Blue mussels *Mytilus edulis* L. and *M. trossulus* Gould in sympatry: assessment of ecological niche divergence using species distribution modeling

Running page head: ~~***"Blue mussels niche divergence"***~~ ***"Mussel species niche divergence"***

V. Khaitov* [1,2], P. Safonov [3], A. Zaichikova [4], M. Katolikova [1], M. Ivanov [1], P. Strelkov [1,5]

[1] St. Petersburg State University, University embankment 7–9, 199034 St. Petersburg, Russia

[2] Kandalaksha State Nature Reserve, Lineynaya 35, 184042 Kandalaksha, Russia

[3] Institute of Cytology RAS, Tikhoretsky 4, 194064 St. Petersburg, Russia

[4] Lomonosov Moscow State University, Leninskie gory 1c12, 119234 Moscow, Russia

[5] Laboratory of Monitoring and Conservation of Natural Arctic Ecosystems, Murmansk Arctic University, Kapitana Egorova 15, 183038 Murmansk, Russia

* Corresponding author

Corresponding author's e-mail: polydora@rambler.ru

Abstract

Species distribution models (SDMs) describing the relationship between species occurrence and environmental parameters can be used to assess the ecological niche of a species. Usually applied to morphologically distinct species, SDMs are also a promising tool for describing niche partitioning in coexisting cryptic species. An example of the latter in the marine realm are blue mussels *Mytilus edulis* (*ME*) and *M. trossulus* (*MT*). Despite considerable research effort, little is known about how they share space and resources in sympatry anywhere except the Baltic Sea. Salinity, substrate, surf and proximity to harbors have been suggested as candidate factors of segregation but no ~~conclusion~~ general consensus has been ~~made~~ reached. Here we assessed partial effects of these predictors on divergence of *ME* and *MT* in the White Sea littoral applying SDMs to 570 mussel samples with known taxonomic structure. We found that each of the predictors influenced spatial segregation. The most expected habitat of *ME* was a bottom substrate in a wind-exposed location with ~~a “normal” salinity (24 ppt)~~ an average White Sea salinity (24 psu) away from ports and large rivers, while for *MT* it was an algal substrate in a wind-protected area with a lower salinity close to ports and large rivers. We also ~~addressed the question whether~~ found that the species segregation by substrate was density-dependent: ~~and found that~~ the degree of segregation positively depended on *ME* abundance, which indicates that *ME* outcompetes *MT* on bottom substrates. ~~We discuss whether the predictors used in our study can drive the segregation of these species outside the White Sea.~~ We suggest that the same predictors can drive the segregation of these two species outside the White Sea.

Key words: *Mytilus*; cryptic species; species distribution models; ecological niche divergence

1. INTRODUCTION

Species distribution models (SDMs) are a numerical tool describing the relationship between species occurrence and environmental parameters. ~~Using SDMs, it is possible~~They can be used to predict distribution patterns of species in space and time and to assess their ecological niche ~~in a formal way~~ (Elith & Leathwick 2009). Joint application of SDMs to ~~several~~ coexisting species, i.e. a community, ~~allows one~~ makes it possible to describe the partitioning of their ecological niches ~~between them~~ ~~This is referred to as Joint Species Distribution Modeling (JSDM)~~ (Ovaskainen & Abrego 2020). In other words, ~~SDM/JSDMs~~ SDMs may describe the axes in ecological space along which coexisting species are segregated. SDMs can be built using various approaches, from regular multiple regressions to advanced machine learning (Elith et al. 2006, Caradima et al. 2019, Poggiato et al. 2021).

~~SDMs are usually applied to "good", i.e. morphologically distinct species (e.g. Reiss et al. 2011, Lindegren et al. 2022), which can be easily involved in routine studies requiring numerous samples.~~ SDMs are usually applied to morphologically distinct species (e.g. Reiss et al. 2011, Lindegren et al. 2022), which are distinguished in routine biodiversity assessment studies. However, there is increasing evidence about coexistence of cryptic species (Bickford et al. 2007, Geller et al. 2010, Struck et al. 2018) and infraspecific taxa (Dufresnes et al. 2023). Cryptic invasions usually stand behind sympatry (Morais & Reichard 2017). ~~It is unlikely that any c~~Coexisting taxa are unlikely to have identical ecological phenotypes, and an ecological niche partitioning between them can be expected (Sáez & Lozano 2005). The question how such taxa share space and resources in sympatry can be answered using ~~SDM/JSDM~~ SDMs (Peterson et al. 2019). ~~Strictly speaking, when SDMs are applied to coexisting cryptic taxa, the latter are considered as a community.~~ ~~In marine ecology, t~~This approach has already been successfully used in marine ecology (e.g. Dennis & Hellberg 2010, Lowen et al. 2019, Hu et al. 2021). The results indicate that cryptic taxa indeed have distinct ecological phenotypes. Therefore, ecological niche assessment of coexisting cryptic species, particularly those of economic, conservation or ecosystem importance, is a promising research direction.

Blue mussels (*Mytilus edulis* species complex) are ~~T~~the longest-known and best-studied cryptic species in the marine realm ~~are those of the blue mussel (*Mytilus edulis*) complex~~ (Knowlton 1993, Gosling 2021). ~~The mussels are powerful ecosystem engineers in temperate and subpolar seas; they play a major role in coastal communities and are important aquaculture objects (Buschbaum et al. 2009, Gosling 2021).~~ ~~The complex is represented by several six recognized species that~~ Six species that make up this complex ~~are easier to distinguish genetically than morphologically and that~~ hybridize in sympatry and are easier to distinguish genetically than morphologically (~~Koehn 1991,~~ Wenne et al. 2020, Gardner et al. 2021). Blue mussels are powerful ecosystem engineers and important aquaculture objects (Buschbaum et al. 2009, Gosling 2021).

~~In the North Atlantic, t~~The dominant species of blue mussels in the North Atlantic are *M. edulis* (*ME*) and *M. trossulus* (*MT*). ~~On an oceanic scale, t~~Their distribution on the oceanic scale is ~~thought to be~~mostly regulated ~~primarily~~ by temperature and its correlates such as sea ice extent and primary production (Hayhurst & Rawson 2009, Wenne et al. 2020). Both these species occur in the Arctic but *ME* is distributed further south than *MT* in temperate seas ~~does not penetrate as far south into temperate seas as *ME*, appearing to be a more stenothermic, cold-loving species~~ (Wenne et al. 2020). ~~These two species form~~ There are multiple zones of sympatry between *ME* and *MT* (hereafter, contact zones), from Scotland and the Gulf of Maine in the south to Greenland and Spitsbergen in the north (Wenne et al. 2020). *ME* and *MT* ~~are fairly old species dating back to the Pliocene. They~~ evolved in allopatry since Pliocene in the Atlantic and the Pacific Ocean, respectively, and their contact zones are thought to have formed as a result of repeated *MT* invasions from the Pacific Ocean to the Atlantic as well as from one part of the Atlantic into another (Väinölä & Strelkov 2011, Wenne et al. 2020 and references therein).

In contact zones, *ME*, *MT* and their hybrids are often found ~~from~~ in the same ~~samples of mussel~~ settlements (Väinölä & Strelkov 2011, Wenne et al. 2020)~~; such settlements~~, which are ~~hereafter~~ referred to as ~~“~~mixed~~”~~ settlements. Scientists generally agree that sympatric *ME* and *MT* are ecologically distinct ~~in sympatry~~ (Riginos & Cunningham 2005, Katolikova et al. 2016, Michalek et al. 2021) and have a different economic value in aquaculture (Penney et al. 2002, Beaumont et al. 2008), but the data on the factors of their ecological segregation are fragmentary and contradictory.

The greatest progress in comparative ecological studies of *ME* and *MT* in sympatry has been made in the contact zones in the Baltic Sea, in the waters of the Kola Peninsula (White and Barents Seas) and in the West Atlantic (mainly, Gulf of Maine and New Scotland). In the Baltic Sea, the brackish areas of its inner part are inhabited by *MT*, while the saltier areas closer to the North Sea are inhabited by *ME*. In the middle there is the contact zone, where mixed settlements ~~could~~ can be dominated by hybrids, with *MT* gene frequency gradually increasing towards the inner Baltic (Väinölä & Strelkov 2011, Zbawicka et al. 2014, Stuckas et al. 2017). As a result, the species distribution is strongly correlated with salinity, ~~while~~ and the role of other factors is~~, in comparison,~~ negligible (Kijewski et al. 2019).

~~Another situation is observed i~~In the contact zones of the Kola Peninsula and the West Atlantic~~. H~~there are few hybrids ~~are always in the minority~~ in mixed settlements~~ there~~, and the spatial distribution of *ME* and *MT* is ~~mosaic~~ patchy ~~both at the regional (i.e. dozens to hundreds of kilometers) and at the local scale~~ ~~at~~ on different scales, from dozens of kilometers to tens of centimeters. The relationship between the distribution ~~of these species~~ and salinity is not obvious ~~anywhere~~ in these contact zones, and ~~it seems that~~ there ~~is~~ seems to be no simple “single-factor” pattern of species distribution (Riginos & Cunningham

2005, Katolikova et al. 2016, Wenne et al. 2020, Marchenko et al. 2023), ~~and several other factors of ecological segregation have been proposed.~~ Depth, fouling substrate, anthropogenic pollution levels and surf effects have been considered, apart from salinity, as possible factors affecting the segregation of *ME* and *MT* (Bates & Innes 1995, Comesaña et al. 1999, Hellou & Law 2003, Tam & Scrosati 2014, Marchenko et al. 2023) but no consensus has been reached.

In particular, in ~~In~~ the White and the Barents Sea, the frequency of *MT* is greater in port areas, possibly because this species has been introduced into the region with ship traffic in historic times (Väinölä & Strelkov 2011, Katolikova et al. 2016). The only segregation factor explicitly tested in the White Sea is the substrate ~~to which~~ of littoral mussels ~~attach~~ (Katolikova et al. 2016). ~~It has been shown that~~ *MT* is more common on fucoid algae while *ME* mostly lives directly on the bottom~~, on~~ substrates such as mud, sand, stones and gravel. However, segregation across substrates cannot fully explain the local-scale ~~mosaic~~ patchiness ~~in the distribution~~ (Katolikova et al. 2016). In the Barents Sea, no correlation with substrate has been found. However, these species have different depth preferences there~~. The proportion of *ME* increases with depth on littoral-sublittoral vertical transects, so that~~: *ME* appears to be a more sublittoral species and *MT* a more littoral one (Marchenko et al. 2023)~~. In West Atlantic, depth, anthropogenic pollution levels and surf effects have been considered as possible factors affecting the segregation of *ME* and *MT* (Bates & Innes 1995, Comesaña et al. 1999, Hellou & Law 2003, Tam & Scrosati 2014), but no definite conclusions have been made (Riginos & Cunningham 2005, Katolikova et al. 2016).~~

~~To sum up, no simple "single-factor" pattern of species distribution has been revealed in the contact zones of *MT* and *ME* outside the Baltic. Moreover, some of the factors~~ It should be noted that some of the candidate factors of species segregation may ~~potentially~~ be collinear and confound the analysis. Ports are often located in ~~storm-protected~~ sheltered areas~~, usually~~ close to river mouths, ~~so that~~ and the effects of shipping ~~(and other anthropogenic factors)~~, surf and salinity are difficult to distinguish. The effects of depth and substrate may obscure each other since fucoids, common in the littoral, are rare in the sublittoral, where they are replaced by kelps (Druehl & Green 1982).

~~This lack of conclusive evidence is partly due to the fact that until recently scientists could identify cryptic species of blue mussels only with the help of labor-intensive genotyping methods and therefore could not handle large amounts of material (Khaitov et al. 2021). In addition, there were no reliable statistical methods for modeling the distribution of sympatric taxa in the space of multiple factors, i.e. an SDM approach could not be implemented.~~ Comparative ecological studies of *ME* and *MT*, which began as early as the 1980s (see Riginos & Cunningham 2005 for review), have been hampered by two circumstances. Firstly, it was impossible to examine large amounts of material because species identification required labor-intensive genotyping methods (Khaitov et al., 2021). Secondly, until recently there were no reliable statistical methods for modeling the distribution of sympatric taxa in the space of multiple factors and the SDM approach could not be implemented. To our knowledge, ~~this approach~~ it has been applied to *ME* and *MT* only twice, in the above-mentioned studies by Kijewski et al. (2019) and ~~by~~ Wenne et al. (2020). In both studies the machine learning techniques were used to model the macro-geographic distribution of species (technically, of allele frequencies at taxonomically informative genes) in the space of multiple climatic and oceanographic characteristics available from public databases. ~~The authors concluded that temperature and salinity were important factors influencing the geographical distribution of these two species, with *MT* tolerating lower salinities and temperatures than *ME* (Kijewski et al. 2019, Wenne et al. 2020, see also above)~~ These methods have never been applied to modeling species distributions within contact zones.

In our previous studies we found a simple semi-diagnostic trait for *ME* and *MT*~~, namely, the presence or absence of an uninterrupted strip of prismatic layer under the ligament on the inner side of the shell (Zolotarev & Shurova 1997,~~ (Katolikova et al. 2016). Using this trait, one could make reliable interpretations of the taxonomic structure of mixed settlements on the basis of morphotype frequencies in samples, i.e. without genotyping. This procedure was referred to as the “morphotype test” (Khaitov et al. 2021). ~~In the White Sea 74% of *MT* but only 4% of *ME* have the strip (Katolikova et al. 2016), and the proportion of *MT* in samples (thereafter *Ptros*) is linearly dependent on the ratio of morphotypes (Khaitov et al. 2021). To note, hybrids are not considered as a separate category within this approach.~~

The aim of this study was to estimate the divergence of ecological niches between *ME* and *MT* in the White Sea littoral along environmental gradients such as substrate, salinity, surf level, and distance from ports. All these factors have been suggested as potentially influencing segregation of these two species in sympatry. Another candidate factor, depth (Marchenko et al. 2023), was not examined in our study but was controlled by sampling at the same littoral level. To achieve our aim, we examined the variability of the environmental predictors ~~mentioned above~~ and the taxonomic structure of mussel settlements using an extensive material (95 study sites, 570 mussel samples, 55,529 mussels) and assessed the partial influence of the predictors on the distribution of proportion of *MT* using SDMs. Since all predictors were included in one model, collinearity could be controlled. ~~Ideally, a model trained on reliable data should be able to predict the proportion of *MT* in mixed settlements (*Ptros*) in independent data, and we evaluated its predictive power using testing datasets from the White and the Barents Sea.~~ Ideally, a model trained on reliable data should be transferable and work well on independent data, therefore we evaluated the predictive power of our model using testing datasets from the White and the Barents Sea. In addition, to reveal a possible competition between the two species, we checked whether the pattern of ~~species~~ their segregation by substrate was density-dependent~~, i.e. whether there was competition between the two species~~.

2. MATERIALS & METHODS

2.1 Study area

The study area was the Kandalaksha Bay, where all previous *ME* and *MT* studies in the White Sea have been conducted (Katolikova et al. 2016, Khaitov et al. 2018, Khaitov et al. 2023). The Bay, 185 km long, is funnel-shaped, with ~~numerous islands and skerries and~~ a highly indented coastline and numerous islands and skerries (Fig. 1). Climate is continental subarctic with 4-5 months of ice cover and the average monthly sea surface temperature in August of 13.8°C. Mean tidal range is about 2 m. Summer surface salinity is 24 ~~ppt~~ psu in most of the Bay (~~“normal”~~ average salinity for ~~most of~~ the White Sea) and lower in the estuarine areas (Berger & Naumov 2000). Two canals of a hydropower plant and 24 rivers with a catchment area of 141 – 12,830 km^2 (Median 240 km^2; see **Table S1 in Supplement ~~2~~1**) flow into the Bay, with the largest river, the Niva, entering the Bay at its very top. Due to the ~~complex geometry of the~~ indented shoreline and numerous rivers, local surf and salinity gradients are pronounced (Filatov et al. 2007).

Six ports operating oceanic vessels were functioning in the area in the 20th century (**Fig. 1**). Two of them, both at the Bay’s top, are still in operation. The other four have been abandoned (Sailing directions of the White Sea 1932, Krasavtsev 2011) but are occasionally visited by small ships (our observations).

~~Mussels are present everywhere in the waters of the Bay. They are particularly abundant in the littoral fucoid belt (mainly *Fucus vesiculosus* L. and *Ascophyllum nodosum* L.), which is continuous 0.5-1.0 m above mean spring tide depth (Berger et al. 2001). According to the data from~~ In 2002–2013, both ~~mussel~~

~~species~~ *ME* and *MT* were almost ubiquitous in the Bay, but their ratio in settlements varied greatly, with *ME* being generally dominant (Katolikova et al. 2016). Mussels in the Bay are particularly abundant in the littoral fucoid belt (mainly *Fucus vesiculosus* L. and *Ascophyllum nodosum* L.), which is continuous 0.5-1.0 m above mean spring tide depth (Berger et al. 2001).

2.2 Modeling data set

2.2.1 Mussel sampling and processing

Mussels were sampled at 95 sites within the littoral fucoid belt in summer months of 2011-2018 (**Fig. 1**). Data for 17 of these sites were included in the study by Katolikova et al. (2016), the other data are new (**Table S2 in Supplement ~~2~~1**). ~~S~~The sites were chosen to describe the littoral populations of the Bay in as much detail as possible and to account for the heterogeneity of their habitat by substrate type, surf level, and distance from rivers and ports. All samples were taken within the fucoid belt ~~to minimize differences in depth~~. At each site, three samples from fucoid thalli (hereinafter, algal samples) and three samples from bottom substrates (bottom samples) were collected a few meters from each other using 0.25 m^2 and 0.025 m^2 frames, respectively. A greater size of the algal frame was associated with the large size of the fucoids and the need to account for their complex geometry. The frames were placed not randomly ~~but in such~~ but approximately at the same depth and in such a way as to capture the dense mussel aggregations.

We used mussels with a shell length larger than 10 mm ~~to identify~~ for a reliable identification of the shell morphotypes ~~reliably~~ (Khaitov et al. 2021). In the bottom samples all mussels from a frame were used. In the algal samples the procedure was different. One bundle of algae, containing at least a few dozen mussels, was chosen and weighed together with the attached mussels. The rest of the algae from a frame were weighted too. Mussels from the bundle were counted and used for further analysis. The ratio between the counted number of mussels and the bundle weight was applied to the total algal weight to reconstruct the total number of mussels in the sample (**Table S3 in Supplement ~~2~~1**). ~~For 12 sites t~~The information on the total number of mussels in algal samples was lacking for 12 of the sites, and they were excluded from the analyses which required data on mussel abundance ~~(**Model 2**, see below)~~(**Model 2** and **Model 3** below).

~~Shell morphotypes (E-morphotype, characteristic of *ME*, and T-morphotype, characteristic of *MT*) were identified for all selected mussels as in Khaitov et al. (2021).~~ Mussel shell morphotypes were identified for all selected mussels as in Khaitov et al. (2021) based on the presence (T-morphotype) or absence (E-morphotype) of an uninterrupted strip of prismatic layer under the ligament on the inner side of the shell. T-morphotype is characteristic of *MT*, and E-morphotype is characteristic of *ME*. ~~Further, t~~The proportion of morphotypes was converted to the proportion of *MT* (*Ptros*) in each sample, in pooled samples from each substrate from each site ($Ptros_{Algae}$ and $Ptros_{Bottom}$) and in pooled samples from each site ($Ptros_{Site}$), using equation

$$Ptros = \frac{e^{-2.4+5.4PT}}{1 + e^{-2.4+5.4PT}}$$

where *PT* ~~–~~ is a proportion of T-morphotype.

~~This equation, derived from the 24 genotyped samples (in total, 1105 multilocus mussel genotypes) from the Kandalaksha Bay, reliably predicts *Ptros* over the entire salinity range in the White Sea (i.e., up to 24 ppt), but may overestimate *Ptros* at higher salinities, as observed in the Barents Sea (Khaitov et al. 2021).~~

This equation, derived from the 24 genotyped samples (in total, 1105 multilocus mussel genotypes) from the Kandalaksha Bay, accounts for the nearly linear dependence of *Ptros* on *PT* and reliably predicts *Ptros* over the entire salinity range in the White Sea, i.e., up to 24 psu (Khaitov et al. 2021). However, as studies in the Barents Sea have shown, this equation may overestimate *Ptros* at higher salinities, e.g. up to 20% at salinity around 30 psu (Khaitov et al. 2021, Marchenko et al. 2023).

2.2.2 Assessment of environmental parameters

In total, we used seven parameters describing possible influence of rivers, ports, surf and substrate on mussels (**Table 1**). We used three different proxies of salinity (*RiverSize*, *DistRiver* and *Salinity*) because, in our opinion, a single estimate of salinity at low tide could be insufficient to characterize overall salinity and river influence per se at the sampling sites. ~~Salinity~~ *Salinity* was measured directly with an accuracy of 1 ppt using an Atago S/Mill-E refractometer. To classify rivers by size (*RiverSize*), the data from ESM (**Table S1 in Supplement ~~2~~1**) ~~was~~ were used. To calculate *Fetch*, the R-package "waver" (Marchand & Gill 2018) was applied to regional geographic map shape-files.

2.3 Testing datasets

Three datasets were used as testing ones, one from the same area of the White Sea ~~**(Table S4)**~~ and two from the Barents Sea. "Kandalaksha littoral" dataset contained 23 samples from 12 littoral sites in the Kandalaksha Bay **(Fig. S1 in Supplement 2)**. We took only algal samples at four sites, only bottom samples at four other sites and samples from both substrates at the remaining four sites (**Table S4 in Supplement ~~2~~1**). Environmental parameters were assessed in the same way as for the modeling dataset.

"Tyuva littoral" and "Tyuva sublittoral" testing datasets were extracted from the published data of Marchenko et al. (2023). These authors mapped in detail the distribution of *Ptros* in mussel settlements of the 3-km-long Tyuva Inlet in the Kola Bay of the Barents Sea (**Fig. 1**) sampled in 2009-2010. *Ptros* was predicted either by direct genotyping or from morphotype frequencies using the formula derived for local populations existing under salinities higher than in the White Sea (Marchenko et al. 2023). They provided a number of environmental characteristics including depth, ~~Salinity~~ *Salinity*, cover of macrophytes in rank scale, and dominant algal species (usually, kelps in the sublittoral and fucoids on the littoral) for each sampling site. “Tyuva littoral” set contained samples from all 23 littoral sites from the depth range corresponding to the fucoid belt (0.5-1.5 m above mean spring tidal depth, Marchenko et al. ~~[2023]~~ (2023); note that the position of fucoid belt in the Barents Sea differs from that in the White Sea due to the different tidal amplitude). “Tyuva sublittoral” contained samples from all 15 sublittoral sites (depth range from -0.5 to -3.5 m). Since the substrate of mussel fouling was not registered during sampling, we classified samples into bottom and algal ones by the algal cover in the sites (ranks 1-3 and 4-5, correspondingly). The remaining environmental parameters were assessed as for the modeling dataset, with the nearest port in Ekaterininskaya Gavan Bight considered as active and the river Tyuva flowing into the inlet as a large one.

2.4 Statistical analysis

All the data were processed using the statistical programming language R 4.05 (R Core Team 2023).

2.4.1 Dependency of *Ptros* on environmental parameters in modeling dataset (Model 1)

We used generalized additive mixed model (~~GAM~~ GAMM, Wood 2017) as a modeling technique, which has been shown to work well for SDM construction (Elith et al. 2006). ~~Importantly, GAM assumes that the relationship between the dependent variable (in our case *Ptros*) and continuous predictors is not necessarily linear but may be curvilinear (Austin 2002).~~ One of the strengths of this approach is that additive models assume that the relationship between the dependent variable (in our case *Ptros*) and continuous predictors is not necessarily linear but may be curvilinear (Austin 2002). The weakness of the approach is that it does not provide a direct assessment of either relative or absolute importance of factors.

~~GAM~~ GAMM fitted (hereafter, **Model 1**) was based on beta-binomial residuals distribution and the restricted maximum likelihood method for ~~parameters~~ estimation of the parameters. Smoothers for all continuous predictors were fitted using cubic basic splines. Categorical predictors were included as parametric terms in the model. Site was considered as a random factor. The function gam() from the package "mgcv" (Wood 2017) was used to fit the model.

To check for ~~the~~ all predictors' collinearity in ~~the model~~ **Model 1** and other models, we calculated the variance inflation factor (VIF, Fox & Monette 1992) considering the value less than 3.5 as acceptable (Quinn & Keough 2002). ~~Additionally, we calculated Pearson correlation between continuous predictors.~~ To verify that **Model 1** met the assumptions of sampling independence, we examined the presence of residuals' spatial autocorrelation by means of spline correlogram construction (Bjørnstad & Falck 2001) with the function spline.correlog() from the package "ncf" (Bjornstad 2022) and found no evidence of spatial autocorrelation. We also considered the model residuals in relation to year of sampling and found no significant patterns.

Dependence of abundance of mussels of different morphotypes on environmental parameters in a modeling dataset (Model 2)

The relationships between the taxonomic structure and the predictors were further investigated using abundances of mussels of different morphotypes. This means that we equated morphotypes with species. However, this assumption should not have crucially biased the results of the analysis, given the proportional relationship between *PT* and *Ptros* in mussel settlements in the study area (Khaitov et al. 2021). The mean abundances of mussels of E- and T-morphotypes across samples from each site (from both substrates) were log-transformed and used as the dependent variable in **Model 2**. The model was constructed as a generalized additive one (GAM) with Gaussian residuals' distribution and included the factor *Morphotype*, the same set of predictors as in **Model 1** except *Substrate* and *Site*, and interactions between *Morphotype* and other predictors.

2.4.2 Association between *Ptros*, substrate and mussel abundance

The ultimate goal of the analysis was to find out how the segregation of *ME* and *MT* between algal and bottom substrates depended on the abundance of each species on each substrate. For each site we calculated the difference between proportion of *MT* in algal and bottom samples: $Dif = Ptros_{Algae} - Ptros_{Bottom}$. The obtained ~~Dif~~ *Diff* values were used as a dependent variable in ~~**Model 2**~~ **Model 3**, which was constructed as GAM with Gaussian residuals' distribution.

~~Assessing the dependence of *Diff* on $Ptros_{Site}$ and mussel abundances, we could not directly operate with *ME* and *MT* densities because they could be calculated otherwise than through *Ptros*, which would have inevitably resulted in the collinearity of the predictors. Therefore, we performed principal component analysis for the abundance matrix of T- and E-morphotypes on algal and bottom substrates. Thus we used PC1 and PC2 values as independent variables, along with $Ptros_{Site}$, in **Model 2**. This means that we had to equate morphotypes with species in this case. However, this assumption should not have crucially biased the results of the analysis, given the proportional relationship between *PT* and *Ptros* in mussel settlements from the study area (Khaitov et al. 2021). We used VIF to control for the level of collinearity of the final set of predictors considering the value less than 3.5 as acceptable (Quinn & Keough 2002).~~

Assessing the dependence of *Diff* on $Ptros_{Site}$, we could not directly operate with densities of morphotypes because they were collinear on different substrates (see Results). Therefore, we performed principal component analysis for the abundance matrix of T- and E-morphotypes on algal and bottom substrates and used PC1 and PC2 values as proxies of morphotype abundances, along with $Ptros_{Site}$, in **Model 3**.

2.4.3 Assessment of predictive power of Model 1

~~We wanted t~~To check whether **Model 1** could be used to predict the dominant species in bottom and algal samples at a site with known environmental parameters, *MT* (*Ptros*>0.5) or *ME* (*Ptros*<0.5)~~. To do so~~, we used all the parameters from **Model 1** to predict $Ptros_{Algae}$ and $Ptros_{Bottom}$ for each site within the modeling dataset and within each of the three testing datasets. The predicted values were categorized into those greater than 0.5 and those ~~less~~ smaller than 0.5 and considered to be classifiers for detecting *MT*- or *ME*-dominated samples. The receiver operating characteristics (ROC) followed by the analysis of the area under the curve (AUC, Fielding & Bell 1997, Fawcett 2006) were used to evaluate the performance of the models. We considered AUC ~~more~~ greater or equal to 0.7 as acceptable discrimination (Hosmer et al., 2013). Function roc() from the package “pROC” (Robin et al. 2011) was used.

3. RESULTS

Ranges and median values of the continuous predictors are summarized in **Table 1**. While the distribution of *Fetch* and *Salinity* values was highly ~~mosaic~~ variable, the most wind-exposed sites were located on the southeastern coast of the Bay and on open shores of the islands in its top (**Fig. 1** A) while the most desalinated areas were located in the very top of the Bay ~~**(Fig. 1 B)**~~ (Fig. 1 C). Expectedly, *Salinity* tended to decrease towards river mouths (**Fig. S2** A **in Supplement** ~~**1**~~ **3**) and was lower closer to large rivers than to small ones (**Fig. S2** B **in Supplement** ~~**1**~~ **3**). Sites close to ports tended to have lower *Fetch* (**Fig. S2** C**in Supplement** ~~**1**~~ **3**), but no association between *DistPort* and *Salinity* was observed (**Fig. S2** D **in**

Supplement ~~**1**~~ **3**). All Pearson’s correlations between *Salinity*, *DistRiver*, *DistPort* and *Fetch* were rather low (**Table S5 in Supplement** ~~**2**~~**1**), the largest being that between *Fetch* and *DistPort* (r = 0.525).

~~When one examines separate~~ A striking feature visible on maps of *Ptros* distribution across algal and bottom substrates~~,~~ is the universally elevated proportion of *MT* on the former ~~is striking~~ (**Fig. 1** D-G). While spatial distribution of *Ptros* was highly variable ~~mosaic~~, its maximum values on both substrates were observed in the Bay’s top and in some deep inlets, while its minimum values were observed along the open part of the southeastern coast (**Fig. 1** D-G). ~~Associations between *Ptros* and environmental predictors other than substrate could not be discerned on the maps (**Fig. 1**).~~ It is difficult to discern relationships between *Ptros* and any environmental predictors other than substrate in the small-scale maps shown in **Figure 1**. For this purpose, it is necessary to consider **Model 1**.

3.1 Relationship of *Ptros* and environmental parameters evaluated by Model 1

Although some non-zero pairwise correlations between environmental factors were found (see above), VIF values calculated for the predictors were generally low (maximal VIF being that for *Fetch*, 1.76). ~~In our opinion, this~~ This result means that the collinearity between the predictors was negligible, i.e. they did not mask each other’s influence.

Model 1 explained 77% of the total deviance. It revealed a significant dependency of *Ptros* on all predictors except *DistRiver*. Effective degrees of freedom for *DistPort* and *Fetch* were close to one, indicating the linear dependence of *Ptros* on them. ~~On the contrary,~~ In contrast, the dependence on the third continuous predictor, *Salinity*, was curvilinear (**Table 2**).

According to the model, *Ptros* decreased both with *DistPort* ~~(**Fig. 2** A)~~ (**Fig. 2 E**) and with *Fetch* ~~(**Fig. 2**, B)~~ (**Fig. 2 G**). This means that the proportion of *MT* was higher near ports and in surf-protected areas. *PortStatus* also had a significant effect: predicted *Ptros* was higher near active ports than near abandoned ones ~~(**Fig. 2** E, F)~~ (**Fig. 2 I**). The curvilinear dependence of *Ptros* on *Salinity* can be described as follows: predicted *Ptros* decreases with salinity in the range from low to ~~“normal” salinity (~~ about 24 psu ~~ppt~~ (average salinity in the White Sea) and increases again at higher salinities (up to 30 ppt) ~~(**Fig. 2**, D)~~ (Fig. 2 A) . ~~Besides,~~ In addition, predicted *Ptros* was higher near large rivers than near small ones (Fig. 2 I). Finally, *Ptros* was higher on algal substrates than on bottom ones (Fig. 2 I, see also Fig 1 D-F). ~~**Fig. 1** C, D; **Fig. 2** E, F). As mentioned above, distance to the nearest river did not affect *Ptros* (**Fig 2** C).~~

3.2 Dependence of abundance of mussels of different morphotypes on environmental parameters evaluated by Model 2

The results of Model 2 were in complete agreement with those of Model 1 for all the predictors, i.e., *Salinity* (**Fig. 2 B**), *RiverSize* (**Fig. 2 J**), *DistPort* (**Fig. 2 F**), *PortStatus* (**Fig. 2 J**), and *Fetch* (**Fig. 2 H**) (see **Table S6 in Supplement 1** for all model parameters). In addition, they revealed an asymmetry in the responses of the two species to some of these predictors. While the abundance of T-morphotypes did not vary with *Salinity*, that of E-morphotypes dropped at low salinity (**Fig. 2 B**). On the other hand, the abundance of E-morphotypes slightly varied with *Fetch* and *DistPort*, while that of T-morphotypes strongly decreased both with the distance from ports (**Fig. 2 F**) and with surf level (**Fig. 2 H**).

3.~~2~~3 Dependency of *Ptros* on substrate and mussel abundance evaluated by ~~Model 2~~ Model 3

~~In the principal component analysis of the abundance matrix of T- and E- morphotypes on different substrates, PC1 and PC2 explained 62% and 20% of variation, respectively.~~ The principal component analysis of the abundance matrix of T- and E- morphotypes allowed to find high positive correlation of PC1 (explained 62% of total variation) with abundances of T-morphotypes and of PC2 (20% of total variation) with abundances of E-morphotypes ~~was found~~ on both substrates (**Fig. 3** B, C). Thus, the abundance of conspecific morphotypes varied consistently on different substrates (see also ~~**Fig. 1** C, D~~

Fig. 1 D, F). Therefore, PC1 and PC2 can be considered as proxies of *MT* and *ME* abundance, respectively.

Parameters of ~~**Model 2**~~ **Model 3**, which explained 31% of the deviance, are provided in ESM (**Table S7 in Supplement 2**). **Figure 3** demonstrates how the difference between *MT* proportion on algal ~~($Ptros_{Algae}$)~~ and bottom ~~($Ptros_{Bottom}$)~~ substrates (~~*Dif*~~ *Diff*) depends on *MT* prevalence at the site ($Ptros_{Site}$) and mussel abundances in terms of PCs according to the model. The dependence of *Diff* on $Ptros_{Site}$ was significant ($p < 0.001$~~, **Table S6 in Supplement 2**~~) and, expectedly, bell-shaped, with minimal values at sites absolutely dominated by *ME* or *MT* (*Ptros* close to 0 or 1) and maximal at sites with equal presence of both species (**Fig.~~4~~**3 A). Dependence of *Diff* on PC1 was marginally significant ($p = 0.087$) and tended to decrease with increasing PC1 (**Fig.3** B). The dependence of *Diff* on PC2 was significantly positive ($p = 0.011$, **Table S6 in Supplement 2, Fig.4** 3 C). This means that the species were strongly segregated by substrates at sites with a high *ME* abundance but not at sites with a high *MT* abundance.

3.3 Assessment of predictive power of Model 1

~~The ability of~~ **Model 1** ~~to classify samples into *ME*- and *MT*-dominated ones~~ ~~was good~~ fit well for the "Kandalaksha littoral" testing dataset (AUC=0.85 vs AUC=0.84 for modeling dataset)~~,~~. It classified the samples into *ME*- and *MT*-dominated ones fairly well, with the exception ~~only~~ of a few false negatives, ~~(~~i.e., sites unpredictably dominated by *MT*~~)~~ (**Fig. 4** A, B). Its ~~P~~predictive power ~~value of the model~~ for the two testing sets from the Barents Sea was lower but acceptable ~~although not critically so~~: AUC = 0.71 for "Tyuva littoral" and AUC=0.69 for "Tyuva sublittoral". ~~Unlike~~ In constrast to the "Kandalaksha littoral" testing dataset, most of the false results were positive~~, i.e., the model overestimated *Ptros* more often~~.

4. DISCUSSION

~~Having applied the SDM approach to an unprecedentedly extensive material data, we demonstrated that a~~Almost all environmental predictors considered in our study — namely, surf level, distance to the port, status of the port ~~(active vs abandoned)~~, salinity at low tide, size of the nearest river and fouling substrate ~~(fucoid algae vs bottom substrates)~~ — influenced the distribution of *Mytilus edulis* (*ME*) and *M. trossulus* (*MT*) in the White Sea. The differences in the distribution~~, evident at scales ranging from meters to tens of kilometers,~~ reflected the partial divergence of ecological niches of these two species.

~~Below we discuss the species adaptations that may underlie the patterns of *ME* and *MT* distribution against different predictors. Then we consider the possible role of competition in their segregation by substrates. Further, we discuss whether the same set of predictors can drive segregation of these species in other habitats than the littoral fucoid belt, elsewhere than in the White Sea and outside the Kola contact zone. Finally, we review the strengths and weaknesses of our approach to assessing ecological niche partitioning of sympatric mussels.~~

4.1 Ecological niche partitioning between MT and ME in the Kola contact zone

~~Our results show~~ We showed that the most expected habitat ~~for~~ of *ME* in the White Sea littoral ~~is~~ was a bottom substrate in a surf-exposed location with a surface salinity of 24 psu (average ~~a "normal" an average surface salinity~~ for the White Sea) ~~(24 ppt psu)~~ situated away from ports and large rivers. The most expected habitat ~~for~~ of *MT* ~~is~~ was an algal substrate in a ~~wind~~surf-protected location with a salinity lower than ~~"normal"~~ the White Sea average ~~salinity~~ situated close to active ports and large rivers. ~~Only the~~ While differences ~~related to~~ associated with ~~ports and~~ substrate~~s~~ and distance to ports have been previously noted in the White Sea (Väinölä & Strelkov 2011, Katolikova et al. 2016), those associated with salinity and surf exposure were first uncovered in this study.

4.1.1 Segregation by salinity

In the Baltic Sea *MT* is adapted to an extremely low salinity~~, as confirmed by ecophysiological data~~ (Knöbel et al. 2021, Wiesenthal et al. 2025 ~~and references therein~~). Comparative ecophysiological data on *MT* and *ME* elsewhere are inconclusive (Gardner & Thompson 2001, Qiu et al. 2002, Sokolova et al. 2024). Before our study, there has been no convincing evidence of segregation of these species by salinity in contact zones outside the Baltic~~, in particular, in the Kola zone~~ (Moreau et al. 2005, Riginos & Cunningham 2005, Katolikova et al. 2016, Marchenko et al. 2023). For the White Sea, this lack of evidence could be due to at least three reasons. Firstly, the role of salinity ~~in species segregation~~ may be masked by other ~~important~~ factors ~~e.g.~~ such as distance to ports. ~~Secondly, the range of salinity in mussel habitats in the White Sea is relatively narrow as compared to the Baltic Sea.~~ This seems particularly plausible in the light of our data that *MT* in the White Sea is a euryhaline species forming mass settlements in the entire salinity range recorded in our study, while *ME* is much less abundant at salinity below 15 psu (**Fig. 2B**). Secondly, in contrast to the Baltic, there are no broad geographic salinity gradients in the White Sea, only local ones.

The third reason is the curvilinear dependence of the proportion of *MT* in mixed settlements (*Ptros*) on salinity: *Ptros* increases not only when the salinity is reduced but also when it is extremely high for the White Sea (up to 30 ~~ppt~~ psu, **Fig. 2 ~~D~~ B**). This nonlinearity, which may prevent the dependence from being detected, can be explained in two ways. On the one hand, local summer surface salinity above 24 ppt in the Kandalaksha Bay, supposedly associated with irregular episodes of upwelling (Dale & Prego 2003), may be a nonspecific stress for littoral animals adapted to lower salinity~~, while *MT* can tolerate it better, being a more opportunistic species (Katolikova et al. 2016, see also below).~~ MT may be more tolerant to this stress. On the other hand, as shown in detailed studies at the Barents Sea (Khaitov et al.

2021, Marchenko et al. 2023), the method of predicting *Ptros* ("morphotype test") used in our study may ~~slightly~~ overestimate it at salinities close to 30 ppt. Therefore, we cannot rule out the possibility that the increased *Ptros* at sites with a high salinity is an artifact.

4.1.2 Non-random distribution depending on distance to ports

~~It has been suggested that t~~The ~~confinement~~ association of *MT* ~~to~~ with harbors in the White and the Barents Sea ~~is~~ may be associated with its invasion ~~into the region~~ with maritime transport from the western Atlantic in the 20th century (Väinölä & Strelkov 2011~~). This hypothesis agrees with the available genetic data (Väinölä & Strelkov 2011, Wenne et al. 2020, Simon et al. 2021).~~ It has also been hypothesized that *MT* is more resistant to anthropogenic pollution and is better adapted to disturbed habitats than *ME* (Katolikova et al. 2016). Our observation that *MT* frequency is lower near abandoned ports than near active ones is consistent with this hypothesis. However, the propagule pressure of *MT* may have decreased near abandoned harbors in recent decades, which could affect the size of its populations.

4.1.3 Segregation by surf level

~~The fact that *ME* and *MT* are segregated~~ Segregation of *ME* and *MT* by surf level may be due to the well-known differences in the mechanical properties of their shells and the ability to form dense aggregations. *ME* has ~~less flexible,~~ thicker~~,~~ ~~and~~ heavier and less flexible shells (Beaumont et al. 2008, Michalek et al. 2021)~~,~~ and is more inclined to form tight clumps (Liu et al. 2011). These features may be adaptive on exposed coasts. ~~Unfortunately, there are no comparative data on the differences between *ME* and *MT* in byssus secretion and attachment strength, which theoretically might also affect their distribution by surf level as well as across substrates~~ Theoretically, the distribution of *ME* and *MT* by surf level as well as substrate may also be affected by differences in byssus secretion and attachment strength. Unfortunately, there are no comparative data on this topic.

4.1.4 Segregation by substrate

~~The differences in shell structure and aggregation behavior possibly explaining segregation by surf may also explain that~~ Segregation by substrate may be explained by the same differences as segregation by surf level. ~~An ability to form dense aggregations is an adaptation to life on bottom, not on algae.~~ Other things being equal, *MT*, with its thin~~ner~~ fragile shells, should be lighter than *ME* (Michalek et al. 2021) and thus better ~~adapted~~ suited to life on algae. ~~Further~~In addition, fucoid thalli may serve as shock absorbers ~~for fragile *MT*~~ (Katolikova et al. 2016) and provide shelter ~~them~~ from starfish selectively preying on *MT* in mixed settlements (Khaitov et al. 2018, Khaitov et al. 2023). By the same token, denser aggregations formed by *ME* are more adaptive on the bottom than on algae.

4.1.5 Competition for substrate

Whatever historical, physiological, morphological, behavioral and other features influence the segregation of *MT* and *ME* ~~by the environmental factors~~, interspecific competition may also be involved. Assessing the role of mussel abundance in the ~~degree of species~~ segregation across substrates, we found that ~~while~~ *MT* abundance did not significantly affect it~~,~~ but *ME* abundance did: as the latter increased, the degree of segregation increased, too (**Fig. 3** B,C). In our opinion, this pattern results from the divergence of the realized species niches~~.~~: *ME* outcompetes *MT* on bottom substrates displacing it to algal thalli~~, which appear to be a less suitable substrate for *ME* (see above).~~

Spatial segregation of sympatric mussels by substrates, which is apparently density-dependent, is evident at the level of tens of centimeters (Katolikova et al. 2016). Direct analogies for segregation at such a small scale can be found in other attached organisms, terrestrial plants (Raventós et al. 2010). A “biologically generated spatial pattern” model~~,~~ ~~relating inter-specific segregation with the intra-specific clustering in competing species, has been suggested~~ (Pacala & Levin 1997, Amarasekare 2003) relates inter-specific segregation with the intra-specific clustering in competing species. Our findings suggest that this model can also be applied to mussels.

4.2 Predictive power of SDM

The ability of our model to classify sites into *ME*- and *MT*- dominated ones in an independent testing dataset from the White Sea was high (AUC = 0.85). Therefore, we assume that the predictors included in the model explain most of the variation in species distribution within the ~~studied habitat, the~~ littoral fucoid belt in the White Sea. ~~The model also showed a satisfactory performance on independent data from the Tyuva inlet in the Barents Sea (AUC ≈ 0.7), including sublittoral data. This result highlights the versatility of this set of predictors as regulators of *ME* and *MT* distribution in the Kola contact zone.~~

~~The worst predictive value of the model for the Barents Sea data may be due to the following reasons.~~ The worst, though formally satisfactory predictive value of the model for the Barents Sea data (AUC ≈ 0.7) may be due to the following reasons. Firstly, it may be associated with a large depth range of the sampling sites, considering that distribution of *ME* and *MT* in the Tyuva Inlet by depth is non-random (Marchenko et al. 2023)~~, it may be associated with a large depth range of the sampling sites~~. The second reason may be a coarser categorization of the Barents Sea samples into algal and bottom ones. Since fouling substrate was not taken into account during sampling, we predicted it based on the projective cover of algae at the sampling site. Thirdly, we do not know whether the two species are non-randomly distributed across bottom and algal substrates in the sublittoral, where fucoids are replaced by kelps. The fourth reason could be a narrow variation of *DistPort*, *DistRiver*, and *Fetch* in the small Tyuva Inlet in comparison with the Kandalaksha Bay.

Finally, the fact that SDM tended to overestimate *Ptros* in the Barents Sea data (false positive predictions) is consistent with the observation that the proportion of *MT* has been declining in the study area in the 2010s under seemingly stable environmental conditions in terms of predictors included in our model (Marchenko et al. 2023). This observation suggests the presence of some yet ~~unknown~~ unstudied factors regulating the taxonomic structure.

4.3 Ecological niche partitioning between MT and ME in the Kola contact zone as compared to other zones

Blue mussels are a challenging model for studying ecological niche partitioning between ~~“cryptic”~~ cryptic species in sympatry due to their wide distribution, biogeographic history and hybridization. *ME* and *MT* play similar ecological roles in their native oceans, Atlantic and Pacific, respectively (compare Commito & Dankers 2001 and Bodkin et al. 2018) and ~~therefore~~ may inherently have ~~strongly~~ overlapping fundamental ecological niches. Contact zones between these species in the Atlantic can be considered as ecological (and evolutionary) ~~“experiments”~~ experiments, set in strikingly different environments (from Baltic to Spitsbergen) at different times (from late post-glacial to the historical period, Väinölä & Strelkov 2011, Wenne et al. 2020, and references therein). The ~~“design”~~ design of these ~~“experiments”~~ experiments was possibly different too, because in some zones the original settler could be *ME* and in others, *MT*. In addition, competition (~~“ecological character displacement”~~ ecological character displacement, Pfennig & Pfennig 2009), hybridization (~~“reinforcement of prezygotic reproductive isolation”~~ reinforcement of prezygotic reproductive isolation, Lukhtanov 2011) and introgression (~~“adaptive introgression”~~ adaptive introgression, Hedrick 2013) could influence the divergence of their ecological phenotypes differently in different zones. These considerations suggest that the zones should differ, and this hypothesis has been a recurrent theme in genetic research on blue mussel contact zones (Riginos & Cunningham 2005, Bierne et al. 2011, Fraïsse et al. 2016). Nevertheless, we believe that the differences between ~~these two~~ *ME* and *MT* ~~as different species~~ are more fundamental and ~~thus~~ that conspecific ecological phenotypes ~~(“niches”)~~ in different zones should thus be similar, producing comparable patterns in species distributions. Some results of this study support this hypothesis.

The observation that *MT* frequency is elevated in low-salinity habitats not only in the Baltic but also in the White Sea seems to resolve the old conundrum about seemingly contrasting salinity adaptations of the Baltic and other Atlantic *MT* populations (e.g. Riginos & Cunningham 2005, Katolikova et al. 2016, see also above). Further, an increased *MT* frequency has been repeatedly observed in calm and freshened waters e.g. in the tops of fjords near Bergen in Norway (Ridgway & Nævdal 2004) and Uummannaq in Greenland (Wenne et al. 2020) and in Loch Etive in Scotland (Beaumont et al. 2008), which is hardly a coincidence. Our observations indicate that this combination of weak surf and low salinity is also favorable for *MT* in the White Sea.

~~No non-random relationship between the distribution of *ME* and *MT* and any of the predictors significant in the White Sea has been convincingly demonstrated in other contact zones, with the exception of salinity in the Baltic contact zone.~~ With the exception of salinity in the Baltic, none of the predictors affecting segregation of *ME* and *MT* in the White Sea have been convincingly shown to act in other contact zones. Data on surf are inconsistent (compare Bates & Innes 1995, Comesaña et al. 1999, Tam & Scrosati 2014 and this study), ~~while~~ and data on ~~fouling~~ substrates are~~, as far as we know, completely~~ absent. ~~If our assumption is correct and the diverging preferences of *ME* and *MT* for sites differing as to surf and substrates are associated with the differences in their morphology and behavior (see above), then these differences should be manifested universally.~~ If different preferences of *ME* and *MT* for surf level and substrates are indeed associated with their morphological and behavioral differences (Katolikova et al. 2016; this study), these preferences should be omnipresent. Ad hoc studies in other contact zones might throw light on this matter. ~~The intriguing differences between these species in stress tolerance, particularly to anthropogenic pollution (as in harbors), also remain unexplained (see discussion in Brooks et al. 2015 and Beyer et al. 2017).~~ It would also be interesting to examine the differences between *ME* and *MT* in tolerance to stress, particularly anthropogenic pollution, e.g. in harbors (see discussion in Brooks et al. 2015 and Beyer et al. 2017).

~~The classical review on the divergence of ecological niches of *ME* and *MT* in different contact zones (Riginos & Cunningham 2005) is already 20 years old. The time is obviously ripe for a new survey, and our observations from the Kola zone may prove useful.~~

4.4 Strengths and weaknesses of our approaches to the study of sympatric mussels

The methods ~~of taxa identification, environment parameters assessment and modeling~~ used in our study have certain limitations. We identified the mussels ~~using~~ with the help of the "morphotype test", ~~allowing the assessment of the taxonomic structure of mussel settlements without genotyping. This test works well in habitats with salinity below 25 ppt in the Kola contact zone (Khaitov et al. 2021) but~~ which does not permit a direct assessment of species abundances ~~or an identification of~~ and does not identify hybrids as a separate category. ~~The former limitation makes it difficult to account for the role of inter-species competition, which, judging from our experience with different substrates, is important. The latter limitation is alleviated by the fact that hybrids are relatively scarce in the Kola zone (Väinölä & Strelkov 2011, Wenne et al. 2020). However, this is not the case in other contact zones (Väinölä & Strelkov 2011, Wenne et al. 2020), where hybrids may be important ecological actors (e.g. Schwartz et al. 2024).~~The latter limitation is not particularly significant in the Kola zone, where hybrids are relatively scarce

(Väinölä & Strelkov 2011, Wenne et al. 2020), but may compromise studies in other contact zones, where hybrids may play an important ecological role (e.g. Schwartz et al. 2024).

Although *ME* and *MT* differ ~~universally~~ everywhere in morphotype frequencies, the magnitude of the differences varies between contact zones and between habitats with different salinities in the Arctic (Khaitov et al. 2021). This means that the “morphotype test” must be ~~additionally~~ calibrated before use in a new area (see Khaitov et al. 2021 for recommendations). Multilocus genotyping, while still too costly for processing dozens of thousands of specimens needed for SDM, remains the gold standard of taxonomic assessment in blue mussels.

~~It should also be noted that we did not account for all potential predictors affecting species segregation (e.g. depth, Marchenko et al. 2023, or predators, Khaitov et al. 2018, Khaitov et al. 2023).~~

We did not account for some potential predictors affecting species segregation such as depth (Marchenko et al. 2023), predators (Khaitov et al. 2018, Khaitov et al. 2023) or temperature (Kijewski et al. 2019). Moreover, some of our predictors could have been estimated more carefully (for example, bottom salinity at high ~~water~~ tide could be more informative for littoral mussels than salinity at low ~~water~~ tide ~~when they are not submerged~~). However, since most of our predictors were shown to be significant, they should not be ignored in future studies.

The correlative approach used in our study does not allow a direct assessment of either relative or absolute importance of factors. For instance, we cannot say whether salinity or substrate is more crucial. However, the take-home message from our research is that there is no single “leading” factor determining the distribution of *ME* and *MT*, contrary to the idea that has dominated the field since the pioneering studies in the Baltic (e.g. Gardner & Thompson 2001, Ridgway & Nævdal 2004, Riginos & Cunningham 2005, Śmietanka et al. 2014).

The limitations discussed above do not detract from the fact that, as shown in our ~~pioneering~~ study, SDMs ~~are a potentially~~ may be a useful tool for the study of distribution of *ME* and *MT* in sympatry. Their obvious benefits include the possibility to analyze the distribution of the species in the space of multiple predictors simultaneously, the possibility to control the collinearity of the predictors and the lack of necessity to treat dependencies as linear.

~~Promising directions of further research on niche partitioning in sympatric mussel species are, in our opinion, as follows. Firstly, a parallel study in different contact zones would reveal common and zone-specific patterns. Secondly, the use of taxonomic methods allowing direct assessment of abundances of species and their hybrids would elucidate the nature of competition between them all. Incorporation of additional environmental factors, including biotic ones, into SDMs might yield surprising results. Finally, it would be worthwhile to have a closer look at different spatial scales, down to the smallest one, in the segregation of these two mussel species.~~

Promising directions of further research on niche partitioning in sympatric mussel species are, in our opinion, as follows. The use of taxonomic methods allowing direct assessment of abundances of species and their hybrids would elucidate the nature of competition between them all. Incorporation of additional environmental factors into SDMs might yield surprising results. Further, it would be worthwhile to have a closer look at different spatial scales, down to the smallest one, in the segregation of these two mussel species. Finally, a parallel study in different contact zones would reveal common and zone-specific patterns. The classical review on the divergence of ecological niches of *ME* and *MT* in different contact zones (Riginos & Cunningham 2005) is 20 years old. The time is ripe for a new survey, and our observations from the Kola zone may prove useful.

5. ACKNOWLEDGEMENTS

We thank all participants of our expeditions, especially Gita Paskerova and Eugene Genelt-Yanovsky, for their help in fieldwork. We are deeply grateful to Natalia Lentsman for the English language assistance during the preparation of the manuscript. We would like to express our special gratitude to the administration of the Kandalaksha State Nature Reserve, where a significant part of the material was collected, for comprehensive assistance. This study was supported by the Russian Science Foundation, grant no. 19-74-20024.

6. LITERATURE CITED

Amarasekare P (2003) Competitive coexistence in spatially structured environments: A synthesis. Ecology Letters 6:1109–1122.

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological modelling 157:101–118.

Bates J, Innes D (1995) Genetic variation among populations of *Mytilus* spp. in eastern Newfoundland. Marine Biology 124:417–424.

Beaumont AR, Hawkins MP, Doig FL, Davies IM, Snow M (2008) Three species of *Mytilus* and their hybrids identified in a Scottish Loch: natives, relicts and invaders? Journal of Experimental Marine Biology and Ecology 367:100–110.

Berger V, Dahle S, Galaktionov K, Kosobokova X, Naumov A, Rat'kova T, Savinov V, Savinova T (2001) White Sea. Ecology and Environment. St-Petersburg-Tromso: Derzavets Publishers.

Berger VY, Naumov A (2000) General features of the White Sea. Berichte Polarf 359:3–9.

Beyer J, Green NW, Brooks S, Allan IJ, Ruus A, Gomes T, Bråte ILN, Schøyen M (2017) Blue mussels (*Mytilus edulis* spp.) as sentinel organisms in coastal pollution monitoring: A review. Marine Environmental Research 130:338–365.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. Trends in Ecology and Evolution 22:148–155.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. Molecular ecology 20:2044–2072.

Bjornstad ON (2022) ncf: Spatial Covariance Functions. R package version 1.3-2.

Bjørnstad ON, Falck W (2001) Nonparametric spatial covariance functions: Estimation and testing. Environmental and Ecological Statistics 8:53–70.

Bodkin JL, Coletti HA, Ballachey BE, Monson DH, Esler D, Dean TA (2018) Variation in abundance of Pacific blue mussel (*Mytilus trossulus*) in the Northern Gulf of Alaska, 2006–2015. Deep Sea Research Part II: Topical Studies in Oceanography 147:87–97.

Brooks SJ, Farmen E, Heier LS, Blanco-Rayón E, Izagirre U (2015) Differences in copper bioaccumulation and biological responses in three *Mytilus* species. Aquatic Toxicology 160:1–12.

Buschbaum C, Dittmann S, Hong JS, Hwang IS, Strasser M, Thiel M, Valdivia N, Yoon SP, Reise K (2009) Mytilid mussels: Global habitat engineers in coastal sediments. Helgoland Marine Research 63:47–58.

Caradima B, Schuwirth N, Reichert P (2019). From individual to joint species distribution models: A comparison of model complexity and predictive performance. Journal of Biogeography *46*:2260-2274.

Comesaña AS, Toro JE, Innes DJ, Thompson RJ (1999) A molecular approach to the ecology of a mussel (*Mytilus edulis - M. trossulus*) hybrid zone on the east coast of Newfoundland, Canada. Marine Biology 133:213–221.

Commito JA, Dankers NM (2001) Dynamics of spatial and temporal complexity in european and north american soft-bottom mussel beds. In: *Ecological comparisons of sedimentary shores*. Springer, p 39–59

Dale AW, Prego R (2003) Tidal and seasonal nutrient dynamics and budget of the Chupa Estuary, White Sea (Russia). Estuarine, Coastal and Shelf Science 56:377–389.

Dennis AB, Hellberg ME (2010) Ecological partitioning among parapatric cryptic species. Molecular Ecology 19:3206–3225.

Druehl LD, Green JM (1982) Vertical distribution of intertidal seaweeds as related to patterns of submersion and emersion. Marine Ecology Progress Series 9:163–170.

Dufresnes C, Poyarkov N, Jablonski D (2023) Acknowledging more biodiversity without more species. Proceedings of the national Academy of Sciences 120:e2302424120.

Elith J, H. Graham C, P. Anderson R, Dudík M, Ferrier S, Guisan A, J. Hijmans R, Huettmann F, R. Leathwick J, Lehmann A, Li J, G. Lohmann L, A. Loiselle B, Manion G, Moritz C, Nakamura M, Nakazawa Y, McC. M. Overton J, Townsend Peterson A, J. Phillips S, Richardson K, Scachetti-Pereira R, E. Schapire R, Soberón J, Williams S, S. Wisz M, E. Zimmermann N (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annual review of ecology, evolution, and systematics 40:677–697.

Fawcett T (2006) An introduction to ROC analysis. Pattern Recognition Letters 27:861–874.

Fielding AH, Bell JF (1997) A Review of Methods for the Assessment of Prediction Errors in Conservation Presence/Absence Models. Environmental conservation 24:38–49.

Filatov N, Pozdnyakov D, Johannessen OM, Pettersson LH, Bobylev LP (2007) White Sea: its marine environment and ecosystem dynamics influenced by global change. Springer Science & Business Media.

Fox J, Monette G (1992) Generalized collinearity diagnostics. Journal of the American Statistical Association 87:178–183.

Fraïsse C, Belkhir K, Welch JJ, Bierne N (2016) Local interspecies introgression is the main cause of extreme levels of intraspecific differentiation in mussels. Molecular Ecology 25:269–286.

Gardner JPA, Thompson RJ (2001) The effects of coastal and estuarine conditions on the physiology and survivorship of the mussels *Mytilus edulis*, *M. trossulus* and their hybrids. Journal of Experimental Marine Biology and Ecology 265:119–140.

Gardner JP, Oyarzun PA, Toro JE, Wenne R, Zbawicka M (2021) Phylogeography of Southern Hemisphere blue mussels of the genus *Mytilus*: evolution, biosecurity, aquaculture and food labelling. In: *Oceanography and marine biology*. CRC Press, p 139–228

Geller JB, Darling JA, Carlton JT (2010) Genetic perspectives on marine biological invasions. Annual review of marine science 2:367–393.

Gosling E (2021) Marine mussels: ecology, physiology, genetics and culture. John Wiley & Sons.

Hayhurst S, Rawson PD (2009) Species-specific variation in larval survival and patterns of distribution for the blue mussels *Mytilus edulis* and *Mytilus trossulus* in the Gulf of Maine. Journal of Molluscan Studies 75:215–222.

Hedrick PW (2013) Adaptive Introgression in Animals: Examples and Comparison to New Mutation and Standing Variation as Sources of Adaptive Variation. Molecular ecology 22:4606–4618.

Hellou J, Law RJ (2003) Stress on stress response of wild mussels, *Mytilus edulis* and *Mytilus trossulus*, as an indicator of ecosystem health. Environmental Pollution 126:407–416.

Hosmer JrDW, Lemeshow S, Sturdivant RX (2013) Applied logistic regression. John Wiley & Sons.

Hu Z-M, Zhang Q-S, Zhang J, Kass JM, Mammola S, Fresia P, Draisma SG, Assis J, Jueterbock A, Yokota M, Zhang Z (2021) Intraspecific genetic variation matters when predicting seagrass distribution under climate change. Molecular Ecology 30:3840–3855.

Katolikova M, Khaitov V, Väinölä R, Gantsevich M, Strelkov P (2016) Genetic, ecological and morphological distinctness of the blue mussels *Mytilus trossulus* Gould and *M. edulis* L. in the White Sea. PLoS One 11:e0152963.

Khaitov V, Makarycheva A, Gantsevich M, Lentsman N, Skazina M, Gagarina A, Katolikova M, Strelkov P (2018) Discriminating eaters: Sea stars *Asterias rubens* L. feed preferably on *Mytilus trossulus* Gould in mixed stocks of *Mytilus trossulus* and *Mytilus edulis* L. Biological Bulletin 234:85–95.

Khaitov V, Marchenko J, Katolikova M, Väinölä R, Kingston SE, Carlon DB, Gantsevich M, Strelkov P (2021) Species identification based on a semi-diagnostic marker: Evaluation of a simple conchological test for distinguishing blue mussels *Mytilus edulis* L. and *M. trossulus* Gould. PLoS ONE 16:1–27.

Khaitov VM, Makarycheva AY, Nematova RB, Evdokimova AI (2023) Predators regulate the taxonomic structure of mixed *Mytilus edulis* L. and *M. trossulus* Gould settlements in the shallow waters of the White Sea. Proceedings of the Zoological Institute of the Russian Academy of Sciences 327:8–24.

Kijewski T, Zbawicka M, Strand J, Kautsky H, Kotta J, Rätsep M, Wenne R (2019) Random forest assessment of correlation between environmental factors and genetic differentiation of populations: Case of marine mussels *Mytilus* . Oceanologia 61:131–142.

Knöbel L, Nascimento-Schulze JC, Sanders T, Zeus D, Hiebenthal C, Barboza FR, Stuckas H, Melzner F (2021) Salinity Driven Selection and Local Adaptation in Baltic Sea Mytilid Mussels. Frontiers in Marine Science 8:692078.

Knowlton N (1993) Sibling Species in the Sea. Annual review of ecology and systematics 24:189–216.

Koehn RK (1991) The genetics and taxonomy of species in the genus *Mytilus* . Aquaculture 94:125–145.

Krasavtsev LB (2011) Russia's foreign trade through the ports of the White Sea in the early twentieth century. (in Russian). Vestnik Severnogo (Arkticheskogo) federal'nogo universiteta Seriya: Gumanitarnye i social'nye nauki 1:17–24.

Lindegren M, Gabellini AP, Munk P, Edelvang K, Hansen FT (2022) Identifying key processes and drivers affecting the presence of non-indigenous marine species in coastal waters. Biological Invasions 24:2835–2850.

Liu G, Stapleton E, Innes D, Thompson R (2011) Aggregational behavior of the blue mussels *Mytilus edulis* and *Mytilus trossulus*: A potential pre-zygotic reproductive isolation mechanism. Marine Ecology 32:480–487.

Lowen JB, Hart DR, Stanley RRE, Lehnert SJ, Bradbury IR, Dibacco C, Hauser L (2019) Assessing effects of genetic, environmental, and biotic gradients in species distribution modelling. ICES Journal of Marine Science 76:1762–1775.

Lukhtanov V (2011) Dobzhansky's rule and reinforcement of prezygotic reproductive isolation in zones of secondary contact. Biology Bulletin Reviews 1:2–12.

Marchand P, Gill D (2018) waver: Calculate Fetch and Wave Energy. R package version 0.2.1.

Marchenko J, Khaitov V, Katolikova M, Sabirov M, Malavenda S, Gantsevich M, Basova L, Genelt-Yanovsky E, Strelkov P (2023) Patterns of spatial and temporal dynamics of mixed *Mytilus edulis* and *M. trossulus* populations in a small subarctic inlet (Tyuva Inlet, Barents Sea). Frontiers in Marine Science 10:1–15.

Michalek K, Vendrami DLJ, Bekaert M, Green DH, Last KS, Telesca L, Wilding TA, Hoffman JI (2021) *Mytilus trossulus* introgression and consequences for shell traits in longline cultivated mussels. Evolutionary Applications 14:1830–1843.

Morais P, Reichard M (2018) Cryptic invasions: A review. Science of the Total Environment 613:1438-1448.

Moreau V, Tremblay R, Bourget E (2005) Distribution of *Mytilus edulis* and *M. trossulus* on the Gaspe coast in relation to spatial scale. Journal of Shellfish Research 24:545–551.

Ovaskainen O, Abrego N (2020) Joint Distribution Modelling (with applications in R). Cambridge University Press.

Pacala SW, Levin SA (1997) Biologically generated spatial pattern and the coexistence of competing species. In: *Spatial ecology: The role of space in population dynamics and interspecific interactions*. Monographs in Population Biology. Princeton University Press Princeton, NJ, p 204–232

Penney RW, Hart MJ, Templeman N (2002) Comparative growth of cultured blue mussels, *Mytilus edulis*, *M. trossulus* and their hybrids, in naturally occurring mixed-species stocks. Aquaculture Research 33:693–702.

Peterson ML, Doak DF, Morris WF (2019) Incorporating local adaptation into forecasts of species' distribution and abundance under climate change. Global Change Biology 25:775–793.

Pfennig K, Pfennig D (2009) Character displacement: ecological and reproductive responses to a common evolutionary problem. The Quarterly Review of Biology 84:253–276.

Poggiato G, Münkemüller T, Bystrova D, Arbel J, Clark JS, Thuiller W (2021). On the interpretations of joint modeling in community ecology. Trends in ecology & evolution 36:391-401.

Qiu JW, Tremblay R, Bourget E (2002) Ontogenetic changes in hyposaline tolerance in the mussels *Mytilus edulis* and *M. trossulus*: Implications for distribution. Marine Ecology Progress Series 228:143–152.

Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge university press.

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Raventós J, Wiegand T, Luis MD (2010) Evidence for the spatial segregation hypothesis: a test with nine-year survivorship data in a Mediterranean shrubland. Ecology 91:2110–2120.

Reiss H, Cunze H, König K, Neumann K, Kröncke I (2011) Species distribution modelling of marine benthos: A North Sea case study. Marine Ecology Progress Series 442:71–86.

Ridgway G, Nævdal G (2004) Genotypes of *Mytilus* from waters of different salinity around Bergen, Norway. Helgoland Marine Research 58:104–109.

Riginos C, Cunningham CW (2005) Local adaptation and species segregation in two mussel (*Mytilus edulis* x *Mytilus trossulus*) hybrid zones. Molecular Ecology 14:381–400.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12:77.

Sáez AG, Lozano E (2005) Body doubles. Nature 433:111-111.

Sailing directions of the White Sea (1932) Leningrad : izdanie i tipografiya Gidrograficheskogo upravleniya. (in Russian).

Schwartz LC, González VL, Strong EE, Truebano M, Hilbish TJ (2024) Transgressive gene expression and expression plasticity under thermal stress in a stable hybrid zone. Molecular Ecology 33:e17333.

Simon A, Fraïsse C, El Ayari T, Liautard-Haag C, Strelkov P, Welch JJ, Bierne N (2021) How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. Journal of Evolutionary Biology 34:208–223.

Śmietanka B, Burzyński A, Hummel H, Wenne R (2014) Glacial history of the European marine mussels *Mytilus*, inferred from distribution of mitochondrial DNA lineages. Heredity 113:250–258.

Sokolova IM, Kovalev A, Timm S, Marchenko J, Sukhotin A (2024) Species-specific metabolome changes during salinity downshift in sub-Arctic populations of *Mytilus edulis* and *M. trossulus* . Frontiers in Marine Science 11:1403774.

Struck TH, Feder JL, Bendiksby M, Birkeland S, Cerca J, Gusarov VI, Kistenich S, Larsson K-H, Liow LH, Nowak MD, others (2018) Finding evolutionary processes hidden in cryptic species. Trends in Ecology & Evolution 33:153–163.

Stuckas H, Knöbel L, Schade H, Breusing C, Hinrichsen HH, Bartel M, Langguth K, Melzner F (2017) Combining hydrodynamic modelling with genetics: can passive larval drift shape the genetic structure of Baltic *Mytilus* populations? Molecular Ecology 26:2765–2782.

Tam JC, Scrosati RA (2014) Distribution of cryptic mussel species (*Mytilus edulis* and *M. trossulus*) along wave exposure gradients on northwest Atlantic rocky shores. Marine Biology Research 10:51–60.

Väinölä R, Strelkov P (2011) *Mytilus trossulus* in northern Europe. Marine Biology 158:817–833.

Wenne R, Zbawicka M, Bach L, Strelkov P, Gantsevich M, Kukliński P, Kijewski T, McDonald JH, Sundsaasen KK, Árnyasi M, Lien S, Kaasik A, Herkül K, Kotta J (2020) Trans-atlantic distribution and introgression as inferred from single nucleotide polymorphism: mussels *Mytilus* and environmental factors. Genes 11:530.

Wiesenthal AA., Timm S, Sokolova IM (2025) Osmotolerance reflected in mitochondrial respiration of Mytilus populations from three different habitat salinities. Marine Environmental Research, 205: 106968.

Wood SN (2017) Generalized additive models: An introduction with R. Chapman; hall/CRC.

Zbawicka M, Sańko T, Strand J, Wenne R (2014) New SNP markers reveal largely concordant clinal variation across the hybrid zone between *Mytilus* spp. in the Baltic Sea. Aquatic Biology 21:25–36.

Zolotarev V, Shurova N (1997) Relations of prismatic and nacreous layers in the shells of the mussel *Mytilus trossulus*. Russian Journal of Marine Biology 23:26–31.

Table 1. Environmental parameters involved in the study

Environmental parameter/ model predictor	Type	Explanation	Range (median) in the data
		Influence of substrate	
Substrate	Categorical	Algal and Bottom samples for each site are treated separately	Algae VS Bottom
		Influence of rivers	
Salinity	Continuous	Surface salinity (ppm) at the time of sampling, i.e. at low tide.	2-30 (19)
DistRiver	Continuous	The straight line distance (km) between the site and the nearest river mouth by map. The values were log-transformed when used for model fitting.	0-18.5 (4.9)
RiverSize	Categorical	Rivers are categorized according to whether their catchment area is larger or smaller than the median area of all rivers in the region.	Small VS Large
		Influence of ports	
DistPort	Continuous	The straight line distance (km) between the site and the nearest port by map. Log-transformed values were used.	0.1-82.2 (18.7)
PortStatus	Categorical	Ports are categorized according to whether they are active or abandoned	Active VS Abandoned
		Influence of surf	
Fetch	Continuous	Unobstructed length of water surface (km) over which wind from a certain direction can blow. Log-transformed values were used.	0.2-28.8 (3.3)

Table 2. Parameters of smoothers and coefficients of parametric terms for Model 1 describing dependency of proportion of M. trossulus in mixed settlements (Ptros) on environmental predictors. edf – effective degrees of freedom; ref.edf - reference effective degrees of freedom.

Smoother terms	edf	ref.edf	Chi square	p-value
s(*Salinity*)	2.4	9	396.7	0.003
s(*DistRiver*)	0.0	9	0.0	0.672
s(*Fetch*)	0.9	9	88.2	0.042
s(*DistPort*)	1.0	9	276.2	0.002
Random effect s(*Site*)	74.4	92	453.6	<0.0001
Parametric terms	Parameter estimate	SE	z-statistic	p-value
(Intercept)	-1.7	0.1	-11.8	<0.0001
$Substrate_{(Algae)}$	0.9	0.1	14.6	<0.0001
$RiverSize_{(Large)}$	0.4	0.2	2.6	0.009
$PortStatus_{(Active)}$	1.0	0.2	5.7	<0.0001

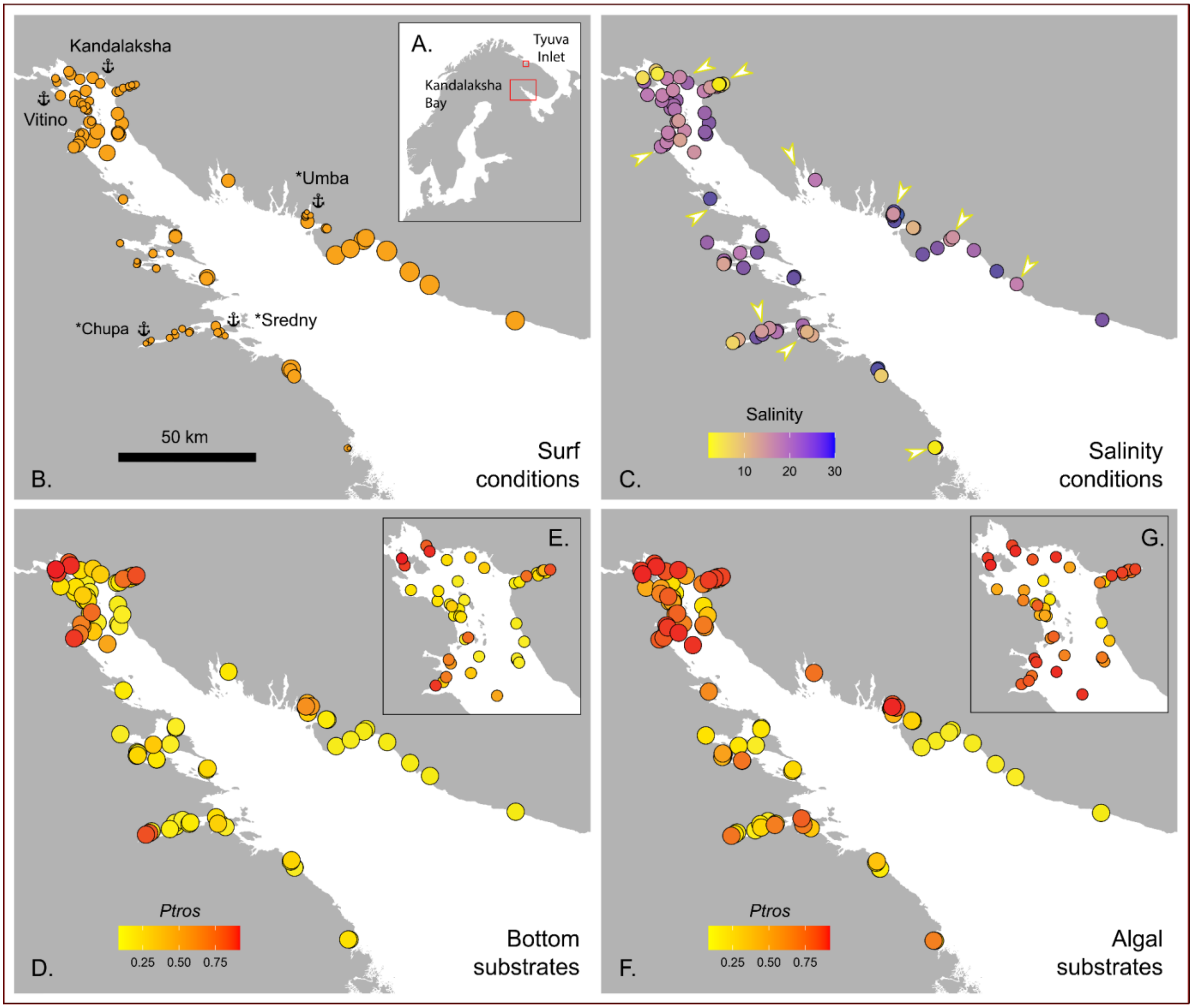
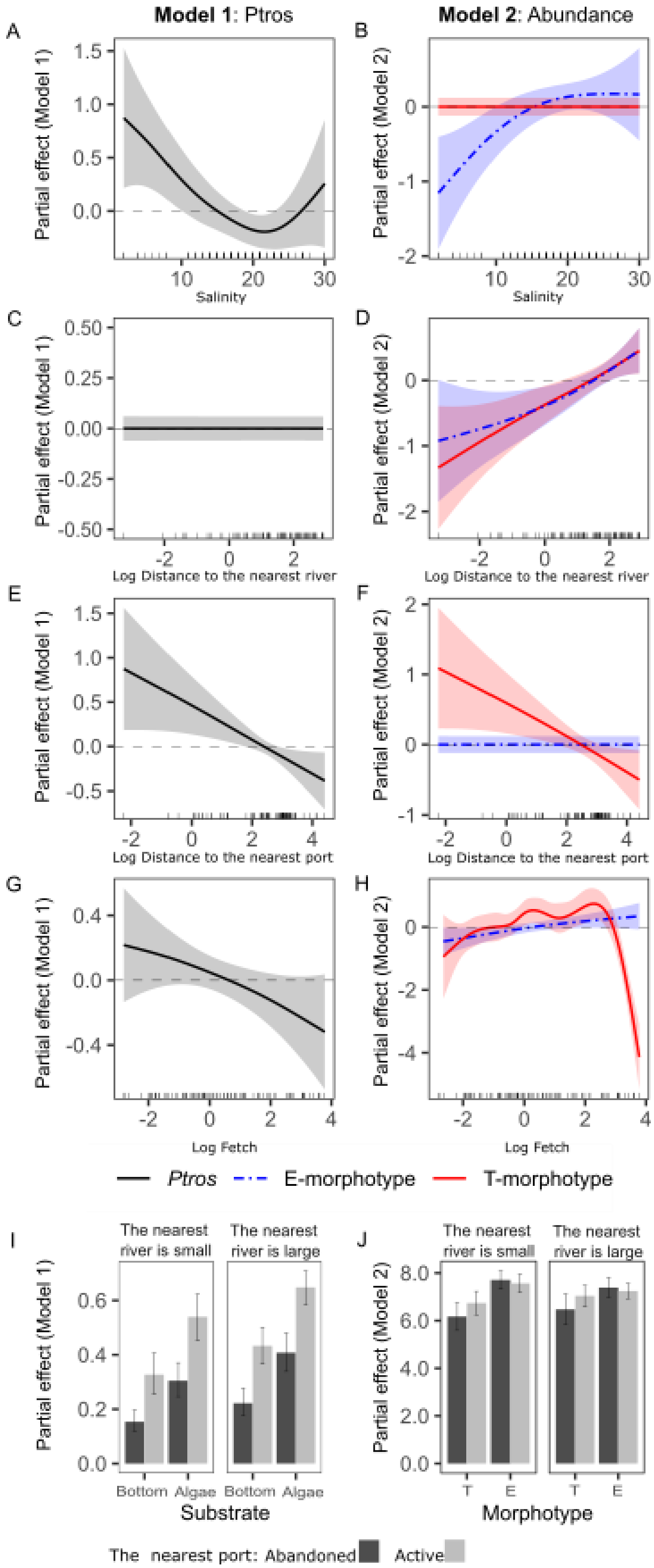


Figure 1. Taxonomic structure of mussel settlements and their habitat characteristics. (A) Map of Northern Europe. Red boxes show the position of Kandalaksha Bay and Tyuva Inlet. Maps on B-G represent the Kandalaksha Bay. Land is grey why the sea is white. B) Surf conditions. Point size is proportional to *Fetch*. ~~Anchors with names mark ports; asterisks mark abandoned ports.~~ Anchors with names mark ports. Asterisks identify whether the port is currently abandoned (C) Salinity conditions. Point filling is proportional to *Salinity*. Arrows mark mouths of large rivers. (D-G) Proportion of *MT* in

bottom ($Ptros_{Bottom}$, D-E) and algal ($Ptros_{Algae}$, F-G) samples. Point filling is proportional to *Ptros*. E and G show the Bay's top in higher resolution.

841



Model 1: Ptros
Model 2: Abundance
A
B
C
D
E
F
G
H
Partial effect (Model 1)
Partial effect (Model 2)
Salinity
Log Distance to the nearest river
Log Distance to the nearest port
Log Fetch
Ptros
E-morphotype
T-morphotype
I
J
The nearest river is small
The nearest river is large
Bottom
Algae
Substrate
T
E
Morphotype
The nearest port: Abandoned
Active

~~**Figure 2**. Partial effects of environmental parameters on proportion of *M. trossulus* in mixed settlements (*Ptros*) evaluated by the GAM fitted (**Model 1**). A-C. Dependency of *Ptros* on distance to the nearest port (*DistPort*, A), wind exposure (*Fetch*, B), distance to the nearest river (*DistRiver*, C) and salinity at low tide (*Salinity*, D). Gray ribbons represent 95% confidence intervals. Dotted horizontal lines indicating zero partial effect are given to show the wiggling of the fitted curves. Points on panels A-D show partial residuals, not raw data. E-F. Dependency of *Ptros* on combinations of categorical predictors. Partial effects of *Substrate* (bottom vs algae) and status of the nearest port (active vs abandoned) when the nearest river is small (E) or large (F). Whiskers represent 95% confidence intervals. Red solid lines indicating a partial effect of 0.5 are provided to facilitate visual comparison of panels E and F.~~

Figure 2. Partial effects of environmental parameters of either proportion of *M. trossulus* in mixed settlements (*Ptros*) or abundance of species-specific morphotypes evaluated by GAM(M)s fitted (**Model 1** and **Model 2** respectively). Dependency of *Ptros* on salinity at low tide (*Salinity*, A), distance to the nearest river (*DistRiver*, **C**), distance to the nearest port (*DistPort*, **E**), wind exposure (*Fetch*, **G**), and discrete predictors: substrate type, nearest port status and size of the nearest river (**I**) . Dependency of E- and T-morphotype abundances on *Salinity* (**B**), *DistRiver* (**D**), *DistPort* (**F**), *Fetch* (**H**), and nearest port status and size of the nearest river (**J**). Ribbons around curves and whiskers represent 95% confidence intervals.

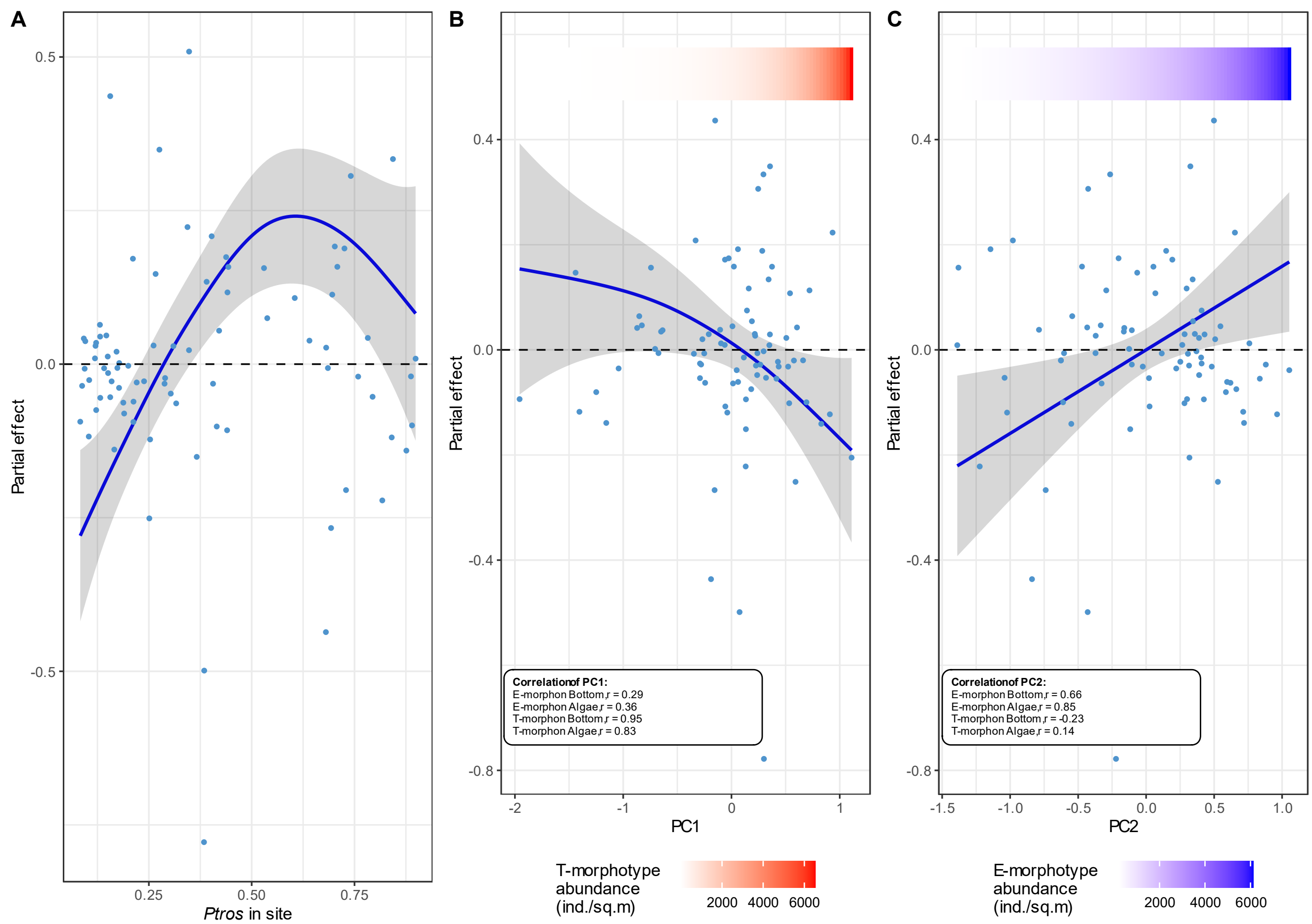


Figure 3. The dependence of difference between proportion of *MT* on algal and bottom substrates (*Diff*) on proportion of *MT* in a site (*Ptros*$_{Site}$) (A) and estimations of total abundance of *MT* (B) and *ME* ~~(D)~~ (C). Principal components from the matrix of T- and E-morphotypes abundances on different substrates are considered as proxies for *MT* and *ME* abundances (PC1 and PC2, respectively). ~~Points reflect partial residuals, not raw data.~~ ~~Gray ribbons~~ Ribbons around curves represent 95% confidence intervals. Colored gradient bars at the top of the figures reflect linear associations between PC1 and T-morphotype (B) and PC2 and E-morphotype abundance (C).

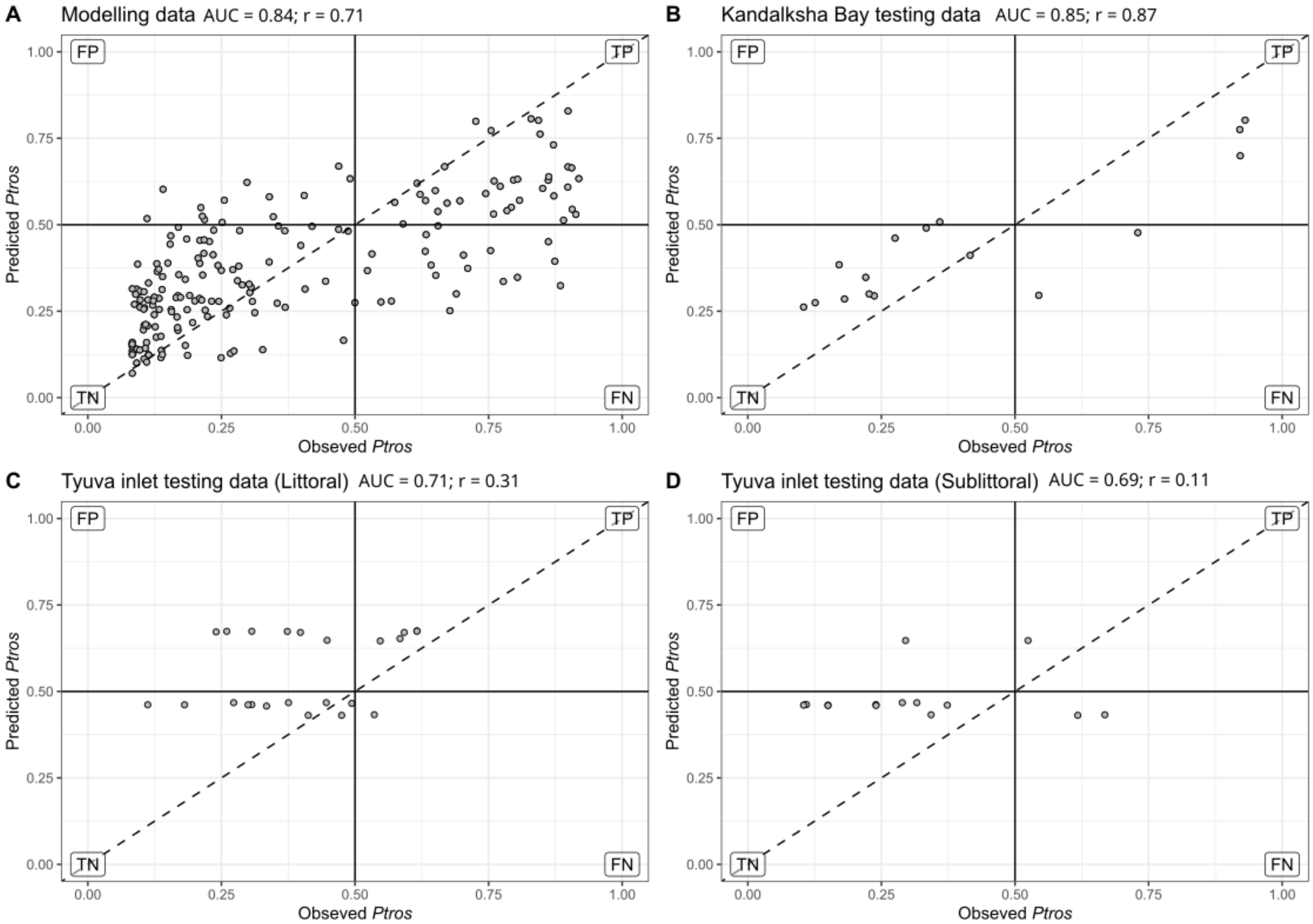


Figure 4. Ability of SDM (**Model 1**) to predict proportion of *MT* (*Ptros*) in mussel samples from the modeling (A) and the testing data sets (B-D). Each plot compares empirical *Ptros* in samples from algal and bottom substrates and *Ptros* predicted by the model within the particular data set. If the empirical and the predicted values were the same, the points would lie on the diagonal (dashed line). ~~Solid lines delineate *MT*- and *ME*-dominated samples on each axis.~~ Labels mark the quadrants with false positive (FP), true positive (TP), true negative (TN) and false negative (FN) predictions in the analysis of the ability of the model to classify samples into *ME*-dominated (Ptros < 0.5) and *MT*-dominated (*Ptros* >0.5) ones. Dataset names are shown in chart headers. Values of AUC for binary classification (*ME*-

dominated vs *MT*-dominated) and Pearson correlations between observed and predicted Ptros are given in figure headings.