MARCO-BOLO

STRENGTHENING BIODIVERSITY OBSERVATION IN SUPPORT OF DECISION MAKING

# Deliverable 5.1
# Scientific document determining biodiversity trends, underlying drivers and essential observations for predictive modelling
# V1.0

17.11.2025

*Andrés Peredo Arce[1], Johannes Nowe[2], Carlota Muñiz[2], Klaas Deneudt[2], Elisabeth Debusschere[2], Daniele Iudicone[3], Patrizio Mariani[4], Peter Haase[1], Carlos Cano-Barbacil[1], Florian Kokoszka[5,3], Sarah Asdar[3], Camil Lefebvre[6], Bruno Buongiorno Nardelli[5], Paola Mercogliano[7], Maurizio Ribera d'Alcalá[3], Francesca Margiotta[3]*

**[1]Senckenberg Gesellschaft fur Naturforschung (SGN), [2]Flanders Marine Institute (VLIZ), [3]Stazione Zoologica Anton Dohrn (SZN), [4]Danmarks Tekniske Universitet (DTU), [5]Consiglio Nazionale delle Ricerche (CNR), [6]École centrale de Nantes (ECNantes), [7]Centro Euro-Mediterraneo sui Cambiamenti Climatici (CCM)**

PUBLIC

Funded by the European Union

Funded by the European Union under the Horizon Europe Programme, Grant Agreement No. 101082021 (MARCO-BOLO). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

UK Research and Innovation

UK participants in MARCO-BOLO are supported by the UKRI's Horizon Europe Guarantee under the Grant No. 10068180 (MS); No. 10063994 (MBA); No. 10048178 (NOC).

MARCO-BOLO
STRENGTHENING BIODIVERSITY OBSERVATION IN SUPPORT OF DECISION MAKING

## Document Information

| | |
|---|---|
| Grant Agreement | 101082021 |
| Project Acronym | MARCO-BOLO |
| Project Title | MARine COastal BiOdiversity Long-term Observations |

| | |
|---|---|
| Deliverable Number | D5.1 |
| Work Package Number | WP5 |
| Deliverable Title | Scientific document determining biodiversity trends, underlying drivers and essential observations for predictive modelling |
| Lead Beneficiary | Senckenberg Gesellschaft für Naturforschung (SGN) |
| Author(s) | Andrés Peredo Arce (SGN), Johannes Nowe (VLIZ), Carlota Muñiz (VLIZ), Klaas Deneudt (VLIZ), Elisabeth Debusschere (VLIZ), Daniele Iudicone (SZN), Patrizio Mariani (DTU), Peter Haase (SGN), Carlos Cano-Barbacil (SGN), Florian Kokoszka (CNR, SZN), Sarah Asdar (CNR), Camil Lefebvre (ECNantes), Bruno Buongiorno Nardelli (CNR), Paola Mercogliano (CCM), Maurizio Ribera d'Alcalá (SZN), Francesca Margiotta (SZN) |
| Due Date | 30.11.2025 |
| Submission Date | 17.11.2025 |
| Dissemination Level | PU[1] |
| Type of Deliverable | R[2] |

| | |
|---|---|
| Version 1.0 | 17.11.2025, Andrés Peredo Arce, Johannes Nowe, Carlota Muñiz, Daniele Iudicone, Patrizio Mariani, Peter Haase, Carlos Cano-Barbacil, Florian Kokoszka, Sarah Asdar, Camil Lefebvre, Bruno Buongiorno Nardelli, Paola Mercogliano, Maurizio Ribera d'Alcalá, Francesca Margiotta |

[1] Dissemination level (DELETE ACCORDINGLY): **PU:** Public, **SEN:** Sensitive, **CL:** EU Classified, information as referred to in European Commission Decision 2015/844

[2] Type of deliverable (DELETE ACCORDINGLY): **R: Document,** Report, **DEM:** Demonstration, pilot, prototype, **DEC:** Website, patent filing videos, **DMP:** Data Management Plan, **Ethics:** Ethics deliverable

## Executive Summary

The main goal of this report is to assess biodiversity trends and their drivers in marine and coastal European ecosystems. To this end, we analysed **community level biodiversity trends** (section 2.1) and **species level habitat use** (section 2.2) during the last decades using data from open sources. Biodiversity trends are paired with environmental parameters and **projected into the future**, using Habitat Suitability Models (section 2.2) and Random Forest Regressions (section 2.3).

The **first study** analysed European time series for six biotic groups (i.e. birds, fish, invertebrates, macroalgae, phytoplankton and zooplankton) to estimate temporal trends (1956-2022) in richness, diversity, and abundance across regions. The data was extracted from open-access databases (BioTIME, EMODnet, REPHY, FishGlob, Continuous Plankton Recorder Survey). Most communities showed no significant change, suggesting **no further widespread biodiversity loss during the observation period**. Positive trends were found for birds and invertebrates in the **Baltic**, and negative ones for fish in the **Atlantic**. However, **uneven data coverage** limits this generalization, highlighting the lack of sufficient rigorous monitoring of biodiversity and of accessibility to existing data.

The **second study** focused on predicting habitat suitability for coastal and marine species protected under Habitat and Bird Directives, under current and future Shared Socioeconomic Pathways (SSP) climate scenarios. Projections are made for four **marine mammal species** (harbour porpoise, harbour seal, common and bottlenose dolphins) in **OSPAR regions II–IV,** using **EurOBIS data (2000–2019)** and environmental predictors. Results revealed **clear spatial and seasonal patterns** in projected suitable habitats: southward shift for porpoises in winter, seals stay coastal, and dolphins concentrate in the Iberian region. **Future projections** suggest an overall reduction in suitable habitats for all four species. While limited by **sampling biases and data inconsistencies**, the study demonstrates the **value of open-access biodiversity data** for large-scale ecological modelling.

The **third study** examined **climate change impacts on coastal ecosystems**, focusing on the **Gulf of Naples** (LTER Mare Chiara site). Combining long-term data, reanalysis, and **machine learning**, the study found **salinity** to be a key driver of **chlorophyll (phytoplankton) variability**, linking land-based freshwater inputs and ocean dynamics. Using Representative Concentration Pathways to 2070 (RCP4.5, RCP8.5), the model predicts **increasing salinity and declining chlorophyll**, largely driven by **reduced rainfall and runoff**. The findings stress the **importance of long-term monitoring** and indicate that **land-driven changes** may affect coastal productivity more than ocean warming alone.

This report highlights the **usefulness of publicly available data** but also its limitations. On the one hand, **existing open data allows us to assess long-term biodiversity trends and to project those trends into the future**. On the other hand, the limited data availability **reduces the representativity of the results and decreases the precision of future projections**. As biodiversity open-source data increases in quantity and quality, the potential of future analyses will grow with it.

With this Report, we aim to provide a broad perspective of the past, present and future of marine and coastal biodiversity across Europe.


## Contents

# 1. Introduction

Assessing the status of ecosystems is of high relevance for conservation, ecology and ecosystem management. Status assessments include a broad variety of approaches including community metrics and indicator species. As the specific metrics and surveying methodologies have been standardised during the last decades, the opportunity to study long term trends, their causes and their future projections arises.

In Europe, a broad variety of projects, research infrastructures and legislations, including the Marine Strategy Framework Directive (MSFD) the Integrated European Long-Term Ecosystem, critical zone and socio-ecological Research Infrastructure (eLTER RI), offer sources of data since many years. Complementarily, scientific knowledge on the cause-effect relationships between environmental predictors and biodiversity has grown during the last decades. At the same time, the robustness of statistical models capable of predicting those relationships grew, being climate models a common example. In sum, there are big datasets of biodiversity descriptors and of ecological parameters for long time periods now available, as well as an improved understanding of their cause-effect relationship, and models able to generate robust future projections of those parameters.

Improvements in time series analyses combined with the increasing amount of biodiversity time series data provide new opportunities for research, conservation and biodiversity assessment. In recent years, the amount of scientific papers addressing biodiversity change from local to global scale has significantly increased. This knowledge gain now allows for calculating trajectories of biodiversity change and assessing the status of ecosystems at a precise moment in time. Combined with environmental driver analyses including pollution, climate change and biological invasions, drivers of biodiversity change could be identified and the respective policies informed. Such analyses could provide evidence that past conservation efforts paid off, as shown, for example, for European freshwater biodiversity which significantly improved over the past decades due to the large-scale implementation of wastewater treatment plants (Haase et al. 2023; Sinclair et al. 2024).

The goal of this report is to assess biodiversity trends and their drivers in marine and coastal European ecosystems. To this end, we analysed community level (section 2.1) and species level (section 2.2) biodiversity trends of the last decades using data from open sources. Biodiversity trends are paired with environmental parameters and projected into the future, using Habitat Suitability Models (section 2.2) and Random Forest Regressions (section 2.3).

# 2. Detection and analysis of biodiversity trends

## 2.1 Time series analysis

### 2.1.1 Objective

Based on long-term ecological time series data from publicly accessible databases (e.g. BioTime; EMODnet), we will quantify trends of coastal and marine biodiversity (abundance, species richness and taxonomic diversity) of six biotic groups (birds, fish, invertebrates, macroalgae, phytoplankton and zooplankton) at European scale.

### 2.1.2 Background

Biodiversity loss is one of the main global-scale challenges our planet is facing. Losses have been described in marine, freshwater and terrestrial ecosystems, in every ocean and continent (Butchart et al. 2010). In European marine and coastal waters, biodiversity loss is caused by human activities: overfishing reduced fish and invertebrate stocks and harmed vulnerable species (Piroddi et al. 2020); trawling destroyed seabed integrity (Eigaard et al. 2017); eutrophication caused harmful algal blooms (Tsikoti & Genitsaris 2021); pollution increased heavy metal levels, intoxicating marine organisms (Tornero & Hanke 2016); oil spills destroyed coastal habitats (Castège et al. 2014); climate change shifted species distribution towards colder waters and enabled biological invasions of species from warmer waters (Chust et al. 2024).

As the causes of biodiversity loss in European marine ecosystems were identified, legislation and policies to counteract the losses were implemented. For example, the Common Fisheries Policy introduces fishing quotas (Froese et al. 2018); deep ocean trawling is now forbidden in some deep-sea ecosystems (Regulation 2016/2336); the Urban Waste Water Treatment Directive fostered significant reductions of nitrogen pollution of coastal and marine ecosystems (Directive 2024/3019), while the Descriptor 8 of the MSFD aims to reduce eutrophication of coastal and marine waters (Grizzetti et al. 2021); the Erika legislative packages increase safety regulations for oil tankers (Regulation 530/2012), resulting in a reduction of oil spills; the European Regulation on Invasive Alien Species sets measures to be taken against invasive species (Regulation 1143/2014); and Marine Protected Areas, which cover more than 12% of the marine EU waters (EEA 2020), limit some of the most prejudicial human activities. Although the results of these policies are sometimes insufficient (Aminian-Biquet et al. 2024; Kleitou et al. 2021), some positive effects on biodiversity are already apparent at local and regional scale (Jacquemont et al. 2022). In sum, there are two opposing processes acting on coastal and marine biodiversity: losses caused by human activities and gains caused by conservation efforts.

To ensure the future of coastal and marine ecosystems, we need to assess long-term trends and the current status of coastal and marine biodiversity, including their drivers. By assessing biodiversity trends, we can better understand if biodiversity loss is worsening, stagnant or improving, and if conservation policies are effective. Finally, we also want to determine if the causes of biodiversity loss

and the effectiveness of conservation policies are similar across European regions and for different groups of organisms. To that end, here we assess the biodiversity trends at continental scale and at regional scale, separately for six biotic groups.

### 2.1.3 Methodology

We created six datasets of time series for European coastal and marine biotic groups: birds, fish, invertebrates, macroalgae, phytoplankton and zooplankton. The time series were extracted from open access databases: BioTIME (version 1.0) (Dornelas et al. 2018), COPEPOD (O'Brien 2014), the European Marine Observation and Data Network (EMODnet) (Beja et al. 2024), REPHY (REPHY 2023), FishGlob (Maureaud et al. 2025) and CPR Survey (Vezzulli & Reid 2003); and a previous study analysing multidecadal biodiversity trends in European marine, freshwater and terrestrial ecosystems (Pilotto et al. 2020). Each of the included time series met the following criteria: (1) had a minimum number of eight sampling years (not necessarily consecutive) with one single annual observation (based on a single sampling event for every biotic group, except for phytoplankton); (2) included abundance estimates, and (3) had consistent sampling site, sampling period (any three consecutive months), protocol and taxonomic resolution over the entire observation period (Table 1.1).

In the case of phytoplankton time series, only years with at least eight sampled months and without gaps longer than three months were considered to avoid the effect of seasonality and blooms. The abundance values of the monthly samples were used to calculate the yearly geometric mean, which was later considered as the yearly abundance estimate. For every biotic group, in the case of months with several samples, the first sample was used.

Table 1.1. Number of time series per biotic group and per data source meeting the criteria

| | Birds | Fish | Invertebrates | Macroalgae | Phytoplankton | Zooplankton | TOTAL |
|---|---|---|---|---|---|---|---|
| **BioTIME 1.0** | | 3 | 157 | 28 | 2 | | **190** |
| **COPEPOD** | | | | | | | **0** |
| **Pilotto et al. 2020** | 2 | 1 | 8 | | | | **1** |
| **EMODnet** | 291 | 413 | 1154 | 14 | 116 | 24 | **2012** |
| **REPHY** | | | | | 89 | | **89** |
| **FishGlob** | | 5 | | | | | **5** |
| **CPR Survey** | | | | | 26 | 26 | **52** |
| **Total** | **293** | **422** | **1319** | **42** | **233** | **50** | **2359** |

The final dataset comprises 2,359 time series, including 552,475 observations of 4,718 coastal and marine taxa in 2,246 different sites (Fig. 1.1). Among these sites 25 are in the Adriatic Sea, two in the Aegean-Levantine Sea, 471 in the Baltic Sea, 54 in the Bay of Biscay and the Iberian Coast, 9 in the Black Sea, 127 in the Celtic Seas, 1,515 in the Greater North Sea, 32 in the Western Mediterranean Sea and one in the White Sea. The time series span from 1956 to 2022 with a mean total duration of 18 years and a mean of 15 sampled years (minimum 8, maximum 59 sampled years).

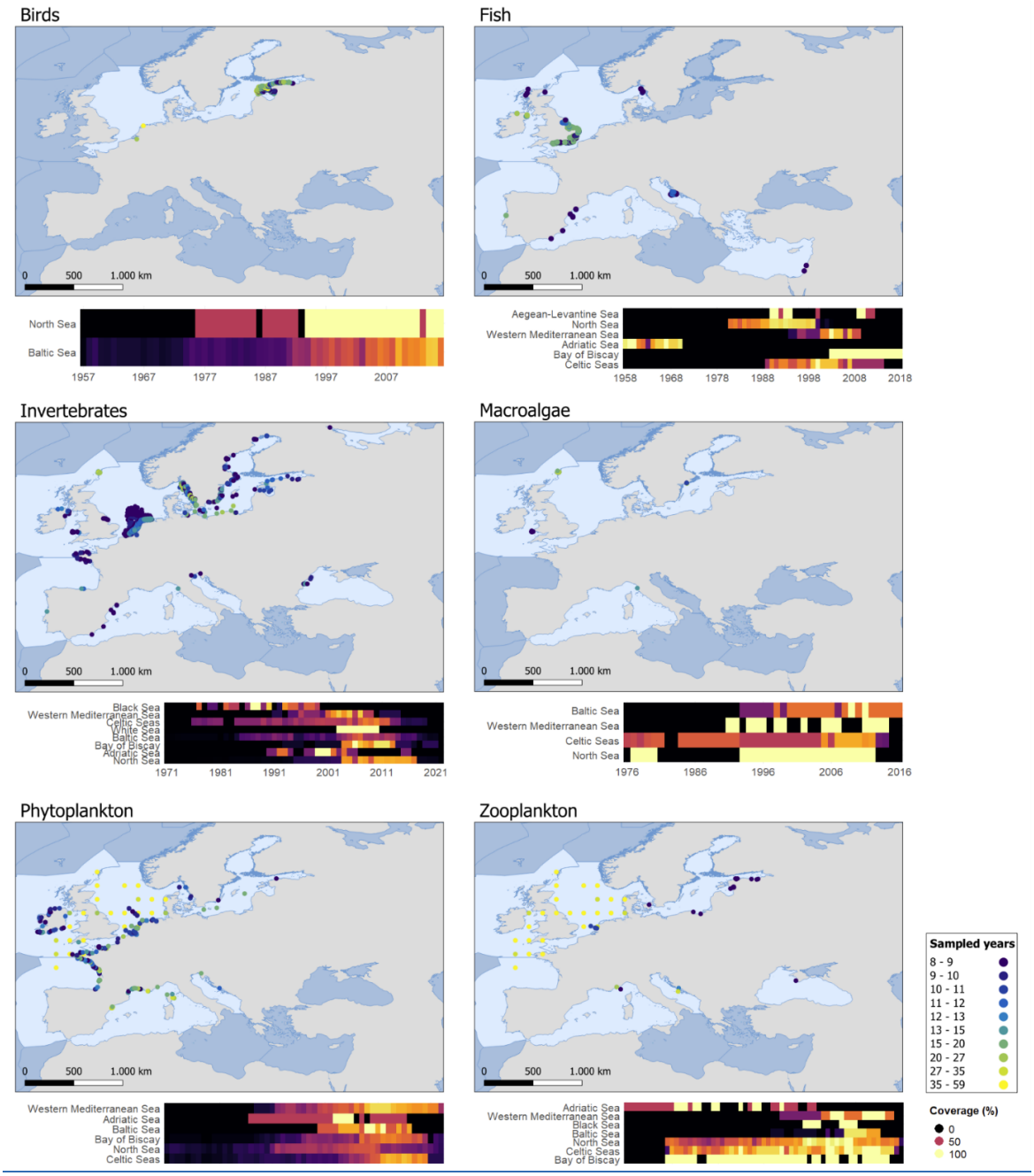



Figure 1.1. Location of each time series (sites) and number of sampled years for each of them, by biotic group.

Taxa names and taxa identification numbers were harmonized using the World Register of Marine Species (WoRMS) (WoRMS Editorial Board 2025). To describe changes in the communities we calculated taxonomic richness, Shannon's diversity and total abundance for each time series and year. Total abundance was $log_{10}$-transformed prior to analysis to reduce skewness.

We used Generalized Least Squares (GLS) models to evaluate time series-level temporal trends in richness, diversity and abundance. For each time series and metric, we fitted a GLS model with year as a continuous fixed effect, while accounting for temporal autocorrelation. The percentage of change per year was calculated by back-transforming model estimates.

To detect differences in the trends across geographical regions, the time series were grouped into five regions (Atlantic, Arctic, Baltic, Black Sea and Mediterranean). The resulting slopes of the GLS models were meta-analysed, including their corresponding sampling variances, the region as moderator and the site and data source as random effects to account for site-level variation and differences in sampling methodologies. The models were fitted via restricted maximum likelihood.

### 2.1.4 Results and discussion

Most of the time series do not show significant changes over time in any biotic group. Overall, more time series show positive trends in each biodiversity metric than negative, but this is not consistent across biotic groups. For example, there are twice as many time series with a negative trend than positive for fish, while there are almost three times more positive than negative trend series for invertebrates (Table 1.2).

Table 1.2. Number and percentage (in brackets) of time series showing significantly positive or negative trends in richness, diversity (Shannon diversity index) and abundance (transformed by the common logarithm), by biotic group. The remaining number and percentage of time series corresponds to the ones that did not significantly change overtime.

| Biotic group | n (time series) | Trend | Richness | Shannon Diversity | Abundance ($log_{10}$) |
|---|---|---|---|---|---|
| **Birds** | 293 | Positive | 95 (32%) | 72 (25%) | 65 (22%) |
| | | Negative | 14 (5%) | 15 (5%) | 37 (13%) |
| **Fish** | 422 | Positive | 15 (4%) | 16 (4%) | 13 (3%) |
| | | Negative | 40 (9%) | 44 (10%) | 38 (9%) |
| **Invertebrates** | 1319 | Positive | 252 (19%) | 190 (14%) | 290 (22%) |
| | | Negative | 95 (7%) | 85 (6%) | 109 (8%) |
| **Macroalgae** | 42 | Positive | 2 (5%) | 5 (12%) | 8 (19%) |
| | | Negative | 9 (21%) | 4 (10%) | 17 (40%) |
| **Phytoplankton** | 233 | Positive | 73 (31%) | 17 (7%) | 36 (15%) |
| | | Negative | 33 (14%) | 17 (7%) | 15 (6%) |
| **Zooplankton** | 50 | Positive | 29 (58%) | 7 (16%) | 2 (4%) |
| | | Negative | 3 (6%) | 8 (14%) | 15 (30%) |
| **Total** | **2359** | **Positive** | **466 (20%)** | **307 (13%)** | **394 (17%)** |
| | | **Negative** | **194 (8%)** | **173 (7%)** | **231 (10%)** |

The frequency distributions of the temporal trends (Fig. 1.2) show that most trends follow a normal distribution centred around zero, in particular the diversity and the abundance trends. This is, in most time series the biodiversity metrics do not change across time, and if they do, the change is relatively small.

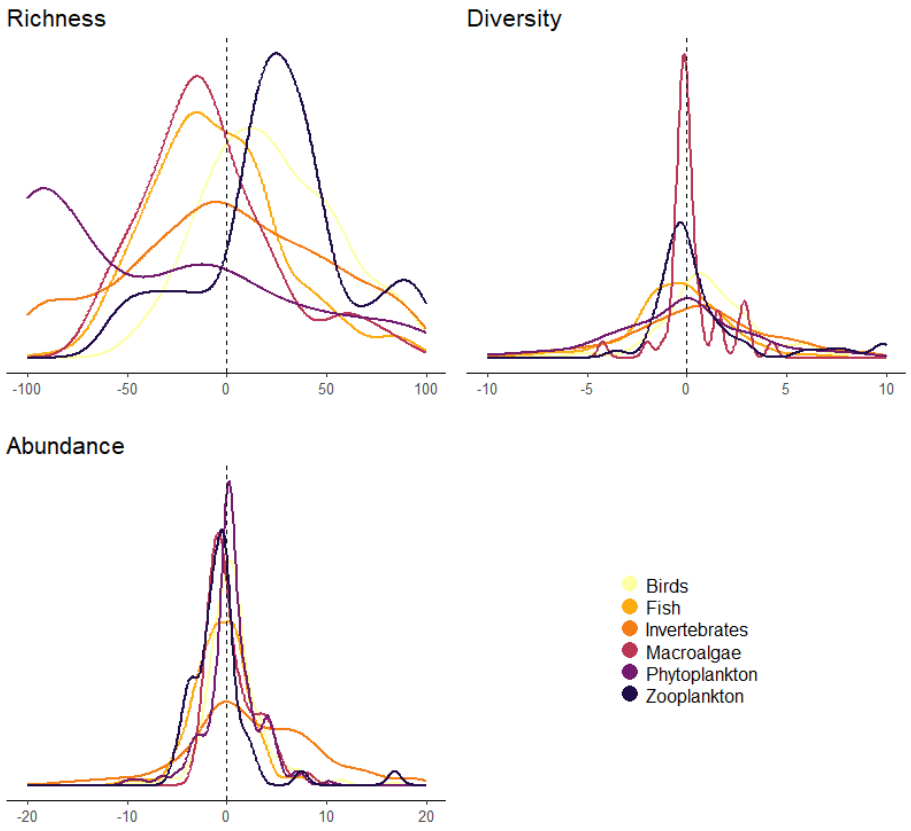



Figure 1.2. Frequency distribution of the change in richness, diversity (Shannon diversity index) and abundance (transformed by the common logarithm) for each biotic group. The horizontal axis shows the percentage of change, from -100 to 100 for richness, -10 to 10 for diversity and -20 to 20 for abundance (-100 meaning complete decline, 0 no change and 100 a value twice as high as the initial value) during the complete sampled period.

The regional meta-analyses show several significant trends across biodiversity metrics and biotic groups (Table 1.3), most of them positive. The clearest regional trends are the increases in bird richness and diversity in the Baltic, the fish abundance decline in the Atlantic, and the increase in invertebrate richness, diversity and abundance in the Baltic. Some indicators show different trends for the same group in different regions: fish abundance (negative trend in the Atlantic but positive in the Mediterranean); and zooplankton abundance (negative trend in the Atlantic and Baltic but positive in the Black Sea).

Table 1.3. Regions showing a significant trend, by biotic group. Metrics and regions not included do not show any significant trend. There were no significant trends for macroalgae. In **bold** the trends with more than 30 time series. n indicates the total number of time series for the biotic group in the region.

| | Metric | Region | n (time series) | Trend |
|---|---|---|---|---|
| **Birds** | Richness | Atlantic | 2 | Positive |
| | | **Baltic** | **291** | **Positive** |
| | Diversity | **Baltic** | **291** | **Positive** |
| **Fish** | Abundance | **Atlantic** | **395** | **Negative** |
| | | Mediterranean | 27 | Positive |
| **Invertebrates** | Richness | **Baltic** | **166** | **Positive** |
| | Diversity | **Baltic** | **166** | **Positive** |
| | | Mediterranean | 11 | Positive |
| | Abundance | **Baltic** | **166** | **Positive** |
| **Phytoplankton** | Diversity | Mediterranean | 24 | Negative |
| | Abundance | Mediterranean | 24 | Positive |
| **Zooplankton** | Richness | Baltic | 14 | Positive |
| | Diversity | Baltic | 14 | Positive |
| | | Black Sea | 1 | Positive |
| | Abundance | Atlantic | 30 | Negative |
| | | Baltic | 14 | Negative |
| | | Black Sea | 1 | Positive |

These results can be summarized as follows: first, there is a high variability in data coverage that limits the scope of this analysis, and second, the biodiversity trends mostly show stability in the communities although there are some differences across biotic groups and regions.

The data coverage is highly variable geographically (more than half of the sites are in the Greater North Sea) and biologically (invertebrates represent more than half of the time series analysed). This limits the reliability of the comparisons between regions or groups and hinders the interpretation of the results.


Accordingly, the higher number of time series for the North Sea and the Baltic Sea make the observed stable trends of biodiversity metrics more reliable compared to the trends for other regions, which are less supported by a sufficient number of time series. Moreover, differences in regional biodiversity trends are commonly a result of differences among regional drivers (e.g. negative anthropic pressures or positive effects of conservation actions). Accordingly, up-scaling our results to continental scale may result in wrong conclusions.

The generally stable biodiversity trends indicate that no widespread loss of biodiversity was detected, but there are some differences between groups and regions. As the reliability of the results is heavily dependent on the number of time series analysed, we will limit the discussion to biotic groups with more data (i.e. invertebrates, fish, phytoplankton and birds).

On the one hand, for invertebrates and birds there are more time series following positive trends, and the significant regional trends (in the Baltic) are also positive. This can indicate that the conservation policies adopted in the Baltic are effective (Cano-Barbacil et al. 2025; Andersen et al. 2017; Sköld et al. 2025).

On the other hand, for fish there are more time series following negative trends, and the significant regional trends are negative in the Atlantic but positive in the Mediterranean. The fishing pressure is still high in the Atlantic (Froese et al. 2018), and its effects could be worse for the most common species (as richness and diversity trends are stable). The positive trend in the Mediterranean is likely observed because most of the time series analysed in the region (18 out of 27 time series) are from an older time period (between 1956 and 1971) when the anthropic pressures were not as intense (Fiorentino & Vitale 2021).

For the phytoplankton the results are mixed, and there is a low number of time series in the only region with significant trends (the Mediterranean). Some of the taxonomical groups within the phytoplankton community can react positively to eutrophication and water warming (Suikkanen et al. 2013), while others could be declining, explaining the increases in abundance but loss of diversity.

Finally, in this section we presented a descriptive analysis of the coastal and marine biodiversity trends based on available public time series data. Overall, we did not detect a widespread biodiversity loss in European coastal and marine communities. However, we caution overinterpretation of this result as our analyses are biased by severe biodiversity time series data gaps regarding both biotic groups and regions. Moreover, compared to linear models, non-linear models may provide more nuanced insights into biodiversity trends, as they can capture fluctuations in trends over time (for example, a decrease followed by an increase). Thus, additional data and analyses are required to provide more robust trends and detailed trajectories as well as correlations of biodiversity trends with trends of their drivers.

## 2.2 Habitat suitability modelling

### 2.2.1 Objective

Based on available large-scale spatial data, we model habitat suitability requirements of coastal and marine species protected under Habitat and Bird Directive legislation to understand their present and predict their future distribution using ensemble models.

### 2.2.2 Background

The European Union's Habitats Directive and Birds Directive (Council directive 92/43/EEC; Directive 2009/147/EC) are essential frameworks in the battle against biodiversity loss. They aim to preserve characteristic habitats with their species of interest. Annexes II and IV of the Habitats Directive lists marine mammals, including harbour porpoises and bottlenose dolphins, that require strict protection. To support effective management and conservation under these directives it is essential to understand both the current distribution of these species and how their habitats may shift in the future under climate change.

Habitat suitability models (HSMs) are commonly used to understand species habitat requirements and distributions. HSMs fit the statistical relationship between observed species occurrences and environmental conditions, allowing predictions of where relative suitable habitat is located across a study area. Numerous modelling algorithms and implementations exist, each with their own strengths and weaknesses. Ensemble models combine these individual models, leveraging their differences, to obtain more robust predictions.

Public biodiversity databases such as (Eur)OBIS and GBIF (EurOBIS; GBIF 2025) provide extensive records of species occurrences. Nonetheless, these data are often collected with varying methods and for specific sampling/research purposes, resulting in data being unevenly distributed in space and time. Comprehensive environmental datasets are openly available from infrastructures such as EMODnet (Beja et al. 2024) (e.g. bathymetry) and the Copernicus Marine Environment Monitoring Service (CMEMS) (Copernicus Marine Service 2025) (e.g. temperature, salinity, productivity) and complementing the biological observations. These combined resources make it possible to assess how environmental drivers shape species distributions today and to project potential changes under future climate scenarios.

### 2.2.3 Methodology

Two modelling approaches are implemented to generate habitat suitability predictions with two temporal scopes: (1) monthly predictions to characterise seasonal dynamics; and (2) decadal predictions to assess long-term trends under Shared Socioeconomic Pathways (SSP, Riahi et al. 2017) climate scenarios (2020-2100).

The study focused on regions II, III and IV of the OSPAR Maritime Area, which include the Greater North Sea, the Celtic Seas, and the Bay of Biscay and Iberian Coast. These regions were chosen because of the availability of both biological and environmental data required for the modelling approach, and because they cover key habitats for many of the species protected under the Habitats Directives.

**Biological data** - Species occurrence data were obtained from EurOBIS, the European node of the Ocean Biodiversity Information System, which provides open-access records of marine species occurrences. We focused on four marine mammal species of conservation interest: harbour porpoise (*Phocoena phocoena*), harbour seal (*Phoca vitulina)*, common dolphin (*Delphinus delphis*), and bottlenose dolphin (*Tursiops truncatus*). These were selected from the species listed in Annex II and IV of the Habitats Directive (f Council Directive 92/43/EEC), with additional filtering to ensure sufficient data availability. Records were downloaded from EurOBIS for the years 2000–2019, and quality control steps were applied to remove strandings, museum specimens, duplicates, and records outside the study area. To reduce bias from oversampled locations or periods, records were thinned so that only one presence per grid cell and time step (month or decade) remained.

Because species occurrence data from public databases are often biased towards areas of higher sampling effort, we applied a background sampling approach that accounts for this. All available records of marine mammals were used to estimate sampling intensity, and background points were drawn with higher probability in areas of higher effort.

**Environmental data** - Sea surface temperature, salinity, and net primary productivity were obtained from the CMEMS (Copernicus Marine Service Information; Global ocean low and mid trophic levels biomass content hindcast) while bathymetry was downloaded from EMODnet. Environmental data were collected at a spatial resolution of 0.083 degrees. To comply with the temporal resolution of the other environmental variables, net primary productivity was aggregated into monthly averages, and both monthly and decadal climatologies were created. For the decadal predictions, environmental layers were downloaded from Bio-ORACLE (Tyberghein et al. 2012; Assis et al. 2024). The present data was downloaded for decades 2000-2010 and 2010-2020. The future projections were downloaded for decades between 2020-2030 and 2090–2100 under a range of climate change scenarios (SSP1-1.9, SSP2-4.5 and SSP5-8.5) (Riahi et al. 2017).

**Modelling approach** - Habitat suitability modelling was carried out using an ensemble approach, combining several different algorithms that are commonly used, including Random Forest, Multivariate Adaptive Regression Splines (MARS), Maximum Entropy (MaxEnt), Extreme Gradient Boosting (XGBoost) and Generalized Additive Models (GAMs). Models were trained using subsets of the biological and environmental data and evaluated using cross-validation to estimate their performance in the absence of a true validation test set.

After training, the final ensemble models were fitted on all available data and used to generate habitat suitability maps across the study area. Predictions were made for both monthly and decadal climatologies of the present period, as well as for future decades under the different SSP climate scenarios. Habitat suitability was expressed on a relative scale from 0 to 1, indicating areas that are less or more suitable for the species.


### 2.2.4 Results and discussion

#### *2.2.4.1 Species occurrences across the study area*

The occurrence data for harbour porpoises (Appendix A, Figure A1) shows clear variation across months, both in terms of spatial coverage and intensity. Some months and areas have far more records than others. This reflects well-known spatial and temporal biases arising from using public datasets originally collected for specific purposes, and not necessarily in a harmonised manner across the spatial and temporal scales covered for these models: for example, there is more sampling effort along shipping routes and close to the coast. When records are aggregated to the decadal scale (as described in section 2.2.3), these differences even out to some degree, reducing the impact of temporal gaps for certain areas. Nevertheless, some regions remain overrepresented to others considering their sampling effort.

Occurrences of harbour seals in the datasets used are mostly restricted to coastal areas around the North Sea, with few records in central parts of the North Sea (Appendix A, Figure A2). Occurrences of common dolphins covered predominantly the Iberian coast and the Bay of Biscay, extending into the Celtic Seas during summer, and with very scarce records present in the North Sea (Appendix A, Figure A3). Bottlenose dolphin occurrences largely resemble those of common dolphins, with very limited records in the North Sea (Appendix A, Figure A4).

Public databases such as (Eur)OBIS include many datasets collected over different time spans, each with their own methodologies and sampling designs. This richness is valuable, but it also introduces inconsistencies when multiple data sources are used for a study. The approach taken here was to simplify these records to a grid-level presence, which inevitably loses detail but makes it possible to combine the available data into a format usable for large-scale modelling.

#### *2.2.4.2 Model evaluation*

Model evaluation suggests that the ensemble approach was able to distinguish presence from background reasonably well based on the Continuous Boyce Index and ROC AUC under cross-validation. The metrics provide an indication of reliability, but it is important to stress that we lack an independent validation dataset. When using these performance metrics with presence-only data, they depend on the amount of background points and the method used to generate them.

#### *2.2.4.3 Monthly habitat suitability predictions*

Monthly habitat suitability predictions are able to inform about habitat suitability and possible usage by a species in the short-term, providing insights valuable to inform human activities and marine spatial planning in current scales. The monthly prediction maps show changes in habitat suitability throughout the year for all four species.

For **harbour porpoises** (Figure 2.1), the Celtic Seas and wider North Sea are generally predicted as suitable throughout the year, with highest habitat suitability during the winter predicted in the southern and more coastal parts of the North Sea, matching well with the occurrences of the species (Appendix A, Figure A1).

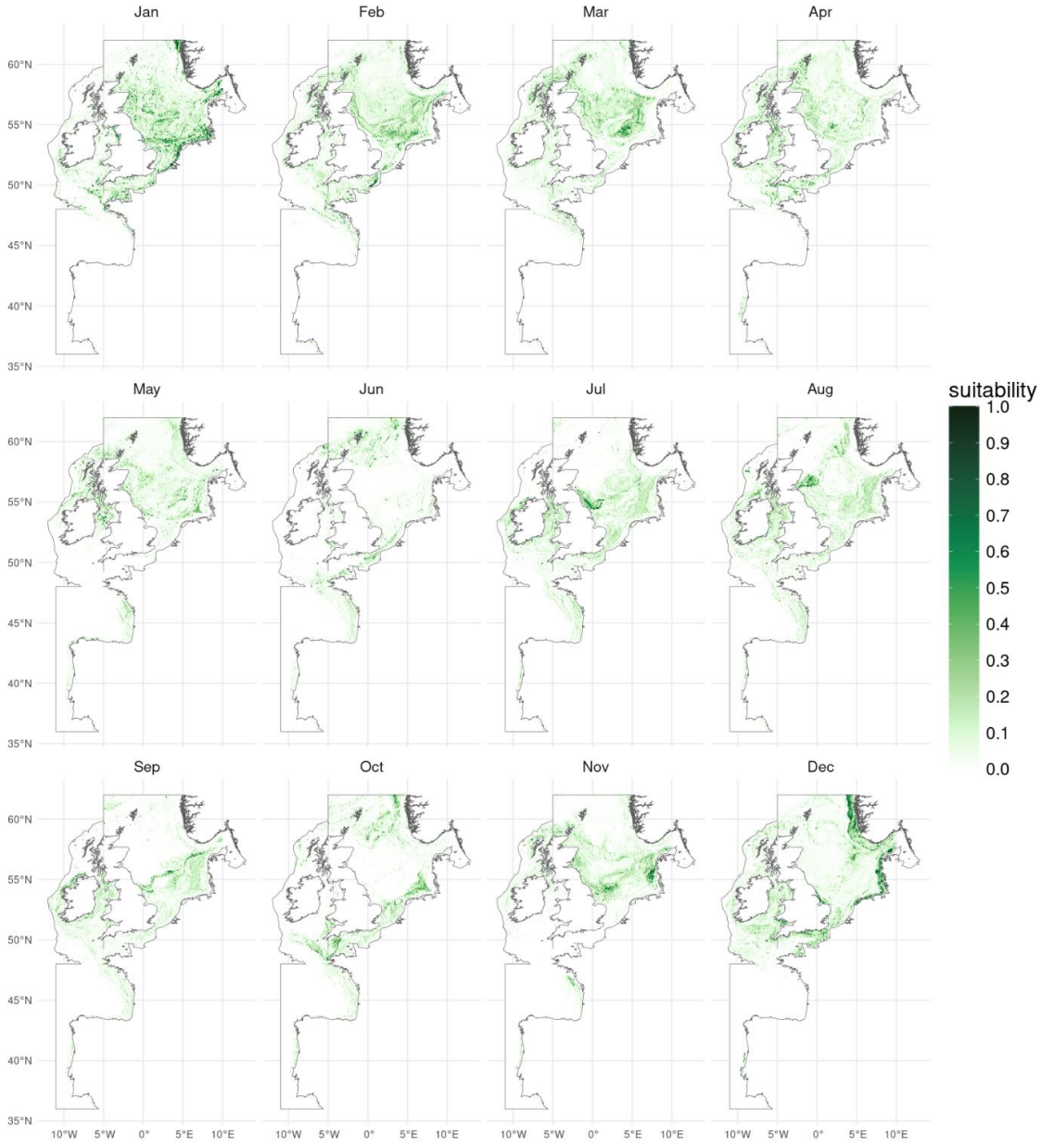



Figure 2.1. Harbour porpoise predicted monthly habitat suitability in OSPAR regions II-IV.

Predicted habitat suitability of **harbour seal** (Figure 2.2) shows highest values restricted to coastal areas, with few obvious seasonal changes, such as higher predicted habitat suitability in the central North Sea during the summer months. Nevertheless, these changes in summer habitat suitability may be largely influenced by some limited occurrences in the dataset further off-shore in central parts of the North Sea (Appendix A, Figure A2) and should be interpreted with caution.

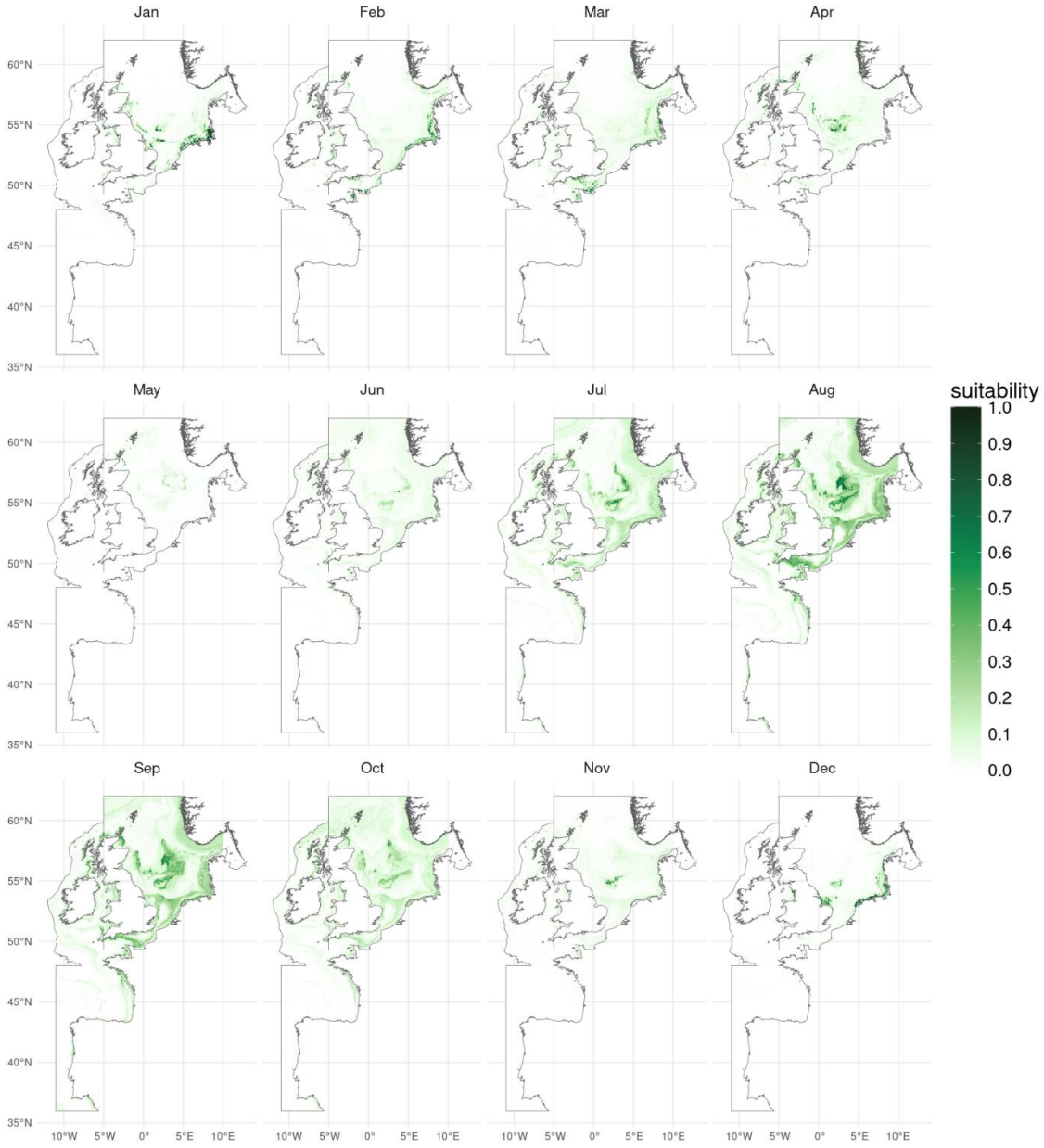



Figure 2.2. Harbour seal predicted monthly habitat suitability in OSPAR regions II-IV.

For **common dolphins** (Figure 2.3), high habitat suitability is predicted along the Iberian coast and the Bay of Biscay throughout the year. During spring and summer, predicted suitable habitat extends further north and across the North Sea, despite occurrences in the dataset being scarce in this area during the summer (Appendix A, Figure A3). Northern records seem to strongly influence the monthly models showing a much higher suitability in these periods.

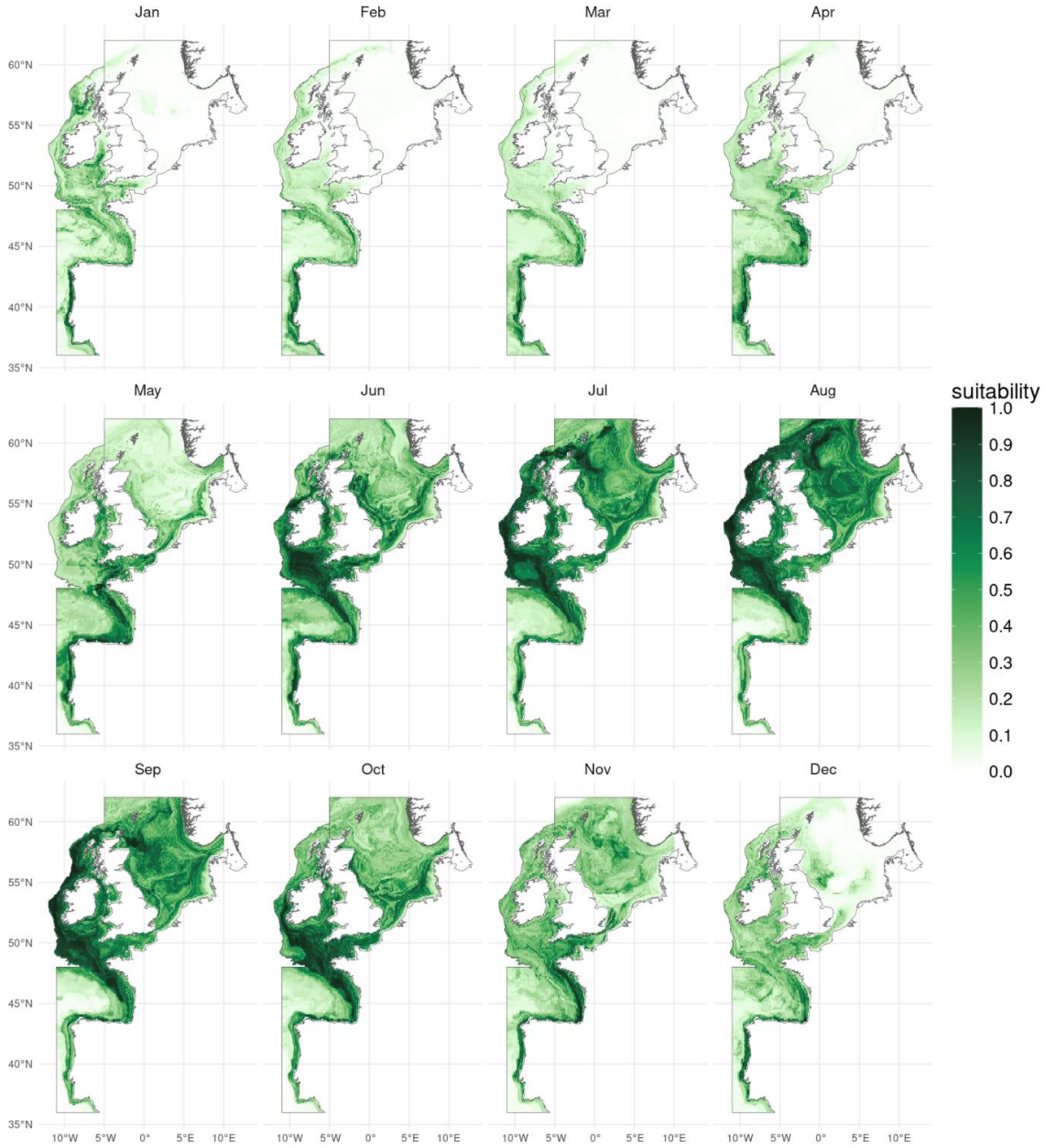



Figure 2.3. Common dolphin predicted monthly habitat suitability in OSPAR regions II-IV.

For **bottlenose dolphins** (Figure 2.4), predicted habitat suitability largely resembles that of common dolphin, with high values along the Iberian coast and the Bay of Biscay throughout the year and suitable habitat extending into the North Sea during the summer. In the same way as with common dolphin, sensitivity to the records in the north (Appendix A, Figure A4) coming out of the incidental sightings dataset may be influencing the predictions.

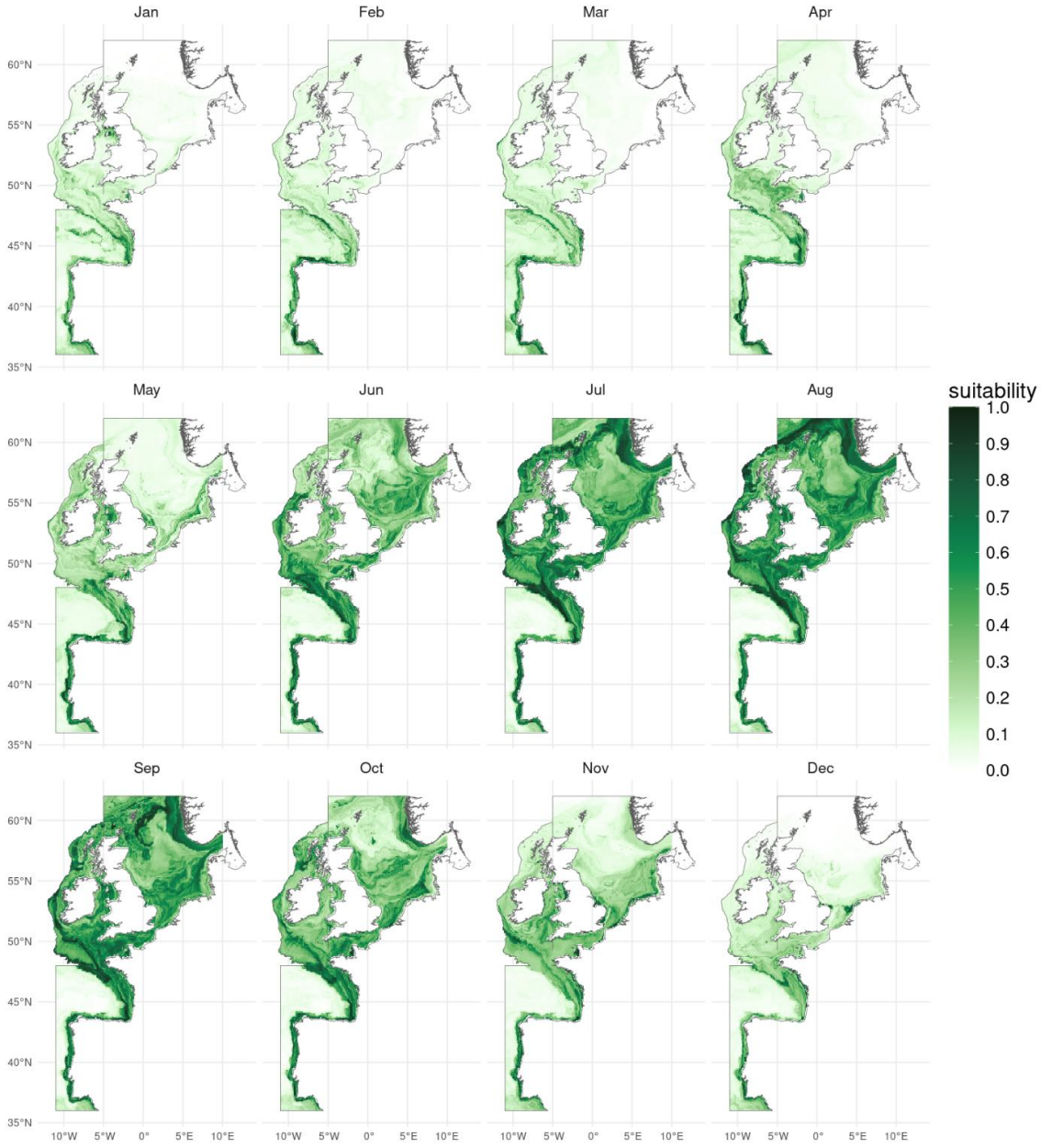



Figure 2.4. Bottlenose dolphin predicted monthly habitat suitability in OSPAR regions II-IV.

#### *2.2.4.4 Decadal habitat suitability predictions for current and future climate scenarios*

While monthly habitat suitability predictions can inform about habitat suitability and possible usage by a species in the short-term, predictions aggregated across longer scales (such as decadal) allow to better understand habitat suitability for a species linked to a future changing climate.

Under present conditions, highest habitat suitability for harbour porpoise is predicted across the North Sea, the English Channel and the Celtic Sea (Appendix A, Figure A5), matching well with the species occurrence records (Appendix A, Figure A1). Future projections under SSP climate scenarios (Figure 2.5) suggest an overall reduction in suitable habitats across decades and under all future SSP

scenarios, possibly linked to increasing sea surface temperatures. The strongest declines are predicted under the SSP5-8.5 scenario, in which the currently highly suitable areas in the southern North Sea (Appendix A, Figure A5) would suffer a large reduction in habitat suitability (Figure 2.5).

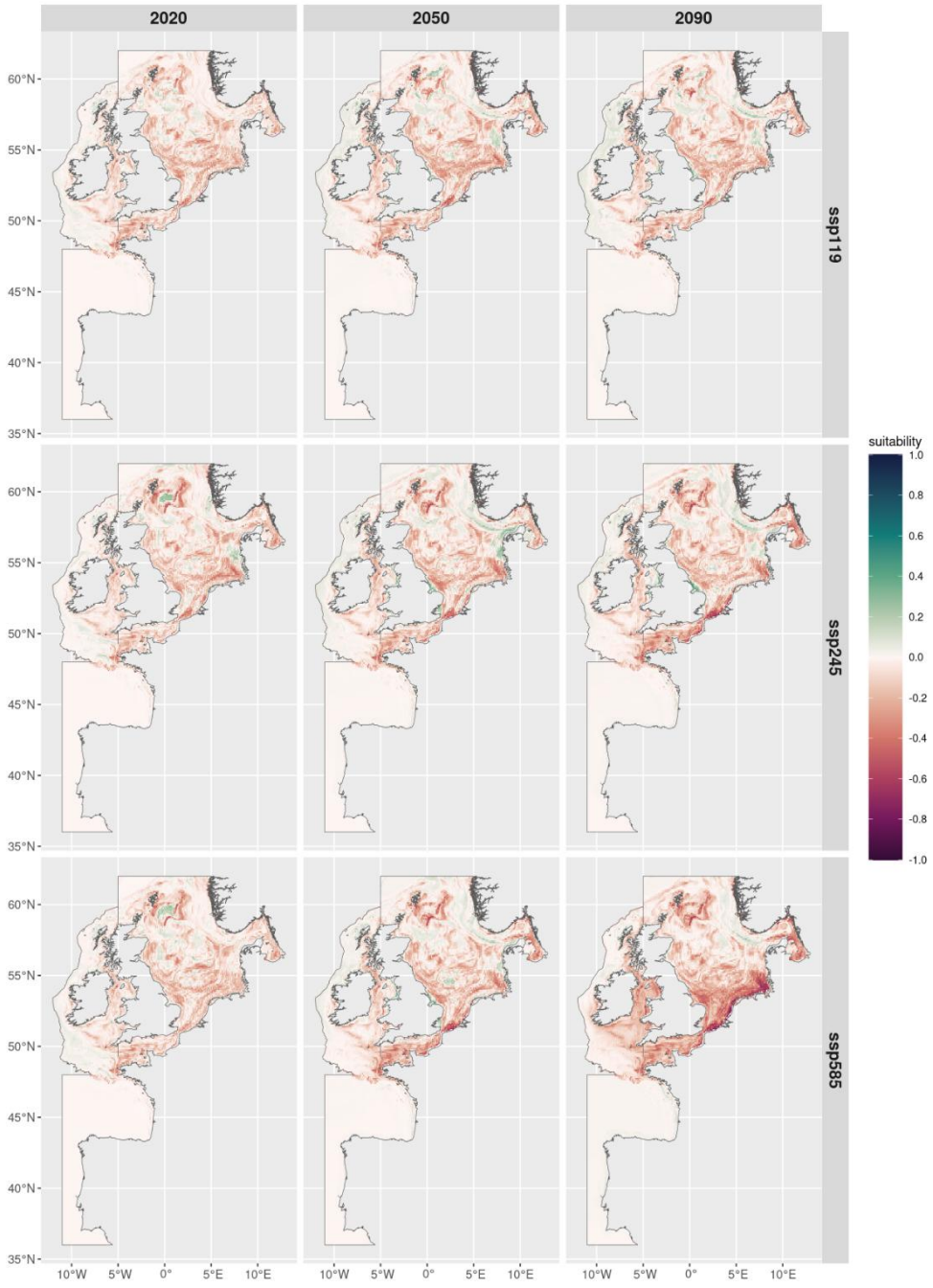



Figure 2.5. Harbour porpoise habitat suitability differences between current decadal predictions and future decadal predictions under climate scenarios SSP119 (top row panels), 245 (middle) and 585 (bottom) for decades 2020-2030 (left panels), 2050-2060 (centre) and 2090-2100 (right). Predicted

habitat suitability for decade 2010-2019 and for all future scenarios are available in Appendix A, Figure A5 and Figure A9, respectively.

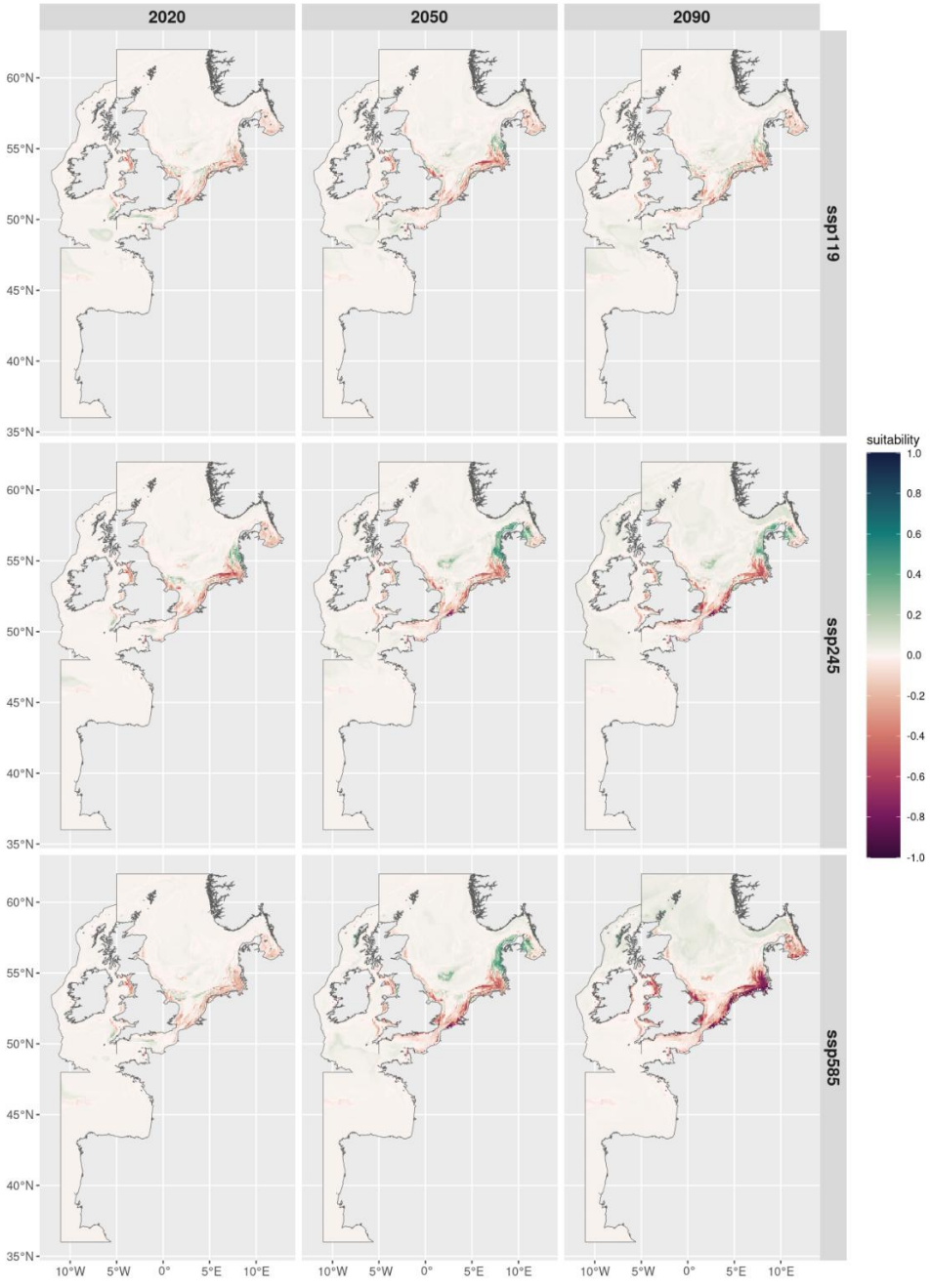



Figure 2.6. Harbour seal habitat suitability differences between current decadal predictions and future decadal predictions under climate scenarios SSP119 (top row panels), 245 (middle row) and 585 (bottom row) for decades 2020-2030 (left panels), 2050-2060 (centre) and 2090-2100 (right). Predicted habitat suitability for decade 2010-2019 and for all future scenarios are available in Appendix A, Figure A6 and Figure A10, respectively.

Predicted habitat suitability for **harbour seal** during the last two decades (Appendix A, Figure A6) is also highest in the southern North Sea, matching well with the species occurrences (Appendix A, Figure A2). In the coming decades, predicted habitat suitability for harbour seals is also expected to decrease across all SSP scenarios in the Southern North Sea, with most marked reductions under the SSP5-8.5 scenario (Figure 2.6). Nevertheless, habitat suitability seems to be predicted to increase in previously less suitable habitats in the eastern North Sea around Denmark (Figure 2.6), likely due to a change to more suitable environmental conditions for the species around that area.

Predicted habitat suitability during the last two decades for **common dolphin** and **bottlenose dolphin** was highest around the Iberian coast and the Bay of Biscay (Appendix A, Figure A7 and Figure A8, respectively). Future projections suggest a marked reduction in habitat suitability around the Iberian coast and the Bay of Biscay for both species (Figure 2.7 and 2.8). Surprisingly, habitat suitability is predicted to increase for common dolphins in some areas further north (i.e. North Sea), which currently has a very low habitat suitability for this species (Appendix A, Figure A7). The changes are even more striking for bottlenose dolphins, with positive changes in habitat suitability in the North Sea and offshore areas of the Atlantic. The changes for both species seem to reflect temperature-driven changes under the different future SSP scenarios.

Overall, habitat suitability for all species modelled (harbour porpoise, harbour seal, common dolphin, and bottlenose dolphin) appears to decrease across all the future SSP scenarios used, with currently suitable areas where these species occur being impacted in the future. Despite some increases in predicted habitat suitability in some specific areas in the future (e.g. for common and bottlenose dolphins), the predictions of these models are only based on environmental conditions and ignore any biological links. Biological variables, such as the associated changes in food availability in new and completely different ecological areas, or the impact on reproductive grounds, are key to understanding the future real impact on the populations. For example, habitat shifts favouring off-shore or open ocean environments, as opposed to coastal areas currently used by the species, may translate into changes in food availability which may prevent the actual use of the predicted suitable habitat in the future, for example. Thus, the outcomes of these models should be interpreted with caution, as they do not incorporate the full complexity of biological systems and therefore are limited in the predictions that they can provide.

These results demonstrate both the potential and the limitations of habitat suitability models based on openly available data. Biases in sampling, differences between datasets, and the aggregation of information to a monthly or decadal resolution all reduce precision. Environmental variables such as temperature or net primary productivity serve as proxies for complex ecological processes. Adding the uncertainty of environmental values over these future climate scenarios, the maps should be viewed as broad indications of trends rather than precise predictions of where species will be found. At the same time, this modelling exercise highlights the value of public biodiversity databases. The ability to build these data-driven models depends on the availability of quality datasets. The more consistent and comprehensive these records become, the more reliable the predictions will be.

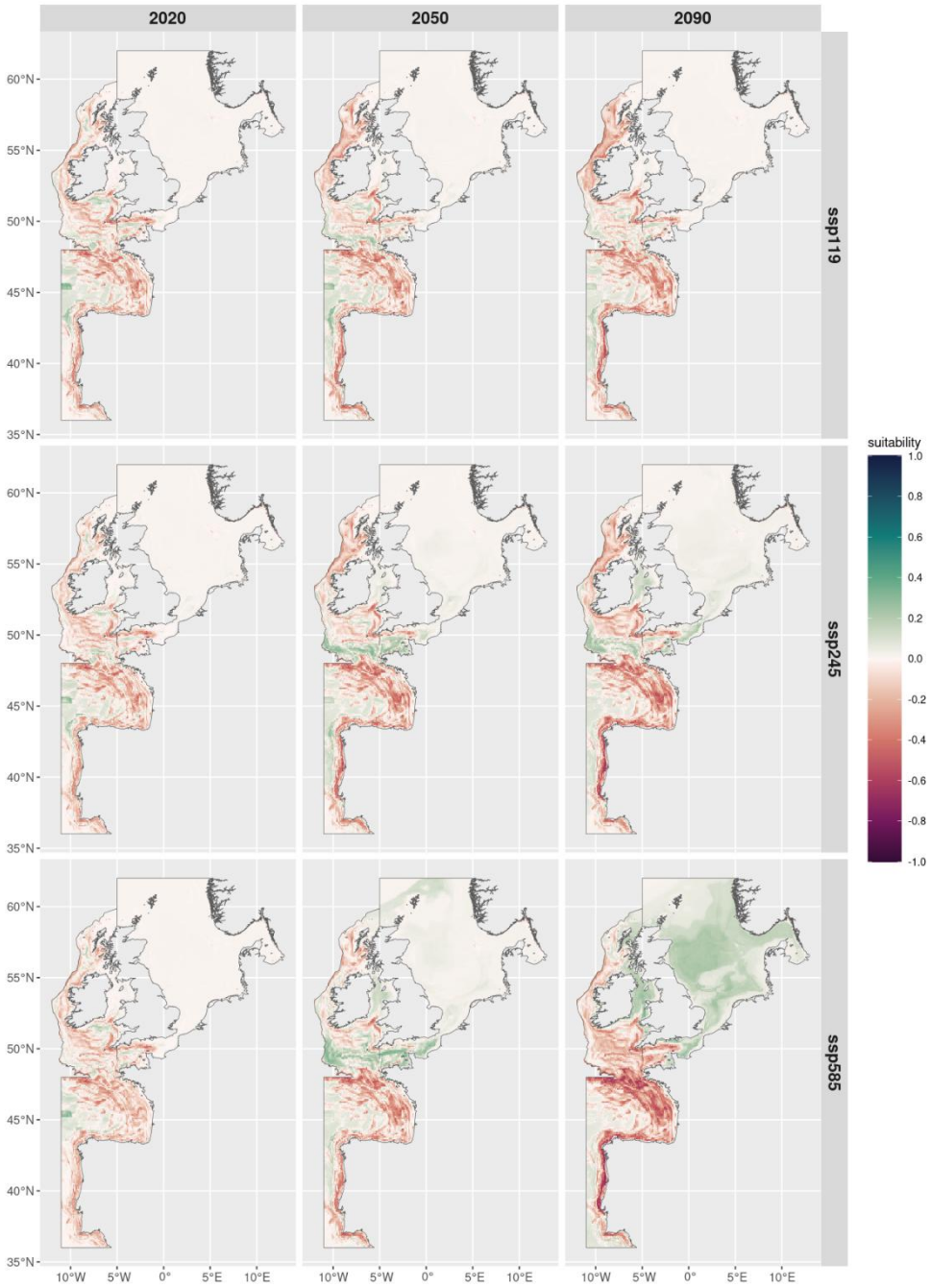



Figure 2.7. Common dolphin habitat suitability differences between current decadal predictions and future decadal predictions under climate scenarios SSP119 (top row panels), 245 (middle row) and 585 (bottom row) for decades 2020-2030 (left panels), 2050-2060 (centre) and 2090-2100 (right). Predicted habitat suitability for decade 2010-2019 and for all future scenarios are available in Appendix A, Figure A7 and Figure A11, respectively.

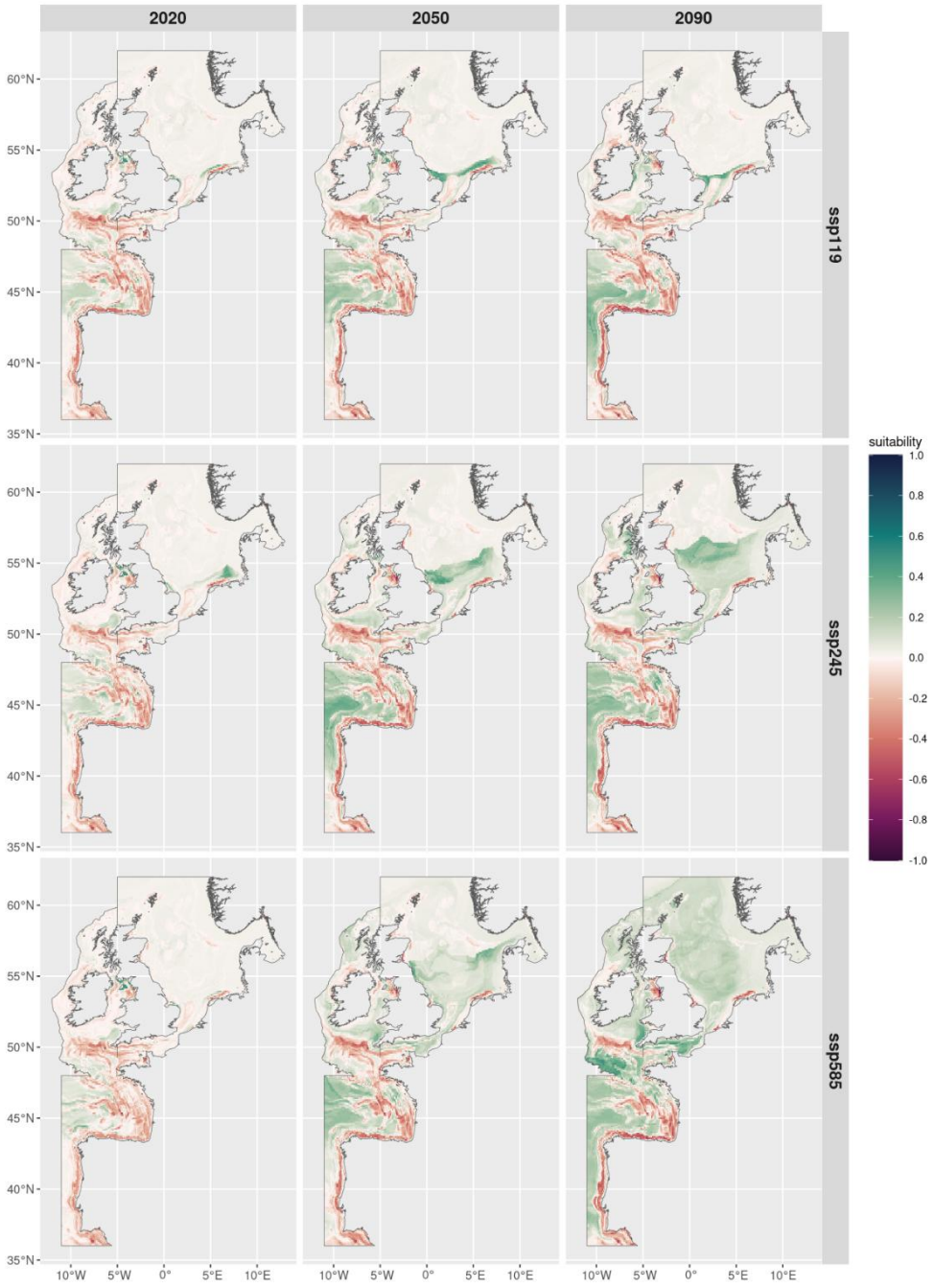



Figure 2.8 Bottlenose dolphin habitat suitability differences between current decadal predictions and future decadal predictions under climate scenarios SSP119 (top row panels), 245 (middle row) and 585 (bottom row) for decades 2020-2030 (left panels), 2050-2060 (centre) and 2090-2100 (right). Predicted habitat suitability for decade 2010-2019 and for all future scenarios are available in Appendix A, Figure A8 and Figure A12, respectively.

## 2.3 Determining causal relationships

### 2.3.1 Objective

The objective of this section is to combine high resolution oceanic numerical modelling with 30-years of atmospheric, oceanographic and ecological data at a coastal LTER site to derive causal relationships between the external forcing and the biological response.

Additionally, we aim to develop AI predicting tools which, in the future, will allow targeting specific biological components of the ecosystems by integrating the outcome of the analysis of local genomic data from the NEREA Augmented Observatory (WP2). Once tuned and adapted, these AI tools can serve as a downscaling toolbox for future projections.

### 2.3.2 Background

A recent study on nutrient dynamics in the Gulf of Naples (GoN) at the Long-Term Ecological Research MareChiara site (LTER-MC) (Romillac et al., 2023) provided a detailed analysis of how anthropogenic pressures and climate variability shape nutrient input and phytoplankton biomass. Based on over three decades of data, it highlighted the interplay between terrestrial nutrient inputs and atmospheric forcing. Further, Kokoszka et al. (2023) observed a long-term shallowing of the mixed layer depth and linked it to changes in wind patterns and freshwater input, making it a key indicator of system responses to climatic changes. Based on these two publications, we designed and run a study to explore the ecosystem evolution under future climate scenarios.

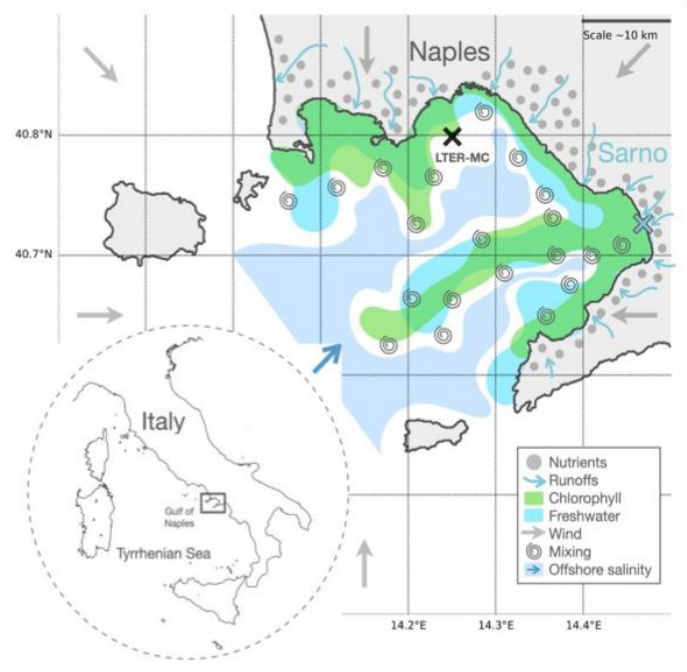



Figure 3.1. The Gulf of Naples, in the Western Mediterranean Sea. Back cross: LTER-MC sampling site; blue cross: Sarno's river mouth.

Figure 3.1 illustrates the interplay between terrestrial influences, atmospheric forcing, and circulation in shaping coastal dynamics and primary productivity in the Gulf of Naples. Key factors considered include freshwater inputs, quantified using proxies such as freshwater content, runoff, and precipitation. Freshwater inflows, particularly from rivers, generate plumes and filaments (Iermano et al. 2012), which are transported by the wind-driven circulation, with stratification controlling their dilution. Vertical mixing disperses these water parcels, alternating with periods when local conditions are shaped by off-shore circulation. Our first aim is to determine whether this dynamic interplay, which governs nutrient distribution and coastal primary production, can be effectively hindcasted or forecasted using a limited set of variables.

Using a hybrid statistical approach that combines mechanistic and machine learning models, we integrate long-term data from in situ observations with reanalysis data. Since chlorophyll-a is a key proxy for phytoplankton productivity (Falkowski & Kiefer 1985) and, indirectly, food web responses, we analyse its variations in relation to freshwater inputs and stratification as indicators of the Gulf of Naples's physical and geochemical dynamics. Leveraging long-term monitoring data, we apply Random Forest algorithms to assess these relationships. The preliminary dynamical analysis relies on the most complete dataset available, while a reduced set of key variables is successively used to set up our predictive model, ensuring both comprehensive analysis and practical forecasting.

We aim, first, to produce a mechanistic understanding of the response of the coastal ecosystem to external forcing and, second, to estimate the possible ecosystem responses to climate change.

### 2.3.3 Methodology

The LTER-MC site (14.25°E, 40.80°N) is located in the Gulf of Naples over a depth of ~75 m (Fig. 3.1). It has been sampled since January 1984, but to ensure standardised data, we focus on the period from January 2001 to February 2020, when chlorophyll-a and nutrient sampling was conducted regularly on a weekly basis. Hourly data for horizontal wind speed components and precipitation was retrieved from VHR-REA_IT Copernicus ERA5 dataset (Raffa et al. 2021). As a measurement of stratification, the Mixed Layer Depth (MLD) was calculated based on Ekman dynamics using wind data. Freshwater Runoff was estimated as the freshwater index FW (Kokoszka et al. 2023), calculated using the averaged river discharge rates $r$ ($m3$ $s$−1) extracted from the European Flood Awareness System provided by Copernicus (Wetterhall 2019). We use high-quality flagged data for three groups of variables: (i) chlorophyll-a (CHL), (ii) physical variables, including temperature (T), salinity (S), precipitation (P), wind velocity (|U|), wind direction (dir), mixed layer depth (MLD), and freshwater runoff (FW), and (iii) biogeochemical variables, including dissolved oxygen (DO), ammonium ($NH_4$), nitrate ($NO_3$), nitrite ($NO_2$), dissolved inorganic nitrogen (DIN), phosphate ($PO_4$), and silicate (Si). We will refer to this ensemble of data as the "Full dataset".

For the 2020-2070 period, no chlorophyll or biogeochemical estimations can be obtained. We retrieved the wind and precipitation variables described above from the VHR-PRO-IT (Very High-Resolution PROjections for ITaly) (Raffa et al. 2023) climate projections based on the IPCC Representative Concentration Pathways RCP4.5 and RCP8.5 (IPCC 2014) over the Italian territory. The MLD was calculated based on the wind variables. The WF was calculated using the estimated

precipitation on the Bay of Naples, and on the Bay and the catchment area to assess inputs from river discharge. This resulted in four datasets, two per RCP scenario, half considering the catchment area and half not considering the catchment area. We refer to these four datasets as the "Projection datasets".

We applied a Random Forest (RF) regression (Breiman 2001) to the "Full dataset" and ran a Feature Importance Analysis to determine the main predictor of the chlorophyll-a. RF is a supervised Machine Learning algorithm based on decision trees that performs pattern recognition for classification and regression tasks. "Training" is the learning phase of the algorithm, realized on 67% of available data the RF classifies a set of variables to predict a target. "Validation" is performed with the 33% of data remaining: this fraction is passed to the RF that produces a prediction that we can compare to the target fraction that was excluded from the learning. The importance of predictors in the fitting is determined as the Gini importance in % and the Pearson coefficient is used to evaluate the strength of the correlations. We used the "Full dataset" with the CHL as the target and the physical and biogeochemical variables as predictors.

Based on our finding that salinity is the strongest predictor of chlorophyll levels, we introduce a parameterization that links salinity to two key factors: (i) its relationship to water column stratification and (ii) cumulative freshwater discharge. We developed a minimal machine learning algorithm that uses climate projections by modelling Salinity and Chlorophyll dynamics using a minimal setup of physical parameters. We build on previously established dependencies: chlorophyll response can be predicted from salinity changes, and salinity changes are driven by freshwater inputs and stratification. The process involves two steps: (a) predicting salinity variations from physical parameters, and (b) modelling chlorophyll response based on salinity changes. We test the use of a minimal set of physical parameters through various simplified trainings—starting with best estimates for S and CHL, then progressively using less accurate ones to assess model degradation.

Finally, we estimate the chlorophyll response using climate projections until 2070. Four RF trainings are generated (two for RCP4.5 and two for RCP8.5, two "with catchment", identified in the following as *) and two "without catchment", identified as **) using the "Projection datasets". To ensure consistency between training and projections, the model training is performed using precipitation data collected between 2001 and 2020, while the projections are generated using the estimated precipitation based on the climate projections for the period between 2001 and 2070. We fit a linear relationship between yearly variations in salinity and chlorophyll to quantify the primary production response to salinity changes, providing a basis for estimating annual carbon uptake in the marine ecosystem. This approach allows quantification of $CO_2$ removal from the atmosphere and Dissolved Inorganic Carbon uptake based on salinity-driven chlorophyll variations. For each distinct training the average slope is calculated between the two RCPs.

### 2.3.4 Results and discussion

#### *2.3.4.1 Chlorophyll-a predictors*

Feature importance analysis (Fig. 3.2a) of the predictor for log-transformed chlorophyll-a concentrations, built using the "Full dataset", identifies surface salinity as the main predictor, contributing over 35% to the model's explanatory power, followed by nitrate, phosphate, and MLD (~10%). The Pearson correlation coefficient (0.751) indicates a strong alignment between predictions and observations, capturing overall chlorophyll trends despite some discrepancies at extreme values.

The time series comparison confirms the model's ability to reproduce seasonal and interannual chlorophyll fluctuations, closely matching observed peaks and troughs, though certain events (e.g., springs 2012 and 2013) remain challenging to replicate. The annual averaged series further demonstrates the model's robustness over nearly two decades, successfully capturing broad temporal patterns, with deviations likely linked to extreme events or episodic nutrient inputs beyond the resolution of weekly predictors.

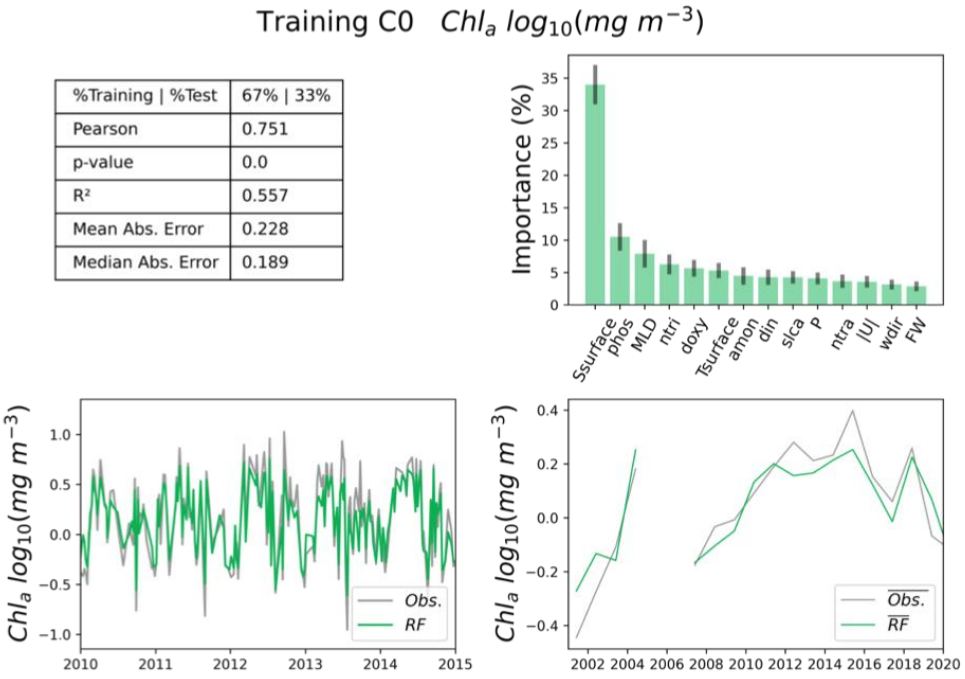



Figure 3.2. Random Forest Model Performance for Chlorophyll-a Prediction in the Gulf of Naples. Training on the historical chlorophyll-a data (obs.) using various environmental predictors. Importance of each predictor variable is indicated with STD error bars (top right), and time series comparisons (bottom left: weekly, arbitrarily shown between 2010-2015; bottom right: inter-annual) illustrate model accuracy and variability. Note the gap 2005-2006 on the inter-annual due to lack of BGC observations in the merged time series.

### 2.3.4.2 Linkage between salinity and physical factors

The simplified model accurately predicts surface salinity, with robustness tested across different training levels, from the best estimator to more simplified approaches (S* and S**, in Figure 3.3). Despite relying on simplified salinity predictions from S* and S**, the RF model still maintains a significant correlation with observations and captures key patterns in the time series (Figure 3.4).

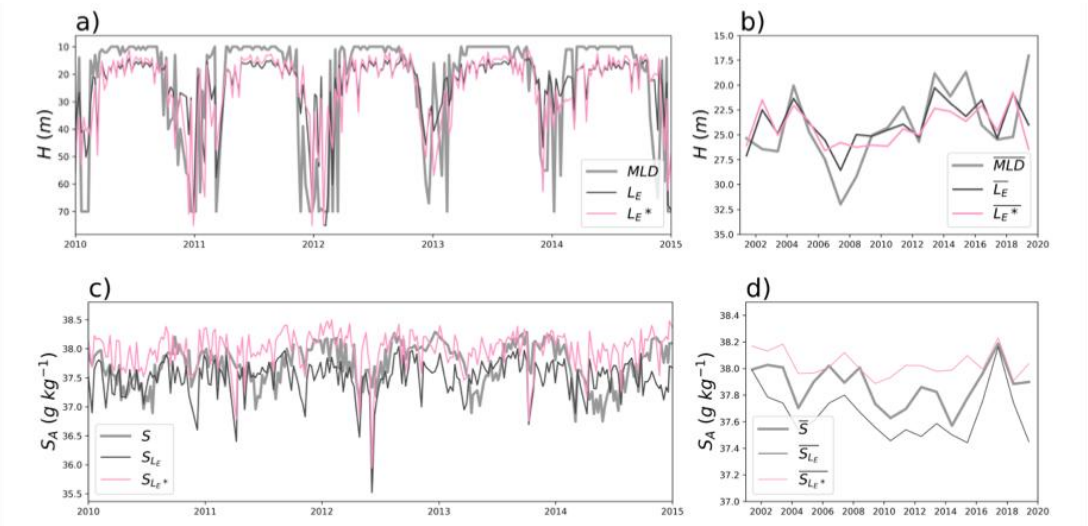



Figure 3.3. Weekly (left) and inter-annual series (right) of Mixed Layer Depth (top) and surface salinity (bottom) in the Gulf of Naples. a,b: Comparison between the estimated Ekman layer and measured Mixed Layer Depth (MLD, thick grey). Two versions are shown: $LE$ (black line) uses best available in situ observations (vertical stratification) together with wind data products, while $LE*$ (pink line) is a simplified version relying on averaged monthly vertical profiles to replace direct observations. c,d: Surface Salinity Parameterizations. Observations (thick grey) are compared with two parametrization outputs: $SLE$ (black, based on $LE$), and $SLE*$ (pink, based on $LE*$).

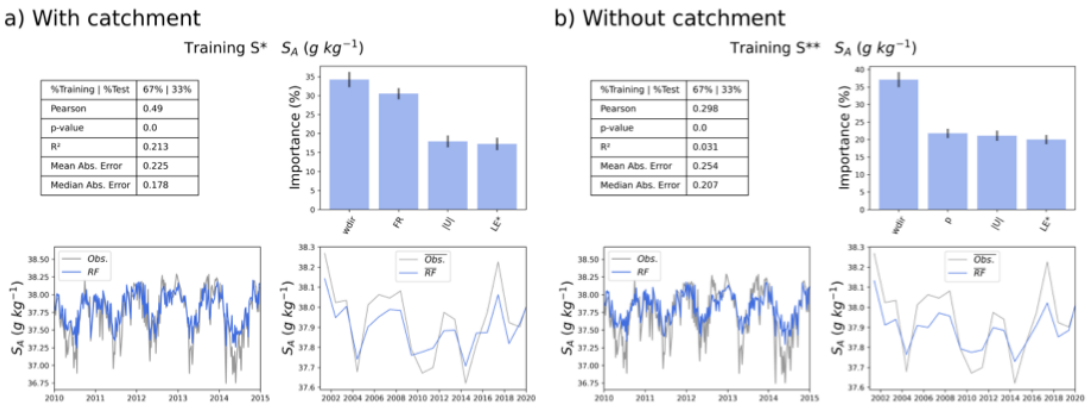



Figure 3.4. Random Forest Model Performance for Salinity using simplified estimators. Training S* and S** use simplified predictors to assess model robustness under data-scarce conditions. Importance of each predictor variable is indicated with STD error bars.

### *2.3.4.3 Future projections*

The projections indicate declining trends in precipitation and wind, which, in our models, result in increasing salinity and a corresponding decrease in chlorophyll levels. Both RCP4.5 and RCP8.5 scenarios show a declining trend in wind speed from 2020 to 2070 (RCP4.5: -0.0018 m $s^{-1}$/year; RCP8.5: -0.0047 m $s^{-1}$/year; Fig. 3.5). The decline is steeper and significant under RCP8.5 (pv = 0.0069), with a visible shift beyond 2045. Similarly, precipitation decreases in both scenarios (Fig. 3.5b), though only RCP4.5 shows a statistically significant trend (RCP4.5: -0.0727 mm $h^{-1}$/year, p = 0.0428; RCP8.5: -0.0644 mm $h^{-1}$/year, pv = 0.1247). Salinity and chlorophyll trends vary depending on the RF training method, distinguishing between runoff-influenced freshwater (FP, Figure 3.5c,d) and direct precipitation without catchment (p, Figure 3.5e,f).

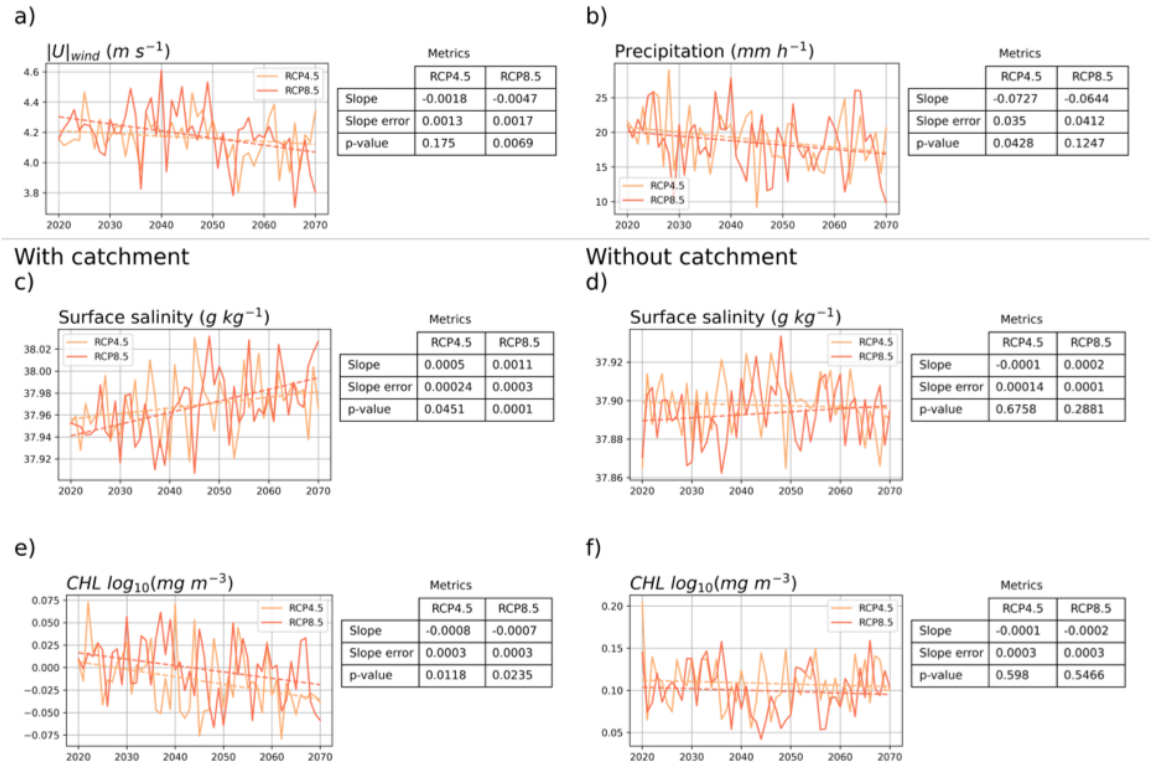



Figure 3.5. Projections of (a) wind and (b) precipitation, under RCP4.5 and RCP8.5 Scenarios (2020-2070). Surface Salinity and Chlorophyll-a are obtained from projections in the training * (c-d), and in the training ** (e-f). Linear trends are emphasized with metrics to assess their significance.

Observed chlorophyll shows a distinct bimodal distribution, with a lower peak in autumn and a higher peak in spring, when productivity increases. Both RF trainings capture this bimodality but yield contrasting future projections depending on whether basin catchment effects are considered. With basin catchment, under RCP4.5, bimodality weakens over time, particularly by the 2060s, as reduced precipitation lowers runoff and nutrient supply. The spring peak diminishes, while the autumn peak becomes more dominant, reflecting an overall decline in chlorophyll-a concentrations. Under RCP8.5, this effect is comparable, leading to a near-monomodal distribution by 2060–2069. The combined impact of reduced runoff and salinification under high emissions thus appears to severely constrain chlorophyll-a production.

The observed variations from year to another, for all scenarios and training (Fig 3.6), show the negative response of chlorophyll to salinity changes. Considering the basin catchment (Fig. 3.6a), a |ΔCHL| of ~0.35 mg $m^{-3}$/year responds to a |ΔS| of 0.1 g $kg^{-1}$/year. Without catchment (Fig. 3.6b), the response is steeper, with a |ΔCHL| of ~0.50 mg $m^{-3}$/year for the same salinity change, indicating greater variability in systems lacking runoff buffering.

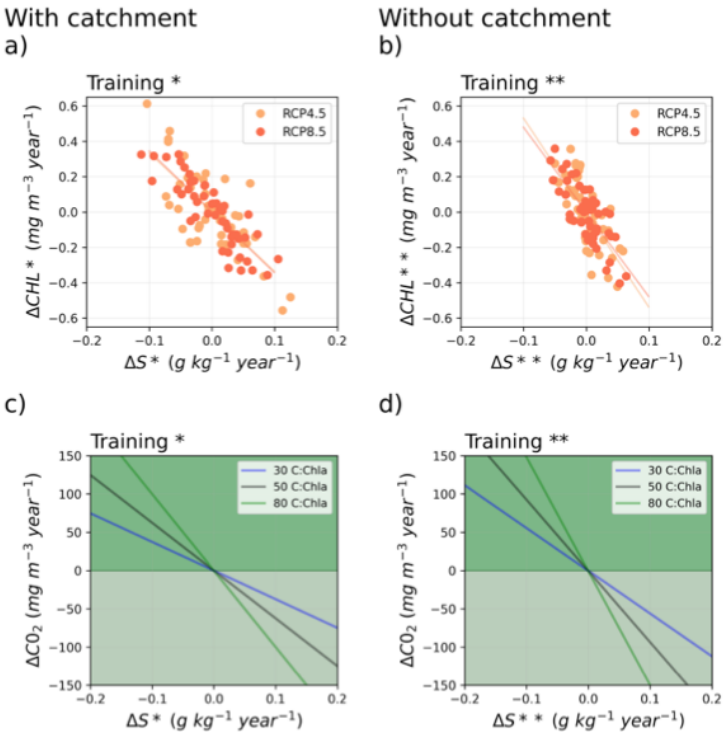



Figure 3.6. Relationship between year-to-year changes in surface salinity and chlorophyll-a concentrations, and associated $CO_2$ fluxes under different conversion scenarios. (a,b) Relationship between interannual differences in salinity (ΔS) and chlorophyll-a concentrations (ΔCHL); (c,d) $CO_2$ variations associated to subsequent primary production changes (ΔPP), for three carbon conversion factors (30, 50, and 80 mg C / mg CHL).


#### *2.3.4.4 Discussion*

The emergence of salinity as a primary predictor of chlorophyll concentration in the Gulf of Naples highlights the role of freshwater inflows and stratification as key modulators of primary productivity.

Our model suggests a progressive increase in salinity and a decline in chlorophyll, particularly in spring, with uncertainties in autumn trends. As expected, land-based changes (e.g., precipitation, basin catchment effects) exert a strong influence on coastal ecosystems. Building on Romillac et al. (2023), we attribute salinity's dominant role as a predictor primarily to nutrient loads carried by runoff, rather than its impact on stratification. While high stratification can enhance phytoplankton accumulation by reducing vertical mixing, light limitation is negligible at the LTER-MC site (75 m depth), even in winter.

The study also quantifies the yearly relationship between salinity and chlorophyll, enabling the development of an index to track primary production responses under climate projections. Our carbon uptake scaling provides a first-order estimation but relies on general carbon conversion factors, omitting phytoplankton adaptation corrections and assessing only temporary inorganic carbon fixation, part of which may be exported beyond the photic zone. The impact of salinity variations on phytoplankton accumulation could also lead to broader ecological shifts, influencing species composition and community structure as less productive conditions favour different phytoplankton assemblages and marine organisms.

Studies (e.g., Zhang et al., 2019) using the Soil and Water Assessment Tool (SWAT) demonstrate that urbanizing watersheds experience increased runoff and peak flows due to reduced infiltration under RCP4.5 and RCP8.5. This effect was not accounted for in our study, where we applied a constant absorption rate of 27%. If absorption capacity declines, runoff buffering weakens, causing greater variability in freshwater inputs that could align more directly with precipitation rates. Finally, even in the case of little changes in annual mean precipitation, the expected increase of extreme precipitation events, where rain is concentrated on short periods, may reduce the actual biomass accumulation since the efficiency of nutrient uptake largely decreases in such short events.

The Random Forest approach demonstrates that salinity variations at a given point can be effectively captured using freshwater runoff and stratification with a minimal set of input variables, independent of *in situ* observations. This highlights the efficiency and robustness of simplified models in accurately representing coastal dynamics. Nevertheless, we expect that the oceanic transport (i.e., the Gulf inner circulation) can play a role as modulator of the impact of the river on the sampling site, given the distance (about 20km) between the latter site and the river mouth. To improve understanding of advection-driven salinity distribution a refined circulation index incorporating currents, eddies, and lateral intrusions is needed (Kokoszka et al. 2023). Preliminary analyses implementing a numerical model of the circulation of the Gulf (ROMS) show that changes of the inner circulation can impact the transport of nutrients to the MC. Nevertheless, they do not improve the predictions of the machine learning tool. Moreover, to further deepen the understanding of the causal relationships between forcings and biological responses, we are now coupling the model for ocean physics with a biogeochemical model (Darwin model, MIT), in synergy with the EU project Biocean5D.


The use of RCPs provides a structured framework for exploring climate futures but also emphasizes the need for continuous updates and intercomparison of new models (O'Neill et al. 2016). Understanding these pathways supports their translation into Shared Socioeconomic Pathways (SSPs), which guide regional adaptation strategies by anticipating ecological impacts on coastal systems like the Gulf of Naples (e.g., Riahi et al. 2017). This study underscores the value of interdisciplinary data integration, aligning the Gulf of Naples with broader efforts to refine climate scenario frameworks (O'Neill et al. 2016). By advancing data-driven modelling approaches, it contributes to the development of Digital Twin Ocean (DTO; Tzachor et al. 2023) initiatives that will finish by joining such as the Digital Twin of the terrestrial water cycle, which leverages high-resolution Earth observations to enhance climate impact assessments and ecosystem management.

This analysis underscores how freshwater inputs, shaped by land use and by the impact of climate change on land, modulate coastal ecosystem responses under climate change. It also highlights the importance of basin catchment modelling for improved coastal ecosystem assessments and projections. In sum, long-term monitoring and interdisciplinary approaches are needed to enhance knowledge of the functioning of ecosystems, hence to better manage them under future climate scenarios.


## 3. Conclusion

In this report we presented three studies dedicated to the identification of biodiversity trends and their underlying drivers based on publicly available data. Moreover, we developed predictive models to project biodiversity trends into the future. Additionally, to ensure the reproducibility of our analyses, all our metadata will be accessible in the near future as part of Marco Bolo Work Package 1, and virtual workflows to reproduce our analyses will be accessible via LifeWatch.

The first study shows overall stable trends for coastal and marine biodiversity across European regions and biotic groups. Some positive trends are detected for specific groups (i.e. birds, invertebrates) and regions (i.e. Baltic), although this cannot be generalized across all the data analysed.

The second study assesses the seasonal distribution of four marine mammal species on the West coast of the European continent. Using decadal habitat suitability models, the distribution of the four species is projected into future decades under different climate scenarios. The projections, specially under the Shared Socioeconomic Pathway SP585, show a generalized reduction in suitable habitats across species, particularly marked in coastal areas.

The third study uses long-term data to determine that freshwater inputs in the Bay of Naples drive water salinity, which is the main predictor of chlorophyll-a in the bay. Using a machine learning approach, a limited set of physical variables was sufficient to produce robust estimations of primary production. Future climate projections show a decrease in rainfall that will lead to an increase in salinity and consequentially to a reduction of chlorophyll-a.

In conclusion, temporal changes in biodiversity are driven by many interacting factors: some of those factors are dependent on human activities, negatively (i.e. climate change, overfishing, pollution) or positively (i.e. establishment of protected areas, habitat restoration). During the last decades, biodiversity in coastal and marine European habitats has been stable, but with variations among sites. Nevertheless, future climate projections (i.e. warmer waters and decreased rainfall) will translate into a change in conditions for coastal and marine European biodiversity. At the same time, other anthropic pressures can worsen the situation, but new legislation and policies can have beneficial effects.

The three studies presented here show the potential of publicly available data combined with statistical modelling and machine learning techniques. Nevertheless, all three studies are subject to strong data limitations, making extrapolations difficult: the data analysed is not homogeneously distributed across regions and biotic groups, and the analyses are restricted to sites and species with high quality datasets.

To produce better analyses and projections, we strongly recommend the **open publication of existing datasets not currently openly available** and to encourage **harmonisation of biodiversity monitoring across the continent**, as well as to invest in **advanced machine learning methodologies**. By understanding the past and present trends and future projections of coastal and marine biodiversity, we can guide conservation efforts to ensure healthy ecosystems for the next decades.

## Bibliography

Aminian-Biquet, J., Gorjanc, S., Sletten, J., Vincent, T., Laznya, A., Vaidianu, N., ... & e Costa, B. H. (2024). Over 80% of the European Union's marine protected area only marginally regulates human activities. *One Earth*, 7(9), 1614-1629. https://doi.org/10.1016/j.oneear.2024.07.010

Assis, J., Fernández Bejarano, S.J., Salazar, V.W., Schepers, L., Gouvêa, L., Fragkopoulou, E., … & De Clerck, O. (2024). Bio-ORACLE v3.0. Pushing marine data layers to the CMIP6 Earth system models of climate change research. *Global Ecology and Biogeography*. https://doi.org/10.1111/geb.13813

Andersen, J. H., Carstensen, J., Conley, D. J., Dromph, K., Fleming-Lehtinen, V., Gustafsson, B. G., ... & Murray, C. (2017). Long-term temporal and spatial trends in eutrophication status of the Baltic Sea. *Biological Reviews*, 92(1), 135-149. https://doi.org/10.1111/brv.12221

Beja, J., Vandepitte, L., Perez, R. P., Weigel, B., Lipizer, M., Vanhoorne, B., & Tyberghein, L. (2024). EMODnet Biology: an EU service for the marine biology community and beyond. In *International Conference on Marine Data and Information Systems IMDIS 2024 MISCELLANEA INGV* (No. 80, p. 200). Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://dx.doi.org/10.13127/misc/80

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Butchart, S. H., Walpole, M., Collen, B., Van Strien, A., Scharlemann, J. P., Almond, R. E., ... & Watson, R. (2010). Global biodiversity: indicators of recent declines. *Science*, 328(5982), 1164-1168. https://doi.org/10.1126/science.1187512

Cano-Barbacil, C., Bowler, D. E., Ballesteros-Pelegrín, G. A., Bertolero, A., Deneudt, K., Genovart, M., … & Haase, P. (2025). Seven decades of coastal bird recovery. [Manuscript submitted for publication]. Global Change Biology, 31(11), e70623. https://doi.org/10.1111/gcb.70623

Castège, I., Milon, E., & Pautrizel, F. (2014). Response of benthic macrofauna to an oil pollution: Lessons from the "Prestige" oil spill on the rocky shore of Guéthary (south of the Bay of Biscay, France). *Deep Sea Research Part II: Topical Studies in Oceanography*, 106, 192-197. https://doi.org/10.1016/j.dsr2.2013.09.035

Chust, G., Villarino, E., McLean, M., Mieszkowska, N., Benedetti-Cecchi, L., Bulleri, F., ... & Lindegren, M. (2024). Cross-basin and cross-taxa patterns of marine community tropicalization and deborealization in warming European seas. *Nature communications*, 15(1), 2126. https://doi.org/10.1038/s41467-024-46526-y

Copernicus Marine Service. (2025). Copernicus Marine Environment Monitoring Service (CMEMS). Available from: https://marine.copernicus.eu/ [Accessed 28 October 2025].

Council Directive. 92/43/EEC. *Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora* (Habitats Directive). *Official Journal of the European Communities*, OJ L206. https://eur-lex.europa.eu/eli/dir/1992/43/oj/eng

Directive 2009/147/EC. *Directive 2009/147/EC of the European Parliament and of the Council of 30 November 2009 on the conservation of wild birds* (Birds Directive). *Official Journal of the European Union*, OJ L20 https://eur-lex.europa.eu/eli/dir/2009/147/oj/eng

Directive. 2016/2336. *Directive (EU) 2024/3019 of the European Parliament and of the Council of 27 November 2024 concerning urban wastewater treatment (recast) (Text with EEA relevance)*. https://eur-lex.europa.eu/eli/dir/2024/3019/oj/eng

Dornelas, M., Antao, L. H., Moyes, F., Bates, A. E., Magurran, A. E., Adam, D., ... & Murphy, G. (2018). BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecology and Biogeography*, *27*(7), 760-786. https://doi.org/10.1111/geb.12729

Downing, J. A. (2014). Productivity of Freshwater Ecosystems and Climate Change. In B. Freedman (Ed.), *Global Environmental Change* (pp. 221–229). Dordrecht: Springer Netherlands.

EEA. (2020). *Spatial Analysis of Marine Protected Area Networks in Europe's Seas III. ETC/ ICM Technical Report 3/2020: European Topic Centre on Inland, Coastal and Marine Waters*. https://www.eionet.europa.eu/etcs/etc-icm/products/etc-icm-reports/etc-icm-report-3-2020-spatial-analysis-of-marine-protected-area-networks-in-europe2019s-seas-iii/@@download/file/Spatial%20Analysis%20of%20Marine%20Protected%20Area%20Networks%20in%20Europe%E2%80%99s%20Seas%20III.pdf

Eigaard, O. R., Bastardie, F., Hintzen, N. T., Buhl-Mortensen, L., Buhl-Mortensen, P., Catarino, R., ... & Rijnsdorp, A. D. (2017). The footprint of bottom trawling in European waters: distribution, intensity, and seabed integrity. *ICES Journal of Marine Science*, 74(3), 847-865. https://doi.org/10.1093/icesjms/fsw194

EurOBIS data. *European node of the Ocean Biodiversity Information System (EurOBIS).* Available from: https://www.eurobis.org [Accessed 28 October 2025].

Froese, R., Winker, H., Coro, G., Demirel, N., Tsikliras, A. C., Dimarchopoulou, D., ... & Matz-Lück, N. (2018). Status and rebuilding of European fisheries. *Marine Policy*, 93, 159-170. https://doi.org/10.1016/j.marpol.2018.04.018

Fiorentino, F., & Vitale, S. (2021). How can we reduce the overexploitation of the Mediterranean resources?. *Frontiers in Marine Science*, 8, 674633. https://doi.org/10.3389/fmars.2021.674633

GBIF: The Global Biodiversity Information Facility (2025) What is GBIF?. Available from: https://www.gbif.org/what-is-gbif [Accessed 28 October 2025].

Global ocean low and mid trophic levels biomass content hindcast. E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). https://doi.org/10.48670/moi-00020 [Accessed 28 October 2025].

Global Ocean Physics Reanalysis. E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). https://doi.org/10.48670/moi-00021 [Accessed 28 October 2025].

Grizzetti, B., Vigiak, O., Udias, A., Aloe, A., Zanni, M., Bouraoui, F., ... & Bielza, M. (2021). How EU policies could reduce nutrient pollution in European inland and coastal waters. *Global Environmental Change*, 69, 102281. https://doi.org/10.1016/j.gloenvcha.2021.102281

Iermano, I., Liguori, G., Iudicone, D., Buongiorno Nardelli, B., Colella, S., Zingone, A., et al. (2012). Filament formation and evolution in buoyant coastal waters: Observation and modelling. *Progress in Oceanography*, 106, 118–137. https://doi.org/10.1016/j.pocean.2012.08.003

Haase, P., Bowler, D. E., Baker, N. J., Bonada, N., Domisch, S., Garcia Marquez, J. R., ... & Welti, E. A. (2023). The recovery of European freshwater biodiversity has come to a halt. *Nature*, 620(7974), 582-588. https://doi.org/10.1038/s41586-023-06400-1

Jacquemont, J., Blasiak, R., Le Cam, C., Le Gouellec, M., & Claudet, J. (2022). Ocean conservation boosts climate change mitigation and adaptation. *One Earth*, 5(10), 1126-1138. https://doi.org/10.1016/j.oneear.2022.09.002

Kokoszka, F., Le Roux, B., Iudicone, D., Conversano, F., & Ribera d'Alcalá, M. (2023). Long-term variability of the coastal ocean stratification in the Gulf of Naples: Two decades of monitoring the marine ecosystem at the LTER – MC site, between land and open Mediterranean Sea. *Marine Ecology*, 44(3), e12725. https://doi.org/10.1111/maec.12725

Kleitou, P., Hall-Spencer, J. M., Savva, I., Kletou, D., Hadjistylli, M., Azzurro, E., ... & Rees, S. E. (2021). The case of lionfish (*Pterois miles*) in the Mediterranean Sea demonstrates limitations in EU legislation to address marine biological invasions. *Journal of Marine Science and Engineering*, 9(3), 325. https://doi.org/10.3390/jmse9030325

Maureaud, A. A., Kitchel, Z., Fredston, A., Guralnick, R., Palacios-Abrantes, J., Palomares, M. L., ... & Mérigot, B. (2025). FISHGLOB: A collaborative infrastructure to bridge the gap between scientific monitoring and marine biodiversity conservation. *Conservation Science and Practice*, *7*(6), e70035. https://doi.org/10.1111/csp2.70035

O'Brien, T.D. (2014). COPEPOD: The Global Plankton Database. An overview of the 2014 database contents, processing methods, and access interface. U.S. Dep. Commerce, NOAA Tech. Memo., NMFS-F/ST-37, 29 pp.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... & Sanderson, B. M. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461-3482. https://doi.org/10.5194/gmd-9-3461-2016

Pilotto, F., Kühn, I., Adrian, R., Alber, R., Alignier, A., Andrews, C., ... & Haase, P. (2020). Meta-analysis of multidecadal biodiversity trends in Europe. *Nature communications*, *11*(1), 3486. https://doi.org/10.1038/s41467-020-17171-y

Piroddi, C., Colloca, F., & Tsikliras, A. C. (2020). The living marine resources in the Mediterranean Sea large marine ecosystem. *Environmental Development*, 36, 100555. https://doi.org/10.1016/j.envdev.2020.100555

Raffa, M., Adinolfi, M., Reder, A., Marras, G. F., Mancini, M., Scipione, G., et al. (2023). Very High Resolution Projections over Italy under different CMIP5 IPCC scenarios. *Scientific Data*, 10(1), 238. https://doi.org/10.1038/s41597-023-02144-9

Raffa M., Reder, A., Marras, G. F., Mancini, M., Scipione, G., Santini, M., & Mercogliano, P. (2021). VHR-REA_IT Dataset: Very High Resolution Dynamical Downscaling of ERA5 Reanalysis over Italy by COSMO-CLM. *Data*, 6(8), 88. https://doi.org/10.3390/data6080088

Regulation. 1143/2014. *Regulation (EU) No 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32014R1143

Regulation. 2016/2336. *Regulation (EU) 2016/2336 of the European Parliament and of the Council of 14 December 2016 establishing specific conditions for fishing for deep-sea stocks in the north-east Atlantic and provisions for fishing in international waters of the north-east Atlantic and repealing Council Regulation (EC) No 2347/2002*. https://eur-lex.europa.eu/eli/reg/2016/2336/oj

Regulation. 530/2012. *Regulation (EU) No 530/2012 of the European Parliament and of the Council of 13 June 2012 on the accelerated phasing-in of double-hull or equivalent design requirements for single-hull oil tankers (recast).* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32012R0530

Reed, D. C., Schmitt, R. J., Burd, A. B., Burkepile, D. E., Kominoski, J. S., McGlathery, K. J., ... & Zinnert, J. C. (2022). Responses of Coastal Ecosystems to Climate Change: Insights from Long-Term Ecological Research. *BioScience*, 72(9), 871–888. https://doi.org/10.1093/biosci/biac006

REPHY - French Observation and Monitoring program for Phytoplankton and Hydrology in coastal waters (2023). REPHY dataset - French Observation and Monitoring program for Phytoplankton and Hydrology in coastal waters. Metropolitan data. SEANOE. https://doi.org/10.17882/47248

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., ... & Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global environmental change*, *42*, 153-168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Romillac, N., Abagnale, M., Kokoszka, F., Passarelli, A., Saggiomo, V., Ribera d'Alcalà, M., & Margiotta, F. (2023). Interplay among anthropogenic impact, climate change, and internal dynamics in driving nutrient and phytoplankton biomass in the Gulf of Naples. *Marine Ecology*, 44(3), e12754. https://doi.org/10.1111/maec.12754

Sinclair, J. S., Welti, E. A., Altermatt, F., Álvarez-Cabria, M., Aroviita, J., Baker, N. J., ... & Haase, P. (2024). Multi-decadal improvements in the ecological quality of European rivers are not consistently reflected in biodiversity metrics. *Nature ecology & evolution*, *8*(3), 430-441. https://doi.org/10.1038/s41559-023-02305-4

Sköld, M., Blomqvist, M., Bradshaw, C., Börjesson, P., Göransson, P., & Wennhage, H. (2025). Long-term recovery and food web response of benthic macrofauna following cessation of bottom trawling in a marine protected area. *Conservation science and practice*, 7(4), e70037. https://doi.org/10.1111/csp2.70037

Suikkanen, S., Pulina, S., Engström-Öst, J., Lehtiniemi, M., Lehtinen, S., & Brutemark, A. (2013). Climate change and eutrophication induced shifts in northern summer plankton communities. PLoS one, 8(6), e66475. https://doi.org/10.1371/journal.pone.0066475

Tornero, V., & Hanke, G. (2016). Chemical contaminants entering the marine environment from sea-based sources: A review with a focus on European seas. *Marine Pollution Bulletin*, 112(1-2), 17-38. https://doi.org/10.1016/j.marpolbul.2016.06.091

Tsikoti, C., & Genitsaris, S. (2021). Review of harmful algal blooms in the coastal Mediterranean Sea, with a focus on Greek waters. *Diversity*, 13(8), 396. https://doi.org/10.3390/d13080396

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. https://doi.org/10.1111/j.1466-8238.2011.00656.x

Vezzulli, L., & Reid, P. C. (2003). The CPR survey (1948–1997): a gridded database browser of plankton abundance in the North Sea. *Progress in Oceanography*, *58*(2-4), 327-336. https://doi.org/10.1016/j.pocean.2003.08.011

Wetterhall, F. (2019). River discharge and related historical data from the European Flood Awareness System [Data set]. ECMWF. https://doi.org/10.24381/CDS.E3458969

WoRMS Editorial Board. (2025). World Register of Marine Species. Available from https://www.marinespecies.org at VLIZ. https://doi.org/10.14284/170

# Appendix A Additional maps

## A1 Seasonal presence maps

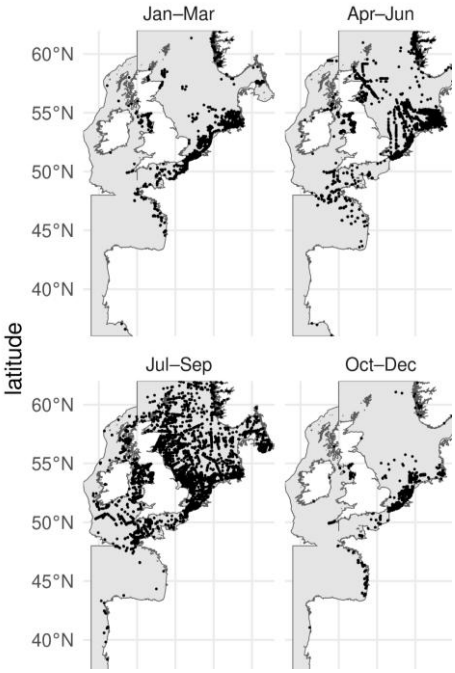



Figure A1. Harbour porpoise occurrences aggregated per season for the period (2000-2019).

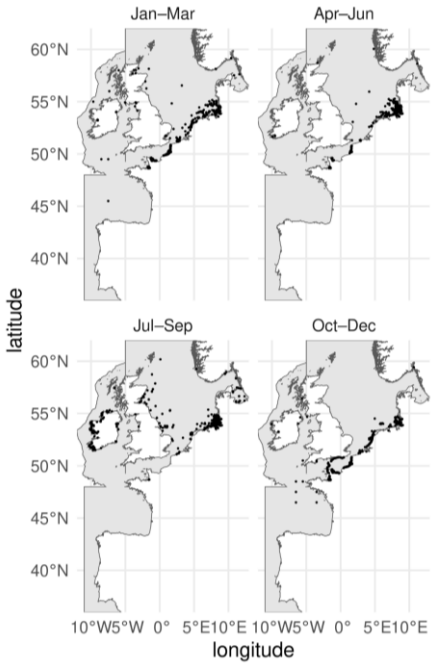



Figure A2. Harbour seal seasonal presence points, aggregated for the period (2000-2019).

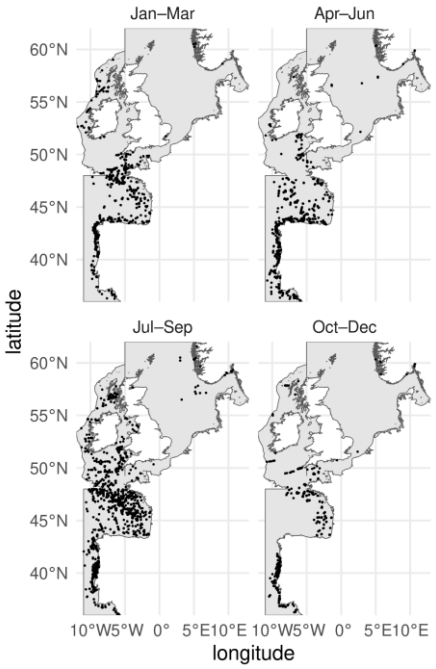



Figure A3. Common dolphin seasonal presence points, aggregated for the period (2000-2019).

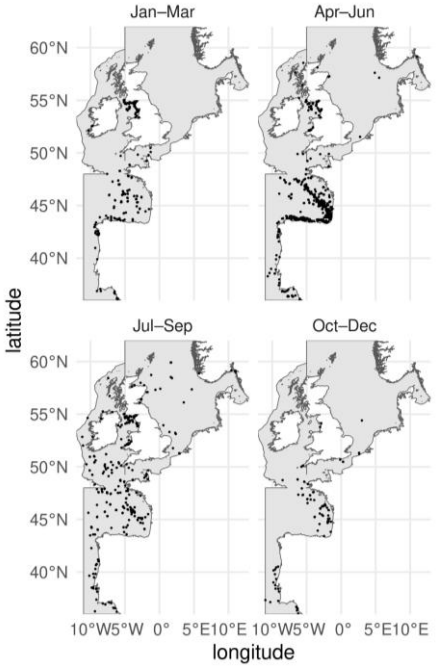



Figure A4. Bottlenose dolphin seasonal presence points, aggregated for the period (2000-2019).

## A2 Decadal predictions present

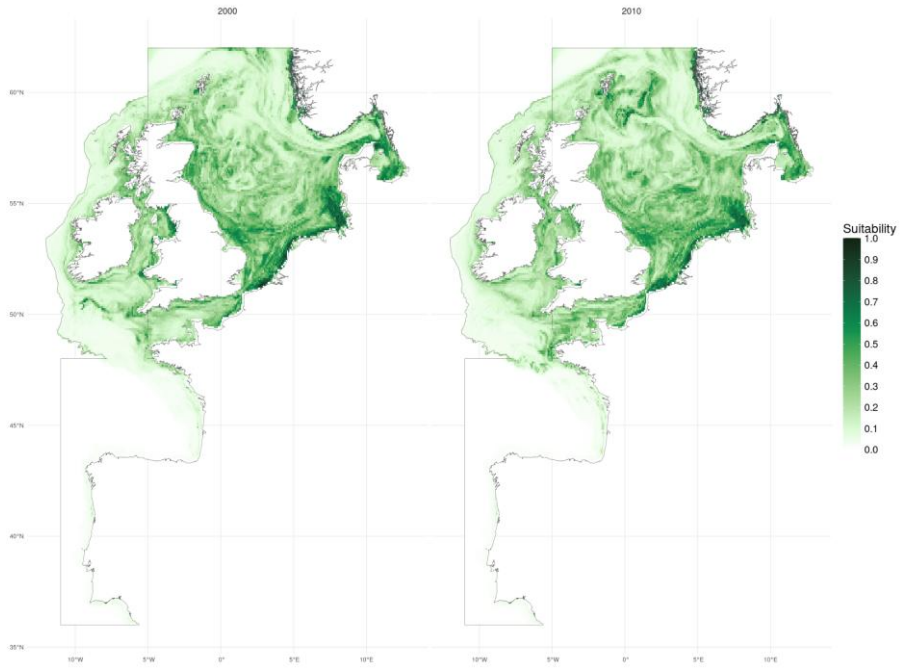



Figure A5. Harbour porpoise decadal predictions for 2000-2009 decade (left) and 2010-2019 decade (right).

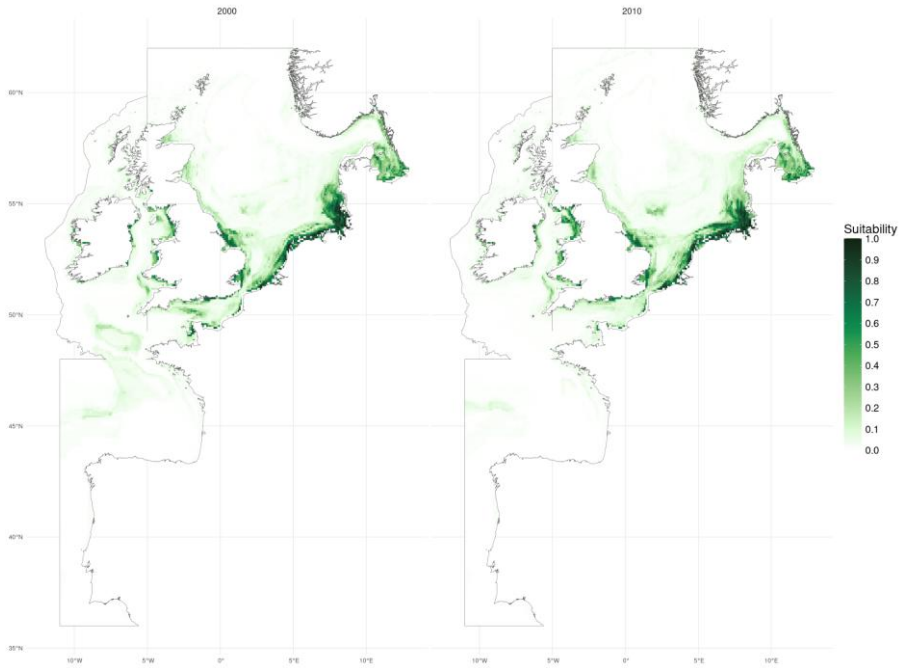



Figure A6. Harbour seal decadal predictions for 2000-2009 decade (left) and 2010-2019 decade (right).

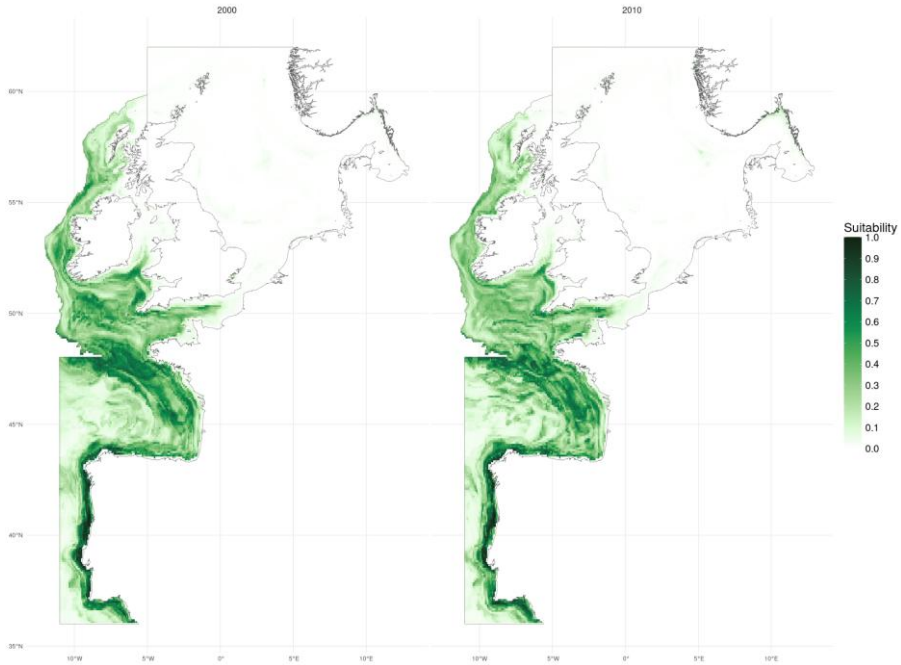



Figure A7. Common dolphin decadal predictions for 2000-2009 decade (left) and 2010-2019 decade (right).

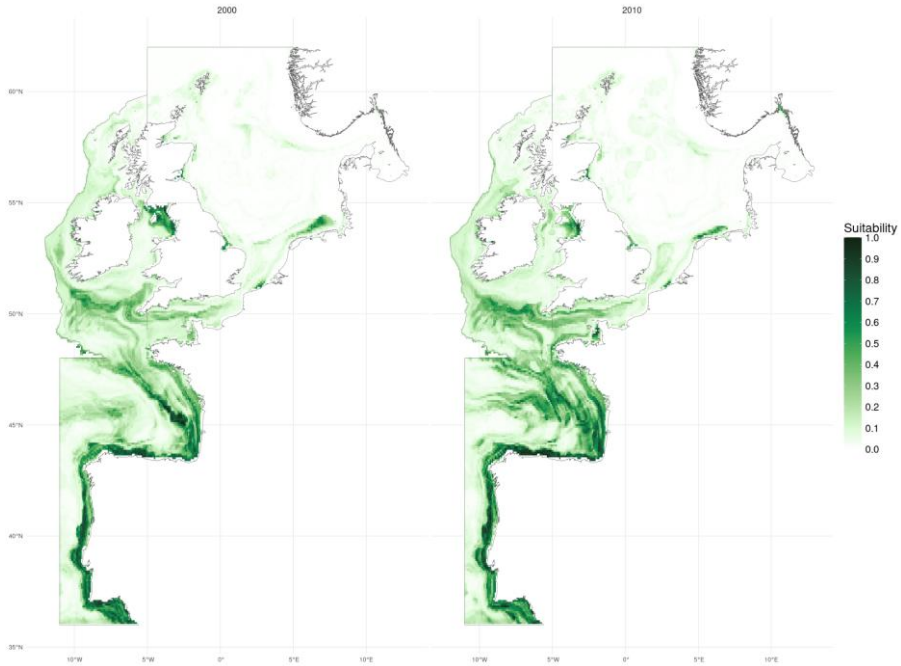



Figure A8. Bottlenose dolphin decadal predictions for 2000-2009 decade (left) and 2010-2019 decade (right).

## A4 Decadal predictions future

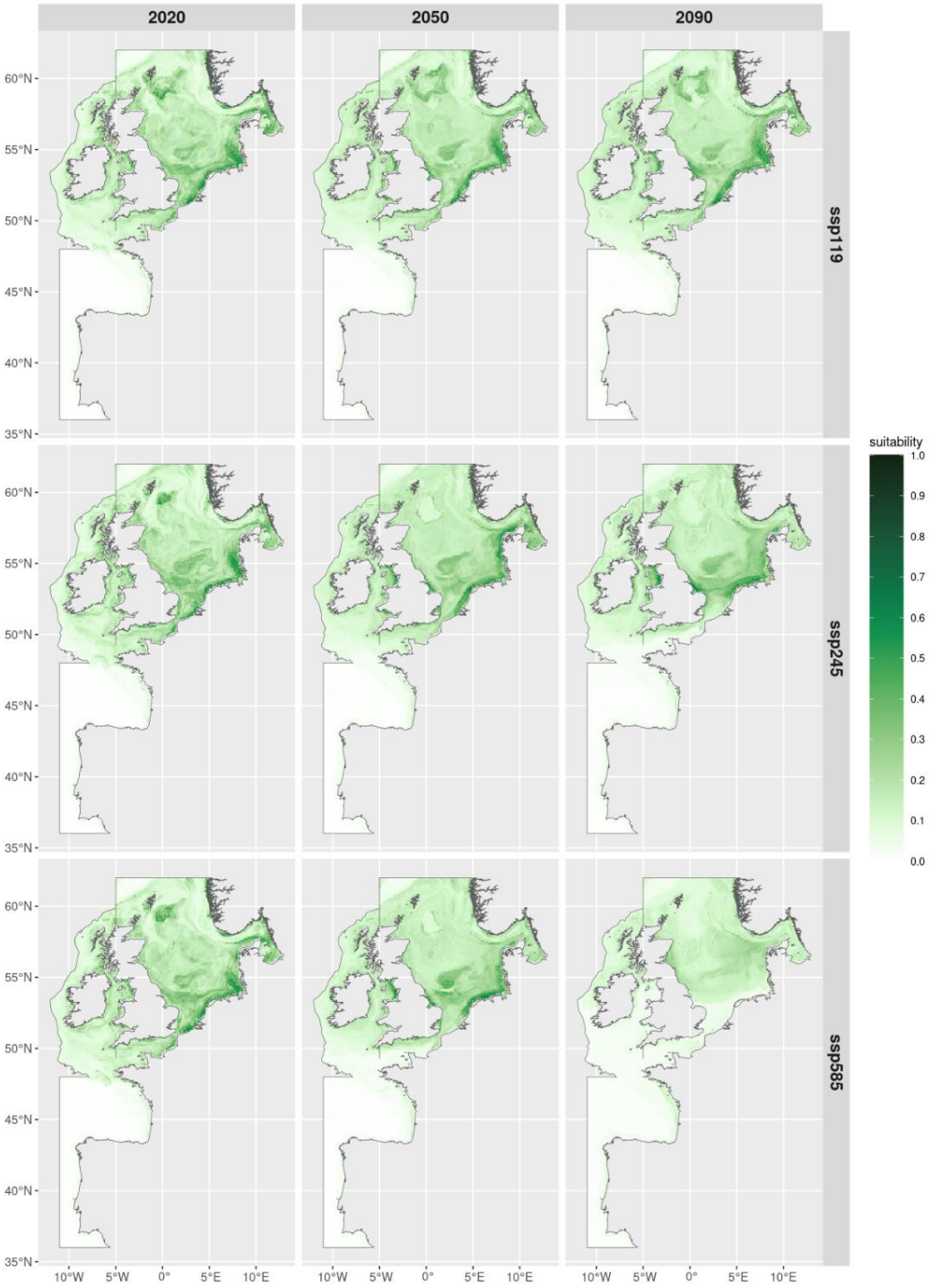



Figure A9. Harbour porpoise future decadal habitat suitability predictions under climate scenarios SSP119, 245 and 585 for decades 2020-2030, 2050-2060 and 2090-2100.

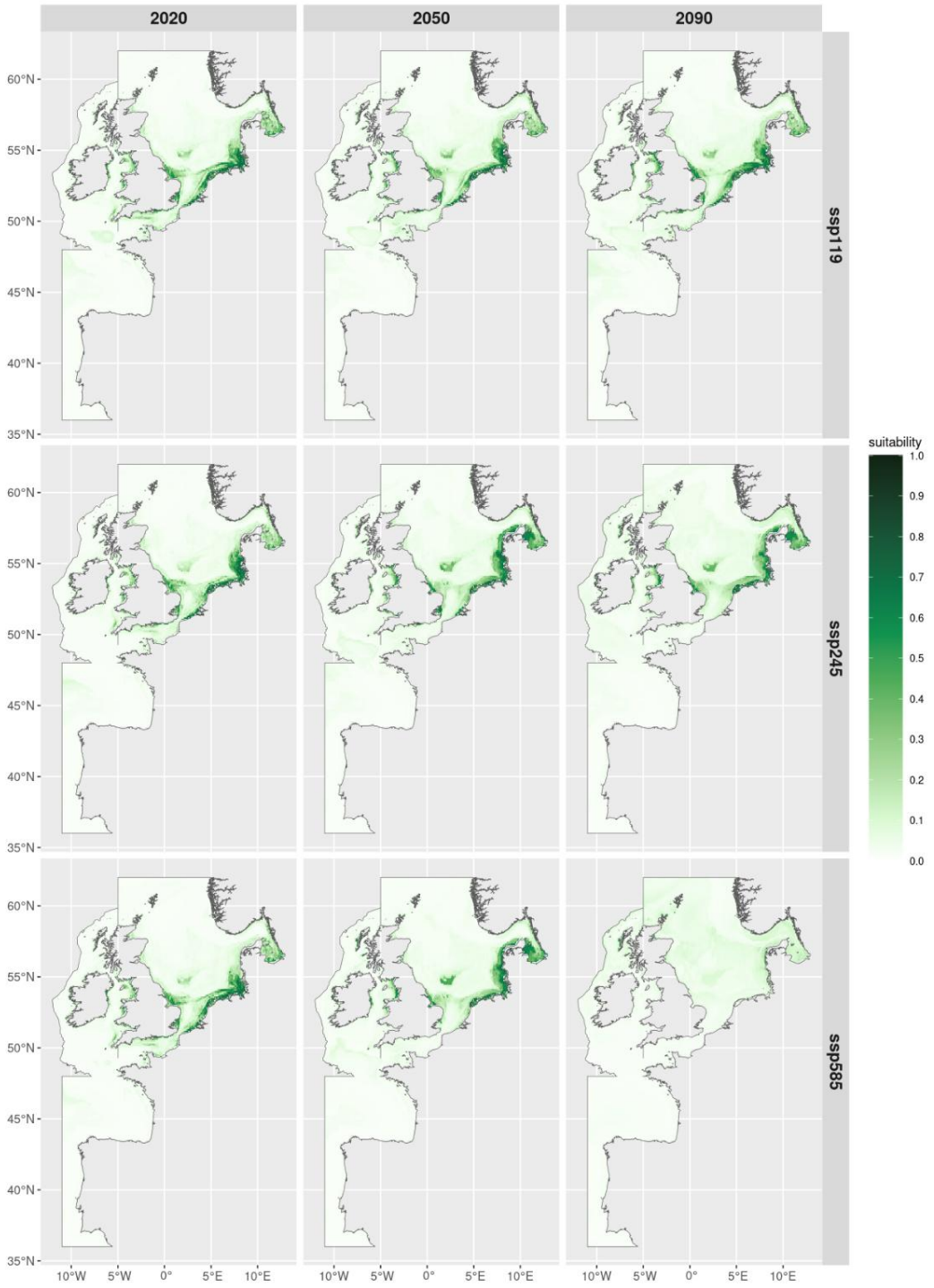



Figure A10. Harbour seal future decadal habitat suitability predictions under climate scenarios SSP119, 245 and 585 for decades 2020-2030, 2050-2060 and 2090-2100.

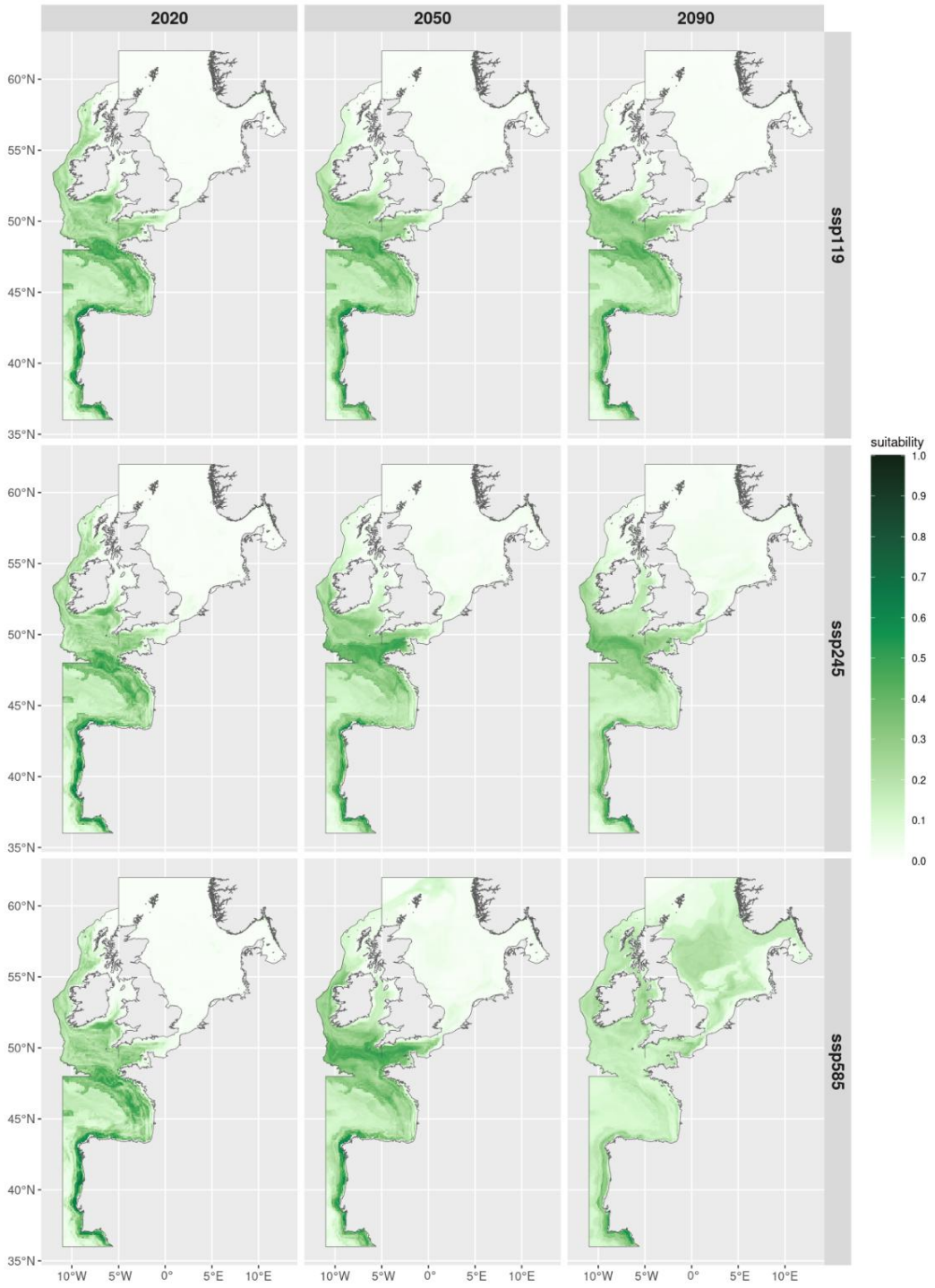



Figure A11. Common dolphin future decadal habitat suitability predictions under climate scenarios SSP119, 245 and 585 for decades 2020-2030, 2050-2060 and 2090-2100.

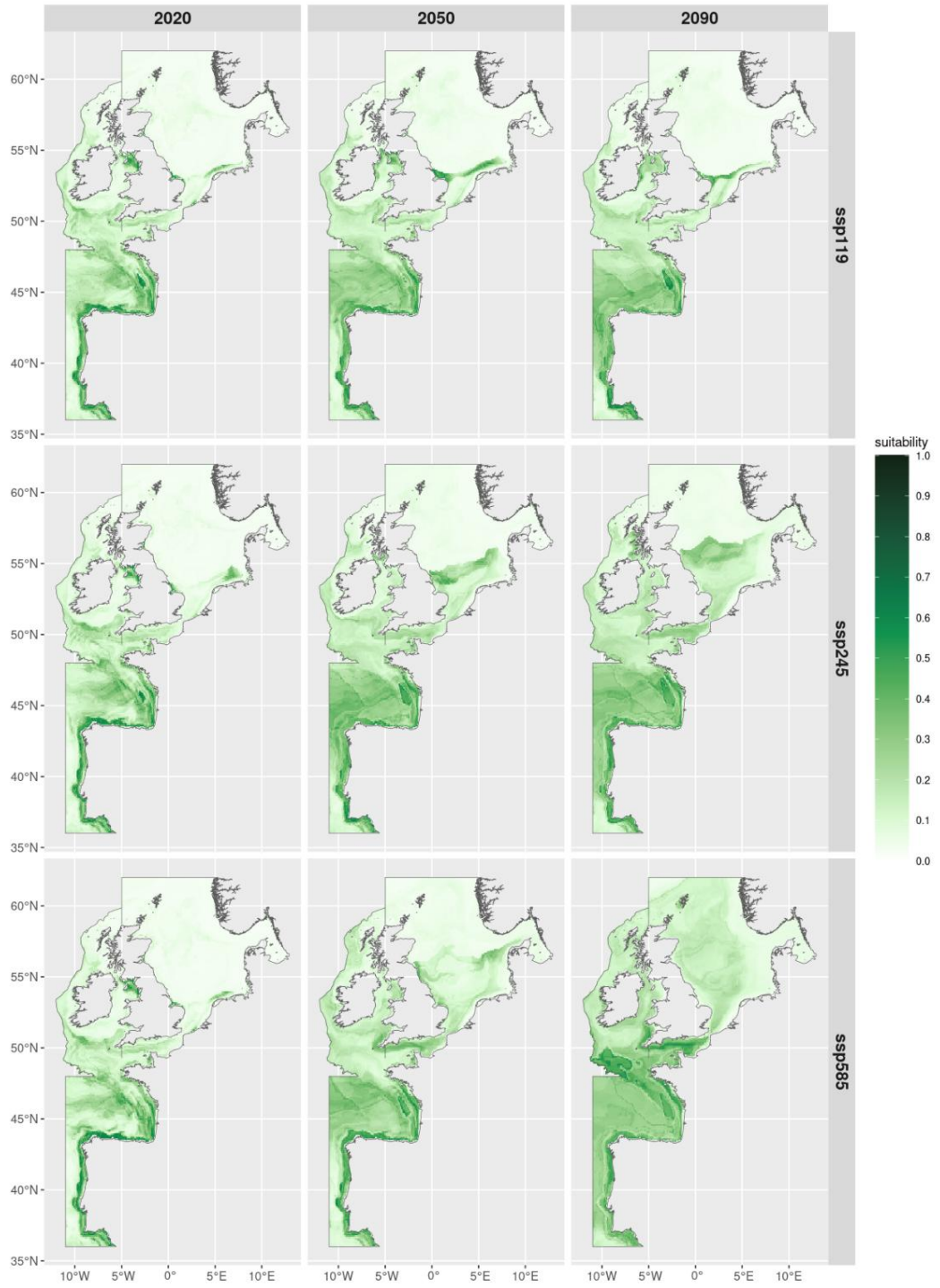



Figure A12. Bottlenose dolphin future decadal habitat suitability predictions under climate scenarios SSP119, 245 and 585 for decades 2020-2030, 2050-2060 and 2090-2100.


# Appendix B Ecological Interactions and Regime Shift Detection for MareChiara Time Series

## B1 Introduction

The primary goal of this research was to analyse the ecological interactions between various biogeochemical parameters at Mare Chiara, with a focus on identifying the main drivers of productivity in this region and understanding their dynamics over time.

Data for the analysis came from the LTER-MC research station in the Gulf of Naples, from 1984 to 2019. This biogeochemical time-series data was first filtered, cleaned and transformed, and subsequently used to train several Random Forest prediction models to estimate the importance of each parameter on chlorophyll production. Having identified the main drivers, the focus shifted to the system's dynamics. An algorithm was developed to detect regime shifts by analyzing variations on the prediction performance of Random Forest models over time.

This report details the methods and results at each step of the analysis, along with a discussion on their strengths and limitations.

## B2 Datasets

### Original Dataset

The original dataset was acquired in the LTER-MC research station in the Gulf of Naples. It includes a time-series of biogeochemical measurements at 10 depths (0, -2, -5, -10, -20, -30, -40, -50, -60, - 70 m) from 1984 to 2019.

The available features are: year, month, day, depth, date, TEMP, QF TEMP, PSAL, QF PSAL, AMON, QF AMON, NTRA, QF NTRA, NTRI, QF NTRI, PHOS, QF PHOS, SLCA, QF SLCA, CHLT QF; where QF stands for quality flag.

### Dataset I: All data

The cleaning process included removing all rows with missing values and the data before 1996 (due to the missing data between 1989 and 1997). The Quality Flags (QF) as it would reduce the dataset further.

New features were then added to the dataset. The surface pressure was assumed to be 1 atm and the pressure at other depths p(z) = 1 + z atm for each measurement. The mixed layer depth was computed for each sample as the depth where the temperature varied 0.8°C from the temperature at 10m. For samples where some measurements were removed in the cleaning process, the reference temperature corresponded to the measurement closest to 10m. The integral gradient of each feature was also obtained for each measurement. For samples with only one measurement, the gradient was considered zero; with two measurements, the gradient was the same for both.


The available features are: year, day, month, date, depth, TEMP, PSAL, AMON, NTRA, NTRI, PHOS, SLCA, CHLT, GRAD TEMP, GRAD PSAL, GRAD AMON, GRAD NTRA, GRAD NTRI, GRAD PHOS, GRAD SLCA, Pressure, MLD.

**Dataset II: Averaged data**

Many biological processes are dependent on depth, either directly or indirectly. Chlorophyll production, for instance, peaks at a subsurface layer depth and decreases as depth increases. Besides, most parameter values show very small variations until the mixed layer depth, leading to the assumption of conservative properties within the mixed layer. Averaging the data within the mixed layer depth eliminates the depth bias, enabling a focus on the overall drivers of productivity in that location, rather than depth dependent drivers. The second dataset was created with each row representing the averaged data from the measurements of each sample, to evaluate how eliminating the depth bias affects the results when compared to the first dataset.

**Dataset III and IV: Transformed data**

Data analysis often benefits from transforming data into a normal distribution, as normality is a requirement for many statistical techniques. The Box-Cox Transformation was applied to most of the numerical features of Datasets I and II to assess the impact of transforming data in the results, and the transformed data was stored in Datasets III and IV, respectively. The Box-Cox Transformation is a power transformation used to normalize positive data. For the features with negative values (such as the gradients), the Cubic-Root Transformation was applied. For the year, the log transformation was used due to the magnitude of the values. For the month, a 1-hot encoding transformation was applied, adding 12 new features corresponding to each month.

The features of the new Dataset III and IV remain the same as the ones of I and II, respectively, with the addition of the 12 month columns.

**Datasets V - VIII: Data limited to "normal" conditions (CHLT < 3 mg/m3)**

More than 90% of the measurements in Dataset I had chlorophyll values below 3 mg/m3, which will be regarded as the threshold for normal conditions. The measurements with chlorophyll values above this threshold are therefore considered peak values or extreme conditions.

Four additional datasets were generated to investigate potential differences in the main drivers of productivity in normal vs extreme conditions. For Dataset V, the data from Dataset I was limited to a maximum chlorophyll production of 3 mg/m3. For Dataset VI, the data from Dataset V was averaged within the MLD. Datasets VII and VIII correspond to Datasets V and VI transformed as described in the previous subsection.

It was not possible to do the same for the peak values due to the reduced amount of data.

**Tab. B1.** Summary of the content of each dataset in the project

| | Limited | Averaged | Transformed |
|---|---|---|---|
| Dataset I | | | |
| Dataset II | | X | |
| Dataset III | | | X |
| Dataset IV | | X | X |
| Dataset V | X | | |
| Dataset VI | X | X | |
| Dataset VII | X | | X |
| Dataset VIII | X | X | X |

### B3 Regression Model: Drivers of chlorophyll production

#### B3.1 Random Forests Regression Model

Random Forests is a widely used ensemble machine learning approach for regression and classification tasks. It combines multiple random decision trees, each trained on random sub- sets of data, and calculates the output as the average prediction. By utilizing several random trees, this method explores more possible predictors, improving the predictive accuracy of the model and reducing overfitting. This approach can handle multiple dimensions (features) and non-linear relationships, but it is less interpretable than single decision trees and requires significant computational resources for training.

The datasets in this project are not excessively large, are high-dimensional and are likely to have complex patterns between features, making Random Forests a suitable choice.


The Random Forests Regression model predicts the value of a dependent variable based on the independent variables, assigning each independent variable a quantitative measure of its importance on the model's outcome. In the context of this project, this approach allows us to understand how each water parameter influences the chlorophyll production, and what are its main drivers.

A baseline model was also computed for each dataset and compared to the RF model, to evaluate the accuracy of the optimal models. This baseline model predicts the chlorophyll concentration for a given sample as the average of the measured values for the same month in the year before. The poor performance of the baseline model highlights the complex interactions between parameters and a need for machine learning approaches.

### B3.2 Training and Testing

A new Random Forests model was trained for each dataset (I to VIII) to evaluate the impact of each modification (data transformation, averaging, peaks removal) on model performance. The *sklearn* package in Python was used to implement and analyze all models.

For each dataset, the dependent variable was isolated from the relevant independent variables, and the data was split into training and testing sets.

The best hyperparameters were obtained with *RandomizedSearchCV* and its fit function by performing a randomized search with cross-validation to identify the optimal hyperparameters for the model. Instead of testing all possible combinations, a specific number of random combinations were tested and compared based on performance to decide on the best set of hyperparameters.

#### B3.2.1 Random training-testing data splitting

Two different approaches to training-testing were compared. The first one splits the data by randomly sampling individual points from the dataset to achieve a specific training-testing ratio. Most of this project uses 90% as training set and 10% as testing set. This ratio is adjusted to improve the performance of the model, with the goal of achieving similar training and testing errors. The cross-validation inside the hyperparameter tuning also sampled the testing-training data according to the same sampling seed.

#### B3.2.2 'Last 2 years' training-testing data splitting

The second approach selects a given number of recent years as testing data and trains on the remaining data. The cross-validation inside the hyperparameter tuning also used groups of consecutive years to decide on the best model. The test score of each of these folds was computed in the end to cross-validate and averaged to produce a metric of the model's overall accuracy.

The different training-testing ratios between the first and the second methods might affect the model performance, specially in smaller datasets, as larger training sets often promote overfitting (see Discussion).

### B3.3 Model performance

The accuracy of the model's predictions was assessed through the mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), R-squared ($R^2$) and the correlation (Corr) between the predicted and real values. Only the last two metrics were considered when comparing models.

The correlation and $R^2$ for each model's prediction are presented in Figure B1 and B2, respectively. In both cases, a higher value is associated with more accurate predictions and more reliable main driver results.

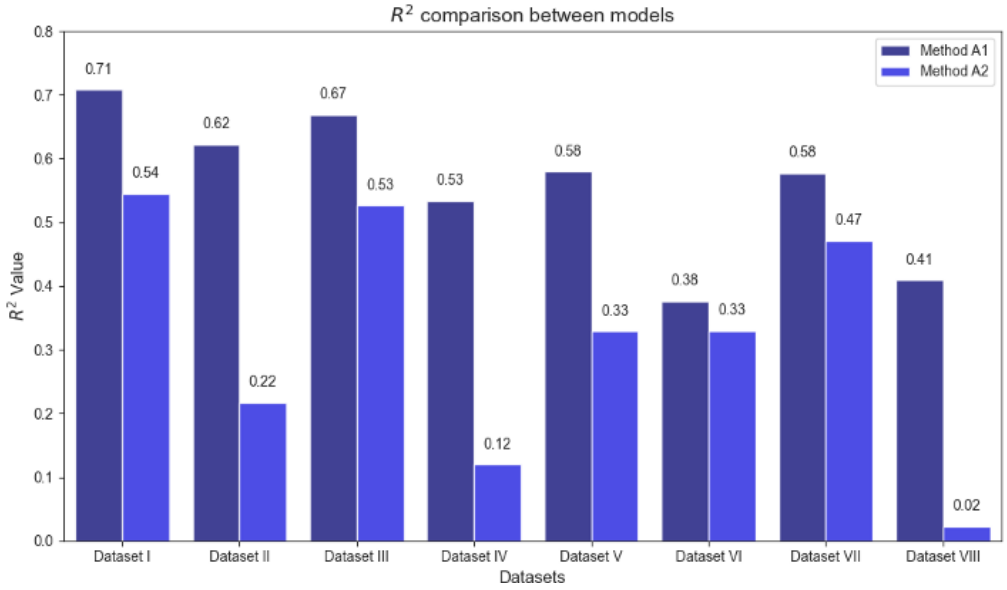



**Fig. B1.** R2 values for models trained with methods A1 and A2 across eight datasets (Dataset I to Dataset VIII). Method A1 is represented with darker bars; method A2 is represented with lighter bars.

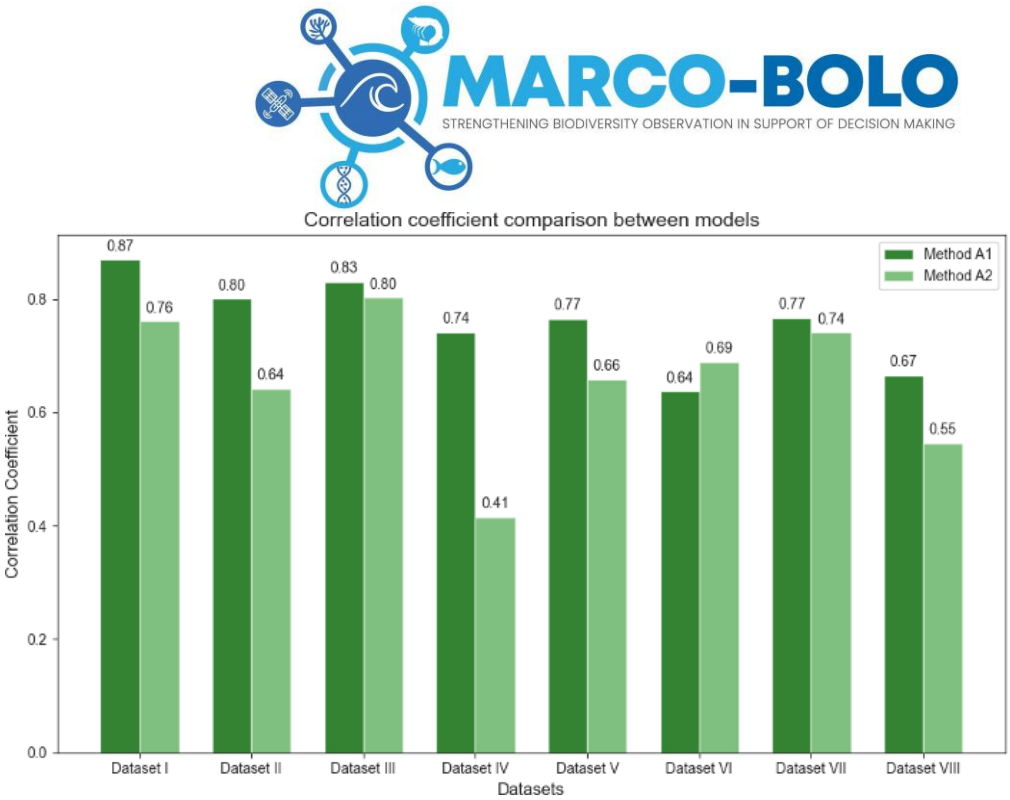



**Fig. B2**. Correlation coefficients for models trained with methods A1 and A2 across eight datasets (Dataset I to Dataset VIII). Method A1 is represented with darker bars; method A2 is represented with lighter bars.

Both Figure B1 and B2 indicate a better performance for Dataset I, where all data is being considered without any additional processing. The performance of the models is very dependent on the size of the dataset being used. When the size is reduced, by removing the peak values or by averaging the data within the MLD, the accuracy of the predictions decreases. The effect of peak removal is seen by comparing the correlation coefficient and the $R^2$ values between Datasets I-IV and Datasets V-VIII. The effect of averaging the data can be seen by comparing Dataset I to Dataset II, III to IV, V to VI and VII to VIII. The exception is the increase in accuracy after averaging the filtered data (Dataset V to VI) with the method M2 model. Regardless of the accuracy loss caused by averaging the data, the results allow more insightful conclusions on the drivers of productivity in the entire water column.

Method A1 (darker bars) generally produces better results than Method A2 (lighter bars). This is likely because A1 uses a random training-testing split, which increases the chances of a predicted value being influenced by data points very close to those in the testing set, making predictions easier. In contrast, method A2 relies on learned behavior from past measurements to make predictions. A higher accuracy with method A2 means that the interactions between variables are similar over time, which results in better predictions. The increase in accuracy with the M2 model after averaging the filtered data (Dataset V to VI) leads to the assumption that the productivity of the entire water column is more stable in time than the productivity at each depth layer. Besides, method A2 seems more sensitive to reduced datasets than A1, likely due to an increased probability of overfitting.

Contrarily to expectations, transforming the data before training the model decreased the accuracy of the predictions. The exceptions are the increase in accuracy with method A2 after transforming


the whole dataset (Dataset I to II) and the filtered data (Dataset V to VII). Generally, data transformation does not benefit the model's performance in this project.

We can show that each dataset modification lowers model performance without substantially changing feature importance results. In all cases, reduced dataset size aggravates the accuracy decline.

### B3.4 Features importance analysis

The analysis of feature importance offers insights into the key drivers of chlorophyll production. In this project, the combination of top features that contribute to at least 50% of the output is considered to represent the relevant drivers. Of those, features with an indi- vidual importance greater than 10% are classified as main drivers. If a main driver has a considerably higher importance than the others, it is regarded as a predominant driver.

The feature importances were obtained using the feature_importances attribute from *sklearn* and the summary_plot function from *shap*. The former indicates the importance of each feature for the predictive regression model, while the latter indicates the importance of each feature for individual predictions. The *shapley* summary plot also illustrates the influence of both the magnitude and the signal of each feature on the output.

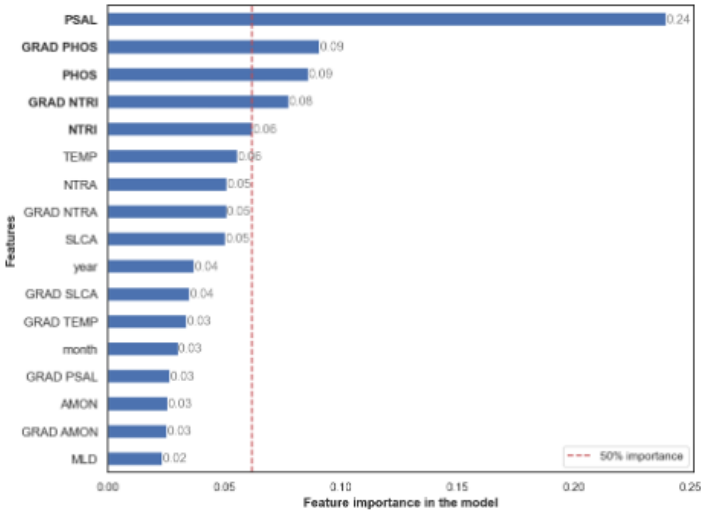
The features importance plots for Dataset I, II and III using method A1 are depicted in Figure B3 as an example of the output of these functions.

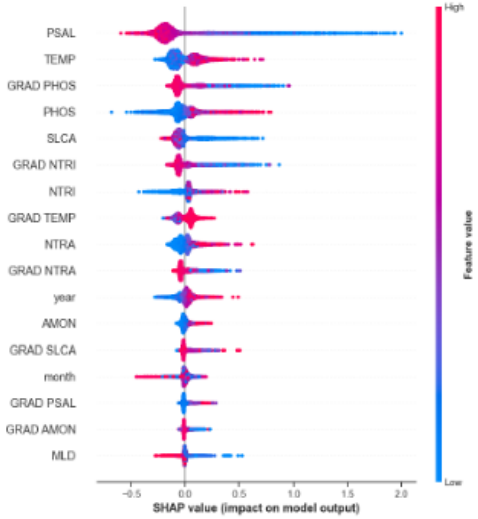
Averaging the data within the MLD provides clearer insights into the main drivers of chlorophyll production as expected. As shown in the center left pane of Figure B3, when analysing individual measurement throughout the water column, salinity stands out as the predominant feature, explaining more than 20% of the output, followed by the phosphate gradient, contributing half as much. However, when considering the entire mixed layer as a whole, salinity remains the most important feature (21%), but the mixed layer depth emerges as the second main driver (14%) and the gradient of phosphate remains as the third main driver (11%). By averaging the data within the mixed layer depth, the dominance of any single feature is reduced, allowing for other important interactions to surface.

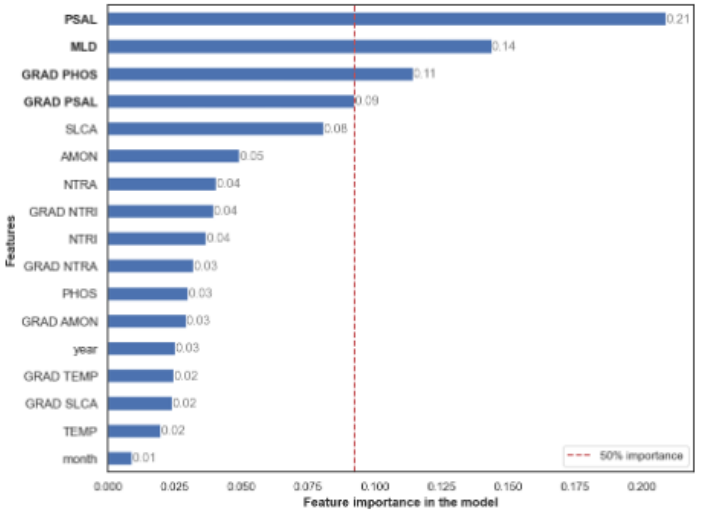
The Shapley Summary Plot on the right pane of Figure B3 illustrates how different feature values affect the output. For instance, low salinity concentrations are highly favorable for chlorophyll production, while high salinity has a negative effect.

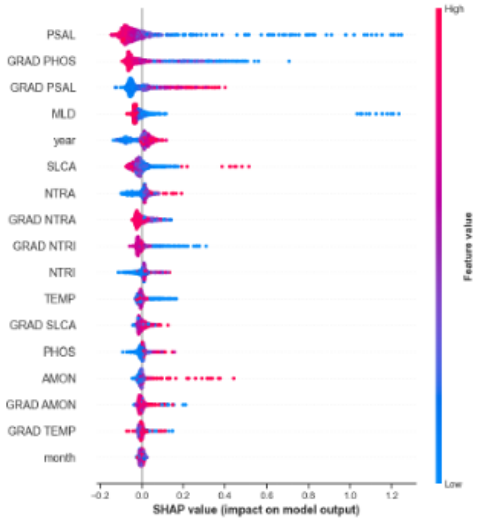
By examining the chlorophyll production at each averaged sample, the main interactions are those which have an impact on the entire mixed layer.

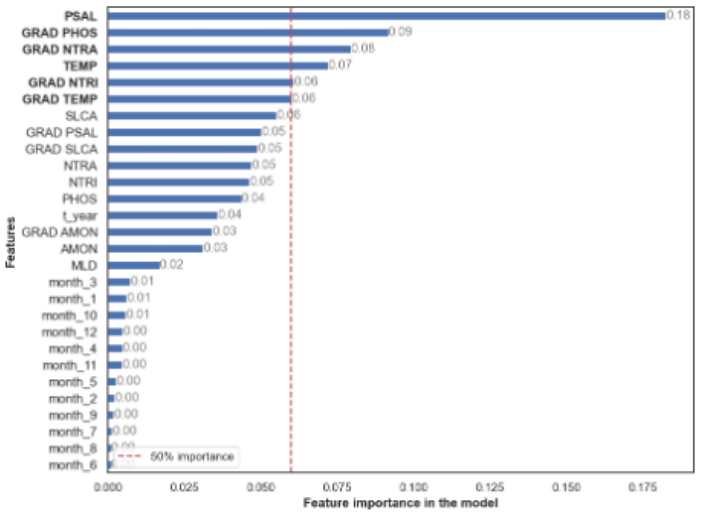
The main drivers of chlorophyll for each model were extracted from these functions. The results are illustrated in Figure B4.


PSAL 0.24
GRAD PHOS 0.09
PHOS 0.09
GRAD NTRI 0.08
NTRI 0.06
TEMP 0.06
NTRA 0.05
GRAD NTRA 0.05
SLCA 0.05
year 0.04
GRAD SLCA 0.04
GRAD TEMP 0.03
month 0.03
GRAD PSAL 0.03
AMON 0.03
GRAD AMON 0.03
MLD 0.02
50% importance
Features
Feature importance in the model


PSAL
TEMP
GRAD PHOS
PHOS
SLCA
GRAD NTRI
NTRI
GRAD TEMP
NTRA
GRAD NTRA
year
AMON
GRAD SLCA
month
GRAD PSAL
GRAD AMON
MLD
High
Low
Feature value
SHAP value (impact on model output)


PSAL 0.21
MLD 0.14
GRAD PHOS 0.11
GRAD PSAL 0.09
SLCA 0.08
AMON 0.05
NTRA 0.04
GRAD NTRI 0.04
NTRI 0.04
GRAD NTRA 0.03
PHOS 0.03
GRAD AMON 0.03
year 0.03
GRAD TEMP 0.02
GRAD SLCA 0.02
TEMP 0.02
month 0.01
50% importance
Features
Feature importance in the model


PSAL
GRAD PHOS
GRAD PSAL
MLD
year
SLCA
NTRA
GRAD NTRA
GRAD NTRI
NTRI
TEMP
GRAD SLCA
PHOS
AMON
GRAD AMON
GRAD TEMP
month
High
Low
Feature value
SHAP value (impact on model output)


PSAL 0.18
GRAD PHOS 0.09
GRAD NTRA 0.08
TEMP 0.07
GRAD NTRI 0.06
GRAD TEMP 0.06
SLCA 0.06
GRAD PSAL 0.05
GRAD SLCA 0.05
NTRA 0.05
NTRI 0.05
PHOS 0.04
t_year 0.04
GRAD AMON 0.03
AMON 0.03
MLD 0.02
month_3 0.01
month_1 0.01
month_10 0.01
month_12 0.00
month_4 0.00
month_11 0.00
month_5 0.00
month_2 0.00
month_9 0.00
month_7 0.00
month_8
month_6
50% importance
Features
Feature importance in the model

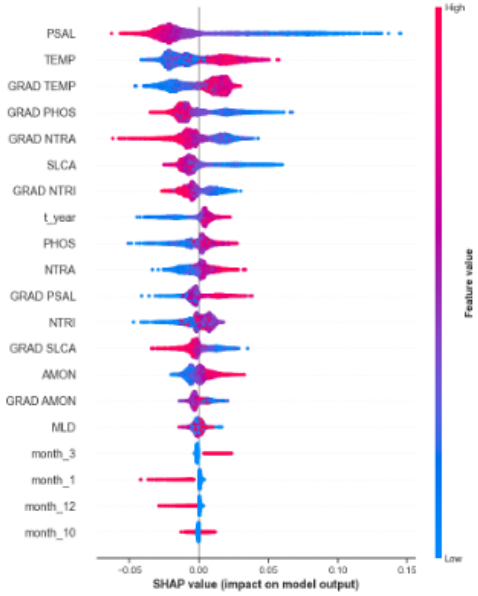
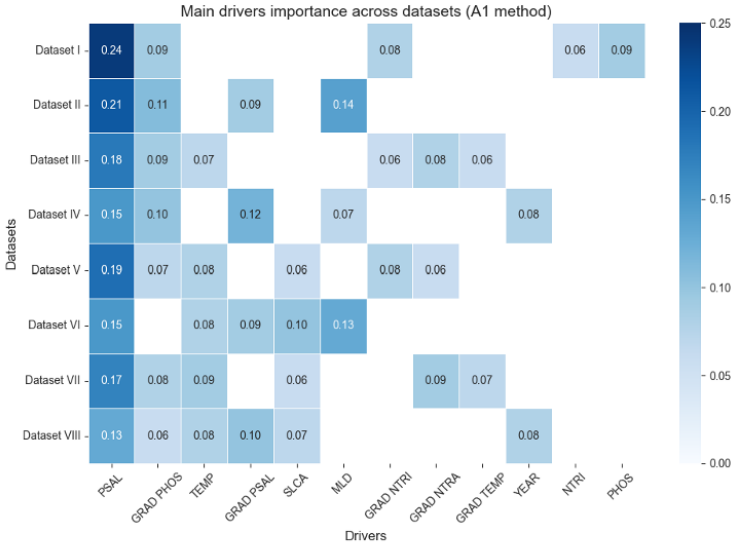

PSAL
TEMP
GRAD TEMP
GRAD PHOS
GRAD NTRA
SLCA
GRAD NTRI
t_year
PHOS
NTRA
GRAD PSAL
NTRI
GRAD SLCA
AMON
GRAD AMON
MLD
month_3
month_1
month_12
month_10
High
Low
Feature value
SHAP value (impact on model output)

**Fig. B3**. Features importance to the model (left) and to individual predictions (right). Results from using approach A1 in Dataset I (top), in Dataset II (center) and in Dataset III (bottom). The features that explain 50 % of the outcome are presented in bold, with the red line delimiting the 50 %.

This shows that salinity is the main driver regardless of the modifications to the dataset or methods used. The phosphate gradient is also a relevant driver for all cases but the ones with Dataset VI, and is a main driver in five of those cases.



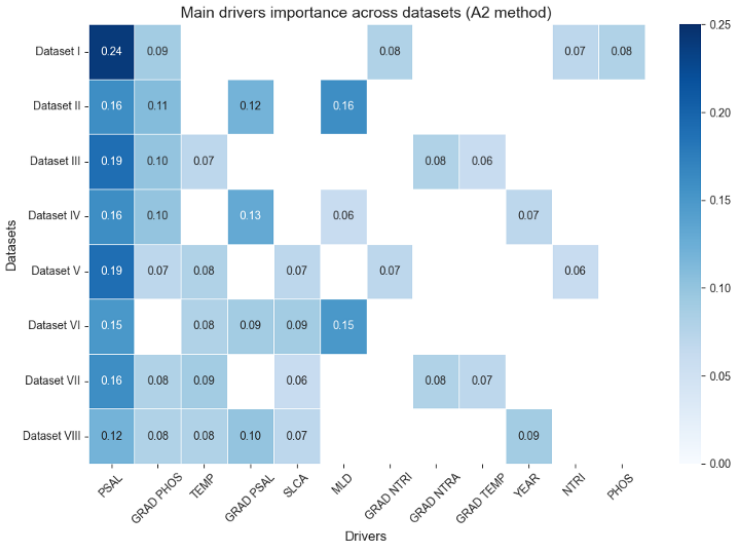




**Fig. B4.** Main drivers for each dataset with respective importances, for models trained with method A1 (top) and A2 (bottom).


Temperature is identified as a relevant driver for five out of eight datasets, including all datasets where the peaks were removed. This indicates that productivity is dependent on temperature in normal conditions, but not in peak conditions.

Likewise, silica is a relevant driver in all cases where the peaks were removed. The productivity of this area in normal conditions is therefore dependent on silica as well. Silica is an essential component of the structure of some phytoplankton species, which could explain this importance.

The gradient of salinity contributes approximately 10% to the overall prediction in all cases where the data is averaged. This suggests that the gradient of salinity influences the overall productivity of the area, rather than directly affecting the chlorophyll concentration at specific depths.

Similarly, the MLD is only selected as a relevant driver in models trained on averaged data (except for Dataset VIII), being a main driver in the non-transformed datasets (II and VI). Besides proving that the transformation disrupts the results, this points to the relevance of the MLD to the productivity of the water column. According to the respective Shapley Summary Plots, a low MLD can have a strong positive effect on the chlorophyll concentration. With MLD as a proxy of turbulence and mixing, this means that stratified waters (reduced mixing and turbulence) can benefit the production of chlorophyll. However, the Shapley Summary Plots also show that higher concentrations are always associated with lower MLD, and vice versa, but the impact of MLD is generally low. The influence of this variable is likely dependent on other features, which would explain this variability.

The gradient of nitrites, the gradient of nitrates and the gradient of temperature are only relevant in cases where the data was not averaged. Even though these drivers never contribute more than 9 % to the overall prediction, they influence the chlorophyll concentration over the water column and should be included in the prediction.

The year of the measurement had a relevant contribution to the prediction in the cases where the data was the most reduced (Dataset IV and VIII). On the contrary, the nitrite concentration is relevant in the cases where there is the most data. The reason behind these two importances is not clear, but is likely related to the size of the datasets relative to the remaining.

Only in models with Dataset I is the concentration of phosphate relevant to the overall prediction of chlorophyll concentration.

### B3.5 Discussion

Initially, this project addressed both the drivers of chlorophyll production in a given time and space, and the drivers of productivity of the entire water column. Additionally, the drivers under normal conditions were analyzed relative to the drivers in any circumstances, to investigate potential differences.

The training of the sixteen models led to the unanimous conclusion that salinity is the primary driver of productivity in this area, followed by the gradient of phosphate. Each modification to the original


dataset revealed different insights, and the interpretation of this combined set of results offers a good understanding of the dynamics in this area.

Under normal conditions, temperature and silica are the most relevant drivers, together with salinity and the phosphate gradient. In the cases where there is no peak removal, these parameters have little influence on both the productivity of the water column and the specific chlorophyll production.

The productivity of the water column is greatly influenced by the salinity gradient, the MLD (particularly in non-transformed data) and the year (particularly in transformed data), besides salinity and phosphate gradient.

Chlorophyll production depends mostly on the gradient of temperature, nitrites and nitrates, and on temperature, together with salinity and the phosphate gradient. When considering all measurements without further modifications, the nitrites and phosphates were relevant as well.

Transforming the data decreased the accuracy of the models without altering the fea- tures importance results, motivating the use of non-transformed data in future analysis.

Regarding the training-testing data splitting methods, the first approach produced more accurate predictions than the second approach. However, this approach is only suitable for analyzing entire datasets to extract the interactions between the system's components. When the aim is to predict values based on past measurements-such as forecast or regime shift detection algorithms- the second approach must be applied, because the first one is not applicable in these scenarios.

The method used to split the data into training and testing sets had little to no impact on the features' importance. This strengthens the results, as two different approaches produced similar/identical conclusions.

No explicit seasonal analysis was conducted. However, the month was included as a month to allow for seasonal patterns to emerge, as an indirect seasonal analysis. The results pointed to inexistent seasonal variations, with very low importances given to the month variables in all cases.

Further analyses include implementing different transformations with the aim of enhancing the performance of the model; taking into consideration the quality flags in the original dataset; conducting explicit seasonal analysis.

### B4 Regime Shift Detection: Changes in drivers over time

#### B4.1 Regime Shift Detection Algorithm

This task requires a time-series; therefore, Datasets II, IV, VI and VII were used. The decision to use four datasets, rather than selecting just one, was made to prevent potential bias in the results. By applying the same method to all four datasets, a comparison of the results could be made, allowing for more robust conclusions.


This algorithm is designed to detect regime shifts in multi-dimensional data by analyzing prediction performance errors. At each time step, a Random Forest Regression model is trained using the data from the current regime (X samples) to predict the chlorophyll concentration for that time step. The same hyperparameters were used for each of the four models (one for each dataset):

params ={'n_estimators': 300, 'min_samples_split': 8, 'min_samples_leaf': 3, 'max_features': 'log2', 'max_depth': 11, 'ccp_alpha': 0, 'bootstrap': False}

based on the hyperparameter tuning results from the models trained in the previous section.

Additionally, a baseline model is created, which predicts the chlorophyll concentration as the measured value in the same month in the previous year. The accuracy of these predictions is evaluated using the same metrics as before: MAE, MSE, RMSE, $R^2$, and correlation coefficient. If the Random Forest model performs worse than a given threshold ($R^2$), the algorithm flags a regime shift. The process continues iteratively until the last sample is reached.

The algorithm records the start date, end date, and feature importances for each identified regime. At the end of the process, a new model is trained for each regime using all the samples within that regime, and the feature importances are recalculated.

It is worth noting that the baseline model was initially designed to be the threshold for detecting regime shifts. However, its very poor performance made it impossible for the RF model to underperform relative to it. As a result, the threshold value was adjusted throughout this study to investigate the stability of the system in terms of regime shifts.

### B4.2 Model performance

#### B4.2.1 Detection error threshold

The first analysis investigated the effect of different detection error thresholds on the results, namely the number of detected regimes. The datasets in use have 23 years of data, so that is the maximum number of regimes that the algorithm can detect.

Figure B5 illustrates the influence of the detection error threshold on the number of detected regimes.

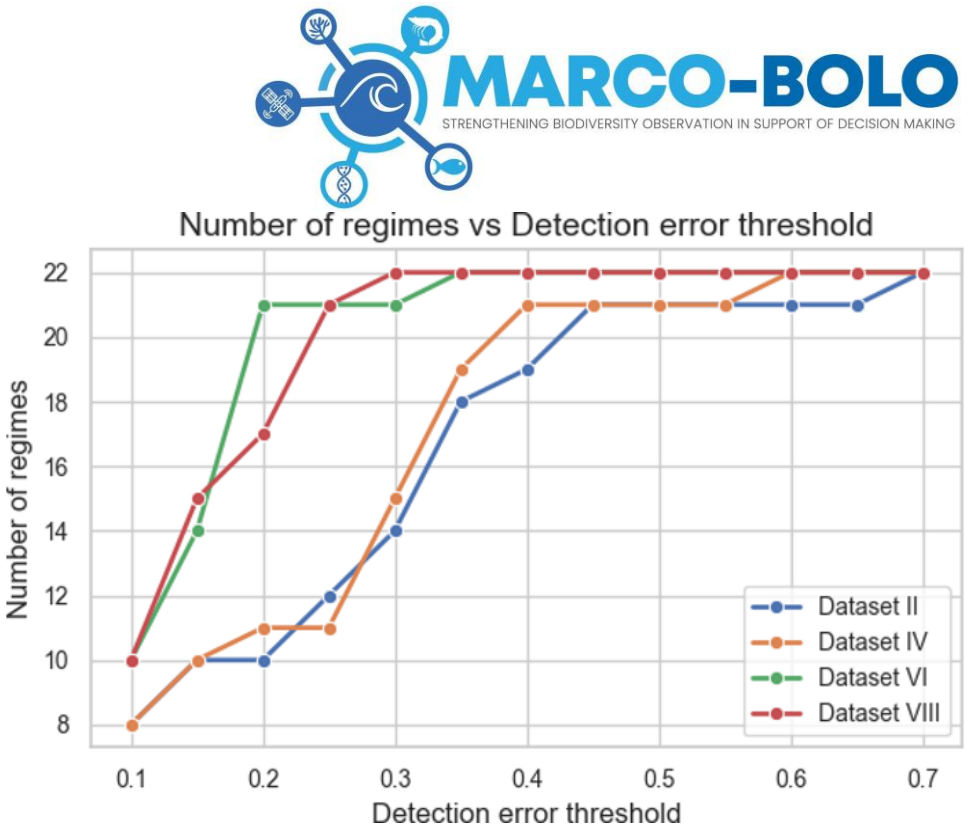



**Fig. B5**. Number of detected regimes in terms of the detection error threshold ($R^2$) used, for the four datasets in use. Dataset IV corresponds to Dataset II after transformation; Dataset VIII corresponds to Dataset VI after transformation

The lines for Datasets II and IV, as well as for Datasets VI and VIII, are closely aligned because they are essentially the same data. Datasets IV and VIII are transformed versions of Datasets II and VI, respectively.

The number of detected regimes is strongly influenced by the detection error threshold. While a dependency on the threshold was anticipated, it was unexpected that the number of regimes would remain so high until the threshold was reduced significantly. For Datasets II and IV, the average duration of a regime increases to 2 years when regime shifts are identified at an $R^2$<0.2. For Datasets VI and VII, this occurs only when the threshold reaches $R^2$<0.1. At detection error thresholds above 0.45, however, the number of detected regimes equals

22, suggesting a new regime for every year. This implies that the interactions between drivers and chlorophyll production vary significantly from year to year, and making predictions based on the previous year will produce poor results.

There are two interpretations for this behavior: (1) the system is inherently stable, with regime shifts only apparent at very low error thresholds, or (2) the system is highly unpredictable, driven by a complex combination of variables. The latter suggests that this approach may not be suitable for detecting regime shifts in such a dynamic environment.

# MARCO-BOLO

STRENGTHENING BIODIVERSITY OBSERVATION IN SUPPORT OF DECISION MAKING

**Project Coordinator**
Nicolas Pade | nicolas.pade@embrc.eu

**Project Manager**
Giulia Vecchi | giulia.vecchi@embrc.eu

**Press and Communications**
Mathilde Vidal | mathilde@erinn.eu

Website: MarcoBolo-Project.eu
Twitter: @MARCOBOLO_EU
LinkedIn: MARCO-BOLO
Bluesky: @marco-bolo.bsky.social
Instagram: marco_bolo_project

Funded by the European Union

Funded by the European Union under the Horizon Europe Programme, Grant Agreement No. 101082021 (MARCO-BOLO). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

UK Research and Innovation

UK participants in MARCO-BOLO are supported by the UKRI's Horizon Europe Guarantee under the Grant No. 10068180 (MS); No. 10063994 (MBA); No. 10048178 (NOC).