Conservation Biology

Contributed Paper

Combining geodiversity with climate and topography to account for threatened species richness

Helena Tukiainen,* ¶§ Joseph J. Bailey,† § Richard Field,† Katja Kangas,‡ and Jan Hjort*

*Geography Research Unit, University of Oulu, P.O. Box 8000, Oulu, FI 90014, Finland
†School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, U.K.
‡Natural Resources Institute Finland (Luke), Economics and Society, University of Oulu, P.O. Box 413, Oulu, FI 90014, Finland

Abstract: *Understanding threatened species diversity is important for long-term conservation planning. Geodiversity—the diversity of Earth surface materials, forms, and processes—may be a useful biodiversity surrogate for conservation and have conservation value itself. Geodiversity and species richness relationships have been demonstrated; establishing whether geodiversity relates to threatened species' diversity and distribution pattern is a logical next step for conservation. We used 4 geodiversity variables (rock-type and soil-type richness, geomorphological diversity, and hydrological feature diversity) and 4 climatic and topographic variables to model threatened species diversity across 31 of Finland's national parks. We also analyzed rarity-weighted richness (a measure of site complementarity) of threatened vascular plants, fungi, bryophytes, and all species combined. Our 1-km*^{2} *resolution data set included 271 threatened species from 16 major taxa. We modeled threatened species richness (raw and rarity weighted) with boosted regression trees. Climatic variables, especially the annual temperature sum above 5 °C, dominated our models, which is consistent with the critical role of temperature in this boreal environment. Geodiversity added significant explanatory power. High geodiversity values were consistently associated with high threatened species richness across taxa. The combined effect of geodiversity variables was even more pronounced in the rarity-weighted richness analyses (except for fungi) than in those for species richness. Geodiversity measures correlated most strongly with species richness (raw and rarity weighted) of threatened vascular plants and bryophytes and were weakest for molluscs, lichens, and mammals. Although simple measures of topography improve biodiversity modeling, our results suggest that geodiversity data relating to geology, landforms, and hydrology are also worth including. This reinforces recent arguments that conserving nature's stage is an important principle in conservation.*

Keywords: biodiversity, conserving nature's stage, geology, geomorphology, heterogeneity, hydrology

Combinación de la Geodiversidad con el Clima y la Topografía para Representar la Riqueza de Especies Amenazadas

Resumen: *Entender la diversidad de especies amenazadas es importante para la planeación de la conservación a largo plazo. La geodiversidad – la diversidad de materiales, formas y procesos en la superficie terrestre – puede ser un sustituto útil de la biodiversidad para la conservación y puede tener un valor de conservación propio. Las relaciones entre la geodiversidad y la riqueza de especies han sido demostradas; el siguiente paso lógico para la conservación es establecer si la geodiversidad se relaciona con la diversidad de especies amenazadas y los patrones de distribución. Usamos cuatro variables de la geodiversidad (riqueza de tipo de roca y de tipo de suelo, diversidad geomorfológica, características de la diversidad hidrológica) y cuatro variables climáticas y topográficas para modelar la diversidad de especies amenazadas en 31 de los parques nacionales de Finlandia. También analizamos la riqueza ponderada con la rareza (una medida de la complementariedad de sitio) de las plantas vasculares, hongos y briofitas amenazadas y todas las especies combinadas. Nuestro conjunto de datos de resolución de 1-km*^{2} *incluía 217 especies amenazadas de*

¶*email helena.tukiainen@oulu.fi.*
§*These authors contributed equally to this article.*
Paper submitted March 1, 2016; revised manuscript accepted July 25, 2016.

Conservation Biology, Volume 31, No. 2, 364–375

DOI: 10.1111/cobi.12799

16 taxones mayores. Modelamos la riqueza de especies amenazadas (cruda y ponderada con la rareza) con árboles de regresión estimulados. Las variables climáticas, especialmente la suma de la temperatura anual sobre los 5 °C, dominaron nuestros modelos, lo que es consistente con el papel crítico de la temperatura en este ambiente boreal. La geodiversidad añadió un poder explicativo. Los altos valores de geodiversidad estuvieron asociados constantemente con la alta riqueza de especies amenazadas en los taxones. El efecto combinado de las variables de la geodiversidad estuvo más pronunciado en los análisis de riqueza sopesados con la rareza (excepto por los hongos) que en aquellos para la riqueza de especies. Las medidas de geodiversidad se correlacionaron más fuertemente con la riqueza de especies (sopesada con la rareza y la crudeza) de las plantas vasculares y las briofitas y fueron más débiles para los moluscos, los líquenes y los mamíferos. Aunque las medidas simples de la topografía mejoran el modelado de la biodiversidad, nuestros resultados sugieren que los datos de geodiversidad relacionados con la geología, las formaciones terrestres y la hidrología también deben ser incluidos. Esto refuerza los argumentos recientes que dicen que conservar el estado de la naturaleza es un principio importante en la conservación.

Palabras Clave: biodiversidad, conservación del estado de la naturaleza, geología, geomorfología, heterogeneidad, hidrología

Introduction

Land-use and climate change threaten species globally (Mantyka-Pringle et al. 2015). It is therefore increasingly important to understand and conserve species' diversity and distributions. One coarse-filter strategy in conservation and protected-area management, conserving nature's stage, centers on the physical structures that underlie biotic processes and recognizes that geodiversity–the diversity of Earth surface forms, materials, and processes (Gray 2013)–itself has conservation value and is related to biodiversity (Anderson & Ferree 2010; Lawler et al. 2015). Thus, proponents suggest that geodiversity be incorporated into biodiversity research and conservation (Lawler et al. 2015). In practical terms, geodiversity data are frequently easier and less expensive to obtain than biodiversity data (Hjort et al. 2012) and advance scientific understanding of the spatial distribution of biodiversity for long-term conservation planning.

Well-mapped abiotic data are commonly used, together with data from ecological communities, to model or predict biodiversity for conservation planning (Albuquerque & Beier 2015*a*). Numerous researchers have explored how abiotic factors are related to species' diversity and distributions (Lawler et al. 2015), and the relationship between environmental heterogeneity and species richness has been established across multiple taxa and spatial scales (Stein et al. 2014). Several researchers have shown that local, idiosyncratic features, such as rock type, and landscape-scale factors, such as energy-related climatological variables, correlate with threatened species' patterns at different spatial scales and differ among taxonomic groups (Berg et al. 2002; Kreft & Jetz 2007; Lawler et al. 2015). Vascular plants are commonly studied and typically have strong relationships with climatic variables and topographic heterogeneity (Field et al. 2009), whereas the effect of abiotic conditions is less known for rare species in other taxonomic groups (e.g., Virkkala et al. 2005; Anderson & Ferree 2010).

Geodiversity represents a more complete characterization of Earth-surface heterogeneity than topographic heterogeneity does. We define *geodiversity* as distinct from widely used topographic measures such as elevation, range in elevation, and slope (herein referred to as *topography*). Geodiversity includes explicit geofeatures such as rock types, soil types, geomorphological landforms, and hydrological features. These can be explicitly incorporated into analyses to capture the aspects of local heterogeneity, such as microclimatic effects and microsite patterns, which are ecologically important but are not captured by climatic and topographic data (Field et al. 2009; Dobrowski 2011). Such local heterogeneity, caused by geology, landforms, and hydrology, relates to extended local-resource gradients, niche space, and habitat variety (Stein et al. 2014).

Conservation professionals have rarely incorporated geodiversity into conservation prioritization efforts (Beier et al. 2015). However, relationships between geodiversity and biodiversity are being demonstrated increasingly (Lawler et al. 2015) and may have considerable implications for conservation. A significant link between plant species richness (dominated by common species) and geodiversity has been identified (Nichols et al. 1998; Hjort et al. 2012). Establishing whether such a relationship exists between geodiversity and threatened species' diversity and distribution patterns is a logical next step. Indeed, quantifying geodiversity may provide greater insight into a landscape's potential to preserve species diversity (Anderson & Ferree 2010; Lawler et al. 2015).

However, numbers of species, and even numbers of threatened species, may not provide optimal measures by which to prioritize sites for conservation (Kirkpatrick 1983; Albuquerque & Beier 2015*a*). For optimal planning, managers often want to identify groups of sites that collectively represent multiple conservation targets (typically threatened species) in small areas (Albuquerque & Beier 2015*a*). Thus, instead of selecting the sites with the greatest species richness or the most threatened species, a set of sites with species assemblages that complement each other and collectively capture the largest number of species is chosen (Albuquerque & Beier 2015*b*). Various site-prioritization methods have been developed; software such as Zonation (Moilanen et al. 2014) and Marxan (Ardron et al. 2010) are quite commonly used. However, a simpler alternative, rarity-weighted richness (RWR), is efficient and reliable and allows the identification of priority sites (Albuquerque & Beier 2015*b*). Albuquerque and Beier (2015*a*, 2015*c*) recently demonstrated that site complementarity can be highly predictable from abiotic characteristics.

We modeled the relationship between the physical environment and both threatened species richness and RWR at 1-km^2 grain size across 31 protected areas of Finland. We used 2 predictor categories: climate and topography (i.e., conventional predictors) and geodiversity. Our main aim was to determine the explanatory power of the geodiversity measures, which in this study are georichness variables. We expected data on geomorphology, geology, and hydrology to be relevant for landscape-scale patterns of threatened species richness and RWR because they should represent local geophysical conditions that are important for the establishment and persistence of threatened species (Rich & Weiss 1991; Engler et al. 2004). In theory, more threatened species should be able to persist where there is greater geological variety because of the broader variety of nutrients, resources, and pH (as previously observed for common species [Hjort et al. 2012]). Our study area included only protected areas, so the levels of human impact were consistently low.

We analyzed data for threatened species of vascular plants, fungi, lichens, beetles (Coleoptera), bryophytes, butterflies and moths (Lepidoptera), molluscs, mammals, and all of these combined. We used threatened species richness and RWR as measures of threatened species diversity and rarity, respectively. We tested the explanatory power of environmental variables in modeling diversity and RWR; assessed the consistency of these relationships across different taxonomic groups; and studied which georichness measures (rock-type richness, soil-type richness, geomorphological diversity, and hydrological feature diversity) added explanatory power to our models. We addressed the following hypotheses, which are not mutually exclusive. Threatened species diversity is strongly related to climate, and in high-latitude environments especially to energy-related climate variables (H_1) (Hawkins et al. 2003; Stein et al. 2014). Different taxonomic groups show different responses to climatic, topographic, and geodiversity predictors (H_2) (Stein et al. 2014). Geodiversity measures improve models for threatened species diversity and RWR (H_3) (Burnett et al. 1998; Anderson & Ferree 2010). Threatened species diversity and RWR can be successfully modeled with climate, topography, and geodiversity variables (H_4) (Pausas et al. 2003; Hjort et al. 2012; Albuquerque & Beier 2015*a*).

Methods

Study Area

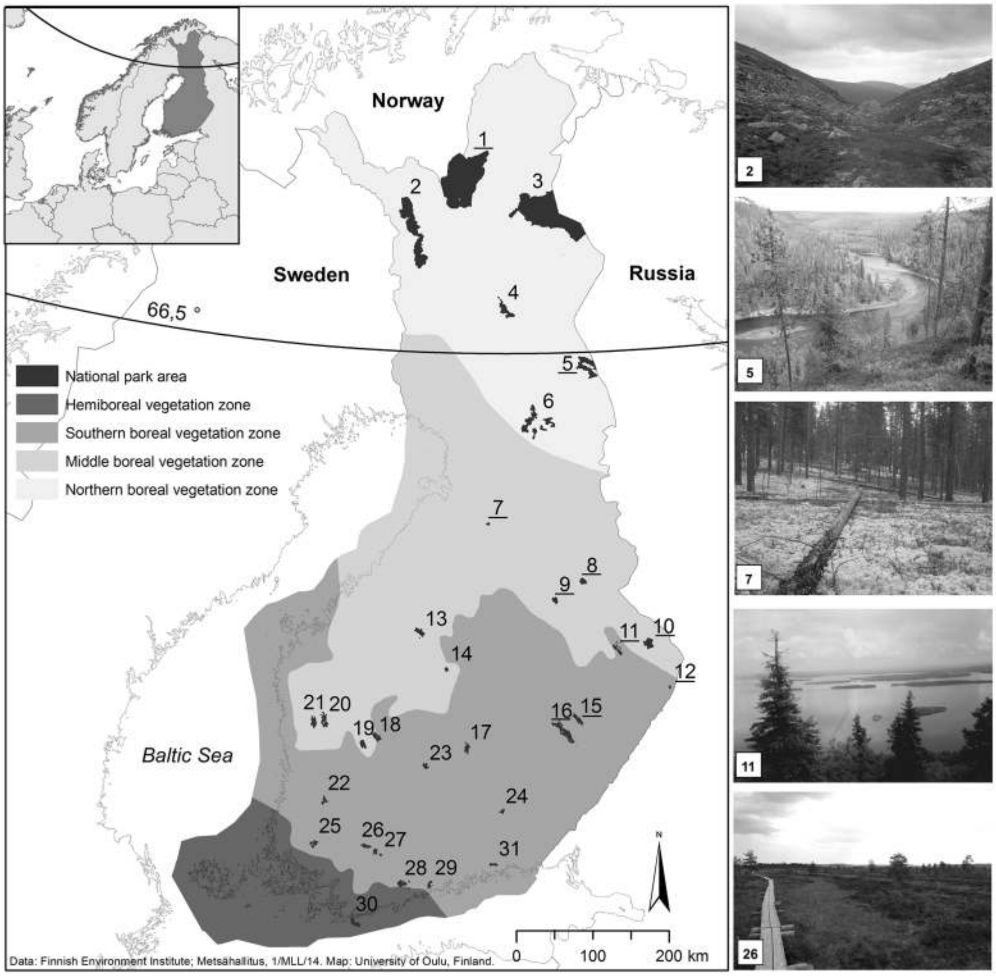
Our study area covered 31 national parks (Fig. 1), extending from southern Finland's coastal archipelago to northern Finland's glacially rounded hills with arctic-alpine conditions. Park area ranged from 6 km^2 (Petkeljärvi) to 2850 km^2 (Lemmenjoki); the total area was 8091 km^2. Finnish national parks follow the definitions and management objectives of the International Union for the Conservation of Nature (IUCN) and natural resources protected area management category II (Heinonen 2013). We superimposed a regular system of 1-km^2 grid cells and retained all cells containing threatened species records that had at least 10% of their area in a national park. For cells without threatened species records, we retained all cells located entirely within national park boundaries. We therefore selected 6571 grid cells, 583 with observations of threatened species and 5988 without.

Biogeographically, the study area covered hemi-, southern-, middle-, and northern-boreal vegetation zones (Ahti et al. 1968) and included a great variety of land-cover types, such as forests, fell areas, and wetlands. Mean annual air temperature varied from −2 °C in the north to approximately 6 °C in the south (Pirinen et al. 2012), and the length of the thermal growing season (>5 °C daily mean temperatures) was from >185 days in the south to <125 days in the north. Mean annual precipitation was moderate through all seasons and ranged from 444 to 739 mm (Table 1) (Pirinen et al. 2012). Finland is part of the Precambrian bedrock block of northern and eastern Europe and consists mainly of crystalline rocks (Atlas of Finland 1990). The soils of Finland originate mainly from during or after the last glacial period and are dominated by ground moraine and peat depositions.

Threatened Species Data

We considered threatened species from the following taxonomic groups: vascular plants, fungi, lichens, bryophytes, beetles, butterflies and moths, molluscs, mammals, 2-winged flies, true bugs, birds, hymenopterans, caddisflies, stoneflies, amphibians, and spiders. The richness of each of the first 8 of these (up to and including mammals) was modeled, as was that of all 16 taxa combined ('all' category). Threatened species were those considered critically endangered, endangered, vulnerable, or near-threatened in Finland according to the IUCN Red List (Rassi et al. 2001). We included a few data-deficient species known to be rare (for details, see Supporting Information).

Geographic coordinates of the records of threatened species were derived from the Hertta database (Finnish Environment Institute 2015). To ensure their spatial accuracy, we used occurrences recorded after the year 2000.



National park name (area km²)

1 = Lemmenjoki (2850)
2 = Pallas-Yllästunturi (1020)
3 = Urho Kekkonen National Park (2550)
4 = Pyhä-Luosto (142)
5 = Oulanka (270)
6 = Syöte (299)
7 = Rokua (9)
8 = Hiidenportti (45)
9 = Tillikkajärvi (34)
10 = Patvinsuo (105)
11 = Koli (30)
12 = Petkeljärvi (6)
13 = Salamajärvi (62)
14 = Pyhä-Häkki (13)
15 = Kolovesi (47)
16 = Linnansaari (96)
17 = Leivonmäki (30)
18 = Helvetinjärvi (50)
19 = Seitseminen (46)
20 = Kauhaneva - Pohjankangas (57)
21 = Lauhanvuori (53)
22 = Puurijärvi-Isosuo (27)
23 = Isojärvi (22)
24 = Repovesi (15)
25 = Kurjenrahka (29)
26 = Torronsuo (30)
27 = Liesjärvi (21)
28 = Nuuksio (45)
29 = Sipoonkorpi (19)
30 = Tammisaari (52)
31 = Valkmusa (17)

Figure 1. Locations and major vegetation zones of the Finnish national parks included in our study of their threatened species diversity and rarity-weighted richness. The parks are split into eastern (underlined) and western parks.

Nearly all (99.3%) the coordinates were recorded with GPS with 100-m accuracy. There was no bias in relation to the number of species occurrences and proximity to recreational routes (this was examined for vascular plants, bryophytes, lichens, and fungi by Siikamäki et al. [2015]).

As well as analyzing raw threatened species richness, we modeled RWR (Williams et al. 1996; Albuquerque & Beier 2015*b*). The rarity value of each species is the inverse of the number of grid cells in which it occurs. The RWR value per grid cell is the sum of the rarity values from each species recorded. Grid cells containing rarer species therefore have higher RWR. We calculated RWR for each taxonomic group with sufficient data (vascular plants, bryophytes, and fungi) and all species combined. For the RWR modeling, we used the same environmental variables as for threatened species richness analyses. We also used these variables for additional analyses of the distribution of each taxon (details given in Supporting Information).

Environmental Variables

We compiled 24 environmental (abiotic) variables for the 6571 1-km^2 cells as potential predictors (Table 1). As well as geodiversity measures (number of types of rock, soil, landform, and hydrological features), these abiotic environmental variables included widely used climatic and topographic variables so we would cover the most likely abiotic correlates of species diversity at the landscape scale (Field et al. 2009; Stein et al. 2014).

We derived topographic variables from a 25-m-resolution digital elevation model (DEM) (NLS 2000). Elevation and slope angle (mean, SD, and range) were calculated per grid cell with ArcMap version 10.2 (ESRI, USA). Topography-derived moisture conditions of the study area were calculated using the topographic wetness index (TWI) (Beven & Kirkby 1979).

Detailed, expert-derived data on all the landforms in each grid cell were available only for 2083 cells, so geomorphological richness for all cells was determined by

Table 1. Details of considered environmental variables for use in boosted regression-tree analyses of spatial richness and rarity-weighted richness patterns for several taxa across Finland's national parks ($n = 6571$).

Environmental variable	*Unit*	*Median (min to max)*	*Source*[a]	*Abbreviation*
Geodiversity				
rock-type richness[b]	number of rock types	1 (1–6)	GSF	rock rich
soil-type richness[b]	number of soil types	2 (1–5)	GSF	soil rich
geomorphological richness[b]	number of geomorphological feature types	6 (0–13)	GAM	GM rich
hydrological feature richness[b]	number of hydrological feature types	1 (0–6)	NLS	hydro rich
Climate[c]				
mean annual air temperature	°C	-1.1 (-2 to 5.9)	FMI[d]	
growing degree days (>5 °C)[b]	degree-days	639 (500.3–1448.3)	FMI	GDD
mean temperature of coldest month (January)	°C	-13.4 (-14.5 to -3)	FMI	
mean temperature of warmest month (July)	°C	13.1 (12.5–17.7)	FMI	
seasonality (mean temperature of July–January)	°C	26.4 (20.1–29.3)	FMI	
mean annual precipitation[b]	mm	540.7 (443.6–739.3)	FMI	MP
potential evapotranspiration[d]	mm year^{-1}	210.1 (194.6–392.6)	FMI[d]	
water balance[d]	mm year^{-1}	327.8 (221–435.9)	FMI[d]	
theoretical solar radiation (mean)[e]	Mj cm^{-2} year^{-1}	0.5 (0.3–0.6)	DEM[e]	
theoretical solar radiation (SD)[e]	Mj cm^{-2} year^{-1}	0.02 (<0.01–0.2)	DEM[e]	
theoretical solar radiation (range)[e]	Mj cm^{-2} year^{-1}	0.2 (<0.01–0.7)	DEM[e]	
Topography				
elevation (mean)	m asl	308.3 (10.2–738.9)	DEM	
elevation (SD)	m	11.1 (0–127.5)	DEM	
elevation (range)[b]	m	48 (0–414)	DEM	ER
slope angle (mean)	degrees	2.9 (<0.01–22.7)	DEM	
slope angle (SD)	degrees	2 (0.02–12.9)	DEM	
slope angle (range)	degrees	12 (0.6–56.7)	DEM	
topographical wetness index (mean)	-	10.8 (6.4–24.3)	DEM	
topographical wetness index (SD)	-	4.6 (0.4–7.7)	DEM	
topographical wetness index (range)[b]	-	23.2 (9.5–32.8)	DEM	TWIR

[a] *Abbreviations: DEM, digital elevation model; GAM, generalized additive model; GSF, Geological Survey of Finland; FMI, Finnish Meteorological Institute; NLS, National Land Survey of Finland.*

[b] *Variables selected for species richness and rarity-weighted richness modeling, based on correlation analysis (see Methods for further details).*

[c] *Climate variables are for 1981–2010.*

[d] *Method following Skov and Svenning (2004).*

[e] *Estimate of potential annual direct incident radiation calculated using ArcGIS 10.2 (McCune & Keon 2002).*

generalized additive modeling (GAM) with the available landform data and the 25-m DEM, as follows (Hjort & Luoto 2012). Landform data were derived from 1:50,000 geomorphological maps and aerial photographs (approximately 30-cm resolution). These were modeled using DEM-based and geographical variables (calibration with 1458 cells and evaluation with 625). The GAM with the final explanatory variables was recalibrated using all 2083 cells and applied to the 6571 grid cells in the study (details given in Supporting Information).

Otherwise, geodiversity per grid cell was calculated following Hjort and Luoto (2012), as follows. Hydrological feature richness was the sum of different hydrological feature types. Features were mapped from the National Land Survey of Finland's database (Table 2) (NLS 2007). Soil- and rock-type richness were measured by summing the number of different soil and rock types, respectively. Soil and rock types were derived from digital soil and bedrock maps, respectively, both produced by the Geological Survey of Finland (Table 2) (GSF 2010*a*,

Table 2. Details of the features or classes of geology, soil, and hydrology on which geodiversity (georichness) variables were calculated for analyzing spatial richness and rarity-weighted richness patterns across Finland's national parks.

Geodiversity variable	*Features or classes*
Rock-type richness	ultramafic intrusive or volcanic rocks
	mafic intrusive or volcanic rocks
	intermediate, intrusive volcanic rocks
	granitic or fesic rocks
	pelitic sedimentary rocks
	conglomerates
	arkosic sedimentary rocks
	black schists
	quartz-rich sedimentary rocks
	sedimentary carbonate rocks or carbonatites
	gneisses and migmatites
	iron ore
	high-grade metamorphic rocks
	metasomatic rocks
	impact melt rocks
	sulphide ore
Soil-type richness	rock (bare rock or thin soil cover; < 1 m)
	till (glacigenic deposits)
	stone and block fields
	sand and gravel
	silt
	clay
	gyttja (lake and sea sediments; > 6% organic material)
	peat
Hydrological feature richness	lakes (>1 ha)
	ponds (<1 ha)
	large rivers (>5 m wide)
	small rivers (2–5 m wide)
	streams (<2 m wide), springs

2010*b*). Climate data for 1981–2010, at 1-km^2 resolution, were derived from the Finnish Meteorological Institute (Pirinen et al. 2012; Table 1).

We removed highly correlated (Spearman's rank correlation coefficient, $|r_S| > 0.7$) climatic and topographic variables after preliminary analysis to avoid multicollinearity. Selection of the final variables was based on their mutual correlations and conceptual relevance and designed to obtain the same number of georichness variables as nongeorichness variables (climate and topography). Therefore, 2 climate (representing energy and moisture availability [Hawkins et al. 2003]) and 2 topographic variables were selected to match the 4 georichness variables (Table 1). Based on the previous steps, we used rock-type richness, soil-type richness, geomorphological diversity, hydrological feature diversity, growing-degree days, mean annual precipitation, elevational range, and range of the TWI.

Analyses

We used boosted regression trees (BRTs) to analyze the patterns of threatened species richness and RWR. Boosted regression trees are an ensemble modeling method in which regression trees are applied (from the classification and regression-tree group of models) and then boosted to combine a collection of models (Elith et al. 2008). The BRT models were fitted in R version 3.1.3 (R Core Development Team 2008) with the gbm package (version 2.1.1) (Ridgeway 2015) and the function gbm.step, which uses regularization methods to discourage overfitting and balance predictive performance with model fit (Hastie et al. 2001). We used a tree complexity of 4, learning rate of 0.001, bag fraction of 0.5, and a Gaussian error distribution. After experimentation, all other arguments used default values. Models were interpreted based on predictors' relative influence (RI) values, which can be thought of as model contributions. Using RI values allows these complex ensemble models to be easily interpreted. Relative influence values are based on weighting the number of times a predictor is used for splitting a tree according to the improvement to the model as a result of each split (Friedman & Meulman 2003).

Although BRTs handle data sets with many zeros reasonably well (Elith et al. 2008), very large proportions of zeros may be problematic. Furthermore, some absences (i.e., grid cells with no threatened species records) represent locations where the conditions are too harsh for any of the threatened species to exist, which favors climatic variables in the modeling. This is appropriate for broad-brush modeling, but geodiversity is hypothesized to be most useful in modeling biodiversity at finer scales (Lawler et al. 2015), allowing explanatory power where the environment is similar in other respects (particularly climatically). Therefore, to reduce the influence of absences and constrain the analyses to climates likely to contain threatened species, we reran the species-richness analyses with data sets in which the absence cells were sampled. We used only cells with threatened species records (presences) and cells immediately surrounding those presences. Analyses performed with the full data set had from 6317 to 6560 absences, depending on the taxonomic group, whereas analyses with the sampled data had from 34 to 339 absences. This method of sampling the absences was intended to focus the resulting models on distinguishing cells that contained threatened species from otherwise similar cells that did not (on the basis that neighboring cells tend to be similar because the environment is spatially autocorrelated) to address the hypothesis that geodiversity is important for distinguishing otherwise similar environments.

We ran BRT models for the full set of threatened species (all 16 taxonomic groups combined) and separately for the threatened species richness of each of

8 taxa with sufficient numbers of recorded threatened species to model individually. Models were run on both the sampled and full data sets for vascular plants, fungi, beetles, and bryophytes. Due to data quantity, models could not be run on the sampled data for lichens, butterflies and moths, molluscs, and mammals. Analyses with RWR values only involved the grid cells with threatened species records, which decreased data quantity and limited analyses to vascular plants, bryophytes, fungi, and all species combined.

Self-statistics (SS) were used to assess internal model fit. We then evaluated our models with 10-fold cross-validation (CV), which, along with SS, is included in the gbm package (Ridgeway 2015). Self-statistics and CV range from 0 to 1; a higher number suggests a better model. The CV procedure randomly selects data from the area within which the model was calibrated, excludes these data from the calibration, and then tests the original model on this held-back portion of data. This is repeated 10 times to give an average correlation between the training and testing data. To test whether model fit reflected more than spatial autocorrelation of the variables, we reassessed the fits of all the BRT models by geographically separating calibration and evaluation data and calculating the root mean-squared error of predicted and actual values for the evaluation data. The division of the grid cells into evaluation and calibration data sets was made at the national park level by dividing the parks in an approximately east–west direction. The first data set consisted of 46% of the grid cells from 10 national parks, which were mostly in the east, and the second data set consisted of grid cells from the 21 remaining national parks (Fig. 1). From these 2 data sets, we chose the one with more presence cells in a given taxonomic group as the calibration data and the other as the evaluation data.

Results

Richness of Threatened Species

The best models performed well (SS or CV value closer to 1) for some taxa (e.g., all species, molluscs, and vascular plants) and poorly (SS or CV value closer to 0) for others (e.g., lepidoptera and mammals) (Table 3 shows the models for full data set). The more difficult task of modeling differences in the number of threatened species between similar (neighboring) cells had lower levels of success (Table 3, sampled data set).

Growing-degree days and mean precipitation were usually the dominant predictors of threatened species richness; elevational range and TWI range were also important for threatened vascular plants and bryophytes, respectively (Table 3). In terms of determining the number of threatened species present in the sampled data set, geodiversity variables contributed relatively more than for the full data set (Table 3); their greatest relative contribution was 24.7% for vascular plants. Of the geodiversity variables, geomorphological richness was the most important for most taxa, whereas rock-type richness was the most important for lichens and vascular plants.

Using the full data set, the combined model influence (contribution) of nongeodiversity variables ranged from 86.6% (butterflies and moths) to 99.4% (molluscs). The combined contribution of geodiversity variables therefore ranged from 0.6% to 13.4%. At this scale, the number of threatened species was thus determined primarily by climate, and most of them occurred in relatively warmer and wetter areas. Geodiversity variables were most important in determining the number of threatened species of Lepidoptera and vascular plants (CV correlation for vascular plants was higher than that for all other species groups) (Table 3). In the sampled models, geodiversity variables had 1.5–2 times more influence and the importance of climate-relevant variables was lower than in the full model.

Rarity-Weighted Richness of Threatened Species

Growing-degree days were strongly positively related to RWR for vascular plants, fungi, and all species combined, whereas this relationship was negative for bryophytes (Table 4). High levels of RWR of threatened species were also associated with high rainfall, TWI range, and hydrological feature richness for fungi; low rainfall, small elevational range, and high hydrological feature richness and soil-type richness for vascular plants; high soil-type richness for bryophytes; and low rainfall and high elevational range and rock richness for total threatened species richness. Internal model fits (SS) were reasonable, whereas CV statistics were weaker than SS, as would be expected. Compared with the threatened species richness models, the combined effect of geodiversity variables was much greater for vascular plants (28.9% greater influence from geodiversity relative to the nonsampled richness model) and bryophytes (22.5%) in the RWR analyses, whereas for fungi and for all species combined, there was little difference.

Discussion

Our results from modeling the diversity of 271 threatened species across Finnish national parks revealed that the number of growing-degree days and mean annual precipitation were of the greatest importance for threatened species richness (both raw and weighted by rarity). Elevational range, a widely used topographic metric, was also important. These results are reassuring given the large body of knowledge built up over centuries with regard to overall species richness (Field et al. 2009; Stein et al. 2014). The use of geodiversity–the diversity of Earth

Table 3. Cross-validation (CV) correlation (corr.) and deviance (dev.), self-statistics (SS), root-mean-square error (RMSE) values, and dominant (greatest relative influence or model contribution) geodiversity (GD) and nongeodiversity (non-GD) predictors^{a} of and total combined geodiversity relative influence (RI [%]) on richness of threatened species from boosted regression-tree modeling with the full (F) and sampled (S) data sets for each taxon^{a}.

Taxon	*Data used (n cells)*	*CV corr.*	*CV dev.*	*SS mean*	*RMSE*	*Non-GD predictor (highest RI%)*	*Non-GD predictor (second highest RI%)*	*Combined non-GD model influence (%)*	*GD predictor (highest RI%)*	*GD predictor (second highest RI%)*	*Combined GD model influence (%)*
All	F (6571)	0.56	1.39	0.66	1.22	MP (39.4)^{b}	GDD (36.9)	93.2	rock rich (3.2)	GM rich (2.7)	6.8
	S (856)	0.41	8.63	0.59	3.12	MP (41.1)^{b}	GDD (28.2)	90.3	rock rich (3.7)	GM rich (3.1)	9.7
Bryophytes	F (6571)	0.40	0.05	0.57	0.14	GDD (32.1)^{b}	MP (26.1)	90.2	GM rich (5.6)	rock rich (2.8)	9.8
	S (191)	0.06	1.20	0.46	0.80	TWIR (27.3)^{b}	GDD (24.4)	78.8	GM rich (9.5)	hydro rich (8.4)	21.2
Beetles	F (6571)	0.34	0.04	0.51	0.15	GDD (63.2)^{b}	MP (20.3)	92.9	GM rich (5.2)	rock rich (0.8)	7.1
	S (110)	0.14	1.64	0.50	0.96	GDD (35.4)^{b}	ER (25.5)	87.7	GM rich (7.7)	rock rich (2.3)	12.3
Fungi	F (6571)	0.38	0.60	0.53	0.71	MP (41.1)^{b}	GDD (36.4)	94.5	GM rich (3.2)	soil rich (1.1)	5.5
	S (593)	0.22	5.11	0.48	1.89	MP (36.3)^{b}	GDD (25.5)	85.6	GM rich (8.2)	hydro rich (3.6)	14.4
Lepidoptera	F (6571)	0.28	0.01	0.42	0.05	MP (62.48)^{b}	GDD (16.4)	86.6	GM rich (9.6)	soil rich (2.3)	13.4
Lichens	F (6571)	0.33	0.23	0.54	0.50	GDD (53.7)^{b}	MP (38.9)	95.7	rock rich (2.7)	GM rich (1.3)	4.3
Mammals	F (6571)	0.22	<0.01	0.47	0.05	GDD (38.0)^{b}	MP (35.6)	93.6	GM rich (3.1)	soil rich (2.7)	6.4
Molluscs	F (6571)	0.58	<0.01	0.59	0.04	GDD (69.5)^{b}	MP (25.3)	99.4	GM rich (0.5)	rock rich (0.1)	0.6
Vascular plants	F (6571)	0.59	0.09	0.72	0.17	GDD (29.3)^{b}	ER (24.3)	88.3	rock rich (4.8)	GM rich (3.5)	11.7
	S (319)	0.44	1.08	0.66	0.79	ER (30.0)^{b}	GDD (19.1)	75.3	hydro rich (11.6)	GM rich (9.1)	24.7

[a] *All predictors in the table showed a positive correlation between the variable and predicted species richness in the models. See Table 1 for definitions of variable abbreviations.*

[b] *Greatest contributing predictor per taxon.*

Table 4. Results of analysis of rarity-weighted richness (RWR) in relation to the geodiversity, climate, and topography. Self-statistics (SS) and cross-validation statistics (CV) from the boosted regression-tree modeling of RWR are also shown.

			Abiotic predictor relative influence (%)[a,b]									
Taxon	*Self-statistic*	*Cross-validation statistic*	*GDD*	*MP*	*ER*	*TWIR*	*CNGD*[c]	*rock rich*	*soil rich*	*GM rich*	*hydro rich*	*CGD*[d]
All	0.58	0.33	27.07	-26.08	27.54	7.93(−)	88.62	6.58	0.55(−)	3.16(−)	1.09	11.38
Bryophytes	0.55	0.41	49.21(−)	-5.55	6.02	6.95	67.73	0.04(−)	30.02	1.84	0.37(−)	32.27
Fungi	0.59	0.28	44.33	18.57	7.58	22.78	93.26	0.07	0.10	0.49	6.08	6.74
Vascular plants	0.62	0.19	27.39	10.92(−)	13.98(−)	7.09	59.38	8.13	12.85	6.38	13.26	40.62

[a] *(−) indicates negative relationship between predictor and rarity-weighted richness.*

[b] *Predictors and abbreviations are more fully described in Tables 1 and 2.*

[c] *CNGD, combined nongeodiversity predictors' absolute model influence.*

[d] *CGD, combined geodiversity predictors' absolute model influence.*

surface materials, forms, and processes–improved model predictions, especially when we sampled the data to include only the range of environments in which threatened species are known to occur. This suggests that climatic gradients may determine the regional species pools, and geophysical factors and local heterogeneity have a greater influence at finer scales. To assess the effect of the sampling strategy, we also ran BRT models using a random subset of absence cells; the results were very similar to those reported above. Our results are to some extent contingent on the identities and relative numbers of variables used in the analysis, although our exploratory data analyses and model checking indicated that our conclusions were robust to these issues.

Our findings are consistent with hypothesis H_1 and with previous studies that indicate thermal conditions and energy availability are among the major limiting factors of species patterns in high latitudes, especially at large geographic extents (Hawkins et al. 2003; Field et al. 2009). This should apply particularly to Finland, which is a long and narrow country that extends over 1000 km north to south and thus has a strong latitudinal gradient in climate. Sampling the data controlled climate to some extent because it removed from consideration the cells that were far from places with threatened species. Even so, the cells spanned nearly the entire length of Finland, and climate remained fairly dominant in the models.

Although the strong modeled effect of growing-degree days was consistent and always positive, threatened species richness of different taxonomic groups had unique relationships with climate, topography, and geodiversity variables, in line with the habitat heterogeneity hypothesis (H_2). This was most prominent in analyses of the sampled data, where absences were restricted to cells neighboring presences. We found support specifically for the hypothesis that geodiversity variables improve models over and above climate and topography (H_3). In the models, geodiversity variables generally had relatively small (consistent with Anderson et al. 2015) but consistent additive effects and improved the diversity models' predictive ability for beetles, bryophytes, and vascular plants. Geodiversity predictors were most important when using the sampled data, rather than the full data set. Thus, although climate is a key driving force, explicitly prioritizing the diversity of geophysical settings may help conserve abiotic and biotic diversity in a dynamic climate.

Of the geodiversity variables, rock-type richness was relatively important for the richness (raw and rarity weighted) of threatened vascular plant species and for threatened lichen species richness. Elsewhere, rock richness has been firmly linked to biodiversity of lichens (Spitale & Nascimbene 2012) and plants (e.g., Pausas et al. 2003; Kougioumoutzis & Tiniakou 2014). Geomorphological richness was consistently significant (if not always strong) in our models, especially for raw richness of threatened species. Geomorphological features provide unique abiotic conditions (Nichols et al. 1998; Hjort et al. 2015) and may promote threatened species diversity (Bétard 2013).

Soil and hydrology both contributed significantly to our models of richness and RWR. Soil-type diversity was important for the RWR of threatened bryophytes and vascular plants, which is consistent with claims that soil-type diversity affects vegetation (e.g., Stein et al. 2014). Although aquatic environments are not the main habitat for many threatened vascular plants in boreal environment, several species occupy habitats near water, such as streams and ponds (e.g., *Saxifraga hirculus*) (Rassi et al. 2010), which may explain the importance of hydrological richness in models for threatened vascular plant species richness (raw and rarity weighted).

We postulated that threatened species can be modeled successfully using climate, topography, and geodiversity variables (H_4). Models based on the full data set performed well, whereas the quality of performance of models based on sampled data differed. These results are consistent with climate being the main control of species patterns at large extents (Hawkins et al. 2003; Field et al. 2009). Climate can be used to provide the first cut for predicting threatened species diversity and distribution at regional and national scales. When we constrained species' absences to within 1 km of the presence of a threatened species, sampling the data removed the most climatically unsuitable grid cells. This reduced the ranges of the climatic variables and made reduction in the explanatory power of climate almost inevitable. This left a data set of cells that were climatically suitable for at least some threatened species and with which it was difficult to predict absences. It is for such cells that geodiversity variables should be, and were, most useful. These variables may represent the next set of abiotic requirements once the climatic-tolerance filter has been passed.

In the analyses of RWR of threatened vascular plant and fungal species, the combined contribution from geodiversity variables exceeded the equivalent contribution in the raw threatened species richness models (Tables 2 & 3). Thus, geodiversity appears to be particularly useful for understanding the patterns of threatened species diversity when site complementarity is considered. Even so, as with the raw richness of threatened species analyses, climate variables dominated our analyses. Growing-degree days were strongly and negatively associated with RWR for bryophytes. The most infrequently occurring bryophytes in Finland therefore appeared to occur in relatively cold areas with short growing seasons and high mineral and nutrient diversity from a greater variety of soils (Rassi et al. 2010). In a practical sense, we see value in conducting analyses such as ours, for threatened species data, to most effectively allocate conservation resources; geodiversity data can be used alongside other climate and topography data to better target areas that support rare and irreplaceable communities (Albuquerque & Beier 2015*c*).

Both abiotic environmental heterogeneity (Stein et al. 2014) and geodiversity (Gray 2013) have been described and measured in a wide variety of ways. It is likely that the best measure to use will be specific to the study area and taxon, but methods should be transparent and transferable where possible. We used simple measures of geodiversity, which provided a time-saving and financially practical way of measuring abiotic environmental heterogeneity of the uniformly sized study units. These geodiversity measures complement broad-brush topographic variables (e.g., elevational range) that are commonly used in analyses of biodiversity and represent a more detailed appraisal of the abiotic environment. Often, geodiversity includes topographic data (e.g., Parks & Mulligan 2010), and this inclusion would considerably increase the effect of geodiversity in our analyses. However, we see value in distinguishing between these different aspects of environmental heterogeneity and in analyzing the effects of explicit measures of geology, landforms, and hydrology.

Consideration of specific geosites, including geological objects or fragments of the geological environment exposed on the land surface, and not just overall geodiversity, could improve our models and understanding of biodiversity (Hjort et al. 2015). It would also be useful to study the relationship between biodiversity and different measures of geodiversity (measures other than richness) because these relationships are likely to depend on the metrics used. The positive relationship between overall species richness and geodiversity is becoming reasonably well established in the literature, but the connection between threatened species and geodiversity deserves more attention in future research. We stress the need for wider incorporation of geodiversity–not just DEM-based topographic variables–in ecology and macroecology and in both conservation theory and management.

Acknowledgments

We thank P. Beier, 3 anonymous referees, C. Lavers, A. Algar, and members of the Biogeography, Ecology and Evolution Research group at the University of Nottingham for valuable comments on previous versions of the manuscript. We also thank Metsähallitus and The Finnish Environment Institute for providing us with the threatened species data and E. Hanski for the classification of the bedrock types. H.T. acknowledges the Geography Graduate School in Finland and Kone Foundation. J.J.B. is supported by a Natural Environment Research Council (NERC, United Kingdom) PhD Studentship (1365737). J.H. acknowledges the Academy of Finland (grant numbers 267995 and 285040). All authors conceived the research. All authors except R.F. amassed the data. J.J.B. and

H.T. did the main analyses and wrote the text with major contributions from R.F. and J.H. All authors commented on and edited the text.

Supporting Information

Information on the taxonomic groups included in the study (Appendix S1), the species list (Appendix S2), the species distribution modeling analyses (Appendix S3), and details of the modeling procedure of geomorphological richness (Appendix S4) are available online. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Ahti T, Hämet-Ahti L, Jalas J. 1968. Vegetation zones and their sections in northwestern Europe. Annales Botanici Fennici **5:**169–211.

Albuquerque F, Beier P. 2015*a*. Using abiotic variables to predict importance of sites for species representation. Conservation Biology **29:**1390–1400.

Albuquerque F, Beier P. 2015*b*. Rarity-weighted richness: a simple and reliable alternative to integer programming and heuristic algorithms for minimum set and maximum coverage problems in conservation planning. PLoS ONE **10**(3)**:**e0119905. DOI: 10.1371/journal.pone.0119905.

Albuquerque F, Beier P. 2015*c*. Global patterns and environmental correlates of high-priority conservation areas for vertebrates. Journal of Biogeography **42:**1397–1405.

Anderson MG, Comer PJ, Beier P, Lawler JJ, Schloss CA, Buttrick S, Albano CM, Faith DP. 2015. Case studies of conservation plans that incorporate geodiversity. Conservation Biology **29:**680–691.

Anderson MG, Ferree CE. 2010. Conserving the stage: climate change and the geophysical underpinnings of species diversity. PLoS ONE **5**(7)**:**e11554. DOI: 10.1371/journal.pone.0011554.

Ardron J, Possingham HP, Klein C. 2010. Marxan good practices handbook, version 2. Pacific Marine Analysis and Research Association. Available from www.pacmara.org (accessed May 1, 2016).

Atlas of Finland. 1990. Geology. National Board of Survey and Geographical Society of Finland, Helsinki.

Beier P, Hunter ML, Anderson M. 2015. Special section: conserving nature's stage. Conservation Biology **29:**613–617.

Berg Å, Ehnström B, Gustafsson L, Hallingbäck T, Jonsell M, Weslien J. 2002. Threatened plant, animal, and fungus species in Swedish forests: distribution and habitat associations. Conservation Biology **8:**718–731.

Bétard F. 2013. Patch–scale relationships between geodiversity and biodiversity in hard rock quarries: case study from a disused quartzite quarry in NW France. Geoheritage **5:**59–71.

Beven KJ, Kirkby MJ. 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin **24:**43–69.

Burnett MR, August PV, Brown JHJ, Killingbeck KT. 1998. The influence of geomorphological heterogeneity on biodiversity: I. A patch-scale perspective. Conservation Biology **12:**363–370.

Dobrowski SZ. 2011. A climatic basis for microrefugia: the influence of terrain on climate. Global Change Biology **17:**1022–1035.

Elith J, Leathwick JR, Hastie T. 2008. A working guide to boosted regression trees. Journal of Animal Ecology **77:**802–813.

Engler R, Guisan A, Rechsteiner L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. Journal of Applied Ecology **41:**263–274.

Field R, et al. 2009. Spatial species-richness gradients across scales: a meta-analysis. Journal of Biogeography **36:**132–147.

Finnish Environment Institute. 2015. Monitoring related to the theme maintaining ecosystem services and biodiversity. Finnish Environment Institute, Helsinki. Available from http://www.syke.fi/en-US/Research__Development/Maintaining_ecosystem_services_and_biodiversity/Monitoring (accessed January 2015).

Friedman JH, Meulman JJ. 2003. Multiple additive regression trees with application in epidemiology. Statistics in Medicine **22:**1365–1381.

Gray M. 2013. Geodiversity: valuing and conserving abiotic nature. 2nd edition. Wiley-Blackwell, Chichester.

GSF (Geological Survey of Finland). 2010*a*. Superficial deposits of Finland 1:200 000. GSF, Espoo.

GSF (Geological Survey of Finland). 2010*b*. Bedrock of Finland 1:200 000. GSF, Espoo.

Hastie T, Tibshirani R, Friedman J. 2001. The elements of statistical learning: data mining, inference, and prediction. Springer-Verlag, New York.

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

Heinonen M, editor. 2013. Applying IUCN protected area management categories in Finland. International Union for Conservation of Nature, Gland, Switzerland. Available from http://www.iucn.org/about/work/programmes/gpap_home/gpap_capacity2/gpap_pub/gpap_catpub/?14652/Applying-IUCN-Protected-Area-Management-Categories-in-Finland (accessed June 2015).

Hjort J, Gordon JE, Gray M, Hunter ML. 2015. Why geodiversity matters in valuing nature's stage. Conservation Biology **29:**630–639.

Hjort J, Heikkinen RK, Luoto M. 2012. Inclusion of explicit measures of geodiversity improve biodiversity models in a boreal landscape. Biodiversity and Conservation **21:**3487–3506.

Hjort J, Luoto M. 2012. Can geodiversity be predicted from space? Geomorphology **153–154:**74–80.

Kirkpatrick JB. 1983. An iterative method for establishing priorities for the selection of nature reserves: an example from Tasmania. Biological Conservation **25:**127–134.

Kougioumoutzis K, Tiniakou A. 2014. Ecological factors driving plant species diversity in the South Aegean Volcanic Arc and other central Aegean islands. Plant Ecology & Diversity **8:**173–186.

Kreft H, Jetz W. 2007. Global patterns and determinants of vascular plant diversity. Proceedings of the National Academy of Sciences United States of America **104:**5925–5930.

Lawler JJ, Ackerly DD, Albano CM, Anderson MG, Dobrowski SZ, Gill JL, Heller NE, Pressey RL, Sanderson EW, Weiss SB. 2015. The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. Conservation Biology **29:**618–629.

Mantyka-Pringle CS, Visconti P, Di Marco M, Martin TG, Rondinini C, Rhodes JR. 2015. Climate change modifies risk of global biodiversity loss due to land-cover change. Biological Conservation **187:**103–111.

McCune B, Keon D. 2002. Equations for potential annual direct incident radiation and heat load. Journal of Vegetation Science **13:**603–606.

Moilanen A, Meller L, Leppanen J, Pouzols FM, Arponen A, Kujao H. 2014. Zonation spatial conservation planning framework and software, version 4. Available from http://cbig.it.helsinki.fi/ (accessed May 2016).

Nichols WF, Killingbeckm KT, August PV. 1998. The influence of geomorphological heterogeneity on biodiversity: II. A landscape perspective. Conservation Biology **12:**371–379.

NLS (National Land Survey of Finland). 2000. Digital elevation model. NLS, Helsinki.

NLS (National Land Survey of Finland). 2007. Topographical database. NLS, Helsinki.

Parks KE, Mulligan M. 2010. On the relationship between a resource based measure of geodiversity and broad scale biodiversity patterns. Biodiversity and Conservation **19:**2751–2766.

Pausas JC, Carreras J, Ferré A, Font X. 2003. Coarse-scale plant species richness in relation to environmental heterogeneity. Journal of Vegetation Science **14:**661–668.

Pirinen P, Simola H, Aalto J, Kaukoranta J, Karlsson P, Ruuhela R. 2012. Climatological statistics of Finland 1981–2010. Finnish Meteorological Institute, Helsinki.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available from http://www.R-project.org.

Rassi P, Alanen A, Kanerva T, Mannerkoski I, editors. 2001. The 2000 Red List of Finnish species. Ministry of the Environment, Finnish Environment Institute, Helsinki.

Rassi P, Hyvärinen E, Juslen A, Mannerkoski I, editors. 2010. The 2010 Red List of Finnish species. Ministry of the Environment, Finnish Environment Institute, Helsinki.

Rich PM, Weiss SB. 1991. Spatial models of microclimate and habitat suitability: lessons from threatened species. Proceedings of the 11th annual ESRI user conference. ESRI Environmental Systems Research Institute, Redlands, California.

Ridgeway G. 2015. Generalized boosted regression models. Documentation on the R Package 'gbm', version 2.1.1. Available from http://cran.r-project.org/web/packages/gbm/gbm.pdf (accessed September 2015).

Siikamäki P, Kangas K, Paasivaara A, Schroderus S. 2015. Biodiversity attracts visitors to national parks. Biodiversity and Conservation **24:**2521–2534.

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

Spitale D, Nascimbene J. 2012. Spatial structure, rock type, and local environmental conditions drive moss and lichen distribution on calcareous boulders. Ecological Research **27:**633–638.

Stein A, Gerstner K, Kreft H. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. Ecology Letters **17:**866–880.

Virkkala R, Luoto M, Heikkinen RK, Leikola N. 2005. Distribution patterns of boreal marshland birds: modelling the relationships to land cover and climate. Journal of Biogeography **32:**1957–1970.

Williams P, Gibbons D, Margules C, Rebelo A, Humphries C, Pressey RL. 1996. A comparison of richness hotspots, rarity hotspots and complementary areas for conserving diversity using British birds. Conservation Biology **10:**155–174.