



**HAL**  
open science

# Cost distances and least cost paths respond differently to cost scenario variations: a sensitivity analysis of ecological connectivity modeling

Paul Savary, Jean-Christophe Foltête, Stéphane Garnier

## ► To cite this version:

Paul Savary, Jean-Christophe Foltête, Stéphane Garnier. Cost distances and least cost paths respond differently to cost scenario variations: a sensitivity analysis of ecological connectivity modeling. *International Journal of Geographical Information Science*, 2022, 36 (8), pp.1652-1676. 10.1080/13658816.2021.2014852 . hal-03508121

**HAL Id: hal-03508121**

**<https://hal.science/hal-03508121>**

Submitted on 3 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Cost distances and least cost paths respond differently to cost scenario variations - a sensitivity analysis of ecological connectivity modeling

Savary, Paul<sup>\*1, 2, 3</sup>, Foltête, Jean-Christophe<sup>2</sup>, and Garnier, Stéphane<sup>3</sup>

<sup>1</sup>ARP-Astrance, 9 Avenue Percier, 75008 Paris, France

<sup>2</sup>ThéMA, UMR 6049 CNRS, Université Bourgogne-Franche-Comté, 32 Rue Mégevand, 25030 Besançon Cedex, France

<sup>3</sup>Biogéosciences, UMR 6282 CNRS, Université Bourgogne-Franche-Comté, 6 Boulevard Gabriel 21000 Dijon, France

## Abstract

Biodiversity conservation measures designed to ensure ecological connectivity depend on the reliable modeling of species movements. Least cost path modeling makes it possible to identify the most likely dispersal paths within a landscape and provide two items of ecological relevance: (i) the spatial location of these least cost paths (LCPs) and (ii) the accumulated cost along them ('cost distance', CD). This spatial analysis requires that cost values be assigned to every type of land cover. The sensitivity of both LCPs and CDs to the cost scenarios has not been comprehensively assessed across realistic landscapes and diverging cost scenarios. We therefore assessed it in diverse landscapes sampled over metropolitan France and with widely diverging cost scenarios. The spatial overlap of the LCPs was more sensitive to the cost scenario than the CD values were. Besides, highly correlated CD matrices could derive from very different cost scenarios. Although the range of the cost values and the properties of each cost scenario significantly influenced the outputs of LCP modeling, landscape composition and configuration variables also explained their variations. Accordingly we provide guidelines for the use of LCP modeling in ecological studies and conservation planning.

**Keywords:** least cost path modeling, sensitivity analysis, ecological connectivity, spatial ecology, landscape ecology

Cet article a été publié en décembre 2021 :

Savary, P., Foltête, J. C. & Garnier, S. 2021. Cost distances and least cost paths respond differently to cost scenario variations - a sensitivity analysis of ecological connectivity modeling. *International Journal of Geographical Information Science*. DOI: 10.1080/13658816.2021.2014852

---

\*Corresponding author: paul.savary univ-fcomte.fr

# 1 Introduction

In the last decades, a variety of spatial models have been put forward for mapping and conserving ecological connectivity, largely benefiting from the development of GIS tools and landscape ecology theories (Rayfield et al., 2011; Zeller et al., 2012). Since the landscape matrix was shown to exert a heterogeneous effect on species' movement depending on its composition and configuration (Ricketts, 2001), landscapes have been represented as cost surfaces, i.e. raster grids on which every pixel value is supposed to reflect resistance to movement. Modeling ecological connectivity on these surfaces then consists in computing the paths followed by individuals for bridging pairs of habitat patches while minimizing movement costs (Adriaensen et al., 2003). These 'least cost paths' (LCP hereafter) provide two kinds of ecological information: (i) their spatial location, i.e. the LCP itself, and (ii) the accumulated cost summed along the LCP, also known as the 'cost distance' (CD hereafter). Although similar connectivity modeling approaches have been developed (Marrotte and Bowman, 2017; McRae, 2006; Panzacchi et al., 2016), the relevance of LCP modeling continues to be reflected by their frequent application in spatial analyses with decision-oriented aims such as wildlife linkage planning (Beier et al., 2008; Carroll et al., 2012; Sawyer et al., 2011), invasive species control (Etherington and Perry, 2016) or in statistical analyses directed at hypothesis testing in ecological studies (Balbi et al., 2019; Mony et al., 2018).

In this approach, cost values are key modeling inputs that most often depend upon the types of land cover in which the pixels lie (Zeller et al., 2012). Their choice is frequently driven by knowledge from literature surveys, field experience, or expert opinion regarding the movement behavior of the study species (Clevenger et al., 2002; Pullinger and Johnson, 2010). Therefore an element of arbitrariness remains at this stage, which could influence the outputs. On the one hand, different cost scenarios can produce substantially different outputs if they differ widely in terms of the absolute values assigned to land cover types and the contrast between them (Gonzales and Gergel, 2007; Murekatete and Shirabe, 2018; Rayfield et al., 2010). On the other hand, the order of the cost values assigned to land cover types is known to influence CDs and LCPs (Beier et al., 2009). This order directly reflects the preference of individuals for land cover types against others. Yet we may wonder whether two highly correlated CD matrices or two sets of spatially equivalent LCPs could derive from cost scenarios ranking land cover types in different orders. Although Murekatete and Shirabe (2018) have assessed the influence of the distribution of cost values on LCPs, their study was based on simulated landscapes and continuous cost ranges. Therefore an assessment of the sensitivity of LCP modeling in real landscapes to changes in the distribution of discrete cost values commonly used by practitioners in conservation modeling should be carried out.

In addition to cost values, the landscape structure in itself has been shown to influence the sensitivity of CD values to cost scenarios in analyses relying on simplified simulated landscapes (Bowman et al., 2020; Cushman et al., 2013; Marrotte and Bowman, 2017; Murekatete and Shirabe, 2018; Rayfield et al., 2010; Simpkins et al., 2017). For example, CD values were less affected by variations of scenarios when habitat areas were highly connected, according to the study of Cushman et al. (2013). However, a similar assessment over a wide range of existing landscapes is still lacking although it could identify the range of realistic landscape contexts in which modeling results are the most dependent on cost scenarios.

Similarly few studies have assessed the sensitivity of the LCP spatial locations to variations of cost scenarios. Beier et al. (2009) modeled a corridor providing a linkage between two Californian protected areas using several expert-based cost scenarios. Although they concluded that this corridor overlapped most of the alternative corridors modeled following alternative scenarios, these scenarios were all somehow similar to a presumably realistic scenario and the cost values ranged from 1 to 10. The relative resistance of land cover types is sometimes known by the experts but the plausibility of the range of cost values cannot be known beforehand. Besides, when this range derives from empirical observations, it is usually much wider (Khimoun et al., 2017; Ruiz-González et al., 2014). In contrast, Pullinger and Johnson (2010) compared the paths followed by woodland caribou according to GPS

tracks to LCPs modeled under several cost scenarios. They reached more pessimistic conclusions suggesting that the spatial location of LCPs is highly sensitive to cost scenarios. Given that a limited number of scenarios were used in these rare studies with inconsistent findings (see [Murekatete and Shirabe \(2018\)](#) for another example), the spatial overlap of LCPs remains to be investigated under more variable scenarios.

In order to bring landscape resistance assumptions closer to ecological reality, a first step of connectivity analysis often consists in inferring cost values from biological data and environmental variables ([Kadoya, 2009](#); [Pressey, 2004](#)). This can be done by converting presence or movement probabilities deriving from species distribution models (SDM) or step selection functions based on telemetry data into cost values ([de la Torre et al., 2019](#); [Duffot et al., 2018](#); [Zeller et al., 2018](#)). Alternatively, statistical approaches have been developed for inferring cost values from the relationship between pairwise CDs and pairwise genetic distances between populations occupying habitat patches ([Peterman et al., 2019](#); [Peterman and Pope, 2020](#); [Zeller et al., 2016](#)). In this case, the cost value scenario leading to the strongest statistical relationship between CD and pairwise genetic distances is supposed to reflect the CD perceived by individuals during their dispersal movements leading to gene flow. The set of inferred cost values associated with land cover types is then the input of LCP modeling. Their absolute values, rank, and contrasts inherently contain information of ecological relevance, making it possible to determine that one land cover type is more resistant than another, or to ascertain how many times more resistant it is, among other interpretations (see [Khimoun et al. \(2017\)](#) or [Ruiz-González et al. \(2014\)](#) for an illustration).

However, the latter statistical approaches may be unable to identify the cost scenario closest to the real ecological situation when competing scenarios lead to highly correlated CD values because they may be equally correlated to the pairwise biological response used for the inference ([Zeller et al., 2016](#)). This reinforces the need to understand how cost value distributions and landscape structure influence the sensitivity of LCP modeling to cost scenarios. Similarly we do not know precisely whether optimization approaches maximizing the strength of a statistical relationship between pairwise biological responses and CDs under several scenarios can lead to reliable predictions of the spatial location of LCPs. Two highly correlated CD matrices computed from the same locations but using different scenarios could lead to spatially distinct LCPs. Using either one or the other set of CD values in a statistical analysis would not significantly affect the output and both cost scenarios could be assigned the same likelihood from an inference using empirical data. In contrast, the choice of one of them would largely influence the LCP, potentially leading to the implementation of spatially inadequate conservation measures. In sum, although it has already been shown that LCPs and CDs are sensitive to both the cost scenarios and the landscape contexts in which they are computed, the relative sensitivities of these outputs and their main drivers need to be investigated simultaneously and in a realistic context if the reliability of connectivity analysis is to be improved.

In this study, we assessed the sensitivity of both LCP spatial locations and corresponding CD values deriving from LCP modeling to variations in cost scenarios. For that purpose, we randomly sampled 77 existing landscapes, geolocated points within them, and computed LCPs and the corresponding CDs under 100 widely diverging cost scenarios. We fixed an arbitrary but plausible scenario that we considered as the ecological 'truth'. We then assessed (i) the correlation between alternative CD matrices and the true CD matrix and (ii) the spatial overlap between the alternative LCP and the true ones. Finally we performed statistical analyses to identify the drivers of the sensitivity of CD values and LCPs to the cost scenario. This novel approach allowed us to identify (i) landscape contexts and (ii) cost scenario characteristics influencing the sensitivity of LCP modeling to cost scenarios.

## 2 Methods

### 2.1 Landscape sampling

With the aim of providing guidelines for LCP modeling in realistic conditions, we randomly sampled 250 landscapes of 30 km  $\times$  30 km with a spatial resolution of 10 m across metropolitan France from the OSO land cover raster map (Inglada et al., 2017). This map is based on remote sensing imagery and initially included 23 land cover types. As this thematic resolution did not reflect the simplified land cover maps commonly used for connectivity modeling and would not have allowed a fine assessment of the influence of the cost assigned to each land cover type, we reclassified it into five land cover types: (1) forests, (2) grasslands and woody perennial crops (grasslands hereafter), (3) annual crops, (4) artificial areas (built-up land, roads and transport infrastructures), and (5) others (water and other land cover types). The spatial and thematic resolutions of this raster layer allowed us to correctly account for the barrier effects of linear landscape features such as transport infrastructures, which can largely influence LCP modeling (Hoover et al., 2020).

We considered the landscape constraints on movement faced by an arbitrary forest species and we therefore only conserved the sampled landscapes with a proportion of forest above 15 %. This proportion of habitat is close to the threshold below which a specialist forest species becomes extinct (Balkenhol et al., 2013; Hanski et al., 1996). To prevent the results from being influenced by the absence of one of the land cover types while allowing the land cover type proportions to vary substantially, we ensured that the proportion of grasslands, crops and artificial areas were above 5%, 5%, and 2% respectively. Finally, we removed coastal landscapes including large maritime areas, which led us to retain 77 landscapes for the analyses (cf. section 3.1).

### 2.2 Cost scenario creation

For comparative purposes, we chose a reference cost scenario ('true scenario' hereafter), in which the cost values associated with forests, grasslands, crops, artificial areas and water and other land cover types were respectively 1, 10, 100, 1000 and 100. They reflected the movement behavior of a forest specialist species. Note that similar cost values have already been used for modeling connectivity for forest species (Gurrutxaga et al., 2010; Schadt et al., 2002) and a similar range (1-1000) has already been inferred from field data (Khimoun et al., 2017; Pérez-Espona et al., 2008; Ruiz-González et al., 2014; Wang et al., 2008).

In order to test for the sensitivity of LCP modeling to cost scenarios, we randomly created 100 widely different alternative cost scenarios. They differed by both the order of the land cover types and the contrast between cost values. We used Shirk et al. (2010) approach to set cost values using the following function:

$$C_i = \left(\frac{Rank_i}{Rank_{max}}\right)^x \times C_{max}$$

where  $C_i$  is the cost value between 1 and  $C_{max} = 1000$  associated with the  $i$ -th land cover type.  $Rank_i$  is the rank of the land cover type  $i$  between 1 and  $Rank_{max} = 4$ . We used  $x$  values equal to 1, 2, 4, 8 or 16. We therefore obtained five series of values: [1, 1, 11, 1000], [1, 4, 11, 1000], [4, 63, 317, 1000], [63, 250, 563, 1000], [250, 500, 750, 1000] (Supporting information, figure 8). Using each of them, we randomly assigned cost values to forests, grasslands, crops and artificial areas before randomly selecting 100 alternative cost scenarios among these combinations. The cost value associated with water and other land cover types was set to 100 in each cost scenario in order to limit the number of combinations to test and because this land cover type was absent from several landscapes.

### 2.3 Least cost path modeling

In every landscape, we randomly selected 50 point locations within forest patches, separated by a distance of more than 500 m. We then computed LCPs between every pair of points in every landscape and under every cost scenario (Figure 1A). We thus obtained in each case a set of LCPs and the corresponding 50  $\times$  50 pairwise CD matrix. We created buffer zones of 200 m on each side of

every spatial line and merged them. Hereafter, we use the term "least cost corridors" for qualifying these polygons of equal width around least cost paths.

## 2.4 Spatial and distance-based comparisons of LCPs

We first measured the proportion of the area of every true least cost corridor between a pair of locations that was overlapped by the corresponding alternative least cost corridor. We averaged the 1225 values obtained when considering every pair of locations, thereby obtaining a spatial overlap measure for each combination of a landscape and an alternative cost scenario. Besides, we assessed the statistical relationship between each alternative CD matrix and the true one by computing their Mantel  $r$  correlation coefficient (Mantel, 1967). This coefficient is commonly used for assessing the relationship between distance matrices.

## 2.5 Landscape structure and cost scenario descriptors

We first aimed to explain the sensitivity of LCP modeling to the cost scenarios according to landscape structure. To that end, we computed the proportion of forests, grasslands, crops, and artificial areas in every landscape (landscape composition variables). We also computed the Shannon index as a landscape composition diversity variable. It was divided by  $\log(n)$  where  $n$  is the number of land cover types so that it ranges from 0 to 1. In order to assess the influence of landscape configuration, we computed several FRAGSTATS configuration metrics (McGarigal, 1995). At the landscape level, we computed the contagion index which reflects the degree of aggregation of the cells of the same land cover type. For the two land cover types that were assigned extreme cost values in the true scenario, i.e. forest and artificial areas, we computed the number of patches, the shape complexity, and the 'clumpy' index of fragmentation. Finally, as a way to assess the global permeability of the landscapes, we computed the sum of the pixel costs on each landscape according to the true scenario.

In parallel, we aimed at explaining the sensitivity of LCP modeling in the different landscapes according to the cost value distribution of each scenario. For that purpose, we first computed binary variables indicating whether each alternative scenario ordered land cover types in the same way as the true scenario does, or whether the contrast of the cost values was similar. In the latter case, we considered that the cost values [1, 4, 101, 1000] were the closest to the true contrast. Finally we identified the pairs of land cover types that were not ordered in the same way as in the true scenario.

## 2.6 Statistical analyses of the drivers of LCP modeling sensitivity to cost scenarios

The values of (i) the Mantel  $r$  correlation coefficients between true and alternative CD matrices (Mantel correlation hereafter) and of (ii) the proportions of spatial overlap of the alternative least cost corridors with the true ones (spatial overlap hereafter) were supposed to reflect the sensitivity of LCP modeling to cost values. The greater the variability of these metrics in a given landscape for the different scenarios, the greater the sensitivity of LCP modeling in this landscape. Besides, for a given cost scenario, these metrics took values across the landscapes reflecting the overall similarity of this scenario to the true scenario. By computing these two metrics in every landscape and for every cost scenario ( $2 \times 77 \times 100$  values), we could compare their sensitivity, assess whether and to what extent the landscape structure explained their sensitivity and identify which characteristics of the cost scenarios make them similar to the true scenario. Accordingly, we first performed separate two-way ANOVA of these two metrics by considering the cost scenario and landscape as the factor variables explaining their variations (Figure 1 B1). This allowed us to quantify the contribution of each factor to the variations of both Mantel correlations and spatial overlaps.

Then we studied whether landscape composition and configuration variables could explain the sensitivity to cost scenarios (Figure 1 B2). For that purpose, we computed the mean and coefficient of variation of the Mantel correlations and of the spatial overlaps for each landscape across the different cost scenarios. Large mean values indicate that, independently of the cost scenario, the Mantel correlations or spatial overlaps tend to be high for a given landscape, whereas large coefficients of variation

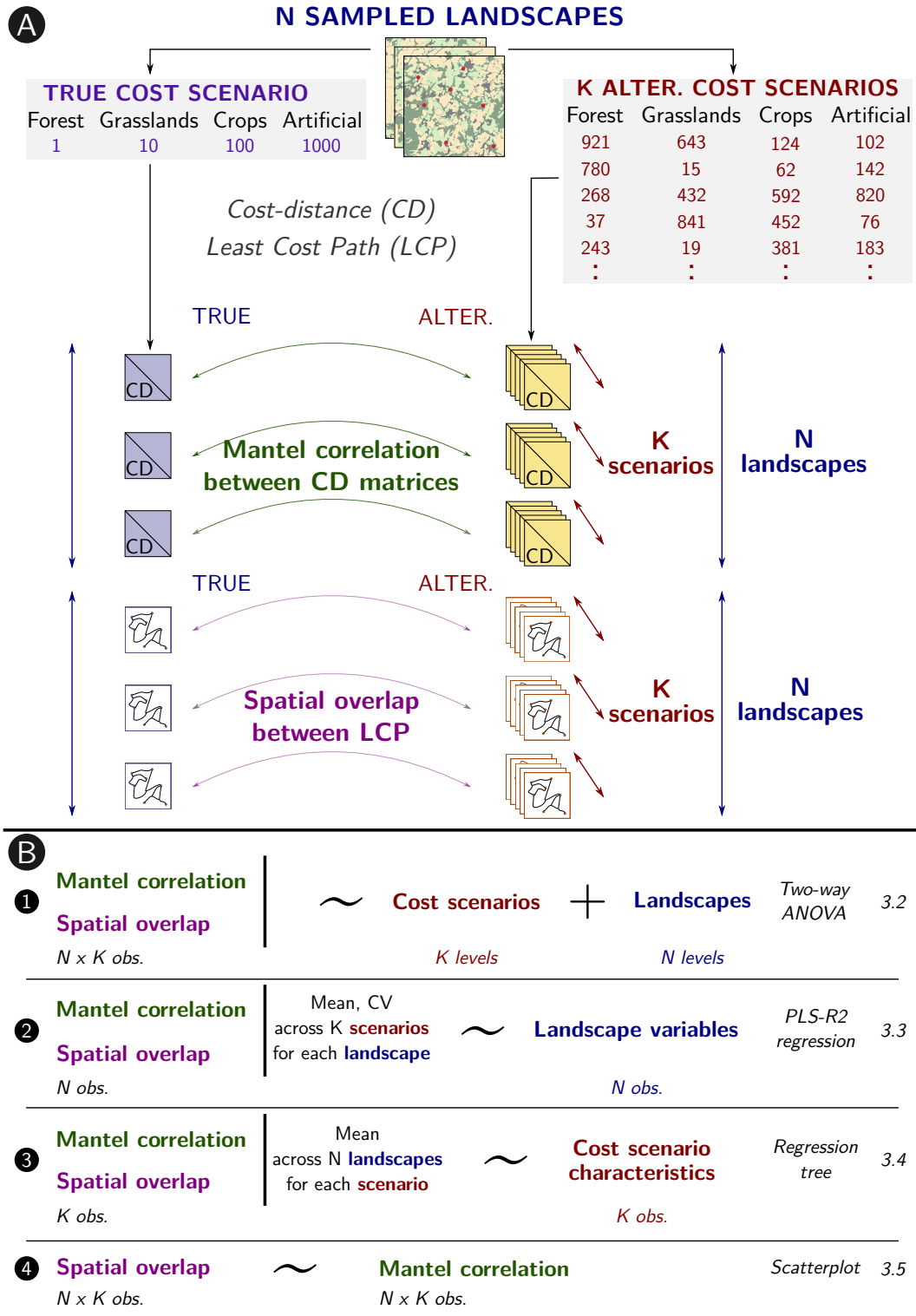


Figure 1: Schematic representation of (A) the spatial and distance-based comparisons of the LCP computed under true and  $K = 100$  alternative cost scenarios within  $N = 77$  landscapes, and of (B) the statistical analyses performed for assessing the sensitivity of LCP modeling to cost scenarios. We performed separate two-way ANOVA for assessing the respective influences of the landscape and cost scenario on (i) the spatial overlap and (ii) the Mantel correlations. Then we assessed the influence of landscape composition and configuration on the sensitivity of LCP modeling to cost scenarios in every landscape (mean and coefficient of variation across scenarios) by carrying out a PLS-R2 regression. We identified the characteristics of the cost scenarios explaining the values of spatial overlap and Mantel correlations averaged across landscapes for every cost scenario with a regression tree. Finally, we assessed the relationship between the spatial overlap and the Mantel correlations for every combination of the landscape and cost scenario by displaying it on a scatterplot. The numbers in italics at the bottom-right of the figure refer to the results section where the corresponding results are described.

indicate that these metrics are highly variable depending on the cost scenario for a given landscape. We then modeled these numeric indicators as a function of the landscape variables. We carried out two separate Partial Least Squares (PLS) regressions with the mean and the coefficient of variation of these two metrics as the response variables (one model for each metric) and the landscape variables as the predictor variables. PLS regressions are an alternative to multiple linear regression and principal component regression (Carrascal et al., 2009; Roy et al., 2015; Wold et al., 2001), particularly suitable for cases in which predictor variables are collinear. This type of regression identifies the factorial space components that simultaneously maximize the explained variance of the response variables and of the predictor variables. This makes it possible to model a set of response variables (PLS-R2). Following Tenenhaus (1998), we computed the  $Q^2$  index to assess the role of every component for improving the prediction of the response variables when performing leave-one-out cross-validation. We described only the results regarding the effects of the components which significantly improved the prediction of the response variables, i.e. when the  $Q^2$  associated with these components were larger than 0.0975.

Finally, we aimed at identifying the characteristics of the cost scenarios driving the sensitivity of LCP modeling to cost scenarios (Figure 1 B3). We computed the mean values of the Mantel correlations and of the spatial overlaps for each cost scenario across the different landscapes. High values indicate that, independently of the landscape, a given cost scenario leads to LCPs and CDs that are very similar to those derived from the true scenario. We expected this similarity to be explained by the raw cost values of each cost scenario, by their orders and contrasts and by their differences with those of the true scenario (cf. previous section). In order to obtain a decision tree showing the cost scenario characteristics leading to the similarity of LCP modeling outputs with the true ones, we built regression trees (Breiman et al., 1984) to explain either the mean Mantel correlation or mean spatial overlap as a function of the cost scenario characteristics. This method involves splitting the predictor space into a limited number of regions called leaves in which the response variable is predicted to take its mean value within the leaf (James et al., 2013). These trees can take both continuous and categorical predictor variables and have been shown to perform better than linear models in the presence of non-linear relationships. They were pruned according to a cost-complexity criterion to prevent overfitting, using `rpart` package (Therneau et al., 2010) in R.

## 3 Results

### 3.1 Structure of the sampled landscapes

After applying our selection criteria to the sampled landscapes, we ended up with 77 landscapes (Figure 2), all very different in terms of both landscape composition (land cover type proportions and diversity) and configuration (fragmentation, number of patches, and contagion)(Table 1). This sample included fine-grained and coarse-grained agricultural landscapes (Figures 2A and 2F, respectively) and widely forested landscapes in both lowlands and highlands (Figures 2G and 2C, respectively).

### 3.2 Relative influences of cost values and landscape structure on the sensitivity of LCP modeling outputs

The Mantel correlations ranged from -0.40 to 0.99 and the spatial overlaps exhibited similar variations (from 7% to 94%) but never reached 100%, their theoretical maximum (Figure 3, see figure 9 for a similar variation using the Spearman Mantel correlation coefficient). This indicates that the wide range of cost scenarios we considered was sufficient for creating contrasted outputs and studying their variability and its drivers. The variance of the spatial overlap for every combination of a landscape and a cost scenario was accounted for much more by the cost scenario than by the landscape considered (F values from the two-way ANOVA: 498.17 vs 33.71, respectively), although both influences were highly significant. Indeed the cost scenario and the landscape together explained 87% of the variance of this metric but applying the same cost scenario to the different landscapes led to lower variations in terms of spatial overlap than applying the different scenarios to the same landscape did (Figures 3A and 3B).

The cost scenario and the landscape together explained a slightly lower yet significant share of



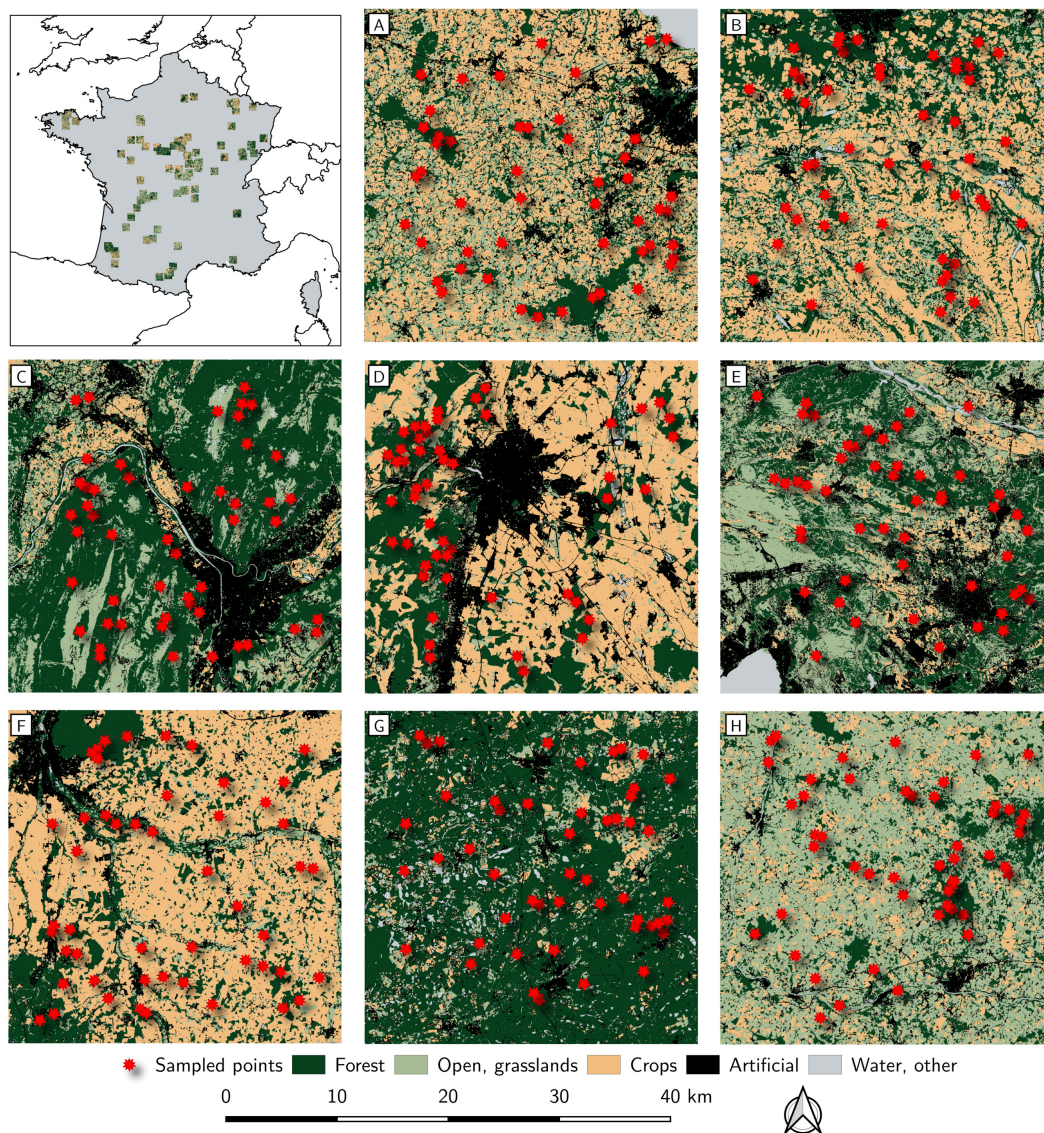


Figure 2: Results of the landscape sampling: 77 landscapes of  $30 \text{ km} \times 30 \text{ km}$  with a spatial resolution of 10 m were randomly sampled across metropolitan France. They were filtered applying the following criteria:  $> 15 \%$  forests,  $> 5 \%$  grasslands,  $> 5 \%$  crops,  $> 2 \%$  artificial areas. Fifty points were randomly sampled in the forest areas, all more than 500 m apart. Eight contrasted examples (A to H) are shown on the map.

Variable	Minimum	Median	Maximum
% forest (15 - 100)	15.61	28.57	79.16
% grasslands (5 - 100)	5.38	26.43	69.66
% crops (5 - 100)	5.08	21.97	59.98
% artificial areas (2 - 100)	2.04	8.42	24.65
Shannon div. index (0 - 1)	0.46	0.75	0.93
Frag. forest patches ('clumpy') (-1 - 1)	0.81	0.93	0.97
Frag. artif. patches ('clumpy') (-1 - 1)	0.67	0.79	0.91
Nb. forest patches (0 - $9 \times 10^6$ )	2467	6563	31918
Nb. artif. patches (0 - $9 \times 10^6$ )	7602	18610	48569
Shape complex. forest patches (>1)	1.24	1.39	1.49
Shape complex. artif. patches (>1)	1.29	1.36	1.43
Contagion (0 - 100)	46.24	58.29	74.60
Total cost ( $\times 10^9$ )(0.009 - 9)	0.41	1.09	2.34

Table 1: Landscape characteristic distributions observed among the 77 sampled landscapes. The possible range of variation of the variables is shown in brackets after the variable name.

variance of the Mantel correlations (79 %). Similarly the magnitude of variation of the Mantel correlations was much lower for a given cost scenario across the landscapes than for a given landscape across the cost scenarios (F values from the two-way ANOVA: 235.39 vs 67.34, respectively, both highly significant, Figures 3C and 3D).

The spatial overlap was more sensitive to the cost scenario than the Mantel correlation, as shown by the rapid decrease of the median spatial overlaps computed with each cost scenario (Figure 3A) compared with the slower decrease of the median Mantel correlations (Figure 3B). In contrast, when considering the distribution of these two metrics for each landscape across the different cost scenarios, we observed less variation for the spatial overlap than for the Mantel correlations (Figures 3B and 3D). Indeed the spatial overlaps across the cost scenarios were overall small for a given landscape (Figure 3B). In contrast, the Mantel correlations were much more variable (Figure 3D) and consistently took large values in some landscapes, whereas they took smaller values with larger variations in others.

### 3.3 Landscape structure influence on the sensitivity of LCP modeling outputs

The PLS regressions identified the landscape variables responsible for the sensitivity of LCP modeling to the cost scenarios. Only the first component of the PLS regression explaining the mean and the coefficient of variation of the spatial overlaps across the cost scenarios had a significant effect ( $Q^2 = 0.34$ , Figure 4A). The mean spatial overlap was highly and positively correlated to the first component of the PLS whereas the coefficient of variation of this variable was only slightly correlated and not significantly explained by this component. The mean spatial overlap was positively influenced by the contagion variable and by the proportion of forests in the landscapes, and negatively influenced by the total cost of the landscape, the Shannon diversity index, the number of patches of artificial area and of forest, and by the proportion of artificial areas (Figure 4A). This means that the spatial overlap was higher in landscapes relatively favorable to species movements, with little diversified land cover types dominated by forests and containing large and aggregated patches.

Similarly only the first component of the PLS regression explaining the mean and the coefficient of variation of the Mantel correlations for a given landscape had a significant effect ( $Q^2 = 0.26$ , Figure 4B). This component was positively correlated with the mean Mantel correlation whereas it was negatively correlated with its coefficient of variation. Mean correlation coefficients were positively influenced by the contagion index. Conversely the coefficients of variation of these coefficients were positively influenced by the total cost of the landscape, the Shannon diversity index, the proportion of crops and artificial areas, the shape complexity of the forest patches, and the Clumpy index of forest and artificial area patches. This means that CD matrices consistently exhibited high correlations with the true CD matrix in landscapes with large contiguous patches. In contrast, alternative CD matrices

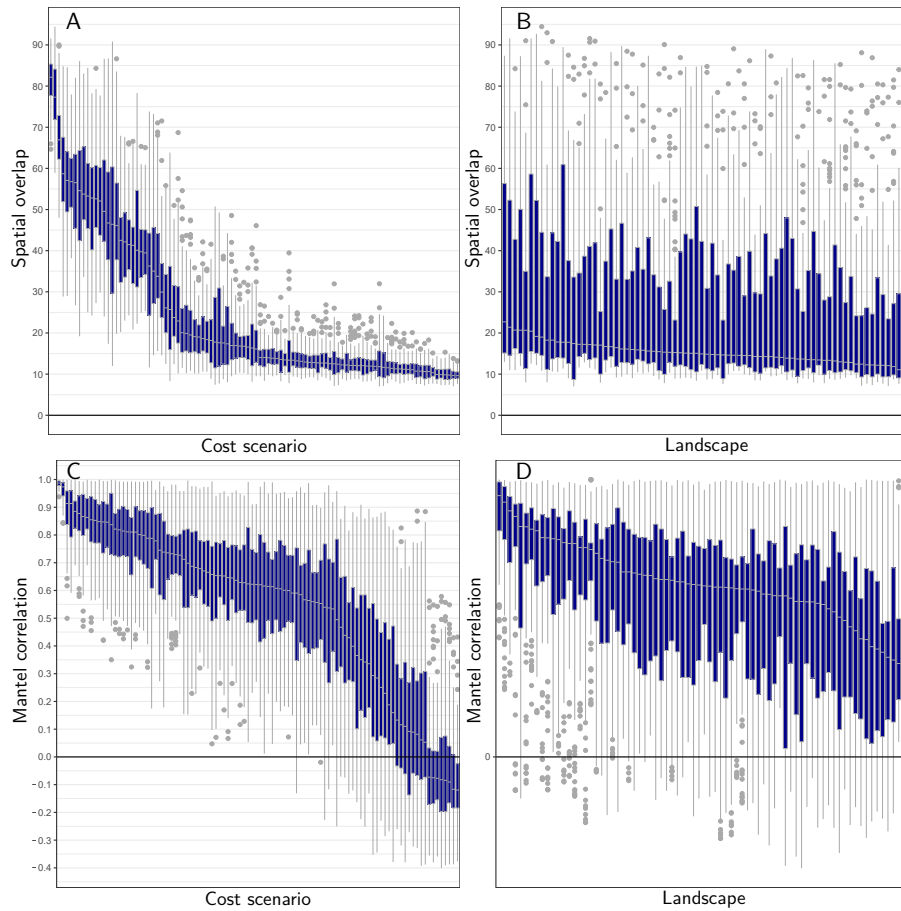


Figure 3: Distribution of the spatial overlap (A, B) or the Mantel correlations (C, D) according to the cost scenario (A, C) or the landscape (B, D). Cost scenarios (A, C) and landscapes (B, D) are placed in decreasing order of median spatial overlap (A, B) or median Mantel correlations (C, D) along the x axis. When the distribution is displayed for each cost scenario (A, C), the 100 boxes are made up of 77 values each corresponding to a landscape, whereas when it is displayed for each landscape (B, D), the 77 boxes are made up of 100 values each corresponding to a cost scenario.

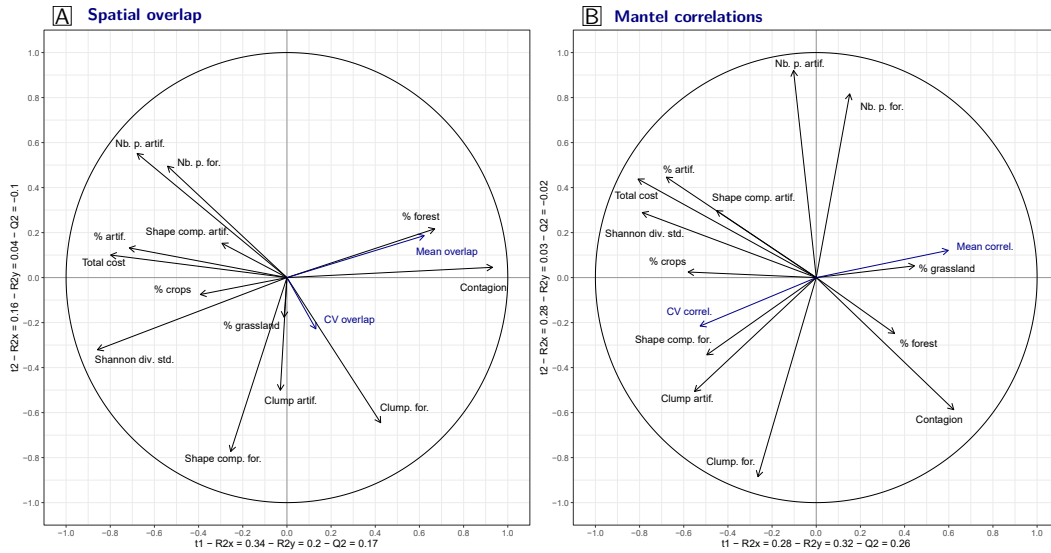


Figure 4: Projection of the response (dark blue labels) and predictor variables (black labels) of the PLS-R2 regression in the factorial space derived from the two first components ( $t_1$ ,  $t_2$ ). The left panel (A) shows the results obtained when modeling the mean and coefficient of variation of the spatial overlap in a given landscape across all the cost scenarios whereas the right panel (B) shows results obtained when modeling the mean and coefficient of variation of the Mantel correlation in a given landscape across all the cost scenarios. The axis labels indicate the percentage of variance of the predictor variables table (R2x) or of the response variables table (R2y) explained by each component ( $t_1$  or  $t_2$ ), as well as the percentage of variance of the response variables table explained by these components when performing a cross-validation (Q2).

tended to be less strongly and more variably correlated with the true CD in diverse landscapes with complex patch shapes and large areas of the least favorable land cover types.

### 3.4 Cost scenario characteristics influence on the sensitivity of LCP modeling outputs

The regression trees identified the cost scenario characteristics explaining their spatial overlaps and Mantel correlations with the true scenario across the different landscapes (Figure 5). The first split of the two regression trees created were conditions regarding the cost value assigned to forest. Indeed, cost scenarios with forest cost values lower than 82 led to mean spatial overlaps averaging 40% (Figure 5A). Conversely the cost scenarios assigning forests a cost value lower than 875 (i.e. different from 1000) led to mean Mantel correlations averaging 0.63 across landscapes (Figure 5B). Interestingly, when the forest cost value was equal to 1000, if the cost values were drawn from either the [63, 250, 563, 1000] or [250, 500, 750, 1000] gradients, the Mantel correlations still averaged 0.55 although such cost scenarios differed largely from the true one. In this case, assigning much lower cost values to grasslands, crops, and artificial areas ([1, 1, 11], [1, 4, 101], or [4, 63, 317]) led to negligible Mantel correlations (Figure 5B). Accordingly, the gradient of values of the cost scenarios was the second most important criterion explaining the Mantel correlations obtained across the landscapes for a given cost scenario.

In contrast, the second most important criterion explaining the spatial overlap was the difference between cost values assigned to forest and grassland, which are the two least resistant land cover types in the true scenario (Figure 5A). Making forests more resistant than grasslands invariably reduced the spatial overlap with the true LCP. Finally, in both trees, the other splitting criteria concerned the costs associated with crops and artificial areas. For example, when forest cost value is both lower than 82 and lower than the grassland cost value, assigning artificial areas a cost value lower than 37 (true cost value: 1000) led to spatial overlaps averaging 33 %, which is a rather large value in light of the range of variation. Similarly, provided the cost value was lower than 284 for forests, greater than 8 for the crops and greater than 82 for artificial areas, the mean Mantel correlation across landscapes averaged over the corresponding scenarios reached 0.8, independently of the order of cost values and the contrast between them. The binary variables comparing each scenario to the true one in terms of

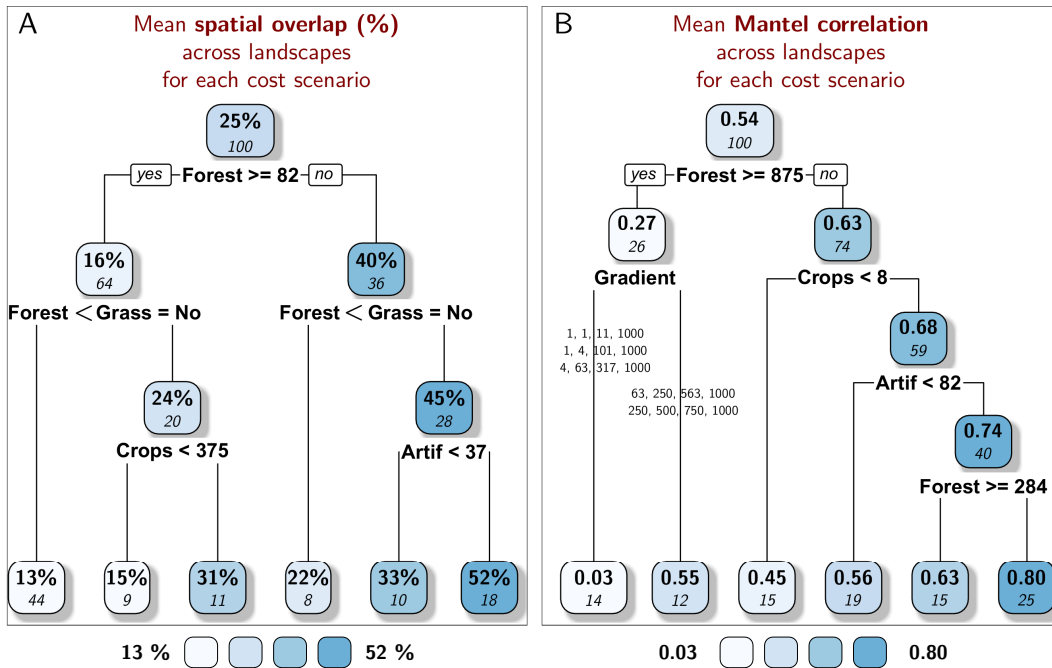


Figure 5: Regression trees explaining the mean spatial overlap (A) or the mean Mantel correlation across landscapes for each cost scenario as a function of the characteristics of the distribution of the cost values. At each node (branch split), the criterion displayed is verified in all the leaves stemming from branches located on the left side of the split, whereas its opposite is verified in all the leaves stemming from branches located on the right side of the split. The colored boxes indicate the mean value of the response variable across all the leaves stemming from a given node (in bold) and the number of observations, i.e. cost scenarios, included in these leaves (in italics).

order and contrasts were not retained in the best trees computed for both metrics.

### 3.5 Relationship between the spatial overlap of LCP and the correlation between CD matrices

From the application of every alternative cost scenario in every landscape (7700 combinations), we observed that the spatial overlaps and the Mantel correlations were somehow related (Spearman's  $r = 0.66$ ) but their relationship was highly non-linear (Figure 6). Spatial overlaps above 65% were only obtained with LCPs whose associated CDs were moderately to highly correlated with the true CD ( $r > 0.5$ , Figure 6). Yet the degree of correlation between CD matrices was a poor proxy for the spatial overlap of LCPs. Indeed, spatial overlaps below 20% were frequently obtained while corresponding CD matrices were highly correlated with the true CD matrix ( $r > 0.9$ , Figure 6 and Table 2).

In addition, spatial overlaps above 80% were mostly observed when land cover types were arranged in the same order of resistance as in the true cost scenario (Figure 6). However, some scenarios incorrectly ordering these land cover types still reached spatial overlaps up to 91%. The cost scenario leading to the largest Mantel correlations was [1, 4, 101, 1000] (for forest, grasslands, crops, and artificial areas, respectively; Mantel  $r$ : mean across landscapes = 0.98, Table 2). This scenario was apparently the most similar to the true one (i.e. [1, 10, 100, 1000]) in terms of order and contrast of the cost values but surprisingly the mean spatial overlap of the corresponding LCP across the 77 landscapes was not the maximum (77.81% vs 81.00%, Table 2). Indeed, the best scenario in terms of spatial overlap was [4, 63, 317, 1000] and also led to CD values highly correlated with the true ones (mean Mantel correlation across landscapes = 0.96).

Although the best cost scenarios in terms of spatial overlap always assigned a larger cost value to grasslands than they did to forests, they did not systematically assign a larger cost value to artificial areas than they did to crops (e.g. scenario [1, 4, 1000, 101]: mean spatial overlap: 68.04%, Table 2). In contrast, in the ten cost scenarios with the strongest Mantel correlations, two cost scenarios assigned a lower cost to grasslands than they did to forests. If the Mantel correlations obtained in these two cases were above 0.85, the corresponding spatial overlaps were nevertheless below 40%.

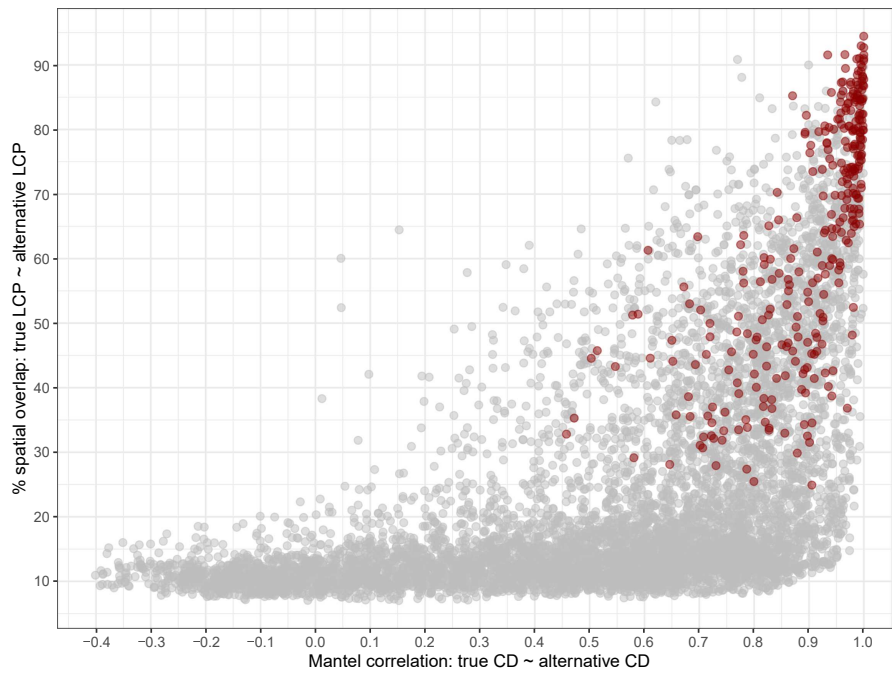


Figure 6: Relationship between the Mantel correlations ( $x$  axis) and the spatial overlap ( $y$  axis). The spatial overlap is expressed as the proportion of the area of the 200 m wide buffer around the true LCP that overlaps the same buffer created from each alternative LCP.  $n = 7700$ , corresponding to the combination of 100 cost scenarios and 77 landscapes. Red dots correspond to alternative cost scenarios ordering the four land cover types in the same way as in the true cost scenario.

Finally, when projected into a spatially-explicit layout, we observed large differences between LCPs resulting from different cost scenarios (Supporting information, figure 7). Interestingly, even highly correlated CD matrices could be derived from LCPs diverging rather markedly from each other.

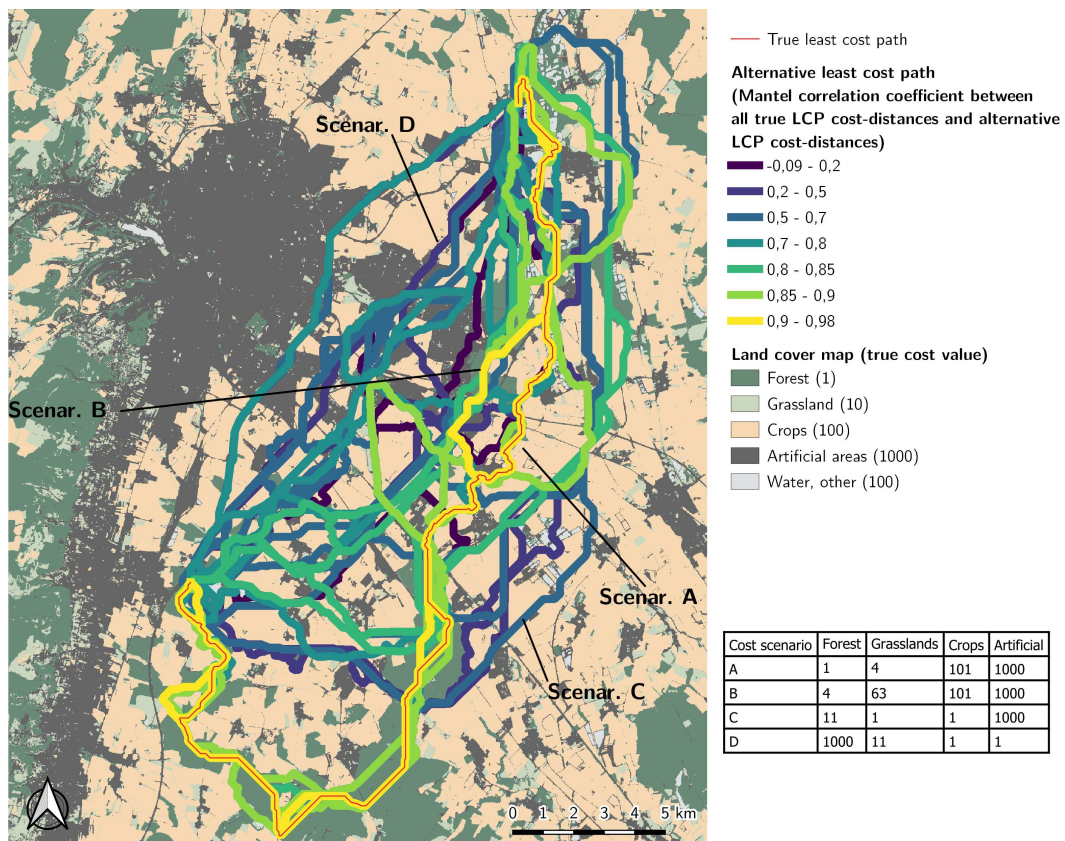


Figure 7: Spatial representation of the influence of cost scenarios on least cost path locations. The least cost corridor between two locations computed applying the true cost values to land cover types is displayed in red on the map. Alternative least cost paths computed using other cost values are displayed with colors reflecting the mean Mantel correlation between the corresponding alternative cost distance matrices and the true cost distance matrix (including all the paths computed on this landscape). Scenarios A, B, C, and D are examples of cost scenarios that diverge to varying degrees from the true one. Their corresponding costs are in the table included in this figure.

Forest	Grasslands	Crops	Artif	For. <Grass.	Order	% overlap	Mantel r
1	4	101	1000	Yes	Same	<b>77.81</b>	<b>0.98</b>
4	63	317	1000	Yes	Same	<b>81.00</b>	<b>0.96</b>
1	4	1000	101	Yes	Diff.	<b>68.04</b>	<b>0.88</b>
4	1	101	1000	No	Diff.	25.78	<b>0.88</b>
4	317	63	1000	Yes	Diff.	<b>58.83</b>	<b>0.87</b>
1	1	11	1000	No	Diff.	36.86	<b>0.87</b>
63	250	563	1000	Yes	Same	<b>56.08</b>	<b>0.86</b>
63	563	1000	250	Yes	Diff.	<b>55.52</b>	<b>0.84</b>
63	563	250	1000	Yes	Diff.	<b>54.59</b>	<b>0.84</b>
1	101	4	1000	Yes	Diff.	<b>54.44</b>	<b>0.83</b>
63	1000	563	250	Yes	Diff.	52.25	0.83
63	1000	250	563	Yes	Diff.	51.60	0.83
1	11	1	1000	Yes	Diff.	45.86	0.81
4	317	1000	63	Yes	Diff.	<b>57.19</b>	0.81
250	500	1000	750	Yes	Diff.	40.09	0.79
250	750	1000	500	Yes	Diff.	41.67	0.79
4	1000	63	317	Yes	Diff.	<b>53.70</b>	0.79
250	500	750	1000	Yes	Same	39.80	0.79
250	1000	750	500	Yes	Diff.	41.89	0.79
250	750	500	1000	Yes	Diff.	40.13	0.78
4	1000	317	63	Yes	Diff.	52.70	0.76
1	11	1000	1	Yes	Diff.	43.99	0.74
1	1000	4	101	Yes	Diff.	48.63	0.67
1	1000	101	4	Yes	Diff.	47.03	0.62

Table 2: Differences between the alternative cost scenarios and the true cost scenario. We included the cost scenarios with the 20 highest mean proportions of spatial overlaps or mean Mantel correlation coefficients (24 scenarios in total). They are ordered in descending order of Mantel correlation coefficient (Mantel  $r$ ). Their spatial overlap with the LCPs derived from the true cost scenario are displayed (% overlap). The 10 largest values of the columns 'Mantel  $r$ ' and '% overlap' are displayed in bold. Values obtained for all LCPs, CD matrices and landscapes have been averaged for each cost scenario. Cost values associated with each of the four land cover types are included in the 'Forest', 'Grasslands', 'Crops', and 'Artif' columns. The 'For. < Grass.' column indicates whether the cost value associated with forest is lower than that associated with grasslands. The 'Order' column indicates whether cost values associated with land cover types in each alternative cost scenario follow the same order as those in the true cost scenario.

## 4 Discussion

Using a wide range of cost scenarios in real diversified landscapes, we analyzed the relative sensitivities of both LCP spatial locations and CD values to the choice of cost values and identified their drivers. As expected, these outputs of LCP modeling were sensitive to the cost scenarios but their sensitivities differed and were not influenced by the same characteristics of the cost scenarios, nor to the same extent according to landscape structure.

### 4.1 Sensitivity of LCP and CD to cost scenarios

*The spatial overlap of LCP is very sensitive to the cost scenario*

An analysis of the spatial sensitivity of LCP modeling to widely diverging scenarios was lacking. Using the spatial overlap between true and alternative LCPs obtained from different cost scenarios as a measure of sensitivity, we showed that LCPs are highly dependent upon the cost scenarios given that just a few scenarios allowed us to reach large proportions of overlap with the true LCPs. Previous studies regarding the spatial overlaps of LCPs provided inconsistent findings. Pullinger and Johnson (2010) showed that LCPs did not follow caribou GPS tracks, whereas Beier et al. (2009) obtained large spatial overlaps of alternative corridors between two protected areas. Our results can explain these opposing conclusions. Although the cost scenario was the main driver of the spatial overlap, in landscapes with large proportions of favorable land cover types, large contiguous patches and a low diversity of land cover types, the spatial overlap between true and alternative LCPs tended to be larger. In this case, using similar scenarios in terms of cost value order and range would lead to large spatial overlaps across the scenarios. This could potentially explain the results of Beier et al. (2009) because



these authors used a set of biologically plausible cost scenarios. In contrast, in fine-grained landscapes with the highest diversity of land cover types and the largest proportions of adverse land cover types, the spatial overlap was consistently low. This situation could reflect the study by [Pullinger and Johnson \(2010\)](#), performed in a landscape in which very contrasted altitude classes were intersected with 10 different land cover types to define a fine-grained permeability map. This also indicates that the thematic resolution of land cover maps could influence the outputs of LCP modeling by influencing their grain and diversity, although spatial resolution is often given more consideration ([Lechner and Rhodes, 2016](#)).

#### *Highly correlated CD matrices can derive from very different cost scenarios*

We studied the sensitivity of the statistical properties of CD to cost scenarios by comparing the Mantel correlations between the true CD matrix and every alternative CD matrix. Our results showed that CD matrices highly correlated to the true one can be obtained using several cost scenarios differing widely from the true cost scenario. Besides, although the sensitivity of this correlation depended more on the characteristics of the cost scenarios, the landscape context was also responsible for the variable sensitivity to cost scenarios observed in the different landscapes. In landscapes with large amounts of favorable land cover types and large patches, whatever the alternative cost scenario, the CD matrices tended to be highly correlated with the true CD matrix. Conversely, in diverse landscapes with patches of complex shapes, correlations with the true CD were lower and much more variable. This result recalls those of [Cushman et al. \(2013\)](#) showing that Euclidean distances and CDs were equivalent for explaining genetic distances when the proportion of habitat in the landscape is high and the contrast between cost values is low.

#### *Highly correlated CD matrices can derive from spatially distinct LCPs*

Spatial overlaps and Mantel correlations exhibited different sensitivity to cost scenarios when considered separately. The main contribution and novelty of our analyses is to provide insights into the relationship between the spatial locations of LCPs and their corresponding CD values, and into the drivers of the mismatches between them. We first showed that two highly correlated CD matrices can derive from paths whose spatial overlap is very low (as low as 15% with correlation coefficients above 0.9). Nevertheless, the reverse does not hold true because large spatial overlaps between paths invariably involve high correlations between CD matrices. This result is explained by the greater sensitivity to cost scenarios of the LCPs than the corresponding CDs.

These contrasted sensitivities partly stem from the fact that LCPs and CDs are not influenced by the same characteristics of cost scenarios. While the relative order of the cost values associated with the least resistant land cover types (forests and grasslands) is a key factor explaining the spatial overlap of a given cost scenario with the true scenario, the correlation of alternative CDs with the true CDs depended more heavily on the gradient of cost values. This seems logical given that the order of the cost values determines whether the path should better cross some land cover types than others, whereas the gradient of cost values determines the CD statistical distribution independently of the spatial location of the corresponding LCP. Interestingly, when the least resistant land cover type was assigned a large cost value (forest: 1000), limiting the contrast with the costs of other land cover types still led to CDs highly correlated with the true CDs. Such strong correlations between CDs deriving from the most homogeneous cost scenarios and the reference CD recall the strong correlations often observed between CDs and Euclidean distances ([Marrotte and Bowman, 2017](#)). This strong correlation has been a reason for preferring the accumulated cost along the LCP (CD) over the length of the LCP as a measure of connectivity ([Etherington and Holland, 2013](#); [Simpkins et al., 2018](#)). However, the dependence of CDs upon Euclidean distances is still a limitation of this measure and can make it difficult to distinguish several CD matrices.

Furthermore, although the cost scenario being at first sight the most similar to the true scenario ([1, 4, 101, 1000] vs [1, 10, 100, 1000]) led to the CDs most strongly correlated with the true CDs, it did not lead to the highest spatial overlap with the true LCPs, obtained with the scenario [4, 63,

317, 1000]. This could be explained by the sensitivity of LCPs to the contrasts of cost values between the least resistant land cover types, the ratios 4/63 and 63/317 being both closer to 1/10 and 10/100 than 1/4 and 4/100.

## 4.2 Implications for cost value inference and LCP modeling

### *Ecological interpretations and use of inferred cost values must be subject to caution*

The statistical distribution of two CD matrices can be almost identical although they correspond to spatially distinct LCPs that derive from cost scenarios implying different ecological interpretations. Assuming that cost value inference from biological data depends essentially upon the statistical properties of CD matrices, care has to be taken when interpreting inferred cost values and using them for mapping LCPs. Similarly the cost scenarios leading to the largest spatial overlap are not necessarily the scenarios whose values are most like the true cost values. Indeed, given that inferred cost values may be closely related to the statistical properties of the CDs, these inferred cost values should better reflect the gradient of cost values and the difference between the lowest and largest cost values than their relative order. Note that this limitation does not concern the cost values inferred from presence or telemetry data. Yet, in the latter case, the method used for converting SDM or step selection function outputs into cost values could significantly affect CD statistical distribution by determining the range and contrasts between these values.

### *Outline solutions for the use of inferred cost values in LCP modeling*

When LCP modeling supports decision-making in conservation, the spatial location of the LCP can be used to design restoration measures such as wildlife crossings for example (Clevenger et al., 2002; Mimet et al., 2016). Such a location optimization based upon LCPs can be suboptimal due to the sensitivity of these paths to the cost scenario. Although it may be problematic when LCP modeling is based upon cost values inferred from the relationship between CDs and biological data, we provide outline solutions to this problem. First, the scenario leading to the highest spatial overlap with the true LCPs was always within the scenarios leading to the CD matrices most closely correlated with the true one. This means that the set of cost scenarios closely reflecting the true landscape constraints on movements share similar statistical properties and could be retained as the 'best ones' in cost value inference. Our results thus call into question the common practice of optimization of a single best cost scenario. Rather than retaining the single 'best' cost scenario from the inference, retaining a set made of several best scenarios could ensure that the 'ecological truth' is part of the inference results. This strategy is not unlike the use of Circuitscape software (McRae and Beier, 2007), which models several alternative paths between locations in a landscape considered as an electric circuit. Similarly, Pinto and Keitt (2009) developed methods for modeling multiple shortest paths between habitat patches and Rayfield et al. (2010) suggested identifying such multiple low-cost routes for coping with the sensitivity of LCP modeling. The Linkage Mapper software (McRae and Kavanagh, 2011) makes this possible by creating least cost corridors of varying width according to the cost surface, which could provide insight into the existence of alternative and equally probable paths around the least cost path (see also Shirabe (2016)).

Yet, instead of modeling several alternative paths under one cost scenario, we here suggest to model several LCPs under a set of highly likely scenarios because this strategy could maximize the likelihood of taking into account the 'true' LCP. It would mirror the growing interest for multi-model inference (Burnham and Anderson, 2004) in ecological science where considering a single best model is often a poor approximation of the stochastic ecological reality. Similarly, the set of highly likely scenarios could be selected on the basis of a model fit criterion, e.g. the AIC (Burnham and Anderson, 2004). We acknowledge that the alternative LCPs thereby identified may occupy very different spatial locations. In such a case, their intersections may be the only information that can be used for conservation purposes.

Another strategy would be to limit the number of cost scenarios to maximize the contrasts between them and their corresponding CD matrices, because statistical inference cannot distinguish them if

they are too strongly correlated (Zeller et al., 2016). Besides, although cost values inferred from such an approach should be used carefully for locating LCPs, the CD matrices derived could be used for estimating the importance of the locations linked by LCPs for the connectivity of a whole network of patches using graph-theoretical connectivity metrics (Foltête et al., 2014). This could represent a reliable alternative to the spatial application of the results of this type of inference. However, note that when CD thresholds are used to define the connections between patches, we can expect the statistical distribution of CD values and consequently the range of the cost values to affect metric calculations.

#### *Methodological perspectives for LCP modeling*

Although the competing cost scenarios can be controlled until the very end of a study, the studied landscape is determined in the early stages. We here showed that in landscapes with high proportions of favorable land cover types, reduced land cover diversity and large contiguous patches, the correlation coefficients between CD matrices deriving from very different cost scenarios consistently reached high values. In such a case, it can be determined beforehand that the reliability of the cost value inference will be reduced, as already shown by Cushman et al. (2013). The sensitivity of LCPs to cost scenarios should therefore be tested prior to any study if the main objective is to infer the resistance to movements. For that purpose, we included the function `link_compar()` which computes the spatial overlap between several sets of LCPs within the `graph4lg` package in R (Savary et al., 2021). This function makes it possible to specify the width of the least cost corridors. Indeed we used here a constant total width of 400 m, which reflects the scale at which conservation measures can be implemented following connectivity modeling (Ford et al., 2020; Spackman and Hughes, 1995) and prevents overestimating the spatial overlap for short LCPs.

The lower sensitivity of LCP modeling outputs to the cost scenarios in landscapes with large proportions of favorable land cover types and large patches may be due to the sampling of points within forests. Although this reflects the fact that connectivity analyses aim at identifying favorable paths between similar habitat patches, it also means that whatever the forest cost value, these areas had to be crossed by LCPs and over larger distances in such landscapes. This could have reduced the differences between LCPs and CD matrices computed under different scenarios. Considering resistance distances using the circuit theory (McRae, 2006) could have decreased the correlations between distance matrices obtained with cost scenarios assigning different cost values to forests. In contrast, using current maps of connectivity obtained from the circuit theory would probably have increased the overall spatial overlap between the most similar cost scenarios due to the consideration of alternative LCPs which are potentially shared across similar scenarios.

Finally we raised concerns about the risk of identifying cost scenarios in data based inference leading to incorrect qualitative and spatial output, while being highly correlated with biological responses. Previous landscape genetic studies investigating the promises and pitfalls of cost surface parametrization from genetic data (Cushman et al., 2013; Graves et al., 2012, 2013; Koen et al., 2012; Spear et al., 2010) should be completed by considering our results.

## **Acknowledgements**

This study is part of a PhD project supported by the ARP-Astrance company under a CIFRE contract supervised and partly funded by the ANRT (Association Nationale de la Recherche et de la Technologie). This work is also part of the project CANON that was supported by the French "Investissements d'Avenir" program, project ISITE-BFC (contract ANR-15-IDEX-0003). We are particularly grateful to ARP-Astrance team for its constant support along the project. Part of the analyses were carried out on the calculation "Mésocentre" facilities of the University of Bourgogne-Franche-Comté. We thank Christopher Sutcliffe for revising the English manuscript

## Declaration of interest statement

The authors declare no conflict of interest.

## Data and Codes Availability Statement

The data that support the findings of this study are openly available in figshare at <https://figshare.com/articles/dataset>

Besides, the functions `sample_raster()`, `graphab_link()`, `mat_cost_dist()` and `link_compar()` respectively used for sampling points within landscapes, computing LCPs between these points, computing cost distances and comparing LCPs spatially have been included into the R package `graph4lg` and are directly available here: <https://cran.r-project.org/web/packages/graph4lg/index.html>. The OSO land cover raster data are available here: <https://www.theia-land.fr/product/carte-doccupation-des-sols-de-la-france-metropolitaine/>.

## A Supplementary figures

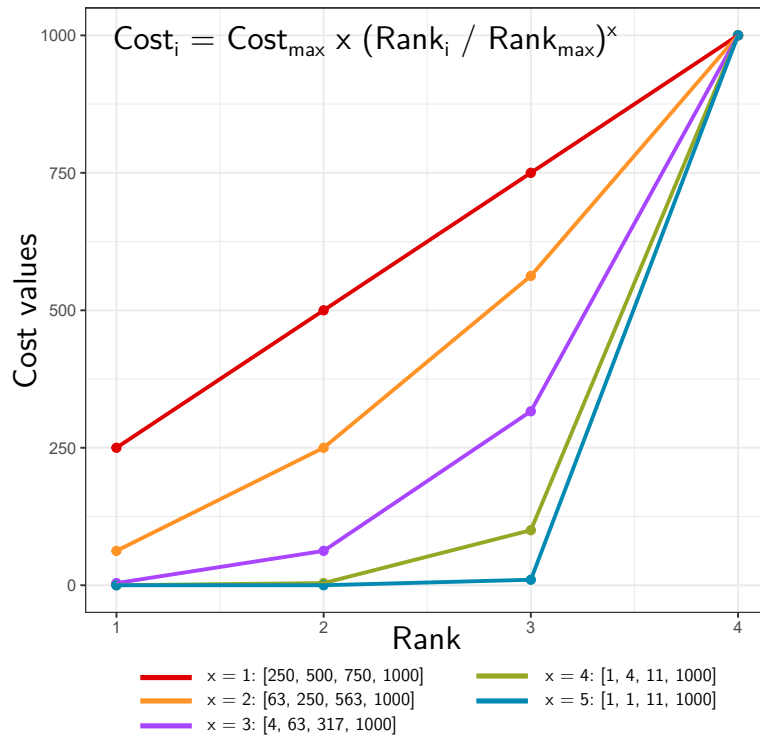


Figure 8: Gradient of cost values according to the alternative cost scenarios created using the method of Shirk et al. (2010).

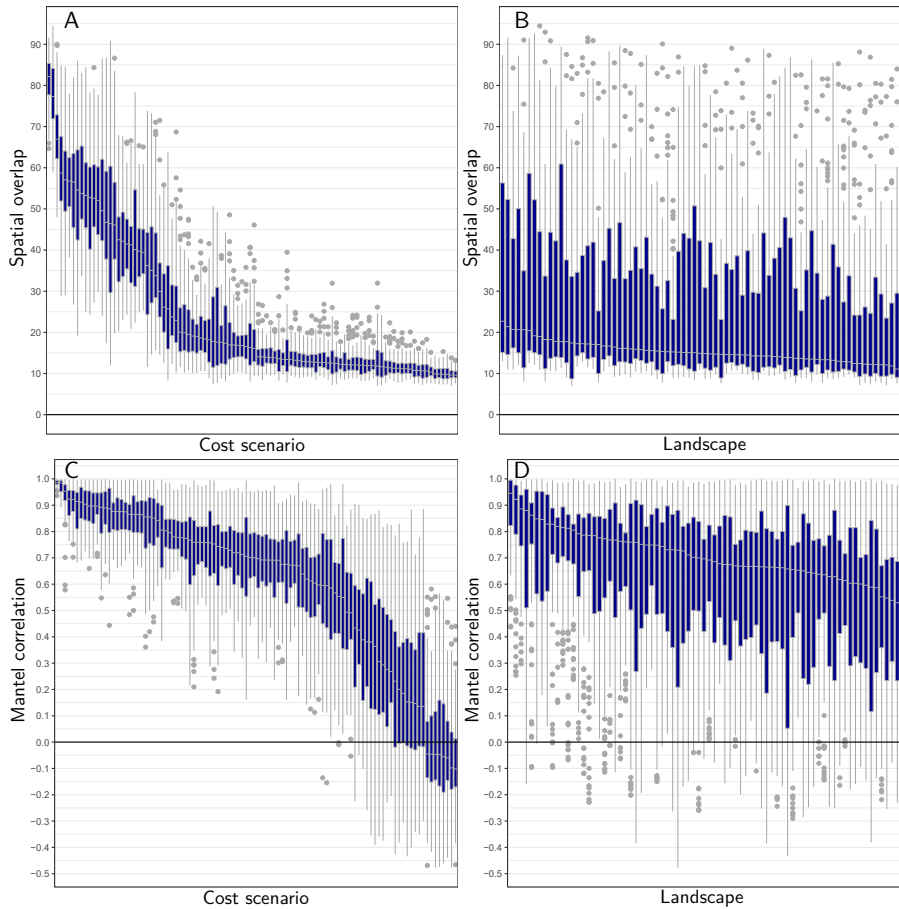


Figure 9: Distribution of the proportions of spatial overlap between the alternative and the true least cost paths (A, B) or the Spearman Mantel  $r$  correlation coefficients between the alternative cost distance matrices and the true one (C, D) according to the cost scenario (A, C) or the landscape (B, D). Cost scenarios (A, C) and landscapes (B, D) are placed in decreasing order of mean proportion of spatial overlap (A, B) or mean correlation coefficient (C, D) along the x axis. When the distribution is displayed for each cost scenario (A, C), the 100 boxes are made up of 77 values each corresponding to a landscape, whereas when it is displayed for each landscape (B, D), the 77 boxes are made up of 100 values each corresponding to a cost scenario.

## References

- Adriaensen, F., Chardon, J., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H., and Matthysen, E. (2003). The application of least-cost modelling as a functional landscape model. *Landscape and Urban Planning*, 64(4):233–247.
- Balbi, M., Petit, E. J., Croci, S., Nabucet, J., Georges, R., Madec, L., and Ernoult, A. (2019). Ecological relevance of least cost path analysis: An easy implementation method for landscape urban planning. *Journal of Environmental Management*, 244:61–68.
- Balkenhol, N., Pardini, R., Cornelius, C., Fernandes, F., and Sommer, S. (2013). Landscape-level comparison of genetic diversity and differentiation in a small mammal inhabiting different fragmented landscapes of the Brazilian Atlantic Forest. *Conservation Genetics*, 14(2):355–367.
- Beier, P., Majka, D. R., and Newell, S. L. (2009). Uncertainty analysis of least-cost modeling for designing wildlife linkages. *Ecological Applications*, 19(8):2067–2077.
- Beier, P., Majka, D. R., and Spencer, W. D. (2008). Forks in the road: choices in procedures for designing wildland linkages. *Conservation Biology*, 22(4):836–851.
- Bowman, J., Adey, E., Angoh, S. Y., Baici, J. E., Brown, M. G., Cordes, C., Dupuis, A. E., Newar, S. L., Scott, L. M., and Solmundson, K. (2020). Effects of cost surface uncertainty on current density estimates from circuit theory. *PeerJ*, 8:e9617.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*, 33(2):261–304.
- Carrascal, L. M., Galván, I., and Gordo, O. (2009). Partial Least Squares regression as an alternative to current regression methods used in ecology. *Oikos*, 118(5):681–690.
- Carroll, C., McRae, B., and Brookes, A. (2012). Use of linkage mapping and centrality analysis across habitat gradients to conserve connectivity of gray wolf populations in western North America. *Conservation Biology*, 26(1):78–87.
- Clevenger, A. P., Wierzchowski, J., Chruszcz, B., and Gunson, K. (2002). GIS-generated, expert-based models for identifying wildlife habitat linkages and planning mitigation passages. *Conservation Biology*, 16(2):503–514.
- Cushman, S. A., Shirk, A. J., and Landguth, E. L. (2013). Landscape genetics and limiting factors. *Conservation Genetics*, 14(2):263–274.
- de la Torre, J. A., Lechner, A. M., Wong, E. P., Magintan, D., Saaban, S., and Campos-Arceiz, A. (2019). Using elephant movements to assess landscape connectivity under peninsular malaysia’s central forest spine land use policy. *Conservation Science and Practice*, 1(12):e133.
- Duflot, R., Avon, C., Roche, P., and Bergès, L. (2018). Combining habitat suitability models and spatial graphs for more effective landscape conservation planning: An applied methodological framework and a species case study. *Journal for Nature Conservation*, 46:38–47.
- Etherington, T. R. and Holland, E. P. (2013). Least-cost path length versus accumulated-cost as connectivity measures. *Landscape Ecology*, 28(7):1223–1229.
- Etherington, T. R. and Perry, G. L. (2016). Visualising continuous intra-landscape isolation with uncertainty using least-cost modelling based catchment areas: common brushtail possums in the Auckland isthmus. *International Journal of Geographical Information Science*, 30(1):36–50.
- Foltête, J.-C., Girardet, X., and Clauzel, C. (2014). A methodological framework for the use of landscape graphs in land-use planning. *Landscape and Urban Planning*, 124:140–150.
- Ford, A. T., Sunter, E. J., Fauvelle, C., Bradshaw, J. L., Ford, B., Hutchen, J., Phillipow, N., and Teichman, K. J. (2020). Effective corridor width: linking the spatial ecology of wildlife with land use policy. *European Journal of Wildlife Research*, 66(4):1–10.
- Gonzales, E. K. and Gergel, S. E. (2007). Testing assumptions of cost surface analysis - a tool for invasive species management. *Landscape Ecology*, 22(8):1155–1168.
- Graves, T. A., Beier, P., and Royle, J. A. (2013). Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. *Molecular Ecology*, 22(15):3888–3903.
- Graves, T. A., Wasserman, T. N., Ribeiro, M. C., Landguth, E. L., Spear, S. F., Balkenhol, N., Higgins, C. B., Fortin, M.-J., Cushman, S. A., and Waits, L. P. (2012). The influence of landscape characteristics and home-range size on the quantification of landscape-genetics relationships. *Landscape Ecology*, 27(2):253–266.

- Gurrutxaga, M., Lozano, P. J., and del Barrio, G. (2010). GIS-based approach for incorporating the connectivity of ecological networks into regional planning. *Journal for Nature Conservation*, 18(4):318–326.
- Hanski, I., Moilanen, A., and Gyllenberg, M. (1996). Minimum viable metapopulation size. *The American Naturalist*, 147(4):527–541.
- Hoover, B., Yaw, S., and Middleton, R. (2020). CostMAP: an open-source software package for developing cost surfaces using a multi-scale search kernel. *International Journal of Geographical Information Science*, 34(3):520–538.
- Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., and Rodes, I. (2017). Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sensing*, 9(1):95.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.
- Kadoya, T. (2009). Assessing functional connectivity using empirical data. *Population ecology*, 51(1):5–15.
- Khimoun, A., Peterman, W., Eraud, C., Faivre, B., Navarro, N., and Garnier, S. (2017). Landscape genetic analyses reveal fine-scale effects of forest fragmentation in an insular tropical bird. *Molecular Ecology*, 26(19):4906–4919.
- Koen, E. L., Bowman, J., and Walpole, A. A. (2012). The effect of cost surface parameterization on landscape resistance estimates. *Molecular Ecology Resources*, 12(4):686–696.
- Lechner, A. M. and Rhodes, J. R. (2016). Recent progress on spatial and thematic resolution in landscape ecology. *Current Landscape Ecology Reports*, 1(2):98–105.
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer research*, 27(2):209–220.
- Marrotte, R. R. and Bowman, J. (2017). The relationship between least-cost and resistance distance. *PLoS ONE*, 12(3):e0174212.
- McGarigal, K. (1995). *FRAGSTATS: spatial pattern analysis program for quantifying landscape structure*, volume 351. US Department of Agriculture, Forest Service, Pacific Northwest Research Station.
- McRae, B. H. (2006). Isolation by resistance. *Evolution*, 60(8):1551–1561.
- McRae, B. H. and Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences*, 104(50):19885–19890.
- McRae, B. H. and Kavanagh, D. M. (2011). *Linkage mapper connectivity analysis software*. The Nature Conservancy, Seattle WA.
- Mimet, A., Clauzel, C., and Foltête, J.-C. (2016). Locating wildlife crossings for multispecies connectivity across linear infrastructures. *Landscape Ecology*, 31(9):1955–1973.
- Mony, C., Abadie, J., Gil-Tena, A., Burel, F., and Ernoult, A. (2018). Effects of connectivity on animal-dispersed forest plant communities in agriculture-dominated landscapes. *Journal of Vegetation Science*, 29(2):167–178.
- Murekatete, R. M. and Shirabe, T. (2018). A spatial and statistical analysis of the impact of transformation of raster cost surfaces on the variation of least-cost paths. *International Journal of Geographical Information Science*, 32(11):2169–2188.
- Panzacchi, M., Van Moorter, B., Strand, O., Saerens, M., Kivimäki, I., St Clair, C. C., Herfindal, I., and Boitani, L. (2016). Predicting the continuum between corridors and barriers to animal movements using step selection functions and randomized shortest paths. *Journal of Animal Ecology*, 85(1):32–42.
- Pérez-Espona, S., Pérez-Barbería, F., McLeod, J., Jiggins, C., Gordon, I., and Pemberton, J. (2008). Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). *Molecular Ecology*, 17(4):981–996.
- Peterman, W. E. and Pope, N. S. (2020). The use and misuse of regression models in landscape genetic analyses. *Molecular Ecology*, 30(1):37–47.
- Peterman, W. E., Winiarski, K. J., Moore, C. E., da Silva Carvalho, C., Gilbert, A. L., and Spear, S. F. (2019). A comparison of popular approaches to optimize landscape resistance surfaces. *Landscape Ecology*, 34(9):2197–2208.
- Pinto, N. and Keitt, T. H. (2009). Beyond the least-cost path: evaluating corridor redundancy using a graph-theoretic approach. *Landscape Ecology*, 24(2):253–266.
- Pressey, R. (2004). Conservation planning and biodiversity: assembling the best data for the job. *Conservation Biology*, 18(6):1677–1681.

- Pullinger, M. G. and Johnson, C. J. (2010). Maintaining or restoring connectivity of modified landscapes: evaluating the least-cost path model with multiple sources of ecological information. *Landscape Ecology*, 25(10):1547–1560.
- Rayfield, B., Fortin, M.-J., and Fall, A. (2010). The sensitivity of least-cost habitat graphs to relative cost surface values. *Landscape Ecology*, 25(4):519–532.
- Rayfield, B., Fortin, M.-J., and Fall, A. (2011). Connectivity for conservation: a framework to classify network measures. *Ecology*, 92(4):847–858.
- Ricketts, T. H. (2001). The matrix matters: effective isolation in fragmented landscapes. *The American Naturalist*, 158(1):87–99.
- Roy, K., Kar, S., and Das, R. N. (2015). Statistical methods in QSAR/QSPR. In *A primer on QSAR/QSPR modeling*, pages 37–59. Springer.
- Ruiz-González, A., Gurrutxaga, M., Cushman, S. A., Madeira, M. J., Randi, E., and Gómez-Moliner, B. J. (2014). Landscape genetics for the empirical assessment of resistance surfaces: the European pine marten (*Martes martes*) as a target-species of a regional ecological network. *PLoS ONE*, 9(10):e110552.
- Savary, P., Foltête, J.-C., Moal, H., Vuidel, G., and Garnier, S. (2021). graph4lg: a package for constructing and analysing graphs for landscape genetics in R. *Methods in Ecology and Evolution*, 12(3):539–547.
- Sawyer, S. C., Epps, C. W., and Brashares, J. S. (2011). Placing linkages among fragmented habitats: do least-cost models reflect how animals use landscapes? *Journal of Applied Ecology*, 48(3):668–678.
- Schadt, S., Knauer, F., Kaczensky, P., Revilla, E., Wiegand, T., and Trepl, L. (2002). Rule-based assessment of suitable habitat and patch connectivity for the Eurasian lynx. *Ecological Applications*, 12(5):1469–1483.
- Shirabe, T. (2016). A method for finding a least-cost wide path in raster space. *International Journal of Geographical Information Science*, 30(8):1469–1485.
- Shirk, A., Wallin, D., Cushman, S., Rice, C., and Warheit, K. (2010). Inferring landscape effects on gene flow: a new model selection framework. *Molecular Ecology*, 19(17):3603–3619.
- Simpkins, C. E., Dennis, T. E., Etherington, T. R., and Perry, G. L. (2017). Effects of uncertain cost-surface specification on landscape connectivity measures. *Ecological Informatics*, 38:1–11.
- Simpkins, C. E., Dennis, T. E., Etherington, T. R., and Perry, G. L. (2018). Assessing the performance of common landscape connectivity metrics using a virtual ecologist approach. *Ecological Modelling*, 367:13–23.
- Spackman, S. C. and Hughes, J. W. (1995). Assessment of minimum stream corridor width for biological conservation: species richness and distribution along mid-order streams in Vermont, USA. *Biological Conservation*, 71(3):325–332.
- Spear, S. F., Balkenhol, N., Fortin, M.-J., McRae, B. H., and Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, 19(17):3576–3591.
- Tenenhaus, M. (1998). *La régression PLS: théorie et pratique*. Editions TECHNIP.
- Therneau, T. M., Atkinson, B., and Ripley, M. B. (2010). The rpart package.
- Wang, Y.-H., Yang, K.-C., Bridgman, C. L., and Lin, L.-K. (2008). Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, 23(8):989–1000.
- Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory systems*, 58(2):109–130.
- Zeller, K. A., Creech, T. G., Millette, K. L., Crowhurst, R. S., Long, R. A., Wagner, H. H., Balkenhol, N., and Landguth, E. L. (2016). Using simulations to evaluate Mantel-based methods for assessing landscape resistance to gene flow. *Ecology and Evolution*, 6(12):4115–4128.
- Zeller, K. A., Jennings, M. K., Vickers, T. W., Ernest, H. B., Cushman, S. A., and 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.
- Zeller, K. A., McGarigal, K., and Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape Ecology*, 27(6):777–797.