# Appendix A - Data Sources and Descriptions of the Datasets



Figure 1: Map of the sampled sites. Point shapes vary according to their contribution to model training set (circles ; used to evaluate model explanatory power) as opposed to the two sites retained for independent model testing (triangles ; used to evaluate model predictive power). Point colours vary according to the presence or absence of the two habitats in each site. The two test sites include the two habitats (i.e. seagrass and bare sand) and were chosen because they occur in environmental conditions that can be considered average at the scale of the region (thereby limiting extrapolation of the model) but still harbour different communities, representative of the known diversity gradient across the region.

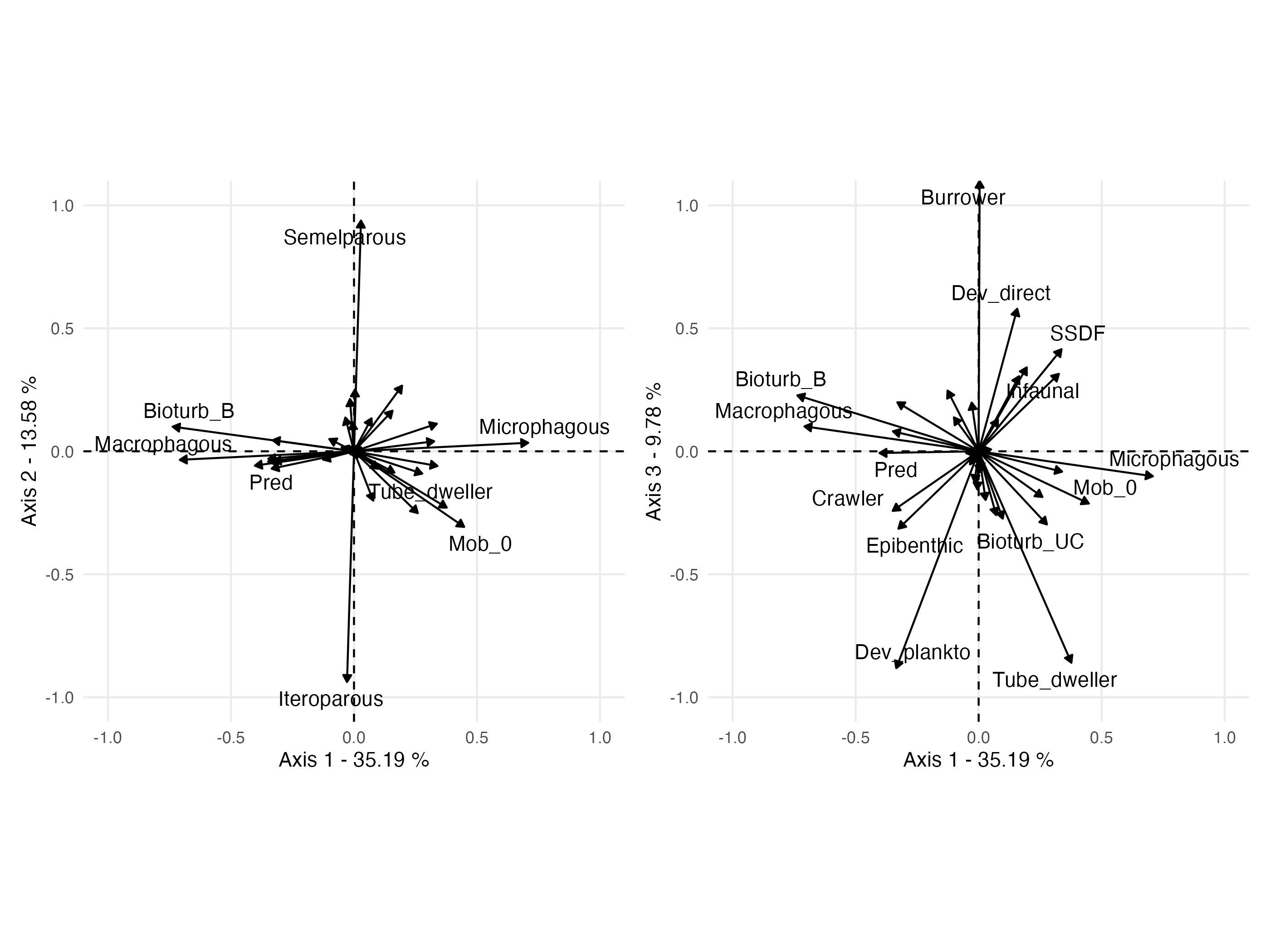


Figure 2: Fuzzy PCA of the species-by-trait matrix. The first three axes represent 58.55% of the total variance. The first axis distinguishes sessile microphagous species (top positive values) from mobile macrophagous predatory species (bottom negative values). The second axis is a gradient of reproductive strategies (semelparous vs iteroparous). The third axis distinguishes burrowers with direct development from tube-dwellers with planktonic development. For abbreviations and meaning of the trait modalities, see Boyé *et al.* ([2019](#ref-Boye_2019a)).

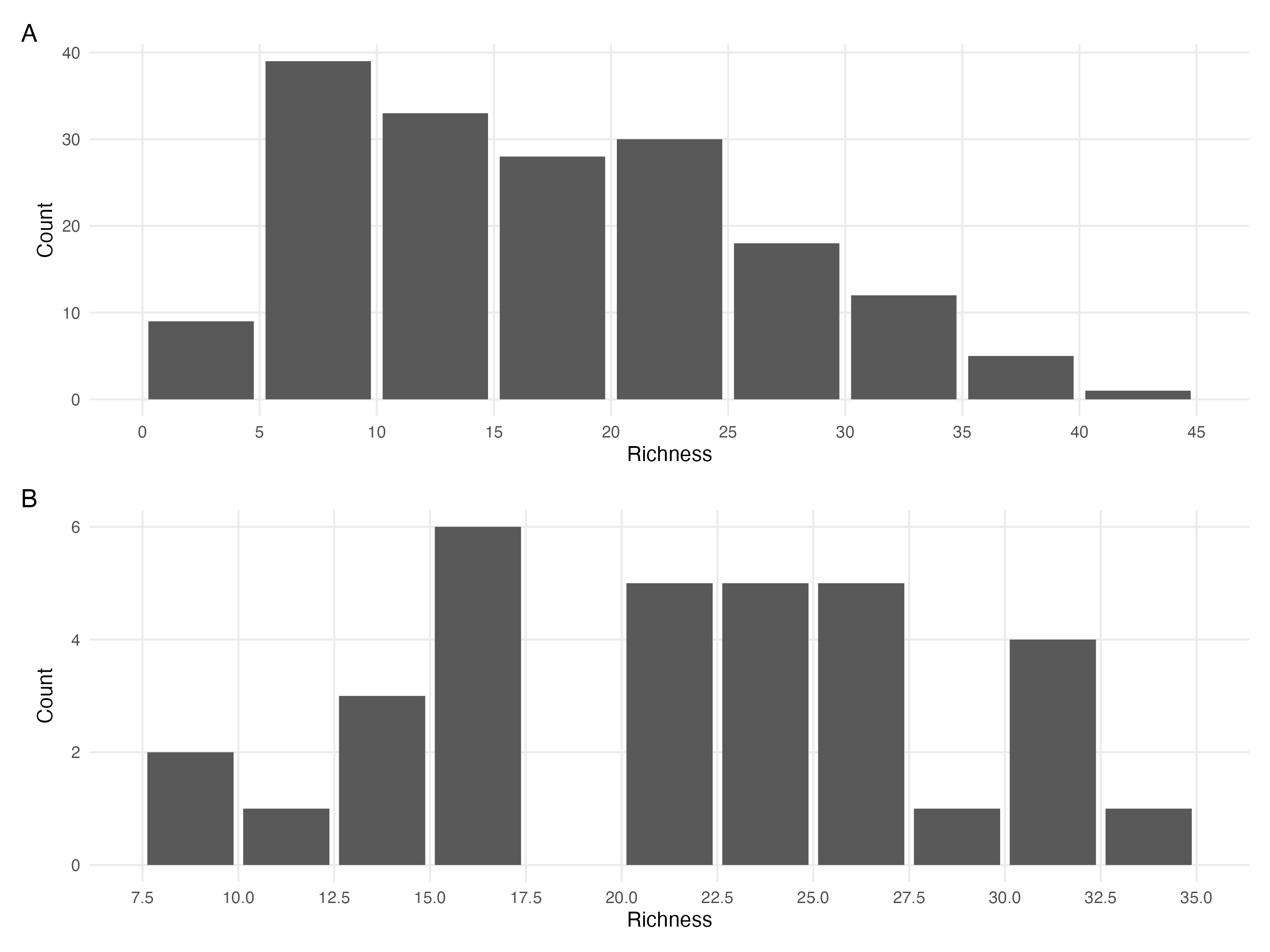


Figure 3: A. Distribution of the number of species in the samples (site times habitat times year) of the train dataset. B. Distribution of the number of species in the samples (site times habitat times year) of the test dataset.

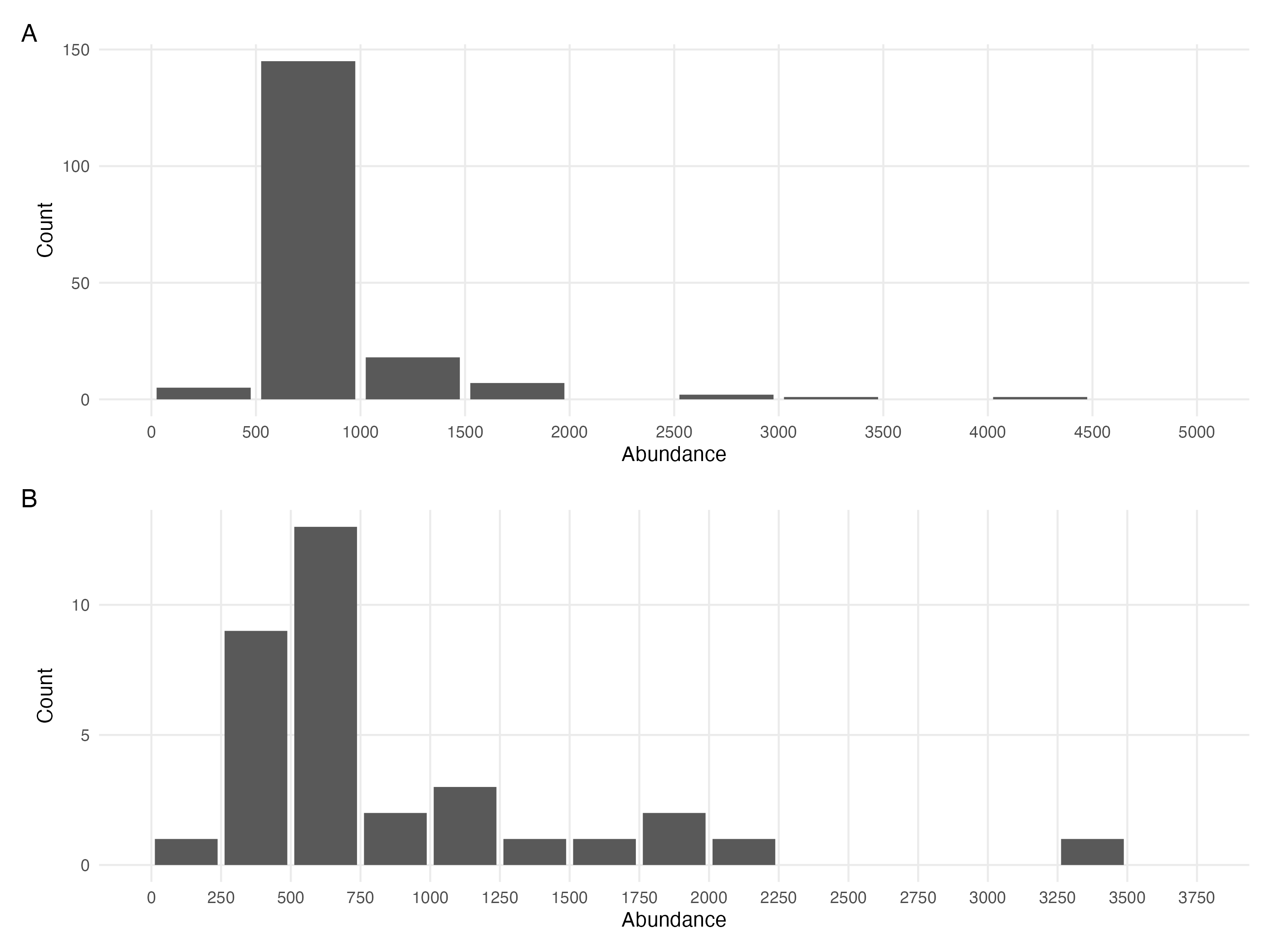


Figure 4: A. Distribution of the total number of individuals (abundance) in the samples (site times habitat times year) of the train dataset. B. Distribution of the total number of individuals (abundance) in the samples (site times habitat times year) of the test dataset.

## Environmental data acquisition

The models used in this study incorporate a dataset consisting of seven environmental variables related to oceanography, hydrography, and granulometry obtenained from Boyé ([2019](#ref-Boye_2019b)) The oceanographic variables include the standard deviation of salinity, surface water temperature, mean velocity of currents, and fetch, which were obtained from the PREVIMER database ([Lecornu & Roeck 2009](#ref-Lecornu_2009)) based on the MARS3D model ([Lazure & Dumas 2008](#ref-Lazure_2008)). The variables were averaged by extracting daily data for the sampled year at the site coordinates and the eight adjacent cells. The fetch was calculated as the average length of nine radiating fetch segments with a maximum distance of 300km. The granulometry variables were derived from sediment cores that were taken along with associated fauna. The cores were dried, separated into 15 fractions, and the Trask index was calculated as the ratio of the 25th to 75th percentile of the grain distribution. Organic matter mass was estimated through the loss of mass after combustion in an oven.

# Appendix B - Model Convergence

## Environmental coefficients

Table 1: Potential scale reduction factors (PSRF) and effective sample sizes (ESS) for environmental regression parameters (i.e beta coefficients) estimated estimated for the four different models (Bench, Ph, TrPh, WhC) fitted either to abundance or presence-absence data. For further details see Fig. S5 to Fig. S12.

| Model | Data Type | Number of coefficients | PSRF (mean sd) | ESS (mean sd) |
| --- | --- | --- | --- | --- |
| Benchmark | Abundance | 1485 | 1.18 0.267 | 701 576 |
| Benchmark | Presence/Absence | 1485 | 1.00 0.002 | 4967 417 |
| Phylogeny | Abundance | 1485 | 1.18 0.204 | 566 420 |
| Phylogeny | Presence/Absence | 1485 | 1.00 0.001 | 4947 408 |
| Traits & Phylogeny | Abundance | 1485 | 1.21 0.317 | 489 358 |
| Traits & Phylogeny | Presence/Absence | 1485 | 1.00 0.008 | 11459 2649 |
| Whole Community | Abundance | 4170 | 1.21 0.287 | 739 631 |
| Whole Community | Presence/Absence | 4170 | 1.00 0.002 | 4962 406 |

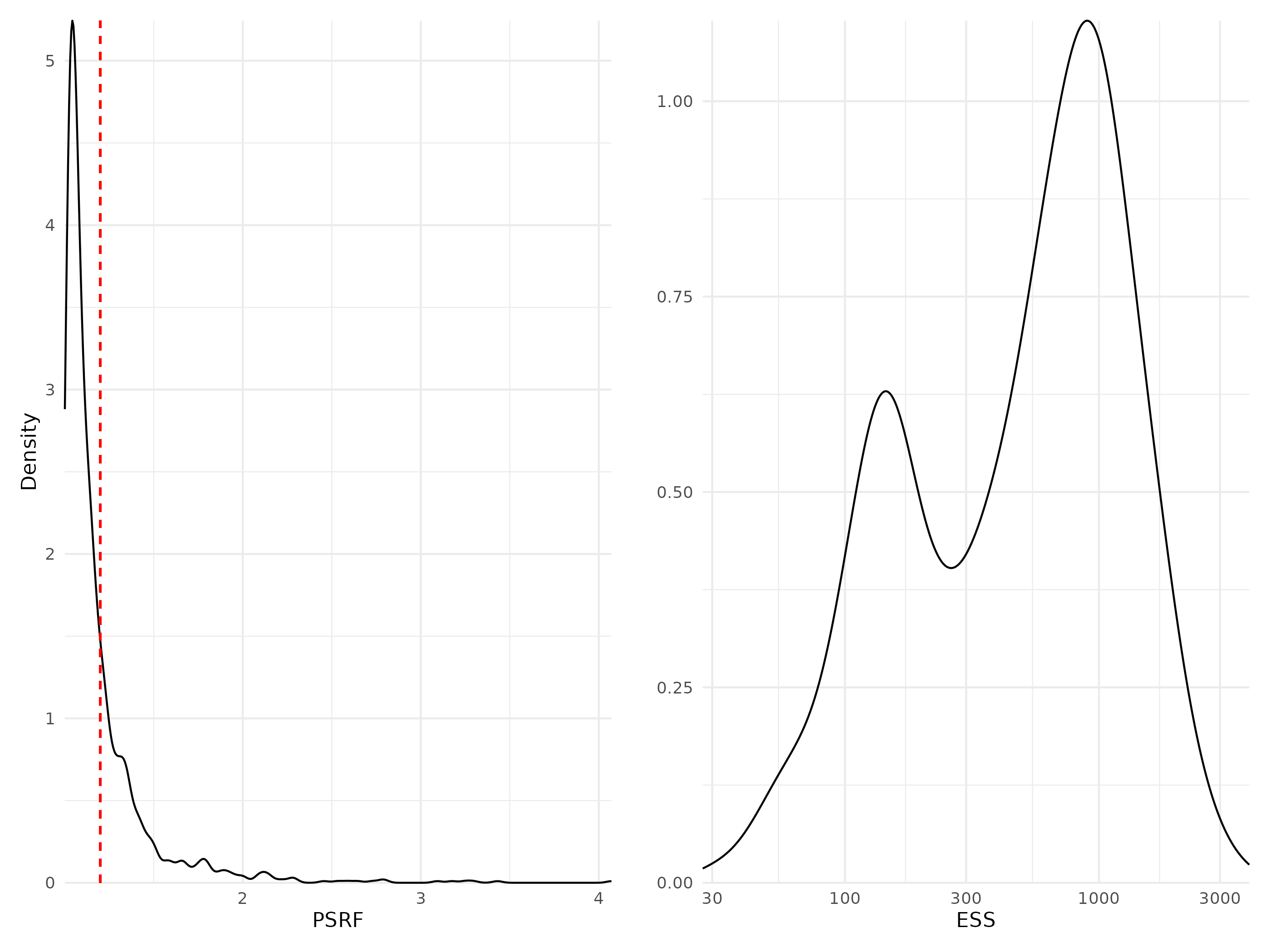


Figure 5: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the benchmark model fitted with abundance data. For PSRF, values greater than 1.2 (dotted red line) indicate potential convergence issues. ESS estimates the number of independent samples used to estimate each parameter (the more the better).

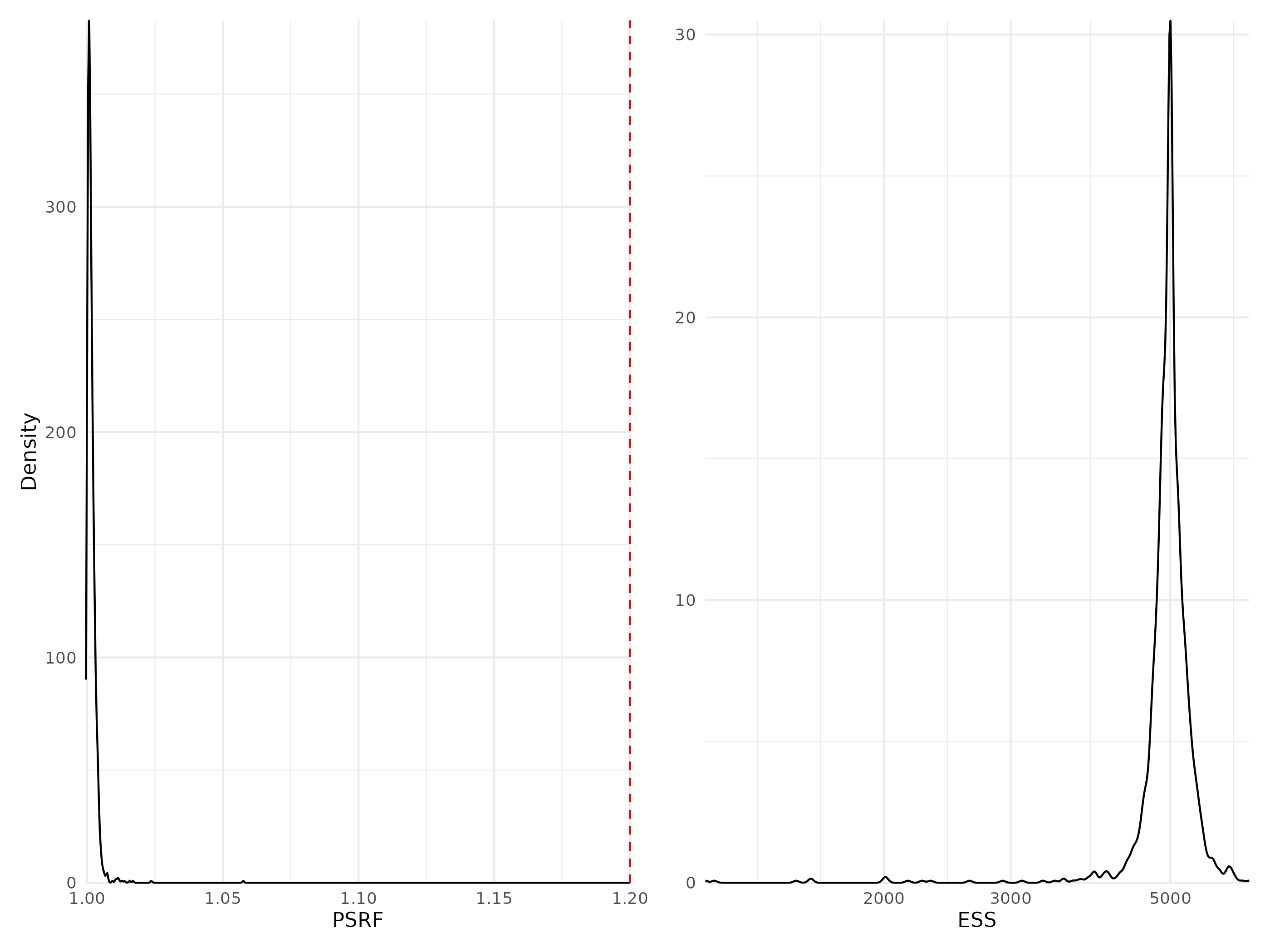


Figure 6: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the benchmark model fitted with presence/absence data. For further details see Fig. S5.



Figure 7: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the phylogeny model fitted with abundance data. For further details see Fig. S5.

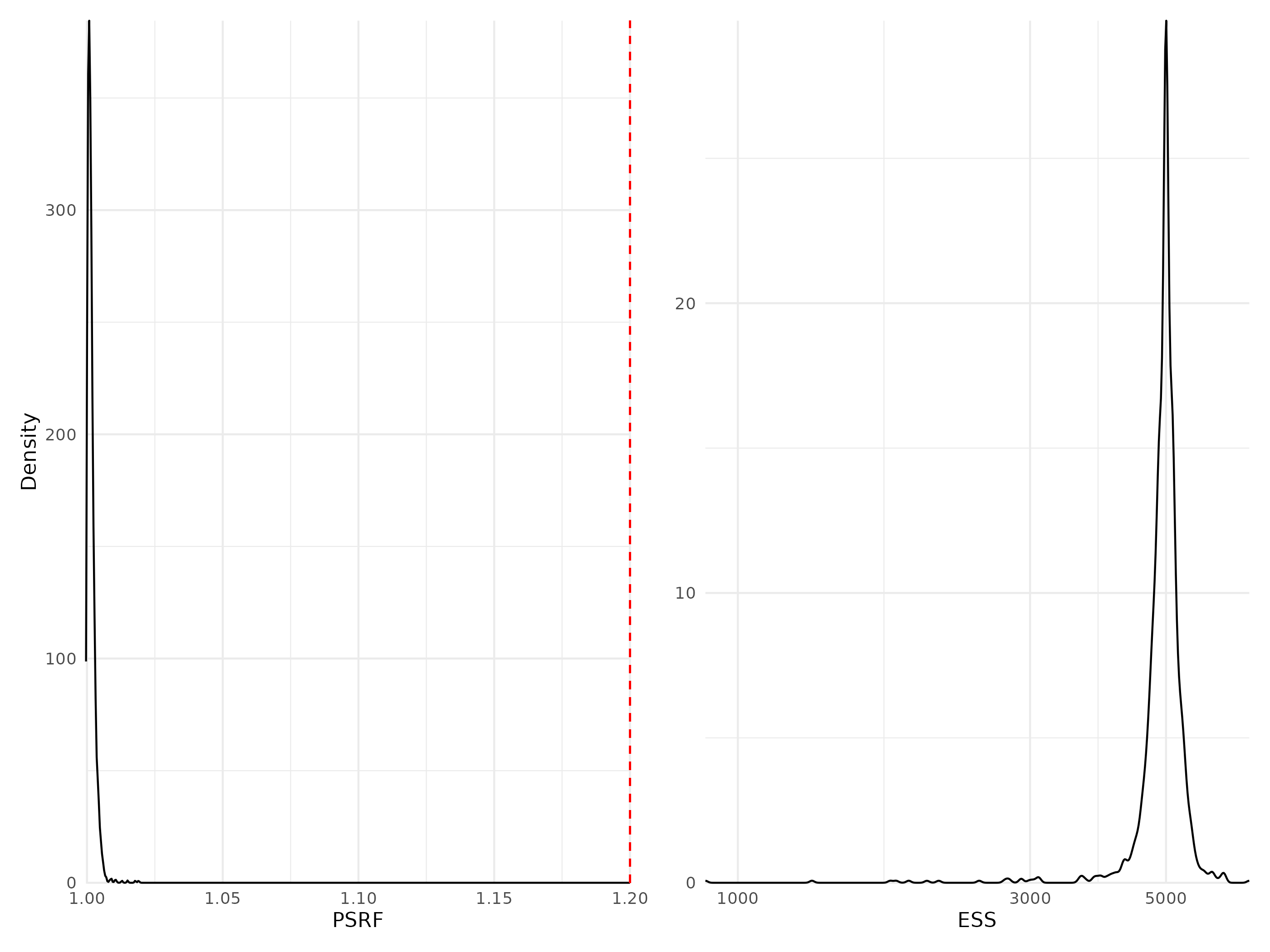


Figure 8: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the phylogeny model fitted with presence/absence data. For further details see Fig. S5.

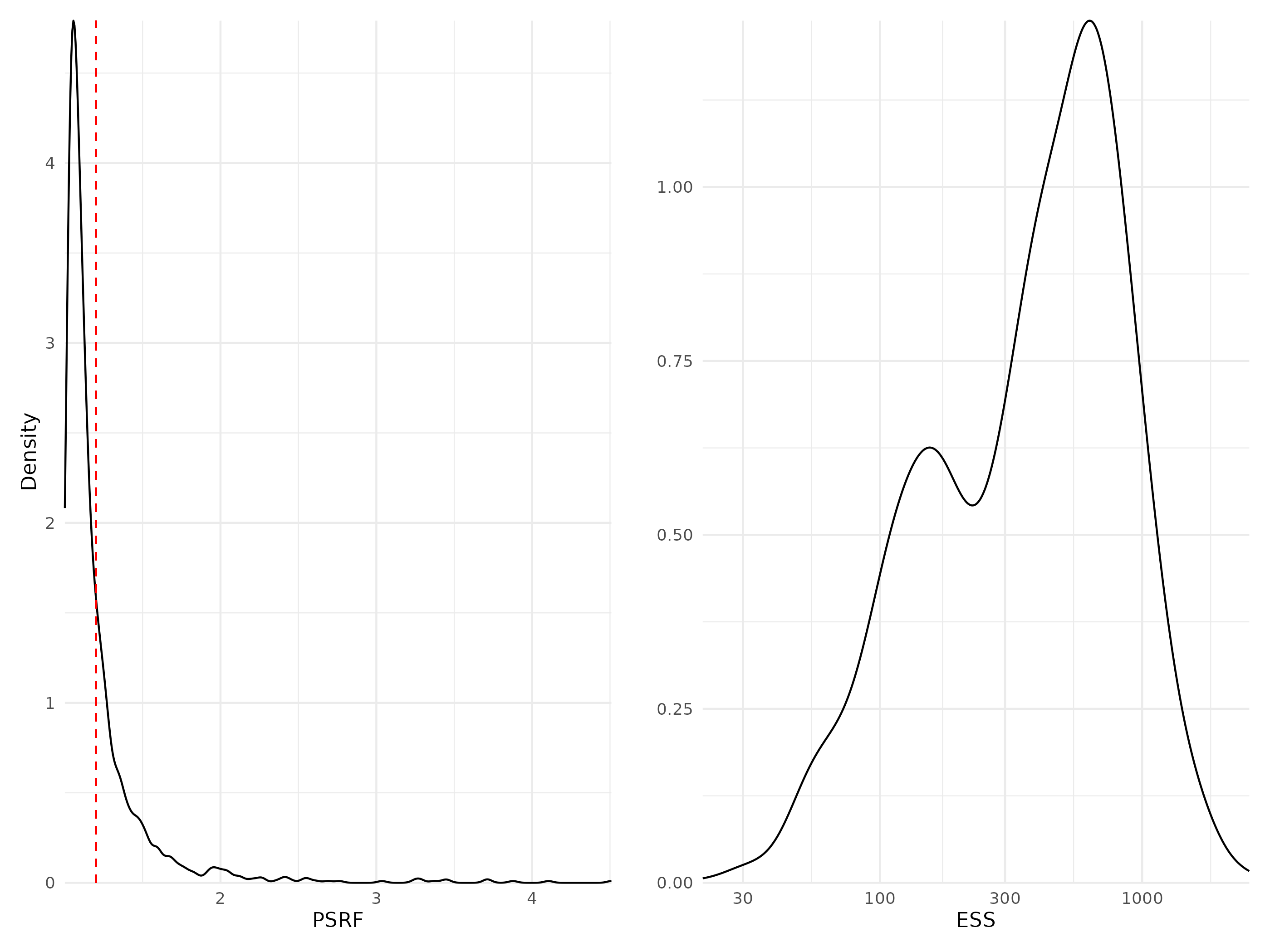


Figure 9: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the traits & phylogeny model fitted with abundance data. For further details see Fig. S5.

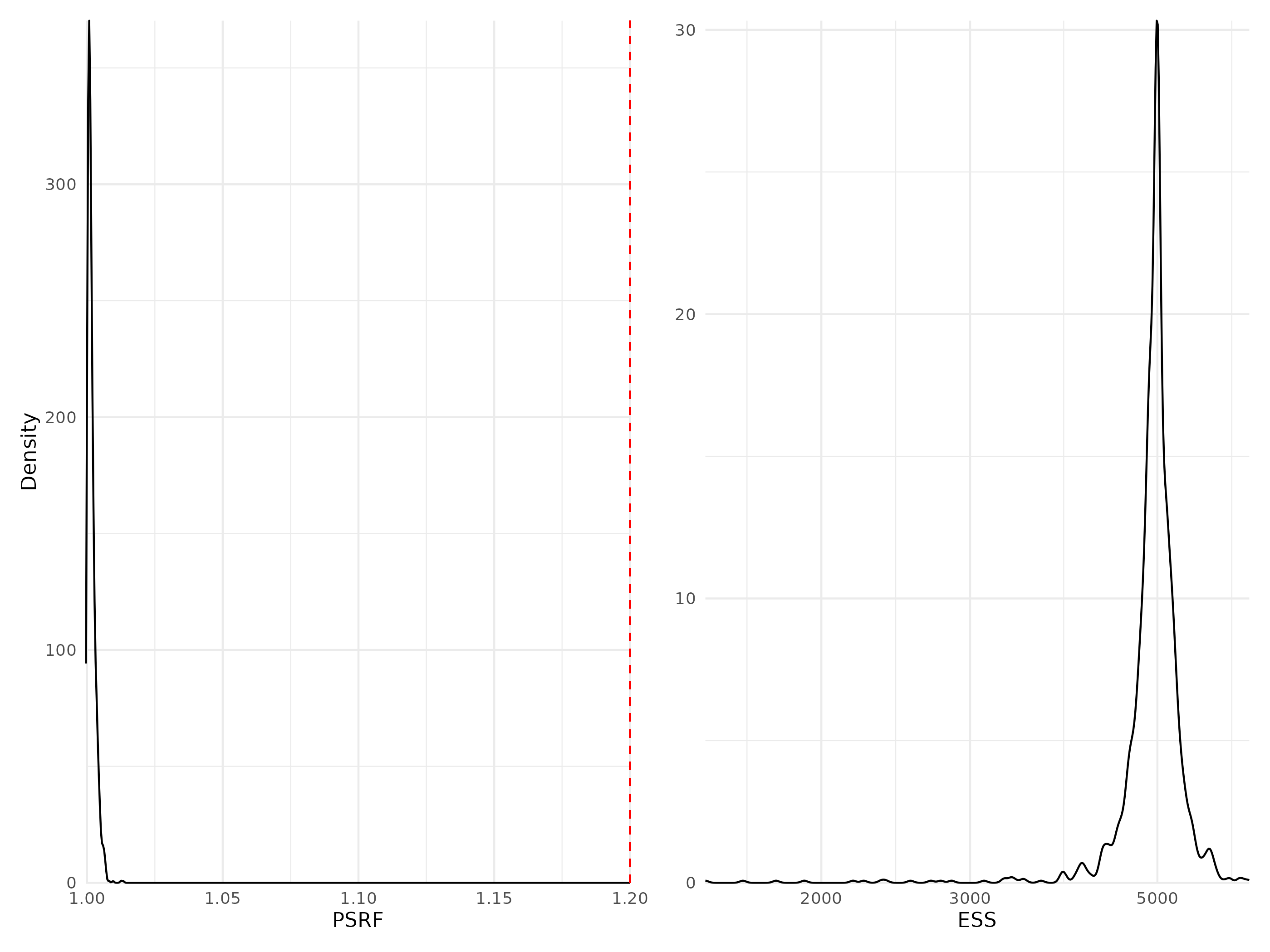


Figure 10: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the traits & phylogeny model fitted with presence/absence data. For further details see Fig. S5.

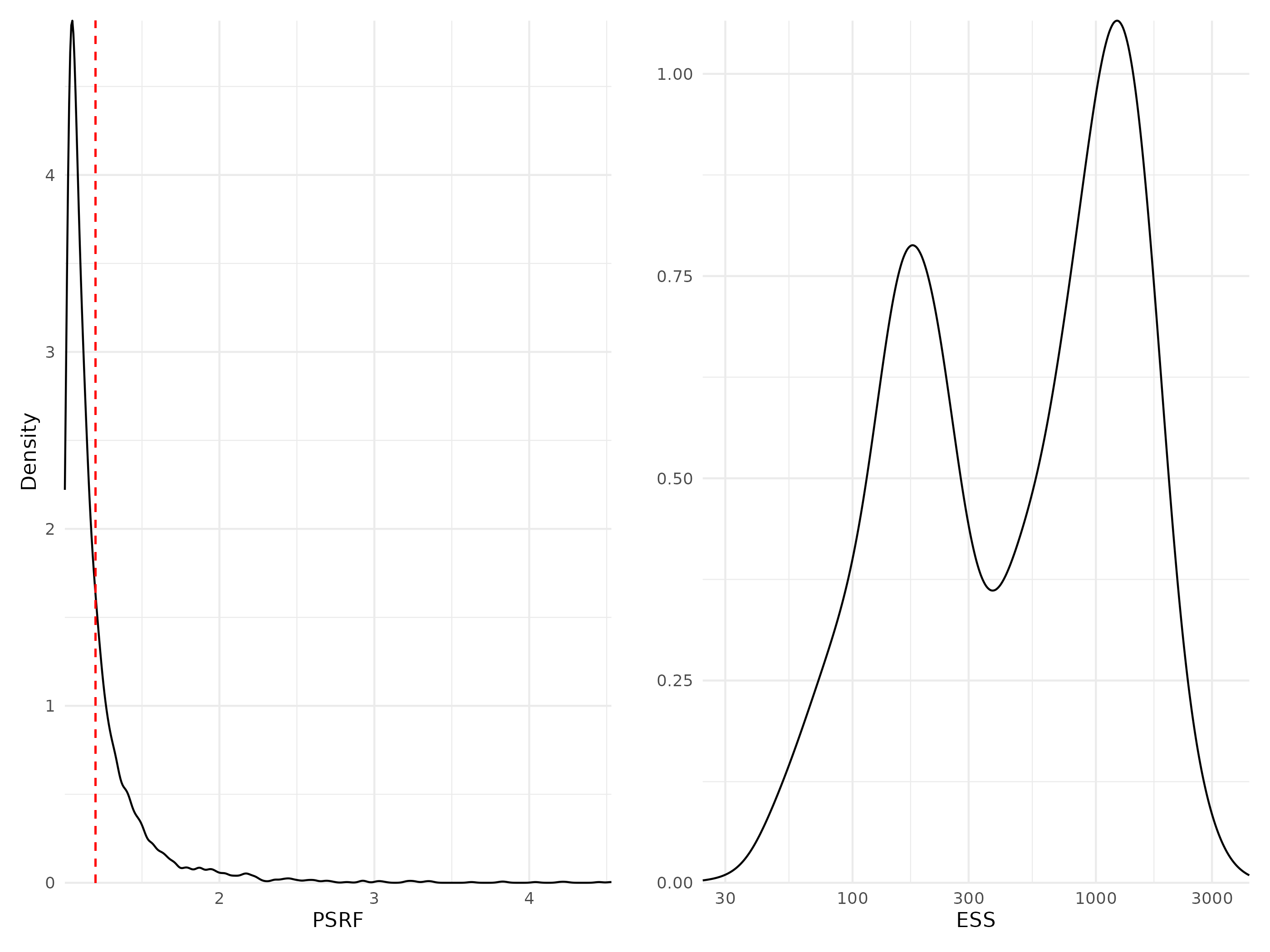


Figure 11: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the whole community model fitted with abundance data. For further details see Fig. S5.

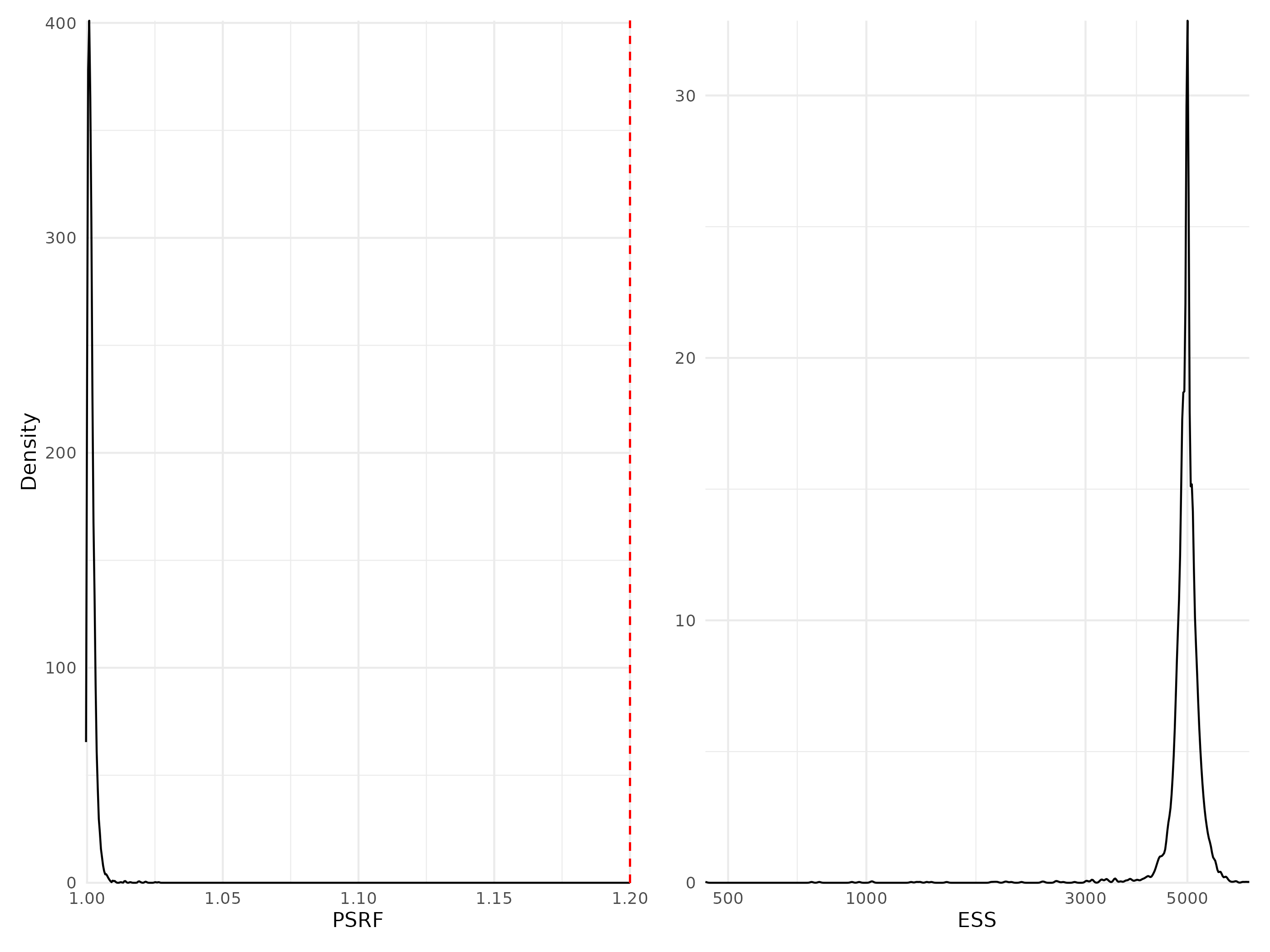


Figure 12: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Beta regression parameters (i.e environmental coefficients) estimated for the whole community model fitted with presence/absence data. For further details see Fig. S5.

## Traits coefficients

Table 2: Potential scale reduction factors (PSRF) and effective sample sizes (ESS) for traits regression parameters (i.e gamma coefficients) estimated for the model including trait information fitted either to abundance or presence-absence data. For further details see Fig. S13 to Fig. S14.

| Model | Data type | Number of coefficients | PSRF (mean sd) | ESS (mean sd) |
| --- | --- | --- | --- | --- |
| Traits & Phylogeny | Abundance | 60 | 1.08 0.092 | 1232 1209 |
| Traits & Phylogeny | Presence/Absence | 60 | 1.00 0.001 | 13227 1897 |

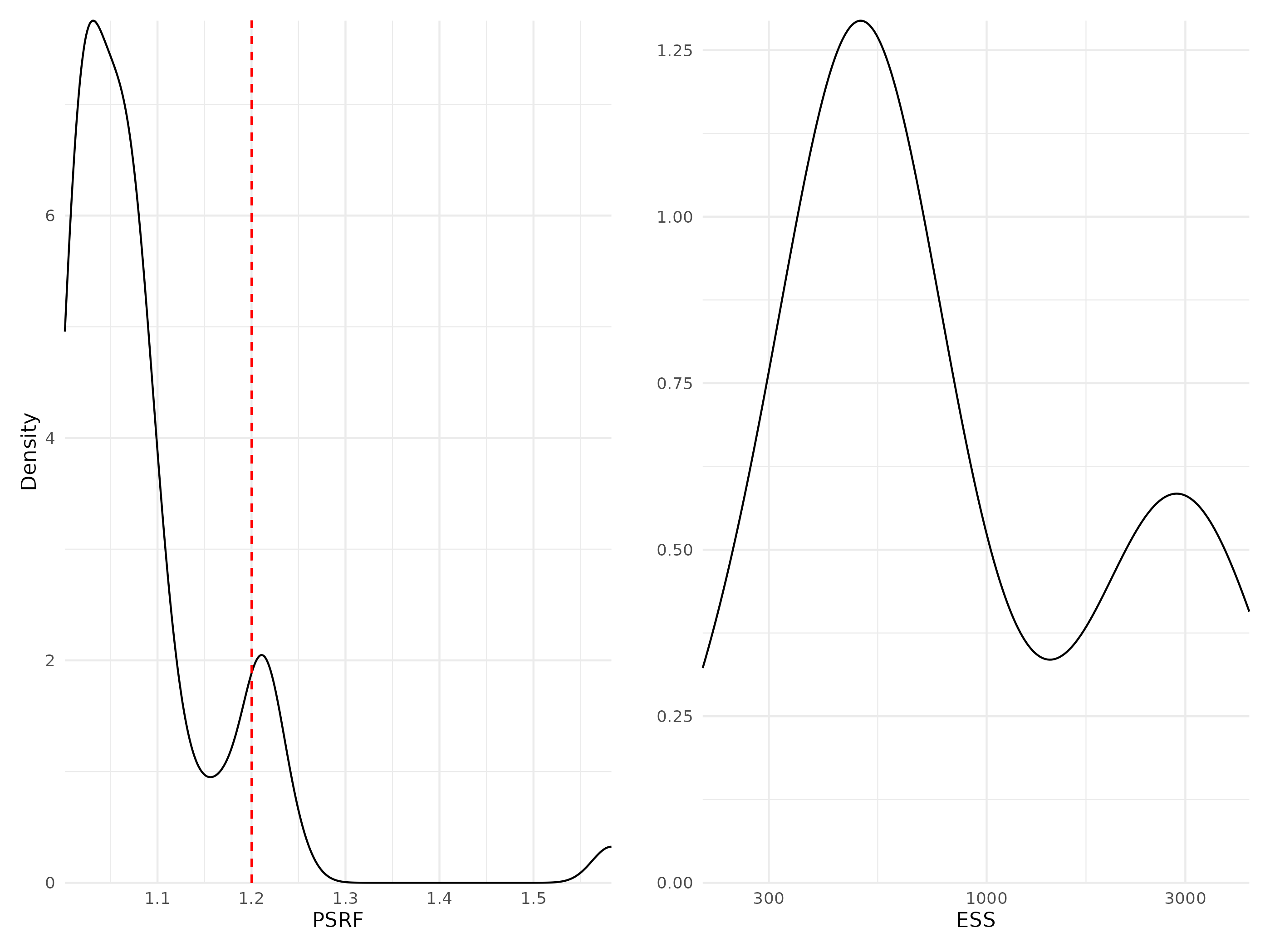


Figure 13: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Gamma regression parameters (i.e coefficients associated with trait-environment interactions, modeling how species traits influence their niches) estimated for the traits & phylogeny model fitted with abundance data. For further details see Fig. S5.

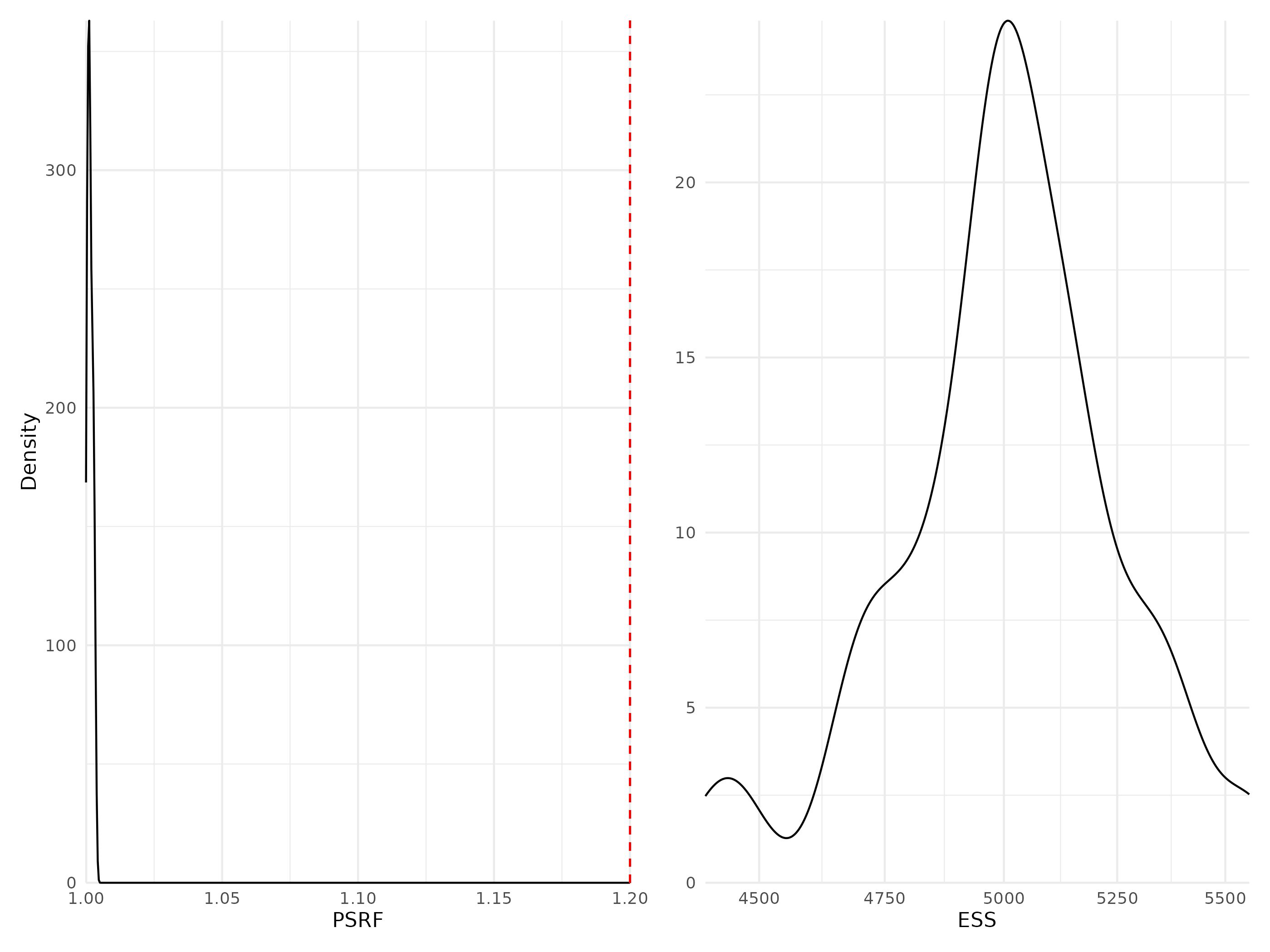


Figure 14: Density curves of potential scale reduction factors (PSRF see Brooks & Gelman ([1998](#ref-Brooks_1998)); left panel) and effective sample sizes (ESS; right panel) for Gamma regression parameters (i.e coefficients associated with trait-environment interactions, modeling how species traits influence their niches) estimated for the traits & phylogeny model fitted with presence/absence data. For further details see Fig. S5.

## Phylogeny coefficients

Table 3: Potential scale reduction factors (PSRF) and effective sample sizes (ESS) for rho regression parameters (i.e phylogeny coefficient) estimated for the two models including phylogenetic information (TrPh and Ph).

| Model | Data type | Number of coefficients | PSRF | ESS |
| --- | --- | --- | --- | --- |
| Phylogeny | Abundance | 1 | 1.07 | 649 |
| Phylogeny | Presence/Absence | 1 | 1.00 | 9349 |
| Traits & Phylogeny | Abundance | 1 | 1.15 | 757 |
| Traits & Phylogeny | Presence/Absence | 1 | 1.00 | 5000 |

## Link between model convergence and species response curves

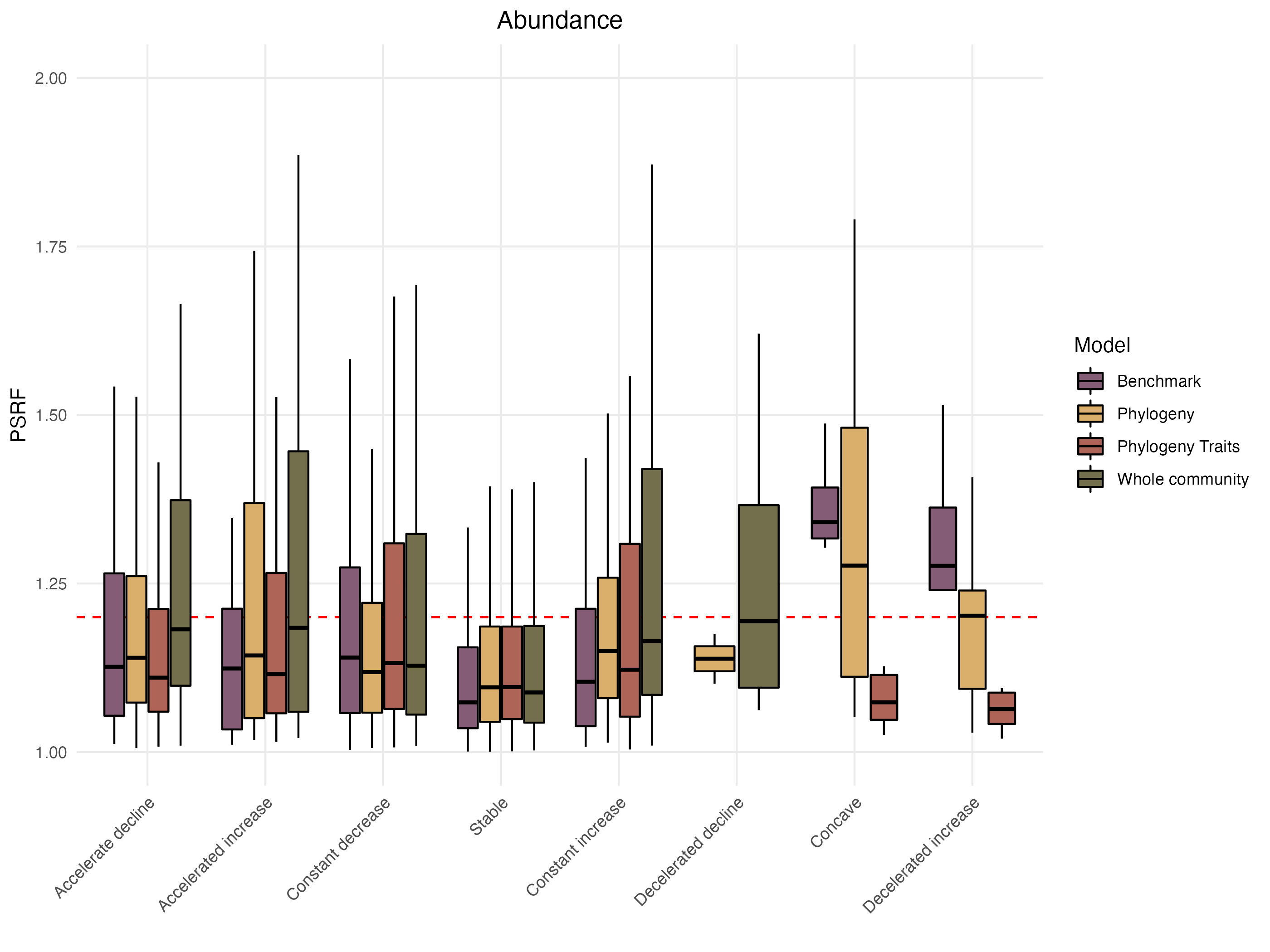


Figure 15: Distribution of the Potential scale reduction factors (PSRF) as a function of the different shapes of the response curves classified following the methodology proposed by Rigal *et al.* ([2020](#ref-Rigal_2020)) methods (see section “Assessing model performance and interpretability” for more details on the calculation methodology). Results for models fitted with abundance data.

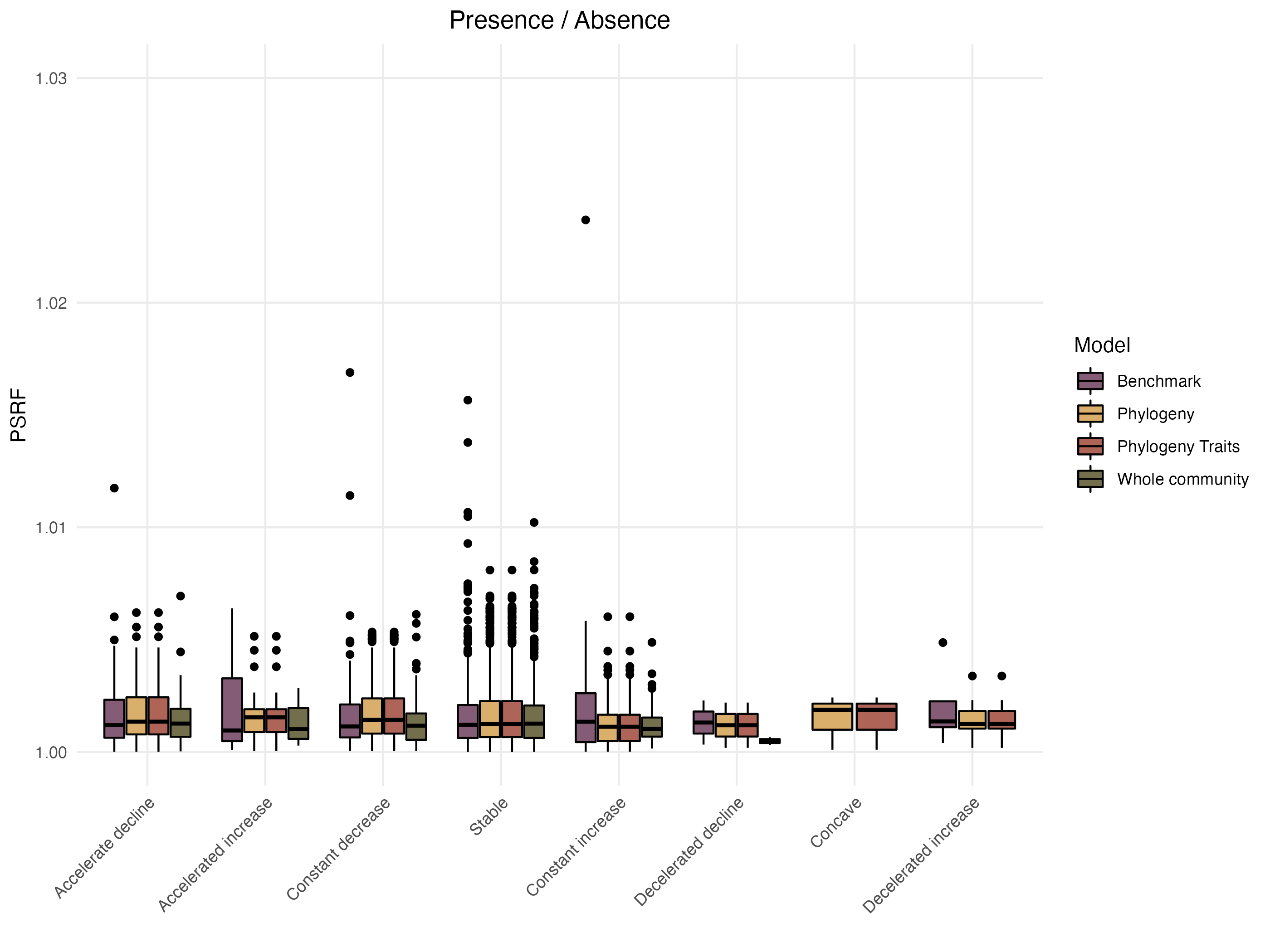


Figure 16: Distribution of the Potential scale reduction factors (PSRF) as a function of the different shapes of the response curves classified following the methodology proposed by Rigal *et al.* ([2020](#ref-Rigal_2020)) methods (see section “Assessing model performance and interpretability” section for more details on the calculation methodology). Results for models fitted with presence/absence data.

# Appendix C - Complementary Results

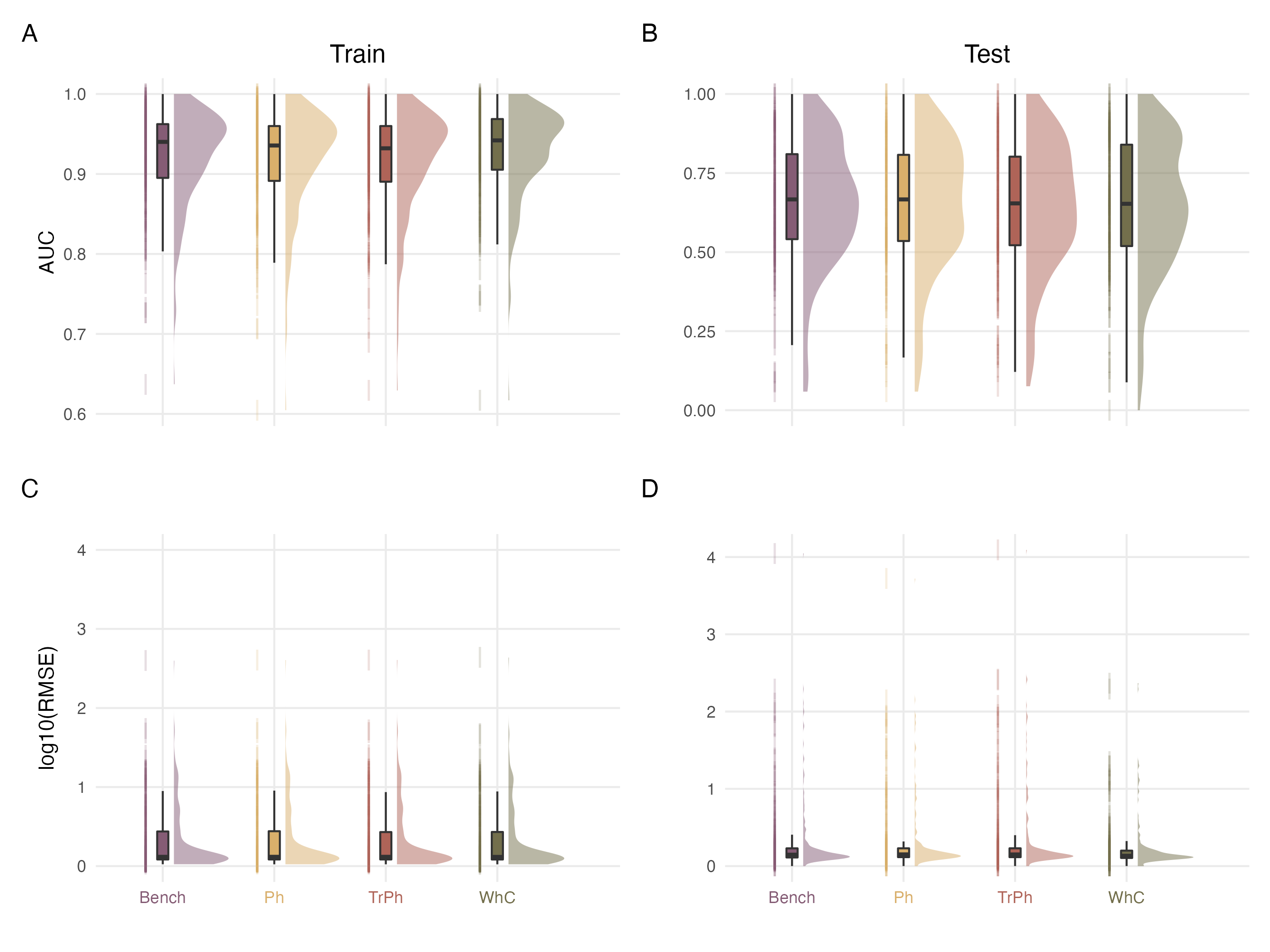


Figure 17: Comparison of explanatory (left column; Train set) and predictive (right column; Test set) performance capacities of the different model structures fitted on presence/absence (top panels) or abundance (bottom panels) data

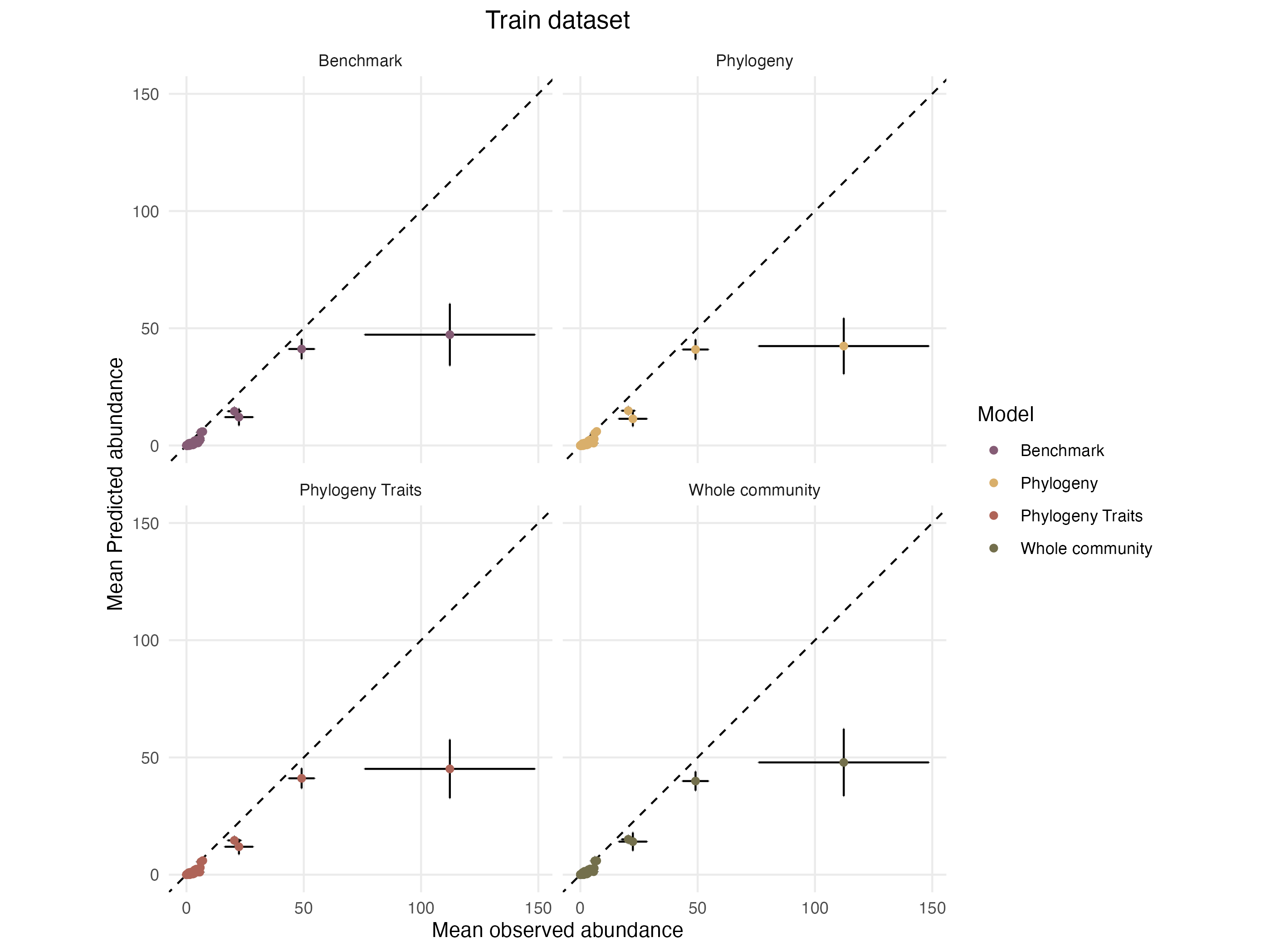


Figure 18: Mean predicted abundances as a function of mean observed abundances in the training dataset. Each species is represented by a dot, the error bars on each point indicates the standard error around the mean relatively to each axis.

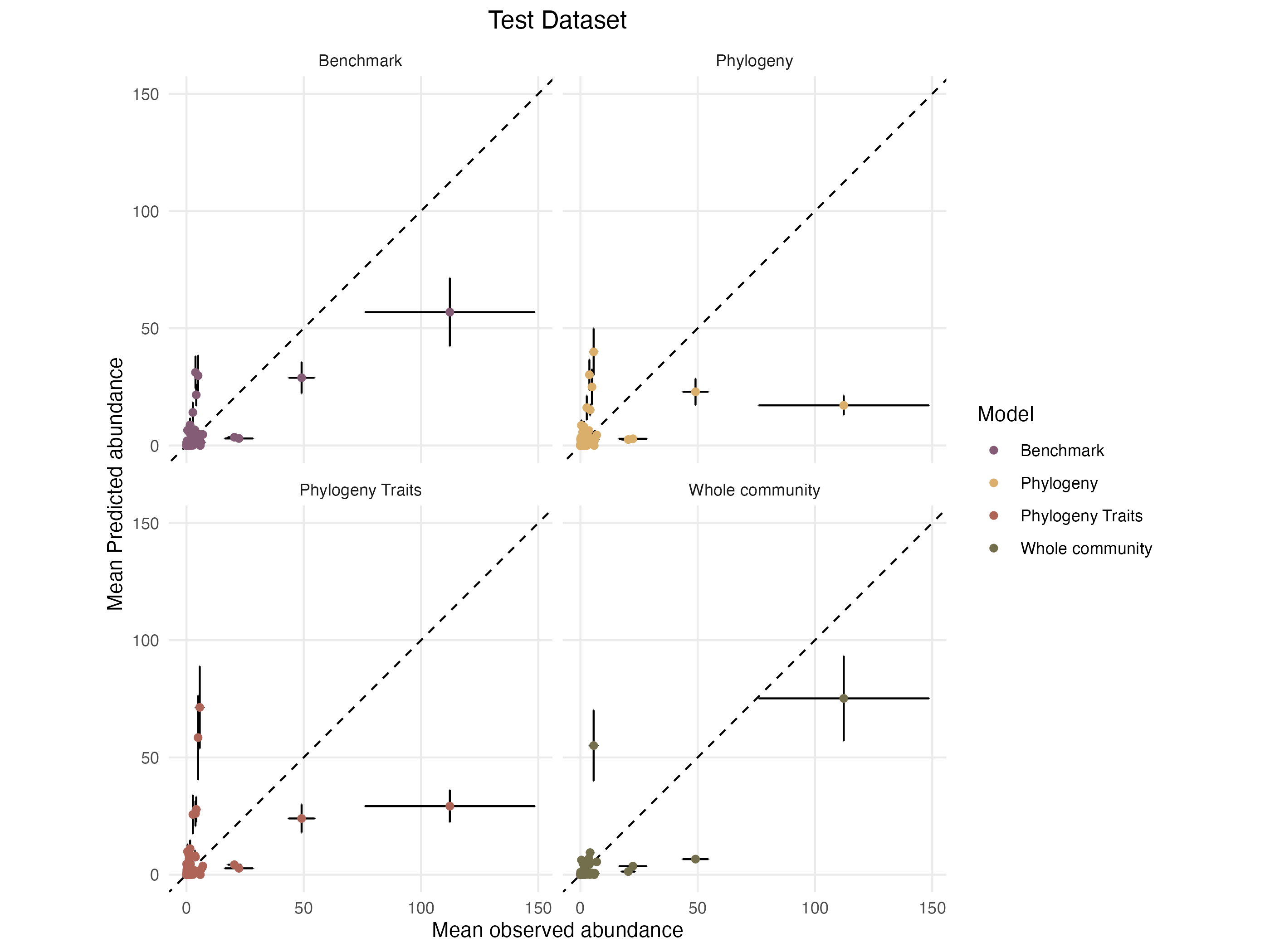


Figure 19: Mean predicted abundances as a function of mean observed abundances in the training dataset. Each species is represented by a dot, the error bars on each point indicates the standard error around the mean relatively to each axis.

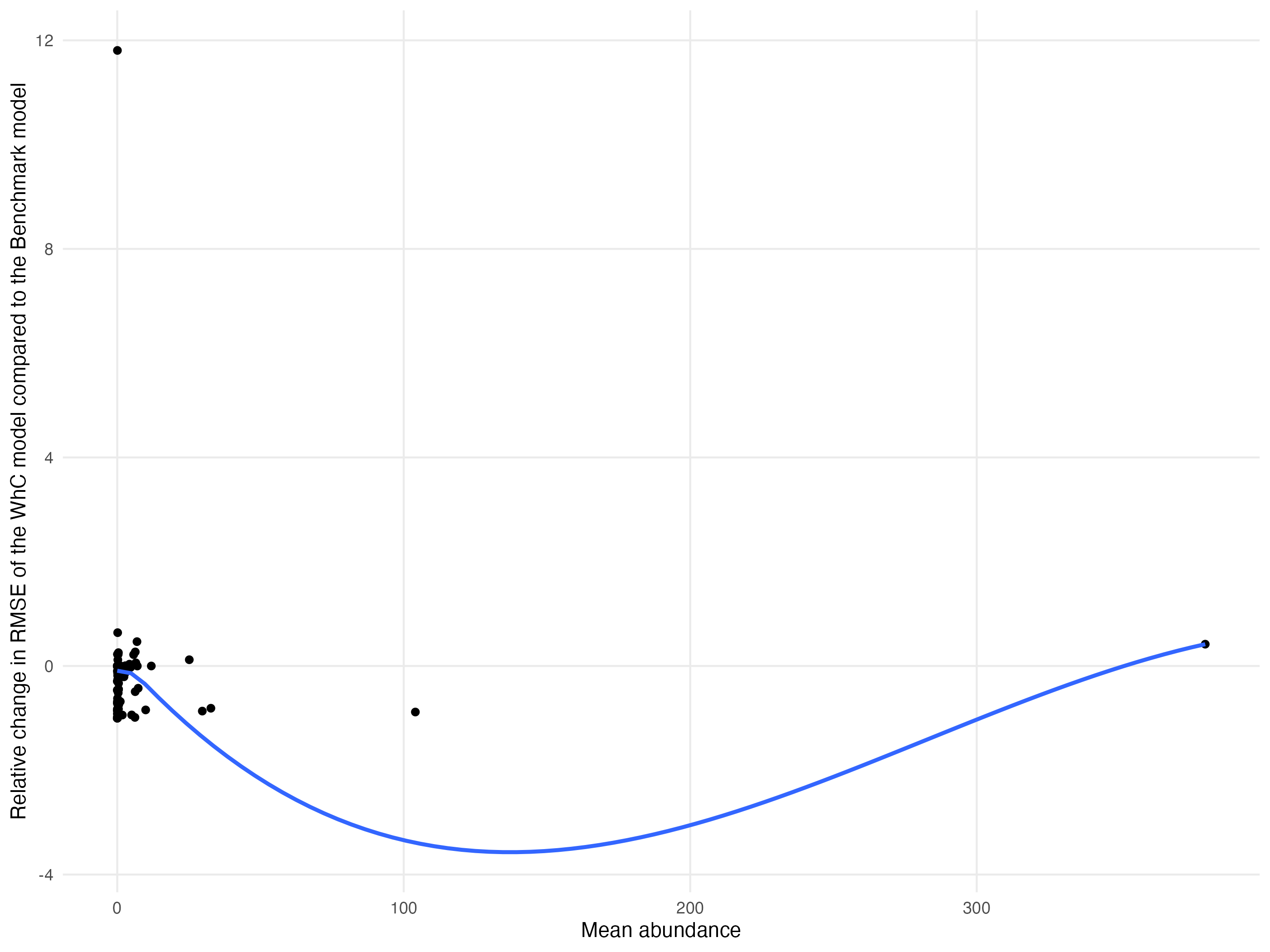


Figure 20: Relationship between the relative improvement in RMSE for the WhC model compared to Bench model and the mean abundance of species in the training dataset. Each dot represents a species. The blue line represents a fit obtained from a LOESS regression.

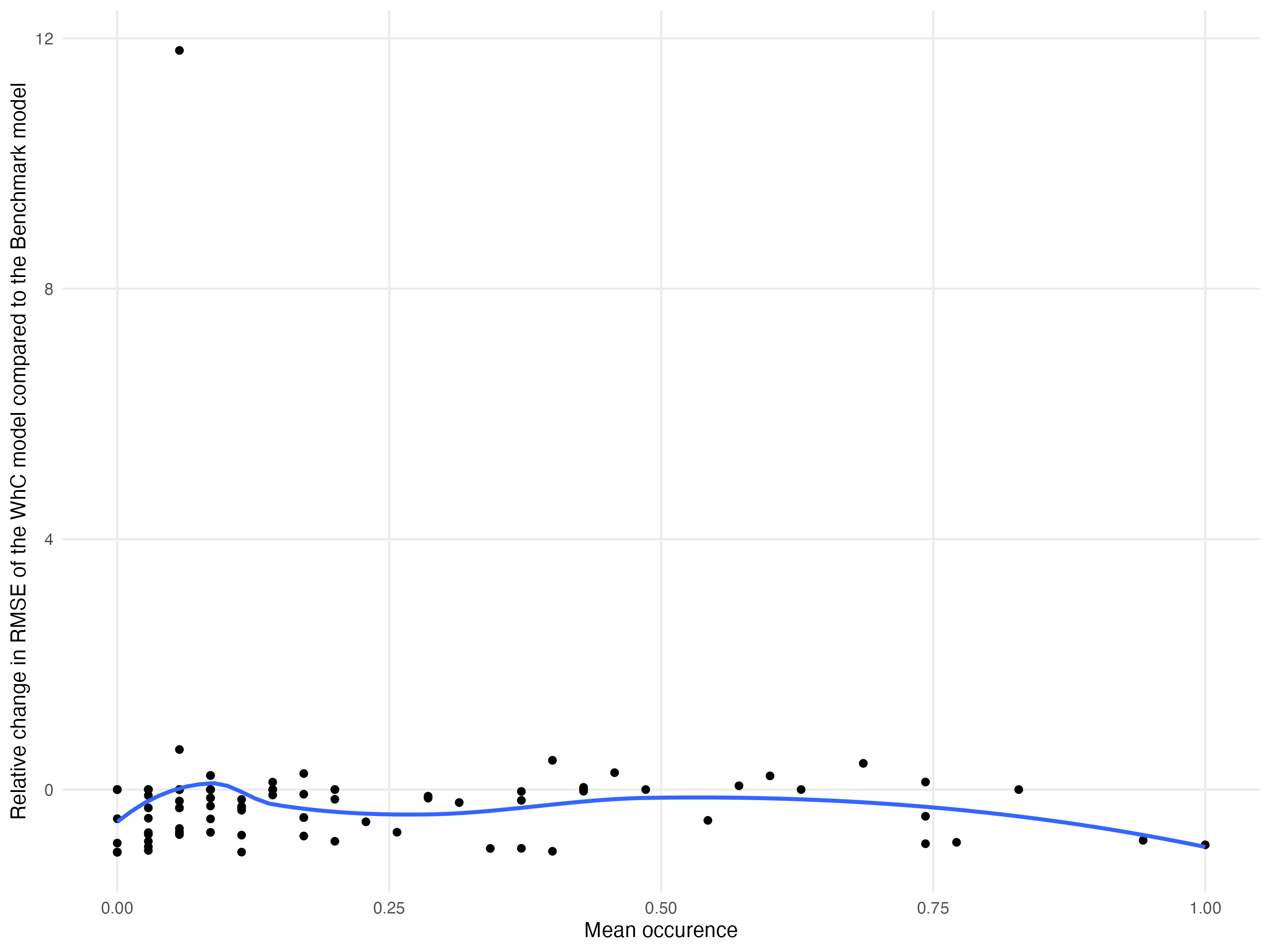


Figure 21: Relationship between the relative improvement in RMSE for the WhC model compared to Bench model and the mean occurrence of species in the training dataset. Each dot represents a species. The blue line represents a fit obtained from a LOESS regression.

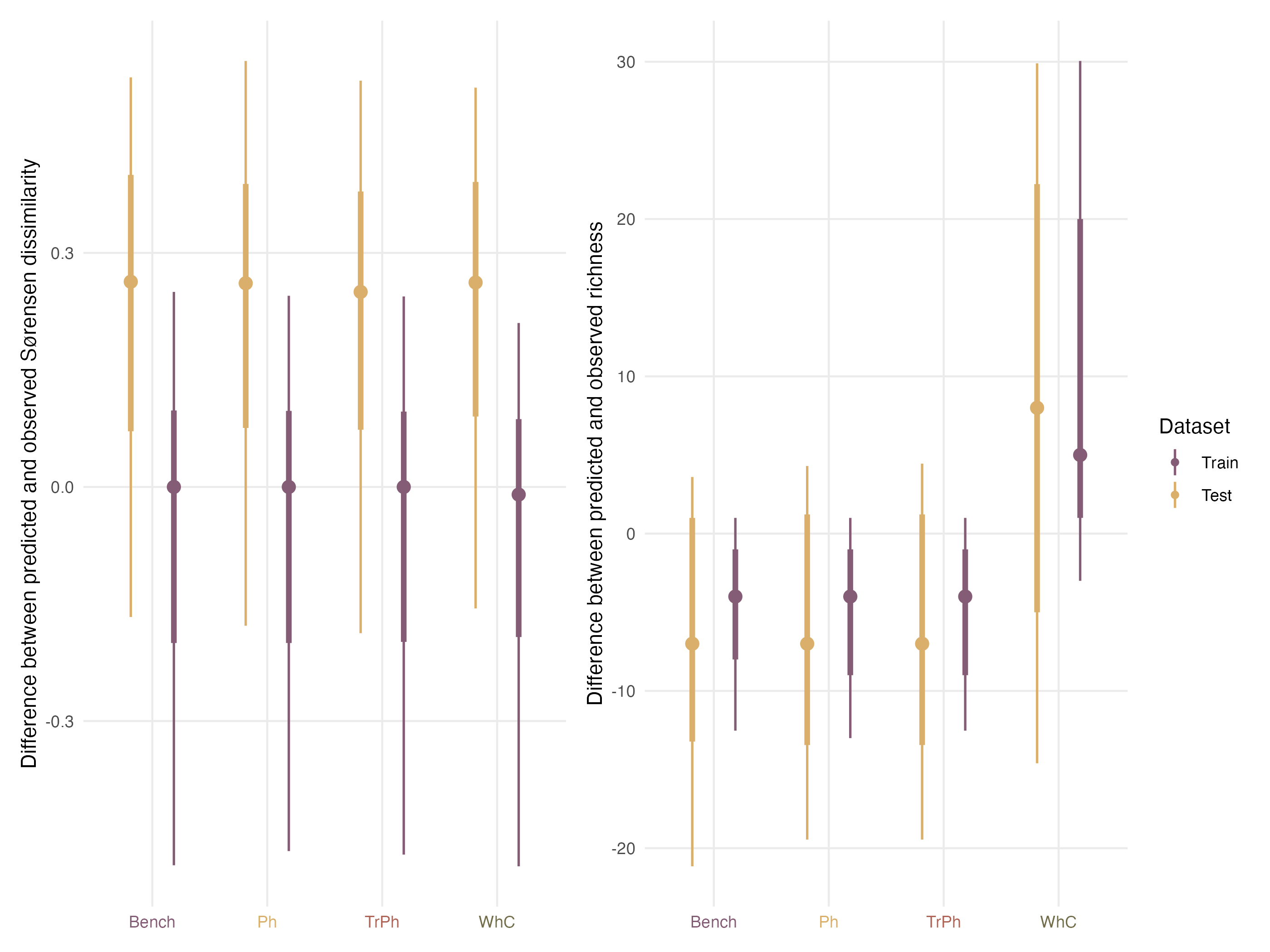


Figure 22: Comparison of model performances with regards to their ability to predict community structures when fitted with presence/absence data for the train (purple) and test (yellow) dataset. The left column indicates for each model the difference between the pairwise dissimilarities computed on the observed assemblages and those computed on the predicted community. The right column presents the differences in species richness between the observed and predicted assemblages.

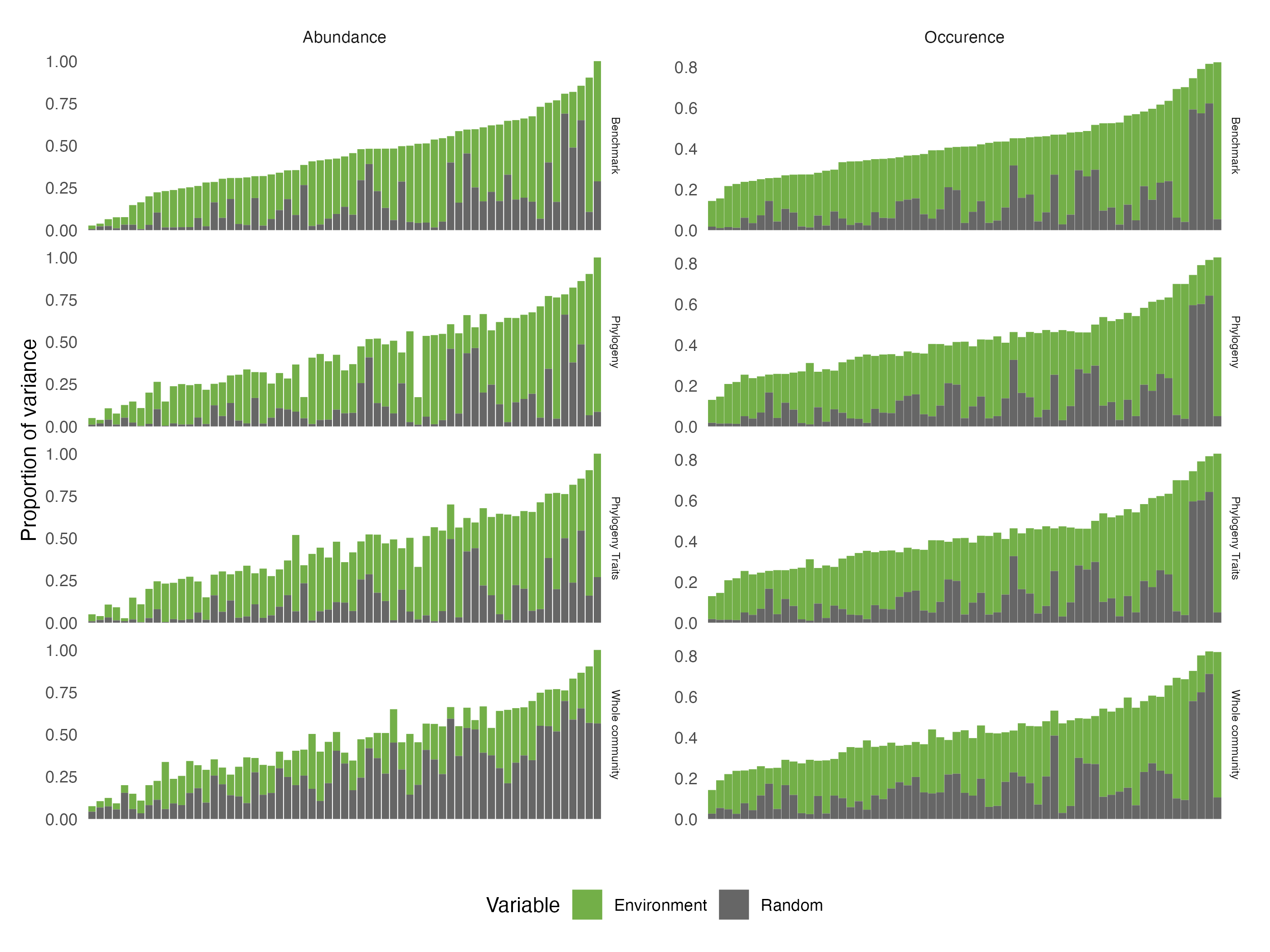


Figure 23: Comparison across the four alternative model structures of the total amount of variance of each species (along the x-axis) that is explained by (1) the environmental variables (Environment) and (2) the three random effects (Random). Results are presented for the models fitted with abundance (left) and presence/absence (right) data. Species are ordered by increasing order of total variance explained by the benchmark model.

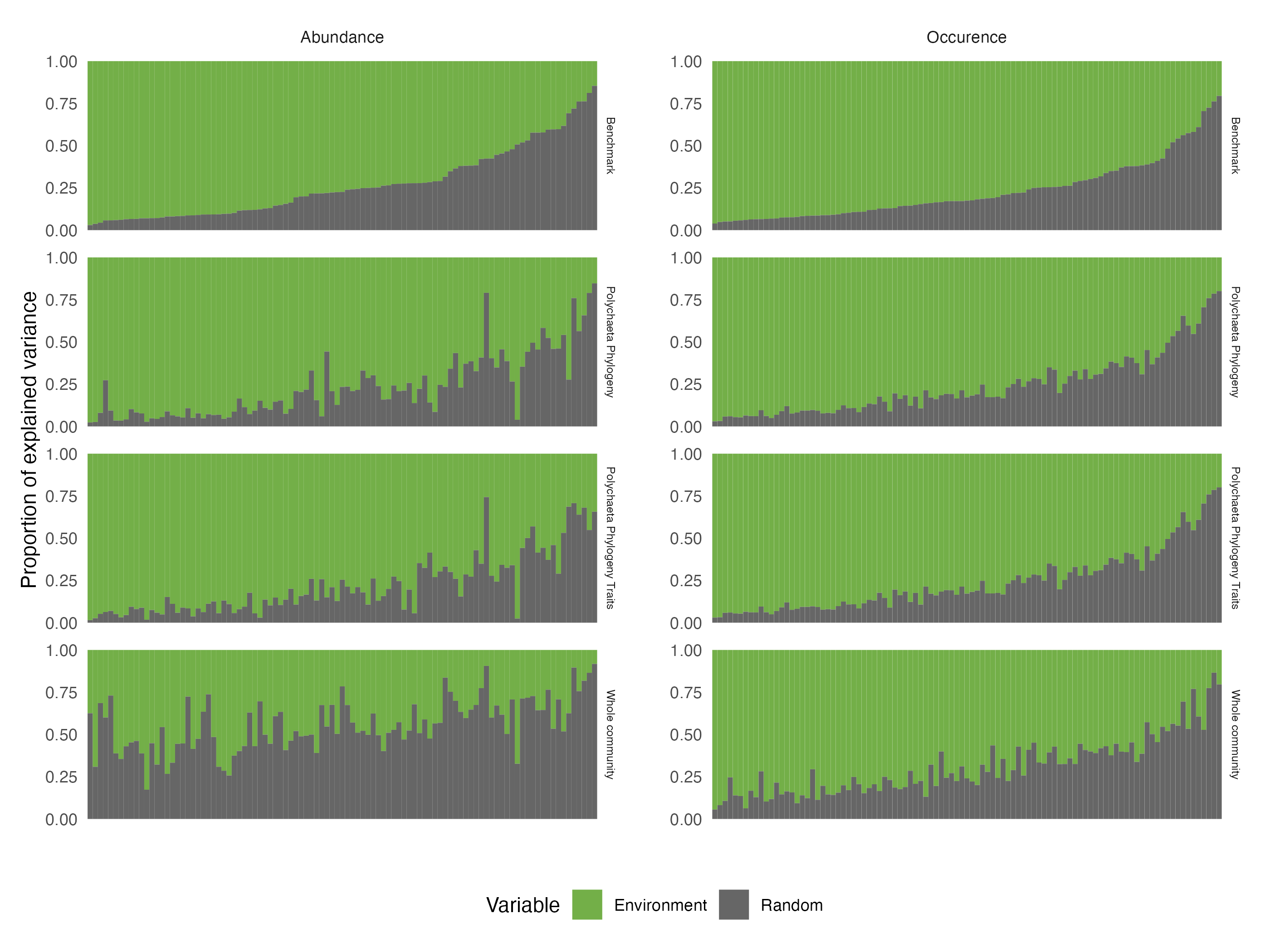


Figure 24: Comparison across the four alternative model structures of the fraction of variance of each species (along the x-axis) that is explained by (1) the environmental variables (Environment) and (2) the three random effects (Random). Results are presented for the models fitted with abundance (left) and presence/absence (right) data. Species are ordered by decreasing order of variance explained by the environment for the benchmark model.

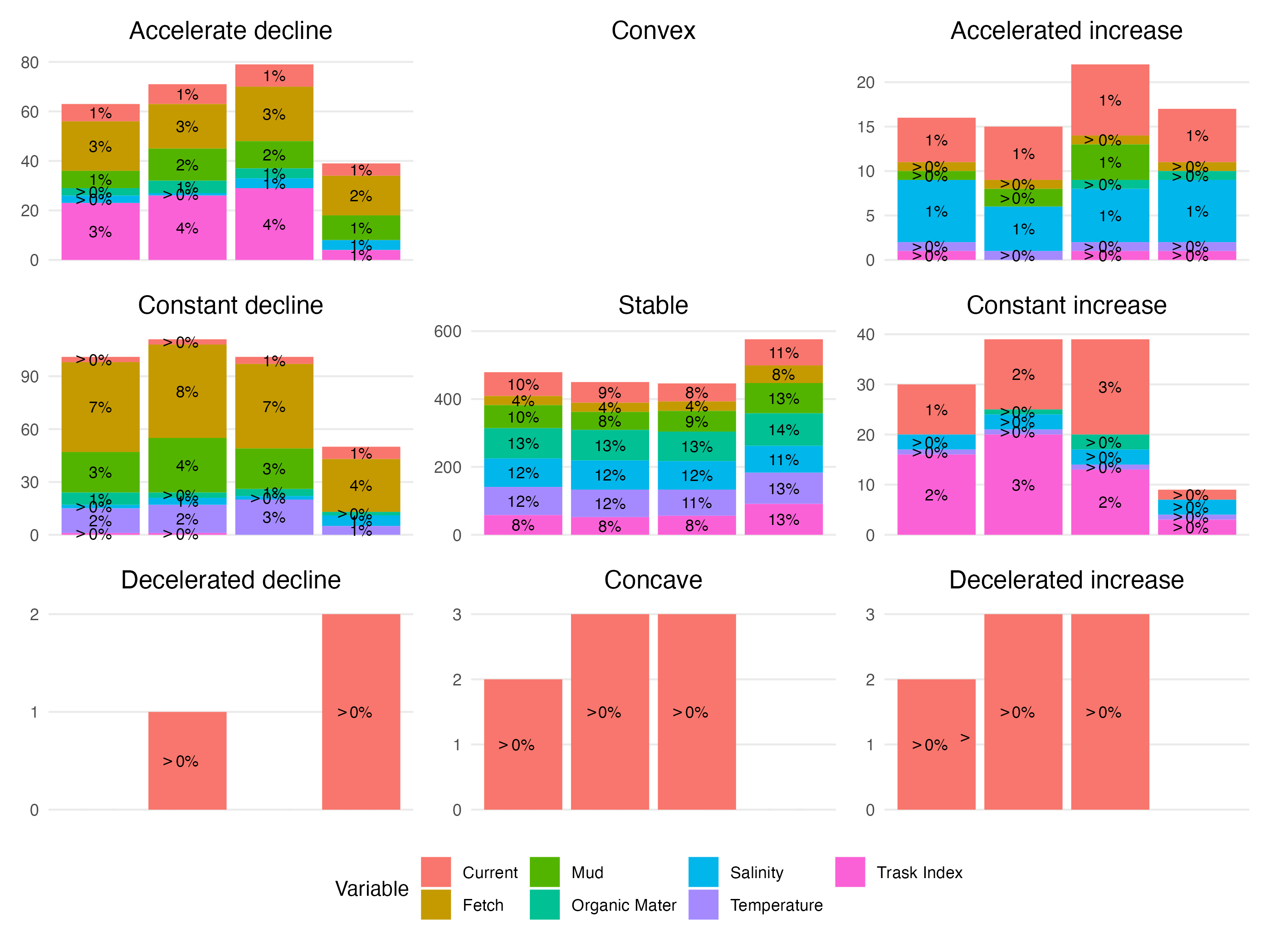


Figure 25: Same figure as Fig. 4 in the main text for models fitted on abundance data. Number (y-axis) and proportion (indicated above individual bars, rounded to the nearest integer) of response curves (i.e. one for each species-predictor combination) according to the nomenclature (nine shapes highlighted by the black curve in each panel) defined by Rigal *et al.* ([2020](#ref-Rigal_2020)). Results are presented for different model structures: from left to right the Benchmark (Bench), the phylogeny (Ph), the traits & phylogeny (TrPh), and the whole community (WhC) models. Each bar is coloured by the relative contribution of each environmental covariate to this particular shape. For illustrative purposes, note that the scale of variation on the y-axis differs across panels.

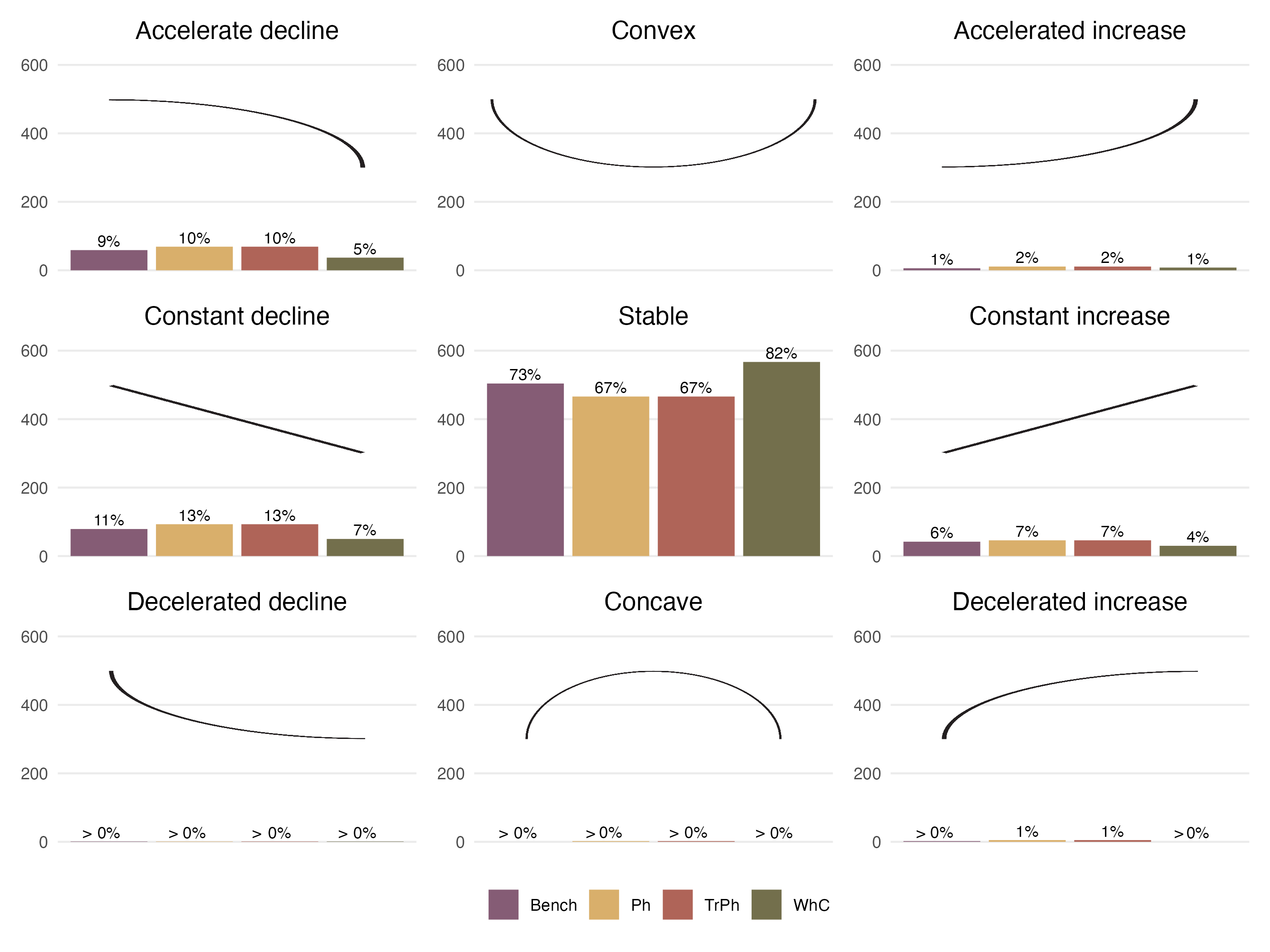


Figure 26: Same figure as Fig. 4 in the main text but for models fitted to presence/absence data. Number (y-axis) and proportion (indicated above individual bars, rounded to the nearest integer) of response curves (i.e. one for each species-predictor combination) according to the nomenclature (nine shapes highlighted by the black curve in each panel) defined by Rigal *et al.* ([2020](#ref-Rigal_2020)). Results are presented for the different structures: purple for the Benchmark (Bench), yellow for phylogeny (Ph), red for traits & phylogeny (TrPh), and green for whole community (WhC) model.

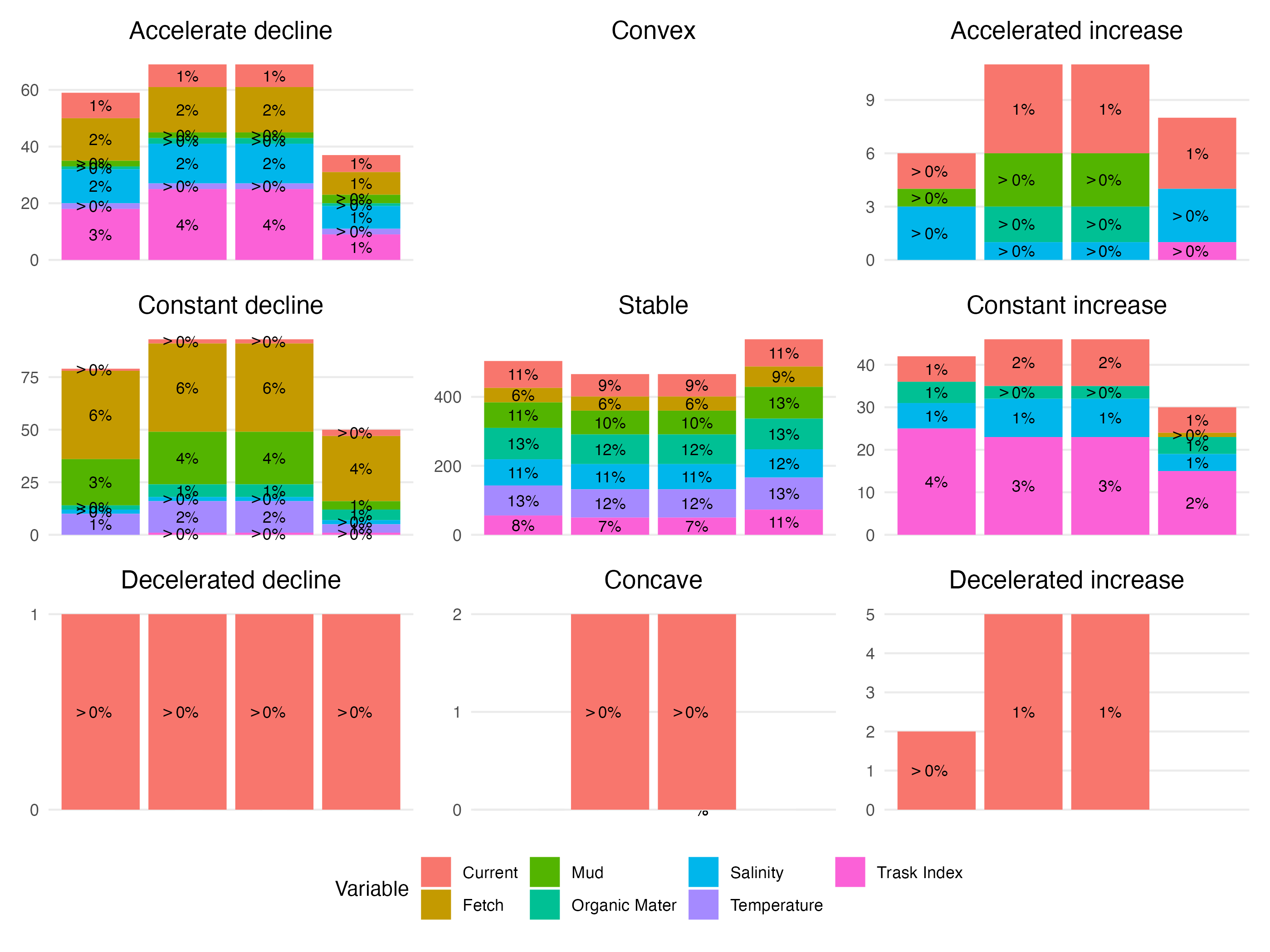


Figure 27: Same figure as +Fig. [25](#fig:supp26) but for models fitted on presence/absence data. Number (y-axis) and proportion (indicated above individual bars, rounded to the nearest integer) of response curves (i.e. one for each species-predictor combination) according to the nomenclature (nine shapes highlighted by the black curve in each panel) defined by Rigal *et al.* ([2020](#ref-Rigal_2020)) for different presence/absence model structures. Each bar is coloured by the relative contribution of each environmental covariate to this particular shape. For illustrative purposes, note that the scale of variation on the y-axis differs across panels.

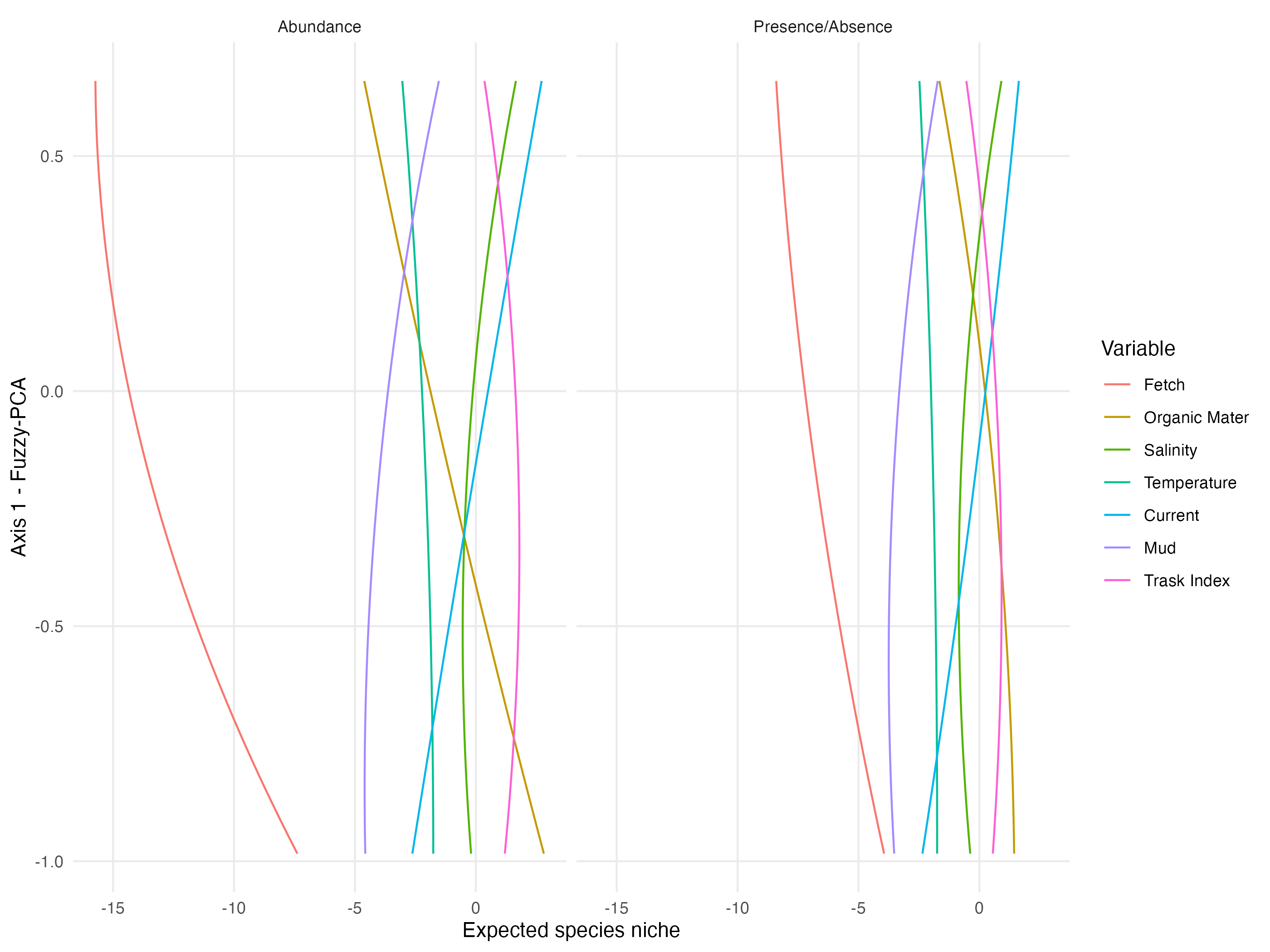


Figure 28: Relationship between species’ position along the first axis of the fuzzy PCA (sessile microphagous-mobile macrophagous gradient) and the different environmental variables used in the models (fitted with abundance data in the left panel, and with presence/absence data in the right panel). Relationships are derived from the regression coefficients estimated for the PhTr model ( coefficients in HMSC; Ovaskainen & Abrego ([2020](#ref-Ovaskainen_2020))).The lines are fitted quadratic regressions representing the average response across the different species. As an example of interpretation, the red lines in both graphs indicate that sessile microphagous species are more negatively influenced (lower abundance, low probability for presence) by fetch than macrophagous mobile species.

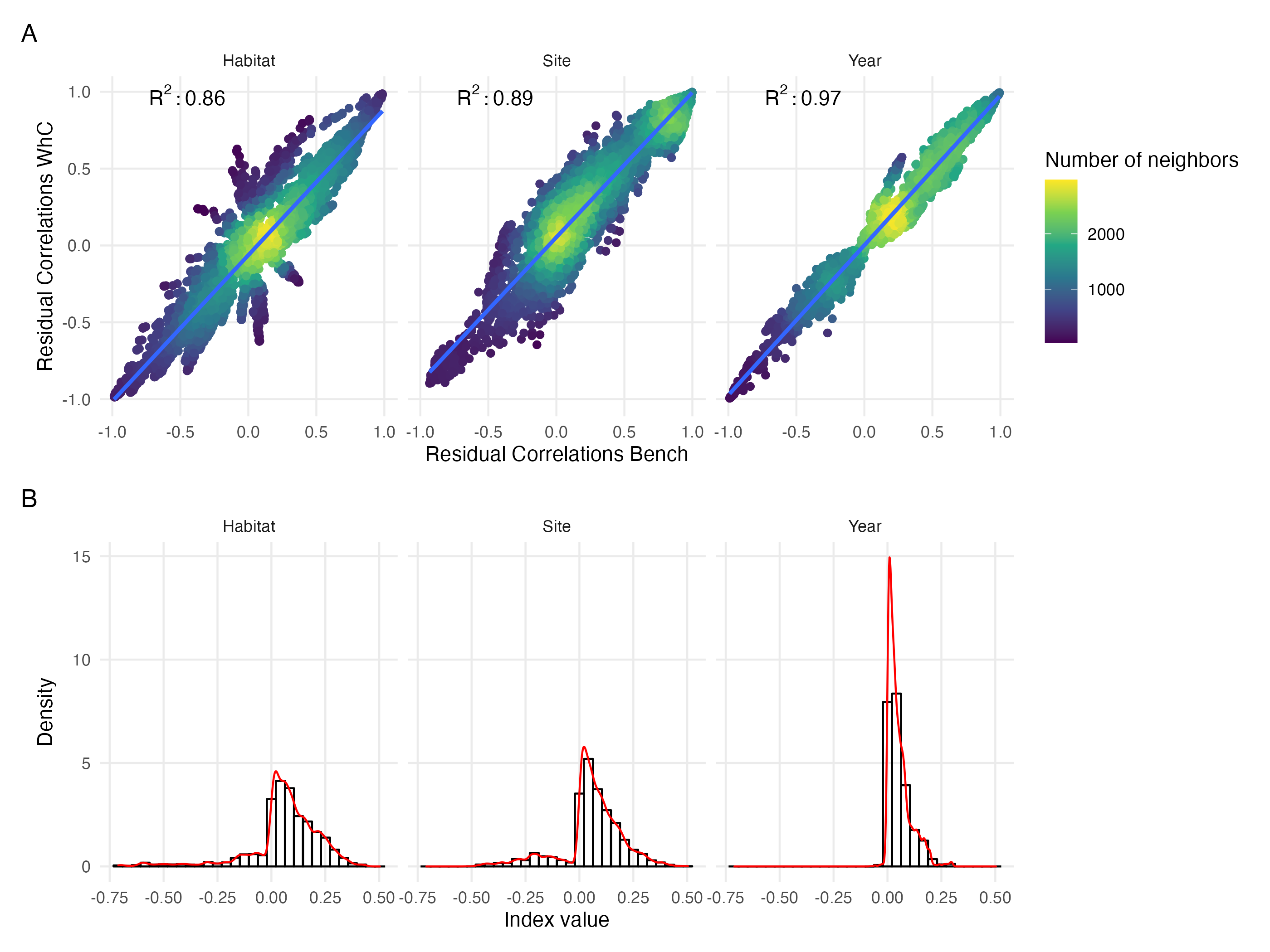


Figure 29: Same figure as Fig. 5 in the main text but for models fitted on presence/absence data. (A) Comparison of residual correlations associated with the three random effects estimated by the Whole Community Model (y-axis) and the Benchmark model (x-axis). The colour scale highlights the density of points in each scatter plot. (B) Distribution of the index measuring change in sign (sign change left to the zero line, no change to the right) and magnitude (higher departure from the zero line indicate higher difference) between residual correlations estimated by the whole community model and the benchmark model for the three random effects (Habitat, Site, Year).

# References

Boyé, A. (2019). Diversité taxinomique et fonctionnelle des habitats benthiques dans l’espace et dans le temps: Une perspective régionale et décennale. PhD thesis.

Boyé, A., Thiébaut, Éric, Grall, J., Legendre, P., Broudin, C., Houbin, C., *et al.* (2019). Trait-based approach to monitoring marine benthic data along 500 km of coastline. *Diversity and Distributions*, 25, 1879–1896.

Brooks, S.P. & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7, 434–455.

Lazure, P. & Dumas, F. (2008). An external–internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS). *Advances in Water Resources*, 31, 233–250.

Lecornu, F. & Roeck, Y.-H.D. (2009). PREVIMER - coastal observations and forecasts. *La Houille Blanche*, 95, 60–63.

Ovaskainen, O. & Abrego, N. (2020). *Joint species distribution modelling: With applications in r*. Ecology, biodiversity and conservation. Cambridge University Press.

Rigal, S., Devictor, V. & Dakos, V. (2020). A method for classifying and comparing non-linear trajectories of ecological variables. *Ecological Indicators*, 112, 106113.