

# BoundaryStats: An R Package to Calculate Boundary Overlap Statistics

by Amy Luo

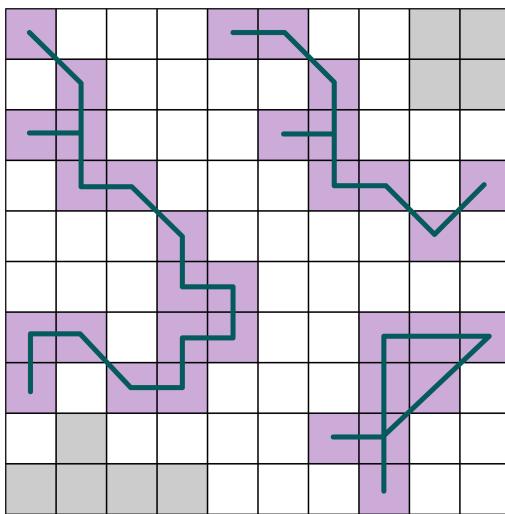
**Abstract** Ecologists and epidemiologists frequently rely on spatially distributed data. Studies in these fields may concern geographic boundaries, as environmental variation can determine the spatial distribution of organismal traits or diseases. In such cases, environmental boundaries produce coincident geographic boundaries in, for example, disease prevalence. Boundary analysis can be used to investigate the co-occurrence of organismal trait or disease boundaries and underlying environmental boundaries. Within boundary analysis, boundary overlap statistics test for the presence of significant geographic boundaries and spatial associations between the boundaries of two variables. There is one pre-existing implementation of boundary overlap statistics, though it is only on Windows and ESRI ArcView, limiting the availability of boundary overlap statistics to researchers. I have created BoundaryStats—an R package available on CRAN—that implements boundary overlap statistics. BoundaryStats is the first open-source, cross-platform implementation of these statistical methods, making the statistics more widely accessible to researchers.

## 1 Introduction

Geographic boundaries are an intrinsic feature of spatial ecology and epidemiology, as the relationships between an underlying environmental variable and organismal traits or disease prevalence often produce coincident geographic boundaries. Boundaries are areas in which spatially distributed variables (e.g., bird plumage coloration, disease prevalence, annual rainfall) rapidly change over a narrow geographic space. They can also represent edges or discontinuities (e.g., neighborhood edges, ecotype boundaries). Boundary zones themselves may be of interest; for example, the temporal boundary dynamics between ecotypes can provide insight into the factors that produce mosaic landscapes (Bowman et al., 2023).

Boundary analysis involves the analysis of spatial boundaries to answer questions about values within a bounded area, patterns of change across a landscape, and associations between the spatial patterns of multiple variables (Jacquez, 2010). Within boundary analysis, boundary overlap statistics can be used to test the association between the boundaries of two spatially distributed variables (Jacquez et al., 2000). These statistics fall within two categories: boundary statistics (i.e., tests for the presence of cohesive boundaries) and boundary overlap statistics (i.e., tests for spatial association between boundaries). BoundaryStats runs two boundary statistics and three boundary overlap statistics, as initially described in Jacquez (1995) and Fortin et al. (1996). The boundary statistics are (1) the length of the longest boundary and (2) the number of cohesive boundaries on the landscape (Fortin et al., 1996). The boundary overlap statistics are (1) the amount of direct overlap between boundaries in variables A and B, (2) the mean minimum distance between boundaries in A and B, and (3) the mean minimum distance from boundaries in A to boundaries in B (Fortin et al., 1996; Jacquez, 1995).

While other spatial statistics account for complications like spatial autocorrelation and environmental heterogeneity (Wagner and Fortin, 2005), boundary overlap statistics can uniquely leverage geographic discontinuities to answer spatial questions. By identifying significant cohesive boundaries, researchers can delineate relevant geographic sampling units (e.g., populations as conservation units for a species or human communities with increased disease risk) (Jacquez, 2010). Associations between the spatial boundaries of two variables can be useful in assessing the extent to which an underlying landscape variable drives the spatial distribution of a dependent variable. For example, ecologists are often interested in whether landscape-level ecological boundaries limit gene flow, thereby producing population structure; if the putative ecological boundary is limiting gene flow, one



**Figure 1:** Example boundary subgraphs. Gray cells are missing values, white cells are non-boundary cells, and purple cells are boundary elements. Subgraphs are each represented using a line connecting all the boundary element cells that comprise them.

would expect concordant geographic boundaries in the ecological variable and population structure (Tarroso et al., 2014; Wagner and Fortin, 2013). The presence of boundaries can similarly limit the distribution of taxonomically similar species (Polakowska et al., 2012). In an epidemiological context, this may look like neighborhood effects on public health outcomes, including COVID-19 infection risk (Hong et al., 2021; van Ham et al., 2012) or spatial relationships between high pollutant density and increased disease risk (Adimalla et al., 2020; Waller et al., 1992).

Currently, there is at least one tool that has implemented boundary overlap statistics: GEM, which was released as an extension of ESRI ArcView and a standalone Windows package. GEM is not available as a cross-platform, free, and open-source software, thereby limiting its accessibility to researchers. **BoundaryStats** implements boundary and boundary overlap statistics in R. It is available to download on CRAN, making the tools more accessible for researchers, especially in epidemiology and spatial ecology.

## 2 Boundary definitions

In this framework, we classify raster cells into a pseudobinary: boundary elements (1), non-boundary cells (0), or missing data (NA). For categorical variables, the algorithm for identifying boundary elements is simple: if any of a cell's neighbors—based on the queen criterion (i.e., eight neighboring cells, including diagonal cells)—belongs to a different category, the cell is classified as a boundary element. For quantitative variables, boundaries exist where two different but internally homogeneous areas neighbor one another (e.g., an area with very low values meets an area with very high values), but these boundaries are generally fuzzy; they represent steep transitions that can still occur over the width of multiple cells. Therefore, we define boundary elements as the cells with the highest boundary intensity values, with the threshold set by the user. Boundary intensity values can be calculated through several different methods, including the magnitude of change across cells or the probability that cells belong to each neighboring spatial group (see details below). Boundaries are defined here as subgraphs of boundary elements, or contiguous cells that are all marked as boundary elements (Figure 1).

Boundary intensities for variables with landscape-level patterns can be calculated in a number of ways, including: lattice- and triangulation-wombling (Fortin et al., 1996; Jacquez, 1995; St-Louis et al., 2004; Strydom and Poisot, 2023), fuzzy set modeling (Jacquez, 1995), Monmonier's algorithm (Manni et al., 2004), spatial Bayesian clustering (Caye et al., 2016;

Safner et al., 2011), agglomeration of inner lines (Wei and Larsen, 2019), and removal of outer lines (Wei and Larsen, 2019). For quantitative variables, BoundaryStats will accept raster objects with the spatial variable directly or boundary intensity values calculated from these or other methods. If given boundary intensity values, boundary elements will be classified directly using the top percent of values. The default proportion of values is 0.2, though this threshold can be changed by the user. When given the variable directly, BoundaryStats will use the Sobel-Feldman operator to calculate the boundary intensity. In accepting either the variables or boundary intensities, there is flexibility for users to define boundaries using relevant metrics for their data.

The Sobel-Feldman operator is commonly used for edge detection in computer vision applications. It approximates the magnitude of the partial derivative (i.e., rate of change) across each cell using the following kernels:

$$G_x = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} * A \quad (1)$$

$$G_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} * A \quad (2)$$

where  $A$  is the input raster cell and its queen neighbors, and  $G_x$  and  $G_y$  are the rates of variable change in the horizontal or vertical directions, respectively. The boundary intensity value is the overall rate of change:

$$G = \sqrt{G_x^2 + G_y^2} \quad (3)$$

### 3 Statistics

BoundaryStats runs two boundary statistics and three boundary overlap statistics, as initially described in Jacquez (1995) and Fortin et al. (1996). Below, I describe these five statistics.

#### 3.1 Number of subgraphs

The first boundary statistic is the number of subgraphs, which describes the number of boundaries on the landscape for a variable. In a raster of boundary elements, it is the number of unique subgraphs, or sets of contiguous boundary element cells (three subgraphs each in Figure 2, Panels A and B).

#### 3.2 Longest subgraph

The other boundary statistic included here is the length of the longest subgraph, or boundary. The function calculates the longest length across each subgraph, then converts the length to distance based on the cell resolution and the projection of the raster. The length of the longest subgraph is then retained (Figure 2, Panel C).

#### 3.3 Direct overlap

The direct overlap statistic is a count of the number of overlapping boundary elements of two variables, when the two raster objects are overlaid (Figure 2, Panel D).

#### 3.4 Mean minimum distance between boundaries

This statistic describes the spatial proximity between boundaries of variables  $x$  and  $y$ , as defined by the mean distance to the nearest boundary element of the other variable. Spatial

relationships between boundaries may not result in direct overlap, so this statistic accounts for potential correlations in non-overlapping boundaries. For each boundary element in  $x$ , the function calculates the distance to the nearest boundary element in  $y$ , then repeats the inverse for each boundary element in  $y$  (Figure 2, Panels E and F). It then takes the mean of these minimum distances across all boundary elements in both raster objects:

$$O_{xy} = \frac{\sum_{i=1}^{N_x} \min(d_i) + \sum_{j=1}^{N_y} \min(d_j)}{N_x + N_y} \quad (4)$$

where  $i$  and  $j$  are boundary elements for variables  $x$  and  $y$ , respectively;  $\min(d_i)$  is the minimum distance between boundary element  $i$  to a boundary element for  $y$ ;  $\min(d_j)$  is the minimum distance between boundary element  $j$  to a boundary element in  $x$ ; and  $N_x$  and  $N_y$  are the number of boundary elements for  $x$  and  $y$ , respectively.

### 3.5 Mean minimum distance from boundary $x$ to boundary $y$

This statistic describes the mean distance from boundary elements in  $x$  to the nearest boundary element in  $y$ . It is an indicator for whether the boundaries in  $x$  depend on  $y$ . The reciprocal nature of the previous statistic implies some reciprocity of effect, as opposed to the unidirectionality implicit here. For each boundary element in the raster for  $x$ , the function calculates the distance to the nearest boundary element of  $y$ , then takes the mean across all boundary elements in  $x$  (Figure 2, Panel E):

$$O_x = \frac{\sum_{i=1}^{N_x} \min(d_i)}{N_x} \quad (5)$$

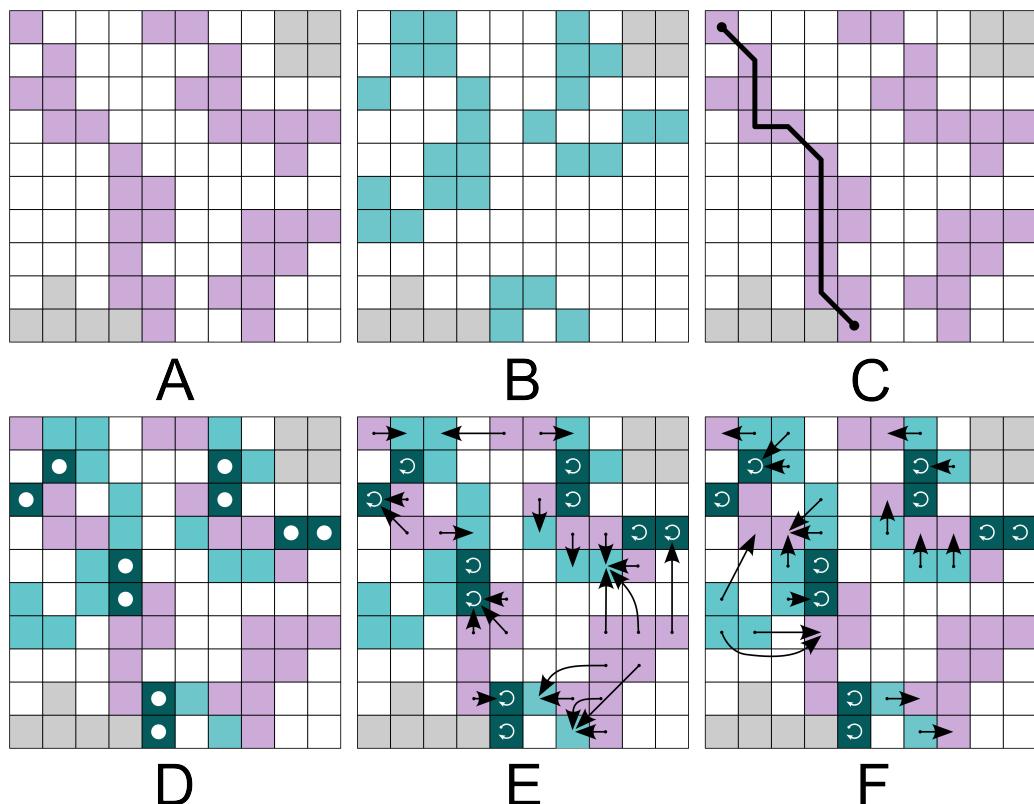
## 4 Neutral models

In addition to calculating each statistic, BoundaryStats uses iterations of a neutral landscape model to determine whether the boundaries in the input landscape differ from a random landscape. Users select a neutral landscape model and number of iterations of that model to produce a null distribution of each statistic, based on the selected model and the structure of the input landscape. BoundaryStats implements three neutral landscape models: stochastic landscapes, Gaussian random fields, and modified random clusters. All three neutral landscape models draw some parameters from the original raster and simulate landscapes with similar parameters. Cells with missing values (i.e., NA values) will be ignored in all models.

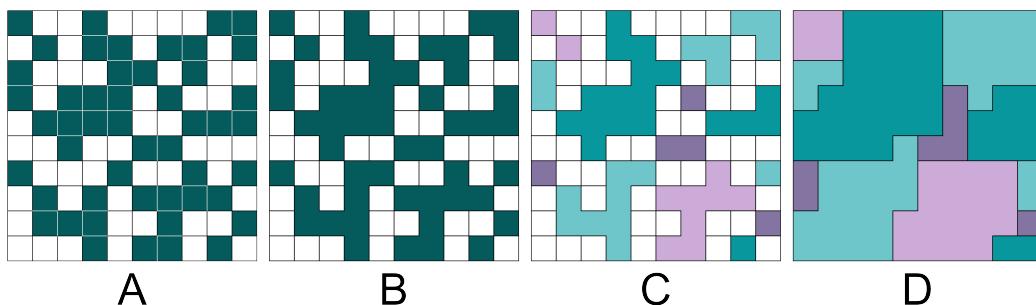
The simplest neutral landscape model is complete stochasticity. While this model is not realistic—a complete lack of spatial autocorrelation is unlikely and may inflate the statistical significance of the observed data—we include it here as a complete null for users who are interested in a lack of spatial autocorrelation. This method takes all the cell values from the input raster and assigns each value to a random cell. Each cell in the simulated raster is assigned a value from the original dataset, with no replacement of values. The simulated raster has the exact same values as the original raster, but values are randomly placed with no spatial autocorrelation.

The next neutral landscape model simulates a Gaussian random field with the same degree of spatial autocorrelation as the input raster. It is suited for continuous or discrete quantitative variables. This method calculates the range of autocorrelation in the original raster by fitting a variogram using functions from [gstat](#) (Pebesma, 2004). The function then simulates a Gaussian random field with the same range of spatial autocorrelation, extent, and resolution as the input raster using methods from the [fields](#) package.

The modified random cluster model is an implementation of the method described by Saura and Martínez-Millan for simulating neutral landscapes for categorical variables ([Saura](#)



**Figure 2:** Example boundaries and statistics. (A) and (B) are boundary elements for hypothetical variables A and B. White cells are non-boundaries, gray are missing values, and purple or teal are boundary elements. (C) Length of the longest subgraph. (D) Produced by overlaying cells in A and B. Dark blue cells, highlighted by white dots, are where the boundary elements overlap. (E) For every boundary element for variable A, the nearest boundary element for variable B. Circular arrows indicate distance to self. (F) For every boundary element for B, the nearest boundary element for A.



**Figure 3:** Modified random cluster procedure, adapted from Saura and Martínez-Millan 2000. (A) Percolated raster with  $p = 0.5$ . (B) Marked cells merged into clusters. (C) Clusters assigned a category. (D) Unmarked cells filled based on neighbors.

and Martínez-Millan, 2000). The first step is a percolated raster (Figure 3, Panel A). Each cell is assigned a value  $0 \leq x \leq 1$  from a uniform distribution, and cells with values above a threshold probability  $p$  are marked.  $p$  is defined by the user, and higher values of  $p$  result in larger cluster sizes in the final simulated raster. Next, contiguous sets of marked cells are grouped into clusters, using the rook criterion (i.e., neighbors are the four edge-touching neighbors) (Figure 3, Panel B). Clusters are then assigned a category (Figure 3, Panel C). Categories from the input raster are chosen one at a time, and random clusters are assigned to that category. When the proportion of that category in the simulated raster reaches the proportion in the original raster, clusters are then assigned to the next category, until all the clusters are assigned. In the last step, the unmarked cells are categorized based on the most frequent category among their neighbors using the queen criterion (Figure 3, Panel D). If there is a tie between two categories, one of the tied categories is picked at random. If all neighbors are unassigned, a random category is picked; probabilities for each category are based on their proportions in the input raster.

## 5 Implementation and example

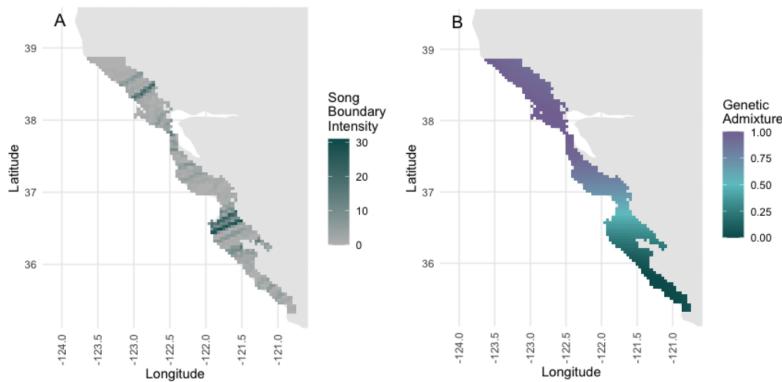
Data in this example are from Luo et al. (2024), in which the authors hypothesized that song divergence is facilitating genetic divergence in white-crowned sparrows through speciation by sexual selection. The data below are song boundaries and genetic admixture interpolations from the study. White-crowned sparrows sing different songs that vary across the landscape, and song boundaries are the spatial transitions between two song groups. The boundary intensities between song groups were estimated using GeoOrigins (Hulme-Beaman et al., 2020), based on the acoustic dissimilarity and spatial relationship between recorded songs. Genetic admixture is the estimated proportion of an individual's genetic material from different populations. In this case, there are two populations (north and south), so the admixture coefficients are the estimated proportions of genetic material from the northern population. The admixture coefficients of individual birds were estimated in fastSTRUCTURE (Raj et al., 2014), and the values were interpolated using local kriging in `gstat`.

### 5.1 Read in data

Read in raster data to `terra` (Hijmans, 2023) `SpatRaster` objects (Figure 4). The two objects need the same projection, extent, and resolution.

```
library(terra)
library(magrittr)

songs <- rast('data/2010_2022_song_boundaries.asc')
genetic <- rast('data/genetic_interpolation.asc') %>%
```



**Figure 4:** Maps of (A) song boundary intensity and (B) genetic admixture between two populations.

```
resample(., songs)
songs <- crop(songs, genetic) %>%
  mask(., genetic)
```

## 5.2 Calculate spatial boundaries for variables

The first step of the analysis is to define which cells are boundary elements (i.e., part of a boundary) using the `define_boundary` function. By default, the function takes quantitative variables (`cat = FALSE`). For quantitative variables, boundary intensity can be calculated however the user chooses; if the input raster already contains boundary intensities, the argument `calculate_intensity` should be set to `FALSE` (default). Users can also set `calculate_intensity` to `TRUE` to use the Sobel-Feldman operator to calculate the boundary intensities.

The song raster already contains boundary intensity values, from which boundary elements can be directly determined. But the values in the genetic raster are the trait data, so boundary intensity needs to be calculated from the genetic admixture coefficient values (`calculate_intensity = TRUE`).

```
library(BoundaryStats)

song_boundaries <- define_boundary(songs)
genetic_boundaries <- define_boundary(genetic, calculate_intensity = TRUE)
```

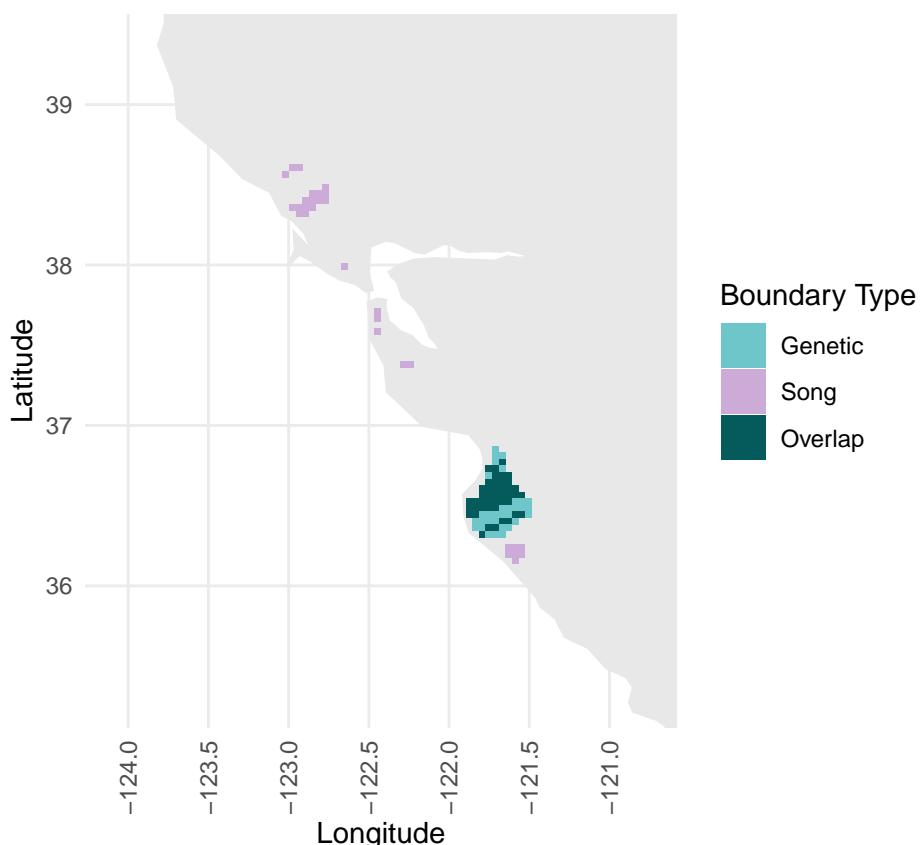
## 5.3 Plot boundary overlap

This optional step is to visualize where the boundaries of the two variables are overlapping using `plot_boundary` (Figure 5). The function is a wrapper function for `ggplot` from `ggplot2` (Wickham, 2016), and the colors and trait names can optionally be customized. If `output_raster` is `TRUE` (default is `FALSE`), then the function will return a `SpatRaster` object with one layer that includes boundary elements for each trait and where the boundary elements overlap.

```
plot_boundary(genetic_boundaries, song_boundaries, trait_names = c("Genetic", "Song"))
```

## 5.4 Create null distributions for statistics

For both boundary statistics, use the function `boundary_null_distrib`. For the three overlap statistics, use the function `overlap_null_distrib`. Both functions simulate random



**Figure 5:** Output of the `plot_boundary` function.

iterations of a raster based on the specified neutral landscape model and input data. Statistics are calculated for each iteration, and custom null probability distributions are calculated based on the iterations. The resulting objects will be used for the statistical tests described in the section below.

Both functions take the `SpatRaster` object(s), a neutral landscape model, and the number of iterations. Further arguments may be required, depending on these arguments. For `overlap_null_distrib`, separate models can be specified for the two variables, and the variable in the first argument is assumed to depend on the variable in the second argument. The argument `rand_both` specifies whether the function should simulate random landscapes for the second `SpatRaster` object; since some hypotheses assume  $x$  depends on a specific underlying distribution of boundaries in  $y$ , users can choose to keep boundaries for  $y$  static for each iteration. For this example, the genetic boundary is hypothesized to depend on song boundaries. Therefore, the `SpatRaster` object containing the genetic admixture interpolation is the first argument, and I keep the song boundaries static (`rand_both = FALSE`).

```
song_boundary_null <- boundary_null_distrib(songs, calculate_intensity = FALSE, cat = FALSE,
                                              n_iterations = 100, threshold = 0.2, model = "gaussian")
genetic_boundary_null <- boundary_null_distrib(genetic, calculate_intensity = TRUE,
                                                cat = FALSE, n_iterations = 100, threshold = 0.2, model = "gaussian")

boundary_overlap_null <- overlap_null_distrib(genetic, songs, rand_both = FALSE,
                                               n_iterations = 100, x_calculate_intensity = TRUE, threshold = 0.2, x_model = "gaussian")
```

## 5.5 Run statistical tests

The two functions for boundary statistics require only the raster with boundary elements and the matching null distribution object, produced by `boundary_null_distrib`.

```

n_boundaries(song_boundaries, song_boundary_null)

#> n_boundary      p-value
#>          13           0

longest_boundary(song_boundaries, song_boundary_null)

#> longest_boundary      p-value
#>        45260.75       0.10

n_boundaries(genetic_boundaries, genetic_boundary_null)

#> n_boundary      p-value
#>          1           0

longest_boundary(genetic_boundaries, genetic_boundary_null)

#> longest_boundary      p-value
#>        61454.64       0.00

```

The functions for boundary overlap statistics also take the boundary element rasters and null distribution as arguments. In this case, it requires two boundary element SpatRaster objects, one for each variable. The order of the variables should match the order used in `overlap_null_distrib`. I am interested in whether the genetic boundary depends on song boundaries (i.e., a unidirectional trend), so I am using the `n_overlap_boundaries` (direct overlap) and `average_min_x_to_y` ( $O_x$ ) tests but not the `average_min_distance` ( $O_{xy}$ ) test. The genetic boundary raster is the first argument, and the song boundary raster is the second argument.

```

n_overlap_boundaries(genetic_boundaries, song_boundaries, boundary_overlap_null)

#> n_overlapping      p-value
#>          44           0

average_min_x_to_y(genetic_boundaries, song_boundaries, boundary_overlap_null)

#> avg_min_x_to_y      p-value
#>        1911.751      0.000

```

## 5.6 Interpretation of example data output

When analyzing the data from Luo et al. (2024), the boundary statistics for the genetic data were significant, suggesting the presence of one cohesive genetic boundary. The boundary statistics for the song data were less clear, as the p-values were at or around 0.05. If the analysis was repeated with more iterations of the neutral landscape model, it may suggest multiple cohesive song boundaries. Results from the boundary overlap statistics show significant direct overlap between the genetic and song boundaries and spatial proximity from the genetic boundary to song boundaries, suggesting a spatial correlation between boundaries of the two variables. While boundary overlap statistics can only demonstrate a correlation between boundaries, the results are generally consistent with the hypothesis that song boundaries are facilitating a coincident genetic boundary.

## 6 Summary

BoundaryStats implements five boundary overlap statistics. Boundary analyses like the boundary overlap statistics implemented here can be used across many contexts that make

use of spatially distributed data. For example, spatial ecologists and epidemiologists can use boundary overlap statistics to assess whether environmental variables are influencing the distribution of organismal traits or disease occurrences. Environmental influences can, in some cases, be detected through the co-occurrence and coincidence of geographic boundaries; environmental boundaries may produce boundaries in the variables of interest. As such, this new open-source, cross-platform implementation will make boundary statistical methods more widely accessible to researchers.

## References

- N. Adimalla, J. Chen, and H. Qian. Spatial characteristics of heavy metal contamination and potential human health risk assessment of urban soils: A case study from an urban region of South India. *Ecotoxicology and Environmental Safety*, 194(110406), 05 2020. doi: 10.1016/j.ecoenv.2020.110406. URL <https://linkinghub.elsevier.com/retrieve/pii/S0147651320302451>. [p2]
- D. M. J. S. Bowman, S. Onde, A. Lucieer, S. Foyster, and L. D. Prior. Forest-sedgeland boundaries are historically stable and resilient to wildfire at Blakes Opening in the Tasmanian Wilderness World Heritage Area, Australia. *Landscape Ecology*, 38:205–222, 2023. doi: 10.1007/s10980-022-01558-x. [p1]
- K. Caye, T. M. Deist, H. Martins, O. Michel, and O. François. TESS3: Fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources*, 16(2): 540–548, 2016. doi: 10.1111/1755-0998.12471. [p2]
- M. Fortin, P. Drapeau, and G. M. Jacquez. Quantification of the spatial co-occurrences of ecological boundaries. *Oikos*, 77(1):51–60, 10 1996. doi: 10.2307/3545584. URL <https://www.jstor.org/stable/3545584?origin=crossref>. [p1, 2, 3]
- R. J. Hijmans. *terra: Spatial data analysis*. 2023. URL <https://CRAN.R-project.org/package=terra>. [p6]
- B. Hong, B. J. Bonczak, A. Gupta, L. E. Thorpe, and C. E. Kontokosta. Exposure density and neighborhood disparities in COVID-19 infection risk. *Proceedings of the National Academy of Sciences*, 118(13):e2021258118, 2021. doi: 10.1073/pnas.2021258118. [p2]
- A. Hulme-Beaman, A. Rudzinski, J. E. J. Cooper, R. F. Lachlan, K. Dobney, and M. G. Thomas. geoorigins: A new method and R package for trait mapping and geographic provenancing of specimens without categorical constraints. *Methods in Ecology and Evolution*, 11(10): 1247–1257, 2020. doi: 10.1111/2041-210X.13444. [p6]
- G. M. Jacquez. The map comparison problem: Tests for the overlap of geographic boundaries. *Statistics in Medicine*, 14(21-22):2343–2361, 1995. doi: 10.1002/sim.4780142107. [p1, 2, 3]
- G. M. Jacquez. Geographic boundary analysis in spatial and spatio-temporal epidemiology: Perspective and prospects. *Spatial and Spatio-temporal Epidemiology*, 1(4):207–218, 2010. doi: 10.1016/j.sste.2010.09.003. [p1]
- G. M. Jacquez, S. Maruca, and M. Fortin. From fields to objects: A review of geographic boundary analysis. *Journal of Geographical Systems*, 2:221–241, 2000. doi: 10.1007/PL00011456. [p1]
- A. R. Luo, S. Lipshutz, J. Phillips, R. T. Brumfield, and E. P. Derryberry. Song and genetic divergence within a subspecies of white-crowned sparrow (*Zonotrichia leucophrys nuttalli*). *PLoS ONE*, 19(5):e0304348, 05 2024. doi: 10.1371/journal.pone.0304348. [p6, 9]
- F. Manni, E. Guérard, and E. Heyer. Geographic patterns of (genetic, morphological, linguistic) variation: How barriers can be detected with Monmonier's algorithm. *Human Biology*, 76(2):173–190, 2004. doi: 10.1353/hub.2004.0034. [p2]

- E. J. Pebesma. Multivariable geostatistics in S: The gstat package. *Computers & Geosciences*, 30(7):683–691, 2004. doi: 10.1016/j.cageo.2004.03.012. [p4]
- A. E. Polakowska, M. Fortin, and A. Couturier. Quantifying the spatial relationship between bird species' distributions and landscape feature boundaries in southern Ontario, Canada. *Landscape Ecology*, 27(10):1481–1493, 12 2012. doi: 10.1007/s10980-012-9804-6. [p2]
- A. Raj, M. Stephens, and J. K. Pritchard. fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, 197(2):573–589, 06 2014. doi: 10.1534/genetics.114.164350. URL <https://academic.oup.com/genetics/article/197/2/573/6074271>. [p6]
- T. Safner, M. P. Miller, B. H. McRae, M. Fortin, and S. Manel. Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. *International Journal of Molecular Sciences*, 12(2):865–889, 01 2011. doi: 10.3390/ijms12020865. URL <http://www.mdpi.com/1422-0067/12/2/865>. [p3]
- S. Saura and J. Martínez-Millan. Landscape patterns simulation with a modified random clusters method. *Landscape Ecology*, 15:661–678, 2000. doi: 10.1023/A:1008107902848. [p4]
- V. St-Louis, M. Fortin, and A. Desrochers. Spatial association between forest heterogeneity and breeding territory boundaries of two forest songbirds. *Landscape Ecology*, 19(6):591–601, 08 2004. doi: 10.1023/B:LAND.0000042849.63040.a9. URL <http://link.springer.com/10.1023/B:LAND.0000042849.63040.a9>. [p2]
- T. Strydom and T. Poisot. Spatialboundaries.jl: Edge detection using spatial wombling. *Ecography*, 2023(5):e06609, 05 2023. doi: 10.1111/ecog.06609. URL <https://onlinelibrary.wiley.com/doi/10.1111/ecog.06609>. [p2]
- P. Tarroso, R. J. Pereira, F. Martínez-Freiría, R. Godinho, and J. Brito. Hybridization at an ecotone: ecological and genetic barriers between three Iberian vipers. *Molecular Ecology*, 23(5):1108–1123, 03 2014. doi: 10.1111/mec.12671. URL <https://onlinelibrary.wiley.com/doi/10.1111/mec.12671>. [p2]
- M. van Ham, D. Manley, N. Bailey, L. Simpson, and D. MacLennan, editors. *Neighbourhood effects research: New perspectives*. Springer Netherlands, 2012. doi: 10.1007/978-94-007-2309-2. URL <http://dx.doi.org/10.1007/978-94-007-2309-2>. [p2]
- H. H. Wagner and M. Fortin. Spatial analysis of landscapes: Concepts and statistics. *Ecology*, 86(8):1975–1987, 2005. doi: 10.1890/04-0914. [p1]
- H. H. Wagner and M. Fortin. A conceptual framework for the spatial analysis of landscape genetic data. *Conservation Genetics*, 14(2):253–261, 2013. doi: 10.1007/s10592-012-0391-5. URL <http://link.springer.com/10.1007/s10592-012-0391-5>. [p2]
- L. A. Waller, B. W. Turnbull, L. C. Clark, and P. Nasca. Chronic disease surveillance and testing of clustering of disease and exposure: Application to leukemia incidence and TCE-contaminated dumpsites in upstate New York. *Environmetrics*, 3(3):281–300, 1992. doi: 10.1002/env.3170030303. [p2]
- X. Wei and C. P. S. Larsen. Methods to detect edge effected reductions in fire frequency in simulated forest landscapes. *ISPRS International Journal of Geo-Information*, 8(6):277, 2019. doi: 10.3390/ijgi8060277. URL <https://www.mdpi.com/2220-9964/8/6/277>. [p3]
- H. Wickham. *ggplot2: Elegant graphics for data analysis*. 2016. URL <https://ggplot2.tidyverse.org>. [p7]

Amy Luo  
University of Tennessee, Knoxville  
Department of Ecology and Evolutionary Biology

*Knoxville, Tennessee*  
ORCID: 0000-0002-9197-4673  
[amy.luo.15@gmail.com](mailto:amy.luo.15@gmail.com)