

# Report on the harmonizing methods of field sampling, sample analysis and spatial modelling\*

## Activity Output T1.2.1

Jonne Kotta, Kristjan Herkül, Anneliis Kõivupuu, Filipp Leontiev, Helen Orav-Kotta, Marina Orlova, Tiia Rosenberg, Robert Szava-Kovats, Annaleena Vaher, Yuriy Zuyev



Tallinn, 2020

## Table of Content

Generic characteristics of the programme area and its nature values .....	3
Biological data .....	3
Abiotic environmental data.....	7
Modeling methods.....	8
Data management and software .....	10
Challenges and solutions related with data and modelling process.....	11
Storage of modelling products .....	11
References.....	12

## 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 northern 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 values are needed.

The data for such maps can be obtained from different sources. We use a macrobenthos database of the Estonian Marine Institute, University of Tartu (data between 2005–2019) and HERTTA database of the Finnish Environment Institute (data between 2000–2019) as data sources for Estonian and Finnish datasets, respectively. In Russia, targeted field works are carried out to get most of all necessary information on the natural environment of the eastern Gulf of Finland. In addition we use data from SPbRC RAS projects carried out between 2012–2014 including the TOPCONS project. Estonian dataset covers both flora and fauna observations while in Finland HERTTA data-base is used only for benthic invertebrate data, Russian data include benthic invertebrates (both benthic and pelagic phase of their life cycle), benthic flora and coastal fishes. An important harmonization goal of the Russian benthic datasets is to broaden the range of environmental niche covered as such aspect would tremendously improve the usage of the datasets in different spatial modelling activities.

Finnish macrophyte data are obtained from the Finnish Inventory Programme for the Underwater Marine Environment (VELMU) and the dataset consisted dive transect points recorded between the years 2004 and 2016. In addition to dive data, Finnish macrophyte models include also drop-video data of partly stratified randomized along environmental variable gradients and partly placed in grids.

Ekman and Van Veen type bottom grab samplers are used for benthic invertebrate samples on soft sediments. Scuba divers collect all the fauna and flora inside a Kautsky sampler that is a 0.04 m<sup>2</sup> metal frame. Benthic samples are sorted in a laboratory and all macrobenthic organisms are identified under a microscope. Abundances of all taxa are quantified. Sampling and analysis follows the guidelines developed for the HELCOM

COMBINE programme (<https://helcom.fi/media/publications/Manual-for-Marine-Monitoring-in-the-COMBINE-Programme-of-HELCOM.pdf>).

Currently, we have selected the following important benthic species or groups of species with different ecosystem functions and recovery potentials to represent benthic nature values: bladder wrack (*Fucus vesiculosus*), clawed fork weed (*Furcellaria lumbricalis*), filamentous algae, eelgrass (*Zostera marina*), vascular plants (excluding *Zostera marina*), charophytes (*Chara* spp., *Tolypella nidifica*, *Nitella* spp.), epibenthic bivalves (*Mytilus trossulus*, *Dreissena polymorpha*, *Mytilopsis leucophaeata*), epibenthic crustacean (*Amphibalanus improvisus*), infaunal bivalves (*Limecola balthica*, *Cerastoderma glaucum*, *Mya areanaria*), infaunal habitat engineering polychaete (*Marenzelleria* spp.), insect larvae. In addition, total species richness is calculated for each sampling station.

Species richness is referred to as the number of species (or in some cases, where species identification is not practical, higher taxa such as genera and families) in a given space, a sampling station in our case. In addition, presence and absence status of modelled species or species groups (presence if any of the included species is present) is defined for every observation point.

Bladder wrack (*Fucus vesiculosus*) is the main habitat-forming perennial macroalgal species in the Gulf of Finland on hard substrate dominated shallows at salinity above 3 PSU. The bladder wrack habitats host a number of seaweed, invertebrate and juvenile fish species (Råberg and Kautsky 2007, Wikström and Kautsky 2007, Kersen et al. 2011) with species richness often exceeding values of other benthic habitats (Kautsky et al. 1992). Bladder wrack grows on hard substrates in sublittoral from shallow water down to 4–5 meters depth, the depth distribution varying spatially and temporally depending on local environmental conditions. Bladder wrack populations have been declining in many areas (e.g. Vahteri and Vuorinen 2016) and the maximum growing depth has diminished remarkably (Torn et al. 2006). The reason for the distribution changes remains still somewhat unclear (see Vahteri and Vuorinen 2016 for discussion) even though eutrophication and restricted light conditions have been usually accounted for the population decline.

*Furcellaria lumbricalis* is a habitat-forming perennial red alga that grows on rocky substrates in the Baltic Sea. The species grows under *F. vesiculosus* canopy but forms a distinct red algal belt in deeper waters often associated with blue mussels (*Mytilus trossulus*). In some parts of the Estonian coast the species also colonizes sandy bottoms where its population is loose-lying (Martin et al. 2006). *Furcellaria lumbricalis* grows down to 15–20 meters in the Gulf of Finland.

The filamentous algal group consists of a number of macroalgal species. Many of them are annual and fast-growing species (Kiirikki and Lehvo 1997). Many species benefit from eutrophication and increase in biomass in eutrophicated conditions. Filamentous algae can also out-compete other macroalgal species (e.g. Thomsen et al. 2012), e.g., by inhibiting the establishment of germlings by covering all available space or by shadowing.

Epibenthic bivalves, blue mussel (*Mytilus trossulus*) and zebra mussel (*Dreissena polymorpha*) and dark false mussel (*Mytilopsis leucophaeata*), grow on hard substrates

along the Gulf of Finland. The distribution of blue mussel is mostly restricted to the western Gulf of Finland where salinity values are the highest. Along with a reduction of salinity towards the eastern Gulf of Finland the biomass and growth of the mussels are considerably reduced (Westerbom et al. 2002). The distribution of zebra mussel on the other hand is limited to the eastern parts of the Gulf of Finland as the species tolerates only very low salinities (Kotta et al. 1998, Antsulevich et al. 2003). The transition zone between the two species lies between the cities Porvoo and Kotka. The dark false mussel in its native area inhabits coastal marine waters and lower reaches of estuaries. However, in its invasion range it is registered until middle reaches of large rivers. In the Baltic Sea it co-occurs with zebra mussels (Darr and Zettler, 2000; Therriault et al. 2004; Verween et al. 2010). As the species originates from subtropical areas its range in the Baltic Sea was expected to be associated with the plume of heated waters of nuclear power plants in Loviisa (Finland) and Kopora Bay (Russia) (Laine, et al. 2006; Орлова и др., 2015). Recent records from Finland, Sweden and Poland in the 2010es indicate that the species is able to disperse into other areas not influenced with artificial heating (Forsström et al., 2016).

Epibenthic bivalves provide an important food source for many birds and fish species and they are efficient filter feeders capable of cleaning coastal ecosystems while feeding plankton. They also produce a three dimensional structure, an habitat for other invertebrates and thereby increase benthic biodiversity (see e.g. Koivisto and Westerbom 2010).

The only seagrass species inhabiting the northern Baltic Sea, eelgrass (*Zostera marina*), forms dense meadows on shallow sandy bottoms. As a marine species with a salinity optimum of >10 PSU (Nejrup and Pedersen 2008), the distribution of eelgrass is limited by low salinity to the western Gulf of Finland and Southwestern Finland. Within the Gulf of Finland, the species reproduces only vegetatively. Increasing nutrient loading is expected to reduce the depth distribution and areal coverage of the species (Boström et al. 2014).

The group of aquatic vascular plants (*Zostera marina* excluded) consists of a number of plant species, inhabiting the shallow coastal areas, generally down to a few meters' depth, due to high demand for light. Vascular plants in the Baltic Sea prefer soft or sandy bottoms in sheltered or moderately exposed areas. Species prefer different environmental conditions and submerged vascular plants can exist from river inlets containing fresh water to sheltered more saline bays. Common aquatic vascular plants in the study area are, for example, *Ceratophyllum demersum*, *Myriophyllum* spp., *Najas* spp., *Potamogeton* spp., *Ranunculus* spp., *Ruppia* spp., *Stuckenia* spp. and *Zannichellia* spp. and they provide reproduction and nursery areas for fishes and habitats for invertebrates.

Charophytes, morphologically complex green algae, grow on shallow soft or sandy bottoms in brackish water. They can form dense meadows on shallow sheltered bays and flads. Some species, like *Chara aspera*, also grow on moderately exposed sandy bottoms. The genera found within the project area include *Chara* spp., *Nitella* spp., *Nitellopsis* spp. and *Tolypella nidifica*.

Infaunal bivalves (*Limecola balthica*, *Cerastoderma glaucum* and *Mya arenaria*) live on sand-dominated bottoms from shallow to deep waters. They are either deposit or filter feeding species. The species' distribution is limited for example by suitable bottom substrate, wave energy on the bottom, food availability and oxygen conditions. Fresh water bivalves are common in the eastern Gulf of Finland where salinity is too low for the aforementioned marine bivalve species.

Infaunal polychaete worms from the genus *Marenzelleria* are euribiotic infaunal habitat engineering invertebrates impacting both sediments and water column due to high densities and their high bioturbation activity (Maximov et al. 2015). The species are found massively on silt-dominated bottoms from slopes to sedimentation areas, including hypoxic areas and areas with Fe-Mn concretions, either in natural and severely disturbed areas due to dredging and dumping. Besides the availability of food, salinity sets the distribution of the species.

Insect larvae, especially chironomids, are important components of food webs in the eastern Gulf of Finland where marine species are absent due to very low salinity.

A harmonized data structure was selected for exchanging and storing biological data from sampling sites. The table structure included general sampling information (sampling ID, sampling date, geographical coordinates, sampler, water depth) and seabed substrate information as separate columns and taxa in separate rows (Table 1).

Table 1. Table structure for storing biotic data from sampling points. The filled in values are dummy example data.

id	Station name	date	sampling gear	depth	latitude	longitude	mud	clay	...	species	variable	unit	value
1	A1	12.07.2019	video	1.7	58.0001	24.0001	0	0	40	<i>Fucus vesiculosus</i>	cover	%	30
1	A1	12.07.2019	video	1.7	58.0001	24.0001	0	0	40	<i>Pilayella littoralis</i>	cover	%	10
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Fucus vesiculosus</i>	biomass	g/m2	96.11
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Pilayella littoralis</i>	biomass	g/m2	12.11
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Gammarus salinus</i>	biomass	g/m2	0.25
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Gammarus zaddachi</i>	biomass	g/m2	0.12
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Jaera albifrons</i>	biomass	g/m2	0.05
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Gammarus salinus</i>	abundance	ind/m2	125
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Gammarus zaddachi</i>	abundance	ind/m2	50
2	B3	02.08.2019	frame	2	59.1111	23.1111	0	0	65	<i>Jaera albifrons</i>	abundance	ind/m2	225
3	C10	28.07.2018	Van Veen	42	58.55555	24.1111	10	5	0	<i>Macoma balthica</i>	biomass	g/m2	52.12
3	C10	28.07.2018	Van Veen	42	58.55555	24.1111	10	5	0	<i>Marenzelleria neglecta</i>	biomass	g/m2	2.11



3	C10	28.07.2018	Van Veen	42	58.55555	24.1111	10	5	0	<i>Manayunkia aestuarina</i>	biomass	g/m2	1.22
3	C10	28.07.2018	Van Veen	42	58.55555	24.1111	10	5	0	<i>Macoma balthica</i>	abundance	ind/m2	120
3	C10	28.07.2018	Van Veen	42	58.55555	24.1111	10	5	0	<i>Marenzelleria neglecta</i>	abundance	ind/m2	30
3	C10	28.07.2018	Van Veen	42	58.55555	24.1111	10	5	0	<i>Manayunkia aestuarina</i>	abundance	ind/m2	10

In order to harmonize the names of species and higher taxa and keep it consistent with the most recent taxonomy, an online Google Sheet taxonomy table was established (Figure 1). The taxonomy table was populated with taxon names from all datasets and the AphiaID and an accepted Latin name from the World Register of Marine Species (<http://www.marinespecies.org/>) was manually added to each taxon. Operational name was also added in order to group some species if needed. An R script was developed to automatically retrieve all taxonomic levels of each entry in the table based on the AphiaID.

	A	B	C	D	E	F	G
1	taxon_data_source	taxon_technical	taxon_normal	taxon_operational	kingdom	AphiaID	comment
6	<i>Aulodrilus limnobioides</i> Bretscher, 1899	aulodrilus_limnobioides_bretscher	<i>Aulodrilus limnobioides</i>	Oligochaeta	animal	1040644	
7	<i>Balanus improvisus</i> Darwin, 1854	balanus_improvisus_darwin_1854	<i>Amphibalanus improvisus</i>	<i>Amphibalanus improvisus</i>	animal	421139	
8	<i>Baltidrilus costatus</i>	baltidrilus_costatus	<i>Baltidrilus costatus</i>	Oligochaeta	animal	743898	
9	<i>Baltidrilus costatus</i> (Claparède, 1863)	baltidrilus_costatus_claparede	<i>Baltidrilus costatus</i>	Oligochaeta	animal	743898	
10	<i>Bivalvia</i> , n/det	bivalvia_n_det	<i>Bivalvia</i>	<i>Bivalvia</i>	animal	105	
11	<i>Bylgides sarsi</i>	bylgides_sarsi	<i>Bylgides sarsi</i>	<i>Bylgides sarsi</i>	animal	130735	
12	<i>Bythinia</i>	bythinia	<i>Bithynia</i>	<i>Bithynia</i> sp.	animal	182698	
13	<i>Ceraclea nigronervosa</i> (Retzius, 1783)	ceraclea_nigronervosa_retzius	<i>Ceraclea nigronervosa</i>	Trichoptera	animal	160252	Fresh-wat
14	Ceratopogonidae	ceratopogonidae	Ceratopogonidae	Ceratopogonidae	animal	150940	
15	<i>Cernosvitoviella</i> sp.	cernosvitoviella_sp	<i>Cernosvitoviella</i> sp.	Oligochaeta	animal	137346	
16	<i>Chaetogaster diaphanus</i>	chaetogaster_diaphanus	<i>Chaetogaster diaphanus</i>	Oligochaeta	animal	137478	
17	<i>Chaetogaster diaphanus</i> (Gruithuiser)	chaetogaster_diaphanus_gruithuiser	<i>Chaetogaster diaphanus</i>	Oligochaeta	animal	137478	
18	Charophyta	charophyta	Charophyta	Charophyta	plant	368664	
19	<i>Chelicorophium curvispinum</i>	chelicorophium_curvispinum	<i>Chelicorophium curvispinum</i>	<i>Chelicorophium curvispinum</i>	animal	148582	
20	<i>Chelicorophium curvispinum</i> (Sars, 1861)	chelicorophium_curvispinum_sars	<i>Chelicorophium curvispinum</i>	<i>Chelicorophium curvispinum</i>	animal	148582	
21	<i>Chelicorophium curvispinum</i>	chelicorophium_curvispinum	<i>Chelicorophium curvispinum</i>	<i>Chelicorophium curvispinum</i>	animal	148582	
22	Chironomidae	chironomidae	Chironomidae	Chironomidae	animal	118100	
23	Chironomidae gen. sp.	chironomidae_gen_sp	Chironomidae	Chironomidae	animal	118100	
24	<i>Chironomus</i> gr. plumosus (Linnaeus)	chironomus_gr_plumosus_linnaeus	<i>Chironomus</i> gr. plumosus	Chironomidae	animal	146802	
25	<i>Cladotanytarsus</i> gr. mancus (Walker)	cladotanytarsus_gr_mancus_walker	<i>Cladotanytarsus</i> gr. mancus	Chironomidae	animal	118100	
26	Coleoptera	coleoptera	Coleoptera	Coleoptera	animal	118085	
27	<i>Cordylophora caspia</i>	cordylophora_caspia	<i>Cordylophora caspia</i>	<i>Cordylophora caspia</i>	animal	117428	
28	<i>Cordylophora caspia</i> (Pallas, 1771)	cordylophora_caspia_pallas_1771	<i>Cordylophora caspia</i>	<i>Cordylophora caspia</i>	animal	117428	

Figure 1. A screenshot of the online taxonomy table.

## Abiotic environmental data

The key abiotic environmental variables (Table 2) include different bathymetrical (depth, slope of sea-bed, topographical position), hydrodynamic (wave exposure, currents), geological (seabed substrate), and physico-chemical (temperature, salinity, transparency, nutrients, ammonium, ice conditions) variables. Following environmental variables are planned to be used in the modelling. The resolution of the layers is set at 100 m where possible.

Table 2. Non-exclusive list of georeferenced environmental variables that are used in the modeling of nature assets in the programme area.

<b>Variable</b>
Water depth
Slope of seabed
Salinity
Wave exposure based on simplified wave model
Chlorophyll a content of sea surface based on satellite imagery
Ice coverage
Water temperature in cold season
Water temperature in warm season
Current velocity
Secchi depth
Concentration of oxygen on the bottom
Concentration of ammonium
Concentration of nitrates
Concentration of phosphates

## 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).



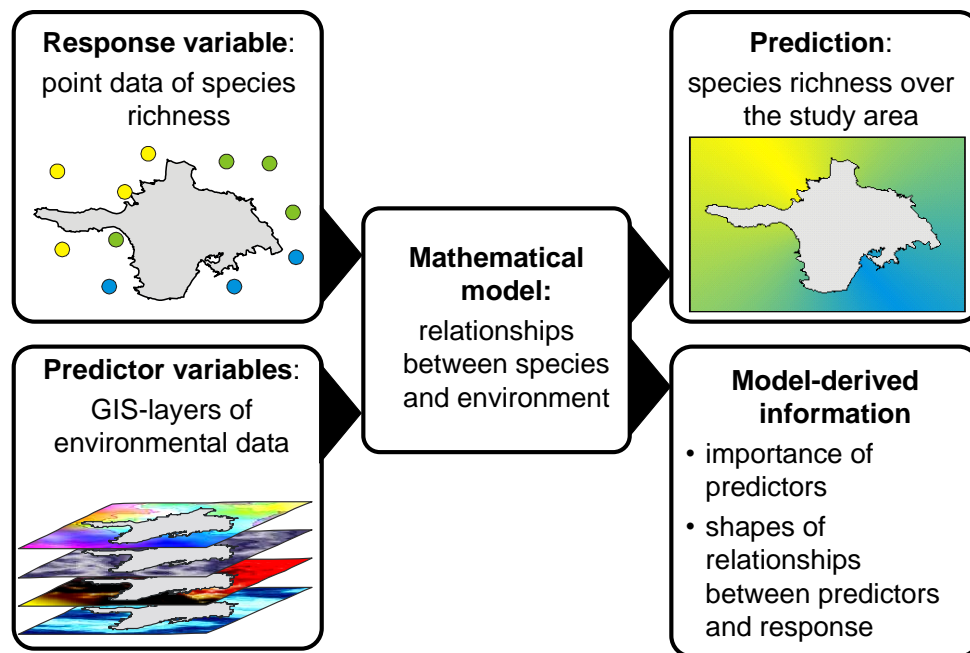


Figure 2. Conceptual scheme of spatial predictive modeling for deriving spatially continuous data from point-wise biological sampling data.

In the ADRIENNE project, the spatial distributions of benthic species, species groups and biodiversity variables are modeled. The key species and species groups, chosen as important ecosystem components, are *Fucus vesiculosus*, *Furcellaria lumbricalis*, filamentous algae, epibenthic bivalves (*Mytilus trossulus*, *Dreissena polymorpha*, *Mytilopsis leucophaeata*), epibenthic crustacean (*Amphibalanus improvisus*), vascular plants (excluding *Zostera marina*), *Zostera marina*, Charophytes (*Chara* spp., *Tolypella nidifica*), infaunal bivalves (e.g. *Limecola balthica*, *Cerastoderma glaucum*, *Mya arenaria*), infaunal polychaeta (*Marenzelleria* ssp.) insect larvae.

Several candidate models are built for each biodiversity variable using boosted regression trees (BRT) and/or random forest (RF) modeling methods. The candidate model with the best predictive performance was chosen to produce the final distribution maps. BRT is an ensemble method that combines the strength of two algorithms: regression trees and boosting (Elith et al. 2008). Regression trees are good at selecting relevant predictor variables and can model interactions. Boosting enables building of a large number of trees in a way that each successive tree adds small modifications in parts of the model space to fit the data better (Friedman et al. 2000). The algorithm keeps adding trees until finding the optimal number of trees that minimizes the predictive deviance of a model. The predictive performance of BRT has been shown to be superior to most other modeling methods (Elith et al. 2006; Reevermann et al. 2012). Important parameters in building BRT models are learning rate, tree complexity, and bag fraction (Elith et al. 2008). Learning rate determines the contribution of each tree to the growing model and tree complexity defines the depth of interactions allowed in a model. Bag fraction determines the proportion of data to be selected randomly at each iteration. Different combinations of these parameters may yield variable predictive performance but generally a lower learning rate and inclusion of interactions gives better results. For each group of species,

BRT models with optimal tree complexity are built. Modelling is done in the statistical software R 4 (R Core Team 2020) using the packages *gbm* (Ridgeway 2007) and *dismo* (Elith and Leathwick 2017) for BRT. Alternatively, random forest method can be applied using a package *randomForest* in R (Liaw and Wiener 2002).

RF is a machine learning method that generates a large number of regression trees, each calibrated on a bootstrap sample of the original data (Breiman et al., 2018). Each node is split using a subset of randomly selected predictors and the tree is grown to the largest possible extent without pruning. For predicting the value of a new data point, the data are run through each of the trees in the forest and each tree provides a value. The model prediction is then calculated as the average value over the predictions of all the trees in the forest (Breiman et al., 2018). Two parameters must be set in RF models: the number of predictor variables to be randomly selected at each node (*mtry*) and the number of trees in a forest (*ntree*). *mtry* was set to one third of the number of predictor variables as suggested by Liaw and Wiener (2002). *ntree* was set to 1000 as 500 trees usually yield stable results (Liaw and Wiener 2002).

The input data is randomly partitioned into calibration and validation datasets. The validation dataset contains data that was not included in model calibration. Calibrated models are used to predict the patterns of nature values with a grid size of 1000 meters covering the whole Gulf of Finland area. Raster layers of predictions are visually assessed to identify possible overfitting and other model- or data-driven artifacts that may not be directly reflected in mathematical validation. Based on both mathematical validation and visual expert assessment, the best performing modeling algorithm are selected. Importance of environmental predictor variables are assessed using percentage relative influence in BRT and RF.

Final modelling products include predicted values of key species, species groups and habitats per each grid cell as well as their uncertainties (standard error).

## Data management and software

Data management, GIS analysis and modelling is done using MS Excel, MS Access, ArcGIS and R. Modeling is done in the statistical software R 4 (R Core Team 2020): the package *gbm* (Ridgeway 2007) and *dismo* (Elith and Leathwick 2016) are used for BRT and *randomForest* (Breiman et al 2018) for RF.

During the Adrienne project modelling workshop through Skype (18.03.2020) it was decided and agreed that the easiest and most reasonable way to share data between the partners is to use Google Drive, Excel, and web-based mapping services. To ensure the harmonization of different datasets and simplify the data management and analysis, template of data format was also uploaded on Google Drive.

## Challenges and solutions related with data and modelling process

To get better overview what kind of biological and environmental data each partner country has, on what resolution, and what data is possible to add or get during the project a table that consisted such an information was circulated and filled among partner countries at the beginning of the project. Biggest challenges with data were related with the harmonizing process between different countries datasets.

During the project modelling workshop in March 2020 it was decided that cross-border modeling was not feasible at the current stage of data compilation as the number of biotic observations in the Russian waters were too low. Currently, the data compilation is in progress and more Russian macrobenthos data will be added soon when the results of extensive sampling of the summer 2020 campaign will become available.

## Storage of modelling products

The map layers of the spatial prediction of key species, species groups and habitats will be published as Web Map Services and publicly displayed in the GIS Assessment Portal (Adrienne section/PlanWise4Blue portal).

## References

- Antsulevich A, Välipakka P, Vaittinenc J (2003) How are the zebra mussels doing in the Gulf of Finland? Proc Estonian Acad Sci Biol Ecol 52:268-283.
- Boström C, Baden S, Bockelmann AC, Dromph K, Fredriksen S, Gustafsson C et al. (2014) Distribution, structure and function of Nordic eelgrass (*Zostera marina*) ecosystems: implications for coastal management and conservation. Aquatic Conserv: Mar Freshw Ecosyst 24:410-434.
- 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/>
- Darr A and Zettler ML (2000) Malakologishe abhandlungen Staatliches Museum fur Tiekunde Dresden. Band 20, Nr 19, Ausgegeben 5: 197-200.
- 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. Annu Rev Ecol Evol Syst 40:677-697.
- Elith J and Leathwick JR (2017) Boosted Regression Trees for ecological modeling. <https://cran.r-project.org/web/packages/dismo/vignettes/brt.pdf> Accessed 30 June 2019.
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151.
- Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77:802-813.
- Friedman JH, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting. Ann Stat 28:337-407.
- Forsström T, Fowler AE, Lindquist M, Vesakoski O (2016) The introduced dark false mussel, *Mytilopsis leucophaeata* (Conrad, 1831) has spread in the northern Baltic Sea. Bio Invasions Records. 5(2): 81-84.
- Guisan A and Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993-1009.
- Guisan A and Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147-186.

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

Kersen P, Kotta J, Bučas M, Kolesova N, & Değere Z (2011) Epiphytes and associated fauna on the brown alga *Fucus vesiculosus* in the Baltic and the North Seas in relation to different abiotic and biotic variables. Marine Ecology 32: 87-95.

Kautsky H, Kautsky L, Kautsky N, Kautsky U, Lindblad C (1992) Studies on the *Fucus vesiculosus* community in the Baltic Sea. Acta Phytogeogr Suec 78:33-48.

Kiirikki M and Lehvo A (1997) Life strategies of filamentous algae in the Northern Baltic Proper. Sarsia 82:259-267.

Koivisto ME and Westerbom M (2010) Habitat structure and complexity as determinants of biodiversity in blue mussel beds on sublittoral rocky shores. Mar Biol 157:1463 -1474.

Kotta J, Futter M, Kaasik A, Liversage K, Rätsep M, Barboza FR, Bergström L, et al (2020) Cleaning up seas using blue growth initiatives: Mussel farming for eutrophication control in the Baltic Sea. Science of the Total Environment 709: 136144.

Kotta J, Orav H, Kotta I (1998) Distribution and filtration activity of the zebra mussel, *Dreissena polymorpha*, in the Gulf of Riga and the Gulf of Finland. Proceedings of the Estonian Academy of Sciences, Biology Ecology 47(1): 32-41.

Kovtun A, Torn K, Martin G, Kullas T, Kotta J, Suursaar Ü (2011) Influence of abiotic environmental conditions on spatial distribution of charophytes in the coastal waters of West Estonian Archipelago, Baltic Sea. Journal of Coastal Research SI64: 412-416.

Laine AO, Mattila J, Lehtikoinen A (2006) First record of the brackish water dreissenid bivalve *Mytilopsis leucophaeata* in the Northern Baltic Sea. Aquatic Invasions. V. 1(1): 38-41.

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

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R. News 2/3, 18e22.

Martin G, Paalme T, Torn K (2006) Growth and production rates of loose-lying and attached forms of the red algae *Furcellaria lumbricalis* and *Coccotylus truncatus* in Kassari Bay, the West Estonian Archipelago Sea. Hydrobiologia 554:107-115.

Maximov AA, Bonsdorff E, Eremina T, Kauppi L, Norkko A, Norkko J (2015) Context-dependent consequences of *Marenzelleria* spp. (Spionidae: Polychaeta) invasion for nutrient cycling in the Northern Baltic Sea. Oceanologia 57: 341-348.

- Möller T and Martin G (2007) Distribution of the eelgrass *Zostera marina* L. in the coastal waters of Estonia, NE Baltic Sea. *Proceedings of the Estonian Academy of Sciences* 56(4): 270-277.
- Nejrup L and Pedersen M (2008) Effects of salinity and water temperature on the ecological performance of *Zostera marina*. *Aquat Bot* 88:239-246.
- R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Råberg S and Kautsky L (2007) A comparative biodiversity study of the associated fauna of perennial fucoids and filamentous algae. *Estuar Coast Shelf S* 73:249-258.
- Revermann R, Schmid H, Zbinden N, Spaar R, Schröder B (2012) Habitat at the mountain tops: how long can Rock Ptarmigan (*Lagopus muta helvetica*) survive rapid climate change in the Swiss Alps? A multi-scale approach. *J Ornithol* 153:891-905.
- Ridgeway G (2007) Generalized Boosted Models: A guide to the gbm package. <http://www.saedsayad.com/docs/gbm2.pdf> Accessed 30 June 2019.
- Therriault TW, Docker MF, Orlova MI, Heath DD, MacIsaac HJ (2004) Molecular resolution of Dreissenidae (Mollusca: Bivalvia) including the first report of *Mytilopsis leucophaeata* in the Black Sea basin. *Mol. Phyl. Evol.* № 30: 479-489.
- Thomsen MS, Wernberg T, Engelen AH, Tuya F, Vanderkift MA, Holmer M, McGlathery KJ, Arenas F, Kotta J, Silliman BR (2012) A meta-analysis of seaweed impacts on seagrasses: generalities and knowledge gaps. *PloS one* 7(1):e28595.
- Torn K, Krause-Jensen D, Martin G (2006) Present and past depth distribution of bladderwrack (*Fucus vesiculosus*) in the Baltic Sea. *Aquat Bot* 84:53-62.
- Vahteri P and Vuorinen I (2016) Continued decline of the bladderwrack, *Fucus vesiculosus*, in the Archipelago Sea, northern Baltic proper. *Boreal Environ Res* 21:373-386.
- Verween A, Vinex M, Degraer S (2010) *Mytilopsis leucophaeata*: The brackish water equivalent of *Dreissena polymorpha*? A review. In G. van der Velde, S Rajagopal and A. bij de Vaate (eds.) *The Zebra Mussel in Europe*. Leiden, The Netherlands: Backhuys. P. 29-43.
- Westerbom M, Kilpi M, Mustonen O (2002) Blue mussels, *Mytilus edulis*, at the edge of the range: population structure, growth and biomass along a salinity gradient in the north-eastern Baltic Sea. *Mar Biol* 140:991-999.
- Wikström SA and Kautsky L (2007) Structure and diversity of invertebrate communities in the presence and absence of canopy-forming *Fucus vesiculosus* in the Baltic Sea. *Estuar Coast Shelf S* 72:168-176.