APPENDIX S2: Estimation of Last Glacial Maximum species ranges by hind-casting

---

This appendix includes the following sections:

- *Species distribution modelling*
- *Spatial agreement among predicted LGM ranges*
- *Correlations among estimates of postglacial accessibility*
- *Restricting species LGM ranges*
- *Discussion of uncertainties*
- *Estimated LGM ranges and accessibility patterns for selected species*

Species distribution modelling

In the present study we predicted areas of suitable climatic conditions for each species during the Last Glacial Maximum (LGM) by calibrating species distribution models on the species' present distribution range in Europe and then projecting them onto LGM climate simulations. We used two different species distribution modelling algorithms: Maxent (Phillips *et al.* 2006) and Bioclim (Guisan & Zimmermann 2000).

Maxent modelling was performed with all background points available in the study area (n = 4878) and with the recommended default settings (Phillips & Dudík 2008). The logistic output estimates probability of presence and ranges from zero to one (Phillips & Dudík 2008), but was converted to a binary prediction of potential presence (suitable climate) or absence (unsuitable climate) using the maximum training sensitivity plus specificity threshold. This threshold generally produces accurate predictions (Jiménez-Valverde & Lobo 2007). The ability of the Maxent models to predict the present observed distribution of species was evaluated using the area under the receiver operating characteristics curve (AUC). The AUC values were computed by training the model on a random 70% subset of the occurrences of a species and testing the model on the remaining 30% (Phillips et al. 2006). AUC = 0.5 indicates a prediction no better than random, while AUC = 1 indicates perfect prediction. On average (± SD) across all study species (*n* = 1016) we obtained an AUC of 0.86 ± 0.08 (median: 0.88, range: 0.59–0.98). The Bioclim models were parameterized using the outlier-corrected minimum and maximum of the observed values for each climatic predictor variable across a species' present range (cf. Skov & Svenning 2004). When all climatic variables in the predictor set fell within the inner range defined by these values, a species' potential presence was predicted.

We used Maxent version 3.0.6 (http://www.cs.princeton.edu/~schapire/maxent/) and implemented the Bioclim modelling in ArcGIS 9.2 (ESRI, Redlands, California) using Python 2.4 (www.python.org).

Spatial agreement among predicted LGM ranges

We assessed the spatial agreement between the binary predictions of species ranges obtained with the two modelling methods (Maxent and Bioclim) and paleoclimatic simulations (Stage 3 and Lmdzhr) by computing the percentage of all AFE cells for a given species that had the same value (always suitable or always unsuitable) in the four implementations. There was high agreement among predictions based on the different models and LGM climate simulations (Table S2.1). This result is in accordance with the previously observed small differences among estimates of LGM distributions for European trees using the same modelling methods and LGM simulations (Svenning *et al.* 2008).

Table S2.1. Spatial agreement (proportion of the study area identically predicted) in estimates of LGM species ranges, comparing pairs of model algorithms and LGM climate simulations. The average (±SD) across all species (*n* = 1016) is reported.

	Spatial agreement
Among models	
Maxent-Current; Bioclim-Current	0.82±0.11
Maxent-Lmdzhr; Bioclim- Lmdzhr	0.87±0.10
Maxent-Stage3; Bioclim-Stage3	0.85±0.11
Among LGM climate	
Maxent-Lmdzhr; Maxent-Stage3	0.86±0.07
Bioclim-Lmdzhr; Bioclim-Stage3	0.86±0.06

Correlations among estimates of postglacial accessibility

The estimates of accessibility to postglacial recolonisation based on LGM ranges estimated with different modelling algorithms and LGM climate simulations were highly correlated (Table S2.2).

Table S2.2. Spearman's rank correlations among postglacial accessibility estimates based on LGM ranges estimates for all pairs of model algorithms and LGM climate simulations. The average (±SD) across all species (*n* = 1016) is reported.

		Maxent		**Bioclim**	
		Lmdzhr	*Stage3*	*Lmdzhr*	*Stage3*
Maxent	*Lmdzhr*	-	0.96±0.03	0.96±0.06	0.93±0.07
	Stage3		-	0.93±0.08	0.94±0.09
Bioclim	*Lmdzhr*			-	0.97±0.03

Restricting species LGM ranges

Species probably did not occupy all areas predicted to have been climatically suitable during the LGM. Optimally, the estimated LGM range for each species should be evaluated against paleoecological records or, alternatively, phylogeographic evidence. Paleoecological records, especially macrofossils, provide clear evidence of species' occurrences within a given time-frame in the past. However, this type of evidence is not available for all species or sites, for several reasons: (i) organic remains (i.e. pollen, leaves) for some species are hardly or not preserved at all; (ii) not all fossils can be distinguished to the species level; (iii) fossil preservation to the present requires certain environmental conditions. To overcome the potential problem of overestimating LGM ranges without having fossil or phylogeographic evidence, we restricted the estimated (unrestricted) LGM ranges for each species according to the most likely regional occurrence of its glacial refugia given by the species' current distribution. Paleoecological and phylogeographic evidence suggests three main glacial refugia for the European biota: the Balkans, Italy, and the Iberian Peninsula (e.g., Lang 1994; Taberlet et al. 1998). We hypothesized that species presently restricted to one of these areas also only occurred there during the LGM (Fig. S2.1a,b,c). For species presently occurring in Central and Northern Europe general expectations of their LGM distributions are difficult to make. We therefore applied some simple, but plausible rules, and restricted species currently occurring in (i) western Europe to central-western Europe, Italy and the Iberian Peninsula (Fig. S2.1d), (ii) central-northern Europe to all of Europe (Fig S2.1e), (iii) Eastern Europe to central-eastern Europe, Balkan and Italy (Fig S2.1f). Based on the restricted and unrestricted LGM ranges we computed species-specific estimates of accessibility to postglacial colonization from these ranges. Fig. S2.2 provides an example of the difference in accessibility computed from the unrestricted and restricted LGM ranges.

Figure S2.1. Six refugia regions (orange) and the areas to which species' LGM ranges were restricted (green and orange) when presently occurring in a given refugia region. LGM ice sheet from Ehlers & Gibbard (2004). Elevation (light to dark) from the SRTM data set as provided in the WORLDCLIM data set (Hijmans et al. 2005). The LGM coastline for Europe (light gray) was estimated by lowering the sea level of the Earth Topography-5' elevation-bathymetry raster (http://www.ngdc.noaa.gov/mgg/fliers/93mgg01.html) by 110 m (Ruddiman 2001).

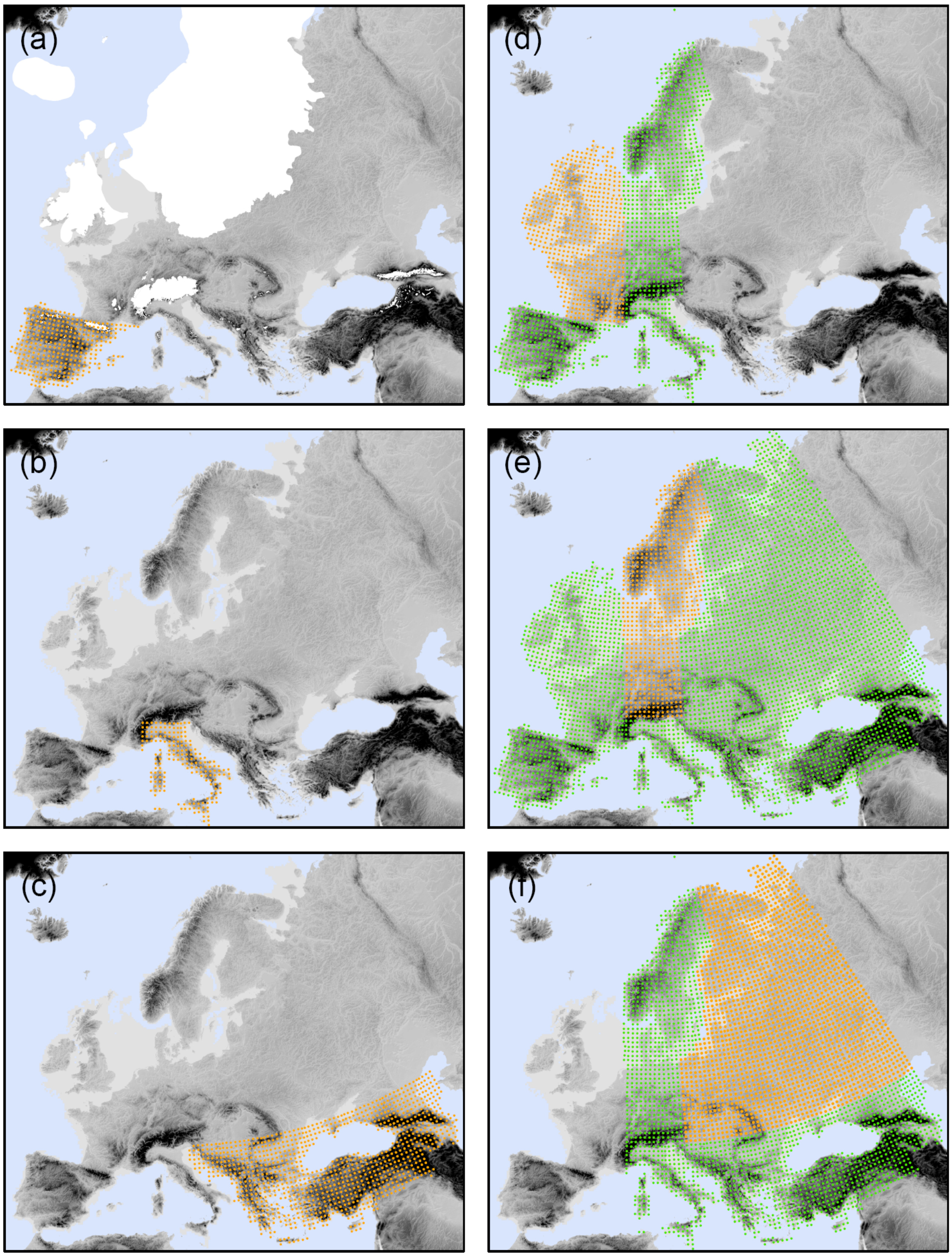
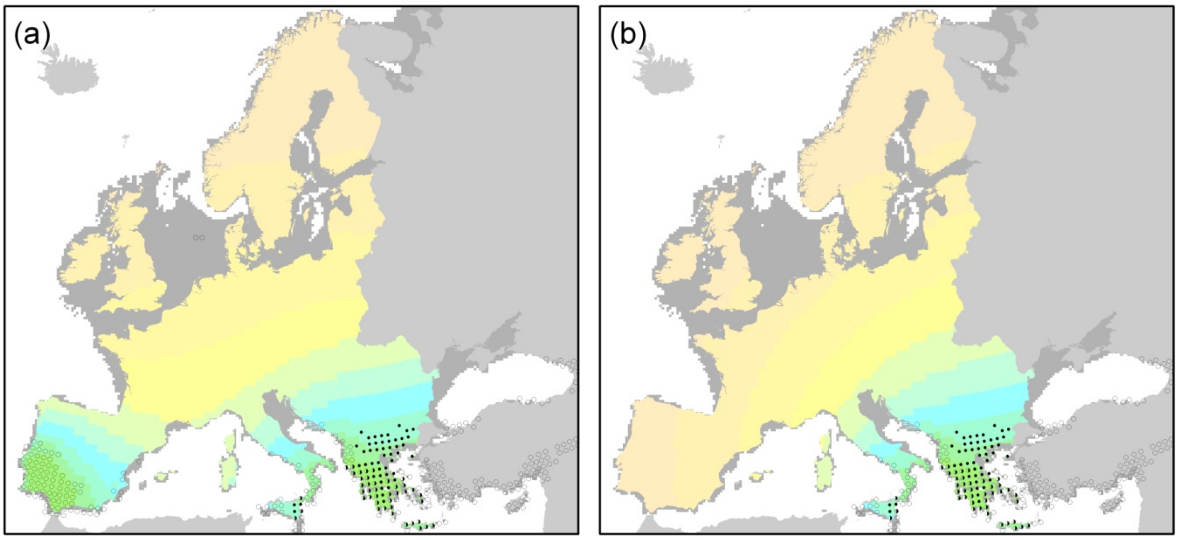


Figure S2.2. The unrestricted (a) and restricted (b) estimated LGM ranges of *Platanus orientalis* L. modelled with Maxent and the LMDZHR simulation and accessibility to postglacial recolonisation calculated from these ranges (green: high, orange: low). Variation in *Platanus orientalis* L. occurrences uniquely explained by climate ($R_L^2{}_{unique(C)}$) or accessibility ($R_L^2{}_{unique(A)}$) were for (a) $R_L^2{}_{unique(C)}$ 9%, $R_L^2{}_{unique(A)}$ 27%, and for (b) $R_L^2{}_{unique(C)}$ 6%, $R_L^2{}_{unique(A)}$ 48%.



Discussion of uncertainties

Several uncertainties related to species distribution modelling and past climate simulations should be kept in mind when interpreting our results (Ramstein *et al.* 2007, Svenning *et al.* 2008, Nogués-Bravo 2009). First, as also discussed in the main text, species distribution modelling relies on the assumption that a species climate niche can be estimated from the climate conditions where the species currently occurs (equilibrium assumption, Guisan & Zimmermann 2000). Some disequilibrium with climate in geographic space does not, however, preclude achieving reliable niche estimates. The quality of the niche estimates and importance of disequilibrium in geographic space depends on the uniqueness of the environmental conditions at the sites from which the species is absent. In order to obtain good estimates of species' niches and thus to minimize uncertainties in the estimates of the LGM ranges we took several steps: (i) We calibrated the models of species climatic tolerances using both native and naturalized occurrences as well as occurrences across all of Europe, including parts of the former Soviet Union, a region that harbours conditions that resample those prevailing in Europe at the LGM (Fløjgaard *et al.* 2009). Hence, occurrence records from this region increase our ability to capture species cold and drought tolerance, and thus improve our estimated LGM ranges. (ii) Different species distribution modelling methods produce different results (Elith *et al.* 2006). We used two modelling methods, maxent and bioclim, which generally provide narrow or broad estimates of species climatic niches, respectively (Elith *et al.* 2006). In addition, especially bioclim mainly depends on the extremes of the species climatic requirements being reached somewhere within a species distribution and thus is relatively robust to disequilibrium with climate in geographic space. Second, hindcasting species climatic niches assumes niche stability through time (see Nogués-Bravo 2009). Although species might evolve in response to climate change, niche stability is generally expected at the rather short time scale (in an evolutionary perspective) investigated in this study. Additionally, the response of European plants to Quaternary climate change was primarily extinction and migration rather than evolution (Lang 1994, Prinzing *et al.* 2001, Svenning 2003, Willis & Niklas 2004). Third, the accuracy of potential LGM ranges depends on how well the climate simulations estimate actual LGM climate conditions. The two simulations we used here are consistent with pollen-based temperature reconstructions, although the simulations represent climate in the warmer end of the confidence intervals of the pollen-based estimates (Ramstein *et al.* 2007). Too warm LGM climate conditions could potentially result in overestimated LGM ranges for warm-adapted species, and thus underestimate postglacial recolonisation distances and the importance of accessibility. Nevertheless, our results were consistent across LGM climate simulations and modelling methods used (see above and appendix S3). Additional sources of uncertainty that could not be integrated in the modelling are the influence of inter-annual climate variability on species distributions (Zimmermann *et al.* 2009, Giesecke *et al.* 2010), and lower CO_2

levels and increased wind speed at the LGM (Svenning *et al.* 2008). Potential changes in inter-annual variability in the past could have allowed the species to occupy more or less habitat than estimated with our models, while lower CO_2 and increased wind speed might have increased vegetation openness (Svenning *et al.* 2008).

Estimated LGM ranges and accessibility patterns for selected species

Paleoecological evidence, especially macrofossils, provides the only certain evidence of a species' occurrence at a given time and place in the past (see the section "Restricting species LGM ranges" above). To show that the modelled distributions used in the present study generally provide good representations of the species' likely LGM occurrences, we here provide maps of modelled LGM ranges for 17 species (restricted LGM ranges as explained above). In the Discussion section of the main text we outline the general agreement between available paleoecological (and phylogeographical) evidence and our modelled LGM ranges, and emphasize the estimated importance of accessibility computed from the estimated LGM ranges. The 17 species were obviously not chosen at random among all analysed species. They have been selected because paleoecological or phylogeographical evidence was available and because they represent examples of species from different biomes providing insight into different types of LGM ranges, range shifts, and accessibility patterns. The species are presented in alphabetical order in Fig. S2.1-Fig. S2.5.

As discussed in the main text, these examples show that we generally achieved reasonable predictions of species' LGM ranges. For example, the boreal tree species *Pinus sylvestris* and *Picea abies* potentially found suitable climate conditions in large areas from Western Europe eastwards to the Russian plains (Fig. S2.3). This wide LGM distribution is supported by several paleoecological studies (Cheddadi *et al*. 2006, Binney *et al.* 2009, Heikkilä *et al*. 2009, Kelly *et al.* 2010), although for some of these areas no fossil remains have been found (e.g., the areas just west of Denmark). Similarly, we might have overestimated the LGM ranges for other species (e.g., predict the species present in areas where paleoecological evidence is missing). Pollen of the genus *Ostrya* is only found in the Balkans (Lang 1994), although suitable climate conditions for *Ostrya carpinifolia* occurred on the Iberian Peninsula according to our modelling (Fig. S2.3). *O. carpinifolia* does not occur naturally on the Iberian Peninsula today. Hence, it is probable that it also did not occur there during the LGM and that we overestimated the LGM range of this species. However, it should be kept in mind that, even for species with a good paleoecological record, the absence of such data at a site does not prove the absence of a species since the lack of pollen or macrofossils can have many reasons (see the section "Restricting species LGM ranges" above). In other cases, we might underestimate areas of suitable climate compared to paleoecological evidence. *Fagus sylvatica* and *Dryas octopetala* are examples for which we might have underestimated the LGM ranges (e.g., models predict the species absent in areas where paleoecological or phylogeographic evidence indicates them as present). For *F. sylvatica*, our models failed to predict areas of suitable climatic conditions in the Carpathians although LGM survival has been suggested in this region (Magri *et al.* 2006) (Fig. S2.2). Fossil evidence shows that *Dryas octopetala* was abundant in central and northern Europe at the end of the last glaciation (Godwin, 1975), but our modelling estimated only scattered occurrences in these regions (Fig. S2.2).

Uncertainties in the potential LGM ranges obviously introduces uncertainty in the geographic pattern of accessibility, and thus in the estimated importance of postglacial dispersal limitation for species' current occurrences. Nevertheless, we found only small discrepancies in our results when comparing different modelling methods or LGM climate simulations (see above and appendix S3). Hence, despite the uncertainties in the LGM range estimates we are confident that our conclusions are robust.

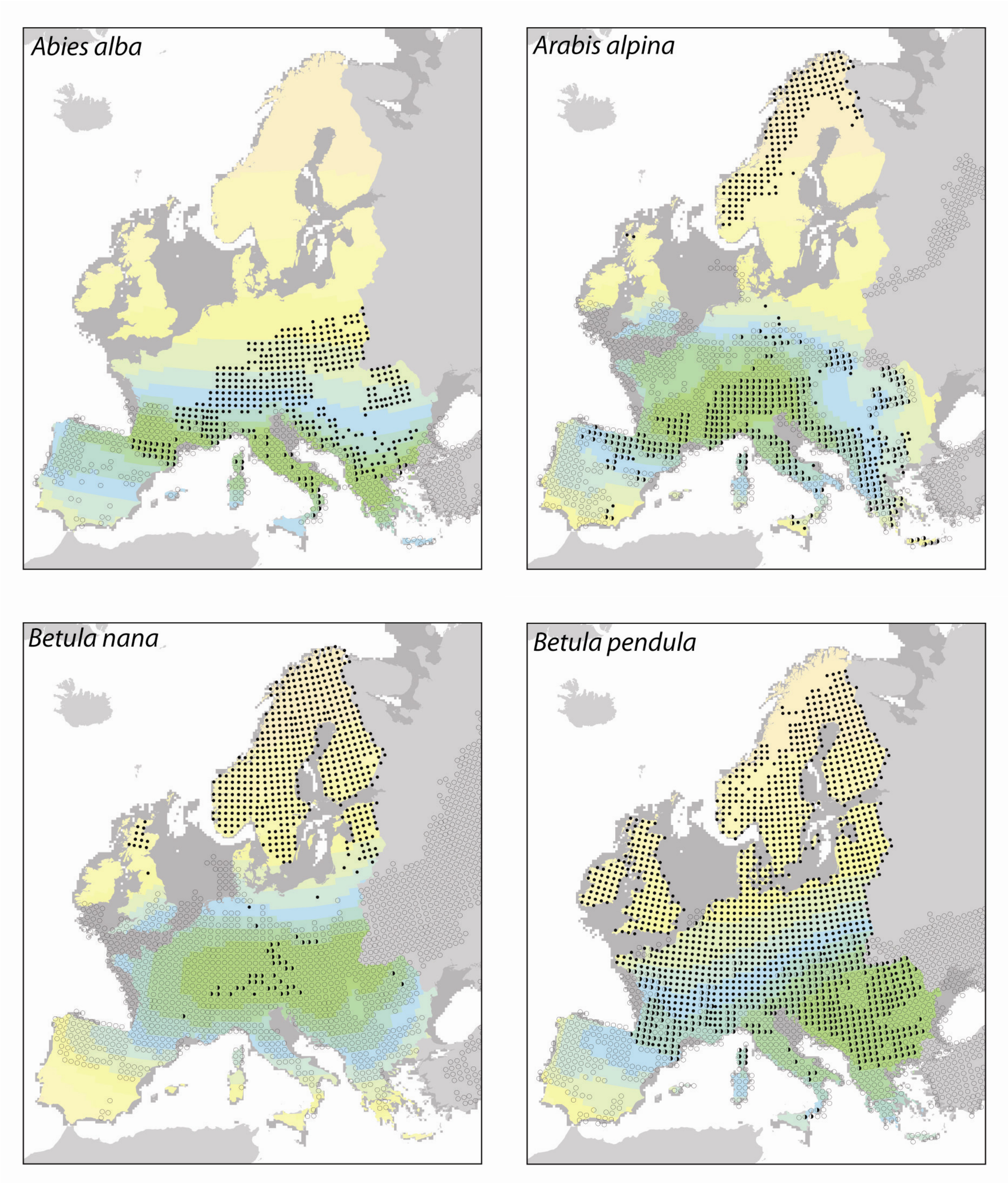


Figure S2.1. Distribution, postglacial accessibility and variation in species occurrences uniquely explained by accessibility ($R_{L}^{2}{}_{unique(A)}$) or climate $R_{L}^{2}{}_{unique(C)}$ for selected species: *Abies alba*; $R_{L}^{2}{}_{unique(C)}$ 26%, $R_{L}^{2}{}_{unique(A)}$ 7%; *Arabis alpina*; $R_{L}^{2}{}_{unique(C)}$ 15%, $R_{L}^{2}{}_{unique(A)}$ 1%; *Betula nana*; $R_{L}^{2}{}_{unique(C)}$ 73%, $R_{L}^{2}{}_{unique(A)}$ 0%, and *Betula pendula*; $R_{L}^{2}{}_{unique(C)}$ 51%, $R_{L}^{2}{}_{unique(A)}$ 0%. Current distribution (black dots), hindcasting-based estimate of LGM distribution (empty circles), overlap between the two (half-filled circles), and accessibility to postglacial colonization from the estimated LGM range (green to orange: high to low). Maps are in the ETRS 1989 Lambert Azimuthal Equal Area projection.

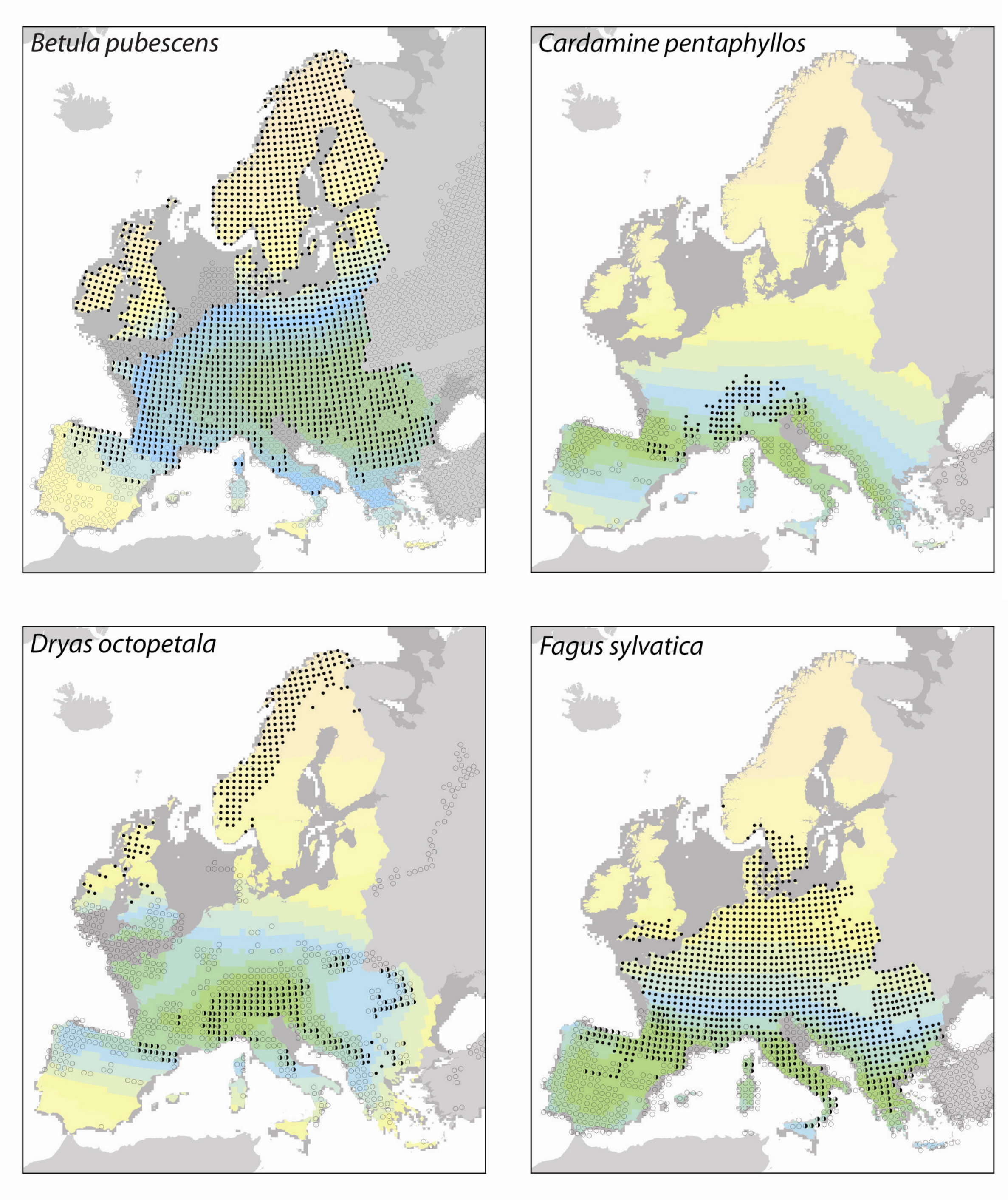


Figure S2.2. Distribution, postglacial accessibility and variation in species occurrences uniquely explained by accessibility ($R_L^2{}_{unique(A)}$) or climate $R_L^2{}_{unique(C)}$ for selected species: *Betula pubescens*; $R_L^2{}_{unique(C)}$ 56%, $R_L^2{}_{unique(A)}$ 0%; *Cardamine penthaphyllos*; $R_L^2{}_{unique(C)}$ 32%, $R_L^2{}_{unique(A)}$ 15%; *Dryas octopetala*; $R_L^2{}_{unique(C)}$ 46%, $R_L^2{}_{unique(A)}$ 0%, and *Fagus sylvatica*; $R_L^2{}_{unique(C)}$ 41%, $R_L^2{}_{unique(A)}$ 5%. Current distribution (black dots), hindcasting-based estimate of LGM distribution (empty circles), overlap between the two (half-filled circles), and accessibility to postglacial colonization from the estimated LGM range (green to orange: high to low). Maps are in the ETRS 1989 Lambert Azimuthal Equal Area projection.

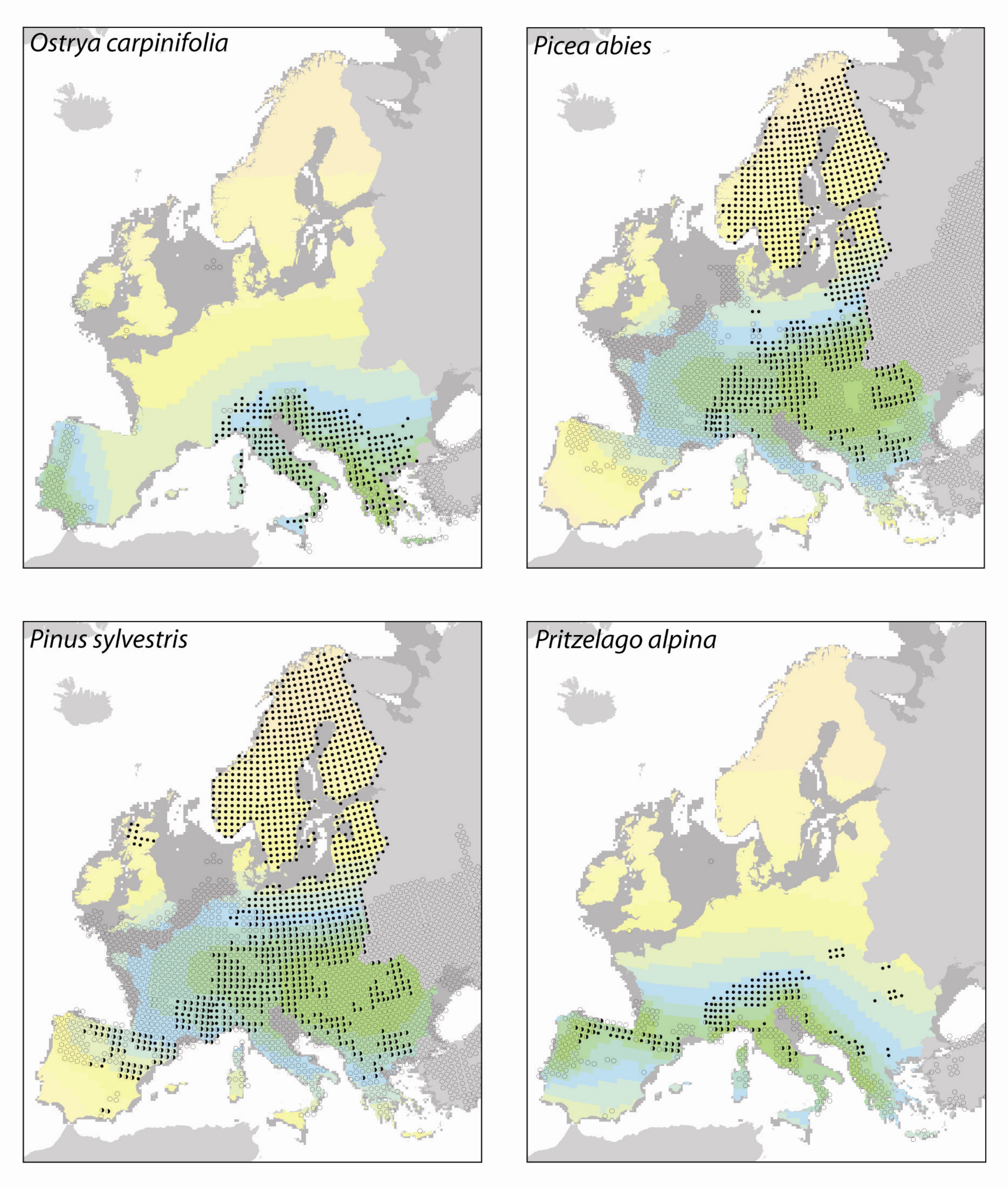


Figure S2.3. Distribution, postglacial accessibility and variation in species occurrences uniquely explained by accessibility ($R_L^{2}{}_{unique(A)}$) or climate $R_L^{2}{}_{unique(C)}$ for selected species: *Ostrya carpinifolia*; $R_L^{2}{}_{unique(C)}$ 12%, $R_L^{2}{}_{unique(A)}$ 21%; *Picea abies*; $R_L^{2}{}_{unique(C)}$ 63%, $R_L^{2}{}_{unique(A)}$ 0%; *Pinus sylvestris*; $R_L^{2}{}_{unique(C)}$ 48%, $R_L^{2}{}_{unique(A)}$ 0%, and *Pritzelago alpina*; $R_L^{2}{}_{unique(C)}$ 26%, $R_L^{2}{}_{unique(A)}$ 19%. Current distribution (black dots), hindcasting-based estimate of LGM distribution (empty circles), overlap between the two (half-filled circles), and accessibility to postglacial colonization from the estimated LGM range (green to orange: high to low). Maps are in the ETRS 1989 Lambert Azimuthal Equal Area projection.

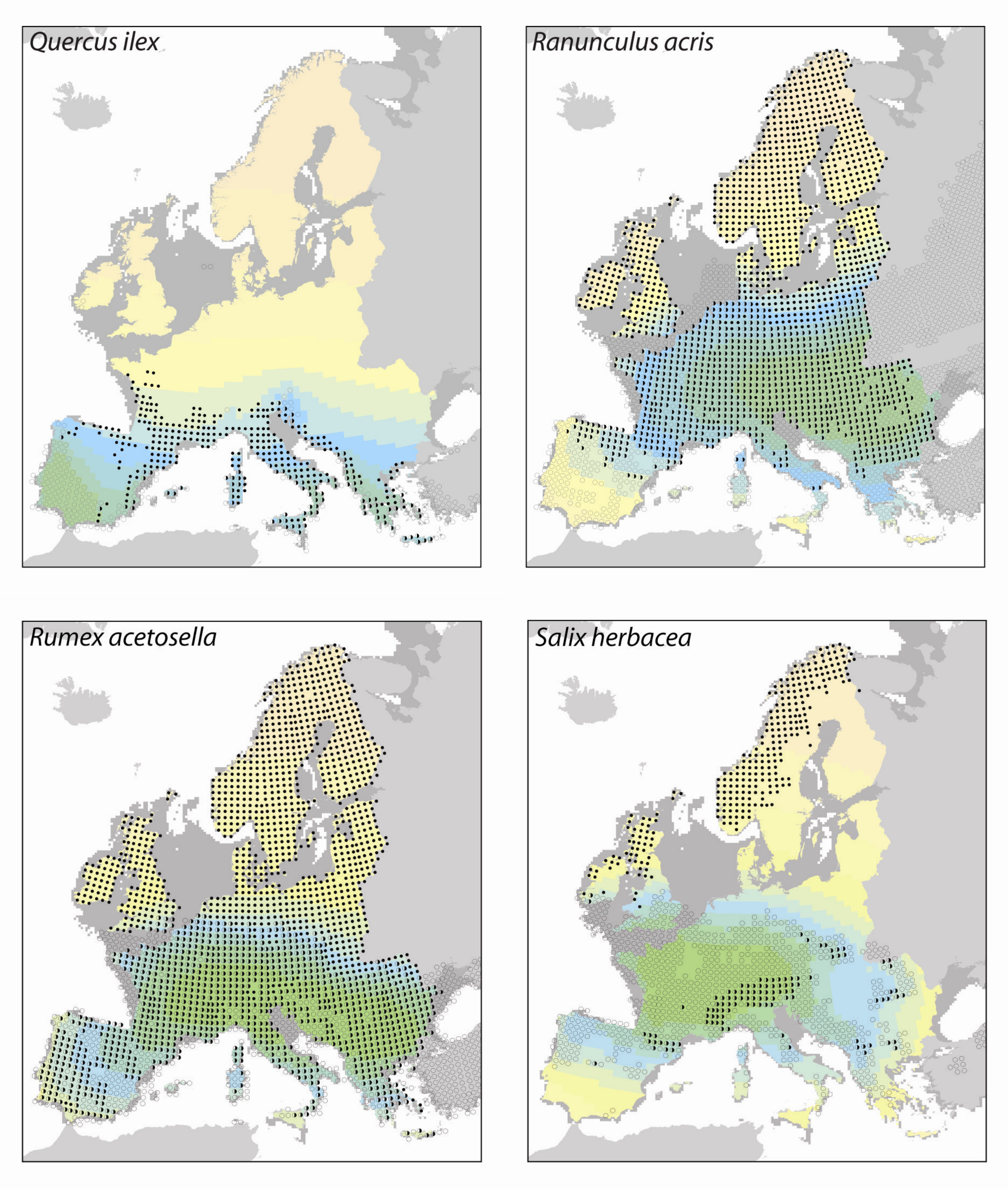


Figure S2.4. Distribution, postglacial accessibility and variation in species occurrences uniquely explained by accessibility ($R_L^{2}{}_{unique(A)}$) or climate $R_L^{2}{}_{unique(C)}$ for selected species: *Quercus ilex*; $R_L^{2}{}_{unique(C)}$ 38%, $R_L^{2}{}_{unique(A)}$ 0%., *Ranunculus acris*; $R_L^{2}{}_{unique(C)}$ 54%, $R_L^{2}{}_{unique(A)}$ 0%; *Rumex acetosella*; $R_L^{2}{}_{unique(C)}$ 30%, $R_L^{2}{}_{unique(A)}$ 0%; *Salix herbacea*; $R_L^{2}{}_{unique(C)}$ 57%, $R_L^{2}{}_{unique(A)}$ 3%. Current distribution (black dots), hindcasting-based estimate of LGM distribution (empty circles), overlap between the two (half-filled circles), and accessibility to postglacial colonization from the estimated LGM range (green to orange: high to low). Maps are in the ETRS 1989 Lambert Azimuthal Equal Area projection.

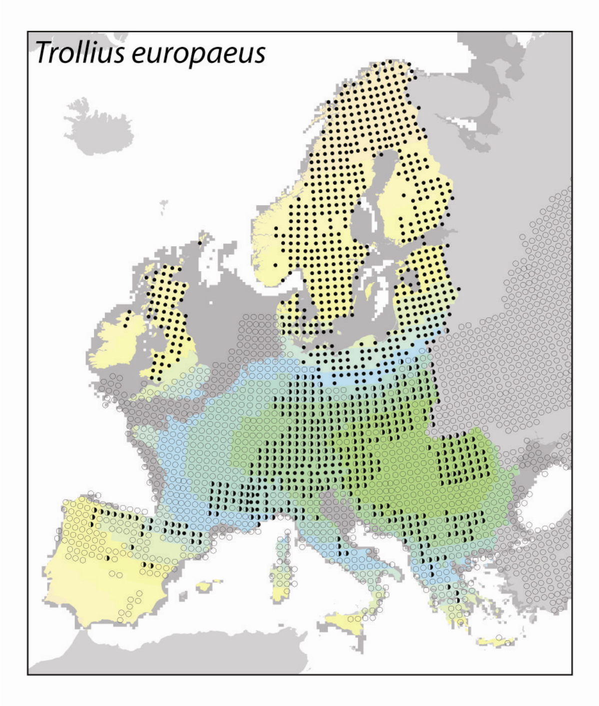


Figure S2.5. Distribution, postglacial accessibility and variation in species occurrences uniquely explained by accessibility ($R_L^2{}_{unique(A)}$) or climate $R_L^2{}_{unique(C)}$ for *Trollius europaeus*; $R_L^2{}_{unique(C)}$ 40%, $R_L^2{}_{unique(A)}$ 0%. Current distribution (black dots), hindcasting-based estimate of LGM distribution (empty circles), overlap between the two (half-filled circles), and accessibility to postglacial colonization from the estimated LGM range (green to orange: high to low). Maps are in the ETRS 1989 Lambert Azimuthal Equal Area projection.

References

Cheddadi, R., G. G. Vendramin, T. Litt, L. Francois, M. Kageyama, S. Lorentz, J. M. Laurent, J. L. de Beaulieu, L. Sadori, A. Jost, and D. Lunt. (2006). Imprints of glacial refugia in the modern genetic diversity of Pinus sylvestris. Glob. Ecol. and Biogeogr. 15: 271-282.

Ehlers J. & Gibbard P.L. (2004). Quaternary glaciations – extent and chronology. Part I: Europe. Elsevier, Amsterdam.
Godwin H. (1975). The history of the British flora: a factual basis for phytogeography., 2 edition. Cambridge University Press, Cambridge

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151.

Fløjgaard, C., S. Normand, F. Skov, and J. C. Svenning. 2009. Ice age distributions of European small mammals: Insights from species distribution modelling. Journal of Biogeography 36:1152-1163.

Giesecke, T., P. A. Miller, M. T. Sykes, A. E. K. Ojala, H. Seppä, and R. H. W. Bradshaw. 2010. The effect of past changes in inter-annual temperature variability on tree distribution limits. Journal of Biogeography 37:1394-1405.

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

Heikkilä, M., S. L. Fontana, and H. Seppä. (2009). Rapid Lateglacial tree population dynamics and ecosystem changes in the eastern Baltic region. SO: Journal of Quaternary Science 24:802-815.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G. & Jarvis A. (2005). Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25, 1965-1978

Jiménez-Valverde A. & Lobo J.M. (2007). Threshold criteria for conversion of probability of species presence to either-or presence-absence. Acta Oecol. 31, 361-369

Kelly, A., D. J. Charman, and R. M. Newnham. (2010). A last glacial maximum pollen record from bodmin moor showing a possible cryptic Northern refugium in Southwest England. J. Quat. Sci. 25: 296-308.

Lang G. (1994). Quartäre Vegetationsgeschichte Europas: Methoden und Ergebnisse. Gustav Fischer Verlag, Jena.

Magri, D., G. G. Vendramin, B. Comps, I. Dupanloup, T. Geburek, D. Gömöry, M. Latalowa, T. Litt, L. Paule, J. M. Roure, I. Tantau, W. O. Van Der Knaap, R. J. Petit, and J. L. De Beaulieu. (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. New Phytol. 171:199-221.

Nogués-Bravo, D. 2009. Predicting the past distribution of species climatic niches. Global Ecology and Biogeography 18:521-531.

Phillips S.J., Anderson R.P. & Schapire R.E. (2006). Maximum entropy modelling of species geographic distributions. Ecol.Modell. 190, 231-259

Phillips S.J. & Dudík M. (2008). Modelling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31, 161-175

Prinzing, A., W. Durka, S. Klotz, and R. Brandl. 2001. The niche of higher plants: evidence for phylogenetic conservatism. Proceedings of the Royal Society B: Biological Sciences 268:2383-2389.

Ramstein, G., M. Kageyama, J. Guiot, H. Wu, C. HÃ©ly, G. Krinner, and S. Brewer. 2007. How cold was Europe at the Last Glacial Maximum? A synthesis of the progress achieved since the first PMIP model-data comparison. Climate of the Past 3:331-339.

Ruddiman W.F. (2001). Earth's Climate: Past and Future. W.H. Freeman & Co., New York.

Skov F. & Svenning J.C. (2004). Potential impact of climatic change on the distribution of forest herbs in Europe. Ecography 27, 366-380

Svenning, J. C. 2003. Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. Ecology Letters 6:646-653.
Svenning, J. C., S. Normand, and M. Kageyama. 2008. Glacial refugia of temperate trees in Europe: Insights from species distribution modelling. Journal of Ecology 96:1117-1127.

Taberlet P., Fumagalli L., Wust-Saucy A.G. & Cosson J.F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. Mol.Ecol. 7, 453-464

Willis, K. J., and K. J. Niklas. 2004. The role of Quaternary environmental change in plant macroevolution: The exception or the rule? Philosophical Transactions of the Royal Society B: Biological Sciences 359:159-172.

Zimmermann, N. E., N. G. Yoccoz, T. C. Edwards, E. S. Meier, W. Thuiller, A. Guisan, D. R. Schmatz, and P. B. Pearman. 2009. Climatic extremes improve predictions of spatial patterns of tree species. Proceedings of the National Academy of Sciences 106:19723-19728.