European Journal of Operational Research 231 (2013) 514–534

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

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

Invited Review

# Mathematical optimization ideas for biodiversity conservation

Alain Billionnet

*Laboratoire CEDRIC, École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise, 1 Square de la Résistance, 91025 Evry Cedex, France*

ARTICLE INFO

*Article history:*
Received 5 January 2012
Accepted 19 March 2013
Available online 27 March 2013

*Keywords:*
Biodiversity protection
Mathematical optimization
Integer programming
Nature reserve
Landscape fragmentation
Genetic diversity

ABSTRACT

Several major environmental issues like biodiversity loss and climate change currently concern the international community. These topics that are related to the development of human societies have become increasingly important since the United Nations Conference on Environment and Development (UNCED) or Earth Summit in Rio de Janeiro in 1992. In this article, we are interested in the first issue. We present here many examples of the help that using mathematical programming can provide to decision-makers in the protection of biodiversity. The examples we have chosen concern the selection of nature reserves, the control of adverse effects caused by landscape fragmentation, including the creation or restoration of biological corridors, the ecological exploitation of forests, the control of invasive species, and the maintenance of genetic diversity. Most of the presented models are – or can be approximated with – linear-, quadratic- or fractional-integer formulations and emphasize spatial aspects of conservation planning. Many of them represent decisions taken in a static context but temporal dimension is also considered. The problems presented are generally difficult combinatorial optimization problems, some are well solved and others less well. Research is still needed to progress in solving them in order to deal with real instances satisfactorily. Moreover, relations between researchers and practitioners have to be strengthened. Furthermore, many recent achievements in the field of robust optimization could probably be successfully used for biodiversity protection, a domain in which many data are uncertain.



## 1. Introduction

Biodiversity, short for biological diversity, represents the diversity of living organisms and ecosystems. It also incorporates the interactions between living organisms and the interactions between living organisms and their environments. It is now accepted that biodiversity – species, genetic and ecosystem – renders important services to human societies and that its preservation is essential. Thus the United Nations General Assembly has approved in late 2010, the creation of the Intergovernmental science-policy Platform on Biodiversity and Ecosystem Services (IPBES). Biodiversity is undergoing significant erosion and that erosion has consequences for the planet as serious, although less known, than those related to climate change. This decline of biodiversity disturbs ecosystem functioning and thus affects the quality of services they provide to human populations concerned. These include, for example, agriculture, food, housing, health, tourism and economy. Biodiversity loss is a particularly serious problem because it is irreversible. According to the latest update of the Red List of threatened plant and animal species established by the International Union for Conservation of Nature (IUCN), about 17,000 species on the 48,000 listed are threatened with extinction (http://www.iucn.org/). The Convention on Biological Diversity (CBD) adopted at the Earth Summit in Rio de Janeiro in 1992, and ratified by about 190 countries had identified five main factors causing biodiversity loss: fragmentation of spaces, overexploitation of species, pollution, invasive species and climate change. In 2002, the signatory countries of the CBD had adopted at the World Summit on Sustainable Development in Johannesburg, a strategic plan to achieve by 2010 a significant reduction in the rate of biodiversity loss. In 2010 the commission found that no country had managed to achieve this goal. At the 10th meeting of the CBD, held in Nagoya in October 2010, a new plan for biodiversity conservation for 2020 was adopted. Available resources to protect biodiversity are obviously limited and it is important to use them effectively. For this, two types of approaches are possible: a direct approach based on the properties and algorithms for mathematical optimization, and simulation approach. They each have their advantages and disadvantages. The approach by simulation, generally simpler to implement, has been widely used to address complex problems in ecology and sustainable development. The approach by mathematical optimization, more difficult to implement, has been less used, but unlike the simulation approach, it allows us to evaluate a large number of options. We illustrate in this article the help that using operational research – mainly mathematical programming – can bring to decision-makers in the implementation of key strategies to protect biodiversity. Among the many optimization problems involved in this area, we chose a few representative

*E-mail address:* Alain.Billionnet@ensiie.fr

0377-2217/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2013.03.025

issues: the selection of nature reserves, the control of adverse effects caused by landscape fragmentation, the rational exploitation of forests, the fight against invasive species, and the maintenance of genetic diversity. The literature is very abundant on these subjects and, in this paper, we mainly limited ourselves to problems that can be formulated, more or less directly, by linear or nonlinear programming with integer or mixed-integer variables. The integer programming approach presents many advantages compared to specific algorithms: simplicity of implementation if an integer programming software is available, reliability of the method, exact or guaranteed approximate solution of the problem, and finally, possibility of easily modifying the model. Many of the considered problems are derived directly or inspired from the literature. Others are generalizations and finally, some of them are new. For problems of the literature, we present the formulations published and sometimes more interesting formulations. Carefully implemented, mathematical programming – also known as mathematical optimization – is a powerful tool that can be used to solve many problems of operational research. There are indeed very efficient algorithms for these problems and many software based on these algorithms are available. This should enable operational research to play an important role in the field of biodiversity conservation, as important as it plays in the fields of transport, energy, telecommunications and manufacturing. Finally, note that mathematical models are one small part of conservation planning. The management context, although important, is little addressed in this review the aim of which is to present mathematical optimization ideas widely applicable to biodiversity protection. Describing the management context and also assessing the utility and applicability of the models would be another work in itself. Regarding these issues, the reader may consult, for example, the following references: (Margules and Pressey, 2000) where six different stages are identified in systematic conservation planning, (Prendergast et al., 1999) where the utility of reserve selection algorithms is examined and the comments on this article by Pressey and Cowling (2001), (Lindenmayer et al., 2006) for a checklist of measures that reflects the multi-scaled nature of conservation approaches on forested lands, (Halkos and Jones, 2012) for an investigation of the influence of social factors on the decision of individuals to contribute an amount for improving environmental protection of biodiversity, (Bergseng and Vatn, 2009) for a discussion about the reasons for conflict in protection of biodiversity in forests, and (Wallace, 2012) for a planning framework in which the planning components are linked through cause-effect relationships and driven by human values.

## 2. Selection of nature reserves

### 2.1. Interest of nature reserves

Many countries have pledged to halt biodiversity loss in the near future and have adopted different strategies for this including the protection of land and sea areas. These protected areas – or reserves – play a decisive role in maintaining biodiversity because they aim directly at the protection of elements which have the strongest risk of extinction. These elements relate to flora, fauna, rocks, minerals and fossils, or major geomorphological sites. The objective is to ensure each threatened species or site has a place where its future is guaranteed. Thus, many governmental and nongovernmental programs seek to restore and protect habitat in order to preserve the species. At the 10th meeting of the CBD, a plan for biodiversity conservation for 2020 was adopted. It contains 20 goals including the restoration of degraded habitats and the establishment of protected areas (terrestrial, marine and coastal). Commenting on the plan, the President of the environmental organization *Conservation International*, said that the problem is not only quantitative but also qualitative and that the most important areas in terms of biodiversity must be protected. The resources available for this protection being obviously limited, it is important to use them efficiently. Until the 1980s, the proposed methods mainly consisted to rank the potential sites in order of interest by using scoring methods. Smith and Theberge (1986) and Cocks and Baird (1989) were some of the first authors to propose the use of mathematical optimization techniques for solving the problem of selecting which sites should ideally be included in a reserve network. Subsequently, many optimization models have been proposed in the literature of operational research and conservation biology to help select sites for designing reserves. These publications are usually theoretical, they are modeling realistic problems and propose algorithms – often heuristics – to solve them. Some authors discuss the applicability of these models (see e.g., Cabeza and Moilanen, 2001). Many objectives can be considered in selecting nature reserves. For example, Juutinen and Mönkkönen (2007) compare the obtained results with two different objectives, the presence of species and species abundance, while varying the relative weights of different species. They carry out their study using actual data for the boreal forest in Finland. Some articles are based on multi-objective mathematical programming (see e.g., Memtsas, 2003). Although many publications present applications of their models to real data (see e.g., Poulin et al., 2006; Toth et al., 2009; Fiorella et al., 2010; Groeneveld, 2010), few of them concern the actual use of these models by an organization to make decisions. Some articles discuss the gap in this field between theory and practice (see e.g., Prendergast et al., 1999; Pressey and Cowling, 2001; Knight et al., 2008; Schindler et al., 2011; Braunisch et al., 2012; Jolibert and Wesselink, 2012). Several software for selecting nature reserves are currently available. *Marxan* (http://www.uq.edu.au/marxan/) finds good solutions to a mathematically well-specified problem. Different optimization techniques are used to drive the optimization phase of this software: integer linear programming to obtain exact optimal solutions and metaheuristics such as genetic algorithms and simulated annealing to obtain approximate solutions. The reader can refer to (Ball et al., 2009) for a comprehensive description of this software including an example of its application to a conservation prioritization for the entire Australian continent. Other valuable software are also available: *Zonation* (http://www.helsinki.fi/bioscience/consplan/software/Zonation/index.html) and *C-Plan*, (http://www.edg.org.au/free-tools/cplan.html).

With regard to *Zonation* and *C-Plan*, the reader may refer to Moilanen et al. (2009) and Pressey et al. (2009), respectively. The reader may also refer to Sarkar et al. (2006), a comprehensive survey on the biodiversity conservation planning tools presented in the conservation biology literature. Among other things this survey reviews the various software tools for conservation planning that have been developed over the past 20 years. Four other interesting references are (Pressey et al., 1996), (Rodrigues and Gaston, 2002b), (Fischer and Church, 2005) and (Vanderkam et al., 2007). In these articles the authors compare exact and heuristics approaches to solve some reserve selection problems.

We present below some selection reserve problems and their formulation by mathematical programming. We first consider the basic problem and some variants. This problem is to select a set of areas, of minimum cost, to protect a set of predefined species. A variant consists in determining, under a budget constraint, a set of areas to protect a maximum number of species. These problems can be modeled easily by 0–1 linear programs and can be solved efficiently by commercial solvers. We then illustrate consideration of spatial constraints. These constraints may affect the compactness of the reserve, its connectivity or its shape. They can also impose a specific role to different areas of the reserve (central zone and buffer zone, for example). Taking into account these spatial constraints often complicates the models because they generate nonlinearities in the associated mathematical programs. The connectivity constraint, common to many operational research problems, is particularly difficult to take into account. We then focus on the definition of reserves taking into account the population of each species that needs protection. In this case, we consider that a species can survive only if the size of its population exceeds a certain threshold. Another objective, also taking into account the population size of each species, is to define a reserve which maximizes the species diversity. Again, the optimal solutions can be difficult to obtain. We also consider the realistic case, where for each species we only know its survival probability in a protected area. The associated problems are often difficult to solve because the expression of the survival probability of a species in a given set of protected areas is generally complicated. In some cases, one should be satisfied with sub-optimal solutions. Finally, we illustrate the temporal dimension which may occur in the design of a reserve. Consideration of time does not necessarily introduces great difficulties in modeling. However it increases the size of the associated programs.

### 2.2. Basic problem and variants

We are interested in a set of species to be protected, $E = \{e_1, e_2, \ldots, e_p\}$, living on a set of parcels – or sites –, $S = \{s_1, s_2, \ldots, s_n\}$, spread over a territory. For each parcel, we know all the species that live there and it is assumed that these species will survive in this parcel if it is protected. The basic problem is to determine the smallest subset of parcels in order to protect all species considered (see e.g., Margules et al., 1988; Possingham et al., 1993; Underhill, 1994; ReVelle et al., 2002). Associate with each parcel $s_i$ a Boolean variable $x_i$ that equals 1 iff this parcel is selected to be protected. Let $P = \{1, \ldots, p\}$ and $N = \{1, \ldots, n\}$. The determination of an optimal reserve, i.e. of an optimal subset of parcels can be formulated by the mathematical program

$$(\mathrm{P1}): \quad \min_{x\in\{0,1\}^n}\left\{\sum_{i\in N}x_i : \sum_{i\in S_k}x_i \geqslant 1 \ (k\in P)\right\}$$

where $S_k$ is the set of indices of parcels that protect the species $e_k$. The objective function corresponds to minimizing the number of parcels used and the $p$ inequality constraints express that for each species considered, at least one parcel containing it should be retained. The 0–1 linear mathematical program (P1) is known in operational research as the *set covering problem* (see e.g., Caprara et al., 2000). There are many variants of (P1). We can try to minimize the total area of the reserve or the total cost when costs are allocated to each parcel, rather than the number of parcels. The objective function to minimize becomes $\sum_{i\in N}a_ix_i$ where $a_i$ is the area or the cost of the parcel $s_i$. The definition of an optimal reserve may also be addressed by seeking to identify a subset of parcels that maximizes the number of protected species (see e.g., Camm et al., 1996; Church et al., 1996; Arthur et al., 1997; Csuti et al., 1997; Rosing et al., 2002). In this case, the number of parcels of the reserve or its area or its cost is limited. This gives the mathematical program

$$(\mathrm{P2}): \quad \max_{x\in\{0,1\}^n, y\in\{0,1\}^p}\left\{\sum_{k\in P}y_k : y_k \leqslant \sum_{i\in S_k}x_i \ (k\in P), \ \sum_{i\in N}a_ix_i \leqslant B\right\}.$$

In an optimal solution of (P2), the Boolean variable $y_k$ takes the value 1 iff the species $e_k$ occurs in at least one of the parcels of the reserve. The constraints of (P1) and (P2) can be adjusted to require that each species is represented in at least $r$ parcels rather than in one. In (P1), the right hand side of the constraints becomes $r$ and in (P2) the constraints $y_k \leqslant \sum_{i\in S_k}x_i$ become $ry_k \leqslant \sum_{i\in S_k}x_i$. Polasky et al. (2001) used this model to study the protection of 415 species of terrestrial vertebrates spread over 289 parcels of the state of Oregon in the United States.

A slightly different problem is to consider, as McBride et al. (2010) do, that the parcels are divided into $q$ disjoint clusters $C_1, C_2, \ldots, C_q$. A cost $c_i$ is associated with the parcel $s_i$ and a cost $d_l$ is associated with the cluster $C_l$. Thus, the protection of a subset, $T$, of parcels of $C_l$ costs $d_l + \sum_{i:s_i\in T}c_i$. The cost $d_l$ is a fixed cost associated with the cluster $C_l$ that corresponds, for example, to the moving of equipment, personnel, to the acquisition of the cluster, etc. As in (P1), we try to select a subset of parcels that protects all species considered, at least cost. Using the Boolean variable $z_l$ that is equal to 1 iff at least one of the parcels of $C_l$ is retained and setting $Q = \{1, \ldots, q\}$, the problem can be formulated by the linear program in 0–1 variables

$$(\mathrm{P3}): \quad \min_{x\in\{0,1\}^n,\ z\in\{0,1\}^q}\left\{\begin{array}{l}\sum_{i\in N}c_ix_i + \sum_{l\in Q}d_lz_l : \sum_{i\in S_k}x_i \geqslant 1 \ (k\in P),\\ z_l \geqslant x_i \ (l\in Q, s_i\in C_l)\end{array}\right\}.$$

According to the last constraint of (P3) and since the objective function is to be minimized, the variable $z_l$ will take the value 1 if at least one of the parcels of $C_l$ is retained and 0 otherwise. Previous models take into account neither the spatial aspect of the reserve nor the population sizes of the species present in the reserve. The solution depends only on the presence or absence of species in the parcels. The spatial aspect of the reserves and the populations sizes are discussed in the sections below.

### 2.3. Taking into account spatial constraints

There are many ways to take into account the spatial aspect of reserves (see e.g., Williams et al., 2005). One might be interested, for example, in the maximum or minimum distance between two sites of the reserve (see e.g., Williams, 2008) or in its total perimeter (see e.g., Önal and Briers, 2002; McDonnell et al., 2002; Fischer and Church, 2003 and Section 4 of this article). The latter criterion allows for groups of sites relatively close, the shape of which approaches a square or circle, but does not control the distance between the groups. One can also look for a reserve forming a convex space, but the obtained reserve can be relatively large to ensure the protection of all species considered (see e.g., Williams, 2003). The proximity between the sites of the reserve can also be measured by the sum of the distances between all pairs of selected sites (see e.g., Nalle et al., 2002; Önal and Briers, 2002). Note that two sites can be considered as close for certain species and distant for others (Cerdeira et al., 2010). The area of the selected parcels can also be an important criterion in the definition of a reserve (Marianov et al., 2008). The reader may consult (Moilanen et al., 2009b), a comprehensive book on the subject, and also (Baskent and Keles, 2005) for a discussion of the inclusion of spatial constraints in forest management. We present below some reserve selection problems with spatial constraints.

#### 2.3.1. Compact reserves

The problem is to select a subset of parcels to protect as many species as possible while respecting a compactness constraint and a budget constraint. As noted above, the compactness constraint can be expressed in many ways. One way consists in imposing to the ratio *total perimeter of the reserve* to *total area of the reserve* to be less than or equal to $\rho$ (see e.g., Öhman and Lämas, 2005). Denote respectively by $l_i$ and $a_i$ the perimeter and the area of the parcel $s_i$ and by $l_{ij}$ the length of the common border between $s_i$ and $s_j$. The total perimeter of the reserve is $L(x) = \sum_{i=1}^{n}l_ix_i - 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}l_{ij}x_ix_j$ and the total area of the reserve is $A(x) = \sum_{i=1}^{n}a_ix_i$. The problem can then be formulated by the nonlinear 0–1 program

$$(\text{P4}): \max_{x\in\{0,1\}^n, y\in\{0,1\}^p} \left\{ \sum_{k\in P} y_k : y_k \leqslant \sum_{i\in S_k} x_i \ (k\in P), \right. \left. L(x) \leqslant \rho A(x),\ \sum_{i\in N} c_i x_i \leqslant B \right\}.$$

(P4) can be easily linearized by replacing the products $x_i x_j$ in $L(x)$ by the variables $z_{ij}$ and by adding the linearization constraints $z_{ij} \leqslant x_i$ and $z_{ij} \leqslant x_j$.

Another way to express the constraint of compactness is to minimize the diameter of the reserve, i.e. the maximum distance between two parcels of the reserve (see e.g., Önal and Briers, 2002). The result is a subset of parcels which the external contour approaches a circle. The problem of defining a compact reserve to protect all species can then be formulated by the mathematical program

$$(\text{P5}): \min_{x\in\{0,1\}^n,\ \alpha\geqslant 0} \left\{ \alpha : \alpha \geqslant d_{ij} x_i x_j \ ((i,j)\in N^2,\ i<j),\ \sum_{i\in S_k} x_i \geqslant 1 \ (k\in P) \right\}$$

where $d_{ij}$ denotes the distance between the parcels $s_i$ and $s_j$. The first constraint expresses that if the parcels $s_i$ and $s_j$ are retained, then the diameter $\alpha$ is greater than or equal to the distance between these two parcels. On the contrary, if at least one of these two parcels is not retained, then the constraint is inactive. The second constraint expresses that for each species, a parcel containing it must be selected. In (P5), the objective function is linear but the constraints are quadratic. A classical linearization of (P5) is given by

$$(\text{P6}): \min_{x\in\{0,1\}^n,\ \alpha\geqslant 0} \left\{ \alpha : \alpha \geqslant d_{ij}(-1 + x_i + x_j) \ ((i,j)\in N^2,\ i<j), \right. \left. \sum_{i\in S_k} x_i \geqslant 1 \ (k\in P) \right\}.$$

The compactness of a reserve can also be measured by the sum of the distances separating each pair of parcels. The problem consists in minimizing a quadratic function of 0–1 variables subject to linear constraints and can then be stated as

$$(\text{P7}): \min_{x\in\{0,1\}^n} \left\{ \sum_{(i,j)\in N^2,\ i<j} d_{ij} x_i x_j : \sum_{i\in S_k} x_i \geqslant 1 \ (k\in P) \right\}.$$

There are several ways to solve (P7). It can be submitted directly to a solver such as CPLEX (2007). One can also use one of the several linearization methods designed for quadratic 0–1 programs (see e.g., Boros and Hammer, 2002; Billionnet, 2010c). One can also reformulate (P7) as the minimization of a convex quadratic function of 0–1 variables subject to linear constraints (Billionnet et al., 2008; Billionnet et al., 2009; Billionnet et al., 2012). An alternative approach would be to reformulate (P7) using the technique proposed by Burer (2009) to model any nonconvex quadratic program having a mix of binary and continuous variables as a linear program over the dual of the cone of copositive matrices. Regarding the recent developments and applications of copositive optimization the reader may consult the survey of Bomze (2012). Note that we can also apply to the resolution of (P7) semidefinite relaxation techniques that are often effective for 0–1 quadratic programs (see e.g., Poljak et al., 1995; Roupin, 2004). Another way to impose a reserve to be compact is to limit, by appropriate constraints, the diameter of the reserve. One can use, for example, the constraints $x_i + x_j \leqslant 1 \quad ((i,j)\in N^2,\ i<j,\ d_{ij} > d_{\max})$ or $\sum_{i:d_{ij}>d_{\max}} x_i \leqslant M(1-x_j) \quad (j\in N)$ where $M$ denotes a constant large enough.

*2.3.2. Connected reserves*

Functional connectivity of a reserve involves the notion of proximity between two parcels and a connected reserve is a set of parcels such that for any couple $(s_i, s_j)$ of this set, it is possible to go from $s_i$ to $s_j$ by progressively moving from a parcel to another one that is close. Note that this notion of proximity between two parcels depends on the species. In other words, in a connected reserve an animal must be able to move throughout the reserve without too much deviation from this reserve (see e.g., Williams, 2002; Cerdeira et al., 2005; Önal and Briers, 2005; Önal and Briers, 2006; Groeneveld, 2010). In the example of Fig. 1, two parcels are considered as close if they are adjacent.

Note that the only connectivity criterion does not control the contour of the reserve obtained. Suppose the area under consideration is represented by a matrix of $m \times n$ identical square parcels $s_{ij}$. Here, we consider that two parcels are functionally close from each other if they are adjacent (share a common side) and we know the species present in each parcel. Specifically, the Boolean coefficient $a_{ijk}$ equals 1 iff the species $e_k$ is present in the parcel $s_{ij}$. We consider here that the species $e_k$ is protected if it is present in at least $b_k$ parcels of the reserve. The problem is to determine an optimal connected reserve to protect all species. As in the other problems, the optimality criterion can be the number of sites making up the reserve, the cost of the reserve or its area. Associate the symmetric directed graph $G = (X, U)$ to the matrix of the $m \times n$ parcels. The vertices of the graph are the couples of indices $(i,j)$ associated with the parcels $s_{ij}$ and there is an arc between two parcels iff they are adjacent (share a common side). The problem may then be stated as follows: find a subset of vertices $S \subset X$ such that each species is present in at least $b_k$ vertices of $S$ and such that the sub-graph induced by $S$ is connected or, equivalently, such that every couple of vertices of $S$ can be linked by a path of $G$ going only through vertices in $S$. This last point can also be formulated by imposing to the sub-graph induced by $S$ to contain a vertex-spanning rooted tree. Let $x_{ij}$ be the Boolean variable that is equal to 1 iff the parcel $s_{ij}$ is retained in the reserve and $y_{ijrs}$ be the Boolean variable that is equal to 1 iff the arc $((i,j),(r,s))$ is chosen to form the rooted tree on the vertices of $S$. Let $M = \{1, \ldots, m\}$ and $N = \{1, \ldots, n\}$. We propose in (Billionnet, 2012) a slight improvement of the model proposed by Önal and Briers (2006). This improvement consists in formulating the problem by the mathematical program

$$(\text{P8}): \min_{\substack{x\in\{0,1\}^{mn} \\ y\in\{0,1\}^{|U|} \\ t\in R_+^{mn}}} \left\{ \begin{array}{l} \sum_{(i,j)\in M\times N} x_{ij} : \sum_{(i,j)\in M\times N} a_{ijk} x_{ij} \geqslant b_k \ (k\in P), \\ y_{ijrs} \leqslant x_{ij}((i,j),(r,s)) \in U, \\ \sum_{(i,j)\in A_{rs}^-} y_{ijrs} \leqslant x_{rs}((r,s)\in M\times N), \\ \sum_{((i,j),(r,s))\in U} y_{ijrs} = \sum_{(i,j)\in M\times N} x_{ij} - 1, \\ t_{rs} \geqslant t_{ij} + 1 - C(1 - y_{ijrs}) \ ((i,j),(r,s)) \in U \end{array} \right\}$$

where $A_{rs}^- = \{(i,j)\in M\times N : ((i,j),(r,s))\in U\}$. The real nonnegative variables $t_{ij}$ $((i,j)\in X)$ are used in the constraints $t_{rs} \geqslant t_{ij} + 1 - C(1 - y_{ijrs})$, which prohibit the formation of a circuit. Indeed, it

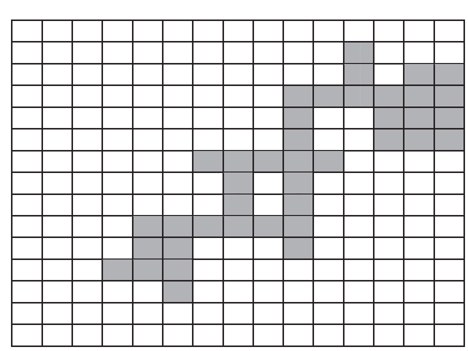

**Fig. 1.** A connected reserve composed of 40 parcels.

is impossible to satisfy such inequalities by assigning $t_{ij}$ values to the vertices of a circuit. Given a subset $S$ of vertices and $x$ its characteristic vector, a vector $y$ of $\{0,1\}^{|U|}$ defines a rooted tree on the subgraph induced by $S$ iff the constraints of (P8) are satisfied. The last set of constraints, where $C$ is a large enough constant, is similar to that used in some formulations of the traveling salesman problem to eliminate sub-tours using a polynomial number of constraints (see e.g., Garfinkel, 1985). To return to the model proposed by Önal and Briers (2006) we should replace in (P8), the second set of constraints by $\sum_{(r,s)\in A^+_{ij}} y_{ijrs} \leqslant 4x_{ij} ((i,j) \in M \times N)$ where $A^+_{ij} = \{(r,s) \in M \times N : ((i,j),(r,s)) \in U\}$ and the last set of constraints, by $z_{ijrs} \geqslant t_{ij} + 1 - C(1 - y_{ijrs})$ $((i,j),(r,s)) \in U)$ and $t_{ij} = \sum_{(r,s)\in A^-_{ij}} z_{rsij} ((i,j) \in M \times N)$ where $z_{ijrs}$ is a nonnegative real variable. The resulting formulation is slightly less effective than (P8) in terms of computation time. Experiments have shown that the size of instances that can be solved by (P8) is limited to a hundred species spread over a hundred of parcels.

A large class of decision and optimization problems consists of finding a connected subgraph of a larger graph satisfying certain cost and revenue requirements. For example, Conrad et al. (2012) present a mixed-integer linear programming model for the connected subgraph problem. This model that they use for optimally design corridors for grizzly bears in the U.S. Northern Rockies is based on a network flow approach. It would be interesting to compare this approach with (P8). Dilkina and Gomes (2010) investigate mathematical formulations and solution techniques to find a connected subgraph that contains a subset of distinguished vertices. They propose several mixed-integer formulations for enforcing the subgraph connectivity requirement. Note that there is a close relationship between the connected subgraph problem and the optimal design of corridors (Section 3.2).

Now we present the heuristic method for this problem proposed in (Billionnet, 2012) that is also based on integer linear programming. The method consists in solving the program (P9) obtained by replacing, in (P8), the constraints prohibiting the circuits by the constraints (C):

$$(\mathrm{P9}): \begin{cases} (\mathrm{P8}) \text{ in which } t_{rs} \geqslant t_{ij} + 1 - C(1 - y_{ijrs}) \\ \quad ((i,j),(r,s)) \in U \text{ is replaced by} \\ (\mathrm{C}): \begin{cases} y_{ijrs} + y_{rsij} \leqslant 1 & ((i,j),(r,s)) \in U, r > i \text{ or } s > j \quad (\mathrm{C}.1) \\ y_{i-1,j,i,j} + y_{i,j,i,j-1} \leqslant 1 & (i,j) \in M \times N,\ i \geqslant 2, j \geqslant 2 \quad (\mathrm{C}.2) \\ y_{i-1,j,i,j} + y_{i,j,i,j+1} \leqslant 1 & (i,j) \in M \times N,\ i \geqslant 2, j \leqslant n-1 \quad (\mathrm{C}.3) \end{cases} \end{cases}$$

Constraints (C.1) prohibit the circuits of length 2 and constraints (C.2)–(C.3), specific to a grid, forbid the circuits of length greater than or equal to 2. Indeed, the selected graph does not contain circuits if we forbid, in addition to circuits of length 2, the paths of length 2 of type $\{(i-1,j),(i,j),(i,j-1)\}$ or $\{(i-1,j),(i,j),(i,j+1)\}$. The first type corresponds to the clockwise circuits and the second type, to the anti-clockwise circuits. Experiments have shown that the solutions of (P9) were obtained much faster than the solutions of (P8) and also that the solutions of (P9) were often very close to the optimal solutions. However, the size of instances that can be solved by (P9) in a reasonable time is limited to about a hundred species over two hundred parcels.

### 2.3.3. Shapes of reserves

The shape of a parcel strongly influences the proportion of interior and edge habitat and it is not the same species that live in these two habitat types. For example, the edge of a wooded area is influenced by environmental conditions of this site that are different from those of the interior (solar radiation, temperature, wind speed, human disturbance, etc.). Thus, for some species, it may be desirable to define reserves with a significant edge habitat (Sisk and Haddad, 2002). The mean shape index (MSI) is an indicator widely used in landscape ecology to measure the complexity of the mean contour of a parcel in comparison with a standard contour (McGarigal and Marks (1994)). It is equal to the perimeter of the parcel divided by the square root of its area, multiplied by 0.25. Thus, the index of a circular parcel is equal to about 0.89 and that of a square parcel is equal to 1. MSI is the average of these indices over all the parcels considered: $\mathrm{MSI} = (1/n)\sum_{i=1}^{n} (0.25p_i/\sqrt{a_i})$ where $p_i$ and $a_i$ are respectively the perimeter and the area of the parcel $s_i$. This ratio increases when the contours of the parcels become more irregular and is minimal when all the parcels are circular. An optimization problem, suggested by Vemema et al. (2005) and associated with this concept of contour, is to select, taking into account an area constraint, a subset of parcels optimal with regard to the MSI value. If $S = \{s_1, s_2, \ldots, s_n\}$ is a set of potential parcels of total area $A$, the problem can be formulated by the linear fractional program in 0–1 variables

$$(\mathrm{P10}): \quad \max_{x\in\{0,1\}^n} \left\{ \frac{\sum_{i\in N}(0.25p_ix_i/\sqrt{a_i})}{\sum_{i\in N} x_i} : A_{\min} \leqslant \sum_{i\in N} a_ix_i \leqslant A_{\max} \right\}$$

where $x_i$ is a Boolean variable that is equal to 1 iff the parcel $s_i$ is selected, and $A_{\min}$ (resp. $A_{\max}$) is the total minimum (resp. maximum) area of the reserve. Note that the objective function of (P10) corresponds to MSI since the quantity $\sum_{i\in N} x_i$ is equal to the number of selected parcels. One way to solve (P10) is to apply Dinkelbach's algorithm (Dinkelbach, 1967) that consists of solving a series of auxiliary problems associated with the fractional problem. In the case of (P10), the auxiliary problem is a linear program in 0–1 variables that consists in maximizing the linear expression $\sum_{i\in N}(0.25p_ix_i/\sqrt{a_i}) - \lambda\sum_{i\in N} x_i$, where $\lambda$ is a real parameter, under the same constraints as those of (P10). The reader can consult (Schaible, 1995; Radzik, 1998) for definitions, properties, and algorithms of fractional programming.

### 2.3.4. Reserves with central area and buffer zone

We end this section on spatial considerations by discussing the selection of nature reserves with central area and buffer zone (see e.g., Williams and ReVelle, 1998; Clemens et al., 1999; Hamaide et al., 2009). In this case, there are two types of species to be protected, endangered and common species, and two types of protected areas, central areas and buffer zones that protect the central area of nuisances caused by human activities outside the reserve. To be protected, an endangered species must be present in a central area and a common species must be present either in a central area or in a buffer zone. It is considered that a selected parcel is part of a central area if all the parcels that surround it are also selected to be either in the central area or in the buffer zone (Fig. 2). The decision maker can define a buffer zone more or less important. Consider the Boolean variable $z_i$ that is equal to 1 iff the parcel $s_i$ is selected to be part of the central area and the Boolean variable $x_i$ that is equal to 1 iff the parcel $s_i$ is selected to be part of the buffer zone or of the central area. Denote by $NC \subseteq N$ the set of indices of the parcels candidates for the central area and by $N_i \subseteq N$ the set formed by the index $i$ and the set of indices of the parcels surrounding the parcel $s_i$ and that must be selected (in the central area or in the buffer zone) to give to parcel $s_i$ the status of central area.

Decompose $P$ into $PR$, the set of indices of rare species, and $PC$ the set of indices of common species. The problem can be formulated by the linear program in 0–1 variables

$$(\mathrm{P11}): \min_{\substack{z\in\{0,1\}^{|NC|} \\ x\in\{0,1\}^{|N|}}} \left\{ \sum_{i\in N} c_ix_i : \sum_{i\in S_k} z_i \geqslant 1 \ (k \in PR), \right.$$
$$\left. \sum_{i\in S_k} x_i \geqslant 1 \ (k \in PC),\ z_i \leqslant x_j \ (i \in NC, j \in N_i) \right\}.$$

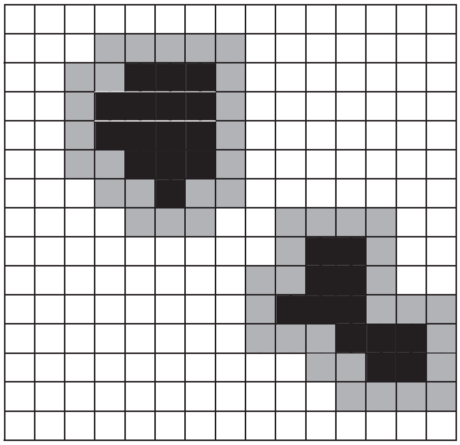

**Fig. 2.** Two disjoint reserves with central area (black) and buffer zone (gray).

The last constraint of (P11) forces the variable $z_i$ to take the value 0 if at least one of the parcels adjacent to $s_i$ if not selected (to be part of the buffer zone or of the central area).

The above models are well suited for species conservation in the short term but less well suited to long-term conservation. Indeed, the reserves determined by these models can be composed of many small parcels widely dispersed and the survival of a species in the long term, in this type of reserve, is not guaranteed, especially if the parcels are located in an inhospitable matrix as a highly urbanized area (see e.g., Crist, 2004).

### 2.4. Proximity of parcels and species populations

Some authors incorporate clusters of close parcels into the concept of nature reserve. These clusters must have a total area and populations of different species sufficiently large. This is the case of Toth et al. (2009) who propose a model that requires the listing of possible clusters. We present here a model that incorporates the concept of clusters in the mathematical program itself. Specifically, a cluster of parcels is a set of parcels such that the maximum distance between two parcels of the set is less than or equal to $d_{\max}$. We assume that a species will persist in the reserve if the reserve involves a cluster of parcels with a total population of the considered species greater than a certain threshold. The objective is to select an optimal subset of parcels, for example of minimum area, to protect all species considered. Let $p_{ik}$ be the population of species $e_k$ in the parcel $s_i$, $b_k$ be the threshold for survival of $e_k$, and $a_i$ be the area of the parcel $s_i$. Let $x_{ik}$ be a Boolean variable that is equal to 1 iff the parcel $s_i$ is selected to ensure the survival of the species $e_k$. Note that a parcel is included in the reserve iff it is used to protect at least one of the species. The problem can be formulated by the program

$$(\text{P12}):\quad \min_{\substack{x\in\{0,1\}^{np}\\ r\in\{0,1\}^n}} \left\{ \begin{array}{l} \sum_{i\in N} a_i r_i : x_{ik} \leqslant r_i \ (i\in N, k\in P), \\ \sum_{i\in N} p_{ik} x_{ik} \geqslant b_k \ (k\in P), \\ \sum_{j\in N, j>i,\ d_{ij}>d_{\max}} x_{jk} \leqslant B_{ik}(1-x_{ik}) \ (i\in N, k\in P) \end{array} \right\}.$$

The objective function sums the areas of the selected parcels. The first constraint requires the Boolean variable $r_i$ to take the value 1 iff the parcel $s_i$ is used to protect at least one of the species. The second constraint expresses that the global population of $e_k$ in the parcels selected to protect $e_k$ – a cluster – must be greater than or equal to $b_k$. The last constraint expresses that the maximum distance between two parcels in a cluster must be less than or equal to $d_{max}$. In this constraint, $d_{ij}$ is the distance between parcels $s_i$ and $s_j$, and $B_{ik}$ is an upper bound on the number of sites selected to protect $e_k$ in an optimal solution, whose subscript is greater than $i$, and located at more than $d_{\max}$ from the site $s_i$. For example, we can fix $B_{ik}$ to the maximum number of sites that a cluster may contain given the value of $d_{\max}$. The experiments that we performed on (P12) showed that hypothetical instances with a hundred of sites and about fifty species could be handled quickly.

### 2.5. Uncertainties in species survival

As we saw in the previous sections a lot of reserve selection problems are based on the knowledge of all the species that survive in a parcel if it is protected. However, in reality, the information is not always so accurate and several publications take into account some uncertainty about the data (see e.g., Arthur et al., 2002; Haight et al., 2002; Drechsler, 2005; Strange et al., 2006; Haight and Travis, 2008). In this section, we consider that the survival of a species in a parcel is defined by a probability (rather than by presence/absence). One way to address the problem of reserve selection in this case is to transform probabilistic information into deterministic information (see e.g., Margules and Stein, 1989; Polasky et al., 2000) but this obviously has many drawbacks.

Denote by $p_{ki}$ the probability that the species $e_k$ survives in the parcel $s_i$ if the parcel is protected. We assume that these events occur independently across species and that the probabilities satisfy $0 \leqslant p_{ki} < 1$. Let $q_{ki} = 1 - p_{ki}$. We can consider the following two objectives (see e.g., ReVelle et al., 2002): (i) identify a subset of parcels of minimum cost and such that, for each species $e_k$, the probability that this species is present in the reserve, i.e. in at least one selected parcels, is greater than or equal to a given value $\alpha_k$ and (ii) determine, under a budget constraint, a subset of parcels such that the total expected number of species present in these parcels is maximal. Let $x_i$ be a Boolean variable that is equal to 1 iff the parcel $s_i$ is selected. The probability that the parcel $s_i$ does not contribute to the presence (and thus to the survival) of the species $e_k$ is equal to $1 - p_{ki}x_i$. It is equal to 1 if the parcel is not selected and to $1 - p_{ki}$ if the parcel is selected. The probability that the species is absent from all selected parcels is therefore equal to $\prod_{i\in N}(1 - p_{ki}x_i)$ and, ultimately, the probability that the species is present in the reserve is equal to $1 - \prod_{i\in N}(1 - p_{ki}x_i)$. In order to require this probability to take a value greater than or equal to $\alpha_k (0 \le \alpha_k < 1)$, we introduce the constraint $1 - \prod_{i\in N}(1 - p_{ki}x_i) \geqslant \alpha_k$. This nonlinear constraint, can be transformed, as for example Haight et al. (2000) do, in a linear constraint using the logarithmic function. The case (i) can be then formulated by the linear program in 0–1 variables

$$(\text{P13}):\quad \min_{x\in\{0,1\}^n} \left\{ \sum_{i\in N} c_i x_i : \sum_{i\in N} x_i \log q_{ki} \leqslant \log(1-\alpha_k)\ (k\in P) \right\}.$$

The case (ii) can be formulated by the nonlinear program in 0–1 variables

$$(\text{P14}):\quad \max_{x\in\{0,1\}^n} \left\{ \sum_{k\in P} \Big(1 - \prod_{i\in N}(1 - p_{ki}x_i)\Big) : \sum_{i\in N} c_i x_i \leqslant B \right\}$$

which is difficult to solve. In case we try to select a subset of $B$ parcels ($c_i = 1$), Polasky et al. (2000) describe a classic greedy heuristic in which a new parcel is added to the already selected set of parcels until $B$ parcels are selected, in such a way that the marginal improvement of the objective function at each stage is maximized. The problem can also be treated by classic metaheuristics. We propose in (Billionnet, 2011a) another way to solve (P14) by first rewriting it as

$$(\text{P15}): \max_{\substack{x\in\{0,1\}^n\\ 0\leqslant v_k<1\ (k\in P)}} \left\{ \begin{array}{l} \sum_{k\in P} v_k : \log(1-v_k) \geqslant \sum_{i\in N} x_i \log q_{ki}\ (k\in P),\\ \sum_{i\in N} c_i x_i \leqslant B \end{array} \right\}.$$

then by considering a linear relaxation of (P15), based on the following property: $\forall(\theta,\beta)\in R_+^2, \frac{1}{\theta}(\beta-\theta)+\log\theta \geqslant \log\beta$. This property results from the fact that $\log x$ is a concave function of $x$ and $1/\theta$ is the value of the derivative of $\log x$ at the point $\theta$. A relaxation of (P15) is given by the mixed-integer linear program

$$(\text{P16}): \max_{\substack{x\in\{0,1\}^n\\ 0\leqslant v_k\leqslant 1\ \ (k\in P)}} \left\{ \sum_{k\in P} v_k : \frac{1-v_k-\theta_r}{\theta_r} + \log\theta_r \geqslant \sum_{i\in N} x_i \log q_{ki}\ (k\in P, r\in H),\ \sum_{i\in N} c_i x_i \leqslant B \right\}$$

where $H = \{1, 2, \ldots, h\}$ and $\theta$ is any vector of $R^h$ such that $0 < \theta_1 < \theta_2 < \cdots < \theta_h = 1$. An optimal solution of (P16) gives both a feasible solution of the initial problem, i.e. the selected parcels, and an upper bound of its optimal value. The feasible solution is given by the vector $x$ and its value is $\sum_{k\in P}(1-\prod_{i\in N}(1-p_{ki}x_i))$. The upper bound is given by the optimal value of (P16). We thus obtain by mixed-integer linear programming an approximate solution and an upper bound to the difference between the value of this approximate solution and the optimal value. The experiments, carried out on instances with 400 parcels and 300 species have shown that by choosing properly the vector $\theta$, we obtained relatively quickly, using the software CPLEX 10.2.0 (CPLEX, 2007), an approximate solution at less than 0.4% from the optimum. Camm et al. (2002) propose for this problem an approach based on an approximation of the logarithmic function by a piecewise linear function, but their approach provides an approximate solution of the problem without any guarantee in relation to the optimum.

A robust approach taking into account some uncertainty about the probabilities $p_{ki}$ is also possible for this problem. Following Bertsimas and Sim (2004), one could determine an optimal reserve ensuring a minimum probability of presence for each species whatever are the values of the probabilities $p_{ki}$ (in a set of allowable values). The problems formulated by (P13) and (P16) are relevant if $p_{ki}$ is the probability that the species $e_k$ is present (and survives) in the parcel $s_i$ when that parcel is protected.

Some authors have focused on the uncertainty regarding the long-term survival of a species that is currently present somewhere in the reserve (see e.g., Margules et al., 1994; Williams and Araújo, 2000, 2002; Araújo et al., 2002). Some authors use metapopulation models to incorporate uncertainty into conservation decision-making processes by considering the extinction risk of the metapopulation of interest. The extinction risk is a function of the location, size and quality of the patches of habitat and the ecology of the species. For example, Nicholson and Possingham (2007) define the probability of extinction of a metapopulation of species $e_k$ as a function of the size and spatial arrangement of the reserves, the mean dispersal distance of species $e_k$, the home range size of species $e_k$, and emigration rate per unit area of species $e_k$. They use information-gap decision theory to explore the impact of uncertainty in conservation decisions. McCarthy (2009) presents an overview of different types of spatial population viability analysis. Game et al. (2008) are interested in the optimal strategy for protection of marine habitats in the presence of uncontrollable disturbance. They propose an interesting probabilistic approach for this problem and apply it to a situation of cyclone disturbance of coral reefs on Australia's Great Barrier Reef.

### 2.6. Selection of reserves for maximizing species diversity

We propose in this section a model for selecting parcels in order to maximize the species diversity in the reserve. There are several ways to measure the species diversity of a population composed of different species. We consider here that it is measured by the Simpson diversity index (Simpson, 1949; Couvet and Teyssèdre-Couvet, 2010), i.e. by the probability that two individuals drawn at random from the population belong to two different species. Species diversity of a population composed of species $e_1, e_2, \ldots, e_p$ is equal to $1-\sum_{k\in P} f_k^2$ where $f_k$ is the frequency of the species $e_k$. As in previous cases, we consider a set of species $E = \{e_1, e_2, \ldots, e_p\}$ distributed over a set of parcels $S = \{s_1, s_2, \ldots, s_n\}$. Let $p_{ki}$ be population of the species $e_k$ in the parcel $s_i$ ($p_{ki} = 0$ if the population of the species $e_k$ in the parcel $s_i$ is not sufficient for the survival of this species in this parcel). The (global) frequency of the species $e_k$ in the set of parcels $S$ is $f_k = \sum_{i=1}^n p_{ki}/\sum_{r=1}^p\sum_{i=1}^n p_{ri}$. Let $c_i$ be the cost associated with the selection of the parcel $s_i$, and $B$ be the available budget. The problem is to identify a subset of parcels that respects the budget constraint and such that the species diversity of species present in the parcels is maximum. We can formulate this problem in several ways. Denoting by $f_k$ the real variable representing the frequency of the species $e_k$ in the selected reserve, we get the mixed-integer nonlinear program

$$(\text{P17}): \min_{\substack{x\in\{0,1\}^n\\ f\in[0,1]^P}} \left\{ \sum_{k\in P} f_k^2 : \sum_{i\in N} p_{ki}x_i = f_k \sum_{(r,i)\in P\times N} p_{ri}x_i\ (k\in P), \sum_{(r,i)\in P\times N} p_{ri}x_i \geqslant 1, \sum_{i\in N} c_i x_i \leqslant B \right\}.$$

The objective function of (P17) is quadratic and convex, the first set of constraints is quadratic and the others are linear. One can easily linearize the quadratic constraints by replacing the product $f_k x_i$ by the variable $z_{ki}$ and adding the usual linearization constraints $z_{ki} \leqslant x_i$, $z_{ki} \leqslant f_k$. Because of the objective function to minimize, constraints $z_{ki} \geqslant f_k - (1-x_i)$ and $z_{ki} \geqslant 0$ are useless. The resulting mathematical program consists in minimizing a convex quadratic function subject to linear constraints and can be therefore solved by classical solvers such as CPLEX (2007). Another formulation is given by the fractional program

$$(\text{P18}): \min_{x\in\{0,1\}^n} \left\{ \sum_{(i,j)\in N^2} \alpha_{ij}x_ix_j \Big/ \sum_{(i,j)\in N^2} \beta_{ij}x_ix_j : \sum_{(r,i)\in P\times N} p_{ri}x_i \geqslant 1, \sum_{i\in N} c_i x_i \leqslant B \right\}$$

where $\alpha_{ij} = \sum_{k\in P} p_{ki}p_{kj}$ and $\beta_{ij} = \left(\sum_{k\in P} p_{ki}\right)\left(\sum_{k\in P} p_{kj}\right)$. The auxiliary problem associated with (P18) consists in minimizing the quadratic function of Boolean variables $\sum_{(i,j)\in N^2} \alpha_{ij}x_ix_j - \lambda\sum_{(i,j)\in N^2} \beta_{ij}x_ix_j$, where $\lambda$ is a real parameter, under the same constraints as those of (P18) (see Section 2.3.3). Although (P18) is more effective than (P17) it only provides exact solutions for small instances (few tens of parcels and species).

The programs (P17) and (P18), like the Simpson index, take into account the number and the frequency of the species but do not take into account the populations of each species. It is easy to overcome this drawback by introducing, for example, the constraints $\sum_{i\in N} p_{ki}x_i \geqslant b_k \quad (k\in P)$ that impose the minimum value $b_k$ to the population of the species $e_k$. Another interesting problem is to select the reserve by using the phylogenetic diversity criterion (see e.g., Rodrigues and Gaston, 2002a; Pardi and Goldman, 2007 and Section 6.4 of this article).

### 2.7. Dynamic reserve selection

In the problems described above, the temporal dimension does not appear while it is of course very important. In this section, we first briefly describe an interesting model (P19), proposed by Toth et al. (2011) that maximizes the expected total biodiversity value of land in both preserved and undeveloped parcels within a competitive land market. This model considers several temporal aspects. For example, the conservation actions may have an impact on the price of the land and on the risk of development outside the reserves. It also takes account of the essential fact that once a land parcel is developed it cannot be purchased for conservation.

Let us consider a set of parcels $S = \{s_1, s_2, \ldots, s_n\}$ and a planning horizon composed of $q$ periods of one year $T = \{1, 2, \ldots, q\}$. For each parcel $s_i$, we know the area $a_i$ in hectares, the biodiversity value $v_i$ measured in biodiversity hectares, and the set of adjacent parcels $S_i$. The model uses the following Boolean variables: $x_{it}$ equals 1 iff parcel $s_i$ is selected for conservation in year $t$, $z_{it}$ equals 1 iff parcel $s_i$ is converted to development in year $t$, and $y_{it}$ equals 1 iff at least one parcel that is adjacent to parcel $s_i$ is selected for conservation in year $t-1$. It also uses the positive real variable $p_{it}$ which is equal to the market value of parcel $s_i$ in year $t$. At the beginning of the planning horizon, $p_{i1}$ is equal to the current market value of parcel $s_i$.

$$
(\mathrm{P19}):\quad \max_{\substack{x\in\{0,1\}^{nq}\\ z\in\{0,1\}^{nq}\\ y\in\{0,1\}^{n(q-1)}\\ p\in R_+^{nq}}}
\begin{cases}
\sum_{i\in N} a_i v_i \sum_{t\in T} x_{it} + \alpha \sum_{i\in N} a_i v_i \left(1 - \sum_{t\in T}(x_{it}+z_{it})\right): \\
\sum_{t\in T}(x_{it}+z_{it}) \leqslant 1 \ (i\in N),\ \sum_{i\in N} p_{it}x_{it} \leqslant B_t \ (t\in T), \\
p_{it} = p_{i,t-1}(1+r+qy_{it}) + E_i \sum_{j\in N-\{i\}} a_j x_{j,t-1}, \\
\sum_{k\in S_i} x_{k,t-1} \geqslant y_{it},\ \sum_{k\in S_i} x_{k,t-1} \leqslant |S_i| y_{it}, \\
p_{it} \geqslant (R_i+\theta_{it}) z_{it},\ \left(1 - \sum_{t'=1}^{t}(x_{it'}+z_{it'})\right) \\
\quad (R_i+\theta_{it}-p_{it}) \geqslant 0 \ (i\in N, t\in T-\{1\})
\end{cases}
$$

The objective function of (P19) maximizes the total amount of land in preserved and undeveloped parcels weighted by each parcel's biodiversity value. $\alpha$ is a biodiversity correction coefficient for unprotected but undeveloped parcels. The first constraint allows a parcel to be either developed or preserved at most once during the planning horizon. The second constraint is a budget constraint where $B_t$ is the available budget in period $t$. The third constraint expresses the expected market value of parcel $s_i$ in year $t$ as the sum of two parts: the first one is the market value of parcel $s_i$ in the year $t-1$, $p_{i,t-1}$, compounded by the expected increase (or decrease) in real estate value $r$ as dictated by the general housing market, plus an amenity premium $q$ that is taken into account only if at least one of the parcels that are adjacent to parcel $s_i$ has been purchased for conservation in year $t-1$ (i.e. if $y_{it} = 1$); the second part $E_i\sum_{j\in N-\{i\}} a_j x_{j,t-1}$ is a price increase associated with an increase demand for land resulting from the conservation purchases in year $t-1$. $E_i = a_i/(\eta^d + \eta^s)$ is a parameter associated with parcel $s_i$ which depends on the price elasticities $\eta^d$ and $\eta^s$ of demand and supply, respectively, for housing development. The two constraints that follow control the value of the binary variables $y_{it}$. The two last constraints set the value of the development indicator variable $z_{it}$. $\theta_{it}$ is a cutoff value, i.e. an amount of money by which the market value of parcel $s_i$ ($p_{it}$) must exceed its open space value ($R_i$) in period $t$ before the parcel is considered developed by the model. Some nonlinearities appear in (P19). They result from the product of the adaptative price coefficient $p_{it}$ by the Boolean variables $x$, $y$ or $z$. The classical linearization techniques allow (P19) to be linearized (see linearization of (P17)). Toth et al. (2011) proved that this integer programming approach was computationally tractable for a large and complex problem instance in a real land market on Lopez Island, Washington.

There are many other publications on the temporal dimension in reserve selections. The reader can consult, for example, the references (Costello and Polasky, 2004; Strange et al., 2006; Sabbadin et al., 2007; Rayfield et al., 2008; Dissanayake and Önal, 2011) where the proposed models are often based on dynamic programming.

We now consider the model proposed by Haight et al. (2005) that involves both a temporal dimension and an uncertainty about parcel availability. This model has two objectives: maximize the expected number of species represented in protected parcels and maximize the expected number of people with access to protected sites. To simplify the presentation we shall consider here only the first objective. That is sufficient to illustrate the approach adopted to take account of uncertainty over a two-period horizon. The reader can refer to (Haight et al., 2005) for a comprehensive overview of the model. We consider a planning horizon composed of two periods and among the considered parcels, some are available for protection in the first period and others are not. Each site not protected in the first period has a probability of remaining undeveloped and being available for protection in the second period. This uncertainty is handled by defining a set of development scenarios $\Theta = \{1, \ldots, v\}$. Specifically, we know, for every possible scenario $\theta \in \Theta$, the probability $p_\theta$ it occurs and, for each site $s_i$, the Boolean parameter $u_{i\theta}$ which equals 1 iff the site $s_i$ is available for protection in the second period when the scenario $\theta$ is realized. As in model (P1) we are interested in a set of species $E = \{e_1, e_2, \ldots, e_p\}$ to be protected and living on a set of parcels $S = \{s_1, s_2, \ldots, s_n\}$. For each parcel $s_i$, we know its area $a_i$, all the species that live there, and it is assumed that these species will survive in $s_i$ if $s_i$ is protected. As in (P1), $S_k$ is the set of indices of parcels that protect the species $e_k$. Associate with each parcel $s_i$ a Boolean variable $x_i^1$ that equals 1 iff this parcel is selected to be protected in period one, and a Boolean variable $x_{i\theta}^2$ that equals 1 iff this parcel is selected to be protected in period two in scenario $\theta$. Associate with each species $e_k$ and each scenario $\theta$ a Boolean variable $y_{k\theta}$ that equals one iff species $e_k$ is represented in at least one protected parcel in scenario $\theta$. Let $P = \{1, \ldots, p\}$ and $N = \{1, \ldots, n\}$. The determination of a subset of parcels which maximizes the expected number of species represented can be formulated by the 0–1 linear program

$$
(\mathrm{P20}):\quad \max_{\substack{x^1\in\{0,1\}^{n}\\ x^2\in\{0,1\}^{nv}\\ y\in\{0,1\}^{pv}}}
\begin{cases}
\sum_{\theta\in\Theta} p_\theta \sum_{k\in P} y_{k\theta}: \\
x_i^1 + x_{i\theta}^2 \leqslant 1 \ (i\in N, \theta\in\Theta), \\
\quad x_{i\theta}^2 \leqslant u_{i\theta} \ (i\in N, \theta\in\Theta),\ x_i^1 = 0 \ (i\in L), \\
\sum_{i\in N} a_i x_i^1 + \sum_{i\in N} a_i x_{i\theta}^2 \leqslant B \ (\theta\in\Theta), \\
\quad y_{k\theta} \leqslant \sum_{i\in S_k}\left(x_i^1 + x_{i\theta}^2\right) \ (k\in P, \theta\in\Theta)
\end{cases}
$$

The objective function is the number of species represented under each scenario weighted by the probability of that scenario's occurrence. The first constraint specifies that site $s_i$ can be selected in either period one or period two, but not both, over all scenarios. The second constraint specifies that site $s_i$ can only be selected in period two in scenario $\theta$ if site $s_i$ is available for protection in that scenario. The third constraint specifies that site $s_i$ ($i \in L$), where $L$ is a subset of $S$, cannot be selected in period one. The fourth constraint places an upper bound $B$ on the total area of selected sites in periods one and two under each scenario. The last constraint is similar to the first constraint of (P2) and defines the condition under which species $e_k$ is represented under scenario $\theta$. Let $(x^{1*}, x^{2*}, y^*)$ be an optimal solution of (P20). This indicates a need to acquire, in the first period, the parcels $s_i$ such as $x_i^{1*} = 1$.Then one of the $q$ scenarios occurs. Let $\hat{\theta}$ be this scenario. It is then necessary to acquire in the second period the parcels $s_i$ such that $x_{i\hat{\theta}}^2 = 1$. Finally, the number of protected species is equal to $\sum_{k\in P} y_{k\hat{\theta}}$.

A lot of other work focused on dynamic selection of nature reserves. For example, Strange et al. (2011) study a dynamic spatial conservation prioritization problem where climate change gradually changes the future habitat suitability of a site' current species. The problem is explored using heuristic approaches. Pressey et al. (2007) summarize ideas, techniques and unresolved issues in planning for biodiversity processes (local extinction, recolonization, predation, adjustment of the distribution of species to changing climate, etc.) in the context of dynamics threats (habitat conversion by agriculture, harvesting of timber, invasive plants and animals, etc.).

## 3. Control of adverse effects caused by landscape fragmentation

Fragmentation of spaces has become a central issue in ecology because it is recognized as the main source of biodiversity loss in industrialized countries. It causes a decrease of the habitats and their division into separate fragments, thus affecting the populations of many species: extinction of rare species, increasing of common and sedentary species, mortality when moving between fragments, genetic erosion, etc. (see e.g., Harrison and Bruna, 1999; Collinge, 2000; Fahrig, 2003; Crist, 2004). An important part of this fragmentation is caused by road and rail infrastructure but is also due to urban development, agriculture and forestry.

We present below several optimization problems related to landscape fragmentation. The first type of problem is to select, under constraints, a set of areas, among a set of candidate areas, in order to optimize some standard indicators of fragmentation. These problems can be formulated more or less easily as fractional combinatorial optimization problems for which there are efficient algorithms. For example, a classical iterative algorithm in which each iteration consists in solving a linear program with integer or mixed-integer variables quickly solves instances with hundreds of candidate areas. Note that the approaches that are used could be easily extended to other indicators. The second type of problem is to define and/ or restore biological corridors in order to reduce as far as possible adverse effects due to landscape fragmentation by reconnecting habitats that are no longer connected. The treatment of these problems involves both graph theory (optimal flows) and mathematical programming and are usually effective. We also consider the problem of restoring a network of existing corridors in order to increase as far as possible its permeability. The probabilities of the movements of species in the network are known and the model built in this framework is based on both Markov chains theory and mathematical programming. Instances of medium size can be solved optimally by this approach. One limitation of this model is the practical difficulty in assessing the movements probabilities.

### 3.1. Minimizing fragmentation

Landscape fragmentation, i.e. the breakup of large areas into smaller parcels must be evaluated and controlled because it influences many ecological processes (see e.g., Barrett and Peles, 1994). It can be measured by multiple indicators (see e.g., Hargis et al., 1998; McAlpine and Eyre, 2002; Öhman and Lämas, 2005). Here, we consider as examples two common indicators, the *mean nearest neighbor distance* (MNND) and the *mean proximity index* (MPI), which estimate the relative isolation of the parcels. These indicators are among those that are used in FRAGSTATS (McGarigal and Marks, 1994), software well known in landscape ecology. Optimization problems associated with the notion of fragmentation include, for example, the selection of a subset of parcels optimal for these indicators. The selected parcels will be, for example, protected or restored as habitat patches and the others will be assigned to other uses (urban development, agricultural areas, logging, etc.). Now let us see precisely the two measures of fragmentation that we consider here, for a set $S = \{s_1, s_2, \ldots, s_n\}$ of parcels. $\text{MNND} = (1/n)\sum_{i=1}^{n} \min\{d_{ij} : j = 1, \ldots, n;\ j \neq i\}$ where $d_{ij}$ is the distance between the two parcels $s_i$ and $s_j$. Many studies have shown that the movements of individuals within a species between different parcels affect the dynamics of the population (see e.g., Fahrig, 2007) and it is clear that these movements are more difficult if the values of MNND are high. Minimizing the value of the indicator privileges the grouping of parcels that are not too far from each other. However, the distance between groups of parcels is not controlled. Although MNND is often used, the only consideration of the nearest parcel may not well represent, from an ecological point of view, the vicinity of a parcel. In addition MNND ignores the size of the parcels. To overcome this drawback, we can consider the indicator MPI ($d$) which takes into account all the parcels that are within a certain given radius $d$ from the parcel considered, and their sizes. Setting $R_i(d) = \{j \in \{1, \ldots, n\}: j \neq i,\ d_{ij} \leqslant d\}$ and denoting $a_i$ the area of the parcel $s_i$, MPI ($d$) is equal to $(1/n)\sum_{i=1}^{n}\sum_{j \in R_i(d)}(a_j/d_{ij}^2)$.

Now consider two problems of optimal selection of parcels, inspired from (Vemema et al., 2005). These authors use a genetic algorithm to solve such problems. A solution by mathematical programming is presented in (Billionnet, 2010b). Consider a set of parcels $S = \{s_1, s_2, \ldots, s_n\}$ of total area $A$ and seek to select a subset of $S$ that minimizes MNND under the constraint that the total area of the selected parcels is between $A_{\min}$ and $A_{\max}$ (Fig. 3). By associating with each parcel $s_i$ the Boolean variable $x_i$ that is equal to 1 iff the parcel $s_i$ is selected, the problem can be formulated by the combinatorial optimization problem

$$(\text{P21}): \quad \min_{x \in \{0,1\}^n} \left\{ \sum_{i \in N(x)} \min_{j \in N_i(x)} d_{ij} \Big/ \sum_{j \in N} x_j : A_{\min} \leqslant \sum_{j \in N} a_j x_j \leqslant A_{\max} \right\}$$

where $N = \{1, \ldots, n\}$, $N(x) = \{i \in N: x_i = 1\}$, $N_i\ (x) = \{j \in N: j \neq i,\ x_j = 1\}$ $(i \in N)$. Let $E = \{(i,j) \in N^2,\ i \neq j\}$ and $N_i = \{j \in N: j \neq i\}$. For a given index $i$, the quantity $\min_{j \in N_i(x)} d_{ij}$ is equal to the distance between the parcel $s_i$ and its nearest neighbor among the selected parcels. $\sum_{i \in N(x)} \min_{j \in N_i(x)} d_{ij}$ represents the sum of these distances for all the selected parcels. By dividing the latter quantity by the number of selected parcels, i.e. $\sum_{j \in N} x_j$, we obtain the desired MNND. (P21) can be rewritten as the 0–1 linear-fractional program

$$(\text{P22}): \min_{\substack{x \in \{0,1\}^n \\ y \in \{0,1\}^{|E|}}} \left\{ \frac{\sum_{(i,j) \in E} d_{ij} y_{ij}}{\sum_{i \in N} x_i} : \sum_{j \in N_i} y_{ij} = x_i\ (i \in N),\ y_{ij} \leqslant x_j\ ((i,j) \in E), \right.$$
$$\left. A_{\min} \leqslant \sum_{j \in N} a_j x_j \leqslant A_{\max} \right\}.$$

MNND is equal to the objective function of (P22) subject to the first two constraints. Indeed, according to the first constraint, for each index $i$, if $x_i = 0$ then all the variables $y_{ij}$ take the value 0 and if $x_i = 1$ then only one variable $y_{ij}$ takes the value 1. Moreover, the second constraint forces the variables $y_{ij}$ to take the value 0 if $x_i$ or $x_j$ is set to 0. The auxiliary problem (See Section 2.3.3) associated with (P22) is a linear program in 0–1 variables that consists in minimizing the expression $\sum_{(i,j) \in E} d_{ij} y_{ij} - \lambda \sum_{i \in N} x_i$, where $\lambda$ is a real parameter, under the same constraints as those of (P22). Again, we refer to (Schaible, 1995; Radzik, 1998) for the treatment of fractional programs. One can also seek to minimize MNND while imposing a maximum value, $v$, to the mean shape index (MSI) presented in Section 2.3.3. To do this, simply add to (P22) the linear constraint $0.25\sum_{i \in N}(p_i/\sqrt{a_i})x_i \leqslant v\sum_{i \in N} x_i$.

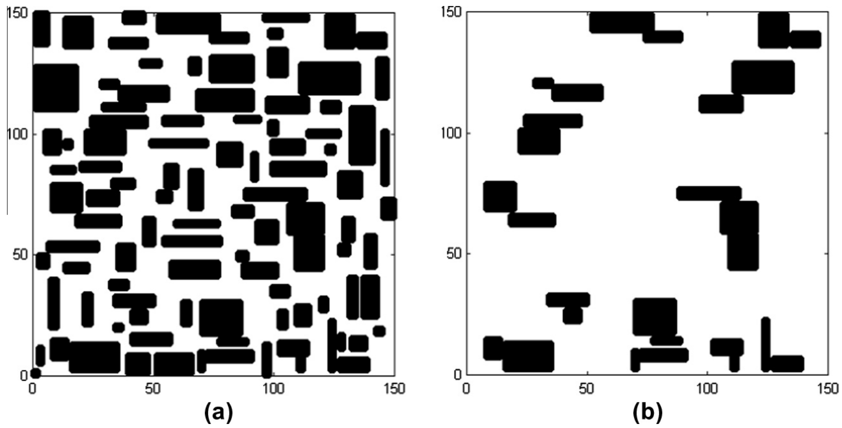


**Fig. 3.** (Billionnet, 2010b). Minimizing MNND. (a) A square area of 15 × 15 km including 100 rectangular parcels (total perimeter: 348.20 km, total area: 65.85 km$^2$, MNND: 3.21 hm). (b) For a total area between 40% and 50% of the initial area, the optimal value of MNND is equal to 1.63. The solution is composed of 12 relatively isolated groups, each group comprising 2–4 close parcels.

The second problem that we present here is to select a subset of parcels that maximizes the value of the indicator MPI ($d$) and which total area is between $A_{\min}$ and $A_{\max}$. It can be formulated by the program

$$(\text{P23}): \quad \max_{x\in\{0,1\}^n}\left\{\sum_{i\in N(x)}\sum_{j\in N_i(x,d)}\frac{a_j}{d_{ij}^2}\Big/\sum_{i\in N}x_i : A_{\min}\leqslant\sum_{j\in N}a_jx_j\leqslant A_{\max}\right\}$$

where $N_i(x,d) = \{j \in N{:}j \neq i,\ x_j = 1,\ d_{ij} \leqslant d\}$. (P23) is equivalent to the fractional program

$$(\text{P24}): \quad \max_{\substack{x\in\{0,1\}^n\\ v\in R_+^n}}\left\{\begin{array}{l}\frac{\sum_{i\in N}v_i}{\sum_{i\in N}x_i}: v_i\leqslant\sum_{j\in N_i(d)}\frac{a_jx_j}{d_{ij}^2}\ (i\in N),\ v_i\leqslant C_ix_i\ (i\in N),\\ A_{\min}\leqslant\sum_{j\in N}a_jx_j\leqslant A_{\max}\end{array}\right\}.$$

where $C_i$ is a constant greater than the value of $\sum_{j\in N_i(d)}\left(a_jx_j/d_{ij}^2\right)$ at the optimum of (P23).

MPI ($d$) is equal to the objective function of (P24) subject to the first two constraints. Indeed, according to these two constraints, $v_i$ will take the value $\min\left\{\sum_{j\in N_i(d)}\left(a_jx_j/d_{ij}^2\right), C_ix_i\right\}$ at the optimum since the objective function is to be maximized. The variable $v_i$ will therefore take the value $\sum_{j\in N_i(d)}\left(a_jx_j/d_{ij}^2\right)$ if $x_i = 1$, and the value 0 if $x_i = 0$. The auxiliary problem associated with (P24) is a mixed-integer linear program that consists in maximizing the expression $\sum_{i\in N}v_i - \lambda\sum_{i\in N}x_i$ under the same constraints as those of (P24). The solution of programs (P22) and (P24) is very fast since, for all values of the couple ($A_{\min}$,$A_{\max}$) we have considered, their solution requires less than three seconds for a hundred parcels, on a PC with an Intel Core Duo 2 gigahertz processor. In these experiments, the solver CPLEX 10.2.0 (CPLEX, 2007) has been used to solve the auxiliary problems in the course of executing Dinkelbach's algorithm (Dinkelbach, 1967).

The formulations (P22) and (P24) were tested on sets of rectangular parcels. Note that these formulations remain valid in more realistic situations where the parcels are irregular polygons.

### 3.2. Connecting a set of nature reserves by biological corridors

As discussed above, landscape fragmentation is a major cause of biodiversity loss. The viability of the species involved in this fragmentation depends strongly on how the fragments are connected. The preservation and restoration of ecological continuity, i.e. of biological corridors, is a way to halt the decline of biodiversity. These corridors are routes used by the animal species to move, reproduce, escape, migrate, and so on. They depend on the species. This ability to move is essential (see e.g., Clergeau and Désiré, 1999; Ménard and Clergeau, 2001; Holzgang et al., 2001; Fahrig, 2007) and the restoration of these corridors is a key protection strategy for the species threatened by the fragmentation of their habitat. The reader can also consult the reference (Bond, 2003) which deals with principles for the design and evaluation of biological corridors, and the reference (Williams and Snyder, 2005) which addresses the question of where restoration should take place to efficiently reconnect habitat with a landscape-spanning corridor. Building upon findings in percolation theory, the authors develop a shortest-path optimization methodology for establishing such corridors.

We present in this section a problem of restoration of corridors inspired by (Williams, 1998). Consider a landscape represented by a matrix of $m \times n$ identical square parcels. This landscape has a number of reserves $R_1, R_2, \ldots, R_T$, each reserve being formed by a connected subset of parcels and all pairs of reserves are not connected.

Among the parcels that do not belong to a reserve, some can be restored and some cannot. The cost of restoration of the parcel $s_{ij}$ is denoted by $c_{ij}$. In cases where $s_{ij}$ is already a parcel habitat but not integrated in a reserve, this cost is zero. A large cost is assigned to a parcel that cannot be restored. The problem is to determine the parcels to restore in order to connect all reserves at least cost. Two reserves $R_i$ and $R_j$ are considered as connected if an animal can move from $R_i$ to $R_j$ passing only in restored parcels (or belonging to a reserve) and moving gradually from a parcel to an adjacent parcel, two parcels being adjacent if they share a common side

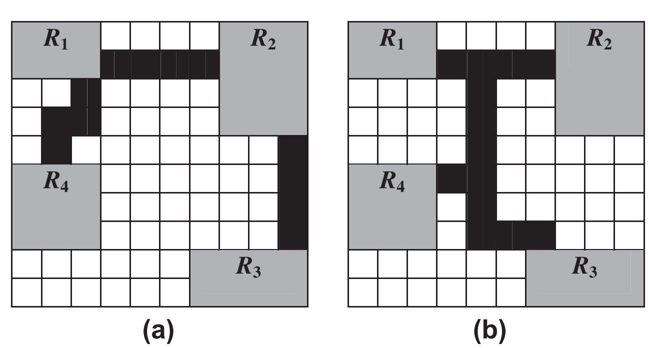


**Fig. 4.** (a) The restoration of 12 parcels connect the four reserves $R_1$, $R_2$, $R_3$, $R_4$. The length of the corridor connecting $R_3$–$R_4$ is 18. (b) The restoration of 13 parcels connect the four reserves. The length of the corridor connecting $R_3$–$R_4$ is equal to 6.

(Fig. 4). Associate with the matrix of parcels a graph $G = (S,U)$. $S = \{(i,j) \in M \times N\}$, where $M = \{1,\dots,m\}$ and $N = \{1,\dots,n\}$, is the set of vertices (corresponding to the parcels) and $((i,j),(k,l))$ is an arc of $G$ iff the parcels $s_{ij}$ and $s_{kl}$ are adjacent. For each reserve $R_k$, choose one of its parcels, $s_{i_k j_k}$, to represent it. The problem can be formulated as follows: determine a subset of vertices of minimum cost $\bar{S} \subset S$ and a subset of arcs $A \subset U$ such that the graph $G = (\bar{S} \cup_{k=1,\dots,T} (i_k, j_k), A)$ admits a path from the parcel representing the reserve $R_r$ to the parcel representing the reserve $R_T$, and this for all $r \in \{1,\dots,T-1\}$. This problem corresponds to the Steiner tree problem (see e.g., Hwang et al., 1992) which we give below a flow type formulation (Williams, 1998). Let $x_{ijkl}$ be the Boolean variable that is equal to 1 iff at least one of the $T$-1 paths from $R_r$ ($r = 1,\dots,T-1$) to $R_T$ uses the arc $((i,j),(k,l))$, and $y^r_{ijkl}$ be the Boolean variable that is equal to 1 iff the path from $R_r$ to $R_T$ uses the arc $((i,j),(k,l))$. Denoting by $A_{ij}$ the set of vertices corresponding to the parcels adjacent to the parcel $s_{ij}$ and setting $d^r_{ij} = 1$ if the parcel $s_{ij}$ represents the reserve $R_r$, $d^r_{ij} = -1$ if the parcel $s_{ij}$ represents the reserve $R_T$ and $d^r_{ij} = 0$ otherwise, we obtain the formulation of the problem by the linear program in 0–1 variables

$$(\text{P25}): \min_{\substack{y \in \{0,1\}^{(T-1)\times|U|} \\ x \in \{0,1\}^{|U|}}} \left\{ \begin{array}{l} \sum_{(i,j)\in S} c_{ij} \sum_{(k,l)\in A_{ij}} x_{ijkl} : x_{ijkl} \geqslant y^r_{ijkl}\ ((i,j),(k,l)) \in U, \\ r = 1,\dots,T-1, \\ \sum_{(k,l)\in A_{ij}} y^r_{ijkl} - \sum_{(k,l)\in A_{ij}} y^r_{klij} = d^r_{ij}\ ((i,j) \in S, \\ r = 1,\dots,T-1) \end{array} \right\}.$$

Note that $c_{ij} = 0$ if $s_{ij}$ belongs to one of the $T$ reserves. The first constraint of (P25) expresses that if the path from $R_r$ to $R_T$ uses the arc $((i,j),(k,l))$ then the variable $x_{ijkl}$ must take the value 1. The second constraint is a constraint of conservation at each vertex. For example, consider a vertex $(i,j)$ which does not represent a reserve: if an arc of the path $r$ reaches the vertex $(i,j)$, then an arc of the same path must also start from this vertex. This type of model is used by Suter et al. (2008) to define a network of corridors suitable for grizzly bear movements in the U.S. Northern Rockies. To control corridor lengths for the couples of reserves $(R_r,R_T)$ ($r = 1,\dots,T-1$), we can simply add the constraint $\sum_{((i,j),(k,l))\in U, s_{ij}\notin R_r\cup R_T} y^r_{ijkl} \leqslant L^{rT}_{\max}$, $L^{rT}_{\max}$ being the maximum length allowed for the corridor connecting $R_r$ to $R_T$ (Williams, 1998). As mentioned in that article, a disadvantage of this model is that the other corridor lengths cannot be controlled. We propose below a slightly different formulation that does not have this disadvantage. We keep the variables $y^r_{ijkl}$ and we replace the Boolean variables $x_{ijkl}$ with the Boolean variables $w_{ij}$ that are equal to 1 iff at least one of the $T$-1 paths from $R_r$ ($r = 1,\dots,T-1$) to $R_T$ goes through the node $(i,j)$. We get the program

$$(\text{P26}): \min_{\substack{y \in \{0,1\}^{(T-1)\times|U|}, \\ w \in \{0,1\}^{|S|}}} \left\{ \begin{array}{l} \sum_{(i,j)\in S} c_{ij} w_{ij} : w_{ij} \geqslant \sum_{(k,l)\in A_{ij}} y^r_{ijkl}\ ((i,j) \in S, \\ r = 1,\dots,T-1), \\ \sum_{(k,l)\in A_{ij}} y^r_{ijkl} - \sum_{(k,l)\in A_{ij}} y^r_{klij} = d^r_{ij}\ ((i,j) \in S, \\ r = 1,\dots,T-1) \end{array} \right\}.$$

(P26) has fewer variables and fewer constraints than (P25). By taking into account additional constraints, a limit can be imposed on the length of the corridor linking any two reserves. As in the previous model we can introduce a constraint on the length of the corridor between $R_r$ and $R_T$. To limit at $L^{st}_{\max}$ the length of the corridor between two arbitrary reserves $R_s$ and $R_t$, different from $R_T$, we define a new Boolean variable $y^{st}_{ijkl}$, which is equal to 1 iff the path from $R_s$ to $R_t$ uses the arc $((i,j),(k,l))$, and we introduce the additional constraints $\sum_{(k,l)\in A_{ij}} y^{st}_{ijkl} - \sum_{(k,l)\in A_{ij}} y^{st}_{klij} = d^{st}_{ij} \quad ((i,j) \in S)$ where $d^{st}_{ij} = 1$ if $s_{ij}$ represents $R_s$, $d^{st}_{ij} = -1$ if $s_{ij}$ represents $R_t$ and $d^{st}_{ij} = 0$ otherwise, $w_{ij} \geqslant \sum_{(k,l)\in A_{ij}} y^{st}_{ijkl} \quad ((i,j) \in S)$ and $\sum_{((i,j),(k,l))\in U, s_{ij}\notin R_s\cup R_t} y^{st}_{ijkl} \leqslant L^{st}_{\max}$. We tested (P26), with additional constraints to impose a maximum length to a certain set of corridors, on a hypothetical instance including $20 \times 20$ identical square parcels and seven reserves. For all additional constraints we considered, the corresponding version of (P26) has been solved in a few seconds of computation. In formulations (P25) and (P26), the parcel set is based on a grid of identical square cells. In more realistic situations, the land parcels are typically irregular polygons. In this case (P25) and (P26) remain valid but the corresponding graph $G = (S,U)$ is different from that based on regular grids, particularly concerning the vertex degrees.

Other authors examined this type of problem. For example, Alvarez-Miranda et al. (2013) consider the following problem: Given a connected node-weighted graph, with a root node $r$, and a (possibly empty) set of nodes $R$, find a connected subgraph rooted at $r$ that connects all nodes in $R$ with maximum total weight. They propose non-polynomial-sized integer linear programming formulations which are based on node variables only.

### 3.3. Restoration of a network of corridors to increase its permeability

We present in this section a problem of restoring a network of corridors inspired by references (Finke et al., 2008; Finke and Sonnenschein, 2008). The problem is to determine the developments to be made in the corridors of a network in order to increase its permeability as much as possible while taking into account a budget constraint. In the two references cited above, the problem is addressed by simulation methods. We present the solution we propose in (Billionnet, 2010a) and that is based on mixed-integer linear programming. Consider a network of corridors and a set of species having all the same behavior in the network. The network can be represented by a graph which the $n$ vertices are the $n$ parcels (or reserves) $s_1, s_2, \dots, s_n$ corresponding to habitats and which the $m$ edges correspond to corridors connecting the parcels. The road and rail infrastructure, urbanization, agriculture and other phenomena have more or less degraded these corridors and the problem is to restore them as efficiently as possible. One can, for example, create wildlife crossings above or below the road and rail infrastructure (see e.g., Ng et al., 2004; Trocmé, 2005). The aim is to restore the corridors, under a budget constraint, to maximize the permeability of the network. This permeability is measured by the expected value of the distance traveled in the network by the species considered. When an animal lies in the parcel $s_i$, it is assumed that it chooses randomly and with equal probability one of the corridors leading to this parcel, i.e. the corridor $[s_i,s_j]$ with the probability $1/d_i$ where $d_i$ is the degree of the node $s_i$. Then it tries to use this corridor with the probability $q_{ij}$. If it decides to use the corridor $[s_i,s_j]$, it is assumed that it managed to reach the other end, i.e. the parcel $s_j$ with the probability $r_{ij}$ and that it fails to reach it, getting killed before that happens, with the probability $1 - r_{ij}$. The restoration of a corridor $[s_i,s_j]$ increases the probabilities $q_{ij}$ and $r_{ij}$, and the more resources devoted to the restoration, the higher the values of these probabilities will be. We assume that we know, for each corridor $[s_i,s_j]$, the possible values of the probabilities $q_{ij}$ and $r_{ij}$ together with the associated costs. This assumes that we know $q^k_{ij}$ (resp. $r^k_{ij}$), the value of the probability $q_{ij}$ (respectively $r_{ij}$) if the investment of level $k$ ($k = 1,\dots,K$) is carried out in the corridor $[s_i,s_j]$. We associate to the network a Markov chain which set of states consists of $n$ parcels ($n$ transient states) and a state $(n+1)$ (absorption) for the death of the animal. The transition probabilities, $p_{ij}$, from state $i$ to state $j$ are the unknowns. Let $\Pi = \begin{pmatrix} Z & D \\ 0 & 1 \end{pmatrix}$ be the transition probabilities matrix, $Z$ corresponding to transient states and $D$, to the absorbing state. Let $N$ be the matrix which general term, $n_{ij}$, is equal to the average number of passages by state $j$, before absorption, for an animal starting from state $i$. According to the Markov chains theory, $N = (I - Z)^{-1}$ where $I$ is the identity matrix of dimension $n \times n$ (see e.g., Grinstead and Snell, 1997). The choice of investments to make in each corridor to maximize the expected value of the total distance traveled by $n$ animals, one animal being initially in each of the $n$ parcels, can be formulated by the mathematical program

$$(\text{P27}): \max_{\Pi \in \widetilde{\Pi},\ w \in R_+^n} \left\{ \sum_{(i,j)\in R} l_{ij}(w_i p_{ij} + w_j p_{ji}) : w_i - \sum_{j=1}^{n} w_j p_{ji} = 1 \ (i = 1, \ldots, n) \right\}$$

where $R = \{(i,j) \in \{1, \ldots, n\}^2 : [s_i, s_j]$ is a corridor, $i < j\}$, $l_{ij}$ is the length of the corridor $[s_i, s_j]$, $w_i$ is a real variable that represents the expression $\sum_{j=1}^{n} n_{ji}$, and $\widetilde{\Pi}$ is a set of stochastic matrices, of general term $p_{ij}$, and feasible for the problem. Let $S = \{(i,j) \in \{1, \ldots, n\}^2 : [s_i, s_j]$ is a corridor$\}$. Using the Boolean variable $x_{ij}^k \quad ((i,j) \in S, k \in K)$ that is equal to 1 iff the investment of level $k$ is made in the corridor $[s_i, s_j]$ and denoting by $c_{ij}^k$ the cost of this investment, the problem can be formulated by the program

$$(\text{P28}): \max_{x \in \{0,1\}^{2mK}, w \in R_+^n} \left\{ \begin{array}{l} \sum_{(i,j)\in R} l_{ij}(e_{ij} + e_{ji}) : \sum_{(i,j)\in R, k \in K} c_{ij}^k x_{ij}^k \leqslant B, \\ \sum_{k \in K} x_{ij}^k = 1 \ (i,j) \in R, \\ e_{ij} = \frac{1}{d_i} w_i \sum_{k\in K} q_{ij}^k r_{ij}^k x_{ij}^k \ ((i,j) \in S), \\ x_{ij}^k = x_{ji}^k \ ((i,j) \in R, k \in K), \\ \frac{1}{d_i} w_i \sum_{j:(i,j)\in S, k\in K} q_{ij}^k x_{ij}^k = 1 + \sum_{j:(j,i)\in S} e_{ji} \ (i = 1, \ldots, n) \end{array} \right\}$$

where $e_{ij}$ represents the expected value of the total number of travels in the corridor $[s_i, s_j]$ going from the parcel $s_i$ to the parcel $s_j$. (P28) can be transformed into a mixed-integer linear program in a classic way by linearizing the quadratic expressions $w_i x_{ij}^k$, products of a real variable by a Boolean variable. Some experiments have shown that (P28) can be solved in about two minutes of computing time for a network with 10 parcels and 15 corridors in which seven investment levels are possible (Billionnet, 2010a).

## 4. Sustainable use of forests

Following the Food and Agriculture Organization of the United Nations, ''sustainable forest management aims to ensure that the goods and services derived from the forest meet present-day needs while at the same time securing their continued availability and contribution to long-term development. In its broadest sense, forest management encompasses the administrative, legal, technical, economic, social and environmental aspects of the conservation and use of forests''. Publications on the subject that introduce the concept of optimization are numerous and the reader can refer to Martell et al. (1998), Weintraub and Murray (2006), and Bettinger et al. (2007) for an overview of such problems and also to the book of Weintraub et al. (2007) which contains a dozen chapters related to optimization problems in forestry. About half of these chapters take into account biodiversity protection at least to an extent. Among the problems presented in the literature, many of them have a spatial aspect. For example, Toth and McDill (2008) study a forest management problem aiming to produce habitat composed of mature large parcels taking into account the total perimeter of these parcels. The spatial constraint taken into account by Goycoolea et al. (2005) is to limit, for various environmental reasons, the area of the blocks of adjacent parcels cut at each period. These authors are interested in effective formulations of this problem by integer linear programming involving the notion of clique in a graph. Exact methods for this problem had already been proposed by McDill et al. (2002) and Crowe et al. (2003).

We present below in more details several issues of sustainable use of the forest. The first problem concerns the choice of a strategy for forest use in order to ensure the best possible protection for some species. In this example, a first type of species prefers cut forest stands and a second type of species prefers the edges between cut and uncut stands. The optimal strategy is determined by integer linear programming. This problem allows us to illustrate the importance of the formulation in integer linear programming. We also present a dynamic version of this problem. We then present a similar problem but with a kind of species whose population expected in a stand depends on the connectivity of this stand with other stands. This problem also allows us to illustrate the approximate solution of a nonlinear program using the approximation of a nonlinear function by a piecewise linear function. In the last problem, which comprises a temporal dimension, two objectives are pursued: the first concerns the marketing of wood and the second is an environmental objective based on the spatial distribution of trees by age group. This problem can be formulated by a 0–1 linear program. The main modeling difficulty comes from the fact that the edges between two adjacent stands is taken into account only if the difference between the ages of the trees present in these two stands is greater than a certain threshold.

### 4.1. Forest management to protect some species

We present in this section a forest management problem proposed by Hof and Joyce (1993), included in (Hof and Bevers, 1998), and aimed at protecting certain species as best as possible. Consider a set of identical square forest parcels represented by a $m \times n$ matrix and two types of species $e_1$ et $e_2$. The habitat of the species $e_1$ is mainly in the cut parcels and the habitat of the species $e_2$ consists mainly of edges between cut parcels and uncut parcels. We consider the area outside the matrix as a cut zone. The expected population of the species $e_1$ in each cut (resp. uncut) parcel $s_{ij}$ is equal to $t_{ij}$ (resp. 0), and the expected population of the species $e_2$ is equal to $gL$ where $g$ is the population of the species $e_2$ expected for each km of edge and $L$, the total length of the edge, given the cuts made. For example, the northern goshawk likes habitat in the vicinity of edges where there are open areas where it can hunt small mammals living in the same habitat (Hof and Bevers, 1998). The $m \times n$ parcels are assumed to be initially in old growth state. The problem is to determine the parcels to cut and the parcels to be left in old growth to maximize the weighted sum of the populations of both species. Associate with each parcel $s_{ij}$ the Boolean variable $x_{ij}$ that equals 1 iff the parcel is uncut. Hof and Bevers (1998) formulate the problem by the linear program in binary variables $x$ and continuous variables $d$

$$(\text{P29}): \max_{d \in R_+^{mn},\ x \in \{0,1\}^{mn}} \left\{ \begin{array}{l} w_1 \sum_{(i,j)\in M\times N} t_{ij}(1 - x_{ij}) + w_2 g l \sum_{(i,j)\in M\times N} (4x_{ij} - d_{ij}) : \\ d_{ij} \geqslant \sum_{(k,l)\in A_{ij}} x_{kl} - |A_{ij}|(1 - x_{ij}) \ ((i,j) \in M \times N) \end{array} \right\}$$

where $M = \{1, \ldots, m\}$, $N = \{1, \ldots, n\}$, $w_1$ and $w_2$ are the weights associated with the two species, $l$ is the length of the side of each parcel, and $A_{ij}$ designates the set of couples $(k,l)$ such that the parcel $s_{kl}$ is adjacent to the parcel $s_{ij}$. Since the objective function of (P29) is to be maximized, the variable $d_{ij}$ takes the smallest possible value: if $x_{ij} = 0$ then, according to the constraint, $d_{ij}$ takes the value 0, and if $x_{ij} = 1$, $d_{ij}$ is equal to $\sum_{(k,l)\in A_{ij}} x_{kl}$, i.e. to the number of uncut parcels among the parcels adjacent to $s_{ij}$. We propose in (Billionnet, 2010d) to formulate the problem as the maximization of a 0–1 quadratic function without constraints

$$(\mathrm{P30}):\quad \max_{x\in\{0,1\}^{mn}} \left\{ \begin{array}{l} \sum_{(i,j)\in M\times N} (w_1 t_{ij}(1-x_{ij}) + w_2 gla_{ij}x_{ij}) \\ +w_2 gl \sum_{(i,j,k,l)\in S} (x_{ij}+x_{kl}-2x_{ij}x_{kl}) \end{array} \right\}$$

where $S = \{(i,j,k,l) \in (M \times N)^2\colon (k,l) = (i+1,j) \text{ or } (i,j+1)\}$, $a_{ij} = 1$ if $(i,j)$ belongs to $\{\{1,m\} \times \{2,\ldots,n-1\} \cup \{2,\ldots,m-1\} \times \{1,n\}\}$, $a_{ij} = 2$ if $(i,j)$ belongs to $\{(1,1),(1,n),(m,1),(m,n)\}$ and $a_{ij} = 0$ otherwise. Note that in the objective function of (P30), the quantity $x_{ij} + x_{kl} - 2x_{ij}x_{kl}$ equals 0 if $x_{ij} = x_{kl} = 1$ or $x_{ij} = x_{kl} = 0$, and equals 1 if only one of the two variables $x_{ij}$ or $x_{kl}$ equals 1. The advantage of this formulation lies in the fact that the constraints matrix associated with its classical linearization is totally unimodular (TU), which is not the case of the constraints matrix associated with the program (P29). The classical linearization of (P30) consists in replacing the products $x_{ij}x_{kl}$ by the variables $y_{ijkl}$ and in adding the linear constraints $1 - x_{ij} - x_{kl} + y_{ijkl} \geqslant 0$ and $y_{ijkl} \geqslant 0$ to force the equality $y_{ijkl} = x_{ij}x_{kl}$. We can prove that the constraints matrix of this linearization is TU, taking into account that the vertex-edge incidence matrix of a bipartite graph is TU (see e.g., Nemhauser and Wolsey, 1988). Large instances of the problem can therefore be easily solved by this formulation.

It is easy to introduce a temporal dimension in the model (P30). Consider a planning horizon consisting of a set $T$ of $q$ periods. The problem is to determine the parcels to cut in each period to maximize the weighted sum of the populations of both species at the end of the planning horizon. Consider the Boolean variable $x_{ijt}$ that equals 0 iff the parcel $s_{ij}$ is cut in period $t$ and the Boolean variable $z_{ij}$ that equals 1 iff the parcel $s_{ij}$ remains uncut at the end of the planning horizon. We assume in this problem that the available budget in period $t$ is equal to $B_t$ and that the cost associated with cutting the parcel $s_{ij}$ is equal to $c_{ij}$. The problem can be formulated by the quadratic program in 0–1 variables

$$(\mathrm{P31}):\quad \max_{\substack{x\in\{0,1\}^{mnq}\\ z\in\{0,1\}^{mn}}} \left\{ \begin{array}{l} \sum_{(i,j)\in M\times N} (w_1 t_{ij}(1-z_{ij}) \\ +w_2 gla_{ij}z_{ij}) + w_2 gl \sum_{(i,j,k,l)\in S} (z_{ij}+z_{kl}-2z_{ij}z_{kl}) : \\ z_{ij} = 1 - \sum_{t\in T}(1-x_{ijt})\ ((i,j)\in M\times N), \\ \sum_{(i,j)\in M\times N} c_{ij}(1-x_{ijt}) \leqslant B_t\ (t\in T) \end{array} \right\}$$

The objective function is similar to that of (P30). The fist constraint expresses the variables $z$ in function of the variables $x$, and at the same time, the fact that a parcel can be cut at most once during the planning horizon. The last constraint is the budget constraint. Although the quadratic program (P31) can be linearized in the same way as (P30), the obtained constraints matrix is no longer TU. Computational experiences we have carried out on the linearization of (P31) showed that hypothetical instances built on a matrix of $50 \times 50$ parcels with a planning horizon of 10 periods could be solved very quickly. Note that if carry-over of funds between the periods is allowed, the budget constraint can be replaced by $\sum_{(i,j)\in M\times N} c_{ij}(1-x_{ijt}) + F_t = \overline{B}_t\ (t\in T),\ \overline{B}_t = F_{t-1} + B_t\ (t = 2,\ldots,q)$, and $\overline{B}_1 = B_1$ where $F_t$ is a slack variable that represents the amount of unused funds in period $t$, whereas $\overline{B}_t$ is a variable that represents the budget that is actually available in period $t$ (Toth et al., 2011).

Now consider the same problem as (P30), i.e. without temporal dimension, but replacing the species $e_2$ by the species $e_3$ which expected population is a little more complicated to calculate. If the parcel $s_{ij}$ is cut, the population expected in this parcel is zero. Otherwise, it is equal to $u_{ij}PR_{ij}$ where $PR_{ij}$ is the connectivity of the parcel $s_{ij}$ with other uncut parcels. This connectivity is measured by the probability that the parcel $s_{ij}$ is connected to the rest of the uncut parcels. Expected population of $e_3$ in the uncut parcels $s_{ij}$ is equal to $u_{ij}$ if $PR_{ij} = 1$. We assume that we know the probabilities $p_{ijkl}$ $(0 \leqslant p_{ijkl} < 1)$ that the parcel $s_{ij}$ is connected to the parcel $s_{kl}$ and that all these probabilities are independent. They depend on a number of factors like for example the distance between the two parcels. We get therefore $PR_{ij} = 1 - \prod_{(k,l)\in M\times N}(1 - pr_{ijkl}x_{kl})$ with $pr_{ijij} = 0$. Moreover, it is also assumed that a minimum number $\alpha$ of parcels must be uncut to ensure the survival of the species $e_3$. The problem of determining what parcels must be cut to maximize the weighted sum of the expected species populations can be formulated by the mixed-integer mathematical program

$$(\mathrm{P32}):\ \max_{\substack{x\in\{0,1\}^{mn}\\ y\in\{0,1\}\\ s_3\in R_+}} \left\{ \begin{array}{l} w_1 \sum_{(i,j)\in M\times N} t_{ij}(1-x_{ij}) + w_3 s_3 : s_3 \leqslant \sum_{(i,j)\in M\times N} u_{ij}PR_{ij}x_{ij}, \\ s_3 \leqslant Cy,\ y \leqslant \frac{1}{\alpha} \sum_{(i,j)\in M\times N} x_{ij}, \\ PR_{ij} = 1 - \prod_{(k,l)\in M\times N} (1 - pr_{ijkl}x_{kl})\ ((i,j)\in M\times N) \end{array} \right\}$$

where $C$ is a constant large enough. When the probabilities $pr_{ijkl}$ have a particular form, Hof and Joyce (1993) propose a tight approximate solution of (P32) based on a linear approximation of the products $PR_{ij}x_{ij}$. In the event these probabilities are arbitrary, we show (Billionnet, 2011b) that by using the same technique as in Section 2.5, we obtain a relaxation of (P32), the mixed-integer linear program

$$(\mathrm{P33}):\ \max_{\substack{x\in\{0,1\}^{mn}\\ y\in\{0,1\}\\ s_3\in R_+\\ v_{ij}\in R_+^{mn}}} \left\{ \begin{array}{l} w_1 \sum_{(i,j)\in M\times N} t_{ij}(1-x_{ij}) + w_3 s_3 : s_3 \leqslant \sum_{(i,j)\in M\times N} u_{ij} v_{ij}, \\ s_3 \leqslant Cy,\ y \leqslant \frac{1}{\alpha} \sum_{(i,j)\in M\times N} x_{ij}, \\ \frac{1}{\theta_r}(1 - v_{ij}) + \log\theta_r - 1 \geqslant \sum_{(k,l)\in M\times N} \log(1 - pr_{ijkl})x_{kl} \\ ((i,j)\in M\times N, r\in H), \\ v_{ij} \leqslant x_{ij}\ ((i,j)\in M\times N) \end{array} \right\}$$

where the variable $v_{ij}$ represents the product $PR_{ij}x_{ij}$. The last constraint of (P33) forces the variable $v_{ij}$ to take the value 0 when $x_{ij}$ is 0. Since the objective function of (P33) is to be maximized, the two fist constraints force the variable $s_3$ to take, at the optimum, the value $\min\left\{\sum_{(i,j)\in M\times N} u_{ij}v_{ij}, Cy\right\}$. The third constraint forces the variable $y$ to take the value 0 if $\sum_{(i,j)\in M\times N} x_{ij} < \alpha$, i.e. if the number of uncut parcels is less than $\alpha$. The optimal value of (P33) provides a tight upper bound of the optimal value of the problem, and the values of $x_{ij}$ in an optimal solution of (P33) give an approximate solution of the problem. Some experiments conducted with CPLEX (2007) on instances with 900 parcels have shown that (P33) provided, in a few minutes of computation on a PC with an Intel Core Duo 2 gigahertz, approximate solutions at less than 0.5% from the optimum.

### 4.2. Commercial exploitation of forests taking into account environmental objectives

We present in this section a forest management problem inspired from (Bertomeu and Romero, 2001). The considered forest is composed of $n$ parcels, $s_1, s_2, \ldots, s_n$, the parcel $s_i$ involving trees of age $a(i)$. Horizon management is composed of $h$ periods of $t$ years. We seek to exploit the forest, i.e. to make one and only one cut of each parcel during the management period, in order to maximize the discounted benefit realized from the marketing of wood while taking into account an environmental objective. The environmental objective is to try to get at the end of the planning horizon, a forest where all ages are represented, where the number of parcels is the same for each age, and where the length of the edge between two parcels for which the age difference exceeds a certain threshold is maximum. To account for the first two aspects of the environmental objective, the same number of parcels are exploited at each period of the planning horizon. However, one can exploit a parcel if its trees have reached a certain maturity. Let $x_{ik}$ be the Boolean variable that is equal to 1 iff the parcel $s_i$ is cut during the period $k$ (in the middle of the period). The age of the trees of the parcel $s_i$ at the end of the planning horizon is equal to $F_i = \sum_{k=1}^{h}(0.5t + (h-k)t)x_{ik}$. The constraint on the maturity of the trees is expressed by $\sum_{i=1}^{n}\sum_{k\in K(i)} x_{ik} = 0$ where $K(i)$ is the set of periods where the age of the trees of the parcel $s_i$ is below the threshold. At the end of the planning horizon, the length of edge, $l_{ij}$, between two adjacent parcels $s_i$ and $s_j$ is to consider if $|F_i - F_j| \geqslant \delta_{\min}$. To express the total length of edge, we use the two Boolean variables $y_{ij}$ and $z_{ij}$, and seek to maximize the expression $\sum_{(i,j)\in S} l_{ij}(y_{ij} + z_{ij})$ under the constraints $F_i + M(1 - y_{ij}) \geqslant F_j + \delta_{\min}$ and $F_j + M(1 - z_{ij}) \geqslant F_i + \delta_{\min}$, where $M$ denotes a large enough constant and $S$, the pairs of indices of adjacent parcels. The economic objective is expressed by $\sum_{i=1}^{n}\sum_{k=1}^{h} NV_{ik}x_{ik}$ where $NV_{ik}$ is the discounted profit corresponding to the exploitation of the parcel $s_i$ at the period $k$. More precisely $NV_{ik}$ equals the area of the parcel $s_i$, in hectares, multiplied by the benefit obtained from one hectare of trees which age is that of the trees of the parcel $s_i$ at the period $k$, multiplied by a discount factor. If this parcel $s_i$ is cut at the period $k$, the age of the trees in this parcel is equal to $I_i + t(0.5 + k - 1)$ where $I_i$ is the initial age of the trees in the parcel. The problem can finally be formulated by the linear program in 0–1 variables

$$(\mathrm{P34}):\quad \max_{\substack{x\in\{0,1\}^{nh}\\ y\in\{0,1\}^{|S|}\\ z\in\{0,1\}^{|S|}}} \left\{ \begin{array}{l} w_1 \sum_{(i,j)\in S} l_{ij}(y_{ij}+z_{ij}) + w_2 \sum_{i=1}^{n}\sum_{k=1}^{h} NV_{ik}x_{ik} : \\ F_i = \sum_{k=1}^{h}(0.5t + (h-k)t)x_{ik}\ (i=1,\ldots,n), \\ \sum_{i=1}^{n} x_{ik} = \lfloor n/h \rfloor\ (k=1,\ldots,h), \\ \sum_{k=1}^{h} x_{ik} = 1\ (i=1,\ldots,n),\ \sum_{i=1}^{n}\sum_{k\in K(i)} x_{ij} = 0, \\ F_i + M(1-y_{ij}) \geqslant F_j + \delta_{\min},\ F_j + M(1-z_{ij}) \\ \geqslant F_i + \delta_{\min}\ ((i,j)\in S) \end{array} \right\}$$

where $S = \{(i,j) \in \{1,\ldots,n\}^2 : i < j,\ s_i$ and $s_j$ share a common edge$\}$, and $w_1$ and $w_2$ are weighting coefficients.

## 5. Fight against invasive species

Some species of plants and animals introduced are very invasive and it is now accepted that they have a significant impact on biodiversity (competition with other species, release of toxic substances, genetic disturbance, epidemics, etc.). Some optimization models were developed in the effective fight against invasive species. The reader can refer, for example, to the very good article written by Epanchin-Niell and Hastings (2010). The authors review studies that address economically optimal control of established invasive species. They describe three important components for determining optimal invasion management: invasion dynamics, costs of control efforts and a monetary measure of invasion damages. Another interesting reference in connection with this question is (De Lara and Doyen, 2008), where different biological models are reviewed.

We present below a simple problem proposed by Hof (1998) to illustrate the help that operational research can provide to treat this type of phenomenon. Consider a forest area represented by a matrix of $m \times n$ identical square parcels and a parasite present in some of these parcels. This parasite disperses by neighborhood and its population is steadily increasing. It is possible to completely eliminate the parasite from a parcel by performing some management actions in this parcel (poisoning, burning, use of attractants, sanitation cutting). The problem is to identify the parcels to be treated so as to minimize the spread of the parasite during a planning horizon consisting of $T$ periods. The number of parcels that may be treated at each period is limited to $K$. Denote by $v_{ijt}$ the population of the parasite in the parcel $s_{ij}$ at period $t$. At period $t+1$, this population grew, became equal to $v_{ijt}(1+r)$ where the rate $r$ is a given positive coefficient, and has spread to the adjacent parcels. Denote by $p_{klij}$ the proportion of the population of the parcel $s_{kl}$ which diffuses into the parcel $s_{ij}$ between $t$ and $t+1$. We have therefore, for any parcel $s_{ij}$ and for all $t > 1$, $v_{ijt} = \sum_{(k,l)\in M\times N} p_{klij}(1+r)v_{klt-1}$ where $v_{kl1}$ is the initial population of the parasite in the parcel $s_{kl}$. Let $x_{ijt}$ be a Boolean variable that is equal to 1 iff the parcel $s_{ij}$ is treated at the period $t$. A parcel is treated only once, and once treated at time $t$, the parasite is definitely eliminated, that is to say that it is eliminated for period $t$ and all subsequent periods. Hof and Bevers (2002) choose to minimize the total population of the parasite on the planning horizon, i.e. the expression $\sum_{(i,j)\in M\times N,\ t=1,\ldots,T} v_{ijt}$. They formulate the problem by the mixed-integer linear program

$$(\mathrm{P35}):\quad \min_{\substack{x\in\{0,1\}^{mnT}\\ v\in R_+^{mnT}}} \left\{ \begin{array}{l} \sum_{\substack{(i,j)\in M\times N\\ t=1,\ldots,T}} v_{ijt} : \sum_{(i,j)\in M\times N} x_{ijt} \leqslant K\ (t=2,\ldots,T), \\ v_{ijt} + b_{ijt}\sum_{k=2}^{t} x_{ijk} \geqslant \sum_{(k,l)\in M\times N} p_{klij}(1+r)v_{klt-1} \\ ((i,j)\in M\times N, t=2,\ldots,T) \end{array} \right\}$$

where $b_{ijt}$ is an upper bound of the population of the parasite in the parcel $s_{ij}$ at the period $t$, whatever the treatments carried out. Since the objective function of (P35) is to be minimized, the variable $v_{ijt}$ will take the smallest possible value at the optimum. If $\sum_{k=2}^{t} x_{ijk} > 0$, i.e. if the parcel $s_{ij}$ has been treated during one of the periods 2, ..., $t$, the second constraint is satisfied by setting $v_{ijt}$ to 0. Otherwise $v_{ijt}$ equals $\sum_{(k,l)\in M\times N} p_{klij}(1+r)v_{klt-1}$. We propose a formulation slightly different from (P35) and that has proved to be more effective in terms of computation time for all the instances we considered. This formulation uses the Boolean variable $y_{ijt}$ that is equal to 1 iff the parcel $s_{ij}$ is treated at period $t$ or before period $t$, and can be stated as

$$(\mathrm{P36}):\quad \min_{\substack{y\in\{0,1\}^{mnT}\\ v\in R_+^{mnT}}} \left\{ \begin{array}{l} \sum_{\substack{(i,j)\in M\times N\\ t=1,\ldots,T}} v_{ijt} : \sum_{(i,j)\in M\times N} (y_{ijt} - y_{ijt-1}) \leqslant K \\ (t=2,\ldots,T), \\ v_{ijt} + b_{ijt}y_{ijt} \geqslant \sum_{(k,l)\in M\times N} p_{klij}(1+r)v_{klt-1} \\ ((i,j)\in M\times N, t=2,\ldots,T), \\ y_{ij1} = 0\ ((i,j)\in M\times N), \\ y_{ijt} \geqslant y_{ijt-1}\ ((i,j)\in M\times N, t=2,\ldots,T) \end{array} \right\}.$$

Now, we present a bit more elaborate model proposed by Büyüktahtakin et al. (2011). As in the previous model, the problem consists in determining which parcels will be treated in each of the $T$ periods. The objective function is the total damage caused by the invasive species in the entire area during the entire time interval. The damage $D_{ijtk}$ caused to the threatened resource $k \in \{1,\ldots,R\}$ in parcel $s_{ij}$ in period $t$ is equal to $w_k R_{ijk} v_{ijt}$. The weight $w_k$ affected to resource $k$ depends on the management priority of this resource, $R_{ijk}$ represents the percentage of resource $k$ in the parcel $s_{ij}$ which is affected by the invasive species, and $v_{ijt}$ is the population density of the invasive species in parcel $s_{ij}$ in period $t$. So, the objective function to be minimized is $\sum_{(i,j)\in M\times N}\sum_{t=1}^{T}\sum_{k=1}^{R} D_{ijtk}$. Using the Boolean decision variables $x_{ijt}$ that are equal to 1 iff the parcel $s_{ij}$ is treated in period $t$, the change in population density in two consecutive periods can be formulated as

$$(\text{Eq.}): \quad \begin{cases} v_{ijt} = (v_{ijt-1} + g(v_{ijt-1}, \bar{v}_{ijt-1}, K_{ij}))(1 - x_{ijt}) \\ \text{if} \quad (v_{ijt-1} + g(v_{ijt-1}, \bar{v}_{ijt-1}, K_{ij})) \leqslant V_{ij} \\ v_{ijt} = (v_{ijt-1} + g(v_{ijt-1}, \bar{v}_{ijt-1}, K_{ij}))(1 - \lambda x_{ijt}) \quad \text{otherwise} \end{cases}$$

where $K_{ij}$ is the carrying capacity (environment's maximal load) in parcel $s_{ij}$, $\bar{v}_{ijt-1}$ is the vector of the population densities of the considered species in the eight parcels surrounding parcel $s_{ij}$ in period $t-1$, $V_{ij}$ is the value of the critical population density in parcel $s_{ij}$, and the function $g$ has a logistic growth form. In both cases, without treatment $(x_{ijt} = 0)$ $\quad v_{ijt} = v_{ijt-1} + g(v_{ijt-1}, \bar{v}_{ijt-1}, K_{ij})$. If treatment is applied ($x_{ijt}$ = 1) then in the first case the parcel population is eradicated and in the second case the population decreases according to the kill rate $\lambda$. The land management faces budget and labor constraints: $\sum_{(i,j)\in M\times N}\left(c_{ij}^1 + c_{ij}^2 v_{ijt}\right)x_{ijt} \leqslant B_t$ where $B_t$ is the available budget in period $t$ and $c_{ij}^1$ and $c_{ij}^2$ are two parameters depending of the parcel; $\sum_{(i,j)\in M\times N}\left(l_{ij}^1 + l_{ij}^2 v_{ijt}\right)x_{ijt} \leqslant L_t$ where $L_t$ is the maximum available labor in period $t$ and $l_{ij}^1$ and $l_{ij}^2$ are two parameters depending of the parcel. In summary, by introducing the Boolean variables $\delta_{ijt}$, the problem can be formulated as

$$(\text{P37}): \quad \min_{\substack{x\in\{0,1\}^{mnT} \\ v\in R_+^{mnT} \\ \delta\in\{0,1\}^{mnT}}} \begin{cases} \sum_{(i,j)\in M\times N} \sum_{t=1}^{T} \sum_{k=1}^{R} w_k R_{ijk} v_{ijt}: \\ \delta_{ijt}(v_{ijt-1} + g(v_{ijt-1}, \bar{v}_{ijt-1}, K_{ij})) \leqslant V_{ij} \quad ((i,j) \\ \quad \in M\times N, t = 2, \ldots, T), \\ v_{ijt} = (v_{ijt-1} + g(v_{ijt-1}, \bar{v}_{ijt-1}, K_{ij})) \\ \quad (1 - x_{ijt}(\delta_{ijt} + \lambda(1 - \delta_{ijt}))) \\ ((i,j) \in M\times N, t = 2, \ldots, T), \\ \sum_{(i,j)\in M\times N} \left(c_{ij}^1 + c_{ij}^2 v_{ijt}\right)x_{ijt} \leqslant B_t \quad (t = 1, \ldots, T), \\ \sum_{(i,j)\in M\times N} \left(l_{ij}^1 + l_{ij}^2 v_{ijt}\right)x_{ijt} \leqslant L_t \quad (t = 1, \ldots, T) \end{cases}$$

Because of the objective function to minimize, the first constraint requires the Boolean variable $\delta_{ijt}$ to take the value 0 if $(v_{ijt-1} + g(v_{ijt-1}, \bar{v}_{ijt-1}, K_{ij})) > V_{ij}$ and the value 1 otherwise. The second constraint expresses $v_{ijt}$ as stated by (Eq.). This model is applied by Büyüktahtakin et al. (2011) for the case of control of the invasive grass, *Pennisetum ciliare*, in the Arizona desert. According to these authors, the large size of the problem makes the application of direct optimization methods impossible. Instead, they analyse and compare the most frequently suggested strategies and their consequences.

Other publications deal with optimization models for controlling invasive species. For example, Hastings et al. (2006) develop simple approaches based on linear programming for determining the optimal removal strategies of different stage or age classes for control of invasive species that are still in a density-independent phase of growth. They illustrate the application of their method to the specific example of invasive *Spartina alterniflora* in Willapa Bay, WA. Wainger et al. (2008) present an extensive study applied to a weed, downy brome (*bromus tectorum*), which invades some mountainous areas of the western United States. Chalak et al. (2011) analyse a spatially explicit dynamic process of controlling invasive species in a stochastic setting. Epanchin-Niell and Wilen (2012) examines the spatial nature of optimal bioinvasion control. They develop a spatially explicit two-dimensional model (an integer program) of species spread that allows for differential control across space and time, and solve for optimal spatial-dynamic control strategies. Leung et al. (2002) approach the issue by stochastic dynamic programming.

## 6. Protection of genetic diversity

First recall some basic notions of genetics (see e.g., Griffiths et al., 2000; Petit et al., 2007). The gene is the unit of genetic inheritance. It is characterized by its precise location – or locus – on a chromosome. A well known example is the gene for eye color or blood type in humans. A gene also corresponds to a sequence of nucleotides, i.e. to unit fragments of DNA. An allele is a version of a gene. Eye color in humans depends on the alleles carried by the individual on the gene concerned. Genetic diversity is the existence of different alleles of a gene. In diploid cells ($n$ pairs of chromosomes), i.e. in the cells of humans and most animals, each individual has two alleles – corresponding to the two chromosomes of a pair – for each gene and an individual received, at random, one of the two alleles of each parent, corresponding to one of two chromosomes of the parents. These alleles may be identical or not. Take the example of blood type with, for example, the following situation: an individual (A/B) randomly inherited the A allele from one parent who was (A/A) and the B allele from the other parent who was (B/O). An individual whose gene has identical alleles on both chromosomes is called homozygous for that gene. Otherwise, we say that the individual is heterozygous for this gene. In the latter case, there are two possibilities for the phenotype (observable state of a character): one allele is dominant and the other is recessive or it is not the case. If more than two alleles correspond to the same gene, the gene is said to be polyallelic. Genetic diversity is a factor allowing species to adapt to changes in their environment especially in the context of climate and global changes. Many articles in the conservation biology literature dedicated to species conservation programs deal with optimization problems arising in genetic diversity. A first type of problem is to determine in a given population, one or more sub-populations with certain properties and maximizing genetic diversity (see e.g., Glover et al., 1995; Allen et al., 2010; Caballero et al., 2010) or (Rodrigues and Gaston, 2002a; Moulton et al., 2007; Bordewich et al., 2009) regarding an important aspect of genetic diversity, phylogenetic diversity. A second problem concerns the management of breeding populations, usually in captivity, in order to maintain the greatest possible genetic diversity (see e.g., Fernandez and Toro, 1999; Fernandez et al., 2001; Fernandez et al., 2008; Vales-Alonso et al., 2003; Xu et al., 2011).

We present below some of these problems. Both of these problems are to select among a set of individuals or species, a subset of maximum biological diversity. The diversity between individuals is measured either by differences between DNA sequences or by kinship coefficients. These problems are easily formulated as quadratic programs in 0–1 variables but large instances are difficult to solve despite the wealth of research on this topic conducted to date. The third problem concerns the maintenance of allelic diversity from one generation to the next in genetic conservation programs. The last problem is to determine the conservation policies to apply to a set of taxa in order to maximize the expected value of the phylogenetic diversity of the set. For these two problems, we show how to obtain very good approximate solutions, with guarantees with respect to the optimum, using a method similar to that used for the problems, despite the differences, discussed in Sections 2.5 and 4.1.

### 6.1. Selecting a species group of maximal genetic diversity: application to cranes

We present in this section a relatively simple problem proposed by Glover et al. (1995) to select some species to be protected –

from a variety of endangered species – in order to maximize biological diversity of the subset of protected species. Let $E = \{e_1, e_2, \ldots, e_p\}$ be a set of endangered species present in a certain region. The financial means available to protect these species being limited, it is impossible to protect all of them. It is assumed here that one has a total budget $B$ and we know the cost $c_i$ of protecting the species $e_i$. To select an optimal subset of species, it is necessary to measure the dissimilarity between two species and then to deduce the diversity of a group of species. The dissimilarity between two species $e_i$ and $e_j$ is measured here by a distance $d_{ij}$ satisfying the properties $d_{ij} \geqslant 0$, $d_{ij} = d_{ji}$ and $d_{ii} = 0$. This distance is a function of the differences between the two species and in relation to a number of features. The diversity of a subset $S$ of species is defined by the sum of the distances between each pair of species of this set: $D(S) = \sum_{i<j:(e_i,e_j)\in S^2} d_{ij}$. The problem is to determine the subset, $S$, of species to be protected that checks some constraints and maximizes $D(S)$. Associate with each species $e_i$, the Boolean variable $x_i$ that is equal to 1 iff the species $e_i$ is retained to be protected. To determine under a budget constraint a subset $S$ of maximum biodiversity can be formulated by the quadratic program in 0–1 variables

$$(\mathrm{P38}): \quad \max_{x\in\{0,1\}^n} \left\{ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d_{ij}x_ix_j : \sum_{i=1}^{n} c_ix_i \leqslant B \right\}$$

which is known as the quadratic 0–1 knapsack problem. Glover et al. (1995) apply this model to the protection of cranes. They consider 14 species of endangered cranes, mainly due to the destruction of wetlands. The hunt for the Nordic races, and pesticides in Africa, also contribute to their decline. There are several approaches to measure the dissimilarity between two species. Glover et al. (1995) use data on the differences between DNA sequences associated with the two species and obtained by a molecular biology technique, DNA hybridization. Regarding the 14 crane species considered, Glover et al. (1995) use genetic distances established by Krajewski (1989). There are several ways to solve (P38). The program can be directly submitted to a solver such as CPLEX (2007). One can also use one of several methods that have been proposed over the last thirty years for quadratic programming in 0–1 variables (see Section 2.3.1) or specific algorithms developed for the quadratic knapsack problem (see e.g., Kellerer et al., 2004). Glover et al. (1995) use a classical linearization of (P38). Lozano et al. (2011) propose an iterative greedy algorithm for this classical problem of combinatorial optimization that has proved to be effective compared to other metaheuristics.

### 6.2. Maintaining allelic diversity in conserved populations

We present a common problem encountered in genetic conservation programs recalled by Vales-Alonso et al. (2003). It consists in determining the optimal contribution (number of offspring) from each parent to minimize the loss of alleles in the generated population. These authors deal the problem with a heuristic method based on a parallel simulated annealing. We propose an approach that provides an approximate solution with a very good guarantee. Consider a population consisting of $P$ individuals, $I_1, I_2, \ldots, I_P$, belonging to a species and a set of $M$ polyallelic genes $g_1, g_2, \ldots, g_M$ to take into account. For each gene $g_i$, the alleles $a_{i_1}, a_{i_2}, \ldots, a_{i_{t(i)}}$ are possible. We consider the case of diploid cells (humans and most animals) where, for each gene, all individuals have two alleles, each allele corresponding to a chromosome of the pair considered. These alleles may be identical or not. Each individual randomly received one of the two alleles from each parent (see Fig. 5). One seeks to determine the optimal number of offspring of each individual to minimize the expected number of alleles missing in the next generation while maintaining a population of constant size. Let $x_j$ be the integer variable representing the number of offspring of the individual $I_j$. Denote by $E^1_{i_k}$ (resp. $E^2_{i_k}$) the set of indices of individuals in the initial population for which allele $a_{i_k}$ appears only once (resp. two times) for the gene $g_i$. Let us evaluate the probability $p_{i_k}$ of extinction of the allele $a_{i_k}$ in the next generation: $p_{i_k} = 0$ if the allele $a_{i_k}$ is not in the initial population or if the variable $x_j$ takes a positive value for at least one index $j$ of the set $E^2_{i_k}$. In the opposite case $p_{i_k} = \prod_{j\in E^1_{i_k}} 0.5^{x_j}$. By denoting $N_m$ (resp. $N_f$) the number of males (resp. females) of the initial population, minimizing the expected number of alleles missing in the next generation, while maintaining the same number of individuals, consists in minimizing the expression $\sum_{i=1}^{M}\sum_{k=1}^{t(i)} p_{i_k}$ under the constraints $\sum_{j=1}^{N_m} x_j = P$ and $\sum_{j=1}^{N_f} x_j = P$. Recall that the initial population size is equal to $N_m + N_f = P$. By setting $IK = \{(i,k): i = 1, \ldots, M,\ k = 1, \ldots, t(i)\}$ and using the nonnegative variables $z_{ik}$ $((i,k) \in IK)$ for expressing the probabilities of extinction of the alleles $a_{i_k}$, the problem can be formulated by the mathematical program

$$(\mathrm{P39}): \quad \min_{\substack{x\in N^P \\ z_{ik}\geqslant 0\ ((i,k)\in IK)}} \left\{ \sum_{(i,k)\in IK} z_{ik} : z_{ik} \geqslant \prod_{j\in E^1_{i_k}} 0.5^{x_j} - \sum_{j\in E^2_{i_k}} x_j\ ((i,k)\in IK), \quad \sum_{j=1}^{N_m} x_j = \sum_{j=1}^{N_f} x_j = P \right\}$$

or by the mathematical program

$$(\mathrm{P40}): \quad \min_{\substack{x\in N^P \\ z_{ik}\geqslant 0\ ((i,k)\in IK) \\ t_{ik}\geqslant 0\ ((i,k)\in IK)}} \left\{ \begin{array}{l} \displaystyle\sum_{(i,k)\in IK} z_{ik} : z_{ik} \geqslant t_{ik} - \sum_{j\in E^2_{i_k}} x_j\ ((i,k)\in IK), \\ \displaystyle t_{ik} \geqslant \prod_{j\in E^1_{i_k}} 0.5^{x_j}\ ((i,k)\in IK), \sum_{j=1}^{N_m} x_j = \sum_{j=1}^{N_f} x_j = P \end{array} \right\}.$$

Since the objective function of (P40) is to be minimized, the variables $z_{ik}$ and $t_{ik}$ take the smallest possible values at the optimum. If $\sum_{j\in E^2_{i_k}} x_j > 0$, then the quantity $t_{ik} - \sum_{j\in E^2_{i_k}} x_j$ is negative and $z_{ik}$ takes the value 0. Otherwise, $z_{ik}$ equals $t_{ik}$, i.e. $\prod_{j\in E^1_{i_k}} 0.5^{x_j}$. The last constraint of (P40) expresses that the total number of offspring of the males of the initial population, like that of the females, must be equal to the initial size of the population, i.e. to $P$. Taking the logarithm of both sides of the second constraint and applying the same technique as in Section 2.5, we obtain a guaranteed approximate solution of the problem by solving the mixed-integer linear program

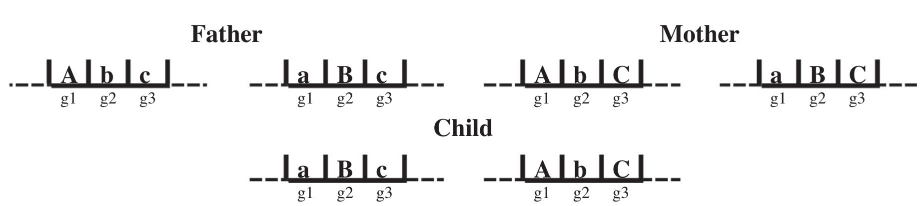


**Fig. 5.** Transmission of three bi-allelic genes g1 (A/ a), g2 (B/ b) and g3 (C/ c) on a pair of chromosomes.

$$(\text{P41}):\quad \min_{\substack{x\in N^P\\ z_{ik}\geqslant 0\ ((i,k)\in IK)\\ t_{ik}\geqslant 0\ ((i,k)\in IK)}} \left\{ \begin{array}{l} \sum_{(i,k)\in IK} z_{ik} : z_{ik} \geqslant t_{ik} - \sum_{j\in E^2_{i_k}} x_j\ ((i,k)\in IK),\\ \log\theta_r + \frac{t_{ik}-\theta_r}{\theta_r} \geqslant \sum_{j\in E^1_{i_k}} x_j \log 0.5\ ((i,k)\in IK, r\in H),\\ \sum_{j=1}^{N_m} x_j = \sum_{j=1}^{N_f} x_j = P \end{array} \right\}.$$

Experiments have shown that instances of (P41) with one hundred individuals, several hundred genes, and two alleles by gene could be solve quickly.

### *6.3. Minimizing the mean kinship of populations: application to the California condor*

Consider the $p$th pair of chromosomes of two individuals $I_i$ and $I_j$ belonging to the population $\{I_1, I_2, \ldots, I_p\}$ and randomly draw one chromosome of $I_i$ and one chromosome of $I_j$ in this pair. The probability that these two chromosomes are identical (that of a common ancestor) is called the kinship coefficient. For example, the kinship coefficient of two brothers is equal to 0.25 and the kinship coefficient of two children $I_i$ and $I_j$, from the couples of individuals (A, B) and (A, C), respectively, is equal to 0.125. Now consider a population of $P$ individuals with known kinship coefficients between each pair. One seeks to remove a group of $d$ individuals of this population to form a new population while minimizing the sum of the average kinship of the two subpopulations formed. Denoting by $k_{ij}$ the kinship coefficient between individuals $I_i$ and $I_j$, the mean kinship of a population with $P$ individuals is equal to $(1/P^2)\sum_{i=1}^{P}\sum_{j=1}^{P}k_{ij}$. This problem is presented by Allen et al. (2010) for the California condor, one of the largest birds in the world. This species was highly endangered in 1985 as only a few individuals remained in the wild. It was saved in extremis by captive breeding and reintroduced into the wild successfully. Using the Boolean variable $x_i$ that is equal to 1 iff the individual $I_i$ is taken from the original population, Allen et al. (2010) formulate the problem by the quadratic program in 0–1 variables

$$(\text{P42}):\quad \min_{x\in\{0,1\}^P} \left\{ \begin{array}{l} \frac{d^2+(P-d)^2}{d^2(P-d)^2} \sum_{i=1}^{P}\sum_{j=1}^{P} k_{ij}x_ix_j\\ +\frac{1}{(P-d)^2}\left(\sum_{i=1}^{P}\sum_{j=1}^{P}k_{ij} - 2\sum_{i=1}^{P}x_i\sum_{j=1}^{P}k_{ij}\right) : \sum_{i=1}^{P}x_i = d \end{array} \right\}.$$

They propose the classic linearization of (P42) that consists in replacing the products $x_ix_j$ by the variables $y_{ij}$ and adding the constraints $1 - x_i - x_j + y_{ij} \geqslant 0$ and $y_{ij} \geqslant 0$. They apply this method to a population of $P = 150$ California condors for which they know the pedigree but find the optimal solution only for the values of $d$ less than or equal to 10. For larger values of $d$, they propose to address the problem through a local improvement heuristic. We noticed that a linearization based on the ideas of Glover (1975) was more efficient from a computation time point of view but this approach does not allow to address the problem for values of $d$ greater than or equal to 11. It is also conceivable to try to solve this difficult problem by one of several methods that have been proposed over the last thirty years for quadratic 0–1 programming (see Section 2.3.1). The direct solution of the 0–1 quadratic program (P42) – whose continuous relaxation is convex – provides relatively quickly a good solution with acceptable guaranteed quality. Note that (P42) is a discrete version of the optimization of a quadratic function over the simplex. Bomze and de Klerk (2002) proposed a polynomial-time approximation scheme for this problem called the standard quadratic optimization problem.

**Table 1**
Presence or absence of six anatomical characters for five species.

| Species characters | Salmon | Xenopus | Crocodile | Duck | Man |
|---|---|---|---|---|---|
| Vertebral column | Yes | Yes | Yes | Yes | Yes |
| Choanes | No | Yes | Yes | Yes | Yes |
| Gizzard | No | No | Yes | Yes | No |
| Hairs | No | No | No | No | Yes |
| Feathers | No | No | No | Yes | No |
| # Heart cavities | 2 | 3 | 4 | 4 | 4 |

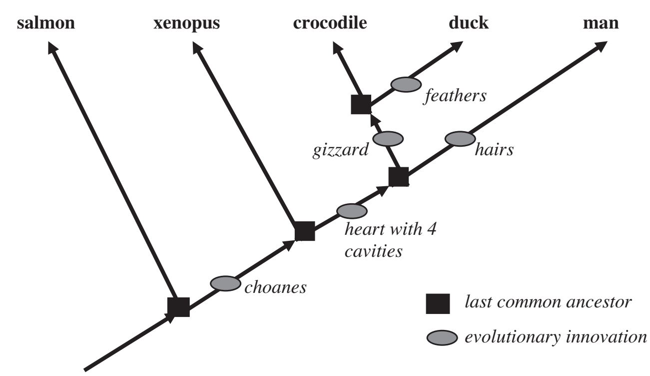


**Fig. 6.** Phylogenetic tree corresponding to Table 1.

### *6.4. Maximizing phylogenetic diversity*

We can consider a phylogenetic tree as a rooted tree $A = (X, Y, U)$ where $X = \{1, \ldots, p\}$ is the set of leaf vertices, $Y = \{p + 1, \ldots, m\}$ is the set of non-leaf vertices, and $U$ is the set of arcs. The leaf vertices correspond to the species – or taxa – considered. The out-degree of each non-leaf vertex is greater than or equal to 2, except possibly that of the root.

With each arc $u$ of the rooted tree is associated a nonnegative value $l_u$. A phylogenetic tree is based on the analysis of characters of the studied species, the branches representing the transformations of these characters. It is built on the assumption of parsimony, i.e. of a minimum number of character transformations (see e.g., Lin et al., 2007). It expresses the kinship between living organisms and indicates who is close to who (see Table 1 and Fig. 6 for an example from (Bergeron et al., 2002)). The phylogenetic diversity of a set of taxa $S \subseteq X$, that we denote by PD ($S$), is the sum of the values of the arcs of the minimum sub-tree connecting the vertices of $S$ and the root (Faith, 1992). This definition is a rigorous tool to make decisions concerning the preservation of biodiversity. It is adopted, for example, by the Zoological Society of London (http://www.zsl.org/) and the World Wide Fund for Nature (http://www.wwf.org/). In connection with this phylogenetic diversity, Weitzman (1998) defines the problem that he called Noah's Ark Problem (NAP): with each taxon $i$ is associated an initial survival probability $a_i$ that can be increased to $b_i$ at cost $c_i$. NAP consists in determining a subset of taxa $S$ which, under the budget constraint $\sum_{i\in S}c_i \leqslant B$, maximizes the expected value of PD ($S$), i.e. the quantity $\sum_{u\in U}l_u(1 - \prod_{i\in X_u\cap S}(1 - b_i)\prod_{i\in X_u - S}(1 - a_i))$ where $X_u$ is the set of leaves reachable from the arc $u$. Several special cases of NAP have been considered in the literature. For example, if $a_i = 0$, $b_i = 1$ and $c_i = 1$, NAP consists in finding a subset $S \subseteq X$ of $B$ leaves that maximizes PD ($S$), and can be solved polynomially by a greedy algorithm (Pardi and Goldman, 2005; Steel, 2005). If $c_i = 1$ and if $(1 - b_i)/(1 - a_i)$ is equal to a constant between 0 and 1 and independent of $i$, the problem can also be solved by a polynomial greedy algorithm (Hartmann and Steel, 2006). For other results on complexity and polynomial approximation of NAP, the reader can refer to (Hartmann and Steel, 2006; Hickey et al., 2008; Pardi and Goldman, 2007). Pardi (2009) proposes a generalization of the problem. One has a total budget of $B$ units and one can allocate to each taxon $i$ an amount of $k$ units, $k \in \{0, 1, \ldots, B\}$. There are thus, for each taxon $i$, $(B + 1)$ possible survival probabilities. The author proposes a solution of the problem by dynamic programming. His method is effective in practice on some instances but its polynomiality is not proved.

Here, we present a generalization of the problem by Billionnet (2013). For each taxon $i$, $n_i$ conservation policies are possible. The $k$th policy applied to $i$ has a cost $c_{ik}$ and leads to a survival probability equal to $p_{ik}$. We assume $c_{i1} < c_{i2} < \cdots < c_{in_i}$ and $0 \leqslant p_{i1} < p_{i2} < \cdots < p_{ini} < 1$. The problem is to determine the policy to apply to each taxon to maximize the expected value of the phylogenetic diversity of all taxa while meeting a budget constraint. Denoting $k(i)$ the policy applied to the taxon $i$, one obtains an expected phylogenetic diversity equal to $\sum_{u \in U} \lambda_u (1 - \prod_{i \in X_u} (1 - p_{i,k(i)}))$. It is easy to show that this problem is NP-hard from the 0–1 linear knapsack problem (Hartmann and Steel, 2006). Now, formulate the generalization of the Noah's Ark problem by a mathematical program. For each taxon $i$ and each policy $k$, define the Boolean variable $x_{ik}$ which is equal to 1 iff the $k$th policy is applied to the conservation of the taxon $i$. Let $IK = \{(i, k) : i \in X, k = 1, \ldots, n_i\}$. The problem can be formulated by the nonlinear program in 0–1 variables

$$(\text{P43}): \max_{x_{ik} \in \{0,1\} ((i,k) \in IK)} \left\{ \sum_{u \in U} \lambda_u \left( 1 - \prod_{i \in X_u} \left( 1 - \sum_{k=1}^{n_i} p_{ik} x_{ik} \right) \right) : \sum_{(i,k) \in IK} c_{ik} x_{ik} \leqslant B, \sum_{k=1}^{n_i} x_{ik} = 1 \ (i \in X) \right\}.$$

The first constraint of (P43) is a budgetary constraint and the second one expresses that a single conservation policy must be applied to each taxon. Solving (P43) is difficult due to the nonlinearity of the objective function. Let us see the formulation (P44) of the problem that consists in expressing the probability that the arc $u$ does not intervene (in the computation of the expected phylogenetic diversity) in function of the probabilities that the arcs directly following $u$ in the rooted tree does not intervene either. Associate with each arc $u$ the real variable $z_u$ which is equal to the probability that $u$ does not intervene. Let $A_u$ be the set of arcs that follow directly the arc $u$, and $U_X$ be the pendant arcs of $U$, i.e. the arcs which end is a leaf. For all non-pendant arcs, we have the equality $z_u = \prod_{j \in A_u} z_j$. For each pendant arc $u$ of the tree, denote by ext $(u)$ the corresponding leaf. The problem can then be formulated by

$$(\text{P44}): \max_{x_{ik} \in \{0,1\} ((i,k) \in IK)} \left\{ \begin{array}{l} \sum_{u \in U} \lambda_u (1 - z_u) : \sum_{(i,k) \in IK} c_{ik} x_{ik} \leqslant B, \\ \sum_{k=1}^{n_i} x_{ik} = 1 \ (i \in X), \\ z_u = \prod_{k \in A_u} z_k \ (u \in U - U_X), \\ z_u = \ 1 - \sum_{k=1}^{n_i} p_{ik} x_{ik} \ (u \in U_X, i = \text{ext}(u)) \end{array} \right\}$$

where for all pendant arcs $u$, $z_u$ is the extinction probability of the species ext $(u)$. Using the logarithmic function, we get the equivalent program, which the nonlinearity appears only in the constraint $y_u = \log z_u$,

$$(\text{P45}): \max_{x_{ik} \in \{0,1\} ((i,k) \in IK)} \left\{ \begin{array}{l} \sum_{u \in U} \lambda_u (1 - z_u) : \sum_{(i,k) \in IK} c_{ik} x_{ik} \leqslant B, \ \sum_{k=1}^{n_i} x_{ik} = 1 \\ \quad (i \in X), y_u = \log z_u \ (u \in U), \\ y_u = \sum_{k \in A_u} y_k \ (u \in U - U_X), \\ y_u = \sum_{k=1}^{n_i} \log(1 - p_{ik}) x_{ik} \ (u \in U_X, \ i = \text{ext}(u)) \end{array} \right\}.$$

Then using the same technique as in Section 2.5, we can obtain an approximate solution of (P45) with a good guarantee. Experiments have shown that a few minutes of computation were normally needed to treat phylogenetic trees with thousands of nodes and about 10 conservation policies for each taxon.

## 7. Conclusion

Biodiversity has become an important environmental issue after the Earth Summit in Rio de Janeiro in 1992 and the international community has pledged to reduce seriously its erosion. However, the funding allocated to the protection of biodiversity are extremely limited and it is therefore necessary to use them as effectively as possible. For this purpose, mathematical optimization is therefore a natural tool. Many articles in the literature of operations research or biological conservation deal with this subject. To illustrate help that mathematical optimization can bring to the protection of biodiversity, we have chosen to present, in some detail, a few problems appearing in important areas of biodiversity protection as the selection of nature reserves, the control of adverse effects caused by landscape fragmentation, the ecological exploitation of forests, the control of invasive species and the maintenance of genetic diversity. For lack of space, we do not review, far from it, all the literature regarding the optimization approach applied to the protection of biodiversity. Among the problems presented, some are well solved and others less well. For example, the methods currently proposed to select a connected reserve (Section 2.3.2), to identify a subset of parcels such that the associated species diversity is maximum (Section 2.6), to connect an optimally set of reserves by a network of corridors (Section 3.2), to select the investments to be carried in a network of biological corridors to enhance its permeability (Section 3.3), or to partition a population into two subpopulations of minimum average kinship (Section 6.3) do not allow instances of large size to be treated accurately. Research is still needed to progress in solving these difficult problems in order to deal with real instances satisfactorily. Also note that many theoretical studies have not led to real actions for conservation. This is what Knight et al. (2008) call the "research-implementation gap", a widespread phenomenon, far beyond the field of biodiversity protection. To reduce this gap, the authors recommend researchers to identify problems with the help of conservation practitioners, to ask questions in a broader context of conservation management and to take more account of the social dimension action of conservation. In conclusion we can say that mathematical optimization is an essential tool for efficient protection of biodiversity, that many studies have been devoted to this issue and that some of them have already led to practical decisions. However much remains to be done in defining and solving realistic models while trying to establish close relations between researchers and practitioners.

## Acknowledgement

I would like to thank the editor and referees for their many comments and suggestions that helped me to improve the first version of the article.

## References


Allen, S.D., Fathi, Y., Gross, K., Mace, M., 2010. An optimal and near-optimal strategy to selecting individuals for transfer in captive breeding programs. Biological Conservation 143, 2858–2863.

Alvarez-Miranda, E., Ljubic, I., Mutzel, P., 2013. The rooted maximum node-weight connected subgraph problem. In: The Tenth International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming, May 18–22, 2013, Yorktown Heights, NY, USA, LNCS.

Araújo, M.B., Williams, P.H., Fuller, R., 2002. Dynamics of extinction and the selection of nature reserves. Proceedings of the Royal Society of London B269, 1971–1980.

Arthur, J.L., Hachey, M., Sahr, K., Huso, M., Kiester, A.R., 1997. Finding all optimal solutions to the reserve site selection problem: formulation and computational analysis. Environmental and Ecological Statistics 4, 153–165.

Arthur, J.L., Haight, R.G., Montgomery, C.A., Polasky, S., 2002. Analysis of the threshold and expected coverage approaches to the probabilistic reserve site selection problem. Environmental Modeling and Assessment 7, 81–89.

Ball, I.R., Possingham, H.P., Watts, M.E., 2009. Marxan and relatives: software for spatial conservation prioritization. In: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), Spatial Conservation Prioritization. Oxford University Press.

Barrett, G.W., Peles, J.D., 1994. Optimizing habitat fragmentation: an agrolandscape perspective. Landscape and Urban Planning 28, 99–105.

Baskent, E.Z., Keles, S., 2005. Spatial forest planning: a review. Ecological Modelling 188, 145–173.

Bergeron, J., Beaujard, P., David, B., Dupuis, M., Hyon, A., Bednarek-Maitrepierre, I., Margerie, D., Margerie, M., Semelin, B., 2002. Sciences de la vie et de la terre. Hatier, Paris.

Bergseng, E., Vatn, A., 2009. Why protection of biodiversity creates conflict – some evidence from the Nordic countries. Journal of Forest Economics 15, 147–165.

Bertomeu, M., Romero, C., 2001. Managing forest biodiversity: a zero-one goal programming approach. Agricultural Systems 68, 197–213.

Bertsimas, D., Sim, M., 2004. The price of robustness. Operations Research 52, 35–53.

Bettinger, P., Boston, K., Kim, Y.-H., Zhu, J., 2007. Landscape-level optimization using tabu search and stand density-related forest management prescriptions. European Journal of Operational Research 176, 1265–1282.

Billionnet, A., 2010a. Integer programming for optimizing habitat network permeability. Management of Environmental Quality 21, 570–588.

Billionnet, A., 2010b. Optimal selection of forest patches using integer and fractional programming. Operational Research: An International Journal 10, 1–26.

Billionnet, A., 2010c. Quadratic optimization in 0-1 variables. In: Paschos, V.Th. (Ed.), Concepts of Combinatorial Optimization. ISTE Ltd and John Wiley & Sons Inc (chapter 7).

Billionnet, A., 2010d. Solving a cut problem in bipartite graphs by linear programming: application to a forest management problem. Applied Mathematical Modelling 34, 1042–1050.

Billionnet, A., 2011a. Solving the probabilistic reserve selection problem. Ecological Modelling 222, 546–554.

Billionnet, A., 2011b. Spatial optimization of wildlife populations with probabilistic habitat connections. Forest Science 57, 336–342.

Billionnet, A., 2012. Designing an optimal connected nature reserve. Applied Mathematical Modelling 36, 2213–2223.

Billionnet, A., 2013. Solution of the generalized Noah's ark problem. Systematic Biology 62, 147–156.

Billionnet, A., Elloumi, S., Plateau, M.-C., 2008. Quadratic 0–1 programming: tightening linear or quadratic convex reformulation by use of relaxations. RAIRO – Operations Research 42, 103–121.

Billionnet, A., Elloumi, S., Plateau, M.-C., 2009. Improving the performance of standard solvers for quadratic 0–1 programs by a tight convex reformulation: the QCR method. Discrete Applied Mathematics 157, 1185–1197.

Billionnet, A., Elloumi, S., Lambert, A., 2012. Extending the QCR method to general mixed-integer programs. Mathematical Programming 131, 381–401.

Bomze, I.M., 2012. Copositive optimization – recent developments and applications. European Journal of Operational research 216, 509–520.

Bomze, I.M., de Klerk, E., 2002. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. Journal of Global Optimization 24, 163–185.

Bond, M., 2003. Principles of Wildlife Corridor Design. Center for Biological Diversity <http://www.biologicaldiversity.org/publications/papers/wild-corridors.pdf>.

Bordewich, M., Semple, C., Spillner, A., 2009. Optimizing phylogenetic diversity across two trees. Applied Mathematics letters 22, 638–641.

Boros, E., Hammer, P.L., 2002. Pseudo-Boolean optimization. Discrete Applied Mathematics 123, 155–225.

Braunisch, V., Home, R., Pellet, J., Arlettaz, R., 2012. Conservation science relevant to action: a research agenda identified and prioritized by practitioners. Biological Conservation 153, 201–210.

Burer, S., 2009. On the copositive representation of binary and continuous nonconvex quadratic programs. Mathematical Programming 120, 479–495.

Büyüktahtakin, I.E., Feng, Z., Frisvold, G., Szidarovszky, F., Olsson, A., 2011. A dynamic model of controlling invasive species. Computers and Mathematics with Applications 62, 3326–3333.

Caballero, A., Rodriguez-Ramilo, S.T., Avila, V., Fernandez, J., 2010. Management of genetic diversity of subdivised populations in conservation programmes. Conservation Genetics 11, 409–419.

Cabeza, M., Moilanen, A., 2001. Design of reserve networks and the persistence of biodiversity. TRENDS in Ecology & Evolution 16, 242–248.

Camm, J.D., Polasky, S., Solow, A.R., Csuti, A.R., 1996. A note on optimal algorithms for reserve site selection. Biological Conservation 78, 353–355.

Camm, J.D., Norman, S.K., Polasky, S., Solow, A.R., 2002. Nature reserve site selection to maximize expected species covered. Operations Research 50, 946–955.

Caprara, A., Toth, P., Fischetti, M., 2000. Algorithms for the set covering problem. Annals of Operations Research 98, 353–371.

Cerdeira, J.O., Gaston, K.J., Pinto, L.S., 2005. Connectivity in priority area selection for conservation. Environmental Modeling and Assessment 10, 183–192.

Cerdeira, J.O., Pinto, L.S., Cabeza, M., Gaston, K.J., 2010. Species specific connectivity in reserve-network design using graphs. Biological Conservation 143, 408–415.

Chalak, M., Pannell, D., Polyakov, M., 2011. Economics of controlling invasive species: a stochastic optimization model for a spatial-dynamic process. Agricultural & Applied Economics Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24–26.

Church, R.L., Stoms, D.M., Davis, F.W., 1996. Reserve selection as a maximal covering location problem. Biological Conservation 76, 105–112.

Clemens, M.A., ReVelle, C.S., Williams, J.C., 1999. Reserve design for species preservation. European Journal of Operational Research 112, 273–283.

Clergeau, P., Désiré, G., 1999. Biodiversité, paysage et aménagement: du corridor à la zone de connexion biologique. Mappemonde 55, 19–23.

Cocks, K.D., Baird, I.A., 1989. Using mathematical programming to address the multiple reserve selection problem: an example from the Eyre Peninsula, South Australia. Biological Conservation 49, 113–130.

Collinge, S.K., 2000. Effects of grassland fragmentation on insect species loss, colonization, and movement patterns. Ecology 81, 2211–2226.

Conrad, J.M., Gomes, C.P., van Hoeve, W.-J., Sabharwal, A., Suter, J.F., 2012. Wildlife corridors as a connected subgraph problem. Journal of Environmental Economics and Management 63, 1–18.

Costello, C., Polasky, S., 2004. Dynamic reserve site selection. Resource and Energy Economics 26, 157–174.

Couvet, D., & Teyssèdre-Couvet A., 2010. Ecologie et biodiversité. Belin, Paris (chapter 6).

CPLEX, 2007. ILOG CPLEX 10.2.0 Reference Manual, ILOG CPLEX Division, Gentilly, France (now IBM ILOG CPLEX Optimizer).

Crist, M., 2004. Landscape Connectivity: An Essential Element of Land Management. Science & Policy Brief, The Wilderness Society, Washington, USA.

Crowe, K., Nelson, J., Boyland, M., 2003. Solving the area-restricted harvest-scheduling model using the branch and bound algorithm. Canadian Journal of Forest Research 33, 1804–1814.

Csuti, B., Polasky, S., Williams, P., Pressey, R., Camm, J., Kershaw, M., Kiester, R., Downs, B., Hamilton, R., Huso, M., Sahr, K., 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. Biological Conservation 80, 83–97.

De Lara, M., Doyen, L., 2008. Sustainable Management of Natural Resources, Mathematical Models and Methods. Springer-Verlag, Berlin Heidelberg (chapters 1 and 4).

Dilkina, B., and Gomes, C.P., 2010. Solving connected subgraph problems in wildlife conservation. In: 7th International Conference on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems. Bologna, Italy.

Dinkelbach, W., 1967. On nonlinear fractional programming. Management Science 13, 492–498.

Dissanayake, S.T.M., Önal, H., 2011. Amenity driven price effects and conservation reserve site selection: a dynamic linear integer programming approach. Ecological Economics 70, 2225–2235.

Drechsler, M., 2005. Probabilistic approaches to scheduling reserve selection. Biological Conservation 122, 253–262.

Epanchin-Niell, R.S., Hastings, A., 2010. Controlling established invaders: integrating economics and spread dynamics to determine optimal management. Ecology Letters 13, 528–541.

Epanchin-Niell, R.S., Wilen, J.E., 2012. Optimal spatial control of biological invasions. Journal of Environmental Economics and management 63, 260–270.

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

Fahrig, L., 2007. Non-optimal animal movement in human-altered landscapes. Functional Ecology 21, 1003–1015.

Faith, D.P., 1992. Conservation evaluation and phylogenetic diversity. Biological Conservation 61, 1–10.

Fernandez, J., Toro, M.A., 1999. The use of mathematical programming to control inbreeding in selection schemes. Journal of Animal Breeding and Genetics 116, 447–466.

Fernandez, J., Toro, M.A., Caballero, A., 2001. Practical implementation of optimal management strategies in conservation programmes: a mate selection method. Animal Biodiversity and Conservation 24, 17–24.

Fernandez, J., Toro, M.A., Caballero, A., 2008. Management of subdivised populations in conservation programs: development of a novel dynamic system. Genetics 179, 683–692.

Finke, J., Sonnenschein, M., 2008. Simulation and optimization of habitat network permeability. Information Technologies in Environmental Engineering 2008, 12–18.

Finke, J., Strein, M., Sonnenschein, M., 2008. A simulation framework for modeling anthropogenic disturbances in habitat networks. Ecological Informatics 3, 26–34.
Fiorella, K., Cameron, A., Sechrest, W., Winfree, R., Kremen, C., 2010. Methodological considerations in reserve system selection: a case study of Malagasy lemurs. Biological Conservation 143, 963–973.
Fischer, D.T., Church, R.L., 2003. Clustering and compactness in reserve site selection: an extension of the biodiversity management area selection model. Forest Science 49, 555–565.
Fischer, D.T., Church, R.L., 2005. The SITES reserve selection system: a critical review. Environmental Modeling and Assessment 10, 215–228.
Game, E.T., McDonald-Madden, E., Puotinen, M.L., Possingham, H.P., 2008. Should we protect the strong or the weak? Risk, resilience, and the selection of marine protected areas. Conservation Biology 22, 1619–1629.
Garfinkel, R.S., 1985. In: Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G., Shmoys, D.B., (Eds.), The Traveling Salesman Problem. John Wiley & Sons, New York (chapter 2).
Glover, F., 1975. Improved linear integer programming formulations of nonlinear integer problems. Management Science 22, 455–460.
Glover, F., Ching-Chung, K., Dhir, K.S., 1995. A discrete optimization model for preserving biological diversity. Applied Mathematical Modelling 19, 696–701.
Goycoolea, M., Murray, A.T., Barahona, F., Epstein, R., Weintraub, A., 2005. Harvest scheduling subject to maximum area restrictions: exploring exact approaches. Operations Research 53, 490–500.
Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C., Gelbart, W.M., 2000. An Introduction to Genetic Analysis. W.H. Freeman, New York, USA.
Grinstead, M., Snell, J.L., 1997. Introduction to Probability. American Mathematical Society.
Groeneveld, R.A., 2010. Species-specific spatial characteristics in reserve site selection. Ecological Economics 69, 2307–2314.
Haight, R.G., Travis, L.E., 2008. Reserve design to maximize species persistence. Environmental Modeling and Assessment 13, 243–253.
Haight, R.G., ReVelle, C.S., Snyder, S.A., 2000. An integer optimization approach to a probabilistic reserve site selection problem. Operations Research 48, 697–708.
Haight, R.G., Cypher, B., Kelly, P.A., Phillips, S., Possingham, H.P., Ralls, K., Starfield, A.M., White, P.J., Williams, D., 2002. Optimizing habitat protection using demographic models of population viability. Conservation Biology 16, 1386–1397.
Haight, R.G., Snyder, S.A., ReVelle, C.S., 2005. Metropolitan open-space protection with uncertain site availability. Conservation Biology 19, 327–337.
Halkos, G.E., Jones, N., 2012. Modeling the effect of social factors on improving biodiversity protection. Ecological Economics 78, 90–99.
Hamaide, B., Williams, J.C., ReVelle, C.S., 2009. Cost-efficient reserve site selection favoring persistence of threatened and endangered species. Geographical Analysis 41, 66–84.
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.
Harrison, S., Bruna, E., 1999. Habitat fragmentation and large-scale conservation: what do we know for sure? Ecography 22, 225–232.
Hartmann, K., Steel, M., 2006. Maximizing phylogenetic diversity in biodiversity conservation: Greedy solutions to the Noah's Ark Probem. Systematic Biology, 644–651.
Hastings, A., Hall, R.J., Taylor, C.M., 2006. A simple approach to optimal control of invasive species. Theoretical Population Biology 70, 431–435.
Hickey, G., Carmi, P., Maheshwari, A., Zeh, N., 2008. NAPX: a polynomial time approximation scheme for the Noah's ark problem. In: WABI '08 Proceedings of the 8th International Workshop on Algorithms in Bioinformatics. Springer-Verlag, pp. 76–86.
Hof, J., 1998. Optimizing spatial and dynamic population-based control strategies for invading forest pests. Natural Resource Modeling 11, 197–216.
Hof, J., Bevers, M., 1998. Spatial Optimization for Managed Ecosystems. Columbia University Press, New York.
Hof, J., Bevers, M., 2002. Spatial Optimization in Ecological Applications. Columbia University Press, New York.
Hof, J., Joyce, L.A., 1993. A mixed integer linear programming approach for spatially optimizing wildlife and timber in managed forest ecosystems. Forest Science 39, 816–834.
Holzgang, O., Pfister, H.P., Heynen, D., Blant, M., Righetti, A., Berthoud, G., Marchesi, P., Maddalena, T., Müri, H., Wendelspiess, M., Dändliker, G., Mollet, P., Bornhauser-Sieber, U., 2001. Les corridors faunistiques en Suisse. Cahier de l'environnement n° 326, Office fédéral de l'environnement, des forêts et du paysage (OFEFP), Société suisse de Biologie de la Faune (SSBF) & Station ornithologique suisse de Sempach, Bern, p. 120.
Hwang, F.K., Richards, D.S., Winter, P., 1992. The Steiner Tree Problem. Annals of Discrete Mathematics. North Holland, Amsterdam.
Jolibert, C., Wesselink, A., 2012. Research impacts and impact on research in biodiversity conservation: the influence of stakeholder engagement. Environmental Science and Policy 22, 100–111.
Juutinen, A., Mönkkönen, M., 2007. Alternative targets and economics efficiency of selecting protected areas for biodiversity conservation in boreal forest. Environmental and Resource Economics 37, 713–732.
Kellerer, H., Pferschy, U., Pisinger, D., 2004. Knapsack Problems. Springer.
Knight, A.T., Cowling, R.M., Rouget, M., Balmford, A., Lombard, A.T., Campbell, B.M., 2008. Knowing but not doing: selecting priority conservation areas and the research–implementation gap. Conservation Biology 22, 610–617.
Krajewski, C., 1989. Phylogenetic relationships among cranes (Gruiformes: Gruidae) based on DNA hybridization. The Auk 106, 603–618.
Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A., Lamberti, G., 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. Proceedings of the Royal Society of London, Biological Sciences 269, 2407–2413.
Lindenmayer, D.B., Franklin, J.F., Fischer, J., 2006. General management principles and a checklist of strategies to guide forest biodiversity conservation. Biological Conservation 131, 433–445.
Lin, Y.-M., Fang, S.-C., Thorne, J.L., 2007. A tabu search algorithm for maximum parsimony phylogeny inference. European Journal of Operational Research 176, 1908–1917.
Lozano, M., Molina, D., Garcia-Martinez, C., 2011. Iterated greedy for the maximum diversity problem. European Journal of Operational Research 214, 31–38.
Margules, C.R., Pressey, R.L., 2000. Systematic conservation planning. Nature 405, 243–253.
Margules, C.R., Stein, J.L., 1989. Patterns in the distribution of species and the selection of nature reserves: an example from Eucalyptus forests in southeastern New South Wales. Biological Conservation 50, 219–238.
Margules, C.R., Nicholls, A., Pressey, R., 1988. Selecting networks of reserves to maximise biological diversity. Biological Conservation 43, 63–76.
Margules, C.R., Nicholls, A.O., Usher, M.B., 1994. Apparent species turnover, probability of extinction and the selection of nature reserves: a case study of the ingleborough limestone pavements. Conservation Biology 8, 398–409.
Marianov, V., ReVelle, C., Snyder, S., 2008. Selecting compact habitat reserves for species with differential habitat size needs. Computers & Operations Research 35, 475–487.
Martell, D.L., Gunn, E.A., Weintraub, A., 1998. Forest management challenges for operational researchers. European Journal of Operational Research 104, 1–17.
McAlpine, C.A., Eyre, T.J., 2002. Testing landscape metrics as indicators of habitat loss and fragmentation in continuous eucalypt forests (Queensland, Australia). Landscape Ecology 17, 711–728.
McBride, M.F., Wilson, K.A., Burger, J., Fang, Y.-C., Lulow, M., Olson, D., O'Connell, M., Possingham, H.P., 2010. Mathematical problem definition for ecological restoration planning. Ecological Modelling 221, 2243–2250.
McCarthy, M.A., 2009. Spatial population viability analysis. In: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), Spatial Conservation Prioritization. Oxford University Press.
McDill, M.E., Rabin, S.A., Braze, J., 2002. Harvest scheduling with area-based adjacency constraints. Forest Science 48, 631–642.
McDonnell, M., Possingham, H., Ball, I., Cousins, E., 2002. Mathematical methods for spatially cohesive reserve design. Environmental Modeling and Assessment 7, 107–114.
McGarigal, K., Marks, B.J., 1994. Fragstats: Spatial pattern analysis program for quantifying landscape structure. Technical report, Forest Science Department, Oregon State University.
Memtsas, D.P., 2003. Multiobjective programming methods in the reserve selection problem. European Journal of Operational Research 150, 640–652.
Ménard, P., Clergeau, P., 2001. La notion de zone de connexion biologique, son application en aménagement du territoire. Mappemonde 64, 24–29.
Moilanen, A., Kujala, H., Leathwick, J.R., 2009. The zonation framework and software conservation prioritization. In: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), Spatial Conservation Prioritization. Oxford University Press.
Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), 2009b. Spatial Conservation Prioritization. Oxford University Press.
Moulton, V., Semple, C., Steel, M., 2007. Optimizing phylogenetic diversity under constraints. Journal of Theoretical Biology 246, 186–194.
Nalle, D.J., Arthur, J.L., Sessions, J., 2002. Designing compact and contiguous reserve networks with a hybrid heuristic algorithm. Forest Science 48, 59–68.
Nemhauser, G., Wolsey, L.A., 1988. Combinatorial Optimization. Wiley-Interscience, New York.
Ng, S.J., Dole, J.W., Sauvajot, R.M., Riley, S.P.D., Valone, T.J., 2004. Use of highway undercrossings by wildlife in Southern California. Biological Conservation 115, 499–507.
Nicholson, E., Possingham, H.P., 2007. Making conservation decisions under uncertainty for the persistence of multiple species. Ecological Applications 17, 251–265.
Öhman, K., Lämas, T., 2005. Reducing forest fragmentation in long-term forest planning by using the shape index. Forest Ecology and Management 212, 346–357.
Önal, H., Briers, R.A., 2002. Incorporating spatial criteria in optimum reserve network selection. Proceedings of the Royal Society of London B 269, 2437–2441.
Önal, H., Briers, R.A., 2005. Designing a conservation reserve network with minimal fragmentation: a linear integer programming approach. Environmental Modeling and Assessment 10, 193–202.
Önal, H., Briers, R.A., 2006. Optimal selection of a connected reserve network. Operations Research 54, 379–388.
Pardi, F., 2009. Algorithms on phylogenetic trees. Ph.D. Thesis, University of Cambridge, St/ Catharine's College, 2009.
Pardi, F., Goldman, N., 2005. Species choice for comparative genomics: being greedy works. PLoS Genetics 1, e71.
Pardi, F., Goldman, N., 2007. Resource-aware taxon selection for maximizing phylogenetic diversity. Systematic Biology 56, 431–444.
Petit, J.-M., Arico, S., Julien, R., 2007. Mini manuel de génétique. Dunod, Paris.

Polasky, S., Camm, J.D., Solow, A.R., Csuti, B., White, D., Ding, R., 2000. Choosing reserve networks with incomplete species information. Biological Conservation 94, 1–10.
Polasky, S., Camm, J.D., Garber-Yonts, B., 2001. Selecting biological reserves cost-effectively: an application to terrestrial vertebrate conservation in Oregon. Land Economics 77, 68–78.
Poljak, S., Rendl, F., Wolkowicz, H., 1995. A recipe for semidefinite relaxation for (0,1)-quadratic programming. Journal of Global Optimization 7, 51–73.
Possingham, H.P., Day, J., Goldfinch, M., Salzborn, F., 1993. The mathematics of designing a network of protected areas for conservation. In: Sutton, D., Cousins, F., Pierce, C. (Eds.), 12th Australian Operations Research Conference. University of Adelaide, pp. 536–545.
Poulin, M., Bélisle, M., Cabeza, M., 2006. Within-site habitat configuration in reserve design: a case study with a peatland bird. Biological conservation 128, 55–66.
Prendergast, J.R., Quinn, R.M., Lawton, J.H., 1999. The gaps between theory and practice in selecting nature reserves. Conservation Biology 13, 484–492.
Pressey, R.L., Cowling, R.M., 2001. Reserve selection algorithms and the real world. Conservation Biology 15, 275–277.
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., Cabeza, M., Watts, M.E., Cowling, R.M., Wilson, K.A., 2007. Conservation planning in a changing world. Trends in Ecology & Evolution 22, 583–592.
Pressey, R.L., Watts, M.E., Barrett, T.W., Ridges, M.J., 2009. The C-Plan conservation planning system: origins, applications, and possible futures. In: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), Spatial Conservation Prioritization. Oxford University Press.
Radzik, T., 1998. Fractional combinatorial optimization. In: Du, D.-Z., Pardalos, P.M. (Eds.), Handbook of Combinatorial Optimization, vol. 1. Springer.
Rayfield, B., James, P.M.A., Fall, A., Fortina, M.-J., 2008. Comparing static versus dynamic protected areas in the Québec boreal forest. Biological Conservation 141, 438–449.
ReVelle, C.S., Williams, J.C., Boland, J.J., 2002. Counterpart models in facility location science and reserve selection science. Environmental Modeling and Assessment 7, 71–80.
Rodrigues, A.S.L., Gaston, K.J., 2002a. Maximizing phylogenetic diversity in the selection of networks of conservation areas. Biological Conservation 105, 103–111.
Rodrigues, A.S.L., Gaston, K.J., 2002b. Optimisation in reserve selection procedures – why not? Biological Conservation 2002, 123–129.
Rosing, K.E., ReVelle, C.S., Williams, J.C., 2002. Maximizing species representation under limited resources: a new and efficient heuristic. Environmental Modeling and Assessment 7, 91–98.
Roupin, F., 2004. From linear to semidefinite programming: an algorithm to obtain semidefinite relaxations for bivalent quadratic problems. Journal of Combinatorial Optimization 8, 469–493.
Sabbadin, R., Spring, D., Rabier, C.-E., 2007. Dynamic reserve site selection under contagion risk of deforestation. Ecological Modelling 201, 75–81.
Sarkar, S., Pressey, R.L., Faith, D.P., Margules, C.R., Fuller, T., Stoms, D.M., Moffett, A., Wilson, K.A., Williams, K.J., Williams, P.H., Andelman, S., 2006. Biodiversity conservation planning tools: present status and challenges for the future. Annual Review of Environment and Resources 31, 123–159.
Schaible, S., 1995. Fractional programming. In: Horst, R., Pardalos, P.M. (Eds.), Handbook of Global Optimization. Kluwer Academic Publishers, pp. 495–608.
Schindler, S., Curado, N., Nikolov, S.C., Kret, E., Caracamo, B., Catsadorakis, G., Poirazidis, K., Wrbka, T., Kati, V., 2011. From research to implementation: nature conservation in the Eastern Rhodopes mountains (Greece and Bulgaria), European Green Belt. Journal for Nature Conservation 19, 193–201.
Simpson, E.H., 1949. Measurement of diversity. Nature 163, 688.
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. (Eds.), Integrating Landscape Ecology into Natural Resource Management. Cambridge University Press, Cambridge, pp. 208–240.
Smith, P.G.R., Theberge, J.B., 1986. A review of criteria for evaluating natural areas. Environmental Management 10, 715–734.
Steel, M., 2005. Phylogenetic diversity and the greedy algorithm. Systematic Biology 54, 527–529.
Strange, N., Thorsen, B.J., Bladt, J., 2006. Optimal reserve selection in a dynamic world. Biological Conservation 131, 33–41.
Strange, N., Thorsen, B.J., Bladt, J., Wilson, K.A., Rahbek, C., 2011. Conservation policies and planning under climate change. Biological Conservation 144, 2968–2977.
Suter, J.F, Conrad, J., Gomes, C.P., van Hoeve, W.J., Sabharwal, A., 2008. Optimal corridor design for grizzly bear in the U.S. Northern Rockies. In: Annual Meeting of the American Agricultural Economics Association, July 27–29, Orlando, Florida.
Toth, S.F., McDill, M.E., 2008. Promoting large, compact mature forest patches in harvest scheduling models. Environmental Modeling and Assessment 13, 1–15.
Toth, S.F., Haight, R.G., Snyder, S.A., George, S., Miller, J.R., Gregory, M.S., Skibbe, A.M., 2009. Reserve selection with minimum contiguous area restrictions: an application to open space protection planning in suburban Chicago. Biological Conservation 142, 1617–1627.
Toth, S.F., Haight, R.G., Rogers, L.W., 2011. Dynamic reserve selection: optimal land retention with land-price feedbacks. Operations Research 59, 1059–1078.
Trocmé, M., 2005. The Swiss defragmentation program-reconnecting wildlife corridors between the Alps and Jura: an overview. In: Irwin, C.L., Garrett, P., McDermott, K.P., (Eds.), Proceedings of the 2005 International Conference on Ecology and Transportation. Center for Transportation and the Environment, N.C. State University, Raleigh, USA, pp. 144–149.
Underhill, L., 1994. Optimal and suboptimal reserve selection algorithms. Biological Conservation 35, 85–87.
Vales-Alonso, J., Fernandez, J., Gonzalez-Castano, F.J., Caballero, A., 2003. A parallel optimization approach for controlling allele diversity in conservation schemes. Mathematical Biosciences 183, 161–173.
Vanderkam, R.P.D., Wiersma, Y.F., King, D.J., 2007. Heuristic algorithms vs. linear programs for designing efficient conservation reserve networks: evaluation of solution optimality and processing time. Biological conservation 137, 349–358.
Vemema, H.D., Calamai, P.H., Fieguth, P., 2005. Forest structure optimization using evolutionary programming and landscape ecology metrics. European Journal of Operational Research 164, 423–439.
Wainger, L.A., King, D.M., Mack, R.N., Price E.W., Maslin, T., 2008. Prioritizing invasive species management by optimizing production of ecosystem service benefits. Final Report to USDA ERS PREISM Program, USA.
Wallace, K.J., 2012. Values: drivers for planning biodiversity management. Environmental Science & Policy 17, 1–11.
Weintraub, A., Murray, A.T., 2006. Review of combinatorial problems induced by spatial forest harvesting planning. Discrete Applied Mathematics 154, 867–879.
Weintraub, A., Romero, C., Bjørndal, T., Epstein, R. (Eds.), 2007. Handbook of Operations Research in Natural Resources. Springer.
Weitzman, M.L., 1998. The Noah's ark problem. Econometrica 66, 1279–1298.
Williams, J.C., 1998. Delineated protected wildlife corridors with multi-objective programming. Environmental Modeling and Assessment 3, 77–86.
Williams, J.C., 2002. A zero-one programming model for contiguous land acquisition. Geographical Analysis 34, 330–349.
Williams, J.C., 2003. Convex land acquisition with zero-one programming. Environment and Planning B: Planning and Design 30, 255–270.
Williams, J.C., 2008. Optimal reserve site selection with distance requirements. Computers & Operations Research 35, 448–498.
Williams, P.H., Araújo, M.B., 2000. Using probability of persistence to identify important areas for biodiversity conservation. Proceedings of the Royal Society of London B 267, 1959–1966.
Williams, P.H., Araújo, M.B., 2002. Apples, oranges, and probabilities: integrating multiple factors into biodiversity conservation with consistency. Environmental Modeling and Assessment 7, 139–151.
Williams, J.C., ReVelle, C.S., 1998. Reserve assemblage of critical areas: a zero-one programming approach. European Journal of Operational Research 104, 497–509.
Williams, J.C., Snyder, S.A., 2005. Restoring habitat corridors in fragmented landscape using optimization and percolation models. Environmental Modeling and Assessment 10, 239–250.
Williams, J.C., ReVelle, C.S., Levin, S.A., 2005. Spatial attributes and reserve design models: a review. Environmental Modeling and Assessment 10, 163–181.
Xu, P., Wang, L., Beavis, W.D., 2011. An optimization approach to gene stacking. European Journal of Operational Research 214, 168–178.