APPLICATION

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

`ResistanceGA`: An `R` package for the optimization of resistance surfaces using genetic algorithms

William E. Peterman

School of Environment and Natural Resources, The Ohio State University, Columbus, OH, USA

Correspondence
William E. Peterman
Email: Peterman.73@osu.edu

Handling Editor: Simon Jarman

Abstract

1. Understanding how landscape features affect functional connectivity among populations is a cornerstone of spatial ecology and landscape genetic analyses. However, parameterization of resistance surfaces that best describe connectivity is a challenging and often subjective process.

2. `ResistanceGA` is an R package that utilizes a genetic algorithm to optimize resistance surfaces based on pairwise genetic data and effective distances calculated using CIRCUITSCAPE, least cost paths or random-walk commute times. Functions in this package allow for the optimization of categorical and continuous resistance surfaces, and simultaneous optimization of multiple resistance surfaces.

3. `ResistanceGA` provides a coherent framework to optimize resistance surfaces without a priori assumptions, conduct model selection, and make inference about the contribution of each surface to total resistance.

4. `ResistanceGA` fills a void in the landscape genetic toolbox, allowing for unbiased optimization of resistance surfaces and for the simultaneous optimization of multiple resistance surfaces to create novel composite resistance surfaces, but could have broader applicability to other fields of spatial ecological research.

KEYWORDS
commute distance, cost distance, gene flow, genetic algorithm, landscape genetics, least cost path, resistance distance, resistance optimization

1 | INTRODUCTION

First coined in 2003, landscape genetics has experienced rapid growth in both the number of studies and range of analytical methods utilized (Manel, Schwartz, Luikart, & Taberlet, 2003; Storfer, Murphy, Spear, Holderegger, & Waits, 2010). This integrative field draws on landscape ecology, spatial statistics, and population genetics to address a wide range of questions. Landscape genetic studies frequently seek to understand how landscape features affect spatial genetic structure (Manel et al., 2003; Storfer et al., 2007), often with a goal of quantifying the effective distance between sample locations as a function of the landscape matrix (McRae, 2006; Spear, Peterson, Matocq, & Storfer, 2005). In the absence of direct observation of movement or dispersal across the landscape, effective distances are often interpreted as functional connectivity (e.g. Cushman, McKelvey, Hayden, & Schwartz, 2006). However, functional connectivity and effective distance require an appropriately parameterized resistance surface. As defined by Spear, Balkenhol, Fortin, McRae, and Scribner (2010), a resistance surface is a spatial layer that assigns a value to each landscape or environmental feature, with values representing the extent to which that feature impedes or facilitates connectivity for an organism.

Resistance values of surfaces have been determined using a variety of methods, including: habitat suitability models (e.g. Wang, Yang, Bridgman, & Lin, 2008), telemetry (e.g. Driezen, Adriaensen, Rondinini, Doncaster, & Matthysen, 2007) and statistical

models using genetic response data (e.g. Dudaniec et al., 2013). Parameterization of resistance values frequently relies on expert opinion (Zeller, McGarigal, & Whiteley, 2012) and less often on empirical movement studies (e.g. Stevens, Verkenne, Vandewoestijne, Wesselingh, & Baguette, 2006) or spatial predictions of ecological processes (Peterman, Connette, Semlitsch, & Eggert, 2014). All of these are acceptable approaches, but each come with caveats. Of particular concern is that expert opinion often fails to accurately describe the biological or ecological process(es) being modelled (Charney, 2012; Shirk, Wallin, Cushman, Rice, & Warheit, 2010). Even when such processes are known and explicitly modelled, there is no guarantee that these processes will relate meaningfully to the movement of genes across the landscape (Khimoun et al., 2017; Peterman et al., 2014). As such, methods to objectively parameterize resistance surfaces using genetic data are needed.

Unfortunately, assignment of resistance values is often a trial and error process that assesses a limited parameter space. This has led researchers to implement exhaustive search and optimization methods. Wang, Savage, and Shaffer (2009) used an exhaustive search approach

to optimize resistance values of a categorical land cover surface, while Graves, Beier, and Royle (2013) used a search algorithm to maximize Mantel r correlation between inter-individual genetic distance and least cost path distance. Although the optimization procedures of Graves et al. (2013) recovered the maximum Mantel r when it existed, they found that resistance estimates were often imprecise and much smaller than simulated resistance values. They also found that response surfaces were quite flat, making identification of a global optimum difficult. In contrast, Peterman et al. (2014) found well-defined global optima when using Ricker and monomolecular data transformations in combination with optimization algorithms. However, the optimization procedure utilized by Peterman et al. (2014) is limited to continuous resistance surfaces (e.g. temperature, canopy cover) and requires an inefficient search of all possible data transformations.

There are numerous challenges to optimizing resistance surfaces based on pairwise genetic data. Among these challenges, foremost is the high dimensionality that resistance surfaces can have. For instance, if a land use, land cover surface consists of four land cover classes, there are 24 possible ways of ranking these classes even
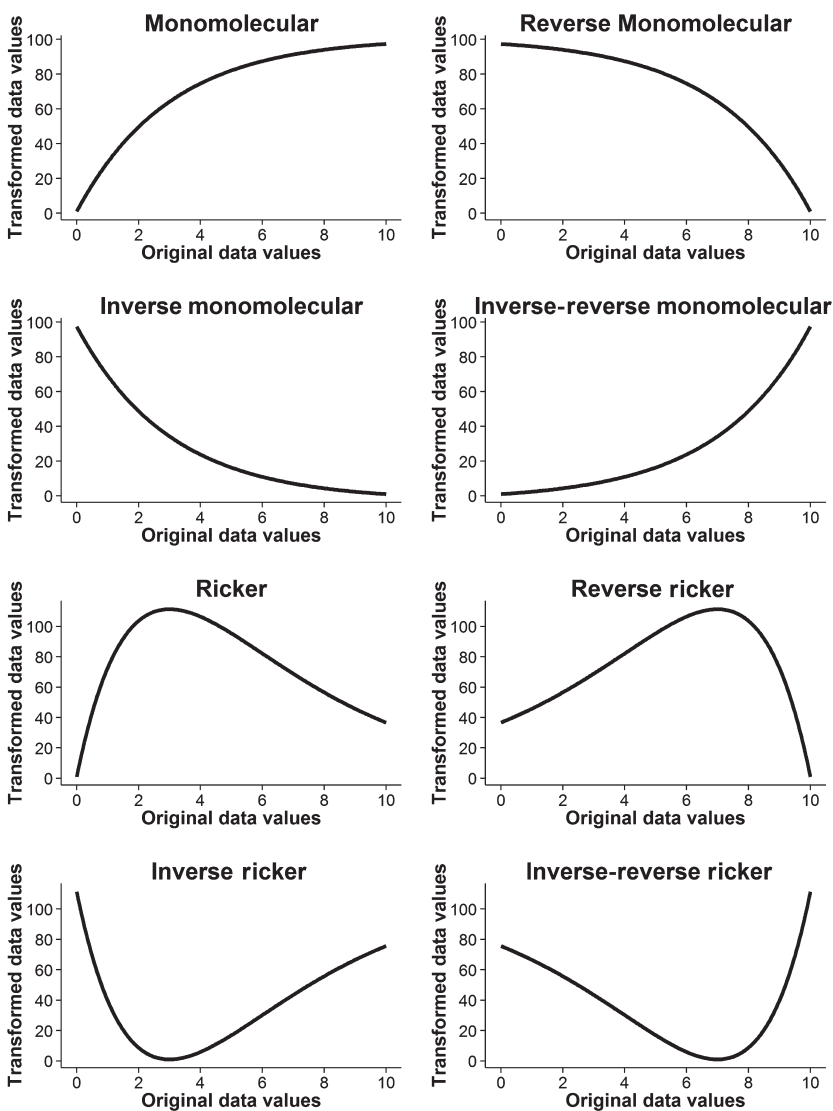

FIGURE 1 There are eight continuous resistance surface data transformations implemented in `ResistanceGA`. Prior to transformation, the original continuous resistance surface had values ranging from 0–10. The shape and magnitude of each transformation are each controlled by a single parameter. All transformations in the figure have a shape parameter value of 3, and maximum value parameter of 100. Linear relationships are not explicitly incorporated, but all monomolecular functions become linear as the shape parameter increases

TABLE 1 Arguments of the required preparation functions and their default settings. Only one of either `CS.prep` or `gdist.prep` needs to be run, depending upon whether optimization will use CIRCUITSCAPE or `gdistance`

Function	Arguments	Defaults	Comments
`CS.prep`	`n.POPS`	Must be defined	
	`response`	NULL	Must be defined to run optimization
	`CS _ Point.File`	Must be defined	
	`CS.program`	"C:/Program Files/Circuitscape/ cs_run.exe"'	Default Windows installation location
	`Neighbor.Connect`	8	
`gdist.prep`	`n.POPS`	Must be defined	
	`response`	NULL	Must be defined to run optimization
	`samples`	Must be defined	
	`transitionFunction`	function(x) 1/M(x)	
	`directions`	8	
	`longlat`	FALSE	
GA.prep	`ASCII.dir`	Must be defined	.asc files should be in their own directory or compiled in a Raster Stack
	`Min.Max`	"max"	Must be "max" when optimization with ga
	`min.cat`	0.0001	
	`max.cat`	2,500	
	`max.cont`	2,500	
	`cont.shape`	NULL	
	`pop.mult`	15	
	`percent.elite`	0.05	
	`type`	"real-valued"	Must be "real-valued"
	`pcrossover`	0.85	
	`pmutation`	0.1	
	`maxiter`	1,000	
	`run`	25	
	`keepBest`	TRUE	
	`population`	"gareal_Population"	
	`selection`	"gareal_lsSelection"	
	`crossover`	"gareal_blxCrossover"	
	`mutation`	"gareal_raMutation"	
	`seed`	NULL	
	`quiet`	FALSE	

before resistance values are assigned. Second, there is currently no closed-form expression to determine the landscape resistance values that describe pairwise effective distances, potentially making optimization intractable with gradient-based algorithms. Finally, landscape features and environmental gradients do not exist in isolation. Therefore, an ideal solution to resistance surface optimization is to simultaneously optimize multiple resistance surfaces to create a composite resistance surface. The `ResistanceGA` package for the R programming environment (R Core Team, 2017) has been developed to address these issues, filling a void in the landscape genetic toolbox. The initial impetus for this package was landscape genetic analyses, but any pairwise measures across the landscape (e.g. movement rates) could be utilized to optimize resistance surfaces

with `ResistanceGA`, potentially making it a valuable, general tool for resistance surface optimization.

2 | DESCRIPTION

2.1 | Genetic algorithms

Genetic algorithms (GAs) provide a powerful and flexible stochastic optimization framework for finding solutions to both discrete and continuous optimization problems (Holland, 1975). Inspired by biological principles, genetic algorithms create a population of individuals (offspring) with traits (parameters to be optimized) encoded on "chromosomes". The genotypes (parameter combinations) of each

TABLE 2 Primary user functions of the `ResistanceGA` package

Function	Returned objects	Description
`all_comb`	`Named list`	Wrapper function used to implement single surface and multisurface optimization, followed by bootstrap analysis of the results
`Combine_Surfaces`	`R` raster object	Combine multiple resistance surfaces into a new composite surface based on specified parameters
`CS.prep`	Named list	This function prepares and bundles the inputs necessary to run CIRCUITSCAPE
`GA.prep`	Named list	This function prepares and compiles the objects and commands needed to execute the genetic algorithm
`gdist.prep`	Named list	This function prepares and bundles the inputs necessary to run `gdistance`
`Grid.Search`	contour plot, `R` data object	For a single continuous surface, this function can be used to visualize the AICc response surface
`k.smooth`	`R` raster object	Applies Gaussian smoothing function to raster surface
`lower`	`R` data object	Convenience function to obtain the lower half of a square distance matrix
`MLPE.lmm`	`lmer` object	Fits a maximum likelihood population effects mixed effects model in `lme4`
`MS_optim`	`GA` object, diagnostic plots, .txt summary file	This is a wrapper function for simultaneously optimizing multiple surfaces
`MS_optim.scale`	`GA` object, diagnostic plots, .txt summary file	This is a wrapper function for simultaneously optimizing multiple surfaces with Gaussian kernel smoothing
`Plot.trans`	`ggplot` object	Implements and plots a continuous surface transformation
`Resist.boot`	`R` data object	Conduct bootstrap analysis with optimized resistance surfaces
`Resistance.tran`	`R` raster object,.asc file	Applies specified transformation to continuous resistance surface and returns a raster object. Optionally, a.asc file can be exported
`Run_CS`	`R` raster object or `R` data object	This function executes CIRCUITSCAPE from R and returns either the lower half of the pairwise resistance distance matrix or the cumulative current map produced by CIRCUITSCAPE
`Run_gdistance`	commuteDistance or costDistance object from `gdistance`	This function executes `gdistance` and returns the either a commuteDistance or costDistance matrix object
`SS_optim`	.tif files,.csv summary tables	This is a wrapper function for optimizing surfaces in isolation. All surfaces in a common directory will be optimized in turn, and numerous summary tables of optimized parameters and objective function values are produced
`SS_optim.scale`	.tif files,.csv summary tables	This is a wrapper function for optimizing surfaces in isolation with Gaussian kernel smoothing. All surfaces in a common directory will be optimized in turn, and summary tables of optimized parameters and objective function values are produced

individual solve the fitness function, and the fittest individuals from each generation survive to reproduce (Sivanandam & Deepa, 2007). The GA evolution process is facilitated by exploration and exploitation (Scrucca, 2013). Exploration of parameter space occurs through random generation of new parameter values resulting from mutation, as well as exchange of genetic information through crossover. Exploitation reduces diversity in the population by selecting the fittest individuals each generation. The population continues to evolve until a sufficient number of generations have passed without an improvement in fitness (Scrucca, 2013).

2.2 | Resistance optimization

`ResistanceGA` utilizes the general-purpose genetic algorithm from the `GA` R package (Scrucca, 2013). Briefly, the optimization proceeds as follows:

1. The original raster surface is imported into `R`. If the surface is continuous, it is rescaled to range from 0–10, preserving the relative spacing between all levels.

2. The evolution process starts by generating a random initial population of size *n* (default = 15× > number of parameters being optimized). If a continuous surface, the selected parameters determine (1) which of eight transformations will be applied (Figure 1); (2) the shape of the transformation; (3) the maximum resistance value. If a categorical surface, each level of the resistance surface is reclassified to the values of the parameters. Starting values for all parameters in the initial population are randomly chosen within the specified range (see Table 1 for defaults).

3. Using the spatial locations where genetic samples have been collected, either CIRCUITSCAPE (McRae, Dickson, Keitt, & Shah, 2008; McRae & Shah, 2009) is called from R to calculate pairwise effective distances across the landscape created in step 2, or

pairwise distances are calculated from least cost paths or random-walk commute times (equivalent to CIRCUITSCAPE resistance distance) using the `R` package `gdistance` (van Etten, 2017).

4. A linear mixed effects model with a maximum likelihood population effects parameterization (MLPE) is fit to the data. Pairwise genetic distance is the response and the scaled and centred pairwise effective distance is the predictor. The MLPE mixed effects parameterization accounts for non-independence among the pairwise data (Clarke, Rothery, & Raybould, 2002), and Shirk, Landguth, and Cushman (2017) recently found that MLPE models performed best in landscape genetic model selection among the seven regression methods assessed.

5. An objective function, specified by the user, is obtained from the fitted MLPE model: log-likelihood, AIC, or marginal R^2 (Nakagawa & Schielzeth, 2013).

6. Steps 2–5 are repeated until the specified number of n individuals have been created. The genetic algorithm then conducts selection on the population, and the individuals with the best objective function values are carried over to the next generation to form the reproducing population (default = top 5% retained each generation). A new population is then formed through mutation and crossover.

7. Steps 2–6 are repeated until the specified number of generations have passed without improvement to the objective function.

2.3 | Continuous surfaces

There are eight transformations that can be applied to continuous surfaces (Figure 1). Transformations are based on Ricker (Equation 1)

and monomolecular (Equation 2) functions (Bolker, 2008), as well as rescaling functions to keep values in positive parameter space. Transformations of resistance surfaces, r, are controlled by shape (x) and magnitude (b) parameters that are varied during optimization (Peterman et al., 2014).

$$y = r \exp^{-bx} \tag{1}$$

$$y = r(1 - \exp^{-bx}) \tag{2}$$

The genetic algorithm searches combinations of transformations, magnitude parameters and shape parameters. Linear transformations are not explicitly included, but all monomolecular functions become linear as the shape parameter increases in value. In this way, linear responses can be modelled without increasing the number of transformations assessed by the genetic algorithm. It is possible to exclude transformations from consideration during optimization.

2.4 | Categorical surfaces

Categorical or feature surfaces, such as land cover or roads, can also be optimized using `ResistanceGA`. A surface is considered categorical in `ResistanceGA` if it contains 15 or fewer levels. To make this process tractable, it is necessary to hold the value of one level constant throughout optimization. Because pairwise resistance values are relative, failing to hold one level constant can result in multiple equivalent solutions, and the algorithm may fail to reach an optimal solution (e.g. relative resistance values of 1, 5 and 10 are equivalent to 2, 10 and 20). See Appendix S1 for more detail about categorical resistance surface optimization.
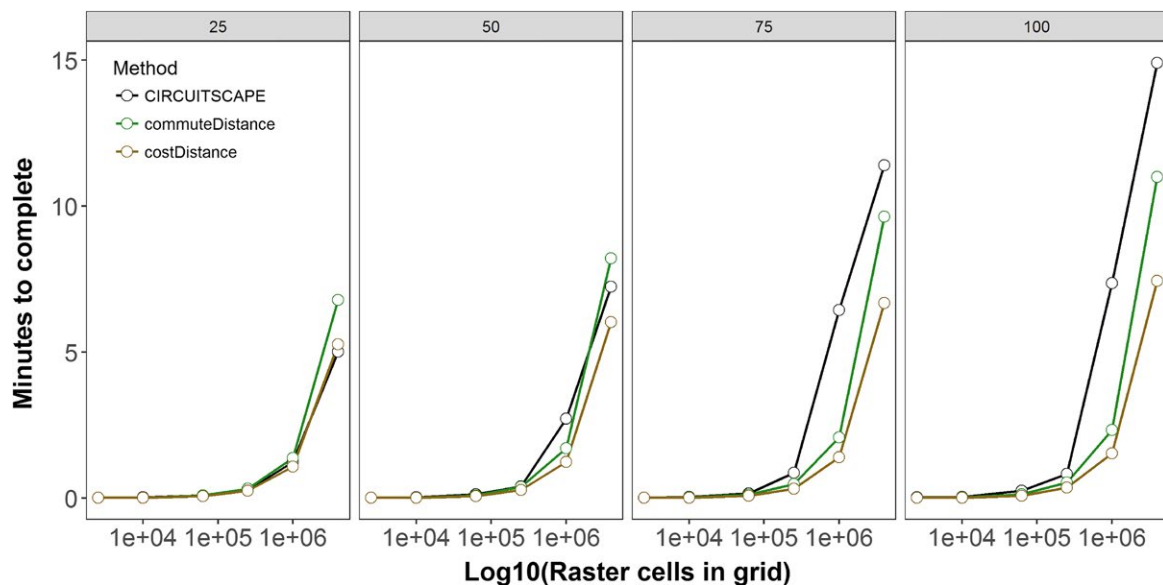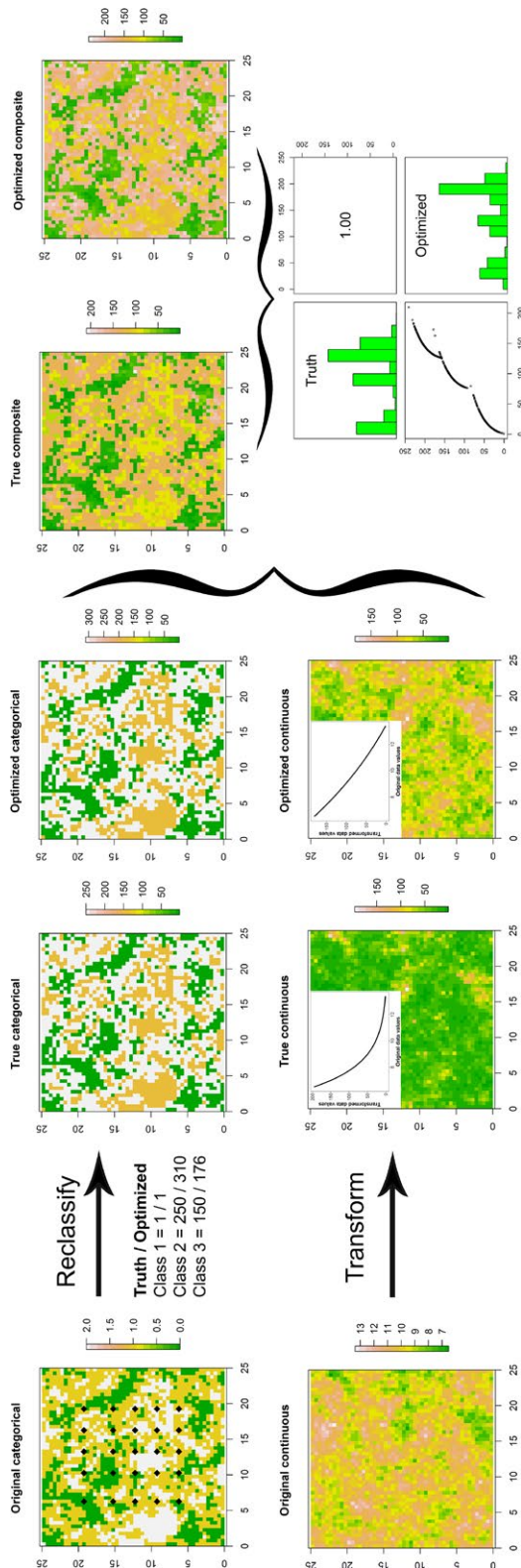

FIGURE 2 Run times for calculating least cost distance (costDistance), random-walk commute time (commuteDistance), or resistance distance (CIRCUITSCAPE) at different sample size and grid size combinations. Sample sizes of 25, 50, 75 and 100 were assessed on grid surfaces that contained 50^2, 100^2, 250^2, 500^2, $1,000^2$ and $2,000^2$ grid cells. Each sample-grid size combination was run ten times and the mean runtime is presented in the figure. Error bars were not visible or meaningful at the scale that the y-axis is presented, but there generally was minimal variation between replicate runs

neighbourhoods (Addicott et al., 1987), wherein a 'neighbourhood' can be determined through the optimization of a kernel smoothing parameter. Using the R package spatstat (Baddeley, Rubak, & Turner, 2015), a Gaussian kernel smoothing is applied to either a continuous or binary resistance surface prior to applying one of the eight transformations. The optimized *SD* of this smoothing is indicative of the ecological neighbourhood for that surface. Scale can be optimized with single surfaces in isolation or separate scales can be optimized for each surface during multisurface optimization.

2.6 | Multisurface resistance optimization

ResistanceGA can simultaneously optimize multiple resistance surfaces to create a novel composite resistance surface. During optimization, each surface in the multisurface analysis is modified using the methods described above. All modified surfaces are then summed together to create a single, composite resistance surface, which is then used to calculate pairwise effective distances.

3 | OVERVIEW OF RESISTANCEGA FUNCTIONS

ResistanceGA optimization functions rely heavily on the R package raster (Hijmans, 2014) to import, export and manipulate spatial raster (*.asc*) files, the gdistance package to calculate cost distances (van Etten, 2017), and lme4 (Bates, Maechler, Bolker, & Walker, 2014) to fit mixed effects models. All functions available in ResistanceGA are summarized in Table 2.

4 | IMPLEMENTATION

2.5 | Scaling resistance surfaces

Scale is a central concept to spatial ecology (Wiens, 1989), but has not been extensively addressed in landscape genetics research (Galpern, Manseau, & Wilson, 2012; Keller, Holderegger, & van Strien, 2013). One way to assess scale is through ecological

Resistance surfaces can be optimized using effective distances calculated using cost distances (least cost path), random-walk commute time, or using circuit-based resistance distances. Cost and commute distances are calculated using the R package gdistance (van Etten, 2014), while electrical current resistances are calculated using CIRCUITSCAPE (v4.0 or greater; McRae & Shah, 2009). Prior
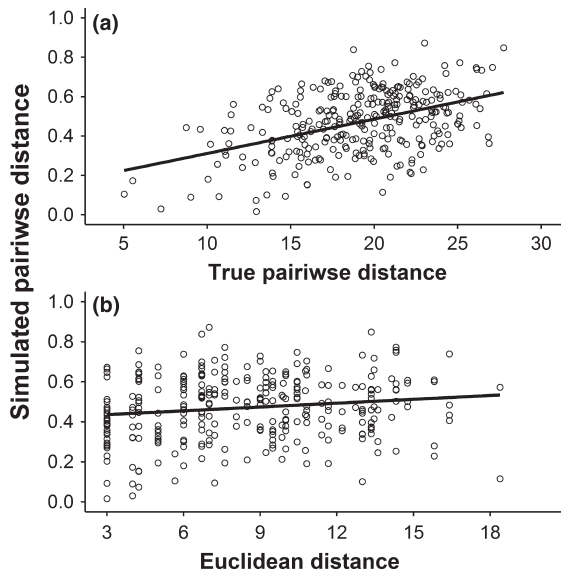
FIGURE 4 Scatterplot showing the correlation between true pairwise response data and simulated pairwise response data with random normal noise added (a), and the correlation between the pairwise simulated response data and the pairwise Euclidean distances (b). The Mantel *r* for each correlation is .45 and .15, respectively

to running any of the optimization functions the `CS.prep/gdist.prep`, and `GA.prep` preparation functions must be run. These functions create and format the inputs and data objects necessary to run CIRCUITSCAPE/`gdistance`, parameterize the genetic algorithm, and fit MLPE mixed effects models. Functions have been developed to require a minimum input from the user, however, all available arguments of the `ga` function can be set by the user to modify the genetic algorithm. The arguments and default settings of the preparation functions are described in Table 1.

Genetic algorithms are stochastic optimization procedures, therefore it is highly advised to run all optimizations at least twice to confirm convergence and parameter estimates. Also, because bounds are placed on the parameter space searched, if the optimized resistance values are at or near the limits set, the optimization should be rerun after expanding the search space. The `all_comb` function has been created to facilitate comprehensive analyses, including replication. This function will carry out single surface optimization followed by multisurface optimization, with all possible combinations of surfaces being assessed. After optimization, a bootstrap analysis is conducted. This bootstrap procedure subsamples the pairwise response and distance matrices generated from each optimized surface (without replacement), refits the MLPE model, and calculates fit statistics for each. The frequency that a model is the top-ranked model is used to assess the level of support for each surface.

Genetic algorithms are effective at finding an optimal solution, but they can be computationally intensive. To ensure that parameter space is adequately searched, the population of individuals produced each generation must be of sufficient size. In `ResistanceGA`, the default setting is to produce a population that is 15 times the number of parameters being optimized. For example, when a continuous surface is being optimized, 45 individuals (3 parameters × 15) will be produced each generation. A typical optimization takes 50–300 generations, resulting in the creation of 2,250–13,500 resistance surfaces and CIRCUITSCAPE/`gdistance` runs. As such, the greatest impediment to optimizing resistance surfaces is time. Both the spatial extent and the number of sample locations affect the time needed to calculate pairwise resistances (Figure 2). Although CIRCUITSCAPE has become a widely adopted and often preferred method of determining pairwise resistances, its use with `ResistanceGA` may be too slow. However, random-walk commute time (function `commuteDistance` in `gdistance`) results in nearly perfectly correlated

TABLE 3 Summary table reporting the true data-generating parameter values of the two surfaces comprising the composite resistance surface and the optimized resistance values determined through three replicate multisurface optimization runs with `ResistanceGA` using the `all_comb` function. Although transformation and resistance values were not exactly recovered during optimization, and different values were obtained between replicate runs, the resultant optimized resistance surfaces closely match the true surface and the absolute mean difference in resistance values of all pixel cells between the true resistance surface and optimized surfaces is minimal

Parameter	Truth	Optimized – replicate 1	Optimized – replicate 2	Optimized – replicate 3
Categorical – class 1	1	1	1	1
Categorical – class 2	250	310.9	412.6	310.0
Categorical – class 3	150	192.2	228.2	175.7
Continuous – transformation[a]	7	5	7	7
Continuous – Shape	2.75	14.9	14.3	14.9
Continuous – Max	150	294.5	481.5	368.1
Percent contribution: Categorical	66.8%	57.7%	58.4%	58.6%
Percent contribution: Continuous	33.2%	42.3%	41.6%	41.8%
Mean resistance value (range)	96.8 (1–209)	142.3 (1–226)	181.3 (1–321)	137 (1–241)
Correlation with truth		1.00	1.00	1.00
Absolute mean difference (±*SD*)		45.85 (±13.27)	84.46 (±34.58)	39.99 (±13.31)

[a]7 is the numeric code for an inverse monomolecular transformation; 5 is the numeric code for a reverse monomolecular transformation.

TABLE 4 Model selection table from complete analysis of sample data included with the `ResistanceGA` package. Surface indicates the surface that was optimized. Surface names separated by a period indicate composite surfaces that resulted from multisurface optimization. The number of parameters fit in each model is indicated by k. AIC is the original AIC value generated from the MLPE mixed effects model, while ΔAICc is the AIC value adjusted for the number of populations sampled and the number of parameters optimized. ω_i is the weight of support for surface i, given the surfaces assessed. R^2m and R^2c are the marginal and conditional R^2 values of the fitted MLPE model, respectively, while LL is the log-likelihood of the MLPE model. In this analysis, the 'categorical.continuous' surface was the true surface. Three replicate optimization runs were conducted, and values in the table are from the run containing the model with the greatest log-likelihood value (replicate 3)

Optimized surface	k	AIC	AICc	ΔAICc	ω_i	R^2m	R^2c	LL
categorical.continuous	4	−293.43	−284.76	0.00	0.548	.151	.151	150.71
categorical	6	−286.17	−284.17	0.59	0.408	.130	.130	147.08
categorical.feature	6	−287.95	−279.29	5.47	0.036	.136	.136	147.98
categorical.continuous.feature	8	−293.47	−276.47	8.29	0.009	.151	.151	150.74
continuous	4	−270.01	−268.01	16.75	0.000	.077	.104	139.00
Distance	2	−264.19	−267.64	17.12	0.000	.054	.097	136.09
feature	3	−264.64	−265.50	19.26	0.000	.053	.095	136.32
continuous.feature	5	−270.22	−265.06	19.70	0.000	.076	.103	139.11
Null	1	−251.75	−255.57	29.19	0.000	.000	.037	128.87

($r ≥ .99$) resistance values that differ only in scaling (van Etten, 2017). On average, across all sample and grid sizes assessed during testing, commute time is 2.1 times faster than CIRCUITSCAPE (Figure 2). To further reduce the optimization time when using `gdistance`, the GA optimization can be parallelized. Additional strategies to reduce runtimes include modifying the connection scheme from the default setting of 8–4, and reducing the resolution of the resistance surfaces (McRae et al., 2008).

5 | WORKED EXAMPLE

The package vignette contains numerous examples demonstrating use and application of functions to optimize resistance surfaces in isolation, in combination, and with a Gaussian kernel smoothing scaling parameter (Appendix S2). The example below uses simulated data provided with the package, and demonstrates how `ResistanceGA` can be used to optimize resistance surfaces, conduct model selection to determine the best supported resistance surface, and make inference about the contribution of each surface to total resistance.

Using the pairwise response data (25 sample locations) and the categorical, continuous, and feature surfaces (each 50^2 pixels) provided with `ResistanceGA`, a true, data-generating, composite resistance surface was created by assigning resistance values of 1, 250 and 150 to the three-class categorical surface and then combining it with the continuous surface that was transformed using an inverse monomolecular function with a shape parameter of 2.75 and maximum value of 200 (Figure 3, Appendix S3). In this composite resistance surface, the categorical layer is responsible for 66.8% of the total resistance, the continuous layer is responsible for 33.2% of the total resistance, and the feature layer

TABLE 5 Summary table reporting the results of a bootstrap analysis conducted using the `Resist.boot` function. Following 1,000 iterations, the optimized categorical.continuous composite surface (the data-generating surface) was ranked as the best supported model in 41% of the bootstrap iterations, which is 5% more support than received by the categorical.continuous.feature composite surface that contained the superfluous feature surface. In combination with the model selection results in Table 3, this provides a preponderance of support for the categorical.continuous composite surface

Optimized surface	Avg. R^2m	Avg. LL	Avg. AIC	Avg. AICc	Avg weight	Avg. rank	Frequency best	k
categorical.continuous	.148	77.44	−146.87	−146.30	0.373	1.43	75.3	6
categorical.continuous.feature	.149	77.45	−146.90	−145.90	0.307	2.40	6.8	8
categorical.feature	.135	76.21	−144.42	−143.85	0.148	3.14	12.8	4
categorical	.131	75.79	−143.59	−143.32	0.133	3.76	3.9	6
continuous	.075	71.57	−135.14	−134.87	0.013	5.51	0.4	4
continuous.feature	.075	71.63	−135.26	−134.85	0.013	5.52	0.7	5
feature	.053	70.14	−132.28	−132.12	0.007	7.03	0.1	3
Distance	.053	70.04	−132.07	−131.99	0.007	7.21	0	2

contributes nothing (determined using the `Combine _ Surfaces` function). Pairwise cost distances were generated using the `commuteDistance` function and random normal error was added (Appendix S3). This "noisy response" data has a Mantel *r* correlation of .45 with the true response data, and a Mantel *r* of .15 with Euclidean distance (Figure 4).

The `all _ comb` function was used to complete three replicate analyses of the three resistance surfaces, including multisurface optimization of all possible combinations of the surfaces. Results were consistent among runs (Table 3). Expanded results for one replicate are hereafter reported. The categorical.continuous surface was top-ranked based on AICc, but indistinguishable from the categorical surface alone (Table 4; ΔAICc = 0.59, ω = 0.41). However, following the bootstrap analysis, the categorical.continuous composite surface was identified as the top model 75% of the time and the categorical surface alone 3.9% of the time (Table 5), thus providing clear and overwhelming support for the categorical.continuous composite. The percent contribution of the optimized surfaces in the composite aligns with the data-generating composite, and the optimized parameter estimates align with the data-generating values (Table 3). The optimized surface is perfectly correlated with the true surface, even though not perfectly optimized (Figure 3, Appendix S3).

6 | DISCUSSION

True optimization of resistance surfaces is an exceedingly difficult task. Numerous approaches have been used in the past (e.g. Dudaniec et al., 2013; Graves et al., 2013; Shirk et al., 2010; Wang et al., 2009), but all either fail to completely search parameter space, rely on a priori assumptions or expert opinion, have minimal capacity to optimize multiple resistance surfaces, or are not generally applicable to both continuous and categorical data types. As demonstrated in the example, `ResistanceGA` addresses all of these limitations. The example also demonstrates the value of model ranking and bootstrap analysis as part of a resistance surface optimization framework. Ultimately, the optimization procedures implemented by ResistanceGA show great potential, but more extensive simulation studies are needed to fully understand the strengths and limitations of ResistanceGA in applied landscape genetic analyses.

7 | OBTAINING ʀᴇꜱɪꜱᴛᴀɴᴄᴇɢᴀ

`ResistanceGA` is hosted on GitHub, and can be downloaded using the `devtools` package:

```
devtools::install.github("wpeterman/ResistanceGA",
                         build_vignettes = TRUE)
```

All examples used version 4.0-0 of ResistanceGA (https://doi.org/10.5281/zenodo.1164840).

DATA ACCESSIBILITY

Example data are provided with `ResistanceGA` on Zenodo (https://doi.org/10.5281/zenodo.1164840) or can be made using the code provided in Appendix S3.

REFERENCES

Addicott, J. F., Aho, J. M., Antolin, M. F., Padilla, D. K., Richardson, J. S., & Soluk, D. A. (1987). Ecological neighborhoods: Scaling environmental patterns. *Oikos*, *49*, 340–346. https://doi.org/10.2307/3565770

Baddeley, A., Rubak, E., & Turner, R. K. (2015) *Spatial point patterns: Methodology and applications with R*. London: CRC Press.

Bates, D. M., Maechler, M., Bolker, B. M., & Walker, S. (2014). Linear mixed-effects models using Eigen and S4. R package version 1.1-6. Retrieved from http://CRAN.R-project.org/package=lme4.

Bolker, B. M. (2008). *Ecological models and data in R*. Princeton, NJ: Princeton University Press.

Charney, N. D. (2012). Evaluating expert opinion and spatial scale in an amphibian model. *Ecological Modelling*, *242*, 37–45. https://doi.org/10.1016/j.ecolmodel.2012.05.026

Clarke, R. T., Rothery, P., & Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, *7*, 361–372. https://doi.org/10.1198/108571102320

Cushman, S. A., McKelvey, K. S., Hayden, J., & Schwartz, M. K. (2006). Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *The American Naturalist*, *168*, 486–499. https://doi.org/10.1086/506976

Driezen, K., Adriaensen, F., Rondinini, C., Doncaster, C. P., & Matthysen, E. (2007). Evaluating least-cost model predictions with empirical dispersal data: A case-study using radiotracking data of hedgehogs (*Erinaceus europaeus*). *Ecological Modelling*, *209*, 314–322. https://doi.org/10.1016/j.ecolmodel.2007.07.002

Dudaniec, R. Y., Rhodes, J. R., Worthington Wilmer, J., Lyons, M., Lee, K. E., McAlpine, C. A., & Carrick, F. N. (2013). Using multilevel models to identify drivers of landscape-genetic structure among management areas. *Molecular Ecology*, *22*, 3752–3765. https://doi.org/10.1111/mec.12359

van Etten, J. (2014). Gdistance: Distances and routes on geographical grids. R package version 1.1-5. Retrieved from http://CRAN.R-project.org/package=gdistance.

van Etten, J. (2017). R package Gdistance: Distances and routes on geographical grids. *Journal of Statistical Software*, *76*, 1–21.

Galpern, P., Manseau, M., & Wilson, P. (2012). Grains of connectivity: Analysis at multiple spatial scales in landscape genetics. *Molecular Ecology*, *21*, 3996–4009. https://doi.org/10.1111/j.1365-294X.2012.05677.x

Graves, T. A., Beier, P., & Royle, J. A. (2013). Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. *Molecular Ecology*, *22*, 3888–3903. https://doi.org/10.1111/mec.12348

Hijmans, R. J. (2014). Raster: Geographic data analysis and modeling. R package version 2.2-31. Retrieved from http://CRAN.R-project.org/package=raster.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence.* Ann Arbor, MI: University of Michigan Press.

Keller, D., Holderegger, R., & van Strien, M. J. (2013). Spatial scale affects landscape genetic analysis of a wetland grasshopper. *Molecular Ecology*, *22*, 2467–2482. https://doi.org/10.1111/mec.12265

Khimoun, A., Peterman, W., Eraud, C., Faivre, B., Navarro, N., & Garnier, S. (2017). Landscape genetic analyses reveal fine-scale effects of forest fragmentation in an insular tropical bird. *Molecular Ecology*, *26*, 4906–4919. https://doi.org/10.1111/mec.14233

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, *18*, 189–197. https://doi.org/10.1016/S0169-5347(03)00008-9

McRae, B. H. (2006). Isolation by resistance. *Evolution*, *60*, 1551–1561. https://doi.org/10.1111/j.0014-3820.2006.tb00500.x

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, *89*, 2712–2724. https://doi.org/10.1890/07-1861.1

McRae, B. H., & Shah, V. B. (2009) *Cicuitscape user's guide.* The University of California, Santa Barbara, CA. Retrieved from www.circuitscape.org

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Peterman, W. E., Connette, G. M., Semlitsch, R. D., & Eggert, L. S. (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Molecular Ecology*, *23*, 2402–2413. https://doi.org/10.1111/mec.12747

R Core Team. (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.

Scrucca, L. (2013). GA: A package for genetic algorithms in R. *Journal of Statistical Software*, *53*, 1–37.

Shirk, A. J., Landguth, E. L., & Cushman, S. A. (2017). A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Molecular Ecology Resources*, *18*, 55–67.

Shirk, A. J., Wallin, D. O., Cushman, S. A., Rice, C. G., & Warheit, K. I. (2010). Inferring landscape effects on gene flow: A new model selection framework. *Molecular Ecology*, *19*, 3603–3619. https://doi.org/10.1111/j.1365-294X.2010.04745.x

Sivanandam, S., & Deepa, S. (2007). *Introduction to genetic algorithms.* Berlin, Heidelberg: Springer.

Spear, S. F., Balkenhol, N., Fortin, M. J., McRae, B. H., & Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Molecular Ecology*, *19*, 3576–3591. https://doi.org/10.1111/j.1365-294X.2010.04657.x

Spear, S. F., Peterson, C. R., Matocq, M. D., & Storfer, A. (2005). Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology*, *14*, 2553–2564. https://doi.org/10.1111/j.1365-294X.2005.02573.x

Stevens, V. M., Verkenne, C., Vandewoestijne, S., Wesselingh, R. A., & Baguette, M. (2006). Gene flow and functional connectivity in the natterjack toad. *Molecular Ecology*, *15*, 2333–2344. https://doi.org/10.1111/j.1365-294X.2006.02936.x

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., … Waits, L. P. (2007). Putting the 'landscape' in landscape genetics. *Heredity*, *98*, 128–142. https://doi.org/10.1038/sj.hdy.6800917

Storfer, A., Murphy, M. A., Spear, S. F., Holderegger, R., & Waits, L. P. (2010). Landscape genetics: Where are we now? *Molecular Ecology*, *19*, 3496–3514. https://doi.org/10.1111/j.1365-294X.2010.04691.x

Wang, I. J., Savage, W. K., & Shaffer, H. B. (2009). Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*). *Molecular Ecology*, *18*, 1365–1374. https://doi.org/10.1111/j.1365-294X.2009.04122.x

Wang, Y.-H., Yang, K.-C., Bridgman, C. L., & Lin, L.-K. (2008). Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, *23*, 989–1000.

Wiens, J. A. (1989). Spatial scaling in ecology. *Functional Ecology*, *3*, 385–397. https://doi.org/10.2307/2389612

Zeller, K., McGarigal, K., & Whiteley, A. (2012). Estimating landscape resistance to movement: A review. *Landscape Ecology*, *27*, 777–797. https://doi.org/10.1007/s10980-012-9737-0

SUPPORTING INFORMATION

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