Landscape Ecol
DOI 10.1007/s10980-015-0214-4

RESEARCH ARTICLE

Empirical validation of landscape resistance models: insights from the Greater Sage-Grouse (*Centrocercus urophasianus*)

**Andrew J. Shirk · Michael A. Schroeder ·
Leslie A. Robb · Samuel A. Cushman**

Received: 12 September 2014 / Accepted: 9 May 2015

Abstract

Context The ability of landscapes to impede species' movement or gene flow may be quantified by resistance models. Few studies have assessed the performance of resistance models parameterized by expert opinion. In addition, resistance models differ in terms of spatial and thematic resolution as well as their focus on the ecology of a particular species or more generally on the degree of human modification of the landscape (i.e. landscape integrity). The effect of these design decisions on model accuracy is poorly understood.

Objectives We sought to understand the influence of expert parameterization, resolution, and specificity (i.e. species-specific or landscape integrity) on the fit of resistance model predictions to empirical landscape patterns.

Methods With genetic and observational data collected from Greater Sage-Grouse (*Centrocercus urophasianus*) in Washington State, USA, we used landscape genetic analysis and logistic regression to evaluate a range of resistance models in terms of their ability to predict empirical patterns of genetic differentiation and lek occupancy.

Results We found that species-specific, fine resolution resistance models generally had stronger relationships to empirical patterns than coarse resolution or landscape integrity models, and that the expert models were less predictive than alternative parameterizations.

Conclusions Our study offers an empirical framework to validate expert resistance models, suggests the need to match the grain of the data to the scale at which the species responds to landscape heterogeneity, and underscores the limitations of landscape integrity models when the species under study does not meet their assumptions.

Keywords *Centrocercus urophasianus* · Greater Sage-Grouse · Landscape genetics · Lek · Resistance · Validation

---

Electronic supplementary material The online version of this article (doi:10.1007/s10980-015-0214-4) contains supplementary material, which is available to authorized users.

A. J. Shirk (✉)
University of Washington Climate Impacts Group,
P.O. Box 355672, Seattle, WA 98195, USA
e-mail: ashirk@uw.edu

M. A. Schroeder
Washington Department of Fish and Wildlife,
P.O. Box 1077, Bridgeport, WA 98813, USA
e-mail: michael.schroeder@dfw.wa.gov

L. A. Robb
P.O. Box 1077, Bridgeport, WA 98813, USA
e-mail: robblar@homenetnw.net

S. A. Cushman
USDA Forest Service, Rocky Mountain Research Station,
2500 S. Pine Knoll Drive, Flagstaff, AZ 86001, USA
e-mail: scushman@fs.fed.us

Introduction

Landscapes resist the movement of individuals and thereby impose demographic and genetic isolation within populations (Wright 1943; McRae 2006). Limited isolation may be beneficial over evolutionary time scales because it creates potential for local adaptations and speciation (Dobzhansky 1940). Over shorter time scales, however, populations isolated by strong barriers experience reduced heterozygosity due to inbreeding, lower allelic diversity due to genetic drift, and increased vulnerability to stochastic events like extreme weather or disease outbreaks (Frankham, Briscoe and Ballou 2002). For these reasons, the extirpation risk of small isolated populations is high. Understanding how landscapes impose isolation on populations is therefore a primary focus of landscape ecology and conservation biology.

A landscape's influence on demographic and genetic isolation may be represented by resistance models, which quantify how landscape features affect movement or gene flow. Resistance models vary in terms of their spatial resolution, number and complexity of landscape features represented (i.e., thematic resolution), and how resistance is assigned to features (Spear et al. 2010). Resistance model parameters may be designed to reflect specific responses of a species (or groups of species with similar ecology) to landscape heterogeneity (e.g. Shirk et al. 2010). Alternatively, landscape integrity (LI) models attempt to quantify connectivity patterns based on anthropogenic modifications to the landscape (Baldwin et al. 2012). Importantly, LI models assume natural landscape variability does not resist movement and all anthropogenic modifications impose isolation on populations. Both species-specific and LI resistance models are now widely used to predict rates of movement and gene flow, yet no study has explored the relative predictive performance of these two approaches.

Many species threatened by extinction or range contractions require immediate conservation action, but the time and resources required to empirically parameterize resistance models to inform connectivity assessments are deemed prohibitive. Instead, species experts often parameterize resistance models. However, few of these expert-opinion models have been subsequently tested with empirical analysis (but see Shirk et al. 2010; Castillo et al. 2014), raising the possibility they may not optimally inform (or even misinform) conservation planning. For this reason, studies that quantify the ability of expert models to adequately reflect empirically observed relationships are needed.

In addition to poorly understood tradeoffs between species-specific and LI models, or expert versus empirically derived models, few studies have explored the influence of spatial and thematic resolution on resistance model accuracy (but see Galpern and Manseau 2013). For studies conducted over broad extents, computational requirements and limited availability of fine resolution data often necessitate resistance models based on spatial data layers of low spatial and thematic resolution. Conversely, studies over limited extents often make use of fine resolution data more tractable. If the species under study responds to landscape heterogeneity at a grain that differs from the grain of the resistance model, a suboptimal model may be identified that does not capture the true connectivity patterns on the landscape and thereby misinform conservation planning (Cushman and McGarigal 2002).

Recent connectivity analyses of Greater Sage-Grouse (*Centrocercus urophasianus*) in the Columbia Plateau, Washington, USA, provide an opportunity to evaluate the relative performance of species-specific versus LI resistance models, parameterizations based on expert opinion or empirical data, and the influence of resolution (both spatial and thematic) on model accuracy. Sage-grouse are a species of conservation concern range-wide, and recovery efforts in the Columbia Plateau are currently focused on increasing habitat suitability, connectivity, and establishing new populations. Three resistance models parameterized by expert opinion (coarse resolution sage-grouse, fine resolution sage-grouse, and fine resolution LI) have been used to inform these conservation efforts, but their ability to predict empirical rates of movement and gene flow has not been assessed.

In this study, we sought to validate expert sage-grouse resistance models with empirical data and compare alternative approaches to modeling landscape resistance that differed by parameterization as well as by the spatial and thematic resolution of input data layers. Specifically, we had four objectives: (1) use empirical landscape patterns derived from genetic and lek occupancy data to evaluate the expert coarse resolution (100 m; low thematic resolution) and fine resolution (30 m; high thematic resolution) sage-grouse resistance models as well as the LI model; (2) validate the expert sage-grouse resistance models by exploring alternative parameterizations focused on components of resistance with the greatest expert uncertainty; (3) ascertain whether greater spatial and thematic resolution improved model predictions; and (4) compare the predictive power of species-specific models (both coarse and fine resolution) to LI models. Sage-grouse are thought to respond to fine-scale landscape heterogeneity and factors that resist movement are not limited to anthropogenic modification of the landscape. Therefore, we anticipated fine resolution, sage-grouse specific models would provide more accurate predictions than coarse resolution models or models based solely on LI.

Methods

Study area

All three resistance models under evaluation encompassed the same study area boundary, which included the historical and current range of Greater Sage-Grouse in Washington. The empirical data used in this analysis were collected over different extents within this study area, such that the genetic data extent was nested within the lek occupancy data extent (Fig. 1).

Most of the Columbia Plateau is a semi-arid desert. Precipitation varies from about 18 cm in the lowest elevations to more than 120 cm in the forested mountains near the periphery. Elevation ranges from 100 m to over 2400 m. Average temperature varies from 3 °C in winter to over 20 °C in summer at the lowest elevations, but is cooler in both seasons at higher elevations (Whiteman et al. 1994). This region was historically dominated by perennial grassland with an overstory of sagebrush (*Artemisia* spp.). Today, the lower elevations of the Columbia Plateau are dominated by irrigated agriculture, intermixed with small cities and towns. At higher elevations, dryland wheat fields predominate. At the periphery of the plateau, the sagebrush biome transitions to conifer forest. A vast network of local roads spans the region, with one interstate highway (I-90) and several other major highways bisecting the study area. The remaining sagebrush habitat is generally limited to scablands where the soil is too shallow and rocky for agriculture. A notable exception occurs in dryland wheat fields that have been replanted with native vegetation as part of the federal Conservation Reserve Program (CRP) or related projects.

Resistance models

We evaluated three expert resistance models in terms of their ability to predict rates of sage-grouse gene flow and lek occupancy. These included a coarse resolution (both spatial and thematic) expert sage-grouse-specific model (CE), a fine resolution expert sage-grouse-specific model (FE), and a fine resolution landscape integrity model (LI). The CE model was derived by expert assignment of resistance values to different pixel types in several categorical raster data layers representing land cover, elevation, slope, interstate highways, major highways, secondary highways, local roads, and housing density. These model input rasters were derived from source data produced between 2000 and 2008, and resampled to 100 m resolution (sources and methods are described in WHCWG 2010). Resistance values assigned to each of these layers ranged from 0 to 1000 (Supplemental Table 1, CE is model 2). Parameterization was based on considerable local expert knowledge of sage-grouse ecology, the Columbia Plateau landscape, and limited empirical analysis (WHCWG 2010). Reclassifying each spatial data input layer based on these resistance parameters, summing them, and adding one to all grid cells produced the final CE resistance model with values ranging from 1 to 2088.

The FE and LI model inputs and parameterization were described in a previous analysis by the same group (WHCWG 2012). The FE model was derived in the same way as the CE expert model, but using spatial data input layers with greater spatial resolution (30 m grid cell size) and thematic resolution (e.g., 30 land cover classes compared to 11 in the CE model). In addition, the FE model included additional data layers representing wind turbines, transmission lines, canals, and railroads that were not used in the CE model. The LI model was based on the same fine-scale spatial data layers as the FE model, but resistance was only assigned to categories associated with human modification of the landscape. These included roads, agriculture, urban areas, population density, transmission lines, wind turbines, canals, and railroads. All fine-scale input raster layers were derived from source data

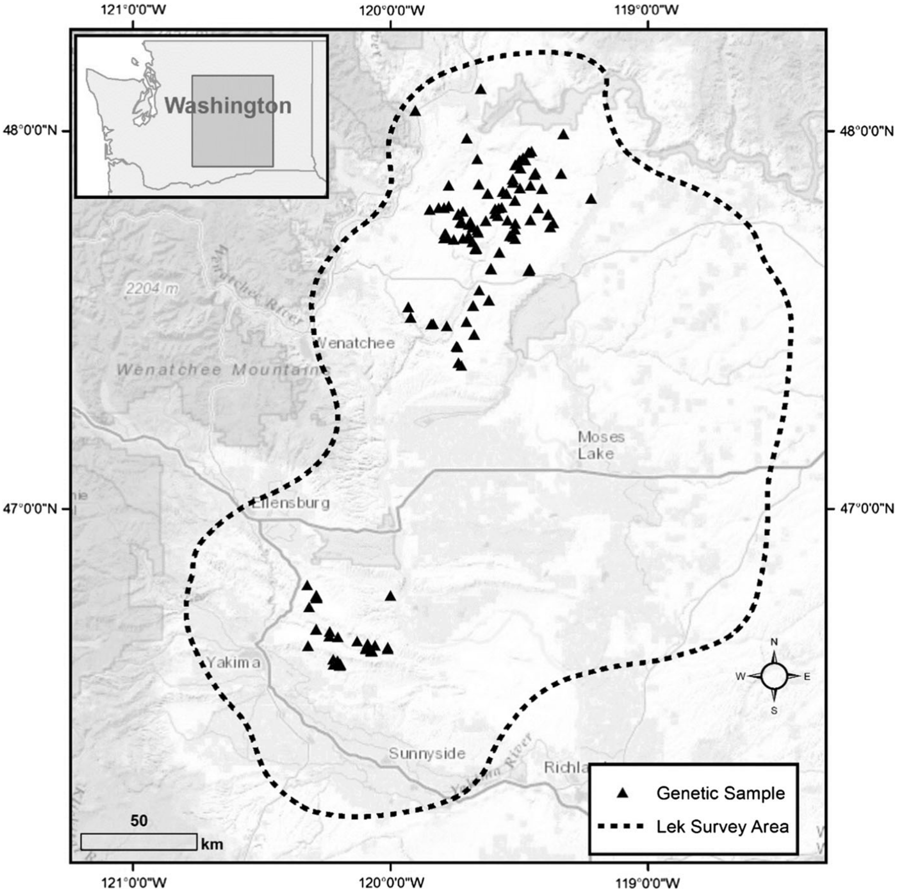


Fig. 1 The study area in the Columbia Plateau of eastern Washington State, USA, including locations of Greater Sage-Grouse genetic samples (*black triangles*) and the extent of lek surveys (*black dashed line*; lek locations not shown due to sensitivity restrictions)

produced between 2000 and 2010 (sources and methods are described in WHCWG 2012). The FE and LI model parameters are provided in Supplemental Table 2 (FE is model 53 and LI is model 82).

We produced alternative sage-grouse resistance models at both coarse and fine resolution using the same spatial input data and approach as described above for the expert models. The alternative models were parameterized by the same biologists that parameterized the three expert models. These biologists were initially asked which CE and FE model parameters they were most uncertain about, and then asked to produce alternative models that varied the uncertain parameters alone or in combination with other uncertain parameters. In most cases, alternative resistance values for these uncertain parameters were either set to zero, doubled from the original value, or quadrupled. In total, we produced 49 alternative models using coarse resolution data (Supplemental Table 1) and 28 alternative models using fine resolution data (Supplemental Table 2). As a null hypothesis, we also evaluated an alternative resistance model that assigned a resistance of one to all cells. This reflects the concept of isolation by distance (Wright 1943), where movement or gene flow between individuals is a function of the Euclidean distance between them, rather than gradients of resistance arising from landscape heterogeneity.

Landscape genetic analysis

Genetic samples were obtained from blood, feathers, egg shells, or remains at mortality sites (predations or accidental deaths) from 1992–2011 throughout the occupied portion of the study area. In the Douglas county population (the northern cluster of genetic samples, Fig. 1), genetic sampling occurred from 1992 to 2010 ($N = 218$). In the Yakima Training Center population (the southern cluster of genetic samples, Fig. 1), genetic sampling occurred from 1999–2011 ($N = 107$). Blood samples were collected from sage-grouse trapped for a radio telemetry study. Feathers, egg shells, and remains were collected from leks and nests. In total, we collected 325 genetic samples.

It is possible genetic differentiation in the Douglas County population could be driven by drift (rather than the landscape, as assumed in our analysis) occurring over the relatively long period of sample collection (18 years, or about six generations). To assess the relative degree of temporal and spatial differentiation in the population, we calculated pairwise F_{ST} values among genetic samples divided into two temporal groups (Douglas County samples collected during the 1990s and the 2000s) and two spatial groups (Douglas County and the Yakima Training Center). We calculated F_{ST} values with the adegenet package (Jombart et al. 2008) in R (R Development Core Team 2013).

DNA extraction and genotyping was performed by the Washington Department of Fish and Wildlife Molecular Genetics Laboratory, Olympia, Washington. Blood and egg shell samples were extracted using the Qiagen DNeasy Kit. Feather samples were extracted using QIAamp Micro kit (Qiagen). We used polymerase chain reaction to amplify 19 microsatellite loci chosen based on successful amplification in other studies (marker names and references are listed in Supplemental Table 3). PCR products were visualized using an ABI-3730 DNA Analyzer with internal size standards (GS500LIZ 3730) and scored using GeneMapper 3.7 software.

We used a causal modeling approach (Cushman et al. 2006; Cushman and Landguth 2010b; Shirk et al. 2010; Cushman et al. 2013a, b) to evaluate the relative support for each resistance model based on the relationship between pairwise individual genetic distances (calculated from the microsatellite genotypes) and effective landscape distances (calculated as cost-weighted distances along the least-cost path, given the resistance hypothesis). We used the adegenet package (Jombart et al. 2008) in R (R Development Core Team 2013) to quantify pairwise individual genetic distances (as a matrix) based on the proportion of shared alleles between all pairs of sampled individuals (a value of 1 would indicate identical alleles and a value of 0 would indicate no shared alleles). To quantify landscape distances for each resistance model, we used ArcGIS 10.0 (ESRI 2008) to calculate matrices with values corresponding to the cost-weighted distance between individuals given the least cost path. We then compared models based on the Mantel correlation (Mantel 1967; Smouse, Long and Sokal 1986) between the genetic and landscape distance matrices while partialling out in turn the effect of each alternative model. A model was deemed causal if two criteria were met. First, it must retain a significant ($\alpha \leq 0.05$) correlation with genetic distance after partialling out the effect of all other models in turn, and second, all other models must have no significant correlation with genetic distance after partialling out the effect of the candidate model.

Recently, there has been some controversy regarding the use of partial Mantel tests for model selection in landscape genetics. Some studies contend have found this approach suffers from high rates of Type I error (Balkenhol et al. 2009; Guillot and Rouset 2013; Graves et al. 2013). However, other studies demonstrate partial Mantel tests are able to identify the causal model from spurious alternatives, even when correlations between models are high (Legendre and Fortin 2010; Cushman and Landguth 2010a, b; Diniz-Filho et al. 2013). The approach we describe above (based on Cushman et al. 2013a, b) is an evolution of the causal modeling with partial Mantel tests that postdates recent criticisms and is designed to reduce Type I error rates.

Lek occupancy analysis

Male sage-grouse congregate on lek sites in the spring to breed. Because lek sites tend to be traditional, occupancy of these locations reflects the broader distribution of sage-grouse. Lek locations can move slightly and satellites may occasionally form, therefore, we combined lek locations into “lek complexes” that accounted for this variability and decreased biases associated with sampling intensity (Schroeder et al. 2000). Lek surveys were conducted at 70 complexes across the region in 1954, 2009, and 2010 (Fig. 1; exact locations are sensitive information and are not shown). Each lek was visited at least three times per year, and the presence or absence of sage-grouse was recorded. Leks are generally located in open areas with high visibility. In addition, the male vocalizations and frequent movements during displays and courtship facilitate detection. For these reasons, the accuracy of determining presence or absence at leks was likely to be very high, though detection rates were not measured.

We evaluated resistance models in terms of their ability to predict whether a historically active lek (based on the 1954 survey) was still currently active in either of the most recent lek surveys in 2009 and 2010. A lek complex was deemed active if at least one sage-grouse was observed during the survey (53.8 % of leks remained active by this definition). The relationship between resistance and lek occupancy is based on two assumptions. First, an active lek complex requires a permeable landscape in adjacent areas so that females can move to nesting sites near leks; the larger the area available for nesting the greater the potential for the lek to be active. Second, leks isolated from the rest of the population by low landscape permeability experience reduced immigration and gene flow, lowering the probability of lek occupancy (Frankham et al. 2002).

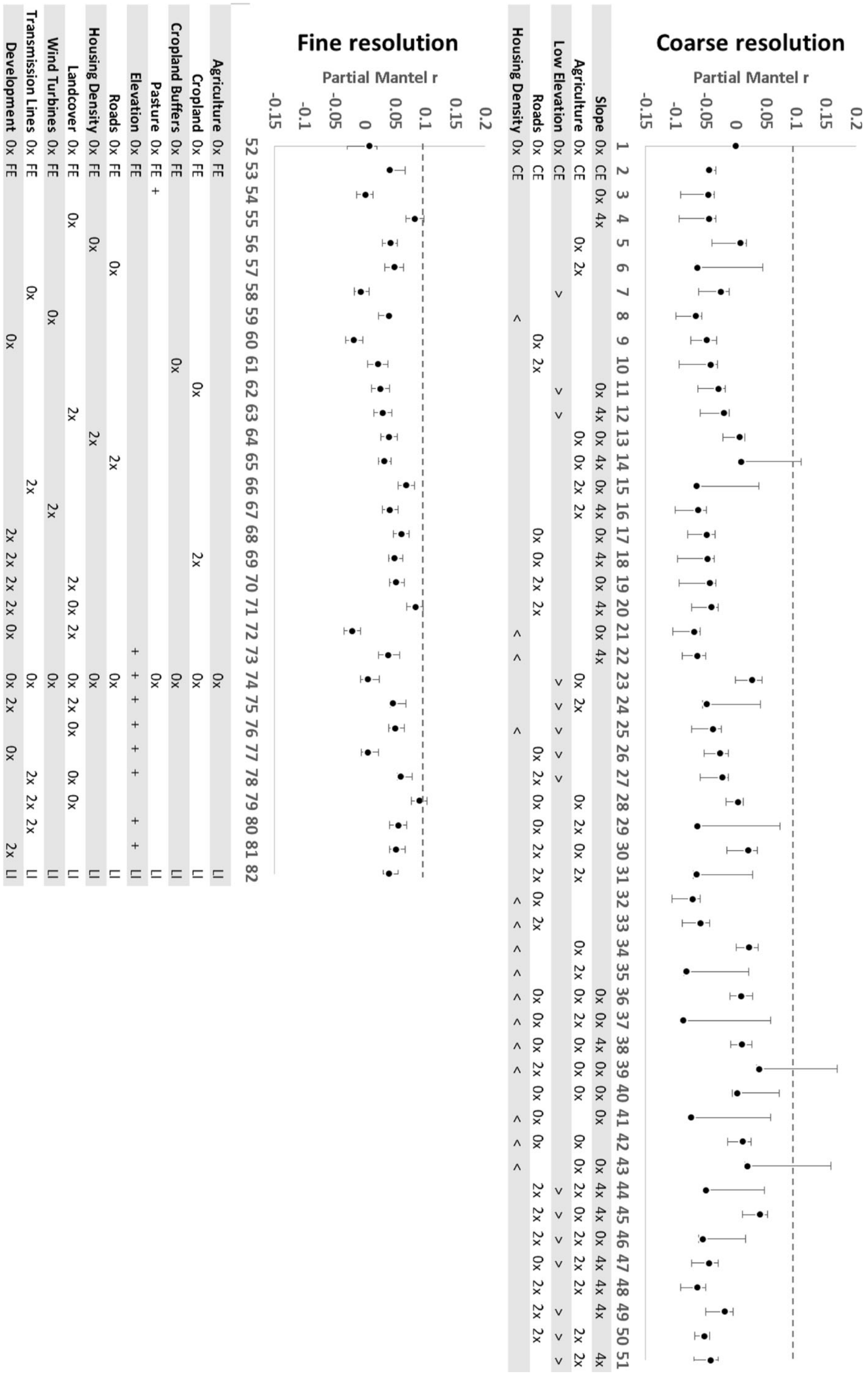
We quantified the available area surrounding each lek and the centrality of each lek within the broader population for each resistance hypothesis. First, we used the costdistance function in ArcGIS 10.0 (ESRI 2008) to calculate a cost-weighted distance surface (i.e. the cumulative cost to move from any given pixel to the nearest lek). We then calculated the 95th percentile of cost-weighted distance for those pixels corresponding to 323 nesting sites (observed throughout the study area from 1992–2010) proximate to leks occupied in the most recent surveys of 2009–2010. Calculating the area (km^2) surrounding each lek based on a boundary drawn at the 95th percentile of nest site cost-weighted distance provided an estimate of the area available for nesting around each lek, given the resistance hypothesis under evaluation. To quantify the centrality of each lek, we calculated the mean cost-weighted distance to all other leks active in the 1954 survey, given the resistance model under evaluation. The lower the mean cost-weighted distance, the more central the lek was in the network of leks that comprise the population.

We modeled the relationship between lek occupancy, available nesting habitat area, and centrality using logistic regression in R (R Development Core Team 2013). For each resistance hypothesis, we fit three models based on available nesting habitat area only, centrality only, or both available nesting habitat area and centrality. We then compared all models based on the difference in Akaike information criterion (AIC; Burnham and Anderson 2002).

Fig. 2 The Mantel correlation between genetic and landscape ▶ distances for Greater Sage-Grouse in Washington State, while partialling out the effect of distance, including 95 % confidence intervals. Higher correlations indicate a stronger relationship between the resistance hypothesis and the empirical pattern of genetic isolation. The *dashed line* serves as a baseline representing the highest correlation among the models evaluated. The models are numbered 1 through 82, and split between coarse resolution models (1–51) in the *top panel* and fine resolution models (52–82) in the *bottom panel*. Models 1 and 52 are null models (i.e., the resistance of all cells equals one, and therefore predict isolation is a function of distance alone, rather than resistance due to landscape heterogeneity) at coarse and fine resolutions, respectively. Models 2 and 53 are the coarse resolution expert (CE) and fine resolution expert (FE) resistance models, respectively. The alternative models differ from the expert models at the corresponding resolution in ways listed below each plot. The notations 0×, 2×, and 4× represents eliminating, doubling, or quadrupling resistance values for a particular spatial data input, respectively. The *lesser than* and *greater than* symbols indicate a reduction or increase, respectively in values for a particular spatial data input. A *plus* symbol indicates resistance was assigned to a layer not used in the expert model. Model 82 was parameterized based on the concept of landscape integrity (i.e., resistance was only assigned to attributes reflecting human modification of the landscape)

Model comparison

To spatially quantify the difference in resistance values between the expert models and empirically supported alternative model (identified based on the synthesis of the landscape genetic and lek occupancy analyses), we subtracted the empirically supported alternative resistance model raster from each expert resistance model (CE, FE, and LI). In the resulting rasters, positive values represented areas where resistance was underestimated by the expert model relative to the empirically supported model and negative values indicated resistance was overestimated. We also compared the expert models (CE, FE, and LI) to the empirically supported model based on their predictions of connectivity between leks. Specifically, for each expert model and the empirically supported alternative model, we used ArcGIS (ESRI 2008) to calculate the cost-weighted distance to the nearest lek. We then subtracted the cost-distance raster corresponding to the



empirically supported model from each of the three expert models. Positive values represented areas where the expert model underestimated connectivity (as measured by cost-weighted distance) to the nearest lek and negative values indicated the expert model overestimated connectivity.

Results

Landscape genetics

The F_{ST} value comparing genetic samples collected in the 1990s to those collected in the 2000s was 0.006. The F_{ST} value comparing genetic samples collected from the YTC to those collected from Douglas County was 0.068. Thus, nearly all of the genetic differentiation in the population was due to spatial structuring rather than temporal differences arising from the 18 year sampling period.

None of the coarse resolution models (including the CE model or alternatives) retained a significant Mantel correlation after partialling out the variance explained by the null model (IBD), which only considered the effect of distance as opposed to the additional resistance imposed by landscape heterogeneity (Fig. 2). Among the fine resolution models, the FE and LI models did not have a significant relationship with the empirical pattern of genetic distances after partialling out the null model. Of the 28 alternative fine resolution models, 14 retained a significant Mantel correlation after partialling out the null model (models 55–57, 66, 68–71, 75, 76, and 78–81). However, only one of these met all the criteria of a causal model within our framework (model 79, Supplemental Table 2). Namely, after partialling out the effect of all other competing models, we still observed a significant partial Mantel correlation between genetic distance and landscape distance for model 79, yet the converse was not true. This fine resolution alternative model differed from the FE model in that it eliminated the resistance of landcover and doubled the resistance of transmission lines.

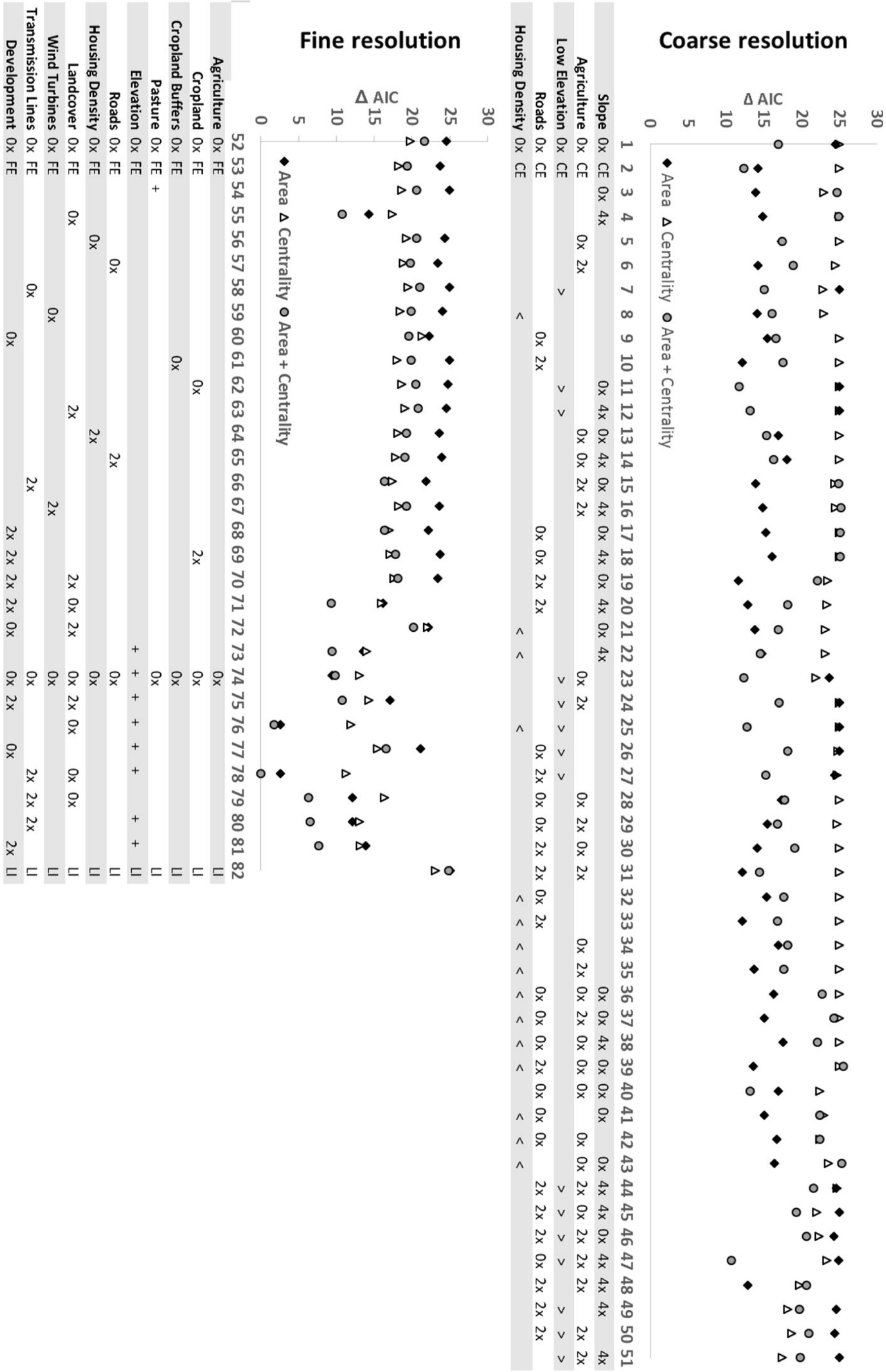
Lek activity

One fine resolution sage-grouse specific model (model 78) was clearly supported relative to all other models based on AIC scores (Fig. 3). This differed from the FE model in that it eliminated resistance due to landcover (also observed in the landscape genetic analysis), doubled the resistance of transmission lines (also observed in the landscape genetic analysis), and added resistance at low elevations. Relating this resistance model to patterns of lek occupancy by logistic regression using terms for both nesting habitat area and centrality produced the lowest AIC score (96.25), however, removing the term for centrality only modestly increased the AIC score (98.82). Including only the centrality term raised the AIC score substantially (107.55) suggesting the relationship was driven primarily by nesting habitat area rather than centrality. The coefficient was 4.8×10^{-7} ($SE = 1.5 \times 10^{-7}$) for the area term and -0.146 ($SE = 0.09$) for the centrality term, indicating larger area available for nesting near leks and lower average cost-distances to other leks predicts greater probability of activity. Model 78 appeared strongly predictive ($AUC = 0.82$, Cohen's $Kappa = 0.616$, classification $accuracy = 0.813$), sensitive ($sensitivity = 0.89$), and specific ($specificity = 0.722$).

Fig. 3 The difference in AIC scores (*delta* AIC) between the model with the lowest AIC and all other models is depicted, based on logistic regression relating lek occupancy to each resistance hypothesis. Three different regressions were performed for each resistance model, including one based on the resistance model's prediction of the area available for nesting adjacent to each lek (*black diamond*), the centrality of the lek within the network of leks (*white triangle*), and the combination of available nesting habitat area and centrality (*grey circle*). The models are numbered 1 through 82, and split between coarse resolution models (1–51) in the top panel and fine resolution models (52–82) in the bottom panel. Models 1 and 52 are null models (i.e., the resistance of all cells equals one, and therefore predict isolation is a function of distance alone, rather than resistance due to landscape heterogeneity) at coarse and fine resolutions, respectively. Models 2 and 53 are the coarse resolution expert (CE) and fine resolution expert (FE) resistance models, respectively. The alternative models differ from the expert models at the corresponding resolution in ways listed below each plot. The notations 0×, 2×, and 4× represents eliminating, doubling, or quadrupling resistance values for a particular spatial data input, respectively. The *lesser than* and *greater than* symbols indicate a reduction or increase, respectively in values for a particular spatial data input. A *plus* symbol indicates resistance was assigned to a layer not used in the expert model. Model 82 was parameterized based on the concept of landscape integrity (i.e., resistance was only assigned to attributes reflecting human modification of the landscape) ▶

Model comparison

The resistance of the FE model differed from the empirically supported model (#78) by as much as 39 units (Fig. 4, top left panel). Resistance was underestimated mainly in low elevations of the study area and along transmission line corridors and overestimated in higher elevations. The CE model showed a similar pattern of error as the FE model, but the magnitude of the difference was generally greater (up to 2035 units; Fig. 4, top middle panel). The comparison to the LI



model revealed the opposite pattern. Low-elevation areas dominated by anthropogenic landscape modifications vastly overestimated resistance (up to 935 units), while the less modified higher elevations underestimated resistance relative to the most-supported model (up to 35 units; Fig. 4, top right panel).

A similar pattern is apparent in the contrasts based on cost-weighted distance to the nearest lek. Both the FE and CE models underestimate connectivity in higher elevations and overestimate connectivity in lower elevations, and the degree of error is greater in the coarse resolution expert model (Fig. 4, bottom left and middle panels). The LI model differs in the opposite way by underestimating connectivity in lower elevations and overestimating connectivity in higher elevations (Fig. 4, bottom right panel).

Discussion

In this study, we used two independent empirical datasets to evaluate the ability of three expert resistance models (CE, FE, and LI) to predict patterns of Greater Sage-Grouse genetic differentiation and lek occupancy. We found statistical support for the expert models was weak compared to an alternative resistance model. Indeed, support for the expert models was equivocal compared to the null hypothesis that distance alone (as opposed to resistance) was the process driving the observed empirical patterns. Differences between the empirically supported alternative model and the expert models suggest transmission lines and low elevation have greater resistance than anticipated. Conversely, the expert models appeared to overestimate the resistance of landcover, particularly non-irrigated agriculture and CRP lands. In general, alternative resistance models based on fine resolution spatial data and tailored to the specific ecology of sage-grouse exhibited much stronger relationships to empirical patterns than models based on coarse resolution spatial data or LI parameters.

The two empirical model validation approaches yielded slightly different inferences regarding the most predictive model of genetic isolation or patterns of lek occupancy. However, these inferences may be synthesized by considering differences in the factors limiting sage-grouse movement or gene flow across the nested extents of our analysis. At the broad extent of the lek occupancy analysis, all major potential sources of resistance in the study area were present. That low elevation and not landcover appeared to drive resistance in the top model over this extent reflects the strong correlation between elevation and sage-grouse habitat quality in the Columbia Plateau. Indeed, elevation is likely a better predictor of resistance than landcover because the lowest elevations capture multiple interacting sources of resistance, including intensive irrigated agriculture (which is not suitable for sage-grouse), very low prevalence of CRP lands (which are suitable), and significant degradation of sagebrush habitat from grazing and the legacy effects of historical land use (neither of which were reflected in the landcover data available). Conversely, the higher elevations near the periphery of the plateau reflect the low resistance of non-irrigated agriculture (dryland wheat) and CRP fields, and the high quality sagebrush habitat there.

Over the more limited extent of the landscape genetic analysis, low elevations were rare. Even though low elevation habitats in this study area may be a strong barrier to gene flow, if they are not present in sufficient quantity in the area of genetic sampling, their resistance may not be observable. This fits with the concept of limiting factors in landscape genetics (Cushman et al. 2013a), which postulates that a landscape factor's resistance to gene flow can only be measured when it's prevalence and configuration in the landscape reaches a point where it forms a dispersal barrier. Thus, the inability to detect the resistance of low elevation could be interpreted as a lack of an ecological relationship, when in fact, it could be merely unobservable due to its rarity. Several recent studies have demonstrated the influence of limiting factors on the detectability of ecological relationships, including analyses of American marten (*Martes americana*) habitat selection (Shirk et al. 2012) and gene flow (Cushman et al. 2011), black bear (*Ursus americanus*) gene flow (Short Bull et al. 2011), and simulations in fractal landscapes (Cushman et al. 2013a). Our results suggest limiting factors should be considered when interpreting and synthesizing results of empirical validation studies. That the top model in the landscape genetic analysis was identical to the top model of the broad-scale lek analysis, minus the influence of elevation, is consistent with the latter being the best predictor among those we examined.

Our study also highlights the need to match the grain of the resistance model input data to the grain at which the species responds to landscape heterogeneity. Though data of coarse spatial and thematic resolution may accurately reflect resistance for some species, our results support our expectation that sage-grouse respond to landscape heterogeneity at a fine grain. Indeed, the top models in each of our empirical

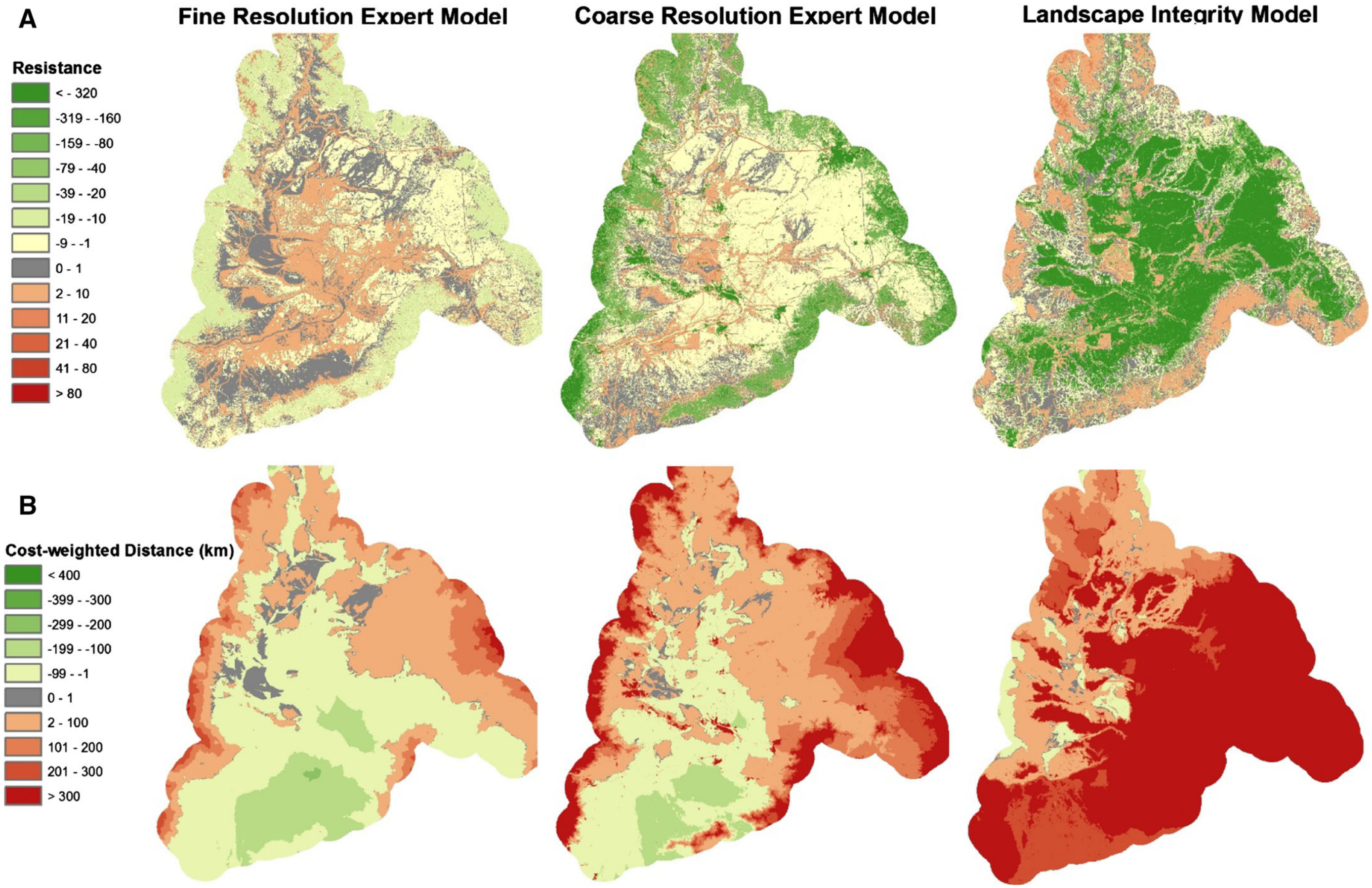


Fig. 4 The top row (**a**) presents the difference in resistance values between the empirically supported model (#78 from Figs. 2 and 3) and the fine resolution expert model, coarse resolution expert model, and landscape integrity models for Greater Sage-Grouse in Washington State. Negative values indicate an overestimation of resistance relative to the empirically supported model, while positive values indicate an underestimation of resistance. The bottom row (**b**) makes the same comparisons but based on the cost-weighted distance to the nearest lek that was active in a 1954 survey. After subtracting each of the three models from the empirically supported model, a negative result indicates underestimation of habitat connectivity while a positive result indicates an overestimation of habitat connectivity

analyses were generally based on fine resolution data. In particular, the ability to attribute resistance to specific types of agriculture and the inclusion of transmission lines in the fine resolution models allowed us to explore resistance models with more accurate predictions of genetic differentiation and lek occupancy compared to the coarse resolution models. This demonstrates how broad-scale connectivity analyses using coarse spatial data may not adequately capture landscape resistance for species that respond to finer resolution habitat attributes. Several recent studies have also demonstrated the importance of matching the resolution of the data and extent of analysis to the ecology of the species under study (e.g. Cushman and McGarigal 2004; Cushman and Landguth 2010a; Shirk et al. 2012). Our results suggest scale is an important consideration when designing and interpreting resistance model validation analyses.

The differences between the empirically supported resistance model and the three expert models in terms of resistance (Fig. 4a) and predicted connectivity (Fig. 4b) suggest caution in the use of expert models for connectivity analyses and conservation planning. The expert sage-grouse models failed to account for the magnitude of resistance due to transmission lines and low elevations, and overestimated the resistance due to certain landcover types, raising the potential to misinform conservation planning related to these sources of resistance. However, many of the parameters in the expert models remained the same (or nearly so) in the empirically derived top model. This affirms the value of expert models in reducing the vast potential hypothesis space to a tractable number of plausible alternative models, while grounding the analysis in a hypothesis testing framework. Other studies have used a similar strategy to improve upon expert models with similar success (e.g. Epps et al. 2007; Shirk et al. 2010). This framework, and our approaches for optimizing resistance models based on genetic and occurrence data, could be applied to other species where validation of expert models is desired.

Our study also demonstrates the limitations of using LI as a proxy for species-specific resistance model parameterization when the species of interest also responds strongly to natural variability in the landscape, or can move efficiently through some human modified habitats. Indeed, because sage-grouse movement and gene flow does not appear to be strongly reduced by dryland wheat and CRP fields in our study area, the LI model vastly underestimated rates of movement in the higher elevations of the Columbia Plateau, where these cover types predominate. The ability of LI models to reflect species' movement and gene flow is largely unexplored, yet LI models are now widely used in connectivity assessments. Understanding the relationship between predicted connectivity based on LI models and the actual connectivity of species is therefore a major research need.

The convergence of two independent empirical validation analyses on a single alternative model (once the confounding influence of limiting factors is taken into account) suggests the need to update previous assumptions about landscape factors that shape sage-grouse lek occupancy and gene flow, particularly with regard to transmission lines. Collisions with transmission lines have been observed as causes of mortality in our study area and elsewhere, and they also provide perches and nesting sites for raptors and corvids. The risk of mortality likely drives behavioral avoidance of transmission lines, with declining habitat use observed 100 m away (Connelly et al. 2011). Our analysis suggests greater attention should be focused on these landscape features in sage-grouse conservation planning, particularly in light of the planned expansion of the energy transmission network throughout much of the sage-grouse habitat in the western United States. In addition, the low resistance of agricultural fields planted to native vegetation as part of the CRP program supports the value of this and related efforts designed to improve habitat suitability and connectivity.

Expert resistance models are widely used to assess habitat connectivity, yet few are subsequently tested with empirical data. Our study underscores the value of empirical validation of expert models, to improve their power to predict habitat connectivity patterns on the landscape and thereby better inform conservation planning. Though our approach involved only a limited exploration of the limitless hypothesis space (centered on the expert model), it was able to clearly identify an alternative model that was able to superior in terms of its ability to predict patterns of differentiation and lek occupancy on the landscape. This model was based on fine-scale data and designed to reflect the ecology of sage-grouse specifically, rather than the concept of landscape integrity. This suggests the value of using fine-scale, species-specific data and parameterization when the species of interest responds to both natural and anthropogenic factors at a fine scale. For species that respond to broad-scale landscape patterns, fine resolution data may not be necessary. Similarly, species that can readily disperse through most habitat types except those that are heavily modified by human activity may be adequately represented by landscape integrity models. More studies are needed to fully characterize the accuracy of expert models in predicting rates of movement and gene flow, as well as the tradeoffs associated with resistance model resolution and parameterization approaches.

References

Baldwin RF, Reed SE, McRae BH, Theobald DM, Sutherland RW (2012) Connectivity restoration in large landscapes: modeling landscape condition and ecological flows. Ecol Restor 30:274–279

Balkenhol N, Waits LP, Dezzani RJ (2009) Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32(5):818–830

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York

Castillo JA, Epps CW, Davis AR, Cushman SA (2014) Landscape effects on gene flow for a climate-sensitive montane species, the American pika. Mol Ecol 23:843–856

Connelly J, Knick S, Braun C, Baker W, Beever E, Christiansen T, Doherty K, Garton E, Hanser S, Johnson D (2011) Conservation of Greater Sage-Grouse. Stud Avian Biol 38:549–646

Cushman SA, McGarigal K (2002) Hierarchical, multi-scale decomposition of species-environment relationships. Landscape Ecol 17(7):637–646

Cushman SA, McGarigal K (2004) Patterns in the species—environment relationship depend on both scale and choice of response variables. Oikos 105:117–124

Cushman SA, Landguth EL (2010a) Scale dependent inference in landscape genetics. Landscape Ecol 25:967–979

Cushman SA, Landguth EL (2010b) Spurious correlations and inference in landscape genetics. Mol Ecol 19:3592–3602

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. Am Nat 168:486–499

Cushman SA, Raphael MG, Ruggiero LF, Shirk AJ, Wasserman TN, O'Doherty EC (2011) Limiting factors and landscape connectivity: the American marten in the Rocky Mountains. Landscape Ecol 26:1137–1149

Cushman SA, Shirk AJ, Landguth EL (2013a) Landscape genetics and limiting factors. Conserv Genet 14(2):263–274

Cushman SA, Wasserman TN, Landguth EL, Shirk AJ (2013b) Re-evaluating causal modeling with mantel tests in landscape genetics. Diversity 5:51–72

Diniz-Filho JAF, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, Telles MPDC, Bini LM (2013) Mantel test in population genetics. Genet mol biol 36(4):475–485

Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. Am Nat 74:312–321

Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS (2007) Optimizing dispersal and corridor models using landscape genetics. J Appl Ecol 44:714–724

ESRI (2008) ArcGIS Desktop: release 10.0. Environmental Systems Research Institute, Redlands, CA

Frankham R, Briscoe DA, Ballou JD (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge

Galpern P, Manseau M (2013) Finding the functional grain: comparing methods for scaling resistance surfaces. Landscape Ecol 28:1269–1281

Graves TA, Beier P, Royle JA (2013) Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. Mol ecol 22(15): 3888–3903

Guillot G, Rousset F (2013) Dismantling the Mantel tests. Methods Ecol Evol 4(4):336–344

Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity 101:92–103

Legendre P, Fortin MJ (2010) Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Mol Ecol Resour 10(5):831–844

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

McRae BH (2006) Isolation by resistance. Evolution 60:1551–1561

R Development Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing. R Development Core Team, Vienna

Schroeder MA, Hays DW, Livingston MF, Stream LE, Jacobson JE, Pierce DJ (2000) Changes in the distribution and abundance of sage-grouse in Washington. Northwest Nat 81:104–112

Shirk AJ, Wallin DO, Cushman SA, Rice CG, Warheit KI (2010) Inferring landscape effects on gene flow: a new model selection framework. Mol Ecol 19:3603–3619

Shirk AJ, Wasserman TN, Cushman SA, Raphael MG (2012) Scale dependency of American marten (*Martes americana*) habitat relations. In: Aubry KB, Zielinski WJ, Raphael MG, Proulx G, Buskirk SW (eds) Biology and conservation of martens, sables, and new synthesis. Cornell University Press, Ithaca

Short Bull RAS, Cushman SA, Mace R, Chilton T, Kendall KC, Landguth EL, Schwartz MK, McKelvey K, Allendorf FW, Luikart G (2011) Why replication is important in landscape genetics: American black bear in the Rocky Mountains. Mol Ecol 20:1092–1107

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Syst Zool 35:627–632

Spear SF, Balkenhol N, Fortin MJ, McRae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Mol Ecol 19:3576–3591

WHCWG (2010) Washington connected landscapes project: analysis of the Columbia Plateau ecoregion. Washington Department of Fish and Wildlife and Washington Department of Transportation, Olympia

WHCWG (2012) Washington connected landscapes project: analysis of the Columbia Plateau ecoregion. Washington Department of Fish and Wildlife and Washington Department of Transportation, Olympia

Whiteman K, Vaccaro J, Gonthier J, Bauer H (1994) The hydrogeologic framework and geochemistry of the Columbia Plateau aquifer system. US Government Printing Office, Washington

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