Mapping the biota and quantifying separate and synergistic impacts of human uses on joint nature assets*

Activity Output T2.1.1

Jonne Kotta, Mihhail Fetissov , Kristjan Herkül, Helen Orav-Kotta , Anneliis Kõivupuu, Tiia Rosenberg, Marge Simo, Robert Christian Szava-Kovats, Annaleena Vaher

Tallinn, 2021

Contents

Importance of mapping the biota

A fundamental component for the design of efficient policy and management actions is transboundary knowledge on the spatial patterns of different species, groups of species and biodiversity elements. This knowledge is a bases for developing well-informed ecosystem-based marine management systems. However, to date such information is not available for the entire Gulf of Finland area. The Activity Output T2.1.1 of the Adrienne project produced harmonized transboundary maps of important biological elements that are crucial for understanding the general spatial patterns of the marine nature assets.

Generic characteristics of the programme area and its nature values

Materials for the ADRIENNE project originates from Estonian, Russian and Finnish marine areas of the Gulf of Finland, located in the North-Eastern Baltic Sea. The Gulf of Finland exhibits strong gradients of wave exposure, depth, and salinity. The sea areas west of the Gulf are exposed to the open Northern Baltic Proper and have a wave fetch of hundreds of kilometres. In contrast, the inner reaches of the bays of the mainland are very sheltered both by the mainland and by islands. Salinity exceeds 7 PSU in the westernmost study area while it falls to almost 0 PSU in the inner parts of bays with riverine inflow. The Gulf of Finland is characterized by a variable fauna and flora, which has species-specific tolerance towards natural and human-induced pressures.

Biological data

In order to assess the impacts of different human pressures on different nature values accurate maps of nature assets are needed.

The data for such maps can be obtained from different sources. We used the biological input data for species distribution modeling originated from the following sources (number of unique geographical sampling locations are shown in brackets):

- Finnish VELMU database (~28000)
- Finnish POHJE database (~1600)
- Benthos database of the Estonian Marine Institute, University of Tartu (~11000)
- Russian benthos data collated by Saint Petersburg Scientific Center RAS (~147)

Collection and analysis of the samples followed the methodologies described in the report of the Output_T1.2.1_Harmonizing_methods_of_field_sampling,_sample_analysis_and_spatial_modelling.

In the current modelling activity, we selected altogether 57 benthic species or groups of species to represent benthic nature values in the Gulf of Finland. These species/groups were selected based on occurrence rate (generally occurrence > 2.5 %) and ecological relevance (habitat formation, important ecological processes).

The modelled species/groups were as follows: bay barnacle (*Amphibalanus improvisus*), wandering pond snail (*Ampullaceana balthica*), *Battersia arctica*, *Ceramium tenuicorne*, lagoon cockle (*Cerastoderma glaucum*), hornwort (*Ceratophyllum demersum*), charophytes (*Chara* sp.), chironomids (*Chironomidae* sp.), dead man's rope (*Chorda filum*), *Cladophora glomerata*, *Cladophora rupestris*, *Coccotylus truncatus*, *Dictyosiphon foeniculaceus*, zebra mussel (*Dreissena polymorpha*), drifting macrophytes, epifaunal bivalves (fresh water), filamentous algae, filamentous brown algae, filamentous green algae, filamentous red algae, bladder wrack (*Fucus* sp.), clawed fork weed (*Furcellaria lumbricalis)*, *Gammarus* sp., *Halicryptus spinulosus*, ragworm (*Hediste diversicolor)*, *Hildenbrandia* sp., hydrozoan, *Idotea* sp., infaunal bivalves, infaunal bivalves (fresh water), *Jaera* sp., Baltic clam (*Limecola balthica*), *Marenzelleria* sp., *Monoporeia affinis*, sand gaper (*Mya arenaria)*, water milfoil (*Myriophyllum* sp.), bay mussel (*Mytilus trossulus*), spiny water nymph (*Najas marina*), oligochaete, other polychaeta, clasping-leaved pondweed (*Potamogeton perfoliatus*), *Pylaiella littoralis* & *Ectocarpus siliculosus*, *Rhodomela confervoides*, beaked tasselweed (*Ruppia maritima)*, *Saduria entomon*, snails, freshwater snails (fresh water), *Stictyosiphon tortilis*, *Stuckenia pectinata*, river nerite (*Theodoxus fluviatilis*), thick brown algae, thick red algae, *Ulva* sp., vascular plants, *Vertebrata fucoides*, horned pondweed (*Zannichellia palustris)*, eelgrass (*Zostera marina*)

Abiotic environmental data

The key abiotic environmental variables (Table 1) used in the modelling task included different bathymetrical, hydrodynamic and physico-chemical variables.

Table 1. List of georeferenced environmental variables that were considered and finally used (in bold) as predictor variables in the models (reference to the data origin in superscript). Due to low confidence or high correlation with some other variable(s), some of the preliminary variables (not in bold) were not included in the final models.

Variable name	Abbrevation
Water depth[1]	**depth**
Seabed slope[1]	
Surface wave exposure[2]	**swm**
Bottom wave exposure[3]	
Salinity in bottom layer[4]	**salinity**
Temperature in bottom layer[4]	**temp**
Water current speed in bottom layer[4]	
Sea ice cover[4]	**ice_conc**
Sea ice thickness[4]	
Chlorophyll a concentration in surface layer[4]	**chl**
Ammonium concentration in bottom layer[4]	
Nitrate concentration in bottom layer[4]	**no3**
Phosphate concentration in bottom layer[4]	**po4**
Dissolved oxygen concentration in bottom layer[4]	
Secchi depth[4]	**secchi_depth**
Wave direction[4]	
Wave height[4]	**wave_height**

[1]Baltic Sea Hydrographic Commission (2013) Baltic Sea Bathymetry Database version 0.9.3. Downloaded from http://data.bshc.pro/

[2]van der Meijs F, Isaeus M (2020) Wave exposure calculations for the Gulf of Finland. AquaBiota Water Research, AquaBiota Report 2020:13. This environmental input layer was a special product of the ADRIENNE project.

[3]Depth-corrected wave exposure calculated from surface wave exposure (source no. 2) based on method by Bekkby T, Isachsen PE, Isaeus M, Bakkestuen V (2008) GIS modeling of wave exposure at the seabed: a depth-attenuated wave exposure model. Marine Geodesy 31, 117-127

[4]Copernicus Marine Environment Monitoring services products BALTICSEA_ANALYSIS_FORECAST_PHYS_003_006, BALTICSEA_REANALYSIS_FORECAST_BIO_003_007, OCEANCOLOUR_BAL_CHL_L3_NRT_OBSERVATIONS_009_049

The variables that were selected for the final models fulfilled the following criteria:
- Ecological importance in relation to the spatial distribution of benthic species
- Data availability in the full extent of the study area
- Availability of data to be used in future scenario modeling. The future scenario modeling will be done in the next stages of the project.

Modeling methods

The most widely used benthic sampling devices such as grabs, trawls and underwater video or photography (Eleftheriou and McIntyre 2005) yield information only from the visited sites (point-wise data), leaving most of the study area unsampled (Herkül et al. 2013). Mathematical predictive modeling based on species–environment relationships (Figure 2) provides a useful framework to synthesize information from scattered samples into coherent seamless maps of distributions of species and habitats, species richness, ecological goods and services (Guisan and Zimmerman 2000; Guisan and Thuiller 2005). These models are numerical methods that relate measurements of biotic variables (e.g. species occurrence or abundance, species richness) to environmental variables (Elith and Leathwick 2009). These relationships are further used to predict the distribution of values of biotic variables across different spatial and/or temporal scales (Elith and Leathwick 2009).

Boosted regression trees (BRT) and random forest (RF) machine learning modeling methods were deployed in this modelling task. BRT was chosen to construct final models due to the possibility of assigning monotonicity to each independent variable. This enables to incorporate the known cause-effect relationships into the modelling framework and thereby improve robustness and precision of the model predictions. Mathematical programming language R (R Core Team 2020) in RStudio software (RStudio Team 2020) was used for data preparation and modeling. An R package *randomForest* (Liaw and Wiener 2002, Breiman et al. 2018) was used for RF and package *gbm* (Greenwell et al 2020) for running BRT models. Model predictions were made in a 1 km grid.

Transboundary biota maps covering the whole programme area

Spatial predictive models produced map layers of 57 different species or group of species covering the entire Gulf of Finland. Probability of occurrence and presence-absence maps (Figure 1) were produced for all species/groups.

Please find some additional species or group of species map layers under the Appendix 1. All the map layers produced were uploaded into the GIS portal environment that is developed under the WP3.
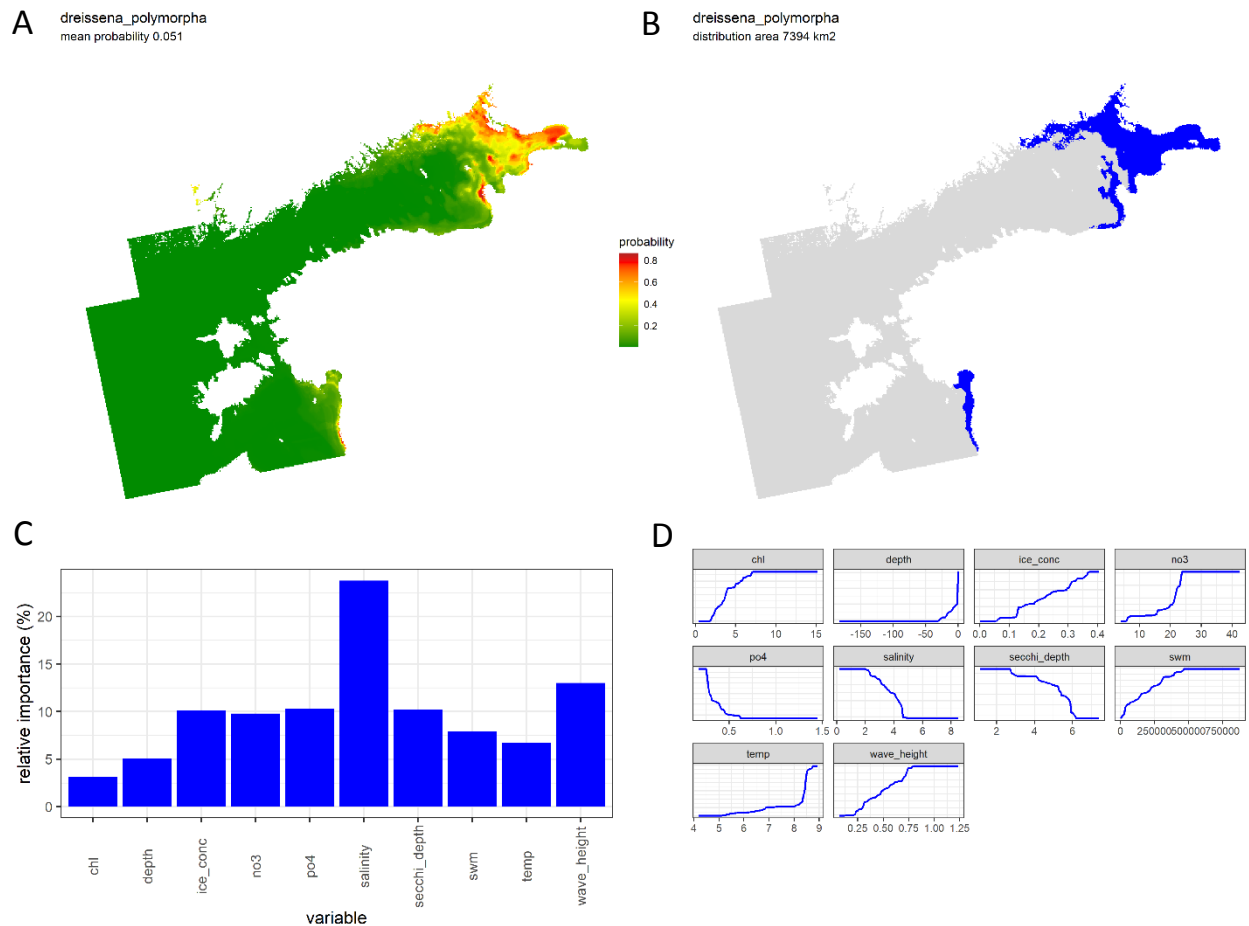

Figure 1. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the zebra mussel (*Dreissena polymorpha*) together with the relative importance of the environmental predictor variables (C) and their response curves (D).

Quantifying separate and synergistic impacts of human uses on joint nature assets

Intensification and diversification of human-induced pressures in marine ecosystems have raised concerns over several sustainability-diminishing consequences, such as hypoxia and overexploitation of resources. Benthic species and habitats are considered to be one of the most important drivers of species richness thus, the modification or loss of habitats can be a serious threat to marine ecosystems (Airoldi and Beck 2007). In the Baltic Sea, the number of habitat-forming species is relatively low therefore in case of a habitat decline, there are not many alternative species to replace the function of that habitat in the ecosystem. Natural stresses adding to the anthropogenic pressures may cause a permanent decline of valuable benthic habitats. Because of its limited water exchange with the ocean, the biota if the Baltic Sea ecosystems are very sensitive to both natural and anthropogenic stresses (Bonsdorff 2006; Zettler et al. 2013). The Baltic Sea is surrounded by nine industrialized countries and over 85 million people live in the Baltic catchment area, (HELCOM 2016) which itself is about four times larger than the surface of the sea, and is therefore currently impacted by multiple anthropogenic pressures.

Successful management, restoration and conservation of intensively used coastal ecosystems demands knowledge of the response of key species and habitats to the increasing pressure of the combined effects of multiple stressors. Over the last decade, however, a large body of literature has evolved that specifically targets interactive effects of multiple pressures on a large variety of ecosystem assets and their services (e.g. Przeslawski et al. 2015; Gunderson et al. 2016). Nevertheless, there exists no suitable models with which to disseminate the complex relationships between pressure, nature assets and ecosystem services to important stakeholders responsible for management of nature assets.

Under the Adrienne project we developed the PlanWise4Blue (PW4B) online GIS based portal. PW4B is an online decision support tool (DST) that performs Cumulative Effect Analysis (CEA) on nature values. The tool incorporates the majority of regional scientific evidence in a way that its algorithm is capable of quantifying both single and synergistic effects of most important human activities on a broad range of nature assets (Kotta et al. 2020).

Data collection

The analysis procedure of this study involves 1) meta-analysis of published or raw data that indicated separate and/or synergistic impacts of the studied human pressures (either from experimental manipulations or ecosystem changes observed before and after impact) and 2) linking the impact data (effect-size estimates) and existing spatial prediction of different nature assets into a cumulative impact assessment framework (e.g. Kotta et al. 2020).

In order to predict the plausible impacts of different human pressures on nature values, the current knowledge from published literature and available datasets needs to be compiled. First, the data for meta-analysis was extracted from scientific articles that fulfilled the following criteria:

1. The study was conducted in the Baltic Sea. In few exceptional cases when limited amount of data was available about certain human impacts from the Baltic sea regions (e.g. wind park development), we gathered data from an area that share similar habitats to the Baltic Sea, such as the North Sea region.

2. There were comparable values from impacted and reference areas i.e. impact/reference comparison could be made both in space (comparison between control and impact sites) or time (comparison over time).

When fulfilling these inclusion criteria, the quantitative statistics of all measurements that were related to the effect of human pressures on the studied habitats were extracted. When possible, we extracted means with measurement unit, standard errors, standard deviations and sample sizes of control and impact values directly from tables and the text of the articles. Alternatively, we used ImageJ software to extract relevant comparisons from figures (Schneider et al. 2012). The extracted quantitative data on the effects of human pressures on benthic habitats were then used to calculate respective effect sizes. Mathematical formulae to calculate effect sizes and their corresponding uncertainty follow Kotta et al. (2020). Then the compiled scientific evidence (experimental and survey data) was uploaded to the PW4B portal to predict the environmental effect of the studied pressures on benthic habitats. The tool uses the nature-value and pressure-specific coefficient of cumulative effect in each region of interest that is then multiplied by the respective value of the nature asset to ascertain the expected changes of this nature asset (Kotta et al. 2020).

Altogether, we collected and summarized information about the following 19 separate or combined anthropogenic pressure types:

Dredging and dumping areas, extraction of minerals, coastal protection, windpark areas, underwater cable areas, harbours, shipping intensity, pelagic trawling, benthic trawling, marine plant harvesting, fish farming, mussel and algal cultivation, wastewater discharge outlet, riverine nutrient inflow, tourism and leisure activities, military activities, species introductions (*Neogobius melanostomus*), species introductions (*Rhithropanopeus harrisii*), modified wave climate.

All the data was combined into one general Excel mastersheet, in total more than 1500 rows.

The separate and combined impacts of human uses mentioned above, were looked on the following separate nature values:

Mammal – populations, Bird - Benthos feeders, Bird - Fish feeders, Bird – Herbivores, Bird - Migration routes, Bird - Wintering areas, Fish - Herring spawning areas, Fish - Pikeperch spawning areas, Fish - Whitefish spawning areas, Fish - Adult populations (benthic), Fish - Adult populations (pelagic), Invertebrates - Large species (benthic), Invertebrates - Large species (pelagic), Habitat – Charophytes, Habitat – Fucus, Habitat – Furcellaria, Habitat - Higher plants, Habitat - Richness flora and fauna, Habitat - Suspension feeders, Habitat – Infauna, Habitat – Zostera, HD – Sandbanks, HD - Mudflats and sandflats, HD – Reefs, Zooplankton, Phytoplankton, Bacteria, Detritus.

Developing the algorithm to calculate separate and synergistic impacts of human uses on nature values

Accurate CEA assessments require solid ecological understanding of cause-effect relationships between pressures and biota and sound estimates of associated uncertainties. Because the total effect is not the sum of single effects but interactions overwhelmingly prevail in nature, it is essential that the synergistic effects of different pressures on nature assets are also quantified and integrated into the assessment. The existing assessments for the Baltic Sea region and for other European waters, however, are not yet able to incorporate this complexity and express impact as the sum of the individual effects of different pressures on different nature assets.

Mathematical formulae to calculate effect sizes and their corresponding uncertainty follow Kotta et al. (2020). The calculation of impact coefficients requires logarithmic mean of a human-induced impact e_i (or series of impacts) and the logarithmic mean of a control e_c.

$$E_i = \ln(e_i) \text{ and } E_c = \ln(e_c) \tag{1}$$

The effect of an individual study (E_S) is defined as the difference between the impact (E_I) and the control (E_C):

$$E_S = E_I - E_C \tag{2}$$

The uncertainty of E_S (U_S) is calculated from the 95% confidence interval of the impact (U_I) and the control (U_C):

$$U_S = \sqrt{U_I{}^2 + U_C{}^2} \tag{3}$$

If necessary, the 95% confidence interval (U) can be calculated from the standard deviation (SD) or standard error (SE):

$$U = SD * t_{0.05(2),N-1}/\sqrt{N} \tag{4}$$

$$U = SE * t_{0.05(2),N-1} \tag{5}$$

where N is the number of samples and $t_{0.05(2),N-1}$ is the t-score.

Demonstrating the results of algorithm, used to calculate impacts of human uses

In order to demonstrate the applicability of the developed algorithm that can calculate separate and synergistic effects of human uses on different nature assets we built 7 scenarios:

1. Current nutrient load
2. Future nutrient reduction (HELCOM MAI target)
3. The presence of non-indigenous species (round goby and mud crab)
4. Projected windparks (according to the Estonian Maritime Spatial Plan)
5. Current nutrient load + non-indigenous species

6. Current nutrient load + non-indigenous species + windparks
7. Future nutrient reduction + non-indigenous species + windparks

An example of results of the above scenario analyses can be found below. In Figure 2 you can see how the abundance of birds in the wintering areas is expected to be changed under the pressures of current nutrient inflow, alien species (round goby and mudcrab) and windparks. Such map layers are very important for nature protection perspective and in Maritime Spatial Planning process as they clearly indicate the magnitude and direction of change of certain nature values under different combinations of human pressures. Resulting maps give a valuable information for policy makers, to make ecosystem based decisions.
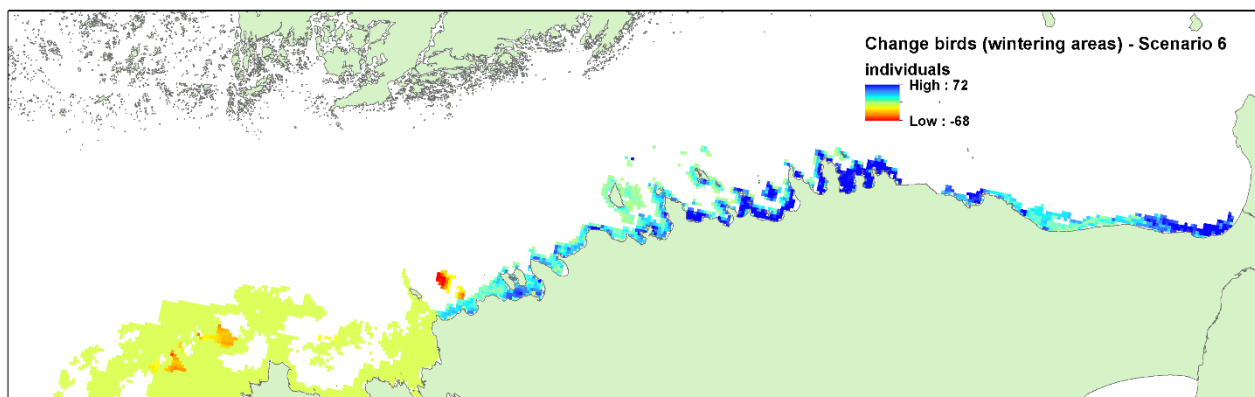
Figure 2. Change in the abundance of birds in their wintering areas under the synergic impact of current nutrient inflow, alien species (round goby and mudcrab) and windparks.

Future steps

All the results that were introduced in the report will be uploaded and integrated into the GIS based online tool (PW4B) during the next periods of the project.

References

Baltic Sea Hydrographic Commission (2013) Baltic Sea Bathymetry Database version 0.9.3. Downloaded from http://data.bshc.pro/

Bekkby T, Isachsen PE, Isaeus M, Bakkestuen V (2008) GIS modeling of wave exposure at the seabed: a depth-attenuated wave exposure model. Marine Geodesy 31, 117-127

Bonsdorff E (2006) Zoobenthic diversity-gradients in the Baltic Sea: continuous post-glacial succession in a stressed ecosystem. Journal of Experimental Marine Biology and Ecology 330, 383–391

Breiman L, Cutler A, Liaw A, Wiener M (2018) randomForest: Breiman and Cutler's random forests for classification and regression. R package version 4.6-14. http://cran.r-project.org/web/packages/randomForest/

Copernicus Marine Environment Monitoring services products BALTICSEA_ANALYSIS_FORECAST_PHYS_003_006, BALTICSEA_REANALYSIS_FORECAST_BIO_003_007, OCEANCOLOUR_BAL_CHL_L3_NRT_OBSERVATIONS_009_049

Depth-corrected wave exposure calculated from surface wave exposure (source no. 2) based on method by Bekkby T, Isachsen PE, Isaeus M, Bakkestuen V (2008) GIS modeling of wave exposure at the seabed: a depth-attenuated wave exposure model. Marine Geodesy 31, 117-127

Eleftheriou A and McIntyre AD (2005) Methods for the Study of Marine Benthos, 3rd edn. Blackwell, Oxford.

Elith J and Leathwick JR (2009) Species distribution models: Ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40, 677-697

Greenwell B, Boehmke B, Cunningham J, GBM Developers (2020) gbm: Generalized Boosted Regression Models. R package version 2.1.8. https://CRAN.R-project.org/package=gbm

Guisan A and Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters 8, 993-1009

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

Gunderson AR, Armstrong EJ, Stillman JH (2016) Multiple stressors in a changing world: the need for an improved perspective on physiological responses to the dynamic marine environment. Annual Review of Marine Science 8, 357-378

HELCOM (2016): Hydrography and oxygen in the deep sea basins. HELCOM Baltic Sea Environment Fact Sheets. Online. 6.7.2017, http://www.helcom.fi/baltic-sea-trends/environment-fact-sheets/

Herkül K, Kotta J, Kutser T, Vahtmäe E (2013) Relating remotely sensed optical variability to marine benthic biodiversity. PLoS ONE 8, e55624.

Kotta J, Fetissov M, Szava-Kovats R, Aps R, Martin G (2021). Online tool to integrate evidence-based knowledge into cumulative effects assessments: Linking human pressures to multiple nature assets. Environmental Advances

Liaw A, Wiener M (2002) Classification and Regression by randomForest. R News 2(3), 18–22

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Przeslawski R, Byrne M, Mellin C (2015) A review and meta-analysis of the effects of multiple abiotic stressors on marine embryos and larvae. Global Change Biology 21, 2122–2140

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RStudio Team (2020) RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. http://www.rstudio.com/

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nature Methods, 9, 671–675

van der Meijs F, Isaeus M (2020) Wave exposure calculations for the Gulf of Finland. AquaBiota Water Research, AquaBiota Report 2020:13. This product was a special order in the ADRIENNE project.

Zettler ML, Karlsson A, Kontula T, Gruszka P, Laine AO, Herkül K, Schiele KS, Maximov A, Haldin J (2013) Biodiversity gradient in the Baltic Sea: a comprehensive inventory of of macrozoobenthos data. Helgoland Marine Research 68, 49–57

Appendix 1

Some additional examples from the produced species/group of species spatial distribution maps that cover the whole project area, Gulf of Finland.
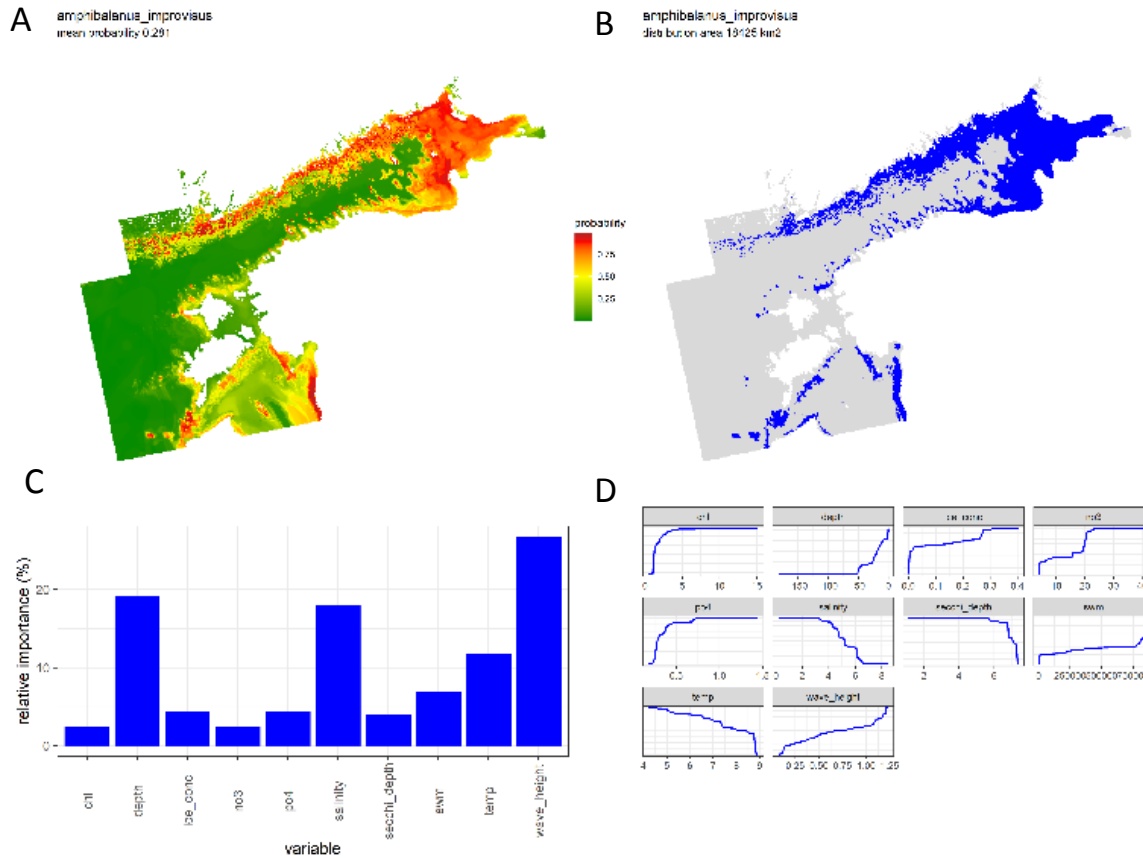
Figure A1. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the bay barnacle (*Amphibalanus improvisus*) together with the relative importance of the environmental predictor variables (C) and their response curves (D).
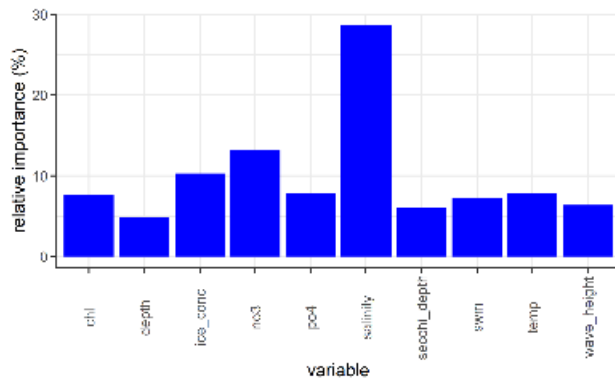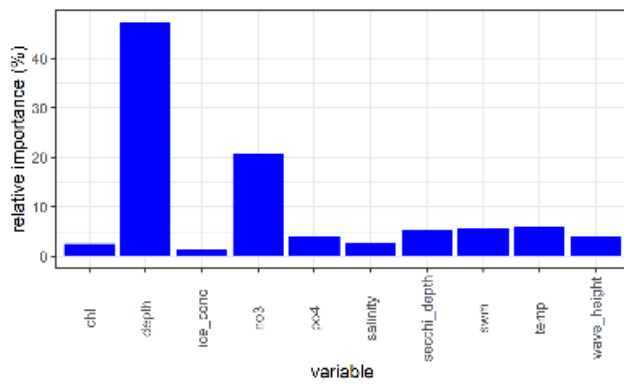
A
charophytes
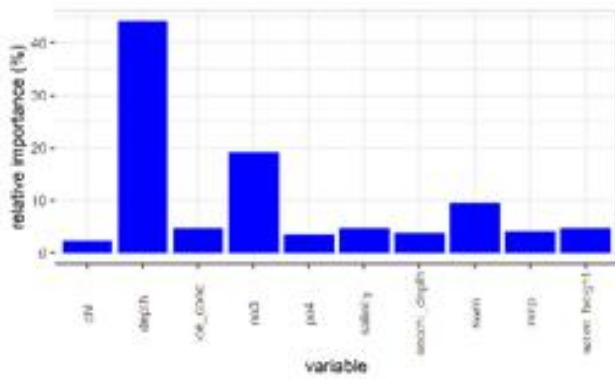mean probability 0.052
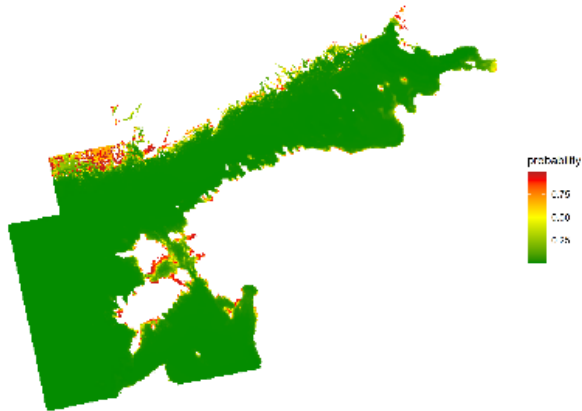
B
charophytes
distribution area 8567 km2
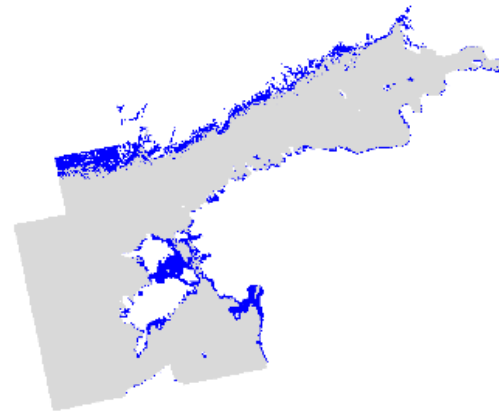
C
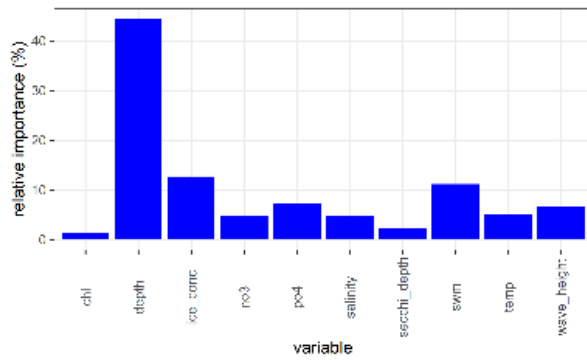
D
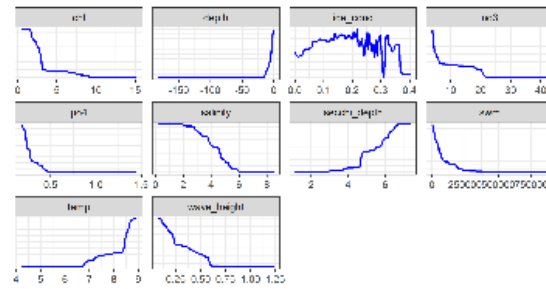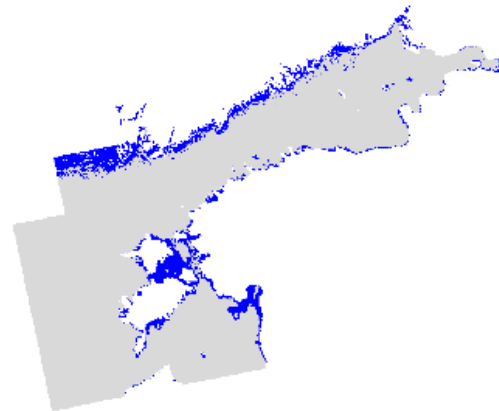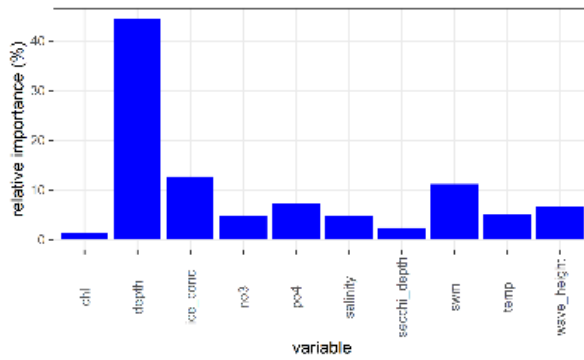
Figure A2. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the bay barnacle (*Amphibalanus improvisus*) together with the relative importance of the environmental predictor variables (C) and their response curves (D).
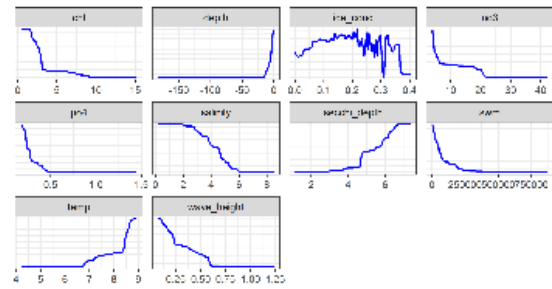
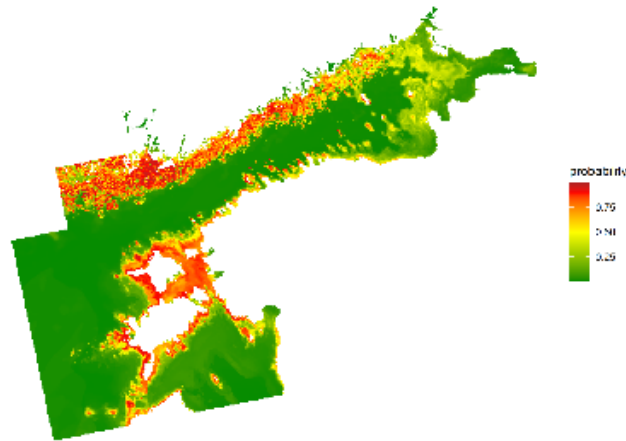**Figure A3.** Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the charophytes together with the relative importance of the environmental predictor variables (C) and their response curves (D).
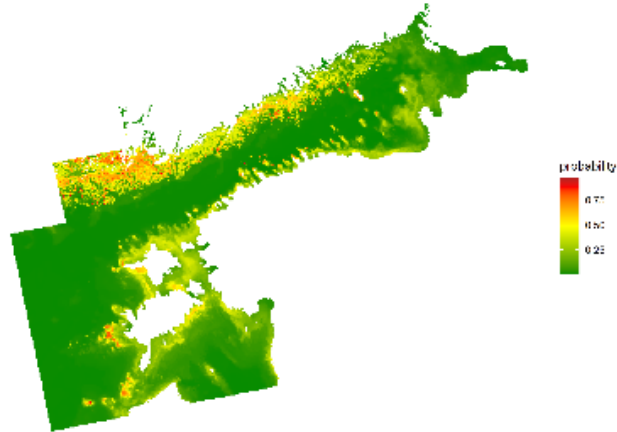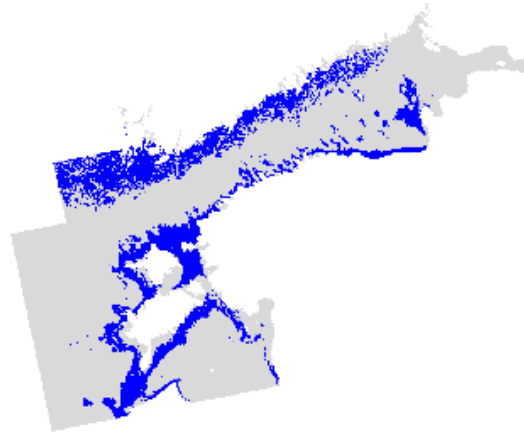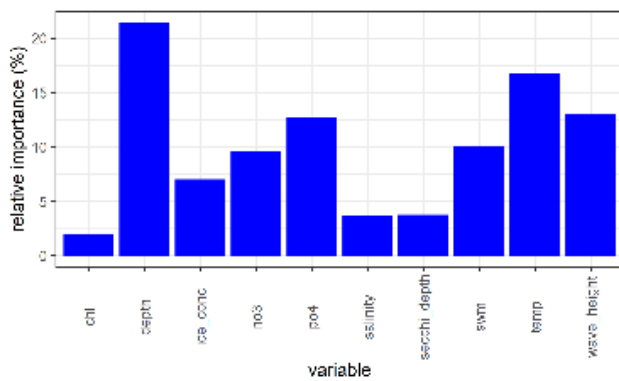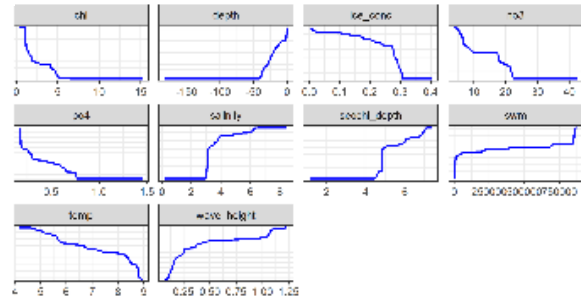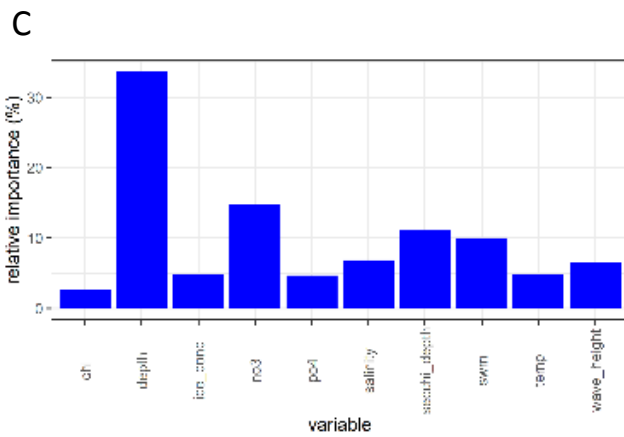
Figure A4. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the bay barnacle (*Amphibalanus improvisus*) together with the relative importance of the environmental predictor variables (C) and their response curves (D).
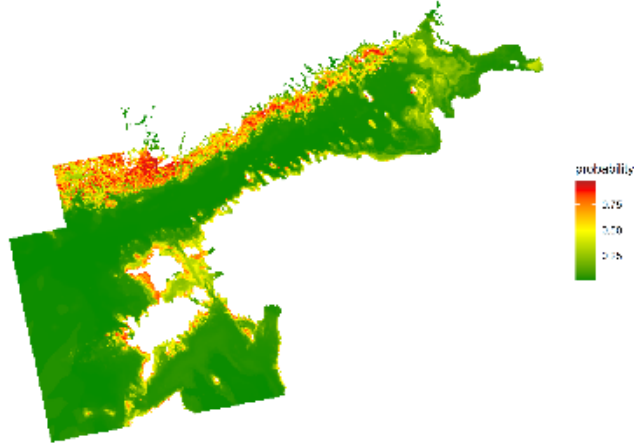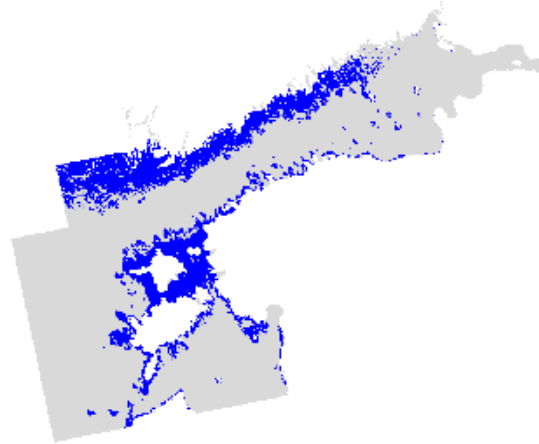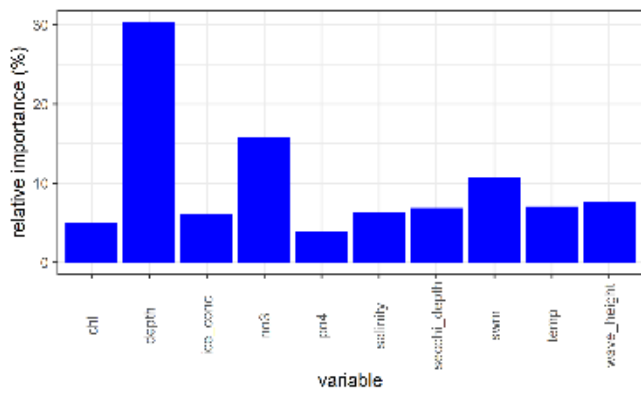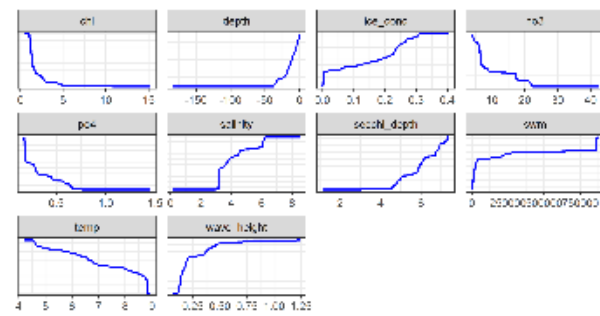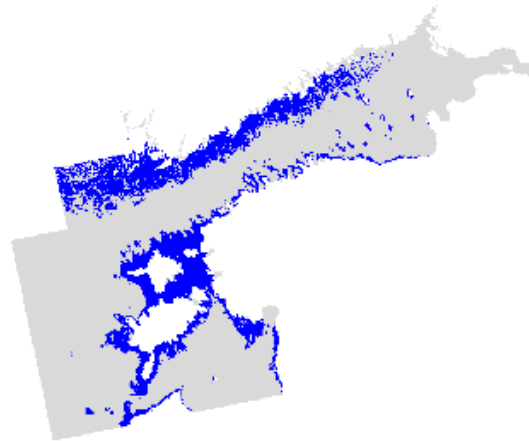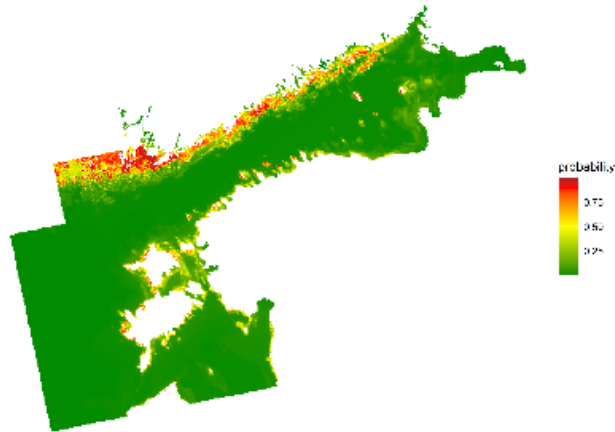
Figure A5. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the *Ceramium tenuicorne* together with the relative importance of the environmental predictor variables (C) and their response curves (D).
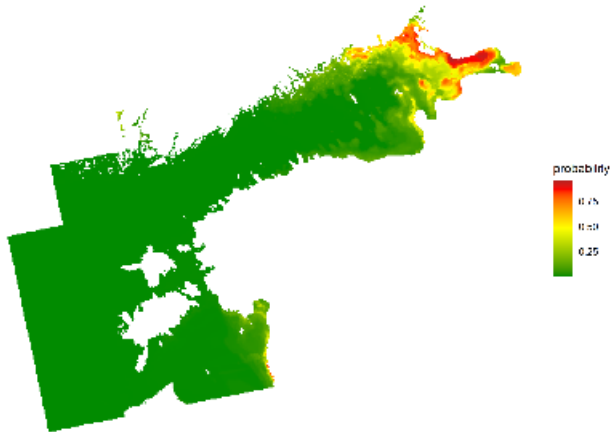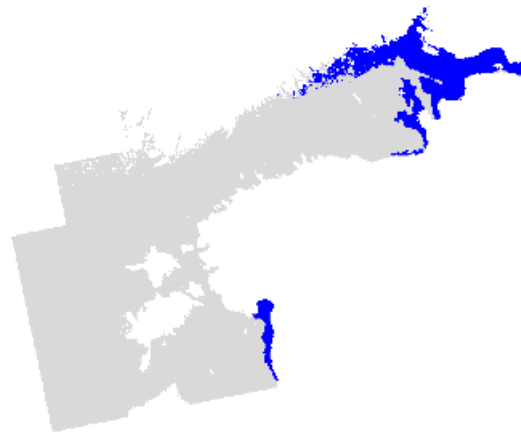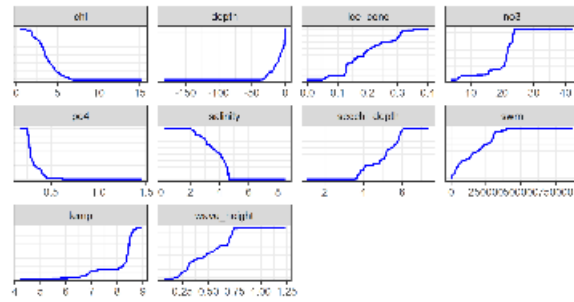
Figure A6. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the epifaunal bivalves (fresh water) together with the relative importance of the environmental predictor variables (C) and their response curves (D).
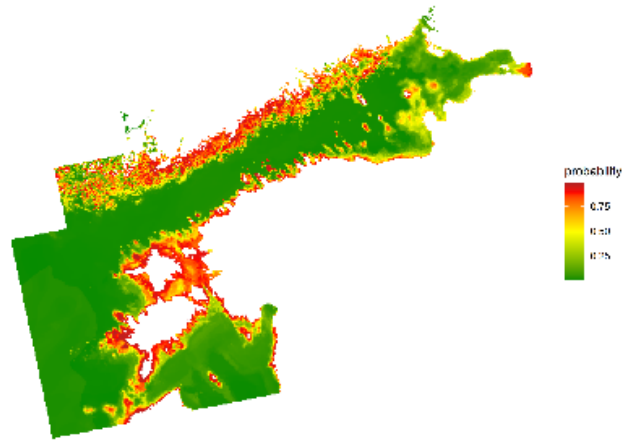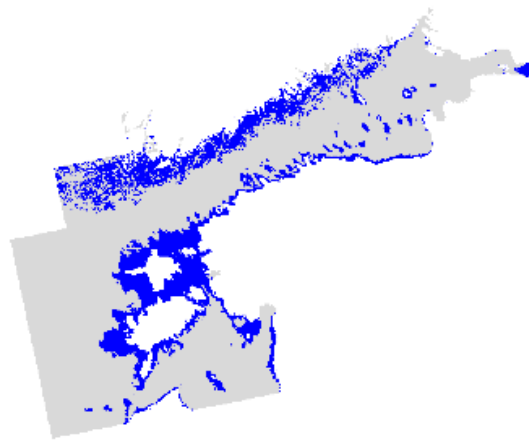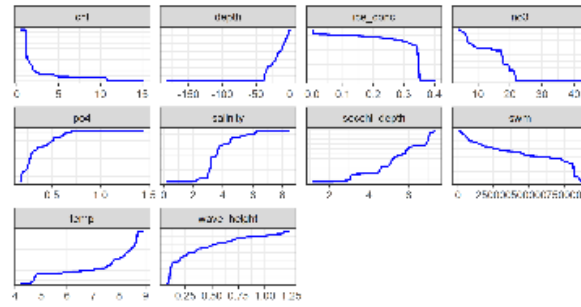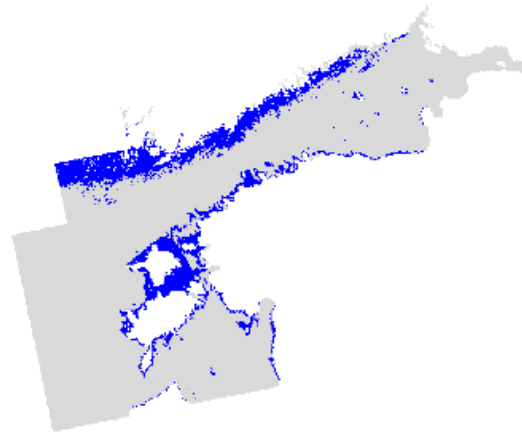
Figure A7. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the filamentous algae together with the relative importance of the environmental predictor variables (C) and their response curves (D).
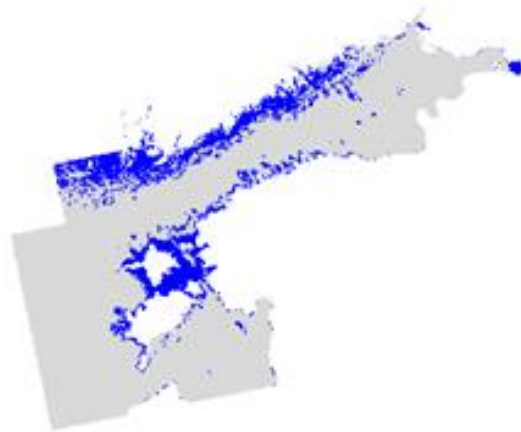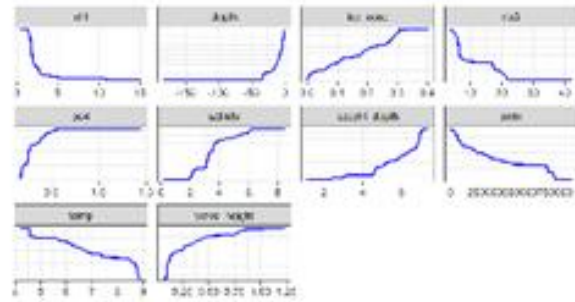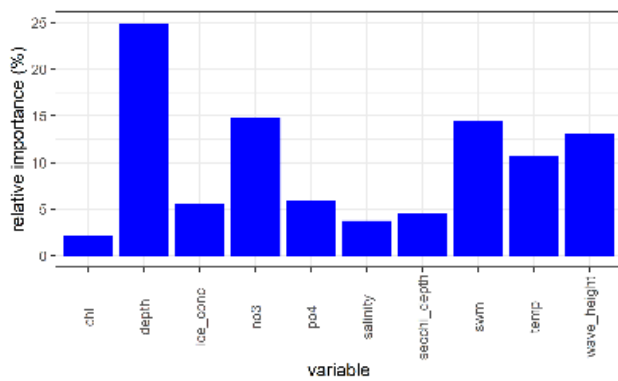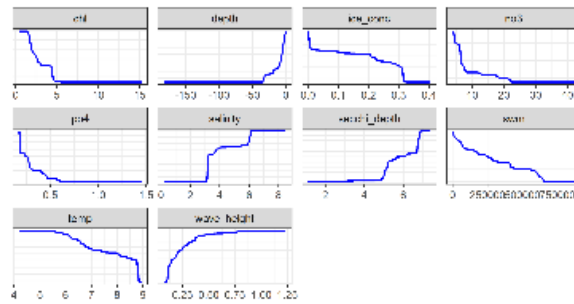
Figure A8. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the filamentous green algae together with the relative importance of the environmental predictor variables (C) and their response curves (D).
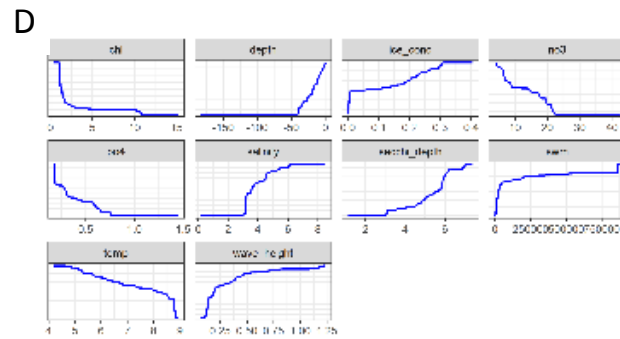
Figure A9. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the filamentous red algae together with the relative importance of the environmental predictor variables (C) and their response curves (D).
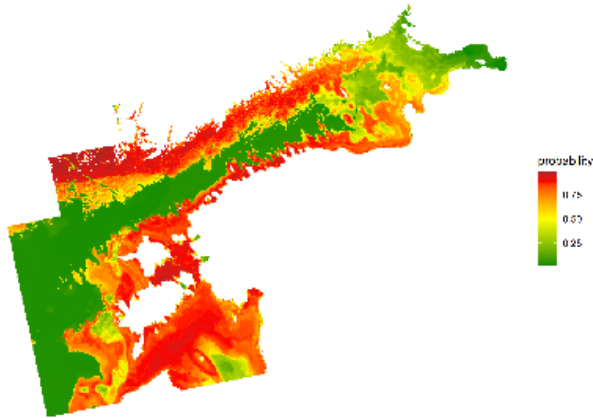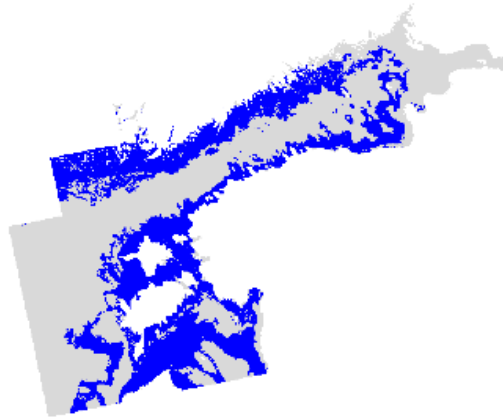
Figure A10. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the bladder wrack (*Fucus* sp.) together with the relative importance of the environmental predictor variables (C) and their response curves (D).
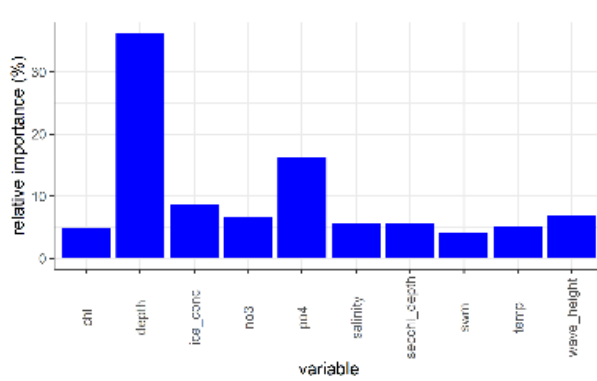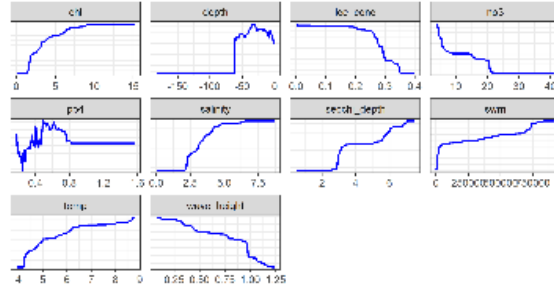
Figure A11. Maps of predictions of the probability of occurrence (A) and presence-absence (B) of the Baltic clam (*Limecola balthica*) together with the relative importance of the environmental predictor variables (C) and their response curves (D).