

RESEARCH ARTICLE

Challenging the geographic bias in recognising large-scale patterns of diversity change

Wenyuan Zhang¹  | Richard Grenyer² | Kevin J. Gaston³ | Ben C. Sheldon¹

¹Edward Grey Institute, Department of Biology, University of Oxford, Oxford, UK

²School of Geography and the Environment, University of Oxford, Oxford, UK

³Environment and Sustainability Institute, University of Exeter, Cornwall, UK

Correspondence

Wenyuan Zhang, Department of Biology, Quebec Centre for Biodiversity Science, McGill University, Montreal, Quebec H3A 1B1, Canada.
Email: wzhang.ecoevo@gmail.com

Present address

Wenyuan Zhang, Department of Biology, Quebec Centre for Biodiversity Science, McGill University, Montreal, Quebec, Canada

Funding information

China Scholarship Council, Grant/Award Number: 201806210038

Editor: Corey T Callaghan

Abstract

Aim: Geographic structure is a fundamental organising principle in ecological and Earth sciences, and our planet is conceptually divided into distinct geographic clusters (e.g. ecoregions and biomes) demarcating unique diversity patterns. Given recent advances in technology and data availability, however, we ask whether geographically clustering diversity time-series should be the default framework to identify meaningful patterns of diversity change.

Location: North America.

Taxon: Aves.

Methods: We first propose a framework that recognises patterns of diversity change based on similarities in the behaviour of diversity time-series, independent of their specific or relative spatial locations. Specifically, we applied an artificial neural network approach, the self-organising map (SOM), to group time-series of over 0.9 million observations from the North American Breeding Birds Survey (BBS) data from 1973 to 2016. We then test whether time-series identified as having similar behaviour are geographically structured.

Results: We find little evidence of strong geographic structure in patterns of diversity change for North American breeding birds. The majority of the recognised diversity time-series patterns tend to be indistinguishable from being independently distributed in space.

Main Conclusions: Our results suggest that geographic proximity may not correspond to shared temporal trends in diversity; assuming that geographic clustering is the basis for analysis may bias diversity trend estimation. We suggest that approaches that consider variability independently of geographic structure can serve as a useful addition to existing organising rules of biodiversity time-series.

KEYWORDS

biodiversity change, functional diversity, nongeographic patterns, pattern recognition, phylogenetic diversity, spatial distribution, taxonomic diversity, temporal trends

1 | INTRODUCTION

Biodiversity describes the variety and heterogeneity of organisms at all levels of the hierarchy of life, from genes to species to ecosystems (Gaston, 2000). It is changing at accelerating rates worldwide due to human activity (Hull et al., 2015; Pecl et al., 2017). This implies an urgent need for appropriate assessments to explain the uneven distributions and patterns of biodiversity in order to decide conservation policies and gain maximal conservation benefits (Cardinale et al., 2012; Jetz et al., 2019). Most research attempting to describe and analyse the multiple dimensions of biodiversity has been from one of two perspectives. First, metrics and indicators are proliferating to capture different facets of, and values derived from, biodiversity (Pascual et al., 2021; Pereira et al., 2013; Skidmore et al., 2021). For example, taxonomic diversity has been assumed to function as a surrogate for different biodiversity facets (Rapacciuolo et al., 2019), mainly owing to the easy quantification and interpretation of species distribution data (Magurran, 2021). However, with the developing recognition that species are not equivalent (e.g. some perform unique functions and some clades carry more evolutionary history than others) (Vane-Wright et al., 1991; Vellend et al., 2011; Winter et al., 2013), measures of functional and phylogenetic diversity have recently been developed (Cadotte et al., 2015; Winter et al., 2013) to incorporate these facets of diversity. To be more specific, functional diversity is a measurement that incorporates different functions that species perform in an ecosystem and is usually characterised in terms of functional traits, that is morphological, physiological, or behavioural characteristics (Petchey & Gaston, 2006). Phylogenetic diversity is a family of metrics derived from the phylogenetic distance among taxa and is sometimes controversially treated as a proxy for functional diversity, because functional traits are often phylogenetically conserved and measured imperfectly (Cadotte et al., 2017; Mazel et al., 2018).

Second, studies have been conducted to investigate the spatial heterogeneity of biodiversity in different geographic clusters (e.g. ecoregions and biomes) within which organisms are assumed to have similar responses to environments (Dobrowski et al., 2021; Yu et al., 2019). For example, regional factors have often been used to infer environmental drivers for different populations in niche modelling and biodiversity mapping (Jetz et al., 2019), and diversity change at a single site has been employed to represent the change at a coarse scale (Antão et al., 2020; Daskalova et al., 2020). Whilst it is becoming widely recognised that species respond differently to the same environmental pressures and the relationship between them may vary at different scales (Jarzyna & Jetz, 2018; Owen et al., 2019; Tucker et al., 2018), macroecological studies like these continue to use predefined geographic clusters to carry out their analyses and recognise diversity patterns. This is likely owing to two reasons. First, it is a classic approach to observe and record biological processes at a local scale and summarise the large-scale biological patterns in a bottom-up way, where an assumption of similarity due to geographic proximity is easily adopted (Díaz & Malhi, 2022; Hughes,

Orr, Yang, & Qiao, 2021; Tscharrntke et al., 2012; Willig et al., 2003). Second, although there has been extensive data collection for biodiversity surveys, it remains challenging to sample a whole given area at large scales to capture all its diversity changes (Hughes, Orr, Ma, et al., 2021; Valdez et al., 2023), especially for long-term biological survey programmes that require substantial coordination and support (Bowler et al., 2022; Zhang et al., 2021). Thus, regional diversity patterns are calculated from samples within them, and geographic structure is used per se.

However, using geographic structure to define and analyse patterns of diversity change may introduce biases for two reasons. First, the spatial distribution of diversity is highly heterogeneous (Gaston, 2000), which means diversity in the same predefined geographic cluster can still vary greatly. Second, given the significance of human effects on nature, divergent diversity trends can be observed in sites that whilst geographically close experience these effects to varying degrees. This suggests events that can change diversity trends may not match simple geographic structures. Therefore, patterns of diversity change that are recognised by different geographical clusters may not effectively capture temporal diversity patterns. However, although such biases may exist, analyses that have estimated diversity change have commonly used geographic structure to delineate large-scale patterns of diversity change, such as ecoregions (Harrison et al., 2018; Sano et al., 2019), bird conservation regions (Jarzyna & Jetz, 2018), country boundaries (Normander et al., 2012) and continents (Blowes et al., 2019; Kerr et al., 2015; Soroye et al., 2020). Little attention has been paid to whether patterns of diversity change can most effectively be recognised by geographically clustering.

Here, we ask: Should geographic structure be used as a default approach to recognise diversity patterns? To address this question, we propose a framework that tests whether times-series identified as having similar behaviour are geographically structured. We first recognise patterns of diversity change based on behaviours of diversity time-series neglecting their geographic locations. As such, diversity time-series are recognised as having the same pattern on the basis of similar trends in their variability rather than their geographic proximity. We then test the spatial dependence of time-series within the same pattern. Therefore, a geographic structure is found when diversity time-series that have a similar pattern of temporal change behaviour tend to be geographically clustered (Figure 1b–e). Otherwise, the time-series should be distributed independently (Figure 1f,g), signalling that geographical recognition of diversity patterns is likely biased. We illustrate this framework with the North American Breeding Bird Survey data both across North America and east of the Mississippi River. The data are collected annually over 5000 survey routes located across the continent of North America since 1966 (Pardieck et al., 2020). To incorporate multiple facets of diversity, we calculate taxonomic, functional and phylogenetic diversity for all time-series. As a comparison, we complement this framework using analysis of a remote-sensing data set (MOD13C2 MODIS/Terra Vegetation Indices) (Didan, 2015) to produce a vegetation index (EVI) for cells across North America occupied by bird survey routes, which we expect a priori to demonstrate very strong spatial structuring.

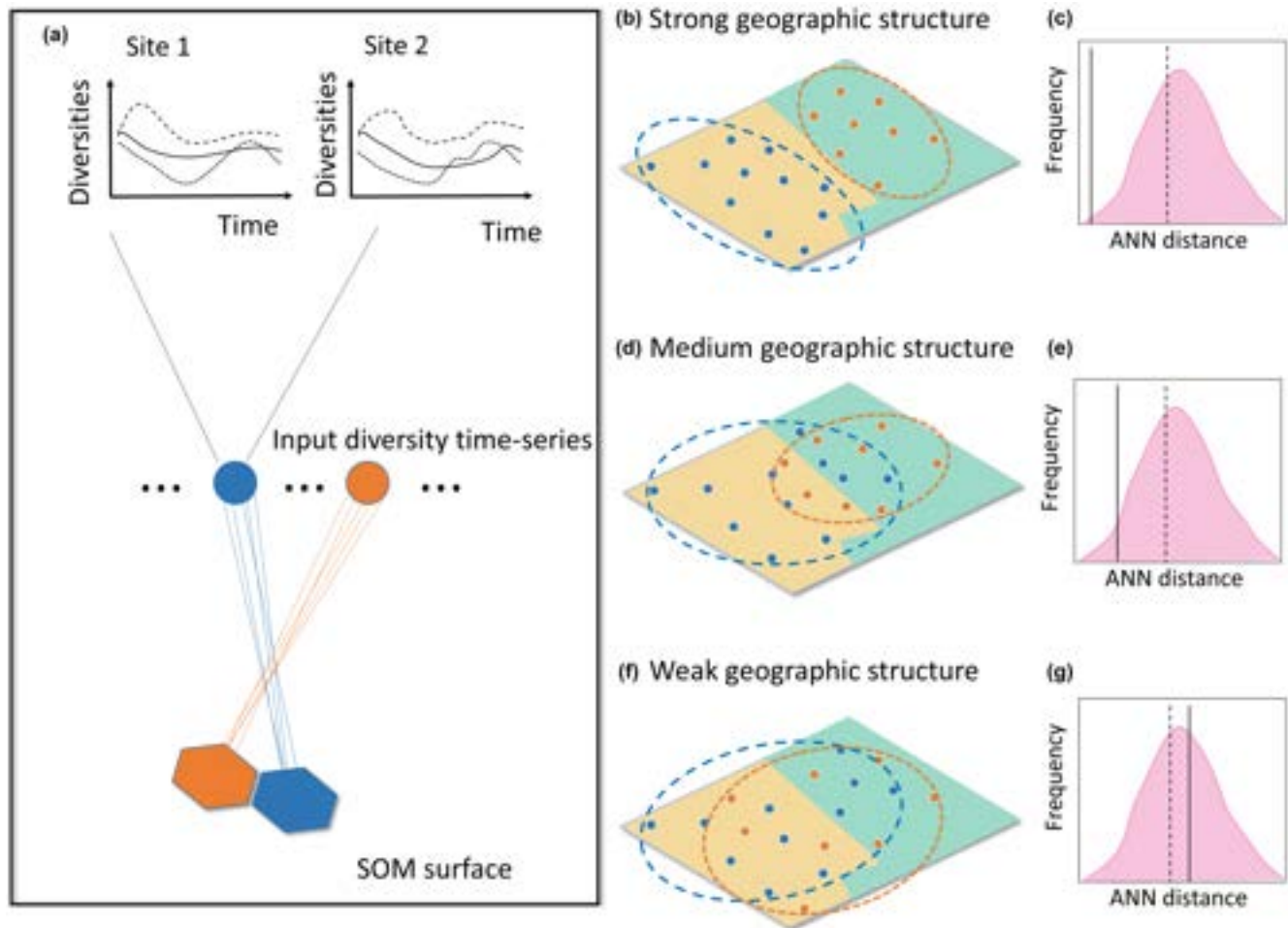


FIGURE 1 Conceptual framework to test spatial dependence. (a) Biodiversity facets from different sites are clustered onto a 2×1 SOM surface. (b and c) Recognised patterns of diversity change show strong geographic structures. (d and e) Recognised patterns of diversity change have medium geographic structures. (f and g) Represent no evidence for strong or medium geographic structures underlying the recognised patterns of diversity change. An average nearest neighbour (ANN) distance is used to measure the geographic proximity of different locations. ANN distance from the diversity time-series falling to the left of the histogram of ANN distances for randomly generated locations supports that geographic structure is important for pattern recognition (c and e), whilst ANN distance from real diversity time-series adhering to recognition of the weak geographic pattern will fall in the middle of the distribution (g). Points in (b), (d), and (f) with the same colour surrounded by a dotted circle represent that they have similar diversity change and are recognised into the same pattern. Solid lines in (c), (e) and (g) represent the ANN distance of diversity time-series, and dashed lines represent the average ANN distance for a set of random samples.

2 | MATERIALS AND METHODS

2.1 | Nongeographic framework

2.1.1 | Pattern recognition

Diversity change time-series are inherently likely to show autocorrelation (Daskalova et al., 2021), and methods used to group diversity time-series should incorporate this. Here, we used self-organising maps (SOMs; T. Kohonen, 1997) to determine the patterns of the different diversity time-series. Self-organising map is based on an artificial neural network and uses an unsupervised learning method, well adapted to complex data analyses. As an informative and intuitive method in feature extraction, SOM has found wide application in a variety of disciplines for clustering, classification, dimensionality

reduction and data visualisation (Liu et al., 2016). Some examples of the application of SOM in ecology are as follows: determining representative species (Park et al., 2006); investigating fish assemblages (Penczak et al., 2012); and modelling temporal evolution of Pacific surface chlorophyll (Huang et al., 2017).

In an SOM algorithm, high-dimensional data sets are projected onto a low-dimensional space (typically two-dimensional), meanwhile preserving the similarity and the difference between the input data vectors. The process is organised by automatically detecting relevant subgroups of similar input vectors and generating neurons (virtual vectors) that describe the coordinates of centres of the subgroups. Neurons in a network respond to a given subgroup of similar input vectors and move closer to each other with the addition of new input vectors to it. Whilst a classic result from a SOM is a topology map (where the data sets to be classified

are mapped onto map nodes), the neurons that are neighbours on the topology map are expected to represent similar patterns; consequently, dissimilar patterns are expected to be distant from each other on the map. The training process iterates and tunes the network based on a preselected number of groups (hence, self-organising).

One of the most challenging steps in using SOM is to choose the number of groups. Topographic and quantisation errors (Teuvo Kohonen, 2012) have been proposed to help determine this for an SOM approach. The topographic error (TE) denotes the continuity of the topology mapping. Here, we calculated TE by the average distance between all pairs of most similar neurons of input vectors. The topographic error therefore can measure the topographic preservation and the accuracy of the mapping in topology.

$$TE = \frac{1}{n} \left(\sum_{i=1}^N \min \|u_i - u\| \right)$$

where N is the number of neurons, and $\|u_i - u\|$ is the topographic distance between a given neuron and all other neurons.

The quantisation error (QE) represents the average distance between each input vector and its neuron.

$$QE = \left(\sum_{i=1}^N \|x_i - u_c\| \right)$$

where N is the number of the input diversity time-series, and $\|x_i - u_c\|$ is the topographic distance between every input vector and its neuron. Usually, small values of TE and QE suggest a good performance of the SOM model. However, QE and TE gradually decrease as the number of neurons increases and thus can only help determine the optimal number of neurons at a local scale (Park et al., 2018). As the data structure is unknown before running a SOM model, the SOM surface is usually set to be similar to a square (i.e. equal length and width) to reduce bias (Park et al., 2018). We selected the number of neurons based on both TE and QE of relatively small values for different numbers of neurons ranging from 4 to 25 with arbitrary intervals to reach square-shaped surface.

Self-organising map models can be applied directly to time-series data (Vesanto & Alhoniemi, 2000). We used the SOM model as a clustering and pattern recognition method to investigate temporal diversity patterns. Specifically, the input diversity time-series were iterated by SOM to find the best matching neuron. Locations with similar diversity variability (i.e. diversity range and variation) were taken as one group. Only once the data were classified on their temporal diversity variability, the geographic distribution of the time-series identified and plotted. We generated nine groups of diversity time-series to explore their temporal change for the relatively smaller TE and QE of these numbers (Figures S1–S17), which can also efficiently extract patterns of diversity. To contrast diversity variability with that from other numbers of groups, we also calculated the SOM models of 5×5 , 5×4 , 4×4 and 2×2 map

size. We employed the 'supersom' function of the 'kohonen' package in R 4.0.1 (<https://www.r-project.org>). All SOM models were performed with a hexagonal structure and Gaussian neighbourhood function. We used the hexagonal structure because nodes in a hexagonal structure have six neighbours and can display greater variance in neighbourhood size, compared to a rectangular structure where nodes have four neighbours. Taxonomic, functional and phylogenetic diversity were used simultaneously with the same weight in each SOM model.

2.1.2 | Spatial dependence

We used a Monte Carlo test with the average nearest neighbour (ANN) distance to test spatial dependence (Clark & Evans, 1954) in membership of the groups of diversity time-series. $ANN = \frac{\sum d_i}{n}$, where d_i corresponds to the distance between feature i and its nearest neighbouring feature, and n is the number of features. The great circle distance on a WGS84 ellipsoid was used for all the distance calculations. Distance in raster data is calculated by the centres of each cell. For a given group of time-series that have the same pattern, we randomly selected an equal number of time-series to the number of time-series of that group from across all time-series analysed. Then, the ANN distance for locations of the randomly selected time-series was calculated. We did not generate random points across the space for randomisation, because the biological surveys are often not randomly located (e.g. avoiding overlap). That process might be explicit (i.e. the people who organise the surveys specify this), or it might not be (e.g. it just results from the decisions individuals make, but they make these in relation to other people's choices). The randomisation process was replicated 1000 times to generate a density histogram showing the distribution of ANN distances of the randomly distributed time-series. A pseudo p -value was calculated to show the probability of incorrectness if rejecting the geographic independence hypothesis that diversity time-series that have the same pattern are randomly distributed:

$$P = \frac{\min(N_{\text{greater}} + 1, N + 1 - N_{\text{greater}})}{N + 1}$$

where N_{greater} is the number of simulated ANN distances greater than the real ANN distance, and N is the number of replications. We did not remove locations that might lead to spatial pseudoreplications as here we aimed to test the geographic proximity of diversity time-series, which would be reflected by the results of the above analysis (e.g. p -value).

2.2 | North American Breeding Bird Survey data

2.2.1 | Data collection

We used the North American Breeding Bird Survey (BBS) data to investigate the biodiversity changes in different diversity facets

(Pardieck et al., 2020). Breeding Bird Survey is a long-term avian monitoring programme that tracks the population dynamics of breeding birds and follows a strict survey protocol allowing for yearly comparisons. Breeding Bird Survey data are collected once per year in June over 5000 survey routes that are located across the continent of North America. Each survey route is approximately 40km long, with 50 stops, and split into five segments. At each stop, the birds will be detected when they are seen or heard within a 400m radius for 3 min. Generally, the trained observers record the stop ID, and the presence and abundance of each species encountered. The sampling for each whole route is conducted over approximately 5h in the early morning.

Breeding Bird Survey data from the early years of the survey are limited due to insufficient sample size and spatial coverage, so we excluded those from 1966 to 1972. An SOM model treats data gaps as form of similarity between time-series (Park et al., 2018), which introduced bias if actual data in these gaps would have been different. To keep the integrity of temporal coverage and obtain enough samples, for the period after 1972, we retained the routes with gaps of less than 3 years' duration and then conducted a multivariate imputation by chained equations (MICE; van Buuren & Groothuis-Oudshoorn, 2011), as implemented by the 'mice' package in R, to fill the gaps after diversity calculation. Multivariate imputation by chained equation has a better prediction quality and produces less biased estimates than other imputation methods in risk models using regression (Ambler et al., 2007). We used a random forest algorithm (Breiman, 2001) for MICE, as it had high performance, required little computational time and was good at dealing with high-dimensional data (Pantanowitz & Marwala, 2009; Penone et al., 2014). We removed from all routes records for unidentified species (66 out of 713 species) and analysed a total of 911,702 records (20,721 per year) from 316 out of 5690 survey routes with less than 3 years' data gaps during 1973–2016. We imputed a sum of 331 data gaps (2.4% of the total number of years across 316 routes) from 214 routes based on all taxonomic, functional and phylogenetic records for each survey route. To reduce the effects of nonrandom spatial sampling, we also implemented our approach for the region east of Mississippi River which resulted in 200 survey routes as a comparison.

2.2.2 | Diversity measures

To compare taxonomic, functional and phylogenetic diversity in a consistent framework, we calculated these using an abundance-weighted measure, Rao's quadratic entropy (Rao, 1982):

$$Q = \sum_{ij} d_{ij} p_i p_j$$

where d_{ij} is the taxonomic, functional or phylogenetic diversity dissimilarity between each pair of species i and j , whilst p_i and p_j are their relative abundance. We then calculated the three diversity indices for a given time-series for every year.

For taxonomic diversity, the distance between all species is the same ($d_{ij} = 1$), and Rao's Q , in this case, is equivalent to the Gini-Simpson index, which usually measures the probability of interspecific encounter (Hurlbert, 1971). For the calculation of functional diversity, we used four categories of 16 traits from Elton Traits 1.0 (Wilman et al., 2014). These traits, which are assumed to represent key Eltonian niche dimensions, comprised: body mass, diet (i.e. the proportional use of invertebrates, vertebrates, carrion, fresh fruits, nectar and pollen, seeds, and other plant materials in species' diet), foraging niche (i.e. prevalence of foraging below water surfaces, on water surface, on terrestrial ground level, in understorey, in mid-canopy, in upper canopy, and aerial) and broad habitat types (i.e. pelagic or not). The traits that we used to calculate functional diversity are correlated with each other to different extents (Figure S2). However, considering there are no traits that are fully correlated and each trait carries unique features shaping avian functions, we retain all traits in the Elton trait database for subsequent analyses. We gave equal weights to each trait category, which resulted in weights for body mass and broad habitat type of 1, and 1/7 for each diet and foraging niche variable. Functional distance was calculated using a multivariate trait dissimilarity of Gower's distance (Gower, 1971) for each pairwise species (Pavoine et al., 2009), with functional diversity calculated as the subtree length on the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) dendrograms of these distances. Phylogenetic distance was calculated using 100 phylogenies sampled from a full pseudo-posterior distribution of phylogenetic trees (<http://birdtree.org>). The mean phylogenetic diversity across these 100 dendrograms was calculated. All three diversity facets were standardised by $SD_i = (D_i - \min(D)) / (\max(D) - \min(D))$ before further analysis, where SD_i is the standardised diversity, D is the diversity and $D = (D_1, \dots, D_n)$.

2.2.3 | Relationship between different diversity facets

To investigate the correlation of functional-taxonomic, phylogenetic-taxonomic and functional-phylogenetic diversity in each group of diversity time-series, we adopted generalised additive models (GAM) and included the year of a given time-series as a random effect. We calculated the deviation explained and 95% CI for all the groups to measure the correlation of different pairs of diversity facets.

2.3 | MODIS/Terra Vegetation Indices

We used the enhanced vegetation index (EVI) from a remote-sensing product (MOD13C2) (Didan, 2015) as a test of the framework in different sample coverages for a system where we expect strong spatial clustering. The EVI provided biomass information and can be used to quantify vegetation greenness. The product provided a monthly EVI in a 5600-m resolution from 2000 to 2021 across the globe. As we aim to test the spatial dependence of EVI time-series, we here

quantified their annual change in September when the activity of vegetation is usually at a maximum and relatively constant over the year (Villamuelas et al., 2016) across North America to simplify the calculation. We applied our framework with EVI time-series that locate within the same cells as bird survey routes (316 time-series).

3 | RESULTS

To test the relative fit of geographic structuring, we assessed the spatial dependence of time-series within each pattern recognised by changes in their variabilities using two data sets with different spatial sample coverages. We found no evidence of stronger geographic structure on diversity change of North American breeding

birds than random. We found that the average distance between the survey routes lay in the middle part of the distance frequency curve calculated from randomly distributed points, although the number of survey routes within eight of the nine groups varied (Figure 2). Similar results were also found for using survey routes within east of the Mississippi River region (Figure 3 and Figures S12–S14) and other numbers of groups of diversity time-series (Figures S3–S6).

The temporal trends, range and correlation of taxonomic, functional and phylogenetic diversity varied across the different groups of diversity time-series (Figure 4 and Figures S7–S10). Taxonomic diversity showed a gradual increase in four patterns (Figure 4G1,G4,G7,G9) and decline or no change in the other groups (Figure 4G2,G3,G5,G6,G8). Functional and taxonomic diversity were positively correlated in most groups ($DE = 0.44$, 95% CI = $0.30 \sim 0.59$;

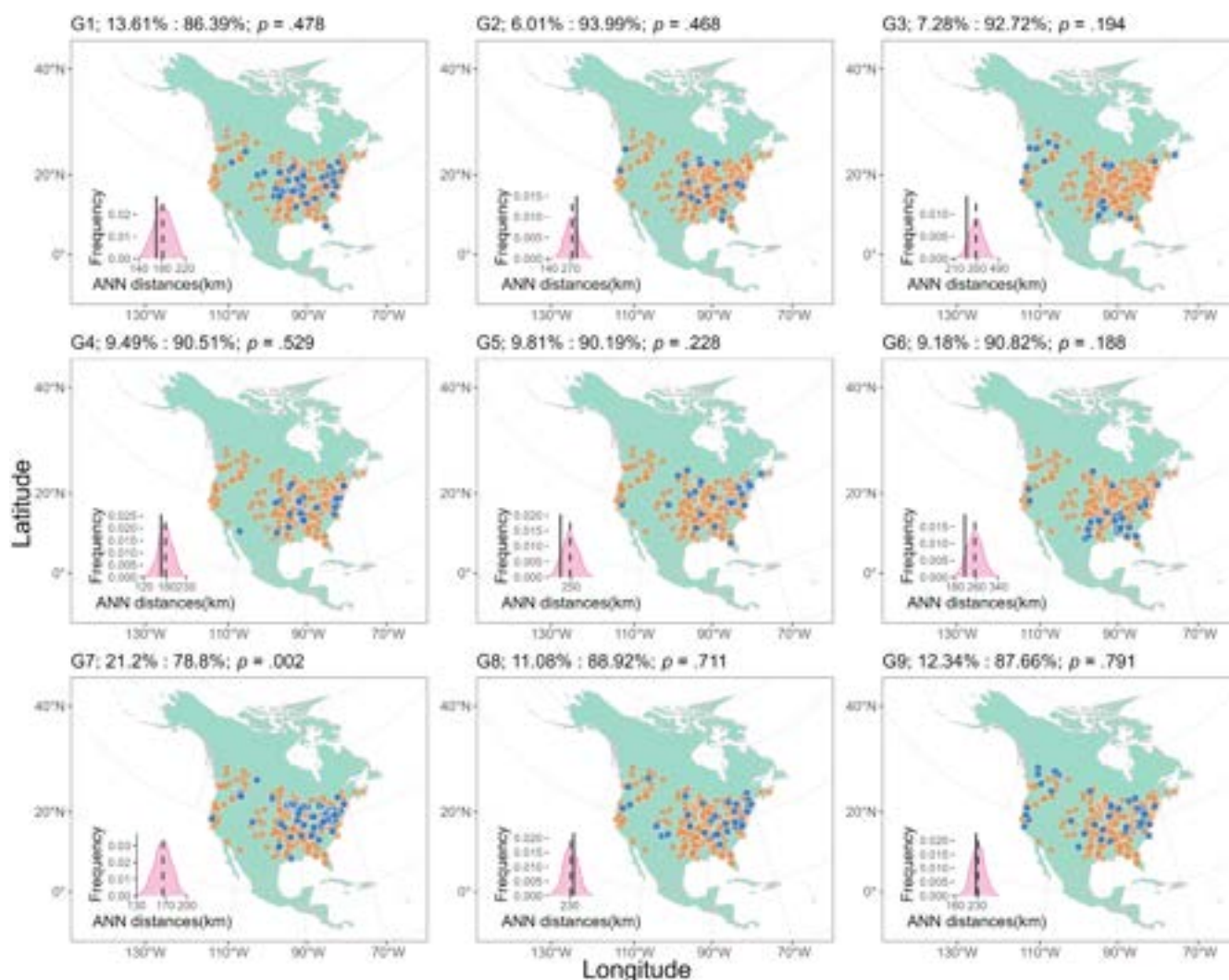


FIGURE 2 Distribution of survey routes in the North American Breeding Birds Survey and spatial independence test. Each panel presents the distribution of diversity time-series that have the same pattern as identified by the SOM. The inset density plot shows the distribution of the ANN distance of 1000 Monte Carlo tests. Solid lines in the density plot represent the average nearest neighbour (ANN) distance of the survey routes, and dashed lines represent the mean value of ANN for 1000 random samples. Blue points represent the start locations of survey routes in the focal SOM group and orange points are the start locations of other survey routes that do not follow into the SOM group. Percentages on the left top of each panel denote the percentages of time-series in that group to all time-series. Pseudo p -values represent the probability of incorrectness if rejecting the geographic independence hypothesis that diversity time-series that have the same pattern are randomly distributed.

Figure S11). Functional diversity showed similar trends to taxonomic diversity, but the change rates were different.

The correlation between taxonomic and phylogenetic diversity varied less in groups of diversity time-series (Figure 4). Phylogenetic diversity had a trend of slowly increasing or staying flat as taxonomic diversity increased ($DE = 0.16$, 95% CI = $0.09 \sim 0.23$). The correlation between functional and phylogenetic diversity differed more in groups of diversity time-series compared with the correlation between taxonomic and functional diversity, or taxonomic and phylogenetic diversity ($DE = 0.41$, 95% CI = $0.32 \sim 0.52$). In most of the groups (seven out of nine groups), functional diversity increased faster than the increase of phylogenetic diversity and could also increase more slowly or even decline with the rise of phylogenetic diversity (Figure 4).

We found EVI time-series that were recognised to have the same pattern of temporal change had a clear geographic structure (Figure 5). We found strong evidence that the distribution of the time-series is distinct from random across all groups (Figure 5 and Figures S15–S17). Hence, this method is able to detect and return strongly spatially clustered signals.

4 | DISCUSSION

Despite the frequent use of geographic structure in studies of the distribution of biodiversity (Blowes et al., 2019; Devictor et al., 2010; Gaston, 2000; Willig et al., 2003), we reveal that using geographic structure as a default rule to recognise diversity patterns may be

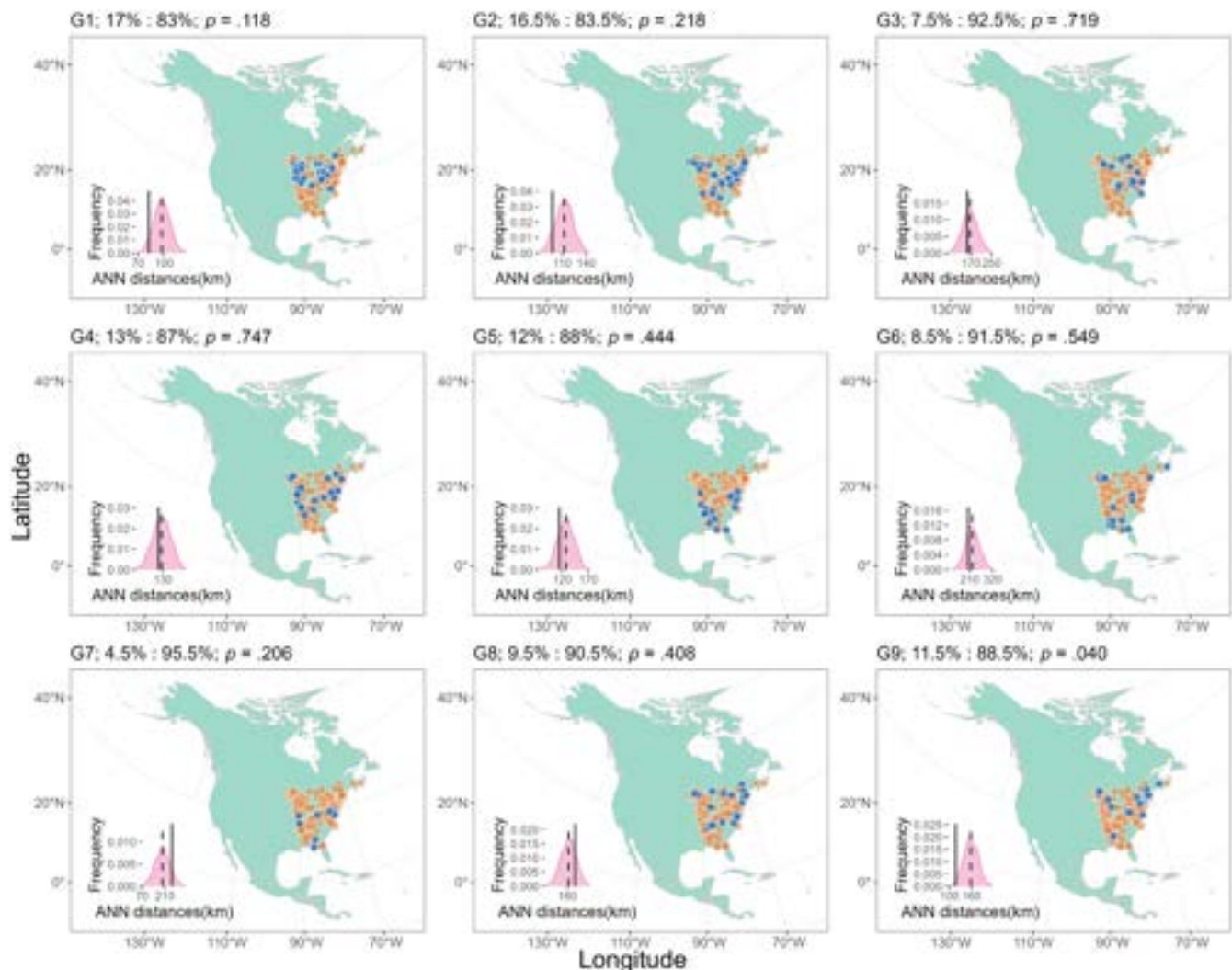


FIGURE 3 Distribution and spatial independence test of survey routes for the North American Breeding Birds Survey to the east of Mississippi River. Each panel presents the distribution of diversity time-series that have the same pattern as identified by the SOM. The inset density plot shows the distribution of the ANN distance of 1000 Monte Carlo tests. Solid lines in the density plot represent the average nearest neighbour (ANN) distance of the survey routes, and dashed lines represent the mean value of ANN for 1000 random samples. Blue points represent the start locations of each survey route in the focal SOM group and orange points are the start locations of other survey routes that do not follow into the SOM group. Percentages in the left top of each panel denote the percentages of time-series in that group to all time-series. Pseudo p -values represent the probability of incorrectness if rejecting the geographic independence hypothesis that diversity time-series that have the same pattern are randomly distributed.

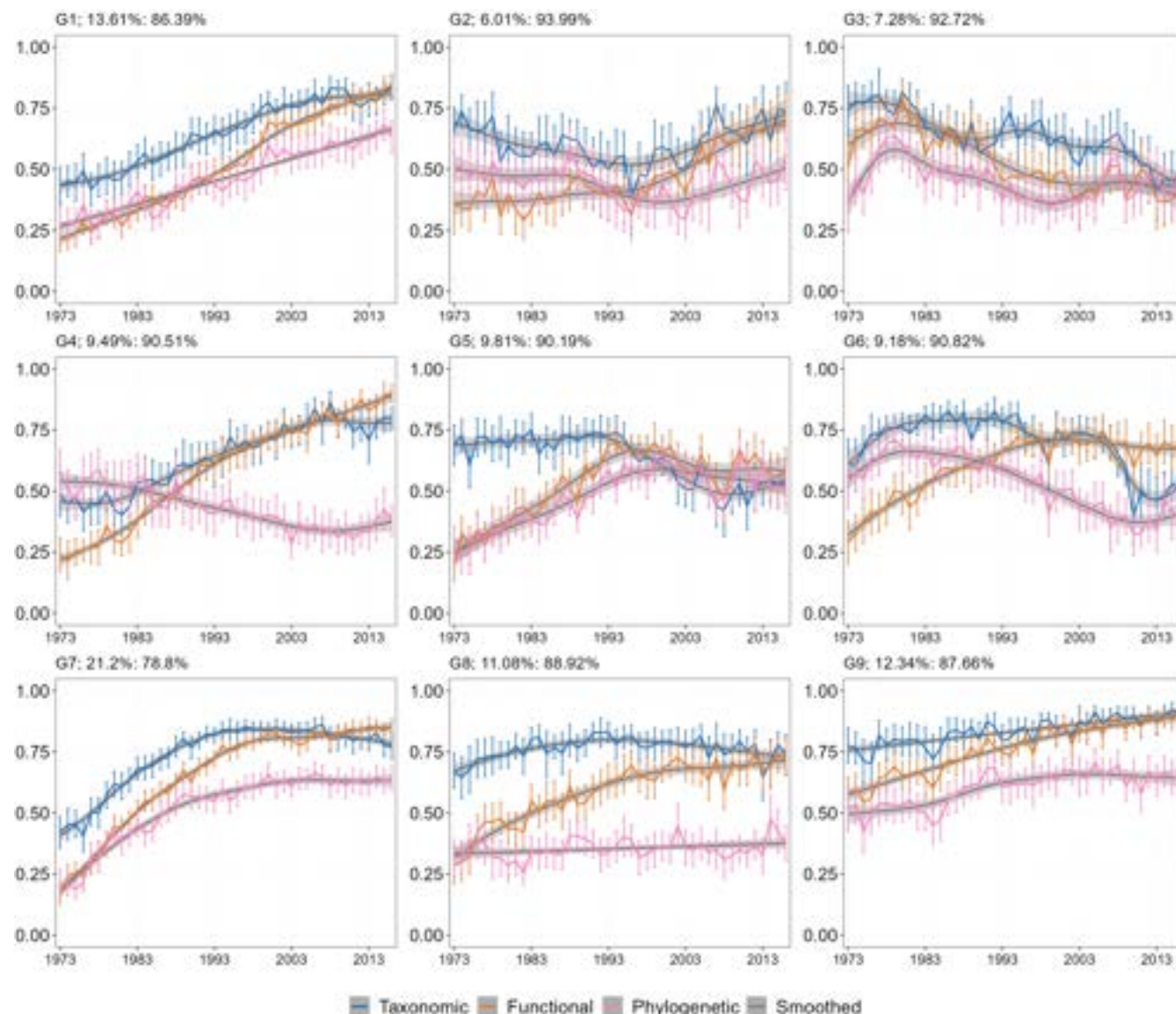


FIGURE 4 Temporal changes of taxonomic, functional and phylogenetic diversity in the North American Breeding Birds Survey. Each panel denotes the change of diversity time-series that have the same pattern. The error bars represent a 95% confidence interval. The non-linear regressions (a loess sliding window with a 33% range width; smoothed line) of the diversity facets were added to describe the major temporal trajectory of each group.

biased; there are other kinds of structure in such data that are important to understand diversity patterns.

We proposed a nongeographic framework to group biodiversity time-series based on a pattern recognition method (i.e. SOM) and highlight potential geographic biases to understand patterns of diversity change. Specifically, we classified the time-series that varied similarly in the temporal trend as having the same pattern and tested whether the geographic distribution of areas of the same pattern is random. Using the North American Breeding Birds data, we demonstrated that diversity time-series that have the same pattern of temporal change could be independent of their geographic locations. This implies that in large-scale studies, geographic proximity may not correspond to shared temporal diversity trends. Ignoring the possible variation between time-series of a given geographic cluster,

for example simply taking all time-series as a whole, might bias the understanding of biodiversity change.

By identifying geographically disparate patterns of similar diversity change, we emphasise the complex resultant structure of biodiversity (Brooks et al., 2006; Gaston, 2000) from two perspectives. First, from a spatial perspective, biodiversity is usually understood as having multiple facets, including taxonomic and species composition, ecosystem function and service potential and evolutionary history, which leads to the complexity in its distribution and structure (Gaston, 2000; Grenyer et al., 2006; Petchey & Gaston, 2006; Tilman, 1982). Communities in neighbouring regions may face different environmental and human pressures, and as a result, have divergent trends in diversity change (Catford et al., 2022). Therefore, the estimation of large-scale diversity patterns may be biased by using

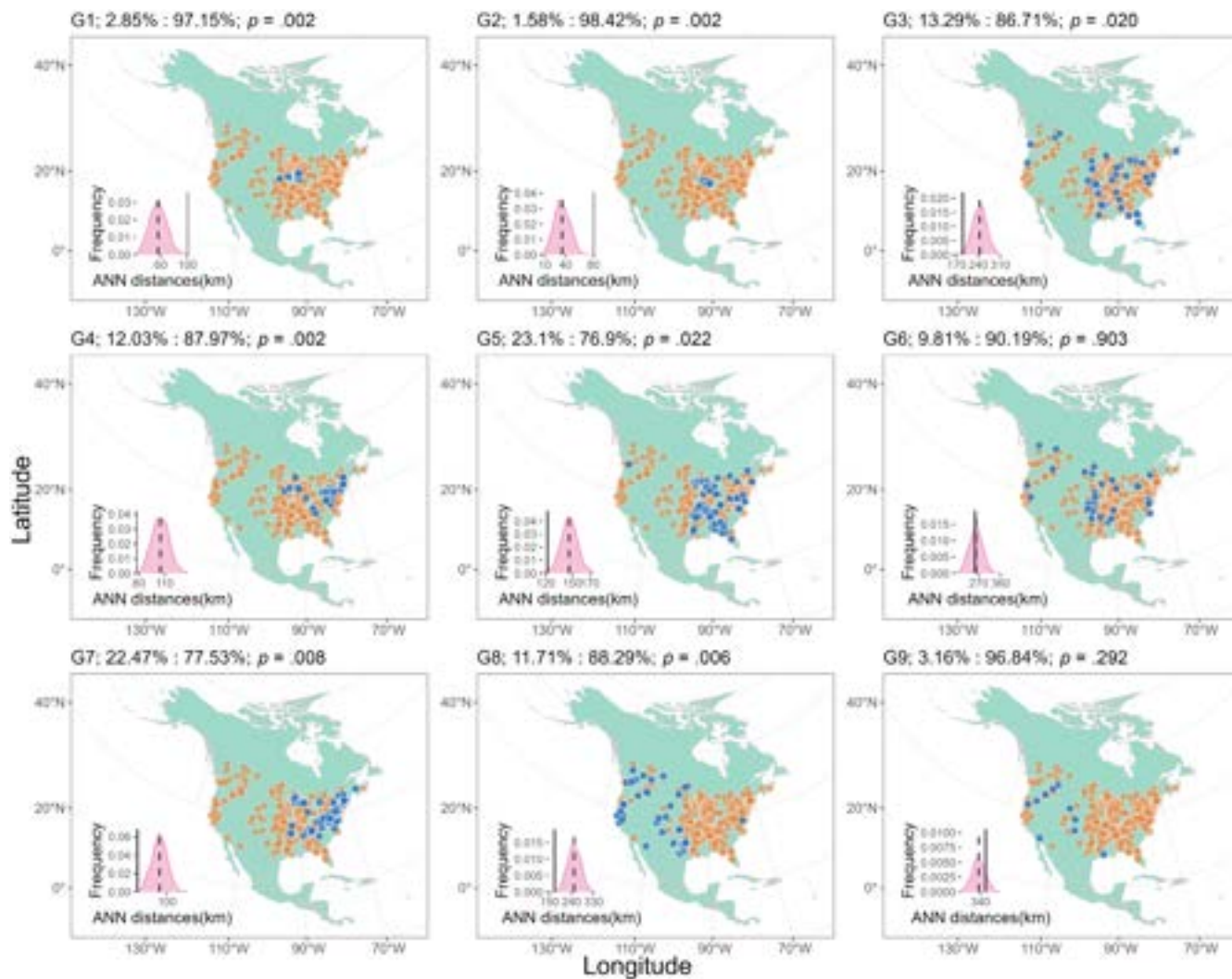


FIGURE 5 Distribution and spatial independence test of enhanced vegetation index (EVI) time-series from the same locations as bird survey routes. Each panel presents the distribution of EVI time-series (blue pixels) that have the same pattern. The inset density plot shows the distribution of the ANN distance of 1000 Monte Carlo tests. Solid lines in the density plot represent the average nearest neighbour (ANN) distance of the locations of EVI time-series, and dashed lines represent the mean value of ANN for 1000 random samples. Blue points represent the start locations of each survey route in the focal SOM group and orange points are the start locations of other survey routes that do not follow into the SOM group. Percentages in the left top of each panel denote the percentages of time-series in that group to all time-series. Pseudo p -values represent the probability of incorrectness if rejecting the geographic independence hypothesis that EVI time-series that have the same pattern are randomly distributed.

geographic structure a priori and not distinguishing the behaviours between diversity time-series. Second, in temporal contexts, biodiversity change is also influenced in complex ways such as the exploitation of biological resources, the introduction of exotic species, climate and land-use change (Mantyka-Pringle et al., 2012; Newbold et al., 2015; Sax & Gaines, 2003). This means despite studies based on static diversity data having found strong geographic structures in their distributions (Bahn & McGill, 2007, 2013), approaches that incorporate the dynamic changes of diversity might still reach juxtaposing conclusions. In addition, the complex change in environments in turn may influence the behaviours of surveyors and estimation of diversity trends (Bowler et al., 2022; Zhang et al., 2021). This is particularly important for recognising patterns of diversity changes considering that the spatial sample coverage of even the most extensive

biodiversity data sets is believed to be less than 7% of the Earth's surface at a 5-km resolution (Hughes, Orr, Ma, et al., 2021), and analyses estimating long-term diversity trends may abandon short time-series or time-series with data gaps (Zhang et al., 2021). Using geographic structure, a priori to seek patterns of dynamic diversity change can result in failures to incorporate diversity changes that were not recorded and bias the diversity estimation. Given the difficulties to improve the spatial coverage of long-term biodiversity surveys, the nongeographic framework that organises diversity time-series independently of their relative or specific locations also adds a novel way to conduct macroecological studies in recognising large-scale diversity patterns for existing diversity time-series data sets.

As for the patterns of North American Breeding bird diversity, we found functional and phylogenetic diversity changed

disproportionately to taxonomic diversity in multiple groups. Specifically, significant changes in functional diversity were accompanied by more minor changes in taxonomic diversity. Similarly, previous studies also found unsynchronised changes between taxonomic and functional diversity (Le Bagousse-Pinguet et al., 2019; Monnet et al., 2014). The correlation between taxonomic and functional diversity was strong and positive in the long term, which was a likely consequence of the replacement of functionally redundant species with unique ones (Beltcik et al., 2020). Furthermore, changes in phylogenetic diversity in the groups were smaller than or even opposite to what would have been expected given the concomitant changes in taxonomic and functional diversity. Correspondingly, phylogenetic diversity showed a tendency to level off and even decline with increasing taxonomic and functional diversity. This phylogenetic homogenisation implies that the bird assemblages have perhaps undergone recruitment of, or even replacement by, closer relatives of species already present. The fast adaptive radiation of colonising species could also contribute to the high levels of taxonomic and functional diversity and low levels of phylogenetic diversity across North America (Jarzyna & Jetz, 2017).

We also tested our framework with an EVI time-series data set. The results showed strong evidence that EVI time-series of the same pattern are geographically structured using the same locations as bird survey routes. This suggests that geographic structure may work well in understanding ecological changes in plant biomass (Holt et al., 2013; Smith et al., 2018) but may not play a meaningful role in delimiting more complex patterns of animal diversity change. Therefore, the results emphasise the geographic biases in recognising patterns of diversity change and the importance of our framework to partition biodiversity change by the similarity of time-series.

Since we focus on the geographic bias in long-term diversity estimations with existing diversity time-series data, there are caveats in interpreting our results. First, temporal or spatial scale-dependences may influence this geographic bias. Future work could explore interplays between different spatial sample coverages and lengths of diversity time-series and how they may influence the geographic bias. Second, our work concentrated on temperate areas whilst diversity dynamics in other biomes can be different given possibly less disturbance in those regions. It is beneficial to explore the extent to which the nongeographic structure is influenced by biomes. Third, in the main analysis, we used North America Breeding Survey data that followed a standard protocol in collecting data. However, studies that might use our framework with other data sets (e.g. citizen data) should be aware of sampling variation, and rarefaction and resampling techniques can be useful for these data sets (Chao et al., 2020; Gotelli & Colwell, 2001).

Future work could also concentrate on identifying drivers of the nongeographic biodiversity patterns extracted from time-series, which can be achieved by determining (non)geographic patterns in the temporal variation of biotic and/or abiotic factors and exploring their potential mismatch with recognised diversity patterns. Then, by analysing the degree of mismatch between diversity and factor patterns, it could be possible to establish networks of ecological interactions and their contributions to the dynamic of diversity change in each pattern. In doing so, recognising nongeographic patterns also

provides a bridge between empirical systems and theoretical interaction frameworks, such as has been used in identifying the relationship between patterns of phytoplankton and eddy lifecycles in oceanography (Huang et al., 2017).

ACKNOWLEDGEMENTS

We thank Owen Lewis, Michael Bonsall and two anonymous reviewers for their suggestions on the paper, and Jie Huang for help in SOM calculation. Wenyuan Zhang is supported by the China Scholarship Council (CSC201806210038).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

BBS data used for these analyses are online <https://doi.org/10.5066/P97WAZE5>, and EVI data sets are available on <https://e4ftl01.cr.usgs.gov/MOLT/MOD13C2.061/>. Codes used for the analyses were uploaded in GitHub (<https://github.com/plmyann/sombio>).

ORCID

Wenyuan Zhang  <https://orcid.org/0000-0002-7102-9922>

REFERENCES

- Ambler, G., Omar, R. Z., & Royston, P. (2007). A comparison of imputation techniques for handling missing predictor values in a risk model with a binary outcome. *Statistical Methods in Medical Research*, 16(3), 277–298. <https://doi.org/10.1177/0962280206074466>
- Antão, L. H., Bates, A. E., Blowes, S. A., Waldock, C., Supp, S. R., Magurran, A. E., Dornelas, M., & Schipper, A. M. (2020). Temperature-related biodiversity change across temperate marine and terrestrial systems. *Nature Ecology & Evolution*, 4(7), 927–933. <https://doi.org/10.1038/s41559-020-1185-7>
- Bahn, V., & McGill, B. J. (2007). Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography*, 16(6), 733–742. <https://doi.org/10.1111/j.1466-8238.2007.00331.x>
- Bahn, V., & McGill, B. J. (2013). Testing the predictive performance of distribution models. *Oikos*, 122(3), 321–331. <https://doi.org/10.1111/j.1600-0706.2012.00299.x>
- Beltcik, M., Lenda, M., Amano, T., & Skórka, P. (2020). Different response of the taxonomic, phylogenetic and functional diversity of birds to forest fragmentation. *Scientific Reports*, 10(1), 20320. <https://doi.org/10.1038/s41598-020-76917-2>
- Blowes, S. A., Supp, S. R., Antão, L. H., Bates, A., Bruelheide, H., Chase, J. M., Moyes, F., Magurran, A., McGill, B., Myers-Smith, I. H., Winter, M., Bjorkman, A. D., Bowler, D. E., Byrnes, J. E. K., Gonzalez, A., Hines, J., Isbell, F., Jones, H. P., Navarro, L. M., ... Dornelas, M. (2019). The geography of biodiversity change in marine and terrestrial assemblages. *Science*, 366(6463), 339–345. <https://doi.org/10.1126/science.aaw1620>
- Bowler, D. E., Callaghan, C. T., Bhandari, N., Henle, K., Benjamin Barth, M., Koppitz, C., Klenke, R., Winter, M., Jansen, F., Bruelheide, H., & Bonn, A. (2022). Temporal trends in the spatial bias of species occurrence records. *Ecography*, 2022(8), e06219. <https://doi.org/10.1111/ecog.06219>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

- Brooks, T. M., Mittermeier, R. A., da Fonseca, G. A. B., Gerlach, J., Hoffmann, M., Lamoreux, J. F., Mittermeier, C. G., Pilgrim, J. D., & Rodrigues, A. S. L. (2006). Global biodiversity conservation priorities. *Science*, 313(5783), 58–61. <https://doi.org/10.1126/science.1127609>
- Cadotte, M. W., Arnillas, C. A., Livingstone, S. W., & Yasui, S.-L. E. (2015). Predicting communities from functional traits. *Trends in Ecology & Evolution*, 30(9), 510–511. <https://doi.org/10.1016/j.tree.2015.07.001>
- Cadotte, M. W., Davies, T. J., & Peres-Neto, P. R. (2017). Why phylogenies do not always predict ecological differences. *Ecological Monographs*, 87(4), 535–551. <https://doi.org/10.1002/ecm.1267>
- Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59–67. <https://doi.org/10.1038/nature11148>
- Catford, J. A., Wilson, J. R. U., Pyšek, P., Hulme, P. E., & Duncan, R. P. (2022). Addressing context dependence in ecology. *Trends in Ecology & Evolution*, 37(2), 158–170. <https://doi.org/10.1016/j.tree.2021.09.007>
- Chao, A., Kubota, Y., Zelený, D., Chiu, C.-H., Li, C.-F., Kusumoto, B., Yasuhara, M., Thorn, S., Wei, C.-L., Costello, M. J., & Colwell, R. K. (2020). Quantifying sample completeness and comparing diversities among assemblages. *Ecological Research*, 35(2), 292–314. <https://doi.org/10.1111/1440-1703.12102>
- Clark, P. J., & Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35(4), 445–453. <https://doi.org/10.2307/1931034>
- Daskalova, G. N., Myers-Smith, I. H., Bjorkman, A. D., Blowes, S. A., Supp, S. R., Magurran, A. E., & Dornelas, M. (2020). Landscape-scale forest loss as a catalyst of population and biodiversity change. *Science*, 368(6497), 1341–1347. <https://doi.org/10.1126/science.aba1289>
- Daskalova, G. N., Phillimore, A. B., & Myers-Smith, I. H. (2021). Accounting for year effects and sampling error in temporal analyses of invertebrate population and biodiversity change: A comment on Seibold et al. 2019. *Insect Conservation and Diversity*, 14(1), 149–154. <https://doi.org/10.1111/icad.12468>
- Devictor, V., Mouillot, D., Meynard, C., Jiguet, F., Thuiller, W., & Mouquet, N. (2010). Spatial mismatch and congruence between taxonomic, phylogenetic and functional diversity: The need for integrative conservation strategies in a changing world. *Ecology Letters*, 13(8), 1030–1040. <https://doi.org/10.1111/j.1461-0248.2010.01493.x>
- Díaz, S., & Malhi, Y. (2022). Biodiversity: Concepts, patterns, trends, and perspectives. *Annual Review of Environment and Resources*, 47(1), 31–63. <https://doi.org/10.1146/annurev-environ-120120-054300>
- Didan, K. (2015). MOD13C2 MODIS/Terra vegetation indices monthly L3 global 0.05Deg CMG V006. NASA EOSDIS land processes DAAC.
- Dobrowski, S. Z., Littlefield, C. E., Lyons, D. S., Hollenberg, C., Carroll, C., Parks, S. A., Abatzoglou, J. T., Hegewisch, K., & Gage, J. (2021). Protected-area targets could be undermined by climate change-driven shifts in ecoregions and biomes. *Communications Earth & Environment*, 2(1), 198. <https://doi.org/10.1038/s43247-021-00270-z>
- Gaston, K. J. (2000). Global patterns in biodiversity. *Nature*, 405, 220–227. <https://doi.org/10.1038/35012228>
- Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4(4), 379–391. <https://doi.org/10.1046/j.1461-0248.2001.00230.x>
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27(4), 857–871. <https://doi.org/10.2307/2528823>
- Grenyer, R., Orme, C. D. L., Jackson, S. F., Thomas, G. H., Davies, R. G., Davies, T. J., Jones, K. E., Olson, V. A., Ridgely, R. S., Rasmussen, P. C., Ding, T. S., Bennett, P. M., Blackburn, T. M., Gaston, K. J., Gittleman, J. L., & Owens, I. P. F. (2006). Global distribution and conservation of rare and threatened vertebrates. *Nature*, 444, 93–96. <https://doi.org/10.1038/nature05237>
- Harrison, T., Gibbs, J., & Winfree, R. (2018). Phylogenetic homogenization of bee communities across ecoregions. *Global Ecology and Biogeography*, 27(12), 1457–1466. <https://doi.org/10.1111/geb.12822>
- Holt, B. G., Lessard, J.-P., Borregaard, M. K., Fritz, S. A., Araújo, M. B., Dimitrov, D., Fabre, P. H., Graham, C. H., Graves, G. R., Jönsson, K. A., Nogués-Bravo, D., Wang, Z., Whittaker, R. J., Fjeldsø, J., & Rahbek, C. (2013). An update of Wallace's zoogeographic regions of the world. *Science*, 339(6115), 74–78. <https://doi.org/10.1126/science.1228282>
- Huang, J., Xu, F., Zhou, K., Xiu, P., & Lin, Y. (2017). Temporal evolution of near-surface chlorophyll over cyclonic eddy lifecycles in the south-eastern Pacific. *Journal of Geophysical Research: Oceans*, 122(8), 6165–6179. <https://doi.org/10.1002/2017JC012915>
- Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, 44(9), 1259–1269. <https://doi.org/10.1111/ecog.05926>
- Hughes, A. C., Orr, M. C., Yang, Q., & Qiao, H. (2021). Effectively and accurately mapping global biodiversity patterns for different regions and taxa. *Global Ecology and Biogeography*, 30(7), 1375–1388. <https://doi.org/10.1111/geb.13304>
- Hull, P. M., Darroch, S. A. F., & Erwin, D. H. (2015). Rarity in mass extinctions and the future of ecosystems. *Nature*, 528, 345–351. <https://doi.org/10.1038/nature16160>
- Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters. *Ecology*, 52(4), 577–586. <https://doi.org/10.2307/1934145>
- Jarzyna, M. A., & Jetz, W. (2017). A near half-century of temporal change in different facets of avian diversity. *Global Change Biology*, 23(8), 2999–3011. <https://doi.org/10.1111/gcb.13571>
- Jarzyna, M. A., & Jetz, W. (2018). Taxonomic and functional diversity change is scale dependent. *Nature Communications*, 9(1), 2565. <https://doi.org/10.1038/s41467-018-04889-z>
- Jetz, W., McGeoch, M. A., Guralnick, R., Ferrier, S., Beck, J., Costello, M. J., Fernandez, M., Geller, G. N., Keil, P., Merow, C., Meyer, C., Muller-Karger, F. E., Pereira, H. M., Regan, E. C., Schmeller, D. S., & Turak, E. (2019). Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology & Evolution*, 3(4), 539–551. <https://doi.org/10.1038/s41559-019-0826-1>
- Kerr, J. T., Pindar, A., Galpern, P., Packer, L., Potts, S. G., Roberts, S. M., Rasmont, P., Schweiger, O., Colla, S. R., Richardson, L. L., Wagner, D. L., Gall, L. F., Sikes, D. S., & Pantoja, A. (2015). Climate change impacts on bumblebees converge across continents. *Science*, 349(6244), 177–180. <https://doi.org/10.1126/science.aaa7031>
- Kohonen, T. (1997). Exploration of very large databases by self-organizing maps. *Proceedings of International Conference on Neural Networks (ICNN'97)*, 1, PL1–PL6. <https://doi.org/10.1109/ICNN.1997.611622>
- Kohonen, T. (2012). *Self-organizing maps* (Vol. 30). Springer Science & Business Media.
- Le Bagousse-Pinguet, Y., Soliveres, S., Gross, N., Torices, R., Berdugo, M., & Maestre, F. T. (2019). Phylogenetic, functional, and taxonomic richness have both positive and negative effects on ecosystem multifunctionality. *Proceedings of the National Academy of Sciences*, 116(17), 8419–8424. <https://doi.org/10.1073/pnas.1815727116>
- Liu, Y., Weisberg, R. H., Vignudelli, S., & Mitchum, G. T. (2016). Patterns of the loop current system and regions of sea surface height variability in the eastern Gulf of Mexico revealed by the self-organizing maps. *Journal of Geophysical Research: Oceans*, 121(4), 2347–2366. <https://doi.org/10.1002/2015JC011493>
- Magurran, A. E. (2021). Measuring biological diversity. *Current Biology*, 31(19), R1174–R1177. <https://doi.org/10.1016/j.cub.2021.07.049>

- Mantyka-Pringle, C. S., Martin, T. G., & Rhodes, J. R. (2012). Interactions between climate and habitat loss effects on biodiversity: A systematic review and meta-analysis. *Global Change Biology*, 18(4), 1239–1252. <https://doi.org/10.1111/j.1365-2486.2011.02593.x>
- Mazel, F., Pennell, M. W., Cadotte, M. W., Diaz, S., Dalla Riva, G. V., Grenyer, R., Leprieux, F., Mooers, A. O., Mouillot, D., Tucker, C. M., & Pearse, W. D. (2018). Prioritizing phylogenetic diversity captures functional diversity unreliably. *Nature Communications*, 9(1), 2888. <https://doi.org/10.1038/s41467-018-05126-3>
- Monnet, A.-C., Jiguet, F., Meynard, C. N., Mouillot, D., Mouquet, N., Thuiller, W., & Devictor, V. (2014). Asynchrony of taxonomic, functional and phylogenetic diversity in birds. *Global Ecology and Biogeography*, 23(7), 780–788. <https://doi.org/10.1111/geb.12179>
- Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., Börger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Diaz, S., Echeverria-Londoño, S., Edgar, M. J., Feldman, A., Garon, M., Harrison, M. L., Alhusseini, T., ... Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520, 45–50. <https://doi.org/10.1038/nature14324>
- Normander, B., Levin, G., Auvinen, A.-P., Bratli, H., Stabbetorp, O., Hedblom, M., Glimskär, A., & Gudmundsson, G. A. (2012). Indicator framework for measuring quantity and quality of biodiversity—Exemplified in the Nordic countries. *Ecological Indicators*, 13(1), 104–116. <https://doi.org/10.1016/j.ecolind.2011.05.017>
- Owen, N. R., Gumbs, R., Gray, C. L., & Faith, D. P. (2019). Global conservation of phylogenetic diversity captures more than just functional diversity. *Nature Communications*, 10(1), 859. <https://doi.org/10.1038/s41467-019-08600-8>
- Pantanowitz, A., & Marwala, T. (2009). Missing data imputation through the use of the random Forest algorithm. In *Advances in computational intelligence* (Vol. 116, pp. 53–62). Springer.
- Pardieck, K. L., Ziolkowski, D. J., Jr., Lutmerding, M., Aponte, V., & Hudson, M.-A. R. (2020). North American breeding bird survey data set 1966–2019. Geological Survey, Patuxent Wildlife Research Center. <https://doi.org/10.5066/P9J6QUF6>
- Park, Y.-S., Chon, T.-S., Bae, M.-J., Kim, D.-H., & Lek, S. (2018). Multivariate data analysis by means of self-organizing maps. In *Ecological informatics* (pp. 251–272). Springer.
- Park, Y.-S., Tison, J., Lek, S., Giraudel, J.-L., Coste, M., & Delmas, F. (2006). Application of a self-organizing map to select representative species in multivariate analysis: A case study determining diatom distribution patterns across France. *Ecological Informatics*, 1(3), 247–257. <https://doi.org/10.1016/j.ecoinf.2006.03.005>
- Pascual, U., Adams, W. M., Diaz, S., Lele, S., Mace, G. M., & Turnhout, E. (2021). Biodiversity and the challenge of pluralism. *Nature Sustainability*, 4(7), 567–572. <https://doi.org/10.1038/s41893-021-00694-7>
- Pavoine, S., Vallet, J., Dufour, A.-B., Gachet, S., & Daniel, H. (2009). On the challenge of treating various types of variables: Application for improving the measurement of functional diversity. *Oikos*, 118(3), 391–402. <https://doi.org/10.1111/j.1600-0706.2008.16668.x>
- Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I. C., Clark, T. D., Colwell, R. K., Danielsen, F., Evengård, B., Falconi, L., Ferrier, S., Frusher, S., Garcia, R. A., Griffis, R. B., Hobday, A. J., Janion-Scheepers, C., Jarzyna, M. A., Jennings, S., ... Williams, S. E. (2017). Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science*, 355(6332), eaai9214. <https://doi.org/10.1126/science.aai9214>
- Penczak, T., Głowacki, Ł., Kruk, A., & Galicka, W. (2012). Implementation of a self-organizing map for investigation of impoundment impact on fish assemblages in a large, lowland river: Long-term study. *Ecological Modelling*, 227, 64–71. <https://doi.org/10.1016/j.ecolmodel.2011.12.006>
- Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., & Costa, G. C. (2014). Imputation of missing data in life-history trait data sets: Which approach performs the best? *Methods in Ecology and Evolution*, 5(9), 961–970. <https://doi.org/10.1111/2041-210x.12232>
- Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H. M., Cardoso, A. C., Coops, N. C., Dulloo, E., Faith, D. P., Freyhof, J., Gregory, R. D., Heip, C., Höft, R., Hurtt, G., Jetz, W., ... Wegmann, M. (2013). Essential biodiversity variables. *Science*, 339(6117), 277–278. <https://doi.org/10.1126/science.1229931>
- Petchey, O. L., & Gaston, K. J. (2006). Functional diversity: Back to basics and looking forward. *Ecology Letters*, 9(6), 741–758. <https://doi.org/10.1111/j.1461-0248.2006.00924.x>
- Rao, C. R. (1982). Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology*, 21(1), 24–43. [https://doi.org/10.1016/0040-5809\(82\)90004-1](https://doi.org/10.1016/0040-5809(82)90004-1)
- Rapacciolo, G., Graham, C. H., Marin, J., Behm, J. E., Costa, G. C., Hedges, S. B., Helmus, M. R., Radeloff, V. C., Young, B. E., & Brooks, T. M. (2019). Species diversity as a surrogate for conservation of phylogenetic and functional diversity in terrestrial vertebrates across the Americas. *Nature Ecology & Evolution*, 3(1), 53–61. <https://doi.org/10.1038/s41559-018-0744-7>
- Sano, E. E., Rodrigues, A. A., Martins, E. S., Bettiol, G. M., Bustamante, M. M. C., Bezerra, A. S., Couto, A. F., Jr., Vasconcelos, V., Schüler, J., & Bolfe, E. L. (2019). Cerrado ecoregions: A spatial framework to assess and prioritize Brazilian savanna environmental diversity for conservation. *Journal of Environmental Management*, 232, 818–828. <https://doi.org/10.1016/j.jenvman.2018.11.108>
- Sax, D. F., & Gaines, S. D. (2003). Species diversity: From global decreases to local increases. *Trends in Ecology & Evolution*, 18(11), 561–566. [https://doi.org/10.1016/S0169-5347\(03\)00224-6](https://doi.org/10.1016/S0169-5347(03)00224-6)
- Skidmore, A. K., Coops, N. C., Neinavaz, E., Ali, A., Schaepman, M. E., Paganini, M., Kissling, W. D., Vihervaara, P., Darvishzadeh, R., Feilhauer, H., Fernandez, M., Fernández, N., Gorelick, N., Geijzendorffer, I., Heiden, U., Heurich, M., Hobern, D., Holzwarth, S., Muller-Karger, F. E., ... Wingate, V. (2021). Priority list of biodiversity metrics to observe from space. *Nature Ecology & Evolution*, 5(7), 896–906. <https://doi.org/10.1038/s41559-021-01451-x>
- Smith, J. R., Letten, A. D., Ke, P.-J., Anderson, C. B., Hendershot, J. N., Dhami, M. K., Dlott, G. A., Grainger, T. N., Howard, M. E., Morrison, B. M. L., Routh, D., San Juan, P. A., Mooney, H. A., Mordecai, E. A., Crowther, T. W., & Daily, G. C. (2018). A global test of ecoregions. *Nature Ecology & Evolution*, 2, 1889–1896. <https://doi.org/10.1038/s41559-018-0709-x>
- Soroye, P., Newbold, T., & Kerr, J. (2020). Climate change contributes to widespread declines among bumble bees across continents. *Science*, 367(6478), 685–688. <https://doi.org/10.1126/science.aax8591>
- Tilman, D. (1982). *Resource competition and community structure*: Princeton university press.
- Tscharntke, T., Tylianakis, J. M., Rand, T. A., Didham, R. K., Fahrig, L., Batáry, P., Bengtsson, J., Clough, Y., Crist, O., Dormann, C. F., Ewers, R. M., Fründ, J., Holt, R. D., Holzschuh, A., Klein, A. M., Kleijn, D., Kremen, C., Landis, D. A., Laurance, W., ... Westphal, C. (2012). Landscape moderation of biodiversity patterns and processes - eight hypotheses. *Biological Reviews*, 87(3), 661–685. <https://doi.org/10.1111/j.1469-185x.2011.00216.x>
- Tucker, C. M., Davies, T. J., Cadotte, M. W., & Pearse, W. D. (2018). On the relationship between phylogenetic diversity and trait diversity. *Ecology*, 99(6), 1473–1479. <https://doi.org/10.1002/ecy.2349>
- Valdez, J. W., Callaghan, C. T., Junker, J., Purvis, A., Hill, S. L. L., & Pereira, H. M. (2023). The undetectability of global biodiversity trends using local species richness. *Ecography*, 2023(3), e06604. <https://doi.org/10.1111/ecog.06604>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 67. <https://doi.org/10.18637/jss.v045.i03>

- Vane-Wright, R. I., Humphries, C. J., & Williams, P. H. (1991). What to protect?—Systematics and the agony of choice. *Biological Conservation*, 55(3), 235–254. [https://doi.org/10.1016/0006-3207\(91\)90030-D](https://doi.org/10.1016/0006-3207(91)90030-D)
- Vellend, M., Cornwell, W. K., Magnuson-Ford, K., & Mooers, A. Ø. (2011). *Measuring phylogenetic biodiversity. Biological diversity: Frontiers in measurement and assessment* (pp. 194–207). Oxford University Press.
- Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, 11(3), 586–600. <https://doi.org/10.1109/72.846731>
- Villamuelas, M., Fernández, N., Albanell, E., Gálvez-Cerón, A., Bartolomé, J., Mentaberre, G., López-Olvera, J. R., Fernández-Aguilar, X., Colom-Cadena, A., López-Martín, J. M., Pérez-Barbería, J., Garel, M., Marco, I., & Serrano, E. (2016). The enhanced vegetation index (EVI) as a proxy for diet quality and composition in a mountain ungulate. *Ecological Indicators*, 61, 658–666. <https://doi.org/10.1016/j.ecolind.2015.10.017>
- Willig, M. R., Kaufman, D. M., & Stevens, R. D. (2003). Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics*, 34(1), 273–309. <https://doi.org/10.1146/annurev.ecolsys.34.012103.144032>
- Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. *Ecology*, 95(7), 2027. <https://doi.org/10.1890/13-1917.1>
- Winter, M., Devictor, V., & Schweiger, O. (2013). Phylogenetic diversity and nature conservation: Where are we? *Trends in Ecology & Evolution*, 28(4), 199–204. <https://doi.org/10.1016/j.tree.2012.10.015>
- Yu, D., Liu, Y., Shi, P., & Wu, J. (2019). Projecting impacts of climate change on global terrestrial ecoregions. *Ecological Indicators*, 103, 114–123. <https://doi.org/10.1016/j.ecolind.2019.04.006>
- Zhang, W., Sheldon, B. C., Grenyer, R., & Gaston, K. J. (2021). Habitat change and biased sampling influence estimation of diversity trends. *Current Biology*, 31(16), 3656–3662.e3653. <https://doi.org/10.1016/j.cub.2021.05.066>

BIOSKETCH

Wenyuan Zhang is interested in understanding how biodiversity is changing across the planet and factors that may influence our understanding on such processes. Particularly, he is interested in exploring biodiversity dynamics and connectivity and how personalised decisions may influence their estimation and conservation. One can find more about Wenyuan's research in <https://wenyuanzhang.com>.

Author contributions: WZ conceived the ideas and all authors contributed to the methodology design; WZ analysed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhang, W., Grenyer, R., Gaston, K. J., & Sheldon, B. C. (2024). Challenging the geographic bias in recognising large-scale patterns of diversity change. *Diversity and Distributions*, 30, 13–25. <https://doi.org/10.1111/ddi.13775>