Environmental Modelling & Software 52 (2014) 176–191

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

A proposal for an integrated modelling framework to characterise habitat pattern

Christine Estreguil [a], Daniele de Rigo [a,b,*], Giovanni Caudullo [a]

[a] European Commission, Joint Research Centre, Institute for Environment and Sustainability, Via E. Fermi 2749, I-21027 Ispra, VA, Italy
[b] Politecnico di Milano, Dipartimento di Elettronica e Informazione, Via Ponzio 34/5, I-20133 Milano, Italy

ARTICLE INFO

Article history:
Received 3 September 2012
Received in revised form
3 September 2013
Accepted 18 October 2013
Available online 21 November 2013

Keywords:
Spatial pattern
Morphology
Edge interface
Connectivity
Habitat reporting
Correlation analysis

ABSTRACT

Harmonized information on habitat pattern, fragmentation and connectivity is one among the reporting needs of the biodiversity policy agenda. This paper presents a generic, reproducible and integrated characterisation of patterns into one modelling framework. Three available conceptual landscape model components are customised, revisited and partly combined to derive a set of indices organized into four families: general landscape composition, habitat morphology, edge interface and connectivity. A harmonized mathematical description is provided for known and suggested new indices. Their unambiguous and easy computability is ensured with the integrated use of publicly available software (GUIDOS free-download software, Conefor Sensinode free software) and of newly programmed tools. An edge interface tool combining morphological analysis and a moving window landscape mosaic tri-dimensional model is presented; a "Power Weighted Probability of Dispersal" (PWPD) function is proposed to make connectivity indices sensitive to the landscape resistance.

The methodology is demonstrated for the focal forest habitat, by using sixty-five in-situ based habitat maps from the EBONE project ("European Biodiversity Observation NEtwork"). Twelve indices are applied. A statistical analysis is then conducted using classical linear correlation and nonlinear Brownian Distance Correlation (Mastrave free software modelling library) as alternative to traditional dimensionality-reduction techniques and with an effort towards reusability in other contexts and reproducible research, by means of concise semantic array programming codelets. The results highlight the less correlated and fundamental pattern components, corroborating the hypothesized hierarchical organization of the indices into four families, and also the feasibility of reducing further the number of indices within each category.

Software availability

The described modelling integration entirely relies on publicly available software (Fig. 3). Key passages are implemented by means of:

- GUIDOS Toolbox. Free-download software available at: http://forest.jrc.ec.europa.eu/download/software/guidos/.
- Conefor Sensinode. Free software (released under GNU GPLv3) available at: http://www.conefor.org/.

The statistical analysis is implemented by means of:

- Mastrave modelling library. Free software (released under GNU GPLv3+): http://mastrave.org/.

The complete modelling steps are summarized in Fig. 3 where straightforward passages integrate the use of well-established GIS tools (ESRI ArcGIS or GRASS GIS) and concise array programming codelets (Mastrave within GNU Octave, Python) A non-monolithic approach led to the use of semantic array programming for expressing less trivial steps as concise data-transformations easy to reuse and adapt (available in the article and online Supplementary materials).

1. Introduction

This research is motivated by the need for a reproducible and concise characterisation of landscape patterns based on key generic

* Corresponding author. European Commission, Joint Research Centre, Institute for Environment and Sustainability, Via E. Fermi 2749, I-21027 Ispra, VA, Italy.
E-mail addresses: christine.estreguil@jrc.ec.europa.eu (C. Estreguil), derigo@maieutike.org, daniele.de-rigo@ext.jrc.ec.europa.eu, derigo@elet.polimi.it (D. de Rigo), giovanni.caudullo@ext.jrc.ec.europa.eu (G. Caudullo).

1364-8152/$ – see front matter
http://dx.doi.org/10.1016/j.envsoft.2013.10.011

ecological principles and integration of available approaches. Despite the plethora of landscape pattern measures available in literature, methodological guidance is still missing on how to conduct pattern assessment (Bogaert, 2003; Riitters et al., 2009) to ease and better support the (non-expert) user community in implementing policy, such as for continental reporting on habitat fragmentation and connectivity in the biodiversity policy agenda (European Commission, 2011; European Environment Agency, 2012; Convention on Biological Diversity, 2010; Forest Europe, 2011). This research used scientifically well-founded landscape ecological (inter-related) principles which exist in literature (Lindenmayer et al., 2008) and policy guidance documents for reporting on impact of fragmentation and climate change (Fischer and Lindenmayer, 2007; Kettunen et al., 2007).

The use and combination of more than one landscape conceptual measure index is strongly recommended to provide more insight for landscape conservation, yet it is rarely done (Lindenmayer et al., 2008). Indeed, no single measure can fully capture the complexity of the spatial arrangement of patches. Often, studies concentrated on measuring one component of spatial pattern, like landscape composition, spatial pattern or connectivity while such components are inextricably linked. On the other hand, the combination of multiple components of a pattern into a single value (Bogaert et al., 2000) or the reduction of the number of indices using factor analyses failed to render the ecological meaning of the designated index to the analyst (Herzog et al., 2001). Where the intrinsic multiplicity of problem dimensions appears so evident, modelling integration should avoid hiding their trade-offs, of possible relevance in the science–policy interface, and should transparently use multiple criteria (de Rigo, 2013) so as to help multifunctional analysis (O'Farrell and Anderson, 2010). Indices should be more effective in terms of their ability to capture different aspects of spatial pattern, their simplicity and their ease of interpretation (Li and Wu, 2004), regardless they are applied at large scale (Riitters et al., 2000, 2009) or at habitat scale (Wrbka et al., 2004). The indices should be organized into landscape pattern components which remain ecologically meaningful and easily understood by the user community – including non-experts. For meaningful inferences in pattern–process correlation analysis, simple measures[1] (patch size, edge, inter-patch distance, proportion) are recommended rather than with complex nonlinear indices (evenness, etc.), as well as relative and well explained range of index values (range from 0 to 1 with clear ecological meanings of the minimum and maximum values). A low redundancy of indices is further required within and across the different landscape components.

To address the integration of indices into a concise modelling frame organized into landscape pattern model components, the current study proposes to concentrate on four key pattern related principles which are listed below and for which three available landscape modelling approaches are potentially relevant. The three approaches that will be tested, revised, customised and programmed when necessary have already been proved valuable at different application scales and geographical regions.

- First, easily computable measures are required to describe a focal habitat in a given landscape in terms of its total amount, its pattern and landscape context. Habitat pattern (spatial arrangement of patches) is inextricably linked to habitat amount in assessments (Koper and Schmiegelow, 2006). Habitat pattern affects the interactions between and within species both within and between patches. The landscape context of habitats, in other terms the interstitial environment between patches or habitat matrix (Dennis et al., 2003) influences habitat content (e.g. vegetation condition).
- Second, morphological shapes of habitat play an ecological role. For example, the geometry of habitat edges (protrusions, corners) and the presence of clumps of habitat in the landscape matter for aggressive edge specialists (Taylor et al., 2008); linear strips of habitat enhance the spatial continuity in a fragmented landscape; interior areas of patches do not experience strong influences from neighbouring patches of other land cover categories (Rutledge, 2003). The mathematical morphological spatial pattern analysis (MSPA) application in the free-download software GUIDOS,[2] developed by Soille and Vogt (2009), provides automatically and unambiguously a segmentation of geometric features from any binary map. It particularly allows the detection of linear connecting pathways between patches and branches at edges as well as disconnected patches. The results are mutually exclusive morphological pattern classes ('core' and non-core as 'perforated,' 'edge,' 'islet', 'connector', and 'branch'). The software was applied for different purposes, among others in the US (Wickham et al., 2010), Europe (Mubareka et al., 2011; Clerici and Vogt, 2013) and Africa (Bucki et al., 2012). The method provides at all scales more precise spatial and thematic pattern classification than the amount-adjacency model based on image convolution from Riitters et al., 2002 (Vogt et al., 2007a,b). Because pattern classes are mapped at pixel level, it is also better suited than aggregated measures over fixed area grid as in traditional patch area and edge based measures (McGarigal et al., 2002). Furthermore, because MSPA uses geodesic distance to implement edge width and derive all non-core classes, edge widths are not rounded to the nearest distance in increments of the cell size as in traditional edge measure like in McGarigal et al., 2002. However, MSPA requires a customisation of entry parameters and outcome classes adapted to the field of application. Its main limitation is the over-simplification of the landscape in a binary model.
- Third, habitat edges are interfaces between two types of habitat. Edge effects may be positive (high biodiversity) or negative (spread of exotic species) features for a landscape. The permeability of edges influences habitat quality for interior-inhabiting species (Ries and Sisk, 2004) depending on the similarity of the adjacent habitat types (Lidicker and Peterson, 1999). The discrimination between natural/semi-natural types of interfaces and more anthropogenic ones are relevant to the edge permeability or "hardness" that is, its resistance to being crossed by focal organisms. For example, human-induced edges are more short-term "hard" landscape features such as woodland-cultivated interfaces, while natural edges are more a long-term "soft" feature (due to soil type, topography, etc.) with high structural diversity (Ries et al., 2004; Ries and Sisk, 2004). The landscape-level mosaic approach from Riitters et al. (2009) describes the landscape mosaic context of a focal land cover class and enables the mapping of edge interface zones at pixel level while other traditional edge contrast measures provide statistics at patch, class or landscape level (McGarigal et al., 2002). It was recently applied to evaluate the anthropogenic risks of grassland and forest habitat degradation from land cover maps over the United States. The model 'integrates' but is not explicit enough on the geometry of patches of a focal land cover (whether the edge is from a patch including interior habitat, a linear connecting path, a protrusion at edge of a patch) and on

[1] Mostly, non-additive measures.

[2] http://forest.jrc.ec.europa.eu/download/software/guidos.

their inter-patch distances. This landscape model is not available as software and needs to be programmed.

- Last, habitat connectivity matters for the migration and survival of species and for the control of invasive species and diseases. It is a combined product of structural and functional connectivity (Kindlmann and Burel, 2008) and includes habitat connectedness for a given taxon, the species perspective of landscape connectivity (connectedness of ecological processes) and the human landscape perspective (physical connectedness of habitat patches). The probability of connectivity index developed by Saura and Torné (2009) is based on structural measures (distances, habitat availability and graph theory) and functional measures (probability of dispersal according to dispersal distance and matrix resistance). The free software Conefor Sensinode[3] allows the index to be computed. This index was recently modified into the equivalent connected area (ECA) index, defined as the size of a single patch (maximally connected) that would provide the same probability of connectivity as the actual habitat pattern in the landscape (Saura et al., 2011a). The index sensitivity to habitat amount and matrix permeability still needs to be addressed.

Recently, the integration of the MSPA application of GUIDOS and the Conefor has been tested by Saura et al. (2011b) for assessing the individual importance of core habitat areas and connectors in providing connectivity. We also propose to test the integrated benefit from the individual and complementary strength of the landscape model components and to revisit or amend related indices. A harmonized mathematical description of indices (based on arrays) will be provided. In addition, particular attention will be paid on the feasibility of unambiguous reproducibility (Morin et al., 2012) of the modelling approach. The “curse of ambiguity” mentioned by Ince et al. (2012) highlights the role of free scientific software; availability of software code and data are the first steps towards reproducible research according to Peng, 2011. GUIDOS and Conefor are two publicly available software packages. GIS and array programming tools will be used when implementing the approach. The required array programming entirely relies on free software. For the GIS data-transformations, an alternative is provided to ESRI ArcGIS by means of equivalent transformations available in the free software GRASS GIS. Since software uncertainty appears as unavoidable (Lehman and Ramil, 2002), conformity of software with the described mathematical formulation (Joppa et al., 2013) is supported by exposing (see Annex 4 of Supplementary materials) concise array programming codelets (de Rigo, 2012c) of less trivial processing steps.

A statistical linear and nonlinear correlation analysis is further proposed for identifying a subset of less correlated indices. The overall objective of the analysis is to highlight the distinctive aspects of less correlated indices, in the view of investigating the significance and applicability of the proposed indices' classification in a few categories defined according to the previously mentioned ecological principles. The proposed analysis is an alternative to traditional dimensionality-reduction techniques. Methods such as Principal Component Analysis and hierarchical clustering of indices are more traditionally used in the landscape ecology field as in McGarigal et al. (2009). The new suggested approach is designed in order for index nonlinearities (Li and Wu, 2004) to be elegantly and concisely taken into account with a wide spectrum of generality — also considering general monotonic transformations in a robust rank-based (Lechner et al., 2013) assessment.

Due to the relatively small set of samples, the analysis purposely did not attempt to derive explicit linear or nonlinear principal components[4] (Kramer, 1991; Dong and McAvoy, 1996; Schölkopf et al., 1998; Scholz and Vigàrio, 2002; Scholz et al., 2008), nor to focus on explicit linear or nonlinear regressors between the indices. Instead, a more general approach is proposed for estimating “the degree of all kinds of possible relationships” (Szekely and Rizzo, 2009a) between indices without deriving actual relationships. The approach is based on the Brownian Distance Correlation (Szekely et al., 2007; Szekely and Rizzo, 2009a,b; Lyons, 2011) and benefits from the conciseness and easy reproducibility of the semantic array programming paradigm as implemented (de Rigo, 2012a) by the Mastrave[5] modelling library (de Rigo, 2012b,c) within the GNU Octave computing environment (Eaton et al., 2008). It allows the contribution of each index to be assessed as a provider of information not derivable from other indices or groups of indices, neither considering linear nor nonlinear relationships, under weak mathematical assumptions.[6] The proposed approach is straightforward to replicate and reuse in different contexts where a similar analysis of heterogeneous indices would be of interest. It is designed as a statistically intensive data-transformation for deriving a concise set of rankings (see Table 8) of less correlated indices, also allowing comparison against aggregations of indices (families). Despite being inherently quantitative, the analysis stresses the importance of subsequent semantic interpretation of the qualitative information it summarizes, aiming at providing evidence for corroborating or rejecting hypotheses and better analysing some less obvious emerging relationships.

As a case study, the models and statistical analysis are investigated from a harmonized in-situ database of habitat maps available across Europe in the European EBONE research project (“European Biodiversity Observation Network” project at www.ebone.wur.nl). It provides a new unique opportunity to derive more general principles getting away from the traditional case study specific in a single geographical location, or from the landscape specific or species-specific landscape ecology investigations (Lindenmayer et al., 2008). The study emphasis is on habitat pattern rather than on habitat function or quality, and on providing guidance in the selection of indices and their harmonized, unambiguous mathematical formulation. The current analysis builds upon preliminary pattern model implementation and results over the same dataset available in Estreguil et al., 2011. It now develops one mathematical framework that classifies known indices and suggests new ones as logical complement. It also introduces the idea of a unique modelling framework for implementation. The paper starts with presenting and integrating the models by providing an abstract mathematical and semantic formulation of indices, and their aggregation in families (as semantic dimensions characterising habitat patterns). It then develops from a concrete case study towards a statistical analysis aiming at the corroboration of the emerging classification of four index families.

2. Data and methods

2.1. General landscape

Because European landscapes are long influenced by humans, native or so called ‘natural’ vegetation is rare, ‘natural’ is a poor facsimile of what was natural for that ecosystem. To describe the general landscape composition, the term ‘natural’ refers

[3] http://www.conefor.org. An updated list of worldwide applications is given in the “Applications” section of the Conefor website.

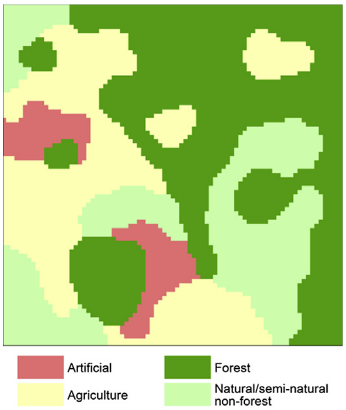
[4] The identification of meaningful components, to which the variety of indices can be mostly reduced, may be classified within the *feature extraction* methodologies (Scholz et al., 2008).

[5] http://mastrave.org.

[6] The indices, considered as real-valued random variables, are expected to have finite second moments (Szekely and Rizzo, 2009a).

Table 1
Data input and notation.

Data input and notation



The land use/land cover of a landscape is simplified into four classes: Focal habitat (Forest), Natural non-focal habitat (Natural non-forest), Agriculture, Artificial.

i, j refer to patches of the focal habitat (from 1 to n, where n is the number of patches in a given landscape of area A_L).

a_i is the area of the i-th focal habitat (Forest) patch.

The share of natural/semi-natural habitats in landscapes is measured in addition to the share of the focal habitat of interest (Table 2).

to natural and semi-natural habitats which are analysed in contrast to more intensively used landscapes such as agricultural and artificial surfaces (Table 1).

2.2. Morphological pattern model

Automatic spatial pattern mapping was tested by implementing the mathematical morphology analysis (MSPA) proposed by the software GUIDOS. This application provides twenty-five classes which are obtained by segmenting an input binary raster map (foreground focal habitat class set to 1 and background non-focal habitat set to 0). They can be reclassified into seven mutually exclusive spatial pattern classes (core, edge, islet, bridge, loop, branch, perforation). The classes were customised for our study. They were partly renamed and aggregated into the five classes as follows; the shares of their focal habitat area were also calculated (Table 3).

'*Interior*' (IF): foreground pixels beyond a distance of a given size parameter s to the background (s), and obtained by erosion of the input map with a Euclidian disk of radius equal to s. The class *Interior* is the *core* MSPA original class. The edge size s is the only parameter fixed by the operator.

'*Islet*' (IS): foreground pixels that do not contain any *core*. *Islets* are areas of small and/or elongated and thin non-core isolated fragments. Islets are potentially vulnerable to disappear due to their shape and size; depending on their landscape context, they offer stepping stones for pollination and dispersal of habitat species between core patches.

'*Boundary*' (BO): pixels on the boundary of a cluster of core pixels, outer side (*edge*) or internal side (*perforation*). Boundaries are edges with a fixed width, they likely host edge habitat which in turn may have effects on interior habitat.

'*Connector*' (CO): foreground pixels with no core that connects at least two different core units (*bridge*) or connects to the same core unit (*loop*). They are structural connections between interior parts of patches and potentially act as pathways for species.

'*Branch*' (BR): foreground pixels with no core that is connected at one end only to a connector or an edge of core. Branches typically identify habitat protrusion at edges (for example, tree encroachment in grasslands after land abandonment).

The aggregation of *connector* and *branch* classes is also proposed to capture *linear* features (LI). *Linear* features represent linear habitat areas physically connected to *interior* patches, in contrast to *islets*.

2.3. Edge interface model

For a focal habitat class, different types of edge interface need to be discriminated in terms of edge morphology (e.g. linear features, edge of interior habitat) and adjacent habitats/land uses (e.g. edges along natural/semi-natural lands, edges along more anthropogenic lands i.e. agricultural and/or artificial). The similarity between the focal habitat and adjacent habitat (e.g. forest–grassland) is also a concern. To achieve this, the combination of the previously introduced morphological model and an available landscape mosaic model was conceptualised, programmed and tested.

In the landscape mosaic model originally developed by Wickham and Norton (1994) and amended in Riitters et al. (2000 and 2009), the landscape context of a focal class can be obtained by applying a 'moving window' approach over a tri-dimensional input raster habitat map (e.g. three input classes U, A, N as artificial, agricultural and natural lands in Figs. 1 and 2).

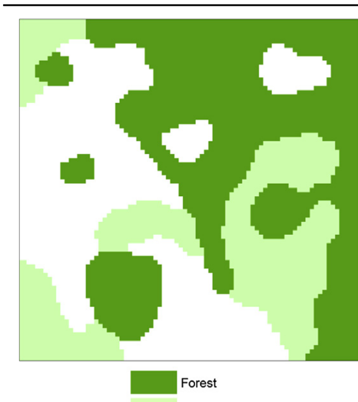
A given piece of land is classified into one of seventeen landscape pattern types on the basis of the proportion of the three main classes in immediate surroundings and pre-defined thresholds (landscape triangle in Fig. 1). The size of the surroundings (landscape context) is defined by the size of the 'moving window' around the piece of land. The output raster map provides seventeen landscape mosaic pattern types. Edge interface zones are mapped for each of the three main classes (e.g. U, A, N). Only core pattern categories (NN, AA, and UU) may include both edges and interior part of patches (Fig. 2 centre). In this study, a tool ('MOSAIC') was developed to implement and amend this conceptual model. The seventeen types were aggregated into four main mosaic pattern categories for each of the input three classes. Further, a "similarity" assumption was made to categorise the permeability or "hardness" of edge interfaces between adjacent classes. Sub-classes of each input class (typically forest–other natural class) were considered similar, making among them soft interfaces.

Fig. 2 (right) illustrates the four mosaic pattern categories for a focal habitat (forest) class belonging to the N category:

- Focal habitats in 'core natural' patterns (NN) are always (100%) adjacent to natural/semi-natural habitats or in the interior part of patches. "Soft" edge interface types are classified as natural.
- Focal habitats in 'mainly natural' patterns (N) are mainly (80%) bordering natural/semi-natural habitats, thus in "soft" edge interface types which are classified as natural.
- Focal habitats in 'mixed natural' patterns (MN = Na + Nu + Nua) are embedded in a predominant natural context (Nx), but are significantly fragmented by agricultural and/or artificial land. "Hard" edge interface types are classified as 'artificial' interface.

Table 2
Indices based on general landscape model.

General landscape model



The landscape composition is summarized by the share in natural habitats and the share of the focal habitat of interest.

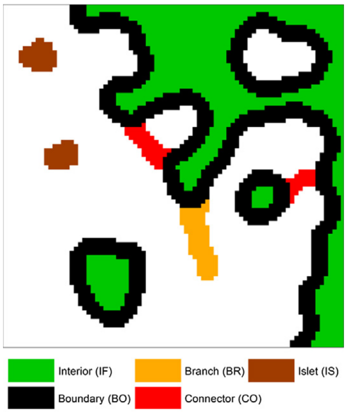
Two indices: Focal habitat (forest) land proportion; Natural land proportion.

Units: dimensionless (percentage)

Range: the index approaches 0% when the focal habitat (or natural land) is rare in the landscape unit and equal 100% when the entire landscape unit consists of focal habitat (or natural land). The index is 0% for no focal habitat (or natural land) in the landscape unit.

Why? Fragmentation involves the reduction of total focal habitat (forest) area. The presence and dominance of natural/semi-natural lands in a landscape make it easier for species to adapt to focal (forest) habitat losses.

Table 3
Indices based on the morphological pattern model.

Morphological pattern model	
	The focal habitat cover is described according to five classes. Interior areas are beyond a fixed distance to the border (edge width); Boundaries of interior areas; Connectors and Branches are Linear features that are always connected to interior areas, while Islets are physically isolated. Indices: focal habitat share in Interior, Boundary, Connector and Branch (Linear feature) and in Islet. Unit: dimensionless (percentage) Range: the index ranges between 0% and 100% and is calculated for each morphological class. The sum of the proportions of each morphological class makes 100% i.e. the whole focal habitat. Why? Fragmentation relates to the ratio of interior versus non-interior habitat. Interior areas provide suitable habitat for interior species. Boundaries are more exposed to the penetration of invasive species. Linear features and islets are key features for habitat provision services but are often vulnerable to disappear due to their shape and size.

■ Focal habitats in 'some natural' patterns (SN = sum of all the other types) are predominantly in human induced contexts (Ux, Ax, Mix) (i.e. forest patch in agricultural landscape), thus with "hard" edge interface types which are classified as 'artificial' interface.

Then, the mosaic model was combined with the morphological model to discriminate between the interior and edge part of patches, and among different edge morphologies (boundary, islet, connector, branch, linear feature). The combination consisted of overlapping each model output pattern maps for the focal class of interest. The 'window' of the mosaic model must be defined as a Euclidian disk of radius *s*, as in the MSPA of the GUIDOS software. By doing so, the landscape mosaic context of each non-interior habitat morphological shape class ($MORPH_{Mosaic}$) was mapped at pixel level, thus enabling to characterise the edge interface as natural (for 'soft' interface with similar adjacent habitats) or as artificial (for 'hard' interface). A new "similarity" index (SI) was proposed to translate the dominance of natural edge interface (landscape mosaic context equal to NN) for each specific habitat morphological shape class (Table 4). For example, for habitat patches with interior areas, SI-BO_{NN} gives the NN proportion in their boundary BO class; i.e. the edge proportion always along natural/semi-natural habitats. Similarly, it can be applied for other edge morphological shapes (SI-IS_{NN}, SI-CO_{NN} SI-LI_{NN}). Alternatively, SI-BO_{MN+SN} refers to artificial edge interfaces and provide the proportion of habitat edges which are along anthropogenic lands.

This model combination further enables to amend the delineation of interior habitat by accounting for both the morphological criteria (as in IF class) and the 'hardness" of edge interfaces. No edge depth (*s*) is then applied when adjacent habitats are similar to the focal class. For example, when forest is adjacent to 'natural' habitats, the interior forest (IF*) class is the IF class enlarged by the NN part of the boundary class (BO_{NN}) and its forest proportion (IF*P) could be calculated (Table 4).

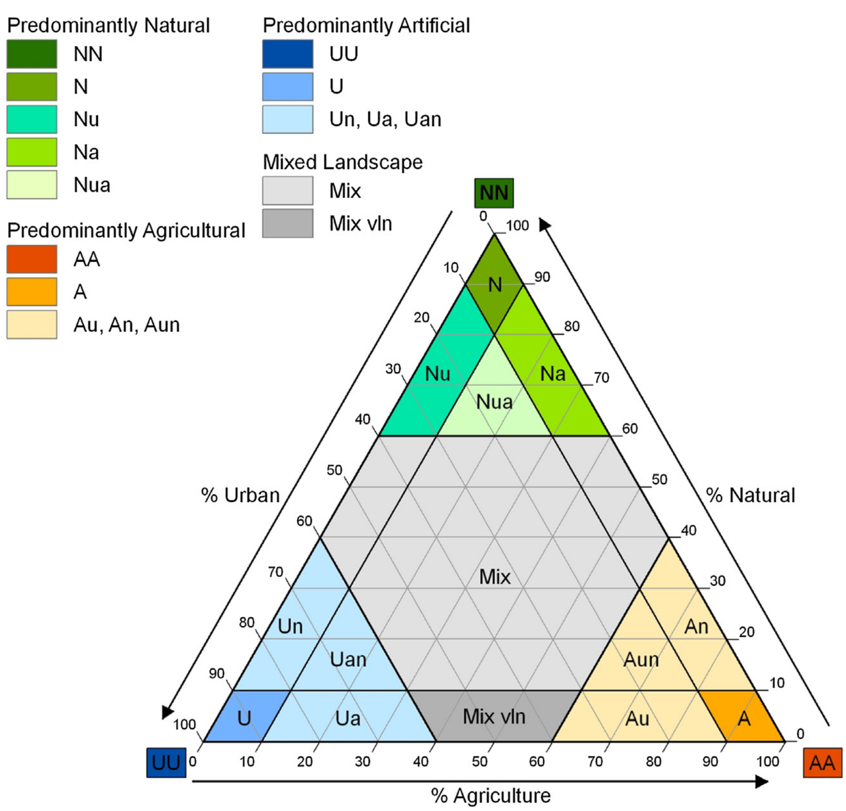


Fig. 1. The fifteen landscape pattern types derived with the landscape mosaic index (from Estreguil et al., 2011).

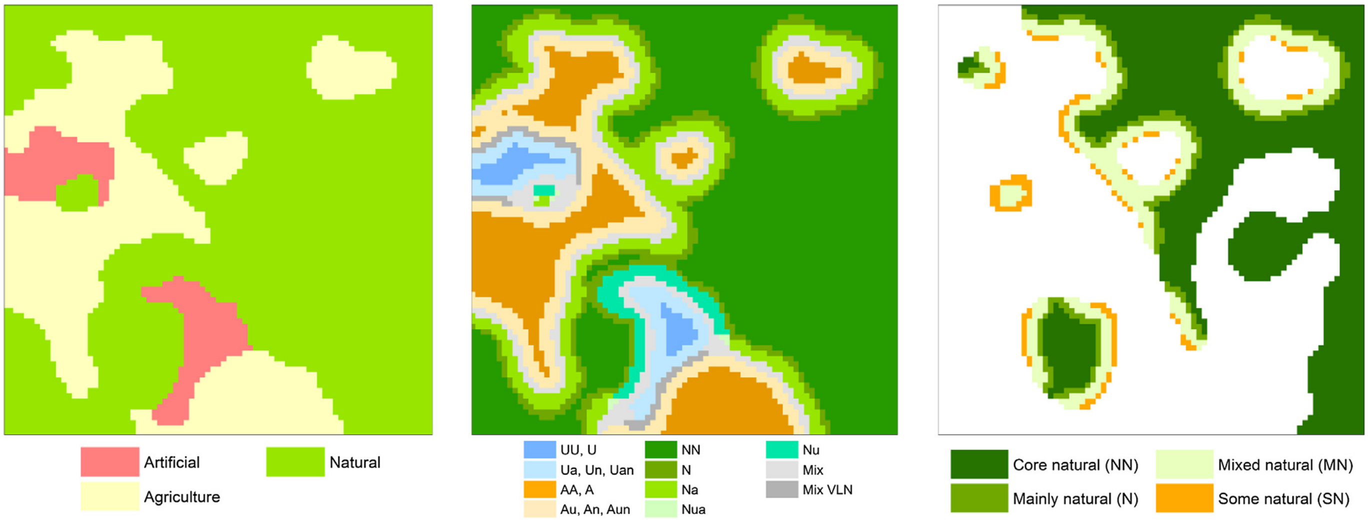
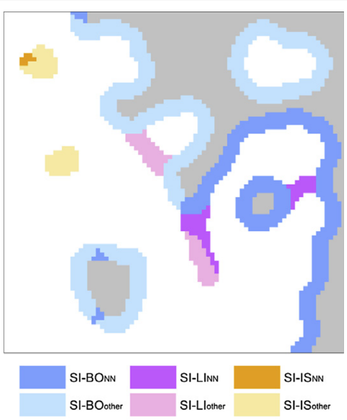


Fig. 2. Processing steps in the landscape mosaic index. Left: aggregation of landscape habitat into three main categories; centre: landscape types as result of mosaic process; right: four main landscape categories mapped for a focal habitat (forest).

Table 4
Indices based on the edge interface model.

Edge interface model



The edge interface model is derived from the integration of the landscape mosaic and the morphological models. Focal habitat edge interfaces are differentiated by morphology (edges of interior, connector and branch as linear features, islets) and characterised according to the similarity of adjacent habitats (focal habitat edges along natural/semi-natural lands as *natural edge interface*, edges along more anthropogenic (i.e. agricultural and/or artificial) lands as *artificial edge interface*).

Three indices: focal habitat edge proportion with a *natural edge interface*, focal habitat edge with an *artificial edge interface*, focal habitat proportion in interior areas and with a natural edge interface.

Units: dimensionless (percentage)
Range: the indices range between 0% and 100%

Why? In temperate regions, fragmentation relates to the shift in land use at edges. The permeability of interfaces for species dispersal depends on the similarity of adjacent habitat types and is likely higher in the case of natural edge interfaces.

2.4. *Connectivity landscape model*

The free software Conefor Sensinode 2.2[7] (Saura and Torné, 2009) computes the area-weighted Probability of Connectivity (PC) index for a focal class in a given landscape, based on topology (inter-patch distances), patch attributes such as area and species specific dispersal ability. In this model, each link between every two patches is characterised by a probability of dispersal, obtained as a function of distance (a decreasing exponential function of either the Euclidean (straight-line) edge-to-edge distance or the effective distance, matching to a 50% probability for a specific average dispersal distance). In PC, for each pair of patches, the weight for areas (intra-patch) is important in comparison to the small value of the p_{ij} component. Our concern was thus about rendering the connectivity index more sensitive to the landscape matrix permeability ("hardness" of dispersion depending on similarities of adjacent habitats in the matrix), and allowing more flexibility in accounting for intra-patch areas and inter-patch dispersal in the index mathematical formulation.

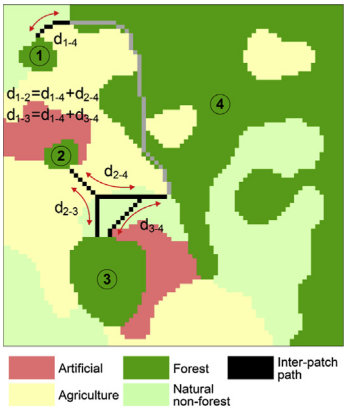
A new family of connectivity indices was formulated. It may be referred to as *Power Weighted Probability of Dispersal* (PWPD), i.e. a power function of the probabilities of dispersal between patches, weighted on the basis of a generic function of the corresponding origin and destination patch areas (Table 5). The numerator accounts for the information conveyed by the probabilities of dispersal and by appropriately weighting them. The denominator analyses the quantities used to derive the weight-functions (i.e. the areas of the i-th and j-th patches between which the ij-th corresponding probability of dispersal is computed) along with A_L to provide a normalization factor. The weight-function of choice can generically transform the information on patch areas to consider for example either their product (as in the case of PC) or only one of them per each pair or even neither of them, in which case the weight-function is a constant function. Auto information, when origin and destination patches are the same patch, can be included or excluded by means of the exponent α which easily allows the Kronecker delta δ_{ij} to be neutralized (the effect of δ_{ij} is to mathematically isolate auto information). Table 5 illustrates the way each component of the PWPD equation – applied to the focal habitat (f), average friction per distance unit (avf) and average dispersal distance (d) as $PWPD^{(f,avf,d)}$ – is instantiated to generate each of the further considered connectivity indices:

- The PC and RPC indices (i.e. square root of PC) apply the product of areas for each pair of patches. As for the equivalent connected area (ECA) index in Saura et al. (2011a), RPC offers a more reasonable and usable range of variation (from 0 to the total focal habitat proportion FP), and an easier and more straightforward interpretation with respect to the focal habitat area proportion. The 'distance' between RPC and FP in the landscape depends on how large are the patches, how far they are from one another and how easy is the dispersal in the landscape matrix in between focal habitat patches. In two landscapes with equal focal habitat area proportion (FP), the more connected landscape will be the one with RPC closest to FP, i.e. with larger and more effectively connected patches.

[7] http://www.conefor.org.

Table 5
List of indices derived from the connectivity model.

Connectivity model



The probability of habitat connectivity in a given landscape is derived on the basis of habitat area, topology and the landscape permeability for a given species dispersal ability. It relates to functional distances between habitat patches.

Four sub-indices derived by Power Weighted Probability of Dispersal (PWPD):
Area-weighted connectivity index set: focal habitat (forest) connectivity (intra-patch and inter-forest patch).
- PC: Probability of Connectivity index.
- RPC: Root Probability of Connectivity index.
- IsoSi: Isolation Sensitive Index.

Unweighted connectivity index: proxy of landscape permeability (only inter-habitat patch functional distance with resistance values for the non-habitat lands.
- APC: Average of Probability of Connectivity.

Units: dimensionless

Range: [0, 1]

Why? Fragmentation is about isolation due to increased functional distances between patches. The lack or loss of connectivity reduces the capability of organisms to move and can interfere with pollination, seed dispersal, wildlife migration and breeding.

Generic connectivity index description	Index
Power Weighted Probability of Dispersal (PWPD): a power function of the probabilities of dispersal between patches, weighted on the basis of a generic function of the corresponding origin and destination patch areas.	$\mathrm{PWPD}^{(f,avf,d)} = \left(\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} f\left(a_i^{(\mathrm{f})}, a_j^{(\mathrm{f})}\right) \cdot (1-\delta_{ij})^{\alpha} \cdot p_{ij}^{(f,avf,d)}}{g\left(a^{(\mathrm{f})}, A_{\mathrm{L}}\right) \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} (1-\delta_{ij})^{\alpha}} \right)^{\beta}$
i, j: focal habitat patches n: number of nodes a_i, a_j: area of patches i and j A_{L}: area of landscape unit $p_{ij} = e^{k \cdot \mathrm{cost}_{ij}}$ probability of dispersal	cost_{ij}: least-cost path from i to j $k = \frac{\ln(0.5)}{avf \cdot d_{50\%}}$ constant of probability avf: average friction per distance unit $d_{50\%}$: cost distance at 50% probability

Name	$f(a_i, a_j)$	$g(a, A_{\mathrm{L}})$	α	β	δ_{ij}	Index
PC	$a_i \cdot a_j$	$\left(\frac{A_{\mathrm{L}}}{n}\right)^2$	0	1	1	$\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_i \cdot a_j \cdot p_{ij}}{A_{\mathrm{L}}^2}$
RPC	$a_i \cdot a_j$	$\left(\frac{A_{\mathrm{L}}}{n}\right)^2$	0	1/2	1	$\sqrt{\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_i \cdot a_j \cdot p_{ij}}{A_{\mathrm{L}}^2}} = \sqrt{\mathrm{PC}}$
IsoSi	a_j	$\frac{A_{\mathrm{L}}}{n} \equiv \left(\frac{A_{\mathrm{L}}}{n}\right)^1$	1	1	1	$\frac{\sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} a_j \cdot p_{ij}}{A_{\mathrm{L}} \cdot (n-1)}$
APC	1	$1 \equiv \left(\frac{A_{\mathrm{L}}}{n}\right)^0$	0	1	1	$\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij}}{n^2}$

- The area-weighted connectivity index adapted from Hanski (1994, 1998),[8] called Isolation Sensitive Index (IsoSi), and similar to PC, only considers the destination patch area, but its mathematical formulation is symmetric with that considering the origin patch area instead. It puts less emphasis on intra-patch connectivity and renders it more sensitive to the inter-patch landscape matrix permeability and possible barrier effects, and more focused on the probability of species movement.
- A new index of connectivity, untitled the unweighted Average of Probability of Connectivity (APC), only accounts for the probability of dispersal between patches based on the matrix resistance, the configuration of focal habitat patches but not their areas. It is normalized with the square of the patch number.

As shown in Table 5, the denominator function $g(a^{(\mathrm{f})}, A_{\mathrm{L}})$ may be simpler.[9] All four indices can be derived using a simplified function $g(A_{\mathrm{L}})$ in the form of $(A_{\mathrm{L}}/n)^{\gamma}$ which does not consider information on focal habitat area $a^{(\mathrm{f})}$. The proposed general formulation allows a richer set of connectivity indices to be described as instances of the same equation, as for example the normalization function in the denominator might combine $a^{(\mathrm{f})}$ and A_{L} or instead only refer to $a^{(\mathrm{f})}$ (see Annex 6.C2 in Supplementary materials). The PWPD formulation can easily include extended definitions on how distances should be computed. For example, switching from a complete description of the focal habitat graph topology to a planar one, or from excluding intra-patch distance to also including it (Foltête et al., 2012), simply requires an appropriate definition of the corresponding probability of connectivity to be adopted, without needing any modification to the PWPD mathematical structure.

2.5. Final model workflow and data available for demonstration

The data and information flow for computing the proposed harmonized set of indices into a concise modelling framework is illustrated in Fig. 3. Directed edges show the chain of data-transformations (de Rigo, 2012c) connecting the initial and derived – intermediate and final – data layers (nodes). The landscape mosaic pattern model was computed using a circular filter (moving windows) operating the sum of the three landscape units and calculating their proportion. This spatial analysis may be implemented in two equivalent ways. Using GIS tools, ESRI ArcGIS 10.0 provides the "Focal Statistics" algorithm which was easy to automate with Python scripting (Van Rossum and Drake, 2011). The free software GRASS GIS (Steiniger and Hay, 2009; Neteler et al., 2012) is also able to operate the mosaic analysis with the module "r.neighbours" (Shapiro, 2010). Using array programming tools, a two-dimensional convolution with a circular constant kernel and subsequent reclassification may ease reproducibility and reusability by means of concise codelets. This second approach is presented in Supplementary materials with the use of semantic array programming (de Rigo, 2012b,c).

The connectivity dispersal distance, which represents a value of movement cost through different habitats, was obtained using least-cost path method with ESRI algorithms "Cost Distance" and "Path Distance" in ArcGIS 10.0 (ESRI, 2011a,b). GRASS GIS has "r.cost" (Awaida and Westervelt, 2011) and "r.drain" (Miller et al., 2004) modules which are equivalent algorithms for deriving least-cost paths.

At this stage, the integration of the whole set of data-transformations into one single toolbox was not possible solely because two steps are processed with stand-alone software (GUIDOS and Conefor). This would be easily solved as discussed in

[8] See Annex 6.C1 in Supplementary materials.
[9] A simplified formulation of the PWPD family (s-PWPD) is proposed in Estreguil et al. (2012). It only considers the landscape unit normalization as in Eq. (2a), Annex 6.C1 in Supplementary materials.

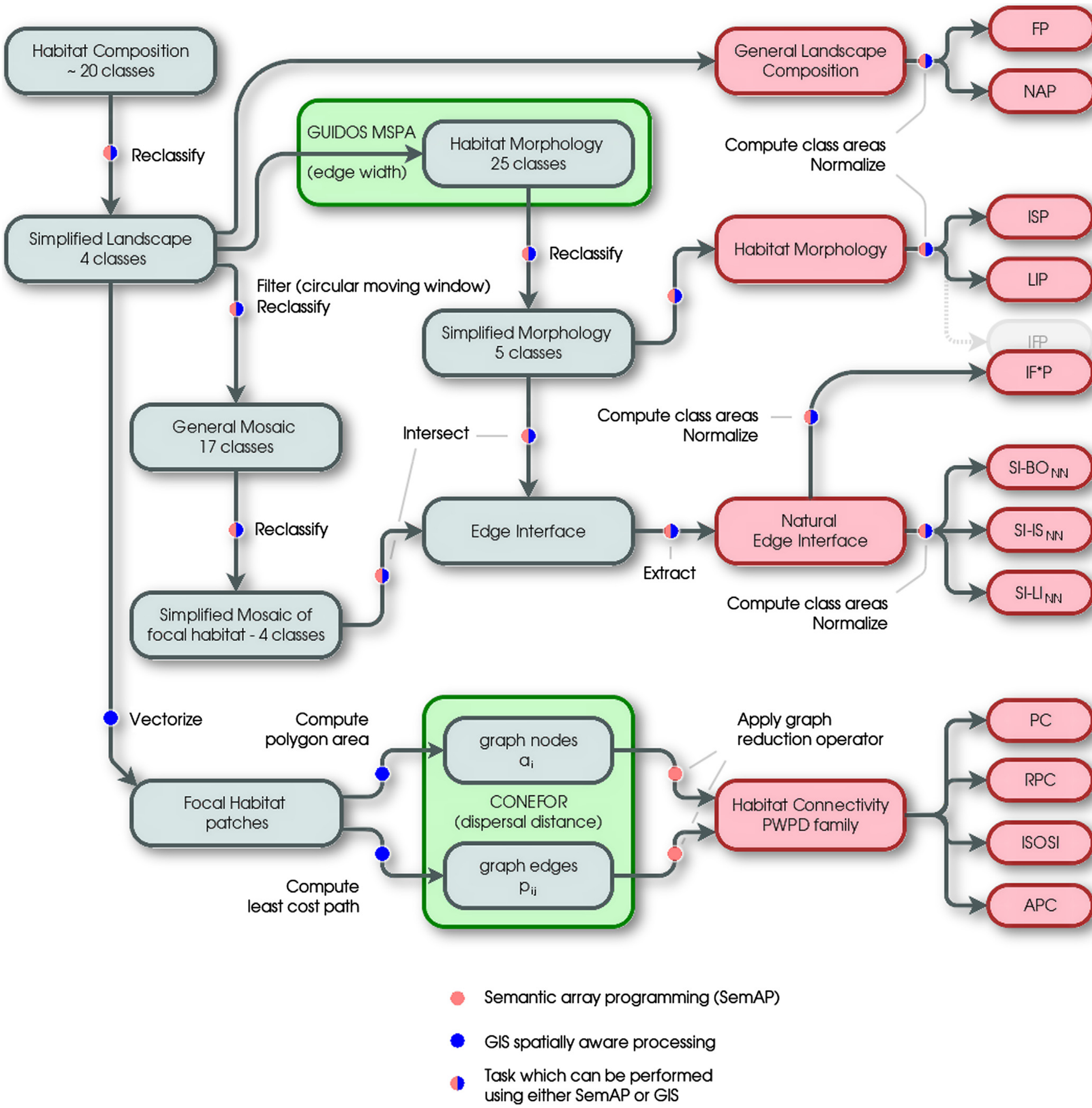


Fig. 3. The data and information flow for computing the proposed harmonized set of indices. The integration with existing software is highlighted. The specific role of Conefor and GUIDOS MSPA is underlined, while the general role of GIS tools and semantic array programming (SemAP) are annotated so to ease discriminating where tasks may be performed equivalently using GIS or SemAP. Most of these tasks are straightforward operations such as reclassification and extraction. Some others less trivial and related to assessing the General/Simplified Mosaic are summarized in Supplementary materials.

the last section of this paper. For the sake of reproducibility, GRASS GIS would be an alternative to ArcGIS which is a proprietary, closed-source tool.

The computation of the models (Fig. 3) was demonstrated on sixty-five samples of 1 km^2 which offered harmonized habitat maps from the EBONE European project. Each map was available as seamless vector layer with 400 m^2 MMU and based on a common classification system into General Habitat Categories (GHCs) (Bunce et al., 2008). GHCs are organized in five super-categories i.e. whether the land element is 'Urban/artificial' (URB), 'Cultivated' (CUL), 'Sparsely vegetated' (SPV) 'Herbaceous' (HER), 'Trees or Shrubs' (TPS). The last three GHC categories are further detailed as grassland, shrub and tree species based on life forms and height. In the TPS class, non-forest phanerophytes habitats with a height varying from 0.3 m to 5 m are discriminated from forest phanerophyte habitats (FPH) which are trees above 5 m height. The later class (FPH) was selected as the focal habitat class (*f*) for this study. Tall phanerophytes (TPH) were also used in the connectivity model.

The GHC's vector maps of all available samples were rasterised at 1 m spatial resolution, For the morphological pattern model, they were reclassified into a binary raster layer forest phanerophytes (FPH)-non-forest. For the landscape mosaic pattern model, the GHCs raster maps were reclassified into a tri-dimensional raster layer according to three main habitat types used as proxies of land use intensity, namely natural/semi-natural (trees/shrubs TPS; herbaceous HER; sparsely vegetated SPV), cultivated (CUL) and urban/artificial types (URB) (Fig. 2). The width of edge habitats (edge depth) vary greatly among species but in forestry, the edge width is generally related to the height and structure of the forest and ranges from narrow (20 m) to wide (160 m) according to Franklin and Forman (1987). The edge width in

Table 6
Indices applied to two samples in Austria (squares codes: AU113 and AU331).

	Indices	AU113	AU331
General	FP	0.627	0.567
	NAP	0.737	0.977
Morphology	IFP	0.642	0.466
	IF*P	0.696	0.688
	ISP	0.041	0.028
	LIP	0.054	0.148
Interface	$SI\text{-}BO_{NN}$	0.204	0.621
	$SI\text{-}IS_{NN}$	0.088	0.504
	$SI\text{-}LI_{NN}$	0.039	0.657
Connectivity	PC	0.349	0.313
	RPC	0.591	0.559
	IsoSi	0.499	0.549
	APC	0.784	0.967

the morphological pattern model as well as the disk radius for the neighbourhood in the landscape mosaic model (s) was set at 25 m.

To compute the four indices of the connectivity model component, the effective inter-patch distance was selected. Demonstration was made for a 500 m average dispersal ability due to the sample size[10] (1 km^2) and review of most frequent upper limit of distance thresholds in seven types of dispersal mode[11] (Vittoz and Engler, 2007). Costs of movement (friction f) were assigned to every habitat types using a logarithmic increment values from forest/tall trees types of habitats (FPH and TPH were allocated lowest friction 1) to urban habitats (highest friction 10.000). The cost distance matching the 50% probability ($cost_{d50\%}$) corresponded to the average dispersal distance ($d_{50\%} = 500$ m) multiplied by the average friction per distance unit (avf). The average friction was set at half a logarithmic scale of frictions, being from 1 to 10,000 ($avf = 100$).

3. Results

The workflow in Fig. 3 for the four-family set of indices (Tables 2–5) was applied to the focal forest phanerophyte (FPH) habitat (f) in the EBONE database. For clarity, the fixed attributes (f, nat, s, avf, d) will be omitted in labelling each index. Analysis units are the 1 km^2 samples, which represent the landscape scale.

3.1. Models implementation for a forest habitat pattern characterisation

Section 3.1 wishes to illustrate the implementation of the modelling approach for two squares. The implementation of the indices over the whole set of sixty-five squares can partly be found in Estreguil et al. (2011).

First, the share of natural/semi-natural habitats (NAP) and the share of the forest habitat in each sample were easily computed. Table 6 provides the values for two samples which have similar forest proportion but different forest distributions and non-forest landscape matrix (Fig. 4: AU331 has a predominant natural landscape with 85 forest nodes while AU113 is a landscape mosaic of natural, urban and cultivated habitats with 33 forest nodes). These two samples are well suited to illustrate the benefit of combining the models, which is to characterise simultaneously the structure (morphology), context (edge interface) and isolation (connectivity) of the forest patches.

Second, for the morphological pattern model component, the binary raster layer forest (FPH)-non-forest was processed with GUIDOS using a narrow forest edge width (s equal to 25 m). The output morphological pattern map has twenty-five forest pattern classes further simplified into five classes (interior, boundary, connector, branch, islet) as described in Section 2.2 and illustrated for two samples in Fig. 4 (top). Their forest shares including also the one for the linear features were calculated (Fig. 5, left and Table 6).

Third, the edge interface pattern model component was run in two steps. The tri-dimensional habitat raster maps (natural, cultivated and urban) were processed to generate the landscape mosaic pattern map (Fig. 4, bottom) where the immediate neighbourhood (disk radius of 25 m for an area of nearly 0.2 ha) around each square metre of land was characterised according to the seventeen mosaic classes. The non-forest classes were masked and the seventeen mosaic classes were aggregated into four main mosaic pattern types ('core natural' NN, 'mainly natural' N, 'mixed natural' MN, or 'some natural' SN) as described in Section 2.3 and illustrated for two samples (AU331 and AU113) in Fig. 4 (bottom). Their forest shares are provided in Fig. 5 (right). Their respective forest mosaic pattern maps in Fig. 5 enable us to visualize differences in forest edge interfaces: natural – 'soft' – type of interfaces (NN at edges), or more 'hard' type of interfaces (mixed natural MN or some natural SN at edges) where forest abut cultivated and/or urban habitats.

The second automatic processing step was on combining the morphological and mosaic pattern maps to provide the edge interface context (natural or artificial) per morphological shapes of edges and deliver the similarity index (Table 6). The differences of the two samples in terms of forest morphology and edge interface can be picked up from the two model components and are illustrated in Fig. 6 and Table 6. Edge interfaces in sample AU331 are more natural for all edge morphologies (boundaries, branches, connectors and islets) than in AU113 (SN shares are very low in Fig. 6 and $SI\text{-}BO_{NN}$ much higher in AU331 (20.4% and 62.1% respectively)). AU331 has however a more fragmented forest landscape morphology (more linear features as connectors and branches, more boundaries, more islets). AU113 has a higher interior forest amount (IFP) distributed in fewer and larger patches than AU331 (Table 6): 63% distributed in nine patches with an average size of 4.4 ha against 46% in seventeen patches with an average size of 1.5 ha. In AU113, forest edges are more exposed to anthropogenic habitats (35.9% compared to 13.3% for $SI\text{-}BO_{MN}$). More than half of forest branches that often represent protrusions at edges, have 'hard' edge interfaces in AU113 (46.9% for $SI\text{-}BR_{SN}$, while only 0.9% in AU331). The natural surroundings of interior forest patches in AU331 also explain the value of IF*P which is significantly higher than IFP.

Last, but not least, forest connectivity indices from the PWPD family (PC, RPC, IsoSi, APC) were calculated with a fixed average dispersal distance of 500 m for an average friction of 100 and habitat friction values that are detailed in Section 2.4 (Table 6 for the two samples).

When RPC and IsoSi are compared to forest proportion (FP), IsoSi is always lower than RPC due to less intra-patch area weight (only the arrival patch area per pair of patches) as expected. Consequently IsoSi is more sensitive to the inter-patch landscape resistance, and possible barrier effects to dispersal and gives more focus on the probability of species movement. In contrast PC and RPC react better to habitat availability (its intra and inter-connectivity). The APC index does not account for forest availability but its configuration and the matrix resistance. Differences in the APC values across the two samples are clearly highlighted. In AU331 the APC is very close to one, which represents a matrix without movement costs, and reflects well the more natural context in this sample when compared to AU113.

3.2. Towards a standardised set of indices

The question is now on how to report the set of indices to comprehensively describe the samples with independent and

[10] The average distance between pairs of uniformly distributed random points in a 1 km × 1 km cell would be 520 m while the mode would be 480 m (see Annex 3, Fig. 2 of Supplementary materials).

[11] While Vittoz and Engler (2007) focus on seeds dispersal modes, they include information on animal vectors of dispersal (zoochory).

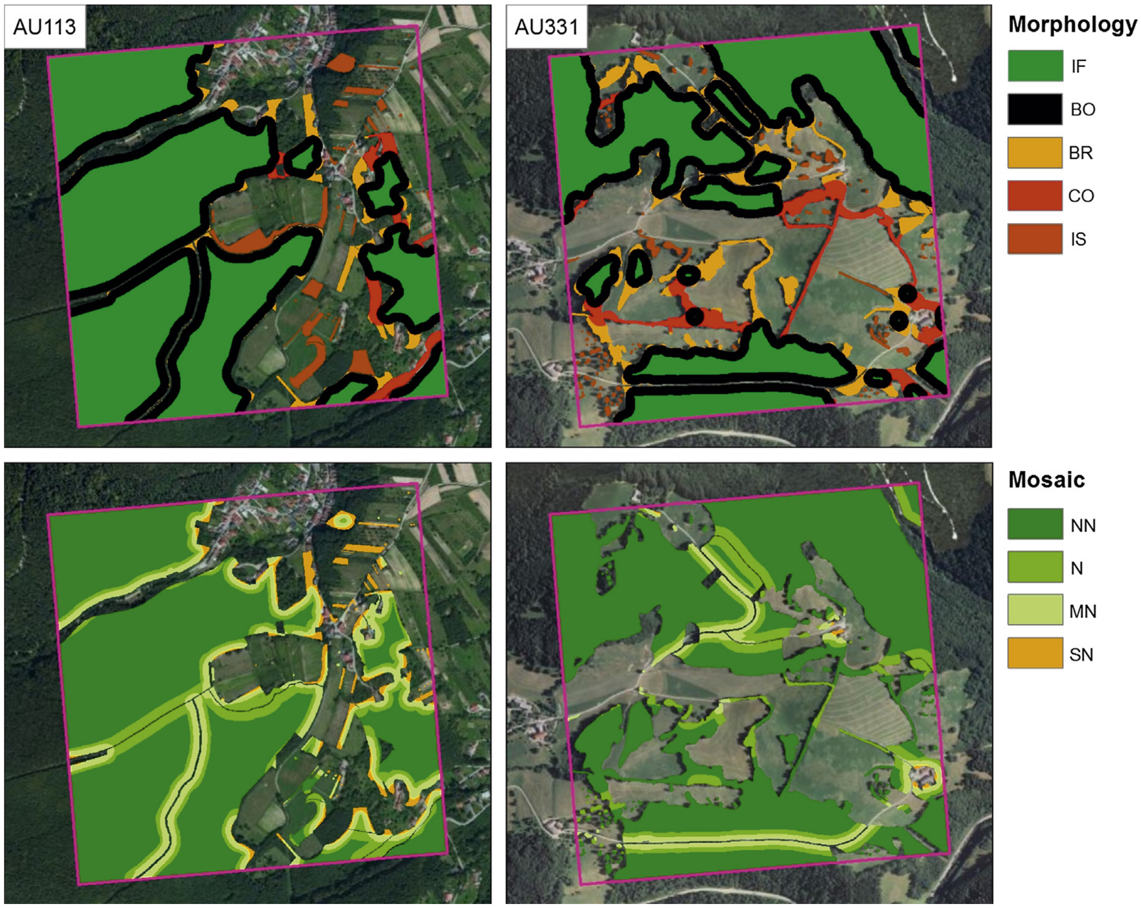


Fig. 4. Morphological (top) and landscape mosaic (bottom) pattern maps for forest for two samples in Austria (squares codes: AU113 and AU331, other squares illustrated in Estreguil et al. (2011)).

fundamental components of pattern and connectivity for a focal habitat. Statistical methods such as factor analysis can be used to organize and reduce the number of indices but the idea is however not to develop composite indexes, nor restrict the analysis to one single preferred measure. As a follow-up of Section 2, the four categories that organize the set of indices will be checked for their redundancy: general landscape indices (NAP, FP), natural edge interface indices (SI-BO_{NN}, SI-IS_{NN}, SI-LI_{NN}), morphological indices (IF*P, ISP, LIP), and connectivity indices (PC, RPC, IsoSi, APC). This reporting choice is oriented towards a natural/semi-natural driven pattern. Alternatively, the reporting choice could be more towards anthropogenic driven pattern with the second category as artificial edge interface indices.

Twelve indices were retained for all samples and their correlation matrix has been computed (it is presented in Annex 2 Table 1 of Supplementary materials). Given the low cardinality of the

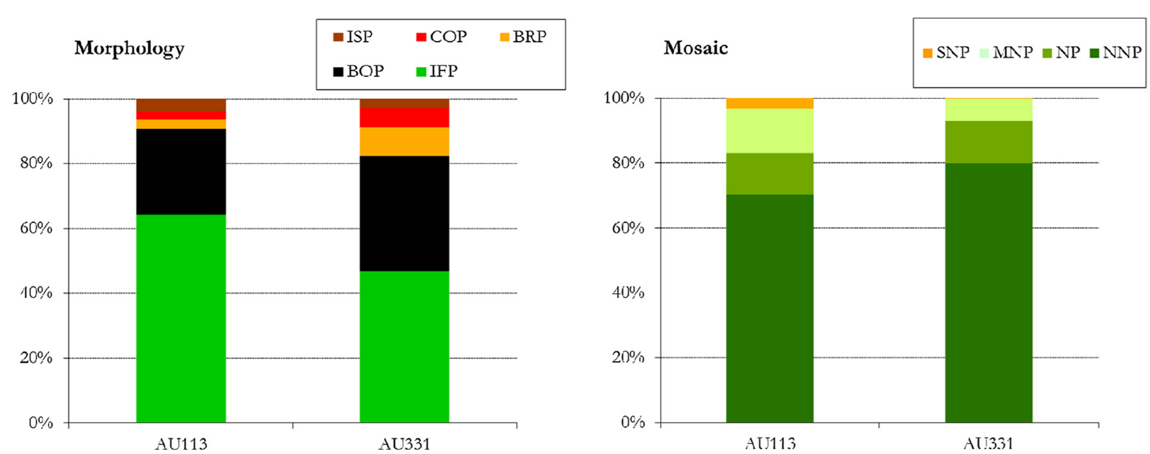


Fig. 5. Forest proportion in each morphological shape class (left) and in each landscape mosaic class (right) for the two samples in Austria (results for the sixty-five samples per biogeographic region in Estreguil et al. (2011)).

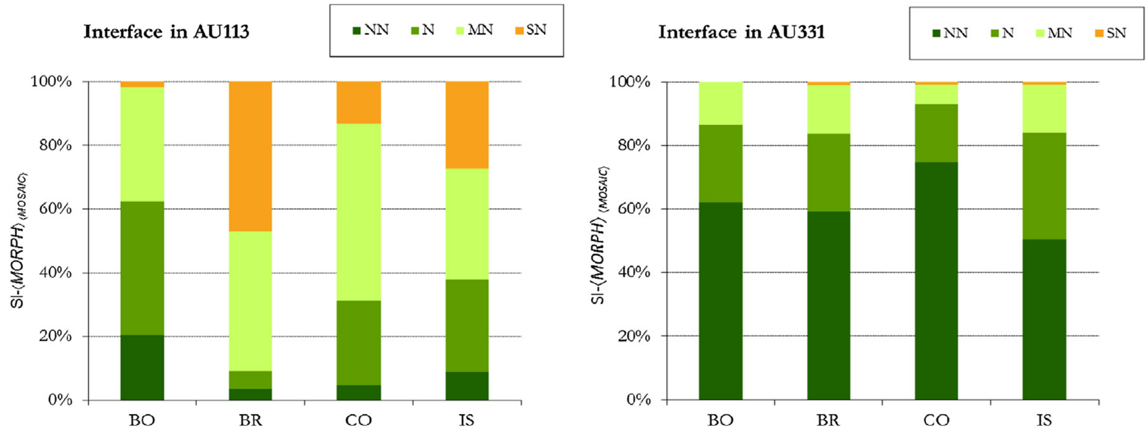


Fig. 6. Edge interfaces of each edge morphological shapes for the two samples in Austria (other samples illustrated in Estreguil et al. (2011)).

dataset, a jack-knife analysis[12] has been carried out to qualitatively assess the sensitivity of the linear correlation statistics. Sixty-five correlation matrices were computed by removing each time a different row of data, corresponding to a different location.

Correlations are generally low or around 50%, apart from three pairs of indices. Within the 'Morphology' set, IF*P and ISP are negatively correlated (ranging in $-0.85 \div -0.79$) due to the fact that forest proportions in interior forest and in islet are interdependent for a selected edge size (the highest IFP, and a fortiori the highest IF*P, the lowest ISP as shown in Annex 3 Fig. 1 of Supplementary materials). As expected, interior forest proportion appears to be correlated to both FP ($0.68 \div 0.71$) and NAP ($0.63 \div 0.69$, see also Annex 2 Table 1 of Supplementary materials). Within the 'interface' set, the correlation for $SI\text{-}BO_{NN}$ and $SI\text{-}LI_{NN}$ ($0.9 \div 0.92$) is more surprising and seems to indicate a correlation between the landscape context of edges and of linear features (see also in Fig. 7). The reference dataset also shows a relatively high correlation (around 0.7) with NAP for the interface indices for edges and linear features, surprisingly not for islets (less than 0.5) suggesting their isolation in less similar adjacent habitats (Annex 2 Table 1 of Supplementary materials).

Correlation is high for FP and IsoSi and RPC as expected (see previous section) and shown in Annex 3 Fig. 1 of Supplementary materials. The correlation between the set of connectivity indices and the forest proportion has also been computed (Annex 2 Table 1 of Supplementary materials). As expected, PC and RPC are highly correlated with FP due to the weight of habitat availability; IsoSi has a lower correlation while APC can be considered as not correlated with any of the indices (Annex 3 Fig. 1 of Supplementary materials). On the basis of those results, we could suggest to select one of the three indices depending on users' needs. When inter-patch connectivity and species movement are at focus, APC may be more appropriate. When fluxes of species (as proxy, species amount depends on patch size) in between patches are as important as the feasibility of movement, IsoSi may be more suited. When intra-patch connectivity is more important than inter-patch connectivity, PC or RPC may be preferred. As shown in Fig. 7, APC does not appear strongly correlated with NAP.

The classical correlation analysis is only able to account for linear relationships among quantities. The twelve indices all range from 0 to 1 and their semantics (the "desirability" of the corresponding habitat pattern) in all cases positively correlates with increasing index's values. Nevertheless, given the complex nonlinear definition of some indices, a generalized nonlinear analysis is recommendable. Nonlinear Principal Component Analysis, NL-PCA (Kramer, 1991; Dong and McAvoy, 1996; Schölkopf et al., 1998) could provide further insights on intrinsic relationships among indices. Hierarchical NL-PCA (Scholz and Vigàrio, 2002; Scholz et al., 2008) is able to constrain a nonlinear decomposition to rank the components' importance similarly to how linear PCA does. Decomposition sensitivity could be assessed via resampling iterations while the discontinuity of decomposition parameterization among iterations could be mitigated by perturbing the parameters selected in previous iterations with general evolutionary techniques. For example, the SIEVE parameter training architecture (Selective Improvement by Evolutionary Variance Extinction, de Rigo et al., 2005) is suitable in iterative nonlinear problems where discontinuities between subsequent iterations may be needed while continuity is generally desirable (de Rigo et al., 2001). Unfortunately, nonlinear decomposition is not unique as linear principal components are. It depends on the nonlinear family chosen for describing the components and on the training methodology for extracting them from the set of indices. The aforementioned mitigations would still preserve a degree of arbitrary choice. This makes the application of NL-PCA (despite a powerful and otherwise recommendable approach) problematic for providing a characterisation of pattern more focused on robustness and generality – also reducing the amount of hypotheses on the structure of indices' relationships – than on explicit derivation of empirical decomposition equations. Furthermore, the relatively small amount of available samples suggests not to attempt to derive explicit empirical equations for either principal components or regressors between the indices. An implicit approach is instead proposed.

In order to account for a wide domain of possible relationships among indices, the linear statistical analysis has been complemented with an analysis of general (nonlinear and even non-monotone) relationships by using Brownian Distance Correlation (Szekely et al., 2007; Szekely and Rizzo, 2009a,b; Lyons, 2011) analysis. Classical Pearson product-moment correlation ranges in $[-1, 1]$ and it is well-known that a null value of that correlation is not a sufficient condition for excluding nonlinear dependencies. Under weak mathematical assumptions,[13] Brownian Distance Correlation (BDC) generalizes (Szekely and Rizzo, 2009a,b; Rémillard, 2009) the idea of correlation by enabling comparisons between multi-dimensional quantities (e.g. categories composed by groups of

[12] See Annex 4, Codelet 2 in Supplementary materials.

[13] See footnote.[5]

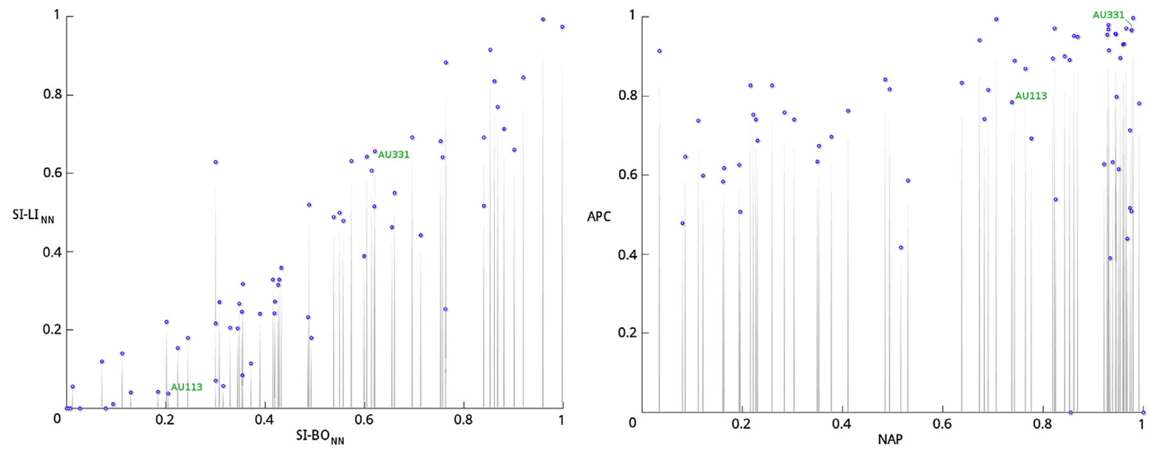


Fig. 7. Scatter charts showing partly unexpected relationships between indices (captured by the proposed correlation analysis). Samples AU113 and AU331 are highlighted: (on the left) correlation between the landscape context of edges SI-BONN and of linear features SI-LINN; (on the right) low correlation between the share of natural/semi-natural habitats (NAP) and the connectivity index APC.

indices) and by providing a statistic in [0, 1] whose value is zero if and only if the analysed quantities are independent. BDC is able to detect nonlinear and nonmonotone dependence. Furthermore, BDC is scale invariant and also rotation invariant (Szekely and Rizzo, 2009a,b).

A jack-knife analysis[14] has been carried out to qualitatively assess the sensitivity of the BDC mono- and multi-dimensional statistics (de Rigo, 2012a), benefiting from concise codelets which follow the semantic array programming paradigm as implemented by the Mastrave modelling library (de Rigo, 2012b,c) within the GNU Octave computing environment (Eaton et al., 2008). The resulting minimum and maximum value of BDC for each couple of indices is in Table 7. For each index, its maximum BDC – either including or excluding the category to which the index belongs – has been summarized to report the highest detected dependence between that index and the others. Less correlated indices are highlighted and ranked (max BDC rank), qualitatively showing the contribution of each index as provider of information not derivable from other indices (see also Table 8). Only this ranking and therefore comparison between estimations with same size is exploited. Also, a certain amount of dependence is theoretically expected among all indices. Therefore, neither testing statistical hypotheses on independence, nor accounting for the sample cardinality of BDC estimations (Szekely and Rizzo, 2009b) is here relevant. In both rankings APC is the least correlated one and is always followed by indices belonging to different categories ($SI\text{-}IS_{NN}$ in the Interface category; NAP in the General category; ISP or LIP in the Morphology category).

The BDC invariance can be extended to cover generic monotone transformations (so that for example PC and RPC indices would be considered equivalent and perfectly correlated, see Annex 2 Table 2 of Supplementary materials, because RPC is a monotonic transformation – the positive square root – of PC). The extension is based on replacing the indices with their rank,[15] leading to the rank-based (Szekely and Rizzo, 2009a; Rémillard, 2009) version of BDC (R-BDC) whose corresponding jack-knife analysis is summarized in Table 10 of Supplementary materials. As for BDC, the maximum R-BDC correlation of each index – either including or excluding the category to which the index belongs – has been summarized.

Less correlated indices are highlighted and ranked, showing that in both rankings APC is the least correlated one and is always followed by indices belonging to different classes ($SI\text{-}IS_{NN}$ in the Interface class; NAP in the General class; ISP or IF*P in the Morphology class). When R-BDC correlations only refer to outside-category indices, IF*P maximum R-BDC (0.743) is slightly less than the maximum R-BDC of NAP (0.746); however the jack-knife variability of the corresponding values is ten times greater.

Since both BDC and R-BDC analyses on the twelve indices corroborated the reasonability of the hypothesized categorisation, a further analysis was performed using the multi-dimensional BDC (Md-BDC) to assess the relationship between isolated indices and subsets of indices grouped by category (Annex 2 Table 3 of Supplementary materials). Euclidean distance was used to compute distances for multi-dimensional groups and the Md-BDC was also computed in its rank-based version (MdR-BDC). For each index, its maximum Md-BDC and MdR-BDC – obviously both computed outside the category to which the index belongs – have been summarized. Again, less correlated indices are highlighted and ranked confirming that in both rankings APC is the least correlated one, always followed by $SI\text{-}IS_{NN}$ in the Interface category and ISP in the Morphology category. The less Md-BDC-correlated index of the General category remains NAP, always preceded by two Interface indices.

Concluding this section with the synoptic Table 8, it is interesting to note that APC is always selected and always followed by one index belonging to different categories, thus confirming the proposed categorisation. However, it is advisable to critically assess what is the overall meaning of the analysis on less BDC-correlated indices.

In principle, an index defined as pure white noise would be highlighted as the one providing "information" completely not available from other indices (so replacing the role here played by APC). It should therefore be evident that the proposed statistical analysis is able to detect loosely coupled indices, irrespective of their possibly nonlinear or even nonmonotone relationship with other indices. However, it is not able to assess the "reasonableness" of these indices. A semantic analysis is always required to corroborate these highlights by discussing the soundness and relevance of both indices and categories of indices (as discussed in Section 2.4 for the selection of the most appropriate connectivity index depending on the interest of the user).

[14] See Annex 4, Codelet 3 in Supplementary materials.

[15] The rank of a vector $X = [X_1 \dots X_n]$ is a permutation $P = \text{rank}(X)$ of the integers 1 … n such that $X(P) = [X_{P1} \dots X_{Pn}]$ is monotonically non-decreasing. In case of repeated values within X, ties in rank should be broken randomly (Szekely and Rizzo, 2009a) for a correct application of R-BDC (so that in this case rank(X) is not unique). See Annex 4, Codelet 3 in Supplementary materials for further details.

Table 7
Variability (jack-knife extremes) of the Brownian Distance Correlation matrix (BDC) between the twelve indices. Distance correlation greater than 0.8 (green cells) and 0.6 (light green cells) is highlighted, as well as distance correlation less than 0.3 (red higher values). For each index, its maximum BDC – either including or excluding the category to which the index belongs – has been summarized and indices have been ranked (max BDC rank) from the less correlated one. The same rank is assigned to indices for which the range of variability of the corresponding maximum BDC overlaps for more than 50%.[16]

Brownian Distance Correlation (BDC)		General		Connectivity				Interface			Morphology		
		NAP	**FP**	**PC**	**RPC**	**ISOSI**	**APC**	**SIBO$_{nn}$**	**SIIS$_{nn}$**	**SILI$_{nn}$**	**IF*P**	**ISP**	**LIP**
General	**NAP**	1	0.602 0.643	0.550 0.579	0.599 0.639	0.571 0.609	0.391 0.434	0.683 0.719	0.428 0.481	0.673 0.704	0.606 0.655	0.405 0.463	0.431 0.480
	FP	0.602 0.643	1	0.965 0.968	0.995 0.996	0.900 0.939	0.309 0.359	0.298 0.339	0.244 0.282	0.296 0.348	0.698 0.726	0.592 0.630	0.245 0.271
Connectivity	**PC**	0.550 0.579	0.965 0.968	1	0.970 0.973	0.864 0.921	0.302 0.360	0.259 0.282	0.186 0.215	0.220 0.261	0.594 0.617	0.496 0.523	0.223 0.266
	RPC	0.599 0.639	0.995 0.996	0.970 0.973	1	0.922 0.959	0.335 0.390	0.281 0.322	0.219 0.259	0.272 0.325	0.684 0.711	0.579 0.616	0.224 0.258
	ISOSI	0.571 0.609	0.900 0.939	0.864 0.921	0.922 0.959	1	0.524 0.571	0.272 0.317	0.273 0.319	0.262 0.315	0.586 0.618	0.499 0.535	0.200 0.217
	APC	0.391 0.434	0.309 0.359	0.302 0.360	0.335 0.390	0.524 0.571	1	0.324 0.369	0.426 0.474	0.274 0.314	0.185 0.217	0.202 0.232	0.361 0.409
Interface	**SIBO$_{nn}$**	0.683 0.719	0.298 0.339	0.259 0.282	0.281 0.322	0.272 0.317	0.324 0.369	1	0.600 0.645	0.882 0.911	0.455 0.513	0.257 0.349	0.744 0.764
	SI-IS$_{nn}$	0.428 0.481	0.244 0.282	0.186 0.215	0.219 0.259	0.273 0.319	0.426 0.474	0.600 0.645	1	0.583 0.631	0.310 0.357	0.233 0.285	0.567 0.657
	SI-LI$_{nn}$	0.673 0.704	0.296 0.348	0.220 0.261	0.272 0.325	0.262 0.315	0.274 0.314	0.882 0.911	0.583 0.631	1	0.505 0.548	0.319 0.380	0.740 0.768
Morphology	**IF*P**	0.606 0.655	0.698 0.726	0.594 0.617	0.684 0.711	0.586 0.618	0.185 0.217	0.455 0.513	0.310 0.357	0.505 0.548	1	0.824 0.878	0.250 0.282
	ISP	0.405 0.463	0.592 0.630	0.496 0.523	0.579 0.616	0.499 0.535	0.202 0.232	0.257 0.349	0.233 0.285	0.319 0.380	0.824 0.878	1	0.223 0.255
	LIP	0.431 0.480	0.245 0.271	0.223 0.266	0.224 0.258	0.200 0.217	0.361 0.409	0.744 0.764	0.567 0.657	0.740 0.768	0.250 0.282	0.223 0.255	1
max BDC rank		0.719 3rd	0.996	0.973	0.996	0.959	0.571 1st	0.911	0.657 2nd	0.911	0.878 5th	0.878 5th	0.768 4th
max BDC rank (outside category)		0.719 4th	0.996	0.968	0.996	0.939	0.474 1st	0.764	0.657 2nd	0.768	0.726 4th	0.630 2nd	0.768

4. Discussion and future developments

This paper is of particular relevance in the context of biodiversity policy reporting on habitat pattern, fragmentation and connectivity. Potential (non-expert) users are either from landscape planning or environmental local, regional, national or international agencies. This paper presents a generic, reproducible and concise characterisation of habitat patterns which is based on a harmonized mathematical description classifying known indices while suggesting new ones as logical complement. It also promotes the derivation of a small, simple and integrated set of indices, easily customised depending on user focus and semantics. The unambiguous, easy computability and reproducibility of the indices are ensured by the integrated use of publicly available software. This effort could be considered in line and complementary to recent 'computation oriented' ones on integrating available landscape measures' calculations, as for example the extensible library by Zaragozí et al. (2012) (Land-metrics Do It Yourself). Similar tools could ease exploiting new measures and methods such as the ones proposed in this paper. At the same time, our proposed methodology for robustly identifying less correlated measures and families of measures could be applied to any set of numerical indices satisfying quite general mathematical assumptions.

The proposed system of pattern characterisation refers to four main families of indices computed in one unique modelling frame: general landscape composition, focal habitat morphology, edge interface and connectivity. The three available pattern model components were revisited to present new indices, programmed when necessary and automated for large data processing. The combination of two model components (GUIDOS/MSPA application and Landscape Mosaic model) is a new tool to provide the edge interface context of identified morphological shapes (edge of interior habitat, connecting linear features and physically isolated islets). The habitat morphology and edge interface index family requires the detection of connecting pathways, a feature which is only available in MSPA and not from traditional patch area and edge measures. Furthermore, a new family of connectivity indices derived from the Conefor Sensinode PC index was proposed as a *Power Weighted Probability of Dispersal* (PWPD) function to allow a weighted sensitivity of connectivity measures to the matrix permeability (proxy). The model components are spatially explicit and only require few simple – still essential – input variables (edge size, dispersal distance, resistance values for habitat matrix), rendering them generic enough and easily adaptable to a variety of focal habitats. While our study demonstrates the feasibility of characterising and measuring habitat pattern, still several critical challenges remain to be addressed in future research. The scaling behaviour of indices should be examined for several Euclidian disk radius (representing both the edge size in habitat morphology and

[16] The overlap percentage between the maximum BDC of two indices simply considers the intervals within the extremes of variability of the corresponding maximum BDC. It is computed by comparing the intersection between the two intervals against the smaller interval. With larger datasets, as far as not only the variability range but even the information provided by the empirical distribution of jack-knife estimations might be considered of interest, nonparametric tests such as the Mann–Whitney *U* test may be preferable.

Table 8
Ranking of maximum BDC family of statistics for each index. The family of statistics includes BDC and rank-based BDC (R-BDC) – either including or excluding the category to which each index belongs – and multi-dimensional BDC (Md-BDC, also including the rank-based version MdR-BDC) between each index and the proposed categories.

Maximum BDC family	General		Connectivity				Interface			Morphology		
	NAP	FP	PC	RPC	ISOSI	APC	$SIBO_{NN}$	$SIIS_{NN}$	$SILI_{NN}$	IF*P	ISP	LIP
max BDC rank	3rd 0.719	11th 0.996	10th 0.973	11th 0.996	9th 0.959	1st 0.571	7th 0.911	2nd 0.657	7th 0.911	5th 0.878	5th 0.878	4th 0.768
max BDC rank (outside category)	4th 0.719	11th 0.996	10th 0.968	11th 0.996	9th 0.939	1st 0.474	6th 0.764	2nd 0.657	6th 0.768	4th 0.726	2nd 0.630	6th 0.768
max R-BDC rank	3rd 0.746	10th 0.995	11th 1.000	11th 1.000	9th 0.968	1st 0.580	7th 0.916	2nd 0.687	7th 0.916	4th 0.840	4th 0.840	4th 0.854
max R-BDC rank (outside category)	4th 0.746	10th 0.995	10th 0.995	10th 0.995	9th 0.955	1st 0.550	7th 0.854	3rd 0.687	6th 0.834	4th 0.743	2nd 0.649	7th 0.854
max Md-BDC rank (outside category)	5th 0.708	12th 0.921	9th 0.779	11th 0.817	8th 0.762	1st 0.453	5th 0.708	2nd 0.488	4th 0.680	5th 0.721	3rd 0.561	9th 0.794
max MdR-BDC rank (outside category)	4th 0.710	12th 0.941	10th 0.912	10th 0.912	8th 0.862	1st 0.497	4th 0.698	2nd 0.577	6th 0.720	7th 0.764	3rd 0.626	9th 0.873

the size of the immediate surroundings in the edge interface context) and for different spatial resolutions of the input map. This point was partially addressed in Ostapowicz et al. (2008). There is also the need to conduct 'neutral model' analysis to set standards for comparing and interpreting the patterns identified by the indices. For example, Riitters et al. (2007) analysed the behaviour of MSPA based habitat morphological shapes on randomly generated habitat maps according to the proportion and morphological configuration of the focal class. Neutral models should be also applied by controlling the composition of the non-focal landscape. Regarding connectivity, we know that the connections between habitat patches are best characterised through a probabilistic model like the probability of connectivity from Saura and Torné (2009), in which there is a certain probability of dispersal among habitat patches, typically modelled as a decreasing function of inter-patch Euclidean or effective distance. The performance and flexibility that offer the PC derived PWPD function should be evaluated against other available indices like fluxes, dispersal success and cell immigration indices similarly as in Saura and Pascual-Hortal (2007).

Future work is needed on implementing all the computation steps of the modelling approach in one toolbox. Full integration is not yet achieved since two steps are processed with stand-alone software which run only via graphical user interface (GUI). The access of GUIDOS and Conefor libraries by Python should be addressed: it would allow data to be processed by command line and a one-step process to be developed.

The integrated modelling exercise was demonstrated with twelve indices applied to sixty-five habitat maps for reporting on the pattern of the forest phanerophyte habitat in European landscapes. The pattern characterisation was corroborated by an in depth statistical jack-knife analysis based on classical linear correlation and nonlinear Brownian Distance Correlation (BDC, R-BDC, Md-BDC, MdR-BDC). This statistical correlation analysis constitutes an alternative to the more traditionally used Principal Component Analysis. The results highlight the less correlated and the fundamental pattern components, corroborating the hypothesized hierarchical organization of a final set of twelve indices into four categories, and also the feasibility of reducing further the number of indices within each category.

After implementation of the models and indices and to add visibility to reporting, the sample based spatial pattern data layers and attributes (indices) were prepared according to a defined common (ESRI shape) data structure and can now be accessed through a web-based mapping client that has been developed (Estreguil et al., 2011). The map viewer allows the user to view the location of EBONE field based samples, to view habitat maps, to query the presence and extent of habitats per sample and per environmental zones, to view habitat pattern maps and related indices on morphology, edge interface and connectivity. For each available sample, the environmental site conditions and management style will be soon available as well as multi-temporal fine-scale habitat maps that would enable to partly address the behaviour or response of indices to spatiotemporal variations.

The proposed pattern characterisation picks-up relevant landscape specificities which are not restricted to basic patch area measures such as in Krauss et al. (2010). It is expected to contribute to studies across temporal and spatial scales on pattern–process relationships and would allow their comparison across regions. We would suggest applying it to targeted habitats that are considered at risk from fragmentation, and particularly vulnerable when combined with effects of climate change as suggested by Kettunen et al. (2007).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.envsoft.2013.10.011.

References

Awaida, A., Westervelt, J., 2011. GRASS GIS manual: r.cost. In: GRASS GIS 6.4.3svn Reference Manual. http://grass.osgeo.org/grass64/manuals/r.cost.html.

Bogaert, J., 2003. Lack of agreement on fragmentation metrics blurs correspondence between fragmentation experiments and predicted effects. Conserv. Ecol. 7 (1). http://www.ecologyandsociety.org/vol7/iss1/resp6/.

Bogaert, J., van Hecke, P., Salvador-Van Eysenrode, D., Impens, I., 2000. Landscape fragmentation assessment using a single measure. Wildl. Soc. Bull. 28 (4), 875–881. http://www.jstor.org/stable/3783842.

Bucki, M., Cuypers, D., Mayaux, P., Achard, F., Estreguil, C., Grassi, G., 2012. Assessing REDD+ performance of countries with low monitoring capacities: the matrix approach. Environ. Res. Lett. 7. http://dx.doi.org/10.1088/1748–9326/7/1/014031, 014031.

Bunce, R.G.H., Metzger, M.J., Jongman, R.H.G., Brandt, J., de Blust, G., Elena-Rossello, R., Groom, G.B., Halada, L., Hofer, G., Howard, D.C., Kovar, P., Mucher, C.A., Padoa-Schioppa, E., Paelinx, D., Palo, A., Perez-Soba, M., Ramos, I.L., Roche, P., Skanes, H., Wrbka, T., 2008. A standardized procedure for surveillance and monitoring European habitats and provision of spatial data. Landscape Ecol. 23 (1), 11–25. http://dx.doi.org/10.1007/s10980-007-9173-8.

Clerici, N., Vogt, P., 2013. Ranking European regions as providers of structural riparian corridors for conservation and management purposes. Int. J. Appl. Earth Observ. Geoinf. 21, 477–483. http://dx.doi.org/10.1016/j.jag.2012.07.001.

Convention on Biological Diversity, 2010. COP 10 decision X/2 – strategic plan for biodiversity 2011–2020. In: Tenth Meeting of the Conference of the Parties to the Convention on Biological Diversity, 18–29 October 2010-Nagoya, Aichi Prefecture, Japan. Aichi Biodiversity Targets. http://www.cbd.int/sp/targets/ (strategic goal B, target 5).

de Rigo, D., 2012a. Detecting General Multi-Dimensional Nonlinear Correlations: the Module "dist_corr" of the Mastrave Modelling Library. Mastrave Project Technical Report http://dx.doi.org/10.6084/m9.figshare.92645 http://mastrave.org/doc/mtv_m/dist_corr.

de Rigo, D., 2012b. Semantic Array Programming with Mastrave – Introduction to Semantic Computational Modelling. The Mastrave Project http://mastrave.org/doc/MTV-1.012-1.

de Rigo, D., 2012c. Semantic array programming for environmental modelling: application of the Mastrave library. In: Seppelt, R., Voinov, A.A., Lange, S., Bankamp, D. (Eds.), 6th International Congress on Environmental Modelling and Software, (iEMSs 2012) "Managing Resources of a Limited Planet", ISBN 978-88-9035-742-8, pp. 1167–1176. http://www.iemss.org/iemss2012/proceedings/D3_1_0715_deRigo.pdf.

de Rigo, D., 2013. Behind the horizon of reproducible integrated environmental modelling at European scale: ethics and practice of scientific knowledge freedom. F1000 Research. To appear as discussion paper.

de Rigo, D., Castelletti, A., Rizzoli, A.E., Soncini-Sessa, R., Weber, E., 2005. A selective improvement technique for fastening Neuro-Dynamic Programming in Water Resources Network Management. In: Proceedings of the 16th IFAC World Congress, vol. 16 (1), ISBN 978-3-902661-75-3, pp. 7–12. http://dx.doi.org/10.3182/20050703-6-CZ-1902.02172. IFAC-PapersOnline, ISSN: 1474-6670.

de Rigo, D., Rizzoli, A.E., Soncini-Sessa, R., Weber, E., Zenesi, P., 2001. Neuro-dynamic programming for the efficient management of reservoir networks. In: Proceedings of MODSIM 2001, International Congress on Modelling and Simulation, vol. 4, ISBN 0-867405252, pp. 1949–1954. http://dx.doi.org/10.5281/zenodo.7481.

Dennis, R.L.H., Shreeve, T.G., Van Dyck, H., 2003. Towards a functional resource-based concept for habitat: a butterfly biology viewpoint. Oikos 102 (2), 417–426.

Dong, D., McAvoy, T.J., 1996. Nonlinear principal component analysis – based on principal curves and neural networks. Comp. Chem. Eng.. ISSN: 0098-1354 20 (1), 65–78. http://dx.doi.org/10.1016/0098-1354(95)00003-K.

Eaton, J.W., Bateman, D., Hauberg, S., 2008. GNU Octave Manual Version 3. Network Theory Limited, ISBN 978-0-9546120-6-1.

ESRI, 2011a. ArcGIS 10.0 Desktop Help Library: How Cost Distance Tools Work. http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_the_cost_distance_tools_work/009z00000025000000/.

ESRI, 2011b. ArcGIS 10.0 Desktop Help Library: How The Path Distance Tools Work. http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_the_path_distance_tools_work/009z00000027000000/.

Estreguil, C., Caudullo, G., de Rigo, D., Whitmore, C., San-Miguel-Ayanz, J., 2012. Reporting on European forest fragmentation: standardized indices and web map services. IEEE Earthzine 5 (2), 384031+, 2nd quarter theme: Forest Resource Information http://www.earthzine.org/?p=384031.

Estreguil, C., Caudullo, G., Whitmore, C., 2011. Habitat Landscape Pattern and Connectivity Indices; Used at Varying Spatial Scales for Harmonized Reporting in the EBONE Project. Alterra Report 2297. 80 p. ISSN 1566-7197. Alterra, Wageningen http://edepot.wur.nl/214000.

European Commission, 2011. Our Life Insurance, Our Natural Capital: An EU Biodiversity Strategy to 2020. COM(2011) 244 final http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0244:FIN:EN:pdf.

European Environment Agency, 2012. Streamlining European Biodiversity Indicators 2020: Building a Future on Lessons Learnt From the SEBI 2010 Process. Technical report No11/2012 http://www.eea.europa.eu/publications/streamlining-european-biodiversity-indicators-2020.

Fischer, J., Lindenmayer, D.B., 2007. Landscape modification and habitat fragmentation: a synthesis. Global Ecol. Biogeogr. 16 (3), 265–280. http://dx.doi.org/10.1111/j.1466-8238.2007.00287.x.

Foltête, J.-C., Clauzel, C., Vuidel, G., 2012. A software tool dedicated to the modelling of landscape networks. Environ. Model. Softw. 38, 316–327. http://dx.doi.org/10.1016/j.envsoft.2012.07.002.

Forest Europe, UNECE and FAO, 2011. State of Europe's forests 2011. Status and trends in sustainable forest management in Europe. In: Ministerial Conference on the Protection of Forests in Europe, ISBN 978-82-92980-05-7. http://www.foresteurope.org/eng/State_of_Europes_Forests_Report_2011/Report/.

Franklin, J.F., Forman, R.T.T., 1987. Creating landscape patterns by forest cutting: ecological consequences and principles. Landsc. Ecol 1 (1), 5–18. http://dx.doi.org/10.1007/BF02275261.

Hanski, I., 1994. Patch-occupancy dynamics in fragmented landscapes. Trends Ecol. Evol. 9 (4), 131–135. http://dx.doi.org/10.1016/0169-5347(94)90177-5.

Hanski, I., 1998. Metapopulation dynamics. Nature 396, 41–49. http://dx.doi.org/10.1038/23876.

Herzog, F., Lausch, A., Muller, E., Thulke, H.H., Steinhardt, U., Lehmann, S., 2001. Landscape metrics for assessment of landscape destruction and rehabilitation. Environ. Manage 27 (1), 91–107. http://dx.doi.org/10.1007/s002670010136.

Ince, D.C., Hatton, L., Graham-Cumming, J., 2012. The case for open computer programs. Nature 482 (7386), 485–488. http://dx.doi.org/10.1038/nature10836.

Joppa, L.N., McInerny, G., Harper, R., Salido, L., Takeda, K., O'Hara, K., Gavaghan, D., Emmott, S., 2013. Troubling trends in scientific software use. Science 340 (6134), 814–815. http://dx.doi.org/10.1126/science.1231535.

Kettunen, M., Terry, A., Tucker, G., Jones, A., 2007. Guidance on the Maintenance of Landscape Features of Major Importance for Wild Flora and Fauna – Guidance on the Implementation of Article 3 of the Birds Directive (79/409/EEC) and Article 10 of the Habitats Directive (92/43/EEC). Institute for European Environmental Policy (IEEP), Brussels, 114 p. and Annexes. Report on the review of measures to promote connectivity. EC Project 'Guidelines: Adaptation, Fragmentation' ENV.B.2/ETU/2006/0042r.

Kindlmann, P., Burel, F., 2008. Connectivity measures: a review. Landscape Ecol. 23 (8), 879–890. http://dx.doi.org/10.1007/s10980-008-9245-4.

Koper, N., Schmiegelow, F.K.K., 2006. A multi-scaled analysis of avian response to habitat amount and fragmentation in the Canadian dry mixed grass prairie. Landsc. Ecol 21 (7), 1045–1059. http://dx.doi.org/10.1007/s10980-006-0004-0.

Kramer, M.A., 1991. Nonlinear principal component analysis using autoassociative neural networks. AIChE J.. ISSN: 1547-5905 37 (2), 233–243. http://dx.doi.org/10.1002/aic.690370209.

Krauss, J., Bommarco, R., Guardiola, M., Heikkinen, R.K., Helm, A., Kuussaari, M., Lindborg, R., Ockinger, E., Meelis, M., Pino, J., Poyry, J., Raatikainen, K.M., Sang, A., Stefanescu, C., Teder, T., Zobel, M., Steffan-Dewenter, S., 2010. Habitat fragmentation causes immediate and time delayed biodiversity loss at different trophic levels. Ecol. Lett. 13 (5), 597–605. http://dx.doi.org/10.1111/j.1461-0248.2010.01457.x.

Lehman, M.M., Ramil, J.F., 2002. Software uncertainty. In: Bustard, D., Liu, W., Sterritt, R. (Eds.), Soft-ware 2002: Computing in an Imperfect World. Vol. 2311 of Lecture Notes in Computer Science. Springer Berlin, Heidelberg, Berlin, Heidelberg, pp. 477–514. Ch. 14.

Lechner, A.M., Reinke, K.J., Wang, Y., Bastin, L., 2013. Interactions between landcover pattern and geospatial processing methods: effects on landscape metrics and classification accuracy. Ecol. Complex. 15, 71–82. http://dx.doi.org/10.1016/j.ecocom.2013.03.003.

Li, H., Wu, J., 2004. Use and misuse of landscape indices. Landscape Ecol. 19 (4), 389–399. http://dx.doi.org/10.1023/B: LAND.0000030441.15628.d6.

Lidicker Jr., W.Z., Peterson, J.A., 1999. Responses of small mammals to habitat edges. In: Barrett, G.W., Peles, J.D. (Eds.), Landscape Ecology of Small Mammals. Springer-Verlag, New York, pp. 211–227.

Lindenmayer, D., Hobbs, R.J., Montague-Drake, R., Alexandra, J., Bennett, A., Burgman, M., Cale, P., Calhoun, A., Cramer, V., Cullen, P., Driscoll, D., Fahrig, L., Fischer, J., Franklin, J., Haila, Y., Hunter, M., Gibbons, P., Lake, S., Luck, G., MacGregor, C., McIntyre, S., MacNally, R., Manning, A., Miller, J., Mooney, H., Noss, R., Possingham, H., Saunders, D., Schmiegelow, F., Scott, M., Simberloff, D., Sisk, T., Tabor, G., Walker, B., Wiens, J., Woinarski, J., Zavaleta, E., 2008. A checklist for ecological management of landscapes for conservation. Ecol. Lett. 11 (1), 78–91. http://dx.doi.org/10.1111/j.1461-0248.2007.01114.x.

Lyons, R., 2011. Distance Covariance in Metric Spaces. Arxiv preprint. arXiv: 1106.5758.

McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002. FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst http://www.umass.edu/landeco/research/fragstats/fragstats.html.

McGarigal, K., Tagil, S., Cushman, S.A., 2009. Surface metrics: an alternative to patch metrics for the quantification of landscape structure. Landscape Ecol. 24 (3), 433–450. http://dx.doi.org/10.1007/s10980-009-9327-y.

Miller, R.S., Franchi, M., Flor, R., 2004. GRASS GIS manual: r.drain. In: GRASS GIS 6.4.3svn Reference Manual. http://grass.osgeo.org/grass64/manuals/r.drain.html.

Morin, A., Urban, J., Adams, P.D., Foster, I., Sali, A., Baker, D., Sliz, P., 2012. Shining light into black boxes. Science 336 (6078), 159–160. http://dx.doi.org/10.1126/science.1218263.

Mubareka, S., Koomen, E., Estreguil, C., Lavalle, C., 2011. Development of a composite index of urban compactness for land use modelling applications. Landscape Urban Plan.. ISSN: 0169-2046 103 (3–4), 303–317. http://dx.doi.org/10.1016/j.landurbplan.2011.08.012.

Neteler, M., Bowman, M.H., Landa, M., Metz, M., 2012. GRASS GIS: a multi-purpose open source GIS. Environ. Model. Softw. 31, 124–130. http://dx.doi.org/10.1016/j.envsoft.2011.11.014.

O'Farrell, P.J., Anderson, P.M.L., 2010. Sustainable multifunctional landscapes: a review to implementation. Curr. Opin. Environ. Sustain. 2 (1–2), 59–65. http://dx.doi.org/10.1016/j.cosust.2010.02.005.

Ostapowicz, K., Vogt, P., Kozak, J., Riitters, K., Estreguil, C., 2008. Impact of scale on morphological spatial pattern of forest. Landscape Ecol. 23 (9), 1107–1117. http://dx.doi.org/10.1007/s10980-008-9271-2.

Peng, R.D., 2011. Reproducible research in computational science. Science 334 (6060), 1226–1227. http://dx.doi.org/10.1126/science.1213847.

Rémillard, B., 2009. Discussion of: Brownian distance covariance. Ann. Appl. Statist. 3 (4), 1295–1298. http://dx.doi.org/10.1214/09-AOAS312F.

Ries, L., Fletcher, R.J., Battin, J., Sisk, T.D., 2004. Ecological responses to habitat edges: mechanisms, models, and variability explained. Annu. Rev. Ecol. Evol. Syst. 35, 491–522. http://dx.doi.org/10.1146/annurev.ecolsys.35.112202.130148.

Ries, L., Sisk, T.D., 2004. A predictive model of edge effects. Ecology 85 (11), 2917–2926. http://dx.doi.org/10.1890/03-8021.

Riitters, K.H., Wickham, J.D., O'Neill, R.V., Jones, K.B., Smith, E.R., 2000. Global-scale patterns of forest fragmentation. Ecol. Soc. (Formerly Conserv. Ecol.) 4 (2), 3. http://www.consecol.org/vol4/iss2/art3/.

Riitters, K.H., Wickham, J.D., O'Neill, R.V., Jones, K.B., Smith, E.R., Coulston, J.W., Wade, T.G., Smith, J.H., 2002. Fragmentation of continental United States forests. Ecosystems 5 (8), 815–822. http://dx.doi.org/10.1007/s10021-002-0209-2.

Riitters, K., Wickham, J., Wade, T., 2009. Evaluating anthropogenic risk of grassland and forest habitat degradation using land-cover data. Landscape Online 13, 1–14. http://dx.doi.org/10.3097/LO.200913.

Riitters, K., Vogt, P., Soille, P., Kozak, J., Estreguil, C., 2007. Neutral model analysis of landscape patterns from mathematical morphology. Landscape Ecol. 22 (7), 1033–1044. http://dx.doi.org/10.1007/s10980-007-9089-3.

Rutledge, D., 2003. Landscape Indices as Measures of the Effects of Fragmentation: Can Pattern Reflect Process? Department of Conservation. New Zealand, Wellington. http://www.doc.govt.nz/documents/science-andtechnical/dsis98.pdf.

Saura, S., Estreguil, C., Mouton, C., Rodríguez-Freire, M., 2011a. Ecological indicators – network analysis to assess landscape connectivity trends: application to European forests (1990–2000). Ecol. Indicat. 11 (2), 407–416. http://dx.doi.org/10.1016/j.ecolind.2010.06.011.

Saura, S., Torné, J., 2009. Conefor Sensinode 2.2: a software package for quantifying the importance of habitat patches for landscape connectivity. Environ. Model. Softw. 24 (1), 135–139. http://dx.doi.org/10.1016/j.envsoft.2008.05.005.

Saura, S., Pascual-Hortal, L., 2007. A new habitat availability index to integrate connectivity in landscape conservation planning: comparison with existing indices and application to a case study. Landscape Urban Plan. 83, 91–103. http://dx.doi.org/10.1016/j.landurbplan.2007.03.005.

Saura, S., Vogt, P., Velázquez, J., Hernando, A., Tejera, R., 2011b. Key structural forest connectors can be identified by combining landscape spatial pattern and network analyses. For. Ecol. Manage. 262 (2), 150–160. http://dx.doi.org/10.1016/j.foreco.2011.03.017.

Schölkopf, B., Smola, A., Müller, K.R., 1998. Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput.. ISSN: 0899-7667 10 (5), 1299–1319. http://dx.doi.org/10.1162/089976698300017467.

Scholz, M., Fraunholz, M., Selbig, J., 2008. Nonlinear principal component analysis: neural network models and applications. In: Lecture Notes in Computational Science and Engineering vol. 58, pp. 44–67. http://dx.doi.org/10.1007/978-3-540-73750-6_2.

Scholz, M., Vigàrio, R., 2002. Nonlinear PCA: a new hierarchical approach. In: Proceedings of the 10th European Symposium on Artificial Neural Networks (ESANN), ISBN 2-930307-02-1, pp. 439–444.

Shapiro, M., 2010. GRASS GIS manual: r.neighbors. In: GRASS GIS 6.4.3svn Reference Manual. http://grass.osgeo.org/grass64/manuals/r.neighbors.html.

Soille, P., Vogt, P., 2009. Morphological segmentation of binary patterns. Pattern Recogn. Lett. 30 (4), 456–459. http://dx.doi.org/10.1016/j.patrec.2008.10.015.

Steiniger, S., Hay, G.J., 2009. Free and open source geographic information tools for landscape ecology. Ecol. Inform. 4 (4), 183–195. http://dx.doi.org/10.1016/j.ecoinf.2009.07.004.

Szekely, G.J., Rizzo, M.L., 2009a. Brownian distance covariance. Ann. Appl. Statist. 3 (4), 1236–1265. http://dx.doi.org/10.1214/09-AOAS312.

Szekely, G.J., Rizzo, M.L., 2009b. Rejoinder: Brownian distance covariance. Ann. Appl. Statist. 3 (4), 1303–1308. http://dx.doi.org/10.1214/09-AOAS312REJ.

Szekely, G.J., Rizzo, M.L., Bakirov, N.K., 2007. Measuring and testing dependence by correlation of distances. Ann. Statist. 35 (6), 2769–2794. http://dx.doi.org/10.1214/009053607000000505.

Taylor, R.S., Oldland, J.M., Clarke, M.F., 2008. Edge geometry influences patch-level habitat use by an edge specialist in south-eastern Australia. Landscape Ecol. 23 (4), 377–389. http://dx.doi.org/10.1007/s10980-008-9196-9.

Van Rossum, G., Drake Jr., F., 2011. The Python Language Reference Manual (version 3.2). Network Theory Limited, ISBN 978-1-906966-14-0.

Vittoz, P., Engler, R., 2007. Seed dispersal distances: a typology based on dispersal modes and plant traits. Bot. Helv. 117 (2), 109–124. http://dx.doi.org/10.1007/s00035-007-0797-8.

Vogt, P., Riitters, K., Estreguil, C., Kozak, J., Wade, T.G., Wickham, T.G., 2007a. Mapping spatial patterns with morphological image processing. Landsc. Ecol 22 (2), 171–177. http://dx.doi.org/10.1007/s10980-006-9013-2.

Vogt, P., Riitters, K., Iwanowski, M., Estreguil, C., Kozak, J., Soille, P., 2007b. Mapping landscape corridors. Ecol. Indic 7 (2), 481–488. http://dx.doi.org/10.1016/j.ecolind.2006.11.001.

Wickham, J.D., Norton, D.J., 1994. Mapping and analyzing landscape patterns. Landscape Ecol. 9 (1), 7–23. http://dx.doi.org/10.1007/BF00135075.

Wickham, J.D., Riitters, K.H., Wade, T.G., Vogt, P., 2010. A national assessment of green infrastructure and change for the conterminous United States using morphological image processing. Landscape Urban Plan.. ISSN: 0169-2046 94 (3–4), 186–195. http://dx.doi.org/10.1016/j.landurbplan.2009.10.003.

Wrbka, T., Erb, K.H., Schulz, N.B., Peterseil, J., Hahn, C., Haberl, H., 2004. Linking pattern and process in cultural landscapes. An empirical study based on spatially explicit indicators. Land Use Policy 21, 289–306. http://dx.doi.org/10.1016/j.landusepol.2003.10.012.

Zaragozí, B., Belda, A., Linares, J., Martínez-Pérez, J.E., Navarro, J.T., Esparza, J., 2012. A free and open source programming library for landscape metrics calculations. Environ. Model. Softw.. ISSN: 1364-8152 31, 131–140. http://dx.doi.org/10.1016/j.envsoft.2011.10.009.