Ecography 33: 1070–1080, 2010
doi: 10.1111/j.1600-0587.2010.06301.x

Subject Editors: Carsten Rahbek and Niklaus Zimmermann. Accepted 18 August 2010

Geography, topography, and history affect realized-to-potential tree species richness patterns in Europe

Jens-Christian Svenning, Matthew C. Fitzpatrick, Signe Normand, Catherine H. Graham, Peter B. Pearman, Louis R. Iverson and Flemming Skov

J.-C. Svenning (svenning@biology.au.dk) and S. Normand, Ecoinformatics and Biodiversity Group, Dept of Biological Sciences, Aarhus Univ., Ny Munkegade 114, DK-8000 Aarhus C, Denmark. – M. C. Fitzpatrick, Univ. of Maryland Center for Environmental Science, Appalachian Lab, Frostburg, MD 21532, USA. – C. H. Graham, Dept of Ecology and Evolution, Stony Brook Univ., NY 11794, USA. – P. B. Pearman, Land Use Dynamics, Swiss Federal Research Inst. WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland. – L. R. Iverson, Northern Research Station, US Forest Service, Delaware, OH 43015, USA. – F. Skov, Dept of Wildlife Ecology and Biodiversity, National Environmental Research Inst., Aarhus Univ., Grenaavej 14, DK-8410 Kalø, Denmark.

Environmental conditions and biotic interactions are generally thought to influence local species richness. However, immigration and the evolutionary and historical factors that shape regional species pools should also contribute to determining local species richness because local communities arise by assembly from regional species pools. Using the European tree flora as our study system, we implemented a novel approach to assess the relative importance of local and regional mechanisms that control local species richness. We first identified species pools that tolerate particular local environments and quantified the proportion of the pool that is present locally, i.e. the realized/potential (R/P) richness ratio. Because no consensus exists on how to estimate potential richness, we estimated it using three different approaches. Using these three estimates separately and in a combined ensemble estimate, we then analyzed the effects of potential drivers on R/P richness ratios. We predicted that the R/P richness ratio would 1) increase with decreasing distance from glacial refugia (accessibility), 2) and be generally low in geographically fragmented southern Europe because of dispersal limitation; 3) increase with actual evapotranspiration because greater availability of water and energy promotes local population persistence; and 4) increase with topographic heterogeneity because it promotes local species coexistence and facilitates long-term species survival. There was considerable variation among the three R/P richness ratio estimates, but we found consistent support for a negative effect of regional geographic fragmentation and a positive topographic effect. We also identified fairly broad support for the predicted effect of accessibility. We conclude that local tree assemblages in Europe often fail to realize a large proportion of the potential richness held in the regional species pool, partially reflecting their geographical, historical, and environmental circumstances. The dispersal-related effects of geographic fragmentation and accessibility exemplify regional controls that combine with local ecological sorting to determine local species richness.

A central question in ecology and evolution addresses the determinants of species richness. The species richness of a locality is generally recognized to be influenced by its environmental conditions and biotic interactions. Nevertheless, because local assemblages arise by assembly from regional species pools, colonization and the evolutionary and historical factors that shape regional species pools should also affect local species richness (Zobel 1997, Ricklefs 2004). While most ecologists and biogeographers acknowledge the roles of both local and regional processes, strong disagreement persists about their relative importance in shaping local assemblages (O'Brien et al. 2000, Francis and Currie 2003, Ricklefs 2004). For example, some authors have interpreted regressions of local species richness on regional richness as evidence for the dependence of local richness on the number of species in the regional species pool (Caley and Schluter 1997), but this approach may not differentiate between local or regional control of local richness (Hillebrand 2005). Further, while numerous studies have analyzed geographic richness patterns and conducted meta-analyses to synthesize these patterns, no consensus has emerged on the importance of local versus regional control of species richness (O'Brien et al. 2000, Francis and Currie 2003, Hawkins et al. 2003, Hortal et al. 2008, Stropp et al. 2009). One likely reason is that local and regional mechanisms are not mutually exclusive but instead act simultaneously, with the outcome depending on the study system. For instance, greater climatic stability may cause local richness to be more a function of local determinism in the tropics than at high-latitudes, which were strongly impacted by the Pleistocene glaciations (cf. McGlone 1996). High-latitude regions can exhibit strong regional effects on species richness because of geographic variability in the impact of the glaciations (Latham and

Ricklefs 1993, Ricklefs 2004, Svenning and Skov 2007a). Nevertheless, tropical rainforest regions may also show geographic variation in species richness driven by regional legacies of past climate change (Stropp et al. 2009). Given this complexity, an analytical framework that integrates both local and regional mechanisms is required for improving our understanding of the causes of variation in local species richness.

In this paper, we implement a novel approach to assess the local vs regional control of local species richness and the probable mechanisms involved. Local environment controls the realized local species pool by ecological sorting. Hence, we first identify species pools that tolerate particular local environments and quantify the proportion of the pool that is present locally. This proportion defines the realized/potential (R/P) richness ratio. Because the R/P richness ratio directly takes ecological sorting into account, analyzing patterns in this diversity measure allows us to focus on a suite of less well-understood factors that potentially impact local richness. We assess two hypotheses regarding the potential regional drivers, and two hypotheses regarding the potential local drivers of geographic variation in the R/P richness ratio. The R/P richness ratio should reflect regional control by 1) decreasing with increasing isolation from glacial refugia (accessibility) when postglacial range expansion of species is dispersal limited (Svenning and Skov 2007a) and 2) by decreasing with increasing fragmentation of a region by physical dispersal barriers such as sea straits and mountain ranges (Nekola and White 1999, Bjorholm et al. 2008). That is, if immigration from the regional species pool controls local richness, then the degree to which local community richness approaches the regional species pool will decrease with the increasing importance of dispersal limitation as a range constraint (cf. MacArthur and Wilson 1967). In contrast, the R/P richness ratio should reflect local control 1) by increasing with increasing availability of water and energy, as these resources should promote the persistence of local populations (O'Brien et al. 2000, Hawkins et al. 2003, Currie et al. 2004), and 2) by increasing with topographic heterogeneity, which provides greater habitat diversity and surface area, thereby promoting local species coexistence (Simpson 1964, O'Brien et al. 2000). Topographic heterogeneity additionally may impact diversity as a more regional-level control by facilitating the long-term local survival of species despite the repeated Pleistocene macroclimatic changes (Bennett et al. 1991, McGlone 1996, Lobo et al. 2001, Svenning and Skov 2007b).

We use the European tree flora as our study system because its current ecology and historical distribution is relatively well understood, making it an ideal system for critical interpretation of the R/P richness ratio as a new diversity measure. Notably, Pleistocene climate changes have strongly influenced European tree diversity and tree species ranges (Bennett et al. 1991, Latham and Ricklefs 1993, Svenning 2003, Svenning and Skov 2004, Leroy and Arpe 2007, Svenning et al. 2008a), and current species richness patterns reflect these historical dynamics (Svenning and Skov 2005, 2007a, b). However, contemporary climate, other environmental factors, and the intense human impact on the European landscapes join these legacies as critical influences on tree species richness patterns in Europe (Svenning and Skov 2005, 2007a, b). Finally, given that our study is the first to investigate the R/P richness ratio, no established methodology exists for estimating it. We explore how three different methods of species pool estimation influence the magnitude and geographic patterns of the R/P richness ratio and inferences on its drivers. We note that the regional species pool in the present study is estimated taking the whole European tree flora into consideration. Species distribution modeling (SDM) offers an obvious approach for estimating potential richness (P) and hence R/P, but there are numerous SDM algorithms with marked differences in how closely they can fit complex species–climate relationships (Guisan and Zimmerman 2000, Elith et al. 2006). We implement an envelope model using a few key bioclimatic predictor variables, as an SDM that may be more likely to represent the potential distribution of a species than would a more flexible fit-maximizing approach. We also implement a state-of-the-art presence/absence data modeling approach in combination with more comprehensive sets of predictor variables. This approach likely fits realized distributions closely, with fewer areas of overprediction, thus providing conservative estimates of P. Finally, as an alternative to the SDM-based approaches, we use quantile regression to estimate potential richness by modeling maximum species richness in relation to environmental predictors (Cade and Noon 2003). Here, potential richness is estimated without the need to estimate the potential distribution of each individual species, but is limited to be no greater than the maximum observed richness and might thus be more constrained by the realized species richness patterns than the simple envelope SDM approach. While the approaches are strongly divergent methodologically, they are also all potentially valid and comparisons among them of their estimates of P and R/P and their spatial patterning remain unexplored.

Materials and methods

Study region and species

We used tree species distribution data from Atlas Florae Europaeae (AFE) that were mapped on an equal-area grid with cells of approximately 50 × 50 km (AFE cells; Jalas and Suominen 1972–1994). The study area (n = 2264 AFE cells) included most of Europe within 34°N–71°N and 11°W–32°W. We excluded the former Soviet Union, where the sampling effort has been relatively limited and variable, and a small number of mostly coastal cells with missing environmental data (Fig. 1). The study species were the 58 species of tall trees (≥ 20 m) in AFE that occur in the study area, approximately 70% of the ≥ 20-m-tall trees in Europe. We use the same dataset as Svenning and Skov (2005), but include several additional species: *Abies nebrodensis*, *Juniperus thurifera*, *Platanus orientalis*, *Quercus trojana*, and *Salix caprea*. Further, several taxonomic changes have been made: *Pinus mugo*, *P. rotundata*, and *P. uncinata* have been placed in *P. mugo*, *Quercus pedunculiflora* has been merged into *Q. robur*, *Q. rotundifolia* into *Q. ilex*, and *Q. sicula* into *Q. pubescens*. We included all native occurrences as well as occurrences of uncertain native status in the modeling, while naturalized occurrences were excluded (status from AFE).

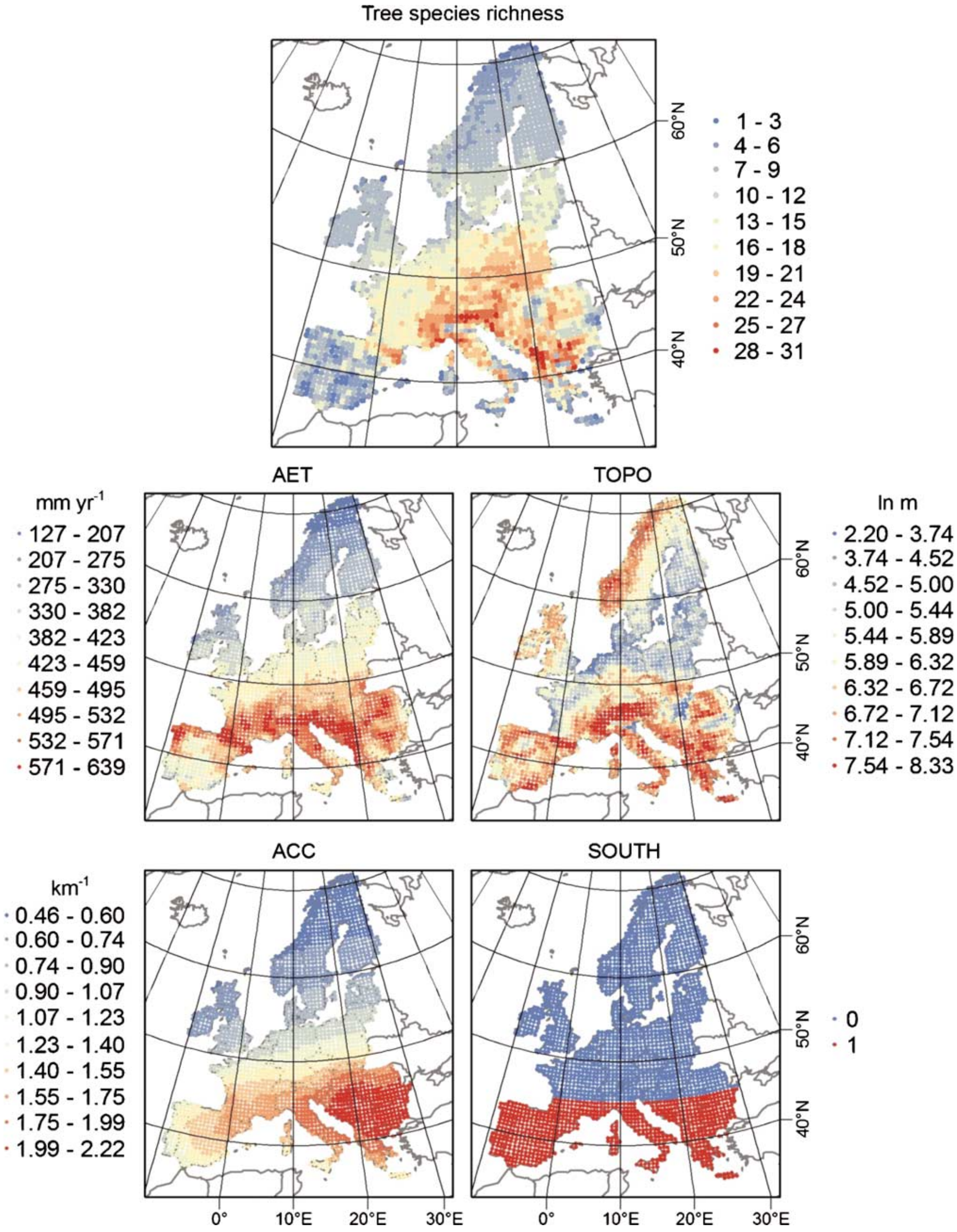

Figure 1. Atlas Florae Europaeae tree species richness and the four predictors of the realized-to-potential species richness ratio used (actual evapotranspiration, AET; log-transformed topographic range, TOPO; accessibility to colonization from the ice age refugia, ACC; an indicator variable for southern Europe [latitude < 45.0°], SOUTH). Ten color classes using Jenk's natural breaks are shown.

Predictor variables

Using the CRU CL 2.0 dataset of observed mean climate from 1961–1990 (New et al. 2002) (<www.cru.uea.ac.uk/cru/data/hrg.htm>) and following Svenning and Skov (2004), we computed two basic climatic variables, annual mean temperature (AMT) and annual precipitation (PANN), and three key bioclimatic variables for plant species ranges, growing-degree days (GDD, 5°C base temperature), water balance (WBAL), and absolute minimum temperature (TMIN). We also included two additional climatic variables that are often important in studies of diversity patterns, namely potential evapotranspiration (PET) and actual evapotranspiration (AET), obtained from the 1920–1980 mean climate data set of Ahn and Tateishi (1994) (<www.grid.unep.ch/GRID_search_details.php?dataid=GNV183>; Fig. 1). We used the Monthly Generalized Global Vegetation Index (GVI) from NESDIS NOAA/9 Weekly GVI Data for 1985–1988 (<ftp://ftp.ngdc.noaa.gov/Solid_Earth/Ecosystems/CEOS_Ecoregions/datasets/a01/>) as a direct measure of vegetation productivity. Three derived soil properties (pH, % sand [SAND], and $CaCO_3$ concentration [$CaCO_3$]) were computed as means of the corresponding top- and subsoil layers extracted from FAO (2003). Finally, we also used two topographic variables, computed from ~1-km resolution digital elevation data from the Shuttle Radar Topography Mission (Farr et al. 2007): slope (SLOPE), computed as the maximum rate of change between each ~1-km pixel and its neighbors, and ln-transformed altitudinal range

(TOPO) within an AFE grid cell, following O'Brien et al. (2000) (Fig. 1). Apart from TOPO, all environmental variables were recomputed as averages per AFE cell.

In addition to these local environmental variables, we also calculated two regional predictor variables. One variable, accessibility (ACC) to postglacial recolonization from the ice-age forest refugia, represents how easily a site could be reached by migration from the ice age refugia. We computed this variable by modifying and extending the approach described in Svenning and Skov (2007a) and Svenning et al. (2008b). For each AFE cell, we computed ACC as $\Sigma(d_{ij}^{-1})$, where d_{ij} is the distance from the focal cell i to a refuge cell j and the sum is taken across all j. Four estimates of the locations of the potential major forest refugia were computed by combining the two different glacial (Last Glacial Maximum, LGM) climate data sets developed in Svenning et al. (2008a) with two variations of the minimum climatic requirements for cool-temperate trees developed by Leroy and Arpe (2007). Areas assessed suitable for cool-temperate trees had GDD $\geq 800°$C, mean temperature of the coldest month $\geq -15°$C, and summer precipitation ≥ 50 mm, or fulfilled just the GDD and temperature requirements. The potential forest refugia were computed at 10' resolution and summarized per AFE cell, by classifying an AFE cell as a potential refuge if it included at least one 10' pixel classified as a potential refuge, for Europe and Asia Minor east to 50.0° longitude. Hence, ACC incorporates postglacial recolonization from refuge areas east of the European study area otherwise not used in this study. Using an ensemble approach (Araujó and New 2007), we represented ACC in the further analyses as the mean of the four estimates, i.e. across two LGM climate scenarios and two models for the climate conditions needed for major tree refugia (Fig. 1). Last, an indicator variable was used to capture the contrast between the geographically fragmented southern Europe and the more continuous and homogeneous northern Europe (SOUTH: 1: latitude $<45.0°$; 0: latitude $\geq 45.0°$; Fig. 1).

Estimating realized-to-potential richness ratio

Potential richness is an estimate of the number of species that could occur in an AFE cell given its environment, i.e. the ecologically relevant part of the regional species pool. Because no consensus exists on how to estimate potential richness, we used three distinct approaches to compute the potential richness of a given AFE cell. Two approaches were based on estimating the potential distributions of individual species using species distribution modeling (SDM: Guisan and Zimmerman 2000) and computing the potential richness by summing these; the third approach was based on modeling potential species richness using quantile regression (Cade and Noon 2003). We evaluated the R/P richness ratio obtained using each of these approaches independently and then, because it was unclear which method would lead to the most accurate estimate, we also combined the models using an ensemble approach, which handles modeling uncertainties by emphasizing consistent patterns (Araujó and New 2007).

One of the SDM approaches (hereafter termed the proSDM approach) used a profile technique based on the modified, fuzzy version of the simple rectilinear bioclimatic envelope approach described in Svenning and Skov (2004). We used only three bioclimatic variables (GDD, TMIN, WBAL) for which there is strong a priori knowledge of their importance in plant physiology to predict species distributions (Svenning and Skov 2004). For each species, all AFE cells in which a given species was recorded as present were selected, and the following values were calculated for each bioclimatic predictor based on its frequency distribution: the 5th percentile (B), the 25th percentile (C), the 75th percentile (D), and the 95th percentile (E). To avoid overestimating the envelopes due to potential outliers, the lower (A) and upper (F) limits of the envelope for a given variable were then calculated as $A = B - 0.1 \times (C - B)$ and $F = E + 0.1 \times (E - D)$, respectively. Finally, a species was considered potentially present in a given AFE cell when the cell's environment fell within the estimated envelope for all three bioclimatic predictors.

The other SDM approach (hereafter termed the paSDM approach) was based on sophisticated presence/absence data modeling (group discrimination techniques) using the BIOMOD framework (Thuiller et al. 2009). BIOMOD contains multiple algorithms for modeling species distributions; for this study, we used both generalized linear models (GLMs) and random forest (RF). Both of these methods are considered robust for capturing species–environment relationships (Elith et al. 2006, Lawler et al. 2006, Iverson et al. 2008), but differ markedly in their approach for doing so. For each species, we used the two modeling algorithms with two sets of predictor variables: a simple set consisting of only the five variables GDD, TMIN, WBAL, pH, and SLOPE, and an expanded set additionally including AMT, PANN, AET, PET, GVI, SAND, and $CaCO_3$. Both predictor sets were larger than the proSDM set because we wanted to represent more comprehensive sets of potential environmental range controls in the paSDM approach. The GLMs were fitted with linear, quadratic, and cubic polynomial terms, and Akaike's information criterion (AIC) was used to select the most parsimonious model (Burnham and Anderson 2002). For RF, we used BIOMOD's default settings and grew 500 classification trees per species. Within BIOMOD, distribution data for each species were partitioned randomly 50 times into calibration (70%) and evaluation (30%) datasets, and models were run on each of the 50 resulting datasets for each species. Model accuracy was calculated using the True Skill Statistic (TSS, Allouche et al. 2006). The TSS, an established approach for assessing the accuracy of weather forecasts, compares the number of correct forecasts, minus those attributable to random guessing, to that of a hypothetical set of perfect forecasts. For a 2×2 confusion matrix, TSS is defined as $\text{TSS} = \text{sensitivity} + \text{specificity} - 1$. It takes into account both omission and commission errors and success as a result of random guessing, and ranges from -1 to $+1$, where $+1$ indicates perfect agreement and values of zero or less indicate a performance no better than random. The TSS has recently been shown to be superior to kappa, as the former is not affected by species prevalence (Allouche et al. 2006). The probability of occurrence was converted into binary presence/absence using the threshold maximizing TSS. To synthesize the four resulting potential richness estimates (2 algorithms $\times$ 2 sets of predictor variables), we took the average, i.e., again using an ensemble approach (Araujó and New 2007). Because the SDM approaches deteriorate with small sample sizes, both proSDM and paSDM were implemented only for species with at least 20 occurrences ($n = 52$ of 58 species), and the R/P richness ratios were consequently computed using the realized richness based on only these species. However, estimates of realized richness based on either the reduced or full species data sets were highly correlated ($r = 0.9997$).

As an alternative to the SDM approaches, we used quantile regression to model the maximum tree species richness across Europe given the environmental conditions (Cade and Noon 2003). We first selected the environmental predictors based on spatial regression modeling implemented using ordinary least-squares (OLS) multiple regression. Based on the literature on plant diversity controls (O'Brien et al. 2000, Hawkins et al. 2003, Svenning et al. 2009), we considered the following suite of 13 models: $\mathrm{PANN} + \mathrm{PET} + \mathrm{PET}^2 + \mathrm{TOPO}$, $\mathrm{PANN} + \mathrm{PET} + \mathrm{PET}^2$, AET + PET + WBAL + TOPO + GVI + pH, AET + PET + WBAL + TOPO, AET + WBAL + TOPO, AET + PET + TOPO, PET + WBAL + TOPO, AET + PET + WBAL, AET + WBAL, AET + PET, PET + WBAL, AET + TOPO, and AET. Model selection was done using AIC-based information-theoretic model selection (Burnham and Anderson 2002). Because significant spatial autocorrelation remained in the regression residuals of the best model, we then repeated the regression modeling using eigenvector-based spatial filtering (Diniz-Filho and Bini 2005) to ensure that spatial autocorrelation did not bias the model selection (Diniz-Filho et al. 2008). The spatial filters were generated as the eigenvectors of a Principal Coordinate Analysis of a pairwise matrix of geographic distances between all AFE cells, following the approach of Diniz-Filho and Bini (2005). Geographic distances were truncated at 600 km because at approximately this distance, positive spatial autocorrelation became negligible (Diniz-Filho and Bini 2005). The spatial filters were then added to the OLS regression models to control for spatial autocorrelation, thus allowing unbiased estimates of regression coefficients and their significance levels. The selection of spatial filters was done for the best unfiltered model, adding successively higher-order filters until spatial autocorrelation was negligible in the regression residuals (Moran's $I < 0.100$ for the 22 default distance classes in SAM [Rangel et al. 2006]). We considered only the first 50 filters in the modeling.

After establishing the best predictive model for tree species richness, we then reran the model using quantile regression (Cade and Noon 2003) to determine the maximum observed richness given the predictor variables in the best spatial regression model. Quantile regression allows analysis of the upper or lower limits of a variable rather than focusing on the mean. By modeling maximum species richness, this approach represents a conservative assessment of how many species could potentially coexist in an AFE cell given its environmental conditions. Similarly, quantile regression has been used to model habitat invasibility, as modeling of the upper limit of the richness of invasive species would be more meaningful than mean richness because many exotics are likely to be dispersal limited and absent from many suitable sites (Brown and Peet 2003). Here, we represented the maximum tree species richness by modeling the 95% quantile. After fitting the 95% quantile regression model, we then projected it back onto the study area to estimate the potential maximum richness per AFE cell. Although the AET + PET + WBAL + TOPO + GVI + pH model clearly emerged as the best (with an Akaike weight of 1.00 (without filters) and 0.98 (with filters)), the adjusted R^2 of the second-best and more parsimonious model, AET + PET + TOPO, was only marginally lower (0.579 vs 0.585 without filters and 0.725 vs 0.726 with filters). Therefore, using an ensemble approach, we calculated the potential maximum richness as the mean of the predictions from these two models.

For each of the three estimates of the potential richness per AFE cell, we computed the R/P richness ratio, resetting the ratio to 1 whenever it exceeded 1 or when the potential richness was estimated to be zero, since the meaningful interpretation in either case would be that as many species as could be present were indeed present. We assessed the consistency of the three estimates in terms of magnitude and pattern, using paired t-tests and correlation, respectively. To develop the ensemble prediction from the three approaches used to calculate R/P ratios, we subjected the three R/P ratio estimates to a Principal Components Analysis (PCA) using their correlation matrix (to ensure that the three estimates received equal weight). We then used the resulting first principal component as the ensemble consensus estimate (PC-R/P) because it was as close to all of the R/P estimates as possible.

Realized-to-potential richness ratio determinants

We used OLS multiple regression to test the four hypothesized determinants of spatial variation in the R/P richness ratio. We also ran the regression models with spatial filtering (Diniz-Filho and Bini 2005) when non-negligible spatial autocorrelation remained in the OLS model residuals (Moran's $i \geq 0.100$ for the 22 default distance classes). We used the minimum spanning tree approach to select the truncation distance (254 km) for the computation of the spatial filters because spatial autocorrelation for PC-R/P became negligible after ~200 km. We implemented spatial filtering using two approaches. First, we used the residual SAC (spatial autocorrelation) approach, in which all filters that were significantly correlated to the OLS model residuals were added using a mixed stepwise selection procedure with probability-to-enter and probability-to-leave equal to 0.05. This is equivalent to the SEVM-v3 approach of Bini et al. (2009), which was highlighted as one of two methods that best simultaneously minimized residual spatial autocorrelation and coefficient shifts. The second approach was the variance inflation factor (VIF) approach, in which spatial filtering was done by successively adding higher-order filters to the OLS regression model until spatial autocorrelation was negligible in the regression residuals (Moran's $i < 0.100$ for the 22 default distance classes; Diniz-Filho and Bini 2005), but with two restrictions: first, a filter was added to the model only if it had a significant effect ($p < 0.05$), and was excluded again if it became non-significant with the addition of a higher-order filter. Second, to avoid introducing multicollinearity problems (Graham 2003), any filter that caused the VIF for any of the four explanatory variables to reach 5.000 or above was not added to the model. Only the first 75 filters were considered in the latter approach because the VIF limitation strongly constrained the scope for filter addition beyond the first 50 filters. For proportional data such as the R/P richness ratios, arcsin square-root transformation may improve normality of residuals. In the present case, the OLS residuals were close to normal (skewness and kurtosis $<|1.0|$) for R/P-proSDM and -QR, while less so for R/P-paSDM (skewness $= -1.50$ and kurtosis $= 4.72$). While transformation improved normality of the latter residuals, it decreased normality of residuals for for R/P-proSDM and -QR. Hence, we report the results based on the raw R/P richness ratio values, noting that transformation only caused minor changes in the importance of predictors.

We carried out the proSDM approach in ArcGIS 9.3 using a script by F. Skov, while for the paSDM approach we used BIOMOD (Thuiller et al. 2009) in R 2.7.2 (R Development Core Team 2008). We used the quantreq package (Koenker 2008) in R 2.8.1 (R Development Core Team 2008) to implement quantile regression. We used SAM 3.0 (Rangel et al. 2006) for computing spatial filters and the spatial R/P richness ratio modeling and used JMP 7.0 (SAS Inst., Cary, NC) for the PCA.

Results

The R/P richness ratio – consistency of the estimates

The R/P richness ratio estimates were particularly high (i.e. close to 1) for the paSDM approach, while the two other approaches provided similar, but much lower estimates (Table 1). In terms of their spatial patterning, the three approaches were only partially consistent (Fig. 2) with low (those involving proSDM) to moderate correlations (Table 1). Nevertheless, the ensemble consensus estimate of the R/P richness ratio (R/P-PC) represented consistent patterns among the three and accounted for 64% of their variation (Table 1). R/P-PC was most correlated to the paSDM- and QR-based estimates, with the major consistent patterns being a low R/P richness ratio on the Iberian Peninsula and a high ratio in the Alps and eastern Europe (Table 1, Fig. 2).

Determinants of the R/P richness ratio pattern

Regarding the drivers of geographic variation in the R/P richness ratio, we predicted that the ratio would 1) increase with increasing accessibility to postglacial recolonization from the ice-age forest refugia (ACC); 2) be generally low in the geographically fragmented southern Europe (SOUTH $= 1$); 3) increase with water-energy (represented by AET); and 4) increase with topographic heterogeneity (TOPO). Several aspects of the results complicate interpretation, however. Most notable were the strong differences among the models in the statistical relationships between the R/P richness ratio and its hypothesized determinants (Table 2). Furthermore, while the residual SAC filtering approach was the only approach that successfully removed residual spatial autocorrelation for all R/P richness ratio estimates, it was potentially compromised by multicollinearity, with high VIFs for all predictor variables (but especially for ACC; Table 2).

The most consistent pattern in the results was that the R/P richness ratio, as predicted, was lower in the geographically fragmented southern Europe (Table 2). The only model that did not show this effect was the model for the proSDM-based estimate under the residual SAC filtering approach. However, here a strong and unique negative relationship between ACC and the R/P richness ratio probably substituted for the south-vs-north contrast (Table 2). In addition, in all models and consistent with our hypothesis, there was a general positive relationship between the R/P richness ratio and TOPO, although this relationship was not always significant (Table 2).

The other hypothesized relationships did not receive as consistent support across all methodologies. The modeling of the ensemble consensus estimate (R/P-PC) indicated that the R/P richness ratio increased with AET, as predicted (Table 2). However, the relationship between AET and R/P richness ratio was always negative for the proSDM-based estimate and absent for the paSDM-based estimate (Table 2). The hypothesized positive relationship between the R/P

Table 1. a) Realized-to-potential tree species richness ratios (R/P) in Europe according to three methods for estimating potential richness: summing potential species distributions according to a bioclimatic profile technique for species distribution modeling (proSDM), or presence/absence-based species distribution modeling (paSDM), or maximum richness estimation based on quantile regression modeling (QR). Different superscript letters indicate significantly different means[†]. b) Pearson correlations of the three ratio estimates. c) The three principal components from a Principal Components Analysis of the three ratio estimates (in parentheses: eigenvalue, percent variance accounted for) and their Pearson correlations of the original ratio estimates. n =2264 AFE cells.

		R/P-proSDM	R/P-paSDM	R/P-QR
a)	Mean $\pm$SD^a	0.623 ± 0.268^a	0.903 ± 0.132^b	0.671 ± 0.201^a
b)	R/P-paSDM	0.405****		
	R/P-QR	0.252****	0.678****	
c)	PC1 (1.92, 63.9)[‡]	0.630****	0.901****	0.841****
	PC2 (0.78, 26.1)	0.769****	−0.156****	−0.410****
	PC3 (0.30, 10.0)	0.107****	−0.405****	0.354****

[†] Assessed by paired t-tests; $p < 0.0001$ in both significant cases, while $p = 0.45$ in the third case; $|t\text{-ratio}| = 0.76–74.9$, DF $= 2263$).
[‡] PC1 $=$R/P$-$PC (ensemble consensus estimate).
****: $p < 0.0001$.

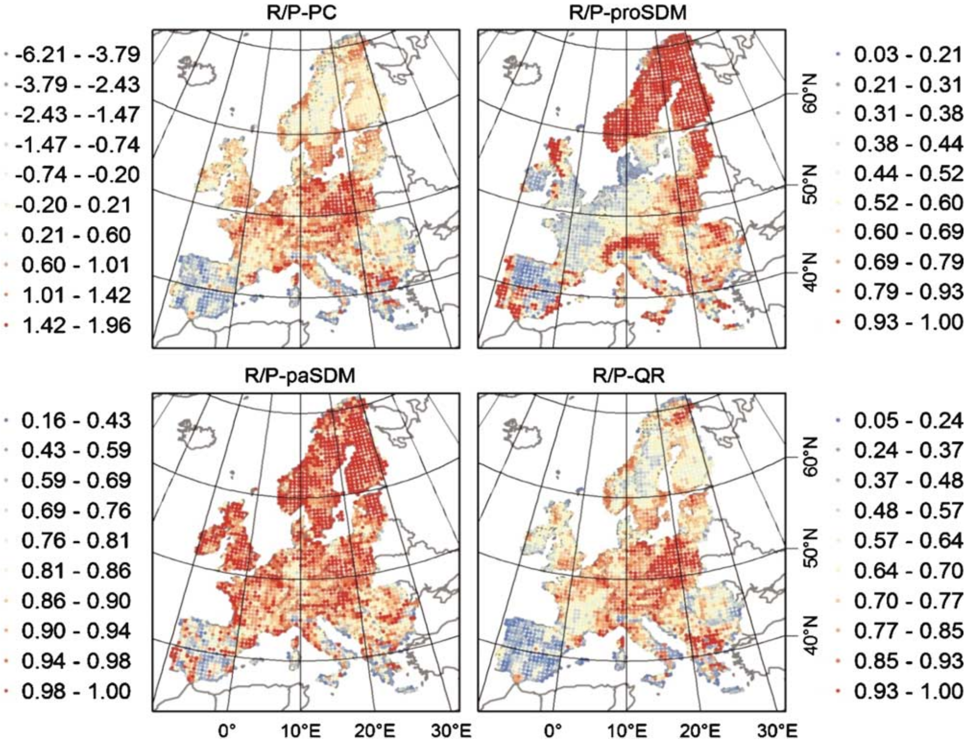

Figure 2. Realized-to-potential richness ratio (R/P) estimates for the European tree flora: an ensemble consensus estimate (R/P-PC, the first principal component of a Principal Components Analysis of the following three estimates), estimates based on species distribution modeling using a profile technique (R/P-proSDM) or an ensemble of models based on group discrimination algorithms for presence/absence data (R/P-paSDM), and an estimate based on using quantile regression to estimate maximum potential species richness (R/P-QR). Ten color classes using Jenk's natural breaks are shown.

richness ratio and ACC emerged consistently only for the quantile regression-based estimate, although it was also positive in the VIF filter approach models for R/P-PC and R/P-proSDM (Table 2). In contrast, the ACC relationship was negative for R/P-paSDM and according to the residual SAC filter model for R/P-proSDM (Table 2).

Discussion

What determines species diversity is a central, but outstanding issue in ecology, evolution, and biogeography (O'Brien et al. 2000, Francis and Currie 2003, Ricklefs 2004). Here we developed a new metric, the potential versus realized species richness (R/P richness) ratio, to evaluate a series of hypotheses on the factors that influence patterns of species distribution and richness in the European tree flora.

Methodological considerations in estimating R/P richness ratios

We identified strong methodological effects on estimates of the extent to which species from the regional species pools were represented in local assemblages (i.e. the R/P richness ratio), both in terms of magnitude and geographic pattern. Notably, the paSDM approach tended to estimate R/P richness ratios close to 1, suggesting that the environmentally relevant regional species pool was generally almost fully represented in each $50 \times 50\text{-km}^2$ AFE cell; in contrast, the two other methodologies provided similar and much lower R/P richness ratio estimates (0.6–0.7). The variability in the R/P richness ratio estimates, in turn, resulted in different conclusions regarding the importance of different local and regional drivers. The limited consistency in magnitude, spatial patterns, and inferred drivers underscores the need for an improved theoretical and methodological basis for estimating species' potential ranges or potential species richness directly.

The tradeoff between overestimating and obtaining narrow realized estimates of species' climatic tolerance is crucial when modeling species' potential ranges. The differences between the paSDM and proSDM illustrate this tradeoff. In this application, we were more confident of the proSDM-based results, as this approach better predicted the naturalized populations of the tree species studied, although it still may have over- or underestimated the potential ranges for some species (Supplementary material). Our implementation of the paSDM approach, which only used AFE records of ''native'' presence or absence, consistently failed to predict the naturalized populations north and west of the native ranges. However, given the input data the paSDM approach ''correctly'' predicted them as absences rather than as locations that naturalized populations could occur. The paSDM techniques could possibly be tuned to better account for such distributions, e.g. by adjusting the threshold set to convert the probability-of-occurrence to a binary prediction. A key issue may be the number of predictor variables, as using just a

Table 2. Multiple linear regressions of four estimates of R/P richness ratios (Fig. 2) against the four explanatory variables (ACC, SOUTH, AET, and TOPO)[a], without a) and including spatial filters using the b) VIF and c) residual SAC approaches. The standardized regression coefficients, model fit (represented by the adjusted R^2), the F ratio, and the maximum Moran's i for the residuals in the 22 default distance classes are given. VIF ranges, 1.248–4.923 (without filters), 1.468–4.999 (filters, VIF approach), and 2.824–26.361 (filters, residual SAC approach). n = 2264 AFE cells.

		R/P-PC	R/P-proSDM	R/P-paSDM	R/P-QR
a)	ACC	0.076[†]	0.041	−0.097*	0.212***
	SOUTH	−0.604***	−0.193***	−0.482***	−0.609***
	AET	0.192***	−0.378***	0.019	0.101**
	TOPO	0.019	0.328***	0.017	0.076***
	R^2_{adj}	0.262	0.209	0.279	0.230
	F	201***	150***	219***	169***
	Moran's i[b]	0.168	0.385	0.091	0.265
b)	ACC	0.079*	0.078*	–[c]	0.221***
	SOUTH	−0.607***	−0.207***	–	−0.643***
	AET	0.167***	−0.454***	–	−0.070*
	TOPO	0.032	0.309***	–	0.135***
	R^2_{adj}	0.326	0.343	–	0.338
	F	73.7***	108.3***	–	89.5***
	Moran's i[b]	0.112	0.269	–	0.188
c)	ACC	−0.036	−0.283***[d]	–[c]	0.137***
	SOUTH	−0.496***	−0.028	–	−0.504***
	AET	0.183***	−0.076	–	0.061
	TOPO	0.180***	0.317***	–	0.264***
	R^2_{adj}	0.553	0.707	–	0.642
	F	32.0***	48.3***	–	38.0***
	Moran's i	0.010	0.025	–	0.012

[a] ACC, accessibility to colonization from the ice age refugia; SOUTH, an indicator variable for southern Europe [latitude <45.0°]; AET, actual evapotranspiration; TOPO, log-transformed topographic range.

[b] The maximum Moran's i was always for the innermost (≤ 184 km) of the 22 distance classes.

[c] Given the little spatial autocorrelation in the residuals of the model without spatial filters, we did not implement the filter models for the paSDM-based analyses.

[d] VIF = 26.4; in all other cases VIF ≤ 13.8.

[†]: $p \leq 0.06$, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

few predictors as in the proSDM implementation will lead to relatively broad estimates of niches and potential distributions, while using rather many predictors as in the paSDM implementation will have the opposite effect, irrespective of the SDM algorithms used. Hence, the number of predictors used is likely to affect R/P richness ratio estimates, making careful consideration of what factors define a species' niche at the scale studied of paramount importance. Unfortunately, this remains a largely unresolved issue, even for a well-studied group like the European trees. Alternative SDM methodologies to those implemented here are being used to handle non-equilibrium distributions. For example, using spatial filters as predictors (spatial eigenvector mapping) can improve predictions in such cases (De Marco et al. 2008; cf. Blach-Overgaard et al. 2010 for an empirical case). However, the inclusion of spatial constraints in the modeling will cause model predictions to represent realized rather than potential distributions, and hence will not be suitable for estimating P and R/R richness ratios. The quantile regression-based approach directly estimates potential richness and thus avoids the complications associated with estimating potential distributions. However, potential richness can never be higher than maximum observed richness in this method (Cade and Noon 2003), thereby somewhat constraining P towards R. Nevertheless, the low R/P-QR values – similar to R/P-proSDM – show that this effect was not strong in our case. We emphasize that also in this approach the number of predictors will determine how closely the P estimate will be constrained by R. Regarding the ensemble consensus estimate, it was particularly representative for the paSDM-based method and must therefore to some extent be similarly compromised. Given these considerations, when interpreting our results we emphasize the most consistent patterns. Following the ensemble approach logic (Araujó and New 2007), we considered relationships that consistently emerged using all four R/P richness ratio estimates as those with the strongest support. However, we also discuss relationships that were consistent for just the proSDM and quantile regression approaches as these seemed to provide the most credible R/P richness ratio estimates (see above).

An additional methodological source of variability also requires consideration, namely the regression modeling used to assess the importance of the potential R/P richness ratio determinants. Spatial autocorrelation in the model residuals and multicollinearity are two factors that may compromise regression results (Graham 2003, Dormann et al. 2007). In the present study, positive residual spatial autocorrelation compromised the OLS regression results. To account for this spatial autocorrelation, we repeated the regression modeling using two implementations of eigenvector-based spatial filtering (Diniz-Filho and Bini 2005). The residual SAC (spatial autocorrelation) approach, as expected, most effectively removed residual spatial autocorrelation, but at the same time introduced multicollinearity, with VIFs for ACC, SOUTH, and AET always > 5 and often > 10. The VIFs were especially high for ACC, reaching 26.4 for the proSDM approach. High VIF values such as these can bias modeling results (Graham 2003). As the alternative (VIF) filter selection approach at least partially handled spatial autocorrelation without introducing strong multicollinearity, we focus on the results based on this method. Nevertheless, the two modeling approaches provided generally similar results.

Local assemblages fail to realize a large proportion of the regional species pool

The proSDM- and quantile regression-based results suggested that, in general, the realized local species richness is considerably less than the environmentally relevant regional species pool, namely 62–67%. These results therefore support the general hypothesis that ecological filtering by the local environment does not exclusively determine local richness, but that other mechanisms also play a role (Caley and Schluter 1997, Zobel 1997, Ricklefs 2004). In relation to the European tree flora, our study extends previous findings that many European tree species only partially fill their potential climatic range (Svenning and Skov 2004). We note that although the proSDM methodology is identical to that of Svenning and Skov (2004), the quantile regression approach is a very different methodology; nevertheless, it also resulted in low R/P richness ratio estimates. Given that local assemblages fail to realize a large proportion of the environmentally relevant regional species pool, understanding factors that determine local representation of regional diversity becomes important (Caley and Schluter 1997).

Determinants of the R/P richness ratio pattern

With respect to the four hypothesized drivers, the most consistent and generally strongest pattern was the southern vs northern Europe contrast, with the R/P richness ratio being lowest in southern Europe, as predicted from the greater geographic fragmentation of this region. This result suggests that mountain ranges and seaways can limit dispersal and thereby influence the extent to which regional species pools are represented in local assemblages. This finding is in agreement with island biogeographic theory, which predicts lower species richness where isolation is greater (MacArthur and Wilson 1967). Numerous studies have found support for isolation effects in island or island-like systems (Kreft et al. 2008), but not previously in a broad-scale continental setting. Greater spatial turnover in species composition in geographically complex regions such as mountainous areas and archipelagos supports the influence of such barriers on individual species ranges (Nekola and White 1999, Bjorholm et al. 2008). Considering European trees, we also note that numerous species have limited distributions within southern Europe, without obvious substitute species in the unoccupied regions. For instance, *Carpinus betulus*, *Castanea sativa*, *Juglans regia*, *Ostrya carpinifolia*, *Platanus orientalis* and *Quercus cerris* are wholly or largely missing from the Iberian Peninsula, and are not replaced by sister taxa in this region (Jalas and Suominen 1972–1994). For some species, the current absence or rarity of native populations on the Iberian Peninsula can be attributed to failure to re-immigrate after local extirpation during the Pleistocene glacials (Postigo Mijarra et al. 2008), probably because of the barriers to migration posed by the Pyrenees and the Mediterranean Sea. We note that the southern vs northern Europe contrast only represent patterns of geographic fragmentation in a very simple manner, and additionally will also incorporate other environmental and historical differences between north and south, e.g. in terms of glacial refugia locations for trees (Bennett et al. 1991, Leroy and Arpe 2007, Svenning et al. 2008a).

Numerous previous studies have documented positive effects of topographic heterogeneity on species richness (Simpson 1964, Kerr and Packer 1997, O'Brien et al. 2000, Kreft et al. 2008). We extend these results to show that topographic heterogeneity enhances the extent to which regional species pools are represented locally. Topographic heterogeneity had a moderately strong and relatively consistent positive effect on the R/P richness ratio, supporting the hypothesis that topographic heterogeneity facilitates a greater local representation of the regional species pool. This higher R/P ratio may be accomplished by promoting local species coexistence due to greater habitat diversity (strong local climatic and edaphic gradients) and surface area (Simpson 1964, O'Brien et al. 2000), as well as long-term persistence of tree species populations under climate change (McGlone 1996, Lobo et al. 2001, Svenning and Skov 2007b). In addition, species losses due to past human impacts may have been smaller in topographically heterogeneous mountain areas, where human land use has been less intensive (Lobo et al. 2001).

The R/P richness ratio generally increased with decreasing distance from ice-age forest refugia, especially for the quantile regression-based results. This finding supports the hypothesis that long-term lags in migration in relation to past climate changes may still limit species distributions and broad-scale richness (Svenning and Skov 2007a, Svenning et al. 2008b), and also affect the richness of local assemblages by limiting the extent to which regional species pools become represented locally. Although the proSDM approach produced low R/P richness ratio estimates for Western Europe, in agreement with the postglacial migration lag hypothesis, the high R/P richness ratio estimates for Scandinavia were contrary to this hypothesis (cf. the ACC pattern in Fig. 1). This pattern may result from a small species pool for the harsh boreal climate, in combination with more numerous and more northerly ice-age refugia for boreal species (Willis and van Andel 2004, Svenning et al. 2008a), facilitating rapid early range expansions and diminishing the limiting effect of postglacial migration lags. A recent study of forest plant species in northern Europe found that species with wide boreal distributions, as opposed to more strictly cool-temperate species, did not have consistently higher prevalence where accessibility to postglacial recolonization was high (Svenning et al. 2008b). In addition, the coarse grain size in this study likely affected our ability to assess accurately the cold climate limits for alpine species. Notably, *Abies alba*, *Larix decidua*, *Pinus cembra*, and *P. mugo*, important alpine tree species that are not currently native in Scandinavia, were mainly predicted to have suitable areas in the southern parts of Scandinavia (cf. Svenning and Skov 2004). However, the cold climate limits of these species have been somewhat underestimated, with *Abies alba* reported to be naturalized in southern Norway and *Larix deciduas, Pinus mugo* to be widely naturalized even to northern Norway, and *Pinus cembra* to have escaped in scattered locations as far north as northern Norway (Lid and Lid 1994). Hence, the high R/P richness ratios for northern Europe according to the proSDM-approach are probably too high.

We found no consistent support for a positive dependency of the R/P richness ratio on AET. In fact, the relationship was often estimated to be negative. This finding contrasts with the hypothesized positive effect of water-energy availability on species coexistence, and on local persistence more generally (O'Brien et al. 2000, Hawkins et al. 2003, Currie et al. 2004). While ecological filtering by the local environment clearly affects realized local tree species richness, we did not find much support for this hypothesized additional environmental effect. Similarly, in their review of the evidence for hypothesized climate-based mechanisms of coarse-scale variation in species richness, Currie et al. (2004) concluded partially in favor of local ecological filtering according to species' tolerance limits but did not find consistent support for two other climate-based mechanisms. Notably, the available data in that study did not support proposed links between energy and the number of individuals or between the number of individuals and species richness, i.e. local persistence (Currie et al. 2004). Regarding the third climate-based mechanism, climatic effects on speciation rates (Currie et al. 2004), geographic variation in speciation rates is unlikely to be of much importance for European tree diversity patterns given the pre-Quaternary age of most species (Mai 1995) and the subsequent strong reshuffling of their ranges induced by the Quaternary climatic oscillations (Bennett et al. 1991, Svenning et al. 2008a). Methodological issues may also contribute to the absence of support for the expected AET effect. Notably, AET may strongly vary within 50 × 50 km grid cells, especially under the warm southern macroclimate. Hence, we cannot rule out that AET may have a stronger effect at a finer spatial resolution.

Conclusions

We believe that the R/P ratio developed here will provide a useful new metric to explore mechanisms driving variation in species diversity. It complements the commonly used metric of species richness by adding additional information on how many species could be in a given location. For Europe (at a $50 \times 50\text{-km}^2$ grain) we provide evidence that local tree assemblages often do not contain all of species from the environmentally relevant regional species pool. Moreover, the proportion of the regional species pool that is represented in local assemblages is predictable from the geographical, historical, and environmental factors. We found consistent support for the hypotheses that geographic fragmentation of a region by physical dispersal barriers lowers the R/P richness ratio, while local topographic heterogeneity increases it. We obtained less conclusive results for the hypotheses that increasing isolation from the glacial refugia should decrease the R/P richness ratio or that the energy–water balance influences this ratio. The effects of geographic fragmentation and accessibility to recolonization clearly exemplify regional controls that supplement ecological sorting by the local environment in determining the local species richness (cf. Zobel 1997, Ricklefs 2004). Importantly, our results varied depending on the method used to estimate the size of the region species pool, highlighting the latter as a crucial, but non-trivial issue. Future studies should therefore focus not only on assessing the generality of the R/P richness ratio relationships found in the present study, but also on strengthening the theoretical and modeling basis for estimating local potential richness.

Acknowledgements – This paper is an outcome of the workshop "The Utility of Species Distribution Models as Tools for Assessing Impacts of Global Change", held in Riederalp, Switzerland, 11–16 August 2008. We thank the Atlas Florae Europaeae project for access to the distribution data. We gratefully acknowledge economic support from the Danish Natural Science Research Council (grant #272-07-0242 to JCS) and the Faculty of Science, Aarhus Univ. (to SN).

References

Ahn, C.-H. and Tateishi, R. 1994. Development of global 30-minute grid potential evapotranspiration data set. – J. Jap. Soc. Photogramm. Remote Sens. 33: 12–21.

Allouche, O. et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). – J. Appl. Ecol. 43: 1223–1232.

Araujó, M. B. and New, M. 2007. Ensemble forecasting of species distributions. – Trends Ecol. Evol. 22: 42–47.

Bennett, K. D. et al. 1991. Quaternary refugia of north European trees. – J. Biogeogr. 18: 103–115.

Bini, L. M. et al. 2009. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. – Ecography 32: 193–204.

Bjorholm, S. et al. 2008. To what extent does Tobler's 1st law of geography apply to macroecology? A case study using American palms (Arecaceae). – BMC Ecol. 8: 11.

Blach-Overgaard, A. et al. 2010. Determinants of palm species distributions across Africa: the relative roles of climate, non-climatic environmental factors, and spatial constraints. – Ecography 33: 380–391.

Brown, R. L. and Peet, R. K. 2003. Diversity and invasibility of southern Appalachian plant communities. – Ecology 84: 32–39.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multi-model inference: a practical information-theoretic approach. – Springer.

Cade, B. S. and Noon, B. R. 2003. A gentle introduction to quantile regression for ecologists. – Front. Ecol. Environ. 1: 412–420.

Caley, M. J. and Schluter, D. 1997. The relationship between local and regional diversity. – Ecology 78: 70–80.

Currie, D. J. et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. – Ecol. Lett. 7: 1121–1134.

De Marco, P. et al. 2008. Spatial analysis improves species distribution modelling during range expansion. – Biol. Lett. 4: 577–580.

Diniz-Filho, J. A. F. and Bini, L. M. 2005. Modelling geographical patterns in species richness using eigenvector-based spatial filters. – Global Ecol. Biogeogr. 14: 177–185.

Diniz-Filho, J. A. F. et al. 2008. Model selection and information theory in geographical ecology. – Global Ecol. Biogeogr. 17: 479–488.

Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609–628.

Elith, J. et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. – Ecography 29: 129–151.

FAO 2003. Digital soil map of the world and derived soil properties. – CD-ROM, Rome, Italy.

Farr, T. G. et al. 2007. The Shuttle Radar Topography Mission. – Rev. Geophys. 45, RG2004, doi:10.1029/2005RG000183.

Francis, A. P. and Currie, D. J. 2003. A globally consistent richness-climate relationship for angiosperms. – Am. Nat. 161: 523–536.

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. – Ecology 84: 2809–2815.

Guisan, A. and Zimmerman, N. E. 2000. Predictive habitat distribution models in ecology. – Ecol. Model. 135: 147–186.

Hawkins, B. A. et al. 2003. Energy, water, and broad-scale geographic patterns of species richness. – Ecology 84: 3105–3117.

Hillebrand, H. 2005. Regressions of local on regional diversity do not reflect the importance of local interactions or saturation of local diversity. – Oikos 110: 195–198.

Hortal, J. et al. 2008. Regional and environmental effects on the species richness of mammal assemblages. – J. Biogeogr. 35: 1202–1214.

Iverson, L. R. et al. 2008. Estimating potential habitat for 134 eastern US tree species under six climate scenarios. – For. Ecol. Manage. 254: 390–406.

Jalas, J. and Suominen, J. (eds) 1972–1994. Atlas Florae Europaeae: distribution of vascular plants in Europe, Vol. 1–10. – The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

Kerr, J. T. and Packer, L. 1997. Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. – Nature 385: 252–254.

Koenker, R. 2008. quantreg: quantile regression. – R package ver. 4.26.

Kreft, H. et al. 2008. Global diversity of island floras from a macroecological perspective. – Ecol. Lett. 11: 116–127.

Latham, R. E. and Ricklefs, R. E. 1993. Global patterns of tree species richness in moist forests: energy-diversity theory does not account for variation in species richness. – Oikos 67: 325–333.

Lawler, J. J. et al. 2006. Predicting climate-induced range shifts: model differences and model reliability. – Global Change Biol. 12: 1568–1584.

Leroy, S. A. G. and Arpe, K. 2007. Glacial refugia for summer-green trees in Europe and south-west Asia as proposed by ECHAM3 time-slice atmospheric model simulations. – J. Biogeogr. 34: 2115–2128.

Lid, J. and Lid, D. T. 1994. Norsk flora. – Det Norske Samlaget.

Lobo, J. M. et al. 2001. Spatial and environmental determinants of vascular plant species richness distribution in the Iberian Peninsula and Balearic Islands. – Biol. J. Linn. Soc. 73: 233–253.

MacArthur, R. H. and Wilson, E. O. 1967. The theory of island biogeography. – Princeton Univ. Press.

Mai, D. H. 1995. Tertiäre Vegetationsgeschichte Europas – Metoden und Ergebnisse. – Gustav Fischer.

McGlone, M. S. 1996. When history matters: scale, time, climate and tree diversity. – Global Ecol. Biogeogr. Lett. 5: 309–314.

Nekola, J. C. and White, P. S. 1999. The distance decay of similarity in biogeography and ecology. – J. Biogeogr. 26: 867–878.

New, M. et al. 2002. A high-resolution data set of surface climate ovre global land areas. – Clim. Res. 21: 1–25.

O'Brien, E. M. et al. 2000. Climatic gradients in woody plant (tree and shrub) diversity: water-energy dynamics, residual variation, and topography. – Oikos 89: 588–600.

Postigo Mijarra, J. M. et al. 2008. Survival and long-term maintenance of Tertiary trees in the Iberian Peninsula during the Pleistocene: first record of *Aesculus* L. (Hippocastanaceae) in Spain. – Veg. Hist. Archaeobot. 17: 351–364.

R Development Core Team 2008. R: a language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna.

Rangel, T. F. et al. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. – Global Ecol. Biogeogr. 15: 321–327.

Ricklefs, R. E. 2004. A comprehensive framework for global patterns in biodiversity. – Ecol. Lett. 7: 1–15.

Simpson, G. G. 1964. Species density of North American recent mammals. – Syst. Zool. 13: 57–73.

Stropp, J. et al. 2009. Disentangling regional and local tree diversity in the Amazon. – Ecography 32: 46–54.

Svenning, J.-C. 2003. Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. – Ecol. Lett. 6: 646–653.

Svenning, J.-C. and Skov, F. 2004. Limited filling of the potential range in European tree species. – Ecol. Lett. 7: 565–573.

Svenning, J.-C. and Skov, F. 2005. The relative roles of environment and history as controls of tree species composition and richness in Europe. – J. Biogeogr. 32: 1019–1033.

Svenning, J.-C. and Skov, F. 2007a. Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? – Ecol. Lett. 10: 453–460.

Svenning, J.-C. and Skov, F. 2007b. Ice age legacies in the geographical distribution of tree species richness in Europe. – Global Ecol. Biogeogr. 16: 234–245.

Svenning, J.-C. et al. 2008a. Glacial refugia of temperate trees in Europe: insights from species distribution modelling. – J. Ecol. 96: 1117–1127.

Svenning, J.-C. et al. 2008b. Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. – Ecography 31: 316–326.

Svenning, J.-C. et al. 2009. Plio-Pleistocene climate change and geographic heterogeneity in plant diversity–environment relationships. – Ecography 32: 13–21.

Thuiller, W. et al. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. – Ecography 32: 1–5.

Willis, K. J. and van Andel, T. H. 2004. Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. – Quat. Sci. Rev. 23: 2369–2387.

Zobel, M. 1997. The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence? – Trends Ecol. Evol. 12: 266–269.

Download the Supplementary material as file E6301 from <www.oikos.ekol.lu.se/appendix>.