Multiscale assessment of functional connectivity: Landscape genetics of eastern indigo snakes in an anthropogenically fragmented landscape in central Florida

Javan M. Bauder\textsuperscript{1,2} | William E. Peterman\textsuperscript{3} | Stephen F. Spear\textsuperscript{4,5,6,7} | Christopher L. Jenkins\textsuperscript{4} | Andrew R. Whiteley\textsuperscript{8} | Kevin McGarigal\textsuperscript{1}

\textsuperscript{1}Department of Environmental Conservation, University of Massachusetts, Amherst, MA, USA
\textsuperscript{2}Illinois Natural History Survey, Prairie Research Institute, University of Illinois, Champaign, IL, USA
\textsuperscript{3}School of Environment and Natural Resources, The Ohio State University, Columbus, OH, USA
\textsuperscript{4}The Orianne Society, Tiger, GA, USA
\textsuperscript{5}Department of Fish and Wildlife Resources, University of Idaho, Moscow, ID, USA
\textsuperscript{6}The Wilds, Cumberland, OH, USA
\textsuperscript{7}Upper Midwest Environmental Sciences Center, U.S. Geological Survey, La Crosse, WI, USA
\textsuperscript{8}Department of Ecosystems and Conservation Sciences and Wildlife Biology Program, University of Montana, Missoula, MT, USA

**Correspondence**
Javan M. Bauder, Illinois Natural History Survey, Prairie Research Institute, University of Illinois, Champaign, IL, USA.
Email: javanvonherp@gmail.com

**Funding information**
The Orianne Society; The Society for the Study of Amphibians and Reptiles; U.S. Fish and Wildlife Service; Illinois Natural History Survey; The Herpetologist’s League; The Explorer’s Club; University of Massachusetts Amherst

**Abstract**
Landscape features can strongly influence gene flow and the strength and direction of these effects may vary across spatial scales. However, few studies have evaluated methodological approaches for selecting spatial scales in landscape genetics analyses, in part because of computational challenges associated with optimizing landscape resistance surfaces (LRS). We used the federally threatened eastern indigo snake (*Drymarchon couperi*) in central Florida as a case study with which to compare the importance of landscape features and their scales of effect in influencing gene flow. We used genetic algorithms (ResistanceGA) to empirically optimize LRS using categorical land cover surfaces, multiscale resource selection surfaces (RSS), and four combinations of landscape covariates measured at multiple spatial scales (multisurface multiscale LRS). We compared LRS where scale was selected using pseudo- and full optimization. Multiscale multiscale LRS received more empirical support than LRS optimized from categorical land cover surfaces, multiscale resource selection surfaces (RSS), and four combinations of landscape covariates measured at multiple spatial scales (multisurface multiscale LRS). We compared LRS where scale was selected using pseudo- and full optimization. Multiscale multiscale LRS received more empirical support than LRS optimized from categorical land cover surfaces or RSS. Multiscale LRS with scale selected using full optimization generally outperformed those with scale selected using pseudo-optimization. Multiscale LRS with large spatial scales (1200–1800 m) received the most empirical support. Our results highlight the importance of considering landscape features across multiple spatial scales in landscape genetic analyses, particularly broad scales relative to species movement potential. Different effects of scale on home range-level movements and dispersal could explain weak associations between habitat suitability and gene flow in other studies. Our results also demonstrate the importance of large tracts of undeveloped upland habitat with heterogenous vegetation communities and low urbanization for promoting indigo snake connectivity.

**Keywords**
connectivity, habitat fragmentation, landscape genetics, multiscale, ResistanceGA, spatial scale
The exchange of genetic information among individuals (i.e., gene flow) strongly influences population and evolutionary processes including population genetic structure and diversity (Bohonak, 1999; Keyghobadi, 2007), inbreeding depression and disease resistance (Clark et al., 2011; Madsen et al., 1999), local adaptation (Garant et al., 2007), and speciation (Crandall et al., 2000). Gene flow is influenced by species’ dispersal potential (Bohonak, 1999) and landscape features including vegetation cover, topography, hydrology, and heterogeneity in such features (Cushman et al., 2006; Short-Bull et al., 2011; Tassone et al., 2021). Anthropogenic landscape alterations, including roads, urbanization, and agriculture, may reduce gene flow (Balkenhol et al., 2013; Clark et al., 2010) potentially leading to reduced genetic diversity and inbreeding depression (Clark et al., 2011; Richardson et al., 2016). However, urbanization may facilitate gene flow depending on the taxa and local landscape context (Johnson & Munshi-South, 2017; Miles et al., 2019). Understanding landscape effects on gene flow is therefore not only important for understanding basic ecological processes but also for implementing species’ conservation programs (Keller et al., 2015).

Ecological relationships are dependent upon spatial scale (Levin, 1992; Wiens, 1989). For example, species' relationships to landscape features may not only vary depending on the spatial extent at which those features are measured (Holland et al., 2004; Thompson & McGarigal, 2002) but also across different ecological response variables (Martin, 2018; Moraga et al., 2019; Zeller et al., 2017). Moreover, the scale of strongest ecological response may not be intuitive given current knowledge of a species' ecology making it important to empirically estimate the scale of effect (Jackson & Fahrig, 2015; Moraga et al., 2019). Accurately modelling relationships between gene flow and landscape features therefore requires considering the scale at which landscape features are measured (Anderson et al., 2010; Cushman & Landguth, 2010; Galpern et al., 2012). While the importance of spatial scale has been historically neglected in landscape genetics (Anderson et al., 2010), an increasing number of studies are evaluating landscape genetics relationships across multiple spatial scales (Galpern, Manseau & Wilson, 2012; Row et al., 2015; Winiarski et al., 2020; Zeller et al., 2017).

Many landscape genetics studies quantify landscape effects on gene flow using resistance surfaces which provide spatially explicit estimates of landscape resistance to gene flow (Spear et al., 2010; Zeller et al., 2012). Resistance surfaces also provide an ideal framework for evaluating the effects of scale on landscape-genetic relationships because candidate resistance surfaces representing one or more landscape features can be measured at different scales and compared using inter-individual or population genetic distance. Multiple approaches have been used to parameterize resistance surfaces including assignment of resistance values through expert opinion, habitat suitability models, or empirical comparisons across multiple candidate surfaces with different resistance values (Dudaniec et al., 2016; Shirk et al., 2010; Zeller et al., 2012). Recent developments have permitted the empirical optimization of resistance values directly from observed genetic distance values thereby avoiding the need to specify candidate surfaces (Peterson, 2018). However, researchers must still select from among many potential spatial scales for each covariate thereby greatly increasing the number of candidate resistance surfaces (Wang et al., 2009).

One common approach to optimizing scale in ecological studies is pseudo-optimization wherein a range of scales are evaluated for each covariate individually and a final multiscale model is fit using each covariate at its single-variable optimal scale (McGarigal et al., 2016). However, pseudo-optimization may suffer from variable omission as variation in the data due to other covariates is explained by a single covariate when selecting that covariate’s scale (Stuber & Gruber, 2020). Full (i.e., true) optimization would simultaneously consider all candidate scales for each covariate yet computational limitations have limited its widespread application (Wang et al., 2009). A potential alternative is to derive multiscale resistance surfaces from multiscale habitat models where some form of scale optimization was used within the multiscale habitat model. While many studies have found habitat suitability to poorly predict genetic connectivity (Mateo-Sanchez et al., 2015b; Reding et al., 2013; Roffler et al., 2016), few studies consider how habitat selection integrated over multiple biological levels of activity (e.g., home range selection and within-home range selection) (DeCesare et al., 2012; Johnson, 1980) may predict genetic connectivity. However, issues of scale selection must still be addressed when developing habitat suitability models with which to evaluate their effects on genetic connectivity.

In this study, we used a terrestrial habitat generalist, the eastern indigo snake (Drymarchon couperi, hereafter indigo snake), as a case study with which to evaluate the consequences of different approaches for representing scale while optimizing genetic resistance surfaces. An imperiled species native to the southeastern United States (U. S. Fish & Wildlife Service, 1978, 2008), indigo snakes provide an excellent system for evaluating this topic. Indigo snakes have high vagility across spatiotemporal scales and show contrasting responses to natural and anthropogenic landscape features which may culminate in scale-specific landscape effects on gene flow. Indigo snakes are large bodied (>2 m) active-foragers (>1 km daily movements), have some of the largest reported home ranges for terrestrial snakes (>500 ha), are surface-active throughout the year, and can disperse up to 22.2 km between populations (Bauder et al., 2016b; Breininger et al., 2011; Hyslop et al., 2014; Stevenson & Hyslop, 2010). Indigo snakes show diverse habitat use but predominately select undeveloped upland habitats with high habitat heterogeneity which increases foraging opportunities and shelter site diversity (Bauder et al., 2018, 2020; Hyslop et al., 2014). In contrast, indigo snakes generally, but not exclusively, avoid anthropogenic habitats including urbanization and agriculture (Bauder et al., 2018, 2020; Hyslop et al., 2014) and suffer higher mortality and reduced population viability in urban environments (Breininger et al., 2004, 2012).

Indigo snake gene flow may be driven by different behavioural mechanisms operating at different spatial scales. Gene flow in many terrestrial taxa is often mediated by emigration (i.e., dispersal), often
by juvenile males (Biek et al., 2006; Kelly et al., 2010; Pusey, 1987). However, gene flow in some snakes is mediated by home range-scale male mate-searching movements (Anderson, 2010; Blouin-Demers & Weatherhead, 2002; Clark et al., 2008; Rivera et al., 2006). Male indigo snakes maintain larger home ranges than females, show little intrasex home range overlap, and increase the frequency and extent of their breeding season movements to locate females (Bauder et al., 2016a, 2016b; Breininger et al., 2011). Gene flow driven primarily by male mate-searching movements may manifest itself through greater empirical support for landscape features measured at home range scales and landscape features corresponding to home range-level habitat suitability. In contrast, dispersal-driven gene flow may manifest itself through greater empirical support for landscape covariates at much broader scales.

We addressed three research questions in our study. First, does a full optimization approach suggest different landscape genetic relationships and optimal spatial scales than pseudo-optimization? We predicted that a fully optimized multiscale resistance surface would have more empirical support than a pseudo-optimized multiscale resistance surface. Second, do the optimal spatial scales correspond with home-range scale movements or dispersal? We used multilevel multiscale resource selection surfaces (RSS) to represent the integrated effects of indigo snake habitat selection across multiple spatial scales and hierarchical behavioural levels (DeCesare et al., 2012; Johnson, 1980). We also conducted a spatial autocorrelation analysis to provide complementary estimates of the scale of indigo snake gene flow. We predicted that home range selection would have a stronger association than within-home range selection. We also predicted that scales approximating indigo snake dispersal distance would have even greater support. Finally, what are the landscape features and their functional forms that most influence indigo snake genetic connectivity? While high vagility may compensate for restrictive effects of urbanization on indigo snake gene flow, we nevertheless expected undeveloped heterogenous upland habitats and urbanization (including roads) to promote and restrict gene flow, respectively.

2 | MATERIALS AND METHODS

2.1 | Study area

We collected our samples across an approximately 50 × 20 km area (1000 km²) along the southern Lake Wales Ridge (Figure 1), an important area for indigo snake conservation (Enge et al., 2013). The Lake Wales Ridge has lost approximately 78%–85% of its native vegetation cover to urbanization, citrus, and improved pasture (Turner et al., 2006; Weekley et al., 2008) and remaining natural areas consist of xeric scrub and sandhill communities, scrubby flatwoods, mesic flatwoods, and seasonal forested and nonforested wetlands (Abrahamson et al., 1984; Myers & Ewel, 1990; Weekley et al., 2008). Areas adjacent to the Lake Wales Ridge in our study area were largely dominated by improved and unimproved cattle pasture, the latter often having substantial vegetation cover, forested and non-forested wetlands, hardwood and cabbage palmetto (Sabal palmetto) hammocks, and other natural vegetation communities. While the extent of our study area was limited relative to the maximum indigo snake dispersal distance (22.2 km), our study area and sampling spanned a diversity of landscape features, including roads and urbanization (Figure 1). A restricted study area also helped reduce sampling gaps that would have resulted from expanding the extent of the study area which could result in misleading inferences regarding the effects of landscape features on gene flow (Anderson et al., 2010; Oyler-McCance et al., 2013).

2.2 | Sample collection and laboratory methods

We collected indigo snake tissue samples (scale clips or shed skins) between 2010 and 2014 throughout our study area primarily during a radio telemetry study wherein 90% of captures were opportunistic (Bauder & Barnhart, 2014). Additional samples, particularly road-killed individuals, were collected by authorized project
partners. We extracted DNA using the Qiagen DNeasy blood and tissue extraction kit. We genotyped individuals at 15 microsatellite loci (Shamblin et al., 2011) divided into three multiplexed panels run on an ABI 3130xl sequencer and scored using genemapper software (see Folt et al., 2019 for detail on PCR conditions and multiplex panels). We reran select samples to verify questionable genotypes and retained samples that amplified at ≥13 loci. We calculated the proportion of null alleles per locus using micro-checker v. 2.2.3 and the Brookfield 1 method (Van Oosterhout et al., 2004) and retained loci with <10% null alleles. Because some of our samples were from shed skins from individuals with unknown identity, we used cervus v.3.0.3 (Kalinowski et al., 2007) and the r (v. 3.4.2, R Core Team, 2017) package allelematch (v. 2.5, Galpern et al., 2012) to test for potential duplicate samples (details provided in Appendix S1). We tested for gametic disequilibrium between all pairs of loci using gnenepop (v. 4.2.1, Rousset, 2008). We tested each locus for deviations from Hardy-Weinberg proportions (HWP) using the hw.test function in the pegas package (v. 0.10, Paradis, 2010), calculated FIS using the F.stat function in the demelerate package (v. 0.9-3, Kraemer & Gerlach, 2017), calculated number of alleles, observed and expected heterozygosity using the adegenet package (v. 2.1.1, Jombart, 2008), and calculated allelic richness using the allel.rich function in the popgenreport package (v. 3.0.0, Adamack & Gruber, 2014). We tested the significance of deviations from HWP and FIS using sequential Bonferroni corrections with \( \alpha = 0.05 \) (Holm, 1979).

2.3 | Genetic distance

We calculated an individual-based genetic distance using a principle components (PC) approach following Shirk et al. (2017) and Shirk et al. (2010; Appendix S2). Shirk et al. (2017) found that genetic distance calculated using >1 PC axes performed best with small sample sizes and weak underlying genetic structure. We therefore calculated the number of significant PC axes using the broken stick and latent root criteria (McGarigal et al., 2000) and retained the smaller number of PC axes. Because the amount of variation explained by successive PC axes decreases, we weighted our retained axes by their eigen values.

2.4 | Spatial autocorrelation analysis

To assess the spatial scale(s) of genetic relatedness, we conducted spatial autocorrelation analyses using genalex 6.5 (Peakall & Smouse, 2006, 2012). We calculated the genetic autocorrelation coefficient (r) using 2, 3, and 4 km distance bins to test the null hypothesis that genotypes are randomly distributed in space within each bin. We calculated bootstrapped 95% CI around r using 9999 bootstrap iterations and calculated the 95% CI around the null hypothesis using 9999 random permutations. We considered spatial autocorrelation significant if r was outside of the 95% CI for the null hypothesis and if the bootstrapped 95% CI did not include zero (Peakall et al., 2003).
We conducted separate analyses for adult males and females (snout-vent length >90 cm) to test for sex-biased dispersal (Banks & Peakall, 2012).

2.5 Landscape covariates and hypotheses

We created a composite land cover map using the Florida Natural Areas Inventory land cover map (Kawula, 2014; Knight, 2010), National Wetlands Inventory data (U. S. Fish and Wildlife Service, 2020), and the National Hydrography data set's GIS flowline data (U. S. Geologic Survey, 2020) following Bauder et al. (2018; see Appendix S3 for additional details). We used the 2016 TIGER roads layer (U. S. Census Bureau, 2016) and reclassified roads as primary, secondary, or tertiary. We obtained the Normalized Vegetation Difference Index (NDVI) data calculated from LANDSAT 8 OLI/TIRS using the U.S. Geologic Survey’s Earth Explorer database via the bulk order service (https://espa.cr.usgs.gov/ordering/new/). We calculated a mean spring NDVI using images from 14 May 2013, 2 April 2015, 6 May 2016, 7 April 2017, and 9 May 2017. All data were converted to 15-m pixel rasters to match the original resolution of our land cover map.

We represented our landscape covariates in two ways. First, we created a categorical land cover surface with six thematic classes: urban, undeveloped upland, wetland, citrus, pasture, and open water. We later added different classes of roads to test hypotheses about the restrictive effects of roads on indigo snake gene flow (Table 1). Second, we smoothed the urban, undeveloped upland, wetland, and pasture cover rasters using Gaussian kernels with 60, 600, 1200, and 1800 m bandwidths. The 60 m bandwidth approximated a mean indigo snake home range (Bauder et al., 2018). We also calculated the standard deviation (SD) of NDVI using 60, 600, 1200, and 1800 m radii uniform kernels to represent habitat heterogeneity (Bauder et al., 2018; 2020). We aggregated all surfaces to 60 m pixels using the aggregate function in RASTER to minimize computational time (Appendix S3). We created 60 m pixel rasters for open water (Water Prop) and roads (Roads Prop) where pixel values represented the proportional area of a 60 m pixel that was open water or road (Appendix S3). We gave primary and secondary road pixels a value of five and tertiary road pixels a value of one when creating Prop_Roads to account for greater indigo snake avoidance of primary and secondary roads (Bauder et al., 2018). Finally, we created a binary 60 m pixel roads raster where a pixel was one if it overlapped a primary or secondary road.

2.6 Resistance surface optimization and evaluation

We optimized our resistance surfaces using the R package RESISTANCEGA (v. 4.1-16, Peterman, 2018) which uses a genetic algorithm to optimize resistance surfaces and the functional form of covariate-resistance relationships. Briefly, RESISTANCEGA applies a nonlinear functional transformation to one or more resistance surfaces, calculates cost distance from the transformed surface, fits a linear mixed-effects model using the MLPE parameterization (Clarke et al., 2002), and uses the log-likelihood as the optimization’s objective function. We measured cost distance using CIRCUSTASCPE v. 5.0.0 written with the JULIA programming language (v. 0.6.4, accessed 1 August 2018 at https://julialang.org/). Additional details of our RESISTANCEGA analyses are provided in Appendix S4.

We compared the empirical support among optimized surfaces using a resampling procedure wherein we randomly subsampled 75% of our samples without replacement, refit each model using its optimized resistance surface, and recorded each model’s AIC adjusted for small-sample sizes (AICc, Burnham & Anderson, 2002). We repeated this procedure 10,000 times and calculated the proportion of times each model was selected as the AICc-best model (s). We also report each model’s marginal $R^2$ (i.e., the proportion of variance explained by fixed-effect factors) for mixed-effects models (Johnson, 2014; Nakagawa & Schielzeth, 2013). Finally, to evaluate the importance of each landscape feature within a multisurface optimization, we calculated the percent contribution of each surface by dividing each transformed resistance surface by the sum of the composite multisurface resistance surface.

2.7 Candidate resistance surfaces

We considered three groups of candidate landscape surfaces with which to evaluate the effects of spatial scale on indigo snake gene flow (Figure 2). Our first group included five categorical land cover surfaces, our original six-class surface and four additional categorical surfaces including different combinations of primary, secondary and tertiary roads: primary roads only, primary and secondary roads combined, primary and secondary roads separate, and primary and secondary roads combined but separate from tertiary roads.

Our second group of candidate landscape surfaces were derived from RSS developed from indigo snake in peninsular Florida (Bauder et al., 2018) and represented home range selection (second-order selection, Johnson, 1980), within-home range selection (third-order selection), or both (scale-integrated resource selection surface [SRSS], DeCesare et al., 2012). We used third-order RSS estimated for breeding and nonbreeding season males and females. We calculated each RSS using either the predicted values from a binomial generalized linear model or the exponential form of the resource selection function (Manly et al., 2002) for a total of 18 RSS/SRSS. We linearly rescaled each RSS/SRSS from 0 to 1 and then subtracted all values from one. Because resistance may have a negative exponential relationship with habitat suitability (Keeley et al., 2016), we allowed RESISTANCEGA to optimize the functional form between RSS values and resistance using an inverse or reverse monomolecular function (Peterman, 2018).

Our third group of candidate landscape surfaces included four different a priori sets of landscape covariates hypothesized to influence indigo snake gene flow. These covariates were used to create
multisurface optimized LRS at one or more spatial scales with scale selected using either pseudo or full optimization. These sets included (i) undeveloped upland, urban, and wetland, (ii) undeveloped upland, urban, and SD NDVI, (iii) undeveloped upland, SD NDVI, and wetland edge, and (iv) undeveloped upland, wetland, and pasture. We included undeveloped upland in each set because of its importance in indigo snake habitat selection (Bauder et al., 2018, 2020; Hyslop et al., 2014). We restricted the allowable functional transformations of each landscape covariate to reflect hypothesized directions of relationship (Table S1). We included Water_Prop in all multisurface optimizations.

We evaluated single-scale representations of each a priori set where each covariate, except Water_Prop, was included at the 600, 1200, or 1800 m scale. We then used pseudo optimization to identify the characteristic scale (Holland et al., 2004) for undeveloped upland, urban, wetland edge, and SD NDVI by running a multisurface optimization for each covariate at its 60, 600, 1200, and 1800 m scale with one of the aforementioned covariates and Water_Prop. Each set therefore contained a single pseudo-optimized multiscale multisurface optimization. Pearson correlation coefficients between landscape surfaces included in the same set were <0.35.

Finally, we used full optimization to simultaneously identify the characteristic scales of each landscape covariate within each a priori set. Specifically, we optimized multiscale multisurface LRS across all possible combinations of each landscape covariate within each a priori set, except Water_Prop, at 600, 1200, and 1800 m scales for a total of 27 optimized LRS within each of our four sets. We then compared across the single-scale, pseudo-optimized multiscale, and fully optimized multiscale LRS in each set. We also conducted a post-hoc analysis to test for effects of roads on indigo snake gene flow. Within each a priori set, we first took the single-scale and multi-scale optimized (pseudo and full optimization) LRS whose cumulative \( \pi \) was \( \geq 0.85 \) to reduce computational time and then reran RESISTANCEGA for those LRS with and without either the binary primary/secondary road surface or the weighted proportional road surface.

We compared optimized LRS across our candidate groups of landscape surfaces by calculating \( \pi \) across the top-ranked LRS from our categorical land cover surface, RSS/SRSS, and each multisurface set. We present the optimized LRS from each of these three candidate groups. We also calculated a model-averaged LRS across multisurface LRS with and without roads whose cumulative \( \pi \) was \( \geq 0.85 \) with each surface weighted by \( \pi \).
3 | RESULTS

We used 102 samples that amplified at ≥13 loci after excluding five potential duplicate shed skin samples. We excluded two loci whose estimated null allele frequencies were >10% (Dry14: 15%, Dry68: 11%). Results of genetic summary statistics and isolation-by-distance tests are presented in Appendix S1 and Table S2. Spatial genetic autocorrelation was positive and significant from approximately 8–12 km (Figure 3; Figures S1 and S2). Adult males (n = 45) showed significant positive genetic autocorrelation at greater distances than adult females (n = 36, approximately 10–12 km vs. approximately 4–8 km, Figure 3; Figures S1 and S2).

3.1 | Categorical surface optimization

The top-ranked categorical surface (π = 0.99) included combined primary and secondary roads and separate tertiary roads (Table 1). Open water had the highest resistance followed in decreasing level of resistance by primary/secondary roads, tertiary roads, urban, wetland, pasture, and citrus (Table 1). The isolation-by-distance model received virtually no empirical support although it indicated significant positive isolation-by-distance (β = 0.033, 95% CI = 0.031–0.035).

3.2 | Resource selection surface optimization

Scale-integrated resource selection surfaces using predicted values from the binomial generalized linear model had a cumulative π of 0.97 out-ranking the Level II and Level III RSS (Table 2). The SRSS including the Level III RSS for breeding season females had the most support (π = 0.54) followed by the SRSS including the Level III RSS for nonbreeding season males (π = 0.28). The inverse-reverse monomolecular transformation was selected for all but the four lowest-ranked optimized surfaces (Figure S3). Surfaces using predicted values from the binomial generalized linear model always outperformed surfaces using the exponential form of the resource selection function (Table 2).

3.3 | Multisurface optimizations

The characteristic scales of our landscape covariates were 1800 m for undeveloped upland, wetland edge, and SD NDVI, 1200 m for pasture, 600 m for urban, and 60 m for wetland (Figure 4; Table S3). When comparing only multiscale LRS with pseudo-optimized multi- and single-scale LRS, the pseudo-optimized LRS outperformed all single-scale LRS in two of our four landscape covariate groups (Table 3). The first exception was the group including undeveloped upland, SD NDVI, and wetland edge where the 1800 m single-scale

![Figure 3](image-url) Figure 3 Correlograms showing the spatial scales of genetic autocorrelation at 3 km distance bins for eastern indigo snakes in central Florida across all individuals (a), males (b), and females (c). Error bars represent bootstrapped 95% CI around r for each distance bin and grey lines are the 95% CI around the null hypothesis of no genetic autocorrelation calculated using randomization tests. Bins with significantly positive values of r are denoted by asterisks.
LRS had overwhelming empirical support ($\pi = 0.94$, Table 3). The second exception was the combination including undeveloped upland, SD NDVI, and urban where the 1800 m single-scale LRS had approximately 7.5 times the empirical support as the multiscale LRS with pseudo-optimized scale ($\pi = 0.29$ and 0.04, respectively, Table 3).

Multiscale LRS with scale selected using full optimization outperformed multiscale LRS with scale selected using pseudo optimization.
TABLE 3  Model rankings from empirically optimized multisurface genetic resistance surfaces for eastern indigo snakes in central Florida across four a priori sets of candidate surfaces. We present the empirical support for each single-scale surface and the top multiscale surfaces with scale selected through pseudo optimization or full optimization. We also present the two multiscale surfaces with and without roads.

<table>
<thead>
<tr>
<th>Optimized surfaces</th>
<th>$K$</th>
<th>$AIC_c$</th>
<th>$\pi$</th>
<th>$mR^2$</th>
<th>$cR^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undeveloped upland + SD NDVI + Urban</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (1200)$^b$</td>
<td>13</td>
<td>-14994.49</td>
<td>0.67</td>
<td>0.32</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (1800)</td>
<td>13</td>
<td>-14988.10</td>
<td>0.29</td>
<td>0.33</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (600)$^a$</td>
<td>13</td>
<td>-14983.19</td>
<td>0.04</td>
<td>0.34</td>
<td>0.83</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (1800) + RDS (Bin)$^b$</td>
<td>15</td>
<td>-14982.44</td>
<td>0.00</td>
<td>0.33</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (1200) + RDS (Pr)$^b$</td>
<td>16</td>
<td>-14977.93</td>
<td>0.00</td>
<td>0.33</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (1200) + RDS (Bin)$^b$</td>
<td>15</td>
<td>-14976.82</td>
<td>0.00</td>
<td>0.34</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (1200) + RDS (Pr)</td>
<td>16</td>
<td>-14975.34</td>
<td>0.00</td>
<td>0.35</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (600) + SDNDVI (600) + URB (1200)</td>
<td>13</td>
<td>-14960.62</td>
<td>0.00</td>
<td>0.31</td>
<td>0.83</td>
</tr>
<tr>
<td>UPL (1200) + SDNDVI (1200) + URB (1200)</td>
<td>13</td>
<td>-14957.98</td>
<td>0.00</td>
<td>0.31</td>
<td>0.80</td>
</tr>
<tr>
<td>Undeveloped upland + SD NDVI + Wetland edge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + WetEd (1800)$^c$</td>
<td>13</td>
<td>-14976.27</td>
<td>0.94</td>
<td>0.28</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (600) + SDNDVI (600) + WetEd (600)</td>
<td>13</td>
<td>-14934.21</td>
<td>0.04</td>
<td>0.28</td>
<td>0.82</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + WetEd (1800) + RDS (Bin)$^c$</td>
<td>15</td>
<td>-14963.24</td>
<td>0.02</td>
<td>0.29</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + WetEd (1800) + RDS(Pr)$^c$</td>
<td>16</td>
<td>-14972.25</td>
<td>0.01</td>
<td>0.29</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1200) + SDNDVI (1200) + WetEd (1200)</td>
<td>13</td>
<td>-14942.02</td>
<td>0.00</td>
<td>0.28</td>
<td>0.81</td>
</tr>
<tr>
<td>Undeveloped upland + Urban + Wetland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPL (1800) + URB (1200) + WET (600)$^b$</td>
<td>13</td>
<td>-14960.95</td>
<td>0.50</td>
<td>0.40</td>
<td>0.82</td>
</tr>
<tr>
<td>UPL (1800) + URB (1200) + WET (600) + RDS (Bin)$^b$</td>
<td>15</td>
<td>-14959.69</td>
<td>0.25</td>
<td>0.40</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1800) + URB (600) + WET (60)$^a$</td>
<td>13</td>
<td>-14950.26</td>
<td>0.16</td>
<td>0.34</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1800) + URB (1800) + WET (1800)</td>
<td>13</td>
<td>-14945.14</td>
<td>0.04</td>
<td>0.40</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (600) + URB (600) + WET (60)</td>
<td>13</td>
<td>-14948.38</td>
<td>0.03</td>
<td>0.38</td>
<td>0.83</td>
</tr>
<tr>
<td>UPL (1800) + URB (1200) + WET (600) + RDS (Pr)$^b$</td>
<td>16</td>
<td>-14936.89</td>
<td>0.00</td>
<td>0.36</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1200) + URB (1200) + WET (1200)</td>
<td>13</td>
<td>-14947.19</td>
<td>0.00</td>
<td>0.37</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1800) + URB (600) + WET (60) + RDS (Bin)$^a$</td>
<td>15</td>
<td>-14931.99</td>
<td>0.00</td>
<td>0.33</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1800) + URB (600) + WET (60) + RDS (Pr)$^a$</td>
<td>16</td>
<td>-14931.91</td>
<td>0.00</td>
<td>0.32</td>
<td>0.80</td>
</tr>
<tr>
<td>Undeveloped upland + Wetland + Pasture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPL (1200) + WET (600) + PAST (1800)$^b$</td>
<td>13</td>
<td>-14960.03</td>
<td>0.62</td>
<td>0.42</td>
<td>0.87</td>
</tr>
<tr>
<td>UPL (1800) + WET (60) + PAST (1200) + RDS (Pr)$^a$</td>
<td>16</td>
<td>-14927.79</td>
<td>0.21</td>
<td>0.38</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1800) + WET (60) + PAST (1200)$^a$</td>
<td>13</td>
<td>-14923.90</td>
<td>0.07</td>
<td>0.30</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1200) + WET (1200) + PAST (1200)</td>
<td>13</td>
<td>-14931.41</td>
<td>0.05</td>
<td>0.34</td>
<td>0.83</td>
</tr>
<tr>
<td>UPL (1800) + WET (1800) + PAST (1800)</td>
<td>13</td>
<td>-14922.42</td>
<td>0.03</td>
<td>0.31</td>
<td>0.82</td>
</tr>
<tr>
<td>UPL (1200) + WET (1200) + PAST (1200) + RDS(Bin)$^b$</td>
<td>15</td>
<td>-14956.97</td>
<td>0.02</td>
<td>0.42</td>
<td>0.87</td>
</tr>
<tr>
<td>UPL (600) + WET (600) + PAST (600)</td>
<td>13</td>
<td>-14892.37</td>
<td>0.00</td>
<td>0.24</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1800) + WET (60) + PAST (1200) + RDS (Bin)$^a$</td>
<td>15</td>
<td>-14917.01</td>
<td>0.00</td>
<td>0.30</td>
<td>0.80</td>
</tr>
<tr>
<td>UPL (1200) + WET (600) + PAST (1800) + RDS(Pr)$^a$</td>
<td>16</td>
<td>-14899.65</td>
<td>0.00</td>
<td>0.35</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Abbreviations: Covariate surfaces include: UPL, undeveloped upland; SDNDVI, standard deviation of the normalized vegetation difference index; URB, urban; WetEd, wetland edge; WET, wetland; PAST, pasture; and roads represented as Bin, binary; or Pr, proportional surfaces; $K$ is the number of model parameters, $\pi$ is the proportion of resampling iterations where the model was the $AIC_c$-best model, $mR^2$ is the marginal $R^2$, and $cR^2$ is the conditional $R^2$.

$^a$Denotes the pseudo-optimized surface.

$^b$The best supported fully optimized surface in a given set.

$^c$Indicates that the best-support surface in a set was both the pseudo-optimized surface and the best supported fully optimized surface. The best-supported optimized surfaces with and without the effects of roads are also presented (see text for details).
and single-scale LSR in three of our four a priori landscape covariate combinations (Table 3; Table S4). The only exception was again the combination including undeveloped upland, SD NDVI, and wetland edge at the 1800 m scale (Table 3). The degree of empirical support for the top-ranked LRS with scale selected using full optimization varied within each group (Table S4). The top-ranked LRS for undeveloped upland, SD NDVI, and urban and undeveloped upland, SD NDVI, and wetland edge had strong empirical support relative to other surfaces within their respective combinations ($\pi = 0.63$ and 0.84, respectively, Table S4). Empirical support was lower for the top models in the combinations with undeveloped upland, urban, and wetland and with undeveloped upland, wetland, and pasture ($\pi \leq 0.23$, Table S4) suggesting more uncertainty with regards to the optimal scale. The top-ranked LRS using full optimization included different combinations of scale than identified by pseudo optimization in the three groups where multiscale LSR had the greatest empirical support.

Re-optimizing the top-ranked LRS including roads almost never improved empirical support for those LRS (Table 3; Tables S5 and S6). The one exception was the undeveloped upland, wetland, and pasture combination with pseudo-optimized scale and the proportional roads surface although empirical support was still low relative to other LRS in that combination ($\pi = 0.13$, Table S6). When we compared all well-supported multisurface LRS (excluding LRS with roads; Table S5), the surface with undeveloped upland (1800 m), SD NDVI (1800 m), and urban (1200 m) had strong empirical support ($\pi = 0.47$, Table S7).

The best-supported multisurface multiscale resistance surface including undeveloped upland, SD NDVI, and urban had the greatest empirical support ($\pi = 0.68$) across the best-supported optimized categorical, RSS/SRSS-based, and multisurface resistance surfaces (Table 4). SD NDVI had the greatest percent contribution (0.51) in this optimized resistance surface with lowest resistance at intermediate values of SD NDVI (Figure 5). Resistance was strongly negatively associated with undeveloped upland (percent contribution = 0.20) and increased almost linearly with increased urban (percent contribution = 0.26). The percent contribution of open water was 0.03, perhaps as an artefact of low prevalence across our study area, and indicated high resistance of open water (Figures 5 and 6).

<table>
<thead>
<tr>
<th>Optimized surface</th>
<th>$K$</th>
<th>$AIC_c$</th>
<th>$\pi$</th>
<th>$mR^2$</th>
<th>$cR^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPL (1800) + SDNDVI (1800) + URB (1200)</td>
<td>13</td>
<td>-14994.5</td>
<td>0.68</td>
<td>0.32</td>
<td>0.81</td>
</tr>
<tr>
<td>UPL (1800) + URB (1200) + WET (600)</td>
<td>13</td>
<td>-14961.0</td>
<td>0.18</td>
<td>0.40</td>
<td>0.82</td>
</tr>
<tr>
<td>Scale-integrated resource selection surface - female breeding (predicted)</td>
<td>4</td>
<td>-14958.2</td>
<td>0.10</td>
<td>0.33</td>
<td>0.85</td>
</tr>
<tr>
<td>UPL (1200) + WET (600) + PAST (1800)</td>
<td>13</td>
<td>-14960.0</td>
<td>0.04</td>
<td>0.42</td>
<td>0.87</td>
</tr>
<tr>
<td>UPL (1800) + SDNDVI (1800) + WetEd (1800)</td>
<td>13</td>
<td>-14976.3</td>
<td>0.00</td>
<td>0.28</td>
<td>0.81</td>
</tr>
<tr>
<td>Categorical, All roads</td>
<td>9</td>
<td>-14915.1</td>
<td>0.00</td>
<td>0.44</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Abbreviations: Covariate surfaces include: UPL, undeveloped upland; SDNDVI, standard deviation of the normalized vegetation difference index; URB, urban; WET, wetland; WetEd, wetland edge; $\pi$, proportion of resampling iterations where the model was the $AIC_c$ best model.

4 | DISCUSSION

4.1 | Consequences of scale optimization approach

Our results highlight the importance of considering the effects of multiple landscape features across a range of spatial scales in landscape genetics analyses. We found that LRS estimated using multiple landscape covariates measured at different spatial scales strongly outperformed LRS estimated from categorical land cover surfaces. RSS/SRSS, and, generally, multiple landscape covariates measured at a single scale. These results are consistent with other landscape genetics studies showing scale-specific relationships between gene flow and landscape features (Galpern, Manseau, Hettinga et al., 2012; Row et al., 2015; Winiarski et al., 2020; Zeller et al., 2017) and broader cautions within ecology against selecting a single scale for analysis (a priori; Jackson & Fahrig, 2015; McGarigal et al., 2016). Our results also show that the manner in which characteristic scales are identified can influence empirical support for LRS. Consistent with our first prediction, LRS using truly optimized scales outperformed LRS using pseudo-optimized scales in three of our four multisurface combinations. Winiarski et al. (2020) also found that LRS with scales selected using full optimization outperformed those using pseudo optimization. Moreover, they also found that optimized single-surface LRS sometimes outperformed multisurface pseudo-optimized LRS. It is also unclear how generalizable our results from this single study area are to other landscapes or taxa particularly as the extent of relevant scales will vary depending on the organism’s ecology. While pseudo optimization is widely used in multiscale ecological analyses (McGarigal et al., 2016), pseudo-optimized scale selection may suffer from omitted variable bias (Stuber & Gruber, 2020). Such bias may at least partly explain our results. While several landscape genetics studies have evaluated the entire parameter space more-or-less thoroughly with regards to resistance values and functional form (Dudaniec et al., 2016; Shirk et al., 2010; Wang et al., 2009), considering scale(s) within the parameter space will markedly increase the parameter space and computational times. However, our approach of combining resistanceEGA optimization with "manual" comparisons across a...
BAUDER ET AL.

discrete range of scales may prove an effective means of optimizing multiscale resistance surfaces. We encourage landscape geneticists to carefully consider the effects of scale in their analyses while also pursuing analytical solutions incorporating multiscale relationships.

4.2 | Optimal scales: Dispersal or home range?

Broadscale landscape covariates received strong empirical support in our analyses while LRS optimized from categorical surfaces or RSS/SRSS received lower support. This pattern is consistent with

---

**FIGURE 5** Functional transformations for landscape covariates within the top-ranked optimized landscape resistance surface within four a priori sets of landscape covariates. We present the scale of each covariate, its percent contribution (PC) to the optimized landscape resistance surface, and the proportion of subsampling iterations across all best-supported optimized LRS where that model was the top-ranked model (π, see Table 4)
FIGURE 6 The top-ranked optimized landscape resistance surfaces for eastern indigo snakes in central Florida from candidate (a) categorical land cover surfaces, (b) scale-integrated resource selection surfaces, and (c) a priori landscape covariate sets. (d) shows a weighted average across optimized surfaces across the top five a priori landscape covariate sets (Table 3). [Colour figure can be viewed at wileyonlinelibrary.com]
other studies reporting poor correlation between gene flow and LRS derived from habitat suitability models (Mateo-Sanchez et al., 2015a; Reding et al., 2013; Roffler et al., 2016; Wasserman et al., 2010). A commonly cited reason for such discrepancies is the potential for individuals to successfully disperse through low-suitability habitats indicating that habitat suitability does not necessarily correlate linearly with dispersal resistance (Elliot et al., 2014; Gaston et al., 2016). However, some studies found that home range-level RSS/SRSS accurately captured dispersal patterns (Fattebert et al., 2015; Zeller et al., 2018). For example, Row et al. (2015) found that resistance surfaces derived from habitat suitability models provided a better description of greater sage grouse (Centrocercus urophasianus) gene flow than those derived from combinations of landscape covariates although they did not optimize independent combinations of covariates. While low habitat selectivity of dispersing individuals could explain the poor performances in our study of our LRS optimized from categorical land cover surfaces or RSS/SRSS, landscape covariate relationships with indigo gene flow were largely consistent with those from habitat selection studies (Bauder et al., 2018, 2020; Hyslop et al., 2014).

The greater support for broadscale landscape covariates in our results may also reflect differences in scales-of-effect between individual habitat selection and multigenerational gene flow (Moraga et al., 2019). The lower empirical support for LRS optimized from categorical land cover surfaces or RSS/SRSS also suggest that indigo snake gene flow is mediated more strongly by dispersal rather than male mate-searching movements. The observed scale of indigo snake spatial genetic autocorrelation (4–12 km) and widespread indigo snake gene flow across peninsular Florida (Folt et al., 2019) further supports this hypothesis. However, additional field research is needed to understand the life history stages responsible for indigo snake gene flow. Simulation (Jackson & Fahrig, 2014) and field (Moraga et al., 2019) studies have also shown that landscape features can influence genetic diversity at broader spatial extents than abundance or occurrence. While habitat selection studies may therefore be useful for selecting particular combinations of landscape covariates for estimating LRS, we recommend caution when selecting covariate combinations based solely on individual habitat selection. We also recommend that researchers consider covariates across at multiple spatial scales including scales several times larger than the mean home range size of their study species.

4.3 Implications for indigo snake gene flow

Our predictions that large amounts of heterogeneous undeveloped upland habitat would promote indigo snake gene flow while urbanization would restrict indigo snake gene flow were largely supported. Our results indicate that, despite the differences in characteristic scale, landscape covariate relationships with indigo snake gene flow and habitat selection were largely consistent. For example, LRS with SD NDVI and urban generally had strong empirical support and Bauder et al. (2018, 2020) found that undeveloped upland, SD NDVI, and urban most strongly influenced indigo snake habitat selection. In contrast, LRS surfaces with wetland and pasture had lower empirical support consistent with the weaker effects of these covariates on indigo snake habitat selection (Bauder et al., 2018, 2020).

Our prediction that roads would restrict indigo snake gene flow was not generally supported which was surprising given previous observations that indigo snakes avoided crossing primary and secondary roads (Bauder et al., 2018) and that urbanization can strongly reduce indigo snake survival (Breininger et al., 2012). We had samples on both sides of multiple major highways (Figure 1) so this result is unlikely due to sampling bias. Roads have also been found to genetically fragment populations in other snake species (Clark et al., 2010). One hypothesis for our lack of strong negative road effects is that roads in our study area have been present for an insufficient length of time to create detectible levels of indigo snake genetic differentiation. However, the primary and secondary roads in our analyses have been present since at least 1952 (U. S. Geologic Survey, 2021) which corresponds to approximately 15–21 indigo snake generations (Breininger et al., 2004). This is probably sufficient time for a detectible genetic signature from roads to emerge if roads indeed strongly restricted indigo snake gene flow (Landguth et al., 2010). An alternative hypothesis is that roads do not strongly limit indigo snake gene flow, perhaps through infrequent but consistent successful road crossings, particularly if individuals were able to cross under culverts or bridges. This hypothesis is consistent with both the high vagility of indigo snakes and extensive genetic similarity across peninsular Florida (Folt et al., 2019). Moreover, genetic connectivity may be maintained through the successful dispersal and reproduction of relatively few individuals per generation (Mills & Allendorf, 1996). However, roads may still negatively affect indigo snake demographic connectivity and therefore increase the risk of localized extinction (Breininger et al., 2004).

4.4 Limitations and conclusions

We acknowledge several limitations in our study design and analyses. First, indigo snakes in our study area were extremely difficult to detect resulting in sampling gaps that did not necessarily represent indigo snake distributional gaps thereby limiting our ability to completely infer the effects of landscape features on gene flow (Anderson et al., 2010; Oyler-McCance et al., 2013). However, our use of bootstrap resampling in evaluating model support should at least partially ameliorate artefacts of the spatial arrangement of our sampling points. Finally, relationships between gene flow and landscape features can vary across landscapes (Richardson et al., 2016; Short-Bull et al., 2011), so we stress caution when transferring our results to other parts of the indigo snake's distribution, particularly the northern part of the indigo snake's distribution where indigo snake habitat associations differ from those in our study area (Hyslop et al., 2014).

Our results reiterate the importance of large tracts of heterogeneous upland habitat for promoting indigo snake population viability.
and connectivity (Bauder et al., 2018; Breininger et al., 2004). While our results suggest that roads do not strongly inhibit indigo snake gene flow, we encourage future research to test this hypothesis in other parts of the indigo snake's distribution and in relation to other road classes (e.g., interstates). We also encourage future research to determine whether indigo snake genetic permeability to roads is due to culverts or bridges, rare successful road crossings, or other mechanisms. Finally, given increasing development pressures across Florida (Turner et al., 2006), we encourage future efforts to provide range-wide spatially-explicit estimates of indigo snake demographic and genetic connectivity to guide conservation planning efforts.

ACKNOWLEDGEMENTS
Funding was provided by the U.S. Fish and Wildlife Service, The Orianne Society, the Department of Environmental Conservation at the University of Massachusetts Amherst, the Illinois Natural History Survey, an E.E. Williams Research Grant from the Herpetologist's League, a Roger Conant Grant in Herpetology from the Society for the Study of Amphibians and Reptiles, and an Exploration Fund Grant from The Explorer's Club. Z. Forsburg, L. Paden, and P. Barnhart assisted with field work and R. Johnson and B. Olson assisted with lab work. The Archbold Biological Station and B. Rothermel provided logistical support. Many private landowners provided access to their properties and G. Graziani, D. Parker, B. Rothermel, and J. Sage provided additional samples. The Ohio Super Computer provided computational resources. This study was conducted under permits from the U.S. Fish and Wildlife Service (TE28025A-1), Florida Fish and Wildlife Conservation Commission (WEX97328), University of Florida Institutional Animal Care and Use Committee (200903450) and the Archbold Biological Station Institutional Animal Care and Use Committee (ABS-AUP-002-R). The comments of three anonymous reviewers greatly improved this manuscript.

AUTHOR CONTRIBUTIONS

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
Microsatellite genotypes, ascii landscape covariate surfaces, and R script are available on FigShare at https://doi.org/10.6084/m9.figshare.c.5282129.

ORCID
Javan M. Bauder https://orcid.org/0000-0002-2055-5324
William E. Peterman https://orcid.org/0000-0001-5229-9268

REFERENCES


**SUPPORTING INFORMATION**

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