scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Multi-scale environmental geodiversity: data for the National Ecological Observatory Network (NEON) with an adaptable workflow

Kelly Kapsar[1,2,3] ✉, Lala Kounta[1,3], Patrick Bills[2,4], Annie Smith[5], Sydne Record[6,7], Angela Strecker[8,9], Benjamin Baiser[10] & Phoebe Zarnetske[1,2,3] ✉

Spatial variability in environmental conditions is fundamental to explaining and predicting patterns of biodiversity. However, most models explaining patterns of biodiversity use mean values collected at one scale rather than measures of spatial heterogeneity, or "geodiversity", of environmental drivers across multiple scales. Furthermore, accessing and using large, high resolution environmental data sets is often challenging for researchers not trained in supercomputing. To help overcome this barrier, we present an open, reproducible, and adaptable workflow to extract and summarize environmental geodiversity at varying spatial scales. We also provide ready-to-use data on climate and elevation geodiversity metrics and mean values at plot, site, and domain scales for National Ecological Observatory Network (NEON) sites across the United States. The final data set contains over 20,400 tabular values organized into 13 spatial layers. These analysis-ready data, and associated R workflow, will offer opportunities to examine the multiscalar effects of climate and elevation geodiversity on ecological processes.

Background & Summary

Uncovering the relationships between environmental drivers (e.g., temperature, precipitation, elevation) and resulting biodiversity and ecosystem functions has long been a key area of inquiry in the field of ecology[1,2]. These biodiversity-environment relationships are driven by processes occurring across local to global scales—ranging from heterogeneity in micro-habitat structure to climate and land-use change[3–5]. For example, scale-dependent climate variables are essential for explaining how microclimate vs. regional climate affect numerous species' distributions[3,6–8].

Of the many environmental variables used to study biodiversity-environment relationships, temperature, precipitation, and elevation metrics are essential for understanding ecosystem processes, predicting responses to climate change, and informing management decisions. Spatial climate data are key drivers for models and analyses, but their utility depends on the scale at which they are collected and applied[9]. Fine-scale data (sub-kilometer) can capture critical features such as elevational gradients, coastal effects, and land use impacts, while coarse-scale (multi-kilometer) data may overlook these local nuances, potentially leading to misinterpretation or error in broader-scale analyses[9]. Similarly, topography plays an important role in determining species

[1]Department of Integrative Biology, Michigan State University, East Lansing, USA. [2]Institute for Biodiversity, Ecology, Evolution, and Macrosystems, Michigan State University, East Lansing, USA. [3]Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, USA. [4]Institute for Cyber Enabled Research, Michigan State University, East Lansing, USA. [5]Washington Department of Natural Resources, Olympia, USA. [6]Department of Wildlife, Fisheries, and Conservation Biology, University of Maine, Orono, USA. [7]Maine Agricultural and Forest Experiment Station US, Orono, USA. [8]Institute for Watershed Studies, Western Washington University, Bellingham, USA. [9]Department of Environmental Sciences, Western Washington University, Bellingham, USA. [10]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, USA. ✉e-mail: kapsarke@msu.edu; plz@msu.edu

distributions[10]. Topographic complexity is associated with increased biodiversity through complex and poorly understood mechanisms that include climate-topography interactions[10,11].

This complexity and heterogeneity of the abiotic environment is known as "geodiversity". While the term is sometimes applied to exclusively refer to geologic diversity (e.g., variation in hydrologic and elevational patterns)[12,13], the definition of geodiversity has also been extended to other abiotic features and processes of Earth including atmosphere, lithosphere, cryosphere, and hydrosphere[14–17].Here, we apply the broader definition, including both climatological and geological aspects of abiotic heterogeneity. In this context, geodiversity provides the abiotic foundation for species' habitats[18], and has scale-dependent effects on patterns of biodiversity, habitat suitability, and species' distributions[14,19,20].

Measuring the effects of geodiversity is a challenging task given that most analyses of environmental drivers of biodiversity rely solely on spatial averages or category-based metrics (e.g., patch metrics on land cover). These approaches inherently collapse heterogeneity into single values, thus limiting the ability to quantify potentially important dimensions of heterogeneity in geodiversity variables[14,20]. Originating in the field of surface metrology, geodiversity metrics are gradient surface metrics that characterize the variability in continuous surfaces beyond just mean and standard deviation, measuring patterns such as the degree of peakedness or the degree of directionality in a surface[21,22]. Measuring the variability of high resolution geodiversity data using geodiversity metrics such as surface skewness (ssk) can, for example, help researchers to quantify the prevalence of different microhabitats or fine-scale climate refugia on a landscape[23,24]. These factors are important for measuring the impacts of climate change on species' presence on a landscape.

Large, coordinated research observatories that span multiple sites and regions can facilitate cross-scale analyses that explain and predict biological patterns and processes, including biodiversity[25–27]. The National Ecological Observatory Network (NEON) is perhaps the largest example of such a network, and was designed with the recognition that emergent ecological phenomena cannot be fully explained by single-scale analyses; instead, cross-scale interactions must be acknowledged and incorporated[28–30]. NEON's hierarchical ecological observatory spans 20 climate domains in the United States, with each domain including multiple terrestrial and freshwater sites. These sites include multiple biotic and abiotic plots nested within different land cover types[31] (Fig. 1). Within these sites, NEON staff collect standardized data on sentinel organisms, airborne remote sensing gathers high resolution information about vegetation and topography, and *in-situ* instruments sense climatic and atmospheric variables. Through this effort, NEON produces 181 open-access data products and 36 different types of physical specimens that are stored in the NEON Biorepository to enable innovations in environmental science[26]. Such a nested spatial sampling design is crucial for identifying both proximal and distal drivers of biodiversity, such as how landscape fragmentation interacts with climate gradients to affect species richness and composition[32].

To analyze biodiversity-environment associations, researchers need access to analysis-ready environmental data spanning multiple metrics and scales[27]. With increasing availability of open network observations and remotely sensed data, abiotic and biotic data spanning scales are becoming more available and useful. However, such data are often difficult to work with due to their large size and require technical expertise in supercomputing and big data analysis to process. These factors can impede research and limit the applicability of these data to ecological analyses[33,34]. We help fill these gaps by providing multi-scale, analysis-ready environmental geodiversity data for climate (temperature and precipitation) and elevation at terrestrial NEON sites across the United States. We follow best practices for NEON-related derived data products[35], and provide the open workflow and vignette in R for replicating the multi-scale mean and geodiversity spatial data sets at other sites around the world.

Methods

Study area. NEON is a United States National Science Foundation funded program designed to promote long-term, multi-scale research on the impacts of climate change, land use change, and invasive species on ecological processes across the United States[31]. NEON's nested design consists of *plots* nested within *sites*, which are nested within *domains*. Within NEON's 20 eco-climatic domains are 47 terrestrial observation sites. Within each terrestrial site, standardized data on sentinel taxa, infectious disease prevalence, and nutrient and carbon cycling are collected from plots using standardized protocols. The plots are placed on different landcover types within the site[36]. For more information on NEON's spatial sampling designs, see Thorpe *et al.*[31].

The example data and workflow presented here use data from multiple scales, including: all 20 NEON domains, all 47 NEON terrestrial sites and, within each site, the small mammal plots (trapping arrays) and NEON towers (which collect meteorological, air chemistry, and eddy covariance data). While the small mammal plots and towers are provided as examples, other plot scale data can be extracted and analyzed for geodiversity as well by following the workflow and provided code.

Geodiversity variables. To capture the spatial heterogeneity of abiotic variables at NEON sites, we calculated metrics of geodiversity on both climate and elevation data using the geodiv R package[37]. Of the approximately 26 metrics available in the geodiv package, we chose a representative sample of 7 metrics for their ecological interpretability[21] (Table 1). These metrics encompass the three surface property "families": amplitude metrics (Sq, Ssk, Sku), surface-bearing metrics (Sbi), and spatial metrics (Sdq, Stdi, Sds), which characterize both landscape composition (in the case of amplitude metrics and surface bearing metrics) and configuration (in the case of spatial metrics)[22]. The metrics are also from three of the four clusters of metrics identified in McGarigal *et al.*[22]: surface roughness (Sq, Sdq), surface height distribution (Sbi, Ssk, Sds), and angular texture (Stdi).

Data sources. *Climate data.* We used a set of 19 bioclimatic variables (BCVs) at high resolution from the Climatologies at high resolution for the earth's land surface areas (CHELSA) initiative[38]. These data have a spatial

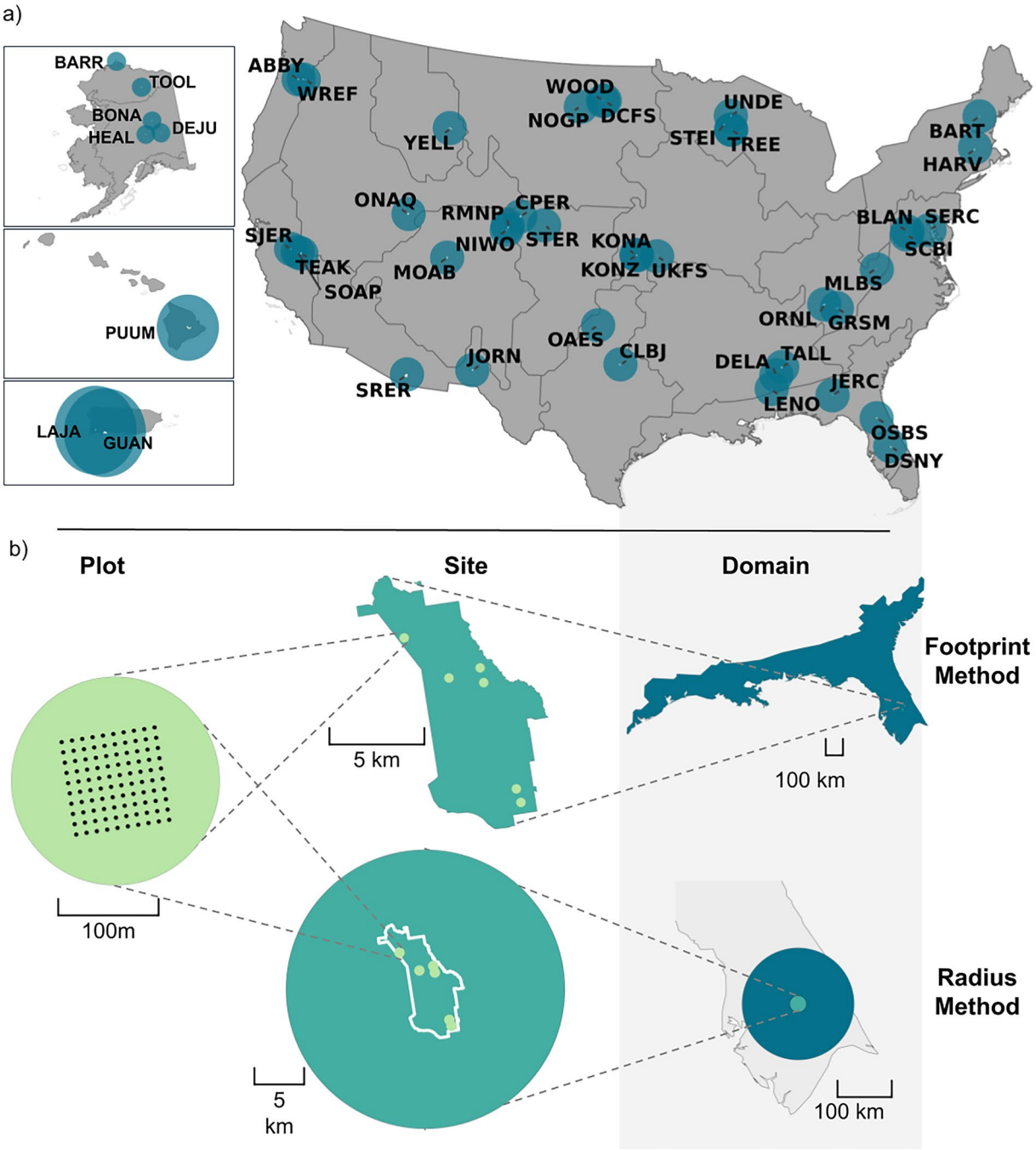


Fig. 1 a) A map of the distribution of NEON domains. Gray outlines represent NEON domain footprints. Blue circles represent NEON domain radii. Light gray background illustrates the use of the Southeast Domain as an example for the schematic representation in the lower portion of the Figure. The contiguous United States is displayed in NAD83/Conus Albers (EPSG 5070); inset maps for Alaska, Hawaii, and Puerto Rico are displayed in region-appropriate equal-area projections (EPSG 3338, 3759, and 32161, respectively) for visual accuracy. All geodiversity analyses were conducted in EPSG 5070. b) A schematic representation of the spatial extents of the data using the Disney (DSNY) NEON site in Florida, USA. Plot level data were generated at a 100 m radius around the center of each small mammal array or the location of a tower at each site. Site and domain data generated using the footprint method (top row) used the site and domain spatial footprints for climate and geodiversity data extraction. Site and domain data generated using the radius method (bottom row) were created by drawing concentric circles at a 15 km radius (site scale) and 100 km radius (domain scale) around either the center of all small mammal plots or the tower location at a given site.

Metric	Name	Surface property family (McGarigal *et al.* 2009 supp)	Category (McGarigal *et al.* 2009; Smith *et al.* 2020)
Sq	Root mean square roughness	Amplitude	Roughness
Sdq	Root mean square slope	Spatial	Roughness
Sbi	Surface bearing index	Surface bearing	Distribution
Ssk	Surface skewness	Amplitude	Distribution
Sku	Surface kurtosis	Amplitude	Distribution
Stdi	Texture direction index	Spatial	Angular
Sds	Summit density	Spatial	Distribution

Table 1. Description of geodiversity metrics, their surface property families and categories.

resolution of 30 arcsec (i.e., less than 1 km). CHELSA data are widely used for climatological and ecological analyses[39,40]. These data consist of the climatological means of the 19 BCVs from the 30-year normal period, 1981–2009, and describe a range of seasonal and annual statistics of precipitation and temperature in a standardized format[41] (see neonEnvData_BCV_description_table.csv in the EDI repository for variable definitions).

Elevation data. To quantify elevation, we used data from NASA's Shuttle Radar Topography Mission (SRTM)[42]. These data represent one of the most complete elevation datasets globally, have high spatial resolution (1 arc-second, ~ 30 m), and extend from 56 degrees South to 60 degrees North latitude.

Spatiotemporal scales. Given the importance of spatial scale for identifying species- and biodiversity-environment relationships, we calculated mean values and geodiversity metrics across multiple scales, varying both grain size and extent (Fig. 1).

For the spatial extent of these data, we measured variables using both the spatial footprints provided by NEON (Fig. 1b, top row) and radii (Fig. 1b, bottom row). The use of both the radius and footprint methods to produce spatial geodiversity metrics provides flexible output useful for multiple applications. For example, the spatial footprint of NEON sites is often constrained by socio-political factors (e.g., land ownership), but environmental variables beyond the footprint boundaries may have an impact on biodiversity within these sites. Therefore, evaluating environmental characteristics solely within the site footprint may overlook neighboring elevation and/or climate factors that the radius method natively incorporates. In other words, radii-based data can help overcome biases that are generated when the environmental characteristics within NEON sites are not necessarily representative of the broader ecological context of which they are a part.

For spatial grain, we varied the spatial grain size of the elevation data from its native resolution (30 m) at fine spatial extents (plot, site) to a coarsened resolution (300 m) at broader spatial extents (site, domain). Elevation metrics were calculated at a 30 m and 300 m grain size at the site extent to allow for comparisons of the effects of grain size at a single extent.

To illustrate the workflow for the *plot* scale, within each NEON site, we extracted climate and elevation data and calculated geodiversity metrics and means at two sets of plot locations: (1) 100 m radius circles around the centroids of each small mammal trapping array; and (2) a 100 m radius circle centered on the point locations of each tower.

For the radius-based extraction method at the *site* and *domain* scales, we constructed 15 km and 100 km radii, respectively, around the center of the minimum enclosing circle of small mammal traps or tower locations within each site. The site scale radius (15 km) was specifically chosen to encompass the entire NEON Airborne Observation Platform flight path at each site. For the footprint-based extraction method, we used the spatial footprint of the sites and domains[43].

Data processing and reproducible workflow. To generate the output file, we first cropped the raw SRTM and CHELSA data to the specified spatial extent (in this case the extent of the NEON domains) and projected to the specified projection (EPSG:5070; Conus Albers; Fig. 2, scripts 1a & 1b). To remove erroneous values in elevation data, we set a minimum elevation of −20 m. All pixels with elevation values less than −20 m were converted to NA prior to extraction and analysis. To ensure that all radii remained equally sized, we did not clip the spatial extents or mask values (for example, for radii extending into Mexico (e.g., SRER; Fig. 1) or into the ocean (e.g.,OSBS)). We then spatially intersected the climate and elevation rasters with our spatial data (see "Spatiotemporal Scales" above) and extracted mean elevation values and climate metric values for all 19 BCVs (Table 1; Fig. 2, script 2). For elevation data and for Bios 1 (annual mean temperature) and 12 (sum of annual precipitation), we also used the geodiv package to calculate each geodiversity metric (Table 1; Fig. 2, script 3).

We conducted all analyses using R version 4.3.2[44]. To accommodate the large file sizes associated with the high resolution elevation data, we processed these data in parallel on MSU's High Performance Computing Cluster. To facilitate reproducibility, all slurm submission files are included in the archived project code. We recognize that there are cloud computing options that avoid the need to download large datasets (e.g., SRTM on Google Earth Engine). We chose not to build this workflow on a proprietary computing platform because companies have the right to begin charging for hosting and compute at any time making such an approach is risky for long-term research efforts and open science.

While the data set presented in this paper is focused on NEON sites as an example, our workflow can be easily adapted to data from any location around the world. To generate output, users must (1) download SRTM and

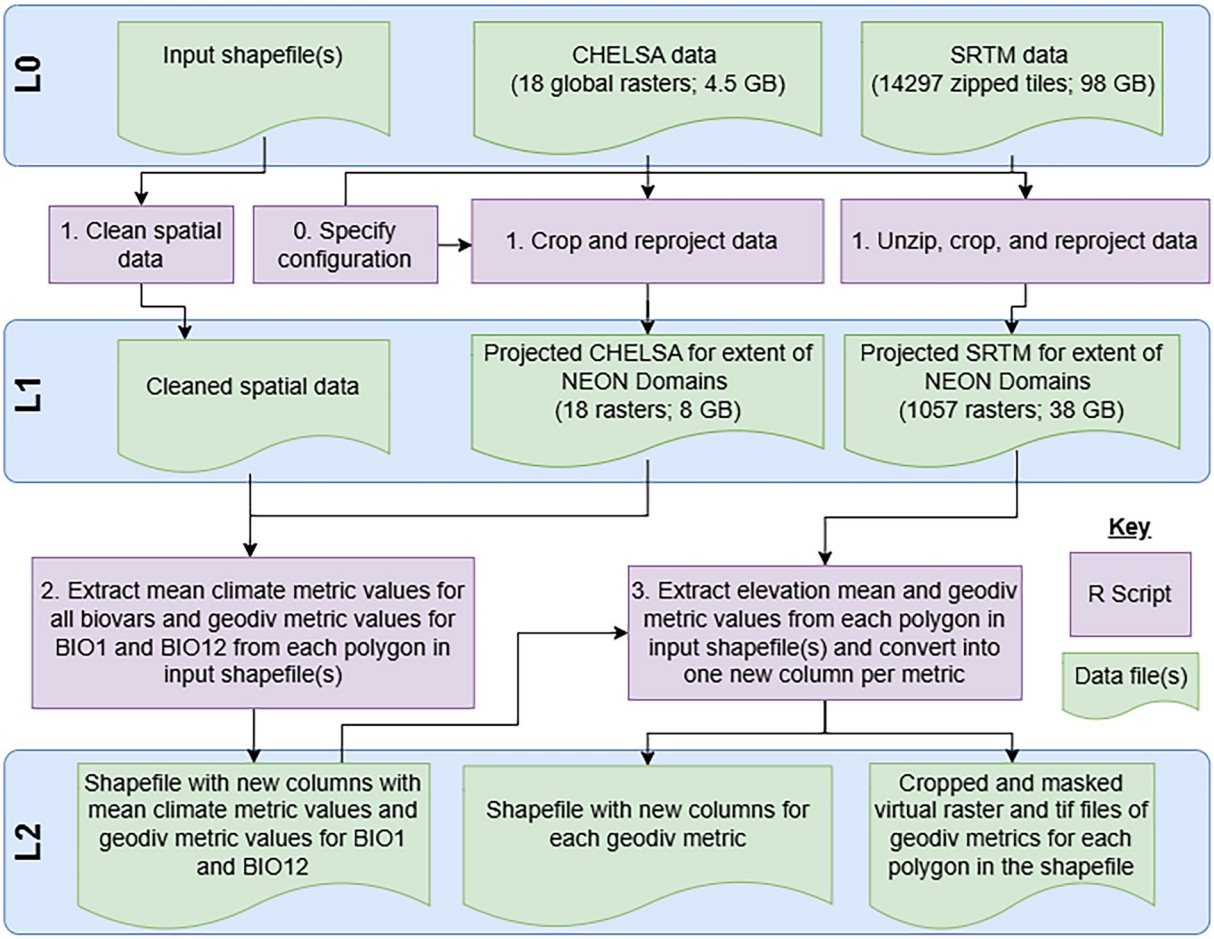


Fig. 2 Schematic workflow for extraction of environmental geodiversity metrics and means at NEON plot, site, and domain scales, using climate and elevation data. Purple boxes represent scripts in the codebase (i.e., archived project code) with script numbers matching those in the codebase itself. Blue boxes represent levels of data organization following the Environmental Data Initiative standards. L0 data are raw, unprocessed data. L1 data are cleaned, and L2 data are integrated and analyzed.

CHELSA data for desired spatial extent, (2) specify the appropriate spatial projection, spatial extent of the data, desired geodiversity metrics, project directory, resolution of the elevation data, and buffer distances for radii at each scale in the config.R file, (3) provide a spatial vector object (with an identification column) for which they would like to calculate climate and elevation metrics, and (4) run the scripts in the specified order (Fig. 2). The code itself will generate the necessary file structure and the final script will produce a geopackage file with calculated metrics. For smaller spatial extents, not all raw climate and elevation files need to be downloaded, likely eliminating the need for the use of a supercomputer to process the data.

Data Records

Output data are archived with the Environmental Data Initiative as geopackage files with one file per unique combination of spatial grain, spatial extent, and sampling design method for a total of 13 geopackage files (Table 2)[45]. Files are named using the following naming convention: "NEON_<focal_location (for radius method)>_<spatial extent>_<analysis method>_<elevation data resolution>.gpkg (Ex: "NEON_mammal_site_radii_elev300m.gpkg". Rows in each file represent individual spatial units (e.g., unique plot, site, domain) and each metric value is contained within its own column. Column names follow the format of "data source_metric". For example, "srtm_sbi" contains the surface bearing index value for the srtm elevation data, and "bio01_mean" contains the spatial mean value of the Bio 1 variable (annual mean temperature). We chose geopackage as the output file format because of its open source, platform-independent nature as well as its flexible ability to store both raster and vector data[46]. Like shapefiles, geopackage files are also compatible with commonly used R packages for geospatial analysis (e.g., sf, terra).

SRTM elevation data were not available for five NEON sites at high latitudes (i.e.,Caribou-Poker Creeks (BONA), Healy (HEAL), Toolik (TOOL), Utqiaġvik (BARR), Delta Junction (DEJU)) and were given NA values for elevation-based metrics. At all sites, geodiversity metrics for climate data at the plot scale also have NA values because the grain size of the climate data was larger than the spatial extent of the plots, meaning that only one climate pixel was present per plot and thus no variability in climate metrics was present to measure.

Focal Area	Spatial area	Spatial Scale		
		Extent	Resolution (Elevation data; meters)	Resolution (Climate data; arcsec)
Tower	Radius	Plot	30	30
Tower	Radius	Site	30	30
Tower	Radius	Site	300	30
Tower	Radius	Domain	30	30
Tower	Radius	Domain	300	30
Small mammal plot	Radius	Plot	30	30
Small mammal plot	Radius	Site	30	30
Small mammal plot	Radius	Site	300	30
Small mammal plot	Radius	Domain	30	30
Small mammal plot	Radius	Domain	300	30
—	Footprint	Site	30	30
—	Footprint	Site	300	30
—	Footprint	Domain	300	30

Table 2. Description of focal area, spatial area, and spatial scale for each of the 13 files in the dataset.

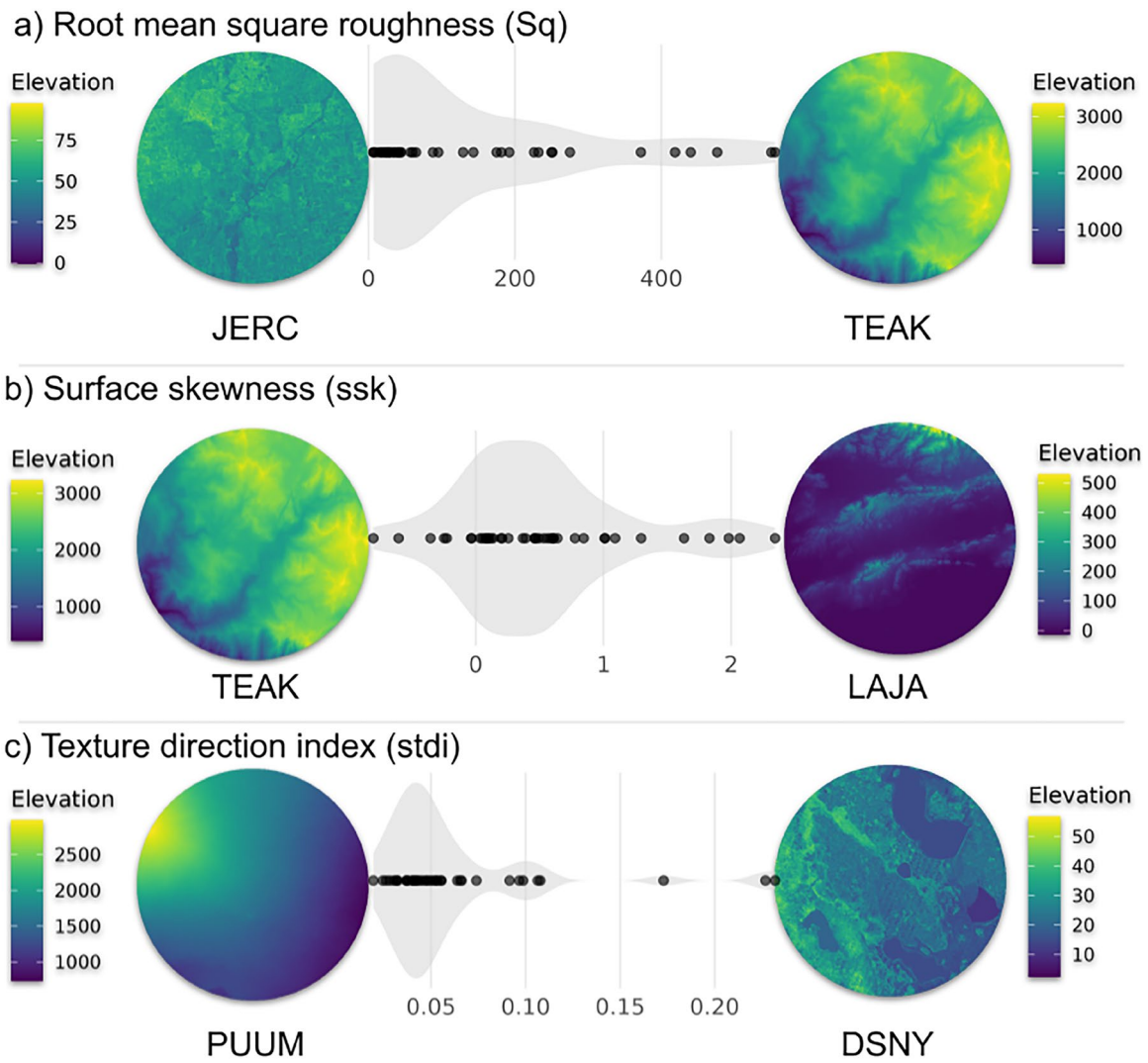


Fig. 3 Violin plots of three elevation geodiversity metrics across NEON sites with example elevation maps for sites at the extreme (maximum and minimum) values of each metric. (**a**) Root mean square roughness (sq) is equivalent to standard deviation with high values indicating a larger variability in elevation and lower values indicating less variability. (**b**) Surface skewness (ssk) identifies peaks and valleys in a distribution. Low values indicate a surface with primarily valleys while high values indicate the presence of peaks. (**c**) Texture direction index (stdi) measures the degree of directionality in the surface. Surfaces with greater overall directionality have a lower stdi value while surfaces less directionality have higher values.

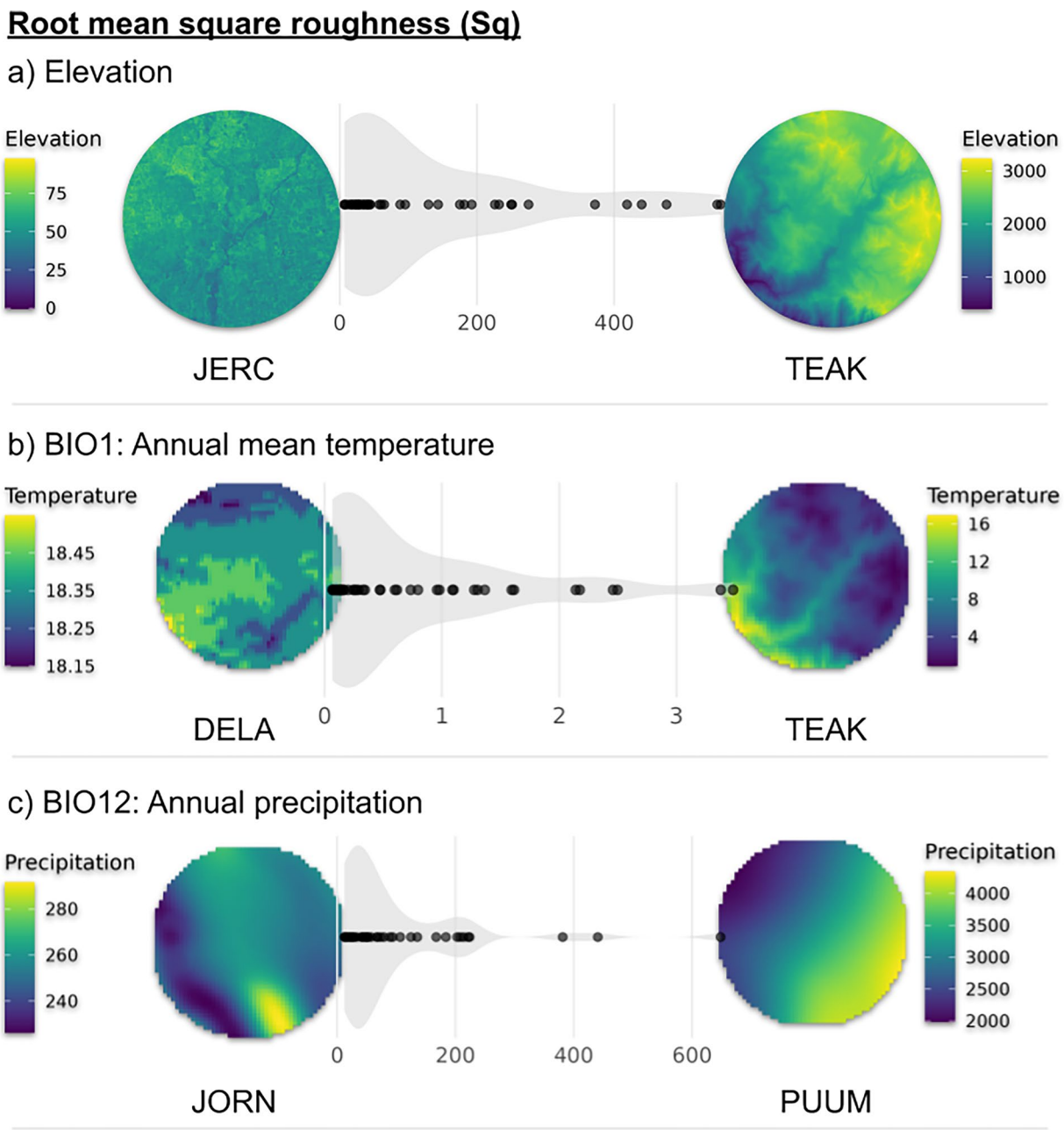


Fig. 4 Violin plots of root mean square roughness (sq) across NEON sites for three variables: (**a**) elevation, (**b**) annual mean temperature (degrees Celsius), and (**c**) annual precipitation (mm). Example maps from the sites with the most extreme values of each variable are plotted on either side of the violin plot. Note that differences in color scales between plots overrepresent variability in areas with less overall variation (e.g., precipitation in JORN).

Technical Validation

To evaluate the validity of output data, we generated a summary table containing the minimum, maximum, and mean values as well as the total number of zero and NA values for each unique scale and metric calculation (archived as neonEnvData_gpkg_summary_data.csv in EDI repository). We also generated plots of the distribution of values for each metric for both climate and elevation data (see examples in Figs. 3, 4). We found that, aside from the expected NA values (described in the previous paragraph), approximately 0.3% of the total data contained unexplained NA values. The vast majority of these unexpected NA values were for the surface texture direction index (stdi) metric. After inspection, we concluded that many of these NA values were likely caused by irregularities in the shape of the polygons (e.g., site footprints) causing challenges with moving window calculations for this spatial metric. Other NA values were likely generated due to pixel edge effects (e.g., plot location fell outside the edge of a pixel in coastal zones).

Data availability

The datasets presented in this paper are archived with the Environmental Data Initiative under a CC-BY license and are available at the following URL: https://doi.org/10.6073/pasta/98440675e6a02f3e33c6a87d93b757e3.

Code availability

The code and vignette presented in this paper are archived alongside the data at the Environmental Data Initiative.

Received: 22 October 2025; Accepted: 4 June 2026;
Published online: 15 July 2026

References

1. Hallett, T. B. *et al*. Why large-scale climate indices seem to predict ecological processes better than local weather. *Nature* **430**, 71–75 (2004).
2. Sandel, B. Towards a taxonomy of spatial scale-dependence. *Ecography* **38**, 358–369 (2015).
3. McGill, B. J. Matters of Scale. *Science* **328**, 575–576 (2010).
4. Soranno, P. A. *et al*. Cross-scale interactions: quantifying multi-scaled cause–effect relationships in macrosystems. *Frontiers in Ecol & Environ* **12**, 65–73 (2014).
5. Mitchell, M. G. E. *et al*. Key questions for understanding drivers of biodiversity-ecosystem service relationships across spatial scales. *Landsc Ecol* **39**, 36 (2024).
6. Levin, S. A. Multiple Scales and the Maintenance of Biodiversity. *Ecosystems* **3**, 498–506 (2000).
7. Garcia, R. A., Cabeza, M., Rahbek, C. & Araújo, M. B. Multiple Dimensions of Climate Change and Their Implications for Biodiversity. *Science* **344**, 1247579 (2014).
8. Moudrý, V. *et al*. Scale mismatches between predictor and response variables in species distribution modelling: A review of practices for appropriate grain selection. *Progress in Physical Geography: Earth and Environment* **47**, 467–482 (2023).
9. Daly, C. Guidelines for assessing the suitability of spatial climate data sets. *Intl Journal of Climatology* **26**, 707–721 (2006).
10. Rahbek, C. *et al*. Humboldt's enigma: What causes global patterns of mountain biodiversity? *Science* **365**, 1108–1113 (2019).
11. Badgley, C. *et al*. Biodiversity and Topographic Complexity: Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution* **32**, 211–226 (2017).
12. Schrodt, F. *et al*. The status and future of essential geodiversity variables. *Phil. Trans. R. Soc. A*. **382**, 20230052 (2024).
13. Gray, M. Defining Geodiversity. in *Geodiversity: Valuing and Conserving Abiotic Nature* (John Wiley & Sons, Incorporated, 2013).
14. Zarnetske, P. L. *et al*. Towards connecting biodiversity and geodiversity across scales with satellite remote sensing. *Global Ecol Biogeogr* **28**, 548–556 (2019).
15. Benito-Calvo, A., Pérez-González, A., Magri, O. & Meza, P. Assessing regional geodiversity: the Iberian Peninsula. *Earth Surf Processes Landf* **34**, 1433–1445 (2009).
16. Parks, K. E. & Mulligan, M. On the relationship between a resource based measure of geodiversity and broad scale biodiversity patterns. *Biodivers Conserv* **19**, 2751–2766 (2010).
17. Record, S. *et al*. Remote sensing of geodiversity as a link to biodiversity. in *Remote sensing of plant biodiversity* 225–253 (Springer, 2020).
18. Hjort, J., Gordon, J. E., Gray, M. & Hunter, M. L. Why geodiversity matters in valuing nature's stage: Why Geodiversity Matters. *Conservation Biology* **29**, 630–639 (2015).
19. Bailey, J. J., Boyd, D. S. & Field, R. Models of upland species' distributions are improved by accounting for geodiversity. *Landscape Ecol* **33**, 2071–2087 (2018).
20. Gerstner, B. E., Blair, M. E., Bills, P. & Zarnetske, P. L. The influence of scale-dependent geodiversity on species distribution models in a biodiversity hotspot. *Philosophical Transactions of the Royal Society A* **382**, (2024).
21. Kedron, P. J., Frazier, A. E., Ovando-Montejo, G. A. & Wang, J. Surface metrics for landscape ecology: a comparison of landscape models across ecoregions and scales. *Landscape Ecol* **33**, 1489–1504 (2018).
22. McGarigal, K., Tagil, S. & Cushman, S. A. Surface metrics: an alternative to patch metrics for the quantification of landscape structure. *Landscape Ecol* **24**, 433–450 (2009).
23. Austin, M. P. & Van Niel, K. P. Improving species distribution models for climate change studies: Variable selection and scale. *Journal of Biogeography* **38**, 1–8 (2011).
24. Keppel, G. *et al*. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography* **21**, 393–404 (2012).
25. Borer, E. T. *et al*. Finding generality in ecology: a model for globally distributed experiments. *Methods Ecol Evol* **5**, 65–73 (2014).
26. Nagy, R. C. *et al*. Harnessing the NEON data revolution to advance open environmental science with a diverse and data-capable community. *Ecosphere* **12**, e03833 (2021).
27. NASEM. *A Vision for Continental-Scale Biology: Research Across Multiple Scales*. 27285 https://nap.nationalacademies.org/catalog/27285, 10.17226/27285 (2025).
28. Schimel, D., Hargrove, W., Hoffman, F. & MacMahon, J. NEON: a hierarchically designed national ecological network. *Frontiers in Ecology and the Environment* **5**, 59–59 (2007).
29. Peters, D. P. *et al*. Living in an increasingly connected world: a framework for continental-scale environmental science. *Frontiers in Ecology and the Environment* **6**, 229–237 (2008).
30. Heffernan, J. B. *et al*. Macrosystems ecology: understanding ecological patterns and processes at continental scales. *Frontiers in Ecol & Environ* **12**, 5–14 (2014).
31. Thorpe, A. S. *et al*. Introduction to the sampling designs of the National Ecological Observatory Network Terrestrial Observation System. *Ecosphere* **7**, e01627 (2016).
32. Uhler, J. *et al*. Relationship of insect biomass and richness with land use along a climate gradient. *Nat Commun* **12**, 5946 (2021).
33. Hampton, S. E. *et al*. Skills and Knowledge for Data-Intensive Environmental Research. *BioScience* **67**, 546–557 (2017).
34. Mahood, A. L. *et al*. Ten simple rules for working with high resolution remote sensing data. *Peer Community Journal* **3**, e4 (2023).
35. Atkins, J. W. *et al*. Recommendations for developing, documenting, and distributing data products derived from NEON data. *Ecosphere* **16**, (2025).
36. Barnett, D. T. *et al*. The terrestrial organism and biogeochemistry spatial sampling design for the National Ecological Observatory Network. *Ecosphere* **10**, e02540 (2019).
37. Smith, A. C. *et al*. The geodiv rpackage: Tools for calculating gradient surface metrics. *Methods Ecol Evol* **12**, 2094–2100 (2021).
38. Karger, D. N. *et al*. Climatologies at high resolution for the earth's land surface areasCHELSA V2.1 (current). 2.1 KB EnviDat, https://doi.org/10.16904/ENVIDAT.228.V2.1 (2021).
39. Bobrowski, M., Weidinger, J. & Schickhoff, U. Is New Always Better? Frontiers in Global Climate Datasets for Modeling Treeline Species in the Himalayas. *Atmosphere* **12**, 543 (2021).
40. Ghosh, S., Matthews, B. & Petchey, O. L. Temperature and biodiversity influence community stability differently in birds and fishes. *Nat Ecol Evol* **8**, 1835–1846 (2024).

41. Karger, D. N. *et al.* Climatologies at high resolution for the earth's land surface areas. *Sci Data* **4**, 170122 (2017).
42. USGS. EROS Archive - Digital Elevation - Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global, https://doi.org/10.5066/F7PR7TFT (2018).
43. NEON (National Ecological Observatory Network). *Spatial Data & Maps*. https://www.neonscience.org/data-samples/data/spatial-data-maps (2025).
44. R Core Team. R: A Language and Environment for Statistical Computing. (2023).
45. Kapsar, K. *et al.* Multi-scale environmental geodiversity: data for the National Ecological Observatory Network (NEON) with an adaptable workflow ver 1. *Environmental Data Initiative* https://doi.org/10.6073/pasta/98440675e6a02f3e33c6a87d93b757e3 (2026).
46. Open Geospatial Consortium. *GeoPackage*. https://www.geopackage.org/ (2018).

Acknowledgements

We thank John Grady for feedback during early iterations of this work. This material is based in part upon work supported by the National Ecological Observatory Network (NEON), a program sponsored by the U.S. National Science Foundation (NSF) and operated under cooperative agreement by Battelle.

Author contributions

P.L.Z., S.R., A.S., A.S. and B.B. conceived of the idea. K.E.K., L.K. and P.B. collected the data, coded the analysis, and produced the results. K.E.K., L.K. and P.L.Z. wrote the manuscript text and all authors reviewed the manuscript.

Funding

This work was funded by the following NSF Division of Environmental Biology grants: #1926567, #1926568, #1926569, #1926610, and #2139918. SR was also supported by U.S. Department of Agriculture National Institute of Food and Agriculture Grant/Award Number MEO-022425. This work was supported in part through computational resources and services provided by the Institute for Cyber-Enabled Research at Michigan State University and used the MSU Data Machine, which is supported through the NSF Campus CyberInfrastructure program through grant #2200792.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-026-07613-5.

Correspondence and requests for materials should be addressed to K.K. or P.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026