



*Escuela Técnica Superior de Ingenieros de Caminos,
Canales y Puertos.*

UNIVERSIDAD DE CANTABRIA



Elaboración de modelos de distribución de especies bentónicas estuarinas

Trabajo realizado por:

Grégoire Dufermont

Dirigido:

Araceli Puente Trueba

Cristina Galván Arbeiza

Titulación:

**Máster Universitario en Gestión Integrada
de Sistemas Hídricos**

Santander, Septiembre de 2022

TRABAJO FIN DE MASTER

RESUMEN

Introducción y objetivos

Los estuarios son importantes centros de biodiversidad debido a su naturaleza, que favorece una gran diversidad de hábitats de transición entre el agua dulce de los ríos y el agua salada del mar. La diferencia de granulometría de los sedimentos, el gradiente de salinidad, el tiempo de emersión e inmersión, la protección o no de las corrientes y el oleaje, todos estos factores forman hábitats muy diversos que albergan especies adaptadas a estos cambios, que se agrupan según su nicho ecológico en las diferentes zonas del estuario. En los estuarios Cantábricos destacan tres comunidades de macroinvertebrados bentónicos: la comunidad de *S. plana* en la zona oligomesohalina, la comunidad de *N. hombergii* en la zona polihalina y la comunidad de *A. alba* en la zona euhalina. Las actividades humanas también influyen en las condiciones del fondo. Por lo tanto, los macroinvertebrados bentónicos se verán afectados por todas estas condiciones ambientales y perturbaciones. Modelos de distribución de especies como GAM, GLM, MAXENT y RF pueden ayudarnos a conocer mejor la distribución de estas especies en estuarios como la Bahía de Santander o las Marismas de Santoña.

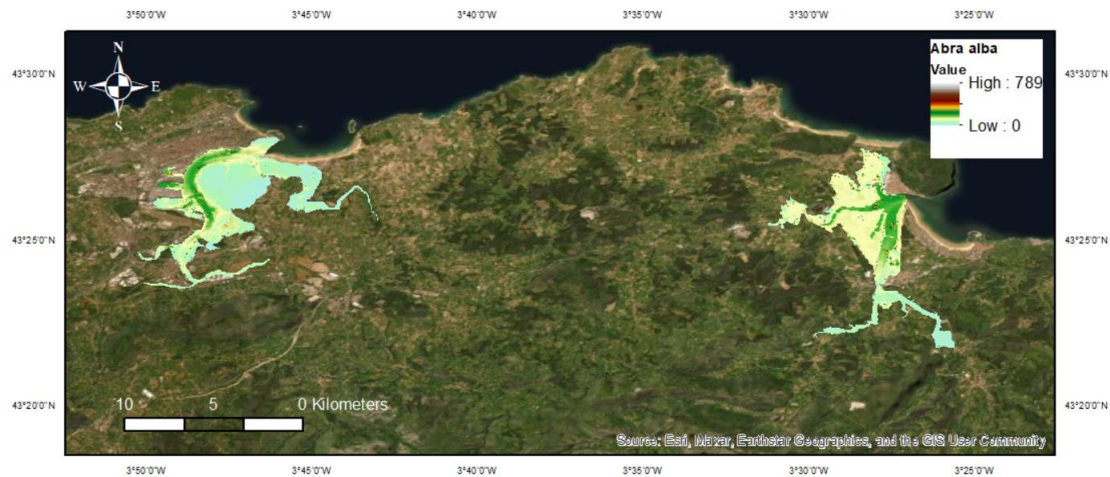
El objetivo de este estudio es, por tanto, modelizar y analizar la distribución de especies representativas de macroinvertebrados bentónicos en los estuarios de la Bahía de Santander y las Marismas de Santoña.

Materiales y métodos

Para lograr este objetivo se ha utilizado el modelo MAXENT en la fase previa para obtener la contribución de las variables en la realización del modelo y una primera evaluación del modelo gracias a la curva ROC y al AUC. Este modelo utilizará datos ambientales así como datos de geolocalización de las especies previamente filtrados y procesados. A continuación, BIOMOD modelará los mapas de distribución de las especies utilizando 4 algoritmos de modelización. Finalmente, los mapas serán procesados en ArcGis para facilitar el análisis. Los modelos se compararán utilizando la métrica AUC, pero también TSS y KAPPA.

Resultados

Los mapas de distribución muestran que las especies de la comunidad de *A. alba* se concentran en zonas de gran profundidad, como la zona portuaria de la bahía de Santander y a lo largo de los canales de las marismas de Santoña, que también se corresponde con la zona euhalina. Se aprecia una pequeña diferencia en la distribución de *M. palmata* causada por la diferencia de adaptación a los sedimentos de la zona, y por tanto una mayor extensión al sur de la bahía de Santander.



Mapa de distribución de *A. alba*

En el caso de la comunidad de *N. hombergii*, las especies son bastante diferentes en su distribución porque no tienen la misma resistencia a la salinidad, por lo que *A. tenuis* y *C. edule* se concentrarán en las zonas internas de los estuarios y cercanas a las desembocaduras de los ríos, mientras que *N. reticulatus* se encontrará en casi toda el estuario. *N. hombergii* estará limitada por la naturaleza de los sedimentos de la zona.

Por último, en el caso de la comunidad de *S. plana*, la especie también se encontrará en la zona interior y en las desembocaduras de los ríos, pero con mayor probabilidad en las desembocaduras donde la salinidad es menor. En particular, *H. diversicolor* se concentra claramente en las costas de los estuarios, ya que prefiere las zonas en las que la frecuencia de la sequía es alta, que es el caso de las zonas del intermareal más alto.

El análisis de comparación de modelos mostró que el ensamblado de los mejores modelos de distribución obtenidos con diferentes algoritmos funciona ligeramente mejor que los modelos individuales.

Discusión y conclusiones

Los resultados mostraron tres variables clave en la distribución de las especies. La salinidad delimitará tres zonas de distribución según la alta salinidad en la zona euhalina, la alta a baja salinidad en la zona de transición polihalina y la baja salinidad en la zona oligo-mesohalina. La batimetría también será clave, ya que podemos ver que algunas especies prefieren profundidades más altas, como la comunidad de *A. alba*, pero otras más bajas, donde la influencia de la marea es más débil, lo que limita la intensidad de las inundaciones y las sequías. Por último, como las especies están en el suelo, la naturaleza del sedimento es importante para entender su distribución. Los modelos han dado buenos resultados, pero se pueden mejorar dando más importancia a la adaptación de los puntos de medición en el estuario a los objetivos del estudio que se quiere realizar y tomando el modelo más adecuado para los medios de que se dispone. Por último, para determinar mejor la disposición de las especies, los futuros estudios de los modelos deberían tener en cuenta las variables antrópicas y bióticas, con el fin de tener en cuenta todos los fenómenos que inciden en su distribución.

SUMMARY

Introduction and objectives

Estuaries are important centres of biodiversity due to their nature, which favours a great diversity of transitional habitats between freshwater from rivers and saltwater from the sea. The difference in sediment granulometry, the salinity gradient, the time of emersion and immersion, the protection or not of currents and waves, all these factors form very diverse habitats that are home to species adapted to these changes, which are grouped according to their ecological niche in restricted areas of the estuary. Three benthic macroinvertebrate communities stand out in the Cantabrian estuaries: the *S. plana* community in the oligo-mesohaline zone, the *N. hombergii* community in the polyhaline zone and the *A. alba* community in the euhaline zone. Human activities also influence bottom conditions. Therefore, benthic macroinvertebrates will be affected by all these environmental conditions and disturbances. Species distribution models such as GAM, GLM, MAXENT and RF can help us to better understand the distribution of these species in bays such as Santander or Santoña.

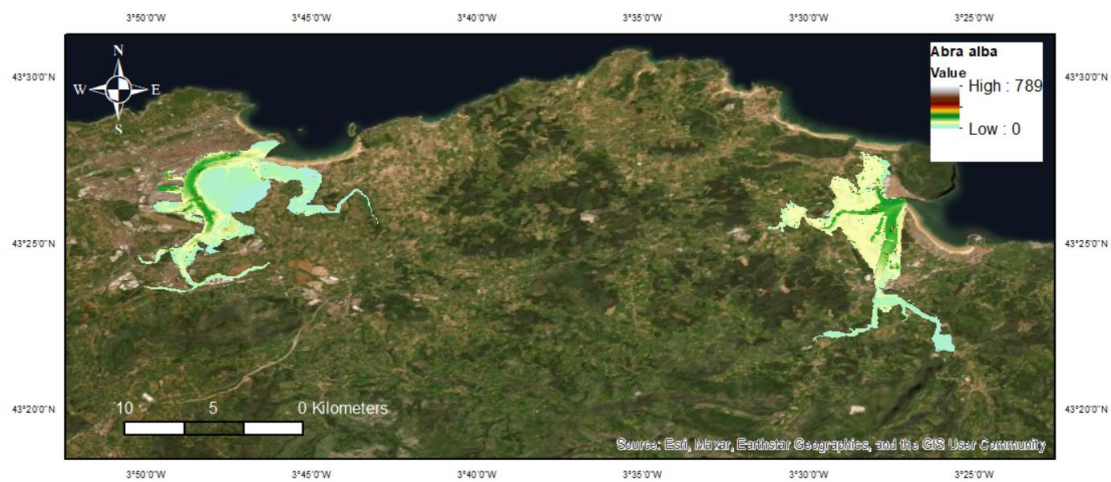
The aim of this study is, therefore, to model and analyse the distribution of representative species of benthic macroinvertebrates in the bays of Santander and Santoña.

Materials and methods

To achieve this objective, the MAXENT model has been used in the previous phase to obtain the contribution of the variables in the realisation of the model and a first evaluation of the model thanks to the ROC curve and the AUC. This model will use environmental data from different studies as well as species geolocation data previously filtered and processed. BIOMOD will then model the species distribution maps using 4 modelling algorithms. Finally, the maps will be processed in ArcGis to facilitate the analysis. The models will be compared using the AUC metric, but also TSS and KAPPA.

Results

The distribution maps show that the species of the *A. alba* community are concentrated in areas of great depth, such as the harbour area of Santander Bay and along the channels of the Santoña marshes, which also corresponds to the euhaline zone. There is a small difference in distribution caused by the difference in adaptation to the sediments of the area, and therefore a greater extension to the south of Santander Bay for *M. palmata*.



Distribution map of A. alba

In the case of the *N. hombergii* community, the species are quite different in their distribution because they do not have the same resistance to salinity, so that *A. tenuis* and *C. edule* will be concentrated in the inner areas of the estuaries with an extension towards the river mouths, while *N. reticulatus* will be found almost everywhere in the bay. *N. hombergii* will be limited by the nature of the sediments in the area.

Finally, in the case of the *S. plana* community, the species will also be found in the inland area and at river mouths, but more likely at river mouths where salinity is lower. In particular, *H. diversicolor* is clearly concentrated on estuarine shores, as it prefers areas where the frequency of drought is high, which is the case in shallow coastal areas.

The model comparison analysis showed that the ensemble performs slightly better than the others.

Discussion and conclusions

The results therefore showed 3 key variables in the distribution of species. Salinity will delimit 3 distribution zones according to high salinity in the euhaline zone, high to low salinity in the polyhaline transition zone and low salinity in the oligo-mesohaline zone. Bathymetry will also be key, as we can see that some species prefer higher depths, such as the *A. alba* community, but others shallower, where the tidal influence is weaker, limiting the intensity of floods and droughts. Finally, as the species are in the soil, the nature of the sediment is important to understand their distribution. The models have given good results, but they can be improved by giving more importance to adapting the measurement points in the estuary to the studies to be carried out and by taking the most appropriate model for the means available. Finally, to better determine the disposition of the species, future model studies should take into account anthropic and biotic variables, in order to take into account all the phenomena that affect their distribution.

Índice de contenido

1.	Introducción	1
2.	Objetivos	4
3.	Materiales y metodología	4
3.1	Área de estudio	5
3.2	Datos biológicos	6
3.3	Datos ambientales	9
3.4	Modelos	10
4.	Resultados	13
4.1	Optimización de la configuración	13
4.2	Comparación de modelos	15
4.3	Análisis de las curvas de respuesta	17
4.4	Análisis de los mapas de distribución	20
5.	Discusión	25
6.	Conclusiones	28
7.	Referencias	30
	Anexo	34
	Datos ambientales	34

Índice de figuras

Figura 1 :	Diagrama resumen de la metodología utilizada	5
Figura 2 :	Localización de la bahía de Santander (izquierda) y las marismas de Santoña (derecha)	5
Figura 3 :	fotos de las especies del estudio, A. alba(A), A. tenuis(B), C. edule(C), C. carinata(D), H. diversicolor(E), M. palmata(F), N. reticulatus(G), N. hombergii(H) y S. plana(I) (www.idscaro.net, v3.boldsystems.org, fr.wikipedia.org, www.aphotomarine.com)	8
Figura 4 :	Resultados de coeficientes KAPPA, AUC y TSS para cada modelo y cada especie	16

Figura 5 : Curvas de respuestas a las variables.....	20
Figura 6 : Mapa de distribución de la especie <i>A. alba</i>	22
Figura 7 : Mapa de distribución de la especie <i>M. palmata</i>	22
Figura 8 : Mapa de distribución de la especie <i>N. reticulatus</i>	23
Figura 9 : Mapa de distribución de la especie <i>A. tenuis</i>	23
Figura 10 : Mapa de distribución de la especie <i>C. edule</i>	23
Figura 11 : Mapa de distribución de la especie <i>N. hombergii</i>	24
Figura 12 : Mapa de distribución de la especie <i>C. carinata</i>	24
Figura 13 : Mapa de distribución de la especie <i>H. diversicolor</i>	25
Figura 14 : Mapa de distribución de la especie <i>S. plana</i>	25

Índice de tablas

Tabla 1 : Abundancia relativa de las especies seleccionadas siguiente la zona de salinidad del estuario	7
Tabla 2 : Resultados de la contribución (%) de las variables ambientales en los modelos de distribución de cada especie construidos con MAXENT. En verde se indican las variables seleccionadas.....	14
Figura 3 : Resultado de la prueba AUC comparando las diferentes configuraciones consideradas entre el multiplicador de regularización y la complejidad.....	15

1. Introducción

Los estuarios se localizan en el límite entre las aguas interiores y el mar, siendo una zona de transición entre el agua dulce y el agua salada y, por tanto, entre las comunidades que albergan. Debido a su naturaleza, albergan una amplia biodiversidad de hábitats y especies y son capaces de proporcionar bienes y servicios ecosistémicos esenciales para el bienestar de la sociedad.

Los ecosistemas estuarinos se caracterizan por la presencia de marcados gradientes ambientales que generan un rico mosaico de hábitats. Uno de los gradientes más importantes es el de la salinidad. La salinidad dentro del estuario muestra un patrón espacial que responde a valores bajos próximos al del agua dulce en las zonas próximas al río y valores altos próximos al agua de mar en la zona de la bocana. Además de este gradiente espacial, los estuarios están sujetos a variaciones temporales de la salinidad, que responden a variaciones interanuales en los aportes fluviales, estacionales e, incluso, diarias asociadas a la marea. Por lo tanto, la mezcla de agua dulce y salada da lugar a condiciones particulares de salinidad que determinan la distribución de las especies en función de su grado de tolerancia o adaptación. Dependiendo de la importancia de la influencia del mar o del río en las diferentes zonas del estuario, la salinidad será mayor o menor, lo que afectará a la importancia de las especies de agua dulce, salada o estuarinas en las comunidades biológicas que colonizan estos espacios. La salinidad es una variable clave para entender la distribución de las especies en el estuario (Galván et al. 2016).

Las condiciones bióticas y abióticas de los estuarios varían considerablemente según el lugar del estuario en el que estamos. Los estuarios son zonas costeras protegidas de la corriente y del oleaje de la marea, lo que favorecerá a ciertas especies fijas o de fondo al permitirles desarrollarse en aguas tranquilas, limitando las fuerzas de arrastre.

La marea modificará la altura del agua en la ría, lo que determinará qué zonas se inundan y cuáles quedan emergidas con el tiempo, y por tanto cambiará la duración de las inundaciones y de la sequía según el lugar de la ría en el que se encuentre. Además,

dependiendo de la amplitud de las mareas, la intensidad de estas inundaciones y sequías variará en intensidad.

La batimetría también influirá en estas condiciones de duración, frecuencia e intensidad, ya que las zonas más altas se verán menos afectadas por la marea, que sólo llegará al final de la subida durante un breve periodo de tiempo antes de volver a bajar. La batimetría es en sí misma un límite para la mayoría de las especies que no pueden desarrollarse a grandes profundidades.

Como las especies bentónicas se fijan al suelo, se entierran o se desplazan por él, la naturaleza del sedimento será importante. Un suelo de grano pequeño, como el fango o la arena, facilitará el enterramiento de las especies, mientras que un suelo de grava no lo hará. Las especies se desarrollarán de forma diferente según su forma de vida en los distintos sedimentos (Galván et al., 2021).

Las especies de macroinvertebrados bentónicos que colonizan los fondos estuarinos pueden agruparse en comunidades ecológicas. En concreto, en los estuarios de Cantabria se reconoce la existencia de 3 comunidades biológicas (Puente et al., 2008). Por un lado, encontramos la comunidad denominada *Scrobicularia plana*. Por otro lado, tenemos la comunidad de *Abra alba*. Por último, se ha definido una tercera comunidad que está dominada por *Nephtys hombergii* y que es muy rica, pero que sólo está presente en tres estuarios de Cantabria, entre ellos la Bahía de Santander y las Marismas de Santoña.

Estas especies se desplazan muy poco y, por lo tanto, están expuestas a diferentes presiones antrópicas y climáticas. Los estuarios son zonas costeras con gran importancia social y económica en torno a los cuales se concentran grandes densidades de población y diversas actividades económicas que ejercen una fuerte presión sobre el ecosistema. Entre las diferentes fuentes de presión antropogénica, las que mayor impacto pueden generar sobre las comunidades de macroinvertebrados bentónicos son la contaminación (Montero et al. 2013 y Birch et al. 2020), los dragados, etc. Las actividades de dragado y navegación perturbarán el fondo, desplazando individuos bentónicos y sedimentos y perturbarán los hábitats cercanos por la llegada de sedimentos que pueden no ser de la misma naturaleza y que perturbarán el hábitat de ciertas especies no adaptadas. Por lo tanto, estas perturbaciones repercutirán en estas comunidades y en su estado.

Muchos estudios se han centrado en los macroinvertebrados bentónicos porque la Directiva Marco del Agua los considera en la clasificación del estado ecológico de las aguas costeras y de transición. Si bien esto puede parecer fácil en aguas costeras, es mucho más difícil en aguas de transición, ya que es una tarea compleja con un alto grado de incertidumbre debido a las condiciones altamente variables (Puente et al. 2008).

Por ello, los macroinvertebrados son un tema importante en la búsqueda de indicadores (Puente et al. 2008) y de métodos de clasificación estuarina. En efecto, son relativamente sedentarios, lo que les dificulta evitar las presiones desplazándose a otros lugares, también tienen un ciclo de vida relativamente largo y, por tanto, están expuestos a las presiones durante este tiempo y, por último, son comunidades ricas con diferentes niveles de tolerancia. Sin embargo, su estudio es complejo debido a la naturaleza del entorno con sus condiciones altamente variables. La modelización de la distribución de los macroinvertebrados puede ayudar a comprender su distribución y a clasificar el medio, pero también a evaluar los impactos del cambio climático en las comunidades.

Para entender la distribución de una especie, o su evolución futura, los modelos de distribución de especies (MDE) son una herramienta muy interesante. Estos modelos se basan en la búsqueda de relaciones entre las condiciones ambientales y los datos de localización de las especies para determinar las condiciones necesarias para el establecimiento y mantenimiento de la especie estudiada (Araújo et al. 2012) y, en consecuencia, determinar su distribución en un área determinada.

Muchas aproximaciones, como son los métodos de regresión como el Modelo Lineal Generalizado (GLM) (McCullagh et al., 1989) y el Modelo Aditivo Generalizado (GAM) (Hastie et al., 1990), los métodos de aprendizaje automático como Random Forest (RF) (Breiman, 2001), y los algoritmos como MAXENT (Phillips et al. 2006), permiten calcular esta distribución en diferentes condiciones ambientales, ya sean climáticas o ambientales. También puede utilizarse para crear curvas de respuesta y para probar las interacciones entre la especie y la variable estudiada (Thuiller et al. 2009). Permiten calcular estas relaciones y, por tanto, la probabilidad de aparición de las especies en distintos puntos de una zona de estudio. Este tipo de modelo se utiliza en diversos estudios, como el análisis de la distribución de una especie, el estudio de los efectos del

cambio climático en dicha distribución o los efectos de la variabilidad de determinadas condiciones ambientales.

Por ello, en este estudio se modelizará la distribución de especies representativas de macroinvertebrados bentónicos para analizar la distribución en el estuario y el papel de cada variable en esta distribución, pero también para comparar los modelos utilizados y determinar qué modelo es el más adecuado para modelar esta distribución.

2. Objetivos

El objetivo general de este estudio es elaborar modelos de distribución de especies bentónicas estuarinas en los estuarios cantábricos. Este objetivo se desglosa en 3 objetivos específicos:

- Comparar la capacidad predictiva de los modelos de distribución de las especies obtenidos con diferentes algoritmos.
- Analizar el papel de las variables ambientales en la distribución de las especies y sus rangos de tolerancia.
- Analizar conjuntamente la distribución de las especies en el entorno y compararla con la distribución de las comunidades biológicas características de los estuarios de Cantabria.

3. Materiales y metodología

Para alcanzar estos objetivos, se ha aplicado una metodología que consta de 3 pasos: (1) Tratamiento de los datos; (2) Optimización de la configuración de los modelos; (3) Modelización de la distribución del hábitat potencial de las especies de interés (Figura 1).

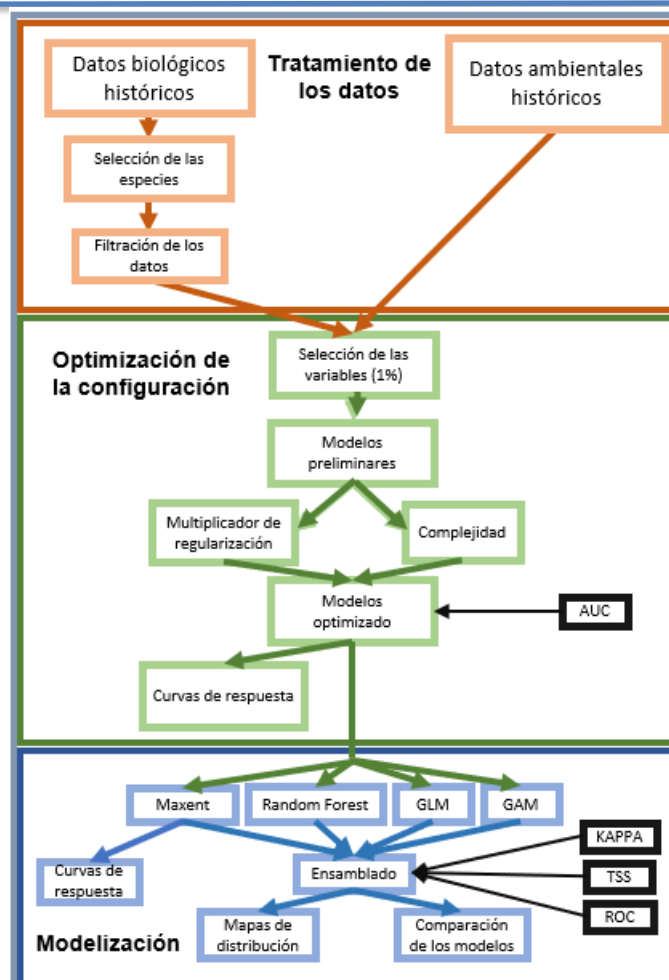


Figura 1 : Diagrama resumen de la metodología utilizada

3.1 Área de estudio

El estudio se centrará en la Bahía de Santander y en las Marismas de Santoña, situadas en la costa norte de España, en Cantabria (Figura 2).



Figura 2 : Localización de la bahía de Santander (izquierda) y de las marismas de Santoña (derecha)

La Bahía de Santander es el mayor estuario del norte de España con una extensión de 22,42 km², una longitud de 9 km y una anchura de 5 km. Está dividida en varias rías: Rías de Solía, Ría de Tijero, Ría de Boo y Ría del Cubas. El principal aporte de agua dulce al estuario es a través del río Miera, con un caudal medio de 8.2 m³/s. Este estuario tiene un alto valor ecológico, social y económico y forma parte de la Zona de Especial Conservación “Dunas del Puntal y Estuario del Miera”.

Las Marismas de Santoña es otro de los estuarios más grandes del norte de España, con una superficie de 19 km². Su principal suministro de agua dulce procede del río Asón, con una aportación media de 16 m³/s. Es un estuario con gran valor ecológico que está protegido por diferentes figuras como son la de Parque Natural, Zona Especial de Conservación y sitio RAMSAR.

3.2 Datos biológicos

Los datos biológicos se obtuvieron en las campañas de verano realizadas en 2009 y 2010 para la Red de Calidad del Litoral de Cantabria. En ambas campañas se muestrearon un total de 31 estaciones (20 en la Bahía de Santander y 11 en las Marismas de Santoña), en cada una de las cuales se identificaron las especies de macroinvertebrados presentes, así como su abundancia y biomasa.

Para la realización de este estudio, se seleccionaron las especies representativas de las tres comunidades de macroinvertebrados presentes en los estuarios de Cantabria (Tabla 1):

- La primera comunidad es la de *Scrobicularia plana*, que es característica de las zonas internas de los estuarios en fondos arenosos y aguas bien oxigenadas. Por lo tanto, se encuentran en la zona oligo-mesohalina, donde la salinidad es baja. En esta comunidad encontramos *Cyathura carinata*, *Scrobicularia plana* y *H. diversicolor*.
- Luego, tenemos la comunidad de *Abra alba* en la zona submareal o en las zonas que muy a menudo están sumergidas de los tramos euhalinos. En esta comunidad encontramos *Abra alba* y *Melinna palmata*.

- Por último, entre estas dos comunidades, tenemos la comunidad dominada por *Nephtys hombergii* que se encuentra en las zonas intermareales polihalinas. Esta comunidad es muy rica pero sólo está presente en 3 estuarios, entre ellos la Bahía de Santander y las marismas de Santoña. Acompañando a *N. hombergii* están *A. tenuis*, *C. edule* y *N. reticulatus*.

Euhalino (<i>Abra alba</i>)		Polihalino (<i>Nephtys hombergii</i>)		Oligo-mesohalino (<i>Scrobicularia plana</i>)	
Especie	Abundancia relativa (%)	Especie	Abundancia relativa (%)	Especie	Abundancia relativa (%)
<i>Melinna palmata</i>	7.1	<i>Nephtys hombergii</i>	12.7	<i>Hediste diversicolor</i>	29
<i>Abra alba</i>	3.1	<i>Abra tenuis</i>	8.4	<i>Scrobicularia plana</i>	24.9
		<i>Cerastoderma edule</i>	5.4	<i>Cyathura carinata</i>	15.3

Tabla 1 : Abundancia relativa de las especies seleccionadas siguiente la zona de salinidad del estuario

Presentación de las especies del estudio

Abra Alba (Figura 3.A) es un molusco bivalvo de la familia Semelidae. Se alimenta de sedimentos mediante largos sifones extensibles (Degraer 2006). Se adapta a todo tipo de sedimentos, pero se encuentra sobre todo en sedimentos finos con un tamaño de grano de entre 50 y 250 μm y un contenido de fango del 10-50%. Es una especie costera común hasta 30 km de la costa (Holtmann 1996)

Al igual que *Abra alba*, *Abra tenuis* (Figura 3.B) es un molusco bivalvo de la familia Semelidae. Vive en sustratos blandos (Montagu, 1803).

Cerastoderma edule (Figura 3.C) es un molusco bivalvo de la familia Cardiidae. Se alimenta de fitoplancton, zooplancton y partículas orgánicas (Fish, 1996). Vive en todo tipo de sedimentos en zonas protegidas de las olas y las corrientes.

Cyathura carinata (Figura 3.D) es un crustáceo isópodo de la familia Anthuridae. Se encuentra en o sobre el lodo de los estuarios y ríos costeros (Naylor, 1972).

Hediste diversicolor (Figura 3.E) es un gusano marino de la familia Nereididae. Es muy tolerante a los cambios ambientales como la temperatura, la salinidad, la oxigenación y la contaminación. Vive en sustratos fangosos a arenosos. Se alimenta de plancton, diatomeas y materia orgánica fragmentada, que puede ser necromasa o excrementos (Campbell et al., 1986).

Melinna palmata (Figura 3.F) es un gusano segmentado de la familia Melinnidae. Son biodegradables. Los individuos pueden alcanzar los 50 mm. La reproducción es iterópara. Son animales sésiles (Grube, 1870).

Nassarius reticulatus (Figura 3.G) es un molusco gasterópodo de la familia Nassariidae. Se alimenta de orangistas muertos. Se encuentra en costas arenosas que no superan los 20 m de profundidad (Hayward et al., 1998).

Nephtys hombergii (Figura 3.H) es un gusano marino poliqueto de la familia Nephtyidae. Es un depredador de moluscos, crustáceos y poliquetos juveniles, pero también de carroña, diatomeas y detritus. Se encuentra cerca de la costa en sedimentos finos con un tamaño de grano inferior a 50 µm (Fish, 1996).

Scrobicularia plana (Figura 3.I) es un molusco bivalvo de la familia Semelidae. Es una especie que se alimenta por filtración. Se puede encontrar bajo la arena o el barro hasta 20 cm de profundidad (da Costa, 1778).

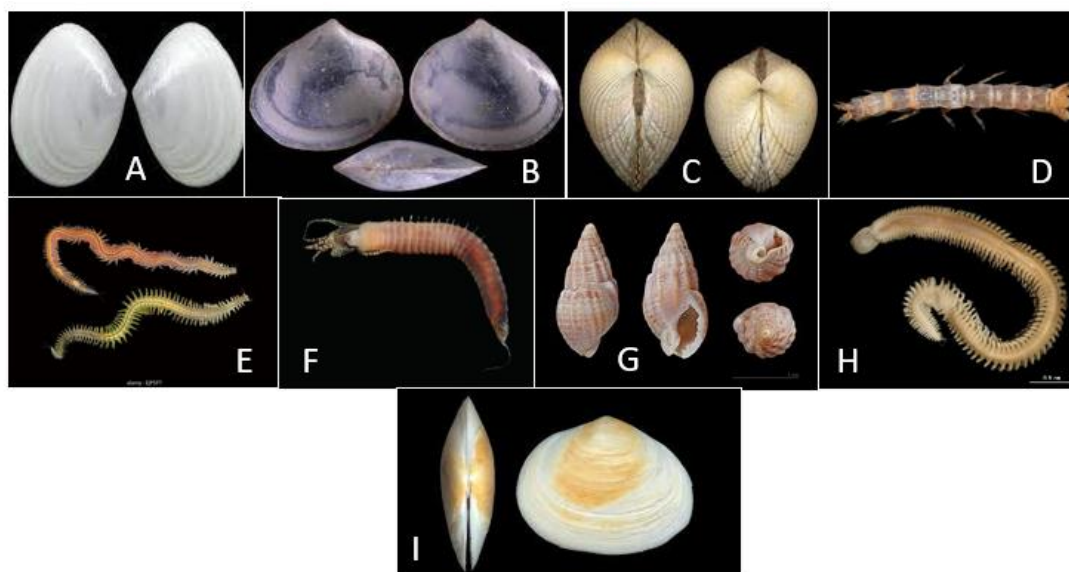


Figura 3 : Fotos de las especies del estudio, *A. alba*(A), *A. tenuis*(B), *C. edule*(C), *C. carinata*(D), *H. diversicolor*(E), *M. palmata*(F), *N. reticulatus*(G), *N. hombergii*(H) y *S. plana*(I) (www.idscaro.net, v3.boldsystems.org, fr.wikipedia.org, www.aphotomarine.com)

3.3 Datos ambientales

En primer lugar, se seleccionaron las variables más importantes de los hábitats bentónicos del estuario: el tipo de sedimento, la batimetría, la altura de columna de agua, la duración y la frecuencia de las inundaciones y las sequías, la salinidad media y el rango de salinidad, el tiempo de emersión y, por último, la fuerza de arrastre en fondo. En segundo lugar, para la caracterización de la variabilidad espacial y/o temporal de estas variables en los dos estuarios objeto de estudio se recopiló información generada en diferentes proyectos y estudios de acuerdo con la metodología desarrollada por Galván et al. (2016, 2021).

Los datos de salinidad fueron calculados por Galván et al (2016). En primer lugar, se creó una base de datos de los aportes fluviales y de la marea. Luego, se seleccionó de los datos históricos un subconjunto de series representativas a corto plazo del caudal del río y del nivel de las mareas. A continuación, estos subconjuntos se combinaron para modelar numéricamente la salinidad. Por último, se reconstruyeron las series de salinidad a largo plazo a partir de los datos del paso anterior utilizando el método análogo.

Para los sedimentos, Galván et al. (2021) agruparon los sedimentos en 3 tipos (arena, fango y grava) y caracterizaron su variabilidad espacial dentro del estuario utilizando datos puntuales recogidos en campo y la técnica de interpolación geoestadística de cokriging (Verfaillie et al., 2006; Jerosch, 2012), en la cual se incorporó la batimetría como co-variable (Galván et al., 2021). Con base en la batimetría se diferenciaron dentro del estuario las zonas intermareales de las submareales, de acuerdo con el nivel de agua más bajo registrado. La duración de la inundación y la altura de la columna de agua se calcularon a partir de la modelización hidrodinámica del nivel del agua utilizando una aproximación equivalente a la usada para la salinidad (Galván et al., 2016).

La caracterización ambiental de todas las variables consideradas se realizó con una resolución espacial de 25x25 metros cubriendo toda la superficie de ambos estuarios.

3.4 Modelos

3.4.1 Optimización de la configuración de los modelos de distribución de especies

La optimización de los parámetros de configuración de los modelos se llevó a cabo con el algoritmo MAXENT a través de la herramienta MAXENT versión 3.4.4 (Phillips et al. 2006). Se seleccionó este algoritmo para realizar el ajuste porque es un modelo ampliamente utilizado por la comunidad científica, es eficiente y no requiere mucho tiempo y esfuerzo de cálculo. No obstante, es necesario ajustar algunos de los parámetros del modelo para obtener resultados más precisos (Phillips et al. 2006). Los parámetros que tienen mayor importancia en la precisión de los resultados son la selección de variables, la complejidad del modelo, la regularización y la eliminación del sesgo de muestreo espacial (Elith et al. 2011, Merow et al. 2013).

Una vez seleccionadas las especies de estudio, se filtraron los datos de localización para eliminar los puntos duplicados y reducir el sesgo espacial de muestreo. Para ello, se utilizó el paquete spThin de R (Aiello-Lammens et al., 2015), que reúne un conjunto de funciones que permiten eliminar algunos puntos de localización de las especies en los que se ha tomado un mayor número de muestras para evitar una representación excesiva en el sistema que pueda causar un sesgo. Se elimina un mínimo de puntos para mantener un máximo de información.

Para seleccionar las variables que mejoran la capacidad predictiva del modelo, se realizó una primera simulación para cada especie con MAXENT incluyendo todas las variables ambientales caracterizadas previamente. Como resultado, se obtuvo la contribución de cada variable en el cálculo de la distribución. Las variables más importantes varían según la especie y muestran cuáles son las más limitantes en el análisis de la distribución. Se seleccionaron aquellas variables con una tasa de contribución superior al 1% (Tanner et al. 2017).

Para aplicar MAXENT, hay que seleccionar otro parámetro, la complejidad o "feature". Esto varía entre varios tipos de características dependiendo del número de

ocurrencias de la especie (Phillips et al. 2008). En el caso del estudio, cada especie tiene entre 10 y 20 puntos de geolocalización, por lo que las “feature” disponibles son “linear”, “category”, “hinge” et “quadratic”. Para comprobar la complejidad, se consideró que la complejidad por defecto tiene en cuenta “linear”, “hinge”, “quadratic” y “product”, mientras que la complejidad modificada elimina el “producto”. Un análisis AUC comparó la fiabilidad del modelo resultante si se utiliza el caso por defecto o una selección particular de características. Del mismo modo, se seleccionó el valor del multiplicador de regularización. Este multiplicador se probó con los siguientes valores 0,75, 1, 2, 5, 10 o 15 (Warren et al. 2011, Shcheglovitova et al. 2013). Los valores más bajos representan modelos más estrictos mientras que los valores más altos representan modelos más flexibles.

Por último, una vez realizada la primera ejecución y seleccionadas las variables, se volvió a ejecutar MAXENT para evaluar el modelo. La evaluación del modelo se llevó a cabo mediante el trazado de la curva ROC (AUC), que evalúa la probabilidad de presencia en comparación con las observaciones de presencia-ausencia (Hanley et al. 1982), y permite ver si las predicciones del modelo no difieren de un caso aleatorio si el AUC es 0 o si las predicciones son perfectas si el AUC es 1. Así, el modelo se clasifica según el AUC, que es poco preciso si el AUC es inferior a 0,7, razonable si el AUC está entre 0,7 y 0,9 y muy bueno si es superior a 0,9 (Araújo et al. 2005).

El 80% de los datos se utilizó para entrenar el modelo y el 20% restante para probar su capacidad de predicción (Richards et al. 2015).

3.4.2 Construcción de modelos de distribución

Una vez establecidos todos los parámetros del modelo, se procedió a construir los modelos de distribución de las especies objetivo utilizando dicha configuración optimizada. Para ello, se utilizó la herramienta BIOMOD en R.

El primer paso fue crear los mapas de pseudo-ausencia de forma aleatoria de 10.000 puntos. Para cada especie, se hicieron 3 mapas para tener variabilidad y tener un resultado más preciso, pero el número de mapas está limitado por el tiempo de simulación. Esto dio lugar a 3 mapas para cada una de las 9 especies, que en total son 27 mapas de pseudo-ausencias. A continuación, el modelo combinó los datos de

presencias y pseudo-ausencias con los datos ambientales para crear los mapas de distribución finales que finalmente fueron proyectados en los dos estuarios de estudio.

Se aplicaron cuatro algoritmos diferentes para generar los modelos de distribución de las especies: dos métodos de regresión, el Modelo Lineal Generalizado (GLM) (McCullagh et al., 1989) y el Modelo Aditivo Generalizado (GAM) (Hastie et al., 1990), un método de aprendizaje automático, Random Forest (RF) (Breiman, 2001), y el algoritmo MAXENT (Phillips et al. 2006). Para cada uno, se realizaron 50 iteraciones para llegar a un resultado. Este resultado es más preciso a medida que aumenta el número de iteraciones hasta alcanzar un nivel de equilibrio. En total, el modelo calculó 2 resultados por cada uno de los 3 mapas por cada modelo, dando lugar a 20 resultados por especie, es decir, 216 resultados en total.

Finalmente, se evaluaron los 20 modelos obtenidos para cada especie y, con base en ello, se ensamblaron para generar un único modelo final de la distribución de cada especie (Araújo et al. 2007). Como se ha visto en el apartado anterior, se ha realizado una primera evaluación de los modelos con el AUC. Además, se han utilizado otras dos métricas de evaluación complementarias, que son la medida KAPPA (Allouche et al. 2006) y el True Skill Statistics (TSS) (Allouche et al. 2006). El modelo final se clasificó como pobre si el TSS es inferior a 0,4, bueno si está entre 0,4 y 0,8 y excelente por encima. Para el coeficiente KAPPA, el modelo es muy débil si KAPPA está entre 0 y 0,2, débil entre 0,21 y 0,4, moderado entre 0,41 y 0,6, fuerte entre 0,61 y 0,8 y perfecto entre 0,81 y 1. BIOMOD ensambló los modelos promediando las probabilidades obtenidas para cada resultado para obtener un único modelo. El ensamblado de los modelos se realizó con aquellos modelos cuyo valor de TSS fue superior a 0,5, con el fin de eliminar los malos resultados que podrían reducir la precisión del resultado final. Los 3 métodos de evaluación seleccionados permitieron comparar la idoneidad de los diferentes modelos.

Por último, estos 9 modelos obtenidos se convirtieron al formato adecuado para su procesamiento en ArcGis con el fin de permitir una mejor visualización de los mapas de distribución y facilitar su análisis. Los valores de probabilidad obtenidos se escalaron a un rango entre 0 y 1000, donde '0' indica una baja o nula probabilidad de presencia y '1000' la mayor probabilidad de presencia. Los mapas obtenidos representan la media de las probabilidades de cada modelo según su número de TSS. Cuanto mayor sea el

número de TSS del modelo, más preciso será y, por tanto, más importante será su peso de probabilidad en la media.

Para estudiar el papel de las variables en la distribución de las especies, se crearon curvas de respuesta utilizando MAXENT.

4. Resultados

En primer lugar, se presentarán los resultados de la optimización de la configuración y se compararán los modelos utilizados y, a continuación, se analizarán las curvas de respuesta y los mapas de distribución.

4.1 Optimización de la configuración

Los resultados de la optimización se presentarán en dos partes, la primera sobre la selección de variables y la segunda sobre la selección de las características a utilizar y el multiplicador de regularización.

4.1.1 Selección de las variables

Los resultados de la contribución de cada variable ambiental al modelo generado con MAXENT para cada especie se presentan en la Tabla 2, con las variables seleccionadas en verde por tener una contribución mayor o igual al 1%.

Puede observarse que *A. alba*, *A. tenuis* y *C. edule* sólo requieren unas pocas variables para su modelización, mientras que los demás utilizan una gran parte de las variables disponibles (Tabla 2). Para *A. alba*, se seleccionaron 6 variables, 5 para *A. tenuis*, 3 para *C. edule*, 8 para *C. carinata*, 9 para *H. diversicolor*, 13 para *M. palmata*, 9 para *N. reticulatus*, 13 para *N. hombergii* y *S. plana*.

Los resultados revelan que algunas de las variables tienen una contribución alta para la mayor parte de las especies, como la salinidad media, mostrando su importancia y papel clave en la distribución de las especies en el estuario (Tabla 2). Del mismo modo, la batimetría y la columna de agua son importantes porque las especies estarán

limitadas por la profundidad en su distribución. La duración de la inundación también es un factor importante para muchas especies, pero sobre todo en la distribución de las especies de la comunidad de *N. hombergii*, que se encuentran en la zona de transición entre la zona submareal, que está continuamente sumergida, y las zonas más próximas a los ríos, que sólo están sumergidas temporalmente.

	A. alba	A. tenuis	C. edule	C. carinata	H. diversicolor	M. palmata	N. reticulatus	N. hombergii	S. plana
<i>batimetria</i>	2.1	20.8	57.9	37	0	0.4	0	19.9	4.1
<i>grava</i>	0	3.1	0	0	0.8	2.3	0	1.7	4.6
<i>arena</i>	0	0	0	0	0	15	0	0.9	1
<i>fango</i>	21.2	30.8	5.1	17.1	19.3	33.2	10.4	30.1	23
<i>salinidad</i>	10.2	0	0.3	16.6	6.3	11.5	41.6	10.8	3.4
<i>mediana</i>									
<i>salinidad</i>	0	0	0.1	18	25.4	1.4	0.6	7.4	19
<i>rango</i>									
<i>columna</i>	47.1	19.6	0.1	0	8.2	7.6	0	1.1	8.5
<i>agua</i>									
<i>fuerza de</i>	17.1	0	0.1	0.2	3	2	6.4	7.7	1.3
<i>arrastre</i>									
<i>duracion</i>	0.2	0	0	1.2	4.2	1.3	3.1	1.2	1.2
<i>sequía</i>									
<i>frecuencia</i>	0	0	0	2.6	4.1	2.2	2.2	2.7	2.1
<i>sequía</i>									
<i>intensidad</i>	0	0	0	1	0.1	5	2.7	2.6	1.3
<i>sequía</i>									
<i>duracion</i>	1	25.4	36.2	0	2.9	2.8	23.8	6.2	11.6
<i>inundación</i>									
<i>frecuencia</i>	0.6	0.1	0	6.3	0.2	1.1	7.6	6.2	0.8
<i>inundación</i>									
<i>intensidad</i>	0.5	0.1	0	0	0.5	0.1	1.6	1.6	0.3
<i>inundación</i>									
<i>tiempo</i>	0	0	0	0.1	24.9	14.1	0	0	17.9
<i>emersion</i>									

Tabla 2 : Resultados de la contribución (%) de las variables ambientales en los modelos de distribución de cada especie construidos con MAXENT. En verde se indican las variables seleccionadas

4.1.2 Configuración de los parámetros

Los resultados muestran que el modelo es muy bueno para los valores de regularización de 0,75 y 1, y luego la calidad disminuye hasta obtener un modelo poco preciso para el valor de regularización más alto de 15 (Figura 3). Además, se observa que la complejidad no tiene un impacto significativo en el rendimiento del modelo para

la mayoría de las especies, ya que los resultados obtenidos son similares, pero aumenta fuertemente la calidad de los resultados para *A. tenuis*, *C. edule* y *C. carinata* con la configuración modificada. Por lo tanto, para la construcción de las curvas de respuesta se utilizará un multiplicador de 0,75 y una complejidad modificada con las características "linear", "hinge" y "quadratic".

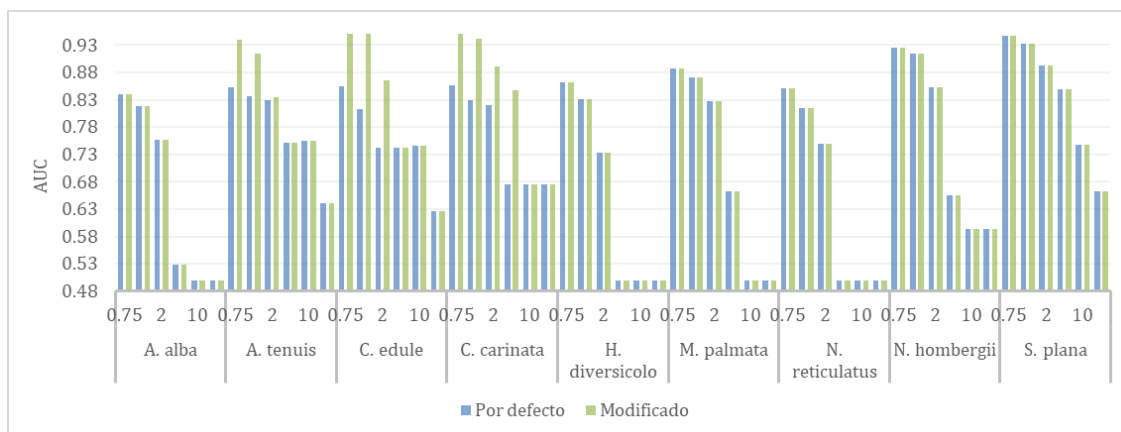


Figura 3 : Resultado de la prueba AUC comparando las diferentes configuraciones consideradas entre el multiplicador de regularización y la complejidad

4.2 Comparación de modelos

Los resultados de los coeficientes AUC, KAPPA y TSS obtenidos para cada modelo muestran una cierta tendencia que se repite en todas las especies (Figura 4). Podemos ver que el modelo GLM tiene coeficientes mucho más bajos que los otros, especialmente para KAPPA, que siempre está muy cerca de 0. Para GAM y RF, tenemos resultados bastante similares con coeficientes que pueden ser más altos para GAM o para RF dependiendo de la especie. Luego, para MAXENT, vemos que los coeficientes TSS y AUC están en los mismos niveles que para GAM y RF, pero su coeficiente KAPPA es mucho menor. Del mismo modo, para todos los modelos se observan coeficientes SST y AUC ligeramente superiores, sin embargo, el KAPPA sigue siendo bastante bajo, aunque más alto que para MAXENT.

La evaluación KAPPA no sigue la misma dinámica que los otros dos coeficientes y, por tanto, complica la selección de un modelo. Por lo tanto, nos concentraremos en el SST y el AUC, que parecen ser más coherentes y, por lo tanto, el modelo que da los mejores resultados es el conjunto de modelos que se utilizará para crear los mapas de distribución.

Se puede observar que los valores de los coeficientes AUC y TSS muestran una buena calidad de los modelos, excepto para *A. alba*, que tiene un AUC ligeramente inferior a 0,7. Los valores de los coeficientes siguen siendo bastante bajos.

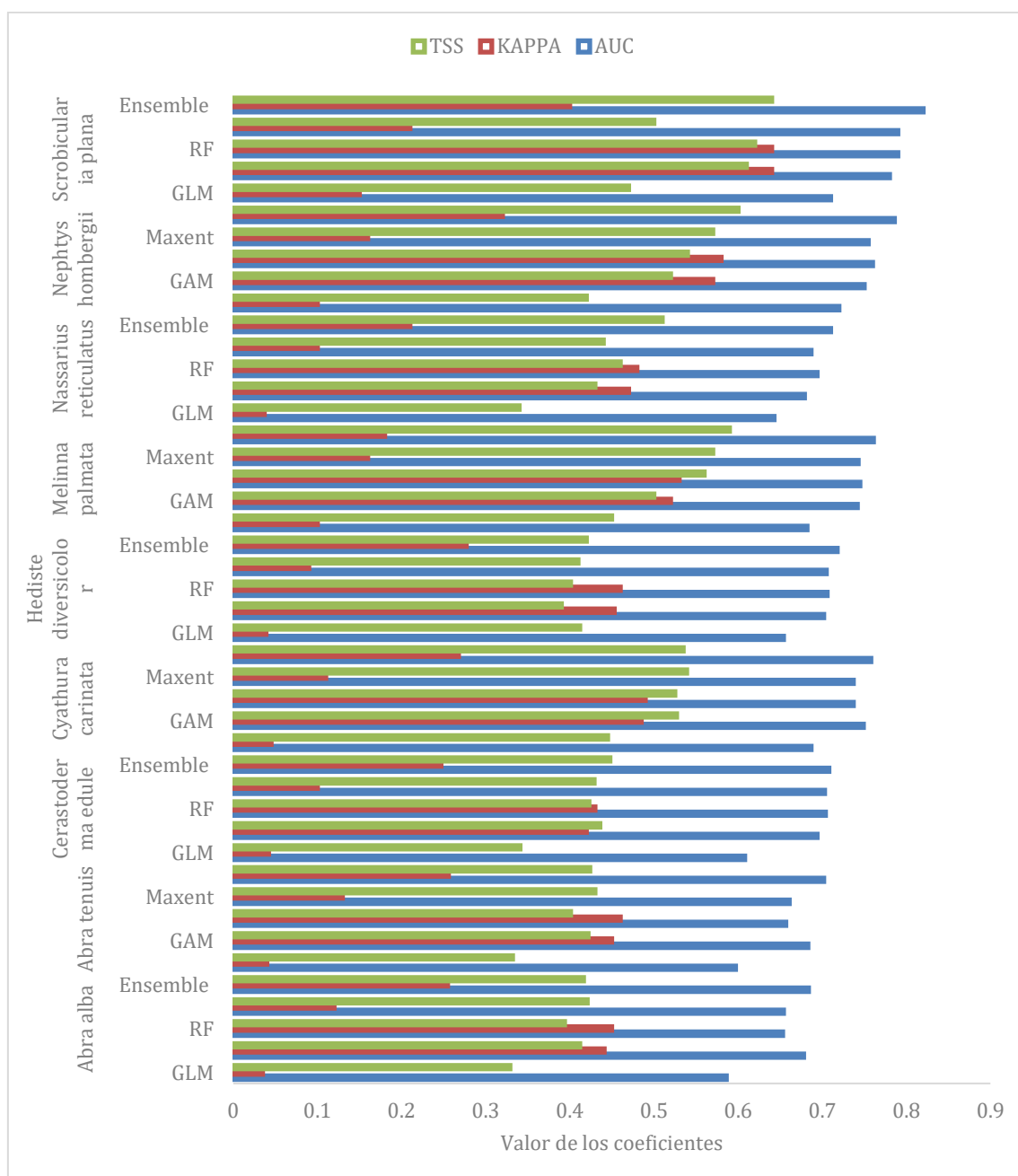


Figura 4 : Resultados de coeficientes KAPPA, AUC y TSS para cada modelo y cada especie

4.3 Análisis de las curvas de respuesta

La primera variable ambiental importante es la naturaleza del sedimento. Esta variable es la que más contribuye a la distribución de las especies en la construcción de los modelos. Se puede ver que la ausencia de fango implica una ausencia de *A. alba* y *M. palmata* (Figura 3). Por debajo del 10% de fango, las especies empiezan a desaparecer. Por lo tanto, las especies prefieren un entorno donde los sedimentos sean muy finos para desarrollarse. Las especies excavadoras *A. alba* y *M. palmata* muestran afinidad por sustratos con un alto contenido en fango (Figura 5.A). Por lo tanto, *M. palmata* estará presente en fondos mixtos de grava y sedimentos finos. Para *N. hombergii*, la naturaleza del sedimento será importante también, ya que se observa que la ausencia de fango conduce a la desaparición de la especie, y lo mismo ocurre con la cantidad de grava, pero con una alta limitación del 25% (Figura 5.B).

En segundo lugar, la fuerza de arrastre en fondo limita la presencia de *A. alba* cuando es demasiado fuerte. Por encima del 45 N/m², la especie desaparece completamente del entorno. A diferencia de *M. palmata* que desaparece cuando la tensión disminuye por debajo del 10% (Figura 5.C).

La salinidad media también es un factor limitante (Figura 5.D). La comunidad de *A. alba* se encuentra en aguas muy saladas y, por tanto, cuando la salinidad disminuye la especie desaparece. La probabilidad comienza a disminuir a partir de un valor inferior de salinidad de 6 y se convierte en cero cuando la salinidad llega a 0, lo que confirma que la especie prefiere las zonas muy saladas del estuario para desarrollarse.

La salinidad tiene un impacto en la distribución de *N. hombergii* y *N. reticulatus*. A diferencia de la comunidad de *A. alba*, que sólo está limitada por la baja salinidad, la comunidad de *N. hombergii* prefiere un determinado rango de salinidad. Para el límite alta salinidad del rango, la especie no desaparece, pero hay una disminución de la población alrededor de 25 de salinidad. Para el límite inferior, la población de *N. hombergii* disminuirá a partir de 6 de salinidad y de 10 para *N. reticulatus*. Esto confirma que la comunidad pertenece a la parte interior del estuario y a la zona intermareal.

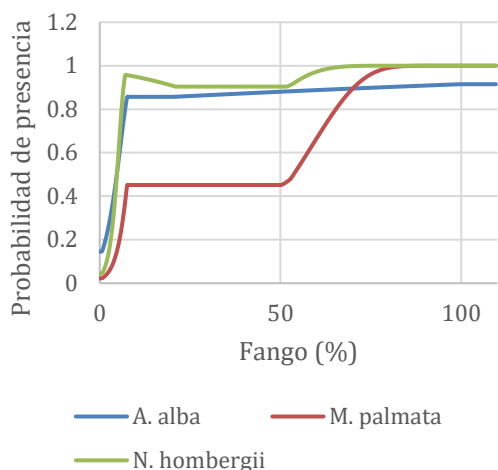
Al contrario que las especies anteriores, las especies de la comunidad de *S. plana* forman parte de la comunidad de especies de zonas oligo-mesohalinas y, por

tanto, presentan un único límite superior a la salinidad, aunque exista un límite inferior porque la salinidad no puede ser inferior a 0. *C. carinata* pierde rápidamente su probabilidad de aparición del 5 a 10 de salinidad, y *H. diversicolor* y *S. plana* con una salinidad superior a 7,5.

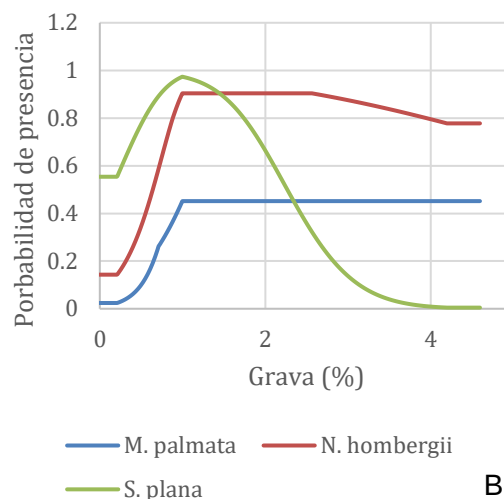
Otro factor limitante es la duración de la inundación, que disminuye la presencia de la especie a medida que aumenta. *A. alba* es una especie que necesita estar sumergida para alimentarse, pero no puede estar inundada durante mucho tiempo (Figura 5.E). También, la duración de la inundación es una variable importante en la distribución de la comunidad de *N. hombergii*, como se desprende de su contribución al modelo (Figura 5.E). Sin embargo, sólo es limitante para 2 especies, *A. tenuis* y *C. edule*. Para ambas especies cuando la duración de la inundación aumenta, la especie desaparece rápidamente. Por lo tanto, se encuentran en zonas con tiempos de inundación muy cortos en los extremos del estuario donde la profundidad es muy escasa.

Esto se confirma con la curva de respuesta de la batimetría. La batimetría es un factor limitante, como lo es para muchas especies que están acostumbradas a aguas poco profundas y desaparecen cuando la profundidad aumenta. La batimetría muestra que *A. tenuis*, *C. edule*, *S. plana* y *C. carinata* colonizan zonas entre los 5 m de altitud y los 5 m de profundidad, con la gran mayoría alrededor del nivel del mar (Figura 5.F) ya que requieren tiempos de inmersión y emersión cortos pero intensos. La batimetría también es limitante para *N. hombergii*, que tiene un rango de tolerancia más amplio entre los 7,5 m de altitud y los 7 m de profundidad. La presencia de *A. alba* comienza a disminuir a partir de los 12,5 m de altitud y se vuelve nula cuando la profundidad supera los 7 m.

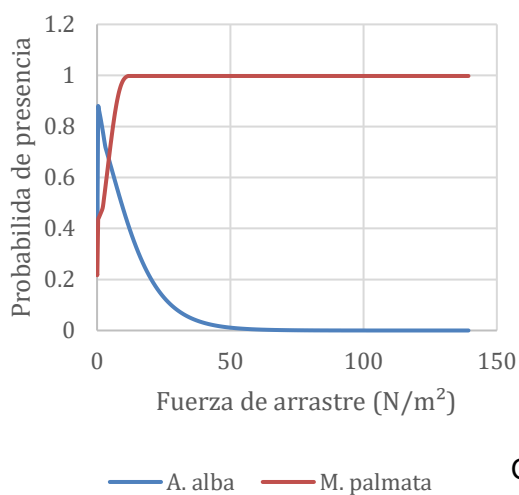
N. reticulatus parece ser mucho más sensible al periodo de sequía y, en particular, a la intensidad ya que podemos ver que su probabilidad de aparición es mayor en zonas de fuerte sequía (Figura 5.G, 5.H y 5.I) y, por tanto, en zonas donde la profundidad es menor y, por tanto, menos frecuente la emergencia y más baja la intensidad de la inundación (Figura 5.J).



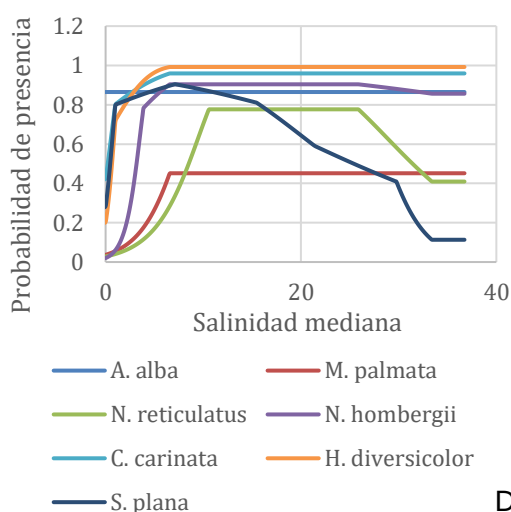
A



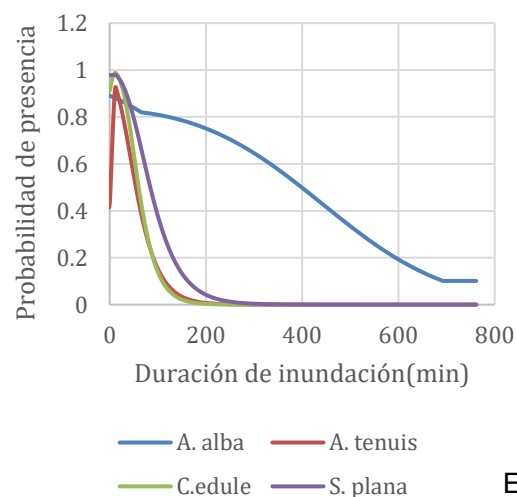
B



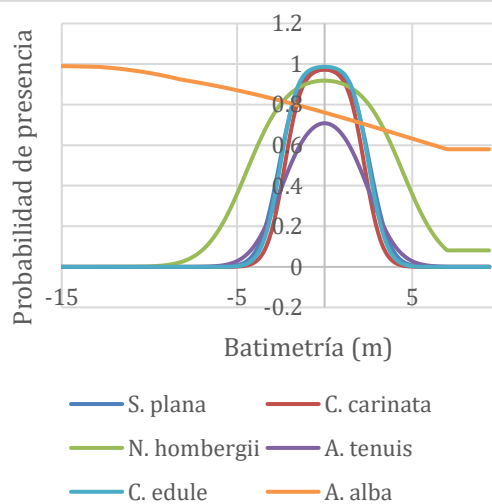
C



D



E



F

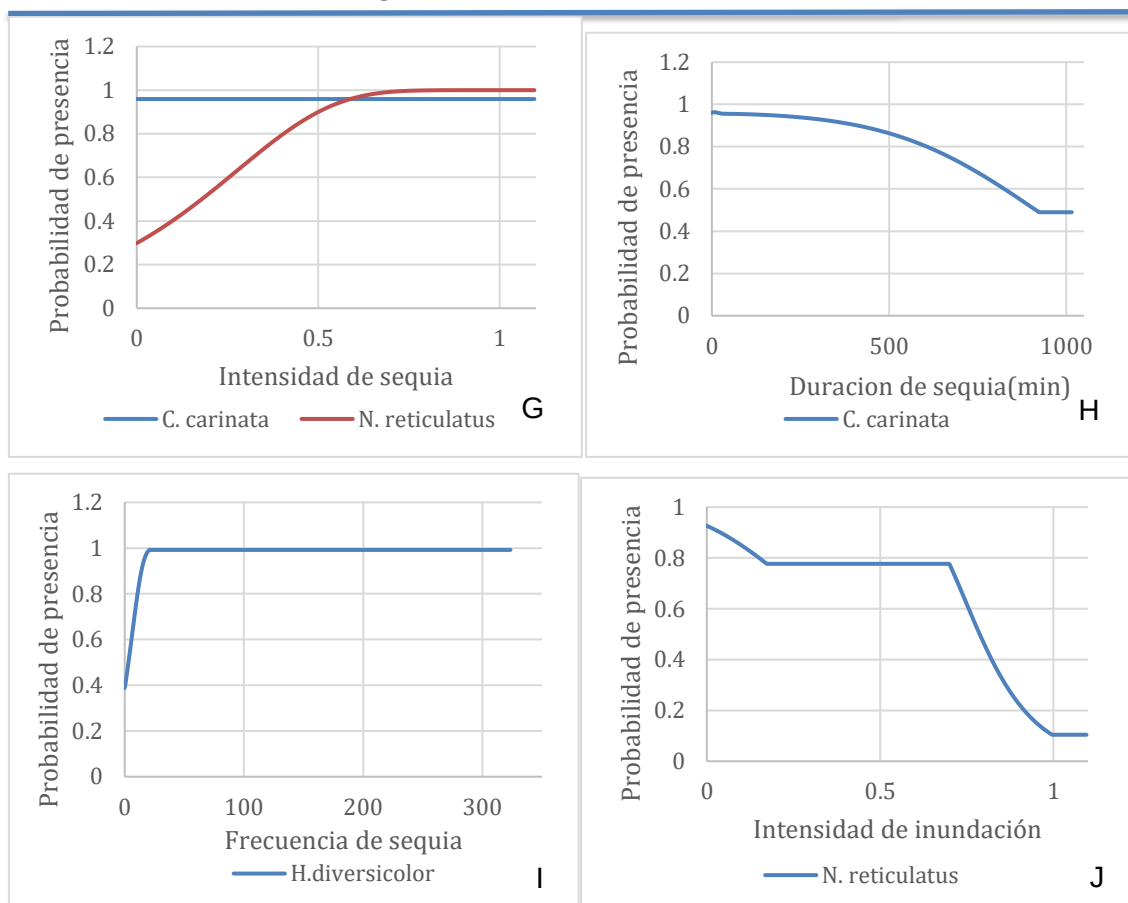


Figura 5 : Curvas de respuestas a las variables ambientales

4.4 Análisis de los mapas de distribución

En la Bahía de Santander, la comunidad de *A. alba* se encuentra en la parte occidental, en el puerto, entre la bocana de la Bahía y la ensenada de San Bartolomé para *A. alba* (Figura 6) y la entrada del mar y el fondo de la ría de San Salvador para *M. palmata* (Figura 7). En las marismas de Santoña, la comunidad se concentra en la desembocadura de la ría y en las zonas más profundas del canal de Hano y de la ría de Treto. Estas áreas corresponden a la zona euhalina (Figura A.1) del estuario directamente conectada al mar y la primera en ser impactada por la marea y la entrada de agua salada. Esta zona también corresponde al canal de navegación del puerto y a los canales naturales de entrada de los ríos, donde la comunidad es característica.

Sin embargo, existen diferencias en la distribución de estas dos especies. En la bahía de Santander, *M. palmata* se extiende más al sur en la ría de Salvador, mientras que *A. alba* se concentra a lo largo del puerto. Esto puede explicarse por la diferencia en la preferencia de los sedimentos, ya que la fuerza de arrastre es la misma en la mayor

parte de la zona (Figura A.3). *A. alba* crece en sedimentos muy finos, pero en este límite entre las dos especies hay una zona de alta concentración de grava que puede haber impedido que *A. alba* se extienda hacia el sur en el estuario (Figura A.5), a diferencia de *M. palmata* que muestra afinidad por la grava.

Las especies de *A. tenuis* y *C. edule*, ambas características de la comunidad de *Nephtys hombergii* y con unas curvas de respuesta a los gradientes ambientales similares, presentan el mismo patrón de distribución espacial dentro del estuario (Figura 9 y 10). Estas especies se distribuyen en la parte central del estuario y cerca de las desembocaduras de los ríos. Estas zonas tienen una cota más baja (Figura A.2) y, por tanto, se inundan con más frecuencia. Estas áreas constituyen la zona intermareal donde la comunidad es característica.

N. reticulatus y *N. hombergii* están ampliamente presentes en las zonas euhalinas. *N. reticulatus* se distribuye por casi toda la superficie del estuario, a diferencia de *C. edule* y *A. tenuis*, que están limitados por su reducida tolerancia a la baja salinidad (Figura A.1), desapareciendo así de las zonas con mayor influencia de los aportes fluviales cerca de la desembocadura del río Miera y del río Asón en la Bahía de Santander y Marismas de Santoña, respectivamente (Figura 8 y 11).

Al igual que *N. reticulatus*, *N. hombergii* está muy extendida en las marismas de Santoña y en el sur (Figura 11). Sin embargo, hemos visto que esta especie está limitada por la naturaleza de los sedimentos y, en particular, por la presencia o ausencia de fango. En el centro de la bahía de Santander, donde debería encontrarse *N. reticulatus*, no hay sedimentos muy finos sino arena, lo que dificulta su desarrollo (Figura A.6). Además, está limitado por la profundidad, a diferencia del más tolerante *N. reticulatus* a la profundidad alta, y, por lo tanto, no se encontrará a lo largo del puerto.

Las especies características de la comunidad *S. plana* se encuentran en las zonas interiores y en las desembocaduras de los ríos (Figura 12, 13 y 14), con mayor probabilidad donde la salinidad es menor (Figura A.1). Algunas de dichas especies como son *S. plana* y *H. diversicolor*, aunque forman parte de la comunidad de la zona oligo-mesohalina, también colonizan la zona polihalina, donde la salinidad es más alta. En el caso de *C. carinata*, su curva de respuesta a la salinidad muestra una disminución en la afinidad por salinidades altas lo cual se observa en el modelo de distribución

generado, donde aparece en tramos polihalinos aunque con una menor probabilidad de ocurrencia.

Las áreas de distribución de *C. carinata* y *S. plana* son bastante similares, con algunas variaciones que se corresponde con la distribución de los distintos tipos de sedimentos. Por otro lado, *H. diversicolor* aparece con baja probabilidad en el centro de los estuarios y con alta probabilidad en las zonas correspondientes al intermareal superior, cerca del límite del estuario con la tierra, donde los períodos de emersión son más largos. Esto se corresponde con la importancia de la frecuencia de las sequías en el modelo de distribución resultante.



Figura 6 : Mapa de distribución de la especie *A. alba*



Figura 7 : Mapa de distribución de la especie *M. palmata*

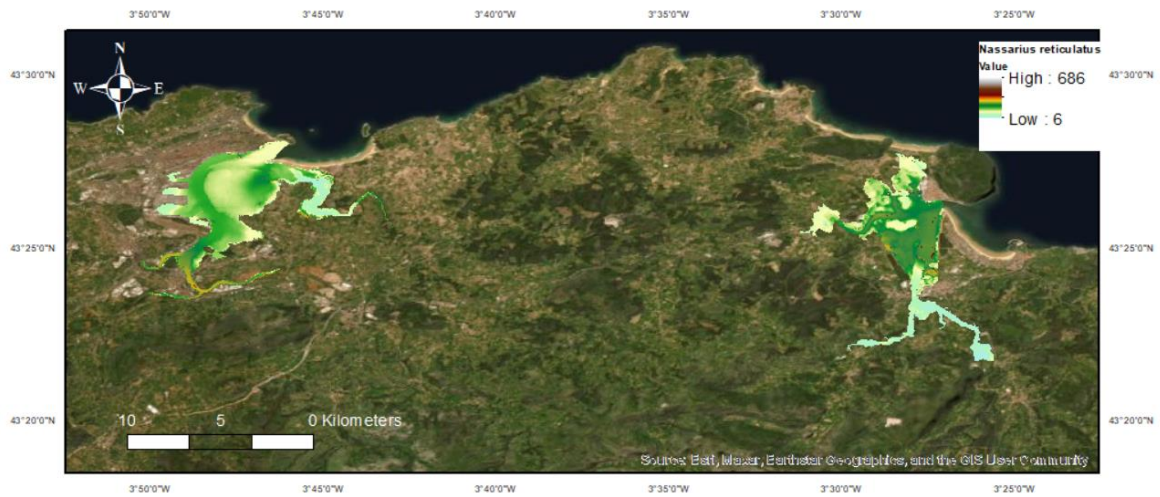


Figura 8 : Mapa de distribución de la especie *N. reticulatus*

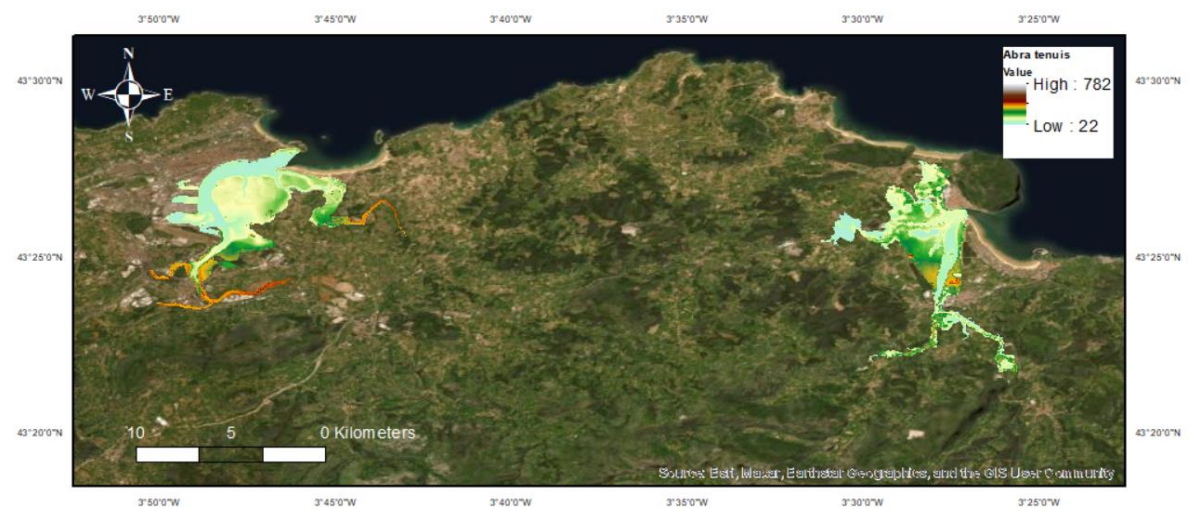


Figura 9 : Mapa de distribución de la especie *A. tenuis*

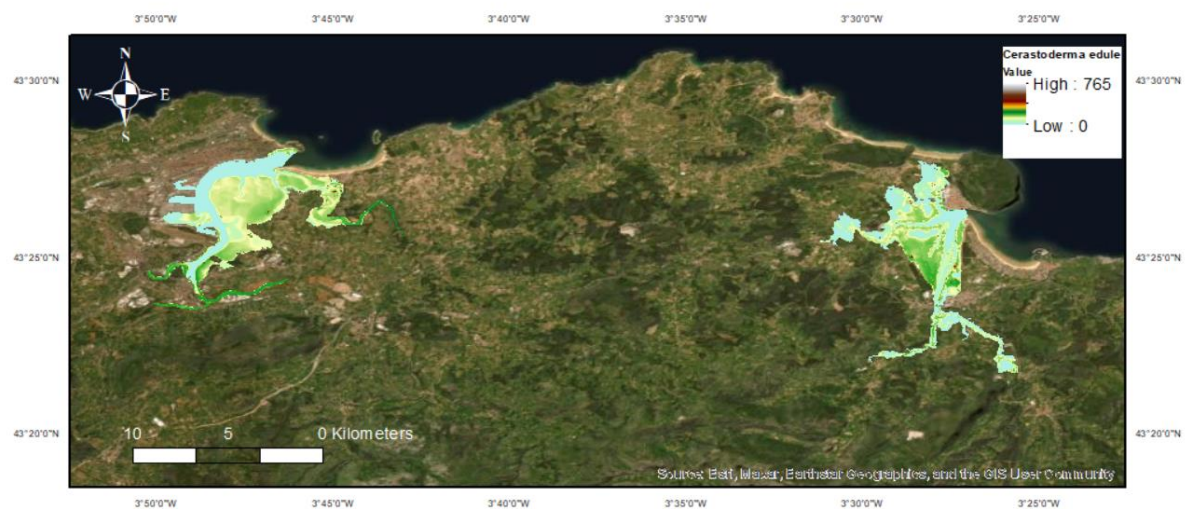


Figura 10 : Mapa de distribución de la especie *C. edule*

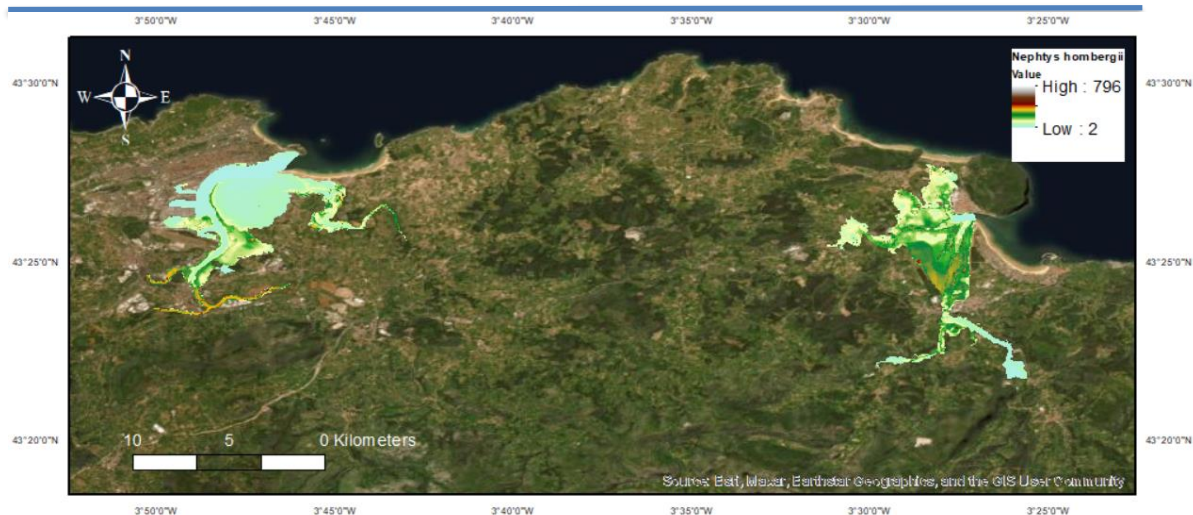


Figura 11 : Mapa de distribución de la especie *N. hombergii*

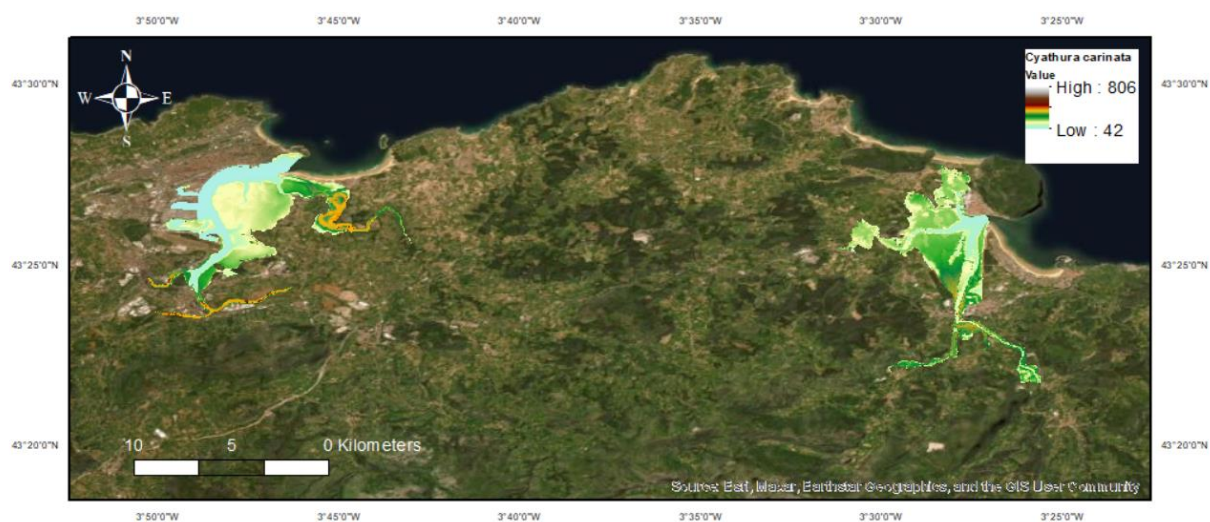


Figura 12 : Mapa de distribución de la especie *C. carinata*

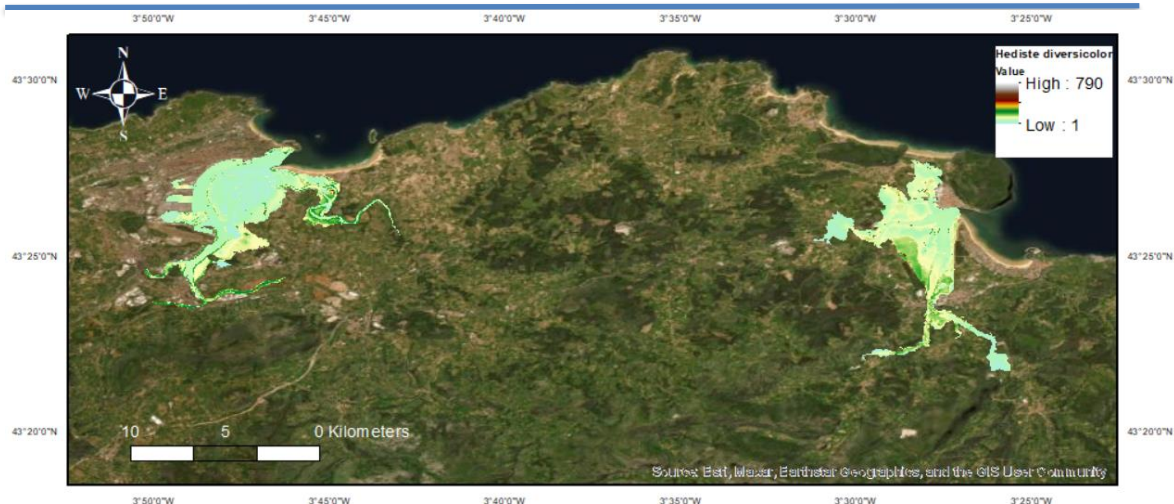


Figura 13 : Mapa de distribución de la especie *H. diversicolor*

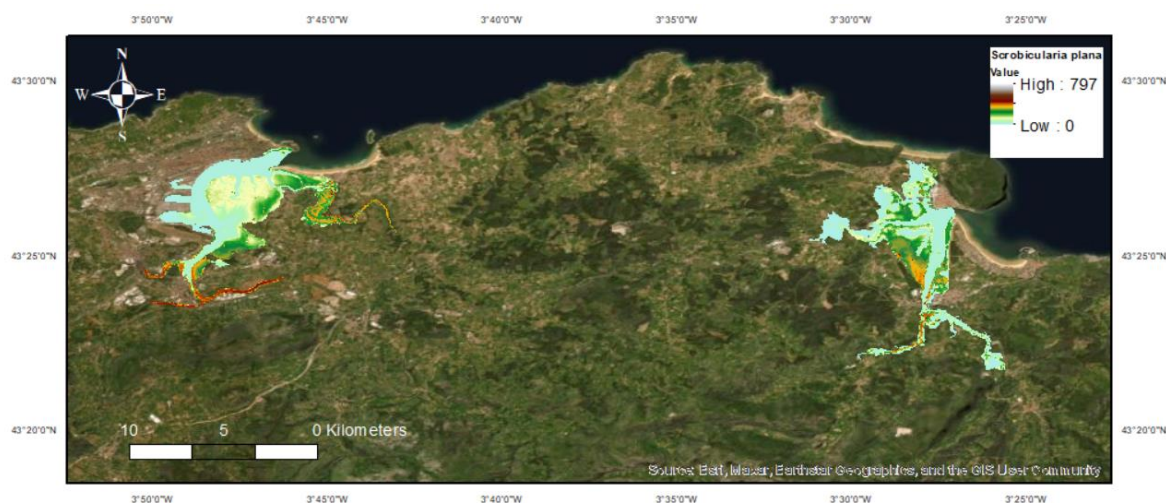


Figura 14 : Mapa de distribución de la especie *S. plana*

5. Discusión

El modelo ha producido resultados que parecen ser buenos y coherentes con la realidad de la distribución de las especies y comunidades estudiadas. Sin embargo, es necesario discutir algunas limitaciones en la metodología y el análisis. En particular, se discutirá la selección de variables, la configuración seleccionada, los resultados de la distribución de especies y la comparación de los modelos.

La selección de las variables es un aspecto muy importante en la modelización, ya que es uno de los primeros pasos en la optimización de la configuración. La selección

mediante el estudio de la contribución de las variables a la distribución de las especies da gran importancia a sus características ecológicas. Aunque para muchos autores el criterio ecológico es el más importante para seleccionar las variables con las que construir un modelo (Fourcade et al. 2018), cabe preguntarse si algunas variables tienen realmente interés en la distribución. Como muestra el análisis de las curvas de respuesta y los mapas de distribución, algunas variables tienen un impacto mucho mayor que otras en la distribución de las especies. Por ejemplo, la salinidad y la batimetría son casi siempre responsables de la distribución y la limitación de la distribución de las especies, y los sedimentos en muchos casos. Sin embargo, la fuerza de arrastre no se ha utilizado como variable explicativa en ninguno de los modelos de distribución generados en este estudio debido a que su variabilidad dentro de la zona de estudio es muy baja y, por tanto, tiene poca capacidad discriminativa. Otra pregunta que surge en relación al criterio para seleccionar las variables explicativas más relevantes es si la selección de variables con un límite del 1% es suficiente o si debería tomarse un umbral más alto para reducir aún más el número de variables y así tener en cuenta sólo las variables realmente importantes en la distribución evitando el ruido de las variables de menor importancia (Martínez et al., 2018).

Para la optimización de la configuración se han estudiado dos factores, la complejidad y el multiplicador de regularización (Elith et al. 2011, Merow et al. 2013). Mientras que esta última mostró su importancia en la optimización de la configuración al dar valores de AUC más altos, la complejidad demostró ser poco útil en la configuración del modelo. De hecho, los resultados obtenidos fueron similares entre la selección manual de los factores de complejidad y la complejidad por defecto para la mayoría de las especies, pero mejor para *A. tenuis*, *C. edule* y *C. carinata*, mientras que normalmente la reducción de la complejidad se supone que da una mejor coincidencia con las respuestas de las especies y por lo tanto da un modelo más cercano a la realidad (Rödger y Engler, 2011; Warren y Seifert, 2011). Esta similitud puede mostrar que este parámetro tiene poco peso en la optimización en comparación con otros factores como la selección de variables y el multiplicador de regularización que mostraron mejores resultados.

Otro punto de discusión es si los resultados obtenidos reflejan la distribución real de las especies. Los resultados obtenidos parecen ser correctos ya que el análisis de distribución parece respetar las condiciones de hábitat de cada especie vistas con las curvas de respuesta. Sin embargo, aunque esta distribución parece correcta, en las

zonas donde no hay datos es cuestionable que la especie esté realmente presente en el caso real. Por ejemplo, se ha visto que algunas especies que deberían tener una alta probabilidad en una zona, no la tienen o que algunas especies tienen una posible presencia en zonas de salinidad donde no deberían encontrarse o sólo muy débilmente.

Para validar o rechazar los resultados, se utilizan métricas ampliamente aceptadas como son el TSS o AUC. Para mejorar la validación de los resultados sería necesario realizar campañas de campo que permitan obtener más puntos de presencia de las especies dentro de los estuarios para verificar la presencia o ausencia de la especie y su abundancia, pero esta técnica sería costosa en promedio, en tiempo y dinero. Los datos disponibles proceden de campañas de medición adaptadas a otros estudios y no a este estudio, por lo que no responden a las necesidades del mismo. Por lo tanto, es importante elegir cuidadosamente dónde colocar los puntos de medición para evitar un sesgo espacial o debilidades.

Por último, la modelización no tiene en cuenta todas las variables que potencialmente condicionan la distribución de las especies. De hecho, el modelo sólo tiene en cuenta las variables abióticas. Sin embargo, las actividades humanas tienen un enorme impacto en el medio ambiente y, por tanto, en la distribución de las especies (Amara, 2010). Del mismo modo, las condiciones bióticas de competencia, depredación y otras relaciones entre especies pueden influir en la distribución de las especies (Lefevre et al. 2010). Por ejemplo, hay especies que pueden estar ausentes en ciertas zonas donde las condiciones ambientales son idóneas para su establecimiento debido a la presencia de un depredador o por la ausencia de presas. Por lo tanto, la aplicación de un método de verificación por campaña de campo no podría validar completamente el modelo, ya que son posibles grandes diferencias teniendo en cuenta o no estas variables. Por lo tanto, el objetivo de futuros estudios para mejorar este modelo será incorporar variables de interacción biótica. Sin embargo, hemos visto que la modelización sólo con variables abióticas es compleja y, por tanto, tener en cuenta estas variables además sería muy complicado y requeriría importantes recursos en términos de tiempo, cálculo y financiación.

La distribución también nos muestra que las especies de diferentes comunidades tienen una distribución más similar que las especies de una misma comunidad. Por ejemplo, *S. plana*, *C. carinata*, *A. tenuis* y *C. edule* tienen una distribución similar. Esto plantea la cuestión de si la distribución de la comunidad para la similitud de la resistencia

a la salinidad es realmente la mejor manera de clasificar estas especies o si se deben tener en cuenta más variables. Por ejemplo, hemos visto que la batimetría también tiene un papel muy importante en la distribución de las especies, principalmente por la profundidad, pero también por la definición de las zonas con inundaciones y sequías más o menos frecuentes, intensas y duraderas. Esto modifica la distribución de las especies a pesar de la salinidad y parece ser otra condición de igual importancia. Por lo tanto, las comunidades deben hacerse teniendo en cuenta ambos factores y no sólo la salinidad.

Para la creación de los mapas de distribución, se han utilizado 4 algoritmos diferentes: GLM, GAM, RF y MAXENT (McCullagh et al., 1989, Hastie et al., 1990, Breiman, 2001 y Phillips et al. 2006). Se han construido mapas utilizando todos estos modelos, y aunque los coeficientes obtenidos muestran una buena calidad, siguen siendo bastante bajos, especialmente para *A. alba* que muestra una calidad pobre en comparación con los demás. La buena calidad de los modelos no sólo se debe a la buena configuración del modelo, la buena filtración o el uso del mejor modelo, sino también a la calidad de los datos de partida. Entre los datos biológicos, cabe preguntarse si el número de puntos de localización de cada especie es suficiente para obtener buenos resultados o si serían necesarias campañas adicionales.

Por último, el número de iteraciones también es importante, modelos como el GLM y el RF dan resultados con bastante rapidez por lo que se puede considerar un gran número de iteraciones que mejoren el resultado. Sin embargo, para el GAM y el MAXENT, el tiempo de cálculo es muy largo y, por lo tanto, calcular con más iteraciones requeriría muchos recursos computacionales. La selección del modelo dependerá, por tanto, de los recursos disponibles.

6. Conclusiones

El objetivo de este estudio fue elaborar modelos de distribución de especies bentónicas estuarina en los estuarios cantábricos. Este objetivo se dividió en tres objetivos específicos que eran analizar el papel de las variables seleccionadas en la distribución de las especies, analizar la distribución espacial de las especies y comparar los modelos generados con diferentes algoritmos.

Para el primer objetivo, se vio que la salinidad era el factor más importante en la distribución de las especies. Actúa como factor limitante para todas las especies y, de hecho, se han identificado tres comunidades biológicas que se asocian con diferentes tramos salinos: la comunidad de *A. alba* que coloniza preferentemente los tramos euhalinos, la comunidad *N. hombergii* que muestra afinidad por los tramos polihalinos y la comunidad *S. plana* que se asocia con los tramos oligo-mesohalinos del estuario. Una vez que se ha tenido en cuenta la salinidad, la batimetría adquiere una gran importancia para un segundo análisis, ya que nos permite ver las diferencias entre las especies que prefieren una menor profundidad para tener un menor tiempo de inmersión mientras que se ven más afectadas por las mareas o viceversa. Por último, el sedimento es otra variable muy importante para la distribución de ciertas especies que prefieren sustratos finos o gruesos.

En cuanto a la distribución de las especies, hemos visto que la comunidad de *A. alba* se distribuye más a lo largo del puerto hasta la Ría de San Salvador en la Bahía de Santander y, en general, a lo largo de los canales de navegación y zonas batimétricas más profundas. La comunidad de *N. hombergii* es un poco más heterogénea, con especies que se distribuyen en la zona euhalina, como la comunidad de *A. alba*, o en la zona interna con salinidades más bajas, dependiendo de su resistencia a la salinidad. Por último, la comunidad de *S. plana* se encontrará principalmente en las zonas de baja salinidad, pero también puede encontrarse en las zonas internas donde la salinidad es un poco más alta.

Entre los 4 modelos utilizados, hemos visto que el ensamblado de los modelos dio un resultado de calidad ligeramente mejor que los otros modelos, seguido de GAM y el RF. Este último tiene la ventaja, además, que las necesidades computacionales son menores, lo que permite un mayor número de iteraciones y, por lo tanto, modelos más fiables.

Este estudio ha proporcionado un enfoque de modelización de la distribución de las especies, que puede mejorarse tomando los datos más adecuados para el estudio y un modelo adaptado a los recursos informáticos y de tiempo disponibles, y utilizarse en estudios de cambio climático o de clasificación medioambiental.

7. Referencias

- Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B., Anderson, R.P., 2015. « spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models ». *Ecography (Cop.)*. 38, 541-545.
- Amara, R. 2010. « Impact de l'anthropisation sur la biodiversité et le fonctionnement des écosystèmes marins. Exemple de la Manche-mer du nord ». *VertigO*, no Hors-série 8 (octobre).
- Araújo, M.B., Peterson, A.T. 2012. « Uses and Misuses of Bioclimatic Envelope Modeling ». *Ecology* 93 (7): 1527-39.
- Birch, G. F., J. -H. Lee, E. Tanner, J. Fortune, N. Munksgaard, J. Whitehead, C. Coughanowr, et al. 2020. « Sediment Metal Enrichment and Ecological Risk Assessment of Ten Ports and Estuaries in the World Harbours Project ». *Marine Pollution Bulletin* 155 (juin): 111129.
- Breiman, L., 2001. « Random Forests ». *Mach. Learn.* 45, 5-32.
- Campbell, A.C., Nicholls J., 1986, « Guide de la faune et de la flore des mers d'Europe », coll. « Les guides du naturaliste », ed. Delachaux & Niestlé, 322p.
- Caspers, H. 1974. « E. Naylor: British Marine Isopods. Keys and Notes for the Identification of the Species ».— *Synopses of the British Fauna N.S. No. 3*. With 24 Fig., 86 Pp. London New York: Academic Press, Publishers for the Linnean Society of London. 1972. £ 1.00, \$ 3.25 ». *Internationale Revue Der Gesamten Hydrobiologie Und Hydrographie* 59 (4): 590-91.
- da Costa, E.M. 1778. « *Historia Naturalis Testaceorum Britanniae* » London : Millan, White, Elmsley & Robson. 254pp, 17pls.
- Degraer, S., éd. 2006. « The Macrobenthos Atlas of the Belgian Part of the North Sea ». Brussel: Belgian Science Policy.
- Fish, J.D., Fish, S., 1996. « A student's guide to the seashore ». Second edition. Cambridge: Cambridge University Press.

- Fourcade, Y., Besnard, A.G., Secondi, J., 2018. « Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics ». *Glob. Ecol. Biogeogr.* 27, 245-256.
- Galván, C., Juanes, J.A., Puente, A. 2010. « Ecological Classification of European Transitional Waters in the North-East Atlantic Eco-Region ». *Estuarine, Coastal and Shelf Science* 87 (3): 442-50.
- Galván, C., Puente, A., Castanedo, S. 2016. « Average vs. extreme salinity conditions: Do they equally affect the distribution of macroinvertebrates in estuarine environments? » *Limnology and Oceanography* 61: 984-1000.
- Gobierno de España, 2005 « Red de calidad de Cantabria de Aguas de Transición y Costeras ».
- Grube, A.E. 1870. « Ueber seinen Aufenthalt in St. Malo, besonders über zwei noch nicht beschriebene Anneliden ». *Jahres-Bericht der Schlesischen Gesellschaft für Vaterländische Cultur, Breslau.* 47: 68-69.
- Hastie, T.J., Tibshirani, R.J., 1990. « Generalized additive models ». Chapman & Hall, Londres (UK).
- Hayward P.J., Nelson-Smith T., Shields C., 1998, « Guide des bords de mer, Mer du Nord, Manche, Atlantique, Méditerranée », ed. Delachaux & Niestlé, 351p
- Holtmann, S.E. 1996. « Atlas of the Zoobenthos of the Dutch Continental Shelf ». Rijswijk: Ministry of Transport, Public Works and Water Management, North Sea Directorate.
- Jerosch, K. 2012. « Geostatistical mapping and spatial variability of surficial sediment types on the beaufort shelf based on grain size data ». *J. Mar. Syst.* 127,513.
- Lefevre, Thierry, Francois Renaud, Marc Selosse, Frederic Thomas, Carlos Bernstein, Olivier Broennimann, S. Charlat, et al. 2010. « Evolution des interactions entre espèces. », janvier.
- Linnean Society of London. 1802. « Transactions of the Linnean Society of London ». Vol. 6. London: [The Society], 1791 1875.
- Martínez, B., Radford, B., Thomsen, M.S., Connell, S.D., Carreño, F., Bradshaw, C.J.A., Fordham, D.A., Russell, B.D., Gurgel, C.F.D., Wernberg, T., 2018. « Distribution models predict large contractions of habitat-forming seaweeds in response to ocean warming ». *Divers. Distrib.* 24, 1350-1366.

- McCullagh, P., Nelder, J.A., 1989. « Generalized Linear Models (Monographs on statistics and applied probability 37) ». Chapman & Hall, Londres (UK).
- Montagu G 1803. « Testacea Britannica, or natural history of British shells, marine, land and the fresh-water, including the most minute: systematically arranged and embellished with figures Romsey », London. 606pp, 16pls
- Montero, N., M. J. Belzunce-Segarra, A. Del Campo, J. M. Garmendia, L. Ferrer, J. Larreta, M. González, M. A. Maidana, M. Espino. 2013. « Integrative Environmental Assessment of the Impact of Pasaia Harbour Activities on the Oiartzun Estuary (Southeastern Bay of Biscay) ». Journal of Marine Systems, XII International Symposium on Oceanography of the Bay of Biscay, 109-110 (janvier): S252-60.
- Naylor, E., 1972. « British Marine Isopods ». Synopses of the British Fauna (NS), 3.
- Phillips, Steven J., Robert P. Anderson, Robert E. Schapire. 2006. « Maximum Entropy Modeling of Species Geographic Distributions ». Ecological Modelling 190 (3): 231-59.
- Puente, A., J.A. Juanes, A. García, C. Álvarez, J.A. Revilla, et I. Carranza. 2008. « Ecological Assessment of Soft Bottom Benthic Communities in Northern Spanish Estuaries ». Ecological Indicators, Assessing the environmental quality status in estuarine and coastal systems: comparing methodologies and indices, 8 (4): 373-88.
- Puente, A., Diaz, R.J. 2008. « Is It Possible to Assess the Ecological Status of Highly Stressed Natural Estuarine Environments Using Macroinvertebrates Indices? » Marine Pollution Bulletin 56 (11): 1880-89.
- Richards, D.R., Friess, D.A., 2015. « A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of socialmedia photographs ». Ecological Indicators, 53: 187-195.
- Rödder, D., Engler, J.O., 2011. « Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks ». Glob. Ecol. Biogeogr. 20, 915-927.
- Shcheglovitova, M., Anderson, R.P., 2013. « Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes ». Ecol. Modell. 269, 9-17.

Tanner, E.P., Papeş, M., Elmore, R.D., Fuhlendorf, S.D., Davis, C.A. 2017. « Incorporating Abundance Information and Guiding Variable Selection for Climate Based Ensemble Forecasting of Species' Distributional Shifts ». PLOS ONE 12 (9): e0184316.

Verfaillie, E., Van Meirvenne, M., and Van Lancker, V. (2006). « Multivariate geostatistics for the predictive modeling of the surficial sand distribution in shelf seas ». Cont. Shelf Res. 26, 2454–2468.

Warren, D.L., Seifert, S.N., 2011. « Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria ». Ecol. Appl. 21, 335-342.

Anexo

Datos ambientales

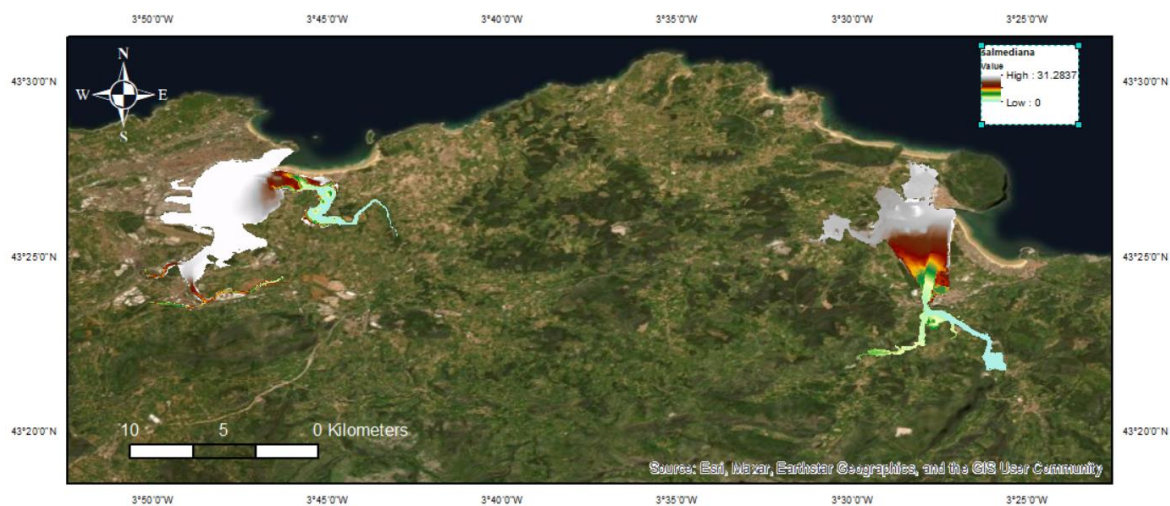


Figura A.1 : Mapa de la salinidad mediana en los estuarios de Santander y las Marismas de Santoña (Galván et al. 2016)



Figura A.2 : Mapa de la batimetría en los estuarios de Santander y las Marismas de Santoña (Galván et al. 2021)

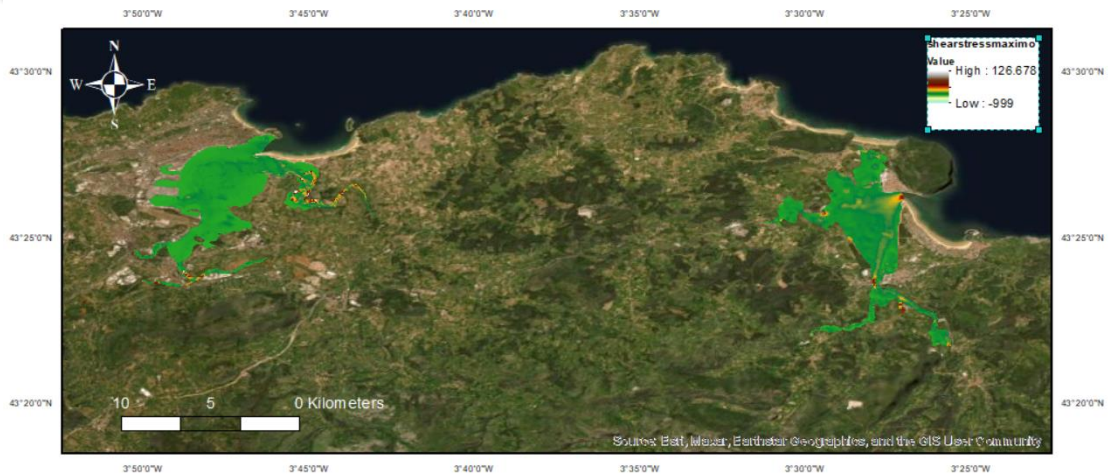


Figura A.3: Mapa de la fuerza de arrastre máxima en los estuarios de Santander y las Marismas de Santoña

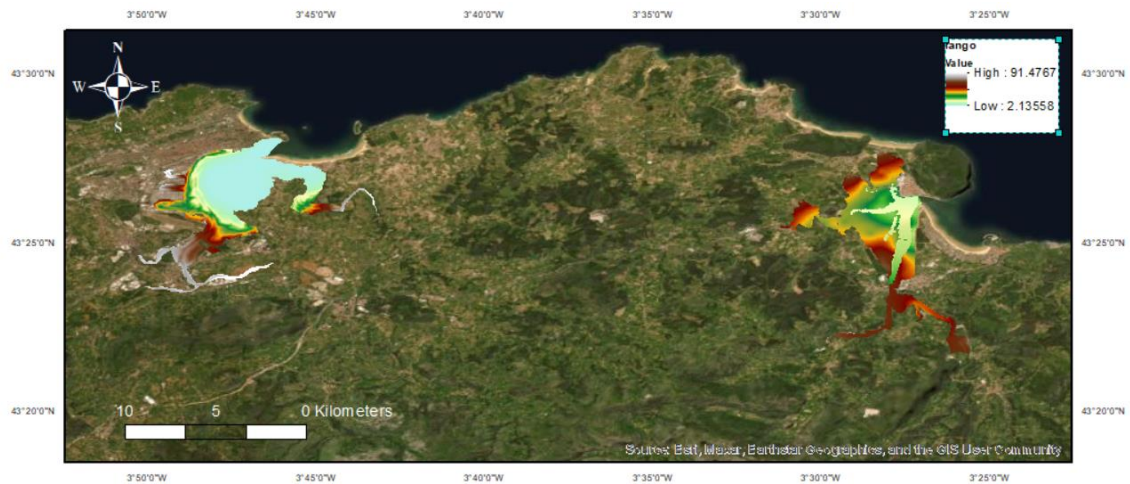


Figura A.4 : Mapa del fango en los estuarios de Santander y las Marismas de Santoña (Galván et al. 2021)

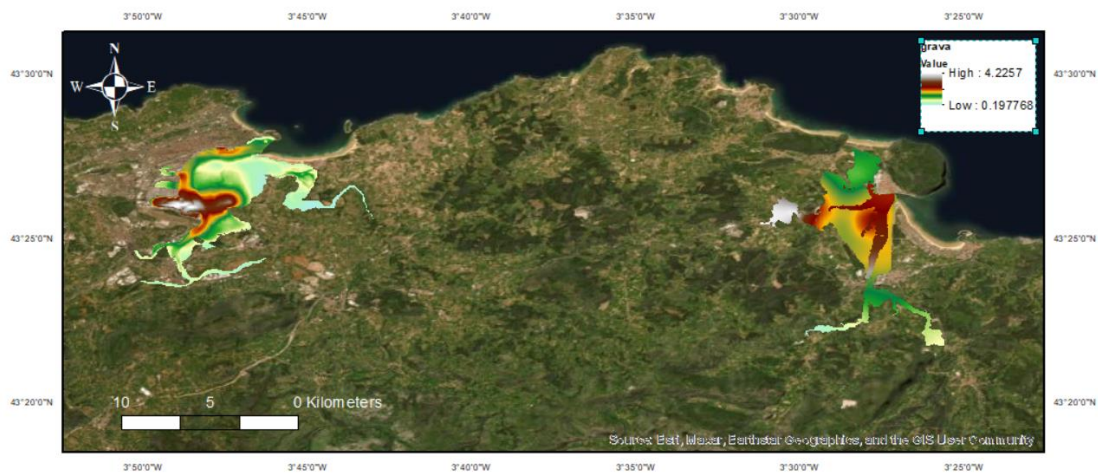


Figura A.5 : Mapa de las gravas en los estuarios de Santander y las Marismas de Santoña (Galván et al. 2021)

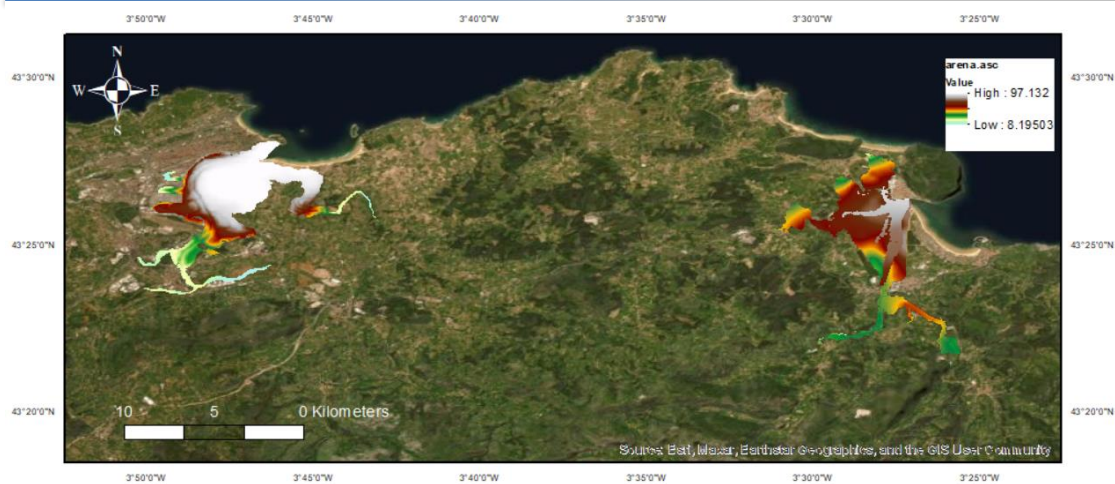


Figura A.6 : Mapa de la arena en los estuarios de Santander y las Marismas de Santoña (Galván et al. 2021)