Author Manuscript

Title: Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models

Authors: D'Amen M., Dubuis A., Fernandes R.F., Pottier J., Pellisser L., Guisan A.,

Journal: Journal of Biogeography

Year: 2015

Issue: 42

Volume: 7

Pages: 1255-1266

DOI: https://doi.org/10.1111/jbi.12485

Original article

LRH: M. D'Amen *et al*.

RRH: Constrained species assemblage prediction from species distribution models

Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models

Manuela D'Amen[1†], Anne Dubuis[1†], Rui F. Fernandes[1], Julien Pottier[2], Loïc Pellisser[1,3] and Antoine Guisan[1, 4 *]

[1]*Department of Ecology and Evolution, University of Lausanne, Biophore building, 1015 Lausanne, Switzerland,* [2]*INRA, Grassland Ecosystem Research Unit (UREP), 5 Chemin de Beaulieu, 63100 Clermont-Ferrand, France,* [3]*Department of Biology, Ecology and Evolution, University of Fribourg, Chemin du Musée 10, 1700 Fribourg, Switzerland,* [4]*Institute of Earth Surface Dynamics, Geopolis, University of Lausanne, 1015 Lausanne, Switzerland*

[*]Correspondence: Antoine Guisan, Department of Ecology and Evolution, University of Lausanne, Biophore building, CH-1015 Lausanne, Switzerland.

E-mail: antoine.guisan@unil.ch

[†] Co-first authors.

ABSTRACT

Aim Modelling species distributions at the community level is required to make effective forecast of global change impacts on diversity and ecosystem functioning. Community predictions may be achieved using macroecological properties of communities (macroecological models, MEM), or by stacking of individual species distribution models (stacked species distribution models, S-SDMs). To obtain more realistic predictions of species assemblages, the SESAM (spatially explicit species assemblage modelling) framework suggests applying successive filters to the initial species source pool, by combining different modelling approaches and rules. Here we provide a first test of this framework in mountain grassland communities.

Location The western Swiss Alps.

Methods Two implementations of the SESAM framework were tested: a 'probability ranking' rule based on species richness predictions and rough probabilities from SDMs, and a 'trait range' rule that uses the predicted upper and lower bound of community-level distribution of three different functional traits (vegetative height, specific leaf area and seed mass) to constrain a pool of species from binary SDMs predictions.

Results We showed that all independent constraints contributed to reduce species richness overprediction. Only the 'probability ranking' rule allowed slight but significant improvements in the predictions of community composition.

Main conclusions We tested various implementations of the SESAM framework by integrating macroecological constraints into S-SDM predictions, and report one that is able to improve compositional predictions. We discuss possible improvements, such as further understanding the causality and precision of environmental predictors, using other assembly rules and testing other types of ecological or functional constraints.

Keywords

Community ecology, functional ecology, macroecological models, MEM, SESAM framework, species distribution models, SDM, stacked-SDM.

INTRODUCTION

Understanding the distribution and composition of species assemblages and being able to predict them in space and time are important for understanding the fate of biodiversity under global change. Different approaches have been proposed to predict the composition of species assemblages, which can work on mechanistic or empirical bases. Neutral views have also been proposed to explain relative abundance patterns in communities (Hubbell, 2001), which were contrasted to niche/trait views (Wennekes *et al.*, 2012). Neutral theory has been challenged for not representing forces that actually operate in nature to shape communities and their composition (e.g. Clark, 2009). Using a more deterministic approach, Shipley *et al.* (2006) proposed the use of predicted community weighted means of functional traits to infer the assemblage composition given species traits through a maximum entropy approach (Shipley *et al.*, 2006, 2011; Sonnier *et al.*, 2010a; see also Laughlin *et al.*, 2012). Mokany *et al*. (2011, 2012) proposed a dynamic framework to model species richness and composition dissimilarity based on species data. A distinct approach, not requiring traits, is to use the empirical relationships between species distribution data and environmental factors to predict community types or axes of compositional variation derived from ordination techniques (Ferrier & Guisan, 2006).

One widely used method is to predict the distributions of individual species with niche-based species distribution models (SDMs; also called ecological niche models, ENMs; see Guisan *et al.,* 2013), and then to stack them to predict species assemblages (stacked-SDM, S-SDM; Dubuis *et al.,* 2011). This method pertains to the category ‘predict first, assemble later’ in Ferrier & Guisan’s (2006) classification of community-level models, and has been tested in recent studies to draw conclusions about species richness (SR), assemblage composition or species turnover under current or future climatic conditions (Baselga & Araújo, 2009, 2010; Aranda & Lobo, 2011; Albouy *et al.*, 2012; Pottier *et al.*, 2013). Stacking individual species predictions can be applied to both rough probabilities (pS-SDM) and binary predictions from SDMs (bS-SDM) (e.g. Dubuis *et al.*, 2011; Calabrese *et al.*, 2014). pS-SDM currently allows the prediction of species richness only, while bS-SDM also provides information on species composition. It has been shown that bS-SDMs tend, on average, to overpredict species richness per unit area (Algar *et al*., 2009; Dubuis *et al*., 2011; Mateo *et al*., 2012), whereas pS-SDMs do not (Dubuis *et al*., 2011; Calabrese *et al*., 2014). Overprediction by bS-SDMs could be expected, as reconstructing communities from SDM predictions implies applying a series of species-specific abiotic filters, without consideration for macroecological constraints on the general properties of the system as a whole (Guisan & Rahbek, 2011). As an alternative explanation, it has also been suggested that overprediction could result from a mathematical artefact if the stacking process is applied to binary SDM predictions, i.e. after thresholding the rough probability of species' predictions (Calabrese *et al*., 2014).

Guisan & Rahbek (2011) proposed a framework – SESAM: spatially explicit species assemblage modelling – that aims to improve predictions of species assemblages. The main idea of the SESAM framework is to reconstruct species assemblages by applying successive filters of the assembly process through four main conceptual steps (Hortal *et al*., 2012). First, the species pool of each modelling unit in the study area must be defined. Second, species are filtered from the species pool according to their suitability to the environmental conditions in the modelling unit, e.g. by fitting SDMs. Third, limits previously set to one or several properties of each assemblage (e.g. richness or functional properties) are used to apply constraints on the assemblage in each unit, based on model predictions. Fourth, the species to be kept in the assemblage are chosen among the potential coexisting species (i.e. those predicted by the S-SDM), through biotic assembly rules. Macroecological constraints can be defined by macroecological models (MEMs), i.e. models of emergent properties or attributes of communities, such as species richness (SR) or other functional characteristics (e.g. functional richness) that are theoretically predictable directly from environmental variables (Francis & Currie, 2003; Moser *et al*., 2005; Sonnier *et al*., 2010b; Dubuis *et al*., 2011, 2013). MEMs, which belong to the 'assemble first, predict later' category of Ferrier & Guisan (2006)'s classification, have been shown to provide less biased predictions of SR than bS-SDMs (Dubuis *et al*., 2011). Yet, no attempt has been made to implement and test the SESAM framework.

In the SESAM framework, assemblage properties are predicted to define constraints to be applied to the assemblage in each unit. In this study, we test three macroecological constraints: (1) richness predicted by the sum of probability S-SDM (pS-SDM); (2) direct predictions of species richness (MEM) (Dubuis *et al.*, 2011); and (3) predicted values of three functional traits (Dubuis *et al.*, 2013). In particular, we test the use of functional traits as macroecological constraints, as they can be predicted spatially (Dubuis *et al.*, 2013) and may provide an understanding of the functional underpinnings of plant communities, allowing generalization beyond species identities (e.g. Hooper *et al.*, 2005; McGill *et al.,* 2006). Functional traits are supposed to enable the refinement of predictions of community composition along environmental gradients, by contrasting trait values for individual species to the ones aggregated at the community level (Shipley *et al.*, 2006; Douma *et al.*, 2012). We consider extremes in trait values to represent a filtering effect, i.e. the trait values that allow a species to be included in a community in a given environment (Keddy, 1992a,b). In order to build macroecological constraints, the same rationale applies to both richness and traits extreme values: limited amount of resources or environmental conditions (e.g. heterogeneity) defines ‘how many’ or ‘what type of’ species can strive in the considered unit. Here, both species richness and the functional characteristics of the community are assumed to be mainly controlled, among other possible factors, by available energy, as expressed by climatic predictors (Wright, 1983; Currie, 1991; Hawkins *et al.*, 2003; Shipley *et al*., 2006; see Guisan & Rahbek, 2011).

By integrating over these sources of information, we set macroecological constraints on the pool of species predicted to potentially co-occur in each site according to SDM predictions only. Doing this, we test – for the first time – a simplified version of the SESAM framework (i.e. without elaborated biotic assembly rules), using outputs from MEMs or pS-SDMs as constraints to limit the number of species predicted by bS-SDMs, this way attempting to improve predictions of community composition. More specifically, we ask the following questions:

1. Does combining different modelling techniques developed for biodiversity prediction improve the predictions of community attributes such as richness, species composition, traits distribution?

2. Does the use of assembly rules (driven either by habitat suitability or functional characteristics) to select the species that enter in the predicted community from SDMs improve the predictions of community richness and composition?

MATERIALS AND METHODS

Vegetation and traits data

The study area is located in the Alps of western Switzerland (http://rechalpvd.unil.ch) and covers *c.* 700 km^2, with elevations ranging from 375 to 3210 m. The species occurrence data used in our analysis originate from fieldwork conducted between 2002 and 2009 in the study area following a random-stratified sampling design and limited to open, non-woody vegetation (for more information see Dubuis *et al*., 2011). A first dataset of 613 vegetation plots of 4 m^2 each was inventoried and used for SDM and MEM calibration ('calibration dataset'). An additional set of 298 plots was identically surveyed to evaluate S-SDMs, and test the efficiency of MEM constraints ('evaluation dataset') (Fig. 1 – Data box). This evaluation dataset was shown to be spatially independent of the first one, and thus valid for model evaluation, by calculating the spatial correlation of SDMs' 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 species were recorded in the study area, with traits data available for a subset of the 189 most frequent species of this pool (Fig. 1; Pottier *et al.*, 2013; Dubuis *et al.*, 2013). We selected three traits (vegetative height, specific leaf area and seed mass) that are expected to represent the key axes of plant ecological strategies following the leaf–height–seed (LHS) scheme of Westoby (1998), already widely used for studying plant assembly rules. In particular, vegetative height (H) and specific leaf area (SLA) were measured on the field (for each species between 4 and 20 individuals were sampled over its entire bioclimatic range). We used the average trait value among all sampled individuals for each species for further analyses (Dubuis *et al*., 2013). Height was measured for each species in the field as the distance between top photosynthetic tissues and the ground, expressed in mm. This trait is related to competitive ability and is correlated with above-ground biomass (Cornelissen *et al*., 2003). SLA was calculated as the ratio of leaf surface to its dry mass and expressed in $mm^2\ mg^{-1}$. SLA is correlated with the relative growth rate and photosynthetic ability of plant species (Cornelissen *et al*., 2003). Seed mass (SM) data originate from literature and field measurements (Pellissier *et al*., 2010) and is expressed in milligrams. This trait is a good predictor of colonization ability of the species and seedling survivorship (Moles & Westoby, 2006). To account for trait range limitation, we calculated percentiles of trait distribution in sites where the 189 species for which trait data were available represented more than 80% of the total vegetation cover (Pakeman & Quested, 2007; see Pottier *et al.*, 2013, Dubuis *et al.*, 2013).

General analytical framework

We tested different implementations of the SESAM framework to predict species composition, by applying two different types of species assembly rules:

1. 'Probability ranking' rule: this rule is based on the assumption that species with the highest habitat suitability are competitively superior. According to this rule, community composition is obtained by selecting the species in decreasing order of their predicted probability of presence from SDMs up to the richness prediction (i.e. predictions from MEM or pS-SDM).

2. 'Trait range' rule: we applied a filter based on important functional characteristics of plant species that relate to competitive and reproductive abilities. We used percentile predictions from MEMs of three functional traits, individual or in combination, as criteria to discard species that do not fall into the predicted functional range of the sites. We implemented this approach with the three percentiles boundaries.

We fitted all the models (both SDMs and MEMs) by applying three modelling techniques in R (2.14.1) with the BIOMOD package (Thuiller *et al.*, 2009): generalized linear models (GLMs), generalized additive models (GAMs) and generalized boosted models (GBMs). The resulting projections were averaged to implement an ensemble forecasting approach.

We applied the SESAM framework following the four step design described by Guisan & Rahbek (2011) and adapted to our study case (Fig. 1).

Step 1 – Species pool

As the first component of the SESAM framework, we considered a unique species pool for all modelling units, defined as the most frequent plant species occurring in our study area (241 species). This pool was used to test the 'probability ranking' rule. A subset of this pool was used to test the 'trait range' rule (189 species).

Step 2 – Abiotic filtering

Single species models were fitted with environmental predictors calculated from temperature and precipitation data recorded by the Swiss network of meteorological stations and from a digital elevation model at 25 m resolution (see Dubuis *et al*., 2011). We used growing degree-days (above 0 °C), moisture index over the growing season (difference between precipitation and potential evapotranspiration), the sum of solar radiations over the year, slope (in degree) and topographic position (unit less, indicating the ridges and valleys). These five variables have been shown to be useful for predicting the topo-climatic distributions of plant species in mountainous environment (Dubuis *et al.*, 2011). The models were evaluated on the evaluation dataset with the area under the curve (AUC) of a receiver operating characteristic plot (ROC) and the true skill statistic (TSS; Allouche *et al.*, 2006). Ensemble predictions were obtained by computing the weighted average of the predictions by the three techniques. To do this, we used weights from the internal cross-validation with both AUC (Swets, 1988) and TSS (Allouche *et al*., 2006) evaluation metrics. The predictive ability of the final ensemble models was then tested with the same metrics using the external evaluation dataset. The raw predictions for the 241 species represent the ‘probability pool’ used in the ‘probability ranking’ rule test. In ‘trait range’ rule tests the projected species distributions for the 189 species were transformed into binary presences and absences using two threshold approaches: (1) the threshold corresponding to equal values of sensitivity and specificity (Liu *et al*., 2005), and (2) the threshold maximizing TSS. The resulting binary projections were stacked to predict assemblages in each of the evaluation plots (bS-SDM). This way, we obtained a pool of species potentially present filtered by topo-climatic factors.

Step 3 – Macroecological constraints

Three different methods were used to define macroecological constraints. First, we summed probabilities from SDMs (Dubuis *et al*., 2011) for the 241 species, obtaining a prediction of richness for each unit (pS-SDM). Second, observed species richness (SR) was calculated as the number of species (among the 241 used in this study) present in each sampling plot. Total SR was predicted with the same environmental predictors and modelling techniques used for SDMs fitted with a Poisson distribution. Also in this case, we applied the ensemble forecasting approach (as described above) to obtain a final richness prediction (‘species richness’ MEM; see Dubuis *et al*., 2011). Finally, we modelled traits values, considering three pairs of percentiles limits: 1st–99th, 5th–95th and 10th–90th. We modelled each trait percentile as a function of the environmental predictors and assuming a normal distribution (‘traits range’ MEM; Dubuis *et al*., 2013). The modelling procedure was the same used for species richness prediction. Prior to modelling, trait data were log-transformed. The predictive power of the SR and traits range models were measured by computing a Spearman rank correlation between the observed and predicted indices values for the evaluation dataset.

Step 4 – Ecological assembly rules

We applied our rules to couple results coming from previous steps. To test the ‘probability ranking’ rule, we determined the community composition by ranking the species in decreasing order of their predicted probability of presence from SDMs up to the richness prediction by pS-SDM or SR-MEM.

We further compared the application of this rule with a random selection of species in the number of the richness predictions, as a null test of composition prediction success. This was performed on the full evaluation dataset of 298 plots not used in model calibration.

In the ‘trait range’ rule, for each site, among the species predicted as present by the binary SDMs (‘traits pool’), we excluded from the final community prediction those species with traits valued outside the predicted functional range predicted by MEMs. In particular, for each percentile pair (1st–99th, 5th–95th, and 10th–90th), we considered the predicted trait values and we excluded all species having traits values outside these quantiles. All seven combinations of the three functional traits were considered (taken singularly, in pairs or all together) to constraint community composition. As a result, we tested a total of 21 macroecological constraints based on traits. The ‘trait range’ rule was applied to the 192 plots of the evaluation dataset for which we had trait data for more than 80% of the vegetation cover for the second test.

Finally, species richness and composition outputs resulting from the SESAM approaches were compared to the evaluation dataset. Assemblage predictions were evaluated with several metrics based on a confusion matrix where all species (species pool: *SP*) are classified into: *TP*: the species observed as well as predicted as present (true positive), *FN*: the species observed as present but predicted as absent (false negative; omission error), *FP*: the species observed as absent but predicted as present (false positive; commission error) and *TN*: the species both observed and predicted as absent (true negative) (see Appendix S1 in Supporting Information). We computed the species richness error (predicted SR – observed SR, expressed as a number of species in Fig. 2), the assemblage prediction success (a), and the Sørensen index, related to Bray–Curtis dissimilarity (b).

$$\text{(a) Prediction success} = \frac{TP + TN}{SP}$$

$$\text{(b) Sørensen index} = \frac{2TP}{2TP+FN+FP}$$

RESULTS

SDMs for most species had an AUC value higher than 0.7 and can therefore be considered as useful for predictions (see Appendix S2). The MEM for species' richness and pS-SDM gave similar results: both predictions showed fair correlations between observed and predicted values of richness in the evaluation dataset (ρ = 0.529 and 0.507, respectively, Spearman rank correlation test). Macroecological models for traits were all above 0.5 (ρ values, Spearman rank correlation test) except for the 1st and 5th percentiles of log(SM) (Appendix S2). The 'trait range' rule was applied by considering all couples of percentile, but as the results are consistent (see Appendix S3), in the following section we only show results coming from the 5th–95th percentiles. The S-SDM built with binary SDMs overpredicted species richness (SR) in all plots (Figs 2a & 3). All filtering types, both coming from the 'probability ranking' rule and the 'trait range' rule contributed on average to reduce SR overprediction, i.e. reduction of SR error (Figs 2a,d & 3), except when using the combination of SLA and SM trait limits as constraining rule.

Considering composition predictions, the prediction success was increased when applying either the 'probability ranking' rule or the 'trait range' rule (Fig. 2b,e), again with the exception of the combination of SLA and SM trait limits. Results from the Sørensen index (Fig. 2c) indicate that the 'probability ranking' rule increased the predictive capability by using both predicted SR from MEM and pS-SDM, as a limit, with the former slightly outperforming the latter. In both cases, the Sørensen index was significantly higher than the one of the simple bS-SDM (Wilcoxon signed rank test, *P*-value < 0.005). On average, this approach was less affected by errors of commission (false positive; Appendix S1) than other approaches and had the highest rate of correctly predicted absences (Fig. 4a). Using SR as a limit (from both MEM and pS-SDM) but choosing species randomly among those predicted yielded the worst assemblage composition predictions (Fig. 2c). We observed a decrease in the ability to correctly predict species identities when using the 'trait range' rule to constraints S-SDM predictions (Fig. 2f). Predicted functional traits did not provide a sufficient constraint to improve composition, and did not allow for a complete reduction of the SR over-prediction. Their use allowed species richness prediction to be improved, but at the cost of slightly decreasing assemblage composition prediction success (Sørensen index) (Fig. 4b). The applications of our rules did not produce a prediction of species assemblage compositions better than an average Sørensen's similarity of 0.5.

Results for community predictions using TSS and the 'trait range' rule were similar to those using AUC and are thus presented in Appendix S1.

DISCUSSION

This study represents the first formal test of the SESAM framework (Guisan & Rahbek, 2011). We have shown different ways to implement the SESAM framework, by integrating stacked predictions from species distribution models (S-SDMs) with richness predictions from macroecological models (MEMs) or from the sum of rough probabilities from S-SDM (pS-SDM). Our results show that the application of macroecological constraints on single species predictions from SDMs improve the overall quality of assemblage' composition estimation. As expected, all the macroecological constraints considered reduced the overprediction of species richness. But more importantly, the sequence of steps of the framework allowed a more accurate prediction of the realized species assemblage as measured with metrics equally weighting commission (false presence) and omission (false absence) errors. This positive result encourages further developments of the SESAM framework to improve the prediction of community attributes.

Among the implementations of the SESAM framework tested here, the application of the 'probability ranking' rule improved the predictions of species richness and composition. First, both ways of producing species richness predictions, i.e. stacking of probabilities from SDMs (pS-SDM), and directly predicting species' richness (MEM), gave more reliable results than the simple binary S-SDMs, a result shown previously (e.g. Dubuis *et al.*, 2011; Calabrese *et al.*, 2014). Second, this approach also produced better predictions of community composition, by selecting single species from the pool predicted by SDMs by decreasing order of predicted probability (until the predicted richness is reached). One possible explanation for this positive result is that the same species that are least likely to be present, i.e. the ones removed by the rule, are also the ones most likely to be overpredicted by bS-SDMs.

The ‘trait range’ rule (as applied here) proved less effective in constraining community predictions, and no specific functional trait or any percentile interval proved more efficient than another in reducing species richness overprediction. Although surprising because MEMs for traits were on average better than those for species richness (see Dubuis *et al*., 2011, 2013), we can hypothesize some explanations for this result: (i) we used trait averages for each species, whereas each of these traits is known *in situ* to exhibit intraspecific variation along environmental gradients (Albert *et al.*, 2010); (ii) the traits that we used have been shown not always to relate significantly to species’ habitat suitability (Thuiller *et al*., 2010); (iii) a larger dataset of traits, as used in trait-based modelling approaches (e.g. Shipley *et al*., 2011), could have been more efficient in setting specific functional limits for the community prediction than the three traits used here. Still, the use of the combination of three traits as a constraint allowed an efficient decrease of species richness’ overprediction, supporting the need to put restraints on species pools based on a simple stacking of species predictions. Roots traits, indicating below-ground competition, could be good additional candidates to complement the functional constraints. These and other possible trait types should be assessed in future studies testing the SESAM framework. A potential limitation to the use of particular functional traits is that they must relate to species’ ability to cope with the environment and be reliably predicted in space by MEM (e.g. Dubuis *et al.*, 2013), which may not always be possible. Finally, we used three different percentiles ranges to depict minimal and maximal trait values as functional constraints, but the results for community predictions were not significantly different, so that we can be confident that our outcomes were not dependent on the percentiles’ choice.

Overall, and even after strongly reducing the species richness overprediction bias, predicted assemblage composition was improved but still remained significantly distinct from the observed ones, a result consistent with those by Aranda & Lobo (2011) and Pottier *et al.* (2013). Even if the individual SDMs have reasonably good independent evaluations, each of them nevertheless contains substantial errors that can be unevenly distributed among species and along environmental gradients (Pottier *et al*., 2013). By stacking SDMs, small errors in many individual species models can accumulate into quite large errors in the community predictions, degrading assemblage predictions accordingly (Pottier *et al.*, 2013). In this regard, the values of the Sørensen index of community similarity obtained in this first formal test of the SESAM framework – above 0.5 – can be considered a reasonable first achievement. A correction for the probability values based on the true species richness has been recently proposed by Calabrese *et al.* (2014). Their maximum likelihood approach, however, still does not allow the determination of which species in the list of probabilities will enter the final community. The error propagation could be even more severe if the single species predictions were binarized before reconstructing the community composition, because the choice of a threshold can matter (Liu *et al*., 2005). Moreover, a statistical bias was recently proposed as the main cause of the general overprediction in richness estimation showed by summing binary SDMs (e.g. Calabrese *et al*., 2014). As just discussed, we acknowledge the fact that stacking binary SDMs could add biases to the community prediction, but on the other hand it has the strength to allow an easy identification of the component species. Predicting assemblage composition over probabilities is still largely wished and applied, especially in conservation studies (e.g. Faleiro *et al.*, 2013; Leach *et al*., 2013). In order to partially control for the additional uncertainty introduced by thresholding, we ran all our analyses using both AUC and TSS threshold maximization metrics. The results of both analyses were consistent and therefore we can be confident that our outcomes are not too sensitive to this threshold choice.

The possibility of predicting species composition in a probabilistic way, without thresholding, holds the promise of reducing methodological biases, but it is still an unresolved issue that will need further developments. In the test of the ‘probability ranking’ rule, we proposed one solution, which avoids the binary transformation of SDM predictions, while still maintaining information about species composition. We did this by selecting a number of species equal to the prediction of species richness on the basis of decreasing probability of presence calculated by SDMs. Predictions of species composition is a great challenge for community ecologists and not many applicable solutions have been proposed (e.g. Webb *et al.*, 2010; Shipley *et al.*, 2011; Laughlin *et al.*, 2012). Our results thus provide new insights to achieve this goal by using SDMs, while avoiding the statistical bias potentially occurring when stacking binary SDM predictions (Calabrese *et al*., 2014). Yet, several issues still need to be resolved; in particular, new approaches are needed to decrease rates of omission error in SDMs and in the resulting community predictions. One route to improve compositional predictions could come from producing single species models that are more efficient at predicting presences correctly (i.e. limiting omission errors by optimizing sensitivity). A source of omission errors in our case may come from limitations related to the environmental predictors and resolution used to build the SDMs (Pradervand *et al.*, 2014). Available predictors can themselves include some level of errors (e.g. from measurement, interpolation, calculation) and other important predictors (see below) may be missing in the underlying SDMs (Austin & Van Neil, 2010). As a result, species' realized niches are likely to be incompletely described and some suitable or unsuitable situations for a species cannot be captured in the model. Two recent papers have shown similar problems of assemblage predictions in the case of butterflies and plants, respectively (Pellissier *et al*., 2012; Pottier *et al*., 2013). In both cases, the sensitivity (true-positive rate) of assemblage predictions was lower at higher elevations, which was probably due to the more fragmented, mosaic-like environmental conditions there and to missing substrate predictors (e.g. rock type, soil depth). Regarding our study area, snow cover and geomorphology (Randin *et al*., 2009), soil moisture and soil temperature (Le Roux *et al*., 2013), as well as edaphic conditions (Dubuis *et al*., 2012) and finer micro-climatic measurements (Pradervand *et al*., 2014), are potential missing predictors that could contribute to improve SDMs and hence the resulting community composition predictions. Yet, these missing predictors are currently not available or only available for some plots, and none of them exist in a spatially explicit way to support the final predictions to be generalized to the whole study area.

Conclusions and future perspectives

In the last decade, the range of possible approaches to model species communities has been expanding. Remarkably, most of the very recent solutions agree on the idea of combining complementary approaches into a single framework, as we did here with SESAM (e.g. Webb *et al.*, 2010; Mokany *et al.*, 2012; Fernandez *et al.*, 2013). A framework approach has a number of highly desirable characteristics, in particular the flexibility to integrate different drivers and processes to represent the complexity of factors that influence community assembly and the possibility to couple strengths of different pre-existing techniques in a unique workflow. Community ecology research is in continuous development and any new technical improvement coming from theoretical advances could be promptly accommodated in a framework approach. For instance, in this study we tested the integration of two types of macroecological models, but other recent implementation could also be used, such as the use of sum of predicted species probabilities (Dubuis *et al.*, 2011; Calabrese *et al.*, 2014). Another innovative way to model species categories would be the species archetypes model (SAM; Dustan *et al.*, 2011; Hui *et al.*, 2013), which predicts communities using a finite mixture of regression model, on the basis of common responses to environmental gradients. Also, macroecological models not based on correlative statistics could be included to explicitly incorporate the mechanisms responsible for the observed distributions (e.g. Gotelli *et al.*, 2009).

Among the great challenges in predictive community ecology is the inclusion of biotic rules. This has been repeatedly attempted in simple SDMs (e.g. by adding other species or simple biotic variables as predictors of the modelled species) with the result of improving significantly the predictions (reviewed in Kissling *et al.*, 2012 and Wisz *et al.*, 2013). In contrast, community-level models most often incorporate the effect of biotic interactions indirectly by considering synthetic community attributes (as we did in this study), while only in a few cases were biotic interactions accounted for in an explicit fashion (e.g. Laughlin *et al.*, 2012; Fernandes *et al.*, 2013; Pellissier *et al.*, 2013). This gap could partly be explained by the shortage of data available to characterize interactions among species in diverse communities (Araújo *et al.*, 2011). A potential way to overcome the lack of biotic interaction information could be the analysis of the spatial patterns of geographical overlap in the distributions of species. These can inform about potential interactions between species, but approaches to control for species habitat requirements should be applied in co-occurrence analyses to correctly infer biotic interactions from observed patterns (e.g. Gotelli *et al*., 1997; Peres-Neto *et al*., 2001; Ovaskainen *et al*., 2010). Because considering each pairwise interaction as a separate process is difficult, some alternative solutions to reduce this complexity have been also suggested, such as the analysis of separate smaller ‘community modules’ (as applied in food web analyses; Gilman *et al*., 2010), or the use of proxies of interactions (‘interaction currencies’) based on measures of non-consumable environmental conditions (described in Kissling *et al*., 2012).

The implementation of the full SESAM framework, i.e. implementing the ‘step 4’ through the definition of biotic assembly rules coming from empirical patterns of co-occurrence or experiments, could represent a promising route to further define the group of species that can coexist at each site, and help decreasing the rate of omission error. This fourth component of the framework has not been tested in an ecologically explicit way in this study, although using ranked probabilities of occurrence per site can be considered a form of implicit biotic rules. Identifying and quantifying other biotic assembly rules that can be applied generally along wide environmental gradients appears still to be difficult given our current state of knowledge and the heterogeneity of approaches used (Götzenberger *et al*., 2012; Kissling *et al*., 2012; Wisz *et al*., 2013), but it constitutes a necessary target if we want to improve our capacity to predict assemblages in space and time.

Further important drivers of community assembly are stochastic processes, associated with environmental disturbance and demographic dynamics within local and regional species pools (Dornelas *et al*., 2006). The potential presence of stochastic effects would deviate the community assemblage process from being fully deterministic, i.e. from yielding a specific community configuration for a given environmental combination and species pool, but instead be probabilistic so that the projections could for instance consist of a density function of various possible end compositions (Ozinga *et al*., 2005; Shipley, 2010; Pellissier *et al*., 2012; Pottier *et al*., 2013). Therefore, assemblage composition will always entail some level of prediction errors. In this regard, what would prove useful in future studies would be to understand and discern better the different sources of errors in the single techniques integrated in the SESAM framework. In particular, it would be useful to assess how errors propagate from individual SDMs to S-SDMs, and what value of the Sørensen index (or other evaluation metric of community similarity) would qualify as a fair value of assemblage prediction. This will help estimate the level of similarity and reliability with which one can ultimately expect species assemblages to be successfully predicted, and how far the latter may contribute to a better understanding and prediction of community assembly in space and time (Hortal *et al*., 2012).

ACKNOWLEDGEMENTS

We are grateful to the numerous people who contributed to the data collection and to Pascal Vittoz and Glenn Litsios for insightful discussions and comments on the manuscript. We would like also to thank B. Cade and M. Geraci for their useful insights on the statistics of trait analyses. This study was supported by the Swiss National Science Foundation (grant no. 31003A-125145 to A. Guisan), by the FP6 Ecochange project of the European Commission (grant GOCE-CT-2007–036866). M. D’Amen was supported by a Marie Curie Intra-European Fellowship within the 7th European Community Framework Programme (FP7-PEOPLE-2012-IEF, SESAM-ZOOL 327987).

REFERENCES

Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudant, A., Boucher, F., Saccone, P. & Lavorel, S. (2010) Intraspecific functional variability: extent, structure and sources of variation. *Journal of Ecology*, **98**, 604–613.

Albouy, C., Guilhaumon, F., Araújo, M.B., Mouillot, D. & Leprieur, F. (2012) Combining projected changes in species richness and composition reveals climate change impacts on coastal Mediterranean fish assemblages. *Global Change Biology*, **18**, 2995–3003.

Algar, A.C., Kharouba, H.M., Young, E.R. & Kerr, J.T. (2009) Predicting the future of species diversity: macroecological theory, climate change, and direct tests of alternative forecasting methods. *Ecography*, **32**, 22–33.

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.

Aranda, S.C. & Lobo, J.M. (2011) How well does presence-only-based species distribution modelling predict assemblage diversity? A case study of the Tenerife flora. *Ecography*, **34**, 31–38.

Araújo, M.B., Rozenfeld, A., Rahbek, C. & Marquet, P.A. (2011) Using species co-occurrence networks to assess the impacts of climate change. *Ecography*, **34**, 897-908.

Austin, M.P. & Van Niel, K.P. (2010) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, **38**, 1-8.

Baselga, A. & Araújo, M.B. (2009) Individualistic vs community modelling of species distributions under climate change. *Ecography*, **32**, 55–65.

Baselga, A. & Araújo, M.B. (2010) Do community-level models describe community variation effectively? *Journal of Biogeography*, **37**, 1842–1850.

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.

Clark, J.S. (2009) Beyond neutral science. *Trends in Ecology and Evolution*, **24**, 8-15.

Cornelissen, J.H.C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D.E., Reich, P.B., ter Steege, H., Morgan, H.D., van der Heijden, M.G.A., Pausas, J.G. & Poorter, H. (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, **51**, 335-380.

Currie, D.J. (1991) Energy and large-scale patterns of animal- and plant-species richness. *The American Naturalist*, **137**, 27–49.

Dornelas, M., Connolly, S.R. & Hughes, T.P. (2006) Coral reef diversity refutes the neutral theory of biodiversity. *Nature*, **440**, 80-82.

Douma, J.C., Witte, J.-P.M., Aerts, R., Bartholomeus, R.P., Ordoñez, J.C., Venterink, H.O., Wassen, M.J. & van Bodegom, P.M. (2012) Towards a functional basis for predicting vegetation patterns; incorporating plant traits in habitat distribution models. *Ecography*, **35**, 294–305.

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., Giovanettina, S., Pellissier, L., Pottier, J., Vittoz, P. & Guisan, A. (2012) Improving the prediction of plant species distribution and community composition by adding edaphic to topo-climatic variables. *Journal of Vegetation Science*, **24**, 593-606.

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

Dunstan, P.K., Foster, S. D. & Darnell, R. (2011) Model based grouping of species across environmental gradients. *Ecological Modelling*, **222**, 955-963.

Faleiro, F.V., Machado, R.B. & Loyola, R.D. (2013) Defining spatial conservation priorities in the face of land-use and climate change. *Biological Conservation*, **158**, 248-257.

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 Change Biology*, **19**, 2596-2607.

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

Francis, A.P. & Currie, D.J. (2003) A globally consistent richness–climate relationship for angiosperms. *The American Naturalist*, **161**, 523–536.

Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010) A framework for community interactions under climate change. *Trends in Ecology and Evolution*, **25**, 325-331.

Gotelli, N.J., Buckley, N.J. & Wiens, J.A. (1997) Co-occurrence of Australian land birds: Diamond's assembly rules revisited. *Oikos*, **80**, 311-324.

Gotelli, N.J., Anderson, M.J., Arita, H.T., Chao, A., Colwell, R.K., Connolly, S.R., Currie, D.J., Dunn, R.R., Graves, G.R., Green, J.L., Grytnes, J., Jiang, Y., Jetz, W., Lyons, S.K., McCain, C.M., Magurran, A.E., Rahbek, C., Rangel, T.F.L.V.B., Soberón, J., Webb, C.O. & Willig, M.R. (2009) Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, **12**, 873-886.

Götzenberger, L., De Bello, F., Anne Bråthen, K., Davison, J., Dubuis, A., Guisan, A., Lepš, J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel, K. & Zobel, M. (2012) Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biological Reviews*, **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. *Journal of Biogeography*, **38**, 1433–1444.

Guisan, A., Tingley, R., Baumgartner, *et al.,* (2013) Predicting species distributions for conservation decisions. *Ecology Letters,* **16**, 1424-1435.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E. & Turner, J.R.G. (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, **84**, 3105–3117.

Hooper, D.U., Chapin, F.S., III, Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J. & Wardle, D.A. (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, **75**, 3–35.

Hortal, J., De Marco, P., Santos, A.M.C. & Diniz-Filho, J.A.F. (2012) Integrating biogeographical processes and local community assembly. *Journal of Biogeography*, **39**, 627–628.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography.* Princeton University Press, Princeton, NJ.

Hui, F.C.K., Warton, D.I., Foster, S.D. & Dunstan, P.K. (2013) To mix or not to mix: comparing the predictive performance of mixture models vs. separate species distribution models. *Ecology*, **94**, 1913–1919.

Keddy, P.A. (1992a) A pragmatic approach to functional ecology. *Functional Ecology,* **6**, 621–626.

Keddy, P.A. (1992b) Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science*, **3**, 157–164.

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G.J., Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning, J.-C., Zimmermann, N.E. & O'Hara, R.B. (2012) Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, **39**, 2163–2178.

Laughlin, D.C., Joshi, C., van Bodegom, P.M., Bastow, Z.A. & Fulé, P.Z. (2012) A predictive model of community assembly that incorporates intraspecific trait variation. *Ecology Letters*, **15**, 1291-1299.

Leach, K., Zalat, S. & Gilbert, F. (2013) Egypt's Protected Area network under future climate change. *Biological Conservation*, **159**, 490-500.

Le Roux, P. C., Lenoir, J., Pellissier, L., Wisz, M. S. & Luoto, M. (2013) Horizontal, but not vertical, biotic interactions affect fine-scale plant distribution patterns in a low-energy system. *Ecology*, **94**, 671-682.

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., 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 and Evolution*, **21**, 178–185.

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

Mokany, K., Harwood, T.D., Williams, K.J. & Ferrier, S. (2012) Dynamic macroecology and the future for biodiversity. *Global Change Biology*, **18**, 3149–3159.

Moles, A.T. & Westoby, M. (2006) Seed size and plant strategy across the whole life cycle. *Oikos*, **113**, 91-105.

Moser, D., Dullinger, S., Englisch, T., Niklfeld, H., Plutzar, C., Sauberer, N., Zechmeister, H.G. & Grabherr, G. (2005) Environmental determinants of vascular plant species richness in the Austrian Alps. *Journal of Biogeography*, **32**, 1117–1127.

Ovaskainen, O., Hottola, J. & Siitonen, J. (2010) Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology*, **91**, 2514–2521.

Ozinga, W.A., Schaminée, J.H.J., Bekker, R.M., Bonn, S., Poschlod, P. & Tackelberg, O. (2005) Predictability of plant species composition from environmental conditions is constrained by dispersal limitation. *Oikos*, **108**, 555–561.

Pakeman, R.J. & Quested, H.M. (2007) Sampling plant functional traits: what proportion of the species need to be measured? *Applied Vegetation Science*, **10**, 91–96.

Pellissier, L., Fournier, B., Guisan, A. & Vittoz, P. (2010) Plant traits co-vary with altitude in grasslands and forests in the European Alps. *Plant Ecology*, **211**, 351–365.

Pellissier, L., Pradervand, J.-N., Pottier, J., Dubuis, A., Maiorano, L. & Guisan, A. (2012) Climate-based empirical models show biased predictions of butterfly communities along environmental gradients. *Ecography*, **35**, 684–692.

Pellissier, L., Rohr, R. P., Ndiribe, C., Pradervand, J.-N., Salamin, N., Guisan, A. & Wisz, M. (2013) Combining food web and species distribution models for improved community projections. *Ecology and Evolution*, **3**, 4572-4583.

Peres-Neto, P.R., Olden, J.D. & Jackson, D.A. (2001) Environmentally constrained null models: site suitability as occupancy criterion. *Oikos*, **93**, 110–120.

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

Pradervand, J.-N., Dubuis, A., Pellissier, L., Guisan, A. & Randin, C.F. (2014) Very high-resolution environmental predictors in species distribution models: moving beyond topography? *Progress in Physical Geography,* **38***,* 79-96.

Randin, C.F., Vuissoz, G., Liston, G.E., Vittoz, P. & Guisan, A. (2009) Introduction of snow and geomorphic disturbance variables into predictive models of alpine plant distribution in the western Swiss Alps. *Arctic, Antarctic, and Alpine Research*, **41**, 347–361.

Shipley, B. (2010) Community assembly, natural selection and maximum entropy models. *Oikos*, **119**, 604-609.

Shipley, B., Vile, D. & Garnier, E. (2006) From plant traits to plant communities: a statistical mechanistic approach to biodiversity. *Science*, **314**, 812–814.

Shipley, B., Laughlin, D.C., Sonnier, G. & Ottinowski, R. (2011) A strong test of a maximum entropy model of trait-based community assembly. *Ecology*, **92**, 507–517.

Sonnier, G., Shipley, B. & Navas, M. (2010a) Plant traits , species pools and the prediction of relative abundance in plant communities : a maximum entropy approach. *Journal of Vegetation Science*, **21**, 318–331.

Sonnier, G., Shipley, B. & Navas, M.L. (2010b) Quantifying relationships between traits and explicitly measured gradients of stress and disturbance in early successional plant communities. *Journal of Vegetation Science*, **21**, 318–331.

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., Albert, C.H., Dubuis, A., Randin, C. & Guisan, A. (2010) Variation in habitat suitability does not always relate to variation in species' plant functional traits. *Biology Letters*, **6**, 120–123.

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.

Wennekes, P., Rosindell, J. & Etienne, R. (2012) The neutral–niche debate: a philosophical perspective. *Acta Biotheoretica*, **60**, 257-271.

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

Wisz, M.S., Pottier, J., Kissling, W.D. *et al*. (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews of the Cambridge Philosophical Society*, **88**, 15–30.

Wright, D.H. (1983) Species-energy theory: an extension of species-area theory. *Oikos*, **41**, 496–506.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Assemblage evaluation metrics and supplementary results.

Appendix S2 Evaluation results for SDMs and MEMs.

Appendix S3 Comparison of the assemblage predictions coming from the application of trait range rule with three pairs of percentiles.

BIOSKETCH

Manuela D'Amen is a postdoc and **Anne Dubuis** is a PhD student in the spatial ecology group at the University of Lausanne (http://www.unil.ch/ecospat). This group, led by **Antoine Guisan**, specializes in spatial modelling of species, diversity and community distributions, using empirical data, statistical models and more dynamic approaches. A strong focus is given on the use of models and their predictions to support conservation management.

Author contributions: A.G. conceived the research idea and designed the data sampling; A.D., J.P., L.P., A.G and many others collected the data; A.D. and M.D. analysed the data with the help of R.F.F., J.P. and A.G. A.D., M.D., and R.F.F. wrote the manuscript with the help of J.P. and L.P.

Editor: Miles Silman

Figure legends

Figure 1 Workflow of the analytical steps followed in the study. *Data box:* We used a calibration and an evaluation datasets derived from field samplings carried out on 613 and 298 (192 with trait data) plots, respectively. These datasets were used to test the 'probability ranking' rule (left side of the figure with dashed arrows) and the 'trait range' rule (left side of the figure with dotted arrows). *Step 1 – species pool:* a total of 241 species collected in the study area were considered the 'species pool' to test the 'probability ranking' rule, (a) and (b). A subset of this species pool (189 species with trait data) was used to test the 'trait range' rule (c) and (d). All models were fitted by an ensemble forecasting approach based on the average of three techniques: generalized linear models (GLM), generalized additive models (GAM), and generalized boosted models (GBM). (e). *Step 2 – abiotic filtering*: distribution of individual species (a) and (c) were modelled and then stacked to create binary stacked species distribution model (S-SDM) predictions to represent a 'probability pool' for the 'probability ranking' rule test (f) and a 'traits pool' for the 'trait range' rule test (g). *Step 3 – macroecological constraints:* three different methods were used to define macroecological constraints, resulting in models with the stacked probabilities from SDMs (h; pS-SDM) and two different macroecological models (MEMs). These were created by modelling directly species richness values (i; SR_MEM) and three pairs of traits percentiles (j; Traits_MEM). *Step 4 – ecological assembly rules*: in the test of the 'probability ranking' rule (k) we limited species richness to fit the MEM or pS-SDM predictions and the species composition was determined (1) as a random selection from the pool or (2) selecting the species in decreasing order of predicted probability. In the test for the 'trait range' rule (l) we used the predicted values of MEM of functional traits (each trait separately and combinations of traits) to discard species functionally outside the assemblage. *Assemblage prediction box:* all the outputs resulting from the different approaches were compared and evaluate using the evaluation dataset (solid arrows).

Figure 2 Boxplots comparing unconstrained stacked species distribution model (S-SDM) predictions to results from the 'probability ranking' rule and random tests when applied constraining richness by the sum of probabilities from SDMs (PRR.pSSDM and rand.pSSDM, respectively) or by macroecological models (PRR.MEM and rand.MEM, respectively) (a, b, c), and to results from the 'trait range' rule test for single traits and all their combinations (d, e, f). The metrics utilized in the comparison are: species richness error, i.e. predicted SR – observed SR (first column); prediction success, i.e. sum of correctly predicted presences and absences divided by the total species number (second column); and Sørensen index, i.e. a statistic used to compare the similarity of two samples (third column). Abbreviations: SR rand, a random choice of species from the probability pool to reach the number predicted by richness model; SR prob, selection of the most probable species to reach the number predicted by richness model; H, height; SLA, specific leaf area of the community; SM, seed mass.

Figure 3 Predictions of species richness on the whole study area produced by (a) the unconstrained stacked species distribution model (S-SDM), and by the application of the SESAM framework implemented with (b) the 'probability ranking' rule implemented with the sum of probabilities from SDMs (pS-SDM), (c) the 'probability ranking' rule implemented with the richness estimation by the macroecological model (MEM) and (d) the 'trait range' rule (using the combination of the three traits as constraints).

Figure 4 Histograms showing the proportion (mean among all plots) of true and false positive, as well as true and false negative for all the implementations of the SESAM framework, compared with the unconstrained sum of binary species distribution model (bS-SDM). In the upper plot results from the 'probability ranking' rule test implemented with macroecological models and sum of probabilities from SDMs (PRR MEM and PRR pS-SDM, respectively) and random selections (rand MEM and rand pS-SDM, respectively). In the lower plot results from the 'trait range' rule test for single traits and all their combinations (H, height; SLA, specific leaf area of the community; SM, seed mass).

Journal of Biogeography

SUPPORTING INFORMATION

Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models

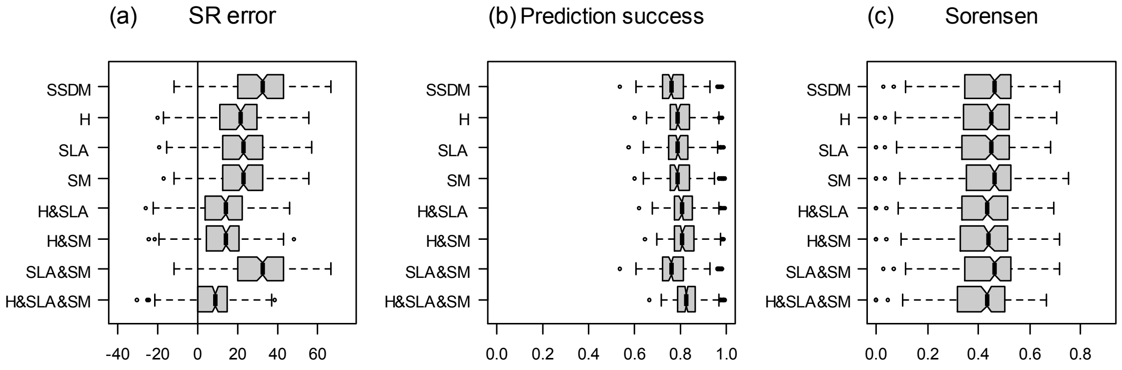
Manuela D'Amen, Anne Dubuis, Rui F. Fernandes, Freitag Julien Pottier, Loïc Pellisser and
Antoine Guisan

Appendix S1 Assemblage evaluation metrics and supplementary results.

Table S1 Confusion matrix used to compute the assemblage evaluation metrics.

		observed	
		0	**1**
predicted	**0**	*TN*	*FN*
	1	*FP*	*TP*

Figure S1 Results from true skill statistic (TSS) thresholding criterion: the boxplots compare results from the 'trait range' rule test for single traits and all their combinations when using TSS to binarize the SDM predictions. The metrics utilized in the comparison are: (a) species richness error, i.e. predicted SR – observed SR (first column), (b) prediction success, i.e. sum of correctly predicted presences and absences divided by the total species number (second column) and (c) Sørensen index, i.e. a statistic used to compare the similarity of two samples (third column). Abbreviations: H, height; SLA, specific leaf area of the community; SM, seed mass.



Appendix S2 Evaluation results for SDMs and MEMs.

Table S2 Summary of the SDMs evaluation results.

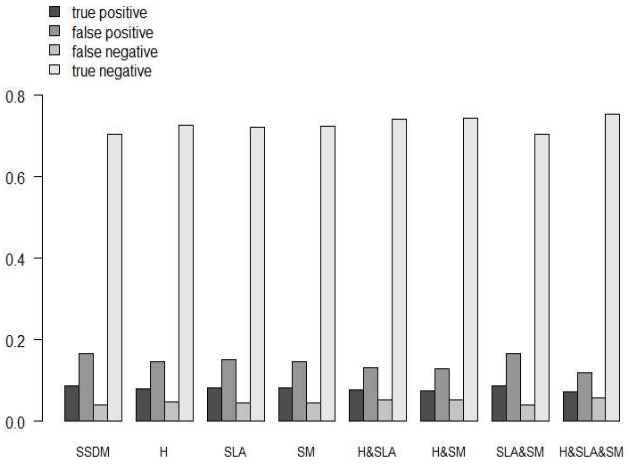
	GAM	**GBM**	**GLM**
AUC mean	0.803	0.783	0.799
AUC stdev	0.078	0.081	0.077
TSS mean	0.541	0.507	0.537
TSS stdev	0.142	0.137	0.142

Table S3 Values of Spearman correlation test between observed and predicted values of trait percentiles.

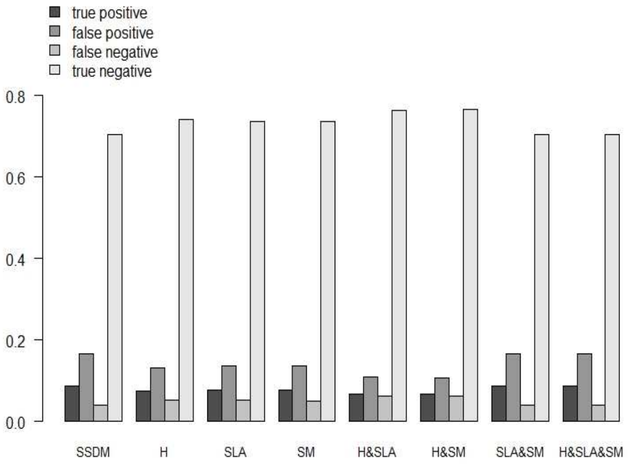
Percentile	**Trait**	**ρ**
1st	H	0.711
1st	SLA	0.759
1st	SM	0.152
99th	H	0.859
99th	SLA	0.584
99th	SM	0.514
5th	H	0.825
5th	SLA	0.803
5th	SM	0.350
95th	H	0.887
95th	SLA	0.652
95th	SM	0.528
10th	H	0.848
10th	SLA	0.814
10th	SM	0.550
90th	H	0.867
90th	SLA	0.677
90th	SM	0.645

Appendix S3 Comparison of the assemblage predictions coming from the application of the trait range rule with three pairs of percentiles. Abbreviations: SSDM, sum of binary SDMs; H, height; SLA, specific leaf area of the community; SM, seed mass.

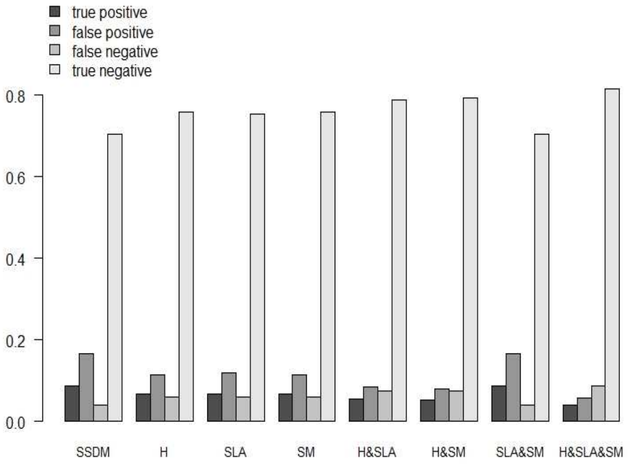
1st–99th

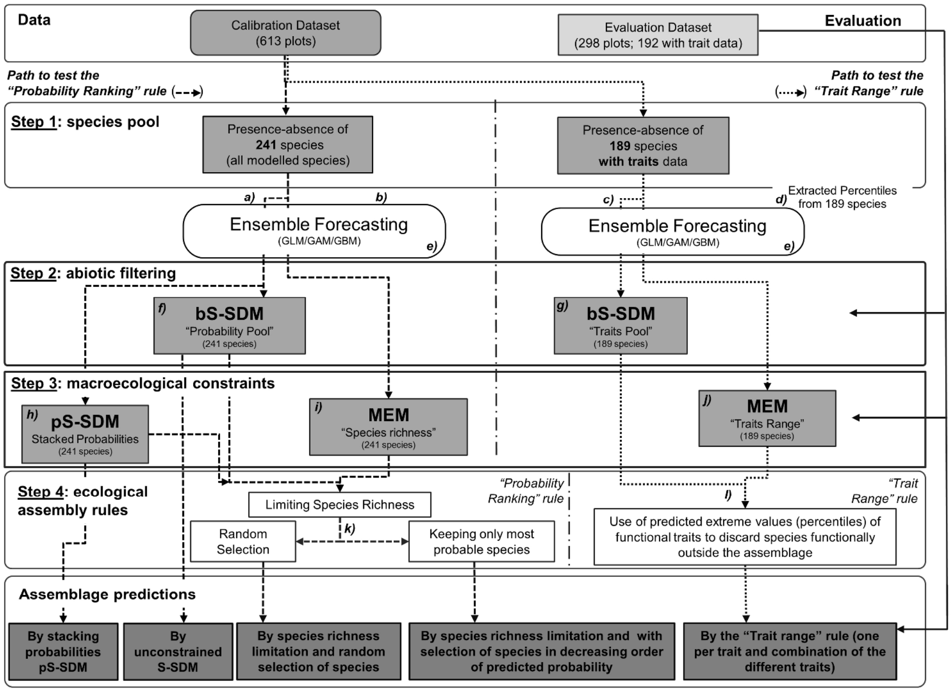


5th–95th

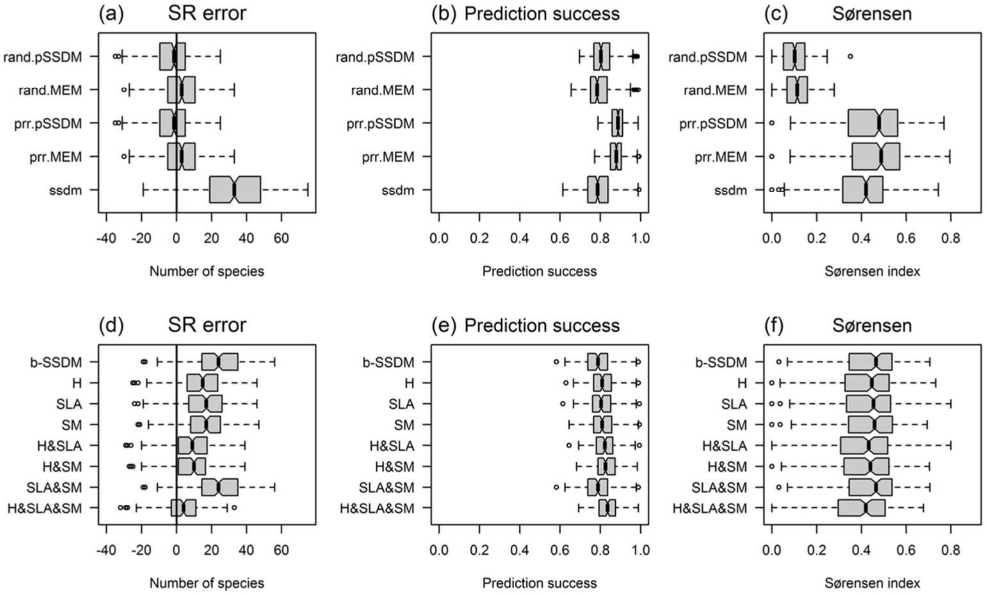


10th–90th

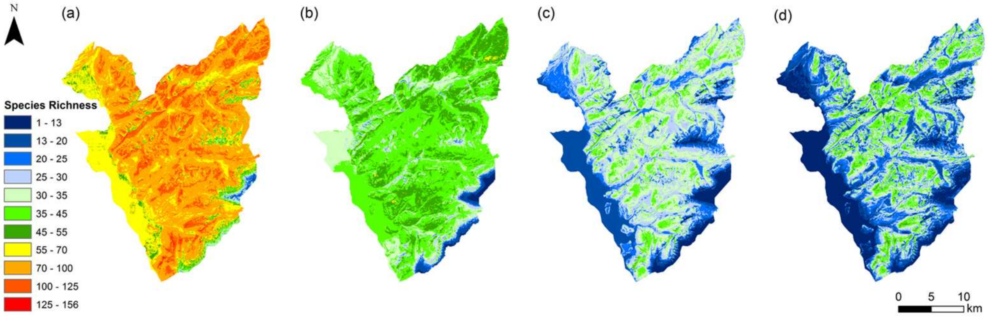




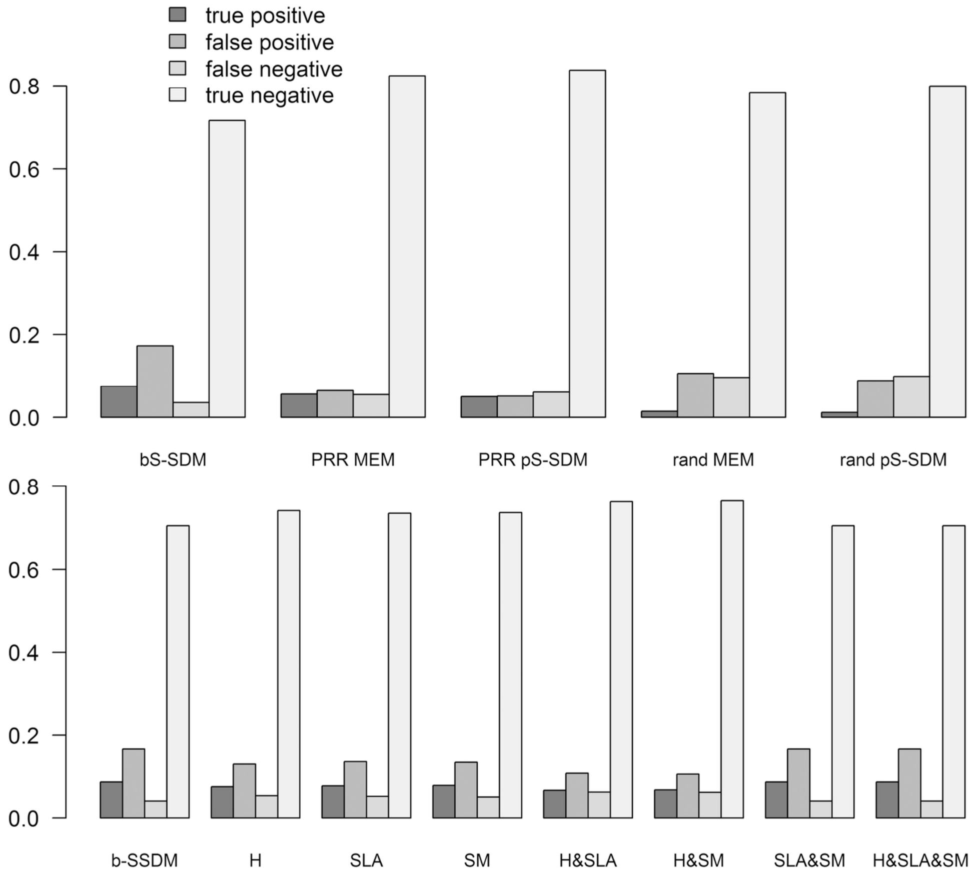
Workflow of the analytical steps followed in the study. Data box: We used a calibration and an evaluation datasets derived from field samplings carried out on 613 and 298 (192 with trait data) plots, respectively. These datasets were used to test the 'probability ranking' rule (left side of the figure with dashed arrows) and the 'trait range' rule (left side of the figure with dotted arrows). Step 1 – species pool: a total of 241 species collected in the study area were considered the 'species pool' to test the 'probability ranking' rule, (a) and (b). A subset of this species pool (189 species with trait data) was used to test the 'trait range' rule (c) and (d). All models were fitted by an ensemble forecasting approach based on the average of three techniques: generalized linear models (GLM), generalized additive models (GAM), and generalized boosted models (GBM). (e). Step 2 – abiotic filtering: distribution of individual species (a) and (c) were modelled and then stacked to create binary stacked species distribution model (S-SDM) predictions to represent a 'probability pool' for the 'probability ranking' rule test (f) and a 'traits pool' for the 'trait range' rule test (g). Step 3 – macroecological constraints: three different methods were used to define macroecological constraints, resulting in models with the stacked probabilities from SDMs (h; pS-SDM) and two different macroecological models (MEMs). These were created by modelling directly species richness values (i; SR_MEM) and three pairs of traits percentiles (j; Traits_MEM). Step 4 – ecological assembly rules: in the test of the 'probability ranking' rule (k) we limited species richness to fit the MEM or pS-SDM predictions and the species composition was determined (1) as a random selection from the pool or (2) selecting the species in decreasing order of predicted probability. In the test for the 'trait range' rule (l) we used the predicted values of MEM of functional traits (each trait separately and combinations of traits) to discard species functionally outside the assemblage. Assemblage prediction box: all the outputs resulting from the different approaches were compared and evaluate using the evaluation dataset (solid arrows).
274x190mm (284 x 284 DPI)



Boxplots comparing unconstrained stacked species distribution model (S-SDM) predictions to results from the 'probability ranking' rule and random tests when applied constraining richness by the sum of probabilities from SDMs (PRR.pSSDM and rand.pSSDM, respectively) or by macroecological models (PRR.MEM and rand.MEM, respectively) (a, b, c), and to results from the 'trait range' rule test for single traits and all their combinations (d, e, f). The metrics utilized in the comparison are: species richness error, i.e. predicted SR – observed SR (first column); prediction success, i.e. sum of correctly predicted presences and absences divided by the total species number (second column); and Sørensen index, i.e. a statistic used to compare the similarity of two samples (third column). Abbreviations: SR rand, a random choice of species from the probability pool to reach the number predicted by richness model; SR prob, selection of the most probable species to reach the number predicted by richness model; H, height; SLA, specific leaf area of the community; SM, seed mass.
106x66mm (300 x 300 DPI)



Predictions of species richness on the whole study area produced by (a) the unconstrained stacked species distribution model (S-SDM), and by the application of the SESAM framework implemented with (b) the ‘probability ranking’ rule implemented with the sum of probabilities from SDMs (pS-SDM), (c) the ‘probability ranking’ rule implemented with the richness estimation by the macroecological model (MEM) and (d) the ‘trait range’ rule (using the combination of the three traits as constraints).
136x44mm (300 x 300 DPI)



Histograms showing the proportion (mean among all plots) of true and false positive, as well as true and false negative for all the implementations of the SESAM framework, compared with the unconstrained sum of binary species distribution model (bS-SDM). In the upper plot results from the 'probability ranking' rule test implemented with macroecological models and sum of probabilities from SDMs (PRR MEM and PRR pS-SDM, respectively) and random selections (rand MEM and rand pS-SDM, respectively). In the lower plot results from the 'trait range' rule test for single traits and all their combinations (H, height; SLA, specific leaf area of the community; SM, seed mass).
152x136mm (300 x 300 DPI)