



O.R. Applications

# Forest structure optimization using evolutionary programming and landscape ecology metrics

Henry David Venema <sup>a,\*</sup>, Paul H. Calamai <sup>b</sup>, Paul Fieguth <sup>b</sup>

<sup>a</sup> *International Institute for Sustainable Development, 161 Portage Avenue, 6th Floor, Winnipeg, Canada R3B 0Y4*

<sup>b</sup> *Department of Systems Design Engineering, University of Waterloo, Canada N2L 3G1*

Received 19 December 2001; accepted 3 November 2003

Available online 1 February 2004

## Abstract

A new design framework is introduced for forest structure optimization based on a genetic algorithm landscape encoding and landscape ecology metrics. Landscape ecology is an important interface between the forest management community, which is a traditional user of operations research methods, and the biological conservation community which is relatively new to OR methods and whose goals are increasingly allied with the spatial ecology concepts emerging from landscape ecology. Deforestation and forest fragmentation are increasingly recognized as the underlying drivers of global biodiversity loss, hence forestry management will need to explicitly incorporate spatial ecology objectives.

A deforestation model is presented which simulates a landscape progressively fragmenting by the incremental removal of forest patches. Principal components analysis (PCA) of multiple deforestation simulations captures the relative influence of the mean proximity index and the mean nearest neighbour distance, two widely used landscape ecology metrics. An evolutionary programming method based on a genetic encoding of landscape structure is used to optimize forest patch selection by maximizing landscape performance with respect to single and multiple landscape ecology metrics weighted according to the PCA.

This optimization approach is envisioned as a key component of a new forestry OR paradigm for designing multi-use landscape systems, incorporating both biodiversity and community needs.

© 2003 Elsevier B.V. All rights reserved.

*Keywords:* Genetic algorithms; Developing countries; Deforestation; Landscape ecology; Biodiversity

## 1. Introduction

The modern consensus between landscape ecologists, conservation biologists and, increasingly,

forest managers asserts that habitat loss and forest fragmentation are the underlying cause of biodiversity loss (Forman, 1995; Forman and Collinge, 1995; Collinge, 2001). A major research challenge linking the forest management and the conservation biology communities is the development of forest management strategies compatible with basic spatial ecology concepts, compatible with remotely sensed

\* Corresponding author. Tel.: +1-204-958-7706; fax: +1-204-958-7710.

E-mail address: [hvenema@iisd.ca](mailto:hvenema@iisd.ca) (H.D. Venema).

data, and relevant to the sustainable development aspirations of the billions of people world-wide dependent on local ecosystems for their subsistence (WRI, 2000). The summary report of The World Commission on Forests and Sustainable Development (IISD, 1999, p. 2) urged forestry managers to adopt an integrated, landscape perspective, “to plan for the use and protection of whole landscapes, not the forest in isolation”.

Martell et al. (1998) reviewed conventional forestry operation research (OR) practice and concluded that OR methods have proven very successful where industrial timber production is the primary concern. However where the protection of rare ecological niches and ecosystem integrity in the face of anthropogenic pressure are primary concerns, the scope for expanding the contribution of OR methods is large. Martell et al. note, for example, that the literature is largely devoid of OR applications to community forestry in the developing world, where management initiatives will typically be small, decentralized and where forest-dependent communities will perceive the forest from a more holistic, multi-use perspective. As the forests and landscapes of the tropical developing world become an increasingly contested resource for carbon sequestration, for biomass energy, for timber, for agricultural land, and as repositories of biodiversity and traditional ecological knowledge (Pandey, 2002), the need for robust design optimization concepts becomes particularly acute.

This paper presents a new landscape planning approach for biodiversity conservation integrating quantitative landscape ecology principles within a spatial optimization framework. Unlike traditional forest management, the objective here is not to identify an economically optimal harvest strategy, but rather to address the essential conservation issue of identifying, protecting, and integrating a network of forest features into a multi-use landscape system. The proposed methodology is compatible with remotely sensed data, an increasingly ubiquitous data source for analysing forest ecosystem health (Heilman et al., 2002) and envisioned as an analytic component of a participatory community forestry planning process that uses traditional and community ecological knowledge to select between alternative landscape

configurations (Martell et al., 1998; Gadgil et al., 1998). The spatial optimization framework is implemented using a genetic algorithm to select among alternatives the configuration of forest features which optimizes the value of a single or linear combination of landscape ecology metrics.

## 2. Spatial ecological design and operations research: A brief review

Diamond (1975) provided the seminal link between spatial pattern and biodiversity, a concept now firmly entrenched in species meta-population dynamics and landscape ecology. Diamond identified the problem of designing biodiversity conservation reserves as one fundamentally governed by geometric design principles and codified these principles as shown in Fig. 1. Ecosystems can

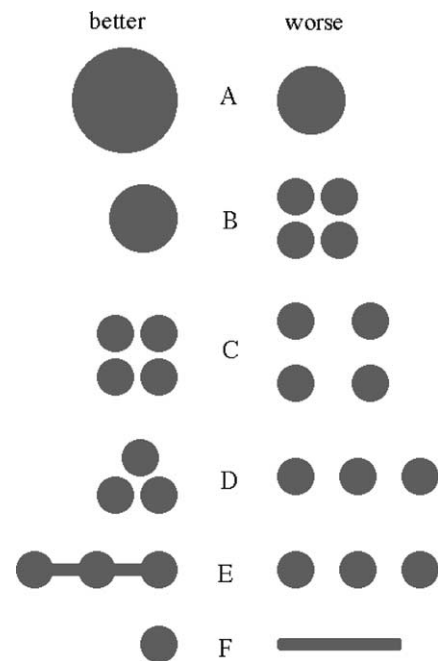


Fig. 1. Diamond's suggested geometric design principles for nature reserves. Case A: large is better than small habitat of the same shape; Case B: contiguous is better than dispersed habitat of the same size; Case C: compact is better than dispersed habitat of the same size; Case D: compact is better than unconnected corridor habitat; Case E: connected corridors are better than unconnected corridors; Case F: compact is better than linear habitat of the same size.

support species populations in either high-density connected or low-density fragmented states. The population of a particular habitat element in a landscape composed of many potential habitats, depends on its neighbouring sites. Once a population at any particular site is extinct, the likelihood of recolonization depends on the aggregate population and proximity of neighbouring sites, producing a positive feedback effect between population density and the resilience of the local population (Hanski et al., 1995). Diamond's hierarchy of principles is not without contention, however, and the "single large or several small" (SLOSS) nature reserve debate remains current in contemporary conservation biology (Possingham et al., 2000). Diamond's ideas have nonetheless had a large influence on the evolution of quantitative landscape ecology (Baschak and Brown, 1994) and has motivated increasingly sophisticated research on landscape fragmentation measurement methods and meta-population dynamics as a function of landscape fragmentation (Verboom et al., 2001).

Despite Diamond's early recognition of the biodiversity conservation problem as fundamentally a spatial design problem, the use of operations research (OR) techniques for optimal landscape design is still in its infancy (Kurtilla, 2001; Possingham et al., 2000). Robust design concepts applicable when remotely sensed data and community participation are the only available inputs for planning forestry management and conservation are a particularly pressing need. Fjelsda (2000) describes conservation biology as a crisis discipline, where urgency and prohibitive costs require decision making with incomplete or proxy data at best. Hof and Bevers (1998) present elegant and complex habitat optimization models that require a high level of species-specific parameterization, and are therefore of limited practicality in frontier forest regions of the developing world where deforestation rates and biodiversity loss are high but detailed ecological surveys scarce. Fjelsda also notes that access to the best possible ecological information will not improve conservation practice if the livelihood issues of the traditional rural communities that share the landscape are not integrated into the plan.

The advent of the landscape-based biodiversity conservation paradigm coincides with an increasing, though still limited, integration of spatial ecology concepts within conventional forestry operations research. Forestry OR applications typically maximize net present value of the harvested timber, without regard to ecological considerations or their proxy, spatial forest structure. Biological conservation concerns have been integrated into these harvest models through adjacency constraints which restrict the harvest of any two spatial forest management units that share a common boundary. Although many variations of adjacency-constrained forest harvest optimization have appeared in the literature, they all share a similar integer linear programming (ILP) formulation and are based on harvest timing decisions; when to harvest each of many individual forest stands over a multi-period planning horizon. Practical sized problems generate a large number of adjacency constraints and limit the application of conventional ILP solution techniques, which has motivated researchers to explore alternative solution heuristics, including Tabu search and simulated annealing (SA) (Kurtilla, 2001). Relegating conservation objectives to a set of adjacency constraints within such models may be appropriate in the context of a homogeneous plantation system of uniform stands with similar ecological value. Such an approach, is of decidedly lesser value in the context of a heterogeneous network of forest patches with varying ecological value and multiple uses—a situation typical of developing country regions under anthropogenic deforestation pressure.

The use of OR methods for capturing spatially-based biodiversity conservation objectives is somewhat better developed in the conservation biology literature where a typical management objective is to identify a network of reserves containing the largest possible subset of threatened and endangered species and prioritize these critical sites for protection. Pressey et al. (1997) present an integer linear programming (ILP) formulation (the standard set-covering problem) of the biodiversity reserve selection problem, minimizing land acquisition costs, while representing at least one occurrence of the habitat

range of a set of threatened species. This approach assumes a priori knowledge of how the species requiring protection are spatially distributed (Reid, 1998).

Possingham et al. (2000) introduced fragmentation concepts to the reserve selection problem by extending the set-covering problem by introducing a second objective; total boundary length minimization of all reserves selected. This formulation preferentially selects contiguous candidate reserves as opposed to isolated “island” reserves. Possingham et al. applied an SA algorithm to solve their multi-objective reserve selection formulation since the boundary length minimization objective violates the strict integer and linearity conditions required for standard ILP solution methods.

The landscape optimization methodology developed in this paper uses quantitative landscape ecology principles. Quantitative landscape ecology assumes a strong link between spatial pattern and ecological function and process, but focuses on measuring forest fragmentation and inferring from the degree of fragmentation, forest ecosystem health, and biodiversity state. Spatial heterogeneity, which represents ecological and function and process (Gustafson, 1998), is quantified by identifying relatively homogeneous landscape patches and examining the statistics of these patches (Forman, 1986; O’Neill et al., 1988; Gustafson and Parker, 1992; Li and Reynolds, 1993; Marks and McGarigal, 1994). Landscape ecology metrics all attempt to capture various aspects of the geometric design principles proposed by Diamond.

Kurtilla observes that although several simulation studies have examined how the impact of varied forest harvest strategies as measured by landscape ecology metrics (Barrett et al., 1998; Gustafson and Crow, 1996) no optimization strategy explicitly based directly on landscape ecology concepts has yet emerged. The optimization approach formulated here assumes no a priori knowledge of species distribution or metapopulation dynamics (as in the reserve selection problem) but is based strictly on the spatial configuration of a set of habitat elements as quantified using landscape ecology metrics.

### 3. Methodology: A deforestation simulation model

#### 3.1. Rationale

Deforestation by incremental degradation is widely acknowledged as the basic process undermining biodiversity. A simulation model of a forested landscape progressively losing structure and integrity from the incremental removal of forest patches is therefore developed to gain insight into the basic dynamics of biodiversity loss in tropical forests of the developing world. The simulation model reveals how landscape ecology metrics respond dynamically to the deforestation signal. We then use the observed metric behavior to develop optimized mitigative strategies in the second stage of the methodology. For this illustrative application we use a small set of landscape ecology metrics, however the methodology is general and can readily be extended to a larger set of metrics. We also introduce the use of principal components analysis (PCA) to analyse the response signal from the landscape ecology metrics to the deforestation process.

Liu et al. (1993) and Ramesh et al. (1997) analysed the deforestation process in the Philippines and the Western Ghats of India. Both studies describe a basic deforestation process characterized by the removal of the smallest and most accessible forest patches, followed by incremental extension of access roads, further encroachments and plantation expansions, generating in turn more fragmentation and eventual deforestation. Forest patches with high perimeter to area ratio’s ( $P/A$ ) were found in both studies to be extremely susceptible to degradation and eventual removal. For comparative purposes we developed two simple models of the deforestation process with different deforestation dynamics; incremental fragmentation by random removal of forest patches, alternatively successive forest patch removal by highest  $P/A$  ratio. Clearly more elaborate deforestation models could be conceived which, for example, include the effects of infrastructure development as observed by Liu et al. However, for purposes of this paper we deemed the simple deforestation model described above more intuitive and appropriate.

3.2. Deforestation signal processing with landscape ecology metrics

A binary forest/non-forest map of Waterloo County, Ontario, Canada derived from Landsat Thematic Mapper (TM) imagery at 25 m pixel resolution served as the base landscape for this study. The upper left panel in Fig. 2 shows the original base landscape. The panel dimensions for all landscape configurations shown in this paper are 10.475 km (north–south) by 10.225 km (east–west).

Quattrochi and Pelletier (1991) review ecological scale and data resolution issues and recommend TM-scale data for local-scale investigations of heterogeneous landscapes such as the base landscape considered here. Although large-scale landscape fragmentation analyses have typically used 1-km resolution AVHRR (advanced very high resolution radiometer) imagery (O’Neil et al., 1997; Riitters et al., 2000) recent attempts at comprehensive fragmentation analyses have used FRAGSTATS metrics and classified Landsat TM imagery (Heilman et al., 2002). Hargis et al. (1998)

also used TM-scale data for their study of landscape ecology metric behavior using simulated landscapes. This simulation study, however, used a single base landscape but simulated the deforestation process by sequentially removing individual forest patches. The simulated deforestation signal thus generated was analysed by calculating four common landscape ecology metrics as each patch was removed.

Fig. 2 illustrates three additional stages of the simulated deforestation process for single simulation sequence of patch removal by highest patch  $P/A$  ratio. One hundred different deforestation sequences were also simulated for the random patch removal deforestation mode. The following metrics were re-calculated on every patch removal for all simulated deforestation sequences; the mean nearest neighbour distance (MNN), the mean proximity index (MPI), the mean shape index (MSI), and the area weighted mean shape index (AWMSI). MNN and MPI are key landscape metrics and estimate the relative isolation of all patches on the landscape (Hargis et al., 1998, 1999; Marks and McGarigal, 1994). The other metrics, MSI and AWMSI, illustrate how measure patch shape complexity can be included within a planning and design framework. Patch shape metrics relate patch area and patch perimeter and have important implications in terms of the quality of available habitat. Depending on the species of concern, it may, for example, be desirable to maximize the occurrence of irregular or elongated patches to maximize the available edge habitat (Sisk and Haddad, 2002).

The definitions and implementation of the four metrics utilized in this study conform to those used in FRAGSTATS (Marks and McGarigal, 1994), a widely used landscape ecology software package.

Suppose we have a landscape comprised of  $n$  patches, with an area  $a_i$  and perimeter  $p_i$ , and with a minimum edge-to-edge distance between patches  $i$  and  $j$  given by  $d_{i,j}$ , we can define MPI, MNN, MSI, AWMSI metrics as follows:

The mean proximity index

$$MPI(d) = \frac{1}{n} \sum_{j=1}^n \sum_{i \in M_j} \frac{a_{ij}}{d_{ij}^2} \tag{1}$$

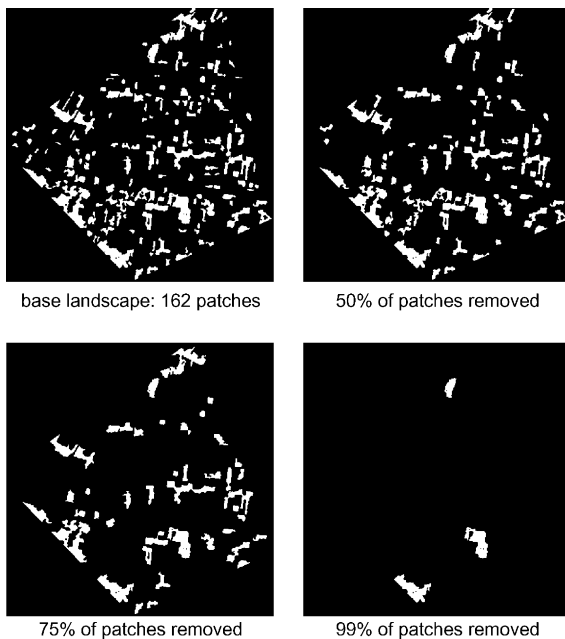


Fig. 2. Largest perimeter/area deforestation mode.

where  $n$  is the total number of patches in the landscape and for each patch  $j$ ,  $a_{ij}$  is the area of the  $i$ th patch of the set  $M_j$ , where  $M_j$  is defined as the set of patches within the threshold distance,  $d$  of the  $j$ th patch ( $d_{ij} \leq d$ ). If all patches have no neighbours within the threshold distance,  $d$ , then  $MPI = 0$ . MPI is dimensionless and increases when nearby patches become larger and less fragmented in distribution (Marks and McGarigal, 1994).

#### The mean nearest neighbour distance

$$MNN = \frac{1}{n} \sum_{j=1}^n \min d_{i,j}, \quad i \neq j. \quad (2)$$

The MNN has dimensions of length, a lower bound of zero and increases without bound as nearest patches become increasingly isolated.

#### The mean shape index (raster version)

$$MSI = \frac{1}{n} \sum_{j=1}^n \frac{0.25 * p_j}{\sqrt{a_j}}. \quad (3)$$

MSI (dimensionless) is minimized if all patches are the simplest possible shape (squares) and increases without bound as patches become more irregularly shaped.

#### The area weighted mean shape index (raster version)

$$AWMSI = \sum_{j=1}^n \frac{0.25 * p_j}{\sqrt{a_j}} \frac{a_j}{a_{\text{total}}} \quad (4)$$

where  $p_j$  and  $a_j$  are the perimeter and area of the  $j$ th patch, respectively, and  $a_{\text{total}}$  is the summed area of all patches ( $\sum_{j=1}^n a_j$ ). The AWMSI (dimensionless) is also minimized if all patches are the simplest possible shape (squares in the raster data case) and also increases without bound as patches become more irregularly shaped.

### 3.3. Simulation results and principle components analysis

The random patch removal deforestation model used a Monte Carlo simulation approach with patches randomly removed from the base landscape until no patches remained. The process was

repeated 100 times. Fig. 3 illustrates the individual traces for each of the 100 deforestation sequences using random patch removal for the MNN (upper panel), and the MPI (lower panel). The four metrics calculated at each patch removal (MNN, MPI, MSI, and AWMSI) were then normalized by their means. Figs. 4 and 5 show simultaneous plots of the normalized metrics for the random patch removal (averaged over all 100 scenarios), and the largest  $P/A$  patch removal deforestation modes, respectively. Although both the MPI and MNN show a jagged, erratic behavior over any particular deforestation sequence, a smooth trace emerges when averaged over 100 different deforestation simulations (Fig. 4). This result has not yet been reported in the landscape ecology literature and certainly suggests that such Monte Carlo-based methods in landscape ecology research may be warranted.

Fig. 5 shows the same normalized metrics as in Fig. 4 for the single deforestation sequence defined by the largest  $P/A$  patch removal. The MNN and MPI show a jagged behavior in this simulations similar to that observed in any individual random patch removal simulation (Fig. 3). Although the shape-based metrics, MSI and AWMSI also show some slight variability across an individual trace (Fig. 5), where multiple simulations are available (Fig. 4), the mean MSI and AWMSI traces are essentially flat. The patch shape distribution of the original landscape determines the behavior of the MSI and AWMSI and is not influenced by the deforestation dynamic.

Landscape ecology studies using dozens of metrics have employed principle components analysis (PCA) to reduce data dimensionality (Griffith et al., 2000). Typically PCA is used to simplify comparisons of static landscapes. This study, however, demonstrates the use of PCA to reduce the dimensionality of the deforestation signal measured dynamically using landscape ecology metrics. In general, PCA transforms a set of correlated variables to an uncorrelated set. PCA is essentially eigen analysis of the covariance structure of the original data (Jackson, 1991; Richards and Jia, 1999). The  $n$  eigenvectors of the  $n$  by  $n$  covariance matrix define  $n$  principle components. The first principle component explains

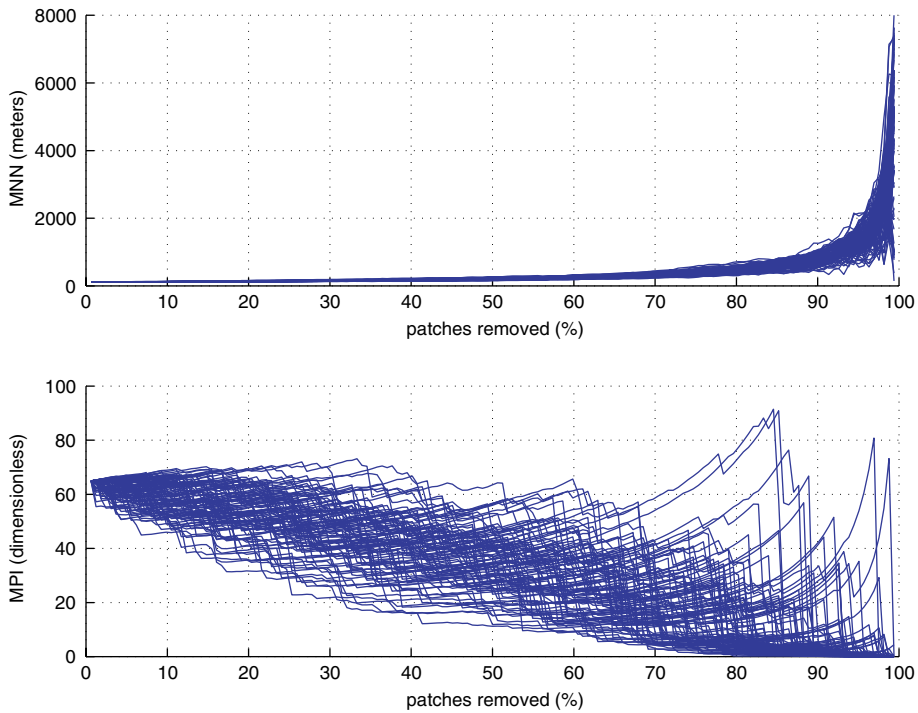


Fig. 3. Simulation-based deforestation signal analysis using the MNN (upper panel) and the MPI (lower panel). The MNN increases fairly smoothly and non-linearly as patches are removed. MPI behavior is erratic for individual traces but by definition must be zero when all patches are removed.

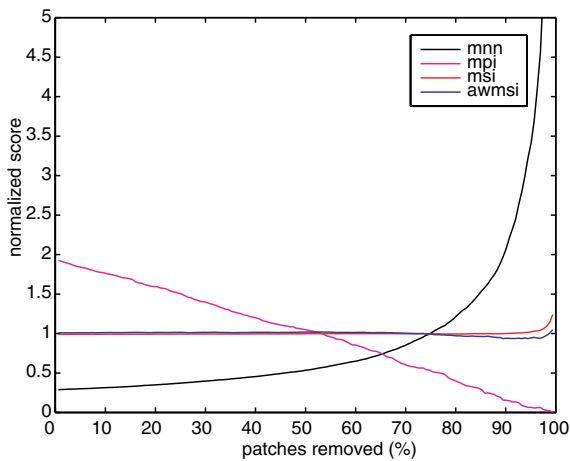


Fig. 4. Normalized landscape ecology metrics, averaged over all random patch removal scenarios. The MNN trace averaged over all 100 simulations shows a smooth, non-linear increase. The averaged MPI trace is smooth and decreases linearly to zero. The averaged MSI and AWMSI metrics are not influenced by the deforestation process.

the most variance in the dataset with subsequent principal components explaining successively less variance. The principal components are defined by linear combinations of the original  $n$  variables.

The normalized data in Figs. 4 and 5 were transformed using PCA. Alternative data scaling methods could have been used in the PCA, however simple normalization about a unit mean is, arguably, the least manipulative of the raw simulation data. Fig. 6 illustrates the exceedance probability of the first component variance for the 100 different random deforestation sequences.

Table 1 shows the variance distribution across all four principle components for the  $P/A$  ratio deforestation mode. Inspecting Fig. 6 and Table 1 indicates that the overwhelming majority of variance for both deforestation modes is explained by the first principal component.

Table 2 lists the linear combination of component weights that define the first principal

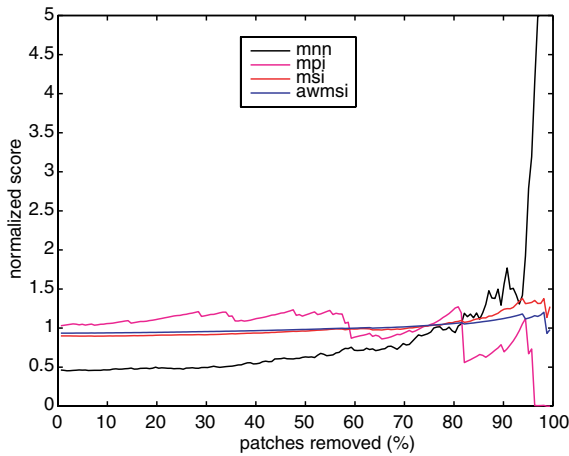


Fig. 5. Normalized metrics for deforestation by  $P/A$  ratio. The single MNN and MPI trace from this deforestation scenario are similar to those from any individual random patch removal scenario in Fig. 3. MNN increases fairly smoothly and non-linearly, while the MPI trace is erratic, reaching zero as all patches are removed. The MSI and AWMSI traces are essentially flat and influenced only by the patch distribution of the original landscape and not the deforestation process.

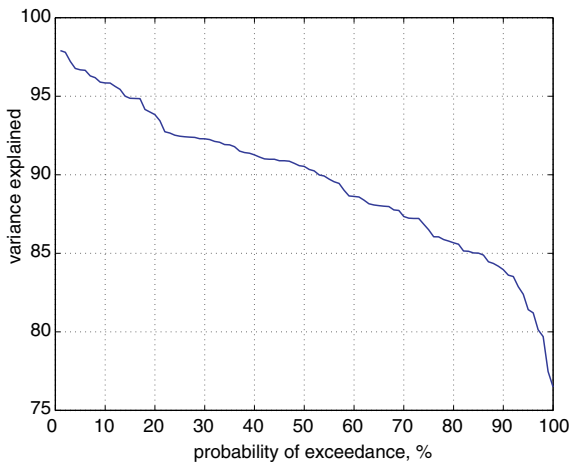


Fig. 6. Random patch removal: first component variance.

Table 1  
Variance distribution for deforestation by largest  $P/A$  patch removal

Principle component	1	2	3	4
Variance explained (%)	98.26	1.32	0.41	0.01

Table 2  
First principle component weights

	MNN	MPI	MSI	AWMSI
Random patch removal	0.9448	-0.3021	0.0128	-0.0118
Largest $P/A$ patch removal	0.9889	-0.1383	0.0508	0.0177

components for both deforestation modes. The component weights for the random patch removal case are the average for the 100 random deforestation scenarios. The very high information yield from the MNN metric is obvious. The first principal component is very heavily weighted by the MNN for both deforestation modes. Fig. 7 illustrates the original data projected along the first principal component compared with the MNN trace for both deforestation modes. The random patch removal deforestation mode shows a slight deviation from the MNN trace—evidence of a weak influence from the MPI component. For the  $P/A$  ratio deforestation mode, the first principal component projection and the original MNN trace are virtually indistinguishable.

The general result of the simulation and PCA analysis is that, of the four landscape ecology metrics considered, MNN and MPI metrics provide a very good estimate of the level of deforestation for both deforestation modes. An intervention design strategy intended to mitigate deforestation and forest fragmentation can therefore utilize this information in formulating objectives as illustrated in the following section.

#### 4. Genetic optimization of forest structure

##### 4.1. Rationale

An important conceptual link between the simulated deforestation process and the landscape optimization framework, is that if a set of landscape ecology metrics can capture the essential dynamic of a landscape progressively losing structure, then optimization with that set of metrics can be used to mitigate the deforestation process by optimally selecting the set of landscape



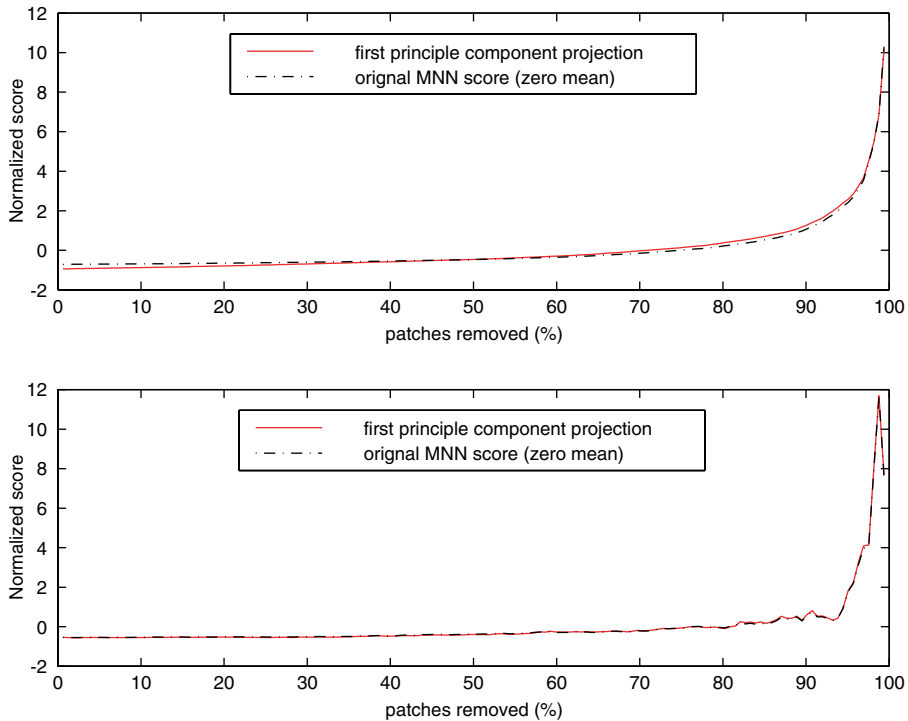


Fig. 7. First principle component projections: random patch removal (upper panel) and largest  $P/A$  patch removal (lower panel).

features for protection and biodiversity conservation. Essentially, measuring the deforestation signal (illustrated in Figs. 3–7) motivates a design optimization approach to identify critical forest landscape elements for protection.

Lindenmayer et al. (1999) provide particularly compelling evidence that retaining remnant forest fragments in a multi-use landscape is critical because the proximity and size of forest patches do reflect actual ecosystem function. Lindenmayer et al. conducted a large study in a 100,000 ha region of predominantly softwood lumber plantation in south-eastern Australia. They surveyed the presence and abundance of 17 different mammalian species in three distinct landscapes and showed that both remnant patch size and relative proximity (landscape isolation characteristics captured by the MNN and MPI metrics) were key determinants of species abundance. Lindenmayer et al. stress that their work has important implications for plantation design; any attempt to

expand plantations (a particularly ubiquitous form of deforestation in developing countries) *must not* clear remnant fragments, particularly larger ones as they are most likely to provide habitat. Christian et al.'s (1998) studies of species occurrence around bioenergy plantation systems in the United States suggest a similar basic dynamic: the existence and protection of remnant native forest patches within the plantation system is extremely important for ensuring habitat maintenance.

The key design implication is that a multi-use landscape system designed with biodiversity conservation objectives should integrate, to the maximum extent possible, a network of remnant forest fragments selected on the basis of their relatively proximity and size. We formulate the essential design problem of optimally identifying a set of existing ecological features for integration within a multi-use landscape using a class of stochastic optimization heuristics known as *genetic algorithms*.

#### 4.2. Genetic algorithm overview

Evolutionary programming techniques such as simulated annealing (SA) and genetic algorithms (GAs) are gaining popularity in a wide range of operations research and engineering design applications (Holland, 1992; Reggiani et al., 2001), primarily because they are well suited to difficult combinatorial problems (such as the set-covering problem). GAs are well suited to parallel computing methods, and produce a ranked hierarchy of feasible solutions as opposed to standard linear programming (LP) and ILP methods which produce a single solution. Ensuring global optimality remains computationally challenging and impractical for many problems.

GAs manipulate a set of candidate solutions, referred to as a *population of individuals* and generate a new population at each iteration of the algorithm. Each individual candidate solution is typically represented as a string of decision variables. The string encoded scheme is context-specific and is usually based on discretized representation of the decision variables as either integers or binary digits. The values of the decision variables are manipulated by subjecting the current population of individual strings or *chromosomes* to a set of standard *genetic operators* that are inspired by Darwinian evolution theory and are referred to as *reproduction*, *crossover*, and *mutation*. The operators involve only random number generation, preferential selection, string copying, and partial string exchanging. The repeated application of these operators, however, evolves an ever improving population of candidate solutions. When some optimization criterion is reached (usually based on the number of algorithm iterations or *generations*), the algorithm terminates.

##### 4.2.1. Population encoding and fitness evaluation

The design problem we address in this paper using several illustrative examples is the optimized selection of a subset of forest patches from the base landscape (shown in the upper left panel in Fig. 2) using landscape ecology metrics and a GA binary string encoding approach. The initial population was generated with each individual's string length equal to the total number of patches in the

original landscape. The position of  $i$ th binary digit (1 or 0) in the string represents the inclusion or exclusion of the  $i$ th patch of the  $I$ th candidate landscape. In essence, each string represents a candidate landscape. A string of all 1's represents the original full landscape, whereas a string of all zero's is a null landscape with all forest patches removed. The relative performance or "fitness" for each candidate landscape is then evaluated by a particular landscape ecology metric or linear combination of metrics. High performing candidate landscapes are preferentially retained and subjected to the standard GA operators: reproduction, crossover, and mutation.

*4.2.1.1. Reproduction.* The reproduction operator preferentially selects high performing candidate strings (in this case landscapes) in any of several ways. One of the most common reproduction schemes, and the approach implemented here, is Monte Carlo-based *roulette wheel selection*. Individuals are extracted from the current population with a probability proportional to the ratio of the individual's fitness to the aggregate fitness of the total population:

$$p(I) = \frac{f(I)}{\sum_{I=1}^{n_{pop}} f(I)} \quad (5)$$

where  $n_{pop}$  is the population size. Constrained problems can be handled by assigning a large negative value to individuals who violate any of several constraints and then adding the magnitude of the largest negative fitness value to all individuals of the current population. Feasible candidates will then have large positive values and infeasible candidates will have zero or very small fitness values, thus rendering their selection probability near zero.

*4.2.1.2. Crossover.* Crossover is the principal GA search mechanism and involves exchanging substrings of randomly paired individuals selected in the previous *reproduction* stage. Each crossover operation generates two new individuals that inherit information (in this case a substring of landscape patch indices) from the parenting individuals. Fig. 8 illustrates a two-point crossover operation.

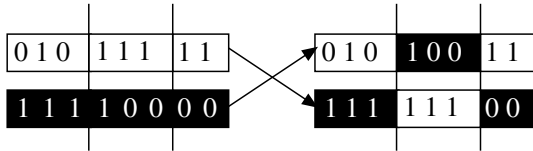


Fig. 8. Two-point GA crossover operation.

**4.2.1.3. Mutation.** The mutation operator has the task of maintaining genetic diversity within the population by ensuring that the entire search space retains some finite probability of being searched. A mutation operator is typically implemented by allowing every design variable to randomly change value with some low probability. In the context of the binary encoded landscape GA implemented here, every candidate landscape could be randomly mutated by changing the bit value of its component patches, essentially randomly turning on or off the inclusion of the *i*th patch.

**4.3. Optimization results**

The genetic algorithm approach is very well suited to extremely complex, non-linear, combinatorial design problems such as landscape ecology-based spatial design. The GA approach is flexible and a wide variety of constraints can be easily modeled. This new GA framework for landscape ecology-based forest structure optimization is illustrated with the following objective function formulations:

- MNN minimization; unconstrained, and with area constraints.
- MNN constrained by MSI.
- PI maximization with varying threshold distance.
- Constrained edge habitat maximization.
- PCA-based optimization.

All the solution landscapes are presented in Figs. 9–13 and include interpretation with respect to the geometric design principles presented in Fig. 1.

**4.3.1. MNN minimization with area constraints**

Given the preceding PCA analysis, the MNN metric evidently provides a clear signal of the deforestation process as modeled by sequential

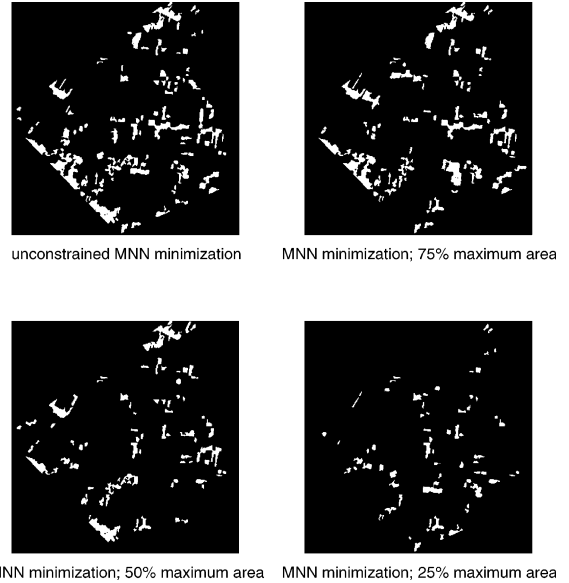


Fig. 9. *Unconstrained MNN minimization:* The solution landscape includes large patches because of their proximity to many small patches resembling Diamond’s case C, where compact configurations are preferred to dispersed. *MNN minimization; 75% maximum area:* constraint. The solution landscape contains fewer large patches than the unconstrained case, those that are retained are more closely associated with clusters of small patches. The solution landscape still favours compact configurations over dispersed and corridor features (Cases C and D). *MNN minimization; 50% maximum area:* The solution landscape now contains only a few large patches and with the total area constraint exhibits more corridor behavior (Cases D and E). *MNN minimization; 25% maximum area:* A break into a distinctly different regime is clearly evident; the solution landscape retains none of the largest patches and forces corridor behavior (Cases D, E, and F).

patch removal (the dominant mode of forest loss in developing countries). A reasonable objective would thus be to identify the set of existing forest patches that produces the lowest possible MNN and to integrate them into a multi-use landscape. The upper left panel of Fig. 9 shows the resulting optimal landscape for the unconstrained MNN minimization case. An interesting observation here is that a carefully selected subset of patches can have a lower MNN distance than the original full landscape, however the large majority of the original set of patches are still retained.

The unconstrained MNN case, which retains most of the original patches, may not be a realistic

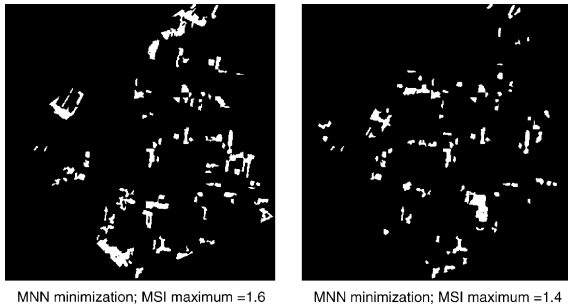


Fig. 10. *MNN minimization; MSI maximum = 1.6*: Compare with Fig. 9; the solution landscape rejects some of the larger complex-shaped patches, favouring more compact patch clusters (Case C). *MNN minimization; MSI maximum = 1.4*: The solution landscape contains few complex patches, favouring clusters and corridors of simple patch shapes (Cases C and D).

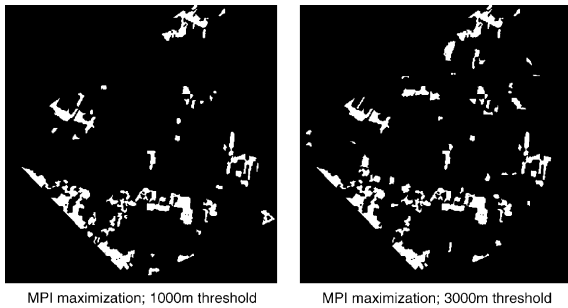


Fig. 11. *MPI maximization; 1000 m focal threshold distance*: The MPI considers both patch size and proximity to other patches; the solution landscape contains many of the largest patches and with the 1000 m focal threshold contains relatively tight clusters (Cases A, B, and C). *MPI maximization; 3000 m focal threshold distance*: The solution landscape is very similar to figure the 100 m case but with more dispersed clusters (Cases A and C).

design condition. The remaining three panels in Fig. 9 illustrate MNN minimization under the constraint that the total area selected must be less than 75%, 50%, and 25% of the original forest area. The constraints were implemented using a penalty function approach within the fitness function calculation of the *I*th landscape:

$$f(I) = -MNN(I) + \text{minimum} \\ (\text{MaxArea} - \text{AreaFraction}(I), 0) * \text{Penalty} \tag{6}$$

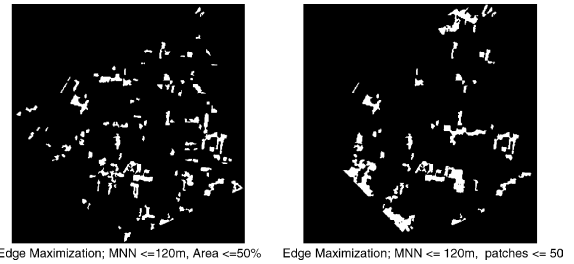


Fig. 12. *Edge habitat maximization; MNN maximum = 120 m, area maximum = 50%*: The solution landscape contains very few large patches and is dominated by clusters and corridors (Cases C and D). *Edge habitat maximization; MNN maximum = 120 m, 50 patches maximum*: The solution landscape is very different from previous case. The constraint on the number of patches favours large patches near clusters of small patches (Cases A and C).

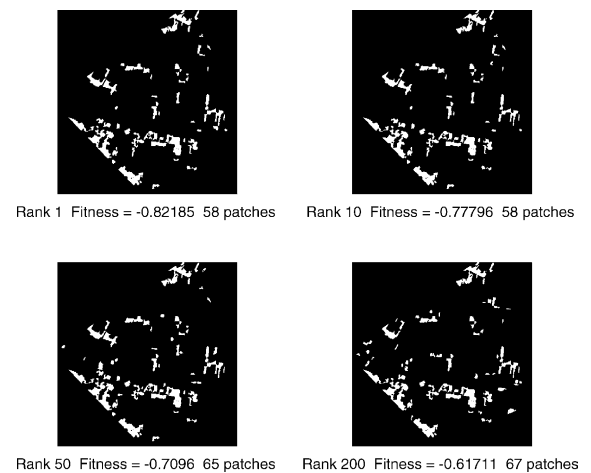


Fig. 13. Principle component weights minimization: alternative solution landscapes.

where MaxArea is the imposed maximum area constraint (a percentage of the original total area), AreaFraction(*I*) is the area percentage of the *I*th landscape, and penalty is a large positive value (1e15). If the constraint is violated, the fitness of the *I*th candidate is a large negative value, and the likelihood of selection in the next generation is very low.

The three area constrained cases show some interesting behavior, reminiscent of the SLOSS (single large or several small) biodiversity reserve design debate that has permeated conservation

biology in recent decades (Possingham et al., 2000). For the 75% and 50% area constrained cases, the optimal solutions exhibit a tendency to cluster around larger patches, a “few large solution”. A transition into a different regime is clearly evident with the 25% maximum area constraint. The solution that minimizes MNN now displays linear and curvilinear features made up of relatively small patches: in essence “corridors” have emerged in this “many small solution”.

#### 4.3.2. MNN minimization with shape constraints

The next two figures depict MNN minimization scenarios subject to constraints on the average patch complexity as measured by the mean shape index (MSI). The fitness function calculation for the  $I$ th landscape is simply:

$$f(I) = -\text{MNN}(I) + \text{minimum} \\ (\text{MaxMSI} - \text{MSI}(I), 0) * \text{Penalty} \quad (7)$$

where MaxMSI is the imposed maximum mean shape index constraint. The first example, the left panel in Fig. 10, shows the imposition of a relatively lax constraint in that the MSI cannot exceed 1.6. This solution has much in common with the unconstrained MNN solution in Fig. 9 except that some of the more complex-shaped patches in the unconstrained case are rejected. The right panel in Fig. 10 shows the solution for a much more stringent case with a maximum MSI constraint of 1.4. The solution landscape in this case displays a tendency towards less complex-shaped patches as well as some curvilinear “corridor” behavior similar to the area constrained case in Fig. 9.

#### 4.3.3. MPI maximization with threshold distance sensitivity

Fig. 11 show optimization scenarios based on MPI maximization with focal distance thresholds at 1000 and 3000 m. The 1000 m case shows the tighter clustering; the MPI value for this scenario is based on the co-occurrence of patches within a 1000 m threshold. The 3000 m case shows a similar but more dispersed landscape structure and includes more distant patches, forming looser clusters, a result consistent with the MPI definition which allows more distance patches to contribute to the MPI calculation.

#### 4.3.4. Constrained edge habitat maximization

The next two examples illustrate the flexibility of the landscape optimization framework. Many conservation planning exercises are concerned with edge and interior habitat. Hof and Joyce (1992), for example, present a mixed integer linear programming model with maximization of edge habitat as an objective. The left panel of Fig. 12 illustrates an edge habitat maximization scenario (as measured by the total perimeter of all patches) subject to a 120 m maximum MNN distance constraint and a 50% maximum total patch area constraint.

The large number of resulting patches (117) may be logistically infeasible, motivating the scenario shown in the right panel of Fig. 12. This scenario also illustrates an edge habitat maximization again subject to a 120 m maximum MNN distance constraint but now also constrained to a maximum of 50 patches. Such design flexibility is particularly important when planning habitat for a target species whose viability is sensitive to the availability of forest edge or forest interior habitat. The design framework could equally accommodate a minimization of edge length with MNN (or MPI) constraints.

#### 4.3.5. PCA-based optimization

The final scenario considered (Fig. 13) illustrates how results from the PCA analysis can be integrated into the optimization framework. In this case, the component weights forming the first principle component for the random patch removal deforestation scenario negatively weight the objective function. The underlying logic advanced here is that if the PCA analysis captures the essential dynamics of a landscape progressively losing structure through incremental patch removal, then optimization should attempt to reverse this underlying dynamic.

The fitness function calculation for the  $I$ th landscape in this case is

$$\text{Min}f(I) = \frac{\text{PCA}_{\text{mnn}} * \text{MNN}(I)}{\text{MNN}_{\text{mean}}} \\ + \frac{\text{PCA}_{\text{mpi}} * \text{MPI}(I)}{\text{MPI}_{\text{mean}}} \quad (8)$$

where the coefficients,  $\text{PCA}_{\text{mnn}}$  and  $\text{PCA}_{\text{mpi}}$  are the average MNN and MPI first principle component

weights from the random patch removal deforestation mode as shown in Table 2. For scaling purposes MNN and MPI are normalized by their means as in Fig. 4. The fitness function should be minimized in this case as depicted in Fig. 7, which shows the behavior of the first principle component.

Four alternative solution landscapes are shown in Fig. 13 and illustrate several features specific to the optimization scenario and the optimization framework in general. Fig. 13 shows 1st, 10th, 50th, and 200th ranked solutions out of, in this case, a database of 1520 landscapes produced when the genetic algorithm terminated. The similarity between all four landscapes is noteworthy, particularly the 1st and 10th which also have the fewest patches. The low patch number can be interpreted as the influence of the MPI in the fitness function. The unconstrained MNN solution landscape (Fig. 9) has, in contrast, 102 patches. The consistency between these alternative configurations suggests that the essence of the conservation strategy should be to focus on the network of small and large patches in the southern third of the landscape along with a few large patches dispersed throughout the landscape.

#### 4.3.6. Future directions

We believe that the ability to probe the landscape database and make quick visual comparisons between alternative configurations, and the ease with which alternative objectives and constraints can be modeled bode well for the integration of landscape optimization within spatial multiple criteria decision models (Malczewski, 1999; Jankowski et al., 2001), particularly in the context of participatory agro-ecosystem planning (Goma et al., 2001) where access to a range of good solutions for further screening is, generally far more important than the optimality of any particular solution (Cocklin, 1989; Pressey et al., 1996). We envision the increased integration of traditional ecological knowledge within the landscape design model for selecting between alternative landscape configurations. Gadgil et al. (1998) cite examples from Asia, Africa, and Mexico where forest-dependent societies traditionally make explicit decisions to protect particular forests

stands for their biodiversity value. Community-based natural resource management decisions such as this will be both ecologically and financially more complicated in the future as forest-dependent societies weigh alternative land and carbon management strategies (Martell et al., 1998; Venema et al., 2000; Pandey, 2002). We believe that landscape-scale forestry OR tools like those introduced in this paper will have much to offer this emerging design paradigm.

## 5. Computational issues

The landscape optimization problem presented in this study is a new formulation for which no benchmarks exist. Preliminary experimentation indicated that the following GA parameters gave satisfactory performance and were used for all optimization scenarios.

- Population size per generation = 40.
- Reproduction rate (fraction of total population mated) = 0.50.
- Mutation rate (bit level) = 0.04.
- Termination criterion: 40 generations without objective function improvement.

The largest computational expense occurs in a pre-processing stage that calculates a full symmetric matrix of nearest edge-to-edge patch distances (required for the MNN and MPI metrics), as well as vectors of patch perimeters and patch areas for the original landscape. With these patch attributes pre-calculated, the fitness function calculation is very efficient.<sup>1</sup>

Fig. 14 shows a comparison between the GA and a random search for the landscape configuration that minimizes the MNN. The unconstrained MNN minimizing landscape found using the GA (upper left panel in Fig. 9) contained 102

<sup>1</sup> Pre-processing of the original landscape (162 patches) required 72.7 s with Matlab 5.2 on a G3 233 MHz processor, a single fitness function evaluation required about 0.025 s. Typical optimization runs required about 2–3 min and were insensitive to the particular optimization metric.

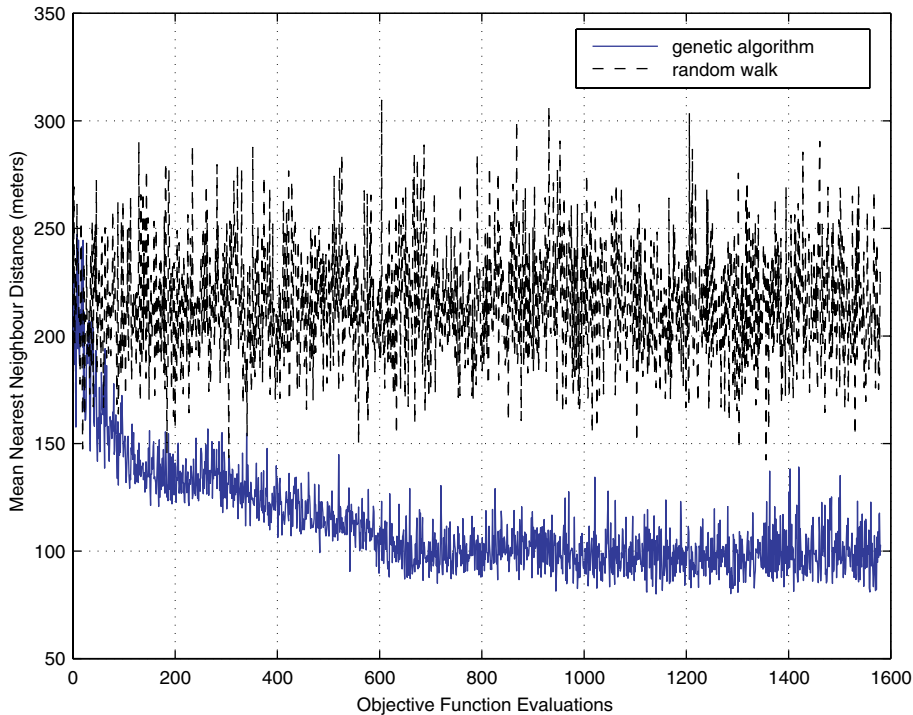


Fig. 14. GA heuristic comparison with a random search: MNN minimization.

of the original 162 patches. The random search shown here was therefore biased in that every landscape evaluated had 102 patches; a complete enumeration would still require  $\frac{162!}{(162-102)!}$  (approximately  $2.42e+205$ ) evaluations,<sup>2</sup> and illustrates the need for heuristic approaches even for such very modest problem sizes. The GA heuristic is proposed here as a flexible, and practical approach, however we strongly encourage comparisons with alternative heuristics.<sup>3</sup>

## 6. Conclusions

Many optimization algorithms have been proposed for forest management and conservation

biology. However as yet, landscape ecology concepts emphasizing the landscape fragmentation and connectivity issues that underlie modern notions of biodiversity conservation, have been weakly integrated. The research presented here introduces landscape ecology design metrics within a spatial optimization framework. The design framework is motivated by a simulation analysis of a deforesting landscape driven by incremental forest patch removal—the underlying dynamic attributed to tropical deforestation and subsequent biodiversity loss. A GA optimization heuristic provides good quality solutions for a wide range of constrained and unconstrained design objectives expressed using landscape ecology metrics. The reader is reminded that this analysis is illustrative and based on an informed though arbitrary selection of metrics. Similar exercises that use larger sets of landscape ecology metrics and evaluate alternative optimization heuristics are encouraged.

This research is not the first to use an explicit optimization approach to address spatial ecology

<sup>2</sup> Assuming (optimistically)  $1e9$  evaluations per second, full enumeration would take  $2.85e192$  years.

<sup>3</sup> The test problem and Matlab code are available from the corresponding author.

issues, however the high level of species-specific parameterization required by existing models is not practical in frontier forest regions of the developing world. Even if species parameters were known, conservation outcomes would not improve if the livelihood issues of the traditional rural communities that share the landscape were not integrated into any management plan. The landscape design paradigm presented in this paper was therefore designed to work when only surrogate remotely sensed data and traditional ecological knowledge are available, a typical situation where deforestation and biodiversity loss are most common, and hence a distinct advantage over more complex meta-population models that require extensive field calibration. In summary, we believe that the landscape design research introduced here is valuable both in its amenability to multi-criteria and participatory approaches that will characterize future forestry OR, and its consistency with the spatial ecology concepts and data products that will form the foundation of biodiversity conservation planning.

### Acknowledgements

Support for this research has been provided by the International Development Research Centre (Canada) through the Doctoral Research Award and through the Natural Science and Engineering Research Council of Canada Individual Discovery Grant 5671 and equipment grant EQPEQ 020, Case File No. EQP0173530.

### References

- Barrett, T.M., Gilles, J.K., Davis, L.S., 1998. Economic and fragmentation effects of clearcut restrictions. *Forest Science* 44 (4), 569–577.
- Baschak, L., Brown, R., 1994. River systems and landscape networks. In: Cook, E., Van Lier, H.N. (Eds.), *Landscape Planning and Ecological Networks*. Elsevier, Amsterdam, pp. 179–199.
- Christian, D.P., Hoffman, W., Hanowski, J.M., Niemi, G., Beyea, J., 1998. Bird and mammal diversity on woody biomass plantations in North America. *Biomass and Bioenergy* 14 (4), 395–402.
- Cocklin, C., 1989. Mathematical programming and resource planning I: The limitations of traditional optimization. *Journal of Environmental Management* 28, 127–141.
- Collinge, S.K., 2001. Spatial ecology and biological conservation. *Biological Conservation* 100, 1–2.
- Diamond, J., 1975. The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation* 7, 128–146.
- Fjelsda, J., 2000. The relevance of systematics in choosing priority areas for global conservation. *Environmental Conservation* 27 (1), 67–75.
- Forman, R.T.T., 1986. In: *Landscape Ecology*. Wiley, New York.
- Forman, R.T.T., 1995. Some general principles of landscape and regional ecology. *Landscape Ecology* 10 (3), 133–142.
- Forman, R.T.T., Collinge, S.K., 1995. The spatial solution to conserving biodiversity in landscapes and regions. In: DeGraaf, R.M., Miller, R.I. (Eds.), *Conservation of Faunal Diversity in Forested Landscapes*. Chapman and Hall, London, pp. 537–568.
- Gadgil, M., Hemam, N.S., Reddy, B.M., 1998. People, refugia and resilience. In: Berkes, F., Folke, C. (Eds.), *Linking Social and Ecological Systems: Management Practices and Social Mechanisms for Building Resilience*. Cambridge University Press, Cambridge, pp. 30–47.
- Goma, H.C., Rahim, K., Nangendo, G., Riley, J., Stein, A., 2001. Participatory studies for agro-ecosystem evaluation. *Agriculture, Ecosystems and Environment* 87, 179–190.
- Griffith, J.A., Martinko, E.A., Price, K.P., 2000. Landscape structure analysis of Kansas at three scales. *Landscape and Urban Planning* 52, 45–61.
- Gustafson, E.J., 1998. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems* 1, 143–156.
- Gustafson, E.J., Crow, T.R., 1996. Simulating the effects of alternative forest management strategies on landscape structure. *Journal of Environmental Management* 46, 77–94.
- Gustafson, E.J., Parker, G.R., 1992. Relationships between landcover proportion and indices of spatial pattern. *Landscape Ecology* 7, 101–110.
- Hanski, I., Poyry, J., Pakkala, T., Kuussaari, M., 1995. Multiple equilibria in metapopulation dynamics. *Nature* 377, 618–621.
- Hargis, C.D., Bissonette, J.A., David, J.L., 1998. The behaviour of landscape metrics commonly used in the study of habitat fragmentation. *Landscape Ecology* 13, 167–186.
- Hargis, C.D., Bissonette, J.A., Turner, D.L., 1999. The influence of forest fragmentation and landscape pattern on American martens. *Journal of Applied Ecology* 36 (1), 157–172.
- Heilman, G.E., Stritthold, J.R., Slosser, N.C., Dellasala, D.A., 2002. Forest fragmentation of the conterminous United States: Assessing forest intactness through road density and spatial characteristics. *BioScience* 52 (5), 411–422.
- Hof, J., Bevers, M., 1998. *Spatial Optimization of Managed Ecosystems*. Columbia University Press, New York.



- Hof, J.G., Joyce, L.A., 1992. Spatial optimization for wildlife and timber in managed forest ecosystems. *Forest Science* 38 (3), 489–508.
- Holland, J.H., 1992. *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA.
- IISD, 1999. *Our forests, our future: Summary Report of the World Commission on Forests and Sustainable Development*. International Institute for Sustainable Development, Winnipeg.
- Jackson, J.E., 1991. *A User's Guide to Principal Components*. Wiley, New York.
- Jankowski, P., Andrienko, N., Andrienko, G., 2001. Map-centred exploratory approach to multiple criteria spatial decision-making. *International Journal of Geographic Information Science* 15 (2), 101–127.
- Kurtilla, M., 2001. The spatial structure of forests in the optimization calculations of forest planning—a landscape ecological perspective. *Forest Ecology and Management* 142, 129–142.
- Li, H., Reynolds, J.F., 1993. A new contagion index to quantify spatial patterns of landscapes. *Landscape Ecology* 8, 155–162.
- Lindenmayer, D.B., Cunningham, R.B., Pope, M.L., 1999. A large scale “experiment” to examine the effects of landscape context and fragmentation on mammals. *Biological Conservation* 88, 387–403.
- Liu, D., Iverson, L.R., Brown, S., 1993. Rates and patterns of deforestation in the Philippines: Application of geographic information systems. *Forest Ecology and Management* 57, 1–16.
- Malczewski, J., 1999. Visualization in multicriteria spatial decision support systems. *Geomatica* 53, 139–147.
- Marks, B.J., McGarigal, K., 1994. *Fragstats: Spatial pattern analysis program for quantifying landscape structure*. Technical report, Forest Science Department, Oregon State University.
- Martell, D.L., Gunn, E.A., Weintraub, A., 1998. Forest management challenges for operational researchers. *European Journal of Operational Research* 104, 1–17.
- O'Neil, R.V., Hunsaker, C.T., Jones, K.B., Ritters, K.H., Wickham, J.D., Schwartz, P.M., Goodman, I.A., Jackson, B.L., Baillargeon, W.S., 1997. Monitoring environmental quality at the landscape scale. *BioScience* 47, 513–519.
- O'Neill, V.R., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., Angelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H., Graham, R.L., 1988. Indices of landscape pattern. *Landscape Ecology* 1 (3), 153–162.
- Pandey, D.N., 2002. Carbon sequestration in agroforestry systems. *Climate Policy* 2, 367–377.
- Possingham, H., Ball, L., Andelman, S., 2000. Mathematical methods for identifying representative reserve networks. In: Ferson, S., Burgman, M. (Eds.), *Quantitative Methods for Conservation Biology*. Springer-Verlag, New York, pp. 291–305.
- Pressey, R.L., Possingham, H.P., Margules, C.R., 1996. Optimality in reserve selection algorithms: When does it matter and how much? *Biological Conservation* 76, 259–267.
- Pressey, R.L., Possingham, H.P., Day, R.J., 1997. Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. *Biological Conservation* 80, 207–219.
- Quattrochi, D.A., Pelletier, R.E., 1991. Remote sensing for analysis of landscapes: An introduction. In: Turner, M., Gardner, R.H. (Eds.), *Quantitative Methods in Landscape Ecology*. Springer-Verlag, New York, pp. 17–76.
- Ramesh, B.R., Menon, S., Bawa, K.S., 1997. A vegetation based approach to biodiversity gap analysis in the Agasthyamalai region, Western Ghats, India. *Ambio* 26 (8), 529–536.
- Reggiani, A., Nijkamp, P., Sabella, E., 2001. New advances in spatial network modelling: Towards evolutionary algorithms. *European Journal of Operational Research* 128, 385–401.
- Reid, W.V., 1998. Biodiversity hotspots. *Trends in Ecology and Evolution* 13, 275–280.
- Richards, J.A., Jia, X., 1999. *Remote Sensing Digital Image Analysis, An Introduction*. Springer-Verlag, New York.
- Ritters, K.H., Wickham, J.D., O'Neil, R.V., Jones, K.B., Smith, E., 2000. Global-scale patterns of forest fragmentation. *Conservation Ecology* 4 (2), 3 [online]. URL: <http://www.consecol.org/vol4/iss2/art3>.
- Sisk, T.D., Haddad, N.M., 2002. Incorporating the effects of habitat edges into landscape models: Effective area models for cross-boundary management. In: Liu, J., Taylor, W.W. (Eds.), *Integrating Landscape Ecology into Natural Resource Management*. Cambridge University Press, Cambridge, pp. 208–240.
- Venema, H.D., Calamai, P.H., Ponnambalam, K., 2000. Multiobjective spatial design principles for rural biomass energy planning. *Journal of Environmental Studies and Policy* 3 (1), 1–19.
- Verboom, J., Foppen, R., Chardon, P., Opdam, P., Luttikhuisen, P., 2001. Introducing the key patch approach for habitat networks with persistent populations: An example for marshland birds. *Biological Conservation* 100, 89–101.
- WRI, 2000. *World Resources 2000–2001—People and ecosystems: The fraying web of life*. World Resources Institute, Washington.