) Fucoid assemblages on the north coast of Spain: Past and present (1977 - 2007). *Botanica Marina*, **55**, 199–207.

De Lange HJ, Lahr J, Van der Pol JJC, Wessels Y, Faber JH (2009) Ecological vulnerability in wildlife: an expert judgment and multicriteria analysis tool using ecological traits to assess relative impact of pollutants. *Environmental toxicology and chemistry / SETAC*, **28**, 2233–2240.

Larkum A, Orth RJ, Duarte CM (2006) *Seagrasses: biology, ecology and conservation* (eds Larkum A, Orth RJ, Duarte CM). Springer, Dordrecht (The Netherlands), 691 pp.

Lauria V, Power AM, Lordan C, Weetman A, Johnson MP (2015) Spatial transferability of habitat suitability models of Nephrops norvegicus among fished areas in the northeast atlantic: Sufficiently stable for marine resource conservation? *PLoS ONE*, **10**, 1–19.

Lerman A, Guidry M, Andersson AJ, Mackenzie FT (2011) Coastal Ocean Last Glacial Maximum to 2100 CO2-Carbonic Acid-Carbonate System: A Modeling Approach. *Aquatic Geochemistry*, **17**, 749–773.

Lima FP, Wethey DS (2012) Three decades of high-resolution coastal sea surface temperatures reveal more than warming. *Nature Communications*, **3**, 1–13.

Lima FP, Ribeiro PA, Queiroz N, Hawkins SJ, Santos AM (2007) Do distributional shifts of northern and southern species of algae match the warming pattern? *Global Change Biology*, **13**, 2592–2604.

Lipkin Y, Beer S, Eshel A (1993) The ability of *Porphyra linearis* (Rhodophyta) to tolerate prolonged periods of desiccation. *Botanica Marina*, **36**, 517–524.

Liquete C, Cid N, Lanzanova D, Grizzetti B, Reynaud A (2016) Perspectives on the link between ecosystem services and biodiversity: The assessment of the nursery function. *Ecological Indicators*, **63**, 249–257.

Little C, Kitching JA (1996) *The biology of rocky shores* (eds Little C, Martin MH, Southwood TRE, Ulfstrand S). Oxford University Press, Great Britain, 240 pp.

Liu C, White M, Newell G (2009) Measuring the accuracy of species distribution models: a review. *Proceedings 18th World IMACs/MODSIM Congress*, 4241–4247.

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.

Lovelock CE, Cahoon DR, Friess DA et al. (2015) The vulnerability of Indo-Pacific mangrove forests to sea-level rise. *Nature*, **526**, 559–563.

Lüning K, Yarish C, Kirkman H (1990) *Seaweeds : their environment, biogeography, and ecophysiology*. Wiley, 527 pp.

M

Marbà N, Díaz-Almela E, Duarte CM (2014) Mediterranean seagrass (*Posidonia oceanica*) loss between 1842 and 2009. *Biological Conservation*, **176**, 183–190.

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, **15**, 59–69.

Marshall NA, Tobin RC, Marshall PA, Gooch M, Hobday AJ (2013) Social Vulnerability of Marine Resource Users to Extreme Weather Events. *Ecosystems*, **16**, 797–809.

Martin CS, Tolley MJ, Farmer LE et al. (2015) A global map to aid the identification and screening of Critical Habitat for marine industries. *Marine Policy*, **53**, 45–53.

Martínez B, Pato LS, Rico JM (2012a) Nutrient uptake and growth responses of three intertidal macroalgae with perennial, opportunistic and summer-annual strategies. *Aquatic Botany*, **96**, 14–22.

Martínez B, Arenas F, Rubal M et al. (2012b) Physical factors driving intertidal macroalgae distribution: Physiological stress of a dominant fucoid at its southern limit. *Oecologia*, **170**, 341–353.

Martínez B, Radford B, Thomsen MS et al. (2018a) Distribution models predict large contractions in habitat-forming seaweeds in response to ocean warming. *Diversity and Distributions*, 1350–1366.

Martínez B, Radford B, Thomsen MS et al. (2018b) Distribution models predict large contractions of habitat-forming seaweeds in response to ocean warming. *Diversity and Distributions*, **24**, 1350–1366.

Mazarrasa I, Marbà N, Garcia-Orellana J, Masqué P, Arias-Ortiz A, Duarte CM (2017) Effect of environmental factors (wave exposure and depth) and anthropogenic pressure in the C sink capacity of *Posidonia oceanica* meadows. *Limnology and*

Oceanography, **62**, 1436–1450.

McCullagh P, Nelder JA (1989) *Generalized linear models*. Chapman and Hall, 526 pp.

Meier ES, Kienast F, Pearman PB et al. (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, **33**, 1038–1048.

Meier HEM, Andersson HC, Eilola K et al. (2011) Hypoxia in future climates: A model ensemble study for the Baltic Sea. *Geophysical Research Letters*, **38**, 1–6.

Méndez-Sandín M, Fernández C (2016) Changes in the structure and dynamics of marine assemblages dominated by *Bifurcaria bifurcata* and *Cystoseira* species over three decades (1977-2007). *Estuarine, Coastal and Shelf Science*, **175**, 46–56.

Mineur F, Arenas F, Assis J et al. (2015) European seaweeds under pressure: Consequences for communities and ecosystem functioning. *Journal of Sea Research*, **98**, 91–108.

Montgomery DC, Peck EA (1982) *Introduction to Linear Regresssion Analysis*. Wiley, New York, 672 pp.

Moon JB, Dewitt TH, Errend MN et al. (2017) Model application niche analysis: Assessing the transferability and generalizability of ecological models. *Ecosphere*, **8**.

Mpakairi KS, Ndaimani H, Tagwireyi P, Gara TW, Zvidzai M, Madhlamoto D (2017) Missing in action: Species competition is a neglected predictor variable in species distribution modelling. *PLoS ONE*, **12**, 1–14.

Muha TP, Rodríguez-Rey M, Rolla M, Tricarico E (2017) Using Environmental DNA to Improve Species Distribution Models for Freshwater Invaders. *Frontiers in Ecology and Evolution*, **5**.

Muller-Karger F, Andréfouët S, Varela R, Thunell R (2005) The Color of the Coastal Ocean and Applications in the Solution of Research and Management Problems. In: *Remote Sensing of Coastal Aquatic Environments - Technologies, Tenchniques and Applications*, pp. 101–127. Springer, Dordrecht (The Netherlands).

N

Naimi B, Araujo MB (2016) sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography*, **39**, 1-8 (ver. 1.0-54).

Naimi B, Skidmore AK, Groen TA, Hamm NAS (2011) Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *Journal of Biogeography*, **38**, 1497–1509.

Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG (2014) Where is positional uncertainty a problem for species distribution modelling? *Ecography*, **37**, 191–203.

Neiva J, Assis J, Fernandes F, Pearson GA, Serrão EA (2014) Species distribution models and mitochondrial DNA phylogeography suggest an extensive biogeographical shift in the high-intertidal seaweed *Pelvetia canaliculata*. *Journal of Biogeography*, **41**, 1137–1148.

Nejrup LB, Pedersen MF (2008) Effects of salinity and water temperature on the ecological performance of *Zostera marina*. *Aquatic Botany*, **88**, 239–246.

Nicastro KR, Zardi GI, Teixeira S, Neiva J, Serrao EA, Pearson GA (2013) Shift happens: trailing edge contraction associated with recent warming trends threatens a distinct genetic lineage in the marine macroalga *Fucus vesiculosus*. *BMC Biology*, **11**, 6.

Norton TA (1970) The marine algae of the County Wexford, Ireland. *British Phycological Journal*, **5**, 257–266.

Norton TA (1977) Experiments on the factors influencing the geographical distribution of Saccorhiza polyschides and *Saccorhiza dermatodea*. *The New Phytologist*, **78**, 625–635.

Norton TA (1978) The factors influencing the distribution of *Saccorhiza polyschides* in the region of Lough Ine. *Journal of the Marine Biological Association of the United Kingdom*, **58**, 527.

O

OBIS (2015) The OBIS Web Portal search interface.

Omstedt A (1987) Water cooling in the entrance of the Baltic Sea. *Tellus*, **39 A**, 254–265.

Ondiviela B, Losada IJ, Lara JL, Maza M, Galván C, Bouma TJ, van Belzen J (2014) The role of seagrasses in coastal protection in a changing climate. *Coastal Engineering*, **87**, 158–168.

P

Pace M, Borg JA, Galdies C, Malhotra A (2017) Influence of wave climate on architecture and landscape characteristics of *Posidonia oceanica* meadows. *Marine Ecology*, **38**, 1–14.

Pearman PB, Guisan A, Broennimann O, Randin CF (2008) Niche dynamics in space and time. *Trends in Ecology and Evolution*, **23**, 149–158.

Pearson K (1901) LIII. *On lines and planes of closest fit to systems of points in space*. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **2**, 559–572.

Pearson G a., Lago-Leston A, Mota C (2009) Frayed at the edges: Selective pressure and adaptive response to abiotic stressors are mismatched in low diversity edge populations. *Journal of Ecology*, **97**, 450–462.

Pereira TR (2014) *Ecology cology of west Iberian kelps: Laminaria ochroleuca and*

Saccorhiza polyschides living near their distributional limits. Universidade do Algarve, 160 pp.

Perez J, Menendez M, Mendez FJ, Losada IJ (2014) Evaluating the performance of CMIP3 and CMIP5 global climate models over the north-east Atlantic region. *Climate Dynamics*, **43**, 2663–2680.

Perez J, Menendez M, Losada IJ (2017) GOW2: A global wave hindcast for coastal applications. *Coastal Engineering*, **124**, 1–11.

Perruche C, Paul J, Drévillon M (2015) *Quality information document for global biogeochemical non assimilative hindcast product*. 1–31 pp.

Peterson A, Herkül K (2017) Mapping benthic biodiversity using georeferenced environmental data and predictive modeling. *Marine Biodiversity*, 1–16.

Peterson AT, Soberón J, Pearson RG, Anderson RP, Martínez-Meyer E, Nakamura M, Bastos Araújo M (2011) *Ecological niches and geographic distributions*, Monographs edn. Princeton University Press, Oxfordshire, 316 pp.

Petitpierre B, Broennimann O, Kueffer C, Daehler C, Guisan A (2017) Selecting predictors to maximize the transferability of species distribution models: lessons from cross-continental plant invasions. *Global Ecology and Biogeography*, **26**, 275–287.

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Piñeiro-Corbeira C, Barreiro R, Cremades J (2016) Decadal changes in the distribution of common intertidal seaweeds in Galicia (NW Iberia). *Marine Environmental Research*, **113**, 106–115.

Piñeiro-Corbeira C, Barreiro R, Cremades J, Arenas F (2018) Seaweed assemblages under a climate change scenario: Functional responses to temperature of eight intertidal seaweeds match recent abundance shifts. *Scientific Reports*, **8**.

Poulos SE, Collins MB (1997) Seasonal variability in sea surface oceanographic conditions in the Aegean Sea (Eastern Mediterranean): an overview. *Journal of Marine Systems*, **13**, 225–244.

Puente A, Guinda X, Juanes JA et al. (2017) The role of physical variables in biodiversity patterns of intertidal macroalgae along European coasts. *Journal of the Marine Biological Association of the United Kingdom*, **97**.

Pulliam HR (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349–361.

Q

Qiao H, Feng X, Escobar LE, Peterson AT, Soberón J, Zhu G, Papeş M (2019) An evaluation of transferability of ecological niche models. *Ecography*, 521–534.

Quintano E, Díez I, Muguerza N, Santolaria A, Gorostiaga JM, Gorostiaga JM (2015) Epiphytic flora on *Gelidium corneum* (Rhodophyta: Gelidiales) in relation to wave exposure and depth. *Scientia Marina*, **79**, 479–486.

Quintano E, Celis-Plá PSM, Martínez B, Díez I, Muguerza N, Figueroa FL, Gorostiaga JM (2019) Ecophysiological responses of a threatened red alga to increased irradiance in an in situ transplant experiment. *Marine Environmental Research*, 0–1.

R

R Core Team (2018) *R: A language and environment for statistical computing*. Computing, R Foundation for Statistical, Vienna, Austria.

Raina N, Rao YSCKKS (2014) Comparative analysis of species distribution modeling of *Daphne papyracea* in Dabka Watershed Nainital District, Uttarakhand. *Journal of Environment and Earth Science*, **4**, 19–30.

Ramos E, Puente A, Juanes JA et al. (2014) Biological validation of physical coastal waters classification along the NE Atlantic region based on rocky macroalgae distribution. *Estuarine, Coastal and Shelf Science*, **147**, 103–112.

Ramos E, Puente A, Guinda X, Juanes JA (2016a) A hierarchical classification system along the NE Atlantic coast: focusing on the local scale (Cantabria, N Spain). *European Journal of Phycology*.

Ramos E, Puente A, Juanes JA (2016b) An ecological classification of rocky shores at a regional scale: a predictive tool for management of conservation values. *Marine Ecology*, **37**, 311–328.

Ramos E, Díaz de Terán JR, Puente A, Juanes JA (2016c) The role of geomorphology in the distribution of intertidal rocky macroalgae in the NE Atlantic region. *Estuarine, Coastal and Shelf Science*, **179**, 90–98.

Ramos E, Puente A, Guinda X, Juanes JA (2017) A hierarchical ecological classification system along the NE Atlantic coast: focusing on the local scale (Cantabria, N Spain). *European Journal of Phycology*, **52**, 75–89.

Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*, **33**, 1689–1703.

Ready J, Kaschner K, South AB et al. (2010) Predicting the distributions of marine organisms at the global scale. *Ecological Modelling*, **221**, 467–478.

Regos A, Gagne L, Alcaraz-Segura D, Honrado JP, Domínguez J (2019) Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Scientific Reports*, **9**, 1–14.

Reguero BG, Menéndez M, Méndez FJ, Mínguez R, Losada IJ (2012) A Global Ocean Wave (GOW) calibrated reanalysis from 1948 onwards. *Coastal Engineering*, **65**, 38–55.

Rengstorf AM, Mohn C, Brown C, Wisz MS, Grehan AJ (2014) Predicting the distribution of deep-sea vulnerable marine ecosystems using high-resolution data: Considerations and novel approaches. *Deep-Sea Research Part I: Oceanographic Research Papers*, **93**, 72–82.

Rhein M, Rintoul SR, Aoki S, Campos E, Chambers D, Freely RA (2013) Climate Change 2013. In: *The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J), pp. 255–316. Cambridge, New York.

Rico JM (1992) *Ecología, fisiología y cultivo de Gelidium latifolium (Grev.) Thur. et Born. de la región de Cabo Peñas (N de España)*. 188 pp.

Riis T, Olesen B, Clayton JS, Lambertini C, Brix H, Sorrell BK (2012) Growth and morphology in relation to temperature and light availability during the establishment of three invasive aquatic plant species. *Aquatic Botany*, **102**, 56–64.

Robertson MP, Caithness N, Villet MH (2001) A PCA-based modelling technique for predicting environmental suitability for organisms from presence records. *Diversity and Distributions*, **7**, 15–27.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, **12**.

Robinson LM, Elith J, Hobday AJ, Pearson RG, Kendall BE, Possingham HP, Richardson AJ (2011) Pushing the limits in marine species distribution modelling: Lessons from the land present challenges and opportunities. *Global Ecology and Biogeography*, **20**, 789–802.

Robinson NM, Nelson WA, Costello MJ, Sutherland JE, Lundquist CJ (2017) A Systematic Review of Marine-Based Species Distribution Models (SDMs) with Recommendations for Best Practice. *Frontiers in Marine Science*, **4**, 1–11.

Rodríguez L, García JJ, Carreño F, Martínez B (2019) Integration of physiological knowledge into hybrid species distribution modelling to improve forecast of distributional shifts of tropical corals. *Diversity and Distributions*, 1–14.

Roulund A, Sutherland J, Todd D, Sterner J (2016) Parametric equations for Shields parameter and wave orbital velocity in combined current and irregular waves. In: *8 th International Conference of scour and erosion* (eds Harris J, Whitehouse R, Moxon S), pp. 313–318. Taylor & Francis Group, Oxford, UK.

Rueness J, Fredriksen S (1989) Field and culture studies of *Gelidium latifolium* (Grev.) Born. & Thur. (Rhodophyta) from Norway. *Dep. of Biology Marine Botany*, **11**, 177–185.

S

Saha S, Moorthi S, Pan HL et al. (2010) The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, **91**, 1015–1057.

Saha S, Moorthi S, Wu X et al. (2014) The NCEP climate forecast system version 2. *Journal of Climate*, **27**, 2185–2208.

Salinas JM, Granado R (1984) *Estudio ficológico de la costa de Vizcaya*. 247 pp.

Sánchez Í, Fernández C, Arrontes J (2005) Long-term changes in the structure of intertidal assemblages after invasion by *Sargassum muticum* (Phaeophyta). *Journal of Phycology*, **41**, 942–949.

Santelices B (1974) *Gelidioid algae, a brief resume of the pertinent literature*. Honolulu, 111 pp.

Sarhan T, Lafuente JG, Vargas M, Vargas JM, Plaza F (2000) Upwelling meechanisms in the northwestern Alboran Sea. *Journal of Marine Systems*, **23**, 317–331.

Sax DF, Early R, Bellemare J (2013) Niche syndromes, species extinction risks, and management under climate change. *Trends in Ecology and Evolution*, **28**, 517–523.

Sbrocco EJ, Barber PH (2013) MARSPEC: ocean climate layers for marine spatial ecology. *Ecology*, **94**, 979.

Schanz A, Asmus H (2003) Impact of hydrodynamics on development and morphology of intertidal seagrasses in the Wadden Sea. *Marine Ecology Progress Series*, **261**, 123–134.

Schibalski A, Lehtonen A, Schröder B (2014) Climate change shifts environmental space and limits transferability of treeline models. *Ecography*, **37**, 321–335.

Schiel DR (2006) Rivets or bolts? When single species count in the function of temperate rocky reef communities. *Journal of Experimental Marine Biology and Ecology*, **338**, 233–252.

Schoener T (1989) Food webs from the small to the large: the Robert H. MacArthur award lecture. *Ecology*, **70**, 1889–1589.

Schuckmann K Von, Traon P Le, Alvarez-fanjul E, Axell L, Balmaseda M (2016) The Copernicus Marine Environment Monitoring Service Ocean State Report. *Journal of operational oceanography*, **9**, 235–320.

Seabra R, Wethey DS, Lima FP, Campo R (2015) Understanding complex biogeographic responses to climate change. *Scientific Reports*, **5**, 1–6.

Segurado P, Araujo MB (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.

Sequeira AMM, Mellin C, Lozano-Montes HM et al. (2016) Transferability of predictive models of coral reef fish species richness. *Journal of Applied Ecology*, **53**, 64–72.

Sequeira AMM, Bouchet PJ, Yates KL, Mengersen K, Caley MJ (2018) Transferring biodiversity models for conservation: Opportunities and challenges. *Methods in Ecology and Evolution*, **9**, 1250–1264.

Short FT, Neckles HA (1999) The effects of global climate change on seagrasses. *Aquatic Botany*, **63**, 169–196.

Simões MVP, Peterson AT (2018) Importance of biotic predictors in estimation of potential invasive areas: the example of the tortoise beetle *Eurypedus nigrosignatus* , in Hispaniola. *PeerJ*, **6**, e6052.

Sjotun K, Fredriksen S (1995) Growth allocation in *Laminaria hyperborea* (Laminariales, Phaeophyceae) in relation to age and wave exposure. *Marine Ecology Progress Series*, **126**, 213–222.

Skene KR (2004) Key differences in photosynthetic characteristics of nine species of intertidal macroalgae are related to their position on the shore. *Canadian Journal of Botany*, **82**, 177–184.

Slangen ABA, Carson M, Katsman CA, van de Wal RSW, Vermeersen LLA, Stammer D (2014) Projecting twenty-first century regional sea-level changes. *Climatic Change*, **124**, 317–332.

Smale DA, Wernberg T (2013) Extreme climatic event drives range contraction of a habitat-forming species. In: *Proceedings of the Royal Society*, pp. 1–9. Royal Society Publishing.

Smale DA, Wernberg T, Yunnie ALE, Vance T (2015) The rise of *Laminaria ochroleuca* in the Western English Channel (UK) and comparisons with its competitor and assemblage dominant *Laminaria hyperborea*. *Marine Ecology*, **36**, 1033–1044.

Soberón J, Peterson AT (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**, 1–10.

Soulsby RL (1983) The bottom boundary layer of shelf seas. In: *Physical Oceanography of Coastal and Shelf Seas*, Vol. 35, pp. 189–266. Elsevier, Amsterdam.

Soulsby RL (1997) *Dynamics of marine sands: a manual for practical applications* (ed Telford T). London, 249 pp.

Soulsby RL, Hamm L, Klopman G, Myrhaug D, Simons RR, Thomas GP (1993) Wave-current interaction within and outside the bottom boundary layer. *Coastal Engineering*, **21**, 41–69.

Spillman CM, Imberger J, Hamilton DP, Hipsey MR, Romero JR (2007) Modelling the effects of Po River discharge, internal nutrient cycling and hydrodynamics on biogeochemistry of the Northern Adriatic Sea. *Journal of Marine Systems*, **68**, 167–200.

Stæhr PA, Pedersen MF, Thomsen MS, Wernberg T, Krause-Jensen D (2000) Invasion of *Sargassum muticum* in Limfjorden (Denmark) and its possible impact on the indigenous macroalgal community. *Marine Ecology Progress Series*, **207**, 79–88.

Stark JD, Donlon CJ, Martin MJ, McCulloch ME (2007) OSTIA : An operational, high resolution, real time, global sea surface temperature analysis system. In: *OCEANS 2007. Marine challenges: coastline to deep sea*, pp. 1–4. Aberdeen, Scotland.

Steen H (2004) Effects of reduced salinity on reproduction and germling development in *Sargassum muticum* (Phaeophyceae, Fucales). *European Journal of Phycology*, **39**, 293–299.

Steen F, Aragay J, Zuljevic A et al. (2017) Tracing the introduction history of the brown seaweed *Dictyota cyanoloma* (Phaeophyceae, Dictyotales) in Europe. *European Journal of Phycology*, **52**, 31–42.

Suárez-Seoane S, Álvarez-Martínez JM, Palacín C, Alonso JC (2018) From general research questions to specific answers: Underspecificity as a source of uncertainty in biological conservation. *Biological Conservation*, **227**, 167–180.

Sutcliffe PR, Mellin C, Pitcher CR, Possingham HP, Caley MJ (2014) Regional-scale patterns and predictors of species richness and abundance across twelve major tropical inter-reef taxa. *Ecography*, **37**, 162–171.

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.

T

Takolander A, Leskinen E, Cabeza M (2017) Synergistic effects of extreme temperature and low salinity on foundational macroalga *Fucus vesiculosus* in the northern Baltic Sea. *Journal of Experimental Marine Biology and Ecology*, **495**, 110–118.

Tamburello L, Bulleri F, Balata D, Benedetti-Cecchi L (2014) The role of overgrazing and anthropogenic disturbance in shaping spatial patterns of distribution of an invasive seaweed. *Journal of Applied Ecology*, **51**, 406–414.

Tapia L, Regos A, Gil-Carrera A, Domínguez J (2017) Assessing the temporal transferability of raptor distribution models: Implications for conservation. *Bird Conservation International*, 1–15.

Taylor KE, Stouffer RJ, Meehl GA, Taylor KE, Stouffer RJ, Meehl GA (2012) An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93**, 485–498.

Teagle H, Hawkins SJ, Moore PJ, Smale DA (2017) The role of kelp species as biogenic habitat formers in coastal marine ecosystems. *Journal of Experimental Marine*

Biology and Ecology, **492**, 81–98.

Thibaud E, Petitpierre B, Broennimann O, Davison AC, Guisan A (2014) Measuring the relative effect of factors affecting species distribution model predictions. *Methods in Ecology and Evolution*, **5**, 947–955.

Thomsen MS, Wernberg T, Stæhr PA, Pedersen MF (2006) Spatio-temporal distribution patterns of the invasive macroalga *Sargassum muticum* within a Danish *Sargassum*-bed. *Helgoland Marine Research*, **60**, 50–58.

Thuiller W, Richardson DM, Py Ek P, Midgley GF, Hughes GO, Rouget M (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.

Thurstan RH, McClenachan L, Crowder LB et al. (2015) Filling historical data gaps to foster solutions in marine conservation. *Ocean and Coastal Management*, **115**, 31–40.

Tingley MW, Monahan WB, Beissinger SR, Moritz C (2009) Birds track their Grinnellian niche through a century of climate change. *Proceedings of the National Academy of Sciences*, **106**, 19637–19643.

Tomás A, Méndez FJ, Losada ÍJ (2012) *Bed shear stress during Sant Esteve´s storm (26th December 2008) along the Catalonia´s coast (NW Mediterranean)*. 45–54 pp.

Tomás A, Méndez FJ, Medina R et al. (2016) A methodology to estimate wave-induced coastal flooding hazard maps in Spain. *Journal of Flood Risk Management*, **9**, 289–305.

Torresan S, Critto A, Rizzi J, Zabeo A, Furlan E, Marcomini A (2016) DESYCO: A decision support system for the regional risk assessment of climate change impacts in coastal zones. *Ocean & Coastal Management*, **120**, 49–63.

Touchette BW (2007) Seagrass-salinity interactions: Physiological mechanisms used by submersed marine angiosperms for a life at sea. *Journal of Experimental Marine Biology and Ecology*, **350**, 194–215.

Tuanmu MN, Viña A, Roloff GJ, Liu W, Ouyang Z, Zhang H, Liu J (2011) Temporal transferability of wildlife habitat models: Implications for habitat monitoring. *Journal of Biogeography*, **38**, 1510–1523.

Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003) Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution*, **18**, 306–314.

Tyberghein L, Verbruggen H, Pauly K, Troupin C, Mineur F, De Clerck O (2012) Bio-ORACLE: A global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, **21**, 272–281.

U

UC (2006) *Programa de seguimiento ambiental del impacto del vertido del buque "Prestige" en las costas de Cantabria*.

V

Valiela I (2001) *Doing science: design, analysis, and communication of scientific research*. Oxford University Press, New York, 303 pp.

Valle M, Chust G, del Campo A, Wisz MS, Olsen SM, Garmendia JM, Borja Á (2014) Projecting future distribution of the seagrass *Zostera noltii* under global warming and sea level rise. *Biological Conservation*, **170**, 74–85.

Vapnik V (1995) *The nature of statistical learning theory*. Springer.

Vasseur DA, Delong JP, Gilbert B et al. (2014) Increased temperature variation poses a greater risk to species than climate warming. In: *Proceedings of the Royal Society B: Biological Sciences 281*.

Vega GC, Pertierra LR, Olalla-Táraga MÁ (2017) MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. *Scientific Data*, **4**, 170078.

Vesanto J (1999) SOM-based data visualization methods. *Intelligent Data Analysis*, **3**, 111–126.

Vesanto J, Himberg J, Alhoniemi E, Parhankangas J (2000) SOM Toolbox for Matlab 5. *Technical Report A57*, **2**, 59.

Vousdoukas MI, Velegrakis a. F, Paul M, Dimitriadis C, Makrykosta E, Koutsoubas D (2012) Field observations and modeling of wave attenuation over colonized beachrocks. *Continental Shelf Research*, **48**, 100–109.

Vranken S, Bosch S, Peña V, Leliaert F, Mineur F, De Clerck O (2018) A risk assessment of aquarium trade introductions of seaweed in European waters. *Biological Invasions*, **20**, 1171–1187.

W

Warren DL, Seifert SN (2010) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335–342.

Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution; international journal of organic evolution*, **62**, 2868–83.

Waycott M, Collier C, Mcmahon K, Ralph P, Mckenzie L, Udy J, Grech A (2007) Vulnerability of seagrasses in the Great Barrier Reef to climate change. In: *Climate*

Change and the Great Barrier Reef (eds Johnson JE, Marshall PA), pp. 193–235. Park Authority and Australian Greenhouse Office, Townsville, QLD, Australia.

Weatherall P, Marks KM, Jakobsson M et al. (2015) A new digital bathymetric model of the world's oceans. *Earth and Space Science*, **2**, 331–345.

Werner FJ, Graiff A, Matthiessen B (2016) Temperature effects on seaweed-sustaining top-down control vary with season. *Oecologia*, **180**, 889–901.

Wethey DS, Woodin S A. (2008) Ecological hindcasting of biogeographic responses to climate change in the European intertidal zone. *Hydrobiologia*, **606**, 139–151.

Wiens JJ, Ackerly DD, Allen AP et al. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310–1324.

Willette DA, Ambrose RF (2009) The distribution and expansion of the invasive seagrass *Halophila stipulacea* in Dominica, West Indies, with a preliminary report from St. Lucia. *Aquatic Botany*, **91**, 137–142.

Wilson KL, Skinner MA, Lotze HK (2019) Projected 21st-century distribution of canopy-forming seaweeds in the Northwest Atlantic with climate change. *Diversity and Distributions*.

de Winter RC, Ruessink BG (2017) Sensitivity analysis of climate change impacts on dune erosion: case study for the Dutch Holland coast. *Climatic Change*, **141**, 685–701.

Wisz MS, Guisan A (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, **9**, 1–13.

Wisz MS, Hijmans RJ, Li J et al. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

Y

Yates KL, Bouchet PJ, Caley MJ et al. (2018) Outstanding Challenges in the Transferability of Ecological Models. *Trends in Ecology & Evolution*, **xx**, 1–13.

Yesson C, Bush LE, Davies AJ, Maggs CA, Brodie J (2015) Large brown seaweeds of the British Isles: Evidence of changes in abundance over four decades. *Estuarine, Coastal and Shelf Science*, **155**, 1–9.

Young M, Ierodiaconou D, Womersley T (2015) Forests of the sea: Predictive habitat modelling to assess the abundance of canopy forming kelp forests on temperate reefs. *Remote Sensing of Environment*, **170**, 178–187.

Z

Zaniewski AE, Lehmann A, Overton JM (2002) Predicting species spatial distributions using presence-only data : a case study of native New Zealand ferns. **157**, 261–280.

Zavatarelli M, Baretta JW, Baretta-Bekker JG, Pinardi N (2000) The dymnamics of the Adriatic Sea ecosystem. An idealised model study. *Deep-Sea Research Part I: Oceanographic Research Papers*, **47**, 937–970.

Zharikov Y, Elner RW, Shepherd PCF, Lank DB (2009) Interplay between physical and predator landscapes affects transferability of shorebird distribution models. *Landscape Ecology*, **24**, 129–144.

Annex A

Ecological typologies of large areas. An application in the Mediterranean Sea

Journal of Environmental Management 205 (2018) 59–72

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Research article

Ecological typologies of large areas. An application in the Mediterranean Sea

Camino F. de la Hoz [a], Elvira Ramos [a], Araceli Puente [a, *], Fernando Méndez [b], Melisa Menéndez [a], José A. Juanes [a], Íñigo J. Losada [a]

[a] Environmental Hydraulics Institute, Universidad de Cantabria, Avda. Isabel Torres, 15, Parque Científico y Tecnológico de Cantabria, 39011, Santander, Spain
[b] Departamento Ciencias y Tecnicas del Agua y del Medio Ambiente, Universidad de Cantabria, Santander, Spain

ARTICLE INFO

Article history:
Received 31 July 2017
Received in revised form
20 September 2017
Accepted 21 September 2017

Keywords:
Meteo-oceanographic variables
Seagrasses
Environmental management
Regions
SOM
K-means

ABSTRACT

One approach to identifying and mapping the state of marine biophysical conditions is the identification of large-scale ecological units for which conditions are similar and the strategies of management may also be similar. Because biological processes are difficult to directly record over large areas, abiotic characteristics are used as surrogate parameters. In this work, the Mediterranean Sea was classified into homogeneous spatial areas based on abiotic variables. Eight parameters were selected based on salinity, sea surface temperature, photosynthetically active radiation, sea-wave heights and depth variables. The parameters were gathered in grid points of 0.5° spatial resolution in the open sea and 0.125° in coastal areas. The typologies were obtained by data mining the eight parameters throughout the Mediterranean and combining two clustering techniques: self-organizing maps and the k-means algorithm. The result is a division of the Mediterranean Sea into seven typologies. For these typologies, the classification recognizes differences in temperature, salinity and radiation. In addition, it separates coastal from deep areas. The influence of river discharges and the entrance of water from other seas are also reflected. These results are consistent with the ecological requirements of the five studied seagrasses (*Posidonia oceanica*, *Zostera marina*, *Zostera noltei*, *Cymodocea nodosa*, *Halophila stipulacea*), supporting the suitability of the resulting classification and the proposed methodology. The approach thus provides a tool for the sustainable management of large marine areas and the ability to address not only present threats but also future conditions, such as climate change.

1. Introduction

The Mediterranean Sea represents only 0.82% of the ocean surface, but with nearly 17,000 known marine species, it is home to 4–18% of global marine biodiversity and 10–48% of endemic species depending on the groups (Coll et al., 2010). Due to the high biodiversity of the Mediterranean, many habitats and species are under protective regimes (Council of the European Commission, 1992; IUCN, 2015; UNEP-MAP-RAC/SPA, 1995), justifying the generation of a network of protected areas as a tool for marine and coastal ecosystems conservation (Rodríguez-Rodríguez et al., 2016). The lack of knowledge regarding species and habitat distribution could, however, compromise the sustainable management of those connected species as a whole. One approach to reducing this uncertainty is the identification of large-scale ecological units that are more manageable to assist in the management of a biogeographically complex region (Gregr et al., 2012; Last et al., 2010). Notarbartolo di Sciara and Agardy (2010a) established that subdividing the Mediterranean into broadly homogeneous subunits would help in priority setting and planning for marine conservation in the region. This division allows the recognition of different parts of the region, where ecological and abiotic conditions are similar and the management strategies could also be similar (Howell, 2010).

The Mediterranean has been traditionally divided into two large subregions (East and West). This division into two ocean basins is, however, too coarse to detect typologies that influence species distribution. In recent years, more detailed classifications of the Mediterranean Sea have appeared to solve this issue (Table 1). Several studies take into account geomorphological and

---

* Corresponding author.
E-mail address: puentea@unican.es (A. Puente).

https://doi.org/10.1016/j.jenvman.2017.09.058

Table 1
The most widely used classifications existing in the Mediterranean Sea.

Reference	Divisions	Criteria
Bianchi 2007	1. Alborán Sea 2. Algerian and north Tunisia coasts 3. Southern Tyrrhenian Sea 4. Balearic Sea to Sardinia Sea 5. Gulf of Lions and Ligurian Sea 6. Northern Adriatic Sea 7. Central Adriatic Sea 8. Southern Adriatic Sea 9. Ionian Sea 10. Northern Aegean Sea 11. Southern Aegean Sea 12. Levant Sea 13. Straits of Messina	Biogeographic peculiarities
Spalding et al., 2007	1. Adriatic Sea (30) 2. Aegean Sea (31) 3. Levantine Sea (32) 4. Tunisian Plateau/Gulf of Sidra (33) 5. Ionian Sea (34) 6. Western Mediterranean (35) 7. Alborán Sea (36)	Biogeographic basis (range discontinuities, dominant habitats, geomorphological features, currents, temperatures, etc.) Practical utility (nested system, operating globally) Parsimony
European Commission 2008	1. Western Mediterranean 2. Adriatic Sea 3. Ionian Sea and the Central Mediterranean Sea 4. Aegean-Levantine Sea	
Notarbartolo di Sciara and Agardy 2010b	1. Alborán Sea 2. Algero-Provenzal Basin 3. Tyrrhenian Sea 4. Adriatic Sea 5. Ionian Sea 6. Tunisian Plateau/Gulf of Sidra 7. Aegean Sea 8. Levantine Sea	Depth profiles Oceanographic and biogeographical differences
Berline et al., 2014	22 regions	Oceanographic distances (dispersal by ocean circulation)
Rossi et al., 2014	65 provinces with a larval duration of 30 days and 32 provinces for a larval duration of 60 says	Larval dispersal
Gabrie et al., 2012	37 epipelagic regions (level III) 16 bioregions (level II) 5 bioregions (level I)	Statistical analysis
Nieblas et al., 2014	5 subprovinces	Statistical analysis
Reygondeau et al., 2017	12 epipelagic layers	Statistical analysis
Witze 2017	The Ecological Marine Units (EMU) number 9 and 21	Statistical analysis

biogeographic features (Bianchi, 2007; European Commission, 2008; Notarbartolo di Sciara and Agardy, 2010b; Spalding et al., 2007); however, they have found that regional patterns are limited to continuous areas, and areas geographically separated but with similar abiotic conditions cannot be considered to have the same typology. Recent studies have characterized the Mediterranean provinces according to abiotic variables using statistical methods to objectively subdivide the area. Some of these classifications have a clear objective, for example, larval dispersal by oceanic currents (Berline et al., 2014; Rossi et al., 2014), while others are broadly oriented, employing a more complete set of variables (Gabrie et al., 2012; Nieblas et al., 2014; Reygondeau et al., 2017; Witze, 2017).

These works have included variables related to the abiotic conditions of the water, providing a wider view of the Mediterranean Sea than biogeographic approaches, which have been geographically limited to the definition of regional seas. However, a comprehensive study also considering meteo-oceanographic (hereinafter referred to as met-ocean) variables has not yet been performed. The met-ocean variables (e.g., temperature and sea surface waves) are related to the survival of species, especially floral species that support the whole ecosystem. Changes in their distributions will therefore have cascading effects on the rest of the community. An initial key step for a robust characterization considering climate features is the use of long temporal data series from validated sources. Moreover, not only mean values but also extreme and dispersion information parameters of the variables may better capture the functioning and evolution of the system. This approach would allow the detection of boundaries to the distribution of species in a more reliable way and would allow for better predictions of changes in species ranges under future climate conditions.

Therefore, there is a need for a Mediterranean Sea classification that takes into account these aspects and allows homogeneous areas to be distinguished from an ecological point of view, namely, one that is linked to species distributions. Nevertheless, the scarcity and lack of homogeneity in the biological information makes it difficult to address this perspective. The use of abiotic variables as surrogates of ecological processes has been shown to be advantageous to resolving this issue (Huang et al., 2015; Juanes et al., 2016; Last et al., 2010; Ramos et al., 2012; Zacharias and Roff, 2000). For example, Ramos et al. (2012) established physically homogeneous coastal zones in the Northeast Atlantic region based on abiotic variables for the potential distribution of macroalgae. Their results demonstrated that met-ocean variables are strongly related to macroalgae distribution; thus, these types of variables can be used as surrogates of biological information (Ramos et al., 2016b, 2016c, 2014).

The selection of species is also a key step to ensure that they are representative of the ecosystem. A possible approach is the use of ecological engineers because they are good indicators of the health of the ecosystem, such as seagrasses (Roca et al., 2016). Seagrasses are considered a priority habitat for conservation by the EU Habitats Directive (Council of the European Commission, 1992) because they are important nursery grounds for a large number of fish and invertebrate species, thereby contributing to the maintenance of marine biodiversity and providing ecosystem services (Duarte et al., 2013; Green and Short, 2003; Ondiviela et al., 2014).

It is necessary to test whether the abiotic variables considered can be used as surrogates of selected species distribution; thus, to verify the suitability of the classification, biological data are needed (Galván et al., 2010; Ramos et al., 2016a, 2016c, 2014). One of the main problems encountered in large marine areas is the lack of an accurate homogeneous cartography. Although the increase in environmental policies has led to greater efforts in the cartography of marine communities, most submerged habitat types have not yet been comprehensively mapped in the entire Mediterranean Sea (Fraschetti et al., 2008). According to species distribution, the most endangered species are the best mapped in the Mediterranean Sea, such as *Posidonia oceanica* or *Cymodocea nodosa* (Marbà et al., 2013). Nevertheless, these cartographies are mainly regional and placed next to the coast and in the supralittoral and intertidal zones due to the high economic and time costs of sampling in the sublittoral zone (Francour, 1997; Sims et al., 2003; Azzellino et al., 2008; Panigada et al., 2008; Holcer et al., 2013; Belando et al., 2015).

The data must be obtained from the best sources available in the study area to achieve a reliable classification. The available data sources are usually grouped into three categories: direct observations, indirect observations using remote sensing devices, and numerical modeling outputs (considered in this work as hindcast and reanalysis). For direct observations, one of the most traditional ways to collect abiotic data is through in situ measurements, where the quality is strongly dependent on the quality of the analysis results. Despite numerous surveys and observation programs carried out in the Mediterranean, a homogeneous data bank has not been compiled in a systematic way, mainly because the average density of the observations is rather low during autumn and winter and more data are available for the West Basin than for the East Basin (Brasseur et al., 1996). The growing availability and reliability of environmental information through remote sensing techniques in recent years may provide a tool to overcome the limitations of in situ measurements. Satellite imagery has the advantage of increasing data collection frequency and expanding areal coverage, offering repeatable, standardized and verifiable information over various time periods when field sampling is not available (de Barbosa Araujo et al., 2015; Muller-Karger et al., 2005; Turner et al., 2003). As a complement to these sources, numerical modeling outputs provide spatially homogenous data at high spatial and temporal resolution. These datasets historically reconstruct climate conditions providing long-time series; however, they should be validated/calibrated against real observations (Reguero et al., 2012).

The final classification needs to be objective from a statistical point of view. The use of statistical tools reduces the level of subjectivity of the final classification to meet this requirement (Huang et al., 2015; Kleisner et al., 2015; Ramos et al., 2012). Such statistical analyses have been carried out in recent decades using clustering techniques to deal with the large amount of information that is grouped using statistical methods. The self-organizing maps (SOM) and the k-means algorithm are two widely used non-linear clustering techniques. Camus et al. (2011) proposed an approach based on a combination of these techniques, and Ramos et al. (2016b,c) and Galván et al. (2016) applied this approach on the north and northwest coasts of the Iberian Peninsula and in a Cantabrian estuary, respectively. Using the heuristic formula, quantization and topographic errors for the selection of the number of units of the SOM map, objective decision rules can be applied. In addition, the SOM map has shown the capacity to deal with complex environmental data and is efficient and performs well in ecological clustering compared with conventional statistical methods (Gevrey et al., 2003; Ramos et al., 2016c; Worner et al., 2015). Similar methodologies have been successfully applied to typify the diatom communities in a Dutch delta and to characterize the fish populations in lotic systems in Serbia (Goldenberg Vilar et al., 2014; Stojković et al., 2014).

Bearing in mind the importance of the Mediterranean Sea for biodiversity conservation, the development of a statistically objective classification should be a cost-effective method to enhance the policies of environmental management (Last et al., 2010). Thus, the aim of this work is to establish an objective Mediterranean Sea classification based on biologically meaningful abiotic variables acquired in a homogeneous way. Additionally, this study provides the first approach to testing the biological relevance of the selected variables and their parameters in the Mediterranean Sea.

2. Materials and methods

2.1. Study area

The study was conducted in the entire Mediterranean Sea, whose basin is the largest (2969 km^2) and deepest (average 1460 m, maximum 5267 m) of the enclosed seas on Earth (Coll et al., 2010). The Mediterranean is a regional sea located at the crossroads of Africa, Europe, and Asia. It connects through the Strait of Gibraltar to the Atlantic Ocean in the west, through the Dardanelles Sea to the Sea of Marmara and the Black Sea in the northeast, and in the southeast, the Suez Canal links it to the Red Sea and the Indian Ocean. The Strait of Sicily divides the Mediterranean Sea into two main subregions: the West Basin (0.85 million km^2) and the East Basin (1.65 million km^2).

To characterize the whole study area, a mesh was defined in the projected coordinate system ETRS 1989 Zone 30N with a longitude and latitude separation of 0.5°. Coastal zones were detailed with more definition (0.125° longitude and latitude) to capture the variability over the continental shelves. The final domain has 1337 grid points.

2.2. Abiotic data

The selection of variables needed is a key step that depends on the objective and the scale of the work (Juanes et al., 2016). To define the set of abiotic variables able to explain the biological variability of the area, the following criteria were applied: (1) variables should be related to geographical seagrass distribution, (2) data must be homogeneous in the entire study area and (3) parameters have to be representative of the Mediterranean Sea conditions. After the development of a sensitivity analysis through several tests and taking into account the above-mentioned criteria, the following variables were finally selected:

- **Sea surface temperature (SST)**. Temperature affects biochemical processes of organisms and determines different physiological responses, including growth rates and sexual reproduction efforts in seagrasses (Valle et al., 2014).
- **Salinity**. Due to the Mediterranean Sea characteristics, an evaporation basin, salinity is a limitation for species because they have to adapt their survival strategies. Changes in surface salinity affect the surface layer, where all primary production takes place (Vuorinen et al., 2015). In the case of seagrasses, salinity is a major environmental component that can influence the structure and function of their communities in some habitats (Nejrup and Pedersen, 2008; Sánchez-Lizaso et al., 2008; Touchette, 2007).
- **Photosynthetically active radiation (PAR)**. The range of solar radiation that photosynthetic organisms are able to use is decisive in their physiology and biomass production (Best et al., 2001; Riis et al., 2012). However, above certain levels, depending on the species, radiation can have negative effects on seagrasses (photoinhibition) (Larkum et al., 2006).
- **Significant wave height (Hs)**. Sea surface dynamics strongly influence the spatial distribution of seagrasses in exposed or submerged areas (Jensen and Denny, 2015; Ondiviela et al., 2014; Valle et al., 2013).
- **Depth**. Bathymetry provides information on the potential settlement areas of species in relation to the limit of light penetration into the water.

Three main data sources were employed. The salinity concentration was obtained from in situ measurements provided by the World Ocean Database (WOD) of the National Oceanic and Atmospheric Administration (NOAA)-NESDIS National Oceanographic Data Centre (NODC) (Levitus et al., 2013). To obtain comparable measures, the first 40 m mean value was employed, since it is the estimated maximum depth for photosynthetic organisms in the Mediterranean Sea (Templado González, 1995).

Photosynthetically active radiation was provided by satellite remote sensing, derived from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and the Moderate Resolution Imaging Spectroradiometer Aqua (MODIS Aqua) provided by the National Aeronautics and Space Administration (NASA) at level 3 (OB.DAAC, 2013). The SST values were supplied by the Operational Sea surface Temperature and sea-Ice concentration Analysis (OSTIA system) (Roberts-Jones et al., 2012). This dataset was designed to provide the best available estimate of the SST from a data assimilation procedure that includes a set of observations (satellites from infrared, microwave and AVHRR Pathfinder and in situ data) and operational reanalysis.

The sea surface wave dataset was from the Global Ocean Wave product GOW, developed by IHCantabria (Perez et al., 2017). The data were generated with a wave generation and propagation model, and the results were validated with satellite measurements and in situ buoy records. Bathymetry was obtained from the National Geophysical Data Center (National Oceanic and Atmospheric Administration, NOAA) (Amante and Eakins, 2009).

According to the different spatial resolutions of each dataset (Table 2), variables were obtained from the nearest available location to each grid-point of the defined geographical domain.

Parameters of the selected variables describing the average and the extreme behavior of the Mediterranean Sea were obtained, and their mutual influence was analyzed through Pearson's correlation coefficient ($r < 0.8$) (Table 2).

2.3. Classification

The selected datasets were classified following the methodology successfully applied by Ramos et al. (2016a,b,c) and Galván et al. (2016): a combination of the clustering techniques self-organizing maps (SOM) and k-means algorithm. Despite previous works being restricted to coastal areas, in this study, open seawaters were also included to ensure the accurate characterization of the whole Mediterranean Sea. This approach allowed us to account for biogeographic complexity in a large marine area.

The SOM extracts the patterns of high-dimensional data and projects them into a bidimensional organized space, allowing an intuitive visualization of the classification and the transformation of the distributions from the high-dimensional space into probability density functions on the lattice (Kohonen, 1998). The optimal map size (number of units) was chosen based on the heuristic formula proposed by Vesanto et al. (2000), $M = 5\sqrt{N}$, where M is the number of map units and N is the number of points in the mesh. In addition, quantization and topographic errors were calculated to support the number of units chosen. Quantization error is a measure of the map resolution because it is the average distance between each data vector and its best matching unit (Kohonen, 2001). Meanwhile, topographic error measures map quality because it represents the proportion of all data vectors for which the first and second best matching units are not adjacent (Kiviluoto, 1996). Datasets of the eight abiotic parameters characterizing each grid-point were used to run the SOM and were subsequently projected onto a two-dimensional (2D) map. The SOM analysis was conducted using MATLAB 8.1 and the SOM coding solution based on SOM Toolbox for MATLAB 5 (Vesanto et al., 2000). The number of units obtained with the application of the SOM may be high for a manageable classification. Thus, as the second step of the abiotic classification, the trained map was statistically organized in clusters using the k-means algorithm (Hartigan and Wong, 1979) to identify similar clusters representing the typologies. K-means is a nonhierarchical technique that produces clusters defined by a prototype and includes the samples with a closer distribution to its prototype. Once the number of clusters is established, the prototypes (centroids) are initialized, and on each iteration, the value nearest to each centroid is identified and the centroid is then redefined as the mean of the corresponding data. The algorithm is iteratively moved until the intragroup distance is minimal and the process converges. In this work, the minimum Davies-Bouldin Index was applied between 2 and 10 clusters to determine the best number of groups (Davies and Bouldin, 1979).

2.4. Biological suitability

Four native seagrasses (*Posidonia oceanica*, *Zostera marina*, *Zostera noltei*, *Cymodocea nodosa*) were selected due to their importance in the Mediterranean Sea ecosystem. The invasive species *Halophila stipulacea* was also considered to test the fit of non-native species with the methodology proposed.

Biological data were compiled from the following global databases: Global Biodiversity Information Facility (GBIF) (GBIF, 2013), Ocean Biogeographic Information (OBIS) (OBIS, 2015) and Towards COast to COast NETworks project (COCONET, 2016). Filters were applied to obtain the best biological database. First, only biological records collected in the same period as the abiotic variables (1985–2013) were considered to get comparable data. Additionally, over these data, four experts in seagrasses ecology of the University of Cantabria were consulted to detect possible mistakes in the species identification or the coordinate system. To confirm or discard the possible mistakes, literature information was searched to guarantee the accuracy of the presences (Borum et al., 2004; Di Martino et al., 2006; Green and Short, 2003). Finally, to avoid overfitting, the data were reduced to only one record per mesh point. Finally, species information was transformed to the presence/absence of the spatial domain, after which the results were

Table 2
Data acquisition characteristics for the quantification of each environmental variable.

Variables	Parameters	Source	Spatial resolution	Temporal resolution	Period
Salinity	Monthly average	In situ measurements (WOD, NOAA)			2005–2013
SST	Monthly average Annual average of the warmest month Annual standard deviation	OSTIA dataset (NASA)	0.05°	Daily	1985–2013
PAR	Monthly average Annual average of the month with the highest values	Satellite sensor (SeaWifs and Modis Aqua, NASA)	2 Km	Monthly	1999–2004
Hs	Monthly average	Regional hindcast (GOW, IH Cantabria)	0.125°	Hourly	1985–2013
Bathymetry (ETOPO1)		National Geophysical Data Center, NOAA	0.017°		

SST, sea surface temperature; PAR, photosynthetically active radiation; Hs, significant wave height; WOD, World Ocean Database; NOAA, National Oceanic and Atmospheric Administration, NASA, National Aeronautics and Space Administration, OSTIA, Operational Sea surface Temperature and sea-Ice concentration Analysis, GOW, Global Ocean Wave.

projected to SOM and k-means clusters to visualize the species spatial distribution according to abiotic classification.

3. Results

3.1. Abiotic data

The general trend of salinity was to increase eastward, although there were some exceptions. In the Aegean Sea, salinity increased from the north to the south, whereas on the north and west coasts of the Adriatic Sea, salinity was lower than in other nearshore areas. In addition, in the Gulf of Lion, some points with high salinity were observed (Fig. 1).

Regarding the monthly average SST, a clear north-south and west-east gradient was observed, with the coldest temperatures in the Alborán Sea and in the north of Algero-Provencal Basin, Tyrrhenian Sea, Adriatic Sea and Aegean Sea. In contrast, the highest values were situated in the east of the Levantine Basin. The pattern in the annual standard deviation was not so clear. The highest values were observed in semi-enclosed coastal regions, such as the Aegean Sea and the Sea of Marmara, the north of the Adriatic Sea and the Gulf of Gabès. The lowest values were found in the Alborán Sea and the west of Turkey. The annual average of the warmest month, August, presented a similar spatial pattern to the average SST map, with isolated coldest patches (the Alborán Sea, the north of the Algero-Provencal Basin and the Aegean Sea) and the warmest ones in the east of the Levantine Sea, in the Gulf of Sidra and the north of Sicily (Fig. 1).

As occurred with the average SST, monthly average PAR showed a clear north-south gradient. Nevertheless, a difference between basins was found for the June annual average (the month with the highest PAR values): the West Basin exhibited lower values, and the north of the East Basin exhibited higher values (Fig. 1).

As can be expected, significant wave height and depth showed a gradient from coastal areas, where the lowest values were in the open sea. In general, the western Mediterranean Sea has higher waves, specifically over the Balearic Sea located between the eastern coast of Spain, the southern coast of France, and the islands of Corsica and Sardinia. Enclosed seas, such as the Adriatic Sea and the Aegean Sea, exhibited low values. In contrast, the deepest area of the Mediterranean Sea is in the Ionian Sea (Fig. 1).

3.2. Classification

Based on the heuristic formula of Vesanto et al. (2000), the map size selected was 182 units (13 × 14 groups). The trained map had a

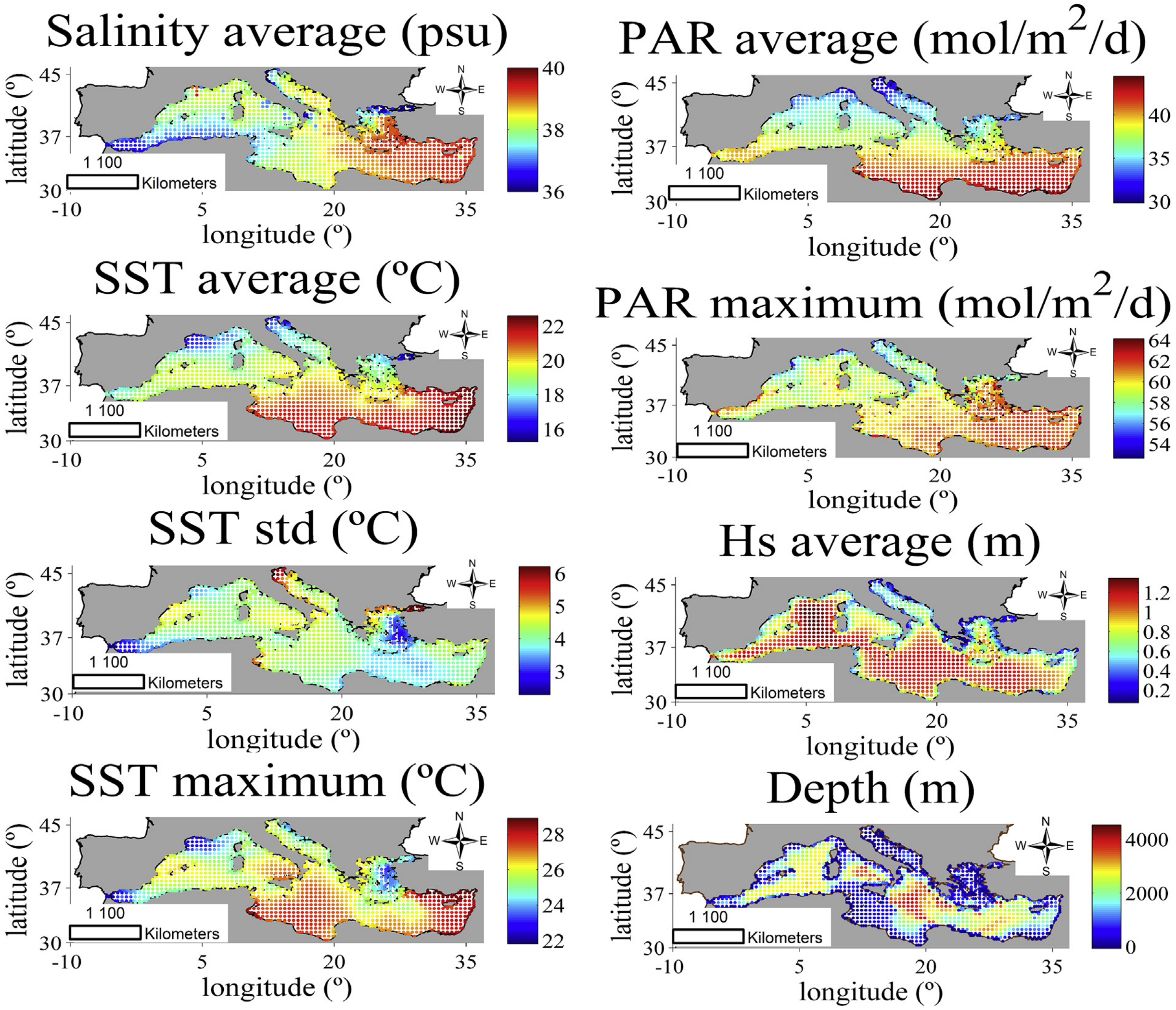


Fig. 1. Spatial distribution maps of the parameters of the abiotic variables along the Mediterranean Sea. From top left column: average salinity, average sea surface temperature (SST), standard deviation SST, maximum SST. From top right column: average photosynthetically active radiation (PAR), maximum PAR, average wave height (Hs) and depth.

quantization error of 0.143 and a topographic error of 0.027. Each unit of the SOM gathered a different number of grid points, with similar characteristics, preserving neighborhood relationships.

The SST average, SST maximum and PAR average showed similar patterns, with the highest values in the bottom right SOM units and the lowest values on the opposite side. In addition, the Hs average and depth had a positive association, with low Hs values in the shallow areas and the highest values in the deep zones (Fig. 2a).

The Davies-Bouldin Index was calculated for two to ten k-means clusters, obtaining values between 0.79 and 1.27. The optimal number of k-means clusters was seven considering the minimum Davies-Bouldin Index (0.79). The seven typologies in the SOM lattice and over the Mediterranean Sea are shown in Fig. 2. Some typologies exhibit a non-continuous geographic dispersal due to the exclusion of abiotic parameters.

The typologies obtained can be classified in two groups: one with low radiation, salinity and SST (typologies “Algerian Basin (I),” “Alborán-Aegean (II),” “North Coast (III)” and “Deep (IV)”) and the other characterized by saltier and hotter waters (typologies “East Basin deep (V),” “South East Basin (VI)” and “East Basin coast (VII)”). Fig. 3 shows the spatial distribution and multivariate distribution of the parameters for each typology. However, some differences among them can be observed:

- **Typology I, Algerian Basin**: mainly restricted to the open sea in the West Basin, with low SST, PAR and salinity values. This typology was deeper and more exposed than typologies “Alborán-Aegean (II)” and “North Coast (III)” but less than typology “Deep (IV).”
- **Typology II, Alborán-Aegean**: characterized by the influence of the cold water entrance with shallower areas, as in the “North Coast (III)” typology. However, this typology presented more stable SST and higher Hs and radiation values.
- **Typology III, North Coast**: encompassed shallow areas with the lowest values for Hs, radiation, SST and salinity. Nevertheless, the SST standard deviation was high.
- **Typology IV, Deep**: the key characteristic of this typology was the highest depth and Hs. Salinity was low as in previous typologies, while the values of the other variables were moderate.

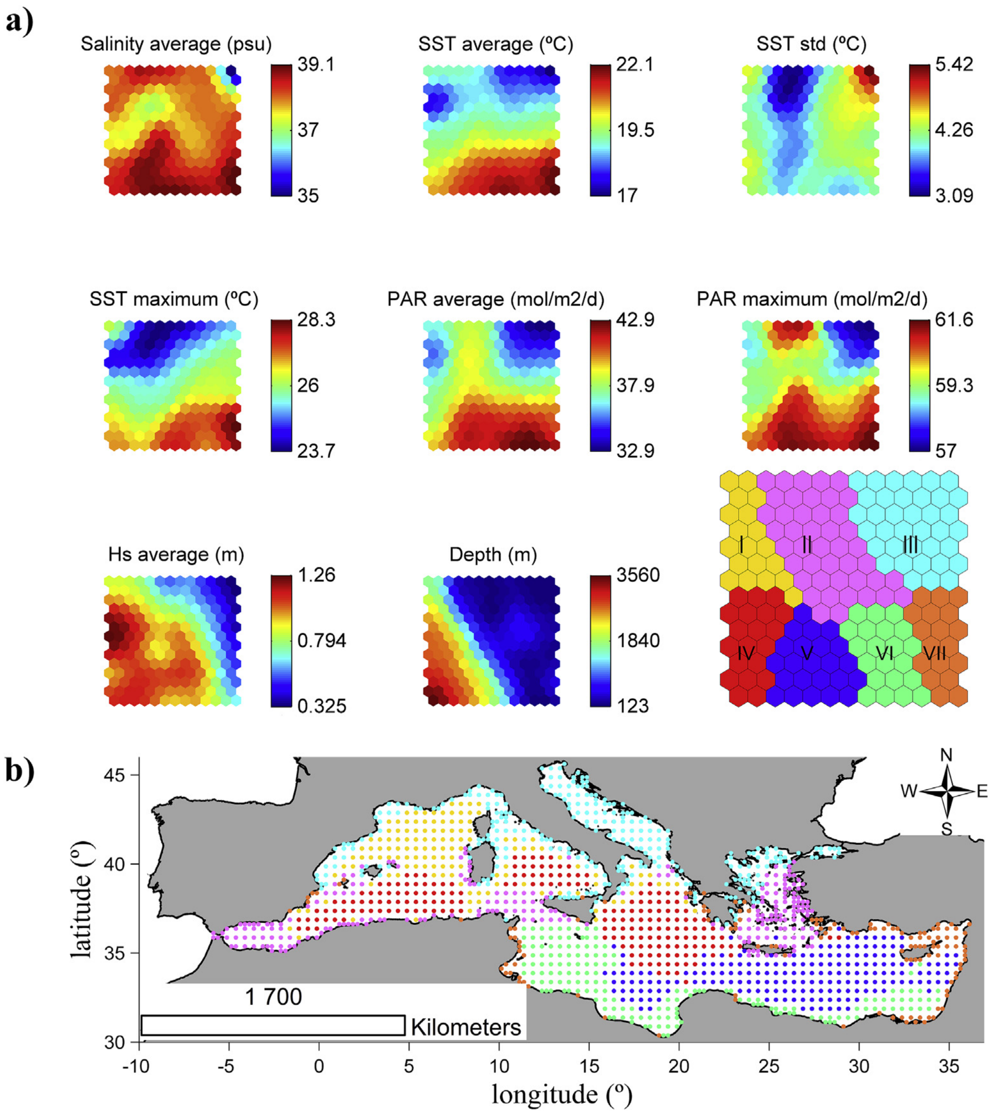


Fig. 2. (a) Gradient analysis of each abiotic parameter on the trained self-organizing map (SOM). Right down, visualization of the k-means results on the SOM plane. (b) Map of the obtained typologies in the Mediterranean Sea by the abiotic classification (based on SOM and k-means statistical analyses results). SST, sea surface temperature; PAR, photosynthetically active radiation; Hs, significant wave height; std, standard deviation.

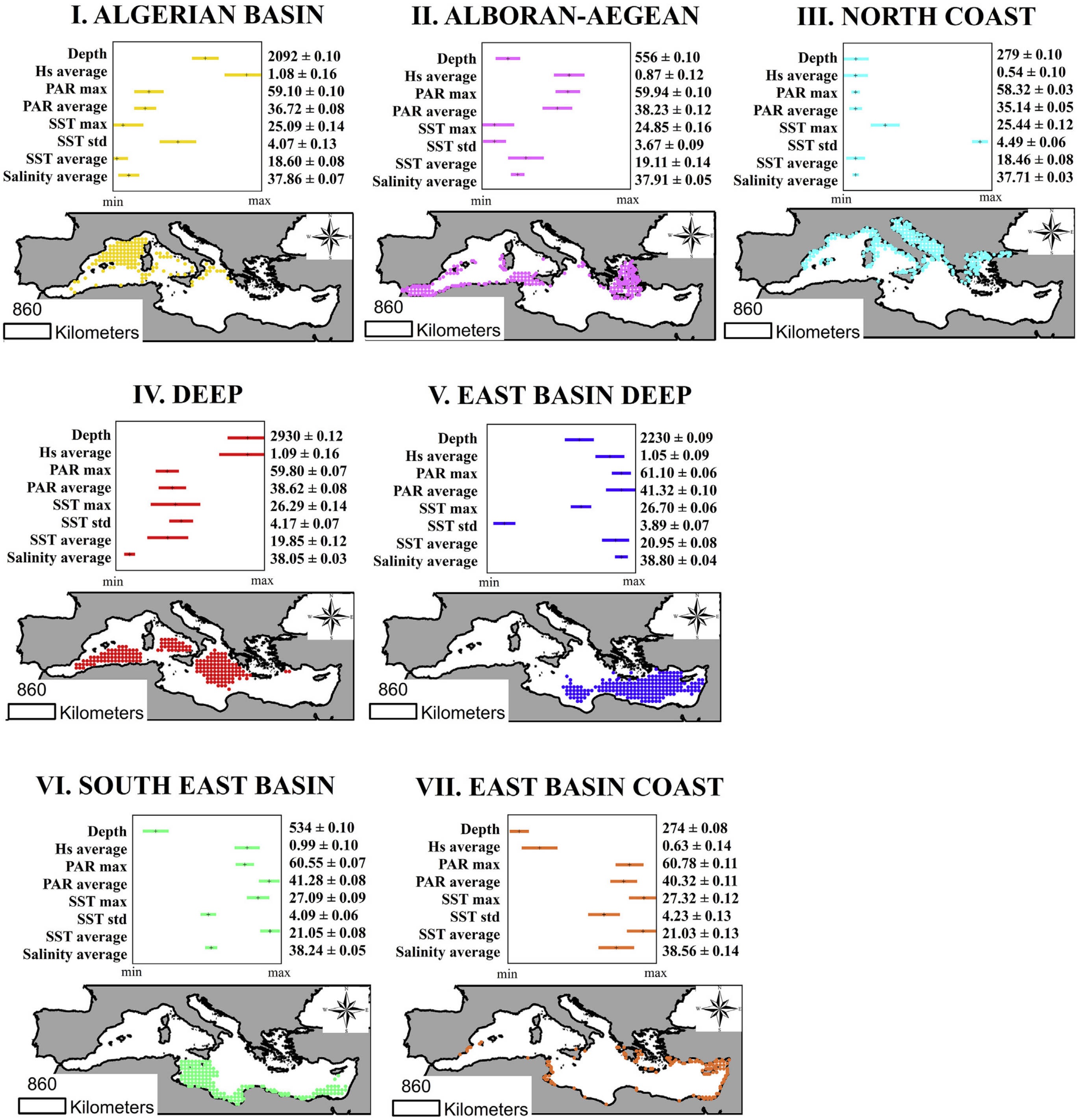


Fig. 3. The spatial distribution and box-plot graphs of the normalised values of the parameters for each typology. The mean value of each parameter is symbolized as a black cross, while the limits of the rectangle represent the standard deviation range. Hs, significant wave height; PAR, photosynthetically active radiation; SST, sea surface temperature; max, maximum; std, standard deviation.

- **Typology V, East Basin deep**: in this typology, the highest radiation and salinity with high SST stability stood out as a difference between typologies "South East Basin (VI)" and "East Basin coast (VII)." Another remarkable characteristic of this typology with regard to other similar typologies ("South East Basin (VI)" and "East Basin coast (VII)") was its great depth.
- **Typology VI, South East Basin**: the main difference from typologies "East Basin deep (V)" and "East Basin coast (VII)" was that it encompassed the shallowest points. The main similarity with them was the high values of radiation and temperature. This typology was characterized by being exposed.
- **Typology VII, East Basin coast**: this typology was the shallowest and the most sheltered within the East Basin. Salinity and temperature were also higher than in the typologies "East Basin deep (V)" and "South East Basin (VI)".

3.3. Biological suitability

Data on the presence of seagrasses in the Mediterranean Sea compiled from the sources described in the section “2.4 Biological suitability” are shown in Fig. 4 for each species separately.

The total and relative frequency values for each species over the SOM units and the k-means clusters are shown in Fig. 5, illustrating their preferences.

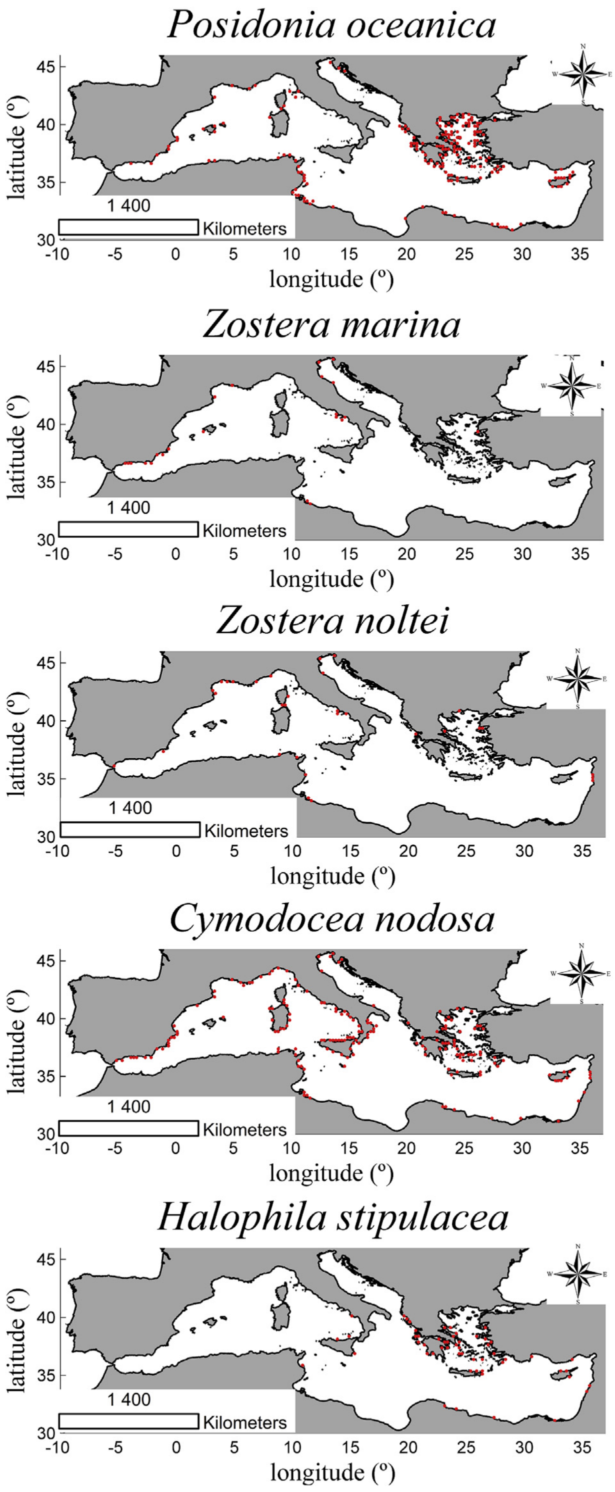


Fig. 4. The spatial distribution of the seagrasses considered along the Mediterranean Sea.

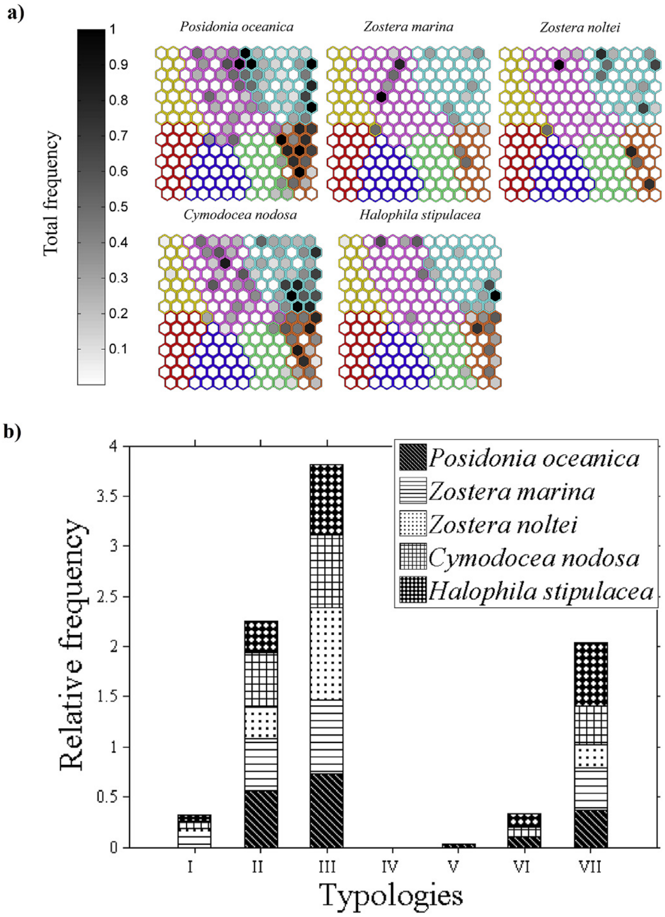


Fig. 5. (a) Visualization of the total frequency of mesh points cover of seagrasses considered (light = low frequency, dark = high frequency). (b) Bar graph for normalized relative frequency values of each species for each typology.

In general, the classification was in agreement with the ecological requirements of the considered seagrasses. The coastal typologies (“North Coast (III)”, “Alborán-Aegean (II)” and “East Basin coast (VII)”) had the highest frequencies according to the seagrass preference in shallow areas. The results showed the highest frequency of typology “North Coast (III)”, which had shallow points with colder waters and less radiation and salinity than the “East Basin coast (VII)” typology. It was also characterized by a shorter wave height than the “East Basin Coast (VII)”. The high frequency of the invasive seagrass *H. stipulacea* compared to the other species was remarkable for the “East Basin coast (VII)” typology. This typology was mainly confined to the East Basin with the highest values of the maximum SST and radiation. However, the frequency of the invasive species was lower than that of the other seagrasses, with the exception of *Z. noltei*, in the typology “Alborán-Aegean (II)”, characterized by cold waters and intermediate values of wave height, radiation and salinity. Regarding native species, *P. oceanica*, *Z. marina* and *C. nodosa* exhibited a higher preference for typologies II and VII than *Z. noltei*, while in the typology “North Coast (III)”, the trend was the opposite. The main difference between these typologies is that “North Coast (III)” presented higher SST standard deviation and lower salinity, radiation and significant wave height values. The typologies where seagrasses showed low frequency (“Algerian Basin (I)”, “Deep (IV)”, “East Basin deep (V)” and “South East Basin (VI)”) were characterized by highly significant wave height or depth values. Thus, regarding the seagrass frequency over the abiotic classification, the species distributions were in agreement with the defined typologies.

4. Discussion

The results presented in this work confirm the ability of the proposed methodology to classify the main regions in the Mediterranean Sea, taking into account its ecological relevance for seagrasses. This approach allows to the detection of areas linked by biological requirements, even if they are geographically separated.

4.1. Abiotic data

The quantification of variables and their parameters in a standardized fashion with the required resolution is a key step achieved with the use of remote sensing and numerical modeling datasets. These datasets provide continuous and uniform information throughout the study area (Devred et al., 2007; Gregr and Bodtker, 2006; Ramos et al., 2016c, 2012; Turner et al., 2003). Therefore, the suitability of the procedure used in this work to characterize the coastal and open waters of a broad area is demonstrated, showing similar patterns to the results obtained in other studies.

The general pattern of increasing the annual average salinity eastward with a mean value in the West Basin of 36.2‰ and of 39‰ in the East Basin, described by other studies (Coll et al., 2010; European Environment Agency (EEA), 2002; Lacoue-Labarthe et al., 2016), is also reflected in this work. In addition, the documented exceptions to this pattern, such as the freshwater influence of the Black Sea in the Aegean Sea and the entrance of less salty waters through the Strait of Gibraltar from the Atlantic Ocean, are captured in the present study (Coll et al., 2010; Georgiou et al., 2015; Poulos and Collins, 1997; Rodríguez, 1982). Furthermore, the results provide evidence of the ability of selected sources to detect local processes, as for the Po River flume on the north and west coast of the Adriatic Sea, where salinity is lower than in neighboring areas (Falcieri et al., 2013). In addition, the presence of saltier and colder waters in the Gulf of Lion is in agreement with the findings of other authors (Millot, 1999; Sarda et al., 2004). These conditions are due to the influence of the Mistral, a strong, dry and cold northwesterly wind that blows from southern France, favoring evaporation and the loss of heat.

Concerning the average SST values, the previously described pattern of cold water entering from the Atlantic and warming up towards the east is also in agreement with our results (Brasseur et al., 1996; Coll et al., 2010; Lacoue-Labarthe et al., 2016; Rodríguez, 1982). The approach described in this study is also able to detect the upwelling zones (the Aegean Sea and the north of the Adriatic Sea), where variability is high, as can be expected due to the characteristics of this phenomenon, which implies that colder water from the bottom moves towards the ocean surface (Marbà et al., 2015; Rodríguez, 1982). This work also detects the influence of the thermocline, which is less structured during summer than in winter. As a result, the exchanges that occur during the warmest months may result in severe local variations, as seen in the presence of patches in the maximum SST distribution (Marbà et al., 2015).

The incident radiation trend is in agreement with the latitudinal gradient north to south. Nevertheless, for the maximum PAR, the gradient is mainly longitudinal, with lower values in the West Basin and higher radiation in the East Basin. Although it has not been possible to find previous studies of PAR distribution in the Mediterranean Sea, these results can be explained by the work of Bat-Oyun et al. (2012), who established the influence of different factors on photosynthetically active radiation. Within these factors, a notable possibility is that the African dust during dry months in the area modifies the general pattern (Bergametti et al., 1992).

The overall patterns in wave energy are in agreement with the water depth, as they are strongly related to coastal points and enclosed seas (Adriatic Sea or Aegean Sea). As described by Cavaleri et al. (2004), the shelf causes less energetic waves, whereas in deeper areas, such as the Algerian and the Ionian basins, the bottom does not limit wave height.

Despite the close agreement of the characterization of the Mediterranean Sea with previous works, some questions require further development. It would be more precise to consider other key variables for species distribution, such as the substrate characteristics (Roff and Taylor, 2000; Templado González, 1995) or the use of bottom shear stress as a measure of the energy of the system (Tomás et al., 2012). Likewise, the parameters considered must be studied in detail to select those that best explain the biological behaviors of species. In this work, the average, maximum and minimum values are used, but it would be interesting to consider the use of parameters that are critical for certain species, such as the number of days in a year that some values are exceeded or the interannual variability of events (Jensen and Denny, 2015; Sánchez Lisazo, 1993; Unsworth et al., 2015). This approach will allow improvements in the power of using abiotic variables as surrogates for species distribution (Seabra et al., 2015), taking into account the met-ocean variables (Puente et al., 2017).

4.2. Classification

The results presented in this work provide evidence for the suitability of the methodological approach to obtain subdivisions for the entire Mediterranean Sea based on accurate and homogeneous abiotic data through the application of a combination of clustering techniques (SOM and k-means). This framework has been successfully applied by Ramos et al. (2016c) and Galván et al. (2016) for regions with different met-ocean conditions.

This combination of clustering techniques represents a considerable improvement compared with the classifications that consider only the k-means algorithm (Gabrie et al., 2012) but at the same time, the methodology is less complex than those presented by other authors (Nieblas et al., 2014; Reygondeau et al., 2017) while also providing good results and easier application.

The typologies obtained were able to depict the general patterns described in the Mediterranean Sea and to detect the main differences between areas. The water-mass exchange barrier existing in the Strait of Sicily (Bonanno et al., 2014; Giaccone and Sortino, 1974; Micheli et al., 2005; Rodríguez, 1982) is clearly distinguished in this work, which supports the ability of the proposed methodology to detect oceanographic differences among the obtained typologies. Therefore, typologies “Algerian Basin (I)”, “Alborán-Aegean (II)”, “North Coast (III)” and “Deep (IV)” can be recognized as representative of the West Basin because they present colder and less salty waters. In contrast, typologies “East Basin deep (V)”, “South East Basin (VI)” and “East Basin coast (VII)” depict the East Basin characteristics, with hotter and saltier waters and higher radiation. In addition to this distinction, the classification obtained has been proven to detect one of the most determinant factors in seagrass survival, depth as a surrogate of light availability. Through this differentiation, coastal areas mainly include the typologies “North Coast (III)”, “East Basin coast (VII)”, “Alborán-Aegean (II)” and “South East Basin (VI)”; meanwhile, the deepest areas are included in the typologies “Deep (IV)” and “East Basin deep (V).” In each of the aforementioned groups (West and East Basins), one of the typologies is characterized by encompassing sheltered points, the typology “North Coast (III)” in the West Basin and the typology “East Basin coast (VII)” in the East Basin. Existing studies have shown that wave energy is a determining factor in the spatial distribution of seagrass meadows (Infantes et al., 2011); thus, the differentiation of these areas from the exposed areas supports the robustness of the methodology used in this study. In addition, the typology "North Coast (III)" is linked to upwelling areas, as the high SST variation reflects. The salinity of this typology is the lowest, related to rivers with high flows in the northwest of the Adriatic Sea and the Gulf of Lion (Po: 1540 m^3/s and Rhone: 1710 m^3/s, respectively). In the Aegean Sea, the influence of the Black Sea water is observed in the salinity values (Falcieri et al., 2013).

The main typologies recognized in this study differed from the geographical limits established in previous classifications (Bianchi, 2007; Notarbartolo di Sciara and Agardy, 2010a; Spalding et al., 2007). These subdivisions were established based on geomorphological and biogeographic features, while we consider met-ocean variables. As a result, the proposed classification, despite not having as clear of boundaries as previous classifications, reflects possible links between areas geographically separated but ecologically related, such as the Alborán Sea and the Aegean Sea or the southern Algerian Basin and the Ionian Sea. Some of these differences are motivated by the inclusion of key variables for species not considered in similar studies (Gabrie et al., 2012; Nieblas et al., 2014; Reygondeau et al., 2017) such as significant wave height or photosynthetically active radiation, which clearly influence species distributions. In addition, the use of extreme and variable parameters considered in this work is remarkable due to their importance in the definition of species thresholds (Makino et al., 2015). This feature allows the characterization of the Mediterranean processes in a more reliable way by taking into consideration the relationships between geographically separated areas.

4.3. Biological suitability

In general, the analysis of taxa and their frequency over typologies confirm the biological relevance of the methodology with respect to seagrass distribution. The ability of the methodology to explain the seagrasses distribution and to distinguish between species requirements allows its use as a powerful tool to manage large regions.

P. oceanica is mainly present in the typologies "North Coast (III)" and "Alborán-Aegean (II)". It is also present in the typology "East Basin coast (VII)" but with a marked decrease in frequency. This finding is consistent with its ecology because they are coastal typologies (Luque and Templado González, 2004; Marín-Guirao et al., 2011; Short et al., 2007; Templado González, 1995) with a temperature range that is in agreement with *P. oceanica* requirements (Bay, 1984; Larkum et al., 2006). Concerning salinity, the typologies "Alborán-Aegean (II)" and "North Coast (III)" present an average value of 37.81‰, which is in agreement with the optimum of 38‰ described by Sánchez-Lizaso et al. (2008). Despite being a stenohaline species, *P. oceanica* is present in the typology "East Basin coast (VII)", where values are higher (average salinity of 38.56‰) but are under the threshold of damage and in the range established by other authors (Larkum et al., 2006; Telesca et al., 2015).

The frequency of occurrence of *Z. marina* in the typologies "North Coast (III)", "Alborán-Aegean (II)" and "East Basin coast (VII)" is very similar to that of *P. oceanica*, with a clear preference for coastal typologies, as it is located from the intertidal zone to a few meters deep (Green and Short, 2003), within the range of temperatures common in these typologies (Lee et al., 2007; Nejrup and Pedersen, 2008). The obtained classification is able to distinguish the preference of *Z. marina* for settlement locations that can exist in areas with a wider range of hydrodynamism than *P. oceanica*, forming vast meadows in sheltered areas and isolated patches in exposed ones (Ondiviela et al., 2005; Short et al., 2010a). In the Mediterranean Sea, *Z. marina* is considered a relict species, seemingly rare in the eastern part (Borum et al., 2004; Green and Short, 2003), which explains its presence in typologies with a salinity higher than its tolerance range of 5‰–35‰ (Short et al., 2010a).

In the Mediterranean, *Z. marina* co-occurs with *Z. noltei*, which is reflected in the classification obtained because both species are present in the same typologies but with some differences in frequency. *Z. noltei* shows a clear preference for the typology "North Coast (III)" because this species can occur from mesohaline to euhaline environments (Fernández-Torquemada and Sánchez-Lizaso, 2011; Luque and Templado González, 2004; Short et al., 2010b) and is adapted to colder waters than *P. oceanica* (Peralta, 2000).

The classification shows the overlap in *C. nodosa*, *H. stipulacea* and *P. oceanica* niches because they are present in the same typologies with some specificities. Meanwhile, *C. nodosa* usually forms mixed meadows with *P. oceanica* because their ecological requirements are very similar (Pérez and Romero, 1994). The origin of *H. stipulacea*, a lessepsian species, determines that it most commonly occurs in areas with characteristics similar to those of the Red Sea. *C. nodosa* is considered a pioneer species; thus, its ranges of tolerance are wider than in other species, as illustrated by the typologies where it is present (Borum et al., 2004; Chefaoui et al., 2016; Fernández-Torquemada and Sánchez-Lizaso, 2011; Olsen et al., 2012).

The alien seagrass *H. stipulacea* has the highest relative frequency in the typology "East Basin coast (VII)", with a primary presence in the East Basin because it entered the Mediterranean Sea through the Suez Canal. Therefore, the typological characteristics are in agreement with the ecology of *H. stipulacea*, whose optimal temperature is very similar to the mean value of the typology "East Basin coast (VII)" (21.5 °C vs 21.03 °C) (Gambi et al., 2009; Short et al., 2010c), and the species copes well with the high salinity of this typology (Lipkin, 1975). Despite this agreement with the typology "East Basin coast (VII)", due to its migrant nature, *H. stipulacea* tolerates a wide range of environmental conditions (Otero et al., 2013), which explains its presence in other typologies, especially in "North Coast (III)".

Although in this work seagrasses have been considered, the proposed methodology could be applied also to other flora or animal species by adjusting the input variables and parameters. For example, taking into account the spatial distribution of the shark *Cetorhinus maximus*, according to the databases of the Materials and methods section, it can be observed its preference for the "North Coast (III)" typology, which is in agreement with the behaviour of this fish, which moves to shallow depths between 0 and 20 m during part of the day to feed on zooplankton *Calanus helgolandicus* (Sims et al., 2005). Additionally, basking shark prefers cold waters with high variability (Cotton et al., 2005), which characterizes this typology.

Our results show a general agreement of the typologies obtained with the considered seagrass ecology, which constitutes a very interesting finding for marine areas management However, only presence data have been considered, meanwhile it will be necessary to have quantitative information and absence data to reach a complete validation. Besides the survey effort in the Mediterranean Sea to map the species presence is not the same in the two basins, and the understanding of their responses to environmental variables and their tolerance thresholds is far from complete (Marbà et al., 2015). Therefore, there remains insufficient knowledge on species encompassing their entire distributional range and their ecology and a complete biological validation of the classification is required.

5. Conclusions

In conclusion, the methodology applied in this work is able to classify the whole Mediterranean Sea in an unbiased manner into biologically relevant typologies, enabling the identification of areas geographically separated but important for species distribution. This approach supports the link between met-ocean variables and species distribution and the importance of including them in ecological modeling. Additionally, this study shows the suitability of the selected clustering methods for grouping abiotic data from reliable sources, which provides an objective procedure for the classification of large areas and its use as surrogates of species distributions.

According to our results, seven areas can be recognized in the Mediterranean Sea. The obtained typologies are able to recognize the oceanographic barrier in the Strait of Sicily and to cluster the results into two groups: one for typologies “South East Basin (VI)”, “East Basin deep (V)” and “East Basin coast (VII)” and the other for typologies “Algerian Basin (I)” and “North Coast (III)”, with differences clearly marked by temperature, salinity and radiation. Within this grouping, it is possible to distinguish coastal (typologies “Alborán-Aegean (II)”, “North Coast (III)”, “South East Basin (VI)” and “East Basin coast (VII)”) and deep areas (typologies “Algerian Basin (I)”, “Deep (IV)” and “East Basin deep (V)”). The influence of river discharges and the entrance of water from other seas are also reflected in typologies “North Coast (III)” and “Algerian Basin (I)”, respectively.

From a biological point of view, the typologies obtained are consistent with the distribution and ecological requirements of the considered seagrasses. Therefore, the link between surrogates and targets can be considered strong. As can be expected, seagrasses show a clear preference for coastal sites, with different frequencies according to their ecology.

The applications of this approach are broad because obtaining spatially and temporal homogeneous abiotic data is easier than a biological approach. Therefore, this approach provides a powerful tool for managers to attain ecosystem-based conservation of coastal and marine waters. Once the present classification and its relationship with species are known, it will be possible to develop hypotheses about changes in the distribution of Mediterranean Sea species due to global warming (Lacoue-Labarthe et al., 2016), which could support the sustainable management of marine biodiversity.

Although requiring further development, the results presented here show the ecological value of the classification obtained with the target species and, consequently, the ability of the methodology proposed to define typologies with biological implications in large areas using abiotic variables as surrogates of biological data. This approach constitutes a powerful tool for managing these zones.

Acknowledgements

This research is included in the framework of the CoCoNET project (Towards COast to COast NETworks of marine protected areas (from the shore to the high and deep sea), coupled with sea-based wind energy potential) and is funded by the National Plan for Research in Science and Technological Innovation from the Spanish Government 2013–2016 (Ministry of Economy and Competitiveness) (C3N project CTM2015-68244-R) and co-financed by the European Regional Development's funds (FEDER). Melisa Menéndez acknowledges financial support from the Ramon y Cajal Program (RYC-2014-16469). Camino F. de la Hoz acknowledges financial support in the form of a predoctoral grant (Ref: BES-2016-076434) from the Spanish Ministry of Economy, Industry and Competitiveness. We also thank the anonymous reviewers for their comments and suggestions.

References

Amante, C., Eakins, B.W., 2009. ETOPO1 1 Arc-minute Global Relief Model: Procedures, Data Sources and Analysis [WWW Document]. NOAA Tech. Memo. NESDIS NGDC-24. Natl. Geophys. Data Center, NOAA. http://dx.doi.org/10.7289/V5C8276M.

Azzellino, A., Gaspari, S., Airoldi, S., Nani, B., 2008. Habitat use and preferences of cetaceans along the continental slope and the adjacent pelagic waters in the western Ligurian Sea. Deep. Res. Part I Oceanogr. Res. Pap. 55, 296–323. http://dx.doi.org/10.1016/j.dsr.2007.11.006.

Bat-Oyun, T., Shinoda, M., Tsubo, M., 2012. Effects of cloud, atmospheric water vapor, and dust on photosynthetically active radiation and total solar radiation in a Mongolian grassland. J. Arid. Land 4, 349–356. http://dx.doi.org/10.3724/SP.J.1227.2012.00349.

Bay, D., 1984. A field study of the growth dynamics and productivity of Posidonia oceanica (L.) delile in Calvi Bay, Corsica. Aquat. Bot. 20, 43–64. http://dx.doi.org/10.1016/0304-3770(84)90026-3.

Belando, M.D., García Muñoz, R., Ramos Segura, A., Franco-Navarro, I., Bernardeau-esteller, J., García Moreno, P., Ruiz, J.M., 2015. Distribution and Abundance of Cymodocea Nodosa Meadows and Pinna Nobilis Populations in the Mar Menor Coastal Lagoon (Murcia, SE of Spain). PeerJ PrePrints.

Bergametti, G., Remoudaki, E., Losno, R., Steiner, E., Chatenet, B., Buat-Menard, P., 1992. Source, transport and deposition of atmospheric phosphorus over the Northwestern Mediterranean. J. Atmos. Chem. 14, 501–513.

Berline, L., Rammou, A.M., Doglioli, A., Molcard, A., Petrenko, A., 2014. A connectivity-based eco-regionalization method of the mediterranean sea. PLoS One 9, 3–11. http://dx.doi.org/10.1371/journal.pone.0111978.

Best, E.P., Buzzelli, C.P., Bartell, S.M., Wetzel, R.L., Boyd, W.A., Doyle, R.D., Campbell, K.R., 2001. Modeling submersed macrophyte growth in relation to underwater light climate: modeling approaches and application potential. Hydrobiologia 444, 43–70. http://dx.doi.org/10.1023/A:1017564632427.

Bianchi, C.N., 2007. Biodiversity issues for the forthcoming tropical Mediterranean Sea. Hydrobiologia 580, 7–21. http://dx.doi.org/10.1007/978-1-4020-6156-1.

Bonanno, A., Placenti, F., Basilone, G., Mifsud, R., Genovese, S., Patti, B., Bitetto, M. Di, Aronica, S., Barra, M., 2014. Variability of water mass properties in the Strait of Sicily in summer period of 1998 – 2013. Ocean. Sci. 10, 759–770. http://dx.doi.org/10.5194/os-10-759-2014.

Borum, J., Duarte, C.M., Krause-Jensen, D., Greve, T.M., 2004. European Seagrasses: an Introduction to Monitoring and Management. Management. https://doi.org/EVK3-CT-2000-00044.

Brasseur, P., Beckers, J.M., Brankart, J.M., Schoenauen, R., 1996. Seasonal temperature and salinity fields in the Mediterranean Sea: climatological analyses of a historical data set. Deep. Res. Part I Oceanogr. Res. Pap. 43, 159–192. http://dx.doi.org/10.1016/0967-0637(96)00012-X.

Camus, P., Mendez, F.J., Medina, R., Cofiño, A.S., 2011. Analysis of clustering and selection algorithms for the study of multivariate wave climate. Coast. Eng. 58, 453–462. http://dx.doi.org/10.1016/j.coastaleng.2011.02.003.

Cavaleri, L., Bertotti, L., Sclavo, M., Athanassoulis, G.A., Stefanakos, Ch.N., Gerostathis, T.P., Gaillard, P., Ravazzola, P., Kontolios, C., Ramieri, E., Dentone, L., Noel, C., Viala, C., Lefèvre, J.M., 2004. Wind and Wave Atlas of the Mediterranean Sea. Western European Armaments Organisation Research Cell.

Chefaoui, R.M., Assis, J., Duarte, C.M., Serrão, E.A., 2016. Large-scale prediction of seagrass distribution integrating landscape metrics and environmental factors: the case of Cymodocea nodosa (Mediterranean–Atlantic). Estuaries Coasts 39, 123–137. http://dx.doi.org/10.1007/s12237-015-9966-y.

COCONET, 2016. Towards COast to COast NETworks of Marine Protected Areas (From the Shore to the High and Deep Sea) Coupled with Sea-based Wind Energy Potential [WWW Document]. FP7 grant Agreem. no.287844. URL. http://coconetgis.ismar.cnr.it/. (Accessed 30 May 2015).

Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Lasram, F.B.R., Aguzzi, J., Ballesteros, E., Bianchi, C.N., Corbera, J., Dailianis, T., Danovaro, R., Estrada, M., Froglia, C., Galil, B.S., Gasol, J.M., Gertwage, R., Gil, J., Guilhaumon, F., Kesner-Reyes, K., Kitsos, M.S., Koukouras, A., Lampadariou, N., Laxamana, E., de la Cuadra, C.M.L.F., Lotze, H.K., Martin, D., Mouillot, D., Oro, D., Raicevich, S., Rius-Barile, J., Saiz-Salinas, J.I., Vicente, C.S., Somot, S., Templado, J., Turon, X., Vafidis, D., Villanueva, R., Voultsiadou, E., 2010. The biodiversity of the Mediterranean Sea: estimates, patterns, and threats. PLoS One 5. http://dx.doi.org/10.1371/journal.pone.0011842.

Cotton, P.A., Sims, D.W., Fanshawe, S., Chadwick, M., 2005. The effects of climate variability on zooplankton and basking shark (Cetorhinus maximus) relative abundance off southwest Britain. Fish. Oceanogr. 14, 151–155. http://dx.doi.org/10.1111/j.1365-2419.2005.00331.x.

Council of the European Commission, 1992. Council directive 92/43/EEC of 21 May 1992 on the Conservation of Natural Habitats and of Wild Fauna and Flora. Official Journal of the European Communities.

Davies, D.L., Bouldin, D.W., 1979. A Cluster Separation Measure. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1, pp. 224–227.

de Barbosa Araujo, C.C., Atkinson, P.M., Dearing, J.A., 2015. Remote sensing of ecosystem services: a systematic review. Ecol. Indic. 52, 430–443. http://dx.doi.org/10.1016/j.ecolind.2015.01.007.

Devred, E., Sathyendranath, S., Platt, T., 2007. Delineation of ecological provinces using ocean colour radiometry. Mar. Ecol. Prog. Ser. 346, 1–13. http://dx.doi.org/10.3354/meps07149.

Di Martino, V., Blundo, M.C., Tita, G., 2006. The mediterranean introduced seagrass

Halophila stipulacea in eastern sicily (Italy): temporal variations of the associated algal assemblage. Life Environ. 56, 223–230.

Duarte, C.M., Losada, I.J., Hendriks, I.E., Mazarrasa, I., Marbà, N., 2013. The role of coastal plant communities for climate change mitigation and adaptation. Nat. Clim. Chang. 3, 961–968. http://dx.doi.org/10.1038/nclimate1970.

European Commission, 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 Establishing a Framework for Community Action in the Field of Marine Environmental Policy (Marine Strategy Framework Directive). Official Journal of the European Communities.

European Environment Agency (EEA), 2002. Europe's Biodiversity — Biogeographical Regions and Seas, EEA Report No 1/2002. Copenhagen.

Falcieri, F.M., Benetazzo, A., Sclavo, M., Russo, A., Carniel, S., 2013. Po River plume pattern variability investigated from model data. Cont. Shelf Res. 1–12. http://dx.doi.org/10.1016/j.csr.2013.11.001.

Fernández-Torquemada, Y., Sánchez-Lizaso, J.L., 2011. Responses of two Mediterranean seagrasses to experimental changes in salinity. Hydrobiologia 669, 21–33. http://dx.doi.org/10.1007/s10750-011-0644-1.

Francour, P., 1997. Fish assemblages of Posidonia oceanica beds at Port Cros (France, NW Mediterranean): assessment of composition and long-term fluctuations by visual census. Mar. Ecol. Della Stn. Zool. Di Napoli I 18, 157–173. http://dx.doi.org/10.1111/j.1439-0485.1997.tb00434.x.

Fraschetti, S., Terlizzi, A., Boero, F., 2008. How many habitats are there in the sea (and where)? J. Exp. Mar. Bio. Ecol. 366, 109–115. http://dx.doi.org/10.1016/j.jembe.2008.07.015.

Gabrie, C., Lagabrielle, E., Bissery, C., Crochelet, E., Meola, B., Webster, C., Claudet, J., Chassanite, A., Marinesque, S., Robert, P., Goutx, M., 2012. The Status of Marine Protected Areas in the Mediterranean Sea.

Galván, C., Juanes, J.A., Puente, A., 2010. Ecological classification of European transitional waters in the North-East Atlantic eco-region. Estuar. Coast. Shelf Sci. 87, 442–450. http://dx.doi.org/10.1016/j.ecss.2010.01.026.

Galván, C., Puente, A., Castanedo, S., Juanes, J.A., 2016. Average vs. extreme salinity conditions: do they equally affect the distribution of macroinvertebrates in estuarine environments? Limnol. Oceanogr. 61, 984–1000. http://dx.doi.org/10.1002/lno.10267.

Gambi, M.C., Barbieri, F., Bianchi, C.N., 2009. New record of the alien seagrass Halophila stipulacea (Hydrocharitaceae) in the western Mediterranean: a further clue to changing Mediterranean Sea biogeography. Mar. Biodivers. Rec. 2, e84. http://dx.doi.org/10.1017/S175526720900058X.

GBIF, 2013. GBIF Backbone Taxonomy, 2013-07-01 [WWW Document]. URL. http://www.gbif.org/species/5328663. (Accessed 23 June 2015). on 2016-01-27.

Georgiou, S., Mantziafou, A., Sofianos, S., Gertman, I., Özsoy, E., Somot, S., Vervatis, V., 2015. Climate variability and deep water mass characteristics in the Aegean Sea. Atmos. Res. 152, 146–158. http://dx.doi.org/10.1016/j.atmosres.2014.07.023.

Gevrey, M., Dimopoulos, I., Lek, S., 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol. Modell. 160, 249–264. http://dx.doi.org/10.1016/S0304-3800(02)00257-0.

Giaccone, G., Sortino, M., 1974. Significato biogeografico della vegetazione marina della Sicilia e delle isole minori nell'area del Mare Mediterraneo. Boll. Stud. Inf. del Giard. Colon. Palermo 26, 130–146.

Goldenberg Vilar, A., Van Dam, H., Vonk, J.A., Van Der Geest, H.G., Admiraal, W., 2014. Typology of diatom communities in the Dutch delta: recognizing patterns of environmental drivers in nutrient rich ditches. Ecol. Indic. 45, 561–569. http://dx.doi.org/10.1016/j.ecolind.2014.05.016.

Green, E.P., Short, F.T., 2003. World Atlas of Seagrasses. UNEP-WCMC by the University of California Press. UNEP World Conservation Monitoring Centre, Berkeley, USA. http://dx.doi.org/10.1515/BOT.2004.029.

Gregr, E.J., Ahrens, A.L., Ian Perry, R., 2012. Reconciling classifications of ecologically and biologically significant areas in the world's oceans. Mar. Policy 36, 716–726. http://dx.doi.org/10.1016/j.marpol.2011.10.009.

Gregr, E.J., Bodtker, K.M., 2006. Adaptive Classification of Marine Ecosystems: Identifying Biologically Meaningful Regions in the Marine Environment. Fisheries Bethesda, Vancouver, Canada. http://dx.doi.org/10.1139/xxxx.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means clustering algorithm. Appl. Stat. 28, 100. http://dx.doi.org/10.2307/2346830.

Holcer, D., Lazar, B., Mackelworth, P., Fortuna, C.M., 2013. Rare or just unknown? The occurrence of the giant devil ray (Mobula mobular) in the Adriatic Sea. J. Appl. Ichthyol. 29, 139–144. http://dx.doi.org/10.1111/jai.12034.

Howell, K.L., 2010. A benthic classification system to aid in the implementation of marine protected area networks in the deep/high seas of the NE Atlantic. Biol. Conserv. 143, 1041–1056. http://dx.doi.org/10.1016/j.biocon.2010.02.001.

Huang, H., Bin, C., Jinlan, L., 2015. The marine spatial classification and the identification of priority conservation areas (PCAs) for marine biodiversity conservation – a case study of the offshore China. Ocean. Coast. Manag. 116, 224–236. http://dx.doi.org/10.1016/j.ocecoaman.2015.06.027.

Infantes, E., Orfila, A., Bouma, T.J., Simarro, G., Terrados, J., 2011. Posidonia oceanica and Cymodocea nodosa seedling tolerance to wave exposure. Limnol. Oceanogr. 56, 2223–2232. http://dx.doi.org/10.4319/lo.2011.56.6.2223.

IUCN, 2015. The IUCN Red List of Threatened Species [WWW Document]. URL, Version 2015-3. http://www.iucnredlist.org/. (Accessed 16 September 2015).

Jensen, M.M., Denny, M.W., 2015. Experimental determination of the hydrodynamic forces responsible for wave impact events. J. Exp. Mar. Bio. Ecol. 469, 123–130. http://dx.doi.org/10.1016/j.jembe.2015.04.013.

Juanes, J.A., Puente, A., Ramos, E., 2016. A global approach to hierarchical classification of coastal waters at different spatial scales: the NEA case. J. Mar. Biol. Assoc. U. K. 1–12. http://dx.doi.org/10.1017/S0025315416000801.

Kiviluoto, K., 1996. Topology preservation in self organizing maps. In: Proceedings of the International Conference on Neural Networks (ICNN'96), vol. 1. IEEE Neural Networks Council, Piscataway, New Jersey, USA, pp. 294–299.

Kleisner, K.M., Coll, M., Lynam, C.P., Bundy, A., Shannon, L., Shin, Y.-J., Boldt, J.L., Maria, F.B., Diallo, I., Fox, C., Gascuel, D., Heymans, J.J., Juan Jordá, M.J., Jouffre, D., Large, S.I., Marshall, K.N., Ojaveer, H., Piroddi, C., Tam, J., Torres, M.A., Travers-Trolet, M., Tsagarakis, K., van der Meeren, G.I., Zador, S., 2015. Evaluating changes in marine communities that provide ecosystem services through comparative assessments of community indicators. Ecosyst. Serv. 1–17. http://dx.doi.org/10.1016/j.ecoser.2015.02.002.

Kohonen, T., 2001. Self-organizing Maps. Berlin.

Kohonen, T., 1998. The self-organizing map. Neurocomputing 21, 1–6. http://dx.doi.org/10.1016/S0925-2312(98)00030-7.

Lacoue-Labarthe, T., Nunes, P.A.L.D., Ziveri, P., Cinar, M., Gazeau, F., Hall-Spencer, J.M., Hilmi, N., Moschella, P., Safa, A., Sauzade, D., Turley, C., 2016. Impacts of ocean acidification in a warming Mediterranean Sea: an overview. Reg. Stud. Mar. Sci. 5, 1–11. http://dx.doi.org/10.1016/j.rsma.2015.12.005.

Larkum, A., Orth, R.J., Duarte, C.M., 2006. Seagrasses: Biology, Ecology and Conservation. Springer, Dordrecht (The Netherlands).

Last, P.R., Lyne, V.D., Williams, A., Davies, C.R., Butler, A.J., Yearsley, G.K., 2010. A hierarchical framework for classifying seabed biodiversity with application to planning and managing Australia's marine biological resources. Biol. Conserv. 143, 1675–1686. http://dx.doi.org/10.1016/j.biocon.2010.04.008.

Lee, K.S., Park, S.R., Kim, Y.K., 2007. Effects of irradiance, temperature, and nutrients on growth dynamics of seagrasses: a review. J. Exp. Mar. Bio. Ecol. 350, 144–175. http://dx.doi.org/10.1016/j.jembe.2007.06.016.

Levitus, S., Antonov, J.I., Baranova, O.K., Boyer, T.P., Coleman, C.L., Garcia, H.E., Grodsky, A.I., Johnson, D.R., Locarnini, R.A., Mishonov, A.V., Reagan, J.R., Sazama, C.L., Seidov, D., Smolyar, I., Yarosh, E.S., Zweng, M.M., 2013. The World Ocean database. Data Sci. J. 12, WDS229–WDS234. http://dx.doi.org/10.2481/dsj.WDS-041.

Lipkin, Y., 1975. Halophila stipulacea, a review of a successful immigration. Aquat. Bot. 1, 203–215. http://dx.doi.org/10.1016/0304-3770(75)90023-6.

Luque, Á.A., Templado González, J., 2004. Praderas y bosques marinos de Andalucía. Consejería de Medio Ambiente, Junta de Andalucía, Sevilla.

Makino, A., Klein, C.J., Possingham, H.P., Yamano, H., Yara, Y., Ariga, T., Matsuhasi, K., Beger, M., 2015. The effect of applying alternate IPCC climate scenarios to Marine Reserve design for range changing species. Conserv. Lett. 8, 320–328. http://dx.doi.org/10.1111/conl.12147.

Marbà, N., Jordà, G., Agustí, S., Girard, C., Duarte, C.M., 2015. Footprints of climate change on Mediterranean Sea biota. Front. Mar. Sci. 2, 1–11. http://dx.doi.org/10.3389/fmars.2015.00056.

Marbà, N., Krause-Jensen, D., Alcoverro, T., Birk, S., Pedersen, A., Neto, J.M., Orfanidis, S., Garmendia, J.M., Muxika, I., Borja, A., Dencheva, K., Duarte, C.M., 2013. Diversity of European seagrass indicators: patterns within and across regions. Hydrobiologia 704, 265–278. http://dx.doi.org/10.1007/s10750-012-1403-7.

Marín-Guirao, L., Sandoval-Gil, J.M., Ruíz, J.M., Sánchez-Lizaso, J.L., 2011. Photosynthesis, growth and survival of the Mediterranean seagrass Posidonia oceanica in response to simulated salinity increases in a laboratory mesocosm system. Estuar. Coast. Shelf Sci. 92, 286–296. http://dx.doi.org/10.1016/j.ecss.2011.01.003.

Micheli, C., Paganin, P., Peirano, A., Caye, G., Meinesz, A., Bianchi, C.N., 2005. Genetic variability of Posidonia oceanica (L.) Delile in relation to local factors and biogeographic patterns. Aquat. Bot. 82, 210–221. http://dx.doi.org/10.1016/j.aquabot.2005.03.002.

Millot, C., 1999. Circulation in the western mediterranean sea. J. Mar. Syst. 20, 423–442.

Muller-Karger, F., Andréfouët, S., Varela, R., Thunell, R., 2005. The Color of the Coastal Ocean and Applications in the Solution of Research and Management Problems, in: Remote Sensing of Coastal Aquatic Environments - Technologies, Tenchniques and Applications. Springer, Dordrecht (The Netherlands), pp. 101–127.

Nejrup, L.B., Pedersen, M.F., 2008. Effects of salinity and water temperature on the ecological performance of Zostera marina. Aquat. Bot. 88, 239–246. http://dx.doi.org/10.1016/j.aquabot.2007.10.006.

Nieblas, A.E., Drushka, K., Reygondeau, G., Rossi, V., Demarcq, H., Dubroca, L., Bonhommeau, S., Madron, X.D. de, Guieu, C., Sempéré, R., Conan, P., Cossa, D., Fraschetti, S., Micheli, F., Levin, N., Giakoumi, S., Katsanevakis, S., Abdulla, A., Portman, M., Notarbartolo-di-Sciara, G., Agardy, T., Katsanevakis, S., Possingham, H., Abdulla, A., Gomei, M., Hyrenbach, D., Notarbartolo-di-Sciara, G., Agardy, T., Bianchi, C., Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Lasram, F., D'Ortenzio, F., D'Alcalà, M.R., Lejeusne, C., Chevaldonne, P., Pergent-Martini, C., Boudouresque, C., Perez, T., Halpern, B., Walbridge, S., Selkoe, K., Kappel, C., Micheli, F., Spalding, M., Fox, H., Allen, G., Davidson, N., Ferdaña, Z., Rice, J., Gjerde, K., Ardron, J., Arico, S., Cresswell, I., Coll, M., Cury, P., Azzurro, E., Bariche, M., Bayadas, G., Edgar, G., Moverley, J., Barrett, N., Peters, D., Reed, C., Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., Kavanaugh, M., Hales, B., Saraceno, M., Spitz, Y., White, A., Giakoumi, S., Sini, M., Gerovasileiou, V., Mazor, T., Beher, J., Millot, C., Millot, C., Robinson, A., Malanotte-Rizzoli, P., Hecht, A., Michelato, A., Roether, W., Sherman, K., Sissenwine, M., Christensen, V., Duda, A., Hempel, G., Lehahn, Y., D'Ovidio, F., Levy, M., Heifetz, E., Rossi, V., Lopez, C., Sudre, J., Hernandez-Garcia, E., Garcon, V., Rossi, V., Ser-Giacomi, E., López, C., Hernández-García, E., Reygondeau, G., Irisson, J.O., Guieu, C., Gasparini, S., Ayata, S., Sydeman, W., Thompson, S., Garcia-Reyes, M., Kahru, M., Peterson, W., Canny, J., Nieto, K., Demarcq, H., McClatchie, S., D'Ovidio, F., Fernández, V., Hernández-García, E., López, C., Tew-Kai, E., Rossi, V., Sudre, J., Weimerskirch, H., Lopez, C., Okubo, A., Weiss, J., Henson, S., Thomas, A., Hartigan, J., Wong, M., Oliver, M., Glenn, S., Kohut, J., Irwin, A., Schofield, O., Beaugrand, G., Ibanez, F., Josey, S., Somot, S., Tsimplis, M., Papadopoulos, V., Josey, S., Bartzokas, A., Somot, S., Ruiz, S., Fernández, V., Dietrich, D., Haney, R., Tintoré, J., Larnicol, G., Ayoub, N., Traon, P. Le, Malanotte-Rizzoli, P., Manca, B., D'Alcalà, M., Theocharis, A., Bergamasco, A., Pinardi, N., Masetti, E., Longhurst, A., Sathyendranath, S., Platt, T., Caverhill, C., Chase, J., Leibold, M., Zavatarelli, M., Mellor, G., Hernández-Carrasco, I., López, C., Hernández-García, E., Turiel, A., D'Ovidio, F., Isern-Fontanet, J., López, C., Hernández-García, E., García-Ladona, E., Cayula, J., Cornillon, P., Hyrenbach, K., Forney, K., Dayton, P., 2014. Defining mediterranean and black sea biogeochemical subprovinces and synthetic ocean indicators using mesoscale oceanographic features. PLoS One 9, e111251. http://dx.doi.org/10.1371/journal.pone.0111251.

Notarbartolo di Sciara, G., Agardy, T., 2010a. The Mediterranean Sea Biodiversity: State of the Ecosystems, Pressures, Impacts and Future Priorities.

Notarbartolo di Sciara, G., Agardy, T., 2010b. Overview of Scientific Findings and Criteria Relevant to Identifying SPAMIs in the Mediterranean Opean Seas, Including the Deep Sea. http://dx.doi.org/10.1002/9780470670590.wbeog931.

OB.DAAC, 2013. Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Ocean Color Data [WWW Document]. URL. http://oceancolor.gsfc.nasa.gov/cms/dataaccess. (Accessed 15 December 2013).

OBIS, 2015. The OBIS Web Portal Search Interface [WWW Document]. URL. http://iobis.org/mapper/. (Accessed 23 June 2015).

Olsen, Y.S., Sánchez-Camacho, M., Marbà, N., Duarte, C.M., 2012. Mediterranean seagrass growth and demography responses to experimental warming. Estuaries Coasts 35, 1205–1213. http://dx.doi.org/10.1007/s12237-012-9521-z.

Ondiviela, B., Alcántara, M.A., Fernández, L., 2005. Contribución al conocimiento de las comunidades de Zostera spp. del Cantábrico: la bahía de Santander y las marismas de Santoña. Locustella, pp. 55–67.

Ondiviela, B., Losada, I.J., Lara, J.L., Maza, M., Galván, C., Bouma, T.J., van Belzen, J., 2014. The role of seagrasses in coastal protection in a changing climate. Coast. Eng. 87, 158–168. http://dx.doi.org/10.1016/j.coastaleng.2013.11.005.

Otero, M., Cebrian, E., Francour, P., Galil, B., Savini, D., 2013. Monitoring Marine Invasive Species in Mediterranean Marine Protected Areas (MPAs): a Strategy and Practical Guide for Managers. IUCN, Malaga, Spain, p. 136.

Panigada, S., Zanardelli, M., MacKenzie, M., Donovan, C., Mélin, F., Hammond, P.S., 2008. Modelling habitat preferences for fin whales and striped dolphins in the Pelagos Sanctuary (Western Mediterranean Sea) with physiographic and remote sensing variables. Remote Sens. Environ. 112, 3400–3412. http://dx.doi.org/10.1016/j.rse.2007.11.017.

Peralta, G., 2000. Estudios sobre el crecimiento en Zostera noltii Hornem.: Dinámica estacional y aspectos ecofisiológicos. University of Cádiz.

Perez, J., Menendez, M., Losada, I.J., 2017. GOW2: a global wave hindcast for coastal applications. Coast. Eng. 124, 1–11. http://dx.doi.org/10.1016/j.coastaleng.2017.03.005.

Pérez, M., Romero, J., 1994. Growth dynamics, production, and nutrient status of the seagrass Cymodocea nodosa in a mediterranean semi-estuarine environment. Mar. Ecol. 15, 51–64. http://dx.doi.org/10.1111/j.1439-0485.1994.tb00041.x.

Poulos, S.E., Collins, M.B., 1997. Seasonal variability in sea surface oceanographic conditions in the Aegean Sea (Eastern Mediterranean): an overview. J. Mar. Syst. 13, 225–244.

Puente, A., Guinda, X., Juanes, J.A., Ramos, E., Echavarri-Erasun, B., de la Hoz, C.F., Degraer, S., Kerckhof, F., Rousou, M., Orav-Kotta, H., Kotta, J., Jourde, J., Pedrotti, M.L., Leclerc, J.-C., Simon, N., Bachelet, G., Arvanitidis, C., Pavloudi, C., Faulwetter, S., Crowe, T., Coughlan, J., Benedetti-Cecchi, L., Dal Bello, M., Magni, P., Como, S., Coppa, S., De Lucia, G., Rugins, T., Jankowska, E., Weslawski, J.M., Warzocha, J., Silva, T., Ribeiro, P., Matos, V., Sousa-Pinto, I., Troncoso, J.S., Peleg, O., Rilov, G., Espinosa, F., Péres Ruzafa, Á., Frost, M., Hummel, H., van Avesaath, P., 2017. The role of physical variables in biodiversity patterns of intertidal macroalgae along European coasts. J. Mar. Biol. Assoc. U. K. 97, 549–560. http://dx.doi.org/10.1017/S0025315416001673.

Ramos, E., Díaz de Terán, J.R., Puente, A., Juanes, J.A., 2016a. The role of geomorphology in the distribution of intertidal rocky macroalgae in the NE Atlantic region. Estuar. Coast. Shelf Sci. 179, 90–98. http://dx.doi.org/10.1016/j.ecss.2015.10.007.

Ramos, E., Juanes, J.A., Galván, C., Neto, J.M., Melo, R., Pedersen, A., Scanlan, C., Wilkes, R., van den Bergh, E., Blomqvist, M., Karup, H.P., Heiber, W., Reitsma, J.M., Ximenes, M.C., Silió, A., Méndez, F., González, B., 2012. Coastal waters classification based on physical attributes along the NE Atlantic region. An approach for rocky macroalgae potential distribution. Estuar. Coast. Shelf Sci. 112, 105–114. http://dx.doi.org/10.1016/j.ecss.2011.11.041.

Ramos, E., Puente, A., Guinda, X., Juanes, J.A., 2016b. A hierarchical classification system along the NE Atlantic coast: focusing on the local scale (Cantabria, N Spain). Eur. J. Phycol. http://dx.doi.org/10.3109/14756366.2014.1003216.

Ramos, E., Puente, A., Juanes, J.A., 2016c. An ecological classification of rocky shores at a regional scale: a predictive tool for management of conservation values. Mar. Ecol. 37, 311–328. http://dx.doi.org/10.1111/maec.12280.

Ramos, E., Puente, A., Juanes, J.A., Neto, J.M., Pedersen, A., Bartsch, I., Scanlan, C., Wilkes, R., Van den Bergh, E., Ar Gall, E., Melo, R., 2014. Biological validation of physical coastal waters classification along the NE Atlantic region based on rocky macroalgae distribution. Estuar. Coast. Shelf Sci. 147, 103–112. http://dx.doi.org/10.1016/j.ecss.2014.05.036.

Reguero, B.G., Menéndez, M., Méndez, F.J., Mínguez, R., Losada, I.J., 2012. A Global Ocean Wave (GOW) calibrated reanalysis from 1948 onwards. Coast. Eng. 65, 38–55. http://dx.doi.org/10.1016/j.coastaleng.2012.03.003.

Reygondeau, G., Guieu, C., Benedetti, F., Irisson, J.O., Ayata, S.D., Gasparini, S., Koubbi, P., 2017. Biogeochemical regions of the Mediterranean Sea: an objective multidimensional and multivariate environmental approach. Prog. Oceanogr. 151, 138–148. http://dx.doi.org/10.1016/j.pocean.2016.11.001.

Riis, T., Olesen, B., Clayton, J.S., Lambertini, C., Brix, H., Sorrell, B.K., 2012. Growth and morphology in relation to temperature and light availability during the establishment of three invasive aquatic plant species. Aquat. Bot. 102, 56–64. http://dx.doi.org/10.1016/j.aquabot.2012.05.002.

Roberts-Jones, J., Fiedler, E.K., Martin, M.J., 2012. Daily, global, high-resolution SST and sea ice reanalysis for 1985-2007 using the OSTIA system. J. Clim. 25, 6215–6232. http://dx.doi.org/10.1175/JCLI-D-11-00648.1.

Roca, G., Alcoverro, T., Krause-Jensen, D., Balsby, T.J.S., Van Katwijk, M.M., Marbà, N., Santos, R., Arthur, R., Mascaró, O., Fernández-Torquemada, Y., Pérez, M., Duarte, C.M., Romero, J., 2016. Response of seagrass indicators to shifts in environmental stressors: a global review and management synthesis. Ecol. Indic. 63, 310–323. http://dx.doi.org/10.1016/j.ecolind.2015.12.007.

Rodríguez-Rodríguez, D., Rodríguez, J., Abdul Malak, D., Nastasi, A., Hernández, P., 2016. Marine protected areas and fisheries restricted areas in the Mediterranean: assessing "actual" marine biodiversity protection coverage at multiple scales. Mar. Policy 64, 24–30. http://dx.doi.org/10.1016/j.marpol.2015.11.006.

Rodríguez, J., 1982. Oceanografía del Mar Mediterráneo. Pirámide.

Roff, J.C., Taylor, M.E., 2000. National frameworks for marine conservation - a hierarchical geophysical approach. Aquat. Conserv. Mar. Freshw. Ecosyst. 10, 209–223.

Rossi, V., Ser-Giacomi, E., López, C., Hernández-García, E., 2014. Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves. Geophys. Res. Lett. 41, 2883–2891. http://dx.doi.org/10.1002/2014GL059540.

Sánchez-Lizaso, J.L., Romero, J., Ruiz, J., Gacia, E., Buceta, J.L., Invers, O., Fernández Torquemada, Y., Mas, J., Ruiz-Mateo, A., Manzanera, M., 2008. Salinity tolerance of the Mediterranean seagrass Posidonia oceanica: recommendations to minimize the impact of brine discharges from desalination plants. Desalination 221, 602–607. http://dx.doi.org/10.1016/j.desal.2007.01.119.

Sánchez Lisazo, J.L., 1993. Estudio de la pradera de Posidonia oceanica (L.) Delile de la reserva marina de Tabarca (Alicante): Fenología y producción primaria. Dep. Ciències Ambient. i Recur. Nat. 121.

Sarda, F., Calafat, A., Flexas, M.M., Tselepides, A., Canals, M., Espino, M., Tursi, A., 2004. An introduction to Mediterranean deep-sea biology. Sci. Mar. 68, 7–38. http://dx.doi.org/10.3989/scimar.2004.68s37.

Seabra, R., Wethey, D.S., Lima, F.P., Campo, R., 2015. Understanding complex biogeographic responses to climate change. Sci. Rep. 12930 http://dx.doi.org/10.1038/srep12930.

Short, F.T., Carruthers, T., Dennison, W., Waycott, M., 2007. Global seagrass distribution and diversity: a bioregional model. J. Exp. Mar. Bio. Ecol. 350, 3–20. http://dx.doi.org/10.1016/j.jembe.2007.06.012.

Short, F.T., Carruthers, T.J.R., Waycott, M., Kendrick, G.A., Fourqurean, J.W., Callabine, A., Kenworthy, W.J., Dennison, W.C., 2010a. Zostera Marina. http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T153538A4516675.en [WWW Document]. IUCN Red List Threat. Species. URL. . (Accessed 10 November 2015).

Short, F.T., Carruthers, T.J.R., Waycott, M., Kendrick, G.A., Fourqurean, J.W., Callabine, A., Kenworthy, W.J., Dennison, W.C., 2010b. Zostera Noltii. http://dx.doi.org/10.2305/IUCN.UK.2010-3.RLTS.T173361A6999224.en [WWW Document]. IUCN Red List Threat. Species. URL. . (Accessed 10 November 2015).

Short, F.T., Carruthers, T.J.R., Waycott, M., Kendrick, G.A., Fourqurean, J.W., Callabine, A., Kenworthy, W.J., Dennison, W.C., 2010c. Halophila Stipulacea [WWW Document]. IUCN Red List Threat. Species.

Sims, D.W., Southall, E.J., Richardson, A.J., Reid, P.C., Metcalfe, J.D., 2003. Foraging and migratory behaviour of basking sharks over seasonal scales: no evidence for winter hibernation. Mar. Ecol. Prog. Ser. 248, 187–196. http://dx.doi.org/10.3354/meps248187.

Sims, D.W., Southall, E.J., Tarling, G.A., Metcalfe, J.D., 2005. Habitat-specific normal and reverse diel vertical migration in the plankton-feeding basking shark. J. Anim. Ecol. 74, 755–761. http://dx.doi.org/10.1111/j.1365-2656.2005.00971.x.

Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lombana, A., Lourie, S.A., Martin, K.D., Mcmanus, E., Molnar, J., Recchia, C.A., Robertson, J., 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. Bioscience 57, 573–583. http://dx.doi.org/10.1641/B570707.

Stojković, M., Milošević, D., Simić, S., Simić, V., 2014. Using a fish-based model to assess the ecological status of lotic systems in Serbia. Water Resour. Manag. 28, 4615–4629. http://dx.doi.org/10.1007/s11269-014-0762-4.

Telesca, L., Belluscio, A., Criscoli, A., Ardizzone, G., Apostolaki, E.T., Fraschetti, S., Gristina, M., Knittweis, L., Martin, C.S., Pergent, G., Alagna, A., Badalamenti, F., Garofalo, G., Gerakaris, V., Louise Pace, M., Pergent-Martini, C., Salomidi, M., 2015. Seagrass meadows (Posidonia oceanica) distribution and trajectories of change. Sci. Rep. 5, 1–14. http://dx.doi.org/10.1038/srep12505.

Templado González, J., 1995. Las comunidades de fanerógamas marinas en el Mediterráneo Occidental. In: La Gestión de Los Espacios Marinos En El Mediterráneo Occidental: Actas de La VII Aula de Ecología. Almería, pp. 29–37.

Tomás, A., Méndez, F.J., Losada, Í.J., 2012. Bed Shear Stress during Sant Esteve's Storm (26th December 2008) along the Catalonia's Coast (NW Mediterranean).

Touchette, B.W., 2007. Seagrass-salinity interactions: physiological mechanisms used by submersed marine angiosperms for a life at sea. J. Exp. Mar. Bio. Ecol. 350, 194–215. http://dx.doi.org/10.1016/j.jembe.2007.05.037.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., Steininger, M., 2003. Remote sensing for biodiversity science and conservation. Trends Ecol. Evol. 18, 306–314. http://dx.doi.org/10.1016/S0169-5347(03)00070-3.

UNEP-MAP-RAC/SPA, 1995. Convention for the Protection of the Marine Environment and the Coastal Region of the Mediterranean.

Unsworth, R.K.F., Collier, C.J., Waycott, M., Mckenzie, L.J., Cullen-Unsworth, L.C., 2015. A framework for the resilience of seagrass ecosystems. Mar. Pollut. Bull. http://dx.doi.org/10.1016/j.marpolbul.2015.08.016.

Valle, M., Chust, G., del Campo, A., Wisz, M.S., Olsen, S.M., Garmendia, J.M., Borja, Á., 2014. Projecting future distribution of the seagrass Zostera noltii under global warming and sea level rise. Biol. Conserv. 170, 74–85. http://dx.doi.org/10.1016/j.biocon.2013.12.017.

Valle, M., van Katwijk, M.M., de Jong, D.J., Bouma, T.J., Schipper, A.M., Chust, G., Benito, B.M., Garmendia, J.M., Borja, Á., 2013. Comparing the performance of species distribution models of Zostera marina: implications for conservation. J. Sea Res. 83, 56–64. http://dx.doi.org/10.1016/j.seares.2013.03.002.

Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J., 2000. SOM Toolbox for Matlab 5. Tech. Rep. A57 2, p. 59.

Vuorinen, I., Hänninen, J., Rajasilta, M., Laine, P., Eklund, J., Montesino-Pouzols, F., Corona, F., Junker, K., Meier, H.E.M., Dippner, J.W., 2015. Scenario simulations of future salinity and ecological consequences in the Baltic Sea and adjacent North Sea areas—implications for environmental monitoring. Ecol. Indic. 50, 196–205. http://dx.doi.org/10.1016/j.ecolind.2014.10.019.

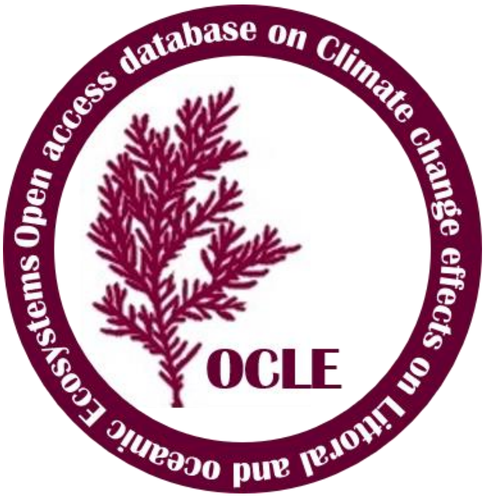
Witze, A., 2017. Ocean ecosystems mapped in unprecedented 3D detail. Nature 541, 10–11.

Worner, S., Eschen, R., Kenis, M., Paini, D., Saikkonen, K., Suiter, K., Venette, R.C., 2015. Detecting and interpreting patterns within regional pest species assemblages using self-organizing maps and other clustering methods. In: Venette, R.C. (Ed.), Pest Risk Modelling and Mapping for Invasive Alien Species, p. 97. Boston, USA. www.cabi.org.

Zacharias, M.A., Roff, J.C., 2000. A hierarchical ecological approach to conserving marine biodiversity. Conserv. Biol. 14, 1327–1334. http://dx.doi.org/10.1046/j.1523-1739.2000.99191.x.

Annex B

OCLE User manual



OPEN ACCESS DATABASE ON CLIMATE CHANGE EFFECTS ON LITTORAL AND OCEANIC ECOSYSTEMS (OCLE)

USER MANUAL

CONTENTS

1. INTRODUCTION

This document is conceived as a guide to use the Open access database on Climate change Effects on Littoral and oceanic Ecosystems (OCLE) website (http://ocle.ihcantabria.com/). OCLE is an ecological-driven database of present and future hazards for marine life in Europe and synthetic information on species distribution. It includes marine dynamics variables and the novelty of the bottom shear stress for the whole Europe. Besides its spatial and temporal resolution is enough to assess parameters relevant to the species distribution. On the other hand, data for several characteristic macroalgae taxa is available along the N and NW Iberian Peninsula. Accuracy of both physical and biological data lies in the quality control processes and the homogeneity of the methodology established to build the data set.

The manual is organized as the website sections to make easy its interpretation.

Please cite the data as indicated:

de la Hoz, C.F., Ramos, E., Acevedo, A., Puente, A., Losada, I.J. & Juanes, J.A., 2018. OCLE: a European open access database on climate change effects on littoral and oceanic ecosystems. Progress in Oceanography, 168, 222-231. https://doi.org/10.1016/j.pocean.2018.09.021.

2. BASIC CONCEPTS

2.1. Significant wave height (Hs)

In a group of N wave heights measured at a point and ordered from the largest to the smallest, Hs is defined as the average of the first (highest) N/3 waves (Dean & Dalrymple, 1991).

2.2. Bottom shear stress

According to Soulsby *et al.* (1993), the bottom shear stress can be understood as the force of friction from the interaction of waves and currents acting on the bed on the sea.

2.3. Representative Concentration Pathway (RCP)

For the 21st century different pathways of greenhouse gas emissions and atmospheric concentrations, air pollutant emissions and land-use have been developed (IPCC, 2014). The RCPs include a stringent mitigation scenario (RCP2.6), two intermediate scenarios (RCP4.5 and RCP6.0), and one scenario with very high greenhouse gas emissions (RCP8.5). Scenarios without additional efforts to constrain emissions ("baseline scenarios") lead to pathways ranging between RCP6.0 and RCP8.5. In OCLE, RCP4.5 has been included as an optimistic scenario and RCP8.5 as a pessimistic one.

2.4. Sea level

In OCLE it was considered the reference sea level of the historical period as 0, therefore, for the projected scenarios the predicted meters of rise are specified, according to Slangen *et al.* (2014).

2.5. Parameter: intertidal or subtidal

In biological contents, *Parameter* indicates the vertical zonation of macroalgae. The *intertidal* zone is submersed and emerged periodically due to tides. The *subtidal* zone is submersed with the upper part at extreme low water level.

3. METHODOLOGY

3.1. Physical data

Regarding physical data, please refer to the Materials and Methods section of de la Hoz et al. (2018).

3.2. Biological data

Biological data has been obtained from the following original IHCantabria works (PhD thesis, Technical Reports, etc):

– Juanes, J.A., Gutiérrez, L. (1992). Cartografía y evaluación de biomasa de *Gelidium sesquipedale* (Clem). Born. et Thur. en la costa oriental de Cantabria (N. de España). *Scientia Marina*, 166, 1-22.

– Juanes, J.A., Gutiérrez, L., Puente, A. (1991). Seguimiento y control de los campos de interés comercial de *Gelidium sesquipedale* en la costa occidental de Cantabria, Universidad de Cantabria - Fundación Leonardo Torres Quevedo. Informe inédito.

– UC (2005). Trabajos efectuados en el convenio para el "Desarrollo de diversos programas de actuación encaminados a corregir los efectos de los vertidos de fuel procedentes del buque Prestige en las costas de Cantabria". Documentos I a VIII. Universidad de Cantabria - Gobierno de Cantabria. Santander.

– IHCantabria (2008). Desarrollo de una metodología para la evaluación del estado de las aguas superficiales del entorno de afección de los saneamientos litorales. Aplicación a los saneamientos litorales de la Confederación Hidrográfica del Norte. Informe inédito. Convenio CEDEX-UC.

– Juanes, J.A., Guinda, X., Puente, A., Revilla, J.A. (2008). Macroalgae, a suitable indicator of the ecological status of coastal rocky communities in the NE Atlantic. *Ecological Indicators*, 8, 351–359.

– Guinda, X., Juanes, J.A., Puente, A., Echavarri-Erasun, B. (2012). Spatial distribution pattern analysis of subtidal macroalgae assemblages by a non-destructive rapid assessment method. *Journal of Sea Research*, 67, 34-43.

– IEO, Ministerio para la Transición Ecológica (2012). Evaluación inicial Estrategia Marina: Ecocartografías macroalgas.

– De Ugarte, A. (2012). Adding value to historical underwater video data for the study of subtidal macroalgae communities. Tesina de Máster Oficial en Gestión Ambiental de Sistemas Hídricos (Universidad de Cantabria).

- Guinda, X., Juanes, J.A., Puente, A. (2014). The CFR index: A validated method for the assessment of macroalgae according to the European Water Framework Directive. *Marine Environmental Research*, 102, 3-10.

- Ramos, E., Puente, A., Juanes, J.A. (2016). An ecological classification of rocky shores at a regional scale: a predictive tool for management of conservation values. *Marine Ecology*, 37, 311–328.

- Ramos, E., Puente, A., Guinda, X., Juanes J.A. (2017). A hierarchical classification system along the NE Atlantic coast: focusing on the local scale (Cantabria, N Spain). *European Journal of Phycology*, 52 (1), 75-89.

- IHCantabria (2008-2018). Análisis histórico de comunidades de macroalgas en La Arnía (Cantabria).

The general methodology applied to obtain the final data visualized in OCLE is summarized in the following steps:

1. To filter data by the selected characteristic taxa.

2. To group data by season and year.

3. To join data to the nearest point of OCLE mesh and to count number of records (i.e. number of presences of each taxa from the different works found for a given mesh point and a previously selected by season and year).

4. OCLE WEBSITE

The OCLE website (http://ocle.ihcantabria.com/) is divided in the following sections:

4.1. Home

The description of the project and its objectives are described in this section.

4.2. Contents

Physical data are available in this section.

4.2.1. Physical

Data selection

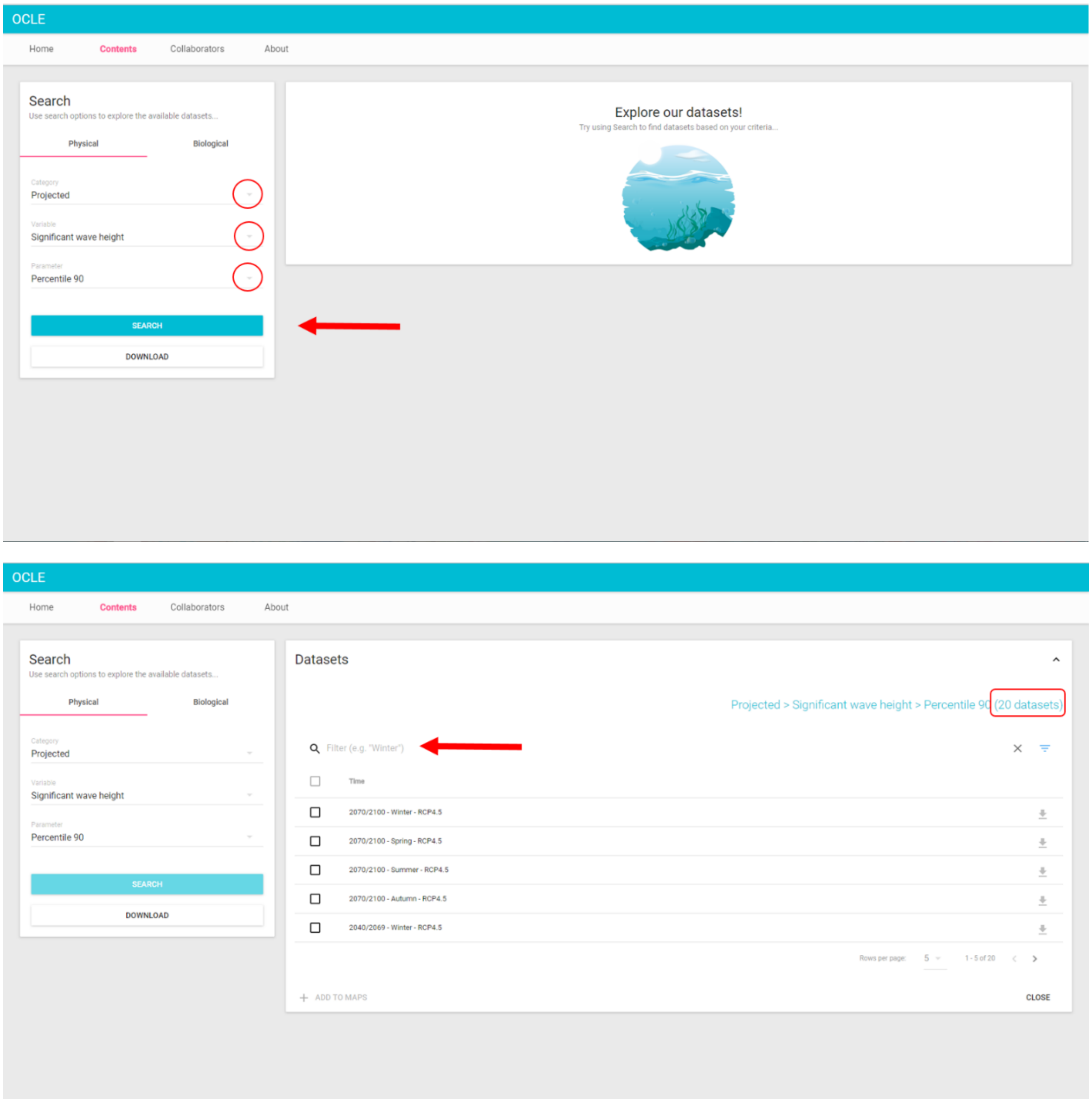
The first step for data selection is the selection of the period, the variable and the parameter of interest, within the options in the drop-down menu (Fig. 1).

By clicking on *Search*, all datasets available appear. It is possible to filter the datasets by season, years, RCP, etc.

Period
Historical
Projected

Variable
Currents
Significant wave height
Attenuation coefficient
Nitrate
pH
Radiation
Salinity
Sea level
Shear stress
Sea surface temperature
Air temperature
Tidal range
Wind speed
Bottom orbital speed
N days/year in P90
N days/year over 2.2N/m^2

Parameter
Maximum
Mean
Minimum
Percentile 10
Percentile 25
Percentile 50
Percentile 75
Percentile 90
Range
Standard Deviation



Search
Use search options to explore the available datasets...

Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

Explore our datasets!
Try using Search to find datasets based on your criteria...

Datasets
Projected > Significant wave height > Percentile 90 (20 datasets)

Filter (e.g. "Winter")

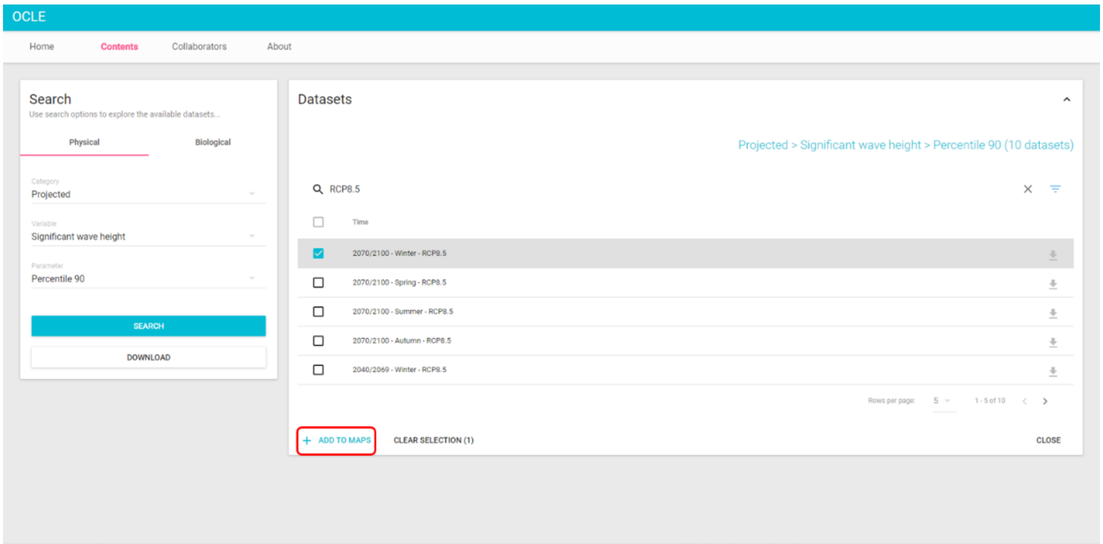
☐	Time
☐	2070/2100 - Winter - RCP4.5
☐	2070/2100 - Spring - RCP4.5
☐	2070/2100 - Summer - RCP4.5
☐	2070/2100 - Autumn - RCP4.5
☐	2040/2069 - Winter - RCP4.5

Rows per page: 5 1 - 5 of 20

Fig. 1. Physical data selection in OCLE website.

Data exploration

The selected datasets can be visualized and explored by clicking in *Add to maps* button (Fig. 2). It is possible to add until six maps each time that can be combined of different parameters and variables. In these maps it is possible to get the value of each mesh point interactively (Fig. 3). Besides, to make easier the comparison of the maps, when the frame of a map is changed by the user, all maps change at the same time.



Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

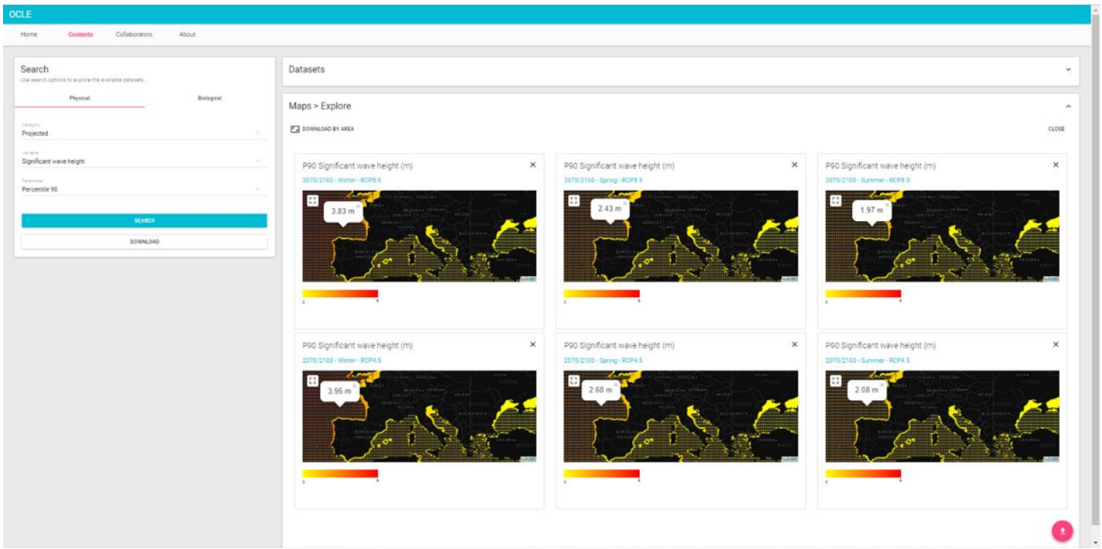
Projected > Significant wave height > Percentile 90 (10 datasets)

RCP8.5

☐ Time
☑ 2070/2100 - Winter - RCP8.5
☐ 2070/2100 - Spring - RCP8.5
☐ 2070/2100 - Summer - RCP8.5
☐ 2070/2100 - Autumn - RCP8.5
☐ 2040/2069 - Winter - RCP8.5

Rows per page: 5 1 - 5 of 10

Fig 2. Add to maps the datasets selected



Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

P90 Significant wave height (m)
2070/2100 - Winter - RCP8.5

P90 Significant wave height (m)
2070/2100 - Spring - RCP8.5

P90 Significant wave height (m)
2070/2100 - Summer - RCP8.5

P90 Significant wave height (m)
2070/2100 - Winter - RCP4.5

P90 Significant wave height (m)
2070/2100 - Spring - RCP4.5

P90 Significant wave height (m)
2070/2100 - Summer - RCP4.5

Fig 3. Interactive exploration of datasets for different RCPs and seasons

A map can be deleted by clicking the *X* in the upper right corner (Fig. 4).

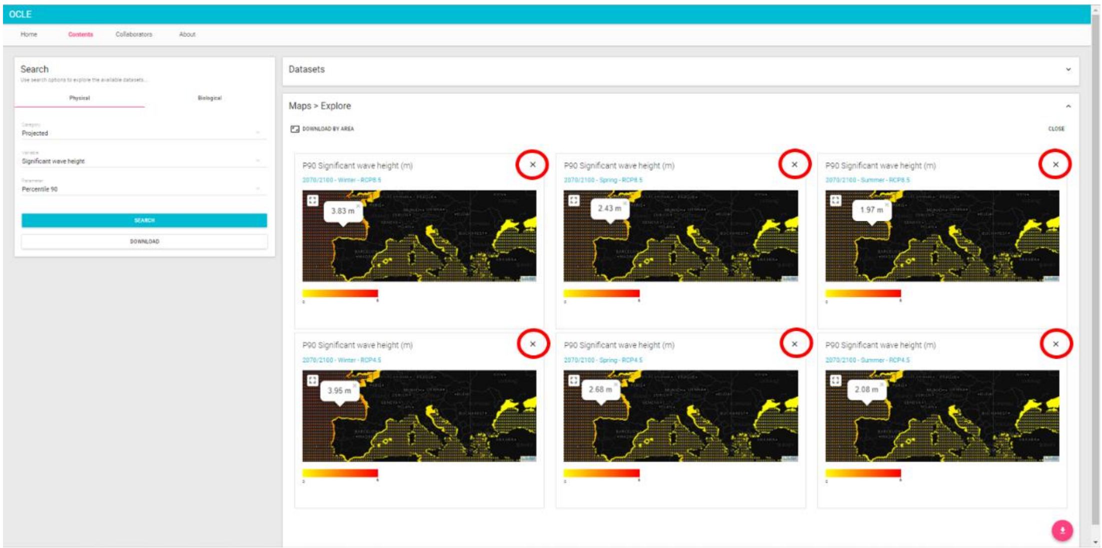


Fig. 4. Delete a map

Data filtering

To avoid computational overcharges, data can be spatially filtered by clicking in the *Download by Area* button in the exploration screen (Fig. 5).

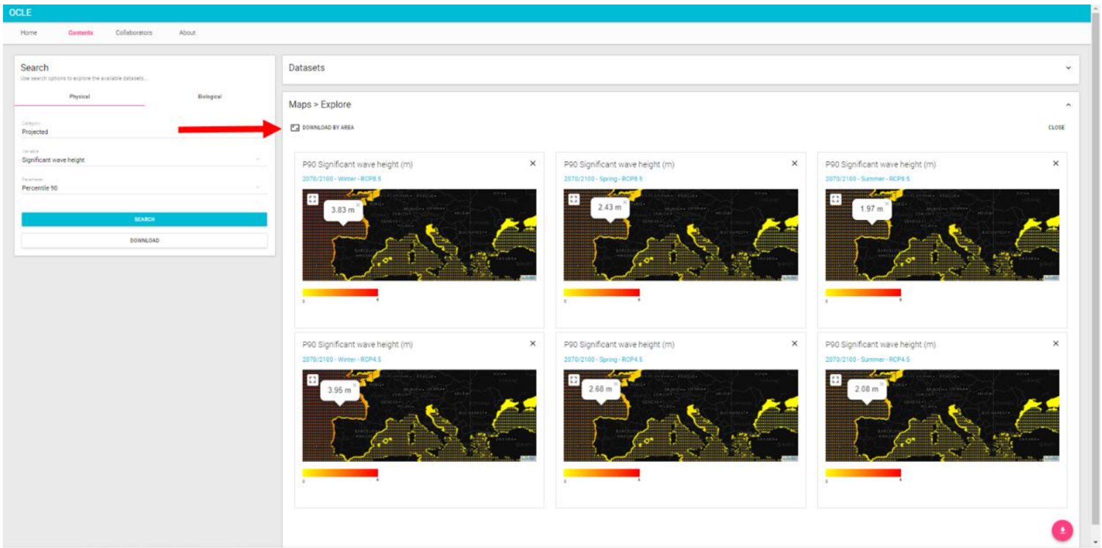
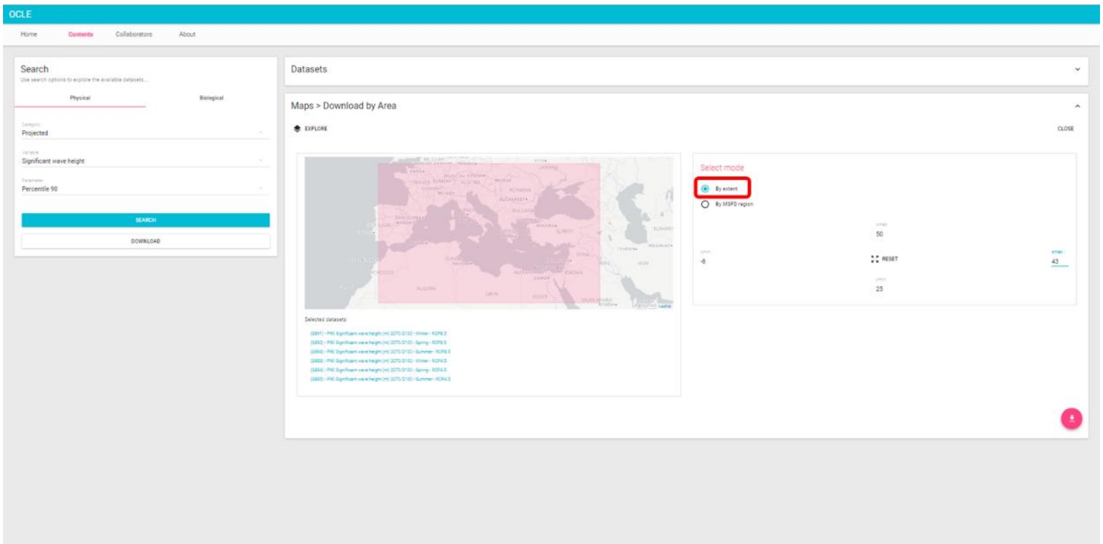


Fig 5. *Download by Area* button.

This filter can be applied in the following ways:

- **By extent**

User can introduce the coordinates of interest in decimal degrees in the datum WGS84 (Fig. 6) by selecting the *By extent* button.



Projected
Significant wave height
Percentile 90

● By extent
○ By MSFD region

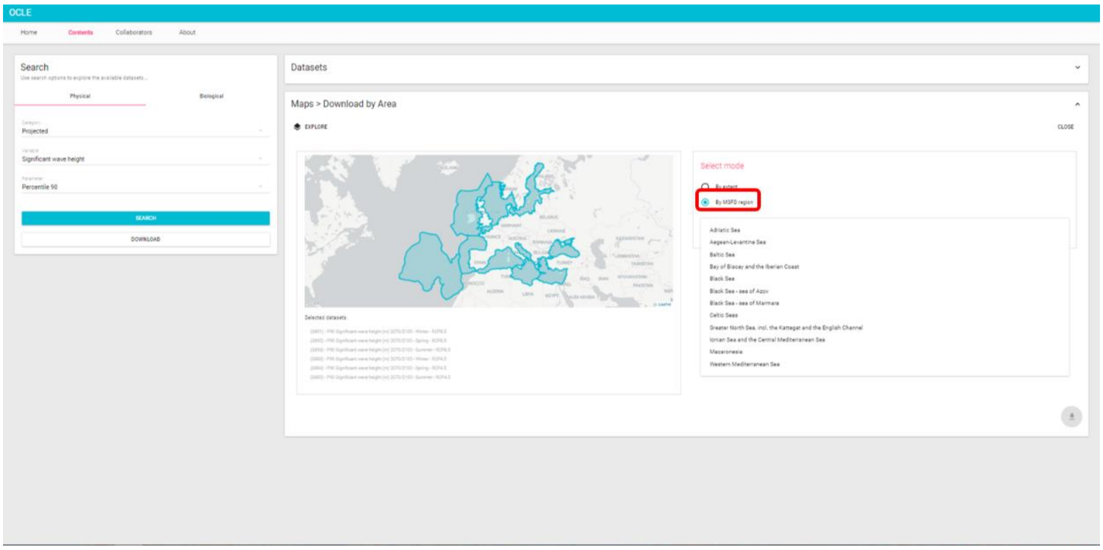
50
-8
43
25

Fig 6. Selection by extent

- **By MSFD regions**

The Marine Strategy Framework Directive (MSFD) regions (European Commission, 2008) are predefined, therefore is possible to select the mesh points that overlap these areas by selecting the *By MSFD regions* button (Fig 7).

MSFD regions
Adriatic Sea
Aegean-Levantine Sea
Baltic Sea
Bay of Biscay and the Iberian Coast
Black Sea
Black Sea – Sea of Azov
Black Sea – Sea of Marmara
Celtic Seas
Greater North Sea, incl. the Kattegat and the English Channel
Ionian Sea and the Central Mediterranean Sea
Macaronesia
Western Mediterranean Sea



Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

○ By extent
● By MSFD region

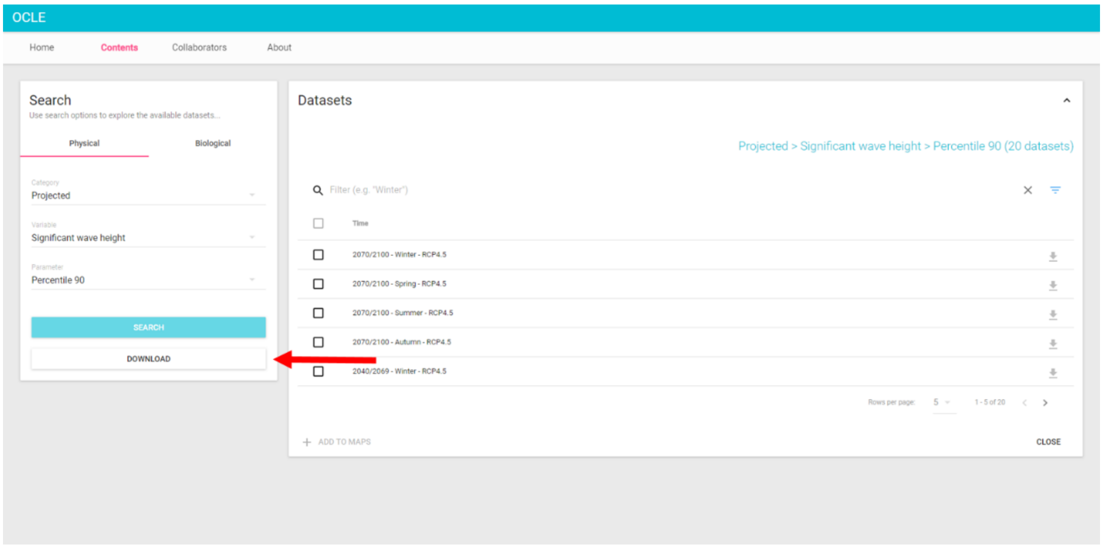
Adriatic Sea
Aegean-Levantine Sea
Baltic Sea
Bay of Biscay and the Iberian Coast
Black Sea
Black Sea - sea of Azov
Black Sea - sea of Marmara
Celtic Seas
Greater North Sea, incl. the Kattegat and the English Channel
Ionian Sea and the Central Mediterranean Sea
Macaronesia
Western Mediterranean Sea

Fig 7. Selection by MSFD regions

Data download

Data can be downloaded in three ways:

1. All the results for a search by clicking in the *download button* without selecting any dataset. A different file is generated for each dataset in the whole mesh (Fig. 8).



Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

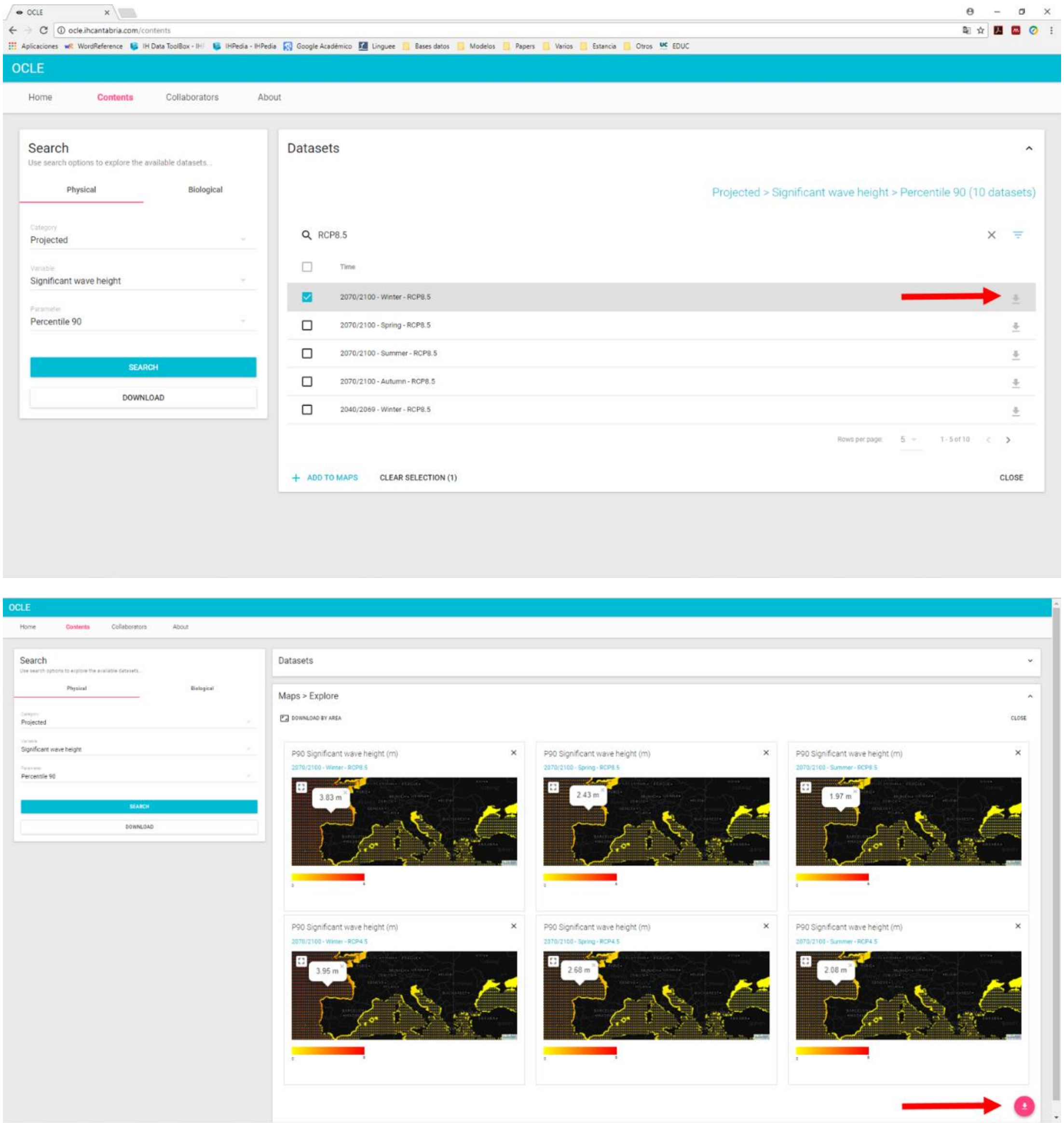
Projected > Significant wave height > Percentile 90 (20 datasets)

☐ Time
☐ 2070/2100 - Winter - RCP4.5
☐ 2070/2100 - Spring - RCP4.5
☐ 2070/2100 - Summer - RCP4.5
☐ 2070/2100 - Autumn - RCP4.5
☐ 2040/2069 - Winter - RCP4.5

Rows per page: 5 1 - 5 of 20

Fig 8. Download all datasets

2. The selected datasets in the whole mesh. By clicking in the *download button*, before or after adding to a map (Fig. 9).



Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

Projected > Significant wave height > Percentile 90 (10 datasets)

RCP8.5

	Time
☐	
☑	2070/2100 - Winter - RCP8.5
☐	2070/2100 - Spring - RCP8.5
☐	2070/2100 - Summer - RCP8.5
☐	2070/2100 - Autumn - RCP8.5
☐	2040/2069 - Winter - RCP8.5

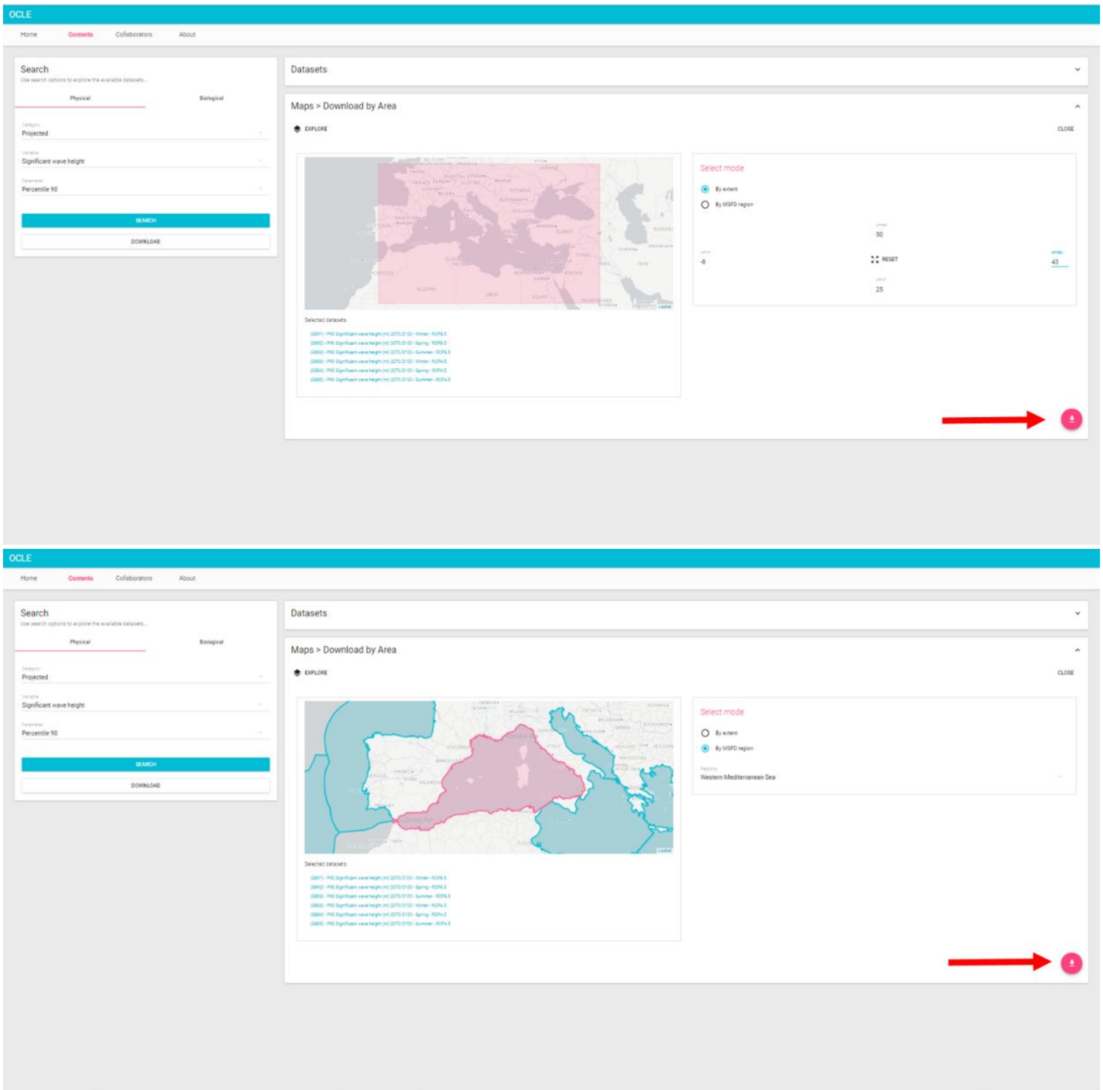
Rows per page: 5 1 - 5 of 10

2070/2100 - Winter - RCP8.5
2070/2100 - Spring - RCP8.5
2070/2100 - Summer - RCP8.5
2070/2100 - Winter - RCP4.5
2070/2100 - Spring - RCP4.5
2070/2100 - Summer - RCP4.5

Fig 9. Download of selected datasets for the whole mesh without adding to a map (upper panel) or after adding to a map (lower panel).

3. The selected datasets spatially filtered.

   The datasets filtered by extent, MSFD regions or the coastal points can be downloaded only for the selected area, using the *download button* (Fig. 10).



Upper panel:

Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

Select mode: ● By extent ○ By MSFD region

50
-8 43
25

Lower panel:

Category: Projected
Variable: Significant wave height
Parameter: Percentile 90

Select mode: ○ By extent ● By MSFD region

Regions: Western Mediterranean Sea

Fig 10. Download of datasets filtered by extent (upper panel) and MSFD region (lower panel).

➢ Data format

Data are in longitude/latitude coordinate reference system (not projected; decimal degrees) and the datum is WGS84.

Units for different variables are shown in Table 1.

Table 1. Variables included in OCLE with their short name and the corresponding units.

Variable	Short name	Units
Sea surface temperature	SST	ºC
Significant wave height	Hs	m
Attenuation coefficient	Kd	m^{-1}
Nitrate	Nit	mol/m^3
Salinity	Sal	psu
Air temperature	AT	ºC
pH	pH	
Wind speed	Ws	m/s
Tidal range	TR	m
Sea level	gost	m
Radiation	rss	W/m^2
Shear stress	Tmax	N/m^2
Bottom orbital speed	Urms	m/s
Currents	Uw	m/s
Bathymetry		m
Substrate		

File Name

The data are stored in ZIP (compressed) files. The ZIP files have names like N_V_T_P_Y-S.ZIP, where each letter indicates:

N -> internal number

V -> variable (with its short name)

T -> temporal period (historical or projected)

P -> parameter according to Table 3 alias

Y -> five years period of data

S -> season (winter was defined as January-February-March, spring as April-May-June, summer as July-August-September and autumn as October-November-December)

Table 3. Parameters included in OCLE with their associated short names.

Parameter	Short name
Maximum	MAX
Mean	MEAN
Minimum	MIN
Percentile 10	P10
Percentile 25	P25
Percentile 50	P50
Percentile 75	P75
Percentile 90	P90
Range	RG
Standard deviation	STD
Number of consecutive days over the percentile 90 of the historic period	nDaysYP90
Number of consecutive days of shear stress over the value 2.2 N/m^2	nDaysYTmax

➢ File format

OCLE data is downloaded in comma-separated values format (.csv) with the following fields (Fig. 11):

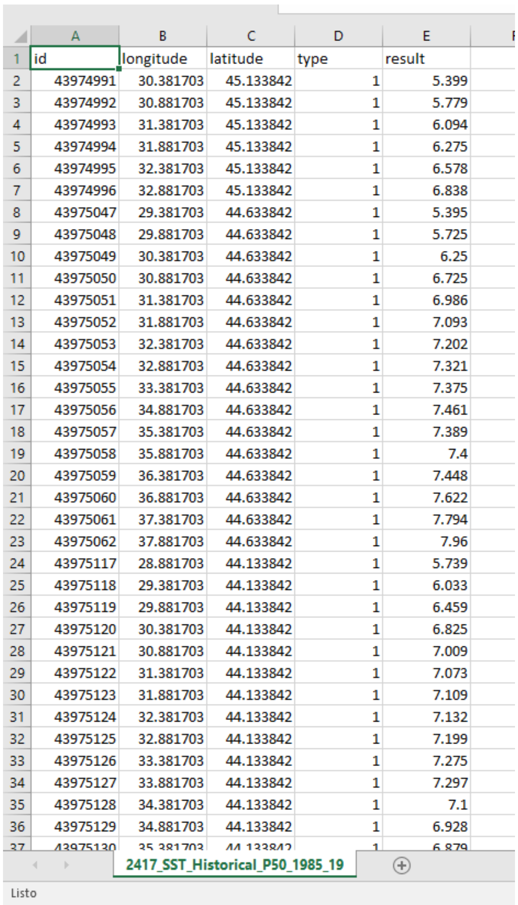
id -> internal identifier

longitude -> longitudinal degrees EPSG projection 4326 WGS 84

latitude -> latitudinal degrees EPSG projection 4326 WGS 84

type -> internal parameter

result -> variable value



	A	B	C	D	E
1	id	longitude	latitude	type	result
2	43974991	30.381703	45.133842	1	5.399
3	43974992	30.881703	45.133842	1	5.779
4	43974993	31.381703	45.133842	1	6.094
5	43974994	31.881703	45.133842	1	6.275
6	43974995	32.381703	45.133842	1	6.578
7	43974996	32.881703	45.133842	1	6.838
8	43975047	29.381703	44.633842	1	5.395
9	43975048	29.881703	44.633842	1	5.725
10	43975049	30.381703	44.633842	1	6.25
11	43975050	30.881703	44.633842	1	6.725
12	43975051	31.381703	44.633842	1	6.986
13	43975052	31.881703	44.633842	1	7.093
14	43975053	32.381703	44.633842	1	7.202
15	43975054	32.881703	44.633842	1	7.321
16	43975055	33.381703	44.633842	1	7.375
17	43975056	34.881703	44.633842	1	7.461
18	43975057	35.381703	44.633842	1	7.389
19	43975058	35.881703	44.633842	1	7.4
20	43975059	36.381703	44.633842	1	7.448
21	43975060	36.881703	44.633842	1	7.622
22	43975061	37.381703	44.633842	1	7.794
23	43975062	37.881703	44.633842	1	7.96
24	43975117	28.881703	44.133842	1	5.739
25	43975118	29.381703	44.133842	1	6.033
26	43975119	29.881703	44.133842	1	6.459
27	43975120	30.381703	44.133842	1	6.825
28	43975121	30.881703	44.133842	1	7.009
29	43975122	31.381703	44.133842	1	7.073
30	43975123	31.881703	44.133842	1	7.109
31	43975124	32.381703	44.133842	1	7.132
32	43975125	32.881703	44.133842	1	7.199
33	43975126	33.381703	44.133842	1	7.275
34	43975127	33.881703	44.133842	1	7.297
35	43975128	34.381703	44.133842	1	7.1
36	43975129	34.881703	44.133842	1	6.928
37	43975130	35.381703	44.133842	1	6.879

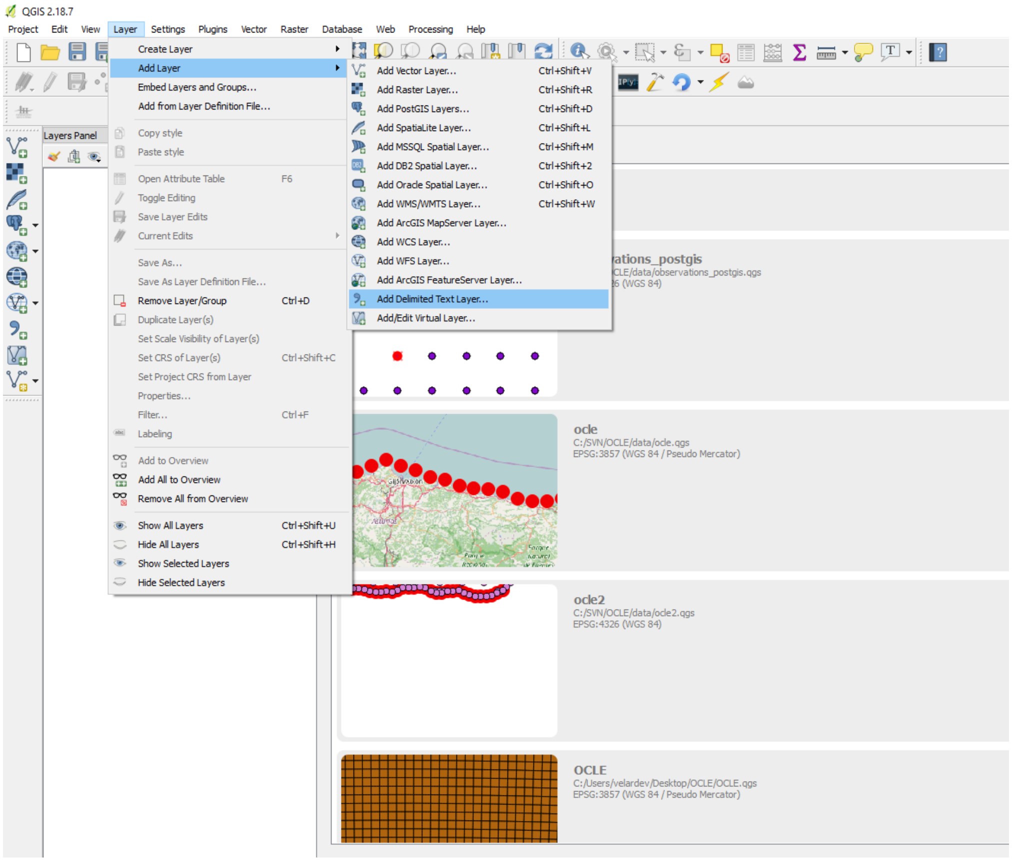
Fig. 11. Example of download file.

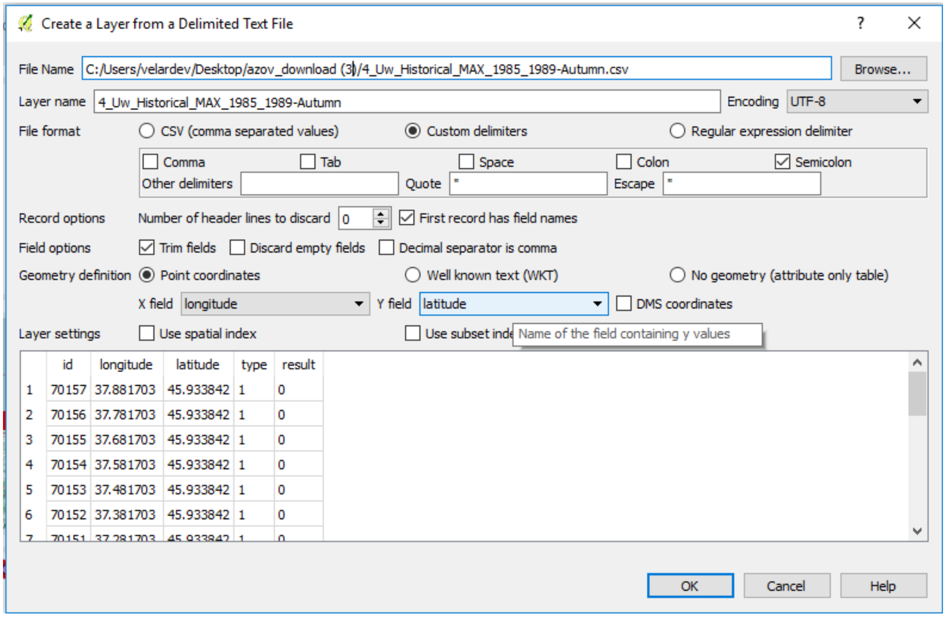
➢ Data use

The OCLE data can be easily imported into most GIS applications.

<u>QGIS</u>

Here is some help as an example in QGIS:





Create a Layer from a Delimited Text File

File Name C:/Users/velardev/Desktop/azov_download (3)/4_Uw_Historical_MAX_1985_1989-Autumn.csv

Layer name 4_Uw_Historical_MAX_1985_1989-Autumn Encoding UTF-8

File format ○ CSV (comma separated values) ● Custom delimiters ○ Regular expression delimiter

☐ Comma ☐ Tab ☐ Space ☐ Colon ☑ Semicolon

Other delimiters Quote " Escape "

Record options Number of header lines to discard 0 ☑ First record has field names

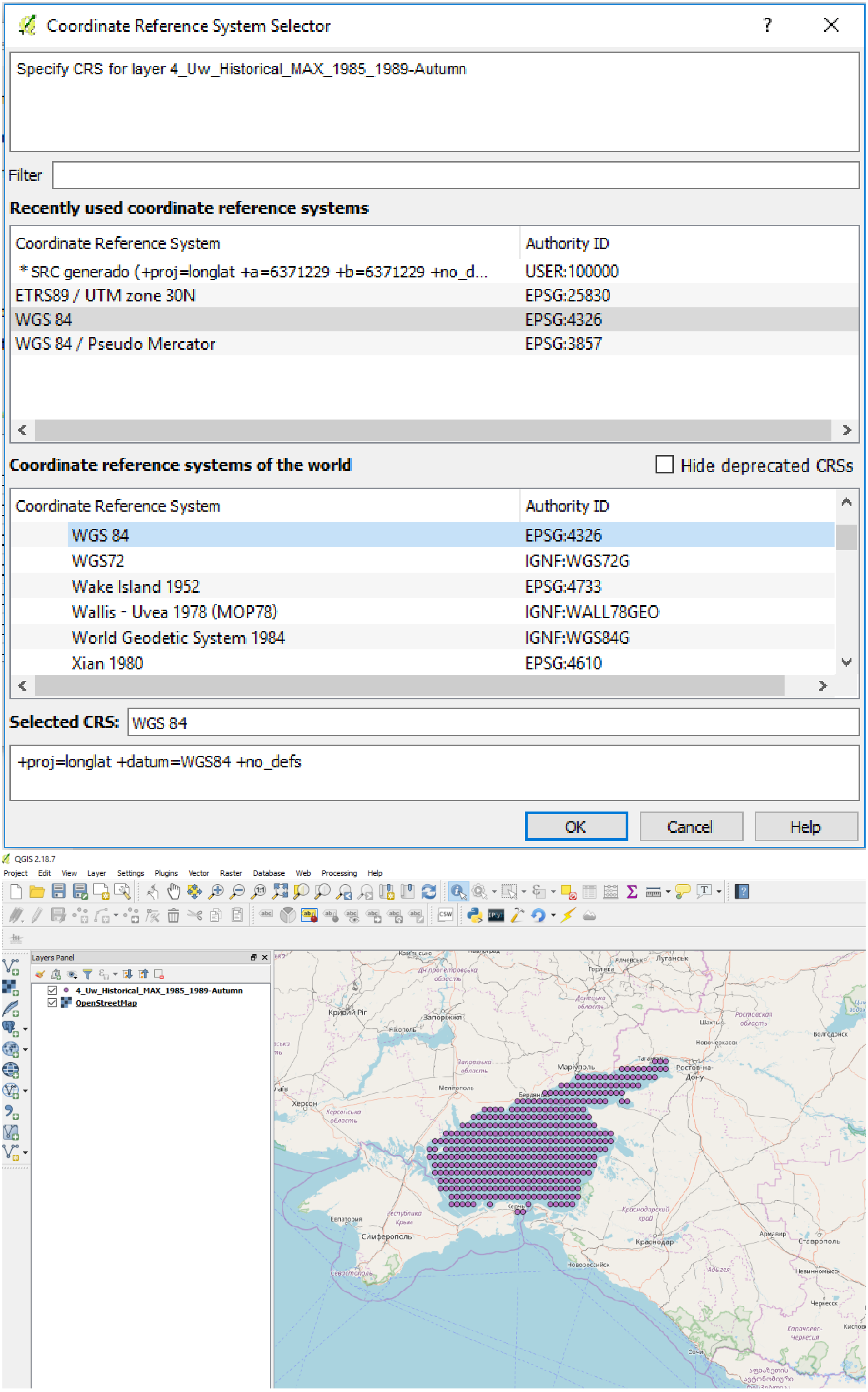
Field options ☑ Trim fields ☐ Discard empty fields ☐ Decimal separator is comma

Geometry definition ● Point coordinates ○ Well known text (WKT) ○ No geometry (attribute only table)

X field longitude Y field latitude ☐ DMS coordinates

Layer settings ☐ Use spatial index ☐ Use subset inde Name of the field containing y values

	id	longitude	latitude	type	result
1	70157	37.881703	45.933842	1	0
2	70156	37.781703	45.933842	1	0
3	70155	37.681703	45.933842	1	0
4	70154	37.581703	45.933842	1	0
5	70153	37.481703	45.933842	1	0
6	70152	37.381703	45.933842	1	0
7	70151	37.281703	45.933842	1	0



Coordinate Reference System Selector

Specify CRS for layer 4_Uw_Historical_MAX_1985_1989-Autumn

Filter

Recently used coordinate reference systems

Coordinate Reference System	Authority ID
* SRC generado (+proj=longlat +a=6371229 +b=6371229 +no_d...	USER:100000
ETRS89 / UTM zone 30N	EPSG:25830
WGS 84	EPSG:4326
WGS 84 / Pseudo Mercator	EPSG:3857

Coordinate reference systems of the world ☐ Hide deprecated CRSs

Coordinate Reference System	Authority ID
WGS 84	EPSG:4326
WGS72	IGNF:WGS72G
Wake Island 1952	EPSG:4733
Wallis - Uvea 1978 (MOP78)	IGNF:WALL78GEO
World Geodetic System 1984	IGNF:WGS84G
Xian 1980	EPSG:4610

Selected CRS: WGS 84

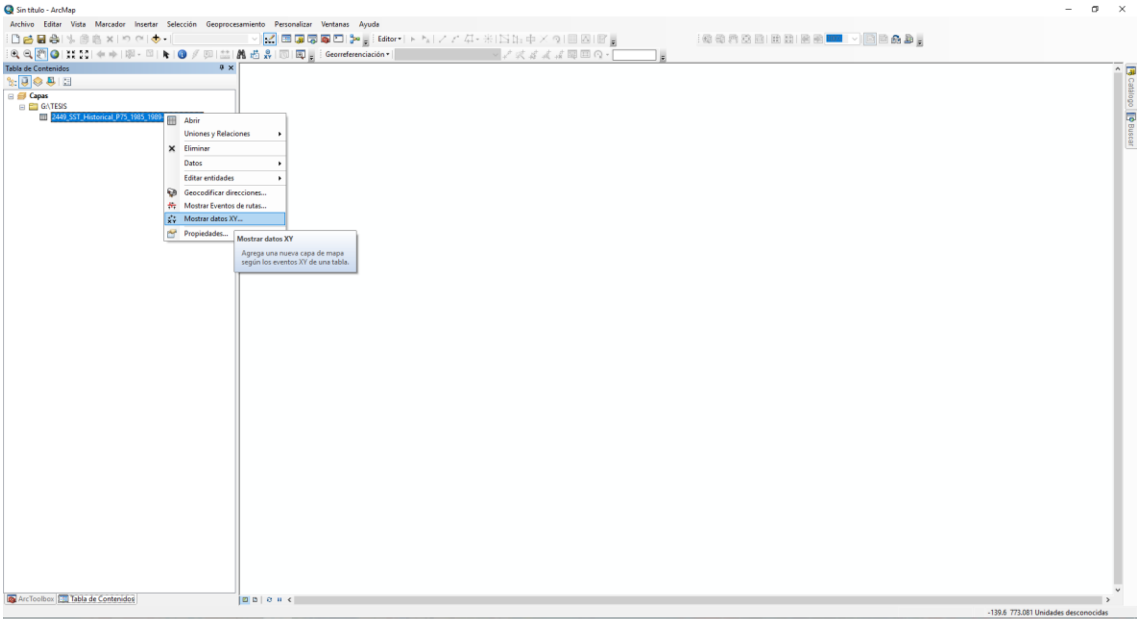
+proj=longlat +datum=WGS84 +no_defs

Fig. 12. Example of steps to follow in QGIS.

ARCGIS

In ArcGis the process to follow is detailed below:

After loading the file, *show XYData*.

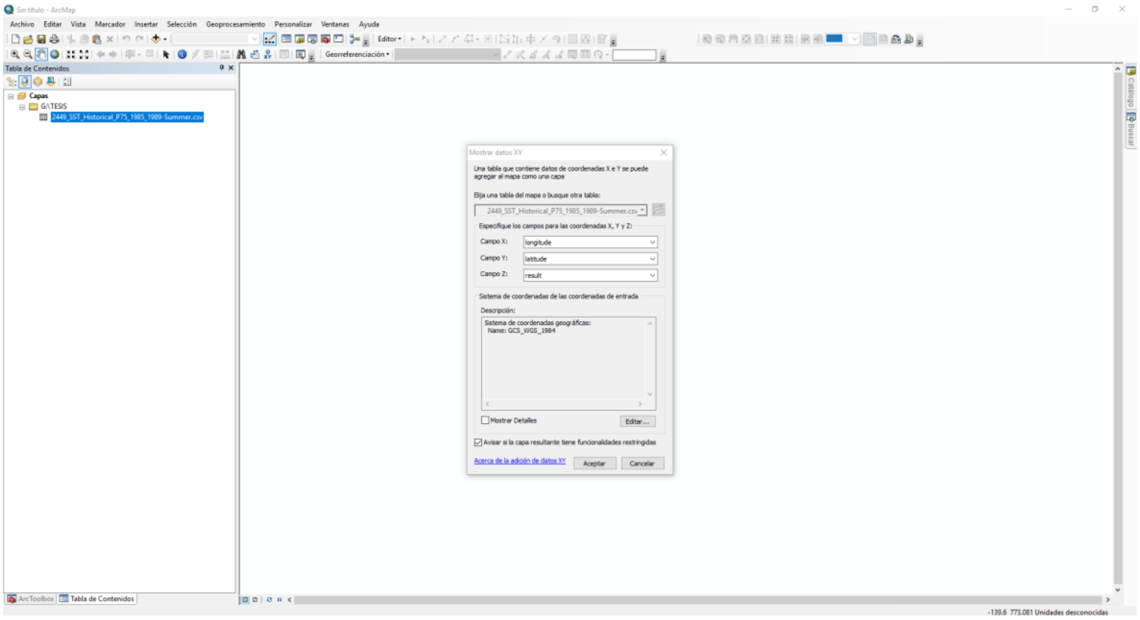


Mostrar datos XY

Agrega una nueva capa de mapa según los eventos XY de una tabla.

Fig. 13. Show XYData.

Define the fields, X, Y and result and the coordinate system (WGS84).



Una tabla que contiene datos de coordenadas X e Y se puede agregar al mapa como una capa

Elija una tabla del mapa o busque otra tabla: 2449_SST_Historical_P75_1985_1989-Summer.csv

Especifique los campos para las coordenadas X, Y y Z:

Campo X: longitude
Campo Y: latitude
Campo Z: result

Sistema de coordenadas de las coordenadas de entrada

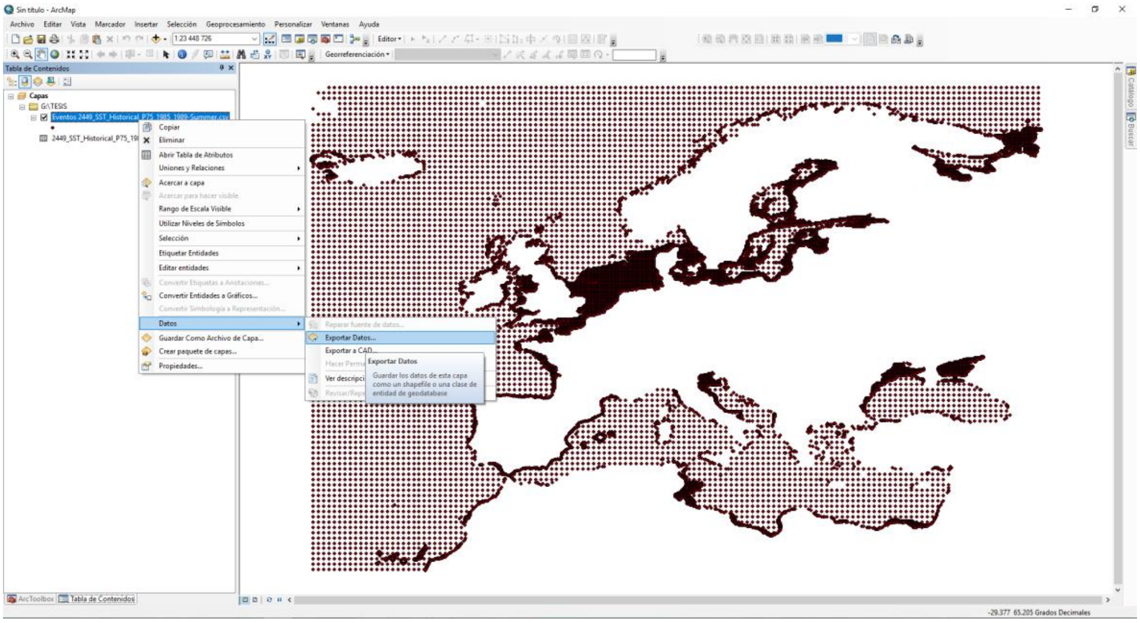
Descripción:
Sistema de coordenadas geográficas:
Name: GCS_WGS_1984

☐ Mostrar Detalles

☑ Avisar si la capa resultante tiene funcionalidades restringidas

Fig. 14. Definition of fields.

Export the data as shapefile



Guardar los datos de esta capa como un shapefile o una clase de entidad de geodatabase

Fig. 15. Export the data as shapefile.

Save in the desired directory and load in the map:

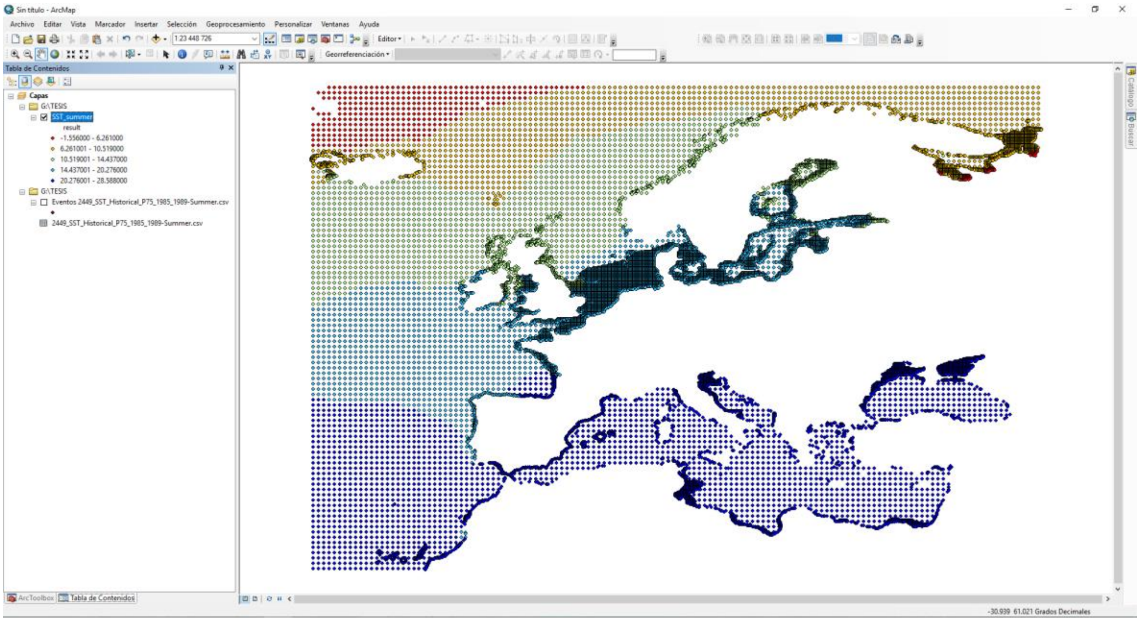


Fig. 16. Save the data.

R software

Data can be also read with R software (R Core Team, 2018). To facilitate data processing, a script is provided in OCLE website to transform data from .csv format to shapefile *FromCsvToShp.R*. The code is detailed below:

```
# FromCsvToShp

# Transform the data from OCLE database to shapefile to be used in GIS

rm(list=ls())

library(readxl)

library(raster)

library(rgdal)

setwd('G:/TESIS/Chapters/1_Hazards/OCLE_database')

# Load the data

Data <- read.csv('2449_SST_Historical_P75_1985_1989-Summer.csv', header = T, sep = ';')

class(Data)

# [1] "data.frame"

# Convert to SpatialPointsDataFrame

coordinates(Data)= c('longitude','latitude')

class(Data)

# [1] "SpatialPointsDataFrame"

# attr(,"package")

# [1] "sp"

# Define the coordinate system

proj4string(Data) = CRS('+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0')

# Save as shapefile

writeOGR(Data, './shp_ejp', 'SST_Historical_P75_1985_1989-Summer', driver="ESRI Shapefile", overwrite_layer=T)
```

4.2.2. Biological

Data selection

The first step for data selection is the selection of the period, the taxa and the parameter of interest, within the options in the drop-down menu (Fig. 17).

By clicking on *Search*, all datasets available appear. It is possible to filter the datasets by season and/or years.

Period
Historical
Projected

Taxa
Bifurcaria bifurcata
Chondrus crispus
Corallina
C. officinalis
Cystoseira baccata
C. tamariscifolia
Ellisolandia elongata
Fucus
F. serratus
F. spiralis
F. vesiculosus
Gelidium
G. corneum
G. spinosum
Halopteris scoparia
Himanthalia elongata
Laminaria
L. hyperborea
L. ochroleuca
Pelvetia canaliculata
Sargassum muticum

Parameter
Intertidal
Subtidal

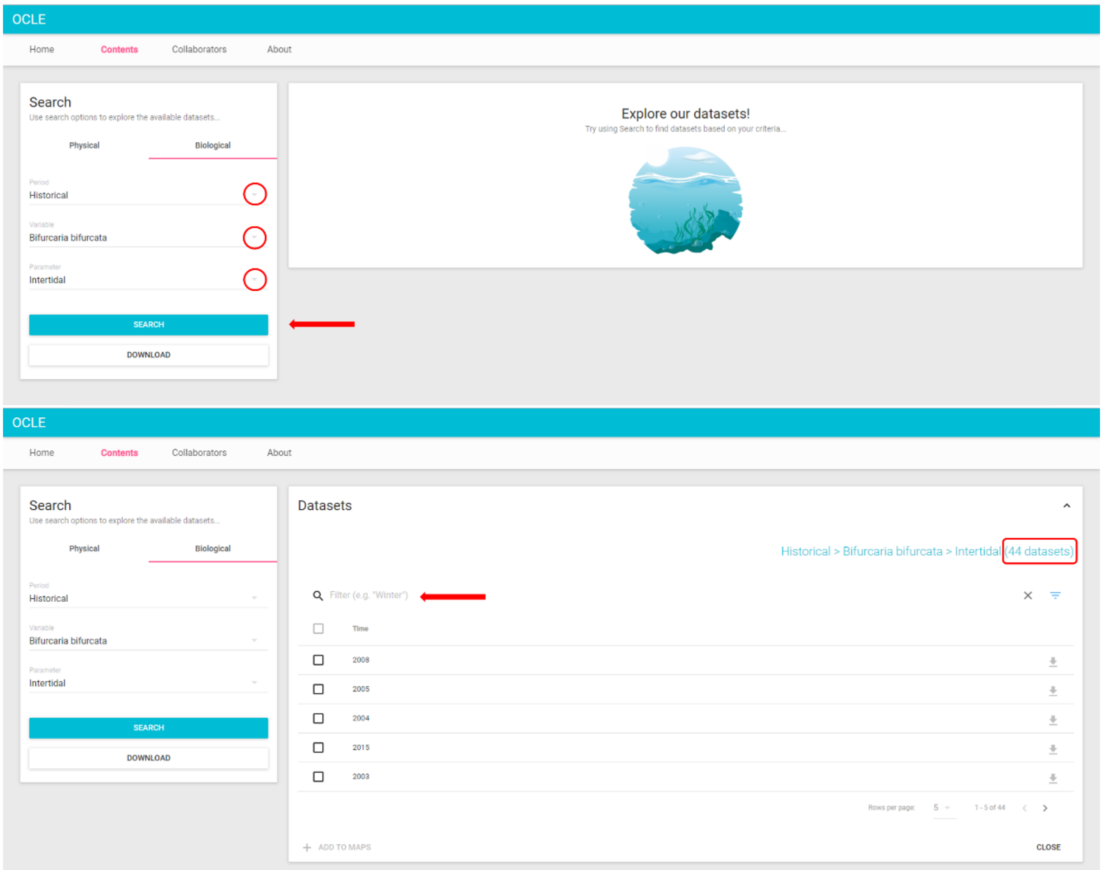


Fig. 17. Physical data selection in OCLE website.

➢ Data exploration

C.f. physical data exploration.

➢ Data download

C.f. physical data download.

➢ Data format

Data are in longitude/latitude coordinate reference system (not projected; decimal degrees) and the datum is WGS84.

Units for different taxa is No. of records.

➢ File Name

The data are stored in ZIP (compressed) files. The ZIP files have names like N_Ta_Te_P_Y-S.ZIP, where each letter indicates:

N -> internal number

Ta -> taxa

Te -> temporal period (historical or projected)

P -> parameter (INT: intertidal, SUB: subtidal)

Y -> year of data

S -> season (winter was defined as January-February-March, spring as April-May-June, summer as July-August-September and autumn as October-November-December)

File format

OCLE data is downloaded in comma-separated values format (.csv) with the following fields (Fig. 18):

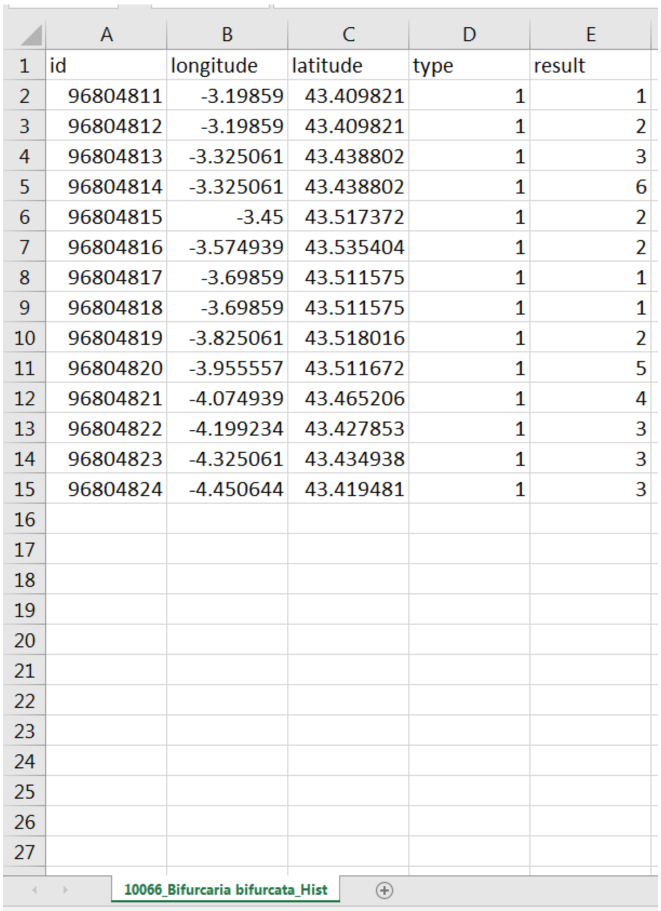
id -> internal identifier

longitude -> longitudinal degrees WGS 84

latitude -> latitudinal degrees WGS 84

type -> internal parameter

result -> number of records



id	longitude	latitude	type	result
96804811	-3.19859	43.409821	1	1
96804812	-3.19859	43.409821	1	2
96804813	-3.325061	43.438802	1	3
96804814	-3.325061	43.438802	1	6
96804815	-3.45	43.517372	1	2
96804816	-3.574939	43.535404	1	2
96804817	-3.69859	43.511575	1	1
96804818	-3.69859	43.511575	1	1
96804819	-3.825061	43.518016	1	2
96804820	-3.955557	43.511672	1	5
96804821	-4.074939	43.465206	1	4
96804822	-4.199234	43.427853	1	3
96804823	-4.325061	43.434938	1	3
96804824	-4.450644	43.419481	1	3

Fig. 18. Example of download file.

Data use

C.f. physical data use.

4.3. About

4.3.1. Datasets

This subsection includes an abstract of the work, the metadata in the Ecological Metadata Language format (Michener, Brunt, Helly, Kirchner, & Stafford, 1997) and the acknowledgements to the institutions that provide the raw data.

4.3.2. Citation

Here is provided the citation of the work in case some data were used for publication, the link to the paper were the database is detailed and the email to ask for the pdf.

4.3.3. Contact

E-mail to contact with developers.

4.3.4. Licence

In this subsection the licence and terms of use of the data can be read. It is mandatory for all OCLE users to read the license. Any question can be consulted by e-mail.

4.3.5. Disclaimer

Disclaimer is included also in the license.

REFERENCES

de la Hoz, C. F., Ramos, E., Acevedo, A., Puente, A., Losada, íñigo J., & Juanes, J. A. (2018). Ocle: a European Open Access Database on Climate Change Effects on Littoral and Oceanic Ecosystems. *Progress in Oceanography*, *168*, 222–231. doi:10.1016/j.pocean.2018.09.021

Dean, R. G. (Robert G., & Dalrymple, R. A. (1991). *Water wave mechanics for engineers and scientists*. World Scientific.

European Commission. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive), 164 Official Journal of the European Communities § (2008).

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)] Geneva, Switzerland.

Michener, W. K., Brunt, J. W., Helly, J. J., Kirchner, T. B., & Stafford, S. G. (1997). Nongeospatial metadata for the ecolgical sciences. *Ecological Applications*, *7*(1), 330–342. doi:10.1890/1051-0761(1997)007[0330:NMFTES]2.0.CO;2

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: Computing, R Foundation for Statistical. Retrieved from http://www.r-project.org

Slangen, A. B. A., Carson, M., Katsman, C. A., van de Wal, R. S. W., Vermeersen, L. L. A., & Stammer, D. (2014). Projecting twenty-first century regional sea-level changes. *Climatic Change*, *124*, 317–332. doi:10.1007/s10584-014-1080-9

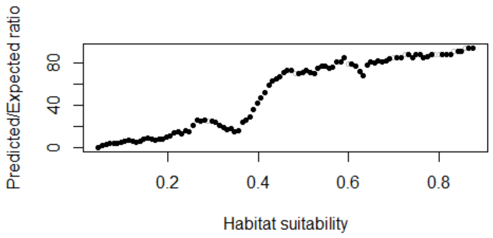
Soulsby, R. L., Hamm, L., Klopman, G., Myrhaug, D., Simons, R. R., & Thomas, G. P. (1993). Wave-current interaction within and outside the bottom boundary layer. *Coastal Engineering*, *21*(1–3), 41–69. doi:10.1016/0378-3839(93)90045-A

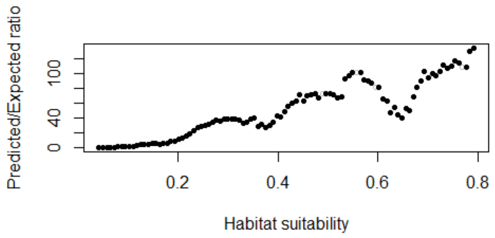
Annex C

Supporting information Chapter V

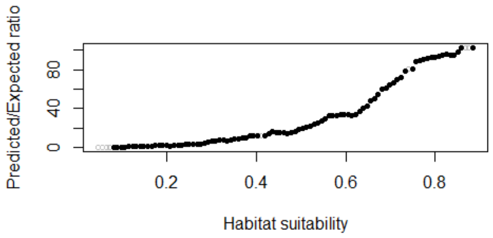
Annex C. Supporting information Chapter V

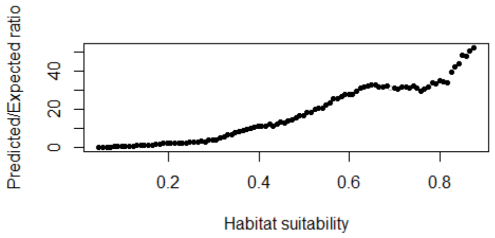
Saccorhiza polyschides ***Gelidium spinosum***





Sargassum muticum ***Pelvetia canaliculata***





Cystoseira baccata

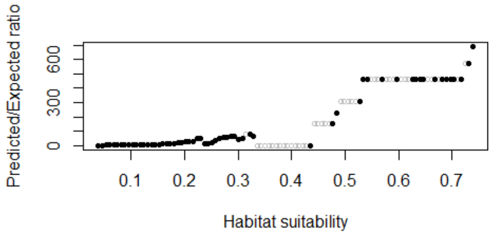


Figure S5.1 .Boyce index plots of the species considered.

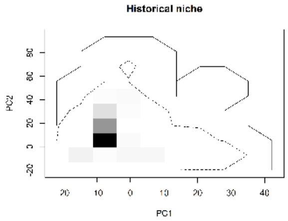
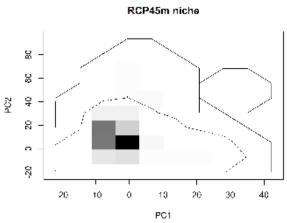
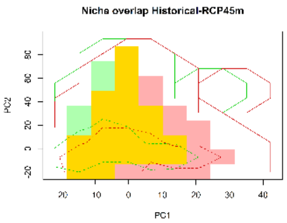
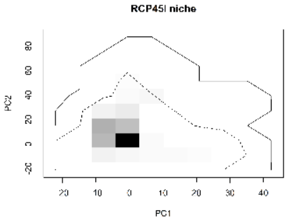
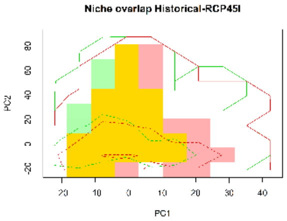
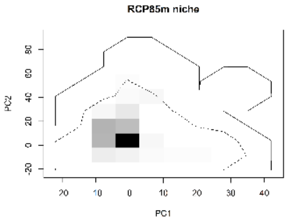
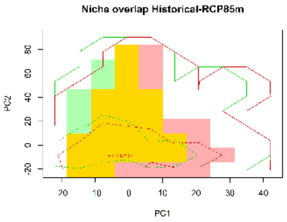
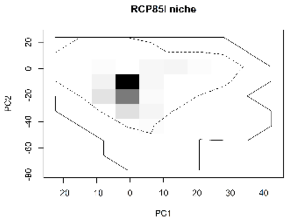
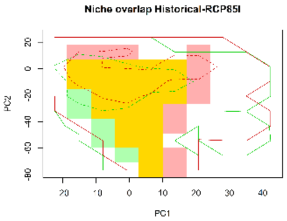
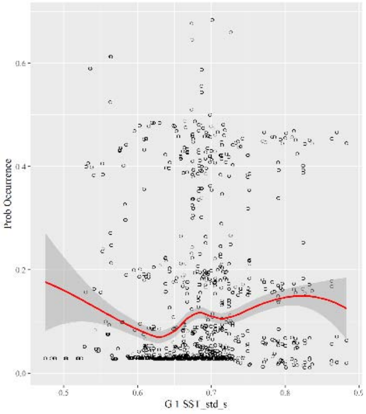
HISTORICAL	RCP	OVERLAP
	45m	45m
		
	45l	45l
		
	85m	85m
		
	85l	85l
		

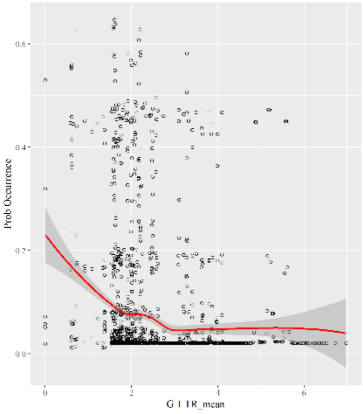
Figure S5.2. Comparison of the realized niche of the historical and RCPs periods and their overlap. The overlapping portion of two niches is showed in yellow, historical conditions in red and RCPs conditions in green.

Saccorhiza polyschides

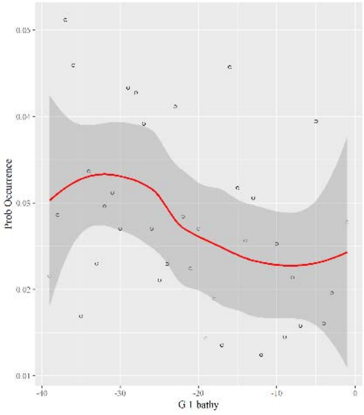
SST std summer



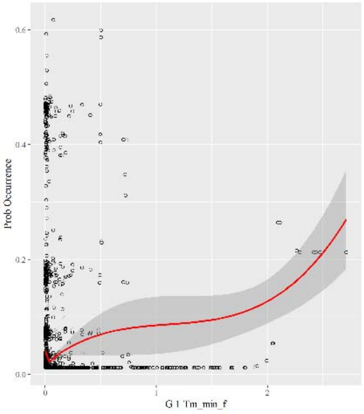
TR mean



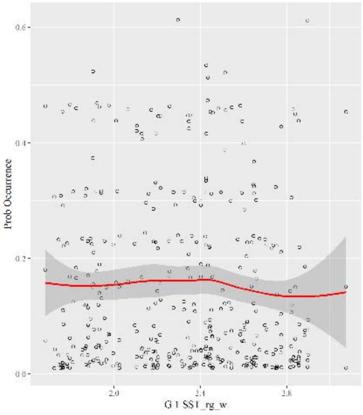
Bathymetry



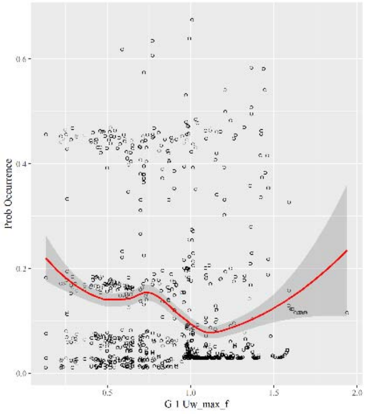
Tm minimum fall



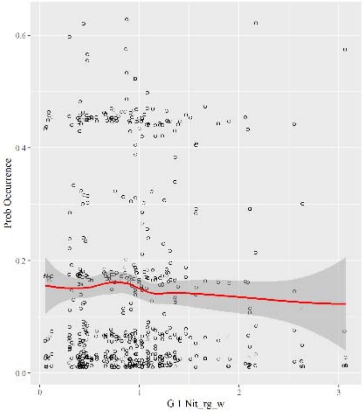
SST range winter



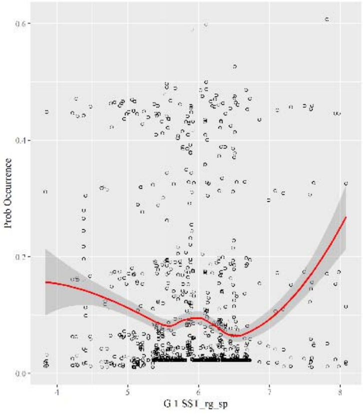
Uw maximum fall



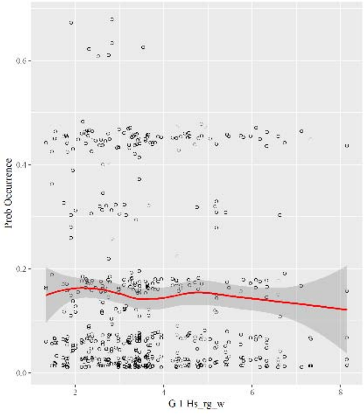
Nitrate range winter



SST range spring



Hs range winter



Continue on the next page

Saccorhiza polyschides

Kd minimum fall

pH range winter

Radiation range summer

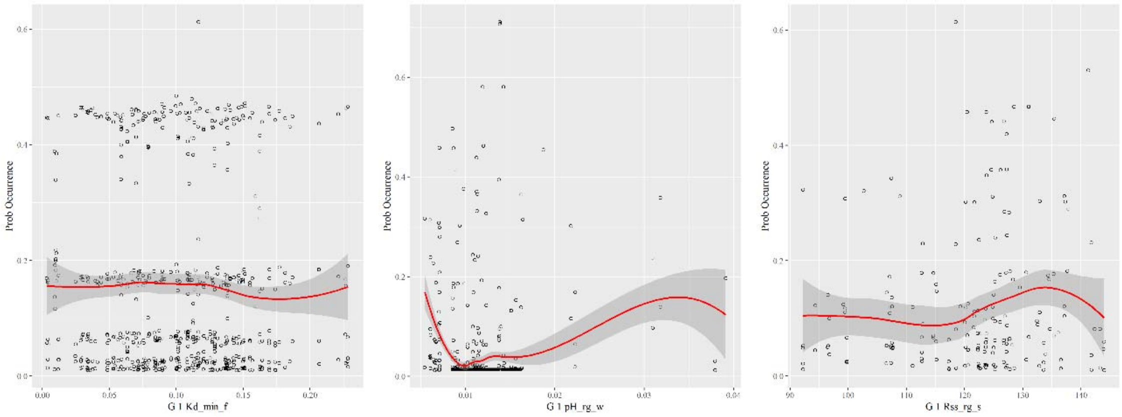
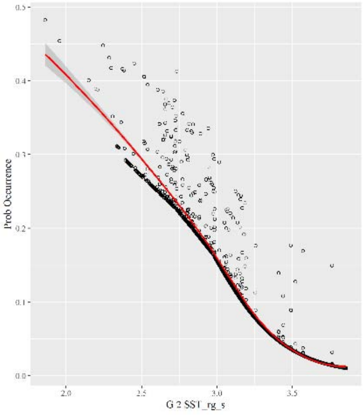
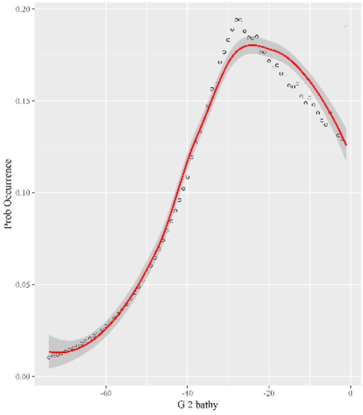
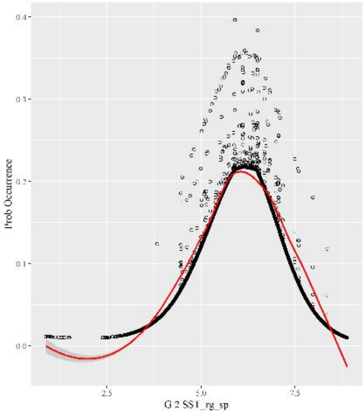
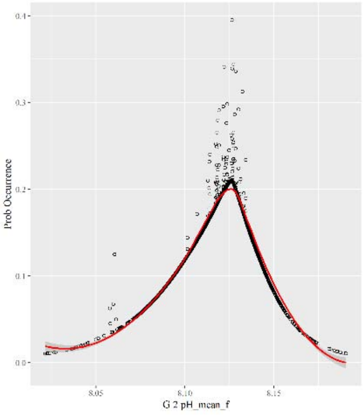
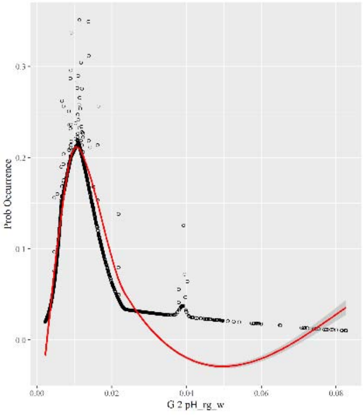
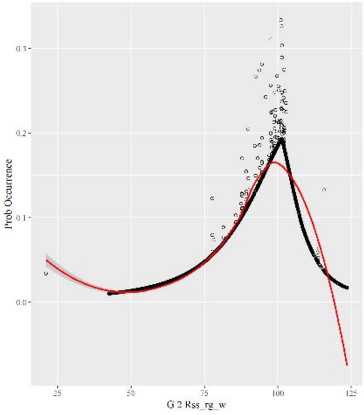
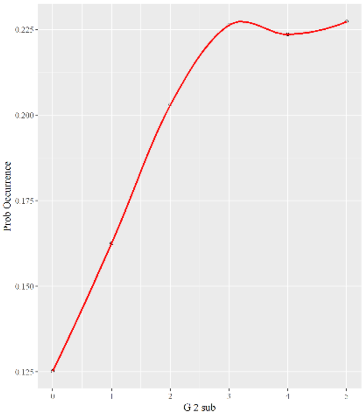
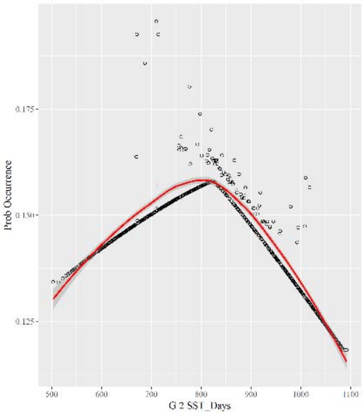
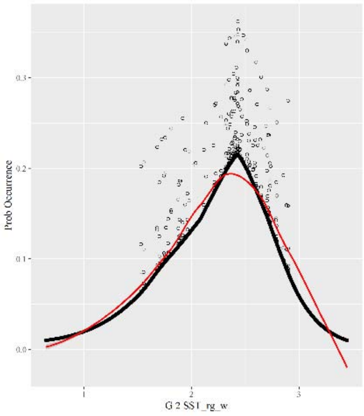


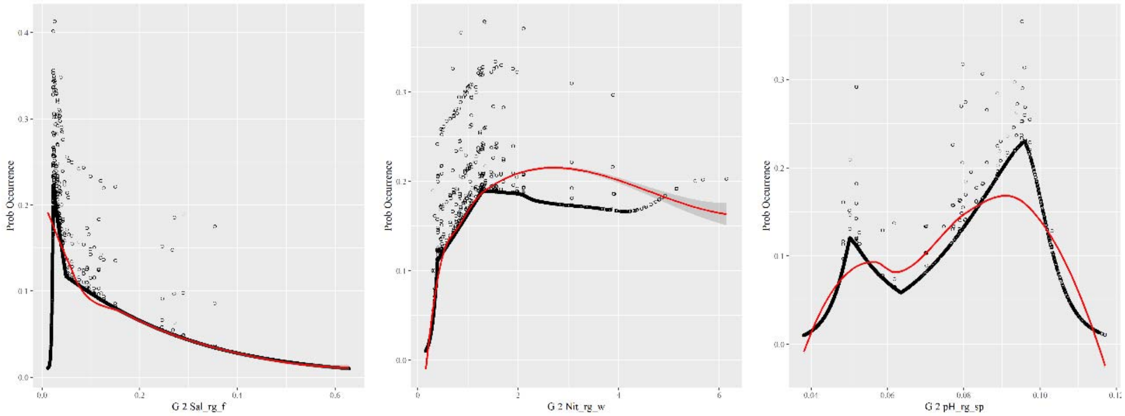
Figure S5.3. Dependence plots of *Saccorhiza polyschides*. In the y-axis the probability of occurrence is represented and in the x-axis the predictors represented, from upper-left to right in importance order. Black dots represent mesh points and their adjusted line is in red.

Gelidium spinosum		
SST range summer	Bathymetry	SST range spring
		
pH mean fall	pH range winter	Radiation range winter
		
Substrate	SST Days	SST range winter
		

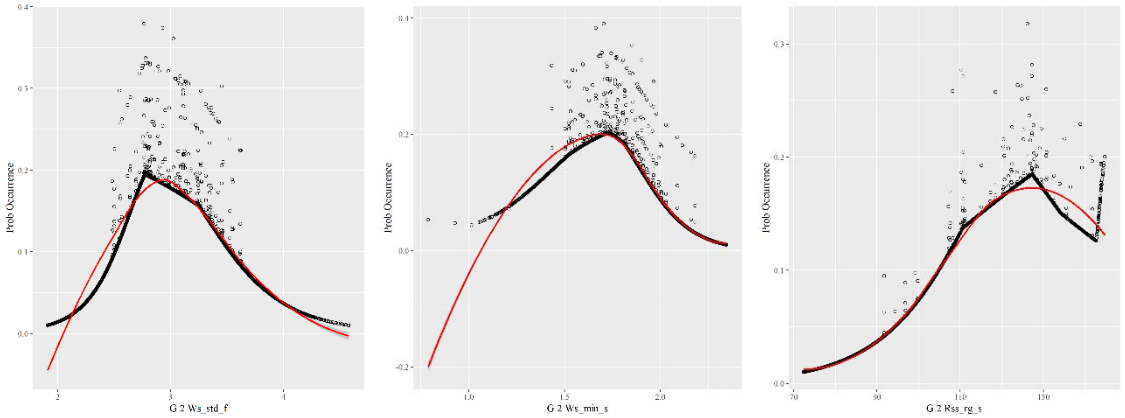
Continue on the next page

Gelidium spinosum

Salinity range fall | Nitrate range winter | pH range spring



Ws std fall | Ws minimum summer | Radiation range summer



Uw percentile 75

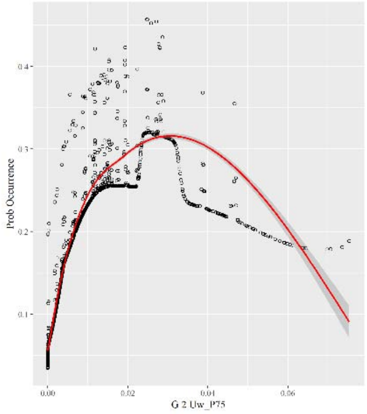
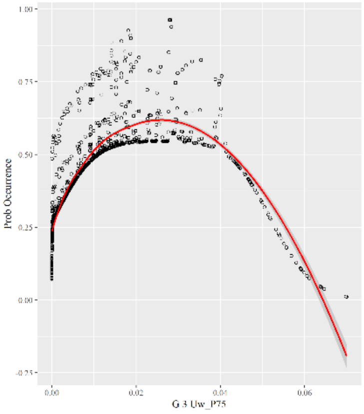


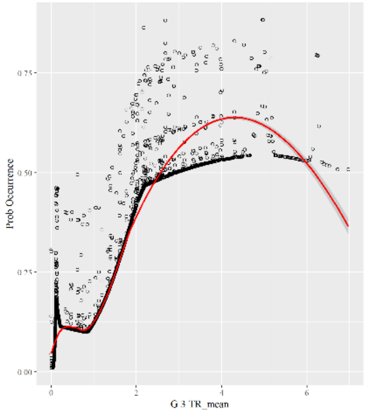
Figure S5.4. Dependence plots of *Gelidium spinosum*. In the y-axis the probability of occurrence is represented and in the x-axis the predictors represented, from upper-left to right in importance order. Black dots represent mesh points and their adjusted line is in red.

Sargassum muticum

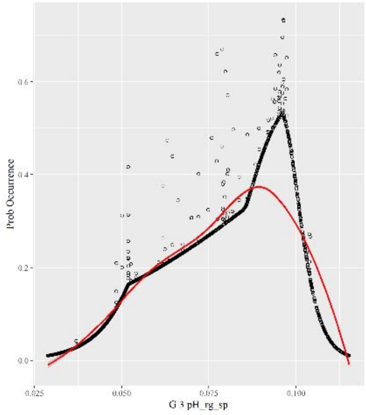
Uw percentile 75



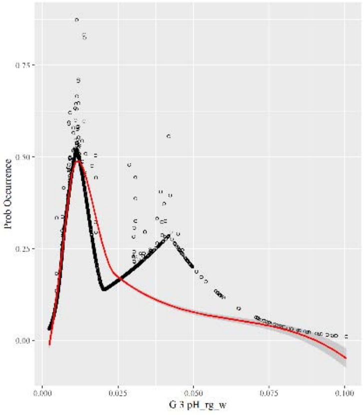
TR mean



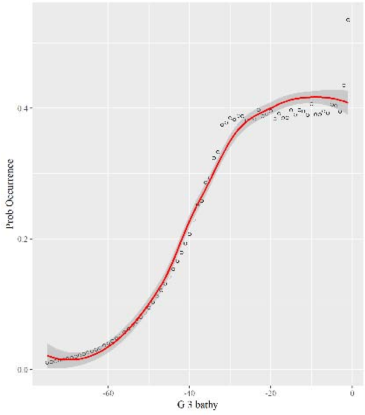
pH range spring



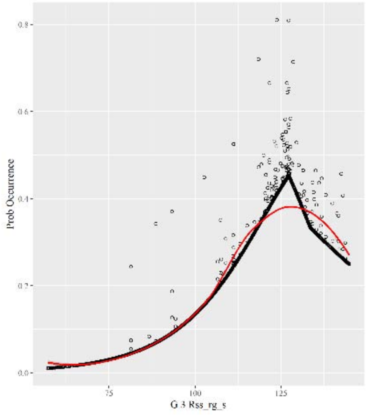
pH range winter



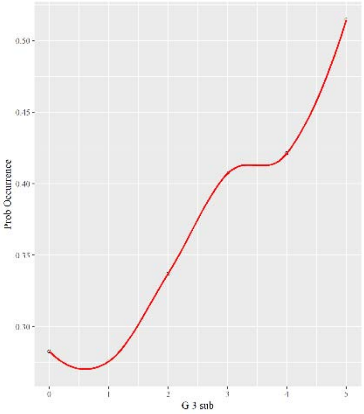
Bathymetry



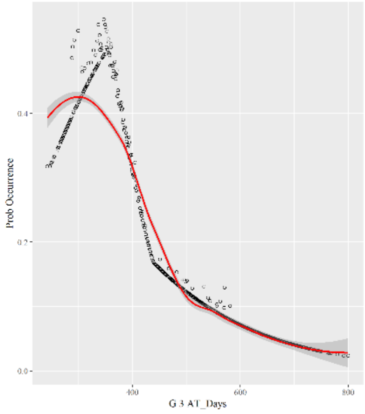
Radiation range summer



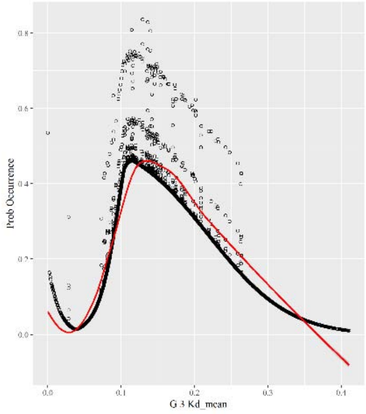
Substrate



AT Days



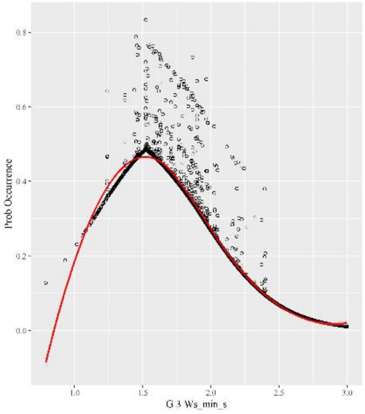
Kd mean



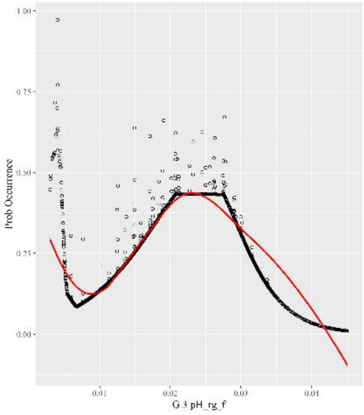
Continue on the next page

Sargassum muticum

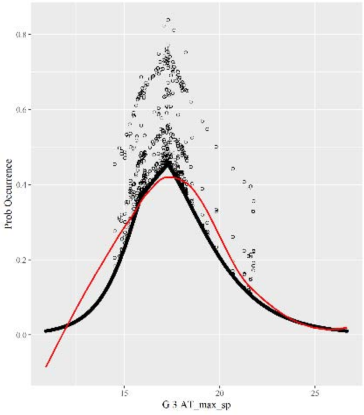
Ws minimum summer



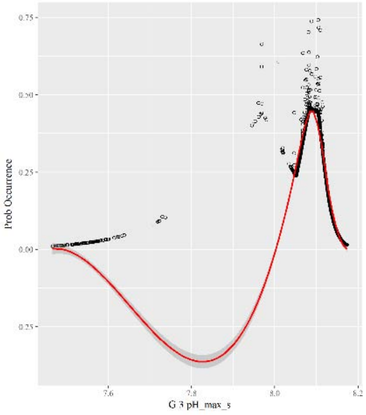
pH range fall



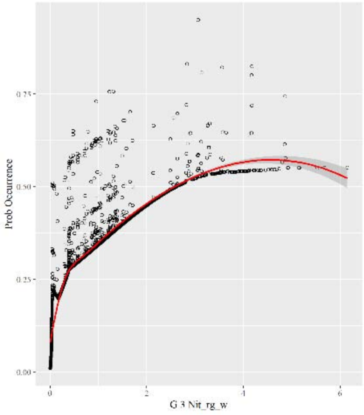
AT maximum spring



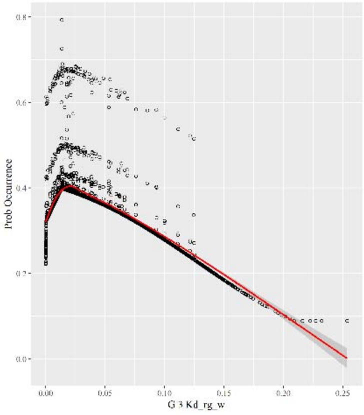
pH maximum summer



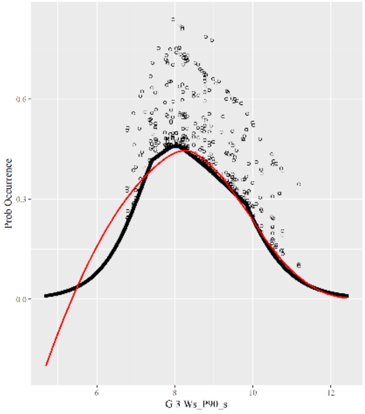
Nitrate range winter



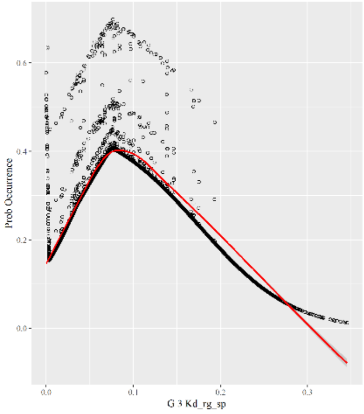
Kd range winter



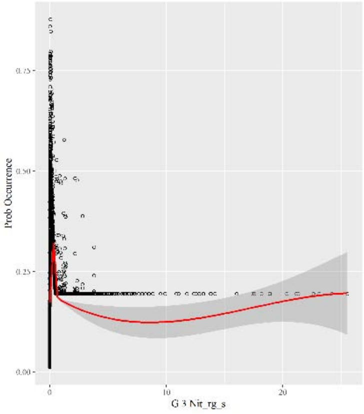
Ws percentile 90 summer



Kd range spring



Nitrate range summer



Continue on the next page

Sargassum muticum

Salinity std winter	Radiation range spring	SST Days

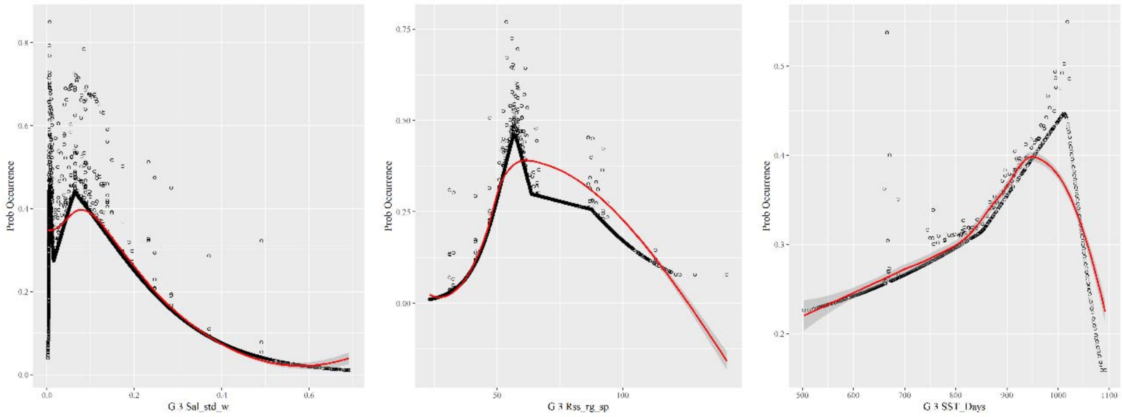
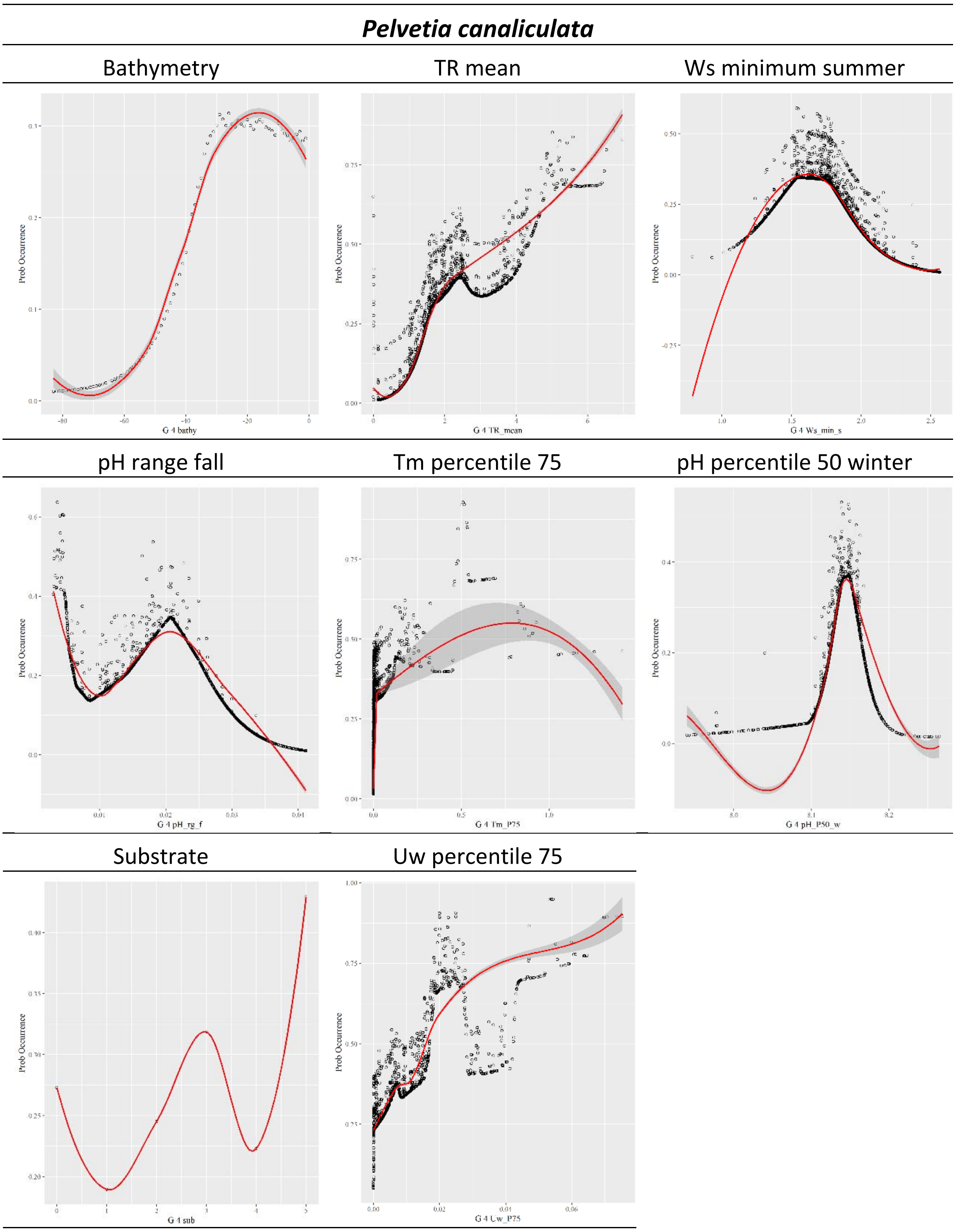


Figure S5.5. Dependence plots of *Sargassum muticum*. In the y-axis the probability of occurrence is represented and in the x-axis the predictors represented, from upper-left to right in importance order. Black dots represent mesh points and their adjusted line is in red.



Pelvetia canaliculata

Bathymetry

TR mean

Ws minimum summer

pH range fall

Tm percentile 75

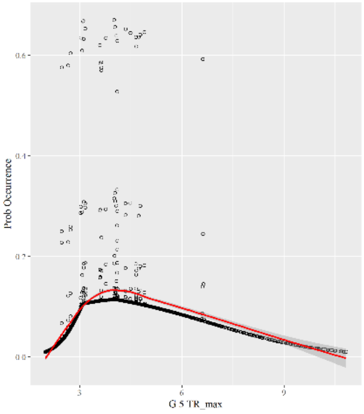
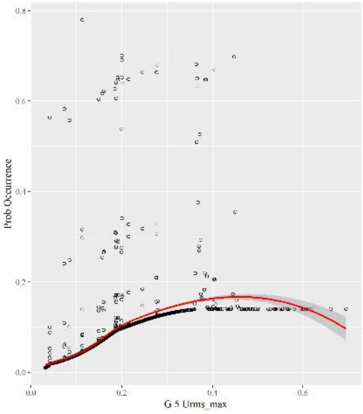
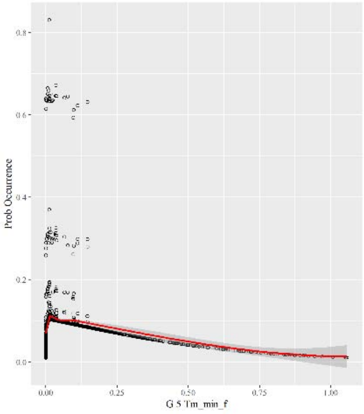
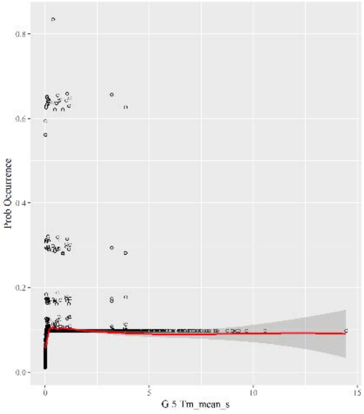
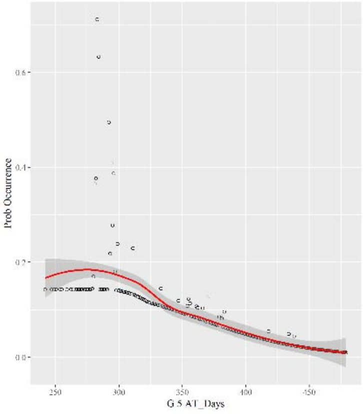
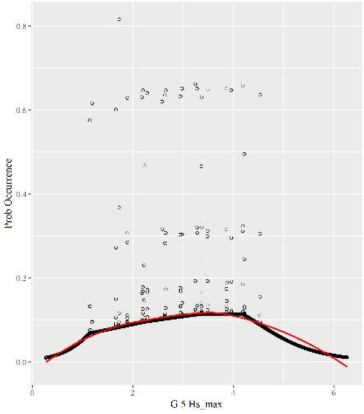
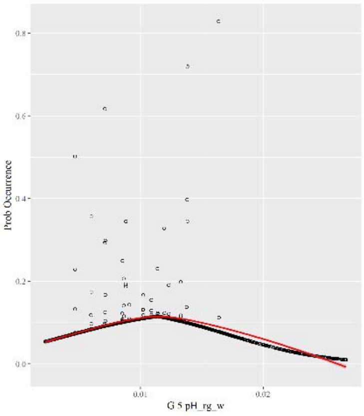
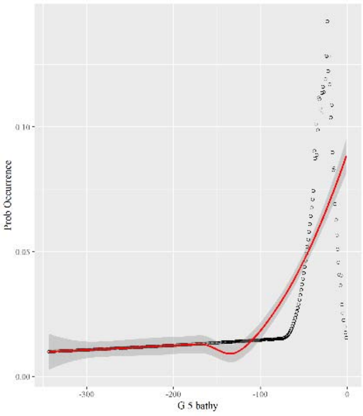
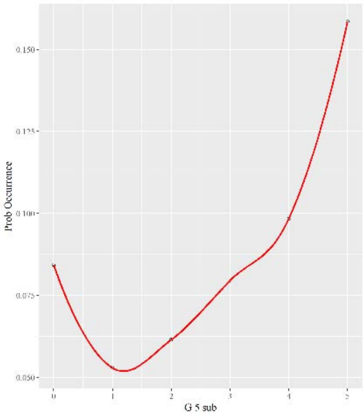
pH percentile 50 winter

Substrate

Uw percentile 75

Figure S5.6. Dependence plots of *Pelvetia canaliculata*. In the y-axis the probability of occurrence is represented and in the x-axis the predictors represented, from upper-left to right in importance order. Black dots represent mesh points and their adjusted line is in red.

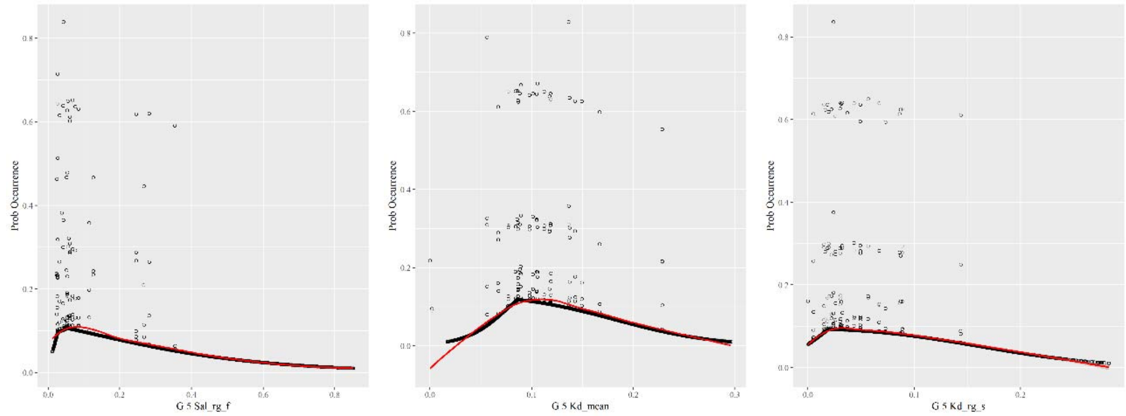
Cystoseira baccata

TR maximum	Urms maximum	Tm minimum fall
		
Tm mean summer	**AT Days**	**Hs maximum**
		
pH range winter	**Bathymetry**	**Substrate**
		

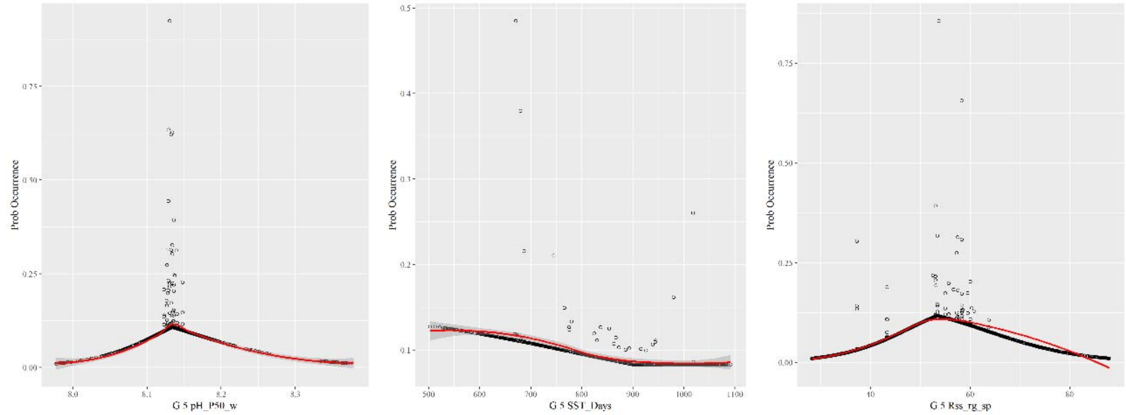
Continue on the next page

Cystoseira baccata

Salinity range fall Kd mean Kd range summer



pH percentile 50 winter SST Days Radiation range spring



Ws minimum spring

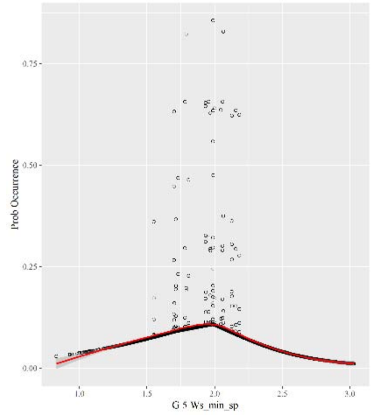


Figure S5.7. Dependence plots of *Cystoseira baccata*. In the y-axis the probability of occurrence is represented and in the x-axis the predictors represented, from upper-left to right are the following: AT_Days, bathy, Hs_max, Kd_mean, Kd_rg_s, pH_P50_w, pH_rg_w, Rss_rg_sp, Sal_rg_f, SST_Days, sub, Tm_mean_s, Tm_min_f, TR_max, Urms_max, Ws_min_sp. Black dots represent mesh points and their adjusted line is in red.

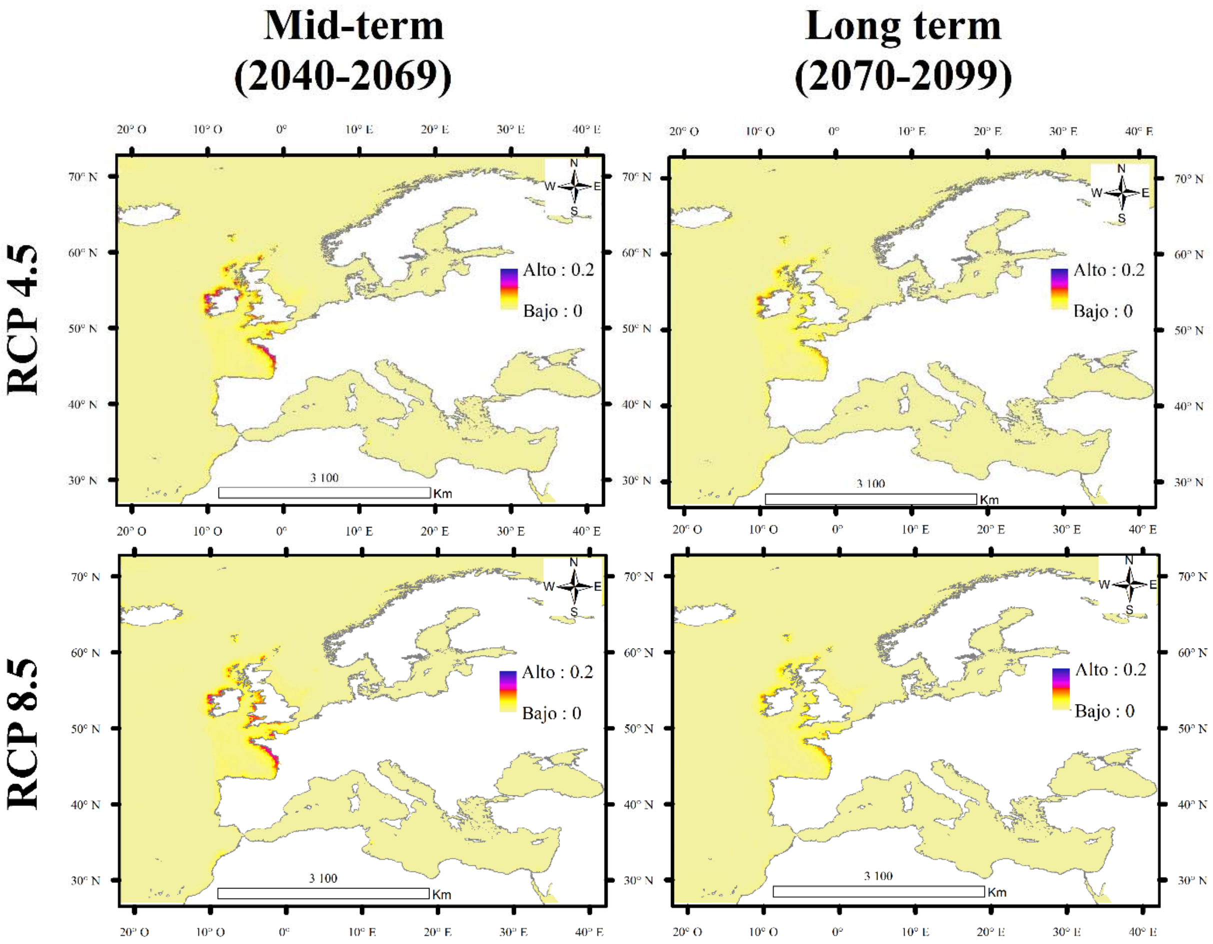


Figure S5.8. Probability of occurrence of *Saccorhiza polyschides* for the Representative Concentration Pathway (RCP) 4.5 and 8.5 in the mid-term (2040-2069) and the long term (2070-2099).

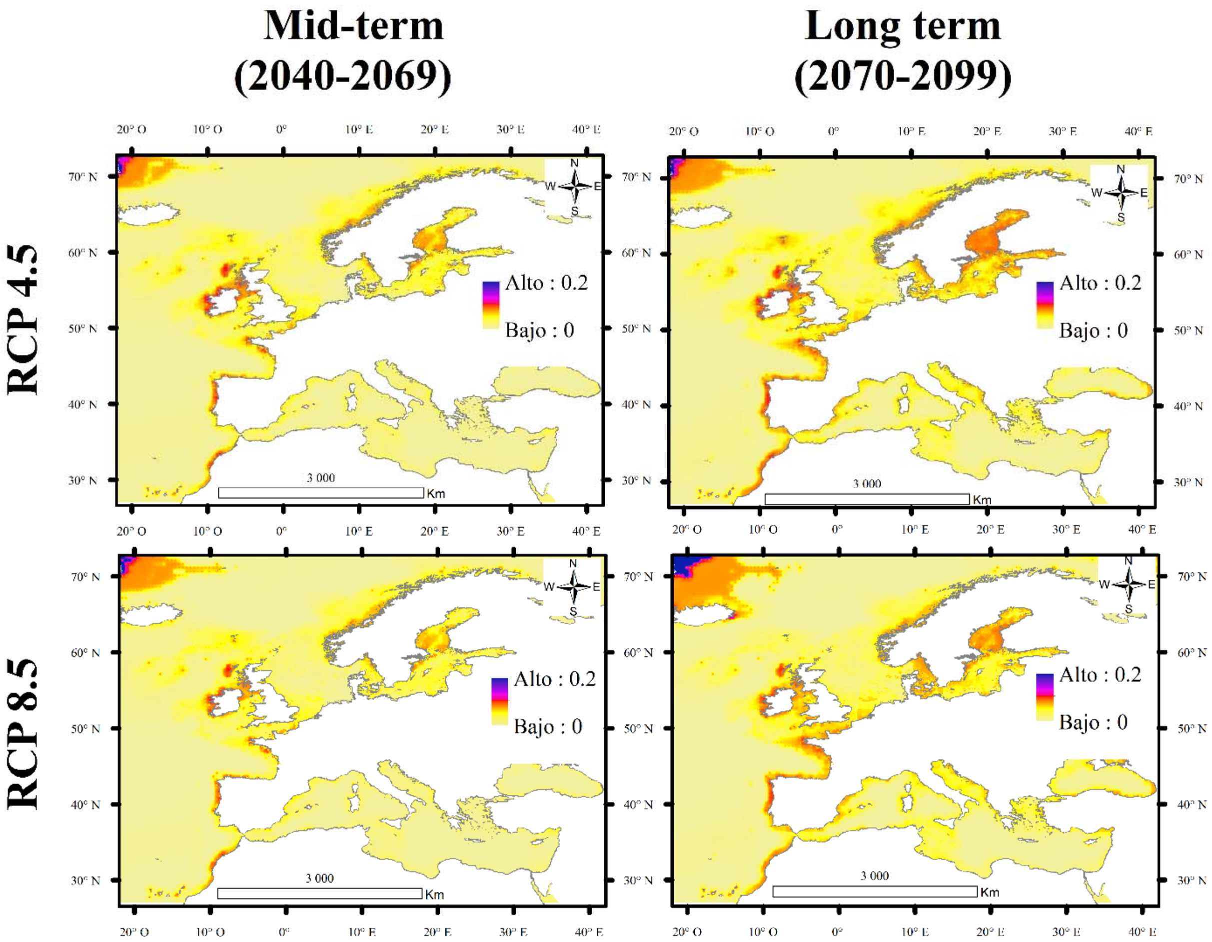


Figure S5.9. Probability of occurrence of *Gelidium spinosum* for the Representative Concentration Pathway (RCP) 4.5 and 8.5 in the mid-term (2040-2069) and the long term (2070-2099).

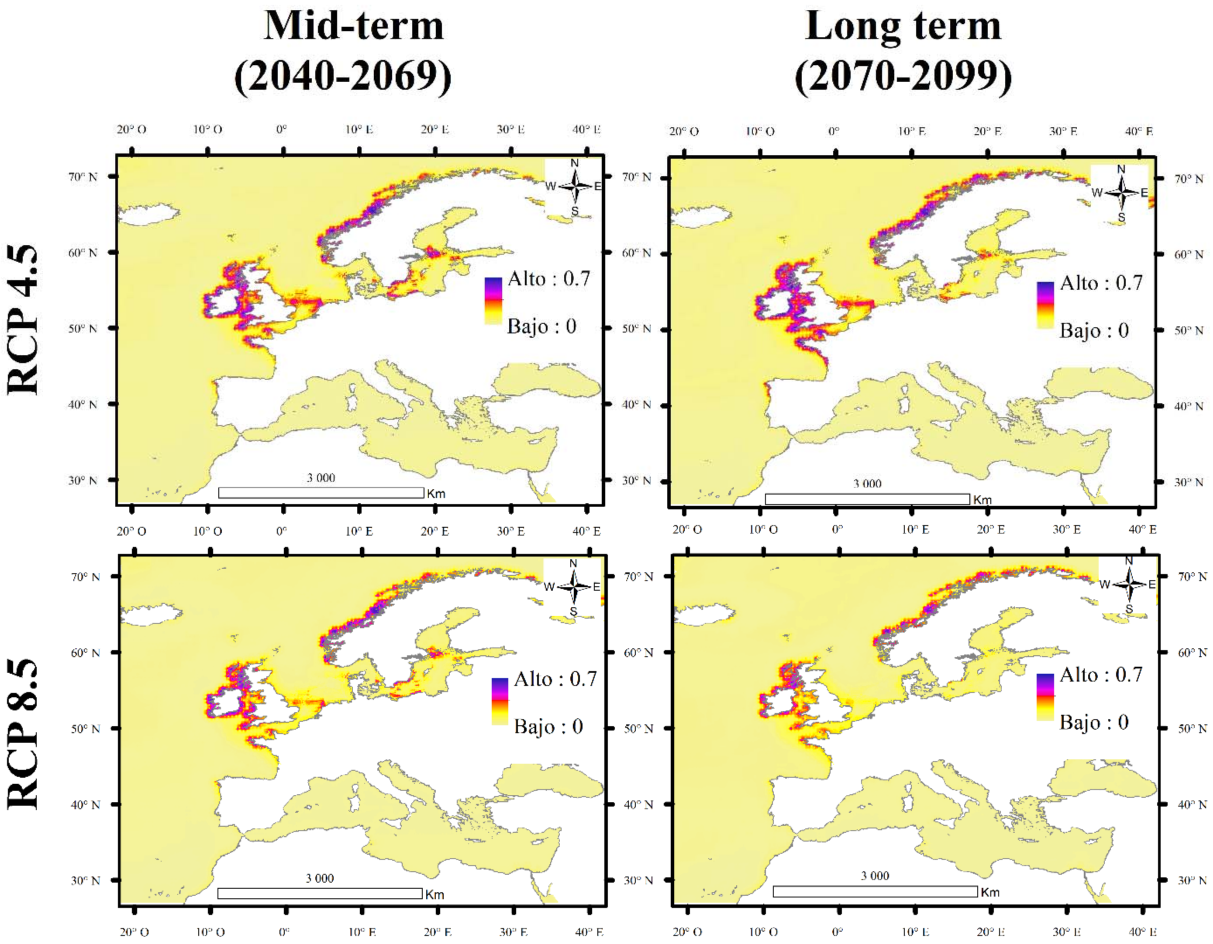


Figure S5.10. Probability of occurrence of *Sargassum muticum* for the Representative Concentration Pathway (RCP) 4.5 and 8.5 in the mid-term (2040-2069) and the long term (2070-2099).

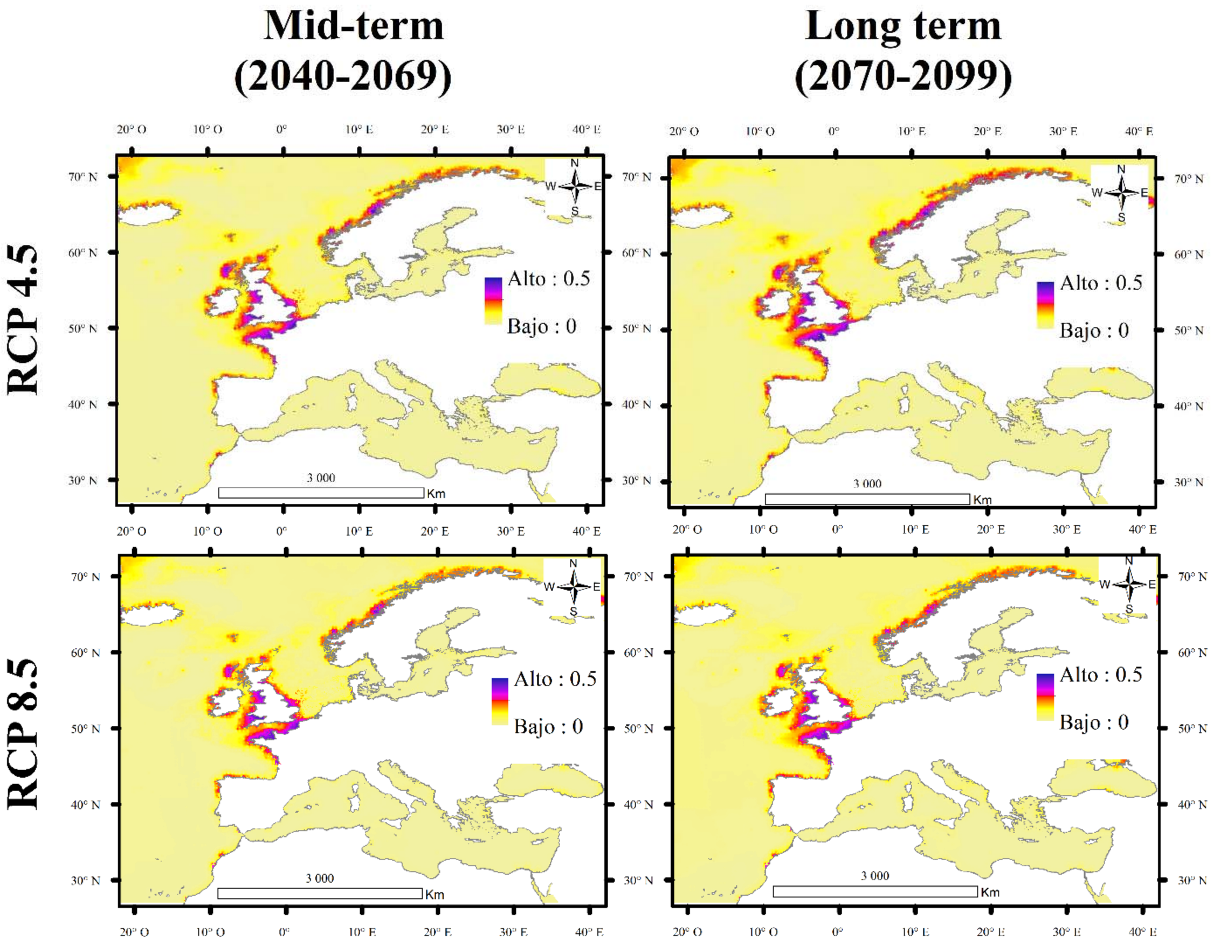


Figure S5.11. Probability of occurrence of *Pelvetia canaliculata* for the Representative Concentration Pathway (RCP) 4.5 and 8.5 in the mid-term (2040-2069) and the long term (2070-2099).

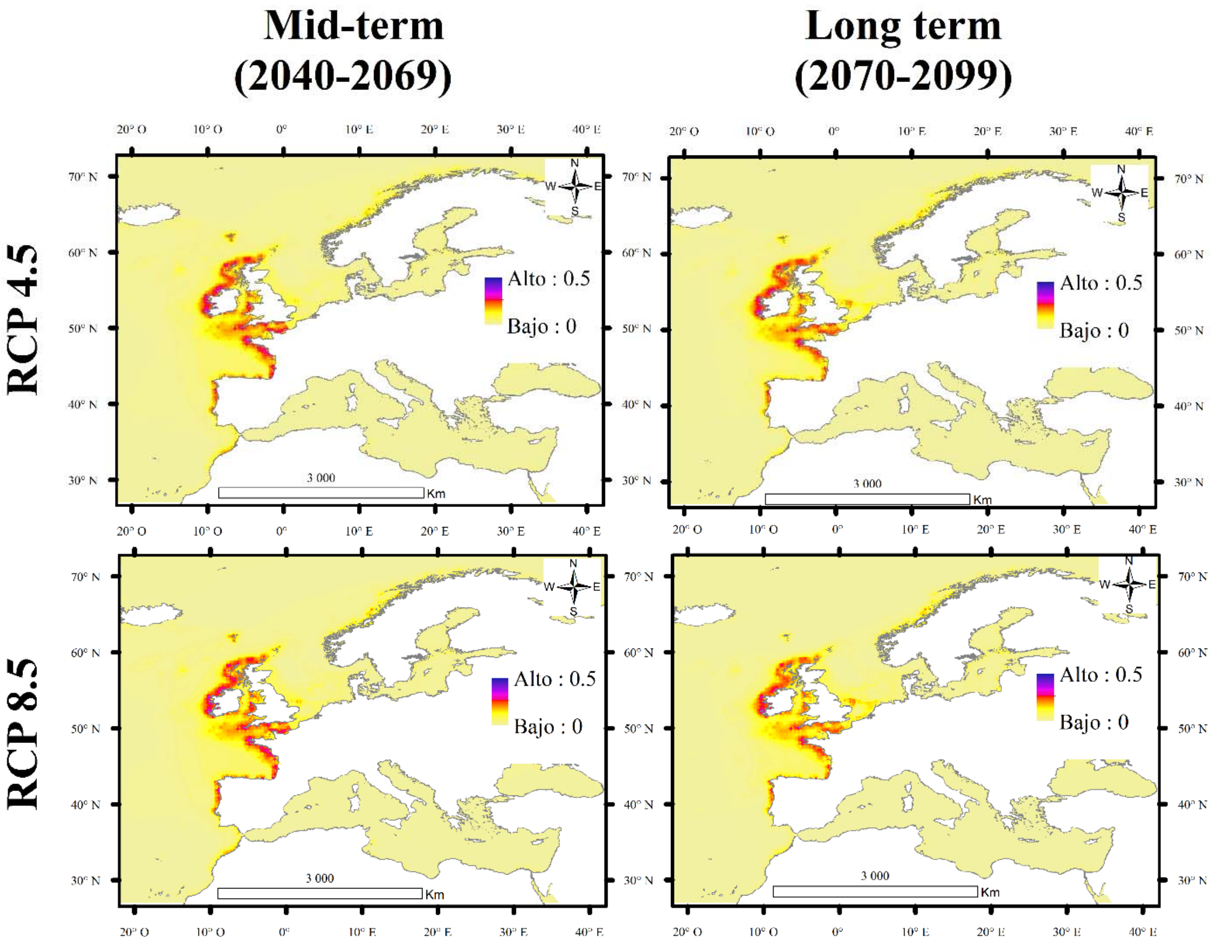
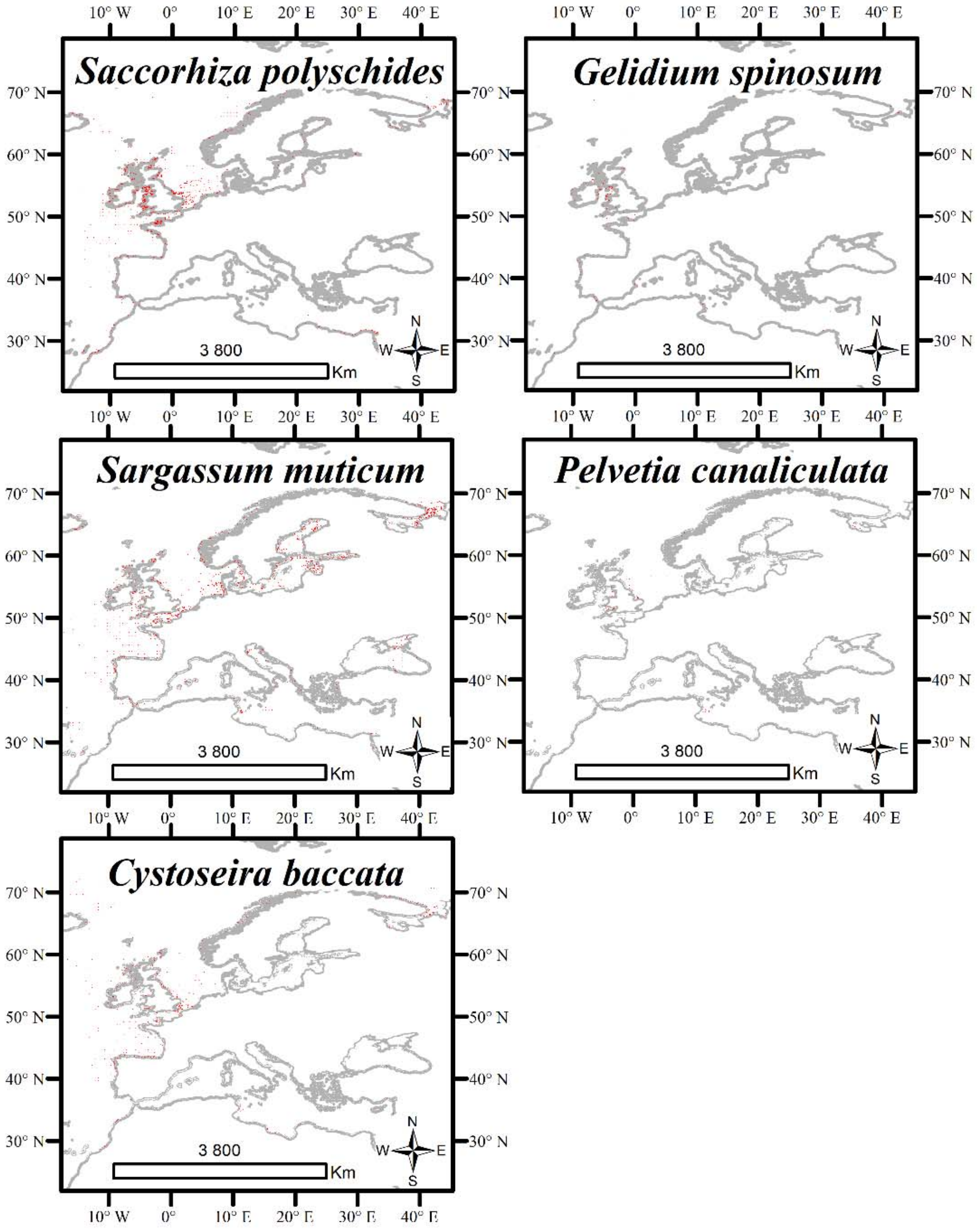


Figure S5.12. Probability of occurrence of *Cystoseira baccata* for the Representative Concentration Pathway (RCP) 4.5 and 8.5 in the mid-term (2040-2069) and the long term (2070-2099).



Saccorhiza polyschides

Gelidium spinosum

Sargassum muticum

Pelvetia canaliculata

Cystoseira baccata

Figure S5.13 Interquartile range (IQR) assessment for considered species. Red points represent outliers, areas where the climate is non-analogous, therefore predictions in those areas have to be considered carefully.

Table S5.1. Parameters considered for each species and their contribution (in percentage) in brackets.

Saccorhiza polyschides	*Gelidium spinosum*	*Sargassum muticum*	*Pelvetia canaliculata*	*Cystoseira baccata*
1. SST_std_s (26.03)	1. SST_rg_s (16.13)	1. Uw_P75 (17.39)	1. bathy (22.48)	1. TR_max (14)
2. TR_mean (22.29)	2. Bathy (13.45)	2. TR_mean (15)	2. TR_mean (21.16)	2. Urms_max (13.19)
3. Bathy (19.17)	3. SST_rg_sp (9.67)	3. pH_rg_sp (8.13)	3. Ws_min_s (11.47)	3. Tm_min_f (12.05)
4. Tm_min_f (5.74)	4. pH_mean_f (9.23)	4. pH_rg_w (5.51)	4. pH_rg_f (11.39)	4. Tm_mean_s (9.25)
5. SST_rg_w (4.37)	5. pH_rg_w (8.48)	5. bathy (5.25)	5. Tm_P75 (9.86)	5. AT_Days (8.32)
6. Uw_max_f (4.02)	6. Rss_rg_w (7.13)	6. Rss_rg_s (4.82)	6. pH_P50_w (8.49)	6. Hs_max (8.30)
7. Nit_rg_w (3.56)	7. sub (5.14)	7. sub (4.64)	7. sub (7.59)	7. pH_rg_w (7.64)
8. SST_rg_sp (3.20)	8. SST_Days (4.62)	8. AT_Days (4.01)	8. Uw_P75 (7.56)	8. bathy (5.56)
9. Hs_rg_w (3.17)	9. SST_rg_w (4.40)	9. Kd_mean (3.92)		9. sub (4.52)
10. Kd_min_f (3.04)	10. Sal_rg_f (3.86)	10. Ws_min_s (3.88)		10. Sal_rg_f (3.16)
11. pH_rg_w (2.78)	11. Nit_rg_w (3.84)	11. pH_rg_f (3.78)		11. Kd_mean (2.91)
12. Rss_rg_s (2.72)	12. pH_rg_sp (3.26)	12. AT_max_sp (3.09)		12. Kd_rg_s (2.78)
	13. Ws_std_f (3.19)	13. pH_max_s (3)		13. pH_P50_w (2.57)
	14. Ws_min_s (2.95)	14. Nit_rg_w (2.74)		14. SST_Days (2.54)
	15. Rss_rg_s (2.62)	15. Kd_rg_w (2.73)		15. Rss_rg_sp (1.65)
	16. Uw_P75 (2.02)	16. Ws_P90_s (2.48)		16. Ws_min_sp (1.55)
		17. Kd_rg_sp (2.19)		
		18. Nit_rg_s (2.18)		
		19. Sal_std_w (1.92)		
		20. Rss_rg_sp (1.79)		
		21. SST_Days (1.53)		

Physical variables: AT, air temperature; Bathy, bathymetry; Hs, significant wave height; Rss, radiation; Sal, salinity; SST, sea surface temperature; Sub, substrate; Tm, bottom shear stress; TR, tidal range; Uw, bottom orbital speed; Ws, wind speed; Days, number of days over a threshold; P, percentile; std, standard deviation; rg, range; min, minimum; max, maximum; w, winter; sp, spring; s, summer; f, fall.