Received: 15 July 2021 | Accepted: 23 June 2022

DOI: 10.1111/2041-210X.13969

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

A statistical framework for modelling migration corridors

Tristan A. Nuñez[1,2] | Mark A. Hurley[3] | Tabitha A. Graves[4] | Anna C. Ortega[1,5] | Hall Sawyer[6] | Julien Fattebert[1,7] | Jerod A. Merkle[8] | Matthew J. Kauffman[9]

[1]Wyoming Cooperative Fish and Wildlife Research Unit, Department of Zoology and Physiology, University of Wyoming, Laramie, Wyoming, USA; [2]Biology Department, University of Washington, Seattle, WA, USA; [3]Idaho Department of Fish and Game, Boise, Idaho, USA; [4]U.S. Geological Survey, Northern Rocky Mountain Science Center, West Glacier, Montana, USA; [5]Program in Ecology, University of Wyoming, Laramie, Wyoming, USA; [6]Western EcoSystems Technology (WEST), Inc., Laramie, Wyoming, USA; [7]School of Life Sciences, University of KwaZulu-Natal, Durban, South Africa; [8]Department of Zoology and Physiology, University of Wyoming, Laramie, Wyoming, USA and [9]U.S. Geological Survey, Wyoming Cooperative Fish and Wildlife Research Unit, Zoology and Physiology Department, University of Wyoming, Laramie, Wyoming, USA

Correspondence
Tristan A. Nuñez
Email: tnunez@uw.edu

Funding information
Idaho Department of Fish and Game; U.S. Fish and Wildlife Service, Grant/Award Number: W-160-R-37; U.S. Geological Survey

Handling Editor: Chris Sutherland

Abstract

1. Management of animal populations requires spatially explicit knowledge of movement corridors, such as those used during seasonal migrations. Global Positioning System (GPS) tracking data allow for mapping of corridors from directly observed movements, but such tracking data are absent for many populations.
2. We developed a novel statistical corridor modelling approach that predicts movement corridors from cost-distance models fit directly to migration tracking data. Unlike existing predictive approaches, this does not require the ad hoc transformation of habitat suitability surfaces into resistance surfaces. We tested the ability of the approach to recover parameters used to generate simulated movements. We then used GPS data from three migrating mule deer *Odocoileus hemionus* herds in Idaho and Wyoming to model corridors as a function of elevation, slope, aspect, percent shrub, date of peak green-up, snow-off date and human footprint. We assessed the predictive ability of the fitted models using validation tracks from the same herd as well as from the other herds.
3. The approach reproduced parameters used to generate the simulated movements, predicted the corridors used by migratory populations, and described the direction, magnitude and confidence levels of the effects of environmental variables on corridors. Within-herd validation indicated that fitted corridor models are more accurate at predicting migration corridors than null models, and cross-herd validation indicated that fitted models for some herds accurately predicted the observed migrations of other herds.
4. In addition to the practical benefit of mapping corridors for management, our statistical corridor modelling framework sets the stage for evaluating fundamental questions about the fitness trade-offs, navigation, learning, fidelity and movement constraints that influence migratory and other corridor-generating

behaviour. Models of predictive corridors can inform management and planning for the conservation of migrations across taxa, including the potential restoration of corridors. Our corridor modelling approach is also readily applied to non-migratory animal movements.

KEYWORDS
corridor ecology, corridor prediction, cost-distance modelling, maximum likelihood, migration ecology, movement ecology, optimization, statistical corridor modelling

1 | INTRODUCTION

The diversity of animal movements is declining worldwide, and we need new and better tools to understand which movements are essential and how to best conserve them (Doherty et al., 2021; Tucker et al., 2018). Mapping animal movement corridors is central to addressing this challenge; this task is undertaken by a number of ecological disciplines, including migration ecology, movement ecology, landscape ecology, and conservation biology and planning. Researchers often map corridors directly from observed movement tracks using utilization distributions, linear interpolations or other approaches; this is particularly true of migration ecologists (LaPoint et al., 2013; Sawyer et al., 2009). Alternatively, corridors are mapped using predictive modelling approaches, which map corridors as a function of environmental factors. Examples of predictive models include cost-distance models, circuit theory and agent-based simulations (Adriaensen et al., 2003; McRae et al., 2008; Oloo et al., 2018). Unlike corridor maps made directly from observed movements, predictive corridor models can be applied to areas for which movement data are not available, and provide inference about how the physical environment shapes corridors. These aspects make predictive modelling an attractive additional approach for modelling both migratory and non-migratory movement corridors.

Cost distance and circuit theory, the dominant predictive approaches, traditionally have relied on inverting and transforming habitat suitability scores from resource selection functions to generate the cost or resistance values used in mapping corridors (Zeller et al., 2012). The inherent subjectivity and uncertainty in cost surface parameterization has long been a vexing issue for corridor modellers (Beier et al., 2008, 2009; Keeley et al., 2017). Two important developments in connectivity modelling include the use of step, integrated step or path selection functions (Avgar et al., 2016; Zeller et al., 2016), and using location data only from periods of directed movement (Abrahms et al., 2017; Zeller et al., 2014), as these approaches better characterize movement behaviour. However, these approaches still rely on ad hoc transformations of selection coefficients into the resistance surfaces used in cost-distance or circuit theory algorithms. As a result, recent work has shown that transformations of habitat suitability scores, even when drawn from step selection functions, are often poor predictors of connectivity (Brennan et al., 2018; Keeley et al., 2017). Researchers relying on genetic rather than telemetry data have instead been able to optimize resistance parameters directly, bypassing this issue (Hanks & Hooten, 2013; Peterman, 2018).

A statistical corridor modelling framework is needed to rigorously quantify the influence of environmental variables on movements and enable corridor prediction in areas with insufficient tracking data for empirical corridor modelling. Related studies have used maximum likelihood or Bayesian techniques to directly fit cost-distance models to dispersal events (Graves et al., 2014) and circuit theory models to genetic distance (Hanks & Hooten, 2013; Peterman, 2018), but such model fitting approaches are lacking for movement tracks. In this paper, we introduce a statistical cost-distance corridor modelling approach based on maximum likelihood for fitting corridor models to animal tracking data. We used mathematical optimization to identify which environmental covariates best predict the movement tracks of migrating animals, without relying on the intermediate step of building and transforming resource or step selection models. We used simulations and empirical case studies with mule deer *Odocoileus hemionus* migrations to illustrate and validate our approach. Although we focused on the migratory movements of ungulates, the method is readily applicable to migratory taxa across the globe, including other mammals, birds and marine species. It is also applicable to modelling non-migratory directed movements, such as those made during dispersal events or between discrete habitat patches within a home range.

2 | MATERIALS AND METHODS

2.1 | Modelling objective

Our motivating objective was to use the observed Global Positioning System (GPS) tracks of migrating animals to model migratory routes as a function of environmental variables, and then to use these models to generate predictions of migration corridors. We sought to model the spatial route taken by the animal during its migration, and not the amount of time the animal spends in locations along the route. This makes the quantity we modelled different from commonly used spatial point process, species distribution, or resource selection or utilization models, which model probability of occurrence, density of observations or habitat selection as a function of the absolute or relative amount of time spent in different areas (Aarts et al., 2012; Johnson et al., 2013; Northrup et al., 2021). In addition, our objective was distinct from that of correlated random

walk and related movement models, which are commonly used to characterize the step length and turn angle distributions in the movement trajectories of individual animals and their relationship to utilization distributions, space use, behavioural state or home range dynamics (Calabrese et al., 2016; Gurarie et al., 2016).

The custom R functions used to fit the models are available as a supplement to this manuscript (see Data Availability), and key calculations are illustrated step-by-step in Figure 1.

2.2 | The statistical corridor model

Animal movements along a corridor can be characterized as outcomes of a point process where the probability of moving through a location C is a function of how much more distance is involved in moving through C than the distance needed to take the most direct route between corridor endpoints A and B. In a theoretical, homogeneous landscape where movement costs are only influenced by Euclidean distance, the most direct, least costly route between the corridor endpoints is the line segment between the endpoints, $\overline{AB}$. The degree to which a movement through C deviates from $\overline{AB}$ can be measured by the residual or additional distance, d, needed to move through C instead of using the most direct route.

$$d = AC + BC - AB. \tag{1}$$

2.3 | The cost-distance model

In real, heterogeneous landscapes, the environment shapes the costs and benefits of movement, so we used cost-distance modelling to account for these influences. Cost-distance corridor models are appropriate analyses for modelling migrations because they require predetermined start and endpoints, such as the seasonal ranges that define the start and end of migratory movements (Beier et al., 2008). Cost-distance calculations weight Euclidean distance by the degree to which the environment impedes or facilitates movement between a focal pixel and its neighbouring pixel.

If the data we have are a movement track t, consisting of an ordered series of GPS locations with index $i = 1, \ldots, n_{points}$, collected from a single animal migrating in one direction between its seasonal ranges, then the first and last locations in the ordered series correspond to the endpoints A and B, respectively, from our Euclidean example above. Applying Equation 1 to every pixel of the landscape and recasting the variables in cost-distance terms, we calculated raster D_t, the additional cost-distance surface given the endpoints of track t. We did this by calculating the quantities AC and BC as raster surfaces using Dijkstra's algorithm, as implemented by the accCost() function from the GDISTANCE package in R (Adriaensen et al., 2003; van Etten, 2018). The pixels of D_t represent the lowest additional distance (in cost-distance units) of using a pixel instead of the least-cost path to move between the migration endpoints (Figure 1). The additional cost distance at each point i in track t is d_{ti}, a value that can be extracted from raster D_t.

To implement Dijkstra's algorithm, we transformed the gridded landscape into a graph in which each pixel is connected to each of 16 (King's and Knight's moves) neighbours by an edge. The cost distance of an edge is the ratio of the geographical length of the edge to the conductance value, g, of the environmental attributes of the edge, x.

$$\text{cost distance} = \frac{\text{geographical distance}}{g}. \tag{2}$$

The conductance value, g, is the inverse of cost or resistance; there are computational advantages to using conductance instead of cost in these calculations (van Etten, 2018). We can model the influence of an environmental variable, x, on cost distance by calculating g as the exponentiation of the product of the environmental covariate's value for a particular edge and a coefficient β. We present a single covariate model here, but generalize this to multiple covariates later.

$$\frac{1}{\text{cost}} = g = e^{\beta x}. \tag{3}$$

Conductance values, g, range from 0 (infinite cost) to infinity (no cost) and represent the degree to which the landscape facilitates movement between a focal pixel and its neighbouring pixel. If a variable has no effect, its coefficient is 0, conductance is 1 and resulting cost distances are equivalent to geographical distances. If an environmental variable facilitates movement, its coefficient is positive. If a variable impedes movement, its coefficient is negative.

We calculated the environmental covariate value of the edge, x, using either the mean of the focal and neighbour cell values (an isotropic calculation) or the difference of those values (an anisotropic calculation). Using a difference calculation measures the effect of stepwise (focal-to-neighbour) changes in a variable, such as elevation or temperature, on movement (van Etten, 2018). We denote stepwise difference covariates with the subscript *diff* in the results section.

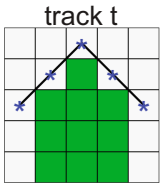
$$x_{\text{edge,mean}} = \frac{x_{\text{focal}} + x_{\text{neighbor}}}{2}, \tag{4}$$

$$x_{\text{edge,difference}} = | x_{\text{neighbor}} - x_{\text{focal}} |. \tag{5}$$

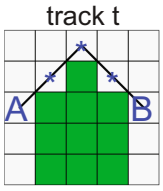
2.4 | The deviation kernel

The probability of an animal moving through location i with a given additional cost distance d while moving between the endpoints of track t can be characterized by an exponential 'deviation kernel', $\lambda e^{-\lambda d_{ti}}$, analogous to the dispersal kernel used to model the probability of an organism dispersing a given distance from its natal location (Graves et al., 2014). Because the additional distance travelled must be ≥ 0, we used the exponential distribution for our deviation kernel, but a half-normal or other distribution could also be used. The exponential distribution has the useful property that the inverse of the rate parameter λ is the mean of the distribution, so the fitted value of $1 / \lambda$ is the mean additional distance, allowing us to directly assess

Our objective is to calculate the likelihood of track t for two different cost distance models, β = 0 and β = -1, where cost = $1/e^{\beta x}$, and x is an environmental covariate



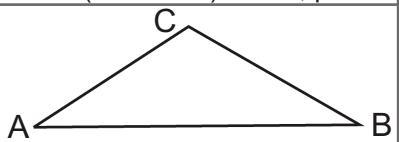
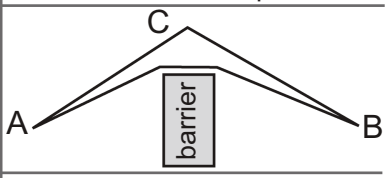
Track t begins at point A and ends at point B



Track t avoids moving through where x =1

covariate x

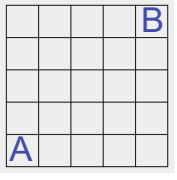
0	0	0	0	0
0	0	1	0	0
0	1	1	1	0
0	1	1	1	0
0	1	1	1	0

	Null (Euclidean) model, β = 0	Avoidance model, β = -1
Our conceptual model defines the probability of moving through any point C on the landscape while moving from A to B probability(C) ∝ AC + BC - AB		
We calculate cost rasters from covariate x and the coefficient β. A cost value of 1 means geographic distance = cost distance cost = $1/e^{\beta x}$	cost 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 If β = 0, x has no effect on cost	cost 1 1 1 1 1 1 1 2.72 1 1 1 2.72 2.72 2.72 1 1 2.72 2.72 2.72 1 1 2.72 2.72 2.72 1 If β = -1, x increases costs
We calculate cost distances from A (AC) and from B (BC) using Dijkstra's algorithm and the cost raster If β = 0, AB = 4; if β = -1, AB = 5.7	AC 2 2.4 2.8 3.8 4.8 1 1.4 2.4 3.4 4.4 A 1 2 3 4 1 1.4 2.4 3.4 4.4 2 2.4 2.8 3.8 4.8 BC 4.8 3.8 2.8 2.4 2 4.4 3.4 2.4 1.4 1 4 3 2 1 B 4.4 3.4 2.4 1.4 1 4.8 3.8 2.8 2.4 2	AC 2 2.4 2.8 3.8 4.8 1 1.4 2.9 4.2 5.2 A 1.5 3.5 5.7 5.7 1 2.1 4.8 7.3 6.7 2 3.1 5.8 8.5 7.7 BC 4.8 3.8 2.8 2.4 2 5.2 4.2 2.9 1.4 1 5.7 5.7 3.5 1.5 B 6.7 7.3 4.8 2.1 1 7.7 8.5 5.8 3.1 2
Raster *D* is the additional distance needed to move through a pixel relative to the least cost path AB D = AC + BC - AB	*D* 2.8 2.2 1.7 2.2 2.8 1.4 0.8 0.8 0.8 1.4 A 0 0 0 B 1.4 0.8 0.8 0.8 1.4 2.8 2.2 1.7 2.2 2.8	*D* 1.2 0.6 0 0.6 1.2 0.6 0 0.1 0 0.6 A 1.5 1.3 1.5 B 2 3.7 3.9 3.7 2 4 5.9 5.9 5.9 4
Raster *L* is the likelihood of moving through a pixel $\lambda e^{-\lambda D}$ is the exponential deviation kernel with scale parameter λ The likelihoods of the points of track t (except track endpoints A and B) are circled $L = \frac{\lambda e^{-\lambda D}}{\sum_S \lambda e^{-\lambda D}}$	*L*, λ = .75 0.01 0.02 0.03 0.02 0.01 0.03 0.05 0.05 0.05 0.03 A 0.09 0.09 0.09 B 0.03 0.05 0.05 0.05 0.03 0.01 0.02 0.03 0.02 0.01	*L*, λ = .75 0.04 0.06 0.09 0.06 0.04 0.06 0.09 0.08 0.09 0.06 A 0.03 0.03 0.03 B 0.02 0.01 0 0.01 0.02 0 0 0 0 0
l_t is the likelihood of track t, reported as a negative log likelihood $l_t = \prod_{i=2}^{n-1} \frac{\lambda e^{-\lambda d_{ti}}}{\sum_S \lambda e^{-\lambda D_t}}$	-log(l_t) = 9.79	-log(l_t) = 7.22
	Inference: model β = -1 is more likely to have generated track t	

Prediction

We can predict a corridor by mapping *L* for new endpoints and a different landscape

New endpoints



New area, covariate x

0	1	0	0	0
0	1	1	0	0
0	1	1	1	0
0	1	1	0	0
0	0	0	0	0

In this example, we use β = -1 and λ = .75

Predicted *D* raster

2.1	1.7	1.3	1.3	B
1.7	1.3	2.8	0.8	0
1.3	2.4	2.1	0.9	0
1.3	2.1	1.1	0	0.6
A	0	0	0.6	1.2

Predicted *L* raster

0.02	0.02	0.03	0.03	B
0.02	0.03	0.01	0.04	0.07
0.03	0.01	0.01	0.04	0.07
0.03	0.01	0.03	0.07	0.05
A	0.07	0.07	0.05	0.03

FIGURE 1 Illustration of likelihood calculations in the statistical corridor modelling framework. In practice, it is critical that the state space S be sufficiently large that *L* values at the edge of the landscape equal or approach zero.

how far on average an animal deviates from the modelled least cost path in cost-distance units.

2.5 | Estimation by maximum likelihood

Adapting the approach of Graves et al. (2014), we use a multinomial probability mass function to model the event of an animal moving through a location with a specific additional cost distance d, in discrete space. Specifically, we calculated the likelihood of observing a point i in track t as the multinomial probability of observing d_{ti} relative to the sum of all the values of D_t.

$$l_{ti} = \frac{\lambda e^{-\lambda d_{ti}}}{\sum_S \lambda e^{-\lambda D_t}}. \quad (6)$$

The denominator $\sum_S \lambda e^{-\lambda D_t}$ represents the sum of the pixel-level probabilities of the entire raster on which the corridor is mapped, which is the state space S, and represents all possible paths between the movement track endpoints. To capture all the possible paths, the state space S (the raster surface) must be sufficiently large so that the pixel-level probabilities at the edge of the raster approach zero.

Calculating the likelihood of a track requires combining the likelihoods of all points within the track. This can be done in one of several ways depending on how the geometry of the d_{ti} calculation is undertaken. These approaches differ in their computational requirements, assumptions regarding the animal's movements and whether they account for the serial nature of movement tracks, and are outlined in detail in Figure S1. To minimize computing time, we used the deviation-from-global-optimum approach, which does not account for serial autocorrelation among points, but instead treats them as independent observations. Readers may be interested in using one of the alternate geometries outlined in Figure S1 if serial autocorrelation is a concern.

If we assume that the d_{ti} values are mutually independent, the likelihood of observing the set of points within track t, l_t, is the product of all of the point-level likelihoods, except for the track endpoints.

$$l_t = \prod_{i=2}^{n_{\text{points}}-1} \frac{\lambda e^{-\lambda d_{ti}}}{\sum_S \lambda e^{-\lambda D_t}}. \quad (7)$$

We also calculated a spatial likelihood surface. L_t represents the predicted corridor in likelihood units. The values of L_t sum to 1, and each pixel represents the likelihood a pixel will be used, conditional on the cost-distance model, movement track endpoints and state space S.

$$L_t = \frac{\lambda e^{-\lambda D_t}}{\sum_S \lambda e^{-\lambda D_t}}. \quad (8)$$

2.6 | Sampling and autocorrelation

We need a spatially representative sample of the true movement route to fully capture the effects of environmental covariates between the track endpoints. Complicating this, GPS collars typically record locations at fixed intervals in time rather than space so that fixes are generally not regularly spaced. During stopover or foraging periods, the animal will remain in the same section of the route, resulting in many closely spaced GPS fixes, whereas rapid, directed movements along the route will be poorly represented by fixes. The spatially irregular sampling of a raw GPS track would give stopover locations higher weight in the likelihood calculations, biasing the fitted route towards stopovers. To address this, in our case study, we thinned GPS tracks so that points were not closer than 2 km. This threshold allowed us to remove duplicate stopover locations, but was still smaller than the resolution of the environmental covariates, so stopover locations were still represented, and the overall architecture of the route (at the resolution of the raster) was not affected.

A closely related concern is whether serial correlation (temporal or spatial) of points in a GPS track results in autocorrelation of the d_{ti} values used in the likelihood calculation, which would violate the assumption that they are mutually independent. Although autocorrelation may not affect coefficient estimates or the predictive accuracy of fitted models, by deflating error estimates it increases the risk of type I errors in inferential applications (Boyce, 2006). For inferential applications, thinning data spatially and diagnostic plotting of d_{ti} values (described below) provide a way of reducing and assessing the degree of autocorrelation, as would use of one of the alternate geometries for calculating d_{ti} outlined in Figure S1.

2.7 | Optimization and software implementation

We used the likelihood equation (Equation 8) as the objective function in a Nelder–Mead optimization routine to identify the most likely values of β (the coefficients) and λ (the shape parameter). We implemented all calculations in version 4.0.0 of R (R Core Team, 2020). In particular, we used the following packages: RASTER for raster data manipulation, GDISTANCE for cost-distance calculations, STATS for the optim function used for optimization and PARALLEL to parallelize the track-level optimizations (Hijmans & van Etten, 2012; R Core Team, 2020; van Etten, 2018).

2.8 | Herd-level inference

Herd-level estimates of β coefficients require extending the likelihood calculation to the entire migratory herd, h, which included a set of migration tracks $t = 1 \ldots n_{tracks}$. We used a two-stage approach for estimating effects at the population level (Fieberg et al., 2010). We fit estimates of $\hat{\beta}_t$ and $\hat{\lambda}_t$ for each track separately, then averaged the track $\hat{\beta}_t$ values to calculate the herd-level $\hat{\beta}_h$ and $\hat{\lambda}_h$, and calculated the standard error for use in calculating confidence intervals:

$$\hat{\beta}_h = \text{mean}\left(\hat{\beta}_{t=1,\ldots,n_{\text{tracks}}}\right), \quad (9)$$

$$SE_{\widehat{\beta}_h} = \frac{sd\left(\widehat{\beta}_{t=1,\ldots,n_{\text{tracks}}}\right)}{\sqrt{n_{\text{tracks}}}}, \tag{10}$$

$$CI_{\widehat{\beta}_h,95\%} = \widehat{\beta}_h \pm 1.96 \cdot SE_{\widehat{\beta}_h}. \tag{11}$$

Following Fieberg et al. (2010), we calculated the herd-level likelihood as the product of the track-level likelihoods with the assumption that they are independent of each other:

$$l_h = \prod_{i=1}^{n_{\text{tracks}}} \prod_{i=2}^{n_{\text{points}}-1} \frac{\lambda e^{-\lambda d_{ti}}}{\sum_S \lambda e^{-\lambda D_t}}. \tag{12}$$

The herd-level likelihood surface is the mean of the individual track probability surfaces:

$$L_h = \left(\sum_{i=1}^{n_{\text{tracks}}} L_t\right) / n_{\text{tracks}}. \tag{13}$$

We explored using a random effects model instead of the two-stage approach, which would require calculating the joint probability of the track-level random effect and the point-level effects from the deviation kernel. This joint probability cannot be calculated analytically, because the probabilities depend on the cost-distance calculations that generate the residual cost surfaces from which the likelihoods are derived. The large number of cost-distance calculations required to numerically approximate the integral for the convolution of the two random variables makes this approach computationally impractical at present. Relative to a mixed effects model, variance estimators will be biased high in the two-stage approach which we took, making our approach comparatively conservative (Fieberg et al., 2010).

2.9 | Model selection

We used Akaike's information criterion (AIC) to identify the most informative model from a set of cost-distance models using different combinations of covariates. We specified that k, the number of estimated parameters, was 1 (for $\widehat{\lambda}$) plus the number of estimated $\widehat{\beta}$ coefficients.

2.10 | Model assumptions

Our framework relies on the following assumptions about the data and animal behaviour: (1) the beginning and ending locations of a migration must be known, (2) movement locations occurring between the beginning and ending locations are all part of the migratory route (i.e. they do not contain back-and-forth, nomadic or exploratory movements), (3) that animals minimize costs and optimize the entire migration route rather than making movement decisions incrementally, and (4) that serial correlation (temporal or spatial) in the d_{ti} values are negligible.

2.11 | Simulation

We simulated movement tracks using different values of β and λ, and assessed the ability of the optimization routine to recover the β and λ values used in the simulation. We used an elevation raster and a human footprint raster as variables in univariate cost-distance models of the form conductance $= e^{\beta * \text{Elevation}}$ and conductance $= e^{\beta * \text{Human Footprint}}$ (Table 1). These rasters had extents of 4 ×4 degrees and a grain of 75 arc-seconds (~2.3 km). We randomly distributed the start and endpoints (20 each) for the tracks in the southern and northern 10% of a 2 ×2 degree box centred in the rasters. We used all combinations of β and log(λ) values (all integers from −4 to 4 for β values, and from −9 to −1 for log(λ) values). For each set of β and λ values, scaled covariate surface and set of track start and endpoints, we generated a L_t surface, and then randomly sampled 20 points in proportion to L_t. We used the model fitting procedure described above to estimate $\widehat{\beta}$ and $\widehat{\lambda}$ values from the resulting point samples.

2.12 | Mule deer migration case studies

We fit cost-distance models to movement tracks from three migratory herds of mule deer (Figure 2). Data from the Tex Creek herd in eastern Idaho were collected between 2007 and 2018. Animals in the Antelope Creek herd, in central Idaho, were collared between 2011 and 2018. The Red Desert to Hoback herd, a portion of the Sublette herd, migrates through western Wyoming, and those data were collected between 2011 and 2018 (Sawyer et al., 2016). Idaho and Wyoming mule deer capture and handling protocols were approved by the Institutional Animal Care and Use Committees of the Idaho Department of Fish and Game Wildlife Health Laboratory and the University of Wyoming (including protocols 20131111KM00040, 20151204KM00135 and 20170215KM00260) and Wyoming Game and Fish Department (Chapters 33–937), respectively.

We only used spring migrations. The beginning and end of each migration were identified from changes in the net squared displacement of a track over time (Merkle et al., 2019). For each track, we thinned points so that consecutive points in the series were at least 2 km apart, ensuring that areas with clusters of GPS points were not overrepresented, and retained only thinned tracks which had at least 10 remaining points. For each herd, we divided the thinned migration tracks into two equal groups of up to 30 tracks, and used one for fitting the maximum likelihood models and the other for model evaluation. This resulted in 30 training and 30 evaluation tracks for the Red Desert to Hoback migration, 21 training and 21 evaluation tracks for the Tex Creek migration, and 14 training and 14 evaluation tracks for the Antelope Creek migration.

2.13 | Environmental covariates

We used Julian date of peak instantaneous rate of green-up (IRG), elevation from a digital elevation model (DEM), slope, aspect, percent

TABLE 1 Environmental covariates used in simulation and model fitting

Acronym	Description	Hypothesized effect	Details and data citation
DEM_{diff}	Elevation	Stepwise change in elevation along a route may increase energetic cost and cost distance	Elevation in the Western United States from the National Elevation Dataset, originally at a 90 m resolution (Leu et al., 2008)
SLOPE SOUTH	Elevation derivatives	Steeper slopes and south-facing aspects, which are associated with shrub habitats, complex topography and less snow depth, respectively, hinder and facilitate movement	Elevation derivatives calculated from the DEM layer are slope (SLOPE) and aspect, which is measured in degrees from due north (SOUTH). Elevation derivatives were calculated using the terrain() function in the RASTER package in R (Hijmans & van Etten, 2012; Leu et al., 2008)
IRG_{diff}	Date of peak IRG	Stepwise change in IRG will increase cost distances, if mule deer surf a resource wave of greening vegetation	Date of peak IRG values were averaged from 2001 to 2014, using spatial data prepared according to Merkle et al. (2016)
HF	Human footprint	Anthropogenic influence impedes movement	The human footprint index is derived by weighted scoring of human habitation, energy, irrigation, and transportation infrastructure, and agricultural use (Leu et al., 2008)
SHRUB	Percent shrub	Shrub vegetation is a primary source of forage for mule deer, and facilitates movement	All shrub species are included. Data are from USGS shrubland map data products, originally at a 30 m resolution (Young, 2017)
$SNOW_{diff}$	Snow-off date	Stepwise change in snow-off date, which results in increased movement through snow, will increase cost distance	The average date at which snow was no longer on the ground from 2001 to 2015, derived from MODIS (O'Leary III et al., 2017)

Abbreviation: IRG, instantaneous rate of green-up.

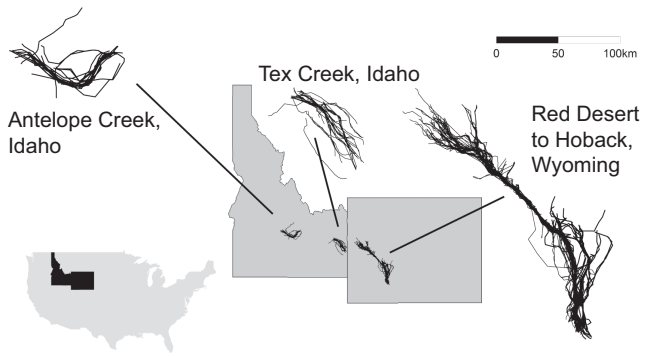


FIGURE 2 Study region and migration tracks used in case study, from Tex Creek (Idaho), Antelope Creek (Idaho) and Red Desert to Hoback (Wyoming) mule deer migrations.

shrub cover, snow-off date and human footprint as covariates (Table 1). Each of these covariates influence mule deer movements or habitat selection (Merkle et al., 2019; Sawyer et al., 2017; Wyckoff et al., 2018). All rasters of covariates were resampled using bilinear interpolation to a 15 arc-second (~0.5 km) unprojected grid in the WGS 84 coordinate reference system (Hijmans & van Etten, 2012). We used a correction for calculating cost distances on unprojected grids in the gdistance package. Each herd had its own analysis area with an extent 30% larger than the combined extent of the tracks; this is the state space S for the herd, described above.

We standardized environmental variables to aid convergence in model fitting. We did this by scaling them by dividing by the root mean square of the variable in each herd's study area. We did not centre the variables so that coefficients could be interpreted relative to the variable's actual zero value. We did not standardize variables used in stepwise difference cost calculations, including elevation, date of peak IRG or snow-off date, as doing so resulted in very small stepwise changes between pixels, large ($>10^3$) fitted coefficients and convergence issues. We found that scaling stepwise difference variables by factors of 10 such that the estimated coefficients ranged between 0.1 and 10, facilitated convergence and interpretability. We scaled date of peak IRG and snow-off date by 100, and elevation by 10. IRG, elevation and snow-off date were strongly correlated (Pearson's p values >0.7) with each other.

2.14 | Model fitting

We fit the models at a 120 arc-second (~3.7 km) pixel size using (1) a predictive candidate model set and (2) an explanatory candidate model set (Tredennick et al., 2021). For predictive modelling, we included all covariates, including highly correlated ones, in the full model SLOPE + SOUTH + SHRUB + HF + IRG + DEM + SNOW. We then used AIC to identify the best model from a candidate set of the full model and all nested models. For explanatory modelling with the goal of ecological interpretation, we excluded any pairs of variables with Pearson's r values >0.7. Using AIC, we selected the best model from a candidate model set consisting of (1) the three full models, each of which included all four non-correlated covariates (SLOPE,

SOUTH, SHRUB and HF) and only one of the correlated stepwise difference covariates (IRG, DEM and SNOW), and (2) all models nested within each of the three full models.

2.15 | Model evaluation and cross-validation

We evaluated the predictive ability of the best-performing predictive model for each herd against the evaluation tracks of the same herd, and against all tracks of the other two herds. We also compared these predictions against a null model prediction consisting of a straight-line path (with $\log(\lambda) = -5$). To do so, we began using either the null model or the top-performing predictive cost-distance model fitted to the training tracks of one herd (see Table S1) to project likelihood corridors, L_h, for all three herds. We did this using the estimated $\hat{\beta}_h$ and $\hat{\lambda}_h$ values to predict likelihood corridors, L_t, between the endpoints of each track. We took the mean of the L_t surfaces to generate L_h. To determine the endpoints for the corridor projections, we used the endpoints of all tracks (training and evaluation tracks combined) of the other two herds as well as just the evaluation tracks of the training herd, ensuring that a model's training tracks were not used in its evaluation. For each model, we then quantified the proportion of the area of L_h needed to contain 95% of points in each track in the evaluation herd, providing a track-level assessment of the predictive ability of the models across herds (Figure S7). We then plotted the percentile of each predicted L_h against the proportion of validation track GPS points contained in each L_h percentile (Figure 4). This measures the ability of a fitted model to predict movement track locations, and allows comparison of the predictive ability of different models. We plotted the predicted L_h surfaces and the evaluation tracks to visualize their spatial differences (Figure 5). We also mapped the predicted L_h surfaces for the training tracks to visually assess the spatial fit of the training data to the fitted model (Figures S8–S10).

We assessed the sensitivity of our coefficient estimates to the 2 km threshold used to rarefy the movement tracks by fitting the top-performing explanatory model to the Antelope Creek herd tracks, rarefied at 100, 300, 500, 1,000, 1,500, 2,000 and 3,000 m thresholds (Figure S14). To visually check for autocorrelation, we used plots of d_{ti} values (which can be thought of as residuals) against observation order, as well as full and partial autocorrelation function plots (Figure S15). We tested for autocorrelation using Spearman's rank correlation coefficient at a lag of one GPS point, for the top-performing explanatory model for each herd.

3 | RESULTS

3.1 | Recovery of simulated parameters

Our model fitting approach successfully estimated the original parameter values (β and λ) from simulated movement track data, with two important exceptions resulting from parameter values that were inappropriate for the spatial extent and grain of our simulation (Figures S3–S6). First, the combination of very small $\log(\lambda)$ values (large mean residual distances) and large β values (which act to 'shrink' Euclidean distance) resulted in very dispersed point distributions and likelihood surfaces relative to the scale of the simulation area, causing fitted coefficients to be underestimated (Figure S2). For example, for an area where the environmental covariate has a value of 1, a $\log(\lambda)$ value of −9 and a β coefficient of 4 would result in a mean residual distance in Euclidean terms of ~442 km, approximately the size of our simulation area. Second, the combination of large $\log(\lambda)$ values and small β values caused the mean residual distance of the simulated points to fall far below the spatial resolution of the simulation and inflated fitted estimates of $\log(\lambda)$ (although β estimates remained accurate). For example, if $\log(\lambda) = -4$, mean residual distances will be ~54 m, whereas the resolution of our simulation was ~2.3 km. The deviation-from-global-optimum geometry we used, which lacks serial autocorrelation and treats points independently, caused the simulated points to be scattered across the likelihood surface (Figure S2); alternate geometries that incorporate autocorrelation would allow more realistic movement track prediction.

3.2 | Ranking of fitted explanatory models by delta AIC

The best-fitting explanatory models all included human footprint, slope, percent shrub and south-facing aspects, and each included a different one of the three correlated, stepwise difference variables (snow-off date, IRG and elevation; Table 2). The best-fitting predictive models included all variables except human footprint for Antelope Creek and Red Desert to Hoback, and all variables except slope for Tex Creek.

3.3 | Coefficients of best-fitting explanatory models

The direction and significance of the effects of environmental variables on movement costs varied by herd. Only south-facing aspects were significant for all three herds, but they facilitated movements for the Antelope Creek and Tex Creek herds, and impeded movement for the Red Desert to Hoback herd (Figure 3, Table 3). All five variables were significant at the herd level for the Red Desert to Hoback herd, three (SHRUB, SOUTH and SNOW) were significant for the Antelope Creek herd, and only one (SOUTH) was significant for Tex Creek.

Human footprint impeded movements for all herds, although it was significant only for the Red Desert to Hoback herd. Steeper slopes facilitated movements for the Tex Creek and Red Desert to Hoback herds, although were only significant for the Red Desert to Hoback herd. Percent shrub significantly facilitated movements for the Antelope Creek and Red Desert to Hoback herds. Stepwise

TABLE 2 Explanatory model rankings based on delta AIC. Explanatory models do not include correlated variables DEM_{diff}, IRG_{diff} and $SNOW_{diff}$ within the same model

Herd	Model	dAIC
Antelope Creek	$HF + SLOPE + SHRUB + SOUTH + SNOW_{diff}$	0
	$SLOPE + SHRUB + SOUTH + SNOW_{diff}$	17.1
	$HF + SHRUB + SOUTH + SNOW_{diff}$	20
	$HF + SLOPE + SHRUB + SOUTH + IRG_{diff}$	23.9
	$HF + SLOPE + SHRUB + SOUTH + DEM_{diff}$	41.1
Red Desert to Hoback	$HF + SLOPE + SHRUB + SOUTH + IRG_{diff}$	0
	$HF + SLOPE + SHRUB + IRG_{diff}$	18.9
	$SLOPE + SHRUB + SOUTH + IRG_{diff}$	30.1
	$HF + SLOPE + SHRUB + DEM_{diff}$	90.7
	$HF + SLOPE + SHRUB + SOUTH + DEM_{diff}$	99.3
Tex Creek	$HF + SLOPE + SHRUB + SOUTH + DEM_{diff}$	0
	$HF + SLOPE + SHRUB + SOUTH$	10.6
	$HF + SLOPE + SOUTH + SNOW_{diff}$	20.6
	$HF + SHRUB + SOUTH + DEM_{diff}$	31.3
	$HF + SHRUB + SOUTH + SNOW_{diff}$	43.4

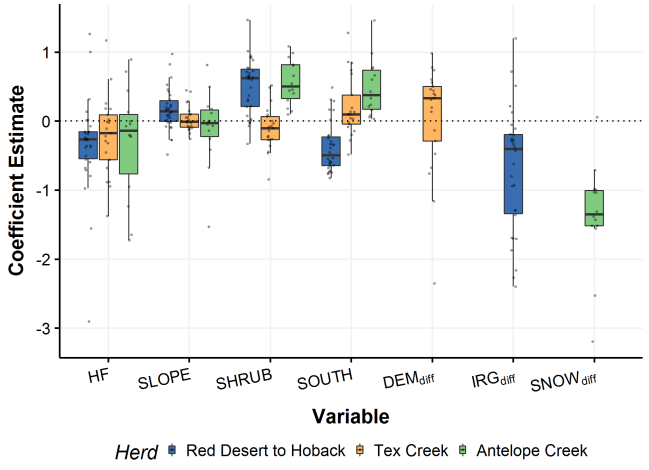


FIGURE 3 Cost-distance coefficient estimates for the top-performing explanatory model for each herd. DEM, IRG and SNOW were correlated, so they were not included together in the same explanatory models. Coefficients above 0 indicate the variable decreases costs, and below 0 indicate increased costs to movement.

changes in snow-off date impeded movement for the Antelope Creek herd, and date of peak IRG impeded movement for the Red Desert to Hoback herd.

3.4 | Predicted corridor surface model evaluation and cross-prediction

Based on the proportion of the study area needed to capture each percentile of the evaluation tracks, we found that for all herds the fitted predictive models predicted the locations of same-herd evaluation tracks better than a null model based on a straight line between track endpoints (Figure 4; Figures S7 and S11–S13). Fitted models had mixed ability to predict the tracks of other herds. The Red Desert to Hoback model performed well in predicting the Antelope Creek herd, and vice versa. Conversely, the Red Desert to Hoback and Antelope Creek herds performed equivalent to or worse than the null model at predicting the Tex Creek tracks, and the Tex Creek model was the least predictive of the other herds' tracks. These general patterns were visually apparent in the spatial patterns of the cross-herd predictions with the evaluation tracks (Figure 5). Coefficient estimates were insensitive to changes in the spatial threshold used to rarefy tracks (Figure S14). At a one-point lag, 2 out of 14 of the Antelope Creek, 4 out of 21 of the Tex Creek and 18 of 30 of the Red Desert to Hoback tracks' d_{ti} values were significantly correlated (Spearman's rank correlation p-values <0.05).

4 | DISCUSSION

4.1 | General significance

Our statistical corridor modelling approach enables hypothesis testing, model selection, corridor prediction and quantifies the influence of environmental factors on migratory or other cost-minimizing movements. Furthermore, ecologists can use it to quantify uncertainty in the spatial location of corridor predictions and in the model coefficients, to incorporate multiple environmental factors that can either increase or decrease movement costs and to measure individual variability in movement behaviour and corridor selection.

Recent corridor modelling studies have focused on post-hoc testing of corridor predictions against movement tracks (Bond et al., 2017; Keeley et al., 2017; McClure et al., 2016; Zeller et al., 2018). These studies have established that corridor predictions are sensitive to the transformation of habitat suitability or selection scores to a cost surface, thereby affecting the degree to

TABLE 3 Fitted coefficients and 95% confidence intervals for each herd's top-performing explanatory model. Significant coefficients (confidence intervals not overlapping 0) are bold

Herd	Model	log(λ)	HF	SLOPE	SHRUB	SOUTH	DEM	SNOW	IRG	Metric
Antelope Creek	HF+ SLOPE+ SHRUB+ SOUTH+ $SNOW_{diff}$	−4.49	0.12	0.18	0.73	0.69	NA	−0.98	NA	Upper 95% C.I.
		−5.29	−0.30	−0.09	**0.56**	**0.48**	NA	**−1.37**	NA	$\hat{\beta}_h$
		−6.08	−0.71	−0.37	0.40	0.27	NA	−1.76	NA	Lower 95% C.I.
Red Desert to Hoback	HF+ SLOPE+ SHRUB+ SOUTH+ IRG_{diff}	−6.19	−0.09	0.28	0.65	−0.25	NA	NA	−0.38	Upper 95% C.I.
		−6.47	**−0.34**	**0.17**	**0.51**	**−0.38**	NA	NA	**−0.70**	$\hat{\beta}_h$
		−6.75	−0.59	0.06	0.36	−0.51	NA	NA	−1.03	Lower 95% C.I.
Tex Creek	HF+ SLOPE+ SHRUB+ SOUTH+ DEM_{diff}	−5.29	0.03	0.12	0.03	0.39	0.04	NA	NA	Upper 95% C.I.
		−5.62	−0.21	0.04	−0.10	**0.21**	0.00	NA	NA	$\hat{\beta}_h$
		−5.96	−0.45	−0.04	−0.24	0.02	−0.03	NA	NA	Lower 95% C.I.

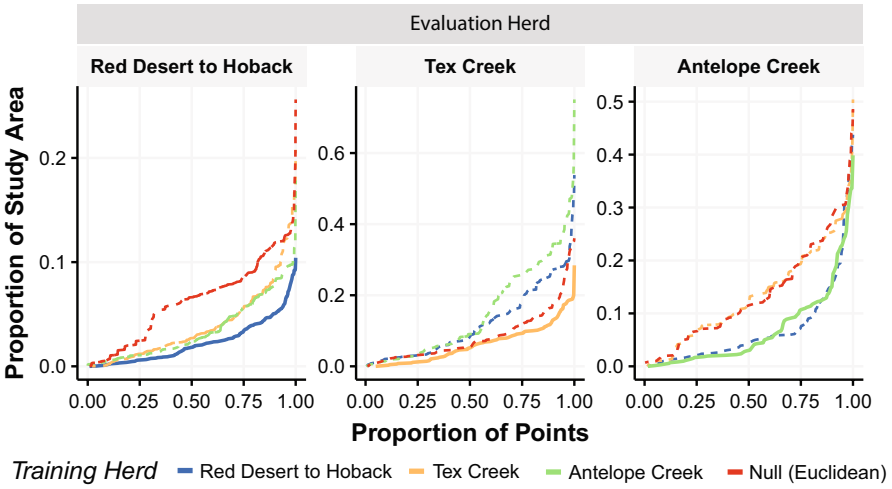


FIGURE 4 Cross-herd predictive ability of the top-performing predictive models (Table S1), measured as the proportion of the predicted corridor likelihood surface needed to contain a given proportion of the evaluation herd's GPS points. The ability of a model to predict migration is better when a higher proportion of the evaluation herd's track points (*x*-axis) are retrieved within in a smaller proportion of the evaluation herd study area (*y*-axis). Within-herd predictions are shown in solid lines, while cross-herd evaluations are shown with dashed lines.

which corridor predictions match movement tracks. Our approach directly fits cost coefficients to the movement data, ensuring that the relationship between geographical distance and environmentally weighted cost distance reflects the animals' actual movements. Cost-distance corridor boundaries are commonly delineated by specifying an arbitrary cost-distance threshold relative to the least cost path. In contrast, our approach allows for the use of probability-based corridor boundary thresholds, because we represent predicted corridors as likelihood surfaces, and these probability-based thresholds derive directly from how points are distributed around the least cost path.

4.2 | Interpretation of simulation results

Our simulations showed that the maximum likelihood optimization routine can recover reasonable simulated coefficients and log(λ) values from the simulated tracks, as long as those values are appropriate for the extent and grain of the study. We can be confident in the coefficients fitted to real movement tracks in the case studies, as the fitted empirical coefficients were largely between −4 and 4, and log(λ) values remained above −7.

4.3 | Interpretation of case study results

The coefficients fitted to the mule deer migration tracks in the explanatory models were consistent with previous mule deer habitat selection studies. Previous studies have found that shrubs and south-facing slopes are important to mule deer habitat (Bishop et al., 2005; Merkle et al., 2019). Conversely, our results indicate that human footprint hinders movements, consistent with existing research on the effect of anthropogenic disturbance on mule deer migrations (Wyckoff et al., 2018). Elevation, snow-off date and date of peak IRG had Pearson's correlation coefficients greater than 0.7, making inference about their independent effects on migration corridors difficult. Deer selected routes that avoided stepwise change in snow-off date and date of peak IRG when these were included

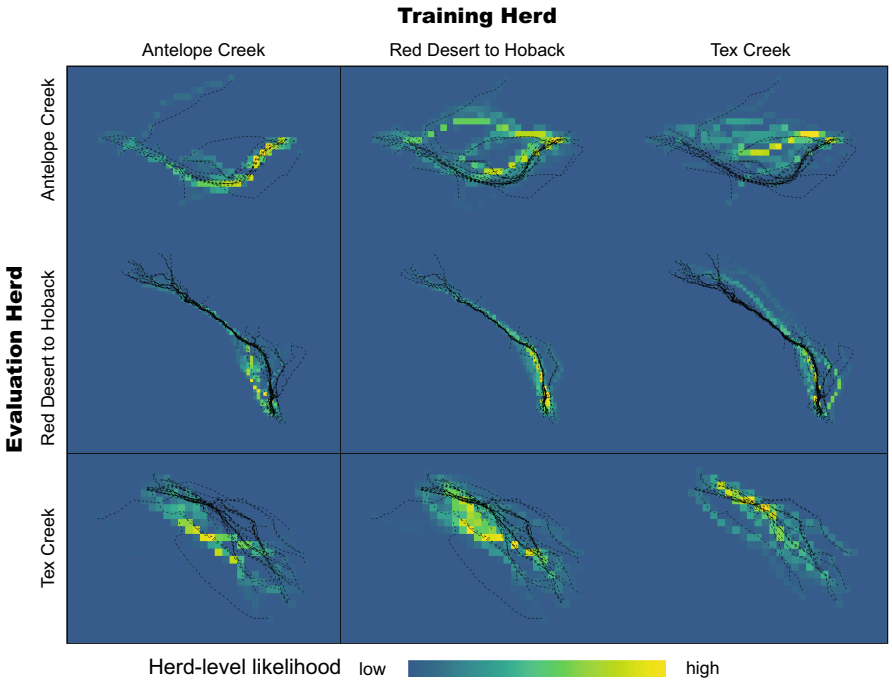


FIGURE 5 Cross-herd cost-distance model predictions using the top-performing predictive models (Table S1). Models were fit using training herd tracks, and then used to create a herd-level likelihood corridor surface using the endpoints of the evaluation herd tracks. The evaluation herd tracks (black dashed lines) are shown over the predicted likelihood corridor surface. When training and evaluation herds are the same, the set of tracks we used for evaluation was different from that which we used for training.

separately as stepwise difference variables, consistent with evidence that mule deer track green-up across landscapes (Aikens et al., 2017; Merkle et al., 2016). Paths with lower stepwise changes in the date of peak IRG or snow-off date along the corridor would make green-wave tracking, or tracking of receding snow, more profitable.

4.4 | Model validation

All within-herd predictions performed better than the null Euclidean model, both at a track level (Figure S7) and at the herd level (Figure 4), illustrating the basic validity of the model. Cross-herd predictive ability was best between the Red Desert and Antelope Creek herds, and worst between those and the Tex Creek herd. This may be due to differences in the spatial configuration of covariates in the migratory landscape or endpoint geometry (Short Bull et al., 2011). The Tex Creek migrations are relatively linear over a homogeneous landscape and thus may be less influenced by the environmental factors included in the models.

4.5 | Modelling assumptions and statistical considerations

Cost-distance models are subject to two primary assumptions, including (1) the animals are able to identify the optimal route between endpoints, for example, as the result of a population learning the optimal route through trial and error over many generations, and (2) the animals will prefer the least-cost, most direct route (in cost-distance terms) between two points in space (Beier et al., 2008). Mule deer typically migrate between distinct seasonal ranges and show strong spatial memory and fidelity to individual migration routes and seasonal ranges that match the framework used here, with start and endpoints (Merkle et al., 2019). Non-migratory directed movements, such as those between distinct habitat patches within a home range, also work well with these assumptions. Cost-distance methods are less appropriate for species, contexts or spatiotemporal scales where animals are not acting to reduce cost distance and moving to a specific destination. For example, it may be less appropriate to fit models to the tracks of nomadic movements, natal dispersal events where the destination is unspecified or area-restricted search foraging movements.

The statistical framework of maximum likelihood also imparts its own assumptions and constraints. Although computationally expedient, the geometry which we used (deviation-from-global-optimum) does not account for the serial nature of movement tracks, potentially leading to autocorrelation in the d_{ti} values within a track. At a one-step lag, we saw a minor amount of autocorrelation for the Antelope Creek and Tex Creek tracks' d_{ti} values. However, there was a substantial amount of autocorrelation in the Red Desert to Hoback d_{ti} values, as these tracks tended to have more points than those of the other herds. Our two-stage approach (which averaged track-level coefficient estimates to arrive at herd-level coefficient estimates and confidence intervals) was robust to the potential for inflated track-level error estimates, but we also could have addressed this by further spatial thinning of the Red Desert to Hoback points, or using an alternate geometry for the d_{ti} calculations.

The herd-level likelihood equations assume that individual tracks are independent of each other. While the effects of within-herd track correlation merit further analysis, this is a reasonable assumption for matriarchal groups within a migratory herd, as groups depart and arrive at distinct winter and summer home ranges at different times. We addressed multicollinearity by conducting model selection separately for the explanatory models (which excluded highly

correlated variables) and predictive models (which did not). Further work is needed to understand the effect of multicollinearity on cost-distance models, and the best way to quantify correlations (Van Strien et al., 2012).

Generating cost-distance surfaces is computationally intensive, and computational constraints limit the spatial resolution and number of tracks that can be included in an analysis. Limiting the number of candidate models and coarsening the resolution greatly decreased the time needed for model fitting. At a 120 arc-second (~3.7 km) resolution, running the model selection for the three herds took approximately 5 days using 40 cores on a 2.4 GHz server. Scaling this approach beyond the scope presented here would benefit from the use of high-performance computing resources.

4.6 | Application and extensions

This methodology enables the direct statistical characterization of how environmental factors impede or promote movement in corridors, and can be applied to fundamental research in movement, migration and corridor ecology. The ability to test hypotheses regarding environment–corridor relationships can be used to assess the consequences of land use or climatic changes occurring in a corridor. Future work could measure the link between modelled cost-distance metrics and the caloric costs of changing movement paths in response to environmental change. Researchers can readily extend this approach to other migratory and non-migratory taxa where animals are acting to reduce movement costs during migration, within-home range or dispersal movements, as long as movements are directional and link discrete start and end locations. For example, one could fit cost-distance models to directed movements within the home range of a non-migratory species, and then predict a corridor between two habitat areas on either side of a highway, or between discrete habitat patches within a home range or in a metapopulation. For predicting corridors for migrations with little or no existing GPS collar data, start and end locations can be generated from seasonal habitat selection models, or from data linking together specific seasonal ranges, such as mark–recapture data (e.g. from ear tags) coming from surveys of seasonal ranges. The approach's predictive ability enables corridor modelling for conservation planning in locations where insufficient data are available for animal-defined empirical corridors, can identify sites for the restoration of corridors that have been blocked, and can assess the effects of management actions or past and future vegetation, land use and climate change on corridors.

AUTHOR CONTRIBUTIONS

Tristan A. Nuñez, Mark A. Hurley and Matthew J. Kauffman conceived the idea and Tristan A. Nuñez, Matthew J. Kauffman and Tabitha A. Graves designed the methodology; Mark A. Hurley, Matthew J. Kauffman, Jerod A. Merkle, Anna C. Ortega and Hall Sawyer collected or contributed movement and spatial covariate data; Tristan A. Nuñez led the analysis and writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

Funding from U.S. DOI Secretarial Order 3362, Idaho Department of Fish and Game and Federal Aid in Wildlife Restoration Grant number W-160-R-37 supported this work. We also greatly appreciate the work of field biologists at the Idaho Department of Fish and Game and Wyoming Game and Fish Department in collecting the movement data, the work of J. Berg, S. Bergen and B. Oates in managing migration data and identifying migration endpoints, and useful feedback from two anonymous reviewers, P. Cross, A. Buerkle, B. Robb and participating state agency partners in the Corridor Mapping Team facilitated by the Wyoming Cooperative Research Unit. Any use of trade, firm and product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The environmental data used are available from the online sources from which they were downloaded; see the 'Methods' section for citations. The Idaho mule deer movement data are considered sensitive and have not been released by the Idaho Department of Fish and Game. The Red Desert to Hoback migration corridors are available at https://doi.org/10.5066/P9O2YM6I (Kauffman et al., 2020). The model fitting R code, including a vignette illustrating this approach, is archived at https://doi.org/10.5281/zenodo.6828966 (Nuñez, 2022).

ORCID

Tristan A. Nuñez https://orcid.org/0000-0002-1776-1442
Tabitha A. Graves https://orcid.org/0000-0001-5145-2400
Anna C. Ortega https://orcid.org/0000-0001-6750-6231
Hall Sawyer https://orcid.org/0000-0002-3789-7558
Julien Fattebert https://orcid.org/0000-0001-5510-6804
Jerod A. Merkle https://orcid.org/0000-0003-0100-1833
Matthew J. Kauffman https://orcid.org/0000-0003-0127-3900

REFERENCES

Aarts, G., Fieberg, J., & Matthiopoulos, J. (2012). Comparative interpretation of count, presence-absence and point methods for species distribution models: *Species distribution as spatial point process. Methods in Ecology and Evolution*, *3*(1), 177–187. https://doi.org/10.1111/j.2041-210X.2011.00141.x

Abrahms, B., Sawyer, S. C., Jordan, N. R., McNutt, J. W., Wilson, A. M., & Brashares, J. S. (2017). Does wildlife resource selection accurately inform corridor conservation? *Journal of Applied Ecology*, *54*(2), 412–422. https://doi.org/10.1111/1365-2664.12714

Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H., & Matthysen, E. (2003). The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning*, *64*(4), 233–247. https://doi.org/10.1016/S0169-2046(02)00242-6

Aikens, E. O., Kauffman, M. J., Merkle, J. A., Dwinnell, S. P. H., Fralick, G. L., & Monteith, K. L. (2017). The greenscape shapes surfing of resource waves in a large migratory herbivore. *Ecology Letters*, *20*(6), 741–750. https://doi.org/10.1111/ele.12772

Avgar, T., Potts, J. R., Lewis, M. A., & Boyce, M. S. (2016). Integrated step selection analysis: Bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution*, *7*(5), 619–630. https://doi.org/10.1111/2041-210X.12528

Beier, P., Majka, D. R., & Newell, S. L. (2009). Uncertainty analysis of least-cost modeling for designing wildlife linkages. *Ecological Applications*, *19*(8), 2067–2077. https://doi.org/10.1890/08-1898.1

Beier, P., Majka, D. R., & Spencer, W. D. (2008). Forks in the road: Choices in procedures for designing wildland linkages: *Design of Wildlife Linkages*. *Conservation Biology*, *22*(4), 836–851. https://doi.org/10.1111/j.1523-1739.2008.00942.x

Bishop, C. J., Unsworth, J. W., & Garton, E. O. (2005). Mule deer survival among adjacent populations in Southwest Idaho. *Journal of Wildlife Management*, *69*(1), 311–321. https://doi.org/10.2193/0022-541X(2005)069%3C0311:MDSAAP%3E2.0.CO;2

Bond, M. L., Bradley, C. M., Kiffner, C., Morrison, T. A., & Lee, D. E. (2017). A multi-method approach to delineate and validate migratory corridors. *Landscape Ecology*, *32*(8), 1705–1721. https://doi.org/10.1007/s10980-017-0537-4

Boyce, M. S. (2006). Scale for resource selection functions. *Diversity and Distributions*, *12*(3), 269–276. https://doi.org/10.1111/j.1366-9516.2006.00243.x

Brennan, A., Hanks, E. M., Merkle, J. A., Cole, E. K., Dewey, S. R., Courtemanch, A. B., & Cross, P. C. (2018). Examining speed versus selection in connectivity models using elk migration as an example. *Landscape Ecology*, *33*, 955–968. https://doi.org/10.1007/s10980-018-0642-z

Calabrese, J. M., Fleming, C. H., & Gurarie, E. (2016). Ctmm: An R package for analyzing animal relocation data as a continuous-time stochastic process. *Methods in Ecology and Evolution*, *7*(9), 1124–1132. https://doi.org/10.1111/2041-210X.12559

Doherty, T. S., Hays, G. C., & Driscoll, D. A. (2021). Human disturbance causes widespread disruption of animal movement. *Nature Ecology & Evolution*, *5*(4), 513–519. https://doi.org/10.1038/s41559-020-01380-1

Fieberg, J., Matthiopoulos, J., Hebblewhite, M., Boyce, M. S., & Frair, J. L. (2010). Correlation and studies of habitat selection: Problem, red herring or opportunity? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1550), 2233–2244. https://doi.org/10.1098/rstb.2010.0079

Graves, T., Chandler, R. B., Royle, J. A., Beier, P., & Kendall, K. C. (2014). Estimating landscape resistance to dispersal. *Landscape Ecology*, *29*(7), 1201–1211. https://doi.org/10.1007/s10980-014-0056-5

Gurarie, E., Bracis, C., Delgado, M., Meckley, T. D., Kojola, I., & Wagner, C. M. (2016). What is the animal doing? Tools for exploring behavioural structure in animal movements. *Journal of Animal Ecology*, *85*(1), 69–84. https://doi.org/10.1111/1365-2656.12379

Hanks, E. M., & Hooten, M. B. (2013). Circuit theory and model-based inference for landscape connectivity. *Journal of the American Statistical Association*, *108*(501), 22–33. https://doi.org/10.1080/01621459.2012.724647

Hijmans, R. J., & van Etten, J. (2012). *Raster: Geographic analysis and modeling with raster data* (R package version 1.2–2) [Computer software]. http://CRAN.R-project.org/package=raster

Johnson, D. S., Hooten, M. B., & Kuhn, C. E. (2013). Estimating animal resource selection from telemetry data using point process models. *Journal of Animal Ecology*, *82*(6), 1155–1164. https://doi.org/10.1111/1365-2656.12087

Kauffman, M. J., Copeland, H. E., Cole, E., Cuzzocreo, M., Dewey, S., Fattebert, J., Gagnon, J., Gelzer, E., Graves, T. A., Hersey, K., Kaiser, R., Meacham, J., Merkle, J., Middleton, A., Nunez, T., Oates, B., Olson, D., Olson, L., Sawyer, H., … Thonhoff, M. (2020). *Ungulate migrations of the western United States, volume 1: U.S. Geological Survey data release*. US Geological Survey. https://doi.org/10.5066/P9O2YM6I

Keeley, A. T. H., Beier, P., Keeley, B. W., & Fagan, M. E. (2017). Habitat suitability is a poor proxy for landscape connectivity during dispersal and mating movements. *Landscape and Urban Planning*, *161*, 90–102. https://doi.org/10.1016/j.landurbplan.2017.01.007

LaPoint, S., Gallery, P., Wikelski, M., & Kays, R. (2013). Animal behavior, cost-based corridor models, and real corridors. *Landscape Ecology*, *28*(8), 1615–1630. https://doi.org/10.1007/s10980-013-9910-0

Leu, M., Hanser, S. E., & Knick, S. T. (2008). The human footprint in the West: A large-scale analysis of anthropogenic impacts. *Ecological Applications*, *18*(5), 1119–1139. https://doi.org/10.1890/07-0480.1

McClure, M. L., Hansen, A. J., & Inman, R. M. (2016). Connecting models to movements: Testing connectivity model predictions against empirical migration and dispersal data. *Landscape Ecology*, *31*(7), 1419–1432. https://doi.org/10.1007/s10980-016-0347-0

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, *89*(10), 2712–2724. https://doi.org/10.1890/07-1861.1

Merkle, J. A., Monteith, K. L., Aikens, E. O., Hayes, M. M., Hersey, K. R., Middleton, A. D., Oates, B. A., Sawyer, H., Scurlock, B. M., & Kauffman, M. J. (2016). Large herbivores surf waves of green-up during spring. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1833), 20160456. https://doi.org/10.1098/rspb.2016.0456

Merkle, J. A., Sawyer, H., Monteith, K. L., Dwinnell, S. P. H., Fralick, G. L., & Kauffman, M. J. (2019). Spatial memory shapes migration and its benefits: Evidence from a large herbivore. *Ecology Letters*, *22*(11), 1797–1805. https://doi.org/10.1111/ele.13362

Northrup, J. M., Wal, E. V., Bonar, M., Fieberg, J., Laforge, M. P., Leclerc, M., Prokopenko, C. M., & Gerber, B. D. (2021). Conceptual and methodological advances in habitat-selection modeling: Guidelines for ecology and evolution. *Ecological Applications*, *32*(1), e02470. https://doi.org/10.1002/eap.2470

Nuñez, T. (2022). *Fitcorridor Intial Release* (v0.1.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.6828967

O'Leary, D., III, Hall, D. K., Medler, M., Matthews, R., & Flower, A. (2017). *Snowmelt timing maps derived from MODIS for North America, 2001-2015*. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1504

Oloo, F., Safi, K., & Aryal, J. (2018). Predicting migratory corridors of white storks, *Ciconia ciconia*, to enhance sustainable wind energy planning: A data-driven agent-based model. *Sustainability*, *10*(5), 1470. https://doi.org/10.3390/su10051470

Peterman, W. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution*, *9*(6), 1638–1647. https://doi.org/10.1111/2041-210X.12984

R Core Team. (2020). *R: A language and environment for statistical computing (4.0.0) [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Sawyer, H., Kauffman, M. J., Nielson, R. M., & Horne, J. S. (2009). Identifying and prioritizing ungulate migration routes for landscape-level conservation. *Ecological Applications*, *19*(8), 2016–2025. https://doi.org/10.1890/08-2034.1

Sawyer, H., Korfanta, N. M., Nielson, R. M., Monteith, K. L., & Strickland, D. (2017). Mule deer and energy development-long-term trends of habituation and abundance. *Global Change Biology*, *23*(11), 4521–4529. https://doi.org/10.1111/gcb.13711

Sawyer, H., Middleton, A. D., Hayes, M. M., Kauffman, M. J., & Monteith, K. L. (2016). The extra mile: Ungulate migration distance alters the use of seasonal range and exposure to anthropogenic risk. *Ecosphere*, *7*(10), e01534. https://doi.org/10.1002/ecs2.1534

Short Bull, R. A., Cushman, S. A., Mace, R., Chilton, T., Kendall, K. C., Landguth, E. L., Schwartz, M. K., Mckelvey, K., Allendorf, F. W., & Luikart, G. (2011). Why replication is important in landscape genetics: American black bear in the Rocky

Mountains. *Molecular Ecology*, *20*(6), 1092–1107. https://doi.org/10.1111/j.1365-294X.2010.04944.x

Tredennick, A. T., Hooker, G., Ellner, S. P., & Adler, P. B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, *102*(6), e03336. https://doi.org/10.1002/ecy.3336

Tucker, M. A., Böhning-Gaese, K., Fagan, W. F., Fryxell, J. M., Van Moorter, B., Alberts, S. C., Ali, A. H., Allen, A. M., Attias, N., Avgar, T., Bartlam-Brooks, H., Bayarbaatar, B., Belant, J. L., Bertassoni, A., Beyer, D., Bidner, L., van Beest, F. M., Blake, S., Blaum, N., … Mueller, T. (2018). Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science*, *359*(6374), 466–469. https://doi.org/10.1126/science.aam9712

van Etten, J. (2018). *gDistance: Distances and Routes on Geographical Grids* [R package version 1.2–2]. https://CRAN.R-project.org/package=gdistance

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*(16), 4010–4023. https://doi.org/10.1111/j.1365-294X.2012.05687.x

Wyckoff, T. B., Sawyer, H., Albeke, S. E., Garman, S. L., & Kauffman, M. J. (2018). Evaluating the influence of energy and residential development on the migratory behavior of mule deer. *Ecosphere*, *9*(2), e02113. https://doi.org/10.1002/ecs2.2113

Young, S. M. (2017). *U.S. Geological Survey shrub/grass products provide new approach to shrubland monitoring* (Report No. 2017–3084; Fact Sheet, p. 4). USGS Publications Warehouse. https://doi.org/10.3133/fs20173084

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*(7), 868–879. https://doi.org/10.1111/ddi.12742

Zeller, K. A., McGarigal, K., Beier, P., Cushman, S. A., Vickers, T. W., & Boyce, W. M. (2014). Sensitivity of landscape resistance estimates based on point selection functions to scale and behavioral state: Pumas as a case study. *Landscape Ecology*, *29*(3), 541–557. https://doi.org/10.1007/s10980-014-9991-4

Zeller, K. A., McGarigal, K., Cushman, S. A., Beier, P., Vickers, T. W., & Boyce, W. M. (2016). Using step and path selection functions for estimating resistance to movement: Pumas as a case study. *Landscape Ecology*, *17*, 1319–1335. https://doi.org/10.1007/s10980-015-0301-6

Zeller, K. A., McGarigal, K., & Whiteley, A. R. (2012). Estimating landscape resistance to movement: A review. *Landscape Ecology*, *22*, 777–797. https://doi.org/10.1007/s10980-012-9737-0

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: Nuñez, T. A., Hurley, M. A., Graves, T. A., Ortega, A. C., Sawyer, H., Fattebert, J., Merkle, J. A., & Kauffman, M. J. (2022). A statistical framework for modelling migration corridors. *Methods in Ecology and Evolution*, *00*, 1–14. https://doi.org/10.1111/2041-210X.13969