Received: 11 March 2023 | Revised: 17 June 2023 | Accepted: 25 July 2023

DOI: 10.1111/mec.17098

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Corridor-based approach with spatial cross-validation reveals scale-dependent effects of geographic distance, human footprint and canopy cover on grizzly bear genetic connectivity

Eric C. Palm[1,2] | Erin L. Landguth[1,3] | Zachary A. Holden[4] | Casey C. Day[1] | Clayton T. Lamb[5] | Paul F. Frame[6] | Andrea T. Morehouse[7] | Garth Mowat[8,9] | Michael F. Proctor[10] | Michael A. Sawaya[11] | Gordon Stenhouse[12] | Jesse Whittington[13] | Katherine A. Zeller[2]

[1]Computational Ecology Lab, School of Public and Community Health Sciences, University of Montana, Missoula, Montana, USA
[2]Rocky Mountain Research Station, Aldo Leopold Wilderness Research Institute, US Forest Service, Missoula, Montana, USA
[3]Center for Population Health Research, School of Public and Community Health Sciences, University of Montana, Missoula, Montana, USA
[4]Northern Region, US Forest Service, Missoula, Montana, USA
[5]Department of Biology, University of British Columbia, Kelowna, British Columbia, Canada
[6]Fish and Wildlife Stewardship Branch, Government of Alberta, Whitecourt, Alberta, Canada
[7]Winisk Research and Consulting, Pincher Creek, Alberta, Canada
[8]Wildlife & Habitat Branch, British Columbia Ministry of Forests, Lands, Natural Resource Operations & Rural Development, Nelson, British Columbia, Canada
[9]Department of Earth, Environmental and Geographic Sciences, UBC Okanagan, Kelowna, British Columbia, Canada
[10]Birchdale Ecological Ltd., Kaslo, British Columbia, Canada
[11]Sinopah Wildlife Research Associates, Missoula, Montana, USA
[12]Grizzly Bear Program, fRI Research, Hinton, Alberta, Canada
[13]Parks Canada, Banff National Park Resource Conservation, Banff, Alberta, Canada

Correspondence
Eric C. Palm, Computational Ecology Lab, School of Public and Community Health Sciences, University of Montana, Missoula, MT 59812, USA.
Email: eric.palm@mso.umt.edu and e2palm@gmail.com

Funding information
Alberta Conservation Association; BC Conservation Corps; BP Canada; Calgary Foundation; Columbia Basin Trust; Counter Assault; Forest Enhancement Society of British Columbia; fRI Research; Government of British Columbia; Habitat Conservation Trust Foundation; Mountain Equipment Co-operative; National Fish and Wildlife Foundation; Parks Canada; Safari Club International Foundation; Sparwood and District Fish and Wildlife Association; Teck Coal; Vanier Canada Graduate Scholarship; Western

Abstract

Understanding how human infrastructure and other landscape attributes affect genetic differentiation in animals is an important step for identifying and maintaining dispersal corridors for these species. We built upon recent advances in the field of landscape genetics by using an individual-based and multiscale approach to predict landscape-level genetic connectivity for grizzly bears (*Ursus arctos*) across ~100,000 km^2 in Canada's southern Rocky Mountains. We used a genetic dataset with 1156 unique individuals genotyped at nine microsatellite loci to identify landscape characteristics that influence grizzly bear gene flow at multiple spatial scales and map predicted genetic connectivity through a matrix of rugged terrain, large protected areas, highways and a growing human footprint. Our corridor-based modelling approach used a machine learning algorithm that objectively parameterized landscape resistance, incorporated spatial cross validation and variable selection and explicitly accounted for isolation by distance. This approach avoided overfitting, discarded variables that did not improve

---

Transportation Institute; Wilburforce Foundation; Woodcock Foundation

Handling Editor: Lisette Waits

model performance across withheld test datasets and spatial predictive capacity compared to random cross-validation. We found that across all spatial scales, geographic distance explained more variation in genetic differentiation in grizzly bears than landscape variables. Human footprint inhibited connectivity across all spatial scales, while open canopies inhibited connectivity at the broadest spatial scale. Our results highlight the negative effect of human footprint on genetic connectivity, provide strong evidence for using spatial cross-validation in landscape genetics analyses and show that multiscale analyses provide additional information on how landscape variables affect genetic differentiation.

KEYWORDS
connectivity, grizzly bear, landscape genetics, machine learning, spatial cross-validation

1 | INTRODUCTION

Connectivity is a fundamental concept in ecology in which landscape structure and species movements interact to influence spatial population dynamics and species distributions (Moilanen & Hanski, 2001; Taylor et al., 1993). Genetic connectivity that promotes species persistence requires sufficient gene flow to maintain genetic diversity as well as avoid harmful effects of local inbreeding and potential spread of disadvantageous alleles across a species' range (adaptive connectivity; Lowe & Allendorf, 2010, Fitzpatrick & Reid, 2019). Understanding how landscape configuration and composition affect gene flow and genetic differentiation is key to informing effective conservation and management decisions for plants and animals in fragmented landscapes (Keyghobadi, 2007).

The field of landscape genetics attempts to characterize whether and to what degree environmental factors influence genetic connectivity (Manel et al., 2003). Most landscape genetics analyses model genetic (dis)similarity between individuals or populations as a function of effective distances (e.g. least-cost path or resistance distances; Manel & Holderegger, 2013) between the spatial locations associated with those same individuals or populations. Improvements in the fit of resistance models using isolation by resistance, or some other measure of landscape structure (McRae, 2006), compared to those with geographic distance alone (i.e. isolation by distance; Wright, 1943), provide evidence that landscape characteristics influence genetic structure.

Recent work indicates that considering multiple spatial scales in landscape genetics approaches provides a more comprehensive understanding of how landscape factors contribute to genetic differentiation, especially for highly mobile species (Balkenhol et al., 2020). Many terrestrial mammals typically disperse relatively short distances and rarely undergo long-distance dispersal (Whitmee & Orme, 2013) and understanding how these dispersal patterns influence genetic connectivity is a major focus of landscape ecology and landscape genetics. Most landscape genetic studies consider only the largest spatial scale of their data, perhaps limiting their ability to distinguish between processes that occur at different scales, such as genetic exchange within home ranges or long-distance dispersal events. Therefore, analyses that only consider one spatial scale may not capture the scale dependence of ecological processes contributing to genetic differentiation.

Another major focus of landscape genetics research is modelling the degree to which human development has affected genetic connectivity and understanding whether any additional landscape resistance from this development is biologically significant and worth mitigating (e.g. Ernest et al., 2014; Thatte et al., 2020). Genetic connectivity maps produced from these models can highlight corridors that may have supported gene flow across a much longer timescale (i.e. generations) than those from movement-based connectivity maps and should correlate with factors that increase survival and maximize fitness (Landguth et al., 2010; Zeller et al., 2017).

Researchers typically parameterize landscape genetics models using resistance values, which represent the degree to which a landscape facilitates or impedes movement (Spear et al., 2010). Modelling approaches that quantify the relative influence of landscape factors on genetic structure often include the creation of resistance surfaces for environmental attributes that are based on expert opinion or spatial predictions from habitat suitability models (Zeller et al., 2012). Because organisms are simultaneously influenced by multiple landscape attributes that are not fully independent of one another, Peterman and Pope (2021) argued that effective distances should therefore be estimated from a single resistance surface that encompasses multiple landscape and environmental characteristics affecting gene flow.

Landscape genetics literature provides many different approaches for parameterizing these predictor variables, transforming them to resistance values and validating models, but there is little consensus on which of these approaches is most appropriate (Vanhove & Launey, 2023). In addition, there may not be existing knowledge on how certain variables, alone or in combination, influence genetic connectivity. The process of selecting a suite of landscape predictors and determining how to transform and combine them into a single resistance surface from which to estimate effective distances can therefore result in many different possible outcomes. Furthermore, models that are not validated using either within-sample or external data may be highly predictive for a specific dataset but are not transferrable to other landscapes (Peterman & Pope, 2021) or teach us little about general limits to connectivity.

Another major confounding factor in landscape genetics studies is spatial autocorrelation, which occurs when nearby samples are genetically more similar than samples separated by greater geographical distances (Legendre, 1993; Meirmans, 2012). Spatial autocorrelation can lead to biased predictions and incorrect inference, especially in heterogeneous landscapes and across large spatial extents (Ploton et al., 2020). Spatial autocorrelation is especially ubiquitous in models for wide-ranging species relating pairwise measures of genetic distances and environmental data, as each spatial location associated with a genetic sample occurs many times within input datasets (*N* genetic samples become $(N*(N-1))/2$ pairwise observations; Franckowiak et al., 2017). Landscape genetics studies have attempted to account for spatial autocorrelation in regressions in various ways, such as including geographic distances between pairwise sampling locations (e.g. and/or a kernel density estimate of sample locations as model variables; Bishop et al., 2021; Pless et al., 2021), or some form of clustering by pairwise geographic distances (Bouyer et al., 2015; Jaffé et al., 2019). Other approaches, such as maximum-likelihood population effects (MLPE), address the non-independence inherent in pairwise distances (genetic, geographic and effective) by using random effects terms for both populations (or both individuals) in each link (Clarke et al., 2002), but do not explicitly account for spatial autocorrelation.

In ecological modelling, spatial cross validation (CV) has gained attention as an important tool that helps account for spatial autocorrelation, allowing practitioners to more reliably evaluate model performance and estimate variable importance (Meyer et al., 2019; Ploton et al., 2020). However, spatial CV has not been used in landscape genetics analyses. Cross-validation procedures generally partition a dataset into subsets, fit the model using all but one subset (i.e. training data) and validate the model on the remaining subset (i.e. test data). In spatial CV ('spatial blocking'), these training and test partitions can be stratified by sample location or the proximity to other samples (Roberts et al., 2017), a process that can reduce spatial dependence and overly optimistic error estimates in model results and increase predictive performance when extrapolating to areas with little or no data (Meyer et al., 2019; Whalen et al., 2022). In analysis based on pairwise distances, spatial CV should cluster observations by accounting for both the start and end locations of each location pair.

A machine learning algorithm paired with spatial CV is an example of a modelling approach that explicitly addresses the two main issues mentioned above: objectively estimating landscape resistance and accounting for spatial autocorrelation. Machine learning models improve upon existing approaches that objectively estimate landscape resistance because they can accommodate highly correlated predictor variables and effectively capture variable interactions and complex nonlinear relationships. One of these existing approaches models genetic distances between pairwise observations directly from raw (untransformed) landscape data that are summarized (e.g. mean) along straight lines between location pairs, and predictions from these models produce a resistance layer that encapsulates the combined effects of individual predictors on genetic differentiation (van Strien et al., 2012). In this corridor-based approach, researchers use the resistance surface to calculate least-cost paths between origin and destination locations, refit the model and repeat the entire process (optimize) until the model performance (measured by root-mean-squared error or a similar metric) no longer improves. Bouyer et al. (2015) applied this iterative optimization in a maximum-likelihood framework, and recent studies modified this approach to use machine learning models (Bishop et al., 2021; Pless et al., 2021). Vanhove and Launey (2023) used simulations to evaluate the performance of several landscape genetics approaches for creating resistance surfaces and found that this machine learning corridor-based approach outperformed MLPE models in multivariate scenarios, more closely matching the expected contribution of individual raster surfaces to the final multivariate resistance surface than MLPE. By also including spatial CV, a process that is easily integrated into machine learning algorithms using R packages such as 'caret' (Kuhn, 2008) and 'tidymodels' (Kuhn & Wickham, 2020), researchers using any of these approaches to create resistance surfaces can account for spatial autocorrelation and improve reliability of spatial predictions.

Grizzly bears are a flagship species for conservation efforts throughout much of their range, and maintaining their genetic connectivity to sustain viable populations is a major conservation challenge (Lamb et al., 2020). In south-east British Columbia (BC) and south-west Alberta (AB), grizzly bear populations are isolated by highways and other forms of human development, which can increase mortality risk and impede genetic and demographic connectivity between areas (e.g. Proctor et al., 2005; Sawaya et al., 2014). Previous research from this area found that grizzly bears living in some human-dominated landscapes rely on connectivity corridors that allow immigration from nearby wilderness areas to offset human-caused mortality and sustain viable populations (Lamb et al., 2020). Within our study area, past studies' estimates of average home range sizes for adult grizzly bears were ~200–500 km^2 for females and ~1000–1400 km^2 for males (Graham & Stenhouse, 2014; Herrero, 2005), while the maximum recorded dispersal distance is 175 km (Proctor et al., 2004).

Existing studies on grizzly bear connectivity have mostly relied on habitat selection models fit GPS collar tracking data (but see Proctor et al., 2012). However, connectivity predictions from these studies may not reflect genetic connectivity, as habitat selection and gene flow occur at different spatiotemporal scales and may be influenced by different underlying processes (Roffler et al., 2016). Compared to GPS telemetry data, genetic data have the advantage of accounting for successful reproduction over generations, and using individual-level rather than population-level genetic data may increase statistical power to identify relationships between landscape variables and connectivity (Kierepka & Latch, 2015; Landguth et al., 2010).

Here, we predicted landscape-level genetic connectivity for grizzly bears (*Ursus arctos*) in Canada's southern Rocky Mountains, building upon the recent advances in the corridor-based method made by Pless et al. (2021) in four ways: first, we used an individual-based approach; second, we incorporated spatial CV; third, we employed variable (feature) selection; and fourth, we conducted the modelling at multiple spatial scales, each with a different maximum pairwise geographic distance. We sought to identify landscape characteristics that limited their genetic connectivity at each scale and to produce continuous wall-to-wall predictive maps of connectivity. Mapping grizzly bear genetic connectivity in this area may help inform ongoing efforts to identify and prioritize mitigation and maintain movement corridors that promote genetic connectivity across subpopulations. We compiled a genetic dataset encompassing over 1000 individuals from across nearly 100,000 km^2, and predicted that across all spatial scales, landscape characteristics would explain more variation in genetic distances than isolation by distance (IBD) alone. Specifically, we tested support for four competing hypotheses of which fundamental factors limit grizzly bear connectivity in our study area: lack of hiding cover, low forage quality, increased ruggedness and natural barriers and increased disturbance. We represented each hypothesis in our models by using one or more environmental predictor variables, as described below in *Methods*. Our models also allowed for all two-way interactions between these variables. We predicted that relatively undeveloped river valleys within mountainous areas would have the highest genetic connectivity (Proctor et al., 2015), as they provide dispersal corridors that minimize energy expenditure (Carnahan et al., 2021) and maximize safety due to low risk from people.

2 | MATERIALS AND METHODS

2.1 | Study area

We modelled connectivity using microsatellite genotype data from grizzly bears across ~100,000 km^2 of south-eastern BC and south-western AB, including a large portion of the southern Canadian Rocky Mountains (Figure 1). The area encompassed several large protected areas, including Banff, Glacier, Kootenay, Waterton and Yoho National Parks, along with several provincial parks. Mountains were heavily glaciated in the north-west but less so farther south. The Rocky Mountain Trench, a deep low-elevation linear valley with human settlements and major roads, ran roughly north–south along the western edge of the study area. The region was characterized by pronounced temperature and precipitation gradients that were highly correlated with elevation. Mountainous areas were characterized by long, cold winters and relatively short summers, and most precipitation occurred in the spring. Forested foothills on the eastern slopes of the mountains gave way to relatively flat, drier areas with aspen parklands, prairies and croplands. A handful of major highways crossed the region, primarily not only along valley bottoms but also across several high mountain passes. Human development was largely confined to valley bottoms and was denser east of the Rocky Mountains.

2.2 | Microsatellite data

Genetic samples were collected from 1992 to 2019 (median year: 2012), primarily using hair snares as part of genetic tagging projects to estimate demographic and genetic parameters (Lamb et al., 2019; Morehouse & Boyce, 2016; Proctor et al., 2010; Sawaya et al., 2012, 2014), but additional samples were collected during live captures or from dead bears (Whittington et al., 2018).

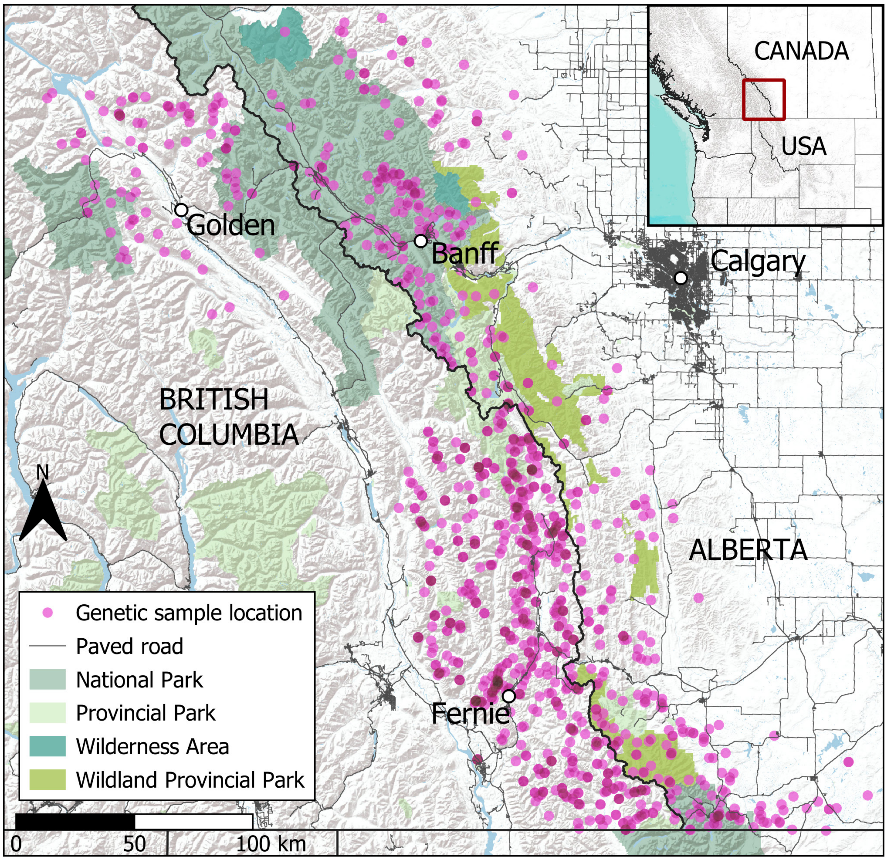

FIGURE 1 Study area and locations of 1161 grizzly bear genetic samples in south-east British Columbia and south-west Alberta, Canada. Darker colours indicate overlapping sample locations.

All microsatellite data were processed at Wildlife Genetics International, a commercial genetics lab in Nelson, BC, Canada. Samples were genotyped at a minimum of 9 or 15 microsatellite loci (see Paetkau, 2003 and Morehouse & Boyce, 2016 for details on DNA extraction, error checking, identification of individuals and data filtering). We conducted exploratory analyses comparing results using 9 versus 15 loci (Figure S1) and proceeded with data from bears genotyped at 9 loci (CXX110, G1A, G1D, G10B, G10H, G10J, G10M, G10P and MU50; Kendall et al., 2009). This allowed us to use data from 1161 unique grizzly bears (vs. 374 unique bears genotyped at 15 loci). For bears where we had GPS collar data ($n = 78$), we used the closest location to the centroid of GPS locations for that bear to best approximate each bear's home range centre. Similarly, in cases with >2 genetic samples, we use the sample location closest to the centroid of the genetic samples. We used the sample location with the earliest date when there were only two genetic samples per individual. We calculated genetic distances between all 673,380 pairwise combinations using 1 – proportion of shared alleles (Bowcock et al., 1994) and Euclidean genetic distance (Excoffier et al., 1992) in the 'gstudio' R package (Dyer, 2012). We chose these metrics because they do not assume Hardy–Weinberg equilibrium and they performed well in a formal comparison of individual-based genetic distances (Shirk et al., 2017). The correlation between the two metrics was .92. We conducted a principal components analysis (PCA) and used the first principal component of these two genetic distances as the response variable in all models. A PCA transforms many correlated variables into fewer uncorrelated variables with minimal information loss. The correlation between the PCA of genetic distances and both independent distances was .98.

We also calculated a suite of genetic diversity indices for grizzly bears across all nine loci and the entire study area. First, we generated a genotype accumulation curve with R package 'poppr', version 2.9.3 (Kamvar et al., 2014), to confirm that our microsatellite panel had the power to identify all the individuals in our sample. This method determines the minimum number of microsatellite loci needed to discriminate among 99% of unique bear genotypes given a random sample of loci. To further confirm our power to identify individual bears from genotypes, we calculated the probability of identity between siblings using the 'popGenUtils' R package (Stourvas, 2022). We assessed deviations from Hardy–Weinberg equilibrium using exact tests, using 1000 Monte Carlo permutations, across all loci using the 'pegas' R package (Paradis, 2010), and deviations from linkage disequilibrium (or non-independence of alleles at each locus) using the index of association in the 'poppr' R package.

2.3 | Environmental data

We used environmental variables that we hypothesized might explain variation in movement and habitat selection of grizzly bears as predictor variables in our model based on past grizzly bear studies in the region (e.g. McClelland et al., 2020; Nielsen et al., 2009; Proctor et al., 2015; Roever et al., 2010). We had four competing hypotheses for which landscape factors would best explain gene flow in our models, each represented by at least one environmental variable (Table 1). Canopy cover represented the hiding cover hypothesis and summer precipitation and enhanced vegetation index (EVI; Huete et al., 2002) represented the forage quality hypothesis. We chose EVI over the normalized difference vegetation index because it is more sensitive in areas with dense vegetation and accounts for canopy background noise. Terrain ruggedness index (TRI) and glacier cover represented ruggedness and natural barriers, and two layers of human disturbance represented disturbance and mortality risk. We used 30-m-resolution canopy cover data from the Global Forest Cover Change Tree Cover Multi-Year Global dataset (Townshend, 2016).

We tested three different layers for human disturbance: a 'built environment' layer for year 2015 (see details below), a 300-m-resolution Canada-specific human footprint index layer for year 2016 (Hirsh-Pearson et al., 2022) and a layer with only paved roads using government data from 2022 (Government of British Columbia, 2017; Government of Canada, 2022). The built environment layer represented the footprints of roads, towns, buildings and mines. To create this layer, we took the 'urban' category from the 30-m-resolution North American Land Cover Change Monitoring System's 2015 land cover layer (NALCMS 2020) and manually added large mine footprints from Landsat imagery from 2012, which was the median year of our genetic samples.

We calculated TRI (Riley et al., 1999) with a 9-pixel window size using 30-m-resolution elevation data from NASA's Shuttle Radar Topography Mission (Farr et al., 2007), and calculated average summer EVI for 2013–2016 in Google Earth Engine using data from 30-m-resolution Landsat 8 Tier 1 8-day EVI composite (Roy et al., 2014). We used the snow and ice category from the 30-m-resolution North American Land Cover Change Monitoring System's 2015 land cover layer to represent glacier cover. We used bilinear resampling to convert all continuous raster data to 180 m resolution, and nearest neighbour resampling for the binary built environment, paved roads and glacier layers.

2.3.1 | Spatial genetic diversity indices

Prior to analyses, we created a Mantel spatial correlogram of genetic distances between all pairs of grizzly bears in our study area, which showed positive autocorrelation in spatial genetic structure up to an interpolated distance of ~120 km, beyond which genetic distances were negatively autocorrelated (Figure S2). We then estimated allelic richness, inbreeding coefficient (fixation index) and the effective number of breeding individuals across the study area in the sGD package (Shirk & Cushman, 2011) using 120 km as the spatial genetic neighbourhood.

TABLE 1 Details of hypotheses for landscape effects on gene flow for 1161 grizzly bears (*Ursus arctos*) in Canada's southern Rocky Mountains.

Hypothesis	Rationale	Variable(s)	Prediction	Reference
Isolation by distance	Null model	Pairwise geographic distance	Greater geographic distance inhibits gene flow	Wright (1943)
Hiding cover	Safer to travel and rest in closed canopies when near humans	Canopy cover	Lower canopy cover inhibits gene flow	Gibeau et al. (2002)
Human disturbance	Bears generally avoid and/or experience higher mortality in areas near people	Built environment (2015) or human footprint index (2016) or paved roads	Higher human disturbance inhibits gene flow	Lamb et al. (2020)
Topography and natural barriers	Rugged terrain and glaciers may limit bear movement	Terrain ruggedness index (TRI)	More rugged terrain inhibits gene flow	Carnahan et al. (2021)
		Glacier cover	Glaciers inhibit gene flow	Proctor et al. (2012)
Forage quality	Bears move between sites with high-quality forage	Enhanced vegetation index (EVI)	Lower EVI (lower forage quality) inhibits gene flow	Proctor et al. (2015)
		Summer precipitation	Drier areas (lower forage quality) inhibit gene flow	Mowat et al. (2013)

2.4 | Landscape genetics analysis

2.4.1 | Multiscale datasets

The clear transition from positive to negative spatial autocorrelation in genetic distances (Figure S2) suggested there were multiscale effects in grizzly bear genetic structure throughout our study area. After creating all pairwise combinations of genetic locations, we created six separate analysis datasets that covered a gradient of spatial scales, each with a different maximum pairwise geographic distance (40, 80, 120, 200, 300 and 440 km).

2.4.2 | Corridor-based modelling approach

We extracted the mean value along buffered straight lines between all pairwise locations for each of the six environmental variables in our models using the 'exactextractr' R package (Baston, 2021). We chose the mean value, as in Pless et al. (2021) and Vanhove and Launey (2023), rather than the median (Bishop et al., 2021) because it was more sensitive for capturing intersections with roads and built environments, which were relatively rare throughout our study area. We tested buffer widths of 0, 1, 3 and 5 km along straight lines and least-cost paths (LCPs; see below for details) to determine their effect on the resulting mean covariate values. Both Pless et al. (2021) and Bishop et al. (2021) used a 1 km raster resolution with no buffer. Next, we ran gradient boosting trees (GBM; 'gbm' R package; Greenwell et al., 2020) and random forest (RF) models ('ranger' R package; Wright & Ziegler, 2017) using the covariates extracted from these straight lines as explanatory variables and genetic distance as the response variable. We included geographic distance in all models, which allowed us to predict its relative influence in final models and helped prevent selection of false-positive variables in the spatial CV and variable selection process described below (Row et al., 2017), and we neutralized the effect of geographic distance when mapping results from our top models by using a constant value (median pairwise geographic distance across all pairwise locations). These maps served as initial composite landscape resistance surfaces from which we calculated LCPs between all pairwise genetic sample locations.

2.4.3 | Spatial cross-validation

We added to Pless et al.'s (2021) machine learning, corridor-based approach by incorporating both spatial CV and variable selection into the modelling process (Figure 2). As part of the straight-line model, we employed spatial cross-validation coupled with a forward variable selection algorithm in 'CAST' (Meyer et al., 2023) and 'caret' R packages to remove uninformative variables and minimize overfitting. Prior to modelling, we created 10 spatial clusters of all 1161 sampling locations (Figure S3) using the 'kmeans' function in R, which uses the algorithm of Hartigan and Wong (1979) to partition the locations in a way that minimizes the sum of squares from the locations to the centroid of their assigned cluster. We used these 10 spatial clusters to create CV folds. For each fold, we iteratively trained models using all pairwise observations except those where either the start or end location corresponded to a location from that fold (spatial cluster). Pairwise paths (either straight lines or LCPs) therefore appeared in multiple test (withheld) datasets in the cross-validation procedure. We compared results from models using spatial CV to those using random CV, where we randomly assigned each pairwise observation to 1 of 10 test datasets of equal size. We also evaluated the performance of models using random

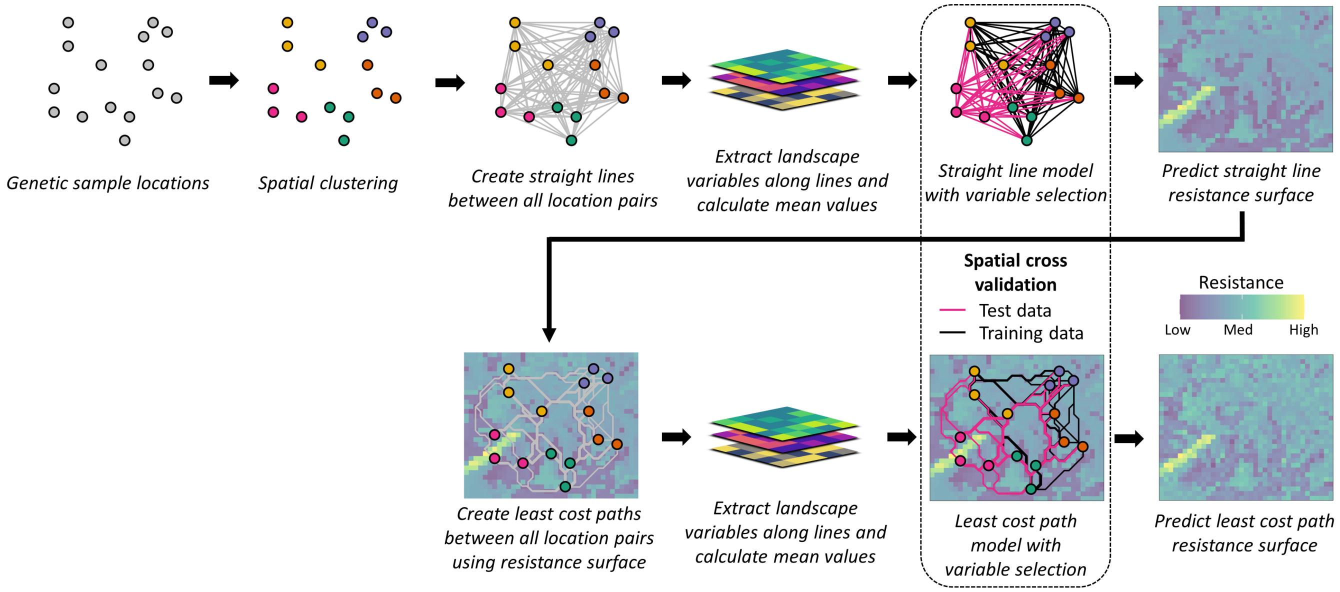

FIGURE 2 Analysis workflow for modelling the influence of landscape characteristics on genetic distance of 1161 grizzly bears (*Ursus arctos*) in Canada's southern Rocky Mountains. The straight-line and least-cost path model steps depict an example where all pairwise observations (lines) associated with the pink locations are withheld as test data in a cross-validation fold, while the model is fit using the training data from the remaining (black) lines. We followed this workflow for all six spatial scales.

versus spatial CV when predicting regions of the study area where data were completely withheld from model training. To do this, we withheld data from regions (separately withholding each spatial test cluster in Figure S3), fit the top variable-selected models for random and spatial CV at the 440-km scale to the remaining data, predicted genetic distances in the withheld datasets using the fitted models and calculated RMSE between predicted versus observed genetic distances.

2.4.4 | Variable selection

We implemented forward variable selection with a two-variable minimum and used the same 10 spatial clusters in the leave-one-cluster-out CV procedure as above. This process started by fitting all possible models with unique combinations of two environmental covariates and evaluating their predictive performance. The algorithm then iteratively increased the number of covariates in each model, tested for improvements over the existing best models and stopped when additional variables no longer improved model performance. We used the resulting models to make predictions for the withheld test observations and calculated the average (±SD) prediction accuracy across all folds. As part of this process, we conducted a random grid search to determine the optimal suite of model hyperparameters to achieve the highest predictive accuracy as measured by root mean squared error of the test dataset ($\mathrm{RMSE}_{\mathrm{test}}$). GBM hyperparameters included shrinkage, number of trees and minimum number of observations per terminal node, while RF hyperparameters included the minimum node size and the number of randomly selected predictors at each tree split. We used the optimal hyperparameters to fit final straight-line models to the entire dataset and make spatial predictions. We repeated the variable selection process in subsequent LCP models. We calculated the relative influence of selected variables in GBM models that included all variables using the method described in Friedman (2001). For RF models, we used the Gini index of classification as a measure of variable importance (Breiman et al., 2017). We explored running separate models for males and females but decided to pool the data because spatial autocorrelation in genetic distances was nearly identical between sexes (Figure S3). Furthermo0re, we assumed that dispersal movements of breeding males would spread their alleles to both male and female offspring alike, minimizing our ability to map sex-specific genetic connectivity.

2.4.5 | Model selection

We calculated LCPs using the spatial predictions from our straight-line models, extracted covariates along these (buffered) paths and used the same CV approach described above to identify the top LCP model. We then generated spatial predictions from the LCP model and repeated the steps of LCP calculation, covariate extraction and modelling using the previous iteration's resistance surface. We used the iteration and associated predicted resistance surface with the lowest $\mathrm{RMSE}_{\mathrm{test}}$ for connectivity modelling. For each spatial scale's best model, we calculated the relative percent difference between each withheld spatial cluster's $\mathrm{RMSE}_{\mathrm{test}}$ and the mean $\mathrm{RMSE}_{\mathrm{test}}$ across all clusters. We used these values to estimate the relative predictive capacity of the models in each spatial region.

2.4.6 | Maps of predicted genetic connectivity

We created resistant kernels on the final resistance surfaces using UNICOR (Landguth et al., 2012) to map predicted genetic connectivity. Given an input resistance surface and a set of source locations, resistant kernels use a cost–distance algorithm to down-weight each pixel by the cumulative resistance from source locations and produce a surface of expected density of dispersing animals (Cushman et al., 2013). A recent simulation study found that resistant kernels outperformed Circuitscape when movement was not directed from or to a specific location, as is the case with grizzly bears in our study system (Unnithan Kumar & Cushman, 2022). We used 2000 source locations distributed randomly throughout the study area.

3 | RESULTS

3.1 | Genetic analysis summary

The genotype accumulation curve showed that six microsatellite loci were needed to identify 99% of the multilocus grizzly bear genotypes in our dataset (Figure S4). All nine microsatellite loci were polymorphic. Three loci (G10B, G10H and G10J) were under the null expectation of Hardy–Weinberg equilibrium (Figure S5). The deficit of heterozygotes in other loci may reflect inbreeding or the presence of null alleles. Observed frequencies of null alleles ranged from 0.001 to 0.022 across all nine loci and did not significantly differ from zero for five loci (CXX110, G10B, G10H, G10J and MU50). Tests for linkage disequilibrium indicated that some loci may be linked, but the value was near 0 (overall standardized index of association = 0.009; *p*-value = .001), suggesting rejection of the null hypothesis of random mating. The highest standardized index of association between all pairwise combinations of loci was 0.02, indicating that only a small proportion variation in one marker is shared with any other marker. Microsatellite loci with linkage disequilibrium could reflect subpopulation structure or genetic drift (Slatkin, 2008). A Mantel test for isolation by distance showed that geographic distances and genetic distances between grizzly bears were significantly correlated ($r = .112$, p-value = .001). We calculated a cumulative $\mathrm{PID}_{\mathrm{sibs}}$ value of 2.5×10^{-4} across all loci, indicating high statistical power to differentiate between closely related individuals and a minimal chance that our results are confounded by multiple individuals having matching multilocus genotypes.

3.2 | Spatial genetic diversity

Genetic distance was positively correlated with geographic distance throughout the study area (Pearson's correlation across all spatial CV test datasets: 120-km scale = .097 ± .029; 440-km scale = .099 ± .018; Figure S6). Genetic diversity measured through allelic richness generally increased from south to north, peaking around the Bow Valley in Banff National Park. Inbreeding was lowest in the far north-west and relatively similar throughout the remainder of the study area (Figure S7).

3.3 | Landscape genetics model results

Across all spatial scales, withheld test datasets in the spatial cross-validation procedure represented 82.3% ± 8.4 (mean ± SD; range = 69.6% – 97.7%) of the total dataset depending on the number of sample locations (individual bears) in each spatial cluster. GBM and RF models had very similar model performance, as measured by the mean $\mathrm{RMSE}_{\mathrm{test}}$ across all 10 spatial CV folds (Table S1). However, accumulated local effects plots depicting average effects of covariates on model predictions across the full range of covariate values (Apley & Zhu, 2020) showed more stable responses for GBM versus RF in portions of covariate ranges with little data (Figure 3; see Figure S8 for a comparison of GBM and RF predictions). Models incorporating spatial CV retained fewer variables in the straight-line and LCP variable selection processes (GBM spatial: paved roads + canopy cover [440-km spatial scale only] + geographic distance; GBM random: all six landscape variables + geographic distance), had lower RMSE across withheld test data (Table S1) and showed more stable covariate responses than those using random CV coupled with variable selection (Figure S9). Variable-selected models at the 440-m scale using spatial CV outperformed those using random CV model at predicting genetic distances for all regions when we completely withheld data from those regions from model training (spatial CV RMSE = .111 ± .003; random CV RMSE = .443 ± .107; Figure S10). We report results and display spatial predictions from the GBM models with spatial CV hereafter. We report the optimal set of hyperparameters for the top GBM models for each spatial scale in Table S2.

In our final models, geographic distance was the most important variable across all spatial scales, indicating that IBD explained a larger portion of variation in genetic distance for grizzly bears than the landscape variables. The 440-km spatial scale model predicted that pairwise geographic distances above ~85 km led to increased genetic differentiation from the mean and predicted that maximum genetic differentiation occurs beyond ~120 km (Figure 3). All models incorporating paved roads outperformed (lowest average $\mathrm{RMSE}_{\mathrm{test}}$ across 10 spatial CV folds) those using either the built environment or human footprint layers. The presence of paved roads between pairs of genetic samples was associated with higher pairwise genetic distance across all spatial scales, indicating that these features consistently impede gene flow. The signal of higher resistance in open canopies and lower resistance in closed canopies first appears in the ALE plots at the 200-km spatial scale and its magnitude increases at larger spatial scales (Figure 3) along with its relative influence in the GBM models that included all variables (Figure 4). Canopy cover was retained in variable selection only in the full dataset including all pairwise geographic distances (440-km spatial scale).

At all spatial scales, final variable-selected models outperformed models that included all landscape variables using $\mathrm{RMSE}_{\mathrm{test}}$ as the metric (Figure S11) and explained more variation in genetic differentiation than models with only geographic distance across all spatial scales (Figure S12). The first LCP iterations of models

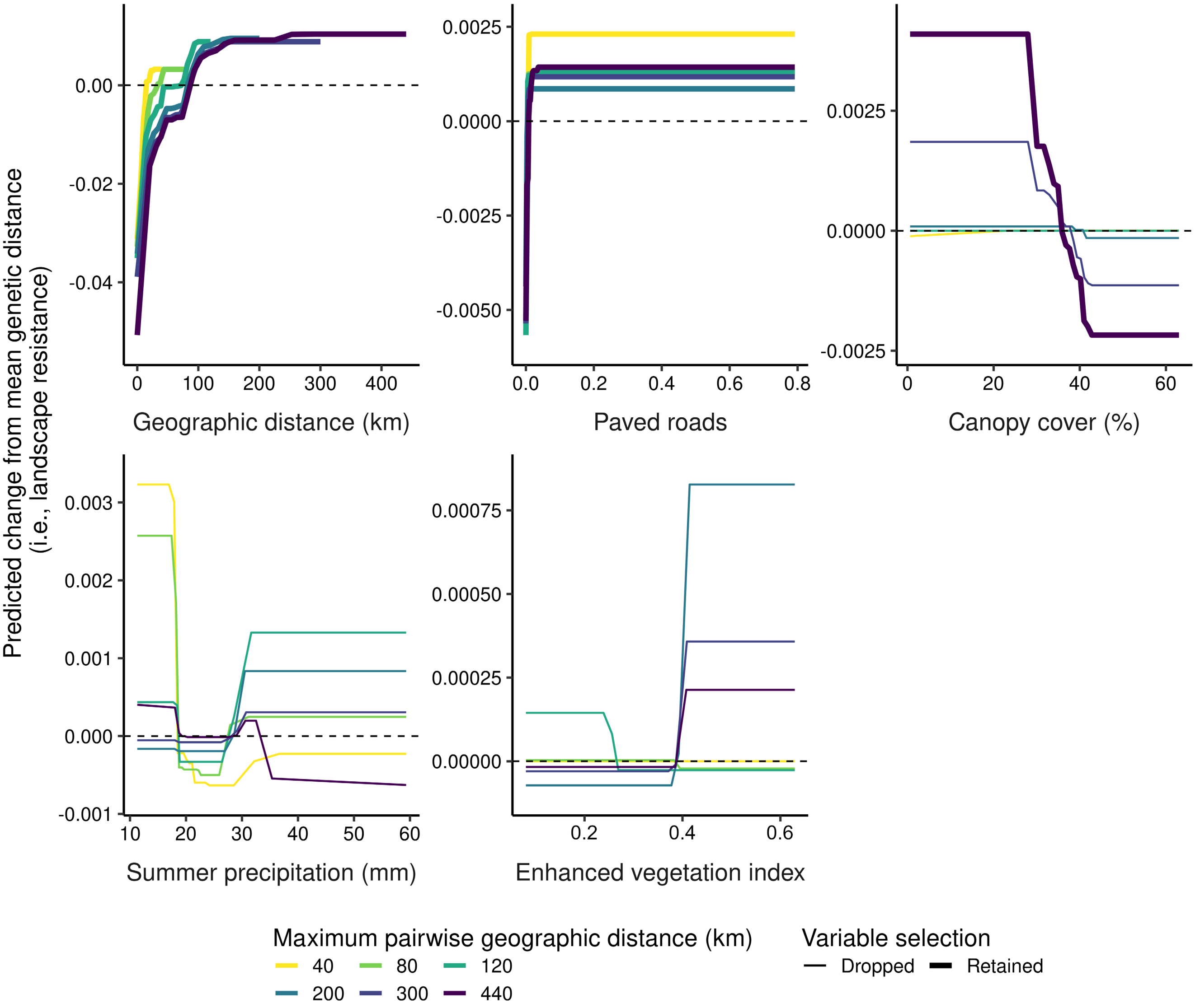

FIGURE 3 Accumulated local effect plots showing the predicted change in genetic distance from GBMs at six different spatial scales using data from 1161 grizzly bears (*Ursus arctos*) in Canada's southern Rocky Mountains across a range of values for different landscape variables. Values on the y-axis depict the main effect of the variable at a certain value compared to the average predicted genetic distance (which is 0) between all pairwise combinations of bears in the entire dataset. Therefore, positive y-axis values indicate higher predicted genetic distance than average or more landscape resistance. Note that the y-axis differs across panels. Terrain ruggedness and glaciers were not included because they were rarely influential and their effects across spatial scales produced no discernible pattern.

with geographic distance and paved roads for the 40-km, 80-km, 120-km, 200-km and 300-km spatial scales performed better than the straight-line models at the same spatial scale, so we used these iterations to predict connectivity via resistant kernels. We did not see further improvement in $\mathrm{RMSE}_{\mathrm{test}}$ values beyond the first LCP iteration. The forward variable selection procedure for both the straight-line and LCP iterations of these models retained only paved roads, along with geographic distance. We used the straight-line model as the final model for the full dataset (440 km). We found that different buffer widths (0, 1, 3 and 5 km) around straight lines and LCPs did not appreciably affect the mean values of extracted covariates (Figure S13). A 1 km buffer had the lowest average $\mathrm{RMSE}_{\mathrm{test}}$ across spatial CV folds in the straight-line models at the 120 and 440 km scales, so we used this buffer size in all subsequent LCP models.

Our GBMs predicted genetic distance best in relatively undisturbed southern areas such as the Kootenay Ranges north-west of Fernie, BC (spatial cluster 8) and BC's Flathead region (cluster 6; Figures S3 and S14). Generally, southern and central areas with roads (clusters 2, 4, 5, 7 and 9) validated better at finer spatial scales and worse at larger spatial scales, while northern areas in Banff National Park (clusters 1 and 3) and BC's Columbia Valley near Golden showed the opposite pattern. Model performance was poorest near the northern portion of BC's Elk Valley (cluster 2) at the four largest spatial scales and in the Columbia Valley at the two finest spatial scales.

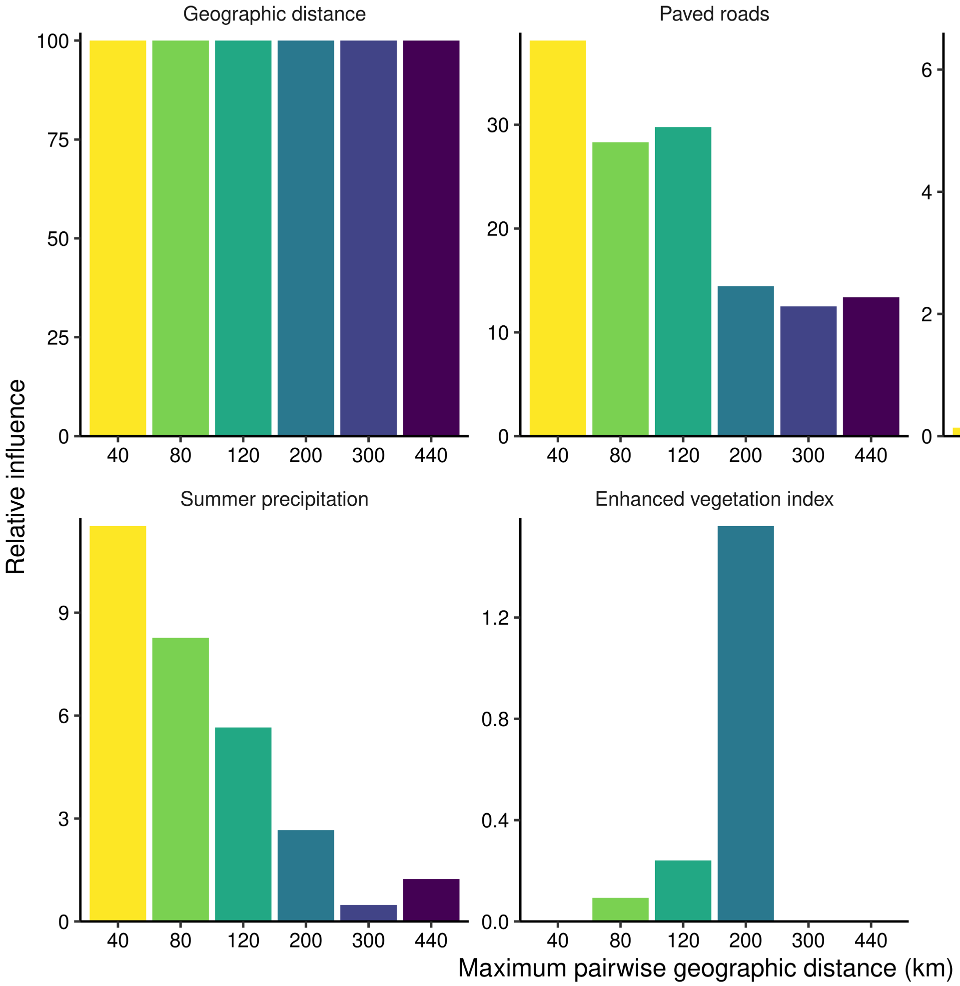

FIGURE 4 Relative influence of variables predicting pairwise genetic distances from GBMs fit using data from 1161 grizzly bears (*Ursus arctos*) in Canada's southern Rocky Mountains, shown by spatial scale. Note that each panel has different y-axis values. Terrain ruggedness and glaciers were not included because they were rarely influential and their relative influence across spatial scales produced no discernible pattern.

In the full dataset, low canopy cover increased genetic differentiation up to ~30% cover, while values above ~38% generally facilitated higher gene flow (Figure 3). Predicted genetic differentiation increased quickly in response to the presence of paved roads, remaining stable as the number of roads intersected by the pairwise corridor increased. Increasing the mean paved roads value along a pairwise corridor from 0 to 1 was roughly equal to adding an additional 10–15 km between animals in terms of genetic connectivity (based on a slope of ~.05 units of predicted genetic distance per 100 km). The Pearson's correlation between predicted and observed genetic distances from our final GBM model ranged from .107 to .140 across all spatial scales (Figure S15).

3.3.1 | Maps of predicted genetic connectivity

Resistant kernels from models with paved roads as the only retained landscape variable (40, 80, 120, 200 and 300 km spatial scales) show highest predicted genetic connectivity throughout the large swath of roadless area (except for a handful of unpaved forestry roads) in the central portion of the study area. Predictions from the 440-km spatial scale whose top model also included canopy cover showed stark declines in genetic connectivity east of the Rocky Mountain front, along the Rocky Mountain Trench from Radium Hot Springs south to the USA border, and north of the TransCanada Highway in Banff National Park (Figure 5). It also highlighted a roughly 50-km-wide pinch point of high connectivity at 49.5° N, near the town of Fernie, BC, between the Rocky Mountain Trench to the west and the prairies to the east.

4 | DISCUSSION

Our large sample size of genotyped bears and broad distribution of genetic samples across a large spatial extent allowed us to test how the relative influence of different landscape variables on patterns of genetic differentiation varies across a gradient of spatial scales. This approach provided a more nuanced understanding of factors

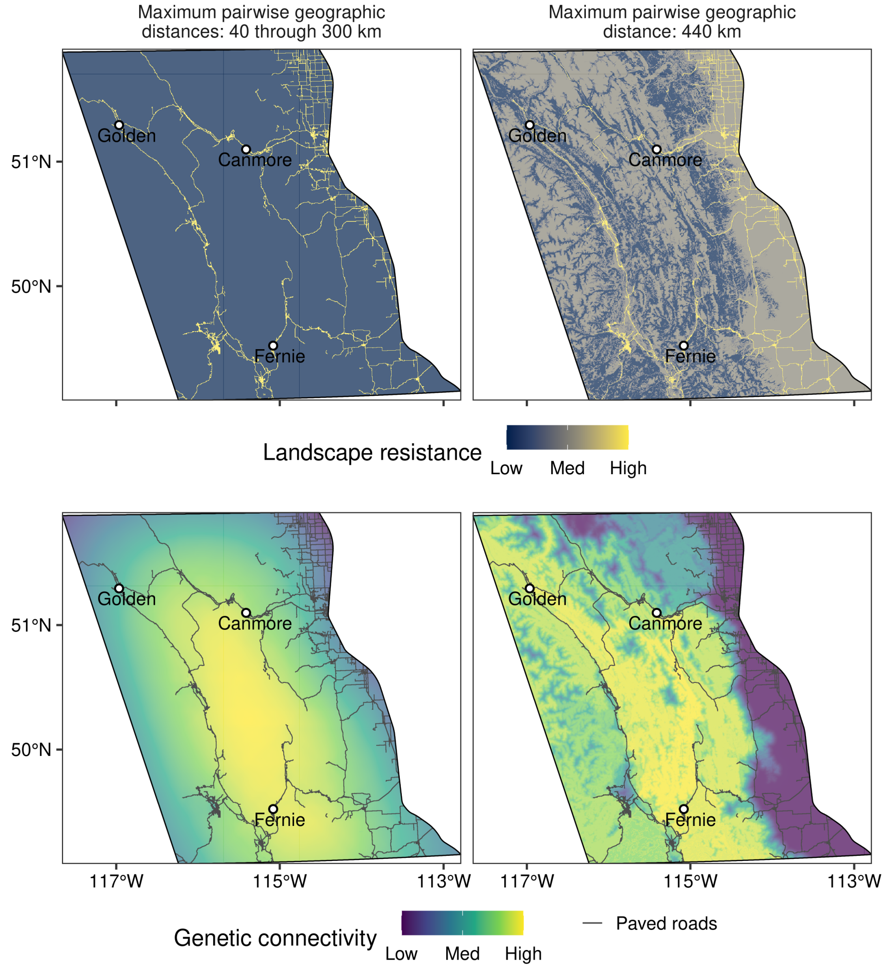

FIGURE 5 Predicted landscape resistance (top) and resistant kernels (bottom) for grizzly bears (*Ursus arctos*) modelled using 1161 genetic samples throughout the southern Canadian Rocky Mountains.

affecting grizzly bear dispersal and gene flow than merely considering a single spatial scale. We found that IBD was the biggest overall driver of genetic differentiation in grizzly bears across all spatial scales among factors we considered, while human disturbance (represented by paved roads) was the most important landscape predictor across all spatial scales. Canopy cover was important at the largest spatial scale. Certain landscape characteristics known to influence grizzly bear habitat selection, such as vegetation greenness, terrain ruggedness, precipitation and glaciers, did not improve predictions of genetic differentiation.

Isolation by distance occurs within a continuously distributed population when movement of genes is spatially restricted (Hardy & Vekemans, 1999) and is an underlying pattern that likely challenged our ability to detect strong relationships between landscape variables and genetic differentiation (Oyler-McCance et al., 2022). Our results were consistent with other landscape genetics studies that showed only incremental improvement in the amount of variation explained by final models when compared to a null expectation of isolation by distance (e.g. Henson et al., 2021; Zeller et al., 2018). Relatively weak effects of landscape variables compared to geographic distance in grizzly bears may be due to many factors, including that grizzly bears are long-lived and wide-ranging habitat generalists or that our analyses may reflect historical signatures of genetic connectivity (Epps & Keyghobadi, 2015). Stronger relationships between certain landscape variables and genetic differentiation may have existed in certain parts of our study area, but our use of spatial cross-validation and variable selection indicated they did not reflect patterns throughout most of the spatial extent. In addition, it is possible that the spatial distribution of our samples (e.g. we did not have genetic sample locations west of the Rocky Mountain Trench) or only using samples genotyped at nine loci limited our power to detect certain patterns, that large population sizes on opposite sides of highway fractures reduced potential effects of genetic drift or that very small number of translocated bears included in the dataset weakened the effects certain landscape characteristics on connectivity. Past research found strong evidence for IBD in grizzly bears, primarily in relatively undisturbed northern areas of their range, although these studies were mostly population based and encompassed a larger spatial extent (Paetkau et al., 1997, 1998; Proctor et al., 2012). For example, Proctor et al. (2012) found that the relative influence of IBD on genetic differentiation was greater in northern BC than southern BC, and that in the southern portion of grizzly bear distribution, the degree of population fragmentation depended on the amount of habitat fragmentation from human disturbance and, secondarily, the presence of glaciers across the continental divide. In south-east BC, south-west AB, northern Idaho and north-west Montana, they found strong evidence of fragmentation from human disturbance and transportation structures and a relatively small, yet statistically significant effect of IBD. The lack of evidence indicating that glaciers increased genetic differentiation in grizzly bears in the southern Canadian Rocky Mountains could reflect that glaciers in this region were smaller than those at the north end of Proctor et al.'s study area.

At finer spatial scales, the relatively strong effect of geographic distance we observed likely reflected movements and mating patterns within and between adjacent home ranges rather than a physical inability to disperse. Grizzly bears have overlapping home ranges (Mace and Waller, 1997) and are promiscuous (Craighead et al., 1995). Dispersal from natal ranges in both sexes often consists of home range shifts and expansions over months to years but can also occur in a single discrete event (McLellan & Hovey, 2001; Proctor et al., 2004). Although females are typically philopatric, male dispersal may be far enough to reduce the probability of inbreeding (McLellan & Hovey, 2001). Given these patterns of space use and dispersal behaviour, gene flow across the entire study area likely could only occur across multiple generations.

Outside of certain forms of human disturbance, there may be few landscape characteristics that inhibit grizzly bear gene flow across most spatial scales in our study area. Human disturbance was the only important landscape predictor of genetic differentiation across all spatial scales except at the broadest scale when canopy cover was also retained in the variable selection procedure. The presence of roads, especially high-traffic roads, has been shown to alter their habitat selection, impede their movements and increase their mortality (Ciarniello et al., 2007; Nielsen et al., 2004; Northrup et al., 2012). Our results indicate that these effects persist out to the broadest spatial scale in our study area and therefore limit landscape connectivity across spatial scales and point to the benefits of mitigating road barriers to improve connectivity at all scales.

Human development may inhibit grizzly bear movement and gene flow through some combination of behavioural avoidance and increased mortality (Lamb et al., 2020; Trombulak & Frissell, 2000). Humans are the main source of grizzly bear mortality (>75%) in the southern portion of their range (McLellan et al., 1999). Grizzly bears tailor their habitat use in space and time to avoid the risks posed by humans (e.g. Gibeau et al., 2002, Roever et al., 2010, Proctor et al., 2019; but see Munro et al., 2006), shaping the landscape of connectivity and survival for this species. The effects of human disturbance on grizzly bear behaviour, movement and survival culminate to modify their genetic connectivity in southern Canada.

When we considered the entire dataset, which included pairwise distances out to 440 km, areas with low canopy cover inhibited gene flow. This pattern was also evident at the 200-km and 300-km spatial scales, but canopy cover was not retained during variable selection in those models. Grizzly bears dispersing relatively short distances may cross larger expanses of open areas such as valley bottoms and agricultural areas by moving quickly to minimize risk, travelling primarily at night (Lamb et al., 2020) or moving through any existing forested corridors. However, as the spatial (and temporal) scale increases and in the rare instances of longer, more directed dispersal, we postulate that the cumulative risk associated with crossing more and more of these open areas may inhibit their connectivity. At this broadest spatial scale, the canopy cover variable may have best captured relatively poor grizzly bear habitats, such as human settlements, agricultural areas, grasslands and prairies, that collectively inhibit long-distance dispersal. High canopy cover areas within our study extent offer cover to help minimize mortality risk yet are often near-edge habitats that provide access to foraging resources.

Our map of predicted genetic connectivity at the largest spatial scale highlighted an important north–south movement corridor near Fernie, BC, where the heavily forested and relatively undisturbed portion of the Rocky Mountains is narrowest. Dispersing bears in this area may be funnelled between the human disturbances in the area (open pit coal mines, cities and expanding rural settlements) and the open canopies and human settlement of the Kootenay River Valley to the west and the foothills and prairies on the eastern slopes of mountains in AB. These results underscore the importance of BC's Elk Valley as a critical corridor, and pinch point, along the Rocky Mountains. They also provide additional evidence supporting ongoing work to create crossing structures along a nearby east–west stretch of Highway 3 between Sparwood, BC, and Lundbreck, AB (Brennan et al., 2022).

Because dispersal directly promotes population connectivity, modelling connectivity using movement data from individual animals is attractive. However, this type of dispersal data typically suffers from very low sample sizes due to the logistical difficulties of tracking animals during this infrequent behaviour (Fagan & Calabrese, 2006). Genetic data can serve as a landscape-scale proxy for dispersal data and highlight connectivity corridors at relatively broad spatiotemporal scales, but they are less appropriate for pinpointing fine-scale linkage areas such as highway crossings than maps derived from detailed tracking data, which for many species (such as grizzly bears) primarily consists of movements within home ranges. In addition, genetic data can be collected non-invasively and across massive spatial extents (Kendall et al., 2009; Lamb et al., 2019; Morehouse & Boyce, 2016). Differences between maps from genetic-based and movement-based models can highlight areas where dispersal does not result in successful breeding due to mortality. Our maps indicate that major roads do increase genetic differentiation but are not complete barriers to gene flow (Proctor et al., 2012), a result consistent with previous work in Banff National Park showing that grizzly bear gene flow does occur across the Trans-Canada Highway with the help of wildlife crossing structures (Sawaya et al., 2014). Past studies confirm that collared grizzly bears do cross highways in our study area (Graham et al., 2010; Proctor et al., 2012), although crossing frequency depends on time of day and vehicle traffic volume (Chruszcz et al., 2003; Waller & Servheen, 2005). It is well documented that grizzly bear mortality is relatively high near people due to conflicts at residences and collisions along highways and railways (e.g. Nielsen et al., 2004; Proctor et al., 2019). Genetic connectivity may thus be impeded due to the combined effect of bears behaviourally avoiding human-dominated areas and dispersing bears dying more in these areas (Lamb et al., 2017). However, management activities that reduce human–bear conflicts may increase connectivity across areas that would otherwise function as fractures between populations (Proctor et al., 2018).

Our modelling approach, which builds upon methods used in several previous studies (Bouyer et al., 2015; Pless et al., 2021; van Strien et al., 2012), can be used to model genetic connectivity in other continuously distributed species and offers several advantages over existing landscape genetics frameworks. First, modelling genetic distance directly from raw landscape covariates eliminated subjectivity in creating an initial resistance surface for calculation of effective distances. Second, our use of spatial CV and variable selection helped account for spatial autocorrelation and strengthened our inference by minimizing overfitting and removing variables that failed to improve model performance throughout the study area (Ploton et al., 2020). Using $\mathrm{RMSE}_{\mathrm{test}}$ values alone, GBM models fit with random CV appeared to outperform those fit using spatial CV (Table S1). However, ALE predictions from these models (Figure S9) highlighted noise and random fluctuations in the training data, illustrating a common problem with machine learning algorithms that use random CV folds with spatial data (Meyer et al., 2019). The similarity in predictor variable space between training and test datasets within random CV folds likely resulted in final models that were better able to reproduce the exact sampling data than ones fit using spatial CV but were not generalizable to the entire study area. Not surprisingly, the spatial CV models easily outperformed the top random CV models at predicting genetic distances involving samples from every region when we completely withheld data from that region during model training. Spatial CV should therefore improve any future predictions in regions where we lacked genetic samples, such as the Purcell Mountains west of the Rocky Mountain trench. Spatial CV also ensured that all pairwise observations involving a genetic location from a withheld spatial cluster appeared in the test dataset only (i.e. not in the training dataset) within each CV fold. Although this approach did not explicitly account for the non-independence in our data, preventing pairwise observations from crossing this ‘training-test divide’ helps avoid inflated performance metrics that are common in machine learning models fit data with strong dependence (Whalen et al., 2022)

Third, machine learning algorithms allowed us to quantify non-linear responses to landscape characteristics without using predefined transformations (e.g. log and quadratic). For example, response curves depicting the effects of canopy cover and human disturbance on grizzly bear genetic differentiation included sharp changes and thresholds that linear modelling frameworks could not accurately characterize. Machine learning algorithms have demonstrated utility in large-scale ecological mapping analyses such as species distribution models (Elith and Leathwick 2009) and resource selection analyses (Shoemaker et al., 2018), yet they have seen limited use in the field of landscape genetics (but see Murphy et al. 2010b, Bishop et al., 2021, Pless et al., 2021, Kittlein et al. 2022).

Ensuring connectivity exists for large carnivores is a major conservation challenge because these species inhabit large spatial extents at relatively low densities, exhibit long-distance dispersal events, are sensitive to human disturbance and pose real and perceived threats to humans (Bartoń et al., 2019; Lamb et al., 2020). Future research comparing connectivity maps produced using genetic data with those using movement and demographic data will further clarify and prioritize management targets for grizzly bears in the region. Our work highlights the utility of cross-jurisdictional collaboration to gain insight into factors driving genetic connectivity for continuously distributed populations in ecologically heterogeneous and human-dominated landscapes. It also highlights the importance of incorporating spatial CV in landscape genetics analyses across all taxa, especially when making predictions over large spatial extents. Finally, we gained additional ecological insights by considering landscape resistance and connectivity at more than one spatial scale and believe there is exciting potential for incorporating multiple spatial scales into future landscape genetics studies.

AUTHOR CONTRIBUTIONS

ELL, ECP and KAZ conceived the study. ECP wrote the manuscript with help from ELL, CTL and KAZ. ECP led the data analysis with help from ELL, ZAH, CCD and KAZ. ATM, CTL, GM, MAS and JW collected genetic data. All authors provided valuable feedback and contributed to the final manuscript.

ACKNOWLEDGEMENTS

We acknowledge support from the Government of British Columbia, Habitat Conservation Trust Foundation, Teck Coal, Vanier Canada Graduate Scholarship, Fish and Wildlife Compensation Program, Forest Enhancement Society of British Columbia, Columbia Basin Trust, Counter Assault, Safari Club International, Sparwood and District Fish and Wildlife Association, BP Canada, and BC Conservation Corps. We thank Parks Canada and the many partners of the fRI Grizzly Bear Project who provided funding over many years. We thank the South Rockies Grizzly Bear Project volunteers for supporting our research project and helping us out in the field. We thank Barb Johnston of the south-west Alberta grizzly bear monitoring project for logistical and financial support. We acknowledge support from the Wilburforce Foundation, Western Transportation Institute, Woodcock Foundation, National Fish and Wildlife Foundation, Alberta Conservation Association, Calgary Foundation and Mountain Equipment Co-operative. We thank Laura Smit, John Paczkowski and Karen Graham for providing data.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

FUNDING INFORMATION

This research was supported in part by the USDA Forest Service, Rocky Mountain Research Station and Aldo Leopold Wilderness Research Institute. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or US Government determination or policy. Computational resources from the University of Montana's Computational Ecology Lab and Griz Shared Computing Cluster contributed to this research (NSF award numbers 2018112 & 1925267).

DATA AVAILABILITY STATEMENT

The following data and code are publicly available at https://doi.org/10.5281/zenodo.8193974: individual genotypes, environmental raster data and R code to create pairwise datasets, extract covariates along straight lines and least-cost paths between example spatial locations, run straight-line and least-cost path GBM models with variable selection and cross-validation, create spatial predictions of landscape resistance from models and code to prepare raster files for resistant kernel connectivity analyses in UNICOR. Spatial coordinates of grizzly bear genetic samples were withheld due to the sensitive nature of this information.

BENEFIT-SHARING STATEMENT

We developed a research collaboration with the biologists from the governments and organizations providing genetic samples, and we included these biologists as co-authors. We have shared the results of this research with the data owners and the broader scientific community (see above), and the research addresses a priority concern, the conservation of grizzly bears. Our group is committed to scientific partnerships across provincial, state and international boundaries, as well as to institutional capacity building.

ORCID

Eric C. Palm https://orcid.org/0000-0002-5330-4804
Erin L. Landguth https://orcid.org/0000-0002-7044-346X
Casey C. Day https://orcid.org/0000-0001-6096-3448
Clayton T. Lamb https://orcid.org/0000-0002-1961-0509
Andrea T. Morehouse https://orcid.org/0000-0002-2015-9938
Michael F. Proctor https://orcid.org/0000-0002-9461-2774
Michael A. Sawaya https://orcid.org/0000-0002-3189-1737
Jesse Whittington https://orcid.org/0000-0002-4129-7491
Katherine A. Zeller https://orcid.org/0000-0002-2913-6660

REFERENCES

Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *82*, 1059–1086.

Balkenhol, N., Schwartz, M. K., Inman, R. M., Copeland, J. P., Squires, J. S., Anderson, N. J., & Waits, L. P. (2020). Landscape genetics of wolverines (*Gulo gulo*): Scale-dependent effects of bioclimatic, topographic, and anthropogenic variables. *Journal of Mammalogy*, *101*, 790–803.

Bartoń, K. A., Zwijacz-Kozica, T., Zięba, F., Sergiel, A., & Selva, N. (2019). Bears without borders: Long-distance movement in human-dominated landscapes. *Global Ecology and Conservation*, *17*, e00541.

Baston, D. (2021). *Exactextractr: Fast extraction from raster datasets using polygons*. R Package version 0.7.2. https://cran.r-project.org/package=exactextractr

Bishop, A. P., Amatulli, G., Hyseni, C., Pless, E., Bateta, R., Okeyo, W. A., Mireji, P. O., Okoth, S., Malele, I., Murilla, G., Aksoy, S., Caccone, A., & Saarman, N. P. (2021). A machine learning approach to integrating genetic and ecological data in tsetse flies (*Glossina pallidipes*) for spatially explicit vector control planning. *Evolutionary Applications*, *14*, 1762–1777.

Bouyer, J., Dicko, A. H., Cecchi, G., Ravel, S., Guerrini, L., Solano, P., Vreysen, M. J. B., de Meeûs, T., & Lancelot, R. (2015). Mapping landscape friction to locate isolated tsetse populations that are candidates for elimination. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 14575–14580.

Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R., & Cavalli-Sforza, L. L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, *368*, 455–457.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification and regression trees* (1st ed.). Routledge.

Brennan, L., Chow, E., & Lamb, C. (2022). Wildlife overpass structure size, distribution, effectiveness, and adherence to expert design recommendations. *PeerJ*, *10*, e14371.

Carnahan, A. M., Van Manen, F. T., Haroldson, M. A., Stenhouse, G. B., & Robbins, C. T. (2021). Quantifying energetic costs and defining energy landscapes experienced by grizzly bears. *Journal of Experimental Biology*, *224*, jeb241083.

Chruszcz, B., Clevenger, A. P., Gunson, K. E., & Gibeau, M. L. (2003). Relationships among grizzly bears, highways, and habitat in the Banff-Bow Valley, Alberta, Canada. *Canadian Journal of Zoology*, *81*, 1378–1391.

Ciarniello, L. M., Boyce, M. S., Seip, D. R., & Heard, D. C. (2007). Grizzly bear habitat selection is scale dependent. *Ecological Applications*, *17*, 1424–1440.

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.

Craighead, J. J., Sumner, J. S., & Mitchell, J. A. (1995). *The grizzly bears of Yellowstone: Their ecology in the Yellowstone ecosystem* (pp. 1959–1992). Island Press.

Cushman, S. A., Lewis, J. S., & Landguth, E. L. (2013). Evaluating the intersection of a regional wildlife connectivity network with highways. *Movement Ecology*, *1*, 1–11.

Dyer, R. (2012). *The gstudio package*. Virginia Commonwealth University.

Epps, C. W., & Keyghobadi, N. (2015). Landscape genetics in a changing world: Disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, *24*, 6021–6040.

Ernest, H. B., Vickers, T. W., Morrison, S. A., Buchalski, M. R., & Boyce, W. M. (2014). Fractured genetic connectivity threatens a Southern California puma (*Puma concolor*) population. *PLoS One*, *9*(10), e107985.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*, 479–491.

Fagan, W. F., & Calabrese, J. M. (2006). Quantifying connectivity: Balancing metric performance with data requirements. In K. R. Crooks & M. Sanjayan (Eds.), *Connectivity conservation* (pp. 297–317). Cambridge University Press.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography Mission. *Reviews of Geophysics*, *45*, RG2004.

Fitzpatrick, S. W., & Reid, B. N. (2019). Does gene flow aggravate or alleviate maladaptation to environmental stress in small populations? *Evolutionary Applications*, *12*, 1402–1416.

Franckowiak, R. P., Panasci, M., Jarvis, K. J., Acuña-Rodriguez, I. S., Landguth, E. L., Fortin, M. J., & Wagner, H. H. (2017). Model selection with multiple regression on distance matrices leads to incorrect inferences. *PLoS One*, *12*, e0175194.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*, 1189–1232. https://doi.org/10.1214/aos/1013203451

Gibeau, M. L., Clevenger, A. P., Herrero, S., & Wierzchowski, J. (2002). Grizzly bear response to human development and activities in the Bow River watershed, Alberta, Canada. *Biological Conservation*, *103*, 227–236.

Government of British Columbia. (2017). *GeoBC Atlas Integrated Transportation Network (dgtl_road_atlas.gdb) Public Delivery Packaged Product Attribute Data Dictionary. GeoBC Data Catalogue*. Retrieved from https://catalogue.data.gov.bc.ca/dataset/digital-road-atlas-dra-masterpartially-attributed-roads

Government of Canada. (2022). National road network – Alberta. Retrieved from https://geo.statcan.gc.ca/nrn_rrn/ab/nrn_rrn_ab_SHAPE.zip

Graham, K., Boulanger, J., Duval, J., & Stenhouse, G. (2010). Spatial and temporal use of roads by grizzly bears in west-central Alberta. *Ursus*, *21*, 43–56.

Graham, K., & Stenhouse, G. B. (2014). Home range, movements, and denning chronology of the grizzly bear (*Ursus arctos*) in west-central Alberta. *The Canadian Field-Naturalist*, *128*, 223–234.

Greenwell, B., Boehmke, B., Cunningham, J., & GBM Developers. (2020). *gbm: Generalized boosted regression models*. R Package version 2.1.8. https://cran.r-project.org/package=gbm

Hardy, O. J., & Vekemans, X. (1999). Isolation by distance in a continuous population: Reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, *83*, 145–154.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. Journal of the Royal Statistical Society. *Series C (Applied Statistics)*, *28*, 100–108.

Henson, L. H., Balkenhol, N., Gustas, R., Adams, M., Walkus, J., Housty, W. G., Stronen, A. V., Moody, J., C. Service, Reece, D., von Holdt, B. M., McKechnie, I., Koop, B. F., & Darimont, C. T. (2021). Convergent geographic patterns between grizzly bear population genetic structure and indigenous language groups in coastal British Columbia, Canada. *Ecology and Society*, *26*, 7.

Herrero, S. (Ed.). (2005). *Biology, demography, ecology and management of grizzly bears in and around Banff National Park and Kananaskis country: The final report of the eastern slopes grizzly bear project*. Faculty of Environmental Design. University of Calgary.

Hirsh-Pearson, K., Johnson, C. J., Schuster, R., Wheate, R. D., & Venter, O. (2022). Canada's human footprint reveals large intact areas juxtaposed against areas under immense anthropogenic pressure. *Facets*, *7*, 398–419.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213.

Jaffé, R., Veiga, J. C., Pope, N. S., Lanes, É. C. M., Carvalho, C. S., Alves, R., Andrade, S. C. S., Arias, M. C., Bonatti, V., Carvalho, A. T., de Castro, M. S., Contrera, F. A. L., Francoy, T. M., Freitas, B. M., Giannini, T. C., Hrncir, M., Martins, C. F., Oliveira, G., Saraiva, A. M., … Imperatriz-Fonseca, V. L. (2019). Landscape genomics to the rescue of a tropical bee threatened by habitat loss and climate change. *Evolutionary Applications*, *12*, 1164–1177.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2014*, 1–14.

Kendall, K. C., Stetz, J. B., Boulanger, J., Macleod, A. C., Paetkau, D., & White, G. C. (2009). Demography and genetic structure of a recovering grizzly bear population. *Journal of Wildlife Management*, *73*, 3–17.

Keyghobadi, N. (2007). The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, *85*, 1049–1064.

Kierepka, E. M., & Latch, E. K. (2015). Performance of partial statistics in individual-based landscape genetics. *Molecular Ecology Resources*, *15*, 512–525.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*, 1–26.

Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles*. https://www.tidymodels.org

Lamb, C. T., Ford, A. T., McLellan, B. N., Proctor, M. F., Mowat, G., Ciarniello, L., Nielsen, S. E., & Boutin, S. (2020). The ecology of human-carnivore coexistence. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(30), 17876–17883.

Lamb, C. T., Ford, A. T., Proctor, M. F., Royle, J. A., Mowat, G., & Boutin, S. (2019). Genetic tagging in the Anthropocene: Scaling ecology from alleles to ecosystems. *Ecological Applications*, *29*, 1–17.

Lamb, C. T., Mowat, G., McLellan, B. N., Nielsen, S. E., & Boutin, S. (2017). Forbidden fruit: Human settlement and abundant fruit create an ecological trap for an apex omnivore. *Journal of Animal Ecology*, *86*, 55–65.

Landguth, E. L., Cushman, S. A., Murphy, M. A., & Luikart, G. (2010). Relationships between migration rates and landscape resistance assessed using individual-based simulations. *Molecular Ecology Resources*, *10*, 854–862.

Landguth, E. L., Hand, B. K., Glassy, J., Cushman, S. A., & Sawaya, M. A. (2012). UNICOR: A species connectivity and corridor network simulator. *Ecography*, *35*, 9–14.

Legendre, P. (1993). Spatial autocorrelation: Trouble or new paradigm? *Ecology*, *74*, 1659–1673.

Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, *19*, 3038–3051.

Mace and Waller. (1997). Spatial and temporal interaction of male and female grizzly bears in northwestern Montana. *The Journal of Wildlife Management*, *61*, 39–52.

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology and Evolution*, *28*, 614–621.

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.

McClelland, C. J. R., Coops, N. C., Kearney, S. P., Burton, A. C., Nielsen, S. E., & Stenhouse, G. B. (2020). Variations in grizzly bear habitat selection in relation to the daily and seasonal availability of annual plant-food resources. *Ecological Informatics*, *58*, 101116.

McLellan, B. N., & Hovey, F. W. (2001). Natal dispersal of grizzly bears. *Canadian Journal of Zoology*, *79*, 838–844.

McLellan, B. N., Hovey, F. W., Mace, R. D., Woods, J. G., Carney, D. W., Gibeau, M. L., Wakkinen, W. L., & Kasworm, W. F. (1999). Rates and causes of grizzly bear mortality in the Interior Mountains of British Columbia, Alberta, Montana, Washington, and Idaho. *The Journal of Wildlife Management*, *63*, 911–920.

McRae, B. H. (2006). Isolation by resistance. *Evolution*, *60*, 1551–1561.

Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, *21*, 2839–2846.

Meyer, H., Mila, C., Ludwig, M., & Linnenbrink, J. (2023). *CAST:"caret" Applications for spatial-temporal models*. R package version 0.5.1. https://CRAN.R-project.org/package=CAST

Meyer, H., Reudenbach, C., Wöllauer, S., & Nauss, T. (2019). Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. *Ecological Modelling*, *411*, 108815.

Moilanen, A., & Hanski, I. (2001). On the use of connectivity measures in spatial. *Ecology*, *95*, 147–151.

Morehouse, A. T., & Boyce, M. S. (2016). Grizzly bears without borders: Spatially explicit capture–recapture in southwestern Alberta. *Journal of Wildlife Management*, *80*, 1152–1166.

Mowat, G., Heard, D. C., & Schwarz, C. J. (2013). Predicting grizzly bear density in western North America. *PLoS One*, *8*, e82757.

Munro, R. H. M., Nielsen, S. E., Price, M. H., Stenhouse, G. B., & Boyce, M. S. (2006). Seasonal and diel patterns of grizzly bear diet and activity in west-central Alberta. *Journal of Mammology*, *87*, 1112–1121.

Nielsen, S. E., Cranston, J., & Stenhouse, G. B. (2009). Identification of priority areas for grizzly bear conservation and recovery in Alberta, Canada. *Journal of Conservation Planning*, *5*, 38–60.

Nielsen, S. E., Herrero, S., Boyce, M. S., MacE, R. D., Benn, B., Gibeau, M. L., & Jevons, S. (2004). Modelling the spatial distribution of human-caused grizzly bear mortalities in the central Rockies ecosystem of Canada. *Biological Conservation*, *120*, 101–113.

Northrup, J. M., Pitt, J., Muhly, T. B., Stenhouse, G. B., Musiani, M., & Boyce, M. S. (2012). Vehicle traffic shapes grizzly bear behaviour on a multiple-use landscape. *Journal of Applied Ecology*, *49*, 1159–1167.

Oyler-McCance, S. J., Cross, T. B., Row, J. R., Schwartz, M. K., Naugle, D. E., Fike, J. A., Winiarski, K., & Fedy, B. C. (2022). New strategies for characterizing genetic structure in wide-ranging, continuously distributed species: A greater Sage-grouse case study. *PLoS One*, *17*, e0274189.

Paetkau, D. (2003). An empirical exploration of data quality in DNA-based population inventories. *Molecular Ecology*, *12*, 1375–1387.

Paetkau, D., Waits, L. P., Clarkson, P. L., Craighead, L., & Strobe, C. (1997). An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*, *147*, 1943–1957.

Paetkau, D., Waits, L. P., Clarkson, P. L., Craighead, L., Vyse, E., Ward, R., & Strobeck, C. (1998). Variation in genetic diversity across the range of north American Brown bears. *Conservation Biology*, *12*, 418–429.

Paradis, E. (2010). pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*, 419–420.

Peterman, W. E., & Pope, N. S. (2021). The use and misuse of regression models in landscape genetic analyses. *Molecular Ecology*, *30*, 37–47.

Pless, E., Saarman, N. P., Powell, J. R., Caccone, A., & Amatulli, G. (2021). A machine-learning approach to map landscape connectivity in *Aedes aegypti* with genetic and environmental data. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, 1–8.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., & Pélissier, R. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, *11*, 1–11.

Proctor, M., Mclellan, B., Boulanger, J., Apps, C., Stenhouse, G., Paetkau, D., & Mowat, G. (2010). Ecological investigations of grizzly bears in Canada using DNA from hair, 1995-2005: A review of methods and progress. *Ursus*, *21*, 169–188.

Proctor, M. F., Kasworm, W. F., Annis, K. M., Teisberg, J. E., Radandt, T. G., Servheen, C., & Franke, W. A. (2018). Conservation of threatened Canada-USA trans-border grizzly bears linked to comprehensive conflict reduction. *Human-Wildlife Interactions*, *12*, 348–372.

Proctor, M. F., McLellan, B. N., Stenhouse, G. B., Mowat, G., Lamb, C. T., & Boyce, M. S. (2019). Effects of roads and motorized human access on grizzly bear populations in British Columbia and Alberta, Canada. *Ursus*, *2019*, 16–39.

Proctor, M. F., McLellan, B. N., Strobeck, C., & Barclay, R. M. R. (2004). Gender-specific dispersal distances of grizzly bears estimated by genetic analysis. *Canadian Journal of Zoology*, *82*, 1108–1118.

Proctor, M. F., Mclellan, B. N., Strobeck, C., & Barclay, R. M. R. (2005). Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerably small populations. *Proceedings of the Royal Society B*, *272*, 2409–2416.

Proctor, M. F., Nielsen, S. E., Kasworm, W. F., Servheen, C., Radandt, T. G., Machutchon, A. G., & Boyce, M. S. (2015). Grizzly bear connectivity mapping in the Canada-United States trans-border region. *The Journal of Wildlife Management*, *79*, 544–558.

Proctor, M. F., Paetkau, D., McLellan, B. N., Stenhouse, G. B., Kendall, K. C., MacE, R. D., Kasworm, W. F., Servheen, C., Lausen, C. L., Gibeau, M. L., Wakkinen, W. L., Haroldson, M. A., Mowat, G., Apps, C. D., Ciarniello, L. M., Barclay, R. M. R., Boyce, M. S., Schwartz, C. C., & Strobeck, C. (2012). Population fragmentation and inter-ecosystem movements of grizzly bears in Western Canada and the northern United States. *Wildlife Monographs*, *180*, 1–46.

Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Science*, *5*, 23–27.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*, 913–929.

Roever, C. L., Boyce, M. S., & Stenhouse, G. B. (2010). Grizzly bear movements relative to roads: Application of step selection functions. *Ecography*, *33*, 1113–1122.

Roffler, G. H., Schwartz, M. K., Pilgrim, K. L., Talbot, S. L., Sage, G. K., Adams, L. G., & Luikhart, G. (2016). Identification of landscape features influencing gene flow: How useful are habitat selection models? *Evolutionary Applications*, *9*(805), 817.

Row, J. R., Knick, S. T., Oyler-McCance, S. J., Lougheed, S. C., & Fedy, B. C. (2017). Developing approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. *Ecology and Evolution*, *7*, 3751–3761.

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., & Scambos, T. A. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172.

Sawaya, M. A., Kalinowski, S. T., & Clevenger, A. P. (2014). Genetic connectivity for two bear species at wildlife crossing structures in Banff National Park. *Proceedings of the Royal Society B*, *281*, 20131705.

Sawaya, M. A., Stetz, J. B., Clevenger, A. P., Gibeau, M. L., & Kalinowski, S. T. (2012). Estimating grizzly and black bear population abundance and trend in Banff National Park using noninvasive genetic sampling. *PLoS One*, *7*, e34777.

Shirk, A. J., & Cushman, S. A. (2011). sGD: Software for estimating spatially explicit indices of genetic diversity. *Molecular Ecology Resources*, *11*, 922–934. https://doi.org/10.1111/j.1755-0998.2011.03035.x

Shirk, A. J., Landguth, E. L., & Cushman, S. A. (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Molecular Ecology Resources*, *17*, 1308–1317.

Shoemaker, K. T., Heffelfinger, L. J., Jackson, N. J., Blum, M. E., Wasley, T., & Stewart, K. M. (2018). A machine-learning approach for extending classical wildlife resource selection analyses. *Ecology and Evolution*, *8*, 3556–3569.

Slatkin, M. (2008). Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, *9*, 477–485.

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.

Stourvas, N. (2022). PopGenUtils. https://github.com/nikostourvas/PopGenUtils

Taylor, P. D., Fahrig, L., Henein, K., & Merriam, G. (1993). Connectivity is a vital element of landscape structure. *Oikos*, *68*, 571–573.

Thatte, P., Chandramouli, A., Tyagi, A., Patel, K., Baro, P., Chhattani, H., & Ramakrishnan, U. (2020). Human footprint differentially impacts genetic connectivity of four wide-ranging mammals in a fragmented landscape. *Diversity and Distributions*, *26*, 299–314.

Townshend, J. R. G. (2016). Global Forest cover change (GFCC) tree cover multi-year global 30 m V003. *NASA EOSDIS Land Processes DAAC*, 176–184. https://doi.org/10.5067/MEaSUREs/GFCC/GFCC30TC.003

Trombulak, S. C., & Frissell, C. A. (2000). Review of ecological effects of roads on terrestrial and aquatic communities. *Conservation Biology*, *14*, 18–30.

Unnithan Kumar, S., & Cushman, S. A. (2022). Connectivity modelling in conservation science: A comparative evaluation. *Scientific Reports*, *12*, 16680.

van Strien, M. J., Keller, D., & Holderegger, R. (2012). A new analytical approach to landscape genetic modelling: Least-cost transect analysis and linear mixed models. *Molecular Ecology*, *21*, 4010–4023.

Vanhove, M., & Launey, S. (2023). Estimating resistance surface on multilayered landscapes using gradient forest and allelic frequencies. *Molecular Ecology Resources*, 1–15.

Waller, J. S., & Servheen, C. (2005). Effects of transportation infrastructure on grizzly bears in northwestern Montana. *Journal of Wildlife Management*, *69*, 985–1000.

Whalen, S., Schreiber, J., Noble, W. S., & Pollard, K. S. (2022). Navigating the pitfalls of applying machine learning in genomics. *Nature Reviews Genetics*, *23*, 169–181.

Whitmee, S., & Orme, C. D. L. (2013). Predicting dispersal distance in mammals: A trait-based approach. *Journal of Animal Ecology*, *82*, 211–221.

Whittington, J., Hebblewhite, M., & Chandler, R. B. (2018). Generalized spatial mark–resight models with an application to grizzly bears. *Journal of Applied Ecology*, *55*, 157–168.

Wright, M. N., & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, *77*, 1–17.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114–138.

Zeller, K. A., Jennings, M. K., Vickers, T. W., Ernest, H. B., Cushman, S. A., & Boyce, W. M. (2018). Are all data types and connectivity models created equal? Validating common connectivity approaches with dispersal data. *Diversity and Distributions*, *24*, 868–879.

Zeller, K. A., McGarigal, K., & Whiteley, A. R. (2012). Estimating landscape resistance to movement: A review. *Landscape Ecology*, *27*, 777–797.

Zeller, K. A., Vickers, T. W., Ernest, H. B., & Boyce, W. M. (2017). Multi-level, multi-scale resource selection functions and resistance surfaces for conservation planning: Pumas as a case study. *PLoS One*, *12*, 1–20.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Palm, E. C., Landguth, E. L., Holden, Z. A., Day, C. C., Lamb, C. T., Frame, P. F., Morehouse, A. T., Mowat, G., Proctor, M. F., Sawaya, M. A., Stenhouse, G., Whittington, J., & Zeller, K. A. (2023). Corridor-based approach with spatial cross-validation reveals scale-dependent effects of geographic distance, human footprint and canopy cover on grizzly bear genetic connectivity. *Molecular Ecology*, *00*, 1–17. https://doi.org/10.1111/mec.17098