The GEODIV R package: Tools for calculating gradient surface metrics

Annie C. Smith[1,2] ⓘ　|　Kyla M. Dahlin[2,3] ⓘ　|　Sydne Record[4] ⓘ　|　Jennifer K. Costanza[5] ⓘ　|　Adam M. Wilson[6] ⓘ　|　Phoebe L. Zarnetske[1,2] ⓘ

[1]Department of Integrative Biology, Michigan State University, East Lansing, MI, USA; [2]Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, MI, USA; [3]Department of Geography, Environment, and Spatial Sciences, Michigan State University, East Lansing, MI, USA; [4]Department of Biology, Bryn Mawr College, Bryn Mawr, PA, USA; [5]USDA Forest Service, Southern Research Station, Research Triangle Park, NC, USA and [6]Departments of Geography and Environment & Sustainability, University at Buffalo, Buffalo, NY, USA

Correspondence
Annie C. Smith
Email: annie.smith@dnr.wa.gov

Abstract

1. The GEODIV R package calculates gradient surface metrics from imagery and other gridded datasets to provide continuous measures of landscape heterogeneity for landscape pattern analysis.

2. GEODIV is the first open-source, command line toolbox for calculating many gradient surface metrics and easily integrates parallel computing for applications with large images or rasters (e.g. remotely sensed data). All functions may be applied either globally to derive a single metric for an entire image or locally to create a texture image over moving windows of a user-defined extent.

3. We present a comprehensive description of the functions available through GEODIV. A supplemental vignette provides an example application of GEODIV to the fields of landscape ecology and biogeography.

4. GEODIV allows users to easily retrieve estimates of spatial heterogeneity for a variety of purposes, enhancing our understanding of how environmental structure influences ecosystem processes. The package works with any continuous imagery and may be widely applied in many fields where estimates of surface complexity are useful.

KEYWORDS

geodiversity, gradient surface model, landscape ecology, R package, remote sensing, spatial heterogeneity, spatial pattern, surface metrics

1 | INTRODUCTION

Landscape patterns shape ecosystem characteristics (Turner, 1989, 2005; Von Humboldt & Bonpland, 2010), including biodiversity, disturbance and energy cycling (Uuemaa et al., 2009). For example, complex landscapes may harbour higher habitat diversity, promoting higher biodiversity (Kumar et al., 2006). Landscape heterogeneity has been explored both for illuminating ecosystem pattern–process

relationships (Kumar et al., 2006; St-Louis et al., 2006; Walz, 2011) and identifying temporal changes in ecosystem patterns (e.g. Stevens et al., 2017). Numerous methods for measuring heterogeneity have been developed and studied, with methods falling broadly into two categories: patch and gradient metrics.

Many studies addressing relationships between landscape heterogeneity and ecological processes have utilized metrics from the patch mosaic model (Turner, 2005). This model conceptualizes landscapes as a set of patches where each patch contains similar ecosystem characteristics (e.g. a burned area). Within this framework, patch metrics are used to describe these patches and their surrounding landscapes. These metrics are derived from categorical representations of land cover or discretized continuous variables which have been split into patches composed of pixels of the same class (McGarigal & Cushman, 2005). Metrics either describe individual patch characteristics, such as edge length (Helzer & Jelinski, 1999), or summarize the spatial configuration of patches across a landscape (Jaeger, 2000). The patch mosaic model and associated metrics are useful for representing categorical variables such as land cover class. However, most landscape features, and associated ecological processes, are continuous. Even categorical data like land cover maps are derived from imagery with continuous values. While the patch mosaic model excels at providing a simplified view of landscape heterogeneity (Turner, 2005), this approach can miss important features of continuous surfaces such as altitudinal temperature changes or the differences between smooth versus hilly landscapes (Cushman et al., 2010; Lausch et al., 2015).

An alternative to the patch mosaic model, the gradient surface model, reflects the continuous nature of landscapes (McGarigal et al., 2009). Gradient surface metrics (GSMs) originated in the fields of physics and materials science, and are used to describe the roughness of machined surfaces. The metrics were not originally developed for use in landscape ecology, and their application to the field is somewhat novel. Several recent papers have illustrated how these metrics link with both patch metrics and ecological processes. For example, average roughness, root mean square roughness and surface kurtosis can align with the Normalized Landscape Shape Index

(NLSI) patch metric, a measure of patch shape complexity (Kedron et al., 2018). The gradient surface model views landscapes at the scale at which data are delivered (i.e. on a grid), allowing for consideration of gradients across an area, or spatially complex landscapes. In this model, GSMs represent an area's heterogeneity within the larger landscape. These metrics are calculated from continuous values rather than discrete patches of categorical values.

Gradient surface metrics can represent more complex aspects of the landscape surface, allowing for novel linkages between ecosystem pattern and process (Kedron et al., 2018; McGarigal et al., 2009). For example, St-Louis et al. (2006) used image texture to characterize habitat structure in New Mexico, explaining 76% of the variability in bird species richness, in part by eliminating errors associated with habitat boundary delineation. GSMs, combined with climate velocities, have also been evaluated for their use in delineating priority conservation areas for climate change (Carroll et al., 2017). However, despite numerous studies focused on developing and applying these metrics, they remain more difficult to apply than patch metrics, primarily due to the challenges associated with their calculation and interpretation (Costanza et al., 2019; Kedron et al., 2018).

With regard to GSM calculation, while several software programs exist for calculating these metrics (Table 1), these programs (e.g. Scanning Probe Image Processor software (SPIP[TM]; Image Metrology, 2019)) are proprietary and/or expensive. As a result, they remain out of reach for many researchers. Additionally, while formulas for GSMs are published, code is not, making it difficult to determine the exact methods behind the metric calculations and hindering data provenance. We note that FRAGSTATS (McGarigal, 1995) is expected to implement a subset of metrics in an upcoming open-source release where calculation will be possible through a graphical user interface (Costanza et al., 2019; Kedron et al., 2018). However, the ability to implement a wide range of surface metrics via the command line where scripts can document their calculation is still lacking. Bringing GSMs into an open access platform that implements more reproducible calculations, such as the widely used R statistical software (R Core Team, 2020), is an essential next step (Table 1). Many of the

TABLE 1 Characteristics of FRAGSTATS, Image Metrology SPIP[TM] and R packages LANDSCAPEMETRICS and GEODIV. Modified from Table 1 of the study by Hesselbarth et al. (2019). Note that patch metrics apply to categorical data, whereas gradient metrics apply to continuous data

Characteristics	FRAGSTATS	Image Metrology SPIP[TM]	LANDSCAPEMETRICS	GEODIV
Metrics for patch or gradient model	Patch (gradient in progress)	Gradient	Patch	Gradient
Open source	No	No	Yes	Yes
Easy integration into scripted workflows	No	No	Yes	Yes
Utility functions	Sampling	Various	Various	Various
Local application of functions over moving windows	NA[*]	No	NA	Yes
Integrated parallel processing	No	No	No	Yes
Compatible across operating systems	No	No	Yes	Yes

*NA = not applicable in this context.

FRAGSTATS patch-based calculations were recently adapted to an R package (LANDSCAPEMETRICS; Hesselbarth et al., 2019), demonstrating the importance of adopting open-source, command line approaches. However, LANDSCAPEMETRICS was developed independent of FRAGSTATS, and it is unclear whether it will incorporate GSMs in the future.

We introduce a new R package, GEODIV, which is available on *R*'s CRAN server and calculates GSMs from gridded data. This package provides functions for calculating both single-value, global, metrics over images, as well as applying metrics locally using moving windows. We introduce functions that calculate GSMs and also provide a tutorial to demonstrate patterns of, and relationships among, metrics in Oregon, USA in a supplemental vignette. This new R package allows researchers to take full advantage of the benefits of more complex heterogeneity metrics. Its ability to work with both rasters and matrices, compatibility across operating systems and capacity to run calculations in parallel to increase computational efficiency enables numerous applications. We also quantify function runtimes and provide suggestions for trade-offs to consider when computational resources are limited. By providing these metrics in an open-source and transparent way, written in a commonly used programming language designed to work both locally and in parallel computing environments, GEODIV includes many critical improvements over available software and will be an important tool for openly reproducible ecological analysis of continuous surfaces from local to global spatial extents.

2 | FUNCTIONS

GEODIV functions include: (a) metric functions, (b) wrapper functions for applying metrics via moving window analysis and (c) utility functions necessary for metric calculations. Metric functions may either be calculated globally, resulting in a single value for an entire raster, or locally, creating a texture image by applying the function over moving windows.

2.1 | Surface metrics

GEODIV includes functions to calculate all metrics (Table 2) discussed in the studies by Kedron et al. (2018) and McGarigal et al. (2009), wherein GSMs were derived using the SPIP^{TM} software. These metrics cluster into four categories based on behavioural similarity: surface roughness, surface value distribution shape, and angular and radial surface texture (McGarigal et al., 2009; Table 2). The variables represent surface heterogeneity, and correlate well with several patch metrics (McGarigal et al., 2009). Metrics representing the surface value distribution are aspatial and represent how the surface value distribution within a defined area differs from a Gaussian distribution. Angular texture metrics describe the direction and magnitude of surface value autocorrelation. Radial surface metrics describe the level of repetition in values radiating out from any location on the surface. Both angular and radial texture metrics are spatial (McGarigal et al., 2009).

Functions are provided to calculate GSMs across different spatial extents and with different computational resources. Individual metric functions calculate metrics globally to generate a single value for the entire raster and provide information on overall landscape heterogeneity. Alternatively, the 'texture_image' and 'focal_metrics' functions calculate metrics locally at a specified spatial grain to quantify spatial heterogeneity across a raster. The 'texture_image' function is faster than the 'focal_metrics' function,

TABLE 2 Descriptions for a subset of gradient surface metric (GSM) functions. Most of the equations for the metrics are from the SPIP^{TM} user guide (Image Metrology, 2019). Functions take rasters and matrices as inputs. For a complete list of GSM functions, benchmarking results and corresponding equations, see Table S1, and Figures S1 and S2. Metric categories are from the study by McGarigal et al. (2009)

Metric	Function name	Description	Category
Average roughness	*Sa*	Absolute deviation of values from the mean value	Roughness
Root mean square roughness	*Sq*	Standard deviation of surface values relative to the mean value	Roughness
Ten-point height	*S10z*	Average height above the mean surface for the five highest local maxima plus the average height below the mean surface for the five lowest local minima	Roughness
Root mean square slope	*Sdq*	Root mean square slope using the two-point method	Roughness
Area root mean square slope	*Sdq6*	Root mean square slope using the seven-point method	Roughness
Surface area ratio	*Sdr*	Ratio of a flat surface to the actual surface	Roughness
Surface bearing index	*Sbi*	Ratio of root mean square roughness (*Sq*) to height at 5% of the bearing area curve	Distribution
Fractal dimension	*Sfd*	3D fractal dimension, calculated using the triangular prism surface area method.	Radial
Dominant texture direction	*Std*	Angle of dominating texture as found from the Fourier spectrum image	Angular
Texture direction index	*Stdi*	Relative dominance of *Std*	Angular

but uses more memory because data are loaded onto multiple cores for processing (Table S3). As a result, 'texture_image' is better for users with access to high-memory machines, or users who require circular windows, which are more complex to calculate. The 'focal_metrics' function is better for users with computational limitations, or for calculations over smaller images with square windows, and is based on the LANDSCAPEMETRICS 'window_lsm' function (Hesselbarth et al., 2019).

Several utility functions that manipulate rasters and matrices to calculate GSMs are also included for transparency and for their general utility. These utility functions (Table S2) include methods for directionally shifting matrix values, fitting and removing best-fit surfaces; calculating surface area; and estimating and plotting summary functions of raster values.

3 | A SIMPLE WORKFLOW FOR GENERATING GSMs FROM RASTERS

Here, we demonstrate how to apply GEODIV functions, using NAIP-derived Normalized Difference Vegetation Index (NDVI) at 15-m resolution covering the 2017 Jolly Mountain fire in Washington state (Figure 1). As described above, there are two application methods: (a) global:functions applied to get a single value representing overall raster heterogeneity, or (b) local: functions applied within moving windows over the raster.

An optional pre-processing step is to remove any overall trend in the raster using the 'remove_plane' function (Box 1). Removing the trend allows local heterogeneity to stand out; otherwise, this heterogeneity might be masked by larger spatial trends in the data.
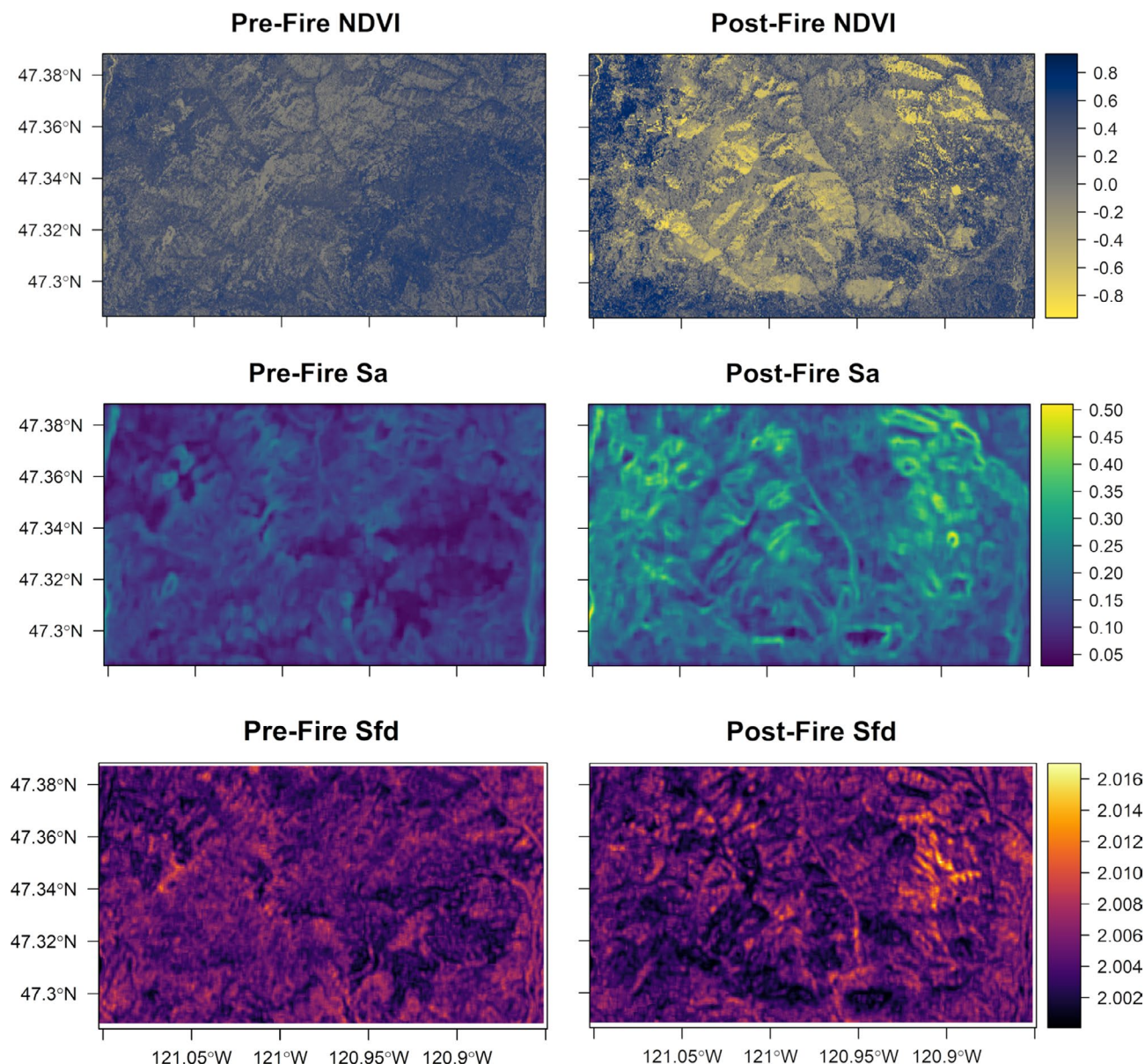

FIGURE 1 Pre- and post-fire Normalized Difference Vegetation Index (NDVI) for the 2017 Jolly Mountain fire in Washington (top panel), and texture images of average roughness (*Sa*; middle panel) and fractal dimension (*Sfd*; bottom panel) created from NDVI. Texture images were created using 30 × 30 pixel (450 m × 450 m) square windows

---

BOX 1 Code for importing NDVI and removing the best-fit surface. The 'remove_plane' function takes input gridded data and returns a grid as well as the polynomial order (between 0 and 3) that minimizes mean absolute error (MAE) between the best-fit surface and original image. However, removing the surface may not be suitable for all applications and users should consider this before applying the function.

Text following '#' are comments in R code and function outputs are italicized.

```
# Import post-fire NDVI
naip2017 <- raster('naip_ndvi2017.tif')
# Remove the best-fit surface
newdata2017 <- remove_plane(naip2017)
"Order" "of polynomial that mini-
mizes errors: 1"
# Calculate average roughness
sa(newdata2017)
0.2966012
# Calculate surface bearing index
sbi(newdata2017)
0.6522381
# Calculate dominant texture direction
std(newdata2017, option=1)
0
# Calculate fractal dimension
sfd(newdata2017)
2.000481
# Create a texture image of post-fire av-
erage roughness calculated with a square
moving window with 30-pixel edges.
texture_image(newdata2017, window_
type='square', size=30, metric='sa')
```

Box 1 demonstrates the global calculation of four metrics (*Sa*, *Sbi*, *Std* and *Sfd*), representing each of the categories described in the study by McGarigal et al. (2009), and the windowed, local, calculation of *Sa* using 'texture_image'. Texture image creation can be time-intensive, so the 'texture_image' function has a logical argument 'parallel', which allows users to perform the calculations in parallel across a specified number of cores (R Core Team, 2020). See Tables S1 and S3 for the computational requirements for all GEODIV functions.

The texture images for average roughness (*Sa*) and fractal dimension (*Sfd*; Figure 1) highlight the utility of GSMs for ecological applications. Disturbed areas are important for increasing variation in otherwise homogeneous landscapes, and disturbance-induced changes in landscape heterogeneity can be assessed to determine impacts on ecosystem services (Turner et al., 2013). In this example, the boundaries of higher severity areas post-fire are clearly delineated with average roughness. Average roughness is the standard deviation of values (Table S1), and highlights these regions with values above ~0.35. Fractal dimension measures the complexity of a self-similar pattern, and here highlights areas with finer scale heterogeneity both pre- and post-fire. *Sa* and *Sfd* provide complementary information on how fire impacted the landscape, showing that multiple metrics may be useful to researchers.

4 | AN ADVANCED VIGNETTE

By assessing heterogeneity using a variety of metrics, researchers can gain a more complete picture of heterogeneity than they would with a single metric (Dahlin, 2016). To more fully demonstrate the utility of GEODIV for this common application, the vignette contains an advanced tutorial that applies all surface metric functions to images across Oregon, USA and examines the patterns of, and relationships among, metrics. The vignette calculates metrics for both elevation data from the Shuttle Radar Topography Mission (SRTM) and a measure of canopy greenness, the Enhanced Vegetation Index (EVI). The vignette shows how to visualize metrics over Oregon to capture different aspects of landscape heterogeneity.

The vignette also examines the correlations among metrics along a transect crossing the state and determines how the metrics cluster using two methods—hierarchical clustering and principal component analysis (PCA). The vignette, associated data and intermediate outputs generated by the vignette are available on figshare (https://doi.org/10.6084/m9.figshare.12834896.v5) and GitHub (https://github.com/bioXgeo/geodiv).

5 | CONCLUSION

Here, we introduced GEODIV, an R package for calculating gradient surface metrics. We provided a brief overview of the package, as well as a simple example of its use. A more detailed example is available in the vignette. The range and simplicity of functions included in GEODIV will allow for a wider application of GSMs in landscape ecology and beyond. As large volumes of imagery become more available and computational limits are reduced, tools like GEODIV will allow ecologists to analyse landscapes in new, open and reproducible ways.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHORS' CONTRIBUTIONS

A.C.S., P.L.Z., K.M.D., S.R., J.K.C. and A.M.W. conceived ideas and methodology; A.C.S. drafted the R package; S.R. drafted the vignette; all authors tested GEODIV; A.C.S. led the manuscript writing. All authors contributed critically to drafts and gave final approval for publication.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13677.

DATA AVAILABILITY STATEMENT

The development version of the GEODIV R package (https://cran.r-project.org/web/packages/geodiv/index.html) is available on GitHub (Smith et al., 2021; http://doi.org/10.5281/zenodo.5093894; https://github.com/bioXgeo/geodiv). All data used in this manuscript are on figshare (https://doi.org/10.6084/m9.figshare.12834896.v5). This research is part of the bioXgeo project (https://github.com/bioXgeo).

ORCID

Annie C. Smith https://orcid.org/0000-0003-0578-9402
Kyla M. Dahlin https://orcid.org/0000-0002-6016-2605
Sydne Record https://orcid.org/0000-0001-7293-2155
Jennifer K. Costanza https://orcid.org/0000-0002-3747-538X
Adam M. Wilson https://orcid.org/0000-0003-3362-7806
Phoebe L. Zarnetske https://orcid.org/0000-0001-6257-6951

REFERENCES

Carroll, C., Roberts, D. R., Michalak, J. L., Lawler, J. J., Nielsen, S. E., Stralberg, D., Hamann, A., Mcrae, B. H., & Wang, T. (2017). Scale-dependent complementarity of climatic velocity and environmental diversity for identifying priority areas for conservation under climate change. *Global Change Biology*, *23*(11), 4508–4520. https://doi.org/10.1111/gcb.13679

Costanza, J. K., Riitters, K., Vogt, P., & Wickham, J. (2019). Describing and analyzing landscape patterns: Where are we now, and where are we going? *Landscape Ecology*, *34*(9), 2049–2055. https://doi.org/10.1007/s10980-019-00889-6

Cushman, S. A., Gutzweiler, K., Evans, J. S., & McGarigal, K. (2010). The gradient paradigm: A conceptual and analytical framework for landscape ecology. In S. A. Cushman, & F. Huettmann (Eds.), *Spatial complexity, informatics, and wildlife conservation* (pp. 83–108). Springer. https://doi.org/10.1007/978-4-431-87771-4_5

Dahlin, K. M. (2016). Spectral diversity area relationships for assessing biodiversity in a wildland–agriculture matrix. *Ecological Applications*, *26*(8), 2758–2768. https://doi.org/10.1002/eap.1390

Helzer, C. J., & Jelinski, D. E. (1999). The relative importance of patch area and perimeter–area ratio to grassland breeding birds. *Ecological Applications*, *9*(4), 1448–1458. https://doi.org/10.2307/2641409

Hesselbarth, M. H., Sciaini, M., With, K. A., Wiegand, K., & Nowosad, J. (2019). landscapemetrics: An open-source R tool to calculate landscape metrics. *Ecography*, *42*(10), 1648–1657. https://doi.org/10.1111/ecog.04617

Image Metrology. (2019). *Scanning probe image processor* [Computer software]. Image metrology APS. Retrieved from http://www.imagemet.com/

Jaeger, J. A. (2000). Landscape division, splitting index, and effective mesh size: New measures of landscape fragmentation. *Landscape Ecology*, *15*(2), 115–130. https://doi.org/10.1023/A:1008129329289

Kedron, P. J., Frazier, A. E., Ovando-Montejo, G. A., & Wang, J. (2018). Surface metrics for landscape ecology: A comparison of landscape models across ecoregions and scales. *Landscape Ecology*, *33*(9), 1489–1504. https://doi.org/10.1007/s10980-018-0685-1

Kumar, S., Stohlgren, T. J., & Chong, G. W. (2006). Spatial heterogeneity influences native and nonnative plant species richness. *Ecology*, *87*(12), 3186–3199. https://doi.org/10.1890/0012-9658%282006%2987%5B3186%3ASHINAN%5D2.0.CO%3B2

Lausch, A., Blaschke, T., Haase, D., Herzog, F., Syrbe, R. U., Tischendorf, L., & Walz, U. (2015). Understanding and quantifying landscape structure – A review on relevant process characteristics, data models and landscape metrics. *Ecological Modelling*, *295*, 31–41. https://doi.org/10.1016/j.ecolmodel.2014.08.018

McGarigal, K. (1995). *FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure* (Vol. 351). US Department of Agriculture, Forest Service, Pacific Northwest Research Station. https://doi.org/10.2737/PNW-GTR-351

McGarigal, K., & Cushman, S. (2005). The gradient concept of landscape structure [Chapter 12]. In J. A. Wiens & M. R. Moss (Eds.), *Issues and perspectives in landscape ecology* (pp. 112–119). Cambridge University Press. https://doi.org/10.1017/CBO9780511614415.013

McGarigal, K., Tagil, S., & Cushman, S. A. (2009). Surface metrics: An alternative to patch metrics for the quantification of landscape structure. *Landscape Ecology*, *24*(3), 433–450. https://doi.org/10.1007/s10980-009-9327-y

R Core Team. (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Smith, A. C., Wilson, A. M., Zarnetske, P., & Record, S. (2021). bioXgeo/geodiv: geodiv (Version v1.0.3). *Zenodo*, https://doi.org/10.5281/zenodo.5093895

Stevens, J. T., Collins, B. M., Miller, J. D., North, M. P., & Stephens, S. L. (2017). Changing spatial patterns of stand-replacing fire in California conifer forests. *Forest Ecology and Management*, *406*, 28–36. https://doi.org/10.1016/j.foreco.2017.08.051

St-Louis, V., Pidgeon, A. M., Radeloff, V. C., Hawbaker, T. J., & Clayton, M. K. (2006). High-resolution image texture as a predictor of bird species richness. *Remote Sensing of Environment*, *105*(4), 299–312. https://doi.org/10.1016/j.rse.2006.07.003

Turner, M. G. (1989). Landscape ecology: The effect of pattern on process. *Annual Review of Ecology and Systematics*, *20*(1), 171–197. https://doi.org/10.1146/annurev.es.20.110189.001131

Turner, M. G. (2005). Landscape ecology: What is the state of the science? *Annual Review of Ecology Evolution and Systematics*, *36*, 319–344. https://doi.org/10.1146/annurev.ecolsys.36.102003.152614

Turner, M. G., Donato, D. C., & Romme, W. H. (2013). Consequences of spatial heterogeneity for ecosystem services in changing forest landscapes: Priorities for future research. *Landscape Ecology*, *28*(6), 1081–1097. https://doi.org/10.1007/s10980-012-9741-4

Von Humboldt, A., & Bonpland, A. (2010). *Essay on the geography of plants*. University of Chicago Press. https://doi.org/10.7208/9780226360683

Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., & Mander, Ü. (2009). Landscape metrics and indices: An overview of their use in landscape research. *Living Reviews in Landscape Research*, *3*(1), 1–28. https://doi.org/10.12942/lrlr-2009-1

Walz, U. (2011). Landscape structure, landscape metrics and biodiversity. *Living Reviews in Landscape Research*, *5*(3), 1–35. https://doi.org/10.12942/lrlr-2011-3

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.