Measuring connectivity in habitat mosaics: the equivalence of two existing network indices and progress beyond them

S. Saura

Phone: +34 91 336 71 22, Fax: +34 91 543 95 57, E-mail: santiago.saura@upm.es

Keywords: Ecological flows, Graph theory, Landscape change, Landscape connectivity, Topological metrics.

Abstract: Network-based methods are being actively developed to respond to the needs for operational assessments of the degree of landscape connectivity and of the impact of landscape changes in ecological flows and related ecological processes. Among these, a recent paper by Matisziw and Murray (2009) presented the C index as an adequate and advantageous way of ranking habitat patches by their importance for the maintenance of landscape connectivity. We show that this index is equivalent and conveys the same information in undirected graphs as the recently described index, the landscape coincidence probability (LCP), which can be readily computed in any landscape network through the Condor Seminode software package. We slightly generalize the LCP definition for cases involving asymmetric dispersal, which makes LCP compatible with C and maintains the equivalence between both indices in directed graphs. We place LCP and C in a broader context of other existing indices and ongoing developments and describe how some of these may be better suited for the analysis of the connectivity in landscape networks and their changes. We conclude by highlighting the need (1) to go beyond the identification of unobstructed movement paths or habitat components (sets of interconnected patches) when pursuing the most appropriate landscape connectivity indices and (2) for increased efforts in assessing and reporting the potential overlaps, coincidences and synergies between the available approaches in order to guide the final user and facilitate index selection in a densely populated metric space.

Introduction

The importance of landscape connectivity for maintaining and facilitating ecological flows and many desirable properties related to them has been highlighted in an almost endless number of researches and conservation initiatives (e.g., Crooks and Sanjay 2006). Population persistence, genetic exchange, wildlife migration and seed and pollen dispersal, among others, are dependent on the degree of landscape connectivity, which hence constitutes an important part of any modern biodiversity conservation or monitoring system. Therefore, the demands and research efforts on adequate and reliable methods to deal with landscape connectivity have largely increased in the recent past (e.g., Hanski and Ovaskainen 2000, Ricotta et al. 2000, Urban and Keitt 2001, Fagan and Calabrese 2006, Vogt et al. 2007, McRae et al. 2008, Saura and Rubio 2010). Indeed, a prerequisite for developing appropriate connectivity indices is that they should rely on methods with a solid analytical and empirical support. In addition, as noted by Fagan and Calabrese (2006), connectivity indices must be pragmatic and based upon data that might actually be attained on a regular basis to allow for their use in operational planning and change assessments at wide scales.

In this context, network-based (or graph theoretical, as they are often called) methods are being increasingly developed and applied for measuring and analyzing functional connectivity in heterogeneous landscapes (e.g., Ricotta et al. 2000, Urban and Keitt 2001, Jordán et al. 2003, Pascual-Hortal and Saura 2006, Bodin and Norberg 2007, McRae et al. 2008, Matisziw and Murray 2009, Urban et al. 2009, Saura and Rubio 2010). In this approach, a landscape is represented as a set of nodes (habitat units) that are connected by links. The definition of nodes and links depends on the scale, amount of available information on the species and the landscape, and the objectives of the analysis; although nodes are usually assumed to correspond to habitat patches, other options are also possible (habitat cells, management or administrative units, protected areas, etc.). In undirected graphs, the possibility or feasibility of movement from patch i to patch j is considered the same as from patch j to i, while directed graphs are used to model asymmetric ecological flows, such as wind-driven dispersal or water flows and population dynamics in river networks (Scholes and Wiens 2003, Schick and Lindsey 2007). Several advantages have been pointed out in the literature advocating for the increased use of these network-based methods (see previous references): (1) they provide a spatially explicit representation of the landscape that is able to evaluate the contribution to connectivity of individual landscape elements (2) some network indices provide similar outcomes to other more complex and biologically detailed spatially explicit metapopulation models in what is usually needed for the actual planning and monitoring (Minor and Urban 2007, Visconti and Elkin 2009), (3) they are adaptable to different degrees of detail in the available information on the species and the landscape; they can accommodate more detailed information when this is available but are still operational with relatively scarce data (Estrada and Bodin 2008, Saura and Rubio 2010), and (4) they benefit from the necessity modifications, of the wealth of algorithms and analytical tools that have been developed in
graph theory to analyze the connectivity of many different types of networks.

However, graphs are just a data structure, as the typical vector or raster structures in geographic information system, and their use is not itself a guarantee of the quality and adequacy of the resultant indices. Indeed, many different outcomes of variable quality can be obtained through a network-based analysis of the landscape. Pascual-Hortal and Saura (2006) noted the weaknesses of several graph-based connectivity indices for (1) successfully prioritizing the landscape elements that are more important for the maintenance or potential improvement of connectivity and for (2) adequately responding to different types of spatial changes that may occur in the landscape. These authors showed that the connectivity indices that performed better in this respect were those based in the concept of measuring habitat availability (reachability) at the landscape scale. Habitat availability indices consider a patch as a space where connectivity exists, integrating in a single measure the connected habitat area that exists within the patches (intrapeck connectivity) with the area made available (reachable) through the connections among different habitat patches (interpatch connectivity) (Pascual-Hortal and Saura 2006, Saura and Pascual-Hortal 2007, Saura and Rubio 2010).

An interesting recent paper by Matiszczik and Murray (2009) proposed a network index (C) for evaluating landscape connectivity and discussed the advantages of such an index for an adequate assessment of connectivity changes in landscape networks. C was compared with the Harary index (H) (Ricotta et al. 2000), but not with other network-based indices that were compared with a considerable analytical detail in Pascual-Hortal and Saura (2006) and which presented a better performance than H according to the set of properties evaluated in that study.

Here we will show that this C index is in fact conveying the same information as the LCP index previously described in Pascual-Hortal and Saura (2006) in undirected graphs, therefore sharing all the advantages and potential limitations reported for LCP in that study. We show how the definition of LCP can be easily generalized and made equivalent to C also in cases involving asymmetric dispersal while maintaining its original definition as a probability. We discuss the scope of application of LCP and C, and briefly describe how other existing indices and recent developments may provide a more convenient or at least complementary way to evaluate landscape connectivity and their changes from a network-based perspective.

The equivalence of the LCP and C indices

The landscape coincidence probability (LCP) is defined in undirected graphs (symmetric dispersal/links) as the probability that two points randomly located within the landscape belong to the same habitat component (Pascual-Hortal and Saura 2006), where component is defined as a set of patches that can be reached from each other through existing links (no path exists between patches belonging to different components). LCP is a generalization of the definition of the Simpson’s diversity index and the degree of coherence index related to it (Jæger 2000, Bogner f et al. 2005). LCP is given by the following expression:

$$LCP = \frac{\text{LCPnum}}{A_L} = \sum_{i=1}^{N} \frac{c_i^2}{A_L}$$

(1)

where NC is the number of components in the landscape, \(c_k\) is the total habitat area in component \(k\) (sum of the habitat area in all the patches belonging to that component), \(A_L\) is a constant corresponding to the total landscape area (comprising both habitat and non-habitat areas), and LCPnum is the numerator of the LCP equation, corresponding to the sum of the squared component areas. In more general terms a patch attribute different from habitat area may be considered, but this is not of particular relevance for the current purposes. Since the C index defined by Matiszczik and Murray (2009) assumed that all the nodes (patches) had the same characteristics (e.g., area), here we assign as well an area of one unit to all the habitat patches, case in which LCPnum simplifies to:

$$LCPnum = \sum_{i=1}^{NC} c_i^2 = \sum_{i=1}^{NC} \phi_i^2$$

(2)

where \(\phi_i\) is the number of patches (nodes) in each component.

On the other hand, the C index proposed by Matiszczik and Murray (2009) as a measure of total network connectivity is defined as:

$$C = \sum_{i=1}^{N} \sum_{j=1}^{N} Z_{ij}$$

(3)

where \(Z_{ij}=1\) if a path exists between patches \(i\) and \(j\) (either a single direct link from \(i\) to \(j\) or a longer path utilizing several intermediate stepping stones patches), including the case \(i=j\), and \(Z_{ij}=0\) otherwise. \(N\) is the total number of patches in the habitat network.

A particular node or habitat patch \(i\) will always be part of just one of the \(NC\) components in the undirected landscape network (if the patch is fully isolated it will be part of a component made up only by itself). If \(p_k\) is the number of patches in component \(k\), it is straightforward that in the sum in equation (3) for the computation of \(C\) a particular patch \(i\) in that component will be connected to a total of \(p_k\) patches (including itself). Therefore, this patch will contribute to the total \(C\) value by an amount of \(p_k\). Any other patch belonging to the same component \(k\) will contribute to \(C\) by the same amount, as it will be connected to the same set of \(p_k\) patches (including itself). Therefore, a set of \(p_k\) patches in a particular component will contribute all together as \(p_k^2\) to the sum in equation (3). It then follows that the total \(C\) value can be computed as the sum of the contributions of each of the \(NC\) components, and that \(C\) is in fact identical to LCPnum:

$$C = \sum_{i=1}^{N} \sum_{j=1}^{N} Z_{ij} = \sum_{k=1}^{NC} p_k^2 = \text{LCPnum}$$

(4)
Then $C$ simply equals to the sum of the squared number of nodes in each component in undirected graphs, which may be easier to compute and may have a more intuitive interpretation than with the original formula and definition by Matsuzawa and Murray (2009). Obviously the minimum $C$ (and $LCPnum$) value is obtained when all the patches belong to a single component, case in which $C$ is equal to $N^2$, as in the two landscapes in Figure 1, with $N = 10$ in both of them.

Because component cannot be defined in the same way in directed than in undirected graph (e.g., in the first case a patch can be connected to a group of patches in one direction but not in the other), for those cases involving asymmetric dispersal $LCP$ cannot be calculated based in equation (1), which assumes undirected networks. However, it is possible to generalize the $LCP$ index, by analogy to the $HC$ (Pascual-Hortal and Saura 2006) and $PC$ (Saura and Pascual-Hortal 2007) indices (both of which can be computed either on directed or undirected network), making $LCP$ compatible with the notation and concept of the $C$ index but still being interpreted as a probability as in its original definition. $LCP$ is then defined as the probability that a particular organism dependent on the focal habitat can move from an initial point $i$ randomly located in the landscape to another destination point $j$ also randomly located within the study area (note that now directionality is introduced in the definition). This requires that both points fall into habitat areas and that, in addition, both points either fall (1) within the same habitat patch or (2) into different but connected patches so that it is possible to move from $i$ to $j$ through the directed links in the network ($Z_{ij} = 1$ using the same notation as above for the $C$ index). This matches with the way of measuring connectivity proposed by Pascual-Hortal and Saura (2006), with the first case corresponding to the connectivity within habitat patches (intrapatch connectivity, $Z_{ij} = 1$ if $i = j$) and the second to the connectivity between habitat patches (interpatch connectivity), both of which have to be considered together for a good reaction of the resultant connectivity index to spatial changes (Pascual-Hortal and Saura 2006):

$\begin{align*}
\text{LCP} &= \frac{LCPnum}{\text{A1}_i^2} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} Z_{ij}}{\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} Z_{ij}} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} Z_{ij}}{\text{A1}_i^2} \quad (5)
\end{align*}$

where $a_{ij}$ is the area of patch $i$ and the rest as defined earlier. As above, when all the habitat patches are considered of equal characteristics as for the $C$ index (i.e., $a_{ij} = 1$ for all the patches), then the $LCPnum$ value given by equation (5) is identical to the $C$ value of equation (3). In addition, equation (5) simplifies to equation (1) in undirected networks (in which $Z_{ij} = 1$). Finally, given equations (4) and (5), it is obvious that the equivalence between $LCPnum$ and $C$ would also hold if the different characteristics of the habitat patches were taken into account in a generalized version of $C$ in which the product $a_{ij} Z_{ij}$ is used instead of just $Z_{ij}$ (or other alternatives that may apply equally to variations of $LCPnum$ and $C$).

The impact of a particular change in the landscape network (e.g., loss of a patch, link or sets of them through time) for the maintenance of connectivity can be quantified through the relative variation in the index value after that change. This was expressed as $dLCP$ in Pascual-Hortal and Saura (2006) and similarly for the $C$ index as $\Delta$ in Matsuzawa and Murray (2009); both will provide the same result for ranking the importance of different landscape elements and spatial changes. $LCP$ and $LCPnum$ also yield the same results in this respect, since $A1_i$ is a constant that remains unchanged after any spatial change that may occur in the landscape, and therefore it is irrelevant to determine the final prioritization ($dLCP$ value) provided by the index.

The values of $LCPnum$ and $dLCP$ in any undirected landscape network can be directly obtained through the current version 2.2 of the Conefor Sensinode software package released in 2007 (Saura and Torne 2009), which is freely available (together with the source codes) from http://www.cone-for.org and was specifically developed for these purposes. Conefor Sensinode computes, among other indices, the $dLCP$ values for every patch in a network, with feasible proc-
cessing in a standard personal computer even for considerably large landscapes (well above 10,000 nodes or habitat patches for LCP). The current 2.2 version does not, however, allow the computation of indices in directed graphs, although this might be incorporated in a future compilation of this software package.

Connectivity indices through network-based methods: beyond landscape components

The characteristics and behavior of a set of network-based indices (including LCP) were analyzed and compared in detail by Pascual-Hortal and Saura (2006). Despite the good performance of LCP to several of the properties evaluated in this study, LCP was not found to be the best performing index. The main limitations of LCP (and C) are its lack of sensitivity to changes in the easiness of movement between the different habitat units, as related for instance to the topological distance (number of links needed to move in the shortest path from one patch to another in the landscape network) or to other distance measure such as geographically, effective (least cost) or genetic distances. LCP (C) is not sensitive to the loss of patches or links when they do not break the habitat network in a large number of components that get fully isolated from each other (Pascual-Hortal and Saura 2006). It also fails to detect the higher importance of those landscape elements that allow shortening the distance that organisms have to traverse to reach other habitat areas (Pascual-Hortal and Saura 2006). For example, in landscape B in Figure 1 all the nodes have the same importance for connectivity according to LCP (dLCP) or C (O), since all the nodes have the same habitat area and the loss of none of them would (if removed one at a time) separate the existing habitat in several differentiated components. However, node 6 would need to be considered more important than 2, because the loss of node 2 has no effect in the length (number of links) of the shortest paths between the other habitat areas while losing 6 would increase the cost of moving between several other pairs of patches: e.g., the topological distance between patches 3 and 10 would increase from two to three links after the loss of patch 6 (Figure 1). For similar reasons as in the previous example, the individual loss of any of the links in landscape B (Figure 1) will remain unreported by LCP or C, which assign a zero importance to any of these links. However, links like the one connecting nodes 4 and 8 in landscape B (Figure 1) are to be considered of some importance for keeping habitat areas functionally closer to each other, since the removal of that link would increase the topological distance from patch 8 to patches 4 (from one to two links), 2 (from three to four links), 1 and 3 (from two to three links) for these latter patches. In the same way, adding in landscape A a direct connection between patches 2 and 3 would reduce the topological distance between them from four to just one link (Figure 1), but this would not produce any variation in the LCP(C) value. The LCP index achieves its maximum value when all the patches can be reached from each other (all of them belong to the same component), but it does not consider the relative difficulty or number of movement steps needed to reach them. LCP and C saturate at that point and do not react to further increases in the number of links and available paths that may actually benefit the frequency and strength of the ecological fluxes in the landscape. For instance, LCP(C) assigns the same degree of connectivity to the two landscapes in Figure 1 because in both of them all the habitat nodes belong to the same component (in both landscapes there is an unobstructed path available for movement between every pair of nodes). LCP or C do not take into account the considerably higher easiness of movement (shorter topological distances) between habitat nodes in landscape B, where stronger and more frequent ecological fluxes and interchanges between the habitat areas are to be expected compared to landscape A. Although the analyses by Pascual-Hortal and Saura (2006) were performed in undirected graphs, the same or analogous considerations apply as well for directed networks.

Other indices like the integral index of connectivity (IIC) are free of the limitations just described for LCP or C, and have been shown to provide a more convenient network-based characterization of landscape connectivity and its changes (Pascual-Hortal and Saura 2006), as applied in a variety of studies (Pascual-Hortal and Saura 2008; Ned 2008; Perotto-Baldivieso et al. 2009; Laitu et al. 2010). The same applies to the probability of connectivity (PC) index (Saura and Pascual-Hortal 2007), which is considered more specifically below. In the examples of Figure 1, both IIC or PC would (a) consider landscape B more connected than A, (b) assign a higher importance to patch 6 than to patch 2 for maintaining connectivity in landscape B, and (c) detect the value as connectivity providers of links such as the one connecting nodes 4 and 8 in landscape B. The closeness centrality is another example of a network index that explicitly takes into account the topological distance between habitat patches (Esteve and Bodin 2008; Junisz and Goetz 2008), indicating lower patch connectivity as this distance increases, as also incorporated by Jordan et al. (2003, 2007) among others.

Matysik and Murray (2009) suggested that the accessibility or easiness of movement between different habitat units should not be taken into account when evaluating landscape connectivity. However, there are numerous empirical evidences that show that the ecological flows related to the degree of landscape connectivity are sensitive to continuous variations in the distance between the habitat units, either measured as a topological distance (number of links), Euclidean (straight line) distance or effective distance (resulting from least-cost paths or other related approaches) (e.g., Nathan and Muller-Landau 2000; Buguenet 2003; McRae and Beier 2007). Moreover, recent research (McRae and Beier 2007; McRae et al. 2008) has shown that not only the distance or easiness of movement along the best (shortest) path matters for ecological connectivity. The contribution of all the possible and multiple pathways existing in the landscape sums up to determine the magnitude of the actual ecological flows between habitat areas and the probability of successful dispersal and mortality of random walkers moving across heterogeneous landscapes. These authors have adopted
methods from electronic circuit theory to improve the characterization of the connections between different habitat areas in the landscape by considering the effects of all possible pathways across a landscape simultaneously, thus taking into account a wider proportion of the landscape than in the conventional least-cost modeling.

Therefore, any connectivity index that is intended to successfully predict and relate to these ecological processes should have some form of distance or accessibility measure as an ingredient. In general, connectivity is not only and should not just be evaluated from the identification of components in the landscape network. For example, Bodin and Norberg (2007) noted the limitations of a component-based connectivity analysis, especially in landscapes with relatively high levels of connectivity. In these circumstances, assessments based only on network components may lack the ability to provide any ranking or discrimination of the value of the different landscape elements for connectivity (see the examples mentioned above for landscape B in Figure 1). They advocated for using, instead of components, landscape compartments that do not need to be fully isolated from each other in order to be identified and discriminated in the landscape. The approach suggested by Bodin and Norberg (2007) can identify groups of patches wherein species interact intensively relative to less frequent and weaker interactions with patches from other groups (e.g., long range and in frequent dispersal that contributes to expand ranges across generation).

Indices like LCP, C or HIC are all based on a binary characterization of the connections (links) between the habitat patches (where each pair of patches is simply connected or not), with no intermediate modulation of the frequency, strength or quality of the individual connections. This may result oversimplified for some applications since not all the usable connections that can be identified in the landscape might be equally effective in conducting the ecological flows and support the same quantities of dispersing individuals. In many real world landscape networks the degree of connectivity between the different habitat patches is variable, but still most of them might be connected to some (even if eventually weak) degree. With this perspective in mind, Sauna and Pascual-Hortal (2007) presented the probability of connectivity (PC) index. PC preserved the original desirable characteristics that are present in HIC but considered a more detailed characterization of the connections between habitat patches in which every link in the graph is characterized by a given probability of dispersal between two patches (weighted graph). PC has been so far applied in a variety of applications related to landscape connectivity (e.g., Neel 2008, Perotto-Baldíevo et al. 2009, Fu et al. 2010, Kreck and Hanser 2010). Recently, PC has also been partially adopted by Watts and Handley (2010) and slightly modified by Sauna et al. (2011) in order to (a) avoid the very low PC values that can be obtained when the focal habitat represents only a minor part of the total landscape area and to (b) allow for a more direct and straightforward comparison and interpretation of the temporal changes in connectivity with those in the amount of habitat in the landscape. In addition, Sauna and Rubio (2010) have shown how indices such as PCo r HIC can be partitioned in three functions that quantify the different ways in which individual landscape elements (patches, links) can contribute to the overall connectivity and availability of habitat in the landscape. These three functions are measured in the same unit and can be directly compared and combined within a unified framework. This avoids (1) the highly problematic and common combination of indices with different derivations when attempting to characterize the distinctive roles and contributions of patches and links in the landscape network and (2) the arbitrary weighting of connectivity considerations when confronted with other ecological criteria and conservation alternatives.

Conclusions

We have contributed to better place the interesting developments and ideas by Matiszisz and Murray (2009) in a wider context of previously presented indices (particularly LCP) and ongoing developments for the analysis of connectivity in landscape networks. Although accounting for the presence or absence of an unobstructed path of movement between each pair of habitat patches is an important part of connectivity analysis, we conclude that an improved network-based assessment should go beyond that evaluation, which in the case of undirected networks is equivalent to the identification of the number and size of the components in the landscape.

Users are currently overwhelmed by a myriad of indices related to connectivity. Hence, the selection of the most appropriate ones for a particular monitoring or conservation problem turns out to be difficult and potentially arbitrary in many situations. This is especially the case when the potential overlaps and synergies between the different available approaches are not reported and specified, and in stead they are viewed as isolated analytical pieces in a disconnected methodological space. In this context, it is particularly important to concentrate efforts in relating the different existing indices and developments (and those that may be produced in the years to come). These efforts should include a sufficient, end-user driven and comparative description of their potential advantages, limitations and possibilities for operational application in on-site monitoring and conservation planning.

Acknowledgements: Funding was provided by the Ministerio de Ciencia e Innovación (Spain) and the European Union (FEDER fund) through the DECOFOR (AGL2009-07140/FOR), MONTES-CONSOLIDAR (CDR2008-00040) and Restauración y Gestión Forestal (PSI-310000-2009) projects. Two anonymous reviewers made helpful comments on an earlier version of this article.

References


Received June 1, 2010
Revised August 23, 2010
Accepted September 8, 2010