Dear Editor,

We are very grateful to you for the favorable assessment of our work and to the reviewers for their valuable advice. Below we provide the responses to the reviewers. References in parentheses are to lines in the text with labeled changes.

Editorial Office comments:  
  
**Please also address the following formatting and style points in your revision:  
(i) Please add the section number for Results and Discussion.**

Done.

**(ii) L414-417: Format the text to be similar with the rest of the manuscript.**

Done

**(iii) Spell out the species name at first mention in the figure/table legends.**

Done

**(iv) Fig. 1: Please add lat/lon coordinates to the left and bottom of the figure (it is enough to have one panel with lat/lon on the frame).**

Done (L 742)  
  
  
-------------------------  
Reviewer 1 report:  
  
Review of Kaitov et al 2024/2025  
  
  
The authors have an impressive dataset that yields insights into the niches of Mytilus trossulus and Mytils edulis. While the data and analysis should be published, I feel the manuscript still needs some revision to present the results in a way that is more coherent and approachable to the reader.  
  
**Line 45 – “stand behind” is an unusual phrase. Recommend they replace it.**

We replaced the this phrase. The sentence now reads:

“Sympatry usually arises as a result of cryptic invasions (Morais & Reichard 2017).” (L. 46)  
  
**Line 60 - Recommend adding “of “ to “Both of these species..**

Done  
  
**Line 76 – “New Scotland”!?!? do they mean “Nova Scotia”.**

Corrected (L.77)  
  
**Lines 94-95 – Is MT more likely to settle on algae? Are they referring to juveniles or adults here?**

We have changed the sentence for clarity.

Now:

”Adult MT are more common on fucoid algae while adult ME mostly live directly on the bottom substrates such as mud, sand, stones and gravel. However, segregation across substrates cannot fully explain the local-scale patchiness (Katolikova et al. 2016).” (L 95-98)

**Lines 179-180 – Maybe I missed it but I feel PtrosBottom and PtrosSite should be explained before they are referenced here.**

To clarify this, we have changed the sentence. Now:

“The proportion of morphotypes was converted to the proportion of MT (Ptros) in each sample, in pooled samples from each substrate from each site (denoted as PtrosAlgae and PtrosBottom) and in pooled samples from each site (PtrosSite)”. (L.178-180).

**Lines 145-148 – Is this information relevant? Delete, move to discussion or explain relevance here.**

We are quite sure that this information is relevant because one of the tasks of this study is to assess the impact of ports (see Introduction). Therefore, we mention them in the Materials and Methods section and show them on the map.

**Line 298 – The allusion to Table 2 to explain that salinity is curvilinear needs additional explanation. (Because the edf = 2.4?)**

We have added some information to the sentence to eliminate uncertainty. Now:

“Effective degrees of freedom (edf) for DistPort and Fetch were close to one, indicating the linear dependence of Ptros on them. In contrast, the dependence on the third continuous predictor, Salinity, was curvilinear (edf > 2, Table 2).” (L.298-301)  
  
**Line 309 – “complete agreement” seems a bit overselling of how well the data fits. Qualify this a bit.**

To qualify this, we changed the sentence. Now:

“The results of Model 2 were in good agreement with those of Model 1 for all the predictors” (L 312-313)

**Lines 328-332 – These relationships need to be better explained in the Results text. The reader still must work pretty hard to make sense of these patterns.**

We added some words to explain this.

This means that the strongest segregation of the two species by substrate was observed in settlements where they occurred in equal proportions. (L 334-335)

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 (to remind, C1 and PC2 are proxies for *MT* and *ME* abundance, respectively). (L 340-342)

**Lines 383 and 390 – I feel these sections could be clearer if the authors started each section by restating the pattern.**

We have followed your advice. The sections now start as follows:

Segregation of ME and MT by surf level (ME is more common in localities exposed to surf, while MT is more common in sheltered areas) may be due to the well-known differences in the mechanical properties of their shells and the ability to form dense aggregations. (L 395-398)

Segregation by substrate (*ME* is more common on bottom substrates, while *MT* is more common on algal ones) may be explained by the same differences as segregation by surf level. (L 405-406)

**Line 401 – “In our opinion” is used here and elsewhere. The phrase is unnecessary and creates ambiguity. Are the patterns real or imagined? If MT is being displaced onto algae they should have the data to show it.**

In this particular paragraph, we would like to leave this phrase as it is because, even though the pattern is real and proven, our explanation of it is speculative. In line 192 and in lines 513-514, we have removed “in our opinion” following your recommendations.

**Lines 403-408 – Connect these statements to the thesis of the paper. They seem almost irrelevant as currently presented.**

We have rephrased these sentences to highlight the applicability of the biologically generated spatial pattern model to competing mussel species. We hope it is now clear that the statements are relevant to the topic.

**Line 720 –Somewhere explain the need/use of edf and ref-edf. Why are they included? (I may be asking to explain something obvious, but I feel details need to be accessible.)**

This is a standard summary for GAM analysis. We have explained edf in the caption to Table 2, which now reads:

Table 2. Parameters of smoothers and coefficients of parametric terms for Model 1 describing dependency of proportion of *Mytilus trossulus* in mixed settlements (Ptros) on environmental predictors. Smoother’s summary: edf – effective degrees of freedom; ref.edf – reference effective degrees of freedom. (L. 735 - 738)

-------------------------  
Reviewer 2 report:  
  
The authors have done a nice job addressing comments from Reviewers. I have a few additional suggestions. Line numbers relate to the track changes version.  
  
**Line 61, instead of aquaculture objects use something like ‘species for aquaculture’**

OK. Now:

Blue mussels are powerful ecosystem engineers and important species for aquaculture (Buschbaum et al. 2009, Gosling 2021). (L 57-58)

**Line 69, since THE Pliocene**

Done (L 65)

Line 91, replace pattern with driver

Done (L 86)

**Line 416, could more concrete information be provided for what inconclusive means? Is it that MT and ME had differential responses? Is it that salinity sometimes affected both species but not always? Do the authors mean the ecophysiological responses of MT and ME to salinity can be context-dependent?**

We actually meant the scarcity of direct ecophysiological data. We have changed the sentence, which now reads:

Comparative ecophysiological data on *MT* and *ME* elsewhere are scarce (Gardner & Thompson 2001, Qiu et al. 2002, Sokolova et al. 2024). (L. 366)

**Line 474, I think the wording here is off. The authors might mean “Our model was able to successfully classify sites into ME- and MT-dominated ones….**

No, we meant that our model demonstrated a high accuracy in sample classification. Within the ROC analysis framework, an AUC of 0.85 is considered indicative of strong model performance. (see for example https://www.evidentlyai.com/classification-metrics/explain-roc-curve).

**Line 481, the worst what? The wording here is off**

Thanks! We changed the text. Now:

“The model's predictive performance for the Barents Sea data (AUC ≈ 0.7), while formally acceptable, was lower than for the White Sea. This may be due to the following reasons.” (L. 431-433)

**Line 489, variation of what? Values?**Thanks! We changed text. Now:

The fourth reason could be a narrow variation in the values of *DistPort*, *DistRiver*, and *Fetch* in the small Tyuva Inlet in comparison with the Kandalaksha Bay. (L 439-440)

-------------------------  
Reviewer 3 report:  
  
**Thank you to the authors for their detailed response to the comments. However, I remain concerned about the revised model (Model 2), which uses a Gaussian distribution. This choice implicitly assumes normality and homoscedasticity of the residuals—assumptions that have not been clearly demonstrated. At minimum, a diagnostic check such as a QQ plot of the residuals and residuals vs linear predictor plots should be provided to support this modeling decision.**

You rightly point out the issues associated with using Gaussian-based models for count data. However, prior to model construction, we log-transformed the data (see Section 2.4.2, line 247), which circumvented these problems. Now we also mention in the legend to Figure 2. (L. 758)

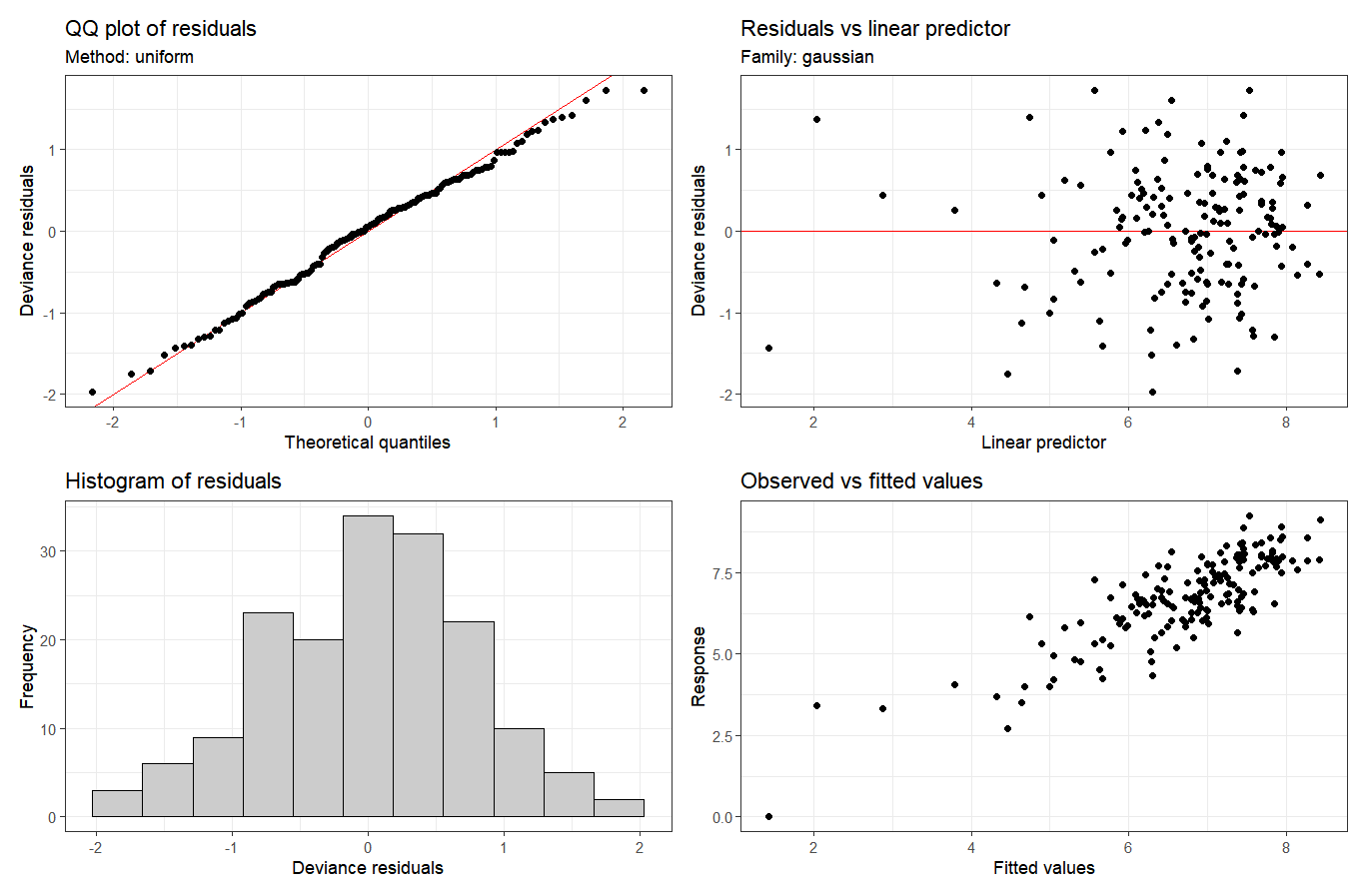
**Moreover, the use of a normal distribution to model count or abundance data is generally inappropriate, as such data are typically non-negative integers and often exhibit skewness or overdispersion. Distributions such as the Poisson or negative binomial are more suitable and widely accepted for this type of response variable.**

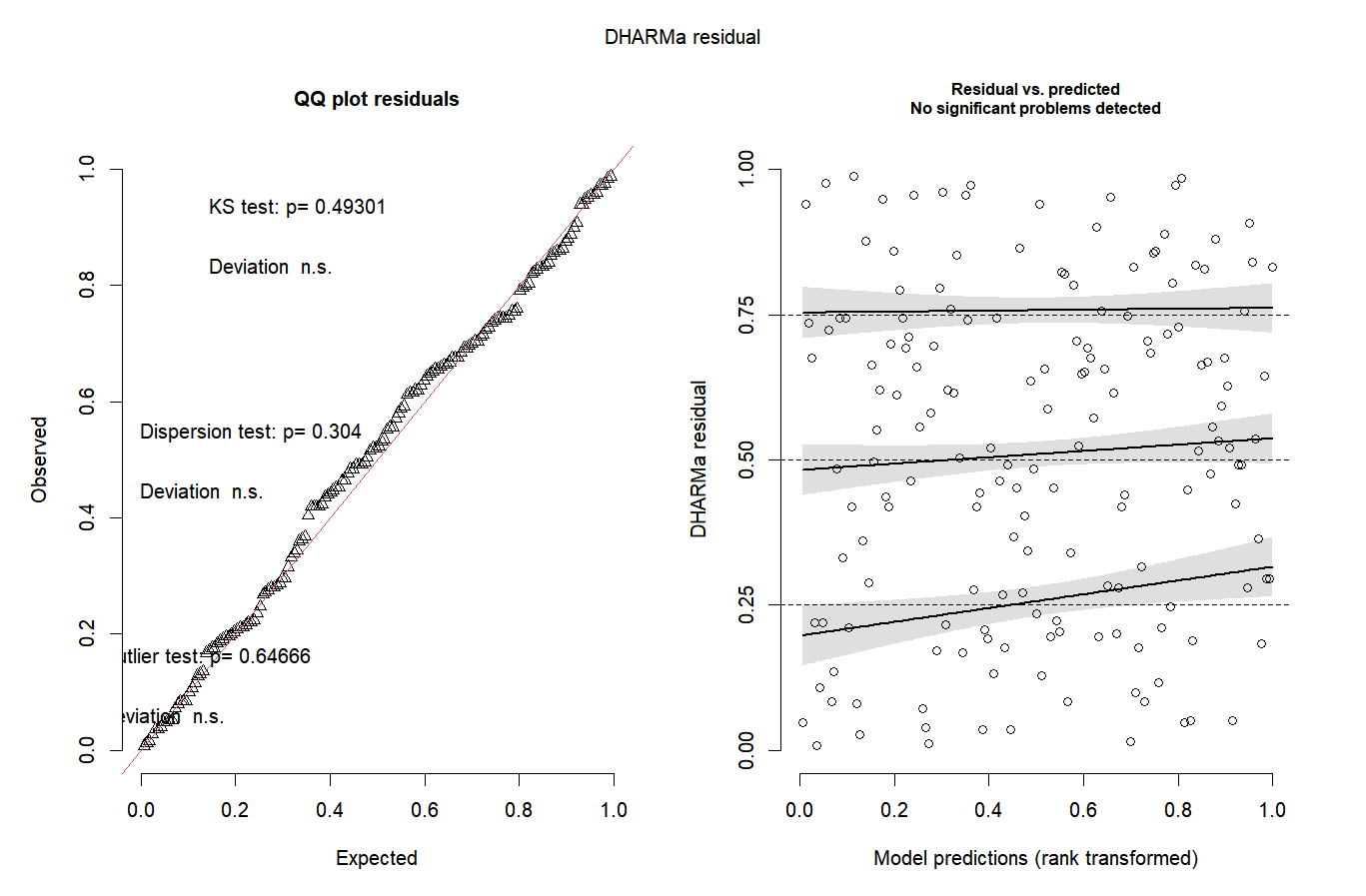
We did, of course, validate Model 2 by examining residual plots. We have added some sentences.

To confirm that **Model 2** satisfies the assumptions of regression analysis, we inspected the residual plots. No discernible patterns were detected in the residuals.

(L. 250-252)

As you requested, we have attached the residual diagnostics to this response letter.





**Lastly, it is surprising that the factor “substrate” is no longer included in Model 2. As noted by the authors in their response, abundance data were collected using different methods depending on the substrate type. This raises concerns about comparability between algal and bare-bottom communities. Including substrate as a random effect could help account for variation introduced by differing sampling approaches and improve the robustness of the model.**

As we mentioned in our previous letter, we couldn't separate the natural differences in settlement organization between algae and bottom from the methodological differences in sampling (these technical differences were unavoidable due to the distinct spatial geometries of the two substrate types). Differences in sampling techniques are not fundamentally important when using Ptros as the dependent variable but crucial when using abundance.

We emphasize that Model 2 is fundamentally different from Model 1. Unlike Model 1, which was applied to the raw data, Model 2 was specifically designed for site-averaged data. This fundamental difference in the approach necessitated consideration of different predictor variables, and we included into the set only the predictors varying between sites (i.e. not Substrate).

We believe that including "Substrate" as a random effect would be inappropriate for these reasons:

1. The "Substrate" factor is an inherently fixed effect. It encompasses all possible substrate types in our study (Algae vs Bottom). Random effects imply the existence of some unmeasured factor levels not present in our data.
2. Had we treated “Substrate” as a random effect, we would have been able to estimate only its associated variance from two levels. This is considered insufficient for random effects, which typically require more than five levels to be reliably estimated (see e.g. Bolker et al., 2009, 'Generalized linear mixed models: a practical guide for ecology and evolution').