

MAPPING BENTHIC BIODIVERSITY USING PREDICTIVE MODELING



Anneliis Peterson and Kristjan Herkül
Estonian Marine Institute, University of Tartu
anneliispeterson@ut.ee



INTRODUCTION

- Biodiversity is critical for maintaining and stabilizing ecosystem processes in changing environments.
- There is a need for high resolution biodiversity maps that cover large sea areas in order to address ecological questions related to biodiversity-ecosystem functioning relationships and also to provide data for marine environmental protection and management decisions.
- Traditional sampling-point-wise field work is not suitable for covering extensive areas in high detail.
- Spatial predictive modeling using biodiversity data from sampling points and georeferenced environmental data layers covering the whole study area is a potential way to create biodiversity maps for large spatial extents
- In this study we aimed at producing benthic biodiversity maps over a relatively large marine area, the whole Estonian sea area, by using a multitude of georeferenced environmental layers and three different modeling algorithms.

MATERIAL & METHODS

- Study area: Estonian marine area, northern Baltic Sea (Fig. 1).
- Biological sampling:
 - biomass samples: 3 474 sites in 2005-2015 (Fig. 1).
 - on soft sediments: bottom grab samplers.
 - on hard surfaces: scuba divers collected all the fauna and flora inside the frame.
 - three macrobenthos biodiversity variables were calculated for each sampling station:
- 1. total species richness
- 2. zoobenthos species richness
- 3. phytobenthos species richness.
- 18 environmental variables (Table 1; all available as raster layers in ESRI ArcGIS file geodatabase).
- Three modeling algorithms were used: random forest (RF), generalized additive model (GAM) and boosted regression trees (BRT).
 - modeling was done in the statistical software R 3.0.3
 - spatial predictions were produced with 100 m grid size covering the whole Estonian sea area

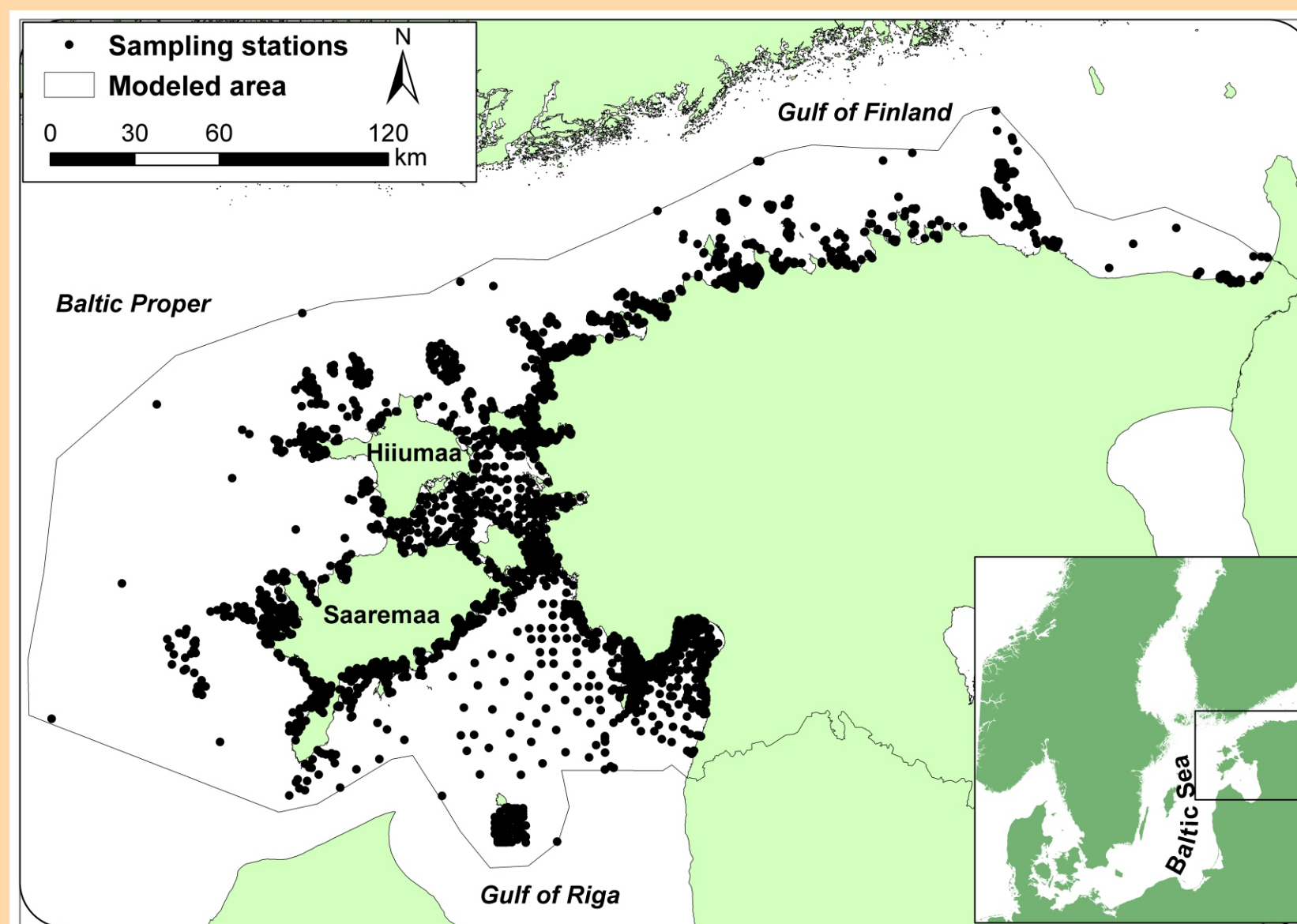


Figure 1. Study area and sampling sites

Table 1. Georeferenced environmental variables that were used in biodiversity modeling.

Variable	Abbreviation
Water depth	depth
Average water depth in 2000 m radius	depth2
Slope of seabed	slope
Slope of seabed in 2000 m radius	slope2
Salinity	salinity
Wave exposure based on simplified wave model	wave
Chlorophyll a (satellite imagery)	chl
Water transparency (attenuation coefficient; satellite imagery)	attenuation
Ice coverage	ice
Water temperature in cold season	tempcold
Water temperature in warm season	tempwarm
Current velocity	current
Orbital speed of water movement at seabed induced by wind waves	orbspeed
Proportion of soft sediment	softsed
Secchi depth	secchi
Concentration of ammonium	ammonium
Concentration of nitrates	nitrate
Concentration of phosphates	phosphate

RESULTS

- All three tested modeling algorithms proved to have high predictive ability.
- RF and BRT predictions had higher correlations with validation data and lower mean absolute errors than those of GAM (Table 2).

Table 2. Model validation statistics (Pearson r and mean absolute error (MAE) of the candidate models. All correlations were statistically significant ($p < 0.05$).

Biodiversity variable	RF		GAM		BRT	
	r	MAE	r	MAE	r	MAE
Total S	0.74	4.08	0.65	4.85	0.73	4.21
Flora S	0.75	1.72	0.68	1.99	0.74	1.77
Fauna S	0.68	2.87	0.58	3.28	0.67	2.94

- Depth and sediments were the most influential environmental variables in all three mathematical models (Fig. 2).

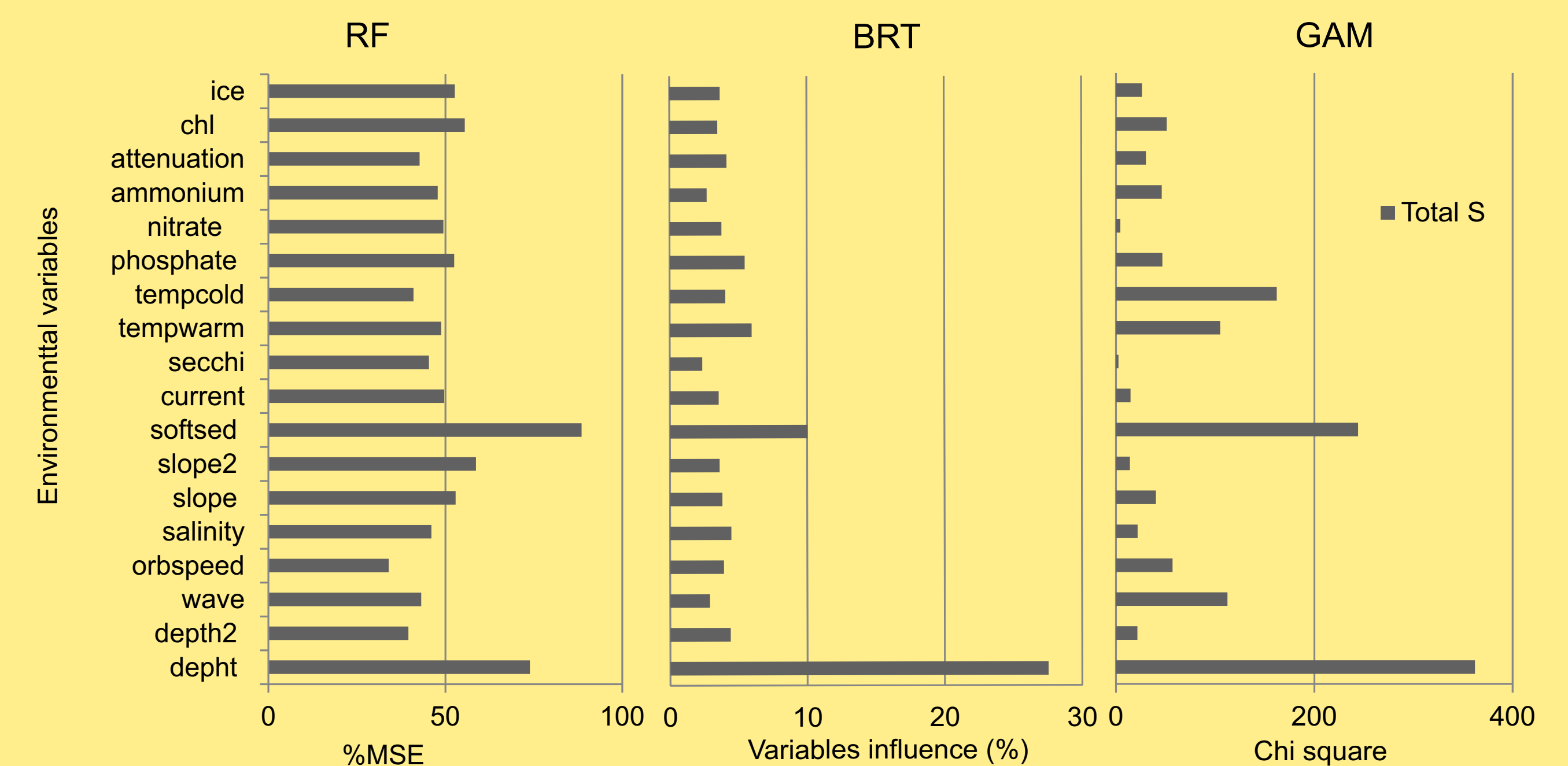


Figure 2. Importance of environmental variables in predicting total species richness in RF, BRT and GAM models. Higher value indicates higher importance.

- Only some insignificant modeling artifacts could be detected in few areas when zoomed in.
- Generally, RF predictions had more signs of overfitting than those of BRT and GAM.

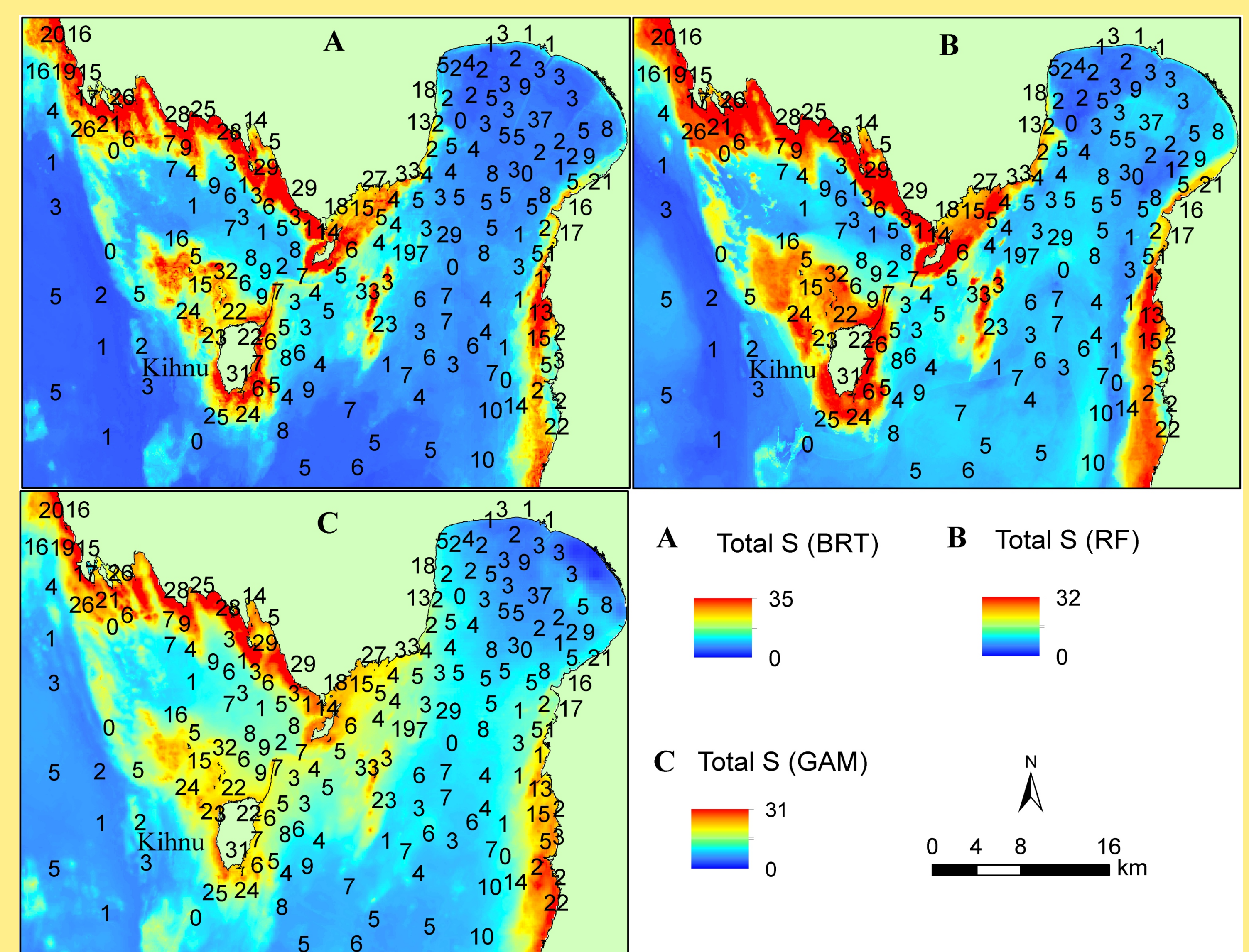


Figure 3. Total species richness in the Pärnu Bay, Gulf of Riga, predicted by BRT, RF, and GAM models. Numbers represent the species richness in the sampling stations.

- Predictions of all models were highly adequate with good agreement between predicted values and observed values and without severe artifacts (Fig. 4).
- Based on mathematical and visual validation, BRT was considered to be the best algorithm in this study.
- Large areas of high benthic biodiversity were located in western Estonia where shallow marine areas prevail.
- This is a first study that accomplishes to deliver high resolution seamless benthic biodiversity maps over a large spatial extent in the northern Baltic Sea.

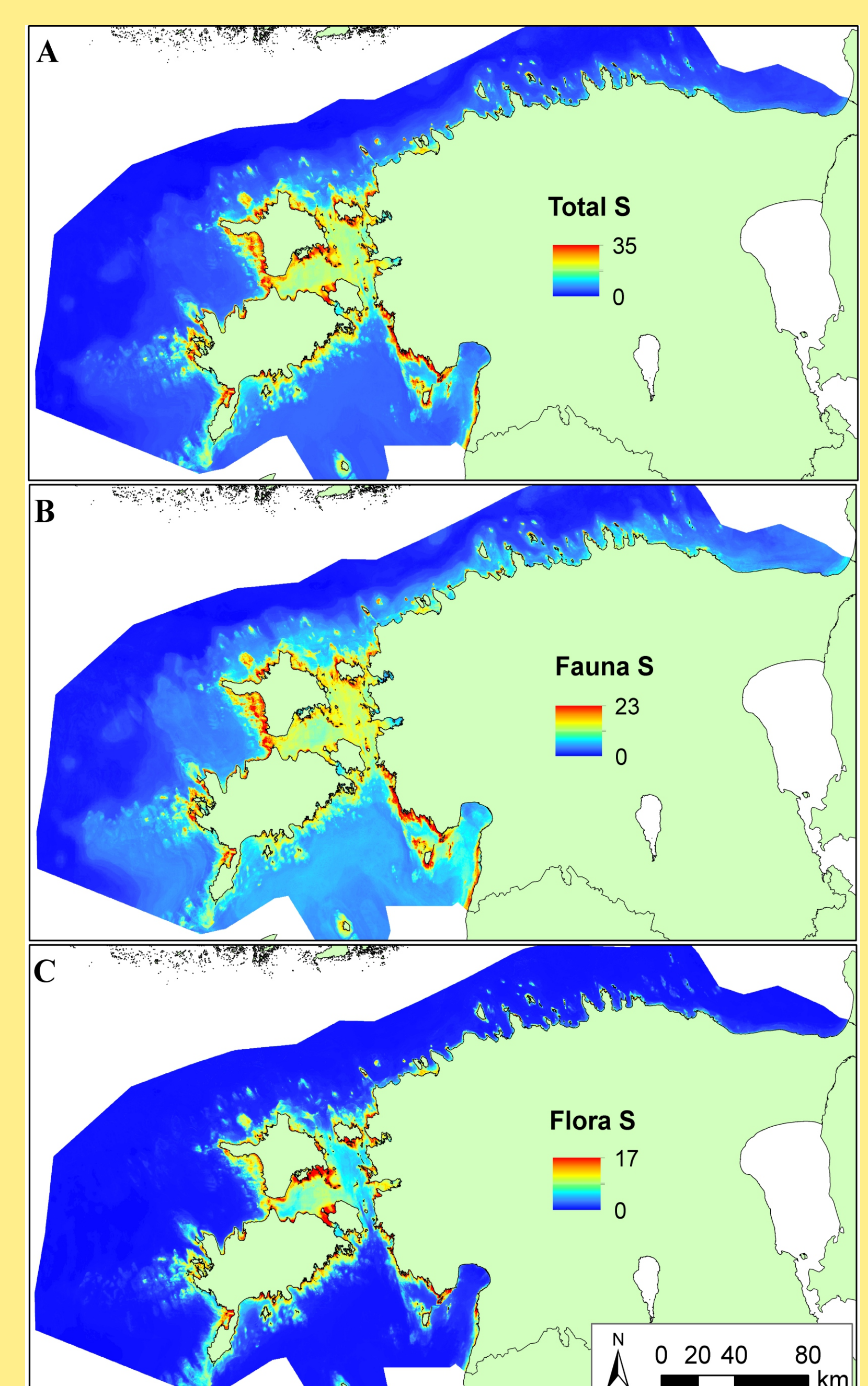


Figure 4. Predictive models of BRT for A: total, B: fauna and C: flora species

ACKNOWLEDGEMENTS

This work was supported by institutional research funding IUT02-20 of the Estonian Ministry of Education and Research.