

USE OF LAND FACETS TO DESIGN CONSERVATION CORRIDORS:  
CONSERVING THE ARENAS, NOT THE ACTORS

Brian M. Brost

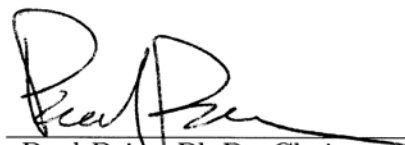
A Thesis

Submitted in Partial Fulfillment  
of the Requirements for the Degree of  
Master of Science  
in Forestry

Northern Arizona University

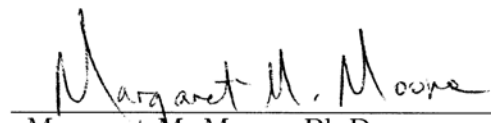
May 2010

Approved:



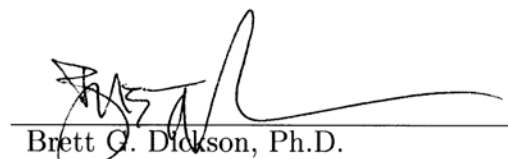
---

Paul Beier, Ph.D., Chair



---

Margaret M. Moore, Ph.D.



---

Brett G. Dickson, Ph.D.

## ABSTRACT

# USE OF LAND FACETS TO DESIGN CONSERVATION CORRIDORS: CONSERVING THE ARENAS, NOT THE ACTORS

BRIAN M. BROST

Extensive shifts in species' geographical distributions have been the most important mechanism through which plants and animals coped with previous large-scale climate changes. Climate-induced movements have also begun during the past century, and it is likely that many species will only persist if they are capable of colonizing newly suitable habitat. Consequently, many advocate the need for conservation corridors and linkages as a means to support species' range shifts.

Least-cost modeling for focal species is the most widely used method for designing corridors and linkages. However, these linkages have been based on today's species' distributions and land cover, factors that will change with large-scale climate change. One method to develop corridors that accommodate species' shifting distributions is to incorporate climate models into their design. But this approach is enormously complex and prone to error propagation because it involves many linked, highly-uncertain components (emissions scenarios, global and regional circulation models, climate envelope models). An alternative is to design linkages for the continuity and interspersed of land facets, or recurring landscape units of relatively uniform topography and soils. The central concept underlying this coarse-filter approach is diverse physical environments support diverse species today, and will interact with future climates to support new assemblages of species in the future. Thus, the goal of designing linkages based on land facets is to conserve the arenas of biological activity rather than the temporary occupants of those arenas.

We illustrate how land facets can be defined in a rule-based and adaptable way, and how they can be used for linkage design in the face of climate change. To demonstrate the flexibility of our procedures, we designed linkages using land facets in three topographically diverse landscapes in Arizona, USA. To determine if these land facets linkages are likely to support movement by species needing connectivity, we compared them to linkages designed for focal species in the same three landscapes.

We also assessed how well the focal species linkages served land facets. We found that linkage designs based on land facets served 25 of 28 focal species as well as or better than the focal species designs. Focal species linkages, on the other hand, provided a similar degree of connectivity as the land facets designs for only 21 of 32 land facets.

In areas where a focal species approach to linkage design is not possible, our results suggest that conservation practitioners can solely implement a land facets approach with some confidence that the linkage design would serve most potential focal species. In areas where focal species designs are possible, we recommend using the land facets approach to complement, rather than replace, focal species approaches.

## ACKNOWLEDGMENTS

This thesis would not have been possible without the help and support of professors, staff, students, friends, and family.

First, thanks to Paul Beier. Paul's door was always open to me over the past three years, and I have grown tremendously as a result of our countless (seriously!) meetings. Thanks also to Margaret Moore and Brett Dickson for their thoughtful and constructive comments. Their suggestions certainly improved the quality of this thesis. The Beier-Chamber's Lab of Conservation Biology and Wildlife Ecology—particularly members Bill Bridgeland, Tabitha Graves, Suzanne Hagell, and Jeff Jenness—also provided helpful feedback and suggestions. Steve Andariese provided valuable IT support and generously allowed me to usurp a graduate lab computer or two for several months. And, of course, thanks to Hondo Brisbin, Donna Peppin, and Larissa Yocom for giving me an extra reason to come into the office each day.

I have Phil Turk and Brent Burch to thank for my development as a quantitative ecologist. Both are enthusiastic instructors and their courses were truly the highlight of my Northern Arizona University experience. Thanks to both for giving me something new to be excited about.

Over the last thirty-one years, my mother and father, Lynn and Mike Brost, were a ceaseless source of encouragement. Their patience and support made this thesis possible.

This research was supported by the United States Department of Agriculture McIntire-Stennis Cooperative Forestry Research Program, The Arizona Board of Forest Research, and a research joint venture agreement (#08-JV-11221636-171) with the United States Department of Agriculture Forest Service Rocky Mountain Research Station.

# Contents

<b>List of Tables</b>	<b>vii</b>
<b>List of Figures</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Use of Land Facets to Design Linkages for Climate Change</b>	<b>4</b>
Introduction . . . . .	5
Methods . . . . .	7
<i>Linkage planning areas</i> . . . . .	8
<i>Defining land facets</i> . . . . .	8
<i>Corridor design using land facets</i> . . . . .	12
<i>Corridor design for high diversity of facets</i> . . . . .	14
<i>Modifying resistance surfaces or corridor termini to produce corridors</i> <i>that better capture the focal land facet</i> . . . . .	15
<i>Linkage design</i> . . . . .	15
Results . . . . .	16
Discussion . . . . .	18
<i>Use of land facets in linkage design</i> . . . . .	21
<b>3 Comparing Linkage Designs Based on Land Facets to Linkage Designs Based on Focal Species</b>	<b>38</b>
Introduction . . . . .	39
Methods . . . . .	41
<i>Linkage planning areas and linkage designs</i> . . . . .	41
<i>Evaluating how well each linkage design served each focal species</i> . . .	42
<i>Evaluating how well each linkage design provided continuity for land</i> <i>facets</i> . . . . .	44
Results . . . . .	46

<i>How well were focal species served by each linkage design?</i> . . . . .	46
<i>How well were land facets served by each linkage design?</i> . . . . .	48
Discussion . . . . .	49
<i>How well do linkages designed for land facets serve focal species?</i> . . .	49
<i>How well do linkages designed for focal species serve land facets?</i> . . .	50
<i>Implications for linkage design</i> . . . . .	51
<b>4 Conclusions</b>	<b>82</b>
<b>Literature Cited</b>	<b>84</b>
<b>Appendix A. R Functions for Defining Land Facets</b>	<b>91</b>
<b>Appendix B. Source Code for R Functions</b>	<b>97</b>

# List of Tables

2.1	Descriptions of planning areas and wildland blocks . . . . .	23
2.2	Characteristics of linkage designs based on land facets . . . . .	24
2.3	Mean and range of topographic attributes encompassed by individual land facets . . . . .	25
3.1	Location, number of focal species and land facets modeled, and size of linkage designs . . . . .	53
3.2	Relative performance of linkage designs with respect to focal species .	54
3.3	Distances between breeding patches for focal species . . . . .	55
3.4	Mean resistance of the resistance profiles and longest high-resistance segments in resistance profiles for land facets . . . . .	56

# List of Figures

2.1	Map of the land facets linkage design for the Black Hills-Munds Mountain planning area . . . . .	27
2.2	Map of the land facets linkage design for the Wickenburg-Hassayampa planning area . . . . .	28
2.3	Map of the land facets linkage design for the Santa Rita-Tumacacori planning area . . . . .	29
2.4	Sequence of operations used to define land facets . . . . .	30
2.5	Example kernel density estimation used to identify outliers . . . . .	31
2.6	Example output of fuzzy <i>c</i> -means cluster analysis depicting composition of land facets within a single topographic position . . . . .	32
2.7	Example cluster validity indices used to determine the optimal fuzzy- <i>c</i> partition for a data set . . . . .	33
2.8	Example corridor termini . . . . .	34
2.9	Sequence of operations used to design a corridor for one land facet . .	35
2.10	Example resistance surface used in least-cost corridor analysis . . . .	36
2.11	Univariate distributions demonstrating the complexity of topographic positions by planning area . . . . .	37
3.1	Map of the land facets and focal species linkage designs for the Black Hills-Munds Mountain planning area . . . . .	59
3.2	Map of the land facets and focal species linkage designs for the Wickenburg-Hassayampa planning area . . . . .	60
3.3	Map of the land facets and focal species linkage designs for the Santa Rita-Tumacacori planning area . . . . .	61
3.4	Metrics used to evaluate the performance of linkage designs with respect to focal species . . . . .	62
3.5	Resistance profiles for species with locally widespread habitat and mountain lion in the Black Hills-Munds Mountain planning area . . .	63



3.6	Resistance profiles for species with locally widespread habitat and desert bighorn sheep in the Wickenburg-Hassayampa planning area . . . . .	64
3.7	Resistance profiles for species with locally widespread habitat in the Santa Rita-Tumacacori planning area . . . . .	66
3.8	Resistance profiles corresponding to the gaps between breeding patches for focal species in the Black Hills-Munds Mountain planning area . . . . .	68
3.9	Resistance profiles corresponding to the gaps between breeding patches for focal species in the Santa Rita-Tumacacori planning area . . . . .	69
3.10	Resistance profiles for land facets in the Black Hills-Munds Mountain planning area . . . . .	72
3.11	Resistance profiles for land facets in the Wickenburg-Hassayampa planning area . . . . .	75
3.12	Resistance profiles for land facets in the Santa Rita-Tumacacori planning area . . . . .	78
3.13	Resistance profiles for high diversity of land facets in the Black Hills-Munds Mountain, Wickenburg-Hassayampa, and Santa Rita-Tumacacori planning areas . . . . .	81

## Preface

This thesis has been prepared in journal format. Consequently, there is some redundancy among the chapters. If the reader is interested in the short version minus the details, I would recommend reading the introductory and concluding chapters (Chapters 1 and 4). Otherwise, Chapters 2 and 3 contain the specifics. In the coming months, manuscripts developed from Chapters 2 and 3 will be submitted to *Ecological Applications* and *Conservation Biology*, respectively.

# Chapter 1

## Introduction

Extensive shifts in species' geographical distributions have been the most important mechanism through which plants and animals coped with previous large-scale climate changes (Graham & Grimm 1990; Huntley 2005). Climate-induced movements have also begun during the past century (Hersteinsson & Macdonald 1992; Grabherr et al. 1994; Parmesan 1996; Pauli et al. 1996; Parmesan et al. 1999; Pounds et al. 1999; Thomas & Lennon 1999). Though some species may be capable of adapting (*sensu* evolution) to future climatic conditions (Skelly et al. 2007; Millar et al. 2007), it is likely that many species will only persist if they are capable of colonizing newly suitable habitat (Williams et al. 2005). Consequently, many advocate the need for conservation corridors and linkages as a means to support species' range shifts (Halpin 1997; Noss 2001; Hannah et al. 2002; Noss & Daly 2006; Mawdsley et al. 2009).

Least-cost modeling for focal species is the most widely used method for designing corridors to connect protected areas (e.g., Walker & Craighead 1997; Singleton et al. 2002; Beier et al. 2006, 2007). However, like most other conservation plans, these corridor designs have been based on today's species' distributions and land cover, both of which will change as climate changes (Hunter et al. 1988; Huntley 2005). Thus, it is uncertain how well focal species linkages will function when some species currently occupying an area may no longer do so in the near future, while others species may be new arrivals.

One method to develop corridors that accommodate species' shifting distributions is to incorporate climate models into their design. Williams et al. (2005) used this approach to identify dispersal chains for individual species through 2050, each chain consisting of temporally and spatially contiguous habitat intended to allow a species to shift its range in response to climate change. Unfortunately, incorporating climate models into corridor design involves many linked, highly-uncertain components (e.g.,

emissions scenarios, global circulation models, regional circulation models, and models of climate envelopes). Therefore, modeling climate and habitat suitability is an enormously complex process, and errors and uncertainty propagate through each of its components. Furthermore, Beale et al. (2008) found that species-climate associations determined from climate envelope modeling performed no better than chance for predicting the current distributions of 68 of 100 European bird species.

Hunter et al. (1988) and Beier and Brost (2010) suggested an alternative coarse-filter strategy to conserve biodiversity in light of climate change. They argued that “basing the coarse-filter approach on physical environments as ‘arenas’ of biological activity, rather than...the temporary occupants of those arenas,” may be a better way to maintain a high level of biodiversity for long-term persistence (Hunter et al. 1988:380). This strategy operates on the premise that diverse physical environments support diverse species (Kirkpatrick & Brown 1994; Faith & Walker 1996; Burnett et al. 1998; Nichols et al. 1998; Cowling et al. 1999; Reyers et al. 2002) today, and will interact with future climates to support new assemblages of species in the future. Protecting diverse physical environments may also ensure the persistence of the ecological and evolutionary processes that maintain and generate biodiversity (Cowling et al. 1999; Noss 2001; Moritz 2002; Cowling et al. 2003; Rouget et al. 2006; Pressey et al. 2007; Klein et al. 2009). Thus, a linkage designed to provide continuity for all physical environments should not only optimize connectivity for the full diversity of plants and animals, but also sustain vital processes.

In Chapter 2 of this thesis, we illustrate this approach by designing linkages based on physical environments in three topographically diverse landscapes in Arizona, USA. We demonstrate how to define *land facets*, or recurring areas of relatively homogenous topography and soils, in a rule-based and adaptable way, and how they can be used for linkage design in the face of climate change. Each linkage includes multiple corridors, one to optimize connectivity for each land facet type. Each of these corridors should facilitate movement of species associated with that facet, today and in the future. To better accommodate rapid, short-distance range shifts, interactions between species, and ecological and evolutionary processes such as speciation, these designs also include a corridor to optimize connectivity for high diversity (i. e., interspersed) of land facets.

Although Beier and Brost (2010) recommend using land facets in conjunction with focal species to design linkages, conservation practitioners may be limited to a land facets approach in areas where species information is poor or maps of land cover are non-existent. Such practitioners might like to know whether a linkage

design based on land facets is likely to support movement by local species needing connectivity. Conversely, practitioners who have previously designed a linkage for focal species might also want to know if the design provides for the continuity and interspersions of land facets, or whether additional analyses are needed to make the design better capture physical environments. In Chapter 3, we address these questions by evaluating how well linkages designed for land facets provide connectivity for focal species, and how well linkages designed for focal species provide continuity of land facets.

## Chapter 2

# Use of Land Facets to Design Linkages for Climate Change

**Abstract.** Least-cost modeling for focal species is the most widely used method for designing conservation corridors and linkages. However, these linkages have been based on today's species' distributions and land cover, factors that will change with large-scale climate change. One method to develop corridors that accommodate species' shifting distributions is to incorporate climate models into their design. But this approach is enormously complex and prone to error propagation because it involves many linked, highly-uncertain components (emissions scenarios, global and regional circulation models, climate envelope models). A simpler alternative is to design linkages for the continuity and interspersion of land facets, or recurring landscape units of relatively uniform topography and soils. This coarse-filter approach aims to conserve the arenas of biological activity rather than the temporary occupants of those arenas. In this chapter, we demonstrate how land facets can be defined in a rule-based and adaptable way, and how they can be used for linkage design in the face of climate change. We used fuzzy *c*-means cluster analysis to define land facets with respect to four topographic variables (elevation, slope angle, solar insolation, and topographic position), and least-cost analysis to design linkages to include one corridor per land facet and another for the interspersion of land facets. To demonstrate the flexibility of our procedures, we designed linkages using land facets in three topographically diverse landscapes in Arizona, USA. Other variables, including soils information, could be used to define land facets in other landscapes. We advocate using land facets to complement, rather than replace, existing focal species approaches to linkage design. This approach can be applied anywhere and is not affected by the bias and patchiness common in species occurrence data.

## Introduction

Extensive shifts in species' geographical distributions have been the most important mechanism through which plants and animals coped with previous large-scale climate changes (Graham & Grimm 1990; Huntley 2005). Climate-induced movements have also begun during the past century (Hersteinsson & Macdonald 1992; Grabherr et al. 1994; Parmesan 1996; Pauli et al. 1996; Parmesan et al. 1999; Pounds et al. 1999; Thomas & Lennon 1999). Though some species may be capable of adapting (*sensu* evolution) to future climatic conditions (Skelly et al. 2007; Millar et al. 2007), it is likely that many species will only persist if they are capable of colonizing newly suitable habitat (Williams et al. 2005). Consequently, many advocate the need for conservation corridors and linkages as a means to support species' range shifts (Halpin 1997; Noss 2001; Hannah et al. 2002; Noss & Daly 2006; Mawdsley et al. 2009).

Least-cost modeling for focal species is the most widely used method for designing corridors to connect wildland blocks (e.g., Walker & Craighead 1997; Singleton et al. 2002; Beier et al. 2006, 2007). The objective of this approach is to identify the swath of land that minimizes the ecological cost of movement through a landscape for a species (Adriaensen et al. 2003; Beier et al. 2008). Each swath of land represents a corridor, and corridors for multiple focal species are combined into a linkage design. Like most other conservation plans, these corridor designs have been based on today's species' distributions and land cover. However, it is likely that some species currently occupying a given area may no longer do so in the near future, while other species may be new arrivals.

One method to develop corridors that accommodate species' shifting distributions is to incorporate climate models into their design. We are aware of two efforts that use this approach, both for the Cape Proteaceae of South Africa. Williams et al. (2005) identified dispersal chains for individual species through 2050, each chain consisting of temporally and spatially contiguous habitat intended to allow a species to shift its range in response to climate change. Phillips et al. (2008) optimized this framework by incorporating network flow concepts into the selection of dispersal chains. Incorporating climate models into corridor design involves many linked components—emissions scenarios, general circulation models, regional circulation models, and models of climate envelopes—each of which unfortunately contains some uncertainty. For example, emissions scenarios differ 6-fold in predicted annual CO<sub>2</sub> emissions by the year 2100 and climate projections differ vastly among the seven commonly-used general circulation models (IPCC 2001; Raper & Giorgi 2005).

Divergence increases further among regional circulation models which project outputs from a general circulation model onto a scale more useful for modeling habitat change. Climate envelope models require additional assumptions and necessarily exclude some important components (e.g., species interactions and altered disturbance regimes) that influence species' distributions (Williams et al. 2005). Therefore, modeling climate and habitat suitability is an enormously complex process, and errors and uncertainty propagate through each of its components. Furthermore, Beale et al. (2008) found that species-climate associations determined from climate envelope modeling performed no better than chance for predicting the current distributions of 68 of 100 European bird species.

Hunter et al. (1988) and Beier and Brost (2010) suggested a coarse-filter strategy to conserve biodiversity in light of climate change. Conventional coarse-filter conservation strategies target biotic communities as the unit of conservation (Noss 1987), but these communities will not respond predictably to climate change. In fact, many communities in existence today are < 8000 years old, each component species having responded individually to past environmental changes (Webb 1987; Hunter et al. 1988). Alternatively, they argued that “basing the coarse-filter approach on physical environments as ‘arenas’ of biological activity, rather than on communities, the temporary occupants of those arenas,” may be a better way to maintain a high level of biodiversity for long-term persistence (Hunter et al. 1988:380).

The central concept underlying this approach is diverse physical environments support diverse species (Kirkpatrick & Brown 1994; Faith & Walker 1996; Burnett et al. 1998; Nichols et al. 1998; Cowling et al. 1999; Reyers et al. 2002) today, and will interact with future climates to support new assemblages of species in the future. Protecting diverse physical environments may also ensure the persistence of the ecological and evolutionary processes that maintain and generate biodiversity (Cowling et al. 1999; Noss 2001; Moritz 2002; Cowling et al. 2003; Rouget et al. 2006; Pressey et al. 2007; Klein et al. 2009). Thus, a linkage designed to provide continuity for all physical environments should not only optimize connectivity for the full diversity of plants and animals, but also sustain vital processes.

Two efforts have considered physical environments in linkage design, but without a clear, objective landscape classification scheme. Beier et al. (2006, 2007) evaluated the topographic composition of preliminary linkages designed to serve multiple focal species, and expanded some of them in an ad-hoc manner to better represent elevation, slope angle, aspect, and landform classes. Rouget et al. (2006) designed conservation corridors to capture large-scale processes, including biotic response to climate change,



by aligning corridors with upland-lowland and macroclimatic gradients.

The objective of this chapter is to demonstrate how physical units can be defined in a rule-based and adaptable way, and how they can be used for linkage design in the face of climate change. Wessels et al. (1999) defined *land facets* as recurring areas of relatively homogenous topography and soils. Because soil maps in our planning areas were incomplete and contained unmapped heterogeneity, in this illustration we define land facets based only on topographic variables. However, our approach uses both categorical and continuous variables and can readily be adapted to accommodate categorical soil variables (e. g., soil type) and continuous soil variables (e. g., soil depth or moisture).

To illustrate our approach, we developed linkage designs based on land facets for three topographically diverse landscapes in Arizona, USA. Like linkages designed for multiple focal species, we designed land facets linkages to include multiple corridors, one to optimize connectivity for each land facet. Each of these corridors is intended to support movement by species associated with that facet, today and in the future. To better accommodate rapid, short-distance range shifts, interactions between species, and ecological and evolutionary processes such as speciation, we also included a corridor to optimize connectivity for the diversity (i. e., high interspersed) of land facets (Cowling et al. 1999; Fairbanks et al. 2001; Rouget et al. 2006). Where necessary, we added reaches of major streams and rivers in the planning area because rivers and drainages promote the movement of animals, sediment, water, and nutrients (Beier & Brost 2010).

## Methods

We used a combination of ArcGIS 9.3 (ESRI, Redlands, California) and R statistical software (R Development Core Team 2009) to define land facets (Fig. 2.4), and developed procedures to delineate corridors entirely within ArcGIS 9.3 (Fig. 2.9). We packaged these procedures into R functions and an ArcGIS extension that allow other users to modify each critical value, such as what fraction of cells to exclude as outliers. In the Methods section, we state the thresholds that produced reasonable results in the three landscapes we analyzed. Users of our packaged tools can specify alternate critical values.

*Linkage planning areas*

We designed linkages based on land facets for three landscapes in Arizona, USA (Table 2.1; Figs. 2.1–2.3). Detailed information regarding each area’s ecological significance, existing conservation investments, threats to connectivity, and patterns of land ownership and land cover are available in Beier et al. (2007).

Beier et al. (2007) delineated two protected wildland blocks to be connected in each landscape (Figs. 2.1–2.3). Each wildland block is a large area without highways or major paved roads and owned by agencies with a mandate to retain the land in natural condition. We refer to the land between and around the wildland blocks as *matrix*. We defined the analysis areas to include both wildland blocks and enough matrix (especially topographically diverse matrix) to allow our procedures to identify highly nonlinear corridors (Beier et al. 2008).

*Defining land facets*

Land facets were defined using three continuous variables, namely elevation, slope angle, and annual solar insolation, and one categorical variable, topographic position (Fig. 2.4). These variables were derived from the 1 arc-second (i.e., 30-m resolution) United States Geological Survey National Elevation Dataset, which typically has a vertical root-mean-square error  $\leq 7$  m (United States Geological Survey 2000. US GeoData Digital Elevation Models factsheet, Reston, VA, USA). Slope angle and solar insolation were computed using the Spatial Analyst extension in ArcGIS 9.3 (solar insolation model developed by Fu & Rich 2002). We computed annual solar insolation by summing instantaneous radiation at half-hour intervals for one day per month over a calendar year. Insolation integrates the effect of latitude, aspect, slope angle, elevation, daily and seasonal changes in sun angle, and topographic shading on incoming solar radiation for a location. Using the CorridorDesigner toolbox to ArcGIS 9.3 (Majka et al. 2007), we assigned each 30 x 30-m grid cell to one of three topographic positions, namely canyon bottom, ridge, or slope, by subtracting the elevation of a focal cell from the mean elevation of cells within a 200-m radius. We defined cells with differences  $\leq -8$  m as canyon bottoms, cells with differences  $\geq +8$  m as ridges, and cells with differences between  $-8$  and  $+8$  m as slopes. Prior to calculating solar insolation, we buffered the analysis area by 5 km to account for shading by distant topography; a 200-m buffer was used to address edge effects on calculations for slope angle and topographic position.

Each of these variables represents an indirect (elevation, slope angle, and topo-

graphic position) or direct (solar insolation) ecological gradient that is biologically important in mountainous landscapes (Davis & Goetz 1990; Parker 1995; Valverde et al. 1996; Pinder et al. 1997; Bolstad et al. 1998; Guisan et al. 1999; Franklin et al. 2000; Pfeffer et al. 2003). Elevation is associated with gradients in temperature and precipitation (Franklin 1995; Urban et al. 2000). Slope angle influences the velocity of water runoff and is therefore related to soil moisture content and soil development (Gallant & Wilson 2000; Lookingbill & Urban 2004). Insolation is related to heat load, photosynthetic potential, evaporation, transpiration, and near surface soil moisture content (Moore et al. 1991; Lookingbill & Urban 2004). Topographic position captures the major components of land curvature, namely canyon bottoms and ridges, and is easily interpreted. It also represents differences in substrate development (canyon bottoms are depositional environments whereas ridges and steep slopes are erosional), and is related to sun and wind exposure (Valverde et al. 1996).

We used numerical clustering procedures to define land facets based on the attributes of cells inside of the wildland blocks only, thereby ensuring that our classification reflected the topographic composition of the wildland blocks. Although some cluster analyses can be performed on continuous and categorical variables simultaneously, we felt that a land facet with homogeneous topographic position would be more interpretable than a land facet that included a mixture of canyon bottoms, ridges, and slopes. Therefore, we defined land facets by first sorting on topographic position, and then clustering on combinations of values for the continuous variables within each topographic position.

We divided slopes into land facets based on all three continuous variables (elevation, slope angle, and insolation), but used only slope angle and elevation to divide ridges and canyon bottoms into land facets. Insolation was not used to identify subclasses of ridges or canyon bottoms because these topographic positions are often symmetrical features. Therefore, a classification that used insolation to define land facets within ridges and canyon bottoms would identify different facets for their opposing sides, such as north-facing and south-facing ridgelines, despite their otherwise similarity. This unnecessarily complicates corridor design because the opposing sides of canyon bottoms and ridges are close to each other and can be treated as a unit for conservation purposes.

We used kernel density estimation to identify outliers, i. e., cells with combinations of values for continuous variables that rarely occurred in the wildland blocks. These cells often occur in isolated patches and are limited to a small portion of the landscape or may only occur in one wildland block. Outliers produce less compact clusters (i. e.,

cells within a cluster span a larger range in attribute space) with a diluted ecological significance. Outliers also shift the position of the cluster centroids (the definition or central concept of the cluster) toward a sparser region of attribute space.

Kernel density estimation is a non-parametric procedure that estimates the probability density function of a random variable (Silverman 1986). We used package *ks* (Duong 2009) in R statistical software for multivariate kernel density estimation. Because our data sets were large, we were forced to group individual cells according to their attribute values into bins of equal interval across the range of each variable. Cells in the canyon bottom and ridge topographic positions were grouped into a 2-dimensional array of  $151^2$  bins; for cells in the slope topographic position, a 3-dimensional array of  $91^3$  bins was used. We selected these array sizes because they present a compromise between resolution and computational efficiency. A diagonal bandwidth matrix, which allows smoothing in the directions parallel to the coordinate axes in the kernel density estimation, was selected using the iterative “plug-in” approach implemented in package *ks* (Duong 2007). Individual cells were then assigned the kernel density estimate of the bins into which they were grouped.

Cells were identified as outliers if the combination of their attributes occurred in the 10<sup>th</sup> percentile “tail” of the multivariate distribution generated from the kernel density estimation (Fig. 2.5). Thus, outliers were identified based not only on their distance from the data’s centroid, but also on the data’s multivariate shape. Because outliers were defined relative to cells inside of the wildland blocks, the proportion of cells in the matrix classified as outliers was higher or lower than 10% (e.g., higher than 10% in Fig. 2.5) when the matrix topography differed from that of the wildland blocks. Although the confusion index (see last paragraph of this subsection) can be used to identify outliers (i.e., extragrades), it leads to less compact clusters than using kernel density estimation for this purpose.

Next, we used fuzzy *c*-means cluster analysis to classify the non-outliers (Fig. 2.6a). Fuzzy *c*-means cluster analysis is an iterative procedure of finding the *c* partitions in a data set that minimizes the within-cluster variances of the classified objects (Bezdek 1981). The number of clusters, *c*, is defined by the user. Fuzzy *c*-means cluster analysis is capable of analyzing large data sets and is not sensitive to case order (i.e., the order of cells in the input file) like two-step cluster analysis (SPSS, Chicago, Illinois). Furthermore, unlike other classification methods that assign each object to one and only one class, fuzzy *c*-means cluster analysis assigns each observation membership to all *c* clusters. Membership ranges between 0 and 1, with larger values indicating higher similarity between an object and a cluster centroid. We used

package *e1071* (Dimitriadou et al. 2009) in R statistical software for fuzzy  $c$ -means cluster analysis.

Membership of cell  $i$  to cluster  $j$  is calculated as

$$\mu_{ij} = \frac{[(d_{ij})^2]^{-1/(\varphi-1)}}{\sum_{j'=1}^c [(d_{ij'})^2]^{-1/(\varphi-1)}}, \quad (2.1)$$

where  $d_{ij}$  is the Euclidean distance (in attribute space) between cell  $i$  and cluster centroid  $j$ ,  $\varphi$  is the “fuzziness” parameter used in the cluster analysis, and  $j'$  is an index over all  $c$  clusters (Burrough et al. 2000). We used  $\varphi = 1.5$  because it represents a compromise between a crisp classification with non-overlapping clusters ( $\varphi = 1$ ) and larger values giving a fuzzier classification (Burrough et al. 2000). The denominator in Eq. (2.1) standardizes the membership values, thus  $\sum_{j=1}^c \mu_{ij} = 1$  for all  $i$ .

Prior to clustering, we standardized variables with respect to the mean and standard deviation of the cells included in the analysis (i.e., cells within a topographic position inside of the wildland blocks). Next, we identified the optimal number of clusters,  $k$ , as the classification that best corresponded to the natural multivariate structure in the continuous variables. To determine  $k$ , we varied  $c$  in the range of  $2 \leq c \leq 10$  and computed eight cluster validity indices for each value of  $c$  (Fig. 2.7). No evidence suggested an optimal solution exceeding five clusters for any of our data sets. Each of the indices is based on the compactness within and/or separation between clusters. Due to the large size of our data sets, we could not use cluster validity criteria such as Dunn’s Index (Xie & Beni 1991) or the Average Silhouette Width Criterion (Campello & Hruschka 2006), each of which requires all pairwise distances between cells. We identified  $k$  as the number of clusters  $c$  that produced the largest marginal improvements in all or most of the eight indices. Situations in which the indices did not clearly indicate a single optimal number of clusters are addressed in the Discussion.

We performed 100 iterations of the fuzzy  $c$ -means cluster analysis for each  $c$  (i.e.,  $2 \leq c \leq 10$ ) to detect cases in which more than one partition for a given  $c$  minimized the within-cluster variance (e.g.,  $c \geq 7$  in the Xie-Beni plot in Fig. 2.7). If this occurred for the optimal number of clusters, the fuzzy- $c$  partition with the best validity index values was selected. For the one landscape with  $> 1.25$  million cells in a topographic position (namely the slope topographic position in the Wickenburg-Hassayampa planning area), each iteration was based on a random sample of 1.25

million cells to avoid memory limitations in R, and a new sample was taken for each iteration. We found that fuzzy  $c$ -means cluster analysis is robust to sample size, even for samples as small as 10% of the original population.

Using Eq. (2.1), fuzzy membership values to the  $k$  optimal clusters were computed for all non-outlier cells in the planning area within the respective topographic position. To gauge how well each cell was classified, we computed a confusion index as the ratio of its first sub-dominant membership value to its dominant membership value (Burrough et al. 2000). If the confusion index is near 0, then the cell is highly associated with the cluster centroid to which it has greatest membership; if the cell's confusion index is near 1, then the difference between the two highest memberships is small, and there is confusion about the most closely associated centroid (Fig. 2.6b). Poorly classified cells are either “intragrades” that are intermediate between centroids, or “extragrades” that belong equally little to all of the centroids. Following Burrough et al. (2000), we considered cells with a confusion index  $> 0.6$  as poorly classified. Cells with a confusion index  $\leq 0.6$  were allocated to the cluster for which they had highest membership, and thus represented the land facets upon which corridors were designed (Fig. 2.6c). Cells with a confusion index  $> 0.6$  (and those identified as outliers) were not allocated to a land facet.

### *Corridor design using land facets*

For each focal land facet type (e.g., low elevation, gentle canyon bottom), corridor termini (starting/ending locations for a corridor) were defined as polygons within the wildland blocks that contained the most occurrences of the focal land facet (Fig. 2.8). We aggregated all cells with at least one occurrence of the focal facet type within a 3-cell radius into polygons, and defined termini as those polygons that were  $\geq$  half the size of the largest polygon in each respective wildland block. Although we used a low density threshold (one cell within a 3-cell radius), we found that the largest polygons always contained a high density of the focal facet type. In situations where the largest polygons do not contain a high density of the focal facet type, our thresholds would select larger polygons more sparsely populated by the focal land facet over smaller but more densely populated polygons.

We designed one corridor per land facet using least-cost corridor analysis (Fig. 2.9). Underlying this approach is a resistance surface wherein the value of a cell represents the difficulty, or cost, in moving through it (Adriaensen et al. 2003). In least-cost modeling for a focal species, resistance is usually estimated as departure

from optimal habitat suitability (Beier et al. 2008; alternatively, it can be estimated from data on movement or gene flow). In the land facets approach to corridor design, we similarly defined resistance as the departure of a cell's attributes from the ideal attributes of the focal land facet (Fig. 2.10). Although the inverse of cluster membership values seems to be a reasonable estimate of resistance, it would only reflect differences in the continuous variables. To create resistance values that additionally reflect differences in topographic position, we used Mahalanobis distance, a multivariate distance measure standardized by the variance-covariance matrix of the independent variables (Clark et al. 1993; Gotelli & Ellison 2004).

Mahalanobis distance is calculated as

$$D^2 = (x - \mu)' \Sigma^{-1} (x - \mu), \quad (2.2)$$

where  $x$  is a vector of attributes associated with each cell in the analysis area,  $\mu$  is a vector representing the attributes of the ideal cell of the focal land facet, and  $\Sigma$  is the variance-covariance matrix of the independent variables. Using the Mahalanobis Distances extension to ArcGIS 9.3 (Jenness Enterprises 2010), we computed Mahalanobis distance on the same variables used to define land facets, and an additional variable, namely the relative density (scaled 0 to 1) of the focal land facet within a 3-cell radius (Fig. 2.8). Values in  $\mu$  for elevation, slope angle, and solar insolation (slope topographic position only) were calculated as the mean of the respective attributes of the cells inside of the wildland blocks allocated to the focal land facet; the value for density was set to 1 because an ideal cell of the focal facet type would be surrounded by other cells of the focal facet type. The variance-covariance matrix ( $\Sigma$ ) was also calculated on the cells inside of the wildland blocks allocated to the focal land facet.

To prevent a corridor from passing through urban or developed areas such as mines, these areas were digitized from aerial photographs (National Agricultural Imagery Program 2007) and assigned “no data” values in the resistance surfaces.

In least-cost modeling, each cell in the planning area is assigned a cost-distance equal to the lowest possible sum of resistance values in a chain of cells to termini in each wildland block (Adriaensen et al. 2003; Beier et al. 2008). The cost-distance maps (one for termini in each wildland block) are summed to produce the cumulative cost surface. A given proportion of cells with the lowest cumulative cost values is a least-cost corridor connecting the edges of termini in both wildland blocks. For each land facet, we examined multiple least-cost corridors containing different proportions

of cells (e. g., 0.5–5.0% in 0.5% increments) and selected the one with an approximate minimum width of 1 km over its length. We used this minimum width because it represents a compromise between narrower corridors that would not serve many species and wider corridors that would be too costly to conserve. Most focal species corridors in these landscapes were also approximately this width (Beier et al. 2007). We used the ‘Cost distance’ tool in the Spatial Analyst extension of ArcGIS 9.3 and the ‘Create corridor slices’ tool in the CorridorDesigner ArcGIS toolbox (Majka et al. 2007) for these procedures.

*Corridor design for high diversity of facets*

A single corridor was also designed to optimize connectivity for high diversity (i. e., high interspersed) of land facets. We used Shannon’s index (Magurran 1988) to measure diversity of land facets in a circular neighborhood with a 5-cell radius. Shannon’s index is calculated as

$$H' = - \sum p_i \ln(p_i), \quad (2.3)$$

where  $p_i$  is the proportion of cells classified as land facet  $i$  relative to cells in the neighborhood classified as any land facet, and the summation is over all land facets present in the landscape. Shannon’s index incorporates richness and evenness into a single measure. Thus, a high Shannon’s index is achieved by not only maximizing the number of land facets within the neighborhood, but also by balancing representation of those facets.

We calculated resistance of a cell as  $(H' + 0.1)^{-1}$ ; adding 0.1 precludes undefined values which could occur if no cells in the neighborhood were classified as a land facet. Thus cells with a high diversity index have low resistance. As in designing corridors for individual land facets, areas unsuitable for providing connectivity were also removed from this surface. We defined the corridor termini by first aggregating into polygons all cells inside the wildland blocks in the upper 50<sup>th</sup> percentile of Shannon’s index values. Of these polygons, we retained those that were  $\geq$  half the size of the largest polygon as termini. Both thresholds were defined with respect to cells in each wildland block separately. As before, least-cost corridor analysis was used to identify the corridor with an approximate minimum width of 1 km over its length.



*Modifying resistance surfaces or corridor termini to produce corridors that better capture the focal land facet*

For six of 35 corridors (Table 2.2), these procedures produced a highly linear corridor when a longer, less-linear corridor would better optimize continuity for the focal land facet (or diversity of facets). This happened when the relatively few matrix cells resembling the focal facet occurred far outside the straight paths between potential termini in opposing wildland blocks. In effect, resistances of cells of the focal land facet or cells resembling the focal facet were not low enough relative to the cost of travel through dissimilar cells to “pull” the corridor toward the low-cost cells. We developed two strategies to address this. Our first strategy was to exaggerate the cost of travel through cells dissimilar to the non-focal land facet by exponentiating the resistance surface by a power of 1.05, and increasing the exponent incrementally, stopping when the corridor shifted to incorporate clusters of low-cost cells. This strategy worked in three of the six cases; the largest exponent was 4. Before this stopping point was reached in the other three cases, the corridor developed wide “balloons” in regions of low resistance with narrow pinchpoints elsewhere. In these three cases, we used an alternative strategy, namely to relax the area threshold (or Shannon’s index threshold for high diversity of land facets) used to define corridor termini. For example, retaining polygons that were  $\geq 25\%$  of the size of the largest polygon often produced termini located in additional sections of the wildland blocks, such that low-cost matrix cells occurred directly between termini in opposing wildland blocks.

In addition to these six cases, two corridors (represented by strand A in Fig. 2.2) occurred in a matrix with few cells resembling the focal land facet types (high-elevation, steep canyon bottoms and ridges). In this situation, the modeled corridors were simply poor corridors for the focal facet types.

*Linkage design*

We created the final linkage design by taking the union of all least-cost corridors (one for each facet type and one for the diversity of facets) and the best riverine or riparian habitat in the analysis area as identified by Beier et al. (2007), who asked local experts to identify the reaches of major streams and rivers with the best perennial flow or (if no stream had such flows) the best riparian habitat. To accommodate edge effects, we buffered corridors by 60 m prior to the union and the resulting linkage design by 90 m. A 200-m buffer was used for riparian habitat (Beier et al. 2007).

## Results

Nine to 12 land facets were defined per landscape; three to five were defined per topographic position in each landscape (Table 2.3). Each land facet could be described by a simple phrase, such as “low elevation, steep canyon bottom” or “high elevation, gentle, hot slope.”

Each linkage design consisted of multiple strands and each strand consisted of one to 11 corridors (Figs. 2.1–2.3). A large degree of overlap occurred among corridors in each of the three landscapes (Table 2.2). On average, the area encompassed by a single corridor increased with distance between wildland blocks (Table 2.2). However, the Black Hills-Munds Mountain planning area contained the smallest and largest corridors; its linkage design was also the largest of the three.

The proportion of cells identified as outliers in each landscape deviated from the a priori 10% threshold described in the Methods (Table 2.2; Fig. 2.5). In the Black Hills-Munds Mountain and Santa Rita-Tumacacori planning areas, the wildland blocks (mountainous) differed sharply from the matrix (dominated by a broad flat valley). Thus, more than 10% of matrix cells (those of the lowest elevation and slope angle) were filtered out as outliers. Conversely, fewer than 10% of cells were defined as outliers in the Wickenburg-Hassayampa planning area because the southern wildland block was topographically similar to the matrix.

The number of poorly classified cells (confusion index  $> 0.6$ ) depended on the density of cells occurring at the medial margins of the clusters in attribute space (Figs. 2.6b–c). On average, 4.9% of cells were identified as poorly classified (Table 2.2). No apparent trend existed among landscapes or topographic positions.

There were more land facets within a topographic position in more complex landscapes. For example, cells in the slope topographic position of the Wickenburg-Hassayampa planning area, which were only classified into three land facets, were narrowly distributed with respect to elevation, slope angle, and insolation (Fig. 2.11, bottom row). Conversely, cells in the slope topographic position of the Black Hills-Munds Mountain planning area, which were classified into five land facets, had a multi-modal distribution with respect to elevation and a wider distribution with respect to slope angle and insolation (Fig. 2.11, bottom row).

In all three planning areas, the wildland blocks were relatively rugged compared to the matrix. Accordingly, the distribution of land facets differed between the wildland blocks and the matrix. For example, the proportion of cells allocated to canyon bottom and ridge land facets was higher in the wildland blocks than in the matrix in

all three landscapes.

Of the three landscapes, the wildland blocks in the Wickenburg-Hassayampa planning area were least similar to each other topographically (Fig. 2.2). The northern wildland block encompassed several mountain ranges, and therefore contained higher elevations, steeper slopes, and greater proportions of canyon bottoms and ridges than the southern wildland block. With exception to the low elevation, gentle, warm slope land facet, all termini in the southern wildland block occurred in areas most similar to the northern wildland block, namely the Big Horn and Harquahala Mountains.

Size of the linkage design depended more on the topography of the planning area than on the distance between wildland blocks or the number of land facets in a landscape. In the Santa Rita-Tumacacori planning area, 12 of 13 corridors contained some part of the topographically diverse San Cayetano Mountains (Fig. 2.3, strands A and B). These mountains lay between the wildland blocks and thus contributed to a relatively compact linkage design. Four of 11 corridors in the Black Hills-Munds Mountain planning area captured the Antelope Hills or Mogollon Rim, even though this indirect route was up to four times the distance between wildland blocks (Fig. 2.1, strands A and B). In the Wickenburg-Hassayampa planning area, all but the corridor for low elevation, gentle, warm slopes (strand D) captured the Vulture, Date Creek, or Weaver Mountains (Fig. 2.2). No cells allocated to the high elevation, steep canyon bottom and ridge land facets occurred in the matrix directly between the wildland blocks in this landscape; therefore, these corridors (strand A) resulted in straight paths between the Harquahala and Date Creek Mountains.

On average, approximately 148 hours were required to run all analyses for a single linkage planning area on a Microsoft Windows XP platform with 3.0 GHz Intel Core 2 Duo processor. The most time consuming step was iterating the fuzzy  $c$ -means cluster analysis for each  $c$  ( $\approx 134$  hours). Time invested in this step could be reduced  $> 80\%$  by performing fewer iterations (we performed 100), and only testing values of  $c \leq 7$  (we tested values as high as 10) as the time required for the cluster analysis increases exponentially with  $c$ . As few as 25 iterations are adequate because in instances where more than one partition was present, the less common partition comprised 32.2% of the iterations on average (range: 6.7–46.7%). Thus, performing 25 iterations yields a 99.99% probability of detecting multiple partitions on average, and an 82% probability for the least-common partition that we observed. Although most steps were relatively mechanical (including an automated script for iterating the cluster analysis), others like choosing the optimal number of clusters, selecting a corridor of the appropriate width, and modifying the resistance surface or potential termini

required user involvement and judgment.

## Discussion

In this chapter, we demonstrated a new approach to designing linkages in the context of climate change. Despite widespread prescriptions for conservation corridors and linkages to aid species' range shifts in response to climate change, few designs explicitly incorporate climate change considerations. Our approach exploits the fact that topography and soils are major drivers of biodiversity, and thus relies only on factors that are stable with respect to climate. Our approach avoids the enormous complexity and the high level of uncertainty associated with modeling climate and habitat suitability. It is also not afflicted by the patchiness and bias common in species occurrence data. In fact, our approach can be applied anywhere because digital elevation models are available for all continents (<http://www.gdem.aster.ersdac.or.jp/>). In areas where soil maps are adequate, soils information should also be used to define land facets.

Our approach is adaptable to different landscapes. The three planning areas used in this illustration varied 4-fold in terms of size, 2-fold in terms of distance between wildland blocks, and encompassed diverse topography (Table 2.1). Despite these differences, all three designs captured areas in the matrix most similar to the wildland blocks, even if doing so required long, nonlinear corridors. In instances where the wildland blocks were dissimilar, corridors connected areas inside the wildland blocks that were most similar across the two blocks.

Because R statistical software is an object-oriented language, this approach is currently limited to data sets (e.g., cells within a topographic position) consisting of approximately  $\leq 10$  million cells with four variables. The advent of 64-bit computer operating systems should mitigate this limitation. Alternatively, only cells inside of the wildland blocks could be imported into R, and the results from the kernel density estimation and fuzzy *c*-means cluster analysis later extrapolated to cells in the matrix using simple tools written for ArcGIS. Sampling cells inside of the wildland blocks would further allow R-based computations on even larger landscapes.

Our clustering procedures cannot test  $c = 1$  (i.e., no partitioning of a data set) because the validity indices either cannot indicate an optimal partition for  $c = 1$  (the Fukuyama-Sugeno and Average Within-Cluster Distance criteria, which are equivalent for  $c = 1$ ; Fig. 2.7) or are undefined for  $c = 1$  (all other indices are based on distances between clusters or an objects two highest membership values). In situations

where our procedures indicate a two cluster solution where one is truly optimal, the resulting land facets should be similar and the corresponding corridors would overlap extensively as if only one land facet was present.

Unlike other methods for modeling connectivity that consider topography (e.g., Rouget et al. 2006, Beier et al. 2007), ours is a rule-based approach. As such, it is structured in a way that allows for uncertainty analysis, an important next step in the land facets approach to linkage design. A strategy similar to Beier et al. (2009) could be adopted to determine how sensitive modeled corridors are to the parameters (e.g.,  $\varphi$  in fuzzy  $c$ -means cluster analysis, size of neighborhoods used to define corridor termini) and decision rules in the approach (e.g., thresholds for defining outliers, poorly classified cells, or corridor termini), as well as to errors in digital elevation models.

We comment on several subjective decisions in our approach:

1. Which variables to use to define land facets: Our flexible procedures can accommodate a variety of topographic variables, placing control in the hands of conservation practitioners and not in the tools. Topographic variables suitable for defining land facets, such as those derived from a digital elevation model, must be mapped continuously over the extent of the analysis area. For a discussion of topographic variables that can be derived from a digital elevation model, see Moore et al. (1991) and Franklin (1995). Where soil maps are complete and do not contain unmapped heterogeneity, we believe soil attributes should help define land facets. Categorical variables (relating to soil or topography) can be integrated into these procedures in the same way we treated topographic position, or they could be converted to continuous variables by using the density of the categories within a neighborhood. Where soil maps are inadequate, land facets defined solely by topographic variables can represent a useful diversity of habitats (Franklin 1995; Hoersch et al. 2002).

To maintain easily interpretable and biologically meaningful land facets, it is best to limit the number of variables (Beier & Brost 2010). Each variable should be viewed in a functional perspective and judged for its influence on the availability and distribution of heat, light, water, or nutrients (Mackey et al. 1988). In mountainous landscapes, elevation contributes meaningfully to the definition of land facets, but in flatter landscapes elevation may not be relevant. If the relative importance of topographic or soil variables is to be established based on correlation with species occurrence, spatial autocorrelation must be considered to avoid inflating the probability of type I error (Legendre 1993).

2. How many land facets to recognize: If two or three values of  $c$  seemed equally apt for the optimal number of clusters, we usually selected the smaller number of clusters within the canyon bottom and ridge topographic positions and a larger number within the slope topographic position. We did this because cells classified as canyon bottoms and ridges were relatively rare, and recognizing a higher number of clusters would have produced facets that were extremely rare. We also evaluated interpretability of classes and draped maps of facets over a topographic hillshade to assess whether the  $c$  clusters corresponded to natural units.
3. How to define outliers: The 10% threshold we used to define outliers separated regions in attribute space densely populated by cells from those more sparsely populated (Fig. 2.5). In other landscapes, examination of 2- or 3-dimensional plots of the cells in attribute space may indicate a more appropriate threshold.
4. How to modify the resistance surface or termini to produce corridors that better capture the focal land facet: When transforming a resistance surface, we recommend starting with a small exponent, such as 1.05, and exploring sequentially larger values (i.e., 1.1, 1.5, 2, ...) if necessary. While larger exponents may produce longer, less-linear corridors when desired, they are also more likely to cause ballooning and pinchpoints in the resulting corridor. Similarly, thresholds used in defining termini should be relaxed only enough to yield additional termini, such that low-cost matrix cells occur between potential termini in opposing wildland blocks. However, it is important to keep in mind that relaxing these thresholds recognizes smaller termini (or termini with lower diversity) that are less able to support area-sensitive species or ecological processes.
5. Applying minimum width to corridors: Identifying the least-cost corridor with an approximate minimum width of 1 km over its length was often challenging because corridors do not have a constant width, being wider in areas of low resistance and narrower in areas of high resistance. When selecting a corridor, we aimed to have no more than 10–20% of its length below 1 km in width; however, in some instances this target was not attainable because doing so caused severe ballooning in other sections of the corridor. In these cases, we attempted to identify the corridor that represented the best compromise between sections that were too narrow and too wide.

*Use of land facets in linkage design*

Several studies address whether the full diversity of physical units represents today's biodiversity. Kirkpatrick and Brown (1994) found a significant correspondence between conservation lands selected based on physical attributes and those selected based on plant species and communities. The number of species represented by the reserve system based on physical attributes increased with area reserved, but still failed to capture many species and communities that had limited distributions. Similarly, Cowling et al. (1999) found that a hypothetical reserve designed using physical units as surrogates for ecological processes (including biotic response to climate change) represented 37% fewer rare plant species than a reserve designed explicitly for those species. Unrepresented species were more likely to require specialized habitat or have distributions that related to historical factors (Lombard et al. 2003). Araujo et al. (2001) showed that areas selected for environmental diversity did not always represent species diversity at a rate higher than that expected by chance. Reyers et al. (2002), on the other hand, found that conservation plans based on physical features represented species well, including rare and endemic species. Their approach based on physical units came at a higher cost to land relative to approaches based on vegetation and species data.

The poor performance of physical units in some of these studies was probably partly due to the coarse scale at which they were mapped. For example, Kirkpatrick and Brown (1994) defined physical units on a 10-km<sup>2</sup> grid and Araujo et al. (2001) used a 2500-km<sup>2</sup> grid. At these scales, each unit contains tremendous topographic heterogeneity and may fail to accurately represent conditions in most or even any portion of the unit. Consequently, inclusion of conditions poorly represented by these broad-scale classifications in conservation plans is left to chance. Such plans may therefore capture less environmental diversity, and by extension species diversity, than those based on finely-defined classes (Reyers et al. 2002). Our approach using 30-m or finer cells could be adapted to reserve design and thus support classes that are finely defined and precisely mapped.

Nonetheless, some habitats will inevitably not be recognized as distinct physical units, thereby causing some species to pass through the coarse filter of the land facets approach to linkage design (Reyers et al. 2002). Conversely, species-based approaches may not represent habitat for some non-modeled species and inherently cannot be used for species with unknown distributions. Therefore, a combination of approaches may best address the shortcomings of species data and the limitations of land facets.

We believe that the use of land facets is a simple and effective strategy to design

linkages for climate change. We advocate using land facets to complement, rather than replace, existing focal species approaches to linkage design. In conjunction, they can provide connectivity for most species and achieve persistence goals by conserving the ecological and evolutionary processes that sustain and generate biodiversity. Brost and Beier (in prep; Chapter 3 of this thesis) found that these land facets linkage designs served most focal species well in these landscapes. If this pattern is generally true, the land facets approach could also be used in lieu of focal species in areas where species models cannot be developed.



Table 2.1. Descriptions of planning areas and wildland blocks used in our analyses.

Planning area	Area (km <sup>2</sup> )	Number of cells in planning area	Minimum distance between wildland blocks (km)	Wildland block 1 (western-most block)			Wildland Block 2 (eastern-most block)			Matrix	
				Area (km <sup>2</sup> )	Elev. range (m)	Major geographic features	Area (km <sup>2</sup> )	Elev. range (m)	Major geographic features	Elev. range (m)	Major geographic features
Black Hills- Munds Mountain (111.9 W, 34.7 N)	3817.4	4241503	16.8	372.8	1080– 2385	Black Hills	374.0	1050– 2090	Mogollon Rim, Munds Mtn., House Mtn., Schnebly Hill, and Horse Mesa	915– 2385	Mogollon Rim and Antelope Hills; Verde River bisects planning area
Wickenburg- Hassayampa (112.8 W, 33.9 N)	9786.6	10874001	26.4	1606.3	360– 1735	Harquahala and Big Horn Mountains; Hassayampa Plain	1682.9	430– 2429	Wickenburg, Weaver, Hieroglyphic, Bradshaw, Buckhorn, and Sheep Mountains	275– 2325	Harcuar and Vulture Mountains; Hassayampa Plain; Hassayampa and Agua Fria Rivers
Santa Rita- Tumacacori (111.0 W, 31.6 N)	2475.9	2706667	13.0	480.6	1050– 1950	Tumacacori, Atascosa, and Pajarito Mountains	429.4	1100– 2885	Santa Rita Mountains	665– 2200	Patagonia and San Cayetano Mtns; Santa Cruz River bisects planning area

Table 2.2. Characteristics of three linkage designs based on land facets. Percent overlap between corridors represents the percent of cells in the linkage design overlapped by two or more corridors. Area of linkage design represents the linkage design after inclusion of riparian habitat and buffering to address edge effects.

Planning area	Percent of cells classified as outliers	Percent of cells poorly classified	Number of corridors	Number of corridors for which procedures were modified		Average area of corridors <sup>a,b</sup> (km <sup>2</sup> ) (range)	Area of union of least-cost corridors <sup>a</sup> (km <sup>2</sup> )	Percent overlap between corridors	Area of linkage design <sup>a</sup> (km <sup>2</sup> )
				Resistance surface	Corridor termini				
Black Hills-Munds Mountain	26	6	12	1	0	43 (16, 116)	370	72	493
Wickenburg-Hassayampa	8	4	10	1	1	53 (36, 78)	360	68	412
Santa Rita-Tumacacori	19	5	13	1	2	32 (22, 44)	187	45	316

<sup>a</sup>Measurement excludes area inside of the wildland blocks.

<sup>b</sup>Measurement excludes 60-m buffer used to address edge effects.

Table 2.3. Mean and range of topographic attributes encompassed by individual land facets in each planning area. Although land facets overlapped in range of elevation, slope angle, and solar insolation, their joint distributions did not (Fig. 2.6). “Hot,” warm,” and “cool” refer to relative amounts of insolation.

Planning area	Land facet description	Mean elevation (m) (range)	Mean slope ( $^{\circ}$ ) (range)	Mean solar insolation (kWh $^{-1}$ ) (range)
Black Hills- Munds Mountain	Canyon bottom: low elevation, gentle	1314.3 (1049.9, 1603.6)	9.8 (0.0, 21.1)	—
	Canyon bottom: mid elevation, steep	1540.6 (1168.0, 1938.0)	27.8 (17.9, 43.6)	—
	Canyon bottom: high elevation, gentle	1829.2 (1544.7, 2090.9)	14.6 (2.3, 28.5)	—
	Ridge: low elevation, gentle	1366.3 (1103.4, 1646.8)	10.6 (0.0, 23.4)	—
	Ridge: mid elevation, steep	1635.6 (1199.1, 2044.8)	29.9 (19.0, 49.2)	—
	Ridge: high elevation, gentle	1922.4 (1625.4, 2326.3)	13.1 (1.0, 27.3)	—
	Slope: low elevation, gentle, warm	1188.5 (1047.1, 1529.4)	4.8 (0.0, 16.5)	1561.5 (1411.7, 1671.4)
	Slope: low elevation, steep, cool	1387.2 (1067.2, 2056.3)	18.2 (10.2, 35.9)	1381.4 (987.6, 1552.0)
	Slope: mid elevation, gentle, warm	1518 (1300.1, 1822.9)	4.8 (0.0, 11.7)	1642.3 (1494.1, 1750.5)
	Slope: mid elevation, steep, warm	1586.0 (1111.3, 2103.9)	18.1 (11.1, 35.6)	1678.7 (1492.1, 1867.0)
	Slope: high elevation, gentle, hot	1986.5 (1690.3, 2384.3)	4.7 (0.0, 15.9)	1729.7 (1546.2, 1886.8)
Wickenburg- Hassayampa	Canyon bottom: low elevation, gentle	785.9 (452.3, 1284.5)	8.5 (0.0, 16.5)	—
	Canyon bottom: low elevation, steep	909.1 (493.1, 1244.0)	22.7 (14.9, 35.5)	—
	Canyon bottom: high elevation, steep	1435.2 (1137.8, 1865.2)	18.3 (4.4, 32.6)	—
	Ridge: low elevation, gentle	827.8 (497.5, 1337.3)	10.4 (0.8, 18.3)	—
	Ridge: low elevation, steep	932.7 (511.0, 1362.9)	25.4 (17.7, 36.7)	—



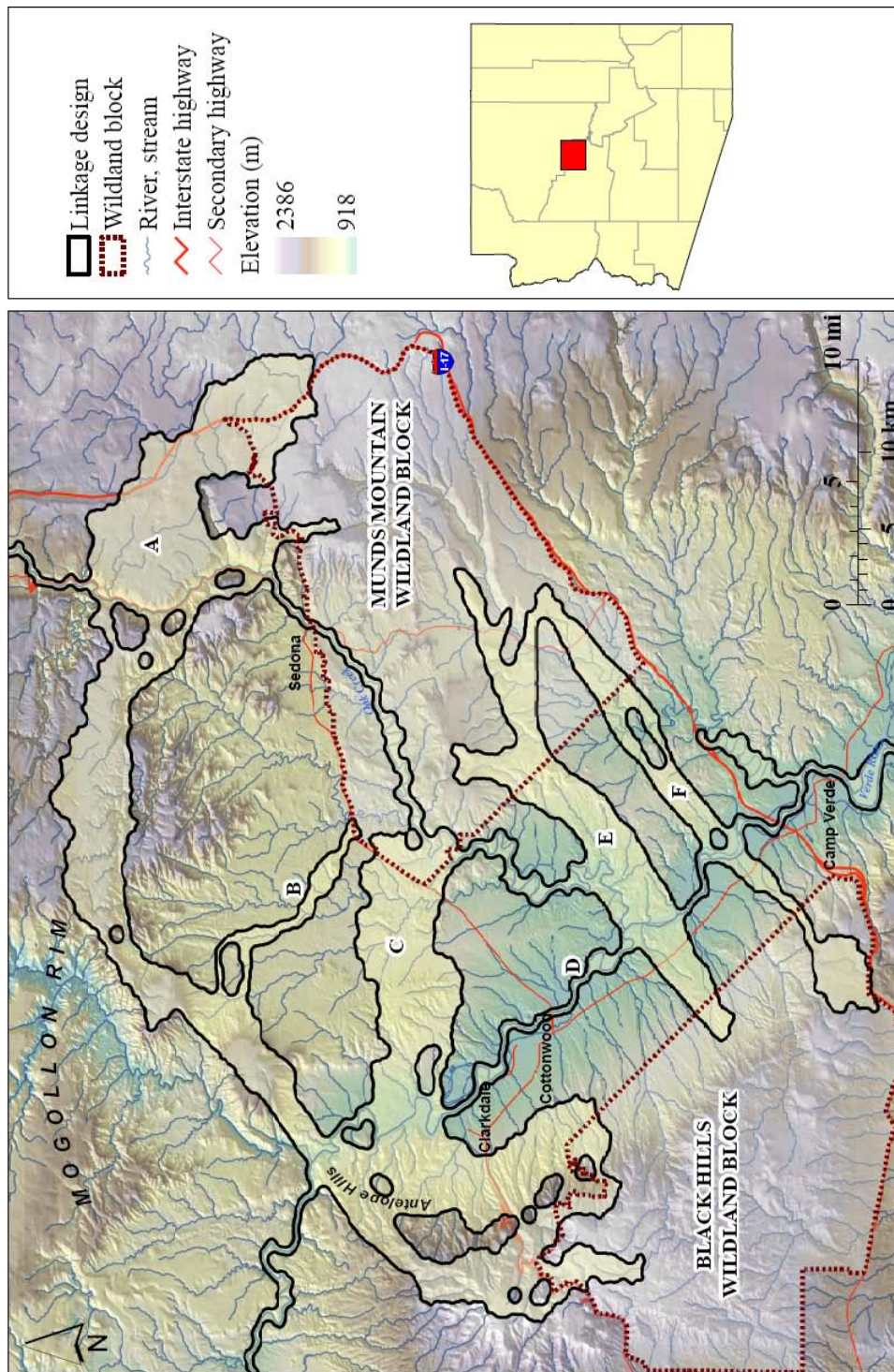


Figure 2.1. Map of the land facets linkage design for the Black Hills-Munds Mountain planning area. Linkage strands consisted of corridors for (A) high elevation, gentle canyon bottoms and ridges; and high elevation, gentle, hot slopes; (B) high diversity of land facets; (C) low elevation, gentle, canyon bottoms and ridges; and low elevation, gentle, warm slopes; (D) riparian habitat; (E) mid elevation, steep canyon bottoms and ridges; low elevation, steep, cool slopes; and mid elevation, steep, warm slopes; and (F) mid elevation, gentle, warm slopes.



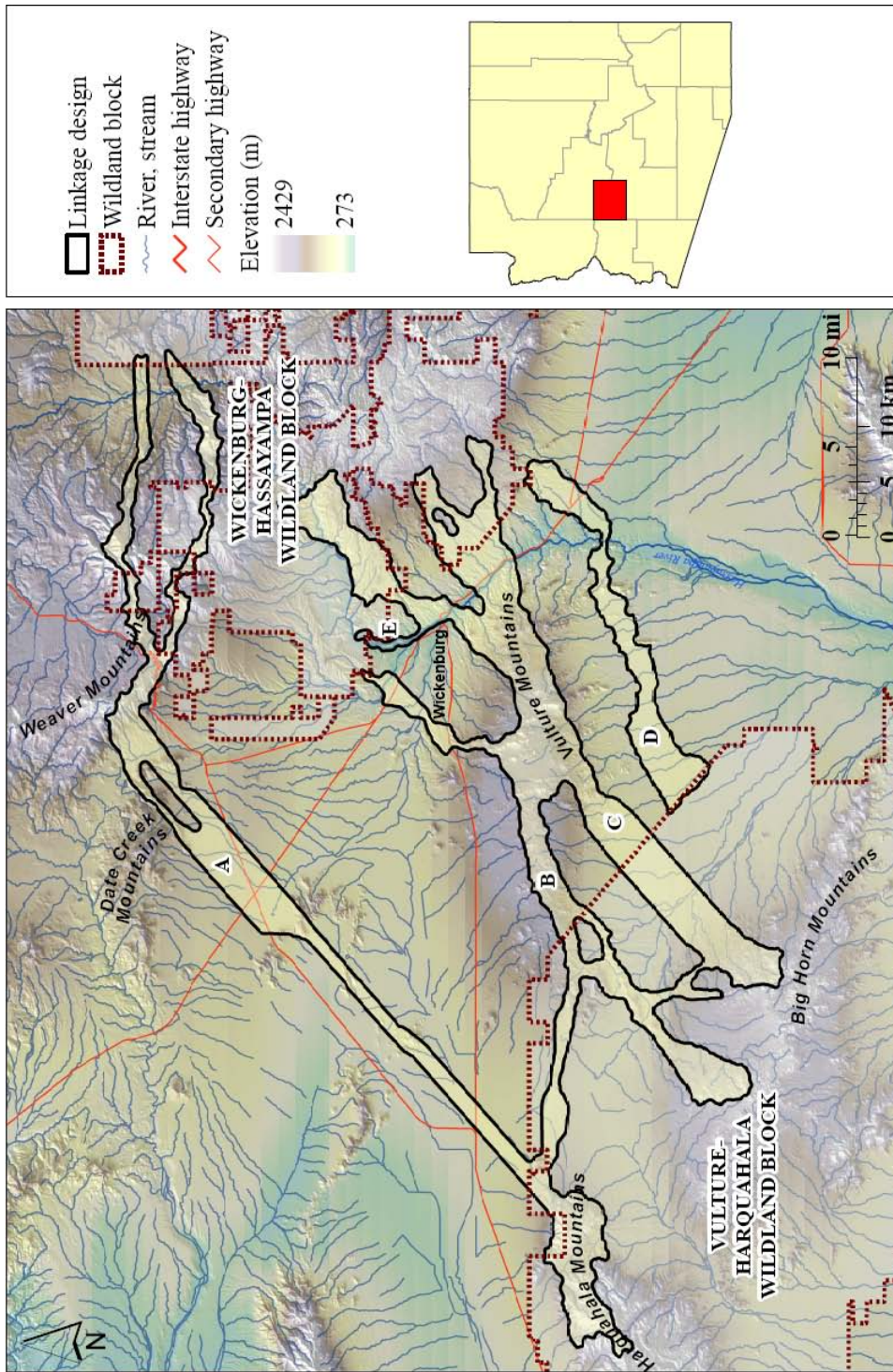


Figure 2.2. Map of the land facets linkage design for the Wickenburg-Hassayampa planning area. Linkage strands consisted of corridors for (A) high elevation, steep canyon bottoms and ridges; (B-C) low elevation, gentle canyon bottoms and ridges; low elevation, steep canyon bottoms and ridges; mid elevation, steep, cool slopes; high elevation, steep, warm slopes; and high diversity of land facets; (D) low elevation, gentle, warm slopes; and (E) riparian habitat.



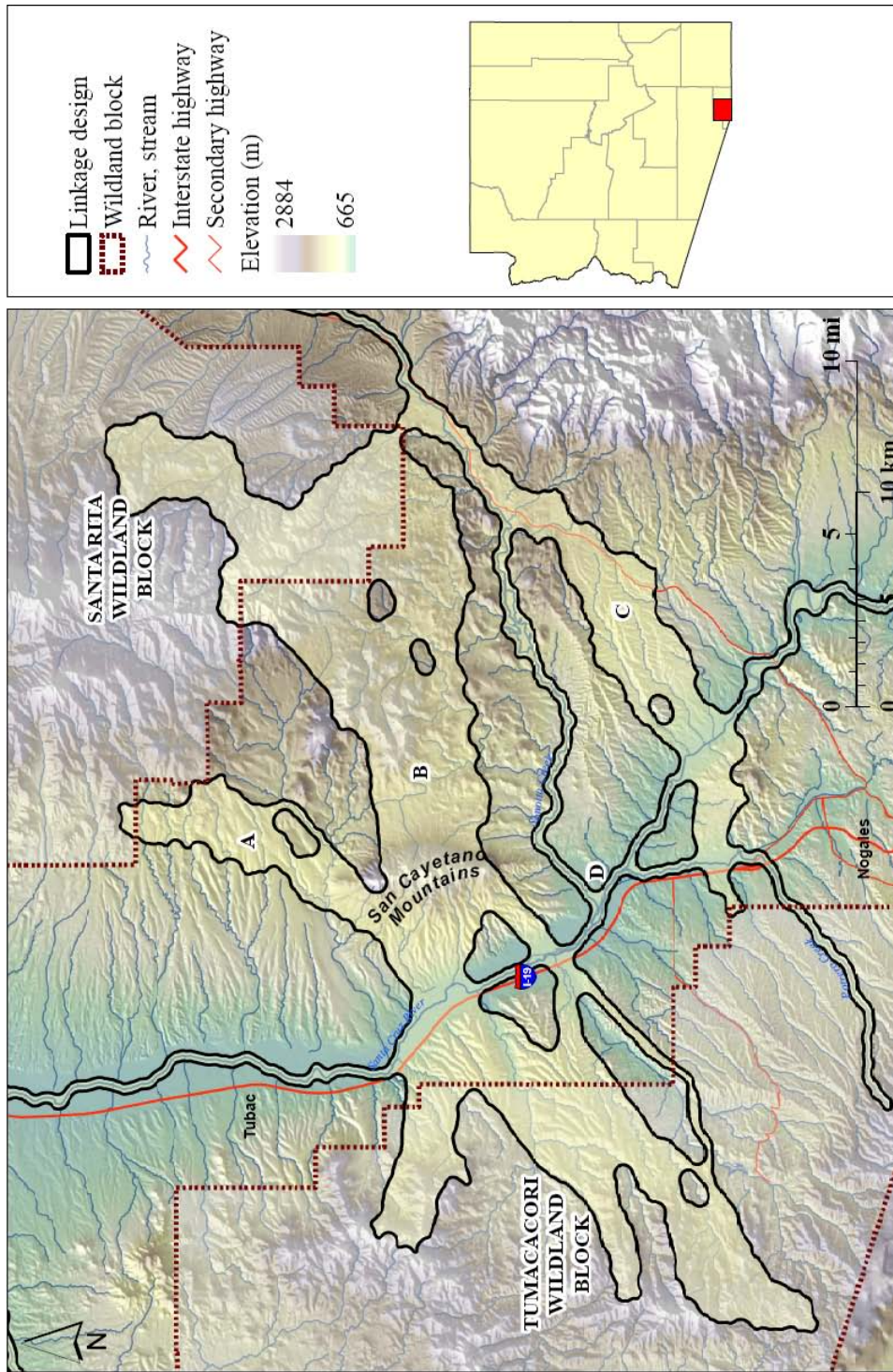


Figure 2.3. Map of the land facets linkage design for the Santa Rita-Tumacacori planning area. Linkage strands consisted of corridors for (A) high elevation, steep canyon bottoms; (B) low elevation, gentle canyon bottoms and ridges; mid elevation, gentle canyon bottoms and ridges; steep canyon bottoms and ridges; high elevation, steep ridges; mid elevation, steep, cool slopes; mid elevation, steep, hot slopes; high elevation, gentle, hot slopes; and high diversity of land facets; (C) low elevation, gentle, warm slopes; and (D) riparian habitat.

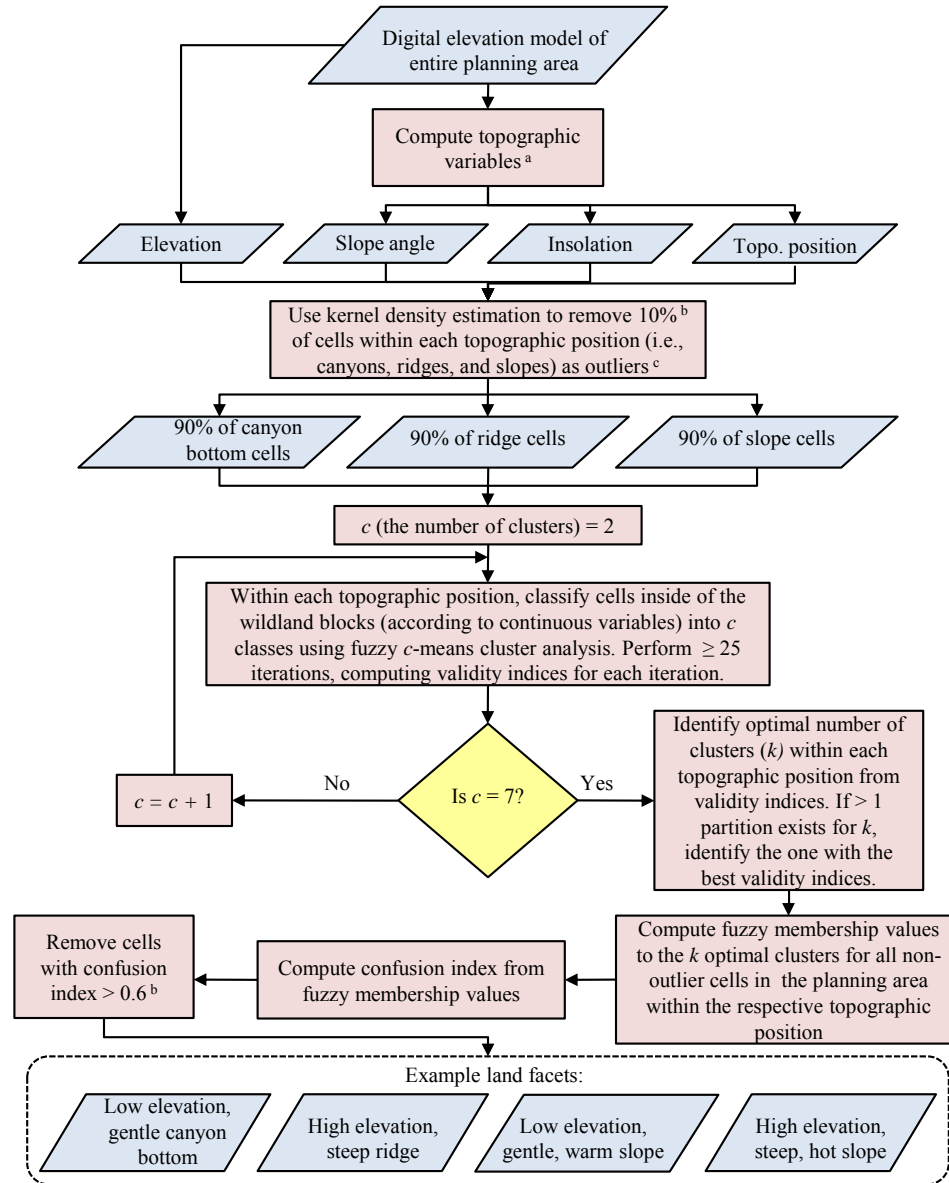


Figure 2.4. Sequence of operations (rectangles) and products (parallelograms) used to define land facets. The first operation occurs in ArcGIS; the remaining operations occur in R statistical software.

<sup>a</sup>The analyst can use other topographic variables or include soil variables.

<sup>b</sup>These parameters can be changed.

<sup>c</sup>We identified outliers and clusters with respect to elevation and slope angle for cells within the canyon bottom and ridge topographic positions. Elevation, slope angle, and solar insolation were used for cells within the slope topographic position. Outliers and clusters were defined with respect to cells inside of the wildland blocks only.



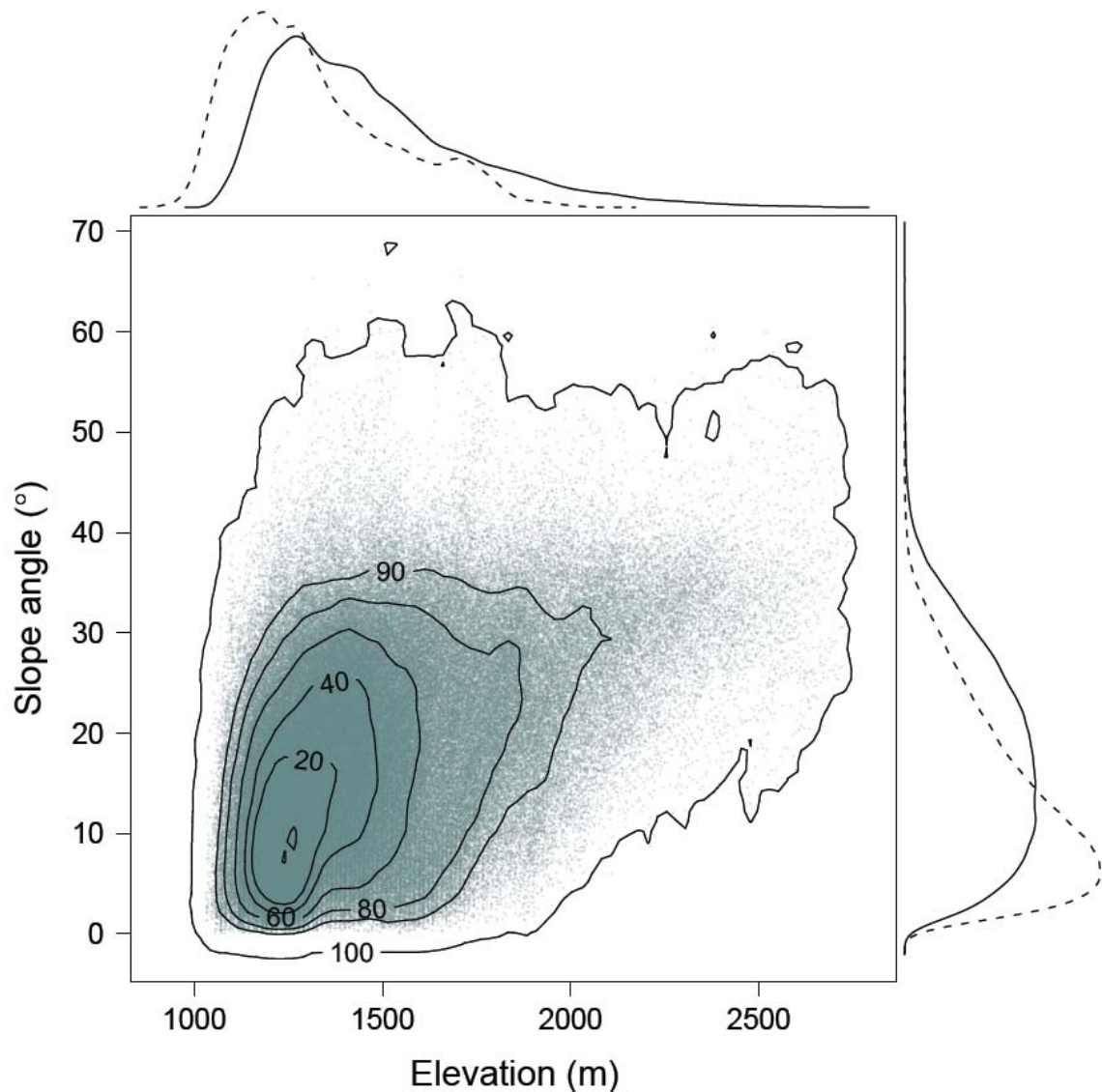


Figure 2.5. Example kernel density estimation for the canyon bottom topographic position in the Santa Rita-Tumacacori planning area. Contour lines in the main plot contain the densest proportion of cells as indicated by the labels. We defined outliers as those cells occurring outside of the 90% contour (i.e., the least dense 10% of cells). The marginal plots show how the distribution of cells in the canyon bottom topographic position inside the wildland blocks (solid line) differed from those in the matrix (dashed line). Because kernel density estimation was based on cells inside of the wildland blocks only, the proportion of cells in the entire planning area classified as outliers deviated from 10% depending on the extent and location of overlap in the marginal distributions.

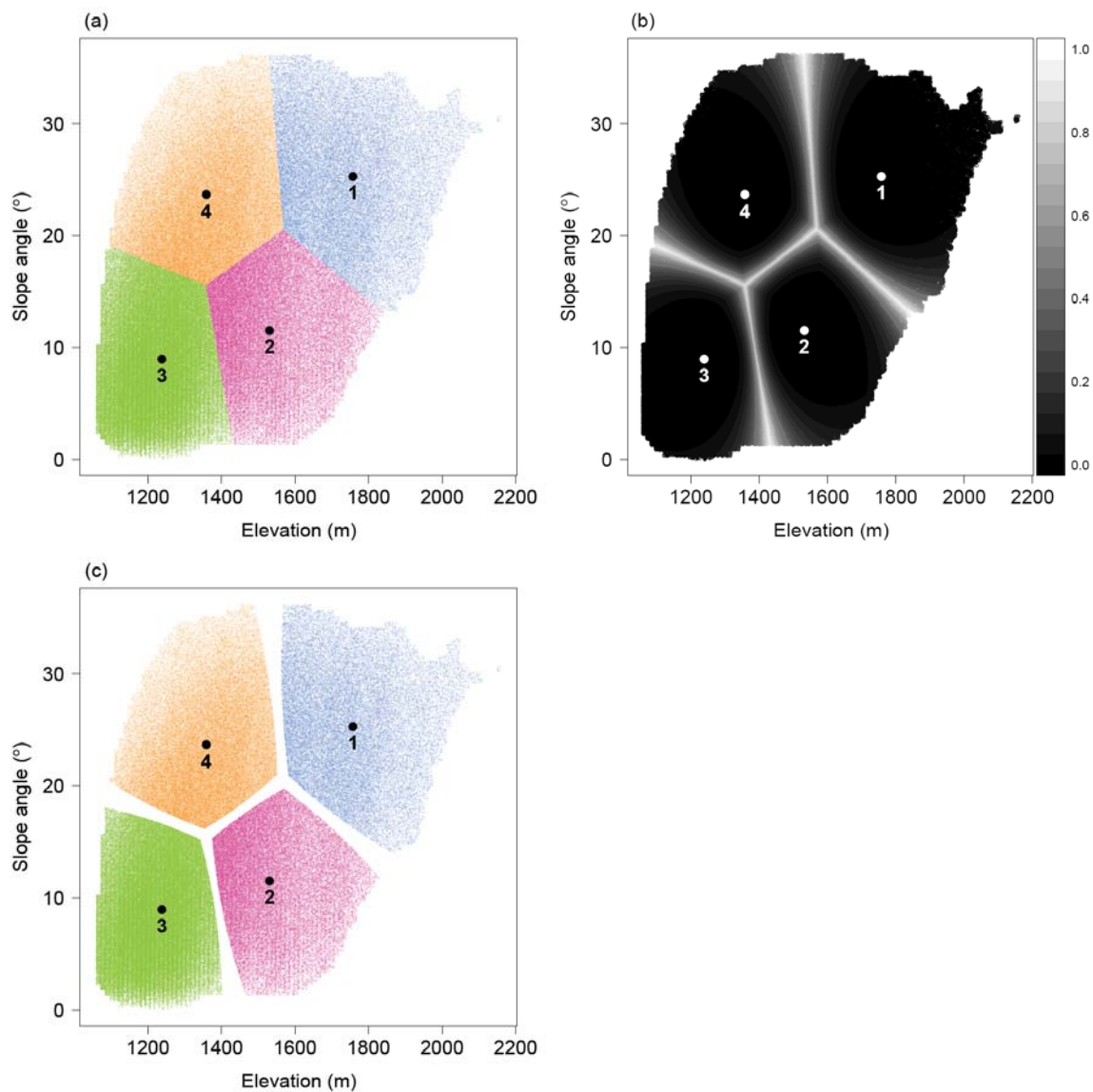


Figure 2.6. Example output of fuzzy  $c$ -means cluster analysis for the canyon bottom topographic position in the Santa Rita-Tumacacori planning area. (a) Plot of non-outlier cells assigned to the cluster centroid for which they have highest membership (ignoring fuzziness of the classification). (b) Confusion index indicating how well each cell is classified. Values near 1 indicate high confusion, whereas values near 0 indicate perfect classification. (c) Plot showing composition of land facets, i.e., cells with confusion index  $< 0.6$ . Despite the “fuzziness” in the classification, the joint distributions of attributes for the resulting land facets are mutually exclusive.

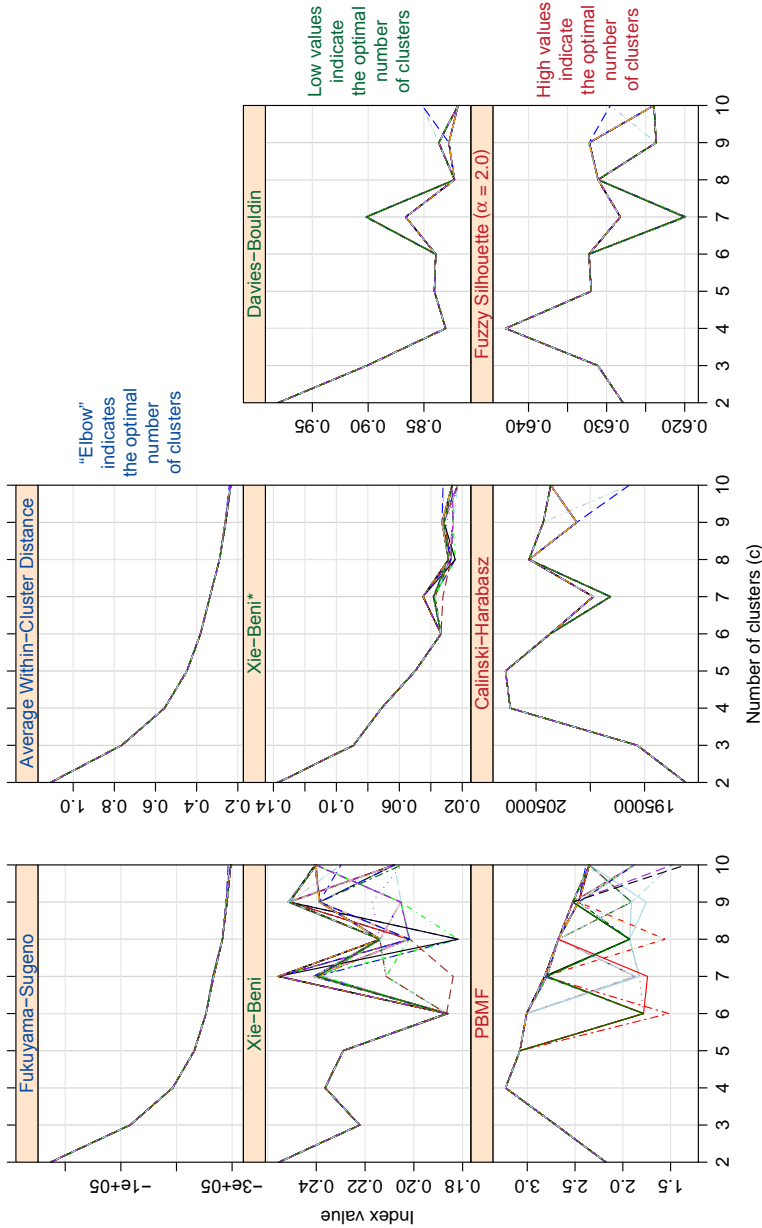


Figure 2.7. Example cluster validity indices for the classification of cells in the canyon bottom topographic position of the Santa Rita-Tumacacori planning area. One hundred iterations of the cluster analysis were performed for each number of clusters (i. e.,  $2 \leq c \leq 10$ ), and indices were computed for each iteration. The five indices in the top two rows are minimized for the optimal fuzzy- $c$  partition, whereas the three indices in the bottom row are maximized (Maulik & Bandyopadhyay 2002; Pakhira et al. 2004; Campello & Hruschka 2006; Hruschka et al. 2006; Celikyilmaz & Turksen 2008; Zhang et al. 2008). The Fukuyama-Sugeno and Average Within-Cluster Distance indices decrease monotonically as a function of  $c$ , and a good partition is indicated by an “elbow” in their plots. Consequently, these indices cannot specify a two cluster optimal partition. The remaining indices are less dependent on  $c$ , especially when  $c \ll n$  (the number of objects). Iterations for a single index diverge where multiple fuzzy- $c$  partitions minimize the within-cluster variance for a given number of clusters. In this case we identified the optimal number of clusters as four.

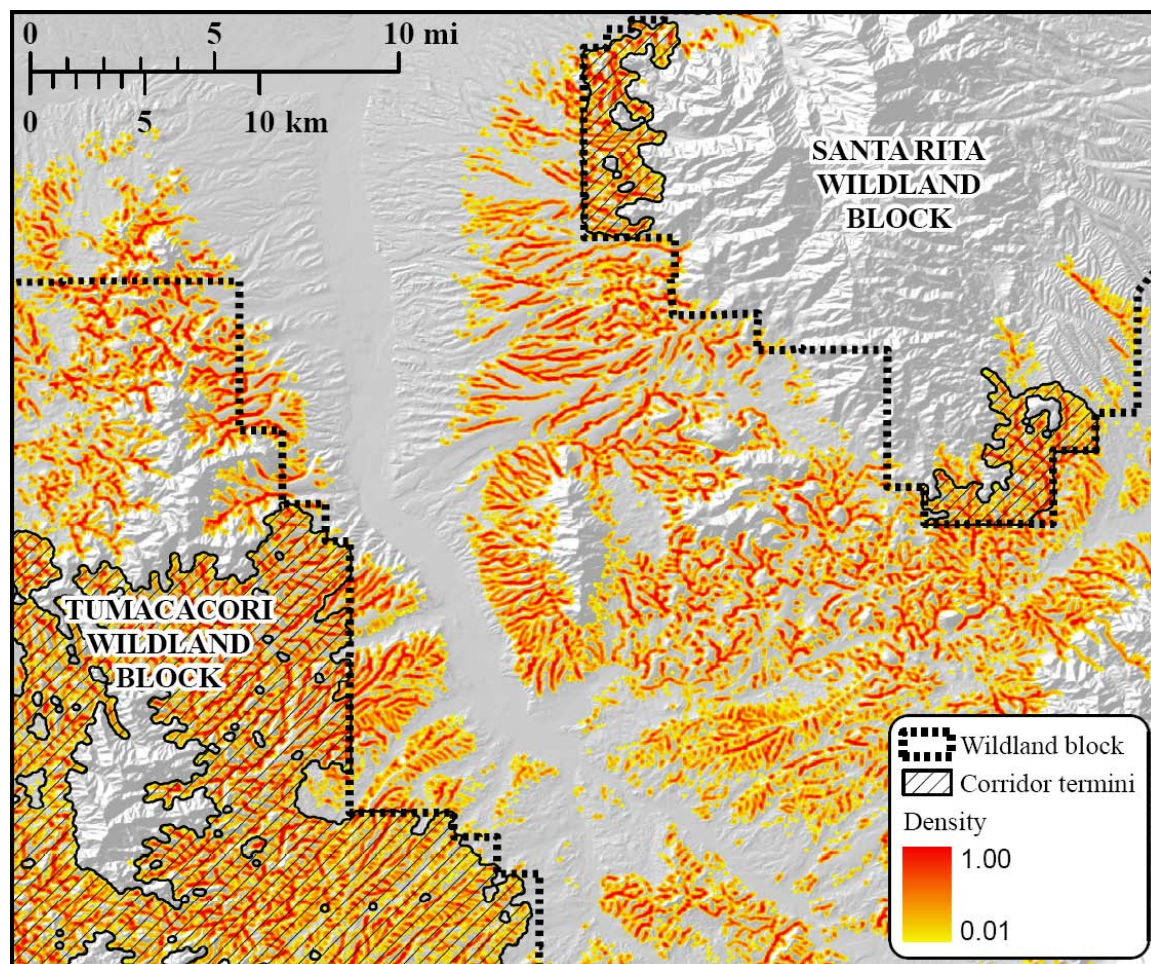


Figure 2.8. Example corridor termini for the low elevation, gentle canyon bottom land facet type in the Santa Rita-Tumacacori planning area. We aggregated all cells with at least one occurrence of the focal facet type within a 3-cell radius (i.e., cells with density values  $> 0$ ) into polygons, and defined termini as those polygons that were  $\geq$  half the size of the largest polygon in each respective wildland block.



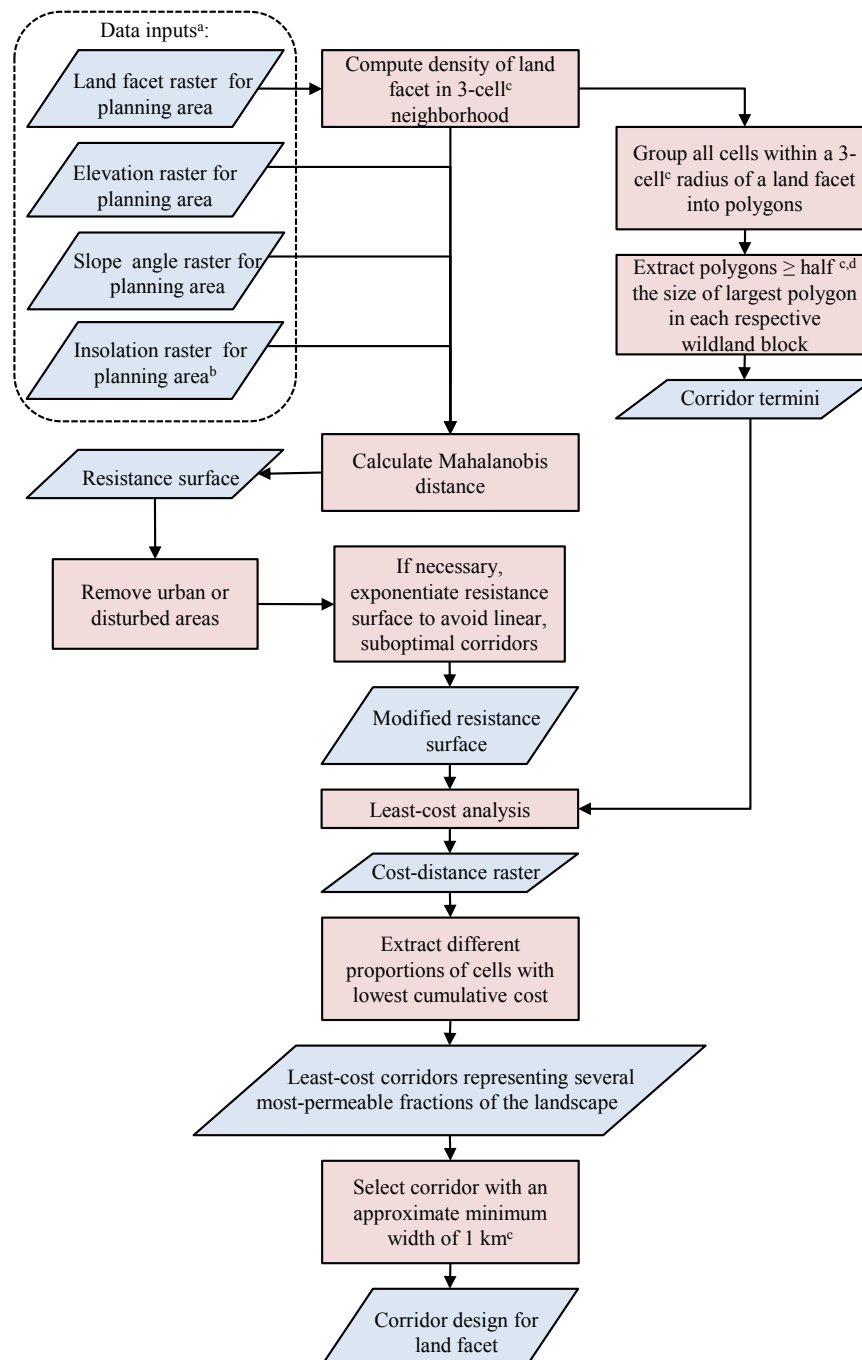


Figure 2.9. Sequence of operations (rectangles) and products (parallelograms) used to design a corridor for one land facet. All operations occur in ArcGIS and the process repeated for each land facet. The resulting corridors, plus a corridor for high diversity of land facets and riparian habitat, are then joined to create the linkage design.

<sup>a</sup>The analyst can use other topographic variables or include soil variables.

<sup>b</sup>We included solar insolation for land facets in the slope topographic position only.

<sup>c</sup>These parameters can be changed.

<sup>d</sup>The area threshold for defining termini can be adjusted to avoid highly-linear corridors.

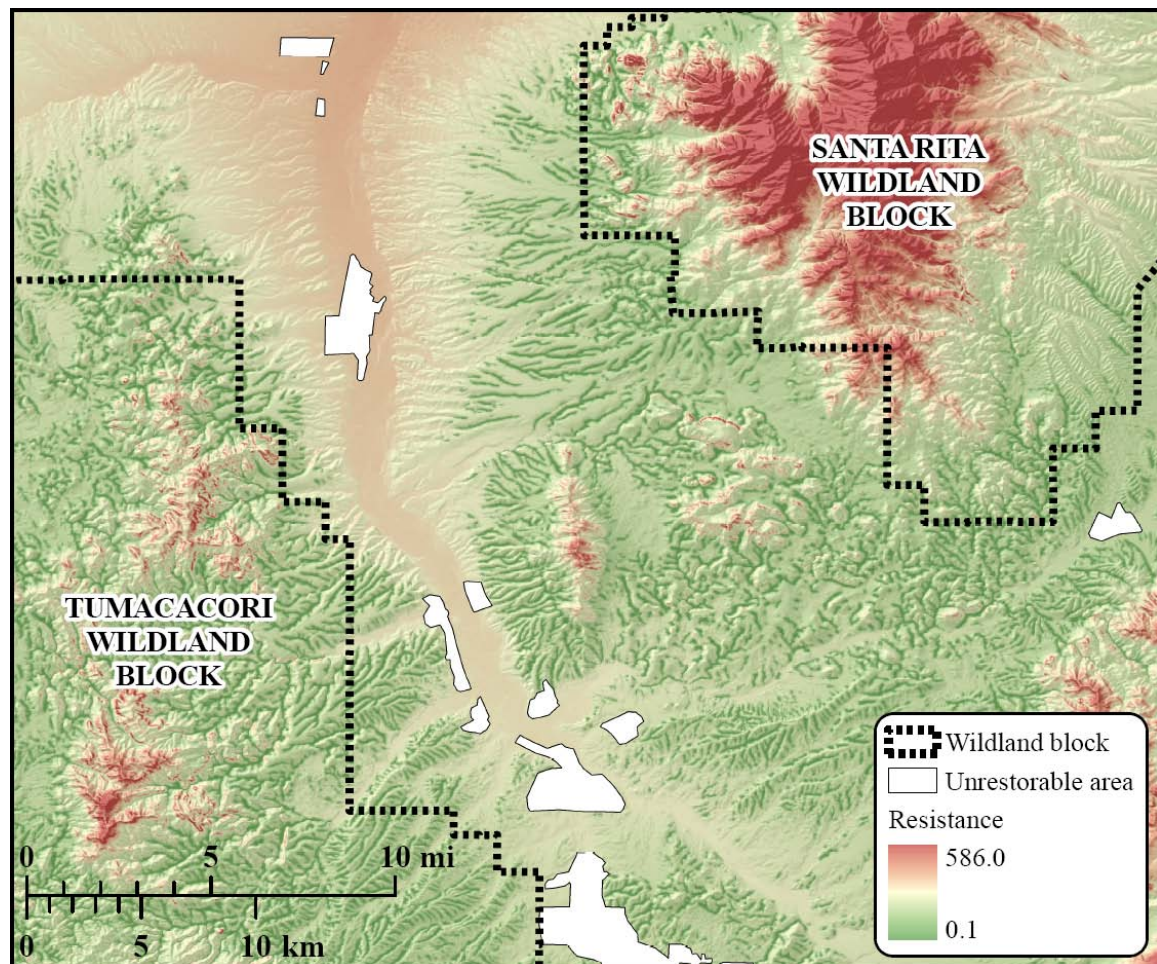


Figure 2.10. Example resistance surface for the low elevation, gentle canyon bottom land facet type in the Santa Rita-Tumacacori planning area. Each cell's resistance, calculated using Mahalanobis distance, represents the departure of its attributes from the ideal attributes of the focal land facet type. Resistance was based on the same variables used to define the land facet, namely elevation and slope angle (solar insolation was additionally used for land facets in the slope topographic position), and the relative density of the focal land facet within a 3-cell radius. In this map, cells that are shaded dark green have a low resistance because they either are the low elevation, gentle canyon bottom facet type, or are similar to this facet type.

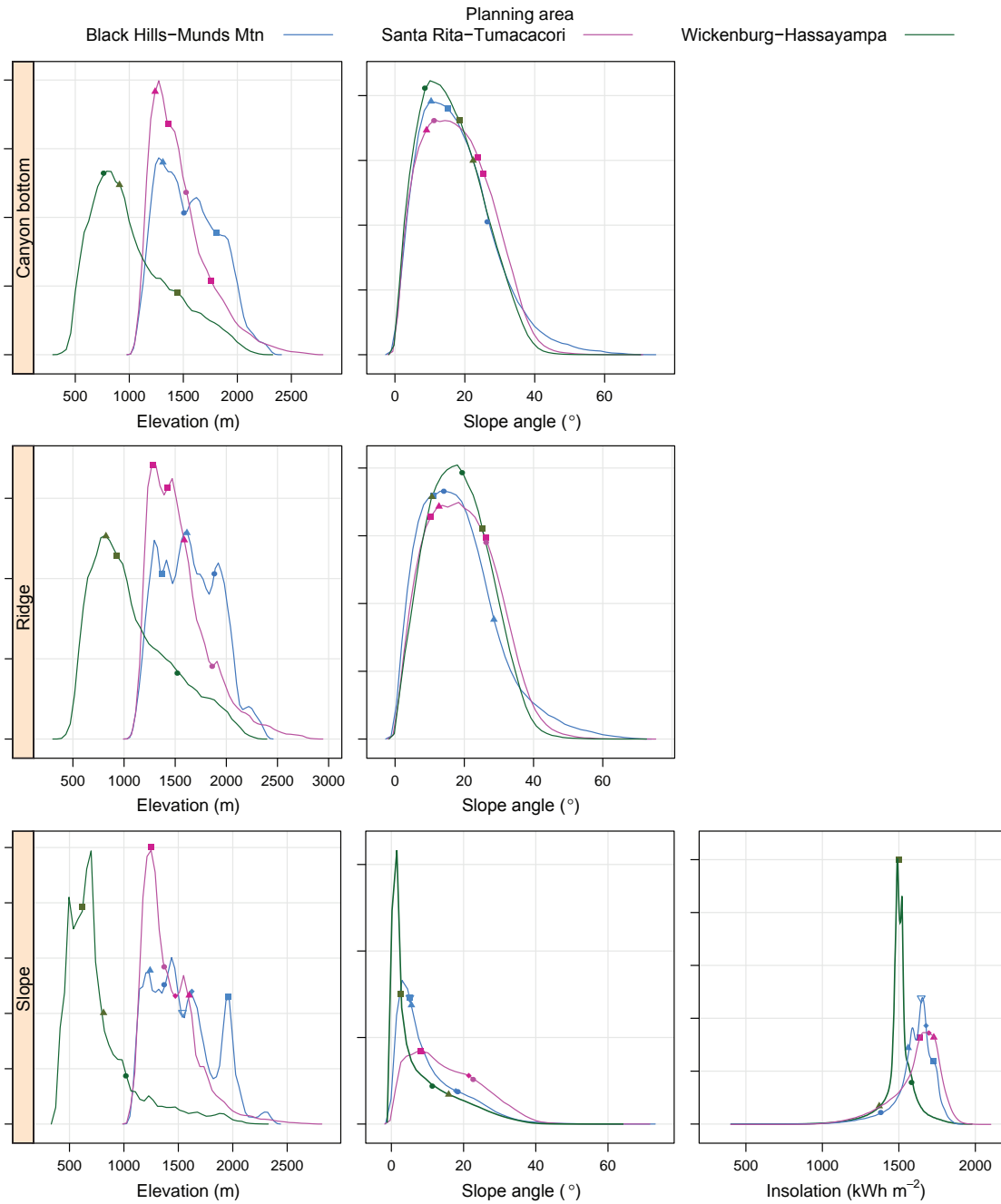


Figure 2.11. Univariate distributions demonstrating the topographic complexity of cells in each topographic position for the three planning areas. Points indicate location of cluster centroids and are symbolized by cluster.

## Chapter 3

# Comparing Linkage Designs Based on Land Facets to Linkage Designs Based on Focal Species

**Abstract.** Least-cost modeling for focal species is the most widely used method for designing conservation corridors and linkages. However, these linkages have been based on today's species' distributions and land cover, factors that will change with large-scale climate change. We recently proposed an alternative approach, namely to design linkages based on land facets, or recurring landscape units of relatively uniform topography and soils. The rationale is that linkage strands with high continuity of individual land facets will facilitate movement of species associated with each facet today and in the future, and linkage strands with high interspersions of facets will facilitate rapid, short-distance range shifts during periods of climate instability. Conservation practitioners and investors might like to know whether a linkage design based on land facets is likely to support movement by species needing connectivity today. Conversely, practitioners who have designed a linkage for focal species might want to know if the design provides continuity and interspersions of land facets, or whether additional analyses are needed to make the design better capture physical environments. To address these questions, we evaluated linkages designed for focal species and land facets in three landscapes in Arizona, USA. We used two variables to measure linkage utility, namely distances between patches of modeled breeding habitat for five to 16 focal species in each linkage, and resistance profiles for focal species and land facets between habitat patches and wildland blocks connected by the linkage. Linkage designs based on land facets served 25 of 28 focal species as well as or better than the focal species designs in these landscapes. The three species



better served by the focal species approach had the most narrowly distributed habitat. Compared to land facets designs, focal species linkages provided a similar degree of connectivity for 12 of 13 land facets in one landscape, but for only half the land facets in the other two landscapes. In areas where a focal species approach to linkage design is not possible, our results suggest that conservation practitioners can solely implement a land facets approach with some confidence that the linkage design would serve most potential focal species. In areas where focal species designs are possible, we recommend using the land facets approach to complement, rather than replace, focal species approaches.

## Introduction

Designing and protecting conservation corridors and linkages is one strategy to conserve connectivity in landscapes increasingly dominated by human activities (Crooks & Sanjayan 2006). It is also the most commonly recommended strategy for biodiversity management in the face of climate change (Heller & Zavaleta 2009). Although many different corridor design methods exist (Noss & Daly 2006), least-cost modeling for focal species is the most widely used method for designing corridors to connect existing wildland blocks (e.g., Walker & Craighead 1997; Singleton et al. 2002; Beier et al. 2006, 2007; other studies summarized in Beier et al. 2008).

The objective of least-cost modeling for focal species is to identify the swath of land connecting two or more wildland blocks that minimizes the resistance to (i.e., difficulty of) movement through a landscape for a species (Adriaensen et al. 2003; Beier et al. 2008). Resistance is a function of raster cell attributes in a geographic information system and is usually estimated as the inverse of habitat suitability (Beier et al. 2008). Corridors for multiple focal species are combined into a preliminary linkage design, which becomes the final linkage design upon modification to accommodate ecological processes, non-modeled species, or edge effects (Beier et al. 2008).

Like most other conservation plans, least-cost corridors for focal species have been based on today's species' distributions and land cover, both of which will change with climate change (Hunter et al. 1988; Huntley 2005). Thus, it is uncertain how well these linkages will function when some species currently occupying an area may no longer do so in the near future, while others species may be new arrivals.

To design linkages only using factors that are stable with respect to climate change, Beier and Brost (2010) proposed an alternative least-cost modeling approach based on land facets, or recurring areas on a landscape that have relatively homogenous

topography and soils. Under this approach, the physical landscape is classified into multiple land facets, and a linkage is designed to optimize their connectivity and interspersions. This strategy operates on the premise that diverse physical environments support diverse species (Kirkpatrick & Brown 1994; Faith & Walker 1996; Burnett et al. 1998; Nichols et al. 1998; Cowling et al. 1999; Reyers et al. 2002), and may also support the ecological and evolutionary processes that maintain and generate biodiversity (Cowling et al. 1999; Noss 2001; Moritz 2002; Cowling et al. 2003; Rouget et al. 2006; Pressey et al. 2007; Klein et al. 2009). Thus, a linkage designed to provide continuity for all land facets should not only optimize connectivity for the full diversity of plants and animals, but also sustain vital processes.

Brost and Beier (Chapter 2) provide procedures to design linkages based on land facets. Each linkage includes multiple corridors, one to optimize connectivity for each facet type. Each of these corridors should facilitate movement of species associated with that facet, today and in the future. To better accommodate rapid, short-distance range shifts, interactions between species, and ecological and evolutionary processes such as speciation, these designs also include a corridor to optimize connectivity for high diversity (i.e., interspersions) of land facets. These procedures produce multi-stranded linkage designs (e.g., Fig. 3.1).

Although Beier and Brost (2010) and Brost and Beier (Chapter 2) recommend using land facets in conjunction with focal species to design linkages, conservation practitioners may be limited to a land facets approach in areas where species information is poor or maps of land cover do not exist. Such practitioners might like to know whether a linkage design based on land facets is likely to support movement by local species needing connectivity. Practitioners who have previously designed a linkage for focal species might also want to know if the design provides for continuity and interspersions of land facets, or whether additional analyses are needed to better capture some physical environments.

The purpose of this chapter is to evaluate how well linkages based on land facets provide connectivity for focal species, and how well linkages designed for focal species provide continuity of land facets. If diverse physical environments support diverse species, linkages designed using land facets should serve species today, and indeed must if we expect them to serve species in the future. Conversely, it has been suggested that linkages designed for diverse focal species should also contain diverse physical environments because factors in many focal species models include land cover (partially determined by physical environment) and topography (Beier et al. 2008). Although both expectations are reasonable, this is the first paper to examine

the issue in the context of linkage design.

We conducted this evaluation in three landscapes in Arizona, USA for which linkages have been designed for both focal species and land facets (Table 3.1; Figs. 3.1–3.3). Because true landscape connectivity (for species or land facets) is not known, we examine the performance of one type of design relative to the other. In other words, to evaluate each approach to linkage design, we used as a benchmark the linkage designed under the alternative strategy in the same landscape. We used two metrics to evaluate how well each linkage design served each focal species, namely resistances along least-cost paths and distances between patches of modeled breeding habitat. Because “breeding patches” cannot be defined for land facets, we used resistance along least-cost paths, and the length of the longest high-resistance section of the least-cost path, to evaluate how well each linkage design provided continuity of land facets.

## Methods

### *Linkage planning areas and linkage designs*

For our evaluation, we selected three areas in Arizona for which linkages have been designed for both land facets and focal species (Table 3.1; Figs. 3.1–3.3). In each planning area, linkages were designed to conserve connectivity between large, publicly-owned wildlands separated by a matrix dominated by unprotected lands. Beier et al. (2007) describe the three linkage designs based on focal species and detail the procedures used to produce them. Each focal species design was the union of five to 16 single-species corridors (Table 3.2; Beier et al. 2007) and therefore had multiple strands (Figs. 3.1–3.3). Beier et al. (2007) also describe each area’s ecological significance, conservation investments, threats to connectivity, and patterns of land ownership and land cover. Brost and Beier (in prep.; Chapter 2) describe the three linkage designs based on land facets and the procedures used to produce them. Each land facets design was the union of nine to 12 corridors for individual land facets and one corridor with high interspersion of facets (Chapter 2). Although Beier and Brost (2010) and Brost and Beier (in prep.; Chapter 2) recommend defining land facets on both soil and topographic attributes, adequate soils information was not available in these landscapes. Therefore, land facets were defined on the basis of elevation, slope angle, solar insolation, and topographic position. Both types of linkages (i.e., focal species designs and land facets designs) were designed using least-cost modeling (Adriaensen et al. 2003; Beier et al. 2008).

*Evaluating how well each linkage design served each focal species*

Patches of habitat large enough to support breeding by a species can serve as stepping stones within a linkage, thereby reducing the amount of unsuitable habitat the species must cross in a single event. To evaluate how well each type of linkage design served each focal species, we developed two metrics to quantify the distribution of breeding patches and the quality of the matrix separating them. Beier et al. (2007) identified breeding patches for each focal species by joining adjacent raster cells of modeled breeding habitat into clusters that exceeded the species' average home range size. They estimated habitat suitability using scientific literature and expert opinion, which also served as a measurement of resistance for least-cost analysis. Suitability and resistance values ranged between 1 (best) and 10 (worst) with a value of 5 being the threshold between breeding and non-breeding habitat.

We quantified the amount and distribution of non-breeding habitat encountered by an animal dispersing from one wildland block to the other by measuring the Euclidean distances between modeled breeding patches for each species in each type of linkage design (Fig. 3.4). Corridor termini (i.e., breeding patches wholly contained within the wildland blocks) served as starting/ending points for the measurements, which were made through the strand of the linkage that minimized the longest distance (i.e., gap) between patches. We digitized distances between breeding patches as a line feature class in ArcGIS 9.3 (ESRI, Redlands, CA), which allowed us to verify that the route minimized the longest gap between patches. We enabled vertex and edge "snapping" in the editing tools to ensure accurate measurements.

To quantify the quality of habitat between breeding patches, we also generated the resistance profile of the least-cost path connecting consecutive patches (Fig. 3.4). A resistance profile is a graph of resistance of each cell in the least-cost path plotted against distance along that path. Least-cost paths are similar to least-cost corridors in that both minimize the cumulative resistance incurred while traversing the matrix; however, a path is only 1 cell (i.e., 30 m) wide. We generated a resistance profile for the six largest gaps between patches of modeled breeding habitat for focal species in each landscape (only one species-landscape combination, Arizona gray squirrel in the Santa Rita-Tumacacori planning area, had > 6 gaps). To aid their comparison, we superimposed (in bold) on each profile a symmetric two-sided moving average (Kutner et al. 2005) of the values in the profile:

$$MA(X_i) = \frac{1}{2q+1} \sum_{j=-q}^q X_{i+j}, \quad (3.1)$$

where  $X$  is the raw resistance value and  $i$  indexes the order of cells in the profile.  $q$  determines the size of the moving average, which was defined as

$$q = \begin{cases} i - 1 & 7 > i > n - 7 \\ 7 & 7 \leq i \leq n - 7 \end{cases}, \quad (3.2)$$

where  $n$  is the number of raster cells in the profile. Although procedures like LOESS (locally weighted regression) or generalized additive models could also be used for smoothing (Crawley 2007), a moving average allowed us to easily maintain the same level smoothing in all profiles. We used the Spatial Analyst extension of ArcGIS 9.3 to identify least-cost paths.

About half of the focal species had “locally widespread” habitat, meaning that  $\geq 90\%$  of the matrix between the wildland blocks consisted of modeled breeding habitat for that species. For these species, either the entire area was one breeding patch and interpatch distance was not defined, or only a single, short gap between breeding patches existed (antelope jackrabbit and jaguar in the Santa Rita-Tumacacori planning area). For one species (mountain lion in the Black Hills-Munds Mountain planning area), no gaps between breeding patches existed under either design, even though the species did not have locally widespread habitat. In another case (desert bighorn sheep in the Wickenburg-Hassayampa planning area), the species only had modeled breeding patches in the wildland blocks, such that the entire matrix was a single gap. For species with locally widespread habitat, mountain lion in the Black Hills-Munds Mountain planning area, and bighorn sheep in the Wickenburg-Hassayampa planning area, we assessed linkage performance solely on the basis of resistance profiles between corridor termini.

When paired resistance profiles (i.e., the profiles for a single species and landscape in both types of linkage designs) were approximately equal in length, we considered the two profiles equivalent if both had resistance values  $< 5$  (i.e., sufficient habitat quality, although not necessarily sufficient area, for breeding) along their entire length, or if both contained similar values (including values  $> 5$ ) along their entire length. We considered the difference between profiles to be biologically significant if one profile was predominantly below 5 resistance units while the other was predominantly above 5 units, or if both profiles had several segments above 5 resistance units but one was typically  $> 2$  resistance units worse than the other. In the few cases when length and resistance differed in more complex ways, we evaluated interactions between length and resistance (relative to the threshold value of 5) in light of species mobility.

We considered and rejected two other simple metrics that could be used to compare linkage designs, namely percent overlap and mean resistance of the profiles. Low values of percent overlap are not meaningful because two non-overlapping linkages often provide similar connectivity (Beier et al. 2009). Mean resistance can yield spurious results because it does not reflect the spatial distribution of habitat quality. For example, a profile with low resistance (4 resistance units) over 90% of its length and a complete barrier (10 units) over the remaining 10% would have the same mean resistance (4.6) as a superior profile in which all cells had a uniform resistance of 4.6 units. For another example, a profile 5 km long with a uniform resistance of 4 units throughout its length has a *higher* mean resistance (4) than an *inferior* 10-km profile with a uniform resistance of 4 units for 5 km and 3 units for the remaining 5 km (mean resistance 3.5). Because we could model breeding patches and compare profiles to the threshold value of 5, we did not have to use mean resistance to compare designs. Although somewhat tedious to measure, the combination of metrics that we used quantified the spatial distribution of habitat patches and the resistance of the matrix separating them. We believe that resistance profiles approximate the resistance an animal would encounter in moving along the path, but recognize that there is information lost when using a path-based evaluation (e.g., a pixel-wide path could be surrounded by inhospitable matrix).

*Evaluating how well each linkage design provided continuity for land facets*

We generated resistance profiles between corridor termini to evaluate how well each type of linkage design served each land facet and the diversity of facets. Corridor termini for land facets were defined as the largest polygons within the wildland blocks that were dominated by the focal land facet (Chapter 2). Resistance for land facets was measured using Mahalanobis distance, a multivariate measure of dissimilarity (Clark et al. 1993; Gotelli & Ellison 2004; Chapter 2) of each cell from an ‘ideal’ or characteristic elevation, slope angle, solar insolation, and relative density of the focal land facet within a circular neighborhood with a 3-cell radius (Chapter 2). Resistance values for land facets reflect the departure of a cell from the prototypical cell of the focal facet type, and are measured in multivariate standardized units (analogous to a standard deviation in a univariate analysis).

For the diversity of land facets, corridor termini were defined as the largest polygons consisting of the most diverse cells (Chapter 2). Shannon’s index (Magurran 1988) was used to measure diversity of land facets within a circular neighborhood

with a 5-cell radius (Chapter 2). Shannon's index ( $H'$ ) incorporates measures of both richness and evenness, and its inverse was used as a resistance surface in least-cost modeling (Chapter 2). Thus, cells with a high diversity index had low resistance. To aid the interpretation of resistance profiles for the diversity of land facets, we converted values in the profiles to the complement of Shannon's evenness ( $E_H$ ; Magurran 1988):

$$1 - E_H = 1 - \frac{H'}{\ln(L)}, \quad (3.3)$$

where  $L$  is the number of land facets in a particular landscape and  $\ln(L)$  is the maximum value of Shannon's index. The subsequent values in the profile are scaled  $[0, 1]$ , where 0 is the lowest possible resistance (all land facets occur in equal proportions) and 1 is the maximum deviation from the optimal Shannon's index value. A symmetric two-sided moving average of the same form used to aid comparison of resistance profiles for focal species was also superimposed (in bold) on the profiles for land facets and the diversity of facets.

The interpretation of resistance values for land facets is less clear than resistance values for focal species because we do not know how Mahalanobis distance or Shannon's index translates into resistance to movement of species associated with land facets. Therefore, the resistance profiles can suggest which linkage design performed better for a particular land facet or the diversity of land facets, but conclusions about the biological significance of such a difference cannot be made.

We quantified the performance of the two types of linkage designs with respect to land facets and the diversity of land facets in two ways. First we compared the mean resistances of the two resistance profiles. Because the resistance scale for land facets lacked a meaningful reference value (such as the threshold of 5 in the resistance scale for focal species), we used mean resistance despite difficulties in interpreting it (see *Evaluating how well each linkage design served each focal species*). To avoid these difficulties, we examined profiles for artifacts that could give rise to spurious inferences (e.g., Table 3.4, footnote a). Second, we measured the length of the longest high-resistance segment of the profiles for land facets. To identify high-resistance segments, we rescaled Mahalanobis distances to  $[0, 1]$  by calculating the p-value associated with each Mahalanobis distance (Farber & Kadmon 2003), and then identified the longest segment of continuous p-values  $< 0.05$  in the resistance profile. Under multivariate normality, Mahalanobis distances are approximately  $\chi^2$  distributed with  $n - 1$  degrees of freedom, where  $n$  is the number of independent variables used to

compute Mahalanobis distance (Clark et al. 1993). For example, a Mahalanobis distance of 6 with 2  $df$  corresponds to a p-value  $\approx 0.05$ . Because our data were not multivariate normal, the p-values should not be interpreted in terms of statistical or biological significance. Nonetheless, p-values closer to 0 indicate higher dissimilarity between a cell and the focal facet type, and the longest segment of p-values  $< 0.05$  is a useful, consistent metric to compare resistance profiles.

When the mean value of paired resistance profiles for a land facet differed  $\leq 5$  units and the longest segments of high resistance differed  $\leq 5$  km, we considered the performance of the designs equivalent for that land facet. If the mean resistance of the profiles differed  $\leq 5$  units but the longest high-resistance segments differed  $\geq 5$  km, we considered the design with the shorter high-resistance segment better. Conversely, if the mean resistance of the profiles differed  $\geq 5$  units but the longest segments of high resistance differed  $\leq 5$  km, we considered the design with the lower mean resistance better unless the resistance profile indicated that the mean resistance was affected by an artifact (one case, Table 3.4 footnote). There were no cases in which mean resistance of a design was  $> 5$  units better but the longest high-resistance segment was  $> 5$  km longer than the alternative design. We considered the performance of the designs equivalent with respect to high diversity of land facets if the means of the paired profiles differed  $\leq 0.1$  (i. e., 10% of the optimal Shannon's index value). If they differed  $> 0.1$ , the linkage design with the lower mean resistance was considered better.

## Results

The linkage designed for land facets was 21% larger than the linkage designed for focal species in the Wickenburg-Hassayampa planning area, and nearly twice as large as the focal species design in the Black Hills-Munds Mountain planning area (Table 3.1). In the Santa Rita-Tumacacori planning area, the linkage designed for focal species was 15% larger than the linkage designed for land facets. Combining the two types of designs would result in a linkage 77% larger, on average, than the focal species linkage designs (Table 3.1).

### *How well were focal species served by each linkage design?*

Of the 28 species-landscape combinations, 16 focal species had locally widespread habitat (Table 3.2). For these species, resistance profiles did not differ substantially



between the focal species and land facets linkage designs (Figs. 3.5–3.7). The largest apparent difference was for badger in the Wickenburg-Hassayampa and Santa Rita-Tumacacori planning areas, for which values in the resistance profiles were often 2.5 units greater (on a scale of 1–10) under the land facets designs than under the focal species designs (Figs. 3.6 and 3.7). Profiles for all other species showed less difference between designs.

Of the 12 remaining species (those whose habitat was not locally widespread), four were served equally well by both designs and five species—especially black bear and elk in the Black Hills-Munds Mountain planning area—were served better by the land facets linkage designs (Table 3.2). The focal species designs provided better continuity of breeding patches for three species, namely Arizona gray squirrel, black-tailed rattlesnake, and tiger rattlesnake in the Santa Rita-Tumacacori planning area. Further details are provided below.

In the Black Hills-Munds Mountain planning area, both types of linkage designs provided continuous habitat between wildland blocks for mountain lion (Fig. 3.5). For black bear, the only gap between breeding patches was 26.0 km long in the focal species design compared to two gaps of 8.6 and 1.9 km in the land facets design (Table 3.3). The distance between elk breeding patches was similar between the two designs; however, the resistance profile was substantially lower in the land facets design (Fig. 3.8). Black bear and elk were served best by the linkage strand for high elevation land facets (Fig. 3.1, strand A), which corresponds to the forest habitat associated with these species (Beier et al. 2007).

In the Wickenburg-Hassayampa planning area, modeled breeding habitat for six of seven focal species in this landscape was locally widespread. For desert bighorn sheep, the land facets design encompassed the Vulture Mountains, the only area in the matrix containing modeled habitat for the species. Therefore, resistance profiles did not differ between designs (Fig. 3.6).

In the Santa Rita-Tumacacori planning area, the maximum distance between breeding patches for white-nosed coati under the land facets design was 0.3 km (31%) less than in the focal species design (Table 3.3). However, this shorter maximum distance was offset by the greater collective length of the two gaps (and their relatively higher resistance profiles) under the land facets design (Fig. 3.9). Maximum distances between breeding patches for Coues' white-tailed deer, mountain lion, and porcupine were about half as long in the land facets design than in the focal species design (Table 3.3). Values in the resistance profiles were less under the land facets design for Coues' white-tailed deer and porcupine, and approximately the same as the

focal species design for mountain lion (Fig. 3.9). The focal species design contained more breeding habitat for Arizona gray squirrel, black-tailed rattlesnake, and tiger rattlesnake, and therefore had shorter distances between patches than the land facets design (0.9–1.2 km less; Table 3.3); however, values in the resistance profiles for these species did not differ substantially between the two designs (Fig. 3.9).

*How well were land facets served by each linkage design?*

The focal species linkage designs performed as well as the land facets designs for approximately half of the land facets in the Black Hills-Munds Mountain and Wickenburg-Hassayampa planning areas, and for all but one facet type in the Santa Rita-Tumacacori planning area (Table 3.4). The focal species designs served the diversity of facets as well as the land facets designs in all three landscapes. A focal species design never provided better continuity for a land facet than a land facets design. Further details are provided below.

In the Black Hills-Munds Mountain planning area, the focal species and land facets designs performed similarly for five of 11 land facets as well as for the diversity of land facets (Table 3.4; Figs. 3.10 and 3.13). For the remaining land facets, mean values of the resistance profiles were 19.9 units (146%) higher and the longest high-resistance segments were 14.8 km (279%) longer, on average, under the focal species linkage design. Least-cost paths for all land facets under the focal species design passed through the black bear corridor (Fig. 3.1, strand 3), a species that was modeled to select canyon bottoms, ridges, and steep slopes (Beier et al. 2007). The least-cost path for the diversity of facets passed through the elk corridor (Fig. 3.1, strand 1), which is the focal species corridor closest to the corridor explicitly designed for high diversity of land facets (Fig. 3.1, strand B).

In the Wickenburg-Hassayampa planning area, the performance of the two types of linkage designs was similar for five of nine land facets as well as for the diversity of facets (Table 3.4; Figs. 3.11 and 3.13). For the remaining land facets, the longest high-resistance segments in the resistance profiles were 10.2 km (162%) longer, on average, under the focal species linkage design; however, mean values of the resistance profiles only differed 4.2 units (15%), on average. For high elevation canyon bottoms and ridges, the focal species design did not provide a direct route to the respective termini in the Wickenburg-Hassayampa wildland block, located at the northeastern terminus of strand A under the land facets design (Fig. 3.2). Consequently, the corresponding least-cost paths under the focal species design required a longer, looping route to

reach the termini than those under the land facets design.

In the Santa Rita-Tumacacori planning area, the focal species and land facets designs performed similarly for 11 of 12 land facets as well as for the diversity of land facets (Table 3.4; Figs. 3.12 and 3.13). Mean resistance of the profile for low elevation, gentle, warm slopes was 7.7 units (335%) higher under the focal species design.

## Discussion

### *How well do linkages designed for land facets serve focal species?*

Linkages designed for land facets served 25 of 28 focal species as well as or better than the focal species designs (Table 3.2). For the 16 species with locally widespread habitat, similar performance under both types of designs was more a virtue of the distribution of these species' habitat than the placement of the linkage design. In fact, any linkage design that excludes urban or disturbed areas would likely perform as well for these species as these two least-cost modeling approaches.

The remaining species had patchily distributed habitat, and thus provide a more meaningful assessment of the land facets approach to linkage design. Among these species, the land facets designs performed as well as or better than the focal species designs for all of the large mammals, and also for porcupine and white-nosed coati (Table 3.2). The land facets design in the Santa Rita-Tumacacori planning area performed worse for Arizona gray squirrel, black-tailed rattlesnake, and tiger rattlesnake. Breeding patches for these three species were also the most narrowly distributed, suggesting that species with limited habitat in the planning area tend to be better served by a focal species approach.

Shorter distances and/or lower resistance between breeding patches do not necessarily translate into increased connectivity, which ultimately depends on the interaction between the linkage design and a species' life history traits (e.g., mobility, behavior, generation time). For example, in the Santa Rita-Tumacacori planning area, shorter distances between breeding patches under the land facets design for Coues' white-tailed deer and mountain lion (Table 3.3) may or may not lead to a substantial increase in connectivity for these highly mobile species (Sweaner et al. 2000; McCoy et al. 2005). In contrast, small increases (0.9–1.2 km) in distances between patches can be important for less-mobile species like Arizona gray squirrel, black-tailed rattlesnake, and tiger rattlesnake. In the Black Hills-Munds Mountain planning area,

substantially shorter distances between breeding patches for black bear and lower resistance between patches for elk under the land facets design would probably provide increased connectivity for these species over the focal species design (Table 3.3; Fig. 3.8).

As demonstrated by our evaluation, focal species corridors do not necessarily minimize distances between breeding patches because least-cost analysis seeks to minimize resistance-weighted distance rather than Euclidean distance between patches. However, we were surprised that for some species the land facets designs simultaneously provided shorter distances between patches and resistance profiles consisting of values similar to or below those of the focal species designs (e.g., black bear and elk in the Black Hills-Munds Mountain planning area; Table 3.2). For these species, the land facets strand that optimized continuity of breeding patches was relatively long but biologically more effective. Perhaps the resistance surfaces used to design least-cost corridors for these species did not adequately represent the difference between breeding habitat and less hospitable matrix. Or, perhaps least-cost models do not accurately depict how a species makes gap-crossing decisions.

*How well do linkages designed for focal species serve land facets?*

Linkage designs for focal species and land facets performed similarly for 21 of 32 land facets (Table 3.4). The designs also performed similarly for the diversity of land facets in all three landscapes. In the Black Hills-Munds Mountain and Wickenburg-Hassayampa planning areas, the focal species designs performed as well as the land facets designs for approximately half of the land facets.

In these two landscapes, some land facets were probably served well because a focal species was associated with that land facet type. For example, the focal species design in the Wickenburg-Hassayampa planning area provided good continuity for low elevation canyon bottoms and ridges because these features were important factors in the habitat models (Beier et al. 2007) for desert bighorn sheep (ridges) and Gila monster (ridges and canyon bottoms).

The focal species design in the Santa Rita-Tumacacori planning area performed as well as the land facets design for 11 of 12 land facets and for high diversity of land facets (Table 3.4). Although this result could be attributed to the number of focal species modeled in this planning area (16, which is more than twice the number of focal species used in the others), it is more likely due to the spatial configuration of the landscape's topography. The San Cayetano Mountains, an isolated group of

peaks lying between the wildland blocks, represents the only mountainous terrain in the matrix. As such, 11 of 12 land facet corridors and the high diversity corridor (Fig. 3.3, strands A and B) contained these mountains. So did corridors for Arizona gray squirrel, black bear, black-tailed rattlesnake, and tiger rattlesnake (Fig. 3.3, strands 3 and 4), resulting in extensive overlap between the two types of linkage designs.

#### *Implications for linkage design*

Our results lend support to the underlying concept of the land facets approach to linkage design, which is that diverse physical environments support diverse biota. In areas where a focal species approach to linkage design is not possible, our results suggest that conservation practitioners can solely implement a land facets approach with some confidence that the linkage design would serve most potential focal species. However, it is clear that linkages designed for land facets will fail to serve some species, especially those with limited habitat in the planning area. Similarly, linkages designed for focal species tend to capture diverse topography. Unfortunately, we do not know how Mahalanobis distance or Shannon's index translates into resistance to movement of species associated with a land facet or the diversity of facets. Nonetheless, compared to land facets designs, focal species designs provided less continuity for many land facets.

We support the recommendation of Beier and Brost (2010) that the land facets approach to linkage design should complement, rather than replace, focal species approaches. But simply combining the two types of linkage designs would produce a very large linkage design that would be expensive to conserve. For example, combining designs in our three landscapes would result in a linkage 30% to 200% larger than the focal species design (Table 3.1). Given the fact that the land facets designs provide as much connectivity as the focal species designs for most species, such linkage designs are needlessly large. Indeed, in two of our areas, the land facets designs served all land facets and focal species as well as or better than the focal species designs. In the third landscape, a conservation planner could efficiently provide connectivity for all species and facets by expanding the land facets design to encompass some of the same breeding patches for Arizona gray squirrel, black-tailed rattlesnake, and tiger rattlesnake that are contained in the focal species design. The new design would only be 3% larger than the original land facets design (Table 3.1).

Before prescribing the use of land facets as a standalone or primary approach to linkage design, additional evaluation is necessary. Evaluations conducted in land-

scapes where many focal species have narrowly distributed habitat would be particularly informative. If further evaluations reinforce the results from our three landscapes, such findings would certainly bolster confidence in the land facets approach to linkage design. This additional information could also be used to develop a general strategy for using land facets in linkage design, which could address the minimum width of corridors or linkage strands, how best to combine linkages based on land facets and focal species, how to limit redundancy in linkage designs (i.e., removal of pixels that do not negatively affect connectivity), and how to enlarge designs to better serve certain species.

Table 3.1. Location, number of focal species and land facets modeled, and size of linkage designs in each of the three planning areas used in our evaluation.

Planning area	Number of focal species modeled	Number of land facets modeled (including diversity of land facets)	Size of focal species linkage design (km <sup>2</sup> )		Size of land facets linkage design (km <sup>2</sup> )		Size of combined linkage designs (km <sup>2</sup> )	Minimum size of linkage design needed to serve all focal species and land facets <sup>a</sup> (km <sup>2</sup> )	Previous column as percent of	
			linkage design (km <sup>2</sup> )	species linkage design (km <sup>2</sup> )	linkage design (km <sup>2</sup> )	facets linkage design (km <sup>2</sup> )			land facets design	focal species design
Black Hills-Munds Mountain (111.9 W, 34.7 N)	5	12	247.9		492.9	492.9	597.8	492.9	100%	199%
Wickenburg-Hassayampa (112.8 W, 33.9 N)	7	10	340.7		411.8	411.8	611.0	411.8	100%	121%
Santa Rita-Tumacacori (111.0 W, 31.6 N)	16	13	363.3		316.4	316.4	473.4	329.0	103%	91%

<sup>a</sup>This design was created by expanding the land facets linkage design so that gaps between breeding patches were no longer than the corresponding gaps in the focal species design. Such expansion was necessary only in the Santa Rita-Tumacacori planning area.

Table 3.2. Relative performance of linkage designs with respect to focal species. Footnotes in the last 2 columns indicate the metrics that differed most between the two types of linkage designs.

Planning area	Locally widespread species for which resistance profiles suggest that the land facets design and focal species design performed equally well	Utility assessed on the basis of distances between modeled breeding patches and resistance profiles		
		Species for which the land facets and focal species design performed equally well	Species for which the land facets design performed better than the focal species design	Species for which the land facets design performed worse than the focal species design
Black Hills-Munds Mountain	Javelina ( <i>Tayassu tajacu</i> )	Mountain lion ( <i>Puma concolor</i> )	Black bear <sup>T</sup> ( <i>Ursus americanus</i> )	
	Mule deer ( <i>Odocoileus hemionus</i> )		Elk <sup>R</sup> ( <i>Cervus elephus</i> )	
Wickenburg-Hassayampa	Badger ( <i>Taxidea taxus</i> )	Desert bighorn sheep ( <i>Ovis canadensis nelsoni</i> )		
	Black-tailed jackrabbit ( <i>Lepus californicus</i> )			
	Desert tortoise ( <i>Gopherus agassizii</i> )			
	Gila monster ( <i>Heloderma suspectum</i> )			
	Javelina ( <i>Tayassu tajacu</i> )			
	Mule deer ( <i>Odocoileus hemionus</i> )			
Santa Rita-Tumacacori	Antelope jackrabbit ( <i>Lepus alleni</i> )	Black bear ( <i>Ursus americanus</i> )	Coues' white-tailed deer <sup>L,R</sup> ( <i>Odocoileus virginianus couesi</i> )	Arizona gray squirrel <sup>L</sup> ( <i>Sciurus arizonensis</i> )
	Badger ( <i>Taxidea taxus</i> )	White-nosed coati ( <i>Nasua narica</i> )		Black-tailed rattlesnake <sup>T</sup> ( <i>Crotalus molossus</i> )
	Desert box turtle ( <i>Terrapene ornata luteola</i> )		Mountain lion <sup>L</sup> ( <i>Puma concolor</i> )	Tiger rattlesnake <sup>T</sup> ( <i>Crotalus tigris</i> )
	Jaguar ( <i>Panthera onca</i> )		Porcupine <sup>L,R</sup> ( <i>Erethizon dorsatum</i> )	
	Javelina ( <i>Tayassu tajacu</i> )			
	Mule deer ( <i>Odocoileus hemionus</i> )			
	Sonoran desert toad ( <i>Bufo alvarius</i> )			
	Sonoran whipsnake ( <i>Masticophis lateralis</i> )			

<sup>T</sup>Total length of gaps between modeled breeding patches.

<sup>L</sup>Length of largest gap between modeled breeding patches.

<sup>R</sup>Values in resistance profiles.



Table 3.3. Distances between breeding patches for focal species. Except for Arizona gray squirrel in the Santa Rita-Tumacacori planning area, all gaps are listed.

Planning area	Focal species	Distances between breeding patches under focal species linkage design (km)	Distances between breeding patches under land facets linkage design (km)
Black Hills-Munds Mountain	Black bear	26.01	8.55
			1.87
	Elk	9.13	8.66
Wickenburg-Hassayampa	Desert bighorn sheep	41.89	42.65
Santa Rita-Tumacacori	Antelope jackrabbit	0.03	0.03
	Arizona gray squirrel <sup>a</sup>	3.37	4.53
		3.21	2.11
		0.98	1.55
		0.90	0.98
		0.90	0.90
		0.86	0.90
	Black-tailed rattlesnake	1.39	1.39
		0.72	1.24
		0.23	1.17
		0.16	0.72
		0.07	0.07
		0.06	
	Black bear	7.87	7.87
	Coues' white-tailed deer	1.45	0.69
			0.67
	Jaguar	0.11	0.19
	Mountain lion	3.70	1.70
	Porcupine	0.63	0.30
		0.03	0.19
	Tiger rattlesnake	0.47	1.38
		0.46	0.73
		0.39	0.70
		0.18	0.11
		0.11	0.03
	White-nosed coati	1.33	1.02
			0.99

<sup>a</sup>Arizona gray squirrel had 22 gaps between breeding patches under both types of linkage designs (mean length of gaps: focal species design = 0.70 km; land facets design = 0.74 km).

Table 3.4. Mean resistance of the resistance profiles and longest high-resistance segments in resistance profiles for land facets under the two types of linkage designs. The “Interpretation” column indicates which type of design performed better for a particular land facet, or if they performed the same. Resistance was calculated as Mahalanobis distance (minimum 0, no theoretical maximum) for land facets, and as the complement of Shannon’s evenness (0 to 1) for land facet diversity.

Planning area	Land facet	Mean resistance		Longest segment of resistance profile w/ p-value < 0.05 (km)		Interpretation (which design was better)
		land facets design	focal species design	land facets design	focal species design	
Black Hills-Munds Mountain	Canyon bottom: low elevation, gentle	19.5	30.6	6.4	17.1	Land facets
	Canyon bottom: mid elevation, steep	33.0	33.0	27.9	27.8	Same
	Canyon bottom: high elevation, gentle	21.2	38.0	16.0	41.7	Land facets
	Ridge: low elevation, gentle	14.6	32.6	5.6	17.3	Land facets
	Ridge: mid elevation, steep	40.3	40.3	28.9	28.9	Same
	Ridge: high elevation, gentle	17.1	46.2	14.4	36.8	Land facets
	Slope: low elevation, gentle, warm	3.1	12.0	0.4	4.1	Land facets
	Slope: low elevation, steep, cool	26.0	27.2	16.6	16.0	Same
	Slope: mid elevation, gentle, warm <sup>a</sup>	56.4	39.4	26.5	25.9	Same
	Slope: mid elevation, steep, warm	33.5	31.6	20.4	21.0	Same
	Slope: high elevation, gentle, hot	22.5	58.1	26.5	41.3	Land facets
	High diversity	0.43	0.48	—	—	Same

<sup>a</sup>Although the mean resistance in the focal species design was 17 units lower than the mean resistance in the land facets design, we considered the performance of the two designs the same because the resistance profiles were nearly identical except for an additional 7 km segment of low resistance that reduced the mean value of the focal species profile (Fig. 3.10). We inspected all profiles; this was the only pair for which the profile caused us to over-ride our impression based on the data in this table.

Table 3.4. (continued)

Planning area	Land facet	Mean resistance		Longest segment of resistance profile w/ p-value < 0.05 (km)		Interpretation (which design was better)
		land facets design	focal species design	land facets design	focal species design	
Wickenburg- Hassayampa	Canyon bottom: low elevation, gentle	29.8	29.8	18.7	18.7	Same
	Canyon bottom: low elevation, steep	29.5	30.0	48.4	64.0	Land facets
	Canyon bottom: high elevation, steep	32.9	38.1	63.1	78.7	Land facets
	Ridge: low elevation, gentle	30.1	30.5	21.8	20.5	Same
	Ridge: low elevation, steep	30.1	31.5	18.1	22.8	Same
	Ridge: high elevation, steep	28.1	36.6	50.0	72.5	Land facets
	Slope: low elevation, gentle, warm	0.6	3.5	0.0	0.6	Same
	Slope: mid elevation, steep, cool	19.0	21.5	6.3	16.5	Land facets
	Slope: high elevation, steep, warm	14.1	16.3	6.6	6.6	Same
	High diversity	0.45	0.55	—	—	Same

Table 3.4. (continued)

Planning area	Land facet	Mean resistance		Longest segment of resistance profile w/ p-value < 0.05		Interpretation (which design was better)
		land facets design	focal species design	land facets design	focal species design	
Santa Rita- Tumacacori	Canyon bottom: low elevation, gentle	25.1	25.1	9.5	7.4	Same
	Canyon bottom: mid elevation, gentle	38.4	38.6	20.8	20.7	Same
	Canyon bottom: mid elevation, steep	26.0	26.0	9.2	9.2	Same
	Canyon bottom: high elevation, steep	47.2	47.8	19.8	21.1	Same
	Ridge: low elevation, gentle	20.0	22.4	6.8	7.8	Same
	Ridge: mid elevation, gentle	33.3	33.6	18.0	17.7	Same
	Ridge: mid elevation, steep	25.9	25.2	8.9	9.0	Same
	Ridge: high elevation, steep	49.4	49.6	18.7	20.2	Same
	Slope: low elevation, gentle, warm	2.3	10.0	0.2	2.0	Land facets
	Slope: mid elevation, steep, cool	31.3	29.9	12.3	11.5	Same
	Slope: mid elevation, steep, hot	29.1	29.0	9.6	9.7	Same
	Slope: high elevation, gentle, hot	25.9	25.8	32.6	31.1	Same
	High diversity	0.47	0.47	—	—	Same

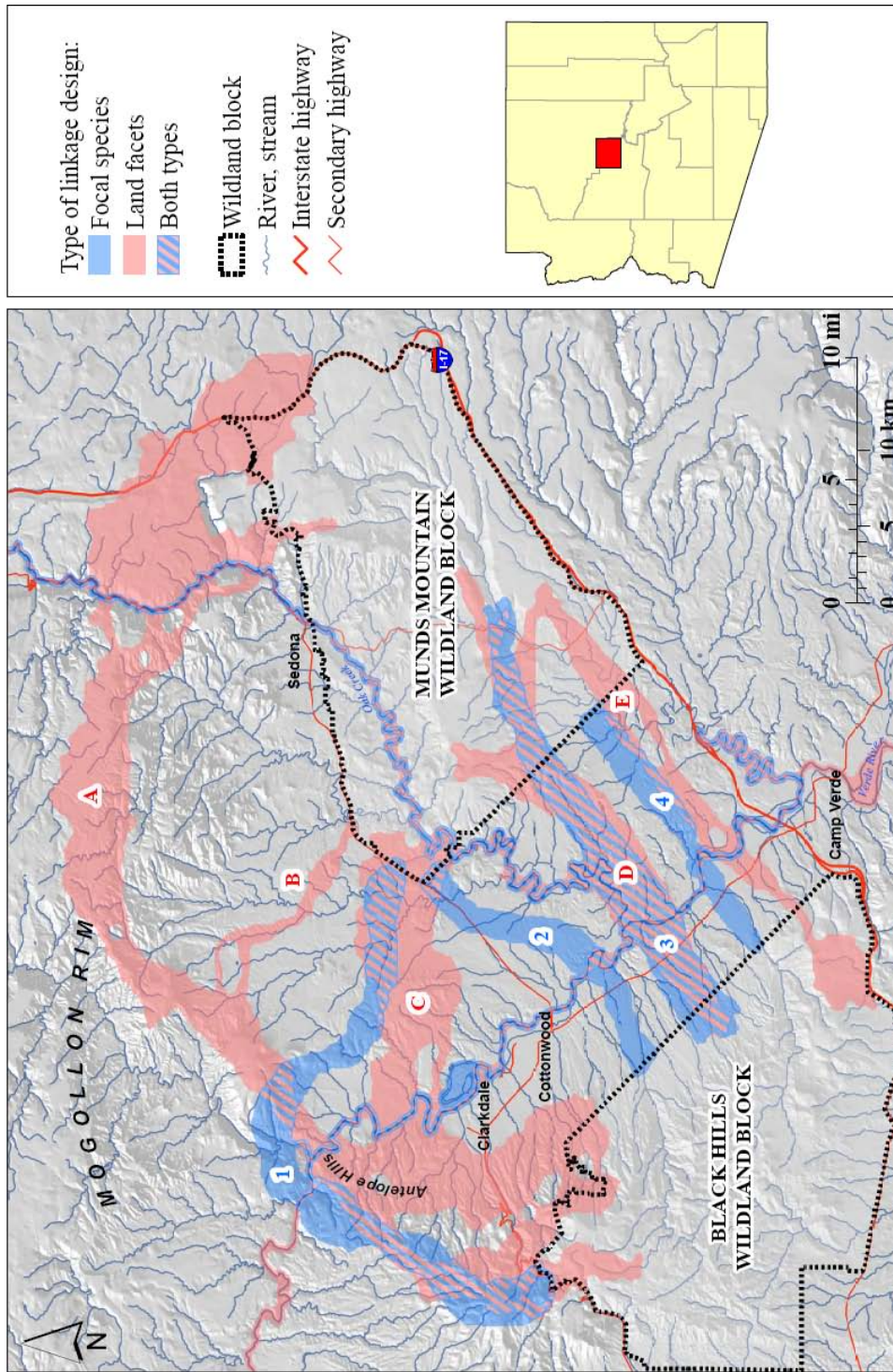


Figure 3.1. Map of the land facets and focal species linkage designs for the Black Hills-Munds Mountain planning area. Linkage strands in the land facets design (pink) consisted of corridors for (A) high elevation, gentle canyon bottoms and ridges; and high elevation, gentle, hot slopes; (B) high diversity of land facets; (C) low elevation, gentle, canyon bottoms and ridges; and low elevation, gentle, warm slopes; (D) mid elevation, steep canyon bottoms and ridges; low elevation, steep, cool slopes; and mid elevation, steep, warm slopes; and (E) mid elevation, gentle, warm slopes. Linkage strands in the focal species design (blue) consisted of corridors for (1) elk, (2) mule deer, (3) black bear, and (4) javelina. Although a corridor was not designed explicitly for mountain lion, linkage strands 1 and 2 contained large amounts of optimal or suitable habitat for this species.



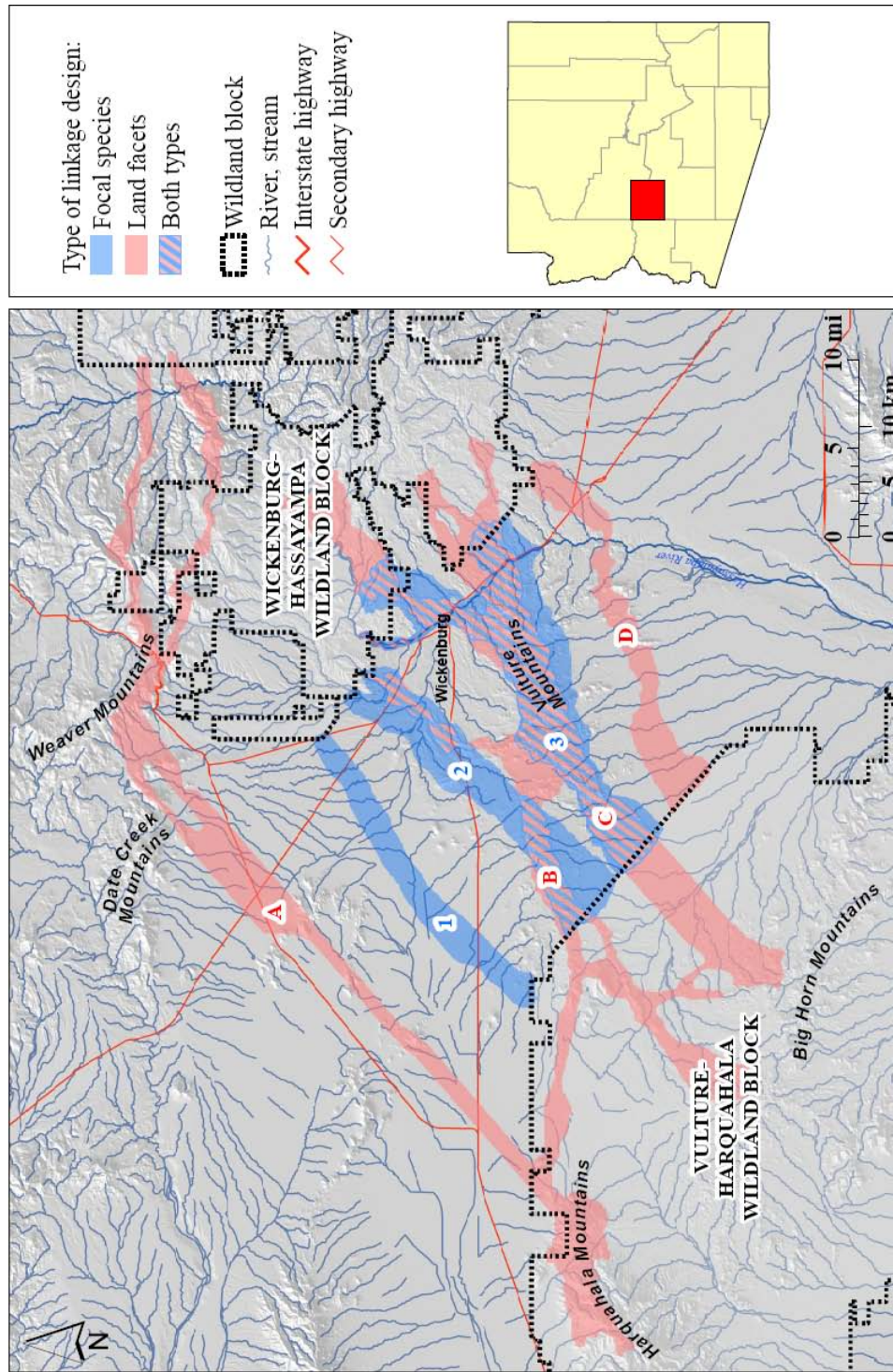


Figure 3.2. Map of the land facets and focal species linkage designs for the Wickenburg-Hassayampa planning area. Linkage strands in the land facets design (pink) consisted of corridors for (A) high elevation, steep canyon bottoms and ridges; (B-C) low elevation, gentle canyon bottoms and ridges; low elevation, steep canyon bottoms and ridges; mid elevation, steep, cool slopes; high elevation, steep, warm slopes; and high diversity of land facets; and (D) low elevation, gentle, warm slopes. Linkage strands in the focal species design (blue) consisted of corridors for (1) badger; (2) black-tailed jackrabbit, javelina, and mule deer; and (3) desert bighorn sheep, desert tortoise, and Gila monster.



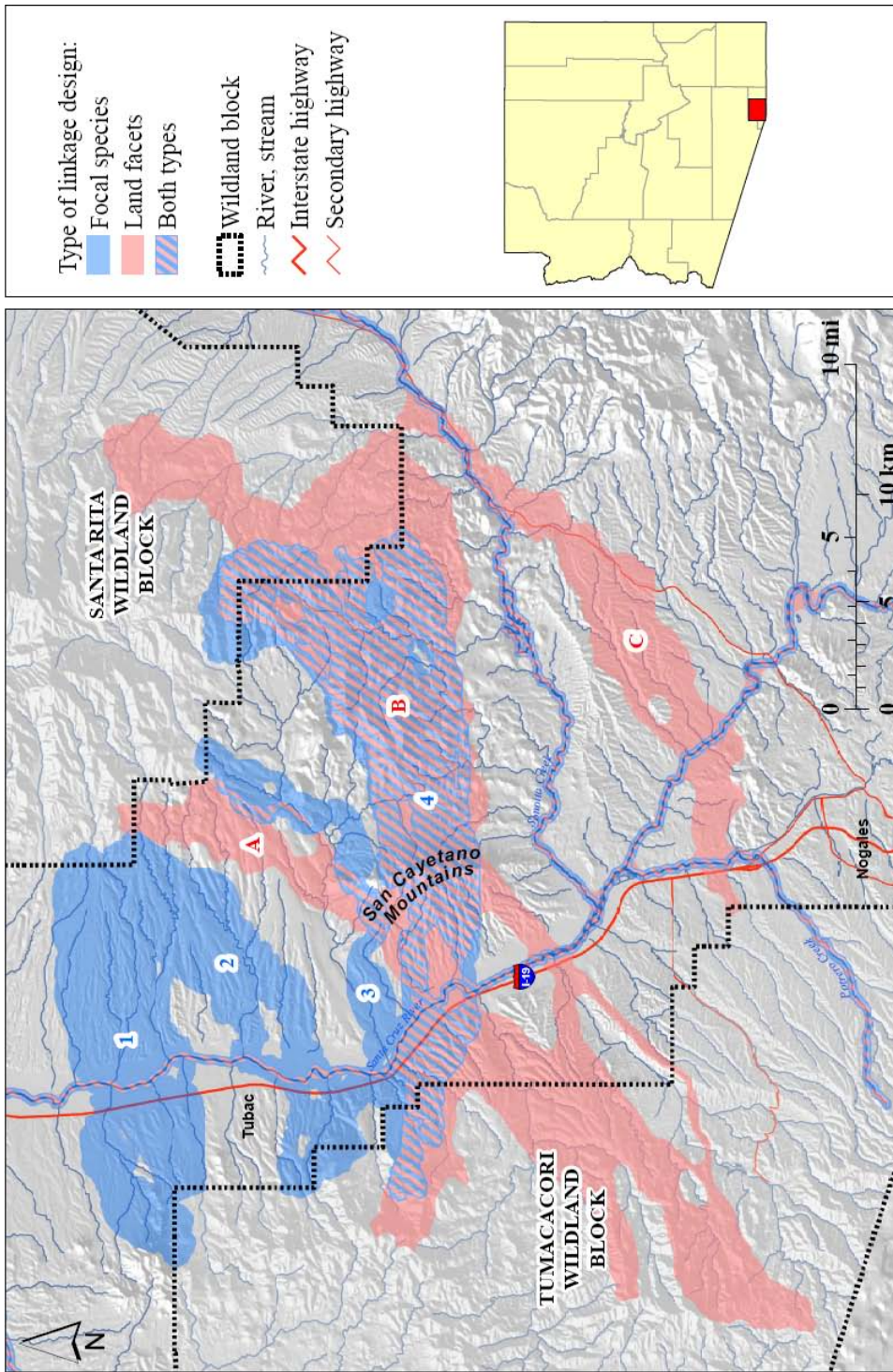


Figure 3.3. Map of the land facets and focal species linkage designs for the Santa Rita-Tumacacori planning area. Linkage strands in the land facets design (pink) consisted of corridors for (A) high elevation, steep canyon bottoms; (B) low elevation, gentle canyon bottoms and ridges; mid elevation, gentle canyon bottoms and ridges; mid elevation, steep canyon bottoms and ridges; high elevation, steep, cool slopes; mid elevation, steep, hot slopes; high elevation, gentle, hot slopes; and high diversity of land facets; and (C) low elevation, gentle, warm slopes. Linkage strands in the focal species design (blue) consisted of corridors for (1) antelope jackrabbit, badger, desert box turtle, jaguar, javelina, mountain lion, mule deer, Sonoran desert toad, and Sonoran whipsnake; (2) Coues' white-tailed deer, porcupine, and white-nosed coati; (3) tiger rattlesnake; and (4) Arizona gray squirrel, black bear, and black-tailed rattlesnake.

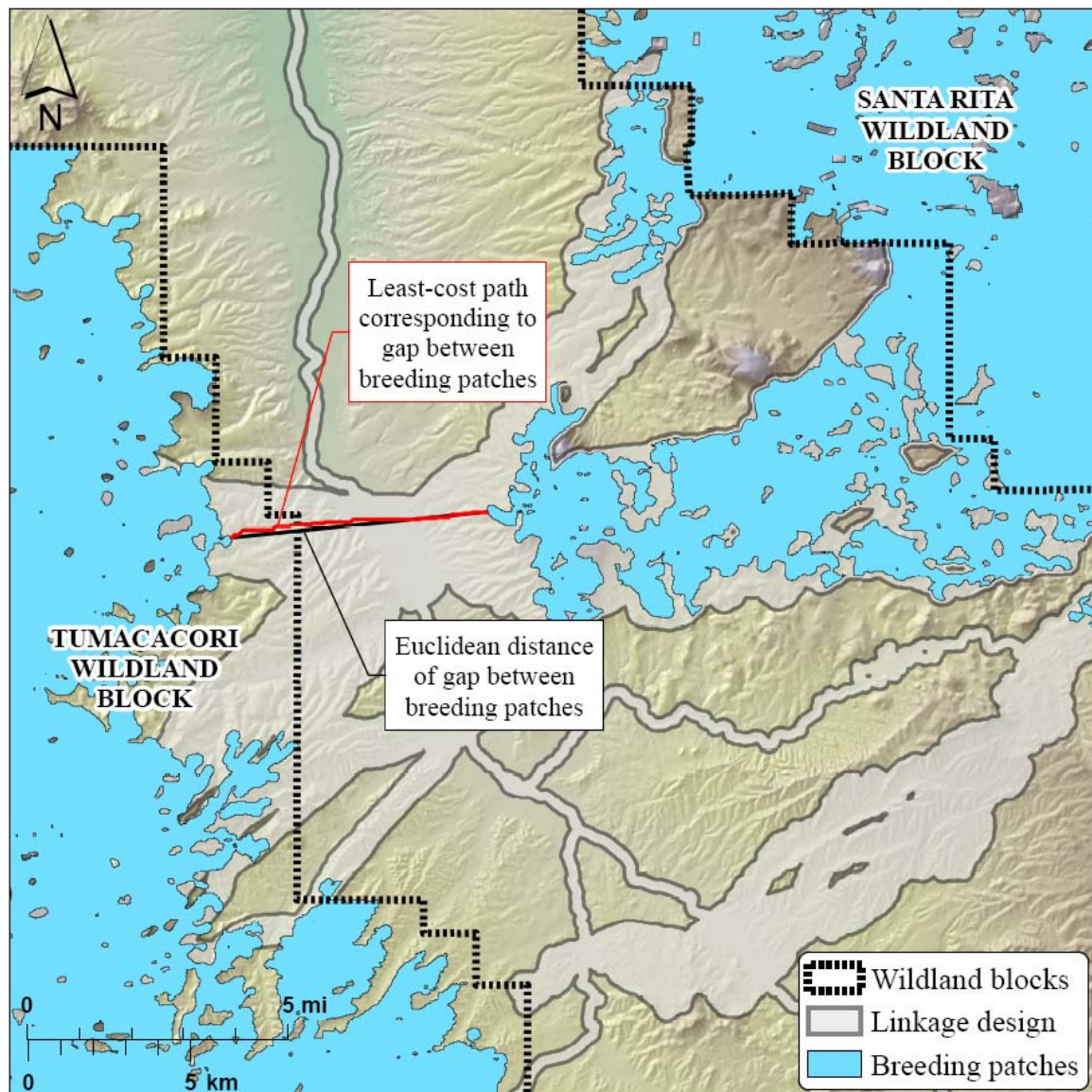


Figure 3.4. Metrics used to evaluate the performance of linkage designs with respect to focal species. This illustration shows the distance between breeding patches (black line) and the associated least-cost path (red line) for black bear in the Santa Rita-Tumacacori planning area under the land facets linkage design. The resistance profile corresponding to the least-cost path is a graph of the resistance of each cell in the least-cost path plotted against distance along that path (see Fig. 3.9). Both metrics were also calculated for the focal species design for comparison.



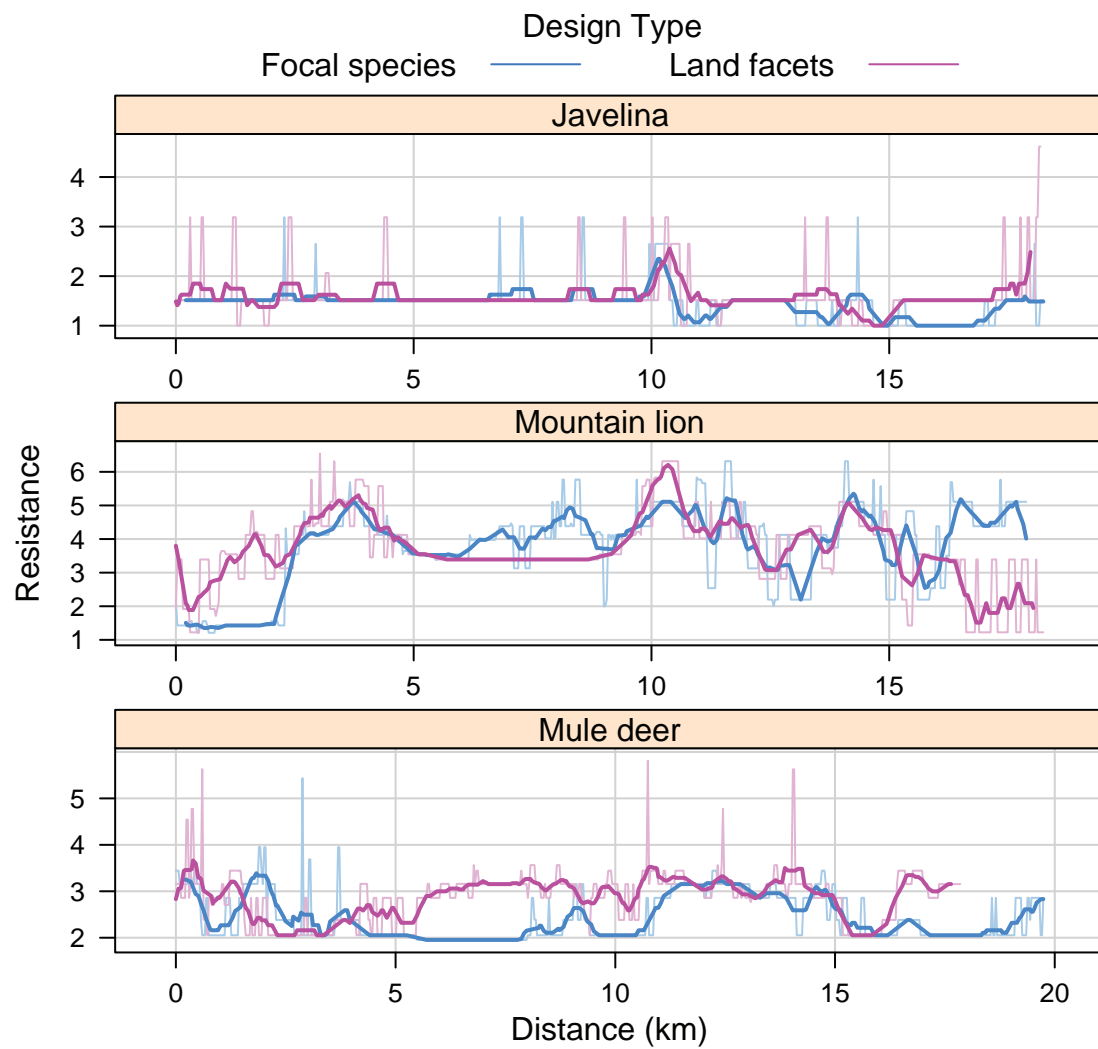


Figure 3.5. Resistance profiles for species with locally widespread habitat and mountain lion in the Black Hills-Munds Mountain planning area.

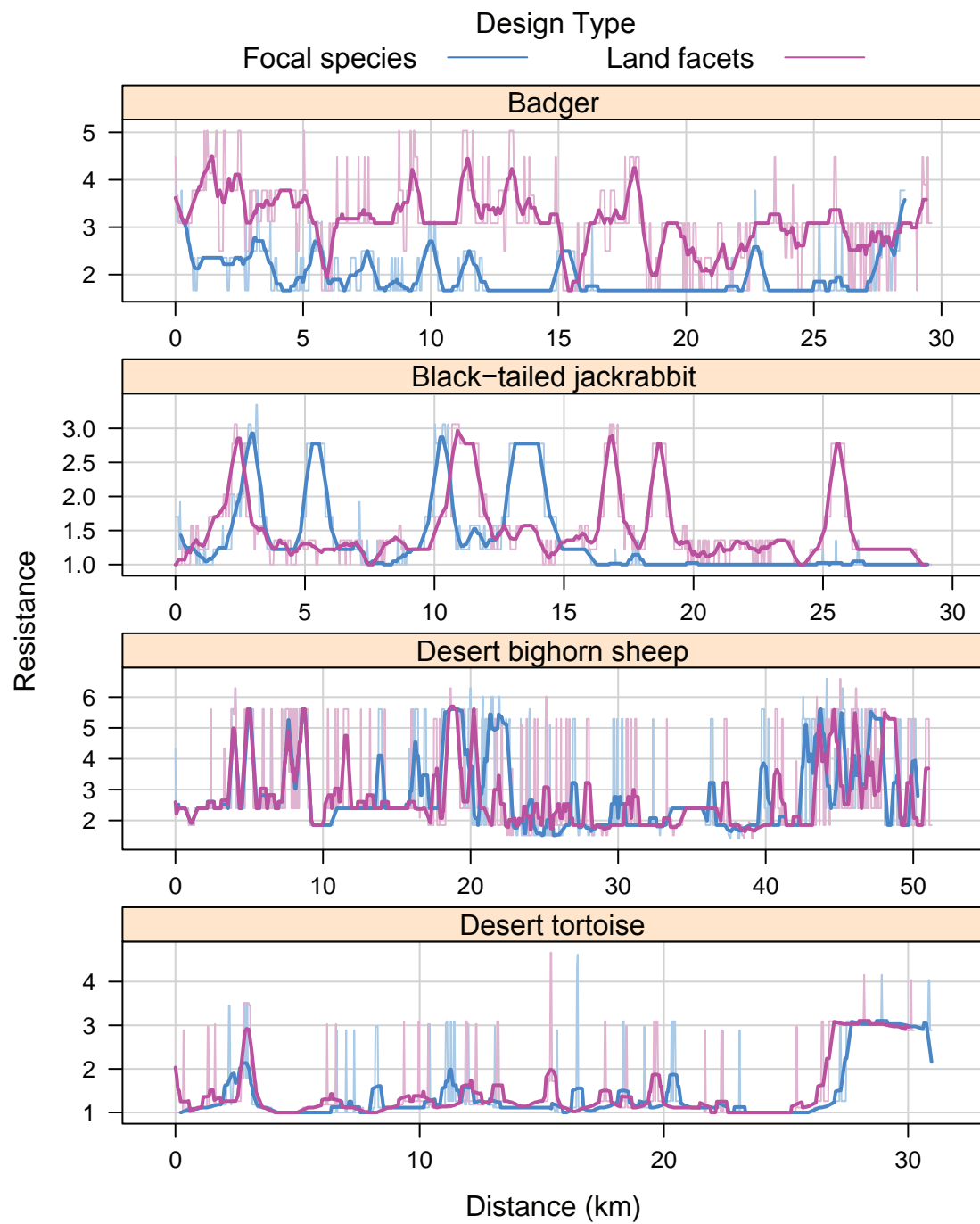


Figure 3.6. Resistance profiles for species with locally widespread habitat and desert bighorn sheep in the Wickenburg-Hassayampa planning area.

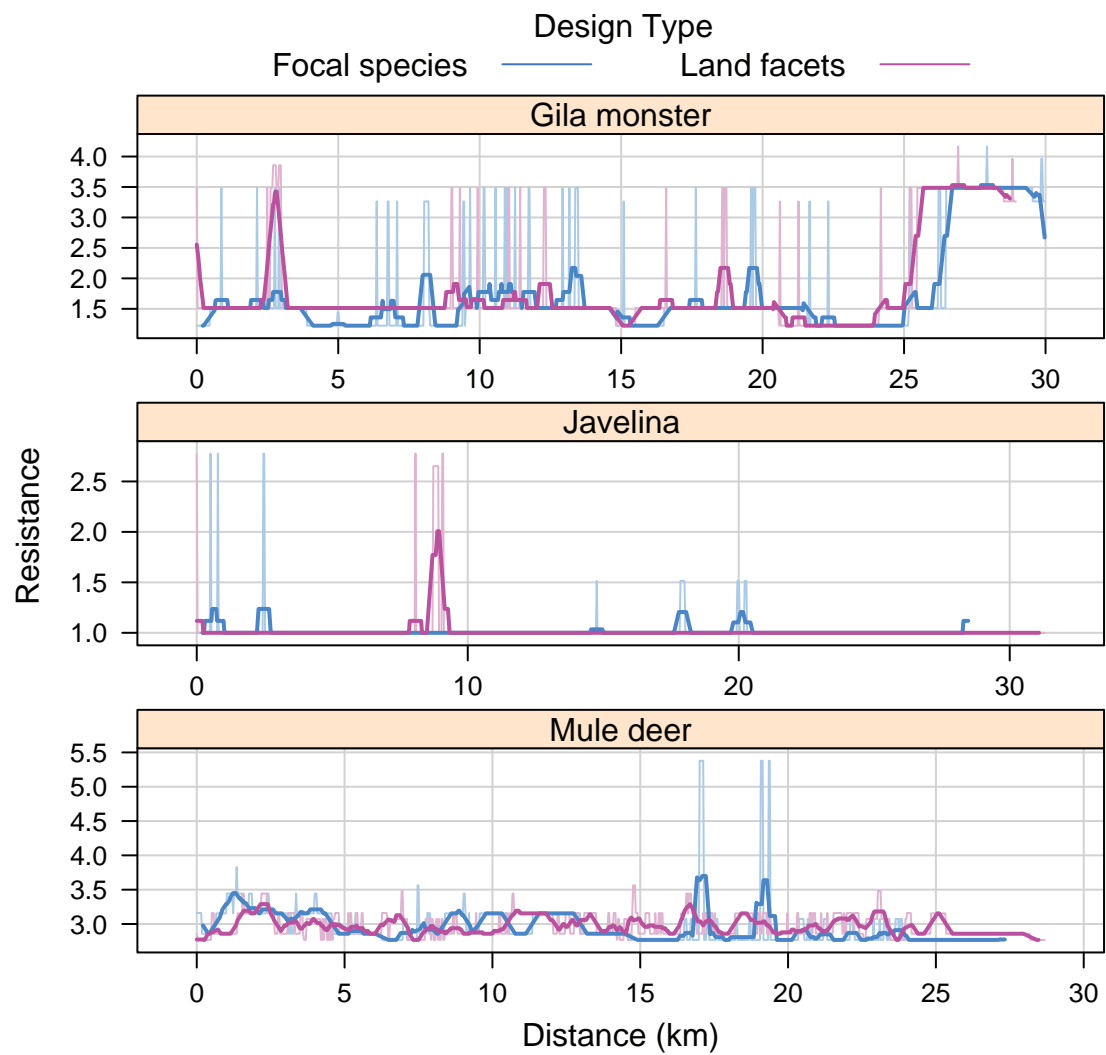


Figure 3.6. (continued)

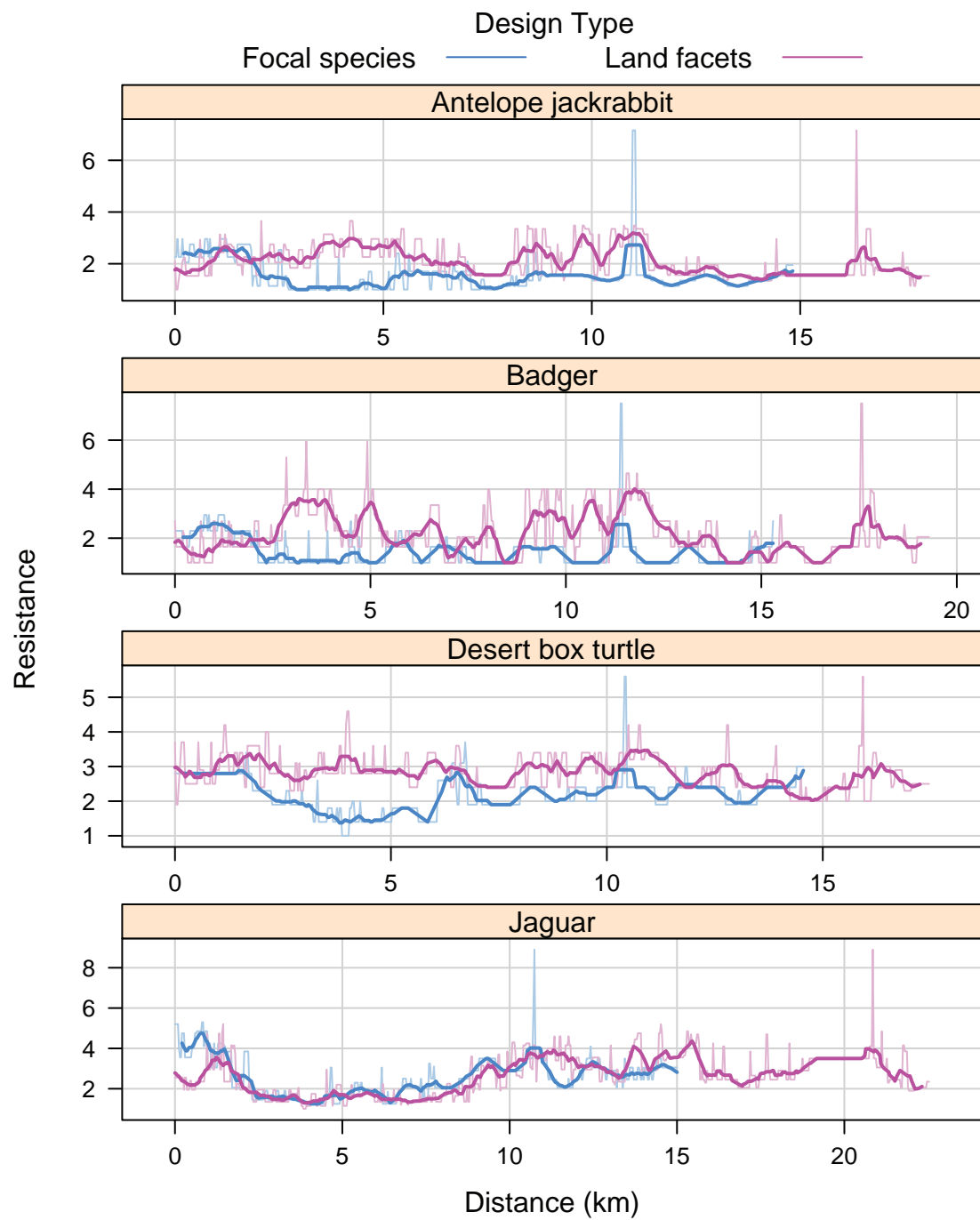


Figure 3.7. Resistance profiles for species with locally widespread habitat in the Santa Rita-Tumacacori planning area.

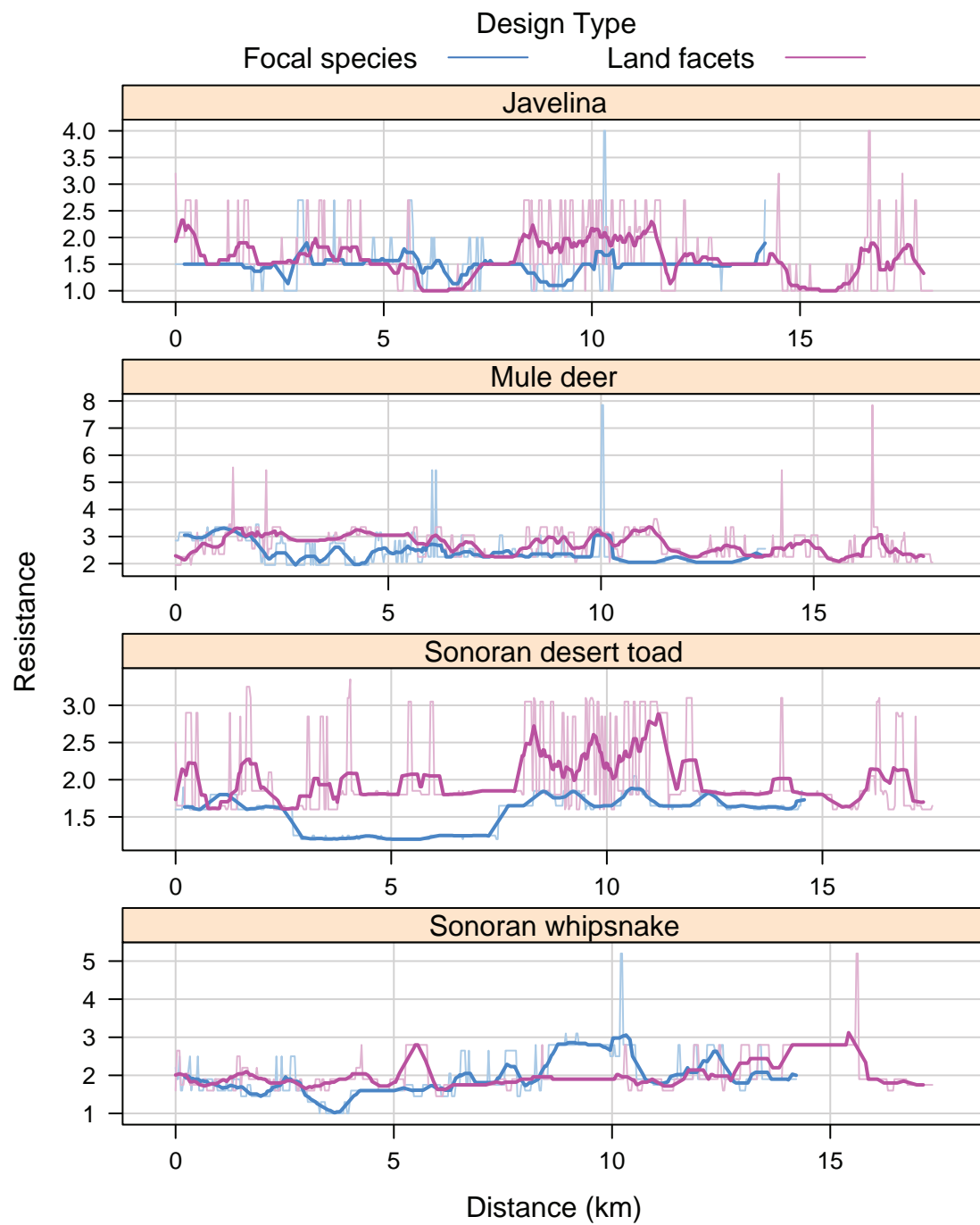


Figure 3.7. (continued)

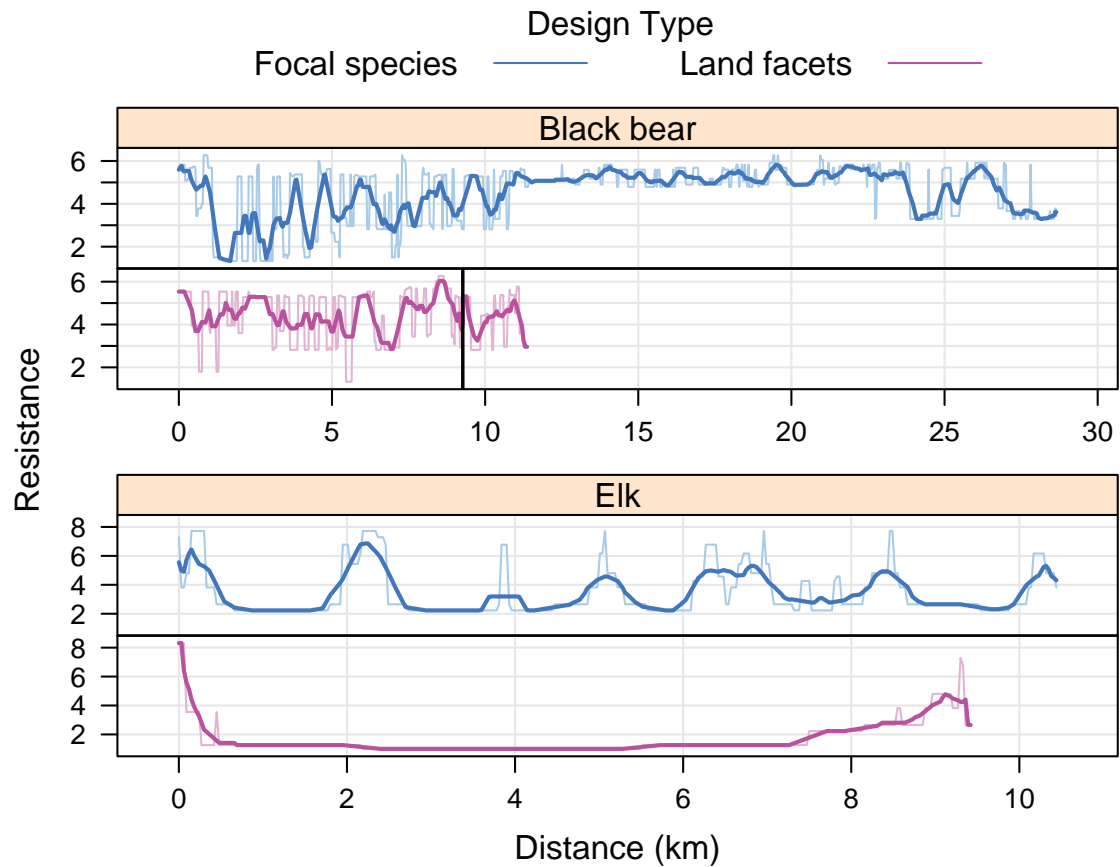


Figure 3.8. Resistance profiles corresponding to the gaps between breeding patches for black bear and elk in the Black Hills-Munds Mountain planning area. Each vertical line indicates a breeding patch between two gaps; line width does not indicate the width of the breeding patch.

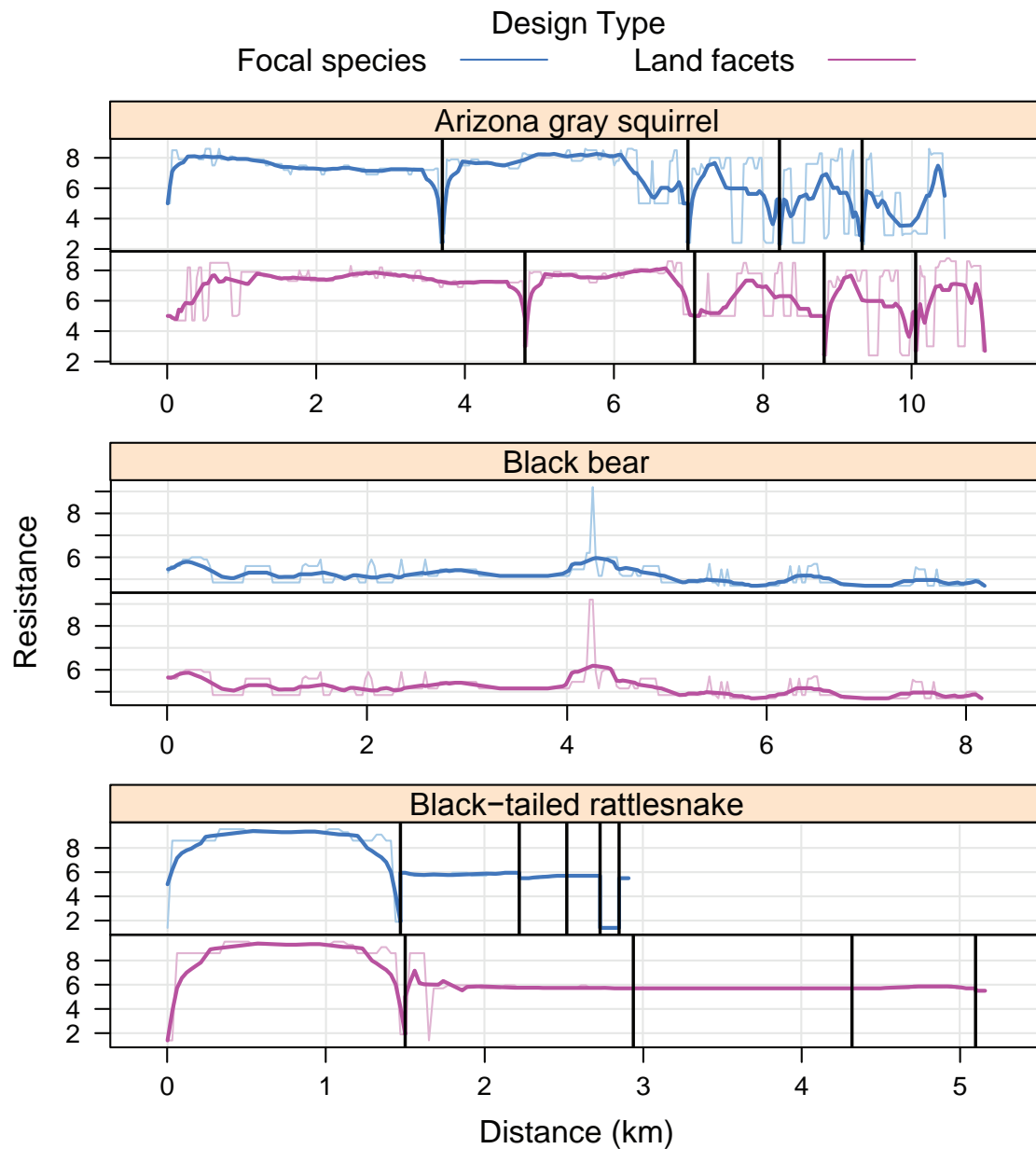


Figure 3.9. Resistance profiles corresponding to the gaps between breeding patches for focal species in the Santa Rita-Tumacacori planning area. Each vertical line indicates a breeding patch between two gaps; line width does not indicate the width of the breeding patch.

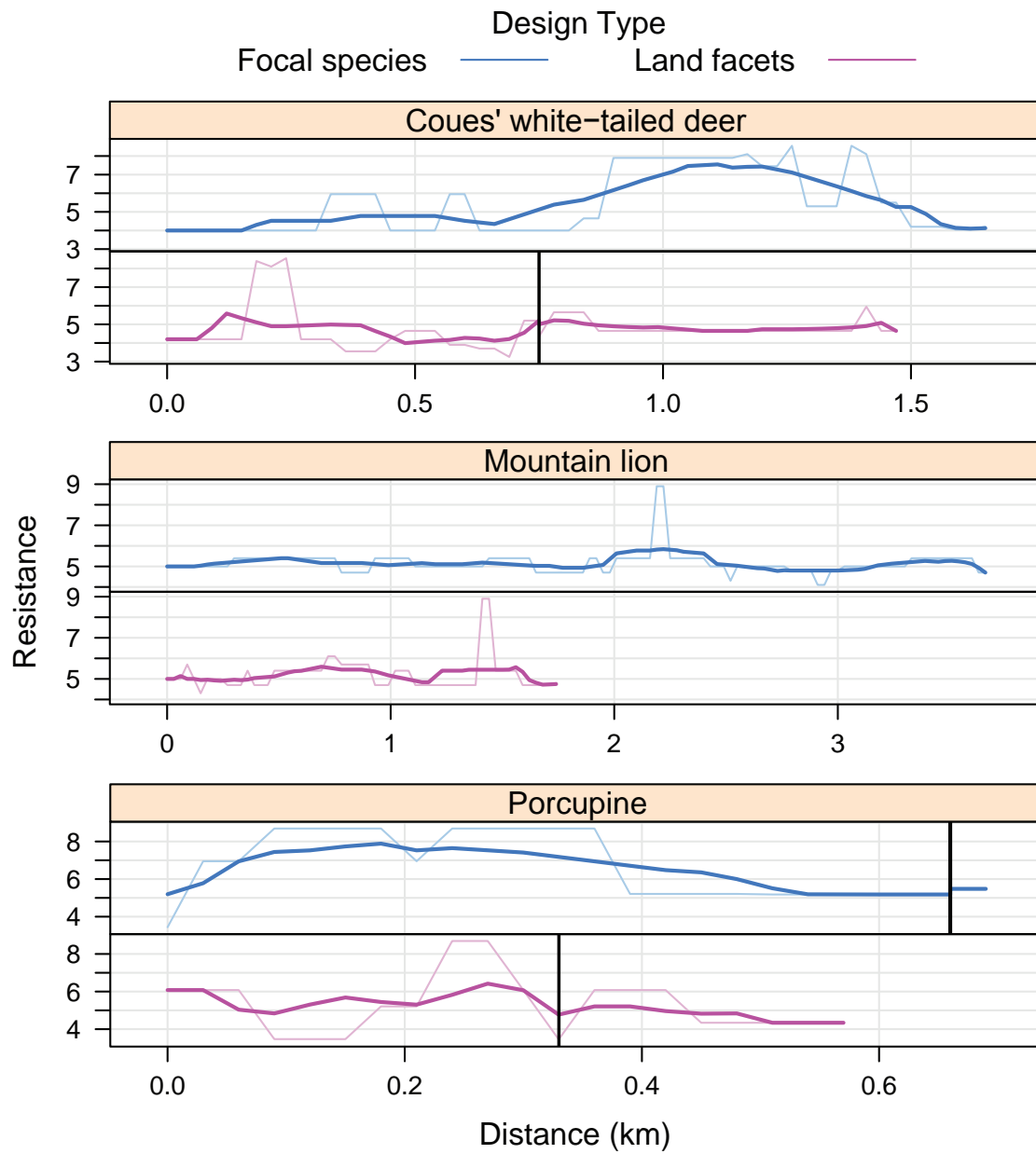


Figure 3.9. (continued)



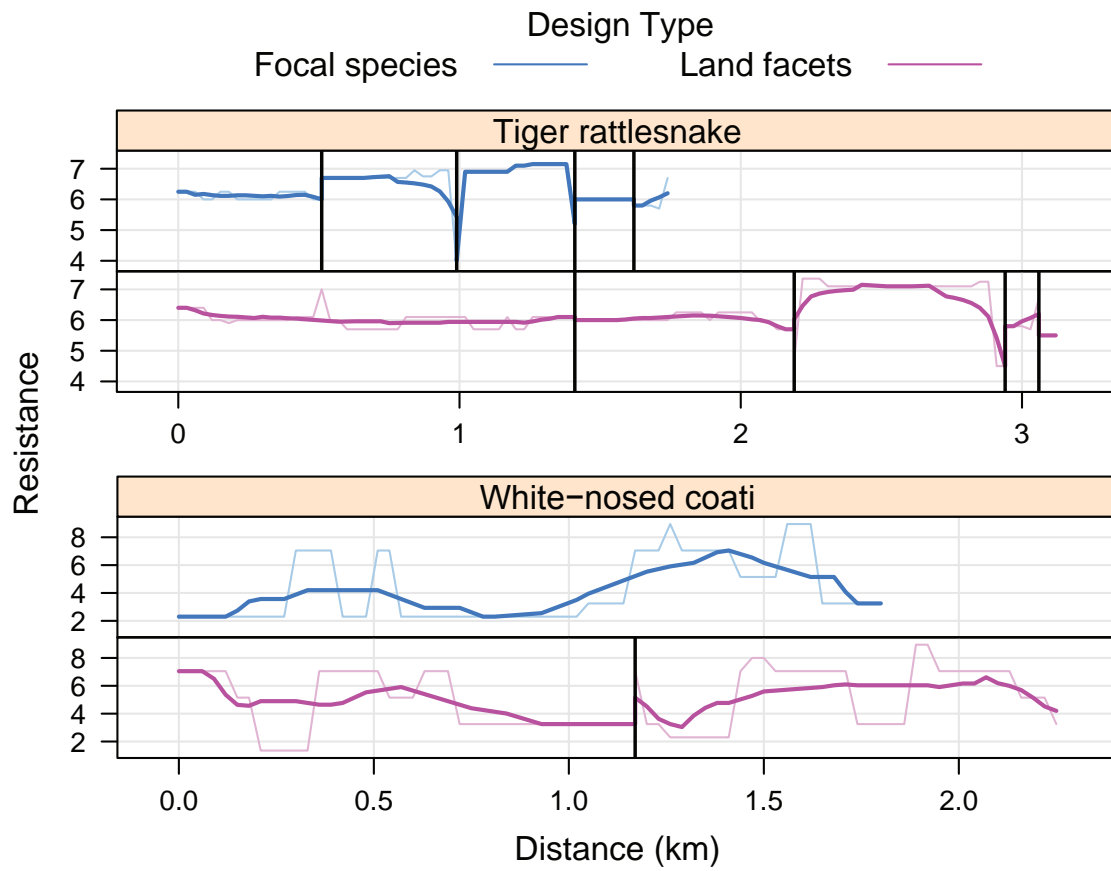


Figure 3.9. (continued)

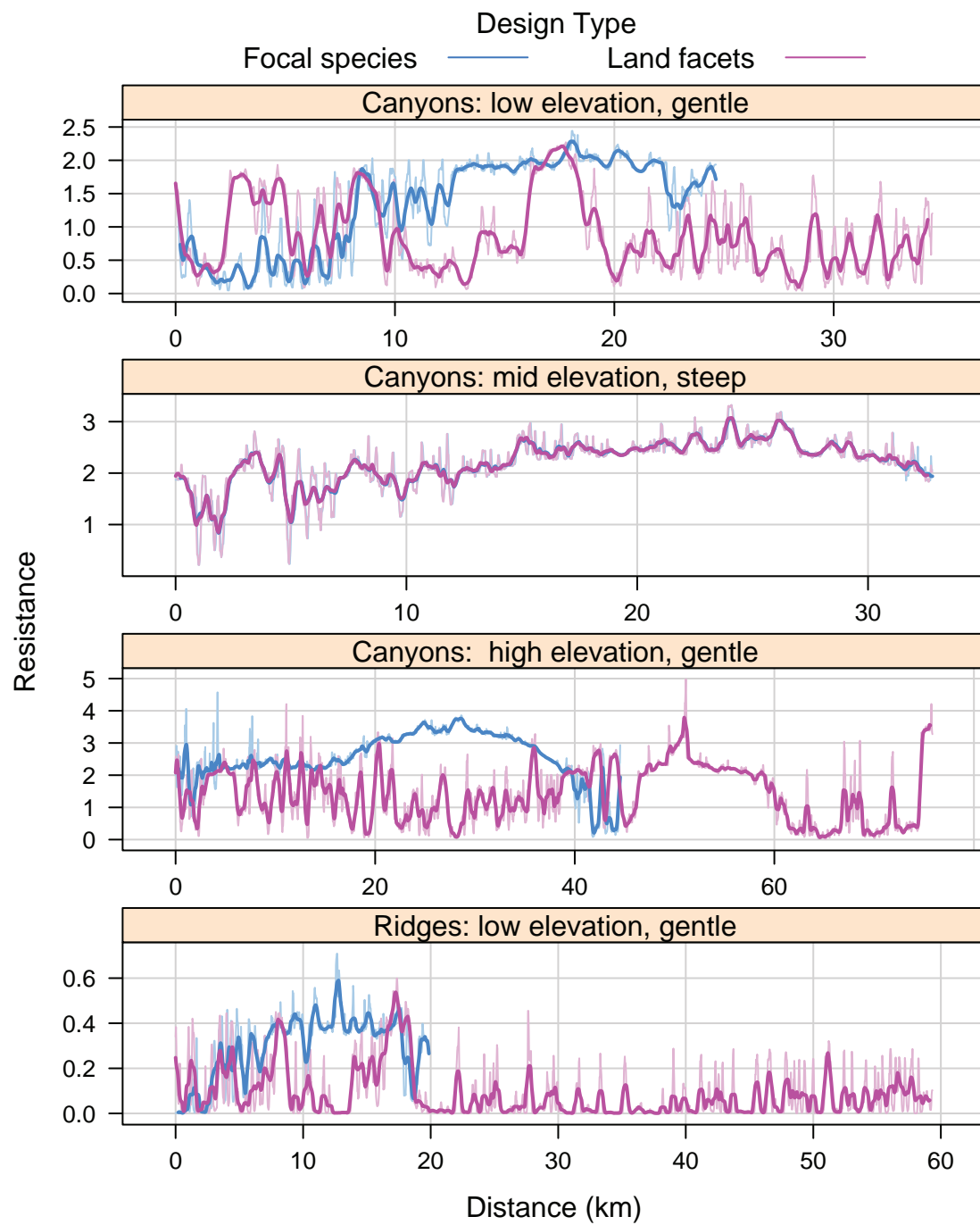


Figure 3.10. Resistance profiles for land facets in the Black Hills-Munds Mountain planning area.

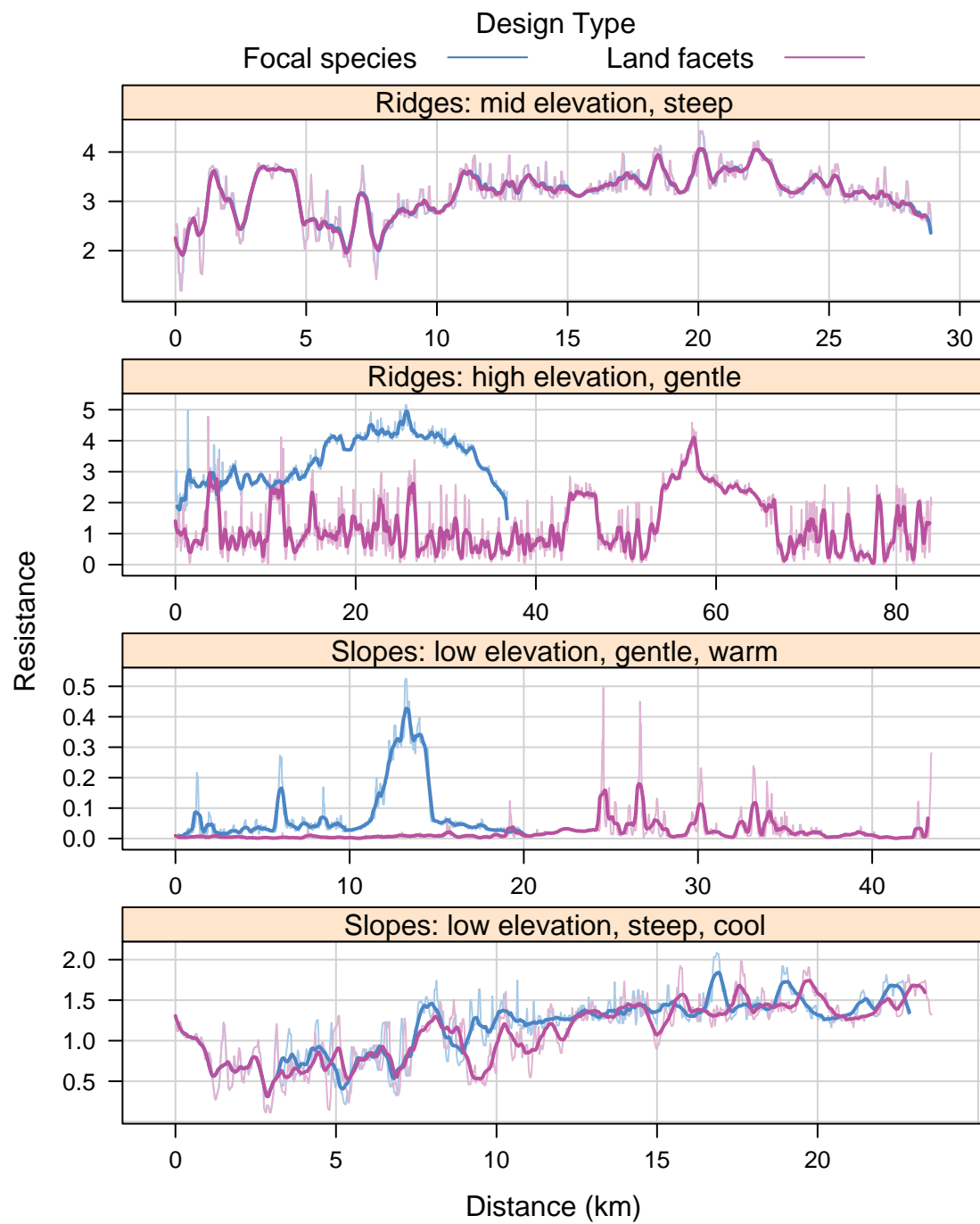


Figure 3.10. (continued)

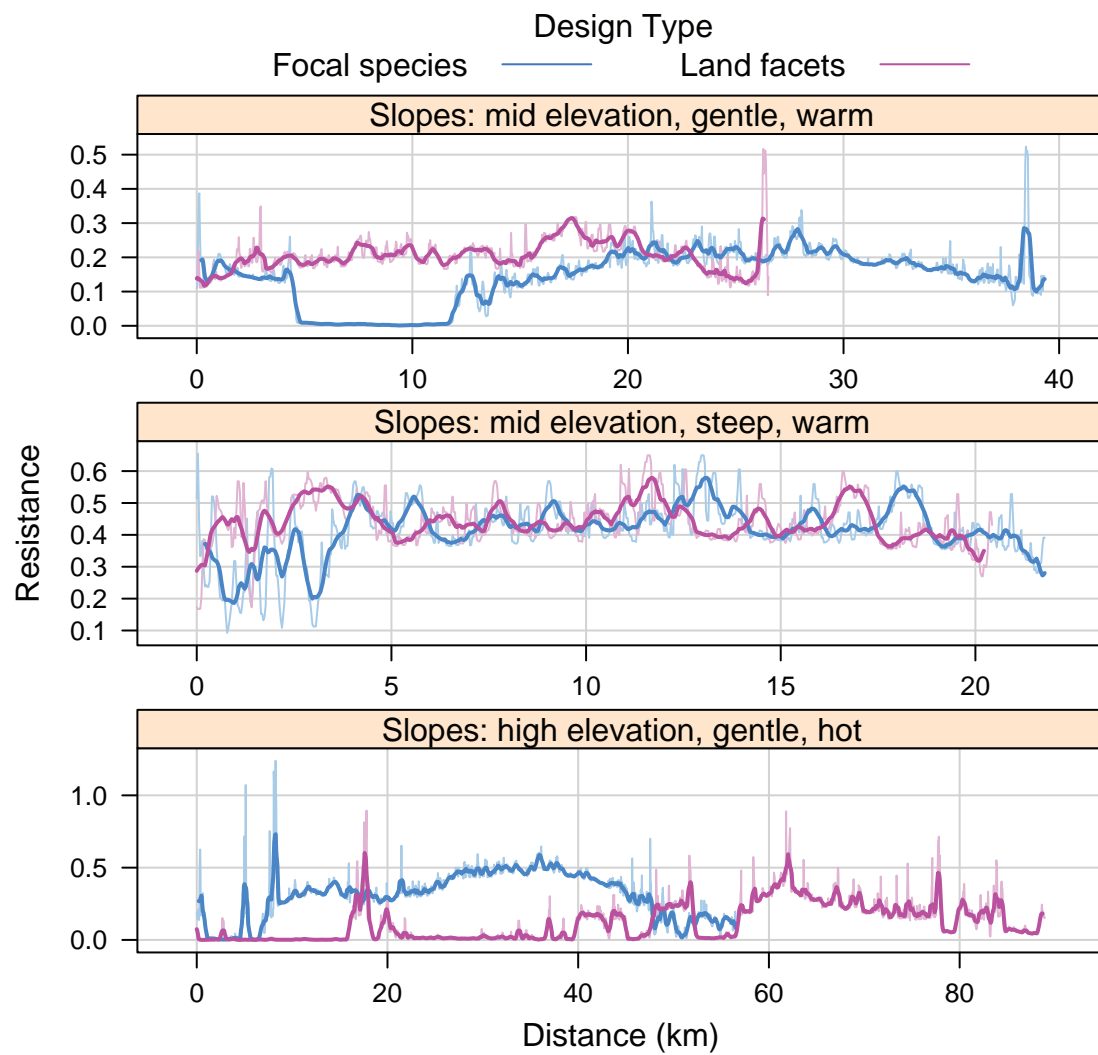


Figure 3.10. (continued)

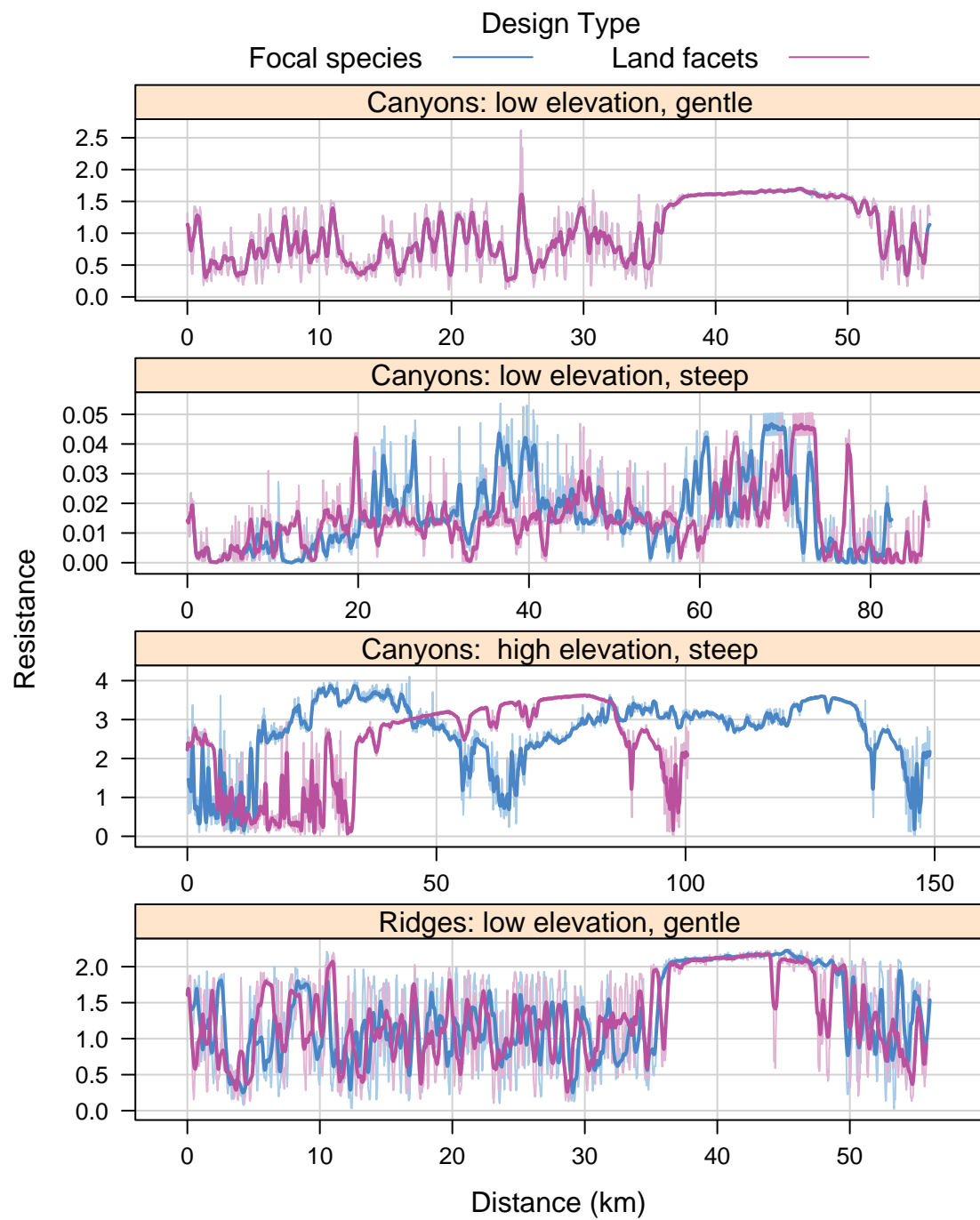


Figure 3.11. Resistance profiles for land facets in the Wickenburg-Hassayampa planning area.

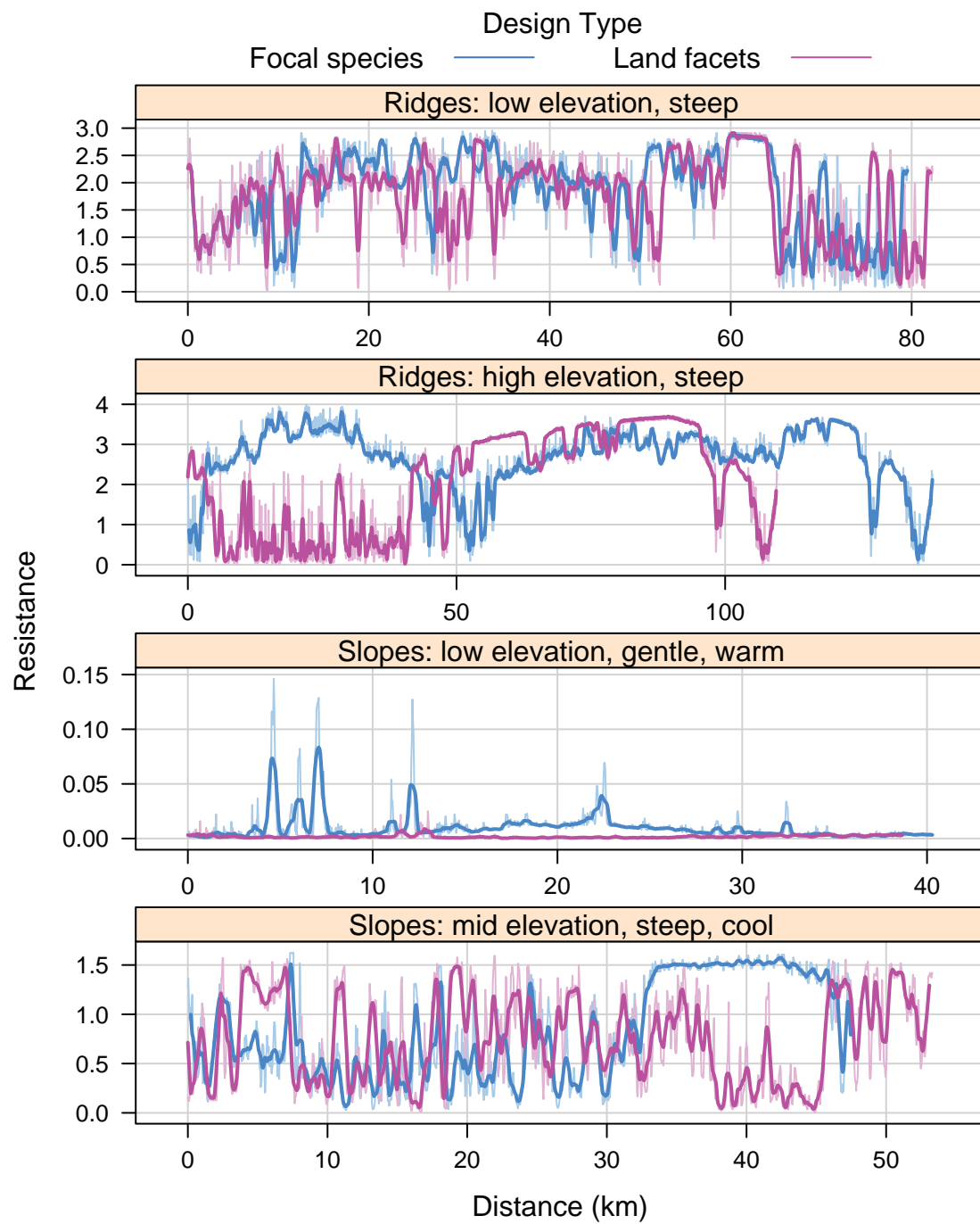


Figure 3.11. (continued)

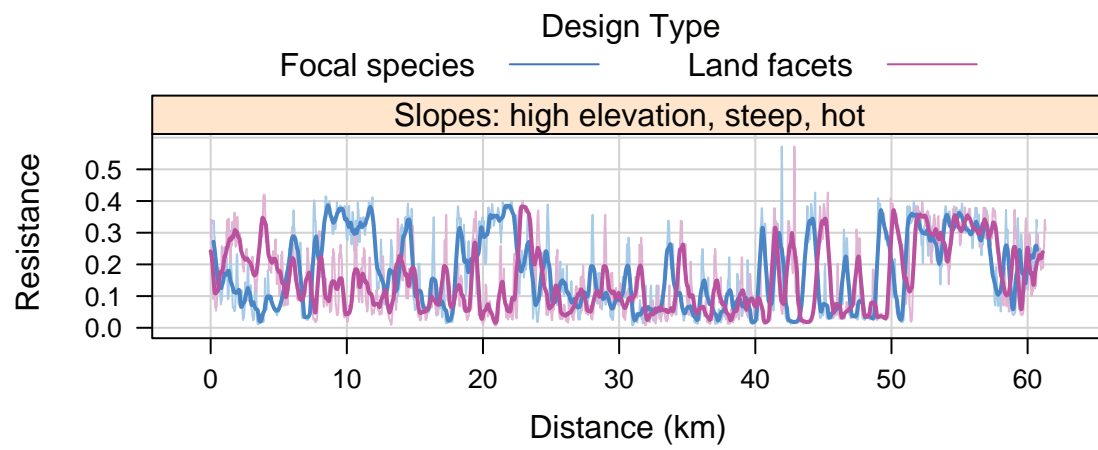


Figure 3.11. (continued)

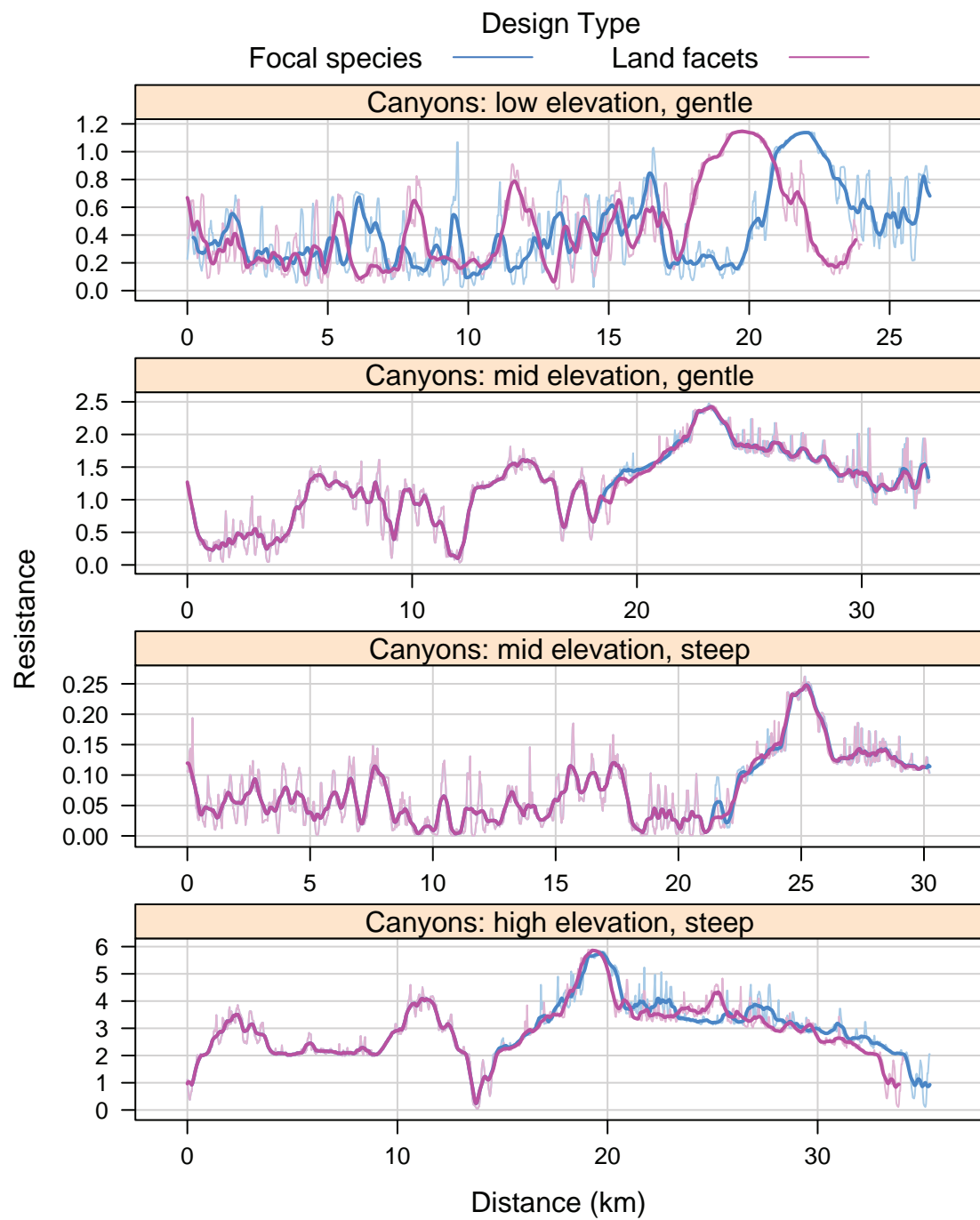


Figure 3.12. Resistance profiles for land facets in the Santa Rita-Tumacacori planning area.



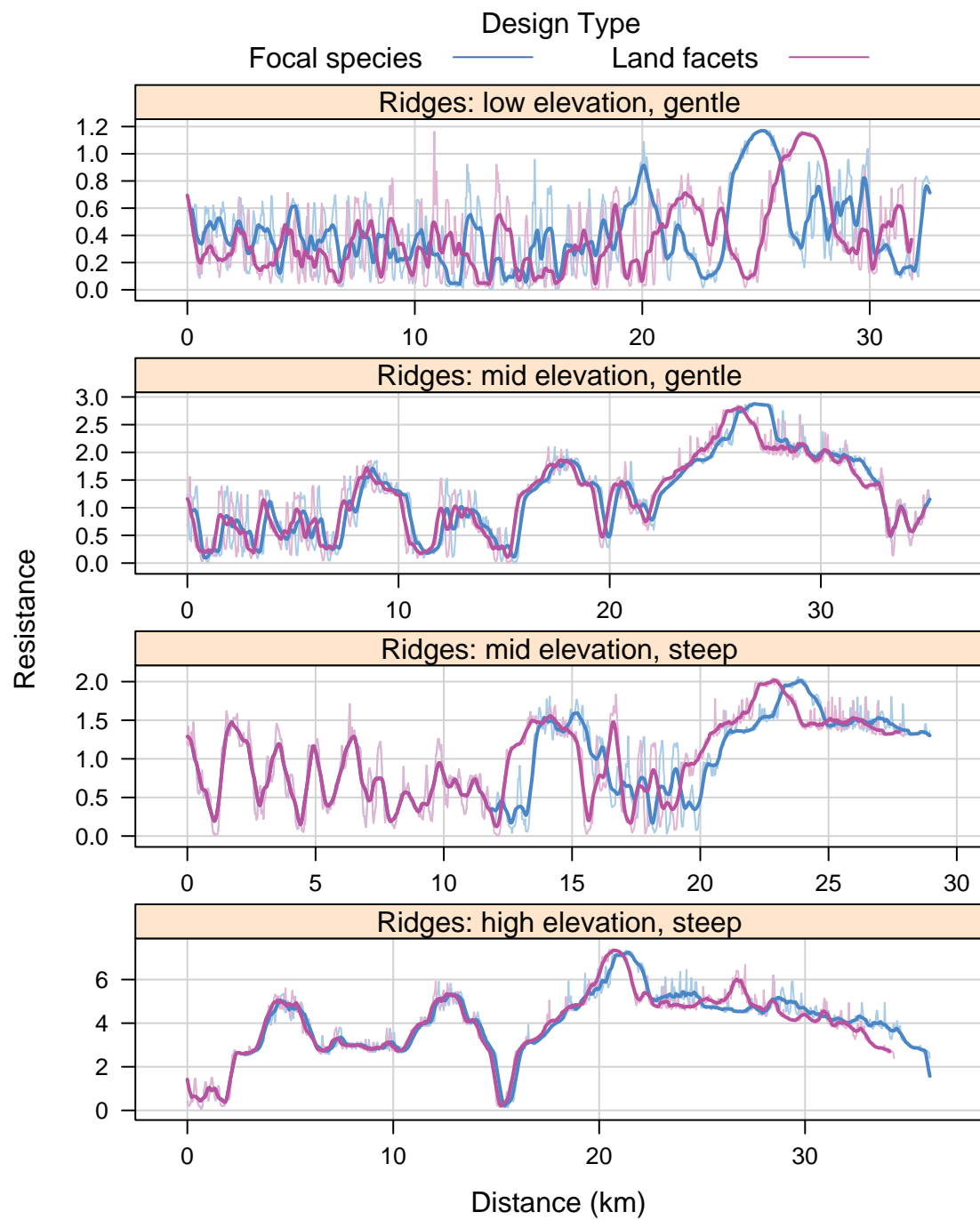


Figure 3.12. (continued)

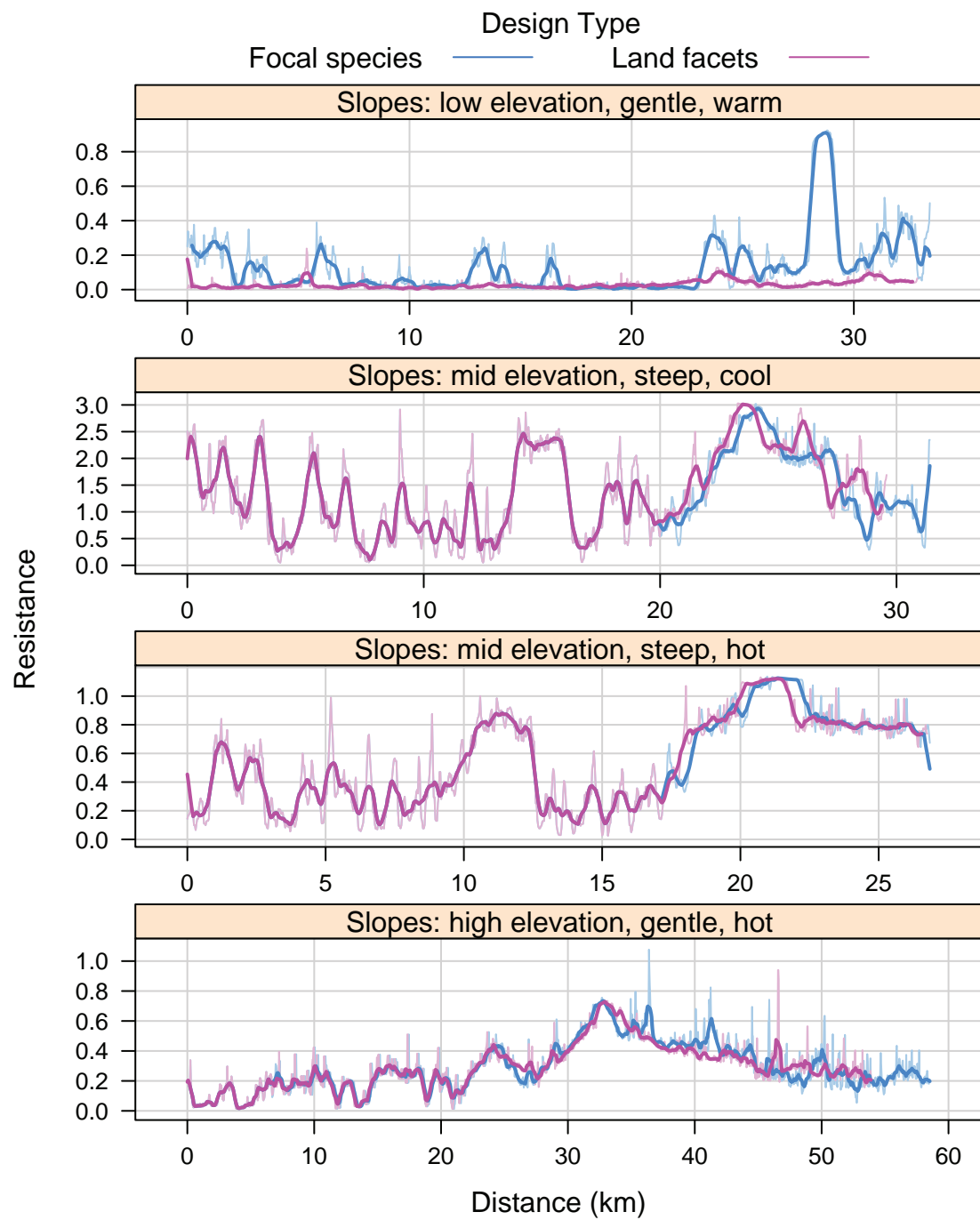


Figure 3.12. (continued)

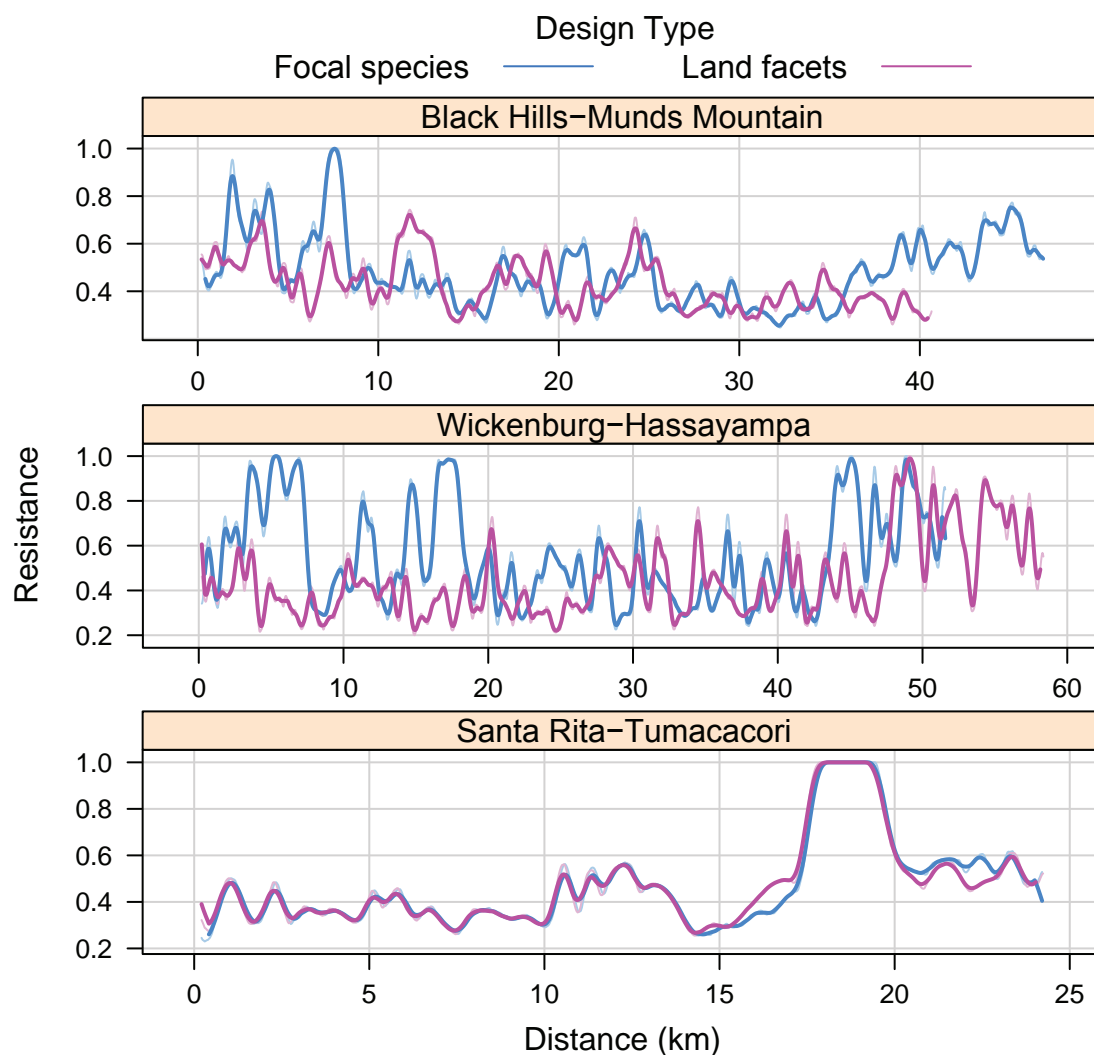


Figure 3.13. Resistance profiles for high diversity of land facets in the Black Hills-Munds Mountain, Wickenburg-Hassayampa, and Santa Rita-Tumacacori planning areas. The values in the profiles are the compliment of Shannon's evenness, where 0 is the lowest possible resistance and 1 is the maximum deviation from the optimal Shannon's index value.

# Chapter 4

## Conclusions

Despite widespread prescriptions for conservation corridors and linkages to aid species' range shifts in response to climate change, few designs explicitly incorporate climate change considerations. Our approach exploits the fact that topography and soils are major drivers of biodiversity, and thus relies only on factors that are stable with respect to climate. It avoids the enormous complexity and the high level of uncertainty associated with incorporating models of climate and habitat suitability into corridor design. The land facets approach to linkage design is adaptable to different landscapes and it is not afflicted by the patchiness and bias common in species occurrence data. In fact, our approach can be applied anywhere because digital elevation models are available for all continents.

In three Arizona landscapes, linkages designed for land facets served 25 of 28 focal species as well as or better than the focal species designs. For the 16 focal species in these landscapes that had locally widespread breeding habitat, similar performance under both types of designs was more a virtue of the distribution of these species' habitat than the placement of the linkage design. Of the 12 remaining species with more sparsely distributed habitat, four were served equally well by both designs and five species were served better by the land facets linkage designs. The three species that had the most narrowly distributed habitat were served better by the focal species designs.

Compared to land facets designs, focal species linkages provided a similar degree of connectivity for only half of the land facets in two of the landscapes. In these landscapes, some land facets were probably served well because a focal species was associated with that land facet type. In the third landscape, the focal species design performed as well as the land facets design for 12 of 13 land facets. The comparable performance of the two types of linkage designs in this planning area is likely due

to the spatial configuration of the landscape's topography. Notably, a focal species design never provided better continuity for a land facet than a land facets design.

These results lend support to the underlying concept of the land facets approach to linkage design, which is that diverse physical environments support diverse biota. However, it is clear that linkages designed for land facets will fail to serve some species, especially those with limited habitat in the planning area. Similarly, linkages designed for focal species tend to capture diverse topography. Nonetheless, focal species designs did not serve many land facets as well as the linkage designs based on land facets. Therefore, we support the recommendation of Beier and Brost (2010) that the land facets approach to linkage design should complement, rather than replace, focal species approaches. But simply combining the two types of linkage designs would produce a needlessly large linkage design that would be expensive to conserve. Indeed, in two of our areas, the land facets design served all land facets and focal species as well as or better than the focal species design. In the third landscape, a conservation planner could efficiently provide connectivity for all species and land facets by expanding the land facets design by merely 3%.

An important next step in the land facets approach to linkage design is to conduct an uncertainty analysis. A strategy similar to Beier et al. (2009) could be adopted to determine how sensitive modeled corridors are to the parameters and decision rules used in this approach, as well as to the error present in digital elevation models. Further evaluation is also necessary before prescribing the use of land facets as a standalone or primary approach to linkage design. If further evaluations reinforce our results, such findings would suggest that the land facets approach could be used in lieu of a focal species approach in areas where species models cannot be developed. This additional information could also be used to develop a general strategy for using land facets in linkage design.

# Literature Cited

- Adriaensen, F., J. P. Chardon, G. De Blust, E. Swinnen, S. Villalba, H. Gulinck, and E. Matthysen. 2003. The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning* **64**:233–247.
- Araujo, M. B., C. J. Humphries, P. J. Densham, R. Lampinen, W. J. M. Hagemeyer, A. J. Mitchell-Jones, and J. P. Gasc. 2001. Would environmental diversity be a good surrogate for species diversity? *Ecography* **24**:103–110.
- Beale, C. M., J. J. Lennon, and A. Gimona. 2008. Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences* **105**:14908–14912.
- Beier, P. and B. Brost. 2010. Use of land facets to plan for climate change: Conserving the arenas, not the actors. *Conservation Biology* **24**:10.1111/j.1523-1739.2009.01422.x.
- Beier, P., K. Penrod, C. Luke, W. Spencer, and C. Cabanero. 2006. South Coast missing linkages: Restoring connectivity to wildlands in the largest metropolitan area in the USA. Pages 555–586 in K. R. Crooks and M. A. Sanjayan, editors. *Connectivity conservation*. Cambridge University Press, Cambridge, United Kingdom.
- Beier, P., D. R. Majka, and T. Bayless. 2007. Linkage designs for Arizona's missing linkages. Arizona Game and Fish Department, Phoenix. Available from: [www.corridordesign.org/arizona](http://www.corridordesign.org/arizona) (accessed September 2009) .
- Beier, P., D. R. Majka, and W. D. Spencer. 2008. Forks in the road: Choices in procedures for designing wildland linkages. *Conservation Biology* **22**:836–851.
- Beier, P., D. R. Majka, and S. L. Newell. 2009. Uncertainty analysis of least-cost modeling for designing wildlife linkages. *Ecological Applications* **19**:2067–2077.
- Bezdek, C. J. 1981. *Pattern recognition with fuzzy objective function algorithms*. Plenum Press, New York.
- Bolstad, P. V., W. Swank, and J. Vose. 1998. Predicting southern Appalachian overstory vegetation with digital terrain data. *Landscape Ecology* **13**:271–283.

- Burnett, M. R., P. V. August, J. H. Brown, Jr., and K. T. Killingbeck. 1998. The influence of geomorphological heterogeneity on biodiversity: I. A patch-scale perspective. *Conservation Biology* **12**:363–370.
- Burrough, P. A., P. F. M. van Gaans, and R. A. MacMillan. 2000. High-resolution landform classification using fuzzy  $k$ -means. *Fuzzy Sets and Systems* **113**:37–52.
- Campello, R. J. G. B. and E. R. Hruschka. 2006. A fuzzy extension of the silhouette width criterion for cluster analysis. *Fuzzy Sets and Systems* **157**:2858–2875.
- Celikyilmaz, A. M. and I. B. Turksen. 2008. Validation criteria for enhanced fuzzy clustering. *Pattern Recognition Letters* **29**:97–108.
- Clark, J. D., J. E. Dunn, and K. G. Smith. 1993. A multivariate model of female black bear habitat use for a geographic information system. *The Journal of Wildlife Management* **57**:519–526.
- Cowling, R. M., R. L. Pressey, A. T. Lombard, P. G. Desmet, and A. G. Ellis. 1999. From representation to persistence: Requirements for a sustainable system of conservation areas in the species-rich Mediterranean-climate desert of southern Africa. *Diversity and Distributions* **5**:51–71.
- Cowling, R. M., R. L. Pressey, M. Rouget, and A. T. Lombard. 2003. A conservation plan for a global biodiversity hotspot—the Cape Floristic Region, South Africa. *Biological Conservation* **112**:191–216.
- Crawley, M. J. 2007. *The R book*. John Wiley and Sons.
- Crooks, K. R. and M. Sanjayan. 2006. *Connectivity conservation*. Cambridge University Press, Cambridge, United Kingdom.
- Davis, F. W. and S. Goetz. 1990. Modeling vegetation pattern using digital terrain data. *Landscape Ecology* **4**:69–80.
- Dimitriadou, E., K. Hornik, F. Leisch, D. Meyer, and A. Weingessel. 2009. *e1071: Misc Functions of the Department of Statistics (e1071)*, TU Wien. R package version 1.5-19.
- Duong, T. 2007. *ks*: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *Journal of Statistical Software* **21**(7).
- Duong, T. 2009. *ks*: Kernel Smoothing. R Package version 1.6.5.
- Fairbanks, D. H. K., B. Reyers, and A. S. van Jaarsveld. 2001. Species and environment representation: Selecting reserves for the retention of avian diversity in KwaZulu-Natal, South Africa. *Biological Conservation* **98**:365–379.
- Faith, D. P. and P. A. Walker. 1996. Environmental diversity: On the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity and Conservation* **5**:399–415.

- Farber, O. and R. Kadmon. 2003. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling* **160**:115–130.
- Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography* **19**:474–499.
- Franklin, J., P. McCullough, and C. Gray. 2000. Terrain variables used for predictive mapping of vegetation communities in southern California. Pages 331–353 in J. P. Wilson and J. C. Gallant, editors. *Terrain analysis: Principles and applications*. John Wiley and Sons, New York.
- Fu, P. and P. M. Rich. 2002. A geometric solar radiation model with applications in agriculture and forestry. *Computers and Electronics in Agriculture* **37**:25–35.
- Gallant, J. C. and J. P. Wilson. 2000. Primary topographic attributes. Pages 29–50 in J. P. Wilson and J. C. Gallant, editors. *Terrain analysis: Principles and applications*. John Wiley and Sons, New York.
- Gotelli, N. J. and A. M. Ellison. 2004. *A primer of ecological statistics*. Sinauer Associates, Sunderland, MA.
- Grabherr, G., M. Gottfried, and H. Pauli. 1994. Climate effects on mountain plants. *Nature* **369**:448.
- Graham, R. W. and E. C. Grimm. 1990. Effects of global climate change on the patterns of terrestrial biological communities. *Trends in Ecology and Evolution* **5**:289–292.
- Guisan, A., S. B. Weiss, and A. D. Weiss. 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* **143**:107–122.
- Halpin, P. N. 1997. Global climate change and natural-area protection: Management responses and research directions. *Ecological Applications* **7**:828–843.
- Hannah, L., G. F. Midgley, T. Lovejoy, W. J. Bond, M. Bush, J. C. Lovett, D. Scott, and F. I. Woodward. 2002. Conservation of biodiversity in a changing climate. *Conservation Biology* **16**:264–268.
- Heller, N. E. and E. S. Zavaleta. 2009. Biodiversity management in the face of climate change: A review of 22 years of recommendations. *Biological Conservation* **142**:14–32.
- Hersteinsson, P. and D. W. Macdonald. 1992. Interspecific competition and the geographical distribution of red and arctic foxes *Vulpes vulpes* and *Alopex lagopus*. *Oikos* **64**:505–515.



- Hoersch, B., G. Braun, and U. Schmidt. 2002. Relation between landform and vegetation in alpine regions of Wallis, Switzerland: A multiscale remote sensing and GIS approach. *Computers, Environment and Urban Systems* **26**:113–139.
- Hruschka, E. R., R. J. G. B. Campello, and L. N. de Castro. 2006. Evolving clusters in gene-expression data. *Information Sciences* **176**:1898–1927.
- Hunter, Jr., M. L., G. L. Jacobson, Jr., and T. Webb, III. 1988. Paleoecology and the coarse-filter approach to maintaining biological diversity. *Conservation Biology* **2**:375–385.
- Huntley, B. 2005. North temperate responses. Pages 109–124 in T. E. Lovejoy and L. Hannah, editors. *Climate change and biodiversity*. Yale University Press, New Haven, CT.
- Intergovernmental Panel on Climate Change. 2001. *Climate change 2001: The scientific basis*. IPCC, Geneva.
- Jenness Enterprises. 2010. Mahalanobis Distances extension for ArcGIS. Available from: [www.jennessent.com](http://www.jennessent.com) (accessed january 2010).
- Kirkpatrick, J. B. and M. J. Brown. 1994. A comparison of direct and environmental domain approaches to planning reservation of forest higher plant communities and species in Tasmania. *Conservation Biology* **8**:217–224.
- Klein, C., K. Wilson, M. Watts, J. Stein, S. Berry, J. Carwardine, M. S. Smith, B. Mackey, and H. Possingham. 2009. Incorporating ecological and evolutionary processes into continental-scale conservation planning. *Ecological Applications* **19**:206–217.
- Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. 2005. *Applied linear statistical models*. McGraw-Hill/Irwin, New York, NY.
- Legendre, P. 1993. Spatial autocorrelation: Trouble or new paradigm? *Ecology* **74**:1659–1673.
- Lombard, A. T., R. M. Cowling, R. L. Pressey, and A. G. Rebelo. 2003. Effectiveness of land classes as surrogates for species in conservation planning for the Cape Floristic Region. *Biological Conservation* **112**:45–62.
- Lookingbill, T. and D. Urban. 2004. An empirical approach towards improved spatial estimates of soil moisture for vegetation analysis. *Landscape Ecology* **19**:417–433.
- Mackey, B. G., H. A. Nix, M. F. Hutchinson, J. P. MacMahon, and P. M. Fleming. 1988. Assessing representativeness of places for conservation reservation and heritage listing. *Environmental Management* **12**:501–514.
- Magurran, A. E. 1988. *Ecological diversity and its measurement*. Princeton University Press, Princeton, NJ.

- Majka, D. J., J. Jenness, and P. Beier. 2007. CorridorDesigner: ArcGIS tools for designing and evaluating corridors. Available from: [www.corridordesign.org](http://www.corridordesign.org) (accessed june 2009).
- Maulik, U. and S. Bandyopadhyay. 2002. Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**:1650–1654.
- Mawdsley, J. R., R. O'Malley, and D. S. Ojima. 2009. A review of climate-change adaptation strategies for wildlife management and biodiversity conservation. *Conservation Biology* **23**:1080–1089.
- McCoy, E. J., D. G. Hewitt, and F. C. Bryant. 2005. Dispersal by yearling male white-tailed deer and implications for management. *Journal of Wildlife Management* **69**:366–376.
- Millar, C. I., N. L. Stephenson, and S. L. Stephens. 2007. Climate change and forests of the future: Managing in the face of uncertainty. *Ecological Applications* **17**:2145–2151.
- Moore, I. D., R. B. Grayson, and A. R. Ladson. 1991. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes* **5**:3–30.
- Moritz, C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology* **51**:238–254.
- Nichols, W. F., K. T. Killingbeck, and P. V. August. 1998. The influence of geomorphological heterogeneity on biodiversity: II. A landscape perspective. *Conservation Biology* **12**:371–379.
- Noss, R. F. 1987. From plant communities to landscapes in conservation inventories: A look at The Nature Conservancy (USA). *Biological Conservation* **41**:11–37.
- Noss, R. F. 2001. Beyond Kyoto: Forest management in a time of rapid climate change. *Conservation Biology* **15**:578–590.
- Noss, R. F. and K. M. Daly. 2006. Incorporating connectivity into broad-scale conservation planning. Pages 587–619 in K. R. Crooks and M. A. Sanjayan, editors. *Connectivity conservation*. Cambridge University Press, Cambridge, United Kingdom.
- Pakhira, M. K., S. Bandyopadhyay, and U. Maulik. 2004. Validity index for crisp and fuzzy clusters. *Pattern Recognition* **37**:487–501.
- Parker, A. J. 1995. Comparative gradient structure and forest cover types in Lassen Volcanic and Yosemite National Parks, California. *Bulletin of the Torrey Botanical Club* **122**:58–68.

- Parmesan, C. 1996. Climate and species' range. *Nature* **382**:765–766.
- Parmesan, C., N. Ryrholm, C. Stefanescu, J. K. Hill, C. D. Thomas, H. Descimon, B. Huntley, L. Kaila, J. Kullberg, T. Tammaru, W. J. Tennent, J. A. Thomas, and M. Warren. 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* **399**:579–583.
- Pauli, H., M. Gottfried, and G. Grabherr. 1996. Effects of climate change on mountain ecosystems—upward shifting of alpine plants. *World Resource Review* **8**:382–390.
- Pfeffer, K., E. J. Pebesma, and P. A. Burrough. 2003. Mapping alpine vegetation using vegetation observations and topographic attributes. *Landscape Ecology* **18**:759–776.
- Phillips, S. J., P. Williams, G. Midgley, and A. Archer. 2008. Optimizing dispersal corridors for the Cape Proteaceae using network flow. *Ecological Applications* **18**:1200–1211.
- Pinder, III, J. E., G. C. Kroh, J. D. White, and A. M. Basham May. 1997. The relationships between vegetation type and topography in Lassen Volcanic National Park. *Plant Ecology* **131**:17–29.
- Pounds, J. A., M. P. L. Fogden, and J. H. Campbell. 1999. Biological response to climate change on a tropical mountain. *Nature* **398**:611–615.
- Pressey, R. L., M. Cabeza, M. E. Watts, R. M. Cowling, and K. A. Wilson. 2007. Conservation planning in a changing world. *Trends in Ecology and Evolution* **22**:583–592.
- Raper, S. C. B. and F. Giorgi. 2005. Climate change projections and models. Pages 199–210 in T. E. Lovejoy and L. Hannah, editors. *Climate change and biodiversity*. Yale University Press, New Haven, CT.
- Reyers, B., K. J. Wessels, and A. S. van Jaarsveld. 2002. An assessment of biodiversity surrogacy options in the Limpopo Province of South Africa. *African Zoology* **37**:185–195.
- Rouget, M., R. M. Cowling, A. T. Lombard, A. T. Knight, and G. I. H. Kerley. 2006. Designing large-scale conservation corridors for pattern and process. *Conservation Biology* **20**:549–561.
- Silverman, B. W. 1986. *Density estimation for statistics and data analysis*. Chapman and Hall, London.
- Singleton, P. H., W. Gaines, and J. F. Lehmkuhl. 2002. Landscape permeability for large carnivores in Washington: A geographic information system weighted-distance and least-cost corridor assessment. USDA Forest Service Research Paper. PNW-RP 549. Pacific Northwest Field Station, OR.

- Skelly, D. K., L. N. Joseph, H. P. Possingham, L. K. Freidenburg, T. J. Farrugia, M. T. Kinnison, and A. P. Hendry. 2007. Evolutionary responses to climate change. *Conservation Biology* **21**:1353–1355.
- Sweanor, L. L., K. A. Logan, and M. G. Hornocker. 2000. Cougar dispersal patterns, metapopulation dynamics, and conservation. *Conservation Biology* **14**:798–808.
- Thomas, C. D. and J. J. Lennon. 1999. Birds extend their ranges northwards. *Nature* **399**:213.
- Urban, D. L., C. Miller, P. N. Halpin, and N. L. Stephenson. 2000. Forest gradient response in Sierran landscapes: The physical template. *Landscape Ecology* **15**:603–620.
- Valverde, P. L., J. A. Zavala-Hurtado, C. Montana, and E. Ezcurra. 1996. Numerical analyses of vegetation based on environmental relationships in the southern Chihuahuan Desert. *The Southwestern Naturalist* **41**:424–433.
- Walker, R. and L. Craighead. 1997. Analyzing wildlife movement corridors in Montana using GIS. In *Proceedings of the ESRI European User Conference, Copenhagen*, pp. 1–18.
- Webb, III, T. 1987. The appearance and disappearance of major vegetational assemblages: Long-term vegetational dynamics in eastern North America. *Vegetatio* **69**:177–187.
- Wessels, K. J., S. Freitag, and A. S. van Jaarsveld. 1999. The use of land facets as biodiversity surrogates during reserve selection at a local scale. *Biological Conservation* **89**:21–38.
- Williams, P., L. Hannah, S. Andelman, G. Midgley, M. Araujo, G. Hughes, L. Manne, E. Martinez-Meyer, and R. Pearson. 2005. Planning for climate change: Identifying minimum-dispersal corridors for the Cape Proteaceae. *Conservation Biology* **19**:1063–1074.
- Xie, X. L. and G. Beni. 1991. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**:841–847.
- Zhang, Y., W. Wang, X. Zhang, and Y. Li. 2008. A cluster validity index for fuzzy clustering. *Information Sciences* **178**:1205–1218.

Appendix A. Documentation for functions developed in the R statistical language to define land facets. Specifically, the functions identify outliers and the optimal fuzzy-*c* partition. Although the functions could be used as standalone tools, they were developed to work in concert with the Land Facet CorridorDesigner extension to ArcGIS 9.3 developed by J. Jenness. This extension implements the GIS-based procedures developed to design corridors based on land facets. R functions are presented below in the order in which they are intended to be used, and documentation is provided in a format similar to R reference manuals. R code in the “Examples” section of each function’s documentation builds on the sample code provided for the previous function. These functions depend on R packages `ks`, `e1071`, `lattice`, `clusterSim`, and `nnclust`.

---

`LF.kde`

*Kernel Density Estimation*

---

### Description

This function is a wrapper for the `Hpi.diag` and `kde` functions in package `ks` (Duong 2009). It selects a bandwidth matrix via a plug-in approach, which is subsequently used for kernel density estimation. Both operations are performed on a “binned” data set, which is necessary due to the large size of raster data sets. This function automatically plots the kernel density estimation and outputs information on the bins used for the kernel density estimation.

### Usage

`LF.kde(x, gridsize)`

### Arguments

<code>x</code>	the matrix or data frame containing topographic and/or soil data. Consists of one row per raster cell and one column per topographic or soil variable.
<code>gridsize</code>	binning grid size. The number of bins of equal interval created across the range of each variable. Due to memory constraints, <code>gridsize=151</code> is appropriate for 2-dimensional data (i. e., kernel density estimation is performed on a 2-dimensional array of $151^2$ bins). For 3-dimensional data, <code>gridsize=91</code> is appropriate (i. e., kernel density estimation is performed on a 3-dimensional array of $91^3$ bins). Larger dimensions require a smaller <code>gridsize</code> .

### Details

Functional for 1- to 6-dimensional data sets.

This function automatically plots the objects in `x` and the density estimation. For 1-dimensional data sets, the plot shows the univariate density curve and a rug plot of objects in `x`. For 2-dimensional data sets, density contours are overlaid onto the bivariate plot of objects in `x`. The contours have an interval of 10% up to 70%, and a 5% interval thereafter. The contour labeled “10” contains the densest 10% of objects. For 3-dimensional data, this function plots the objects in `x` and the 3-dimensional density contours. For simplicity, 25, 50, 75, 90, and 100% density contours are plotted. Red contours correspond to regions of higher density. This function also outputs to the

workspace `bin_width.csv`, a comma delimited file that contains the (multidimensional) half-width of bins used in the kernel density estimation. This file is necessary for compatibility with the Land Facet CorridorDesigner extension to ArcGIS.

### Value

<code>x</code>	dataframe containing topographic and/or soil data—same as input.
<code>width</code>	(multidimensional) half-width of bins used in the kernel density estimation. Same as values outputted in <code>bin_width.csv</code> .
<code>eval.points</code>	points at which the density estimate is evaluated.
<code>estimate</code>	density estimate at <code>eval.points</code> .
<code>H</code>	bandwidth matrix.
<code>h</code>	scalar bandwidth (1-dimension only).
<code>names</code>	names of variables in <code>x</code> .
<code>w</code>	weights.

### References

Duong, T. 2009. `ks`: Kernel smoothing. R package version 1.6.5.

### Examples

```
#An artificial data set
data <- as.data.frame(cbind(elevation=rnorm(10000), slope=rnorm(10000)))

#Inspect data
head(data)

#Perform kernel density estimation
kde <- LF.kde(x=data, gridsize=151)
```

---

`LF.outlier`

*Identify Outliers*

---

### Description

This function interpolates the density estimate of bins to individual objects in the data set and uses function `contourLevels` of package `ks` (Duong 2009) to identify outliers based on a density threshold. This density threshold can be estimated from the plot generated by function `LF.kde`. Function `LF.outlier` also outputs to the workspace information about the location of bins at or below the specified density threshold.

**Usage**

```
LF.outlier(x, threshold=90)
```

**Arguments**

**x** an object returned from function `LF.kde`.

**threshold** threshold corresponding to the density contour (from contour plot of `LF.kde`) beyond which observations are identified as outliers. For example, a threshold of 90 identifies the 10% of cells beyond the 90% contour as outliers (i. e., the 10% of cells with the lowest kernel density estimates would be identified as outliers).

**Details**

For compatibility with the Land Facet CorridorDesigner extension to ArcGIS, this function outputs to the workspace `grid.csv`, a comma delimited file that contains the location of bins within the density contour corresponding to the specified threshold.

**Value**

**outlier** a vector of length `nrow(x)` consisting of values 0 (indicates non-outlier) or 1 (indicates outlier).

**density** a vector of length `nrow(x)` containing the interpolated density for each object in `x`.

**References**

Duong, T. 2009. ks: Kernel smoothing. R package version 1.6.5.

**Examples**

```
#Identify outliers using a 90% threshold
interp <- LF.outlier(x=kde, threshold=90)

#Inspect interpolated density estimate of objects
head(interp$density)

#Tabulate number of outliers/non-outliers
table(interp$outlier)

#Create "outlier" column in data indicating which objects are outliers/non-outliers
data$outlier <- interp$outlier

#Inspect data
head(data)

#Plot non-outliers
plot(data[data$outlier==0,1:2], pch="." , col="gray")

#Overlay kernel density estimation contours to double check results
plot(kde, cont=c(seq(0,70,10), seq(70,100,5)), add = TRUE)
```

---

**LF.cluster***Fuzzy c-means Cluster Analysis*

---

**Description**

This function is a wrapper for the `cmeans` function in package `e1071` (Dimitriadou et al. 2009), plus additional functionality to perform fuzzy *c*-means cluster analysis on multiple values of *c* (the user-specified number of clusters into which the data should be classified), perform multiple clustering iterations for each *c*, and compute cluster validity indices. This function also automatically generates plots of the cluster centroids and validity indices for each iteration.

**Usage**

```
LF.cluster(x, nclust=c(2:7), niter=30, psi=1.5, max=1000000)
```

**Arguments**

- |               |   |
|---------------|---|
| <b>x</b>      | the matrix or data frame containing topographic and/or soil data, minus objects identified as outliers by function <code>LF.outlier</code> . Consists of one row per non-outlier raster cell and one column per topographic or soil variable.     |
| <b>nclust</b> | values of <i>c</i> for which to perform the cluster analysis.   |
| <b>niter</b>  | number of iterations of the cluster analysis to perform for each value of <i>c</i> .  |
| <b>psi</b>    | a weighting parameter that determines the overlap between clusters. A crisp classification (i.e., no overlap between classes) corresponds to <code>psi = 1</code> ; larger values of <code>psi</code> produce increasingly fuzzy classifications. |
| <b>max</b>    | the maximum number of observations upon which to perform the cluster analysis.<br>Provided to avoid memory limitations encountered on large data sets.  |

**Details**

This function could take hours to complete. For example, testing `c=c(2:7)` and `niter=30` on a data set containing 2 attributes for 256205 objects took 6 hours to complete on a Microsoft Windows XP platform with 3.0 GHz Intel Core 2 Duo processor. Computation time increases with *c* and the size of *x*.

This function automatically generates two plots. The first of these is a plot of cluster centroids for each iteration of each *c*, which is helpful for identifying instances where multiple partitions minimize the within-cluster variances for a given *c*. The second plot is of the validity indices for each iteration of each *c*, which is necessary for identifying *k*, the optimal number of clusters for the given data set.

**Value**

- |                  |  |          |                     |          |            |
|------------------|--|----------|---------------------|----------|------------|
| <b>centroids</b> | location of cluster centroids for each iteration of each <i>c</i> .  |          |                     |          |            |
| <b>validity</b>  | validity indices for each iteration of each <i>c</i> . Contents: <table border="0" style="margin-left: 20px;"> <tr> <td style="vertical-align: top; padding-right: 10px;"><b>c</b></td> <td>number of clusters.</td> </tr> <tr> <td style="vertical-align: top; padding-right: 10px;"><b>z</b></td> <td>iteration.</td> </tr> </table> | <b>c</b> | number of clusters. | <b>z</b> | iteration. |
| <b>c</b>         | number of clusters.  |          |                     |          |            |
| <b>z</b>         | iteration.   |          |                     |          |            |



FS	Fukuyama-Sugeno index (Zhang et al. 2008).
AWCD	Average Within Cluster Distance (Campello & Hruschka 2006).
XB	Xie-Beni index (Xie & Beni 1991).
XB_star	Xie-Beni* index (Celikyilmaz & Turksen 2008).
DB	Davies-Bouldin index (Davies & Bouldin 1979).
PBM	PBM index (Pakhira et al. 2004; Zhang et al. 2008).
CH	Calinski-Harabasz index (Calinski & Harabasz 1974).
FSil	Fuzzy Silhouette index (Campello & Hruschka 2006).

## References

- Calinski, R. B., and J. Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics* 3:1-27.
- Campello, R. J. G. B., and E. R. Hruschka. 2006. A fuzzy extension of the silhouette width criterion for cluster analysis. *Fuzzy Sets and Systems* 157:2858-2875.
- Celikyilmaz, A., and I. B. Turksen. 2008. Validation criteria for enhanced fuzzy clustering. *Pattern Recognition Letters* 29:97-108.
- Davies, D. L., and D. W. Bouldin. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1:224-227.
- Dimitriadou, E., K. Hornik, F. Leisch, D. Meyer, and A. Weingessel. 2009. e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R package version 1.5-19.
- Pakhira, M. K., S. Bandyopadhyay, and U. Maulik. 2004. Validity index for crisp and fuzzy clusters. *Pattern Recognition* 37:487-501.
- Xie, X. L., and G. Beni. 1991. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13:841-847.
- Zhang, Y., W. Wang, X. Zhange, and Y. Li. 2008. A cluster validity index for fuzzy clustering. *Information Sciences* 178:1205-1218.

## Examples

```
#Perform cluster analysis on non-outliers in data set
clust <- LF.cluster(x=data[data$outlier==0,1:2],
  nclust=c(2:7), niter=2, psi=1.5, max=1000000)

#Standardized cluster centroids
clust$centroids

#Validity indices
clust$validity

#Cluster centroids for 4 cluster solution
clust$centroids[clust$centroids$c==4,]
```

---

**LF.export***Export Information for Land Facet CorridorDesigner*

---

**Description**

Outputs to the workspace three comma delimited files necessary for compatibility with the Land Facet CorridorDesigner extension to ArcGIS: (1) cluster centroids for the optimal fuzzy-*c* partition, (2) mean and standard deviation of attributes (necessary for standardizing data), and (3) value for **psi** used in function **LF.cluster**.

**Usage**

```
LF.export(x, k=4, iter=1)
```

**Arguments**

<b>x</b>	object returned from <b>LF.cluster</b> .
<b>k</b>	optimal number of clusters.
<b>iter</b>	iteration having optimal validity indices for <b>k</b> clusters. Necessary only for instances where multiple solutions minimize the within-cluster variances for <b>k</b> partitions.

**Details**

Outputs to the workspace: (1) centroids.csv, which contains the location of cluster centroids for the optimal fuzzy-*c* partition; (2) params.csv, which contains the mean and standard deviations of the attributes in the data set; and (3) psi.csv, which contains the value of **psi** used in function **LF.cluster**.

**Examples**

```
#Export cluster centroids of optimal partition
LF.export(x=clust, k=4)
```

Appendix B. Source code (version 0.1) for functions developed in the R statistical language to define land facets.

### Function LF.kde and its dependencies

```
LF.kde <- function(x, gridsize){
  if(ncol(x) > 6) { #Error message for > 6 dimensions
    print(noquote("ERROR: x must have fewer than 6 dimesions"))
    return() #Process returns NULL value
  }

  if(ncol(x)>1){ #For >1-D Data
    Hpi <- Hpi.diag(x=x, pilot="samse", binned=TRUE,
      bgridsize=rep(gridsize,ncol(x))) #Estimate bandwidth matrix
    kde.binned <-kde(x, H=Hpi, binned=TRUE,
      bgridsize=rep(gridsize,ncol(x))) #Kernel density estimation
    if(ncol(x)==2){ #Plot kernel density estimation
      plot(x, col=rgb(102,139,139,50,maxColorValue=255), pch=16,
        cex=0.25)
      plot(kde.binned, add=TRUE, cont=c(seq(0,70,10), seq(70,100,5)))
    }
    if(ncol(x)==3){ #Plot kernel density estimation
      plot(kde.binned, drawpoints=TRUE,
        ptc=rgb(102,139,139,50,maxColorValue=255),
        cont=c(25,50,75,90,100), pch=16, size=0.25)
    }
  }

  if(ncol(x)==1){ #For 1-D data
    lab <- names(x)
    y <- x[,1]
    Hpi <- hpi(x=as.matrix(y), binned=TRUE,
      bgridsize=rep(gridsize,1)) #Estimate bandwidth
    kde.binned <-kde(y, H=Hpi, binned=TRUE,
      bgridsize=rep(gridsize,1)) #Kernel density estimation
    kde.binned$names <- lab
    plot(kde.binned, drawpoints=TRUE, #Plot kernel density estimation
      ptc=rgb(102,139,139,50,maxColorValue=255), xlab=lab)
  }
  #Identify half-width of bins used in kernel density estimation
  kde.binned$width <- LF.binwidth(kde.binned)

  kde.binned
}

#Function used in LF.kde to identify (multidimensional) half-width of bins
#used in kernel density estimation
LF.binwidth <- function(x){
  if(is.numeric(x$x)) pts <- matrix(unlist(x$eval.points), ncol=1,
    byrow=FALSE) #Convert list of eval.points to matrix
  if(!is.numeric(x$x)) pts <- matrix(unlist(x$eval.points),
    ncol=length(x$eval.points), byrow=FALSE)
  width <- matrix(NA, ncol=ncol(pts))
}
```

```

for(i in 1:ncol(pts)){
  d <- mean(nnfind(pts[,i])$dist)/2 #Distance between nearest neighbors/2
  width[1,i] <- d
}
colnames(width) <- x$names
write.csv(width, "bin_width.csv", row.names=FALSE)
width
}

```

### Function LF.outlier

```

LF.outlier <- function(x, threshold=90) {
  lev <- contourLevels(x, prob=((100-threshold)/100)) #Density threshold for
  #defining outliers

  if(is.numeric(x$x)) pts <- matrix(unlist(x$eval.points), ncol=1,
    byrow=FALSE) #Convert list of eval.points to matrix
  if(!is.numeric(x$x)) pts <- matrix(unlist(x$eval.points),
    ncol=length(x$eval.points), byrow=FALSE)
  dims <- ncol(pts) #Dimension of kernel density estimation
  size <- nrow(pts) #Gridsize
  bins <- matrix(c(combn(rep(1:size,dims),dims)), ncol=dims, byrow=TRUE)
  bins <- unique(bins) #Location of bins indexed according to position of values in
  #x$eval.points

  grid <- matrix(NA, nrow=nrow(bins), ncol=ncol(bins))
  for(i in 1:ncol(bins)){ #Loop to identify location of bins
    grid[,i] <- pts[bins[,i],i]
  }
  nn <- nnfind(grid, as.matrix(x$x))$neighbour #Find nearest bin
  #centroid to each object

  density <- numeric(nrow(bins))
  for(i in 1:nrow(bins)){ #Identify kernel density estimate of each bin
    if(ncol(bins)==1) density[i] <- x$estimate[bins[i]]
    if(ncol(bins)==2) density[i] <- x$estimate[bins[i,1], bins[i,2]]
    if(ncol(bins)==3) density[i] <- x$estimate[bins[i,1], bins[i,2],
      bins[i,3]]
    if(ncol(bins)==4) density[i] <- x$estimate[bins[i,1], bins[i,2],
      bins[i,3], bins[i,4]]
    if(ncol(bins)==5) density[i] <- x$estimate[bins[i,1], bins[i,2],
      bins[i,3], bins[i,4], bins[i,5]]
    if(ncol(bins)==6) density[i] <- x$estimate[bins[i,1], bins[i,2],
      bins[i,3], bins[i,4], bins[i,5], bins[i,6]]
  }

  index <- which(density>lev) #Identify which bins are non-outliers
  bins <- as.matrix(bins[index,]) #Location of non-outlier bins indexed
  #according to position of values in x$eval.points
  grid <- matrix(NA, nrow=nrow(bins), ncol=ncol(bins))
  for(i in 1:ncol(bins)){ #Loop to identify location of non-outlier bins
    grid[,i] <- pts[bins[,i],i]
  }
}

```

```

}
colnames(grid) <- x$names
write.csv(grid, "grid.csv", row.names=FALSE)

density <- density[nn] #KDE of each object
outlier <- ifelse(density > lev, 0, 1) #Identify which objects are outliers
list(outlier=outlier, density=density)
}

```

### Function LF.cluster and its dependencies

```

LF.cluster <- function(x, nclust=c(2:7), niter=30, psi=1.5, max=1000000){
  stand <- apply(x, 2, function(x) rbind(mean(x), sd(x))) #Mean and
    #SD of variables
  for(i in 1:ncol(x)){ #Standardize data
    x[,i] <- (x[,i]-stand[1,i])/stand[2,i]
  }

  for(c in nclust){ #Loop for multiple values of c
    print(noquote(paste("Number of clusters:",c, " Time started:",
      (Sys.time()))))
    for(z in 1:niter){ #Loop for multiple iterations for each c
      if(nrow(x)>max){ #Tale SRS wor of n=max from data, if necessary
        index <- sample(x, max)
      }
      else if(nrow(x)<=max){ #Otherwise, use all data for cluster analysis
        index <- 1:nrow(x)
      }

      fcm <- cmeans(x[index,], c, iter.max=1000, dist="euclidean",
        method="cmeans", m=psi) #Cluster analysis

      if(c==min(nclust) & z==1){ #Catalogue centroids from each iteration
        centroids <- as.data.frame(cbind(c=c,iteration=z,
          fcm$centers[order(fcm$centers[,1]),]))
      }
      if(c>min(nclust) | z>1){
        centroids <- rbind(centroids, cbind(c, iteration=z,
          fcm$centers[order(fcm$centers[,1]),]))
      }

      #Calculate validity indices
      FS <- fuksug(x=x[index,], centroids=fcm$centers, memb=
        fcm$membership, psi=psi) #Fukuyama-Sugeno Index
      AWCD <- awcd(x=x[index,], centroids=fcm$centers, memb=
        fcm$membership, psi=psi) #Average Within Cluster
        #Distance Index
      XB <- xiebeni(x=x[index,], centroids=fcm$centers, memb=
        fcm$membership, psi=psi) #Xie-Beni and Xie-Beni* Indices
      DB <- index.DB(x=x[index,], cl=fcm$cluster,
        centrotypes="centroids", p=2, q=1)#Davies-Bouldin Index
      PBM <- pbm(x=x[index,], centroids=fcm$centers,
        memb=fcm$membership, psi=psi) #PBM Index
    }
  }
}

```

```

CH <- index.G1(x=x[index,], cl=fcm$cluster,
               centrotypes="centroids") #Calinski-Harabasz Index
FUZZY.SIL <- fuzzy.sil(x=x[index,], centroids=fcm$centers,
                      memb=fcm$membership, cluster=fcm$cluster,
                      psi=psi, alpha=2) #Fuzzy Silhouette Index

if(c==min(nclust) & z==1){ #Catalogue Indices
  validity <- as.data.frame(cbind(c, z, FS, AWCD, XB$xb,
                                   XB$xb.star, DB$DB, PBM, CH, FUZZY.SIL$fuzzy.sil))
  names(validity) <- c("c", "iteration", "FS", "AWCD",
                      "XB", "XB_star", "DB", "PBM", "CH", "FSil")
}
if(c>min(nclust) | z>1){
  validity <- rbind(validity, c(c, z, FS, AWCD, XB$xb,
                                 XB$xb.star, DB$DB, PBM, CH, FUZZY.SIL$fuzzy.sil))
}
}

for(i in 1:ncol(centroids[, -c(1:2)])){ #Unstandardized cluster centroids
  centroids[,i+2] <- (centroids[,i+2] * stand[2,i] + stand[1,i])
}

#Generate plots
reps <- ceiling(niter/6) #Number of times to repeat colors and line styles below
my.colors <- rep(c("light blue", "red", "grey", "green", "pink", "blue",
                  "black", "orange", "dark green", "purple", "brown"), reps)
palette(my.colors)
linetype <- rep(1:6, reps)
max.n <- max(nclust) #Maximum number of clusters

#Plot cluster centroids
clust.lab <- paste(centroids$c, "Clusters", sep=" ")
p1 <- xyplot(centroids[,4]~centroids[,3]|clust.lab, group=iteration,
             data=centroids, as.table=TRUE, ylab=names(x)[2],
             xlab=names(x)[1], col=centroids$iteration, pch=19, type=c("p", "g"),
             panel=function(x, y,...){
               panel.text(x, y, labels=centroids$iteration,
                         col=centroids$iteration, pos=2, offset=0.5)
               panel.xyplot(x, y, col=centroids$iteration,...)
             })

#Plot validity indices
p2<- xyplot(PBM+CH+FSil+XB+XB_star+DB+FS+AWCD~c,
            groups=iteration, data=validity, type=c("l", "g"), col=my.colors,
            lty=linetype, xlim=c(2,max.n), xlab=list("Number of clusters (c)",
            cex=1.0), ylab=list("Index value", cex=1.0), layout=c(3,3),
            scales=list(x=list(relation="same", at=c(2:max.n), tck=c(1,1),
            cex=1.0, alternating=FALSE), y=list(relation="free", cex=1.0,
            tck=c(1,1))), par.strip.text=list(cex=1.0),
            strip=strip.custom(factor.levels=expression("PBMF",
            "Calinski-Harabasz", paste("Fuzzy Silhouette
            (", alpha, " = 2.0)", sep = ""), "Xie-Beni", "Xie-Beni*",
            "Davies-Bouldin", "Fukuyama-Sugeno", "Average Within-Cluster

```

```

Distance")), as.table=FALSE, between=list(x=0.5), key =
list(space="right", title="Iteration", cex=1, cex.title=1.25,
columns=1, lines=TRUE, text=list(as.character(1:niter)),
col=my.colors[1:niter], lty=linetype[1:niter]),
panel=function(x,y,...){
  panel.abline(v=c(3:(max.n-1)), col="light grey")
  panel.xyplot(x,y,...)
})

print(noquote(paste("Finally done:", Sys.time()))))
par(ask=TRUE) print(p1) print(p2) #Print plots to screen
list(validity=validity, centroids=centroids, stand=stand, psi=psi)
}

#Function used in LF.cluster to create and apply distance function for
#N-dimensional data
LF.dist <- function(x, y){
  if(is.vector(x)) x <- t(as.matrix(x))
  if(is.vector(y)) y <- t(as.matrix(y))
  fxn <- "sqrt((x[,1]-y[,1])^2" #Function for 1-dimensional data
  if(ncol(x)>1){ #Add to fxn for >1-dimensional data
    for(i in 2:ncol(x)){
      fxn <- paste(fxn, "+(x[,",i,"]-y[,",i,"])^2", sep="")
    }
  }
  fxn <- paste(fxn,")", sep="")
  dist.fxn <- function(x,y)
  eval(parse(text=fxn))
  d <- as.numeric(dist.fxn(x, y))
  d
}

#Function used in LF.cluster to calculate Fukuyama-Sugeno Index (Zahid et al. 1999)
fukusug <- function(x, centroids, memb, psi){
  rec <- numeric()
  for(i in 1:nrow(centroids)){
    cdist <- LF.dist(x, centroids[i,])^2 #Squared distances between objects
    #and centroid i for J_m, a measure of compactness
    mdist <- LF.dist(apply(x, 2, mean), centroids[i,])^2 #Squared distances
    #between mean of objects and centroid i for K_m, a measure of separation
    wdist <- (memb[,i]^psi) * (cdist-mdist) #Membership values weighted by
    #difference between compactness and separation measures
    rec <- sum(rec, wdist) #Sum of weighted squared distances
  }
  as.numeric(rec)
}

#Function used in LF.kde to calculate Xie-Beni and Xie-Beni* Indices
#(Zahid et al. 1999; Xie & Beni 1999; Celikyilmaz & Turksen 2008)
xiebeni <- function(x, centroids, memb, psi){
  rec <- numeric()
  for(i in 1:nrow(centroids)){
    cdist <- LF.dist(x, centroids[i,])^2 #Squared distances between objects
    #and centroid i for J_m, a measure of compactness

```

```

    wdist <- (memb[,i]^psi) * (cdist) #Membership values weighted
    #by compactness
    rec <- c(rec, sum(wdist)) #Sum of weighted squared distances
  }
  numerator <- sum(rec)
  cmin <- min(dist(centroids))^2 #Minimum distance between centroids squared
  denominator <- nrow(x)*cmin
  xb <- numerator/denominator #Xie-Beni Index

  numerator <- max(rec)
  xb.star <- numerator/denominator #Xie-Beni* Index

  list(xb=as.numeric(xb), xb.star=as.numeric(xb.star))
}

#Function used in LF.kde to calculate Average Within Cluster Distance
#(Campello & Hruschka 2006)
awcd <- function(x, centroids, memb, psi){
  rec <- numeric()
  for(i in 1:nrow(centroids)){
    cdist <- LF.dist(x, centroids[i,])^2 #Squared distances between objects
    #and centroid i for J_m, a measure of compactness
    wdist <- (memb[,i]^psi) * (cdist) #Membership values weighted by
    #compactness measure
    numerator <- sum(wdist)
    denominator <- sum(memb[,i]^psi)
    rec <- sum(rec, numerator/denominator) #Sum of weighted squared
    #distances
  }
  awcd <- 1/nrow(centroids)*rec as.numeric(awcd)
}

#Function used in LF.kde to calculate PBM Index (Pakhira et al. 2004;
#Zhang et al. 2008)
pbm <- function(x, centroids, memb, psi){
  E_1 <- sum(LF.dist(x, apply(x, 2, mean))) #Sum of distances between objects
  #and centroid of data
  J_m <- 0
  for(i in 1:nrow(centroids)){
    cdist <- LF.dist(x, centroids[i,]) #Squared distances between objects and
    #centroid i for J_m, a measure of compactness
    wdist <- (memb[,i]^psi) * (cdist) #Membership values weighted by
    #compactness measure
    J_m <- J_m + sum(wdist) #Sum of weighted squared distances
  }
  D_c <- max(dist(centroids))
  pbm <- ((1/nrow(centroids))*(E_1/J_m)*D_c)^2 as.numeric(pbm)
}

#Function used in LF.kde to calculate Fuzzy Silhouette Index
#(Campello & Hruschka 2006; Hruschka et al. 2006)
fuzzy.sil <- function(x, centroids, memb, cluster, psi, alpha){
  d <- matrix(NA, nrow=nrow(x), ncol=nrow(centroids)) #Receptacle for
  #distances between each object and all centroids

```



```

for(i in 1:nrow(centroids)){
  d[,i] <- LF.dist(x, centroids[i,]) #Distance between each object
    #and centroid i
}
s_i <- numeric(nrow(x)) #Receptacle for silhouettes of each object
for(i in 1:nrow(centroids)){
  indx <- which(cluster==i)
  a_i <- d[indx, i] #Distance between each object and its corresponding
    #cluster centroid
  if(nrow(centroids)==2) b_i <- d[indx,-i] #Distance between each object
    #and its next closest cluster centroid
  if(nrow(centroids)> 2) b_i <- apply(d[indx,-i], 1, min) #Distance
    #between each object and its next closest centroid
  s_i[indx] <- (b_i-a_i)/apply(cbind(b_i, a_i), 1, max) #Calculate
    #silhouette for object i
}
d <- numeric(nrow(x)) #Receptacle for differences between highest and second
  #highest membership
max.memb <- apply(memb, 1, max) #Highest membership value for each object
for(i in 1:nrow(centroids)){
  indx <- which(cluster==i)
  if(nrow(centroids)==2) sec.memb <- memb[indx,-i] #Second highest
    #membership for each object
  if(nrow(centroids)>2) sec.memb <- apply(memb[indx,-i], 1, max)
    #Second highest membership for each object
  d[indx] <- max.memb[indx] - sec.memb #Difference between highest and
    #second highest membership values
}
fuzzy.sil <- sum(d^alpha * s_i)/sum(d^alpha) #Calculate fuzzy
  #silhouette criterion
list(fuzzy.sil=as.numeric(fuzzy.sil), alpha=alpha)
}

```

### Function LF.export

```

LF.export <- function(x, k, iter=1){
  for(i in 1:ncol(x$centroids[, -c(1:2)])){ #Unstandardized cluster centroids
    x$centroids[,i+2] <- (x$centroids[,i+2] - x$stand[1,i])/x$stand[2,i]
  }
  write.csv(x$centroids[x$centroids$c==k & x$centroids$iteration==iter,
    -c(1:2)], "centroids.csv", row.names=FALSE) #Write cluster
    #centroids for optimal fuzzy-c partition
  row.names(x$stand) <- c("mean", "sd")
  write.csv(x$stand, "params.csv") #Write values for mean and SD of variables
  junk <- as.matrix(x$psi)
  colnames(junk) <- "psi"
  write.csv(junk, "psi.csv", row.names=FALSE) #Write value for psi used in
    #LF.cluster
}

```