

Supplementary Materials I – Summary GLMs Statistics and Models Goodness of Fit

1) Bryophytes species richness model

Predictors are described in Table 1 in the main text of Marignani et. al.

To model bryophytes richness, a Poisson Generalised linear model with a “log” link function was applied.

Full GLM formula: $\text{glm}(\text{bryophytes_richness} \sim X + Y + \text{Elevation} + \text{Distance_sea} + \text{Vegetation_type} + \text{Bare_soil} + \text{Alien_presence} + \text{Alien_cover} + \text{ratio_alien/total}, \text{family} = \text{poisson}(\text{log}))$

The Minimum Adequate Model, i.e. defined as the model that contains the minimum number of predictors that satisfy some criterion, for example, the model that only contains predictors that are significant at some pre-specified probability level, was obtained via iterative step-forward procedure. Specifically, reduced models were ranked using the Akaike Information Criterion (AIC, Akaike 1974). The best model (MAM) that minimise the AIC was found through exhaustive screening of all the possible candidates (Fig. 1). The relative importance of predictors in all the evaluated reduced models was graphically assessed (Fig. 2).

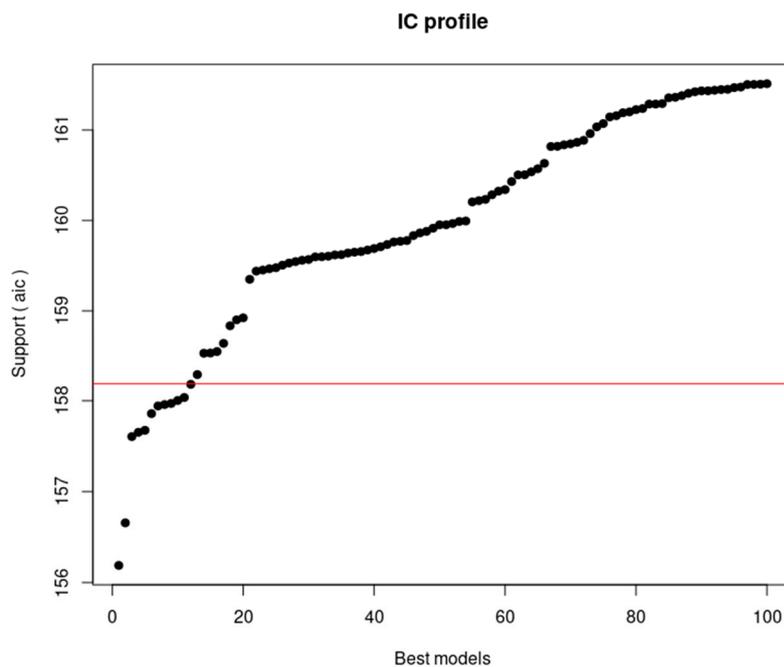


Figure S1: The AIC profile for the reduced models. A horizontal line delineates models that are less than 2 AIC units away from the best model. The selected MAM was the one minimising the AIC.

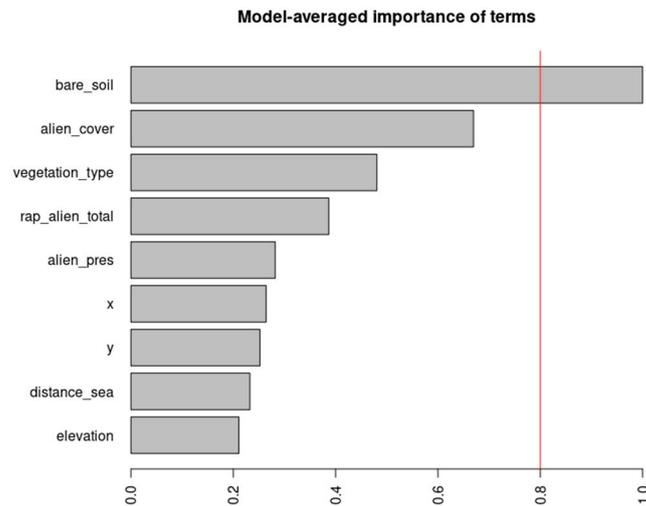


Figure S2: plots the relative importance of model terms, i.e. the overall support for each variable across all models. A vertical line is arbitrary drawn at 80%

Minimal Adequate Model selected (formula): `glm(bryophytes_richness ~ vegetation_type + bare_soil + alien_pres, family = poisson(log))`

Summary statistics for this model (e.g., coefficient signs and amount of single-term explained variance) was described in Table 4 of the main text in Marignani et. al. As a Goodness of fit measure for the MAM, we used the D^2 Adjusted. For generalized linear models, this is the equivalent of the amount of deviance accounted by the set of predictors retained (Guisan & Zimmermann 2000, Barbosa et al. 2014). Adjusted D^2 takes into account the number of observations and the number of model parameters, thus allowing direct comparison among different models (Guisan & Zimmermann 2000). For the Bryophytes richness MAM, the amount of explained variance was ~26%

```
> Dsquared(MAM_bryophytes_richness)
[1] 0.2613
```

The significant main term in the MAM are graphically displayed in Fig. 3.

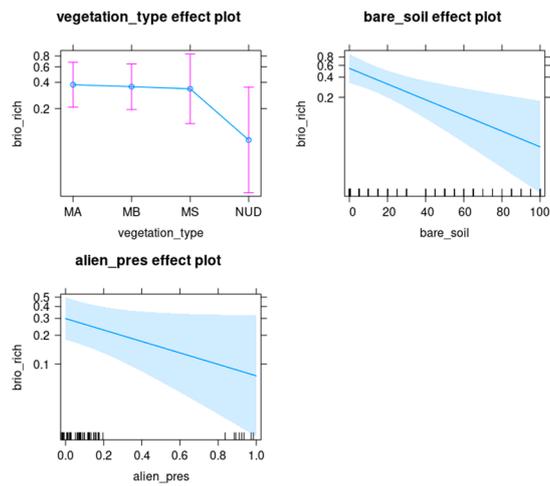


Figure S3: Effects of the MAM retained predictors versus the response variable (briophyte richness)

2) Bryophytes functional diversity model

The same framework as for the modelling approach described above was applied to determine the influence of environmental and biological factors on functional bryophytes diversity. A Gaussian model was selected in this case, with an “identity” link function.

Full GLM formula:

```
glm(funRao~x+y+elevation+distance_sea+vegetation_type+bare_soil+alien_pres+alien_cover+rap_alien_total, family(gaussian))
```

The best model (MAM) that minimise the AIC was found through exhaustive screening of all the possible candidates (Fig. 4). The relative importance of predictors in all the evaluated reduced models was graphically assessed (Fig. 5).

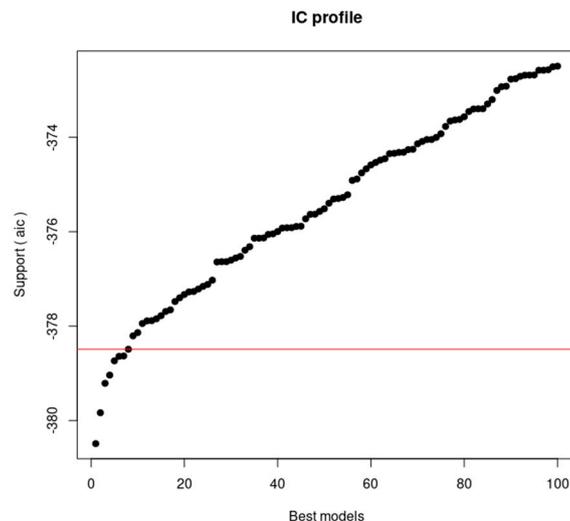


Figure S4: The AIC profile for the reduced models. A horizontal line delineates models that are less than 2 AIC units away from the best model. The selected MAM was the one minimising the AIC.

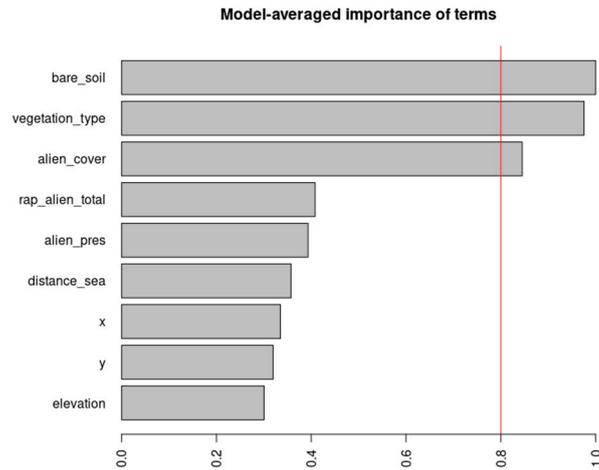


Figure S5: plots the relative importance of model terms, i.e. the overall support for each variable across all models. A vertical line is arbitrary drawn at 80%

Call:

Minimal Adequate Model selected (formula): `glm(functional_Rao ~ distance_sea + vegetation_type + bare_soil + alien_pres + alien_cover)`

Summary statistics for this model (e.g., coefficient signs and amount of single-term explained variance) was described in Table 5 of the main text in Marignani et. al. For the Bryophytes functional diversity, the amount of explained variance by the MAM was ~26%

`> Dsquared(modello_funz_red)`

`[1] 0.2611`

The significant main term in the MAM are graphically displayed in Fig. 6.

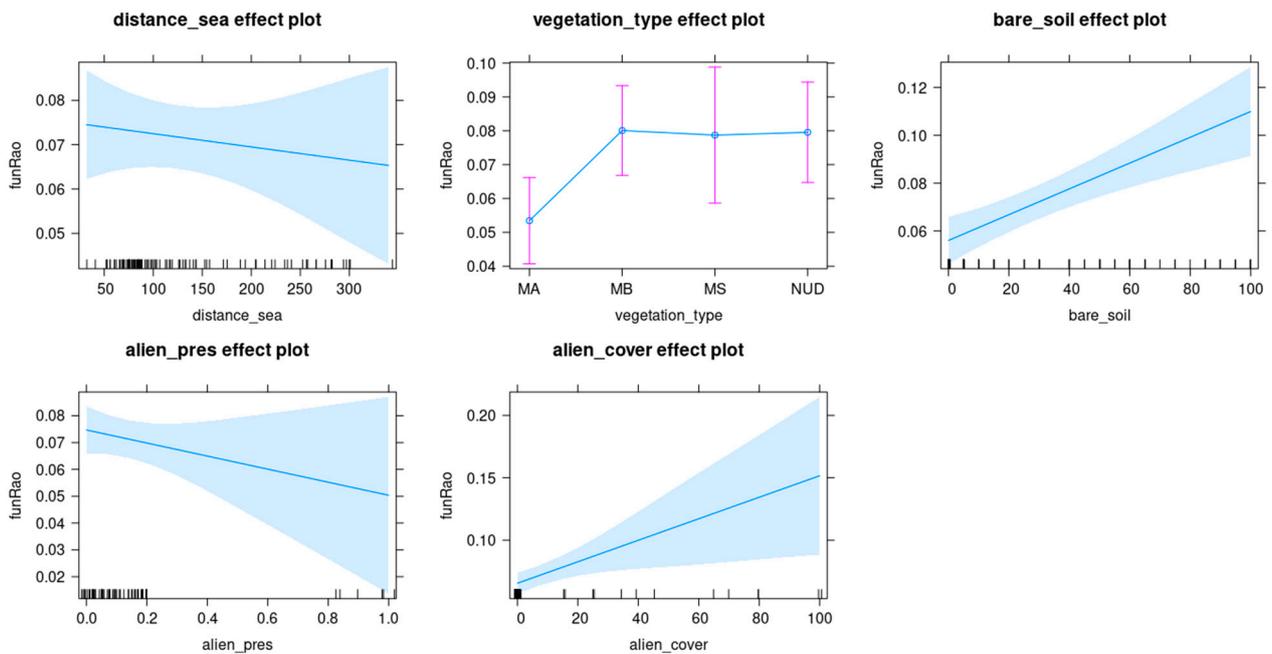


Figure S6: Effects of the MAM retained predictors versus the response variable (briophytes functional diversity RaoQ)

Bibliography

Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **1974**, 19, 716–723.

Barbosa A.M., Brown J.A. & Real R. modEvA – an R package for model evaluation and analysis. R package, version 0.1. **2014**

Guisan, A. & Zimmermann, N.E. Predictive habitat distribution models in ecology. *Ecological Modelling* **2000**,135: 147-186

Calcagno Vincent. glmulti: Model Selection and Multimodel Inference Made Easy. R package version 1.0.7.1. 2019 <https://CRAN.R-project.org/package=glmulti> accessed on February **2020**

Whittingham MJ, Stephens PJ,Bradbury RB,Freckleton RP. Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, **2006**, 75, 1182–1189