Improving spatial predictions of taxonomic, functional and phylogenetic diversity

Manuela D'Amen †1, Rubén G. Mateo †1, Julien Pottier 2, Wilfried Thuiller 3, Luigi Maiorano4, Loïc Pellissier5, Charlotte Ndiribe6, Nicolas Salamin1 and Antoine Guisan 1,7

† Co-first authors

1. Department of Ecology & Evolution, Biophore, University of Lausanne, CH-1015 Switzerland
2. INRA, UR874 (Unité de Recherche sur l'Ecosystème Prairial), 5 chemin de Beaulieu, F-63039 Clermont-Ferrand, France.
3. Université Grenoble Alpes, CNRS, Laboratoire d'Ecologie Alpine (LECA), Grenoble F-38000, France
4. Department of Biology and Biotechnologies "Charles Darwin", University of Rome "La Sapienza", Rome, Italy.
5. Department of Environmental Systems Science, institute of Terrestrial Ecosystems at ETH Zurich and at WSL, Zurich, Switzerland.
6. Department of Cell Biology and Genetics. University of Lagos, Akoka, Lagos, Nigeria.
7. Institute of Earth Surface Dynamics, Geopolis, University of Lausanne, 1015 Lausanne, Switzerland

Corresponding author: Manuela D’Amen, Department of Ecology & Evolution, Biophore, University of Lausanne, CH-1015 Switzerland e-mail: manuela.damen@msn.com

Running head: Spatial predictions of multiple plant diversity facets

Summary

1. In this study, we compare two community modelling approaches to determine their ability to predict the taxonomic, functional and phylogenetic properties of plant assemblages along a broad elevation gradient and at a fine resolution. The first method is the standard stacking individual species distribution modelling (SSDM) approach, which applies a simple environmental filter to predict species assemblages. The second method couples the SSDM and macroecological modelling (MEM - SSDM-MEM) approaches to impose a limit on the number of species co-occurring at each site. Because the detection of diversity patterns can be influenced by different levels of phylogenetic or functional trees, we also examine whether performing our analyses from broad to more exact structures in the trees influences the performance of the two modelling approaches when calculating diversity indices.

2. We found that coupling the SSDM with the MEM improves the overall predictions for the three diversity facets compared with those of the SSDM alone. The accuracy of the SSDM predictions for the diversity indices varied greatly along the elevation gradient, and when considering broad to more exact structure in the functional and phylogenetic trees, the SSDM-MEM predictions were more stable.

3. SSDM-MEM moderately but significantly improved the prediction of taxonomic diversity, which was mainly driven by the corrected number of predicted species. The performance of both modelling frameworks increased when predicting the functional and phylogenetic diversity indices. In particular, fair predictions of the taxonomic composition by SSDM-MEM led to increasingly accurate predictions of the functional and phylogenetic indices, suggesting that the compositional errors were associated with species that were functionally or phylogenetically close to the correct ones; however, this did not always hold for the SSDM predictions.

4. *Synthesis.* In this study, we tested the use of a recently published approach that couples species distribution and macroecological models to provide the first predictions of the distribution of multiple facets of plant diversity: taxonomic, functional and phylogenetic. Moderate but significant improvements were obtained; thus, our results open promising avenues for improving our ability to predict the different facets of biodiversity in space and time across broad environmental gradients when functional and phylogenetic information is available.

Key words. Assemblage modelling, vegetation, ecological assembly rules, macroecological models, similarity effect, Sørensen dissimilarity index, stacked species distribution models.

Introduction

Significant research efforts are allocated to assessing and modelling the impact of global changes on biodiversity over a wide range of scales and for different ecosystems and taxa (Bellard *et al.* 2012). An integrative assessment requires the simultaneous investigation of the multiple facets of biodiversity: taxonomic, functional and phylogenetic (Thuiller *et al.* 2015). The most frequently considered element of biodiversity is the taxonomic facet, which is associated with species richness and composition and requires the counting and identification of species that co-occur in a given unit area (alpha diversity; e.g., Ferrier & Guisan 2006; White & Kerr 2006; Mateo *et al.* 2012). The capacity to model functional and phylogenetic diversity has been explored more recently and is an important area of research (Dubuis *et al.* 2013; Ndiribe *et al.* 2014; Rosauer *et al.* 2014; Chalmandrier *et al.* 2015b; Thuiller *et al.* 2015; Jarzyna & Jetz 2016). Phylogenetic diversity is a measure that accounts for phylogenetic relationships among taxa, and it facilitates investigations into the impacts of evolutionary history in the assemblage of communities (Faith 1992; Webb *et al.*, 2002). Finally, functional diversity represents the variance of species functional traits within a given assemblage (*sensu* Violle *et al.* 2007) and may provide insights into the processes that shape local assemblages (Lavorel & Garnier 2002; McGill *et al.* 2006).

Biodiversity patterns can be predicted via modelling approaches of varying complexity and different emphases on the processes that generate community structure (e.g., Webb *et al.* 2010; Mokany *et al.* 2011; Fernandes *et al.* 2013, see D’Amen *et al.* 2015c for a review). A standard approach to obtaining compositional information is to reconstruct communities by stacking individual species predictions obtained from species distribution models (stacked species distribution models, SSDMs, Pineda & Lobo, 2009; Mateo *et al.* 2012). This strategy is based on the assumption that communities originate from the coincidental assemblage of individualistic ecological responses of species. Because individual species distribution models

(SDMs) mainly use abiotic environmental data to perform inferences regarding species range limits and habitat suitability (Franklin 1995; Guisan & Zimmermann 2000; Elith & Leathwick 2009), the derived SSDMs can account for the effect of environmental drivers in the community predictions but may not include the effects of historical factors, biotic interactions and dispersal limitations (Thuiller *et al.* 2013). These models show varying abilities in predicting observed species richness patterns or assemblage compositions, but for certain groups they have a tendency to over-predict the number of species in a community (e.g., Guisan & Rahbek 2011; Pineda & Lobo 2012; Calabrese *et al.* 2014; but see D'Amen, Pradervand & Guisan 2015b and Mateo *et al.* 2016).

Recently, the novel community modelling framework SESAM has been proposed to reconstruct species assemblages by integrating the different drivers of the assembly process, including species pool definitions, habitat filtering, macroecological constraints and ecological assembly rules (Guisan & Rahbek 2011). To date, the SESAM framework has been mostly implemented by applying the habitat filter through stacking simple static SDMs and constraining predictions for different assemblage properties via coupling with macroecological models (MEMs) (e.g., as species richness or functional limits) (D'Amen *et al*. 2015a). A key question in such SESAM implementation is related to the species that should be selected from the pool of species predicted by the SSDM to produce the MEM prediction. One possible solution is to rank species by their SSDM-predicted probability at each site and retain only the highest probability species in the final assemblage prediction until the MEM prediction is reached ("probability ranking" rule, PRR, D'Amen *et al*. 2015a). This approach can be considered a putative ecological assembly rule if the probability of presence at each site can be assumed to be a good proxy for species competitive strength at that site. Because our implementation only represents a part of SESAM, in the following sections we will refer to it as SSDM-MEM.

The SSDM-MEM approach has been shown to improve upon the SSDM approach by limiting species richness over-predictions (D'Amen *et al.* 2015a). However, although many studies have compared different approaches to predicting the taxonomic component of biodiversity (see D'Amen *et al.* 2015a), no such comparisons have been performed for the functional and phylogenetic facets. In particular, SSDM-MEM has never been tested for the prediction of these components (D'Amen, Pradervand & Guisan 2015b). Here, we compare the two community-modelling approaches - simple SSDM and SSDM-MEM - to determine their ability to predict the three main facets of biodiversity: taxonomic, functional, and phylogenetic. We use fine-resolution data of plant communities along an elevation gradient and utilize a simplified concept of "community", which we define as taxonomic assemblages of species inhabiting the same plot. We also calculate and predict diversity indices for these three components because they are important aggregated measures in ecology and conservation biology (Corbelli *et al.* 2015; Jarzyna & Jetz 2016).

First, we test whether the performance of SSDMs can be improved by implementing the SSDM-MEM approach. We expect taxonomic diversity to be predicted more accurately by using SSDM-MEM because this method makes it possible to limit richness over-predictions (D'Amen *et al*. 2015c). Functional diversity is directly linked to the response of species to the environment (Thuiller *et al*. 2015); thus, we can expect here simple SSDMs that account for the abiotic environmental drivers in the species sorting to show sufficient/good performances. Conversely, we expect that phylogenetic diversity will be predicted with less accuracy than functional diversity because it is likely to be influenced by additional processes, such as evolutionary history, which are not considered here (Mouquet et al. 2012; Münkemüller et al. 2015).

Second, we test whether differences in the predictive performance of the two community modelling approaches can be detected along the elevation gradient. Elevation gradients include variations in several environmental factors (e.g., temperature, precipitation, topography, and soil) that directly influence the presence of organisms. Therefore, these gradients are useful for testing general hypotheses regarding the main drivers that shape diversity patterns (e.g., Callaway *et al*. 2002; Michalet *et al*. 2014) and testing community-modelling approaches (e.g., Dubuis et al. 2011, Mateo et al. 2012).

Finally, recent reports have indicated that the detection of diversity patterns can be influenced by the different levels of the phylogenetic or functional tree when calculating functional or phylogenetic indices (i.e., the "*similarity effect*" in Chalmandrier *et al.* 2015b). In fact, community drivers can be applied differently across descendant clades (Vamosi *et al.* 2009). Habitat filtering is primarily used when considering the assembly of distantly related lineages, whereas biotic interactions have been shown to be more important for retrieving phylogenetic and functional structures at the tips of the tree, i.e., among closely related taxa (Crisp *et al.* 2009; Ndiribe *et al.* 2013). Based on such evidence, we also test how assigning greater weight to ancient or recent divergences in the functional and phylogenetic trees influences the performance of the two modelling approaches in predicting diversity indices (Chalmandrier *et al.* 2015b).

Materials and Methods

Species, traits, and phylogenetic data

The study area covers approximately 700 km^2 of a mountain region located in the Western Swiss Alps, and it is characterized by a large elevation gradient ranging from 375 m to 3210 m a.s.l. (Appendix S1 in Supporting Information). Exhaustive floristic inventories consisted of a set of 613 plots of 4 m^2 ranging from 700 to 3000 m a.s.l., and they were distributed within the study area according to a stratified-random sampling design to evenly cover the range of habitat conditions (Hirzel & Guisan 2002). We did not consider the low elevation band (lower than 700 m a.s.l.) to avoid areas of intense human pressure. The presence of plant species in this dataset was used for to calibrate the SDM and MEM. An additional set of 298 plots representing the validation sites was surveyed using methods that were identical to those applied to the training dataset. The validation sites were previously shown to be spatially independent from the training sites and valid for model evaluation by calculating the spatial correlation of the SSDM residuals between the calibration and the evaluation datasets based on neighbourhood graphs and Moran's I coefficient (Pottier *et al.* 2013). A total of 241 vascular plant species were recorded. We considered 175 species that presented more than 30 occurrences and showed a low and quite uniform prevalence in the study area, with 70% of the species having a prevalence ≤ 0.1 and 84% having a prevalence ≤ 0.2 (species list is available in Appendix S2). No significant relationship was observed between the prevalence and elevation. To build a functional tree, we also field-sampled two uncorrelated plant traits associated with the performance of plant species during the persistence phase of their life cycle (Westoby *et al.* 1998): vegetative height (VH in mm) and specific leaf area (SLA in $mm^2 mg^{-1}$) (see Dubuis *et al.* 2013, for details on the measurements in the field). Vegetative height is a stature trait associated with the plants' ability to compete for light, and specific leaf area is related to the plants' ability to capture, use, and release resources to their environment, thereby providing a good estimate of the position of the species along the leaf economic spectrum. Each trait was log-transformed to conform to normality and scaled between 0 and 1. We then constructed a functional tree by calculating all of the pairwise dissimilarities (Gower's distances) between observations in the data set (function daisy {cluster} in R). A phylogenetic tree for these species is available in Ndiribe *et al.* (2013).

Biodiversity modelling: taxonomic component

We modelled species richness and composition using two community modelling approaches.

1) SSDM: individual species distribution models were fitted with climatic and topographic predictors calculated from data recorded by the Swiss network of meteorological stations and obtained from a digital elevation model at 25 m resolution. We selected five abiotic topoclimatic variables that previous studies reported to be relevant predictors of the distribution of plant species in this mountain environment (Dubuis *et al.* 2011): growing degree days (above 0 °C), moisture index throughout the growing season (difference between precipitation and potential evapotranspiration), solar radiation sum for the entire year, slope (in degrees), and topography (indicating the ridges and valleys). We used three modelling techniques: a generalized linear model (GLM), generalized additive model (GAM), and generalized boosted model (GBM). We created a weighted average ensemble model of the three techniques for each species and used weights from the internal cross-validation with the true skill statistic (TSS, Allouche, Tsoar & Kadmon 2006) evaluation metrics. All of the models and the ensemble were tested for their predictive ability on the evaluation dataset using both the area under the curve (AUC) of a receiver operating characteristic (ROC) plot (Swets 1988) and the TSS metric. The potential species distributions obtained for the 175 species were binarized (presence/absence) using two threshold approaches: *i)* the threshold corresponding to equal values of sensitivity and specificity (Liu *et al.* 2005); and *ii)* the threshold maximizing the TSS. The binary models were stacked to predict the assemblage richness and composition in the evaluation plots (SSDM predictions). We fitted all of the models in R (2.14.1) using the biomod package (Thuiller et al., 2009).

2) SSDM-MEM: for this approach, we began with the ensemble predictions for the SDMs that define the potential pool of species by abiotic (i.e., topoclimatic) drivers considering the same set of 175 species. We defined macroecological drivers (MEM) by modelling the observed species richness with the same environmental predictors and techniques used for the SDMs and establishing a Poisson distribution (MEM). This model

differs slightly from the one fitted by Dubuis *et al.* (2011) due to a smaller subset of species and different predictors being used in our model. We further applied the ensemble forecasting approach (as described above) to obtain a final richness prediction. We then coupled the raw ensemble predictions from the SSDMs and the richness predictions by applying the "*probability ranking*" rule (PRR) (D'Amen *et al.* 2015c). Using this procedure for each site, we selected a number of species equal to the MEM richness predictions based on the decreasing probability of presence calculated by the SDMs, thereby obtaining the binary compositional predictions.

Biodiversity modelling: functional and phylogenetic components

We used the SSDM and SSDM-MEM predictions of assemblage composition to reconstruct the functional and phylogenetic diversity facets of communities. The functional and phylogenetic diversity patterns in the evaluation dataset were estimated based on the observed and predicted assemblages. In addition, the functional and phylogenetic diversity indices were calculated using the approach proposed by Chalmandrier *et al.* (2015a), which builds on a multiplicative decomposition framework (Pavoine, Love & Bonsall 2009; Chao, Chiu & Jost 2010; Leinster & Cobbold 2012) and allows for variation in the *dominance* of species in the assemblage (importance assigned to dominant vs. rare species) and the *similarity effect* (effects of considering different scales in the tree from broad to more exact structures – see the Introduction). However, because our analyses were not based on abundance data, we fixed the first parameter in the analyses. We included and measured the strength of the *similarity effect* in the analyses of the diversity indices by applying the δ transformation (Pagel 1997) to the functional and phylogenetic trees. The parameter δ scales the overall path lengths in the tree, such as the distance from the root to the species as well as the shared path lengths. We considered nine δ values (0.01, 0.05, 0.1, 0.5, 1, 2.5, 5, 7.5, and 10) to assign higher importance to ancient branches or recent divergences by distorting the trees. Specifically, a tree that is “stretched” (i.e., distorted by the δ transformation) close to its roots assigns more weight to large distances, whereas a tree stretched close to its tips assigns more weight to small distances. This investigation involved nine estimations of the functional and phylogenetic diversity patterns in the observed data and the SSDM and SSDM-MEM predictions.

Evaluation

We calculated the taxonomic, functional and phylogenetic dissimilarities between the predictions by the SSDM and SSDM-MEM approaches and the evaluation data (Sørensen dissimilarity index). We also disentangled the two components of the Sørensen dissimilarity index, *richness difference* and *species replacement,* between the observed and predicted communities using the `betapart` R package (Baselga & Orme 2012; Legendre 2014). For taxonomic diversity, we calculated both the sensitivity (the proportion of species correctly predicted as present among the species observed as present) and specificity (the proportion of species correctly predicted as absent among the species observed as absent). Finally, we evaluated the performance of the SSDM and SSDM-MEM approaches in predicting the species richness and the functional and phylogenetic diversity indices by calculating i) Spearman's correlation (we use the non-parametric test because the variables are not normally distributed), ii) the root mean square error (RMSE, Potts & Elith 2006), and iii) the average error (AVE, Potts & Elith 2006):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$$

$$AVE_{error} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$$

These formulas are based on the sample size (n) and the discrepancy between the observed (y_i) and predicted ($\hat{y}_i$) values.

To explore variations in the predictive accuracy along the elevation gradient, all of the above statistics were also calculated by grouping the evaluation plots within 500 m wide elevation bands. Correlations for different bands were compared based on Z scores by applying Fisher's transformation (Fisher 1921). Finally, the same statistics were calculated for each value of the parameter δ to test the *similarity effect* in the predictions of functional and phylogenetic diversity indices.

Results

Accuracy of taxonomic predictions

Species distribution models for most species had an AUC value higher than 0.7; therefore, they can be considered useful for predictions. In particular, the mean and standard deviation of the AUC scores for the different techniques were GAM: 0.803 ± 0.078; GLM: 0.799 ± 0.077; and GBM: 0.783 ± 0.081 (see Appendix S2 for the evaluation statistics for all species by both AUC and TSS). The MEM prediction showed a fair correlation value between the observed and predicted values of the species richness in the evaluation dataset ($\rho = 0.529$, Spearman's rank correlation test), and the species richness error was centred on zero (Appendix S3). The SSDM based on the two threshold approaches produced similar results ($\rho = 0.504$ and 0.507 Spearman rank correlation tests); thus, we have only presented the results derived from binarization using the AUC statistic. Compositional predictions from the SSDM achieved higher sensitivity scores than the SSDM-MEM, whereas the opposite was observed for specificity (Appendix S4). In other words, the SSDM was better for predicting the presence of a species, whereas the SSDM-MEM was better for predicting the absence of a species. The SSDM-MEM predictions of community composition produced a moderate but significant decrease in the Sørensen dissimilarity index compared with the SSDM predictions (t-test $p < 0.001$) for the entire elevational gradient (Appendix S5). The RMSE for the richness component of taxonomic diversity was higher with the SSDM than the SSDM-MEM (28.52 and 8.72, respectively). Moreover, the average error (AVE) was low for the SSDM-MEM (1.09) but high for the SSDM (-23.94), thereby confirming the high overestimation of the number of species in the latter approach without the MEM constraint.

Accuracy of functional and phylogenetic index predictions and the similarity effect

According to the Sørensen dissimilarity index, the SSDM-MEM significantly increased the predictive capacity of the SSDM for the phylogenetic facet (mean of Sørensen dissimilarity index: SSDM-MEM = 0.27, SSDM = 0.35; Wilcoxon signed rank test, $p < 0.001$) (Fig. 1 and Appendix S5). The two components of Sørensen's dissimilarity index showed similar patterns for all of the biodiversity facets and indicated that the SSDM-MEM framework provided significantly improved predictions compared with the SSDM for the richness component but worse predictions for species replacements in communities (Fig. 1 and Appendix S5). The difference in the prediction errors of functional and phylogenetic diversity showed that the former was better predicted than the latter by both modelling approaches (Wilcoxon test, $p < 0.001$) (Fig. 2, Appendix S6). The functional and phylogenetic diversity predictions derived from the SSDM-MEM framework out-performed those derived from the SSDM. In fact, the SSDM-MEM predictions obtained a lower average RMSE (functional index: mean RMSE for

SSDM-MEM = 2.68 and for SSDM = 6.88, phylogenetic index: mean RMSE for SSDM-MEM = 3.48 and for SSDM = 9.30). According to the AVE, the predictions derived from the SSDM-MEM accurately retrieved both the functional and phylogenetic diversity indices, whereas the predictions derived from the SSDM produced greater errors, which were more pronounced in the phylogenetic component (functional index: mean AVE for SSDM-MEM = 0.41 and for S-SDM = -5.76, phylogenetic index: mean AVE for SSDM-MEM = 0.02 and for S-SDM = -7.81) (Fig. 2). Considering the similarity effect, assigning more weight to ancient divergences ($\delta \leq 0.01$) lowered the differences in the errors between the SSDM-MEM and SSDM for functional and phylogenetic diversity. Both the RMSE and AVE increased for a medium to strong similarity effect, although the increase was more pronounced for the SSDM predictions than for the SSDM-MEM predictions (Fig. 2; Appendix S6).

Accuracy of predictions along the elevation gradient

Overall, the SSDM-MEM predictions were better than the SSDM predictions across the elevation gradient for the *dissimilarity* and (especially) the *richness difference* index, as expected. For the SSDM-MEM predictions, the *Sørensen dissimilarity index* and its components increased with elevation except for the *richness difference* component of taxonomic diversity, which was low and constant across the whole gradient (Fig. 1). This trend was not observed for the SSDM predictions (Fig. 1). Nevertheless, we observed the reverse pattern for the *species replacement* component when considering taxonomic diversity (Fig. 1). For the taxonomic component of the *Sørensen dissimilarity index*, the accuracy improvement produced by the SSDM-MEM was not evenly distributed across the elevation gradient. The SSDM-MEM produced a 20% improvement in performance over the SSDM in the low elevation bands (up to 1500 m); however, this improving trend decreased as the elevation increased to 2500 m and was no longer noticeable at higher elevations (Fig. 1). The improvement produced by the SSDM-MEM for the taxonomic component considering the *richness difference* index was high and ranged from 65% at lower elevations to 50% at higher elevations.

The SSDM-MEM approach was better at reproducing species patterns along elevations for the species richness and functional and phylogenetic diversity indices and presented the lowest deviations from the observed values (Appendix S7). Comparing the accuracy across elevation bands, the predictions for diversity indices in the 2000-2500 m band showed significantly higher ρ correlation coefficients than the other elevation bands based on both modelling approaches and all levels of the *similarity effect* (all δ values for the tree transformations) (Fig. 3). The RMSE and AVE values were not significantly dissimilar across the elevation bands for the SSDM-MEM predictions of the functional and phylogenetic diversity indices (Fig. 3). However, the SSDM prediction errors varied for moderate to high similarity effects, with the highest error in the 1500-2000 m band and a decreasing error trend observed from low to high elevations (Fig. 3).

Discussion

This study is the first to compare different modelling approaches to predict the spatial patterns of multiple biodiversity facets. Our results show that the SSDM coupled with the MEM (SSDM-MEM) can moderately but significantly improve the predictions of the taxonomic, functional, and phylogenetic diversity of plant assemblages. Moreover, the accuracy of the SSDM predictions for the diversity indices varied greatly along the elevation gradient and for different intensities of the *similarity effect*, whereas the SSDM-MEM predictions were more stable. The predictive improvements observed via the application of the SSDM-MEM were likely related to the implementation of the following two steps: 1) the correction of species overestimations produced by the SSDM by imposing a limit on the number of species in each site, which produced a cascade effect on the quality of the derived functional and phylogenetic indices; and 2) the application of the "probability ranking" rule to sort the species that can co-occur at the studied site (D'Amen *et al.* 2015c).

As expected, the species composition was the most difficult diversity component to predict (Thuiller *et al.* 2015). Considering the direct dissimilarity in the comparison between the observed and predicted assemblages, the SSDM-MEM improved the *richness difference* component of the Sørensen index, which we used as a measure of the capacity of the model to predict the number of species correctly (Baselga 2010; Legendre 2014). Thus, the improved predictions of the taxonomic component by the SSDM-MEM were likely driven by the corrected number of species predicted and the removal of the least probable species. However, this removal likely also produced a degree of degradation of the *species replacement* component, i.e., the capacity of the model to correctly predict the species present in the observed community (see Baselga 2010). These results are consistent with our findings regarding the sensitivity and specificity indices in the compositional predictions, with the SSDM-MEM improving the specificity but producing worse scores for the sensitivity index. Overall, the balance between the improved accuracy of the *richness difference* and the decreased accuracy of the *species replacement* indicates that the SSDM-MEM approach can produce a moderate improvement in the SSDM predictions but not for all of the components.

The results reported above suggest that in our study, certain species were not correctly removed by the *probability ranking* rule (PPR) when attempting to control for SSDM overprediction. The PPR has been assumed to represent a putative ecological assembly rule that translates competitive strength as a function of environmental suitability, meaning that species will be more competitive in their most favourable habitats (D'Amen *et al*. 2015a). Species with the highest probability of occurrence are expected to be better adapted to the environmental conditions of the site, thereby leading to higher competitive ability. A more methodological justification of the PPR performance is that by removing the species with the lowest probability of occurring at a site, the species with the highest probability of being overpredicted by the SSDMs are also removed. A potential problem with the PPR as currently implemented is that it may depend on the species prevalence in the study area. The most frequent species are likely to have higher a probability of occurrence than the least frequent species (pers. obs.), which may in turn affect the ranking of the species and produce communities that are always composed of the most common species, although good evaluation scores are still obtained. This bias should be analysed in future studies by testing whether and how rescaling the probability outputs from the SSDM-MEM to correct for prevalence (e.g., via changing the weighting of the presences and absences) might improve the predictions. Alternative rules that also consider historical or anthropogenic factors could be implemented in further developments of the SSDM-MEM framework to improve the species selections from the abiotic species pool (SSDM) to match the richness predictions by the MEM.

Upon “aggregating” the diversity components to estimate the functional and phylogenetic diversity indices, the performance of the models is increased, which is likely because these indices can be predicted equally well for different sets of species with similar functional or phylogenetic characteristics. However, in such cases, the taxonomic diversity component would lose prediction accuracy. The SSDM-MEM predictions of the species composition led to accurate predictions of the functional and phylogenetic indices, suggesting that the species selected incorrectly in the final community composition were at least functionally or phylogenetically close to the correct ones, although this was not always true for the SSDM predictions. The similarity effect (Chalmandrier *et al.* 2015b) exerted a substantial impact on the detection of patterns in the functional and phylogenetic diversity indices in our plant communities. The prediction errors varied with the weight assigned to large *vs.* small species similarities when calculating diversity indices, with a much stronger variation in accuracy observed for the SSDM compared with the SSDM-MEM. In particular, the predictions for the phylogenetic and functional diversity indices were more accurate for a weak similarity effect. The error in the SSDM predictions increased in proportion as the strength of the similarity effect increased. This result can be interpreted as related to the increasing importance of biotic interactions *vs.* environmental abiotic filtering when the assembly of increasingly closely related lineages is considered (Ndiribe *et al.* 2013). Because the SSDM primarily accounts for the effect of abiotic drivers in shaping the community structure, it performs better when greater emphasis is placed on large species similarities. Conversely, the SSDM-MEM framework could predict recent and profound node structures of the communities, meaning that it shows good performance when a greater emphasis is placed on large or small species similarities. This result is likely because of the SSDM-MEM framework's ability to set a macroecological constraint on community richness. The SSDM produces considerable errors when predicting phylogenetically closely related species, which may indicate that the latent underlying mechanisms related to the phylogenetic community structure may not have been considered in the modelling approach.

The SSDM-MEM framework was the best technique for reproducing the observed diversity patterns along elevation. A varying degree of prediction errors in species richness along elevation has been previously reported for the SSDM. This variation was mainly caused by the differences in the degree of overprediction along this gradient (Pottier *et al.* 2013). However, we observed a different trend when considering the Sørensen dissimilarity index. The SSDM-MEM index predictions showed an overall tendency towards lower accuracy with increasing elevation, especially for the functional component. This trend was not noticeable for the SSDM predictions.

According to the stress gradient hypothesis (Normand *et al.* 2009), the influence of the environmental filter and biotic interactions is expected to vary with elevation. Biotic interactions should be more intense or frequent under low-stress conditions (i.e., middle to lower elevations), whereas the effect of the environment should be more important under high-stress conditions (i.e., higher elevations; Bertness & Callaway 1994; Brooker & Callaghan 1998; Pottier *et al.* 2013). The performance of the SSDMs in our study system conformed to balance between the two assembly drivers, with improved performance at higher elevations because of the stronger relationship between the species distribution and the environment (Pottier *et al.* 2013). The decreased performance at middle elevation sites may have been caused by the increased importance of biotic interactions at these sites compared with the performance at higher elevation sites because the sites are generally more productive at middle elevations; therefore, species interactions may occur more frequently and result in the exclusion of certain of species. An alternative explanation for this result might be the disequilibrium between the species distribution data and the topoclimatic data. Under mild climatic conditions, other factors mediated by human-related (e.g., grazing or agricultural fertilization) or stochastic processes may play a more prominent role in determining the composition of the extant communities.

The reverse trend was observed for the SSDM-MEM predictions when considering the results of the *Sørensen dissimilarity index*. We found that the SSDM-MEM had a lower performance in predicting plant diversity at high elevations, where the most important community driver is the environment alone. This reduction in performance was especially evident in the SSDM-MEM predictions of functional diversity, which is the diversity facet that is expected to be most strongly influenced by the environment. However, we may have also obtained different values for the maximum species richness for each elevation band. The plots with the lowest richness were located at high elevations, which might have caused the increased performance of the SSDM in this band. This behaviour was not observed for the SSDM-MEM, thereby supporting its more stable performance, even at different richness values. However, this potential bias does not affect the comparison of the modelling methods within each elevation band.

The patterns described above for the functional and phylogenetic diversity indices were only observed when considering moderately to more closely related lineages (i.e., for a medium to strong similarity effect), and stronger differences in performance were proportional to the strength of the similarity effect. However, with low similarity effects, the prediction error was more constant along elevation. An explanation for this result may be that an environmental filter is the main assembly driver for distantly related lineages across the entire elevation gradient, which also enables the SSDM approach to retrieve community patterns at middle elevations.

Certain peculiar characteristics of the studied system and approaches could have influenced the results. We focused on a simplified taxonomic assemblage of grassland species, and including more complex communities, such as those with prey-predator interactions, would have required a much more complex modelling strategy. Moreover, we analysed communities at a very fine spatial resolution. Biotic interactions are expected to have greater importance at this resolution (Pearson & Dawson 2003), which may in turn cause a stronger over-prediction by the SSDMs. The accuracy of taxonomic predictions should thus improve toward lower spatial resolutions for this approach (Thuiller *et al.* 2015). The influence of scale on the SSDM-MEM should be assessed in future studies, specifically to ascertain whether the benefits of this approach can be observed at different resolutions and extents. In addition, the SSDM-MEM method can be improved by including more dynamic (e.g., Keith *et al.* 2008) and/or mechanistic SDMs (e.g., Kearney & Porter 2009) when the data required to fit the models are available for all species (Guisan & Rahbek 2011). Finally, concerning the SSDM performance, it was shown elsewhere that the threshold selection may strongly influence the reliability of the predicted richness and composition (Benito et al 2013). In this paper, we presented the results from two thresholding methods, but we had also calculated a large range of other species-specific thresholds, none of which affected our results (unpubl. data).

Biodiversity integrates biotic components from genes to ecosystems. Therefore, measuring such a broad concept represents a great challenge (Purvis & Hector, 2000). Here, our goal was to test the capacity of a novel community modelling approach to predict three key facets of biodiversity - taxonomic, functional and phylogenetic - at the alpha diversity level. The approach to coupling SSDMs with richness predictions through a simple probability ranking rule produced a moderate but significant improvement over the most common and simpler approach of stacking individual species predictions. This result opens promising avenues towards improving biodiversity predictions in space and time and across broad environmental gradients using functional and phylogenetic information. However, future studies on the SSDM-MEM approach could further be expanded to test additional components of biodiversity. For example, when considering broader regions than our study area, the collective diversity may be better determined by differences between regional species pools (Guisan & Rahbek 2011) and by the turnover in biological composition between locations (i.e., beta diversity) than by the site-level diversity (i.e. alpha diversity). Measuring beta diversity could therefore be a critical complement to the alpha measures performed here. Certain modelling approaches have already been developed to predict this component as an emergent property of biodiversity (e.g., Ferrier *et al*. 2007), but often at the expense of information on species identities at the local level. Recent frameworks have been proposed that combine different levels of modelling to predict both alpha and beta diversities. For instance the dyamicFOAM framework (Mokani *et al.* 2011) combines correlative richness (alpha-diversity) models and models of compositional turnover (beta-diversity) to generate compositional data for meta-communities and gamma –diversity. A comparison of the performance of the SSDM-MEM method with such beta-diversity-level approaches represents an interesting research perspective for improving spatial modelling research and conservation applications.

References

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223-1232.

Bakkenes, M., Alkemade, J.R.M., Ihle, F., Leemans, R. & Latour, J.B. (2002) Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, **8**, 390-407.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, **19**, 134–143.

Baselga, A. & Orme, C.D.L. (2012) betapart: an R package for the study of beta diversity. *Methods in Ecology and Evolution*, **3**, 808-812.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W. & Courchamp, F. (2012) Impacts of climate change on the future of biodiversity. *Ecology Letters*, **15**, 365-377.

Bertness, M. & Callaway, R.M. (1994) Positive interactions in communities. *Trends in Ecology and Evolution*, **9**, 191–193.

Boulangeat, I., Gravel, D. & Thuiller, W. (2012) Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters,* **15**, 584–593.

Brooker, R.W. & Callaghan, T.V. (1998) The balance between positive and negative plant interactions and its relationship to environmental gradients: a model. *Oikos*, **81**, 196–207.

Calabrese, J.M., Certain, G., Kraan, C. & Dormann, C.F. (2014) Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*, **23**, 99–112

Callaway, R.M., Brooker, R.W., Choler, P., Kikvidze, Z., Lortie, C.J., Michalet, R., Paolini, L., Pugnaire, F.I., Newingham, B., Aschehoug, E.T., Armas, C., Kikodze, D. & Cook, B.J. (2002) Positive interactions among alpine plants increase with stress: a global experiment. *Nature*, **417**, 844–848.

Chalmandrier, L., Münkemüller, T., Devictor, V., Lavergne, S., & Thuiller, W. (2015a) Decomposing changes in phylogenetic and functional diversity over space and time. *Methods in Ecology and Evolution*, **6**, 109-118.

Chalmandrier, L., Münkemüller, T., Lavergne, S. & Thuiller, W. (2015b) Effects of species' similarity and dominance on the functional and phylogenetic structure of a plant meta-community. *Ecology*, **96**, 143-153.

Chao, A., Chiu, C.-H. & Jost, L. (2010) Phylogenetic diversity measures based on Hill numbers. Philosophical Transactions of the Royal Society B: Biological Sciences, **365**, 3599–3609.

Corbelli, J.M., Zurita, G.A., Filloy, J., Galvis, J.P., Vespa, N.I. & Bellocq, I. (2015) Integrating Taxonomic, Functional and Phylogenetic Beta Diversities: Interactive Effects with the Biome and Land Use across Taxa. *PloS ONE*, **10**, e0126854.

Crisp, M.D., Arroyo, M.T., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S. & Linder, H. P. (2009) Phylogenetic biome conservatism on a global scale. *Nature*, **458**, 754-756.

D'Amen, M., Dubuis, A., Fernandes, R.F., Pottier, J., Pellisser, L. & Guisan, A. (2015a) Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography,* **42**, 1255–1266.

D'Amen, M., Mateo, R.G. & Guisan, A. (2016) Species Assemblages, Macroecology, and Global Change. Encyclopedia of Biodiversity. Elsevier Inc., (This article is a revision of the previous edition by Brian J. McGill, Encyclopedia of Biodiversity (Second Edition), pp 651-666, (c) 2013, Elsevier Inc).

D'Amen M., Pradervand J-N., Guisan, A. (2015b) Predicting richness and composition in mountain insect communities at high resolution: a new test of the SESAM framework. *Global Ecology and Biogeography*, **24**, 1443-1453.

D'Amen M., Rahbek C., Zimmermann N.E., Guisan A. (2015c) Spatial prediction of community structure: state of the art and future perspectives. *Biological Review* doi: 10.1111/brv.12222

Dubuis, A., Pottier, J., Rion, V., Pellissier, L., Theurillat, J.-P. & Guisan, A. (2011) Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. *Diversity and Distributions*, **17**, 1122–1131.

Dubuis, A., Rossier, L., Pottier, J., Pellissier, L., Vittoz, P. & Guisan, A. (2013) Predicting current and future spatial community patterns of plant functional traits. *Ecography,* **36**, 1158–1168.

Elith, J. & Leathwick J.R. (2009) Species Distribution Models: ecological explanation and prediction across Space and Time. *Annual Review in Ecology and Systematics* **40**, 415–436.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1-10.

Fernandes, J.A., Cheung, W.W., Jennings, S., Butenschon, M., de Mora, L., Frolicher, T.L. & Grant, A. (2013) Modelling the effects of climate change on the distribution and production of marine fishes: accounting for trophic interactions in a dynamic bioclimate envelope model. *Global Chang Biology,* **19**, 2596–2607.

Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43**, 393-404.

Fisher, R.A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron,* **1**, 3–32.

Franklin, J. (1995) Predictive vegetation mapping: geographical modelling of biospatial patterns in relation to environmental gradients? *Progress in Physical Geography,* **19**, 474-499.

Guisan, A. & Rahbek, C. (2011) SESAM–a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography*, **38**, 1433–1444.

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

Hirzel, A. & Guisan, A. (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling*, **157**, 331-341.

Jarzyna, M.A. & Jetz, W. (2016) Detecting the Multiple Facets of Biodiversity. *Trends in Ecology & Evolution*, **31**, 527–538.

Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters* **12**, 334-350.

Keith, D.A., Akcakaya, H.R., Thuiller, W., Midgley, G.F., Pearson, R.G., Phillips, S.J., Regan, H.M., Araujo, M.B. & Rebelo, T.G. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560-563.

Lavorel, S. & Garnier, E. (2002) Predicting changes in community composition and ecosystem functioning from plant traits: Revisiting the Holy Grail. *Functional Ecology*, **16**, 545-556.

Legendre, P. (2014) Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography*, **23**, 1324-1334.

Leinster, T. & Cobbold, C.A. (2012) Measuring diversity: the importance of species similarity. *Ecology*, **93**, 477–489.

Liu, C.R., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385–393.

Mateo, R.G., Broennimann, O., Normand, S., Petitpierre, B., Araújo, M.B., Svenning, J.C., Baselga, A., Fernández-González, F., Gómez-Rubio, V., Muñoz, J., Suarez, G.M.., Luoto, M., Guisan, A. & Vanderpoorten, A. (2016) The mossy north: an inverse latitudinal diversity gradient in European bryophytes. *Scientific Reports,* **6**, 25546.

Mateo, R.G., Felicísimo, Á.M., Pottier, J., Guisan, A. & Muñoz, J. (2012) Do Stacked Species Distribution Models Reflect Altitudinal Diversity Patterns? *PloS ONE*, **7**, e32586.

McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, **21**, 178-85.

Michalet, R., Schöb, C., Lortie, C. J., Brooker, R. W. & Callaway, R. M. (2014) Partitioning net interactions among plants along altitudinal gradients to study community responses to climate change. *Functional Ecology*, **28**, 75–86.

Mokany, K., Harwood, T.D., Overton, J.M., Barker, G.M. & Ferrier, S. (2011) Combining alpha – and beta –diversity models to fill gaps in our knowledge of biodiversity. *Ecology Letters,* **14**, 1043–1051.

Mouquet, N., Devictor, V., Meynard, C. N., Munoz, F., Bersier, L.-F., Chave, J., Couteron, P., Dalecky, A., Fontaine, C., Gravel, D., Hardy, O. J., Jabot, F., Lavergne, S., Leibold, M., Mouillot, D., Münkemüller, T., Pavoine, S., Prinzing, A., Rodrigues, A. S.L., Rohr, R. P., Thébault, E. & Thuiller, W. (2012) Ecophylogenetics: advances and perspectives. *Biological Reviews*, **87**, 769–785.

Münkemüller, T., Boucher, F. C., Thuiller, W. & Lavergne, S. (2015) Phylogenetic niche conservatism–common pitfalls and ways forward. Functional *Ecology*, **29**, 627-639.

Ndiribe C., Pellissier L., Antonelli S., Dubuis A., Pottier J., Vittoz P., Guisan A., Salamin N. (2013) Phylogenetic plant community structure along elevation is lineage specific. *Ecology and Evolution* **3**, 4925-493.

Ndiribe, C., Pellissier, L., Dubuis, A., Vittoz, P., Salamin, N. & Guisan, A. (2014). Plant functional and phylogenetic turnover correlate with climate and land use in the Western Swiss Alps. *Journal of Plant Ecology*, **7**, 439–450.

Normand, S., Treier, U.A., Randin, C., Vittoz, P., Guisan, A. & Svenning, J.C. (2009) Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. *Global Ecology and Biogeography*, **18**, 437-449.

Pagel, M. (1997) Inferring evolutionary processes from phylogenies. *Zoologica Scripta*, **26**, 331-348.

Pavoine, S., Love, M.S. & Bonsall, M.B. (2009) Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically-structured species assemblages: application to rockfish (genus: Sebastes) in the Southern California Bight. *Ecology Letters*, **12**, 898-908.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.

Pineda, E. & Lobo, J.M. (2012) The performance of range maps and species distribution models representing the geographic variation of species richness at different resolutions. *Global Ecology and Biogeography*, **21**, 935–944.

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C.F., Vittoz, P. & Guisan, A. (2013) Accuracy of plant assemblage predictions from species distribution models varies along environmental gradients. *Global Ecology and Biogeography*, **22,** 52-63.

Potts, J. M & Elith, J. (2006) Comparing species abundance models. *Ecological Modelling*, **199**, 153–163.

R Development Core Team. (2011) R: A language and environment for statistical computing.

Rosauer, D.F., Ferrier, S., Williams, K. J., Manion, G., Keogh, J.S. & Laffan, S. W. (2014) Phylogenetic dissimilarity modelling: a new approach to analyzing and predicting spatial turnover in the phylogenetic composition of communities. *Ecography*, **37**, 21-32.

Swets, J.A. (1988) Measuring the Accuracy of Diagnostic Systems. *Science*, **240**, 1285-1293.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369-373.

Thuiller, W., Münkemüller, T., Lavergne, S., Mouillot, D., Mouquet, N., Schiffers, K. & Gravel, D. (2013) A road map for integrating eco-evolutionary processes into biodiversity models. *Ecology Letters*, **16**, 94-105.

Thuiller, W., Maiorano, L., Mazel, F., Guilhaumon, F., Ficetola, G.F., Lavergne, S., Renaud, J., Roquet, C. & Mouillot, D. (2015) Conserving the functional and phylogenetic trees of life of European tetrapods. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**, 20140005.

Thuiller, W., Pollock, L.J., Gueguen, M. & Münkemüller, T. (2015) From species distributions to meta-communities. *Ecology Letters*, **18**, 1321-1328.

Vamosi, S.M., Heard, S.B., Vamosi, J.C. & Webb, C.O. (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, **18**, 572-592.

Violle. C., Navas, M-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–92.

Westoby, M. (1998) A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, **199**, 213–227.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882-892.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475-505.

Webb, C.T., Hoeting, J.A., Ames, G.M., Pyne, M.I. & LeRoy Poff, N. (2010) A structured and dynamic framework to advance traits-based theory and prediction in ecology. *Ecology Letters*, **13**, 267–283.

White, P.J. & Kerr, J.T. (2006) Contrasting spatial and temporal global change impacts on butterfly species richness during the 20th century. *Ecography,* **29**, 908–918.

Author's Contributions

MD, RGM, JP, and AG developed the concept of the study, with further contributions by WT, LP and LM, and all contributed to the final design of the analyses. AG provided the plant distribution data and CN and NS developed the phylogenetic tree. MD, RGM, JP analysed the data. MD, RGM, AG drafted the manuscript. All authors contributed to the interpretation of results and to revise the manuscript critically.

SUPPORTING INFORMATION

Additional supporting information can be found in the online version of this article:

Appendix S1: Map of the study area with sampling plots.

Appendix S2: SDM evaluation results for each species.

Appendix S3: Macroecological model results.

Appendix S4: Comparison of the SSDM and SSDM-MEM predictions based on the sensitivity and specificity indices.

Appendix S5. Comparison of the SSDM and SSDM-MEM predictions for the three diversity components based on the Sørensen dissimilarity index and its components.

Appendix S6. Variation of the root mean square error for model predictions for the FD and PD indices at different similarity δ values.

Appendix S7: Values of Spearman's correlation between the observed and predicted FD and PD indices for the SSDM-MEM and SSDM approaches at each similarity parametrization.

Figure 1: Mean values for the Sørensen dissimilarity index and its two components (Sørensen richness difference and Sørensen species replacement) calculated for the predictions of taxonomic (TD), functional (FD), and phylogenetic (PD) diversity from the SSDM and SSDM-MEM models for each elevation band.

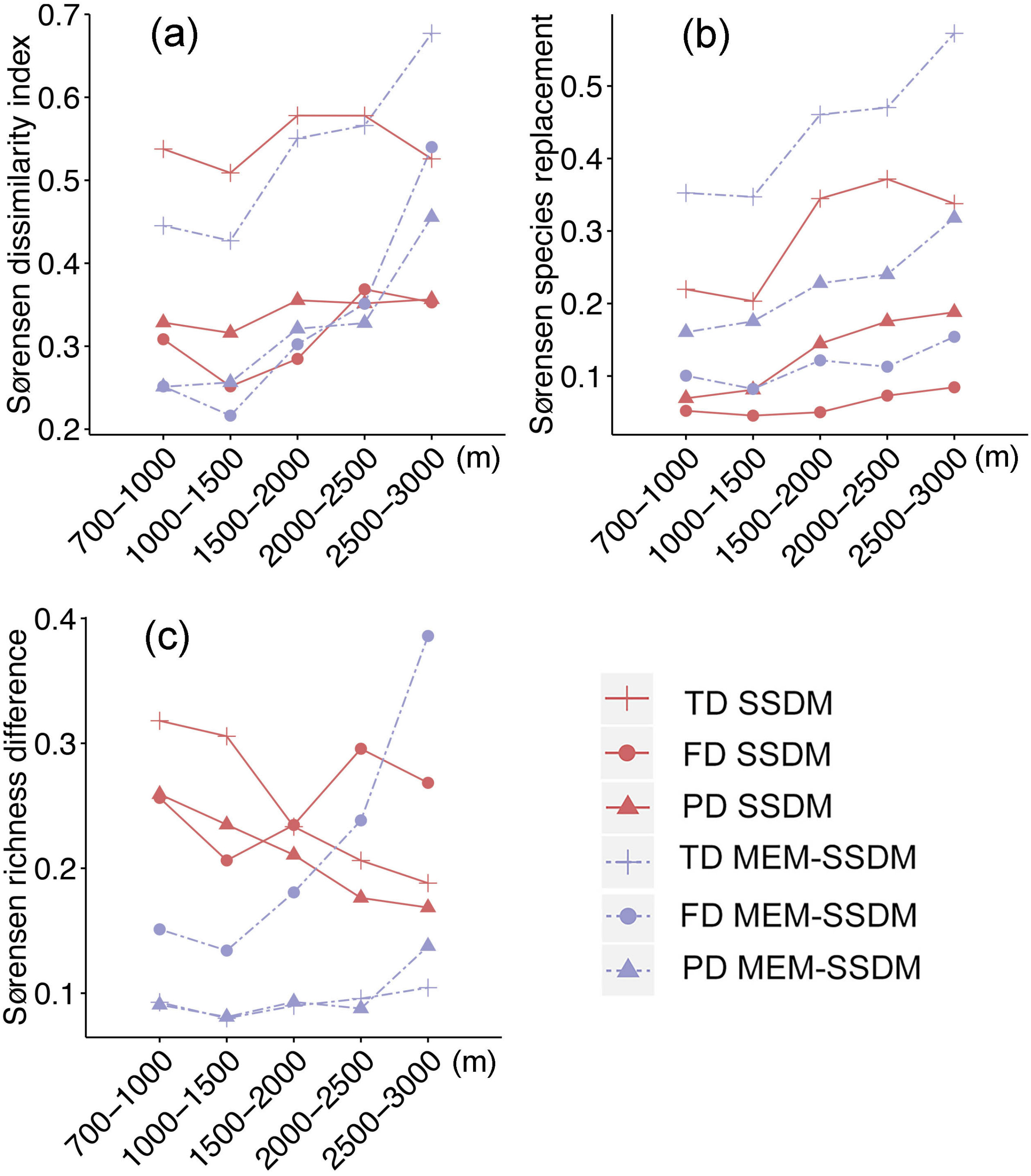

Figure 2: Variation of the average error (AVE) for the SSDM and SSDM-MEM predictions for the functional and phylogenetic diversity indices for different intensities of the similarity effect (different values of the δ parameter transform the functional and phylogenetic trees: values closer to 0 indicate that a greater importance is assigned to ancient branches, whereas values closer to 10 indicate that the transformation assigns more weight to recent divergences).

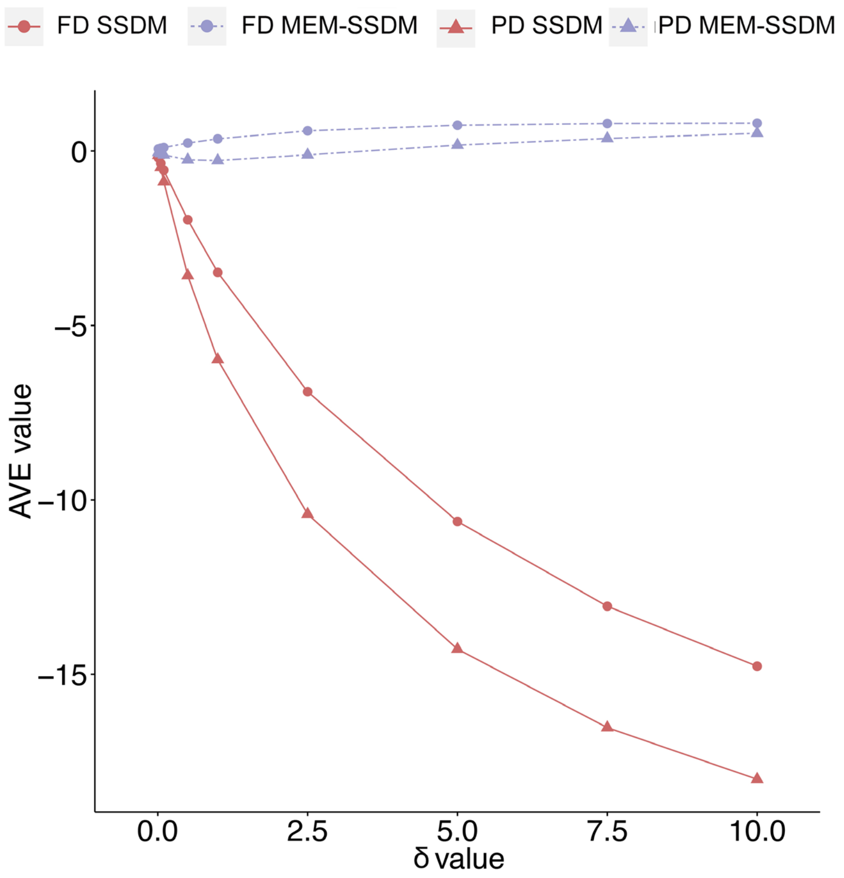

Figure 3: Evaluation statistics (root mean square error, RMSE; and average error, AVE) across elevation bands for the functional diversity index and all similarity δ values, derived from the SSDM and SSDM-MEM predictions.