LETTER

From species distributions to meta-communities

Wilfried Thuiller,[1,2]* Laura J. Pollock,[1,2] Maya Gueguen[1,2] and Tamara Münkemüller[1,2]

[1]Université Grenoble Alpes, Laboratoire d'Écologie Alpine (LECA), Grenoble F-38000, France
[2]CNRS, Laboratoire d'Écologie Alpine (LECA), Grenoble F-38000, France

*Correspondence: E-mail: wilfried.thuiller@ujf-grenoble.fr

Abstract

The extent that biotic interactions and dispersal influence species ranges and diversity patterns across scales remains an open question. Answering this question requires framing an analysis on the frontier between species distribution modelling (SDM), which ignores biotic interactions and dispersal limitation, and community ecology, which provides specific predictions on community and meta-community structure and resulting diversity patterns such as species richness and functional diversity. Using both empirical and simulated datasets, we tested whether predicted occurrences from fine-resolution SDMs provide good estimates of community structure and diversity patterns at resolutions ranging from a resolution typical of studies within reserves (250 m) to that typical of a regional biodiversity study (5 km). For both datasets, we show that the imprint of biotic interactions and dispersal limitation quickly vanishes when spatial resolution is reduced, which demonstrates the value of SDMs for tracking the imprint of community assembly processes across scales.

Keywords

Assembly rules, biotic interactions, dispersal, diversity patterns, spatial scale, species distribution models.

Ecology Letters (2015)

INTRODUCTION

Understanding the factors controlling the spatial distribution of species has long fascinated ecologists (e.g. de Candolle 1855; Chase & Leibold 2003). At the forefront of this research, species distribution models (SDMs) are the pivotal tools for predicting and understanding species distributions by relating species occurrence information to environmental data (Guisan & Thuiller 2005). SDMs have benefited from the development of statistical tools together with the ever-increasing availability of species occurrence and environmental data (Ahmed *et al.* 2015). The focus on the relationship between species distributions and the environment in SDMs has been justified by the long-held belief that the environmental niche of species is the major driver of species' distributions (Soberon & Nakamura 2009; Soberon 2010). These models do not explicitly account for other potential processes such as dispersal limitation and biotic interactions (Thuiller *et al.* 2013), which are the main arguments against their use in ecology and, more importantly, against their use to predict species assemblages under novel conditions (e.g. Davis *et al.* 1998; Pearson & Dawson 2003). Recently, joint species distribution modelling has been proposed as a means of inferring residual correlation between species (Pollock *et al.* 2014) and improving estimates of species distributions by accounting for other species (Clark *et al.* 2014; Pollock *et al.* 2014). Although some biotic interactions are accounted for in joint species distribution models, they are difficult to isolate from all the other reasons on which species presences might be correlated (e.g. missing predictor variables). We also do not have a solid understanding of when such complex models are warranted.

In general, not including biotic interactions, for example, could lead to obvious caveats in the estimation of a species' niche when a better competitor excludes it from an otherwise suitable environmental zone (Araújo & Peterson 2012). The underlying and largely unanswered question is to what extent dispersal limitation and biotic interactions prevent a correct estimation of species niches, and prevent the use of SDMs to correctly infer community structure and species assemblages (Fig. S1, Pottier *et al.* 2013).

Interestingly, this latter question is intimately linked to community ecology, in which communities are thought to assemble through a series of environmental, dispersal and biotic interaction filters (Lortie *et al.* 2004). These filtering processes select the species from the regional pool that have the suitable physiological traits to maintain viable populations in the local environment, that have the dispersal traits to reach the suitable sites, and the fitness and competitive traits to handle biotic interactions (Fig. S1). These processes are thought to be important at different spatial scales ranging from biotic interactions (e.g. competition) acting at a more local scale, dispersal acting at an intermediate scale, to environment and historical processes (e.g. speciation) acting at a regional scale (Weiher & Keddy 1999). In theory, the same hierarchical filters that control community assembly are likely to be the driving forces of the distribution of species' populations and of the spatial variation in species abundances (Soberon 2010; Boulangeat *et al.* 2012). For instance, using a comprehensive framework for investigating mechanisms underlying species distributions and their abundances, Boulangeat *et al.* (2012) concluded that community-scale effects of biotic interactions and local dispersal on plant species distribution in the French Alps were strong predictors of species abundances, while environmental factors and long distance dispersal more likely explained presence-absence distribution of species. Dispersal was deemed important for explaining species' occurrence since plants are generally dispersal limited in those systems. Additionally, with a simple

Lotka–Volterra model with two species, Godsoe *et al.* (2015) recently argued that large-scale phenomena such as priority effects promoted regional coexistence and hence reduced the influence of biotic interactions on species distributions at large spatial scales.

Combining the bottom-up approach from community ecology with its focus on community assembly and the top-down approach from species distribution modelling with its focus on species' ranges might, thus, help us to better understand the scale dependency of the different processes and the resulting diversity patterns (Guisan & Rahbek 2011). At the meta-community scale, a widely accepted idea is that temporal and spatial storage effects allow regional coexistence of species that 'should', in theory, exclude each other in homogeneous environments (Chesson 2000). If this holds true, a given species might still occupy a large part of its suitable environmental space even if dispersal limitations prevent it from occupying all of it. In other words, a species distribution model (SDM) built with both local occurrence data and fine-resolution environmental layers at regional scale, should be able to give a sensible estimation of the niche of the species (Soberon 2010). At low resolution, the match between this estimated niche and the observed distribution of species should match well while, at increasingly higher resolution, more and more gaps in this match will emerge because biotic interactions lead either to local competitive exclusion or to reduced abundances and thus reduced detectability of the species. The resolution at which this switch from a good to a bad match should be expected depends on the mechanisms supporting regional coexistence (Godsoe *et al.* 2015). For example, under a spatial storage effect this depends on the spatial distribution of the environmental variables driving differential responses of species and on species dispersal capacities (Amarasekare 2003).

If SDMs are built for all species of the region and overlaid (i.e. stacked SDMs), they should be able to give some insights on the meta-community structure and predict diversity (e.g. species richness, Parviainen *et al.* 2009; Pottier *et al.* 2013). At a local scale, fine spatial resolution, stacked SDMs should over-predict species richness if biotic interactions and dispersal limitations are important since stacked SDMs will predict the potential species richness from environmental suitability and will not be able to disentangle which species is locally excluded or not, or which did not manage to invade the assemblage (D'Amen *et al.* 2015). The departure of the modelled species richness from the truth could be seen as an indication that underlying mechanisms vary along environmental gradients and/or in space (Götzenberger *et al.* 2012; Pottier *et al.* 2013). However, when model predictions are aggregated sequentially to coarser resolution, they should get closer to the true distribution of a species since dynamics are averaged over several populations. Taking this reasoning further, we argue that estimated species richness from stacked SDMs should become increasingly better predicted at aggregated spatial resolution. The sharper the improvement with the aggregation scheme, the less pervasive the effects of biotic interactions and dispersal limitations on diversity patterns and meta-community structure that are likely to be at scales typically used in biodiversity studies.

Under this reasoning, not only species richness, but also other facets of diversity should respond to spatial aggregation. Indeed, local exclusion between competing species can be analysed through the prism of limiting similarity in terms of resource acquisition (HilleRisLambers *et al.* 2012). Two species with the same environmental niche and the same traits are not supposed to coexist, only one should persist in a given site. In other words, species identity might be more difficult to predict than traits. If our stacked SDMs are not able to correctly predict species richness because of local exclusion, they might better predict the correct trait diversity. With aggregation, we might expect that trait diversity is correctly predicted faster than species richness. The same holds for phylogenetic diversity if we assume that phylogeny can be informative about niche differences (Mouquet *et al.* 2012). However, since phylogenetic diversity does not necessarily reflect traits linked to competition or dispersal, stacked SDMs might predict phylogenetic diversity less well than trait diversity, but better than species richness if phylogeny is somewhat of a surrogate for traits (Münkemüller *et al.* 2015).

More importantly, the resolution at which stacked SDMs are able to retrieve diversity patterns should give us important insights on the pervasiveness of biotic interactions and dispersal limitations (and other underlying processes) compared to the environmental and habitat effects, and guidance on which spatial scale we can safely use SDMs to predict different types of diversity.

When testing this overall hypothetical thinking with empirical data, we are facing two related challenges: (1) Local measurement errors in either occurrence data or environmental variables may lead to erroneous local SDM predictions, while, with spatial aggregation, several data are aggregated and chances are that data errors average out, (2) the aggregation of communities into larger units leads to homogenisation, fewer outliers and lower variance. Thus, an observed fit between measured and predicted diversity patterns with increasing aggregation cannot unequivocally be attributed to dispersal and biotic interaction effects. One way to meet these challenges is to accompany data from field observations with those simulated with process-based models where observation errors can be ruled out and potential homogenisation effects can be estimated with data from randomly assembled communities.

In this article, we tested our hypothetical reasoning using both empirical and simulated data. We asked (1) to what extent stacked SDMs built at fine resolution are able to predict community structure and diversity patterns, (2) and, if they do, at which spatial scale the imprints of dynamics processes (e.g. biotic interactions and dispersal) disappear, (3) whether functional diversity is more easily predicted than phylogenetic diversity and species richness at increasing spatial scale and (4) whether stacked SDMs are able to give insights about the pervasiveness of biotic interactions and dispersal across scales.

Both the empirical and simulated data confirmed our hypotheses by showing that stacked SDMs can reveal the scale dependency of coexistence mechanisms, and that the imprints of biotic interactions and dispersal vanish relatively rapidly when defining communities at incremental resolution.

This result has important consequences not only for the use of SDMs to derive biodiversity patterns, but also on their specific use for testing coexistence mechanisms.

MATERIALS AND METHODS

Empirical datasets

To test whether the distribution of species within a region can help reveal the scale dependency of processes that generate biodiversity patterns, we selected a published dataset for which all species within a region were modelled using fine resolution biotic and abiotic data (Fig. 1). This dataset published in Thuiller *et al*. (2014), consisted of an ensemble of species distribution models built for each of the 2,755 plant species that occur in the French Alps. Models were calibrated at 250 m resolution using precisely located presence and absence data (~ 250 m precision) together with fine resolution climatic, land cover and soil data. Models were found to be very accurate using a combination of evaluation metrics estimated on held-out datasets (Thuiller *et al*. 2014). Predicted probabilities of occurrence at 250 m resolutions were converted into binary

presence-absence information using a threshold that maximised the True Skill Statistics (Thuiller *et al*. 2014).

Trait information was chosen to represent the components of the leaf-height-seed syndrome of plant traits (Westoby 1998): seed mass relates to dispersal distance and establishment success, height is considered as a surrogate of species ability to intercept light, while SLA strongly relates to species relative growth rate (Westoby 1998). In addition, life form was added to reflect survival strategies and longevity. Phylogenetic information was available as a genus-level phylogeny resolved randomly at the species level by applying a birth-death (Yule) bifurcation process within each genus (for more information on the phylogenetic inference, see Thuiller *et al*. 2014).

Presence-absence predictions from the ensemble model were available for each of the species at a 250 m resolution. We then aggregated those presence-absence data obtained from the ensemble models at 500, 1000, 1750, 2500, and 5000 m resolution (from two to twenty fold aggregation). Then, for each aggregated spatial resolution that defined a given community, we stacked the predictions to derive a set of diversity indices at the pixel level.
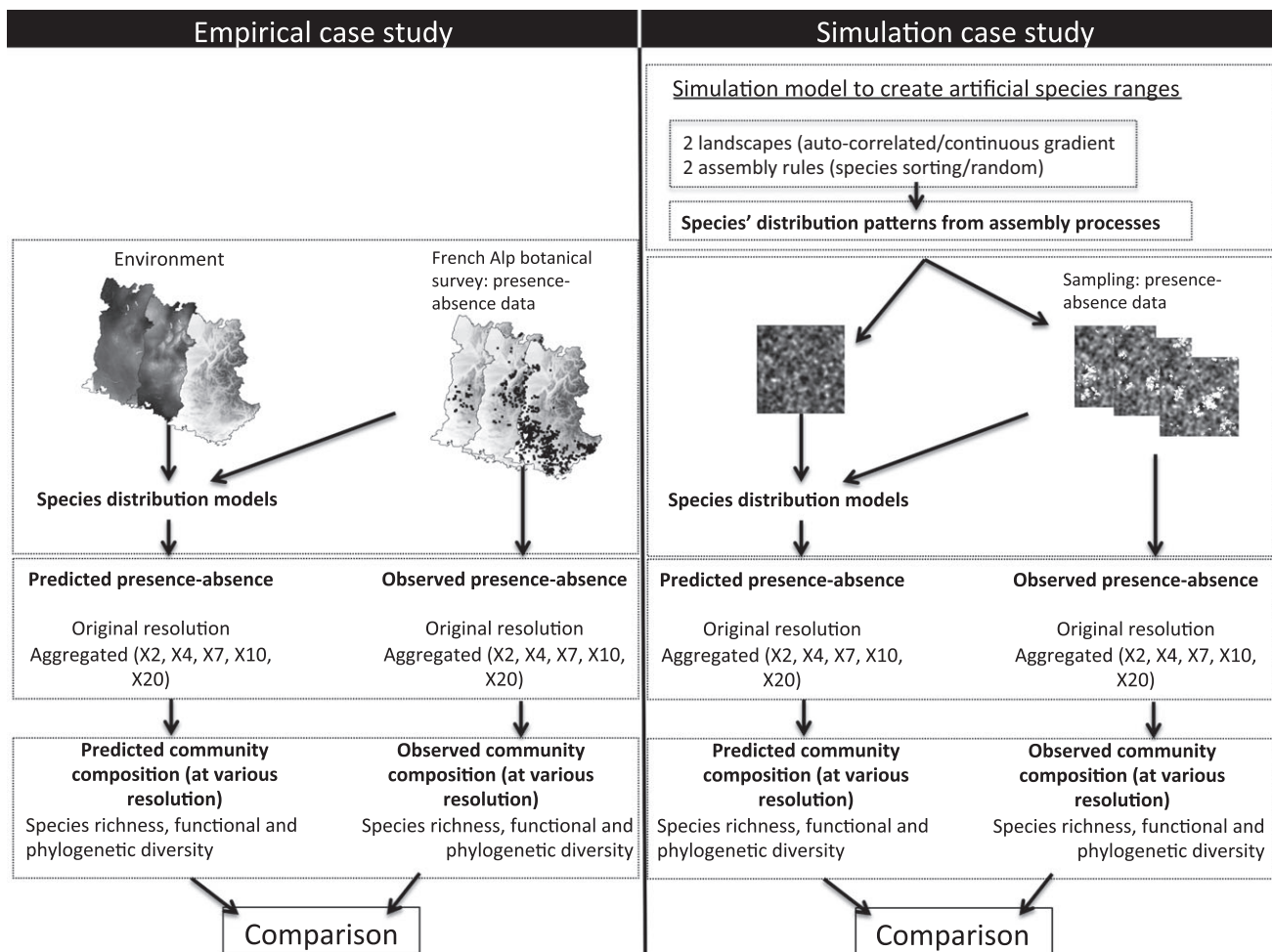

Figure 1 Workflow of the analyses.

Simulation experiment

The SDM-based datasets allowed us to support our hypotheses based on diversity estimates across an entire region. However, the true mechanisms driving the distribution of species and community structure were unknown, and results only provided evidence for potential causal links. Thus, we paired the empirical dataset with virtual data for which the mechanisms of assembly rules and species niches were explicitly modelled and the resulting distribution of species emerges from these defined processes (Fig. 1). These simulations also allowed us to test the effects of observation errors and homogenisation on our results and helped interpreting the results of the case study.

The simulation model used to generate these virtual data is a cellular automaton that simulates stochastic demographic dynamics of individuals dying, reproducing, dispersing and competing for space on a landscape grid of 300×300 grid cells with one sessile individual per cell (see Münkemüller *et al.* 2012). Across this grid, environmental conditions were spatially auto-correlated to create realistic contrasted landscapes. Each species of the species pool was assigned a trait value that represents species environmental preferences (species differ in their niche optimum but have equal niche width, maximal relative performance and fecundity). Successful colonisation depended on this trait. For example, a colonising species had a greater chance of displacing a resident species when many individuals reached that cell and the species' environmental preference closely matches environmental conditions of the cell. We chose a species pool consisting of 50 species with equal abilities to disperse, but environmental preferences that varied between species. Species environmental preferences were distributed equally across the environmental niche space. For initialisation, species were randomly distributed in the landscape, then demographic dynamics drove changes in species' distribution. After 5000 time steps, we sampled our virtual data from the grids of simulated species distributions. We uniformly placed 1764 'sampling sites' of 5×5 grid cells on each grid with a 2-grid cell buffer separating 'sample sites', and recorded species presence and absence and averaged environmental conditions within these sampling sites. The whole simulation was repeated 10 times yielding a dataset structurally equivalent to the empirical dataset. As a proof of concept and for testing for potential effects of homogenisation through aggregation, we generated a second simulated dataset in which species assembly was completely random.

We analysed the two sets of virtual data (species sorting and random assembly) in the exact same way as the empirical data from the Alpine plant case study. First, we modelled the distributions of each of the simulated species using a simple logistic model against the environmental gradient. Second, the built SDMs were then used to predict the probability of presence of each species within the sampling sites. Probabilities of presence were then converted into presence and absence data using a threshold that maximised the True Skill Statistics. Like in the empirical dataset, we obtained a matrix of predicted presence-absence data for all sampling sites. Finally, we aggregated these predicted presence-absence data with spatial aggregation coefficients (2–20-fold) equivalent to those used for the case study and derived a set of diversity indices at each spatial resolution (α-diversities).

Diversity decomposition

To compute diversity metrics incorporating species differences (such as phylogenetic divergence of functional dissimilarity) we used the same metric used in Thuiller *et al.* (2014) that relied on Leinster & Cobbold's (2012) framework that builds on a generalisation of Hill's numbers (Hill 1973). We applied this framework to estimate α-diversity for taxonomic, phylogenetic and functional diversity. α-diversity was estimated as the local diversity within each pixel for each diversity metric (following eqn. 1). The general formula calculates the diversity D for a relative abundance vector p = {pi} of the S species present in the pixel, and a matrix Z containing the similarities Z_{ij} between species i and j:

$$^D(p) = (\sum_{i=1}^{S} p_i (\sum_{j=1}^{S} Z_{ij} p_j)) \tag{1}$$

The α-diversity of each pixel was calculated from the vector of species presence-absence per pixel. The similarity matrix, Z, was calculated as one minus the cophenetic distance between species for phylogenetic diversity and the Gower distance for the four selected traits (SLA, height, seed mass and life form) for trait diversity, divided by the maximum respective distance to have Z bounded by 0 and 1.

α-diversity was estimated for both the simulation experiments and the case study and for all different grain sizes, and for taxonomic, functional and phylogenetic diversity.

Analyses

For both the empirical (Alpine plants) and simulated datasets (species sorting and random assembly), we first tested the ability of stacked SDMs to retrieve the structure of the communities at different spatial resolution. For that, we built on the True Skill Statistic (TSS, Allouche *et al.* 2006) that is usually calculated for each species by comparing observed and predicted ranges. Here, we used it differently. For a given pixel at a given resolution, we calculated an overall TSS for all species together. The misclassification matrix was calculated for each pixel, and filled with the number of species correctly predicted (as a presence or absence) and incorrectly predicted. We calculated the TSS metric for the empirical dataset by comparing the predicted dataset to the observed, while for the simulated dataset, we compared the simulated data to the predicted one. This analysis was carried out for each of the spatial resolutions.

Second, for the empirical dataset, we also investigated whether the TSS between observed community structure and that predicted from stacked SDMs changed with elevation at the different spatial resolutions. This latter analysis was carried out to test the hypothesis that in stressful and harsh conditions (high elevation), species distribution and community structure are strongly driven by abiotic components rather than biotic interactions (Mitchell *et al.* 2009). In this case,

stacked SDMs are expected to predict biodiversity patterns to perform relatively well, compared to low elevation and more productive sites where competitive exclusion might lead to erroneous predictions from stacked SDMs (Pottier *et al.* 2013).

Finally, for both the empirical and simulated datasets, we compared the predictions of community patterns from stacked SDMs to the observed/simulated community ones at the different spatial resolutions using spearman rank correlation for the three diversity facets (taxonomic, functional and phylogenetic).

RESULTS

Empirical dataset

The true skill statistic (TSS) that represents the adequacy of the stacked SDMs to correctly represent species presence and absence per pixel was, on average, very good at all resolutions. However, the average TSS decreased slightly with aggregation contrary to our expectation that TSS would increase (Fig. 2). However, the range of variation in TSS between pixels was remarkably wide at the finest resolution with TSS values up to −0.6, which means that it was possible to almost predict the opposite communities than observed. Conversely, although the mean TSS was slightly lower at the lowest resolution (20-fold aggregation), the range of variation in TSS was quite low (Fig. 2). In other words, the commnity

Figure 2 The true skill statistics (TSS) as a function of the aggregated resolution for modelled and observed distributions in the case study. TSS was calculated for each pixel by comparing correctly and incorrectly predicted species. To be able to compute the TSS, all four values of the misclassification table were needed, which means it could not be calculated for pixels in which we only had occurrence data and thus no values for false positive and true negative components. Percentages of pixels with TSS values are given for each aggregated resolution, and mean values are in red.

structure was generally well predicted (min TSS > 0.5). When analysing how TSS behaved with elevation (Fig. S2), the match between observed and predicted community structure generally increased with the mean elevation of the pixel, which was the case for all resolutions considered (i.e. the definition of a community). The range of variation in TSS was also generally wider at low elevation than at high elevation. At high elevation, in stressful conditions, the community structure was, thus, better predicted than at low elevation.
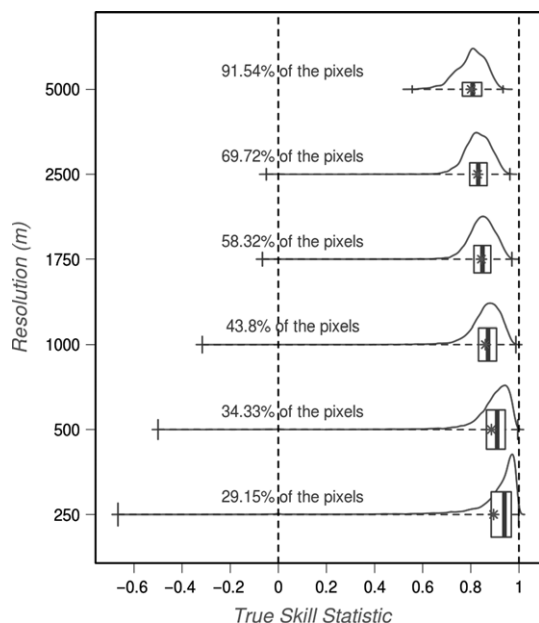
In general, the different facets of diversity were relatively well predicted by the stacked SDMs, and the quality of the prediction increased markedly with aggregation, and, thus, the larger communities (Fig. 3, left column). The shape of the improvement was, however, different between the three facets. While the correlation between observed and predicted species richness at the original fine resolution was close to 0.55, it was at 0.9 for functional α-diversity, and in between for phylogenetic diversity (ρ = 0.75). This correlation reached a plateau close to 1 at fourfold aggregation level (1000 m resolution), while species richness only reached a plateau at 10-fold aggregation level (2500 m resolution).
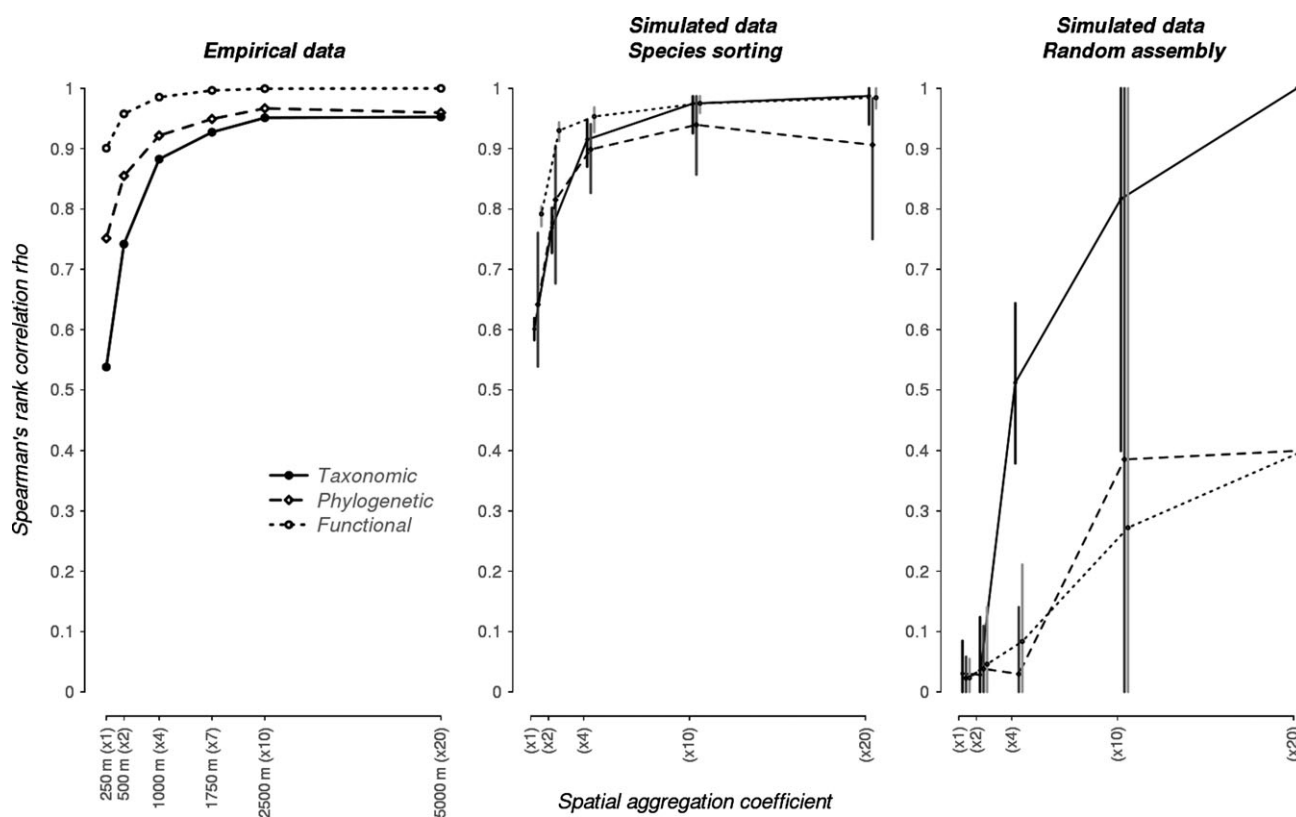
Simulation experiments

As expected, SDMs from the simulated experiment were much better at predicting species distributions (and diversity measures) generated by a species sorting than by a random assembly process. Under a process of species sorting, both functional and phylogenetic α-diversity were initially relatively well predicted at the finest resolution (ρ = 0.79 and 0.64 for functional and phylogenetic diversity respectively), and the correlation was nearly 1 for functional diversity after a fourfold aggregation (Fig. 3). Species richness was predicted moderately well at the original fine resolution (ρ = 0.6) and prediction performance increased more sharply than for the other two diversity facets with aggregation. For random assembly, species richness was only retrieved after a 10- to 20-fold aggregation, probably mostly due to sampling effects. Functional and phylogenetic diversity were never correctly predicted at any spatial resolution. Overall, results from the empirical dataset closely matched the simulated data under a species-sorting assembly process and an auto-correlated environmental gradient, but differed markedly with the simulations under random assembly.

DISCUSSION

In this article, we tackled the challenge of testing whether modelled ecological niches at fine resolution are able to correctly predict community structure and diversity patterns. We analysed this at incremental scales and by using different biodiversity facets that are supposed to provide complementary aspects of the structure of assemblages and community assembly rules (Mouquet *et al.* 2012; Cadotte *et al.* 2013; Chalmandrier *et al.* 2015). The key results from our study are that, indeed, fine resolution stacked SDMs predict community structure and community diversity patterns very well, that the fit between observed and predicted biodiversity patterns increases sharply with aggregated resolution and that, in con-

Figure 3 Spearman rank correlation (*rho*) between observed and predicted diversity metrics estimated from stacked-SDM for the empirical datasets (first column) and the simulated datasets through an auto-correlated gradient under species-sorting assembly (second column) and under random assembly (last column) for the different aggregated spatial resolution.

cordance with theoretical expectations, this increase depends on the facet of diversity considered. These results suggest that the imprint of competitive exclusion and dispersal quickly vanishes from diversity patterns and community structure at broader resolution. Finally, although they only focused on pairwise interactions and climate, we corroborate findings by Araújo & Rozenfeld (2014), from which they conclude that the spatial signature of competition is unlikely to be pervasive beyond local and regional scales. We confirm these results from the French Alps case study with those from the simulation case study using an auto-correlated landscape. Specifically, these simulations suggest that neither observation errors nor homogenisation effects were driving our results because there was no observation error, and results could not be produced for data that were simulated without competition (i.e. there was a much shallower increase in the fit between observed and predicted biodiversity patterns with aggregated resolution under random assembly). Results from simulations with a continuous gradient remain similar, although community structure was retrieved more slowly with decreasing spatial resolution (Fig. S3). This is in accordance with our expectation since the French Alps are highly heterogeneous with contrasted auto-correlated gradients. Thus, the combined analysis of field and simulated data strongly suggests that an overall dominant environmental filter in combination with spatial and temporal storage effects are the most likely candidates driving meta-community assembly. An alternative explanation for such a result might also be that the environmental

factors co-vary with biotic processes. If such covariance exists, then it may be captured by SDM models and will ultimately mask the effects of biotic interactions at all scales. Although such covariance might indeed exist in nature in very simplistic systems, this is, however, unlikely to be the case in highly heterogeneous landscapes such like the French Alps when considering that thousands of our plant species might interact for various aspects such like light, soil resources or mycorrhiza that are not necessarily related to environmental drivers.

Interestingly, for both case studies, functional diversity is the best predicted diversity pattern from stacked SDMs, even at the finest resolution where the effects of local interactions are expected to be the strongest. Functional diversity is directly linked to the response of species to the environment. The four plant functional traits used here, especially height and specific leaf area, are traits known to be tightly linked to temperature and humidity (Westoby 1998; de Bello *et al.* 2013). This is especially true in harsh environment like the French Alps at both the intra and interspecific level (Albert *et al.* 2010). The decrease in plant height with increasing elevation is a well-known physiological response of plant to frost damage and extreme wind (Körner 1999). The same applies to specific leaf area (SLA) known to decrease with elevation (i.e. temperature), associated with stress tolerance and an efficient strategy for nutrient conservation (low specific leaf area and high leaf dry matter content, Choler 2005). In other words, functional diversity is usually better predicted than taxonomic diversity. This is because the environment first stratifies the

species with the correct traits, something that we could call functional trait sorting (de Bello *et al.* 2012), before species interactions result in the exclusion of some of these species. The correct distribution of traits is relatively easily predicted by stacked SDMs and becomes very well predicted already at relatively small grain sizes, while the number of species fulfilling these trait requirements and being able to coexist is more difficult to predict. The same holds for the simulated data since the trait is directly linked to the response of species to the environment. Competition for space, limiting similarity and founder effects are likely to act locally, and explain the poor prediction of species richness at fine resolution. Interestingly, species richness requires more than a fourfold aggregation (1000–1750 m resolution) to achieve the same predictive accuracy as functional diversity at a 250 m resolution (confirmed by the TSS analyses). This likely means that the local competition, dispersal limitation and stochastic events allow coexistence of ecologically suitable and similar niches at intermediate scales.

The most striking result is that the simulation experiments under a species-sorting process yielded a nearly identical pattern with even the same scaling factor. We can obviously not conclude that there is a universal scaling law to retrieve biodiversity patterns from modelled ecological niches at fine resolution, but this is a noticeable result that asks to be confirmed or tested in other studies. Moreover, simulation experiments under random assembly give very different results. As aggregation of random communities leads to very strong homogenisation (i.e. neighbouring communities are dissimilar), this strongly suggests that the observed sharp increase in the fit between observed and predicted biodiversity patterns is due to ecological processes and not due to homogenisation effects or measurement errors.

Results are very similar when it comes to phylogenetic diversity. Interestingly, phylogenetic diversity is better predicted than species richness across scale but always lags behind functional diversity for both the empirical and simulated datasets. This means that although the distribution of ecological niches of the entire region represents the functional adaptation of species to environment well (i.e. functional diversity is well predicted even at fine resolution), this is slightly less true for phylogenetic diversity. In other words, phylogenetic diversity probably represents some, but not all functional aspects. A similar result was recently found for plant grasslands in an Alpine watershed, where it was shown that different environments along gradients filtered different species according to their functional traits, while the same competitive lineages dominated communities across the gradients (Chalmandrier *et al.* 2015).

Our results corroborate recent findings by Pottier *et al.* (2013). They showed that species richness (in absolute values) was more strongly over-predicted at low elevation than at high elevation where it was slightly underestimated. They argued that important environmental filters were missing to regulate predicted species richness at high elevation from stacked SDMs, while non-equilibrium situations with climate (i.e. local exclusion) might explain the under-prediction at low elevation. Our empirical case study with a larger extent and our simulated data allow us to go a step further. The fact that functional and phylogenetic diversity are always better predicted than species richness along elevation (Fig. S4) supports the idea that competitive exclusion is much stronger at low elevation—the right traits are predicted, the right lineages, but not the right number of species. Competitive exclusion thus removes species with similar traits but with lower fitness differences at low elevation (Kraft & Ackerly 2010). At high elevation, when conditions get more extreme, the pool of species having suitable traits and environmental tolerances is much lower, which limits species exclusion and produces better predictions from stacked SDMs.

CONCLUSION

In this article, we asked whether modelled distribution of species and their aggregation at different spatial resolutions could help reveal when community assembly processes might impact biodiversity patterns. The answer is yes. This is especially true when empirical results are compared to simulated datasets, and when ecological niche modelling is carried out at a very fine resolution using high-quality biotic and abiotic data. More importantly, the joint analysis of taxonomic, functional and phylogenetic diversity at incremental spatial resolutions provides a unique perspective on these potential community assembly processes. We suggest that the imprint of biotic interactions and dispersal on biodiversity distribution is relatively limited in scale, diminishes relatively rapidly and has nearly vanished at resolutions typical of biodiversity analyses. Although this result might only hold for spatially heterogeneous systems that are not dominated by a single superior competitor, it confirms that adequate and well-designed species distribution models (SDMs), calibrated at fine resolution, are able to make inference and predictions on other facets of diversity such as functional and phylogenetic diversities.

AUTHORSHIP

W.T. and TM designed the study, WT carried out the species distribution models and wrote the initial draft of the paper, TM built and ran the simulation model, and provided results, MG carried out the spatial aggregation, subsequent analyses and plots, LJP and TM helped WT in interpreting the results and participated to the paper writing.

REFERENCES

Ahmed, S.E., McInerny, G., O'Hara, K., Harper, R., Salido, L., Emmott, S. *et al.* (2015). Scientists and software - surveying the species distribution modelling community. *Divers. Distrib.*, 21, 258–267.

Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudan, A., Boucher, F., Saccone, P. *et al.* (2010). Intraspecific functional variability: extent, structure and sources of variation within an alpine valley. *J. Ecol.*, 98, 604–613.

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.

Amarasekare, P. (2003). Competitive coexistence in spatially structured environments: a synthesis. *Ecol. Lett.*, 6, 1109–1122.

Araújo, M.B. & Peterson, A.T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93, 1527–1539.

Araújo, M.B. & Rozenfeld, A. (2014). The geographic scaling of biotic interactions. *Ecography*, 37, 406–415.

de Bello, F., Price, J.N., Munkemuller, T., Liira, J., Zobel, M., Thuiller, W. *et al.* (2012). Functional species pool framework to test for biotic effects on community assembly. *Ecology*, 93, 2263–2273.

de Bello, F., Lavorel, S., Lavergne, S., Albert, C.H., Boulangeat, I., Mazel, F. *et al.* (2013). Hierarchical effects of environmental filters on the functional structure of plant communities: a case study in the French Alps. *Ecography*, 36, 393–402.

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

Cadotte, M., Albert, C.H. & Walker, S.C. (2013). The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecol. Lett.*, 16, 1234–1244.

de Candolle, A.I. (1855). *Géographie Botanique Raisonnée*. Masson, Paris.

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

Chase, J.M. & Leibold, M.A. (2003). *Ecological Niches*. Chicago University Press, Chicago.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.*, 31, 343–366.

Choler, P. (2005). Consistent shifts in Alpine plant traits along a mesotopographical gradient. *Arctic, Antartic, Alpine Res*, 37, 444–453.

Clark, J.S., Gelfand, A.E., Woodall, C.W. & Zhu, K. (2014). More than the sum of the parts: forest climate response from joint species distribution models. *Ecol. Appl.*, 24, 990–999.

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

Davis, A.J., Jenkinson, L.S., Lawton, J.H., Shorrocks, B. & Wood, S. (1998). Making mistakes when predicting shifts in species range in response to global warming. *Nature*, 391, 783–786.

Godsoe, W., Murray, R. & Plank, M.J. (2015). The effect of competition on species' distributions depends on coexistence, rather than scale alone. *Ecography*, DOI: 10.1111/ecog.01134.

Götzenberger, L., de Bello, F., Bråthen, K.A., Davison, J., Dubuis, A., Guisan, A. *et al.* (2012). Ecological assembly rules in plant communities —approaches, patterns and prospects. *Biol. Rev.*, 87, 111–127.

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

Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.*, 8, 993–1009.

Hill, M.O. (1973). Diversity and evenness – unifying notation and its consequences. *Ecology*, 54, 427–432.

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M. & Mayfield, M.M. (2012). Rethinking Community Assembly through the Lens of Coexistence Theory. *Annu. Rev. Ecol. Evol. Syst.*, 43, 227–248.

Körner, C. (1999). *Alpine Plant Life*. Springer-Verlag, Berlin.

Kraft, N.J.B. & Ackerly, D.D. (2010). Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecol. Monogr.*, 80, 401–422.

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

Lortie, C.J., Brooker, R.W., Callaway, R.M., Choler, P., Kikvidze, Z., Michalet, R. *et al.* (2004). Rethinking plant community theory. *Oikos*, 107, 433–438.

Mitchell, M.G.E., Cahill, J.F. & Hik, D.S. (2009). Plant interactions are unimportant in a subarctic-alpine plant community. *Ecology*, 90, 2360–2367.

Mouquet, N., Devictor, V., Meynard, C.N., Munoz, F., Bersier, L.-F., Chave, J. *et al.* (2012). Ecophylogenetics: advances and perspectives. *Biol. Rev.*, 87, 769–785.

Münkemüller, T., de Bello, F., Meynard, C.N., Gravel, D., Lavergne, S., Mouillot, D. *et al.* (2012). From diversity indices to community assembly processes: a test with simulated data. *Ecography*, 35, 468–480.

Münkemüller, T., Boucher, F., Thuiller, W. & Lavergne, S. (2015). Common conceptual and methodological pitfalls in the analysis of phylogenetic niche conservatism. *Funct. Ecol.*, 29, 627–639.

Parviainen, M., Marmion, M., Luoto, M., Thuiller, W. & Heikkinen, R.K. (2009). Using summed individual species models and state-of-the-art modelling techniques to identify threatened plant species hotspots. *Biol. Conserv.*, 142, 2501–2509.

Pearson, R.G. & Dawson, T.P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.*, 12, 361–372.

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M. *et al.* (2014). Understanding co-occurrence bymodelling species simultaneously with a Joint Species DistributionModel (JSDM). *Methods Ecol. Evol.*, 5, 397–406.

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C.F. *et al.* (2013). The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Glob. Ecol. Biogeogr.*, 22, 52–63.

Soberon, J.M. (2010). Niche and area of distribution modeling: a population ecology perspective. *Ecography*, 33, 159–167.

Soberon, J. & Nakamura, M. (2009). Niches and distributional areas: concepts, methods, and assumptions. *Proc. Natl Acad. Sci. USA*, 106, 19644–19650.

Thuiller, W., Münkemüller, T., Lavergne, S., Mouillot, D., Mouquet, N., Schiffers, K. *et al.* (2013). A road map for integrating eco-evolutionary processes into biodiversity models. *Ecol. Lett.*, 16, 94–105.

Thuiller, W., Guéguen, M., Georges, D., Bonet, R., Chalmandrier, L., Garraud, L. *et al.* (2014). Are different facets of plant diversity well protected against climate and land cover changes? A test study in the French Alps. *Ecography*, 37, 1254–1266.

Weiher, E. & Keddy, P. (eds.) (1999). *Ecological Assembly Rules: Perspectives, Advances, Retreats*. Cambridge University Press, Cambridge.

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