LQGraph: A Software Package for Optimizing Connectivity in Conservation Planning

Trevon Fuller and Sahotra Sarkar*

Biodiversity and Biocultural Conservation Laboratory, Section of Integrative Biology, University of Texas at Austin, 1 University Station, #C0930, Austin, TX 78712-1180, USA

Word count: 3641

Abstract

LQGraph implements methods for optimizing the connectivity of sites administered to protect biodiversity (a conservation area network). The methods are suitable for existing protected areas (for example, national parks) or a proposed network. We model the landscape using graphs. The user provides a quality score for each site in the landscape outside the conservation areas. Based on these scores, LQGraph finds contiguity areas of maximum quality to link the conservation areas. The contiguity areas can be filtered to prioritize those that connect the conservation areas with the minimum number of sites or area. Compared with existing methods for establishing connectivity in a conservation area network, LQGraph can analyze substantially larger networks. LQGraph also identifies sites that efficiently isolate conservation areas, which is useful for halting the spread of pathogens or invasive species. The software provides routines for constructing the required input files and for visualizing several properties of the landscape, including site quality and the shortest paths between conservation areas. LQGraph and the accompanying documentation can be freely downloaded from the worldwide web.

Keywords: connectivity; conservation area networks; reserve selection; graph theory; ecological restoration

---

*Corresponding author: Tel.: +1 512 232 7122; fax: +1 512 471 4806.
E-mail address: sarkar@mail.utexas.edu (S. Sarkar).

Software availability

Name of software: LQGraph
Developers: Trevon Fuller and Sahotra Sarkar
Contact address: Biodiversity and Biocultural Conservation Laboratory, Section of Integrative Biology, University of Texas at Austin, 1 University Station, #C0930, Austin, TX 78712-1180.
E-mail: sarkar@mail.utexas.edu
Year first available: 2005
Hardware required: Any IBM compatible PC
Software required: Microsoft Windows XP
Programming language: Microsoft Visual C++ .NET 2003
Program size: 880 KB
Availability: Freely downloadable with manual and supporting materials from http://uts.cc.utexas.edu/~consbio/Cons/ResNet.html
Available since: August 23, 2005.
Online documentation: http://uts.cc.utexas.edu/~consbio/Cons/ResNet.html

1. Introduction

A landscape is functionally connected if species can disperse between habitat sites; otherwise, it is fragmented (Merriam, 1984). Though the effects of fragmentation *per se* are controversial (Fahrig, 2003), the reduced dispersal associated with fragmentation can harm plant (Murphy and Lovett-Doust, 2004) and animal species (Bowne and Bowers, 2004). Connectivity areas such as corridors increase the movement of birds and seeds of threatened plants between habitat sites (Levey et al., 2005). Species richness of many taxa, including bats, beetles, birds, butterflies, frogs, and plants is higher when habitat sites are connected (Banks et al., 2005; Drinnan, 2005; Harvey et al., 2005). Linear habitat features such as live fences show high richness of animal species, possibly serving as stepping stones between tropical forest remnants (Estrada and Coates-Estrada, 2005). Provided that connecting areas are selected carefully, it is plausible that many species in conservation areas will use them for migration or egress in the event of local environmental stress. Current methods of landscape management often do not provide functional connectivity for many species (Creagan and Osborne, 2005). Thus, methods for establishing connectivity in fragmented landscapes are important for conservation planning.

This note describes a software tool for optimizing connectivity of a conservation area network (CAN), that is, a set of sites administered to protect biodiversity. During the design of a CAN, connectivity may be incorporated in two ways. First, planners can use a hierarchical approach in which the biodiversity content of a proposed site is assessed before its spatial properties (Garson et al., 2002). New sites could be selected so as to optimize some spatial property (such as the summed inter-site distance) of the existing sites plus the new sites (Onal and Briers, 2002). Second, planners can use multi-criteria analysis to optimize both biodiversity representation and connectivity of a network (Sarkar et al., 2004). After selecting a CAN, it may be beneficial to augment its connectivity by selecting new sites to link the old sites in an uninterrupted chain of contiguity areas (such as traditional corridors). LQGraph implements the latter approach.

We model the landscape as a graph, which is a discrete mathematical structure consisting of a set of vertices and edges. (For a history of graph-theoretic models in ecology, see Fuller, [2004]). In LQGraph, the user provides contiguity and quality scores for each site in the study region. Based on these scores, LQGraph constructs a set of high-quality paths between conservation areas. The user may want to filter the set of paths to select those that connect the conservation areas most efficiently. To this end, LQGraph can prioritize contiguity areas to link conservation areas via paths that: (i) have maximum quality and (ii) occupy the minimum amount of land.

2. Background

To use LQGraph, the study region must be divided into cells, the most basic units of the subsequent analysis. The set, L, of all cells in the study area is partitioned into the set A of cells available for conservation planning and the set B of unavailable cells. A has subsets P, the set of cells that can potentially be included in a CAN, and B, the set of existing CAN cells. Each cell $a \in A$ has a list of adjacent cells, $Adj[a]$, a "contiguity" score, $contiguity[a] \in [0,1]$, where 1 is the best score in the ranking system and 0 the worst and a "landscape quality" score, $landscape\_quality[a]$, normalized in the same way. The user supplies these quality scores in the input file. A place prioritization algorithm selects from A the set S of cells in the CAN, which satisfy biodiversity representation goals such as the protection of a targeted percentage of each species' habitat. A target of 10 % is conventionally used in conservation planning (IUCN 1994). The set $CA = \{CA_1, CA_2, \ldots, CA_n\}$ of conservation areas selected is a partition of S such that each subset $CA_j \in CA$ contains the cells in one conservation area, $CA_j \cap CA_k = \emptyset, 1 \le j \le n, 1 \le k \le n$, and $\bigcup_{j=1}^{n} CA_j = S$.

2.1. Landscape Contiguity Graph

A mapping $g : S \to V_C$ assigns each cell $s \in S$ to a vertex v of a directed, weighted "contiguity" graph $G_c = (V_c, E_c)$. The purpose for constructing the contiguity graph is to find the shortest path between pairs of cells in separate conservation areas. The edge set E_c contains an ordered pair $e = (u,v)$ if for a given pair of vertices $u,v \in V_c$, $g^{-1}(v) \in Adj[g^{-1}(u)]$. Thus, two vertices of the contiguity graph are linked by an edge if the sites in the landscape that correspond to the vertices are neighbors. A mapping $w_C : E_C \to \Re$ assigns weights to each $e \in E_C$ such that

$$w_C(u,v) = \begin{cases} 1 - contiguity[g^{-1}(v)] & g^{-1}(v) \in P - S \\ 1 & g^{-1}(v) \in S \end{cases}.$$

This mapping ensures that the shortest path between conservation areas will include few cells in conservation areas.

2.2. Landscape Quality Graph

We now construct a "landscape quality graph" G_L to establish connectivity between the conservation areas. Each edge of the landscape quality graph is the shortest path between two conservation areas. We find each such shortest path using Dijkstra's algorithm (Cormen et al., 2001). We then filter the set of shortest paths, SP, to construct a refined set, SPR, such that all paths: (i) begin and end at the center of mass a conservation area and (ii) traverse no intermediate conservation area. The justification for (ii) is that we wish to establish connectivity between conservation areas rather building paths through existing ones. The landscape quality graph $G_L = (V_L, E_L)$ is a condensation on G_c with respect to the partition CA . G_L is an undirected graph such that there is 1 vertex $v \in V_L$ for each conservation area. E_L has an edge (CA_j, CA_k) if SPR contains some shortest path $q = < s, \ldots, t >$ beginning at vertex s and ending at vertex t such that $g^{-1}(s) \in CA_j$ and $g^{-1}(t) \in CA_k$. The weight of the edge is defined by the function $w_L : q \to \Re$ as $\sum_{v \in q} (1 - landscape\_quality[v])$. Thus, high-quality paths have low edge weights.

Next we describe a method for prioritizing some edges of the landscape quality graph. We find a subgraph $H = (V_H, E_H)$ of G_L with the following properties: H is connected, acyclic, $V_H = V_L$, and $\sum_{e \in E_H} w_L(e)$ is as small as possible. The biological significance of H, which is called a "minimum spanning tree" (MST) of the landscape quality graph, is that it links all conservation areas via the smallest possible number of high-quality areas.

Experiments and asymptotic analysis indicate that our connectivity establishment procedure is faster than the method of Willams (1998). In LQGraph, Prim's MST algorithm is implemented with a min-heap priority queue with running time $O(E_L \lg V_L)$ (Siek et al., 2001). Williams models connectivity establishment as an NP-complete Steiner tree problem; this method can solve data sets with at most five vertices and on a desktop PC in reasonable time (Williams, 1998). LQGraph solves much larger data sets quickly (see section 4).

Last, LQGraph has a procedure for finding all MSTs of the landscape quality graph. At each iteration of Prim's algorithm, the vertices are partitioned into the set S of vertices in the incipient MST and $V_L - S$ outside it. An edge is said to "cross a cut" if one of its endpoints is in S and the other is in $V_L - S$. A "light edge" has that property that it crosses a cut and its weight is at least as small as any other edge crossing the cut. Suppose there are k "light" edges during an iteration of Prim's algorithm. We make k copies of the incipient MST, add a different light edge to each copy, then proceed with Prim's algorithm. Since every light edge belongs to some MST (Sedgewick, 1983), all trees found by the procedure will be MSTs (though there may be redundant trees).

3. Program description

LQGraph 1.0 is a dialog-based application with several libraries used for visualization. The main dialog consists of menu items, a settings panel that lists currently activated options, and a progress bar. Optimization and visualization routines are activated by filling out dialog boxes (for details, see Fuller and Sarkar, [2005]).

3.1. Input

LQGraph requires as input two text files. For each cell, the first input file contains the adjacency list and contiguity and landscape quality scores described above. The second input file lists the center of mass of each conservation area, which must be provided by the user. LQGraph selects connectivity areas to link these centers of mass.

3.2. Minimum Spanning Tree

After constructing the contiguity and landscape quality graphs, the user can prioritize edges of the latter. LQGraph prioritizes edges of the landscape quality graph by finding one or all MSTs of each component of the landscape quality graph. A component is a set of conservation areas that can be reached from one another by traversing sites outside the CAN. (G_L will likely have several components if $|B| >> |A|$).

3.2.2. Disconnecting Sets

The minimum disconnecting set is the smallest number of edges that must be removed from a component of the landscape quality graph to fragment it. LQGraph uses an extension of Matula's (1987) algorithm to find all such sets. The biological significance of the disconnecting set is that it represents the most efficient way to isolate conservation areas in the network, which is beneficial for halting the spread of an invasive species or pathogens.

3.2.3. Graph Condensation

To "condense" a component of the landscape quality graph means to remove all edges below a user-specified threshold weight and combine all remaining connected vertices into a single vertex. The biological significance of condensation is that it identifies the number of ecologically independent conservation regions in a network. Since such neighboring conservation areas are likely to be affected by the same environmental changes, they can be thought of as constituting a single conservation region.

3.3. Visualization/Output

LQGraph lets the user visualize various properties of the landscape quality graph, including each cell's contiguity and landscape quality score and the minimum spanning tree(s), disconnecting set(s), and condensations of each component. In addition to output displayed to the screen, by default LQGraph produces TXT files containing *x*- and *y*-coordinate data which can be displayed in ArcGIS. The user can optionally request that LQGraph produce output files in the JPEG or EMF image formats. As an additional option, LQGraph can produce georeferenced output in ASC format, which can be loaded into the open-source packages MapWindow GIS and Quantum GIS.

4. An example

The following procedure summarizes conservation planning with LQGraph: (i) assemble data layers on the locations of biodiversity surrogates (such as threatened species), (ii) prioritize places to protect the biodiversity surrogates, (iii) input cells into LQGraph, (iv) construct the contiguity graph, (v) find a shortest path between all pairs of conservation areas, (vi) construct the landscape quality graph, (vii) find the connected components of the landscape quality graph, (viii) find the MST(s) and disconnecting set(s) of each component. We illustrate steps (ii) – (viii) with an example from the Eastern Himalayas (India).

4.1. Place prioritization

The data set has 365 347 cells at the $0.01^\circ \times 0.01^\circ$ scale, each containing presence/absence records on 47 biodiversity surrogates (Sarkar et al., 2006). We designed a CAN with the ResNet software package (Garson et al., 2002) to protect 10 % of the sites containing each surrogate using the minimum number of cells. The CAN formed the vertices of the landscape quality graph.

4.2. Constructing the contiguity and landscape quality graphs

We generated contiguity and landscape quality graphs using two models of cell cost (see Table 1). In the first model, the edge weight functions w_C and w_L were based on Euclidean distance so that the edges of the landscape quality graph represented the physically shortest paths between conservation areas. Our second model assigned to each cell a random contiguity and landscape quality score.

4.3. Prioritizing connectivity areas

The MST based on Euclidean distance occupied 5.1 % of the ecoregion (Figure 1). This is 249 893.18 km^2 less than the least-cost paths made up of the edges of the landscape quality graph. The MST based on random quality scores occupied 5.75 % of the ecoregion. This is 72 973.88 km^2 less than the edges of the corresponding landscape quality graph. Thus our prioritization method provides connectivity via high quality contiguity areas in a substantially smaller area than the set of all least-cost paths between conservation areas.

4.4. Disconnecting sets

The minimum disconnecting set of the random landscape quality graph had 1 edge, whereas the disconnecting set based on physical distance had 2 edges. Thus, it is easier to isolate conservation areas in the random model.

5. Final notes

LQGraph provides three extensions to existing methods for the design of CANs:

(i) LQGraph is the only free, stand-alone software for connectivity establishment in a CAN. Two existing software packages can construct least-cost paths between habitat sites: ArcInfo (Rouget et al., 2003) and PATHMATRIX, an extension to ArcView 3.x (Ray 2005). However, when the number of sites to be connected is large, least-cost analysis is inadequate because the paths will take up too much of the landscape. LQGraph provides routines for prioritizing some of the least cost paths via MSTs. Unlike LQGraph, none of the commercial packages finds the minimum disconnecting set(s). SDSS, a plug-in for ArcView (Larson and Sengupta, 2004), has a procedure similar to the condensation function in LQGraph. The former is intended for refining the predicted distribution of a species.

(ii) Compared with existing methods for connectivity establishment, the method implemented here can be used to analyze much large CANs (6 430 conservation areas and 327 298 intermediate sites). Connecting habitat with the MST was first proposed by Urban and Keitt (2001).

(iii) LQGraph can find all MSTs and all disconnecting sets of the landscape quality graph. Providing alternative methods for establishing or removing connectivity is beneficial because real world conservation planning often requires the development of hundreds of alternative plans (Sarkar et al., 2004). Alternative MSTs or disconnecting sets are interchangeable with respect to their connectivity-conferring properties but may differ in other criteria relevant for biodiversity conservation (for example, the cost or human population of their constituent cells). There are always other uses for land proposed for a CAN (agriculture, industrial development, etc.). If planners have several alternative methods for incorporating connectivity in the CAN, they can address these competing demands more adequately by performing multi-criteria analysis to select the best network(s) based on biodiversity representation, connectivity, and other criteria (Sarkar and Garson, 2004).

LQGraph selects optimal chains of sites to connect conservation areas by evaluating the quality score assigned to the each site by the user. This assumes that all species perceive a site in the same way and that a single set of connectivity areas is adequate for all species. A more complex alternative approach would be to design a different set of connectivity areas for each species. If this alternative approach is used, all the different sets of connectivity areas must be integrated to provide a final conservation plan for the region (Williams et al. 2005). Our method is simpler; it is also unclear that enough data exist for the systematic use of the more complex alternative.

References

Banks, S. C., Lindenmayer, D. B., Ward, S. J., Taylor, A.C., 2005. The effects of habitat fragmentation via forestry plantation establishment on spatial genotypic structure in the small marsupial carnivore, *Antechinus agilis*. Molecular Ecology 14: 1667 – 1680.

Bowne, D.R., Bowers, M.A., 2004. Interpatch Movements in Spatially Structured Populations: a Literature Review. Landscape Ecology 19, 1 - 20.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C., 2001. Introduction to algorithms. MIT Press, Cambridge, Massachusetts.

Creegan, H. P., Osborne, P. E., 2005. Gap-crossing decisions of woodland songbirds in Scotland: an experimental approach. Journal of Applied Ecology 42, 678 - 687.

Drinnan, I. N., 2005. The search for fragmentation thresholds in a southern Sydney suburb. Biological Conservation 124, 339 - 349.

Estrada, A., Coates-Estrada, R., 2005. Diversity of neotropical migratory landbird species assemblages in forest fragments and man-made vegetation in Los Tuxtlas, Mexico. Biodiversity and Conservation 14, 1719 - 1734.

Fahrig, L., 2003. Effects of habitat fragmentation on biodiversity. Annual Reviews of Ecology, Evolution, and Systematics 34, 487 -515.

Fuller, T., 2004. The use of graph theory in the design of conservation area networks. M.A. Thesis, University of Texas, Austin.

Fuller, T., Sarkar, S., 2005. LQGraph manual Ver 1.0. Biodiversity and Biocultural Conservation Laboratory, University of Texas, Austin. http://uts.cc.utexas.edu/~consbio/Cons/ResNet.html.

Garson, J., Aggarwal, A., Sarkar, S., 2002. ResNet Ver 1.2 Manual. Biodiversity and Biocultural Conservation Laboratory, University of Texas, Austin. http://uts.cc.utexas.edu/~consbio/Cons/ResNet.html.

Harvey, C.A., Villanueva, C., Villacís, J., Chacón, M., Muñoz, D., López, M., Ibrahim, M., Gómez, R., Taylor, R., Martinez, J., Navas, A., Saenz, J., Sánchez, D., Medina, A., Vilchez, S., Hernández, B., Perez, A., Ruiz, F., López, F., Lang, I., Sinclair, F. L., 2005. Contribution of live fences to the ecological integrity of agricultural landscapes. Agriculture, Ecosystems and Environment 11: 200 - 230

IUCN (International Union for the Conservation of Nature), 1994. World Conservation Strategy: Conservation for Sustainable Development. Worldwide Fund for Nature, Gland.

Larson, B.D., Sengupta, R. R., 2004. A spatial decision support system to identify species-specific critical habitats based on size and accessibility using U.S. GAP data. Environmental Modelling and Software 19, 7 -18.

Levey, D. J., Bolker, B. M., Tewksbury, J. J., Sargent, S., Haddad, N. M., 2005. Effects of landscape corridors on seed dispersal by birds. Science 309, 146 - 148.

Matula, D. W., 1987. Determining edge connectivity in *O*(*nm*). 28th Annual Symposium on the Foundations of Computer Science, 249 -251.

Merriam, G., 1984. Connectivity: a fundamental ecological characteristic of landscape pattern. In Brandt, J., Agger, P., (Eds.), Methodology in Landscape Ecological Research and Planning, Roskilde Universitetsforlag, Roskilde, Denmark:, pp. 5 -15.

Murphy, H. T., Lovett-Doust, J., 2004. Context and connectivity in plant metapopulations and landscape mosaics: does the matrix matter? Oikos 105, 3 - 14.

Onal, H., Briers, R. A., 2002. Incorporating spatial criteria in optimum reserve network selection. Proceedings of the Royal Society of London B 269, 2437 - 2441.

Ray, N., 2005. PATHMATRIX: A geographical information system tool to compute effective distances among samples. Molecular Ecology Notes 5, 177 - 180.

Rouget, M., Cowling, R. M., Pressey, R. L., Richardson, D. M., 2003. Identifying spatial components of ecological and evolutionary processes for regional conservation planning in the Capetown Floristic Region, South Africa. Diversity and Distributions 9, 191 - 210.

Sarkar, S., Aggarwal, A., Garson, J., Margules, C., Zeidler, J., 2002. Place prioritization for biodiversity content. Journal of Biosciences 27, 339 - 346.

Sarkar, S., Mayfield, M., Cameron, S., Fuller, T., Garson, J., 2006. Conservation area networks for the Indian ecoregion: systematic methods and future prospects. In: Stevenson, R., Ganeshaiah, K.N., Bawa, K. (Eds.), Ecoinformatics for Biodiversity Conservation: Prospects and Challenges. I.K. International Private Limited, Bangalore, India.

Sarkar, S., Moffett, A., Sierra, R., Fuller, T., Cameron, S., Garson, J., 2004. Incorporating multiple criteria into the design of conservation area networks. Endangered Species Update 21, 100 - 107.
Sedgewick, R., 1983. Algorithms. Addison-Wesley, Reading, Massachusetts.

Siek, J. G., Lee, L.-Q., Lumsdaine, A., 2001. The BOOST graph library. Addison-Wesley, Boston.

Urban, D., Keitt, T., 2001. Landscape connectivity: a graph-theoretic perspective. Ecology 82, 1205 - 1218.

Williams, J. C., 1998. Delineating wildlife corridors with multiobjective programming. Environmental Modeling and Assessment 3, 77 - 86.

Williams, P., Hannah, L., Andelman, S., Midgley, G., Araújo, M., Hughes, G., Manne, L., Martinez-Meyer, E., Pearson, R., 2005. Planning for climate change: identifying minimum dispersal corridors for the Cape Proteaceae. Conservation Biology 19, 1063 - 1074.

Figure Captions

Figure 1: **Minimum Spanning Tree of the Eastern Himalayas Landscape Quality Graph found using LQGraph**.

The cells in black represent a conservation area network ("CAN") that protects 10 % of the sites with 47 biodiversity surrogates. The gray cells ("MST") link the conservation areas via high-quality connectivity areas containing as few sites as possible. Scale: $0.1^\circ \times 0.1^\circ$. (The connectivity areas are difficult to see at the $0.01^\circ \times 0.01^\circ$ scale). The inset shows the study region (in black), which consists of the World Wildlife Federation ecoregions that intersect with the eastern Himalayas in India.

Table 1

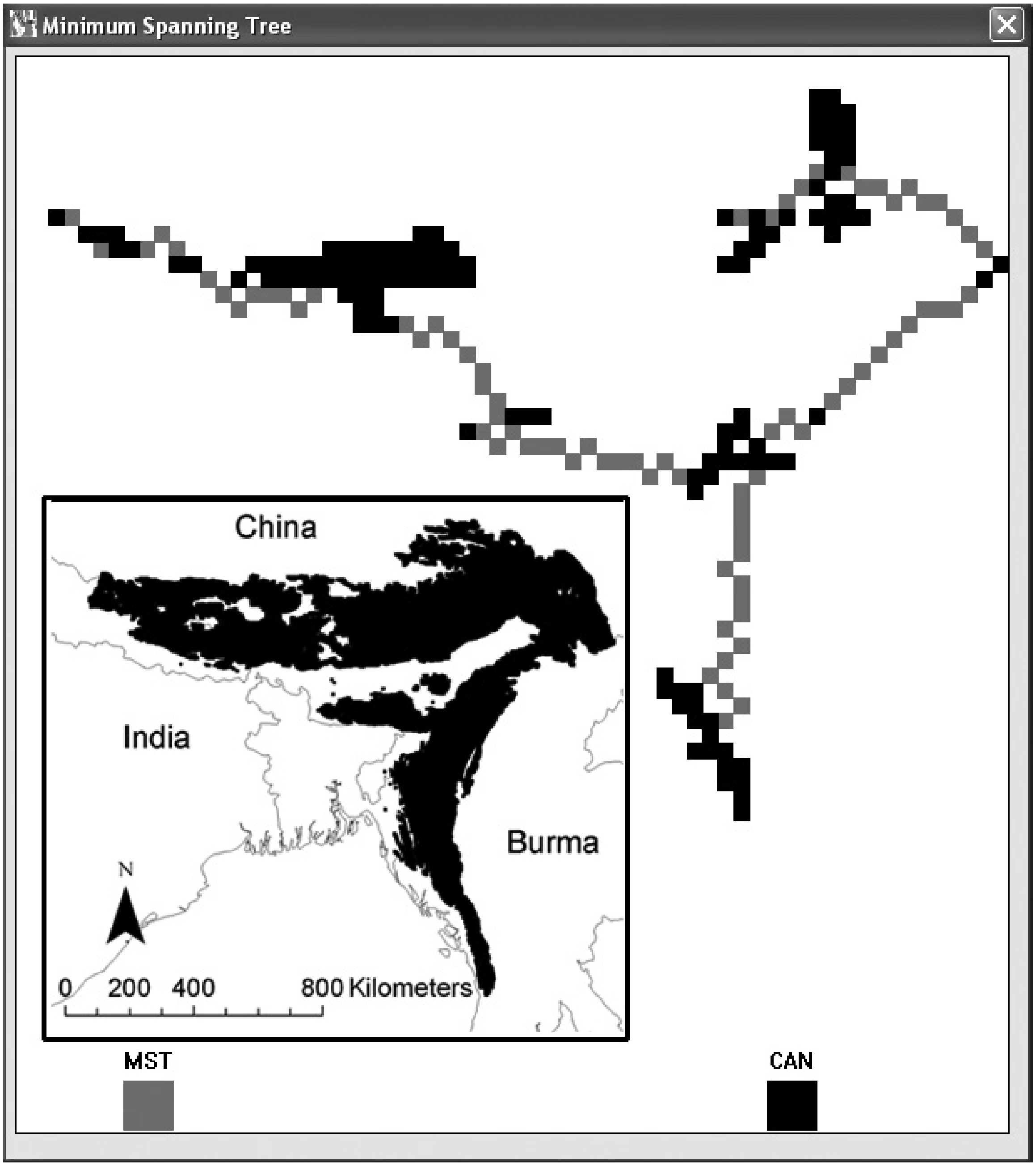
Graph-Theoretic Analysis of the Eastern Himalayas Ecoregion*

Structure	**Vertices**	**Edges**	**Edge area** km^2 **(% of ecoregion)**
Contiguity graph	365 347 same	2 819 420 same	401 834.03 (100) same
Landscape quality graph	6 430 6 430	158 333 135 147	270 376.51 (67.29) 96 074.59 (23.91)
MST**	6 409 6 405	6 408 6 404	20 483.32 (5.1) 23 100.71 (5.75)
Disconnecting set**	3 2	2 1	31.87 (0.000079) 16.73 (0.000042)

*The value in the upper (lower) panel of each cell is for the Euclidean distance (random) model.

**The values are for the largest component of the landscape quality graph.

Figure 1


Minimum Spanning Tree
China
India
Burma
N
0 200 400 800 Kilometers
MST
CAN