



Bailey, Joseph J. and Boyd, Doreen S. and Hjort, Jan and Lavers, Chris P. and Field, Richard (2017) Modelling native and alien vascular plant species richness: at which scales is geodiversity most relevant? *Global Ecology and Biogeography*, 26 (7). pp. 763-776. ISSN 1466-8238

**Access from the University of Nottingham repository:**

[http://eprints.nottingham.ac.uk/39903/8/Bailey\\_et\\_al-2017-Global\\_Ecology\\_and\\_Biogeography.pdf](http://eprints.nottingham.ac.uk/39903/8/Bailey_et_al-2017-Global_Ecology_and_Biogeography.pdf)

**Copyright and reuse:**

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution licence and may be reused according to the conditions of the licence. For more details see: <http://creativecommons.org/licenses/by/2.5/>

**A note on versions:**

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact [eprints@nottingham.ac.uk](mailto:eprints@nottingham.ac.uk)

# Modelling native and alien vascular plant species richness: At which scales is geodiversity most relevant?

Joseph J. Bailey<sup>1</sup>  | Doreen S. Boyd<sup>1</sup>  | Jan Hjort<sup>2</sup> | Chris P. Lavers<sup>1</sup> | Richard Field<sup>1</sup> 

<sup>1</sup>School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

<sup>2</sup>Geography Research Unit, University of Oulu, P.O. Box 8000, FI-90014, Finland

## Correspondence

Joseph J. Bailey, School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom.  
Email: josephjbailey@outlook.com

Editor: Dr. Adriana Ruggiero

## Funding information

This research was supported by the U.K. Natural Environment Research Council (NERC) PhD Studentship 1365737, which was awarded to J.J.B., University of Nottingham, in October 2013 (supervised by R.F. and D.B.). J.H. acknowledges the Academy of Finland (project number 285040).

## Abstract

**Aim:** To explore the scale dependence of relationships between novel measures of geodiversity and species richness of both native and alien vascular plants.

**Location:** Great Britain.

**Time period:** Data collected 1995–2015.

**Major taxa:** Vascular plants.

**Methods:** We calculated the species richness of terrestrial native and alien vascular plants (6,932 species in total) across the island of Great Britain at grain sizes of 1 km<sup>2</sup> ( $n = 219,964$ ) and 100 km<sup>2</sup> ( $n = 2,121$ ) and regional extents of 25–250 km diameter, centred around each 100-km<sup>2</sup> cell. We compiled geodiversity data on landforms, soils, hydrological and geological features using existing national datasets, and used a newly developed geomorphometric method to extract landform coverage data (e.g., hollows, ridges, valleys, peaks). We used these as predictors of species richness alongside climate, commonly used topographic metrics, land-cover variety and human population. We analysed species richness across scales using boosted regression tree (BRT) modelling and compared models with and without geodiversity data.

**Results:** Geodiversity significantly improved models over and above the widely used topographic metrics, particularly at smaller extents and the finer grain size, and slightly more so for native species richness. For each increase in extent, the contribution of climatic variables increased and that of geodiversity decreased. Of the geodiversity variables, automatically extracted landform data added the most explanatory power, but hydrology (rivers, lakes) and materials (soil, superficial deposits, geology) were also important.

**Main conclusions:** Geodiversity improves our understanding of, and our ability to model, the relationship between species richness and abiotic heterogeneity at multiple spatial scales by allowing us to get closer to the real-world physical processes that affect patterns of life. The greatest benefit comes from measuring the constituent parts of geodiversity separately rather than one combined variable (as in most of the few studies to date). Automatically extracted landform data, the use of which is novel in ecology and biogeography, proved particularly valuable in our study.

## KEYWORDS

alien species, biodiversity, conserving nature's stage, environmental heterogeneity, geodiversity, geology, geomorphometry, native species, scale, vascular plants

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Global Ecology and Biogeography* Published by John Wiley & Sons Ltd

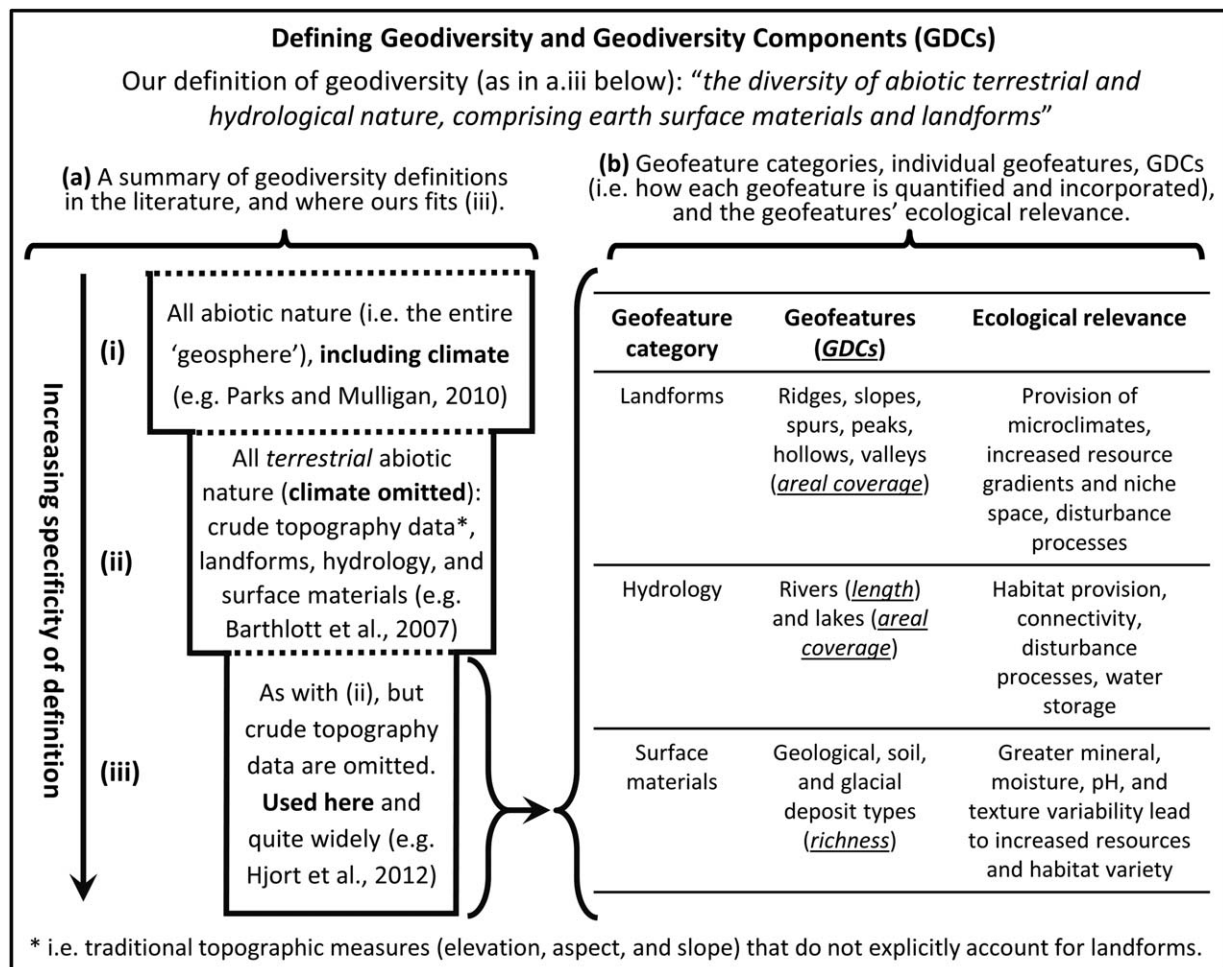
## 1 | INTRODUCTION

Understanding the spatial patterns of biodiversity is important for scientific theory, conservation and management of ecosystem services (Hanski et al., 2012; Lomolino, Riddle, Whittaker, & Brown, 2010). Climatic variables are well known to correlate strongly with species richness over large spatial extents (Hawkins et al., 2003); correlates of species richness at smaller extents (regional and landscape scales) are less well established (Field et al., 2009; Valdés et al., 2015), but environmental heterogeneity is widely thought to be important (Stein, 2015; Stein, Gerstner, & Kreft, 2014). Although a bewildering array of measures of environmental heterogeneity have been used, there is growing interest in geodiversity, both as having value in itself (Gray, 2013) and as a potential correlate and predictor of spatial biodiversity patterns (Lawler et al., 2015).

Geodiversity, which we define as ‘the diversity of abiotic terrestrial and hydrological nature, comprising earth surface materials and landforms’ (Figure 1a), may be an important correlate of biodiversity at landscape and subnational scales (Barthlott et al., 2007; Gray, 2013; Hjort, Heikkinen, & Luoto, 2012; Lawler et al., 2015). Geodiversity

comprises ‘geofeatures’ (Figure 1b), which are the individual landforms and geological types (for example) that constitute the abiotic landscape. Quantification of these geofeatures varies across studies (e.g., Pellitero, Manosso, & Serrano, 2015). We introduce the term ‘geodiversity component’ (GDC; Figure 1b), to refer to the quantified geofeature, whether this be areal coverage (e.g., of a particular landform), richness (e.g., the number of geological types) or length (e.g., of a river). These GDCs together measure ‘geodiversity’ at the scale being studied. The GDCs we use here are intended to capture aspects of the abiotic heterogeneity with which living organisms interact – and thus better and more explicitly measure environmental heterogeneity for the purposes of explaining species richness patterns than crude topographic measures such as mean slope, elevational range or mean aspect (Figure 1). Such topographic measures have been widely used as correlates or predictors of species richness (Stein & Kreft, 2014), and to create a conceptual distinction we omit these from our definition of geodiversity.

A small but rapidly growing number of studies have found that explicit measures of geodiversity add explanatory power to statistical models accounting for spatial biodiversity patterns (e.g., Hjort et al.,



**FIGURE 1** Our definition of ‘geodiversity’, which is amongst the more specific in the context of the wider literature. It omits relatively crude topography and climate data (a) and consists of geodiversity components (GDCs). The GDCs used in our study, and their associated geofeatures and ecological relevance, are listed (b)

2012; Kougioumoutzis & Tiniakou, 2014; Pausas, Carreras, Ferre, & Font, 2003; Tukiainen, Bailey, Field, Kangas, & Hjort, 2016; see the review by Lawler et al., 2015). However, these studies have tended either to consider only one or two aspects of geodiversity or use a single geodiversity variable that simply counts geofeatures to produce an overall measure of georichness (e.g., Hjort et al., 2012; Räsänen et al., 2016). The considerable improvements in explanatory power that these preliminary approaches have achieved indicate the need for fuller analysis of the relationship between biodiversity and geodiversity, and particularly for explicit consideration of the separate components of geodiversity (Beier et al., 2015a). To date, very few studies have attempted this, and even fewer at geographical extents greater than the landscape scale – except for that by Tukiainen et al. (2016), which only analyses threatened species. Therefore, we now have evidence suggesting that geodiversity affects biodiversity, but our understanding of how it does so remains severely limited.

All of the GDCs in Figure 1 measuring geomorphological, hydrological, geological and pedological geofeatures implicitly incorporate local abiotic variability and processes that are considered to have important influences on species richness via local resource availability, habitat diversity and niche variety (Albano, 2015; Bétard, 2013; Burnett, August, Brown, & Killingbeck, 1998; Dufour, Gadallah, Wagner, Guisan, & Buttler, 2006; Hjort, Gordon, Gray, & Hunter, 2015). Processes and abiotic variability related to geofeatures include, but are not limited to: microclimatic and sheltering effects around landforms (e.g., hollows and ridges); erosion, water storage capacity, physical and chemical weathering, pH variability, and mineral and textural variety via geology and soil; and water storage, transfer and connectivity via hydrological features and rock composition and soil texture (Guitet, Péliissier, Brunaux, Jaouen, & Sabatier, 2014; Hjort et al., 2015; Moser et al., 2005). GDCs may also be linked to natural geomorphological and hydrological disturbance processes, which are relevant to vegetation diversity and distributions (e.g., Le Roux, Virtanen, & Luoto, 2013; Randin, Vuissoz, Liston, Vittoz, & Guisan, 2009; Viles, Naylor, Carter, & Chaput, 2008; Virtanen et al., 2010). Much of this information is lost when crude topographic measures, such as elevational range and mean slope, are used.

While we know much about the scale dependence of the relationships between species richness and many of its commonly used correlates (McGill, 2010; Mittelbach et al., 2001; Pausas et al., 2003; Ricklefs, 1987; Rosenzweig, 1995), little is known about the scales at which richness is most strongly correlated with geodiversity. Current thinking is that geodiversity is most relevant to species richness at landscape to regional extents, with climate dominating at broader (e.g., continental) extents and biotic interactions more locally (Lawler et al., 2015). Theoretically, the local and landscape extents are most relevant because the various GDCs may be amongst the most variable predictors at this scale (Tukiainen et al., 2016; Willis & Whittaker, 2002), unlike climate. Therefore, if GDCs are important determinants of the spatial arrangement of biodiversity, we should expect their statistical explanatory power to be strongest at the local and landscape scales. We also know relatively little about the importance of grain size in

modelling species richness. Theoretically, coarser grain sizes may average out fine-scale abiotic environmental heterogeneity over a larger area, thus relating more weakly to species richness, unless these fine-scale data are related to broad environmental gradients (Field et al., 2009; Hawkins et al., 2003).

A key reason for the limited research to date on geodiversity and its relationship with biodiversity is limited data availability. In broad-scale macroecological studies in particular, the widespread use of topographic measures to date, such as topographic range or standard deviation, in statistical models of species richness patterns is explained primarily by the difficulty of obtaining more sophisticated and meaningful environmental heterogeneity variables (e.g., O'Brien, Field, & Whittaker, 2000). However, better data and processing capabilities now allow landscape heterogeneity to be quantified in new ways. Here we take advantage of these developments to move beyond simplistic measures of topographic heterogeneity and derive novel geodiversity variables. In particular, we use 'geomorphon', a recently developed geomorphometric tool for extracting landform data from digital elevation models (Jasiewicz & Stepinski, 2013). This allows low-cost quantification of landform features, which we use to measure landform richness at a spatial resolution of 25 m across the whole island of Great Britain.

Alien and native species richness are likely to relate differently to the abiotic environment (Kumar, Stohlgren, & Chong, 2006; Pyšek et al., 2005), but little work has compared the relationship of alien and native species richness with environmental heterogeneity. Native species have had longer to equilibrate with abiotic environmental conditions (Räsänen et al., 2016), so their richness may be expected to be more closely related to geofeatures and topography. Conversely, geofeatures may account less well for alien species richness, especially of neophytes (species introduced after AD 1500), which are more likely to be found where temperatures are higher and where there is greater human presence and connectivity via transport networks (Celesti-Gradow et al., 2006; Pyšek, 1998). An exception may be waterways – these geofeatures can promote the spread of alien species (Deuschewitz, Lausch, Kühn, & Klotz, 2003). Natural disturbance processes may also create suitable conditions for alien species (Fleishman, Murphy, & Sada, 2006). Broadly, we expect native species to have the strongest relationship with geodiversity, followed by archaeophytes (alien species introduced before AD 1500) and then neophytes.

Overall, despite the clear potential for geodiversity to improve our understanding of spatial biodiversity patterns in relation to environmental heterogeneity, its incorporation into biodiversity modelling is underdeveloped conceptually, spatially and empirically. Outstanding questions include: At what spatial scales and in which types of location is geodiversity most relevant? For which taxa? Does it relate differently to alien species than to native species? Which geofeatures are most important? Here, we begin to address some of these knowledge gaps by analysing the relationships between a wide range of GDCs and the species richness of both native and alien vascular plants across Great Britain. We test the degree to which GDCs add explanatory power over and above widely used topographic and climatic variables at varying spatial scales, using two grain sizes and either seven (small grain

size) or five (large grain size) study-area extents. Our main aims are to determine: (a) the scales at which geodiversity best accounts for species richness patterns; (b) which components of geodiversity account for the most variation in species richness, and how much; and (c) whether geodiversity–species richness relationships differ between native and alien species. Specifically, we tackle the following hypotheses: (H<sub>1</sub>) geodiversity will contribute significantly to biodiversity models, particularly at smaller study-area extents (Hjort et al., 2015; Tukiainen et al., 2016); (H<sub>2</sub>) the most relevant GDCs will vary between native and alien species (Deuschewitz et al., 2003) and, within alien species, between archaeophytes and neophytes.

## 2 | METHODS

### 2.1 | Data

All predictors and predictor sets are summarized in Table 1. Data sources are detailed further in Appendix S1 in the Supporting Information. Data were compiled for each 1 km<sup>2</sup> ( $n = 222,111$ ) and 100 km<sup>2</sup> ( $n = 2,121$ ) British National Grid cell using ArcGIS 10 (and GRASS GIS for geomorphometry, as detailed below) and processed and analysed in R (R Core Team, 2016).

Vegetation data were provided by the Botanical Society of Britain and Ireland (BSBI) via the Distribution Database at two grain sizes: 1 km × 1 km ('monad') and 10 km × 10 km ('hectad') grid cells corresponding to the British National Grid. The BSBI hosts a single database to which data are contributed by its volunteers and coordinators, who are strongly encouraged to use unbiased sampling (Walker, Pearman, Ellis, McIntosh, & Lockton, 2010). We used accepted data records (those verified within the database) collected between January 1995 and September 2015.

Species were defined as native, archaeophyte (probably introduced by humans before AD 1500), or neophyte (after AD 1500). 'Casual aliens' (those that fail to establish) were excluded. Total species richness (all three groups plus uncategorized species or those with no accepted status) and alien species richness (archaeophytes plus neophytes) were also modelled. Status definitions of each species followed the Wild Flower Society (2010), which, in turn, used multiple sources. Grid cells with less than 75% land coverage (considering lakes and ocean) were excluded. The final dataset contained 6,932 species: 1,490 natives and 1,331 aliens comprising 151 archaeophytes and 1,180 neophytes, the rest of the species being unclassified.

Undersampled grid cells were excluded – this removed bias arising from unrealistic species richness values due to undersampling. To determine undersampling, we performed a series of linear regressions that used climate and topography variables (not geodiversity) to account for the species richness of grid cells within a radius of 150 km around each hectad. A cell within this region was flagged as potentially undersampled if its standardized residual was less than  $-1.5$  (i.e., if species richness in that cell was strongly over-predicted). This was repeated for every hectad for both grain sizes. Grid cells flagged as undersampled more than 15% of the time they were analysed were classed as undersampled and removed. Two hectads (0.1%) and 2,147

monads (1%) were removed, leaving 2,121 and 219,964, respectively. This procedure ensures that grid cells are not perceived to be under-sampled when they are simply in harsh environments that would most likely contain few species anyway.

A 25 m × 25 m-resolution digital elevation model (DEM) was produced by resampling the 5 m × 5 m NEXTMap DEM from Intermap (obtained under academic license via the NERC Earth Observation Data Centre; see Table 1). Using the DEM, we performed geomorphometric analyses (see below) and calculated commonly used topographic metrics (mean and standard deviation of elevation and slope). We downloaded c. 1-km<sup>2</sup> resolution climate data from WorldClim (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). We calculated land-cover variety using the number of Corine land-cover classes. The total human population per grid cell was calculated from 2010 census data from Casweb.

We compiled GDCs (Figure 1b) using existing national datasets and automated extraction of landform coverage using geomorphometry (Table 1). Data included geological diversity and superficial deposit diversity derived from 1:50,000 scale shapefiles provided by the British Geological Survey under an academic licence. Soil texture data were from the same source but had a resolution of 1 km<sup>2</sup>. We calculated river length and lake area using OS Strategi GIS data. We used the geomorphometric algorithm 'r.geomorphon' developed by Jasiewicz & Stepinski (2013) in GRASS GIS 7.1 (GRASS Development Team, 2016) to automatically extract landform coverage data from the DEM (Appendix S2). The following landforms and features were mapped in raster format: peaks, ridges, shoulders, spurs, slopes, footslopes, hollows, valleys, flat areas and pits. We did not explicitly quantify mineralogy and pH, but these are implicitly incorporated via geology. Fossils, important for geoheritage and geoconservation (Thomas, 2012), were not included because of their limited theoretical relevance to the biodiversity patterns studied here and a lack of consistent data. Maps of climate, topography and geofeatures are presented in Appendix S3.

### 2.2 | Analysis

We developed species richness models for three predictor sets: (a) geodiversity only, (b) geodiversity variables excluded (leaving standard topographic variables, climate, population and land-cover variety) and (c) all variables. Models were run for all species groups (all species, native, alien, archaeophyte, neophyte) at 12 scales, where 'scale' refers to a unique combination of extent and grain size; we used seven geographical extents [whole of Great Britain ('national') and 25-, 50-, 100-, 150-, 200- and 250-km diameter regions; see bottom of Figure 2] and two grain sizes [1 km<sup>2</sup> (monad) and 100 km<sup>2</sup> (hectad)]. The two smallest extents were not used for the coarser grain size for reasons of sample size. All regional models were run using the centroid of each hectad grid cell ( $n = 2,121$ ) as the central point of each 'region'.

We used boosted regression trees (BRTs) to model species richness in R 3.0.2 (R Core Team, 2016). BRT is a machine-learning method that can be seen as an advanced form of regression modelling (Elith, Leathwick, & Hastie, 2008). Here, with a complex dataset, largely unknown relationships (particularly GDCs) and multiple scales with variable collinearities and interactions, use of a BRT was efficient and

TABLE 1 A summary of the variables within each predictor set

| Predictor set                  | [Predictor sub-set] Variable   | Resolution/scale               | Measurement per 1 km × 1 km and 10 km × 10 km grid cell  | Source (detailed in Appendix S1)  |
|--------------------------------|--|--------------------------------|--|---|
| Geodiversity components (GDCs) | [Landforms] Coverage of ridges, slopes, spurs, peaks, hollows, valleys.                    | 25 m (resampled from 5 m)      | Areal coverage; geomorphon <sup>3</sup> in GRASS GIS 7.1 | NEXMap data (Intermap, 2015 <sup>2</sup> via NEODC, 2015 <sup>3</sup> ) |
|                                | [Materials] Geological richness  | 1:50,000                       | No. of rock types  | British Geological Survey <sup>4</sup>                                  |
|                                | [Materials] Superficial deposit richness   | 1:50,000                       | No. of sup. dep. types                                   | British Geological Survey <sup>4</sup>                                  |
|                                | [Materials] Soil texture richness  | 1:50,000                       | No. of texture types                                     | British Geological Survey <sup>4</sup>                                  |
|                                | [Hydrology] River length   | 1:50,000                       | Total length   | OS Strategi via Edina Digimap <sup>5</sup>                              |
|                                | [Hydrology] Lake area  | 1:50,000                       | Areal coverage   | OS Strategi via Edina Digimap <sup>5</sup>                              |
| Climate                        | Bioclimatic variables*: 1, 2, 4, 6, 12, 15   | 30 arcsec (c. 1 km × 1 km)     | Mean   | WorldClim (Hijmans et al., 2005)  |
| Topography                     | Mean elevation; standard deviation in elevation<br>Mean slope; standard deviation in slope | 25 m (resampled from 5 m)      | Mean   | NEXMap data (Intermap, 2015 <sup>2</sup> via NEODC, 2015 <sup>3</sup> ) |
| Land cover and anthropogenic   | Land cover variety   | 100 m                          | Number of land cover types                               | Corine Landcover (2013) <sup>6</sup>                                    |
|                                | 2010 total human population  | Census lower super output area | Total  | 2010 UK census data (Casweb) <sup>7</sup>                               |

The modelling uses three combinations of these predictor sets: (a) geodiversity only; (b) all predictors except for geodiversity and (c) all predictors combined. Details of the data sources and URLs are provided (Appendix S1).

\*Bioclimatic variables (WorldClim): 1, annual mean temperature; 2, mean diurnal range [mean of monthly (max. temp. – min. temp.)]; 4, temperature seasonality (standard deviation × 100); 6, minimum temperature of the coldest month; 12, annual precipitation; 15, precipitation seasonality (coefficient of variation).

References: <sup>1</sup>Jasiewicz & Stepinski (2013); <sup>2</sup>Intermap (<http://www.intermap.com/data/nextmap>); <sup>3</sup>NERC Earth Observation Data Centre (<http://www.neodc.rl.ac.uk/>); <sup>4</sup>licence no. 2014/128 ED British Geological Survey © NERC. All rights reserved; <sup>5</sup>Ordnance Survey Strategi Data via Edina Digimap (<http://digimap.edina.ac.uk/>); <sup>6</sup>Corine Landcover (<http://www.eea.europa.eu/publications/COR0-landcover>); <sup>7</sup>Casweb (<http://casweb.mimas.ac.uk/>).

appropriate. Additionally, BRTs explicitly consider interactions, which can indicate important combined effects, and handle nonlinearity and collinearity relatively well (Dormann et al., 2013; Elith et al., 2008). However, we also assessed collinearities separately.

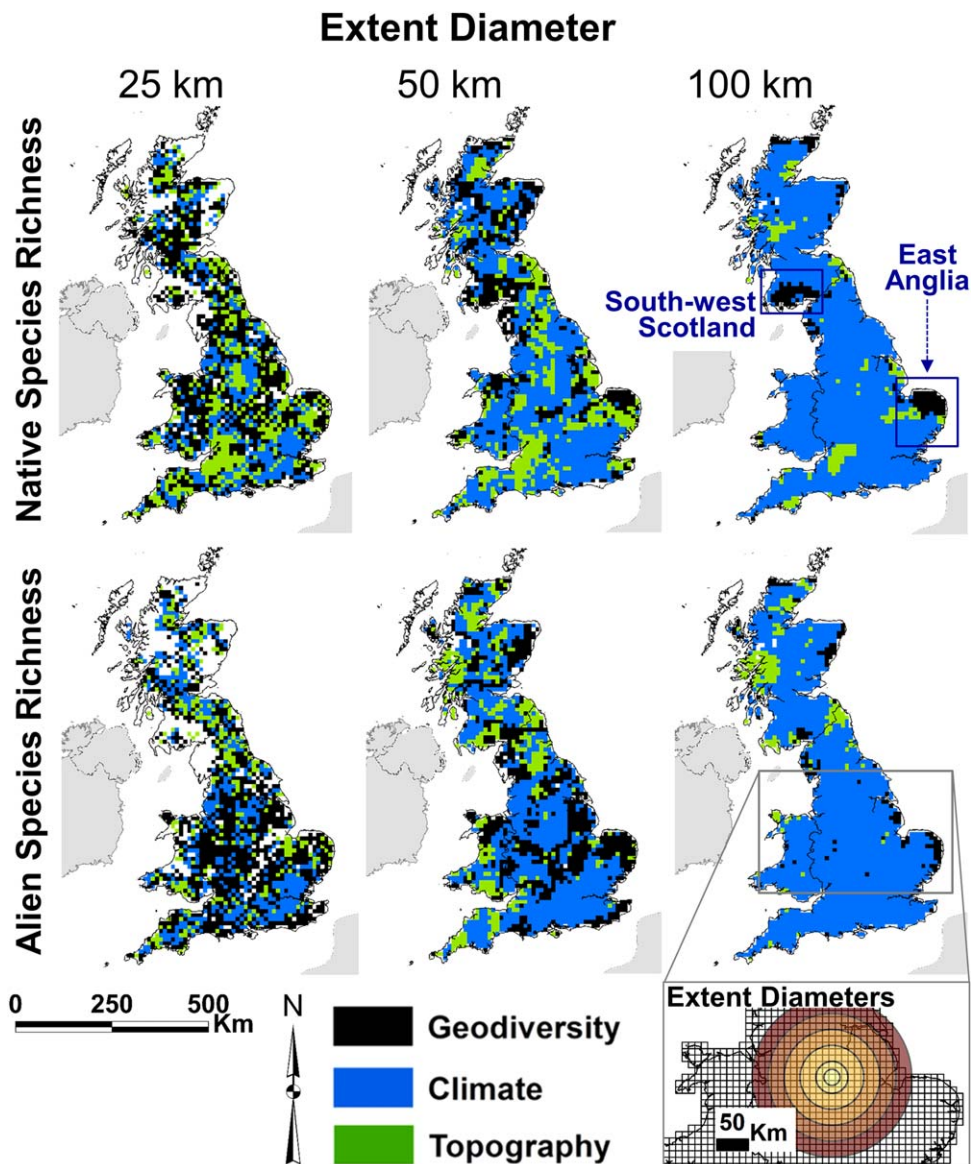
We used `gbm.step` ('gbm 2.1.1' package in R; Ridgeway, 2015) to implement BRT. This function controls the number of terms in order to produce parsimonious models. To quantify modelled effects of individual explanatory variables, the contribution (relative influence) of each predictor was obtained from `gbm.step`. These are scaled to add to 100, where '100' for a predictor would mean that it was the sole contributor to the final model. Where the model contribution reflected a negative relationship with species richness, we then made the value negative for display purposes. Combined model contributions were calculated for the predictor sets and subsets defined in Table 1. We used a tree complexity of 3 (allowing up to three-way interactions; Elith et al., 2008), a bag fraction of 0.5 and a preferred learning rate of 0.05, which was occasionally reduced to 0.01, 0.005 and then 0.001 according to data requirements. Predictors contributing <10% (or sometimes <7.5%) were removed from the initial model, which was rerun with the simplified predictor set to produce the final results (further details are given in Appendix S4).

As well as evaluation using internal fit statistics ('self-statistics'), models were validated using 10-fold cross-validation (CV) in the 'gbm' package. This approach randomly subsamples the data 10 times according to the user-defined bag fraction; our bag fraction was 0.5, so

each time 50% of the data were used to parameterize the model and the other 50% to evaluate it. The final cross-validation correlation statistic is the mean correlation between the training and testing data across 10 runs. Model statistics were compared with and without GDCs using paired-samples t-tests.

### 3 | RESULTS

Geodiversity components (GDCs) made the largest contributions to models at the smallest study extent and smallest grain size (in the geodiversity column of Figure 3, the left-hand blue boxplot is the highest). At this scale, geodiversity was the strongest of all the predictor sets (of all the left-hand blue boxplots in Figure 3, those for geodiversity are the highest). With each increase in extent, the modelled contribution of geodiversity declined substantially relative to the other types of variable. GDCs were not relevant at the larger extents, giving way particularly to climate and human population. Climate was more important for archaeophytes than neophytes. The contribution of 'topography' (the coarse variables typically used in modelling species richness patterns) showed similar patterns to geodiversity, but was less important at smaller scales and declined less sharply as scales increased. Mapping the results (Figure 2) shows the widespread dominance of the geodiversity predictor set at the smaller geographical extents, its importance generally declining relative to climate with increasing extent, except in



**FIGURE 2** The dominant predictor set for native (top row) and alien (bottom row) species richness at the 1-km<sup>2</sup> grain size for three spatial extents. White spaces are where the quantity of data was insufficient to run a reliable model or cells were excluded as they were undersampled. An example of the six extent diameters (25, 50, 100, 150, 200 and 250 km) is shown in the bottom-right, in this case for British National Grid cell SK54, which is one of 2,121 cells around which species richness was analysed at the two grain sizes and six extents

northern East Anglia and south-west Scotland where it persisted for natives. At the national extent, climatic variables were dominant (Table 2), particularly annual mean temperature; human population was also often important, especially for species richness of neophytes.

The contribution of geodiversity to biodiversity models was dominated by landform data from geomorphometry, but hydrology (rivers and lakes), and to a lesser extent materials (soil, superficial deposits and geology), were also important (Figure 3, Appendix S5). Most trends in specific GDCs followed that of geodiversity generally (i.e., declining model contribution with increasing scale). Model contributions from GDCs mostly resulted from positive relationships with biodiversity (Figure 4, Appendix S6), but some relationships were negative. No GDCs were consistently negatively related to species richness. The highest

positive model contributions for natives and aliens came from river length, valley coverage and lake area. Most GDCs were most strongly related to species richness at the smallest extent; the main exceptions were land surface materials, which, for native species, had the highest contributions at the largest extent (Figure 4).

Across models accounting for species richness at all scales, interactions between GDCs and other predictors tended to be uncommon (Appendix S7), with the possible exception of precipitation and hollows for archaeophyte richness. The main interactions varied with scale, but those between climate and topography, and between various climatic variables, frequently tended to be dominant. Also, climate and mean elevation often interacted with human population, especially when modelling alien species richness. Collinearities between GDCs and

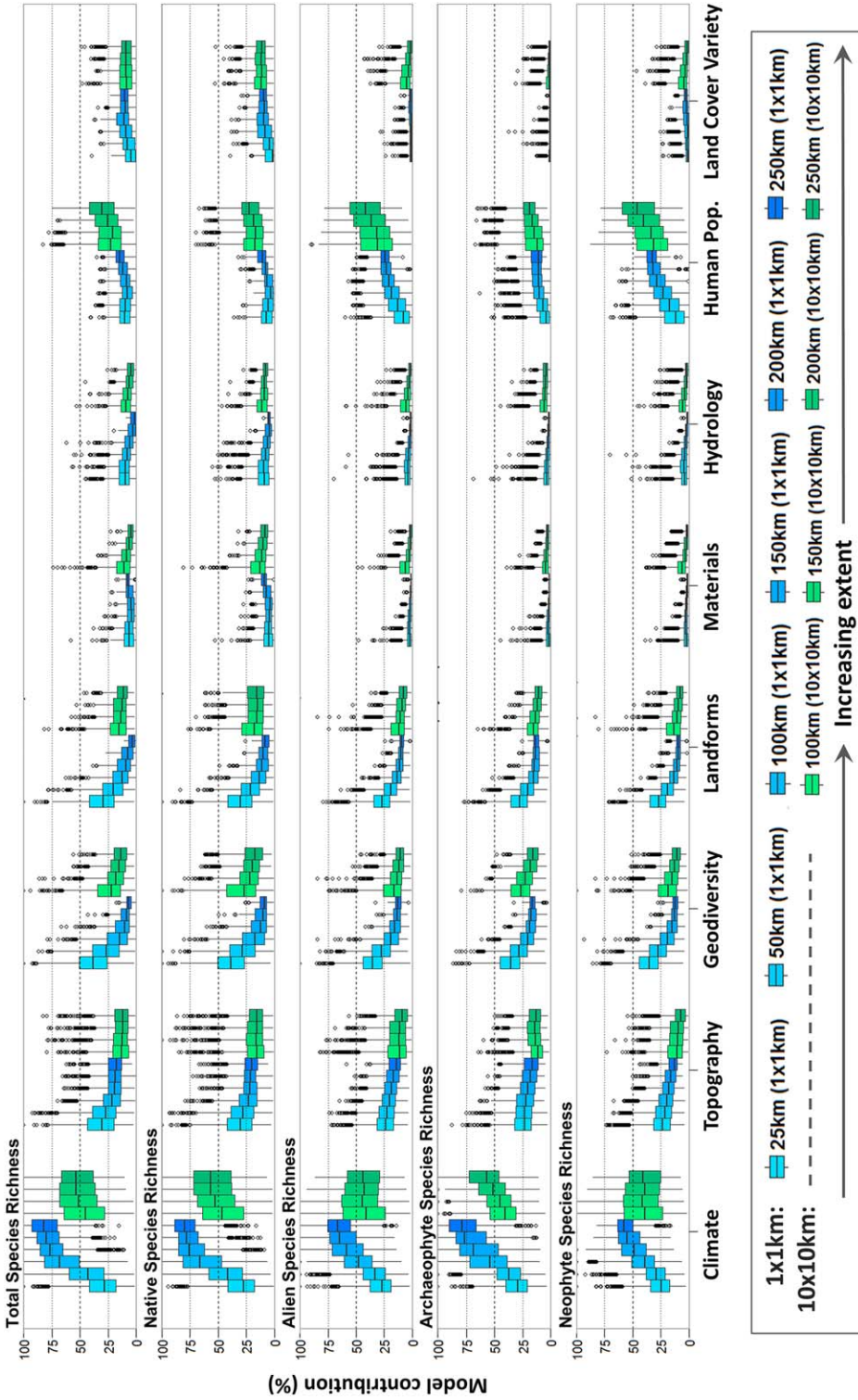


FIGURE 3 Combined model contributions (%) of each predictor set (climate, topography, geodiversity) and geodiversity, as well as isolated predictors human population and land-cover variety. All scales are shown (except national); grain sizes (1 km × 1 km cells in blue, 10 km × 10 km in green) and extent diameters [lighter (further left) = smaller]. A high-resolution version of this figure is included in Appendix S5. This plot shows only model contributions of each predictor set, and gives no indication of directionality. 'Pop', population



TABLE 2 National extent results

| Grain size    | Species group | CV    | SS    | Dominant predictor (% model contribution) | Second highest predictor (% model contribution) | Third highest predictor (% model contribution) |
|---------------|---------------|-------|-------|---|---|--|
| 1 km × 1 km   | All           | 0.376 | 0.380 | Annual mean temperature (27%)             | Precipitation seasonality (21%)                 | Min. temperature coldest month (15%)           |
|               | Native        | 0.366 | 0.369 | Min. temperature coldest month (22%)      | Annual mean temperature (20%)                   | Precipitation seasonality (16%)                |
|               | Alien         | 0.567 | 0.601 | Human population (22%)                    | Annual mean temperature (17%)                   | Precipitation seasonality (14%)                |
|               | Arch.         | 0.541 | 0.569 | Annual mean temperature (28%)             | Mean diurnal range (16%)                        | Annual precipitation (14%)                     |
|               | Neo.          | 0.564 | 0.608 | Human population (30%)                    | Precipitation seasonality (13%)                 | Annual precipitation (11%)                     |
| 10 km × 10 km | All           | 0.717 | 0.737 | Annual mean temperature (43%)             | Human population (37%)                          | Mean diurnal range (6%)                        |
|               | Native        | 0.659 | 0.696 | Annual mean temperature (34%)             | Human population (27%)                          | Mean diurnal range (10%)                       |
|               | Alien         | 0.815 | 0.829 | Human population (45%)                    | Annual mean temperature (40%)                   | Mean diurnal range (7%)                        |
|               | Arch.         | 0.892 | 0.969 | Annual mean temperature (43%)             | Human population (13%)                          | Annual precipitation (10%)                     |
|               | Neo.          | 0.788 | 0.808 | Human population (60%)                    | Annual mean temperature (24%)                   | Mean diurnal range (8%)                        |

Numbers show the combined model contributions (rounded to whole numbers) for each predictor set. Model evaluation (mean cross-validation correlation, CV) and fit statistics (self-statistics, SS) are also presented. Arch = archaeophytes; Neo = neophytes.

topography and climate varied greatly across scales and between predictors (Appendices S8 and S9). Most GDCs were only weakly collinear with topography and climate (e.g., hydrology, rock variety, coverage of hollows, slopes and valleys) and others more strongly (e.g., coverage of peaks, ridges and spurs was moderately related to higher, cooler places at the coarser grain size and nationally), but these collinearities were

still often much weaker than those between and within climate and topography predictors.

Self-statistics and cross-validation statistics were consistently higher (indicating better models) for larger extents, the coarser grain size of 100 km<sup>2</sup> and alien species richness (Figure 5, Appendices S10 and S11). Adding geodiversity often, but not always, resulted in

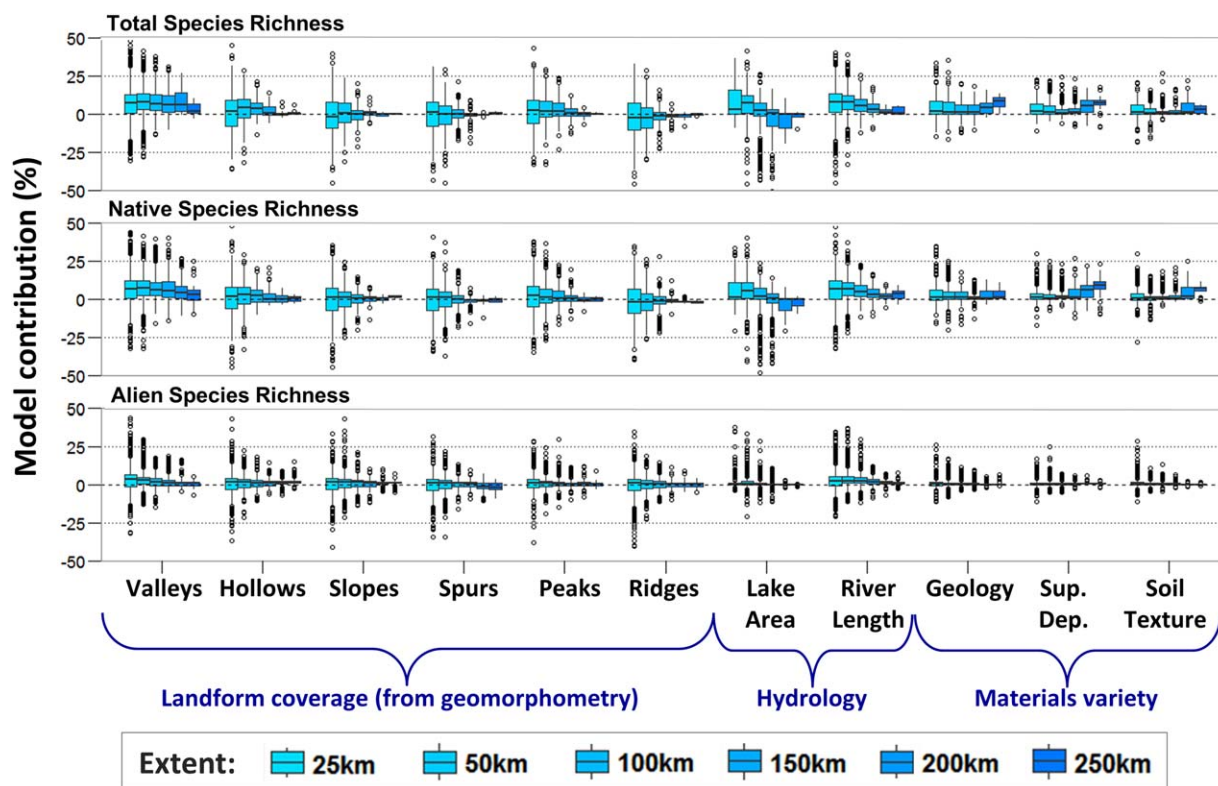
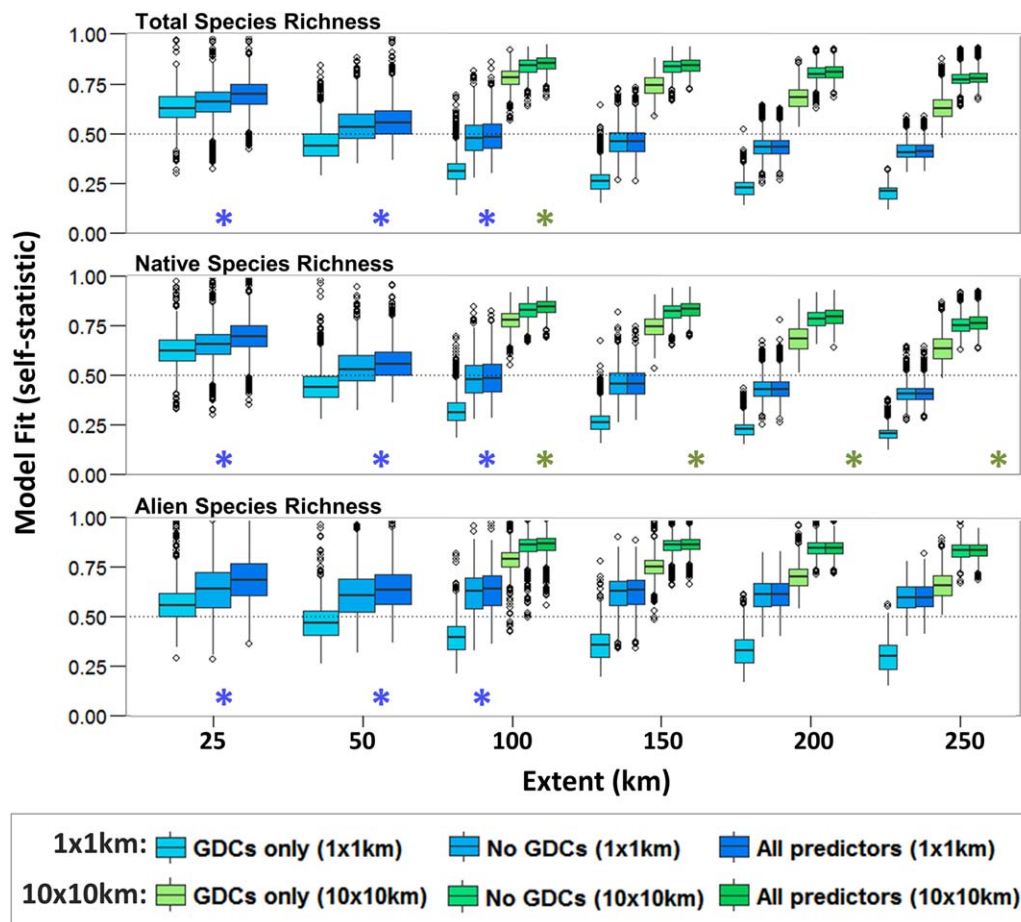


FIGURE 4 Model contributions from individual geodiversity components at the 1 km × 1 km grain size for each extent. These graphs are truncated at +50% and −50%, but only a small minority of points lie beyond these values. A full version of this figure with all species groups is included in Appendix S6. 'Sup. Dep.', superficial deposits



**FIGURE 5** Comparison of model fit statistics (self-statistics, SS) with and without geodiversity according to grain size (colour/shade) and extent (x axis). \*Significant ( $p < .01$ , paired  $t$  test) average model improvement across all models when comparing those without geodiversity (the middle boxplot for each scale) with those with all predictors (the right boxplot). Values are given in Table 3. Archaeophyte and neophyte results can be seen alongside these in Appendix S10, and an equivalent graph for cross-validation statistics is also provided (Appendix S11). GDC, geodiversity component

significantly better models, especially at the smaller extents and for native species richness for both grain sizes (Table 3). Results for total species richness broadly followed those for native species richness (Figure 3), despite the presence of many uncategorized species in the overall richness data. Results for alien species richness tended to follow those for archaeophytes, even though there were relatively few archaeophyte species.

## 4 | DISCUSSION

Geodiversity made a significant addition to models of vascular plant species richness over and above widely used topographic metrics, particularly at smaller geographical extents ( $H_1$ ). At the smallest extent, geodiversity contributed more than any other type of predictor accounting for species richness, while at larger extents climatic variables became increasingly dominant. With respect to individual geodiversity components (GDCs), automatically extracted landform data were of particular explanatory value, demonstrating that species richness–landform relationships can be detected at macroecological scales. These data represent a novel predictor set in macroecology and are rel-

atively easily extracted from widely available DEMs. Our analyses also highlighted the importance of separately analysing individual GDCs rather than lumping them into a single variable for use in biodiversity modelling, as done in most of the limited research to date. Results were broadly similar for alien and native species richness patterns ( $H_2$ ), except that neophytes were more strongly related to human population than were the other plant groups. Results for total species richness were very similar to those for native species richness, despite the presence of many uncategorized species in the overall richness data.

Geodiversity therefore succeeded in capturing unique dimensions of environmental heterogeneity that have theoretical mechanistic links to species richness, and which add explanatory power when modelling species richness patterns of vascular plants (Stein et al., 2014). This is consistent with our first hypothesis ( $H_1$ ), which was based on theorized links between biodiversity and the presence and diversity of both landforms and surface materials – reflecting the presence of more resources and greater habitat and niche variety (Anderson & Ferree, 2010; Hjort et al., 2012; Lawler et al., 2015; Moser et al. 2005), and possibly the results of some disturbance processes (le Roux et al., 2013). Also consistent with  $H_1$  was the decline in magnitude of the contribution of

**TABLE 3** Mean difference in model fit (self-statistics, SS) and evaluation (cross-validation, CV) statistics (also see Figure 5 and Appendices S10 and S11) between models with and without geodiversity (i.e., a positive value indicates an increase in model performance after geodiversity was added)

| Grain size    | Species group   | 25 km           |                 | 50 km           |                 | 100 km          |                 | 150 km          |        |
|---------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--------|
|               |                 | SS              | CV              | SS              | CV              | SS              | CV              | SS              | CV     |
| 1 km × 1 km   | All             | 0.039           | 0.018           | 0.022           | 0.007           | 0.008           | 0.004           | 0.001           | 0.001  |
|               |                 | <i>n</i> = 1775 |                 | <i>n</i> = 2047 |                 | <i>n</i> = 2087 |                 | <i>n</i> = 2091 |        |
|               | Native          | 0.041           | 0.015           | 0.023           | 0.007           | 0.008           | 0.003           | 0.002           | 0.001  |
|               |                 | <i>n</i> = 1730 |                 | <i>n</i> = 2034 |                 | <i>n</i> = 2087 |                 | <i>n</i> = 2091 |        |
|               | Alien           | 0.042           | -0.006          | 0.029           | -0.004          | 0.012           | -0.001          | 0.005           | 0.001  |
|               | <i>n</i> = 1574 |                 | <i>n</i> = 2044 |                 | <i>n</i> = 2089 |                 | <i>n</i> = 2091 |                 |        |
|               | Arch.           | 0.039           | -0.011          | 0.028           | -0.003          | 0.014           | 0.000           | 0.007           | 0.001  |
|               |                 | <i>n</i> = 1325 |                 | <i>n</i> = 1846 |                 | <i>n</i> = 2057 |                 | <i>n</i> = 2084 |        |
|               | Neo.            | 0.038           | -0.006          | 0.028           | -0.003          | 0.011           | -0.002          | -0.001          | -0.005 |
|               |                 | <i>n</i> = 1588 |                 | <i>n</i> = 2042 |                 | <i>n</i> = 2089 |                 | <i>n</i> = 2091 |        |
| 10 km × 10 km | All             |                 |                 |                 |                 | 0.014           | 0.010           | 0.008           | 0.005  |
|               |                 |                 |                 |                 |                 | <i>n</i> = 1697 |                 | <i>n</i> = 2037 |        |
|               | Native          |                 |                 |                 |                 | 0.018           | 0.025           | 0.014           | 0.013  |
|               |                 |                 |                 |                 |                 | <i>n</i> = 1697 |                 | <i>n</i> = 2037 |        |
|               | Alien           |                 |                 |                 |                 | 0.006           | -0.001          | 0.003           | -0.004 |
|               |                 |                 |                 |                 | <i>n</i> = 1692 |                 | <i>n</i> = 2037 |                 |        |
|               | Arch.           |                 |                 |                 |                 | 0.021           | -0.004          | 0.013           | -0.003 |
|               |                 |                 |                 |                 |                 | <i>n</i> = 1688 |                 | <i>n</i> = 2036 |        |
|               | Neo.            |                 |                 |                 |                 | 0.008           | -0.003          | 0.003           | -0.004 |
|               |                 |                 |                 |                 |                 | <i>n</i> = 1691 |                 | <i>n</i> = 2037 |        |

Shaded cells indicate a significant average improvement ( $p < .01$ , paired *t* test) in SS or CV across all models for that scale of at least 0.01. The 200-km and 250-km extents are not shown but continue the pattern of declining values in difference and worsening *p* value. The number of models compared is also shown. Arch = Archaeophyte, Neo = Neophyte

geodiversity with increasing extent, at both grain sizes, as other variables (particularly broad-scale climate) took over. Geodiversity therefore seems to provide a predictor set that can account for the variety of the abiotic environment at these finer extents ('landscape' scale) where broad-scale climate is more constant. At these scales, geodiversity data may be strongly related to microclimate and localized hydrological, edaphic and geological conditions that are relevant to the establishment and persistence of species.

Theoretically, variables measuring fine-resolution environmental heterogeneity may contribute relatively little to models of species richness using large grain sizes because of the tendency for the heterogeneity to average out within grid cells (Field et al., 2009). If so, GDCs such as those measuring landforms should have reduced explanatory power at larger grain sizes, when extent is held constant, while climate- and productivity-related variables may increase. However, for the 100-, 150-, 200- and 250-km geographical extents (for which both grain sizes were assessed), we observed similar geodiversity results for each grain size, often with slightly higher relative geodiversity contributions at the 100-km<sup>2</sup> grain than 1 km<sup>2</sup>. This suggests that the size (extent) of the study area more strongly affects the relative contribution of geodiversity as a biodiversity predictor than does grain size. This may be because the heterogeneity measured by GDCs is correlated with broader environmental gradients, so the averaging of fine-scale variation at larger grains does not affect the explanatory power of GDCs much compared with the large increases in the degree to which broad climatic and topographic gradients are captured at larger geographical extents (Hawkins et al., 2003). Further research is required on this question.

The general lack of strong and frequent collinearities and interactions between most GDCs and both topography and climate in our models suggests largely unique model contributions from geodiversity variables. While crude topographic variables such as mean elevation, elevational range and mean slope can provide useful information, as they did here and in much previous research (e.g., Field et al., 2009; Hjort et al., 2012), these variables are typically strongly collinear with each other and with climate (e.g., Ferrer-Castán & Vetaas, 2005; Koppel, Gillespie, Ormerod, & Fricker, 2016; Appendices S8 and S9). More detailed analyses are needed of how GDCs correlate with other predictors in different places and at different scales. However, if the use of GDCs results in greatly reduced multicollinearity problems compared with the use of crude topographic variables, then our ability to determine cause and effect should be improved; this is consistent with the notion that GDCs relate more directly to mechanisms than do crude topographic variables (Gray, 2013; and see the Introduction). That is, explicit consideration of landscape features in biodiversity modelling may enhance ecological understanding (Hjort et al., 2015), and is also likely to be highly relevant to the modelling of individual species' distributions.

Specific GDCs were important in the species richness models, consistent with the notion that species richness-GDC relationships can be detected at macroecological scales, and add to biodiversity models. These results were far more informative than using a compound measure of geodiversity. For example, we observed some negative relationships between biodiversity and various GDCs (Figure 4), while valley coverage, river length and surface materials had more consistently

positive model contributions. Modelled interactions between valley coverage and river length were neither frequent nor strong, and it is likely that the valley landform data detected smaller geofeatures (e.g., different erosional and depositional features produced by geomorphological processes) that are ecologically important (Hjort et al., 2015) but not represented in the relatively coarse river maps that are generally available and used in this study. Knowledge of surface (soil and superficial deposits) and subsurface (geology) material richness was less useful than expected from previous research (e.g., Anderson & Ferree, 2010; Tukiainen et al., 2016). Perhaps an explicit consideration of the coverage of specific types of rock and soil (and mineralogy more generally) would be revealing. Further research on this would help us to better understand the links between specific GDCs and biodiversity.

The relative contributions of different predictors to alien and native species richness models showed broadly similar patterns across scales, but the magnitudes varied somewhat. The contributions from GDCs, particularly landforms, were greater for native species richness than alien, and native biodiversity models were also most improved by the addition of GDCs (partly consistent with  $H_2$ ). Contributions from GDCs might therefore particularly add important information to native biodiversity models, which sometimes underperform compared with alien species richness models (Deutschewitz et al., 2003; Kumar et al., 2006). This finding is supported by the relative importance of geodiversity in explaining native richness compared with total richness models seen elsewhere (Räsänen et al., 2016).

The main difference between models of alien and native plant species richness lay in the contribution from human population, which was highest for neophytes, then archaeophytes and relatively low for natives. While the relationship between alien species richness and cities or human populations has been known for some time (e.g., Deutschewitz et al., 2003; Kühn, Brandl, & Klotz, 2004; McKinney, 2008; Pyšek, 1998), our results suggest that this relationship is more pronounced for neophytes and that the strength of this relationship is affected by scale (both extent and grain size). In line with known links between riverine habitats and neophyte species richness (Deutschewitz et al., 2003), we observed a substantial contribution of river length to neophyte richness models, and frequent interactions between river length and human population, implying increased human influence along rivers, which in turn may promote neophyte species richness. However, there was no notable relationship between native richness and river length, in contrast to findings elsewhere (Deutschewitz et al., 2003). Other aspects of hydrology (including lake area) tended to be less important in our models than landforms, topography, human population and climate. Overall, transport, for which human population provides a proxy, and river networks (and their interaction) could be promoting the dispersal of alien species and thus increasing neophyte species richness (Hulme, 2009; Pyšek et al., 2010).

#### 4.1 | Geodiversity in biodiversity science and conservation: Opportunities and challenges

We have used geodiversity to try to provide more explicit representation of environmental heterogeneity than crude topographic variables.

Indeed, GDCs more directly measure environmental conditions and processes, such as habitat diversity, resource gradients and microclimatic and sheltering effects (Hjort et al., 2015; Matthews, 2014), thus enabling us to more precisely capture the causal processes behind the biodiversity patterns. With this in mind, geodiversity may have benefits beyond species richness modelling – in species distribution modelling, for example. Investigating where geodiversity is most relevant to patterns of life globally requires research on geodiversity in geographical domains beyond our study area. It has also been suggested that such information might be important in the context of refugia by identifying parts of the landscape that can withstand long-term environmental change by providing stable microclimates (Keppel et al., 2015; Lawler et al., 2015); and geodiverse locations may facilitate the adaptation of species to climate change, as well as their persistence (Albano, 2015; Maclean, Hopkins, Bennie, Lawson, & Wilson, 2015).

There are also practical reasons for improving our understanding of the relationship between geodiversity and biodiversity; the first relates to conservation. The idea known as 'conserving nature's stage' is gaining momentum (see Anderson & Ferree, 2010; Beier, Hunter, & Anderson, 2015b). This suggests that instead of targeting individual species and habitats for conservation we target areas capable of supporting high biodiversity under future environmental changes, by either better maintaining the existing environment or by providing greater environmental heterogeneity. Furthermore, geodiversity has been related to the diversity of threatened species richness and rarity-weighted richness, at least in high latitudes (Tukiainen et al., 2016). Indeed, the geodiversity data used in our study are likely to enhance our understanding of previously demonstrated links between abiotic diversity and site complementarity, which may be a measure of biodiversity that is more relevant to conservation than raw species richness (Albuquerque & Beier, 2015; Beier & Albuquerque, 2015). Areas high in geodiversity are also thought to promote greater resilience to climatic change for biodiversity and for essential provisioning, regulating, cultural and supporting ecosystem services (Brazier, Bruneau, Gordon, & Rennie, 2015; Gordon & Barron, 2013). However, we stress the importance of not overlooking the value of individual species in 'geohomogeneous' (low geodiversity) places or unique geofeatures in species-poor areas. For example, an endangered species in a forest underlain by a single geology and few landforms should not be overlooked, whilst unique and irreplaceable geofeatures (e.g., certain fossils or rare geological units and mineralogy) will not always be relevant to biodiversity and present-day species distributions, but have geoheritage value.

Another practical advantage of geodiversity data is that they are usually cheaper and faster to collect and collate than species occurrence data (Hjort & Luoto, 2010) and, in areas where geodiversity correlates very strongly with biodiversity (e.g., northern Finland; Hjort et al., 2012), geodiversity may represent a useful surrogate for biodiversity. We have compiled a lengthy (but not exhaustive) table of GDC categories, including geology, soil, landforms and hydrology in Appendix S12 and a list of the remotely sensed datasets required, which may be a useful resource for reference. Importantly, a key dataset in our research was the automatically extracted landform data, which only

required open-source GIS software (GRASS), a freely available algorithm within that software (r.geomorphon; Jasiewicz & Stepinski, 2013) and a DEM. Datasets related to land surface materials are less accessible in much of the world, but high-resolution datasets of ecologically relevant soil variables have recently become widely available (Hengl et al., 2014; www.isric.org/content/soilgrids). Other geodiversity data not used in our study might improve models further, including explicit data on pH and mineralogy, for example, or topographic wetness or insolation. Additional sources of geodiversity data, appropriate for smaller study areas than in our study, include those captured on airborne platforms: Hjort & Luoto (2010) used aerial photography combined with field surveys. It would be interesting to know whether such intensive data add explanatory power and ecological meaning to data obtained automatically from geomorphometry. Future sources of such data could include capture by unmanned aerial vehicles (UAVs), release of archival data and increased data sharing on capture (Lowman & Voirin, 2016).

In conclusion, we have shown that geodiversity can add significantly to models of species richness of vascular plants over and above the widely used topographic metrics in our study area. We found some differences in the response of alien and native species richness to geodiversity; further research on this may be beneficial for conservation and management. Our findings demonstrate the largely unexploited potential of explicit geodiversity data, which may aid explanation by more directly measuring causal factors and reducing multicollinearity of explanatory variables. Research on the role of geodiversity across a variety of taxonomic and geographical domains is still in its infancy, and we have pointed to some research needs. Our finding that automatically extracted landform data were valuable should encourage collaboration between geomorphologists, ecologists and biogeographers.

## ACKNOWLEDGMENTS

This research was supported by the U.K. Natural Environment Research Council (NERC) PhD Studentship 1365737, which was awarded to J.J.B., University of Nottingham, in October 2013 (supervised by R.F. and D.B.). J.H. acknowledges the Academy of Finland (project number 285040). We thank two anonymous referees and the editors for their helpful comments on the manuscript. We would like to highlight that this research uses a large amount of data collected and collated by thousands of people over many decades and across a number of organizations. We therefore wish to offer our deepest thanks to the volunteers and coordinators, present and past, of the Botanical Society of Britain and Ireland (BSBI). Secondly, we thank BSBI staff, Tom Humphrey, Alex Lockton, and Kevin Walker, for their prodigious cooperation regarding access to the BSBI's Distribution Database. Our thanks also go to the British Geological Survey (BGS) for the provision of the 1:50,000 geology, superficial deposit and soil shapefiles, which have been constructed over many decades by BGS surveyors. Finally, we wish to thank the GRASS GIS developers for their help with some technical issues early on.

## AUTHOR CONTRIBUTIONS

All authors contributed to the development of ideas for this research. J.J.B. gathered and processed the data and conducted the analyses. J.J.B., R.F. and D.B. wrote the manuscript, with input from J.H. and C.L.

## REFERENCES

- Albano, C. M. (2015). Identification of geophysically diverse locations that may facilitate species' persistence and adaptation to climate change in the southwestern United States. *Landscape Ecology*, *30*, 1023–1037.
- Albuquerque, F. S., & Beier, P. (2015). Global patterns and environmental correlates of high-priority conservation areas for vertebrates. *Journal of Biogeography*, *42*, 1397–1405.
- Anderson, M. G., & Ferree, C. E. (2010). Conserving the stage: Climate change and the geophysical underpinnings of species diversity. *PLoS One*, *5*, e11554.
- Barthlott, W., Hostert, A., Kier, G., Küper, W., Kreft, H., Mutke, J. ... Sommer, J. H. (2007). Geographic patterns of vascular plant diversity at global to continental scales. *Erdkunde*, *61*, 305–315.
- Beier, P., & Albuquerque, F. S. (2015). Environmental diversity as a surrogate for species representation. *Conservation Biology*, *29*, 1401–1410.
- Beier, P., Sutcliffe, P., Hjort, J., Faith, D. P., Pressey, R. L., & Albuquerque, F. S. (2015a). A review of selection-based tests of abiotic surrogates for species representation. *Conservation Biology*, *29*, 668–679.
- Beier, P., Hunter, M. L., & Anderson, M. (2015b). Special section: Conserving nature's stage. *Conservation Biology*, *29*, 613–617.
- Bétard, F. (2013). Patch-Scale Relationships Between Geodiversity and Biodiversity in Hard Rock Quarries: Case Study from a Disused Quartzite Quarry in NW France. *Geoh Heritage*, *5*, 59–71.
- Brazier, V., Bruneau, P. M. C., Gordon, J. E., & Rennie, A. F. (2015). Making space for nature in a changing climate: The role of geodiversity in biodiversity conservation. *Scottish Geographical Journal*, *128*, 211–233.
- Burnett, M., August, P., Brown, J., & Killingbeck, K. (1998). The influence of geomorphological heterogeneity on biodiversity. A patch-scale perspective. *Conservation Biology*, *12*, 363–370.
- Celesti-Grapow, L., Pyšek, P., Jarošík, V., & Blasi, C. (2006). Determinants of native and alien species richness in the urban flora of Rome. *Diversity and Distributions*, *12*, 490–501.
- Deuschewitz, K., Lausch, A., Kühn, I., & Klotz, S. (2003). Native and alien plant species richness in relation to spatial heterogeneity on a regional scale in Germany. *Global Ecology and Biogeography*, *12*, 299–311.
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G. ... Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 027–046.
- Dufour, A., Gadallah, F., Wagner, H. H., Guisan, A., & Buttler, A. (2006). Plant species richness and environmental heterogeneity in a mountain landscape: Effects of variability and spatial configuration. *Ecography*, *29*, 573–584.
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*, 802–813.
- Ferrer-Castán, D., & Vetaas, O.R. (2005). Pteridophyte richness, climate and topography in the Iberian Peninsula: Comparing spatial and non-spatial models of richness patterns. *Global Ecology and Biogeography*, *14*, 155–165.

- Field, R., Hawkins, B. A., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guégan, J., ... Turner, J. R. G. (2009). Spatial species-richness gradients across scales: A meta-analysis. *Journal of Biogeography*, *36*, 132–147.
- Fleishman, E., Murphy, D. D., & Sada, D. W. (2006). Effects of environmental heterogeneity and disturbance on the native and non-native flora of desert springs. *Biological Invasions*, *8*, 1091–1101.
- GRASS Development Team (2016). *Geographic Resources Analysis Support System (GRASS) Software*, Version 7.1. Open Source Geospatial Foundation. Retrieved from <http://grass.osgeo.org>. Accessed December, 2016.
- Gray, M. (2013). *Geodiversity: valuing and conserving abiotic nature* (2nd ed.). Chichester: Wiley-Blackwell.
- Gordon, J. E., & Barron, H. F. (2013). The role of geodiversity in delivering ecosystem services and benefits in Scotland. *Scottish Journal of Geology*, *49*, 41–58.
- Guitet, S., Pélissier, R., Brunaux, O., Jaouen, G., & Sabatier, D. (2014). Geomorphological landscape features explain floristic patterns in French Guiana rainforest. *Biodiversity Conservation*, *24*, 1215–1237.
- Hanski, I., von Hertzen, L., Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., & Vartiainen, E. (2012). Environmental biodiversity, human microbiota, and allergy are interrelated. *Proceedings of the National Academy of Sciences USA*, *109*, 8334–8339.
- Hawkins, B. A., Field, R., Cornell, H. V., Currie, D. J., Guégan, J., Kaufman, D., ... Turner, J. R. G. (2003). Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, *84*, 3105–3117.
- Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., & Gonzalez, M. R. (2014). SoilGrids1km – global soil information based on automated mapping. *PLoS One*, *9*, 1–17.
- 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*, 1965–1968.
- Hjort, J., & Luoto, M. (2010). Geodiversity of high-latitude landscapes in northern Finland. *Geomorphology*, *115*, 109–116.
- Hjort, J., Heikkinen, R. K., & Luoto, M. (2012). Inclusion of explicit measures of geodiversity improve biodiversity models in a boreal landscape. *Biodiversity Conservation*, *21*, 3487–3506.
- Hjort, J., Gordon, J. E., Gray, M., & Hunter, M. L. Jr (2015). Why geodiversity matters in valuing nature's stage. *Conservation Biology*, *29*, 630–639.
- Hulme, P. E. (2009). Trade, transport and trouble: Managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, *46*, 10–18.
- Jasiewicz, J., & Stepinski, T. F. (2013). Geomorphons – a pattern recognition approach to classification and mapping of landforms. *Geomorphology*, *182*, 147–156.
- Keppel, G., Gillespie, T. W., Ormerod, P., & Fricker, G. A. (2016). Habitat diversity predicts orchid diversity in the tropical south-west Pacific. *Journal of Biogeography*, *43*, 2332–2342.
- Keppel, G., Mokany, K., Wardell-Johnson, G. W., Phillips, B. L., Welberge, J. A., & Reside, A. E. (2015). The capacity of refugia for conservation planning under climate change. *Frontiers in Ecology and the Environment*, *13*, 106–112. doi: 10.1111/jbi.12805
- Kougioumoutzis, K., & Tiniakou, A. (2014). Ecological factors driving plant species diversity in the South Aegean Volcanic Arc and other central Aegean islands. *Plant Ecology and Diversity*, *8*, 173–186.
- Kühn, I., Brandl, R., & Klotz, S. (2004). The flora of German cities is naturally species rich. *Evolutionary Ecology Research*, *6*, 749–764.
- Kumar, S., Stohlgren, T. J., & Chong, G.W. (2006). Spatial heterogeneity influences native and nonnative plant species richness. *Ecology*, *87*, 3186–3199.
- Lawler, J. J., Ackerly, D. D., Albano, C. M., Anderson, M. G., Dobrowski, S. Z., Gill, J. L., ... Weiss, S. B. (2015). The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. *Conservation Biology*, *29*, 618–629.
- Le Roux, P. C., Virtanen, R., & Luoto, M. (2013). Geomorphological disturbance is necessary for predicting fine-scale species distributions. *Ecography*, *36*, 800–808.
- Lomolino, M. V., Riddle, B. R., Whittaker, R. J., & Brown, H. J. (2010). *Biogeography (4th ed.)*. Sunderland, MA: Sinauer Associates.
- Lowman, M., & Voirin, B. (2016). Drones – our eyes on the environment. *Frontiers in Ecology and the Environment*, *14*, 231.
- Macleán, I. M. D., Hopkins, J. J., Bennie, J., Lawson, C. R., & Wilson, R. J. (2015). Microclimates buffer the responses of plant communities to climate change. *Global Ecology and Biogeography*, *24*, 1340–1350.
- Matthews, T. J. (2014). Integrating geoconservation and biodiversity conservation: Theoretical foundations and conservation recommendations in a European Union context. *Geoheritage*, *6*, 57–70.
- McGill, B. J. (2010). Matters of scale. *Science*, *328*, 575.
- McKinney, M. L. (2008). Urbanisation as a major cause of biotic homogenisation. *Biological Conservation*, *127*, 247–260.
- Mittelbach, G. G., Steiner, C. F., Scheiner, S. M., Gross, K. L., Reynolds, H. L., Waide, R. B., ... Gough, L. (2001). What is the observed relationship between species richness and productivity? *Ecology*, *82*, 2381–2396.
- Moser, D., Dullinger, S., Englisch, T., Niklfeld, H., Plutzer, C., Sauberer, N., ... Grabherr, G. (2005). Environmental determinants of vascular plant species richness in the Austrian Alps. *Journal of Biogeography*, *32*, 1117–1127.
- O'Brien, E. M., Field, R., & Whittaker, R. J. (2000). Climatic gradients in woody plant (tree and shrub) diversity: Water–energy dynamics, residual variation, and topography. *Oikos*, *89*, 588–600.
- Parks, K. E., & Mulligan, M. (2010). On the relationship between a resource based measure of geodiversity and broad scale biodiversity patterns. *Biodiversity Conservation*, *19*, 2751–2766.
- Pausas, J. G., Carreras, J., Ferre, A., & Font, X. (2003). Coarse-scale plant species richness in relation to environmental heterogeneity. *Journal of Vegetation Science*, *14*, 661–668.
- Pellitero, R., Manosso, F. C., & Serrano, E. (2015). Mid-and large-scale geodiversity calculation in Fuentes Carrionas (NW Spain) and Serra do Cadeado (Paraná, Brazil): Methodology and application for land management. *Geografiska Annaler*, *97A*, 219–235.
- Pyšek, P. (1998). Alien and native species in Central European urban floras: A quantitative comparison. *Journal of Biogeography*, *25*, 155–163.
- Pyšek, P., Jarošík, V., Chytrý, M., Kropáč, Z., Tichý, L., & Wild, J. (2005). Alien plants in temperate weed communities: Prehistoric and recent invaders occupy different habitats. *Ecology*, *86*, 772–785.
- Pyšek, P., Jarošík, V., Hulme, P. E., Kühn, I., Wild, J., Arianoutsou, M., ... Winter, M. (2010). Disentangling the role of environmental and human pressures on biological invasions across Europe. *Proceedings of the National Academy of Sciences USA*, *107*, 12157–12162.
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>. Accessed September, 2016.
- Randin, C. F., Vuissoz, G., Liston, G. E., Vittoz, P., & Guisan, A. (2009). Introduction of snow and geomorphic disturbance variables into predictive models of alpine plant distribution in the Western Swiss Alps. *Arctic, Antarctic, and Alpine Research*, *41*, 347–361.
- Räsänen, A., Kuitunen, M., Hjort, J., Vaso, A., Kuitunen, T., & Lensu, A. (2016). The role of landscape, topography, and geodiversity in

- explaining vascular plant species richness in a fragmented landscape. *Boreal Environment Research*, 21, 53–70.
- Ricklefs, R. E. (1987). Community diversity: Relative roles of local and regional processes. *Science*, 235, 167–171.
- Ridgeway, G. (2015). *gbm: Generalized Boosted Regression Models*. R package version 2.1.1. Retrieved from: <http://CRAN.R-project.org/package=gbm>. Accessed September, 2016.
- Rosenzweig, M. L. (1995). *Species diversity in space and time*. Cambridge: Cambridge University Press.
- Stein, A. (2015). Environmental heterogeneity–species richness relationships from a global perspective. *Frontiers of Biogeography*, 7, 168–173.
- 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.
- Stein, A., & Kreft, H. (2014). Terminology and quantification of environmental heterogeneity in species-richness research. *Biological Reviews*, 90, 815–836.
- Thomas, M.F. (2012). Geodiversity and landscape sensitivity: A geomorphological perspective. *Scottish Geographical Journal*, 128, 1–16.
- Tukiainen, H., Bailey, J. J., Field, R., Kangas, K., & Hjort, J. (2016). Combining geodiversity with climate and topography to account for threatened species richness. *Conservation Biology*, doi: 10.1111/cobi.12799
- Valdés, A., Lenoir, J., Gallet-Moron, E., Andrieu, E., Brunet, J., Chabrierie, O., ... Decocq, G. (2015). The contribution of patch-scale conditions is greater than that of macroclimate in explaining local plant diversity in fragmented forests across Europe. *Global Ecology and Biogeography*, 24, 1094–1105.
- Viles, H., Naylor, L. A., Carter, N. E. A., & Chaput, D. (2008). Biogeomorphological disturbance regimes: Progress in linking ecological and geomorphological systems. *Earth Surface Processes and Landforms*, 33, 1419–1435.
- Virtanen, R., Luoto, M., Rämä, T., Mikkola, K., Hjort, J., Grytnes, J. A., & Birks, H. J. B. (2010). Recent vegetation changes in the high-latitude tree-line ecotone are controlled by geomorphological disturbance, productivity and diversity. *Global Ecology and Biogeography*, 19, 810–821.
- Walker, K., Pearman, D., Ellis, B., McIntosh, J., & Lockton, A. (2010). Recording the British and Irish flora 2010–2020 (Annex 1: Guidance on sampling approaches). Retrieved from: <http://www.bsbi.org.uk/RecordingStrategy.pdf>. Accessed September, 2016.
- Willis, K. J., & Whittaker, R. J. (2002). Species diversity – scale matters. *Science*, 295, 1245–1248.
- Wild Flower Society (2010). List of all British plants 2010 (online). Retrieved from: [http://www.thewildflowersociety.com/wfs\\_list\\_of\\_all\\_plants/main\\_menu\\_2010.htm](http://www.thewildflowersociety.com/wfs_list_of_all_plants/main_menu_2010.htm). Accessed September, 2016.

## BIOSKETCH

JOSEPH J. BAILEY is a PhD candidate at the University of Nottingham funded by the U.K. National Environment Research Council (NERC; Studentship 1365737) and supervised by R.F. and D.B. His PhD comprises multiple individual projects across a number of scales and geographical locations. The common theme is the scrutinization of when and where we can most effectively use geodiversity in modelling biodiversity and species distributions.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Bailey JJ, Boyd DS, Hjort J, Lavers CP, Field R. Modelling native and alien vascular plant species richness: At which scales is geodiversity most relevant? *Global Ecol Biogeogr.* 2017;00:1–14. <https://doi.org/10.1111/geb.12574>