Ecological Modelling 478 (2023) 110281

Contents lists available at ScienceDirect

Ecological Modelling

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

A sequential approach to reserve design with compactness and contiguity considerations

Lakmali Weerasena [a,*], Douglas Shier [b], David Tonkyn [c,1], Mark McFeaters [a], Christopher Collins [a]

[a] Department of Mathematics, University of Tennessee at Chattanooga, 615 McCallie Avenue, Chattanooga, TN 37403-2598, United States of America
[b] School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC 29634, United States of America
[c] Department of Biological Sciences, Clemson University, Clemson, SC 29634, United States of America

ARTICLE INFO

Dataset link: https://github.com/lweeras/MODERN.git

Keywords:
Compactness
Conservation areas
Corridors
Graph density
Optimization models
Reserve site selection

ABSTRACT

Biological conservation depends increasingly on the establishment of protected areas that include as many species as possible, and are extensive, compact, and connected. Various optimization models have been developed to address one or more of these goals; this article develops a set of models that can be used sequentially or separately to promote all four. Assuming that species are mapped across a rectangular grid, we first identify core areas that are extensive and compact by maximizing the density of the graph associated with the reserve. This minimizes the number of boundary edges (suitably normalized) and thus reduces opportunities for organisms to leave the core areas, and external threats to enter. We then identify contiguous corridors between these compact areas based on costs and species conservation goals, allowing the judicious replacement of core elements where possible. This suite of optimization models can assist in designing an efficient and effective system of compact protected areas along with connecting corridors.

1. Introduction

In modern reserve design, conservationists seek to designate areas for protection that, all else being equal, (i) protect as many species as possible, and are (ii) extensive, (iii) compact, and (iv) connected to one another. The arguments are now well known. Reserves that protect more species will be more effective biologically and more easily justified economically and politically. Extensive reserves can support large populations that are less vulnerable to intrinsic factors such as demographic stochasticity (chance variation in births, deaths and sex composition (Lande, 1988)) and the loss of genetic variation (through inbreeding and chance inheritance (Frankham, 2005; O'Grady et al., 2006)) as well as from extrinsic factors such as catastrophes (fires, floods, hurricanes) and climate change (Beier, 2012; Bengtsson et al., 2003). Compact reserves will be less affected by threats that enter through reserve boundaries such as altered microclimates, illegal grazing and wood cutting, disease, poaching, and predation (Murcia, 1995). They may also be easier and cheaper to manage and protect (Bruner et al., 2004). Finally, connected reserves will allow for the movement of individuals between sites, buffering against downward fluctuations in numbers or genetic diversity, sustaining populations and evolutionary processes (Diamond, 1975; McCullough, 1996; Wilson and Willis, 1975), and allowing species to track changing climates (Beier, 2012; Dinerstein et al., 2020; Hannah, 2011) at far less effort and cost to alternatives such as assisted migration (Krosby et al., 2010). These ideas informed and motivated conservation planning from local to global scales (Kingsland, 2002; Olson and Dinerstein, 2002), and the number, cumulative area and connectivity of reserves quickly grew (Jenkins and Joppa, 2009; UNEP-WCMC, accessed 2021). While the biological imperative to create large and connected reserves is now clear, there are significant economic, political, and social costs to setting lands aside from human use. Conservationists need to designate reserves that are as effective and efficient as possible, but initially lacked two essential tools for doing so. First, they had limited knowledge of many species' distributions, especially in remote and species-rich areas such as the tropics, so there was little confidence in the effectiveness of any proposed reserve to protect species (Williams et al., 2002). Second, even when species' distributions were known, conservationists had limited tools to find the most efficient set of sites to preserve these species. There has been continued improvement in both areas, and the science of reserve design now draws on increasingly complete data and sophisticated analytical tools.

* Corresponding author.
E-mail addresses: Lakmali-Weerasena@utc.edu (L. Weerasena), shierd@clemson.edu (D. Shier), tdavid@clemson.edu (D. Tonkyn), gps847@mocs.utc.edu (M. McFeaters), pdq278@mocs.utc.edu (C. Collins).

[1] Regrettably Prof. Tonkyn passed away suddenly on January 2, 2022. This article is dedicated to his contributions to ecology, conservation and population biology.

https://doi.org/10.1016/j.ecolmodel.2023.110281
Received 20 April 2022; Received in revised form 8 December 2022; Accepted 9 January 2023
Available online 18 January 2023
0304-3800/Published by Elsevier B.V.

Early attempts to achieve the first goal, protecting the most species, focused on identifying and protecting sites that were hotspots of various kinds, for example, that were high in endemic species, in taxonomically unique species, in endemics and total species, or in endemics and threats (reviewed in Brooks et al., 2006). These approaches did not use all the available data on species' distributions, and several researchers proposed instead using a greedy algorithm: selecting the most diverse site and then adding, in stepwise fashion, subsequent sites that contain the greatest number of species not yet represented, providing "complementarity", a key early principle in reserve design (reviewed in Margules and Pressey, 2000; Pressey et al., 1993). This approach is easily implemented but does not guarantee an optimum in coverage or efficiency, since the value of any particular site depends ultimately on which other sites are also selected.

Later, methods from mathematical optimization were introduced to regional conservation planning, for example in the US Gap Analysis Program (Scott et al., 1993), to identify sites that would protect all species of interest. An early application (Kiester et al., 1996) mapped all 357 species of vertebrates in Idaho onto a grid of 389 hexagons, each 640 km^2 in area, and asked for each size of a reserve system, of $1, 2, 3, \ldots$ hexagons, which specific hexagons would protect the most species. Global optima were found for sets of up to five hexagons by exhaustive search. Larger sets were identified using an IBM Optimizing Subroutine Library for a "maximal location covering problem" (Church et al., 1996). This approach was extended globally, though with necessarily lower resolution (Rodrigues et al., 2004a,b).

Numerous studies have now treated the reserve design problem as a set- or maximal-set covering problem, respectively seeking the most efficient solution to protect all species or the most species with a reserve system of a given size or budget (Cabeza and Moilanen, 2001). These approaches easily accommodate sites that are already protected or that are weighted to reflect their differing costs or benefits to species conservation. While these approaches do identify sites that are efficient in space and cost, the sites chosen tend to be small and remote from one another and not ideal for the long term. For example, the four hexagons in Idaho with greatest combined diversity of vertebrates include one each from the northern rainforest, central boreal forest, southwest desert, and southeast Snake River basin. To be fair, with larger total reserve sizes (more hexagons), the set-covering analyses find more equivalent solutions, providing "flexibility" in decision making, another early principle in reserve design (Pressey et al., 1993). However, there clearly was a need to formally address these additional goals.

A number of models have now been devised to design reserve systems that are efficient in protecting species and that also satisfy the additional goals of being connected, compact, or both. Connectivity can be measured in a variety of ways (Butler et al., 2022; Keeley et al., 2021; Pascual-Hortal and Saura, 2006); here we focus on structural or habitat connectivity of an ecoscape. Early work was conducted on establishing connected reserves (Önal and Briers, 2002, 2005, 2006; Williams et al., 2002), followed by a more recent treatment (Billionnet, 2012). A variety of optimization models have also been developed for connected reserve design that incorporates aspects of species home range (Gupta et al., 2019) and species relocation (Dissanayake et al., 2012).

In order to model compactness, one line of research focused on minimizing the sum of pairwise distances (or maximum distance) between sites in a given cluster (Nalle et al., 2002a; Önal and Briers, 2002). Alternative measures are the sum of distances (or maximum distance) to an optimally chosen central site within the cluster (Billionnet, 2016; Önal et al., 2016; Wang and Önal, 2016; Williams et al., 2005). Another approach to promote compactness minimizes the perimeter (or scaled perimeter) of the cluster (Daigle et al., 2020; Fischer and Church, 2003; McDonnell et al., 2002; Öhman and Lämås, 2005; Önal and Briers, 2003; Possingham et al., 2000; Wright et al., 1983). A thorough description of alternative compactness measures and their mathematical formulation can be found in Billionnet (2021). Some notable efforts have sought to generate reserves that are both connected and compact (Billionnet, 2016; Cova and Church, 2000; Önal et al., 2016; Wang and Önal, 2016; Williams and ReVelle, 1996; Williams, 1998).

Biobjective approaches have been studied in order to combine two measures of compactness (Nalle et al., 2002a,b). Our previous work (Weerasena et al., 2014) used a hierarchical optimization approach to create compact clusters of sites. The well-known Marxan software (Ball et al., 2009) employs a biobjective approach that tries to simultaneously maximize compactness and minimize total site acquisition cost. Though originally implemented using simulated annealing (Ball et al., 2009), solution by integer linear programming was later added (Beyer et al., 2016).

Connectivity between protected areas is now considered integral to any system of protected areas (Alagador et al., 2012), especially in light of continued climate change, and is mandated in many international conservation targets (Ward et al., 2020). Connectivity between reserves can be provided in various ways. For mobile species, stepping stones along migration pathways may be sufficient, as for example with migratory waterfowl and the US National Wildlife Refuge system (Fischman, 2005). For species that are less mobile or that cannot move through human-dominated landscapes, corridors may need to be continuous. The corridors themselves may serve as conduits for movement, for what Beier and Loe (1992) called "passage species" that simply transit the corridor. Or they may need to be habitats for "corridor dwellers" that live and reproduce in them (Hess and Fischer, 2001). In the first case, corridors can be tacked on afterwards to a system of large, compact nature reserves; in the second, the corridors are better viewed as part of a single, connected reserve. Various methods have been developed to identify efficient or effective corridors between existing reserves, using individual-based models, circuit theory, graph theory, or least-cost measures, among others (reviewed in Hilty et al., 2020, Hall et al., 2021). Notable contributions for addition of corridors using optimization techniques include (Conrad et al., 2012), Gomes (2011), Dilkina et al. (2017), Dissanayake et al. (2012), Gupta et al. (2019), Hamaide et al. (2022), Sessions (1992), Suter et al. (2008), and Williams (1998).

A number of studies have used mathematical optimization to simultaneously identify compact reserves and connecting corridors (Billionnet, 2013; Fuller and Sarkar, 2006). Earlier versions of Marxan could find corridors with the help of manual inputs or specialized scripts, though a new program Marxan Connect (Daigle et al., 2020) now streamlines that process.

The present work continues along such lines by formulating a particular measure of compactness (network density), creating compact clusters of sites that protect species within a given cost budget, and then judiciously adjusting these clusters to enable the provision of (efficient and species-rich) corridors that join them. This yields a system of connected reserves that jointly satisfies the four goals outlined at the beginning of this section. It should be emphasized that we illustrate here a sequential process that first identifies compact clusters of core areas, then constructs corridors that link these clusters, and finally allows adjustment of the original clusters to achieve a feasible solution that respects a budgetary constraint. These three steps, which involve the use of optimization models for their solution, can also be applied individually to either create compact clusters, to connect existing core areas, or to adjust a system of core areas plus connecting corridors. We believe that the flexible approach developed here is especially applicable when natural core areas/reserves are geographically distant from one another so that the addition of (noncompact) single-parcel width corridors makes ecological and economic sense.

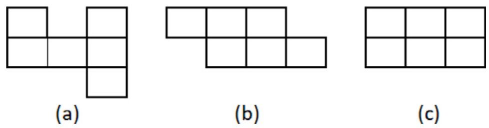

Fig. 1. Potential reserves consisting of six sites within a uniform rectangular grid.

2. Optimization models

2.1. Measuring compactness

For simplicity, we assume that potential reserve sites are defined on a rectangular grid of uniform sites. Fig. 1 shows illustrative examples in which six sites (indicated by the outlined squares) might be selected to define a potential reserve within the entire grid.

It is useful to define the graph $G = (N, E)$ associated with a potential reserve, consisting of the selected sites (nodes) N and in which the graph edges E indicate adjacency of sites in N. Specifically, two sites $a, b \in N$ are adjacent, with $[a, b] \in E$, if they share a common side. Let $n = |N|$ and $m = |E|$ denote, respectively, the number of nodes and the number of edges in G. For example, the graph associated with the reserve shown in Fig. 1(a) has $n = 6$ nodes and $m = 5$ edges. In general, a potential reserve consists of m pairs of adjacent sites and has p exterior defining boundary lines. Since grid cells are uniform, we assume for simplicity that the distance between adjacent sites is 1 and that the distance between nonadjacent sites is given by the minimum number of edges joining them in G.

As discussed in Section 1, compactness can been defined in a number of ways. In our study we concentrate one specific measure of compactness, namely the *density* of the graph $G = (N, E)$, which is defined as $\delta = m/n$. The reason for using the term density here arises from the general concept of graph density, defined as the ratio of the number of edges in a graph to the maximum number of possible edges. Since a reserve graph G is embedded in the plane, the maximum number of edges in G is at most $3n - 6$ (Chartrand and Zhang, 2005). Consequently, the ratio m/n does indeed measure the density of a reserve graph. It is useful to note that for any regular grid of sites (e.g., triangular, rectangular, hexagonal), we must have $m \le 3n - 6$ so that the density $\delta = m/n$ will be bounded above by 3.

We argue that maximizing the density of the associated graph is a reasonable goal for achieving compactness. Indeed maximizing the graph density δ is closely related to minimizing the perimeter p, the sum of lengths of all boundary edges of the reserve. Namely, for a graph $G = (N, E)$ embedded in a uniform rectangular grid (all unit length cell sides), there are four lines surrounding each cell giving a total count of $4n$ lines defined by the figure. This is also equal to the number of boundary edges p plus twice the number of graph edges m, giving $4n = p + 2m$, and so $4 = \frac{p}{n} + 2\delta$. Therefore maximizing the density δ is equivalent to minimizing a *normalized* form of the perimeter p. Biologically, maximizing the density of a reserve graph then minimizes the proportion of boundaries across which organisms can leave a reserve, or external threats can enter. By such normalization, the density measure allows us to compare reserve configurations containing different numbers of sites n. Previous investigations have measured compactness using a different normalization, by comparing the perimeter to the circumference of a circle with the same area (Öhman and Lämås, 2005; Possingham et al., 2000). Billionnet (2021) considers compactness to be measured by the ratio of the perimeter to the area of the reserve; since we assume uniform grid cells, minimizing this measure is then equivalent to maximizing the graph density.

As an illustration, the densities of the reserves depicted in Fig. 1 are (respectively) $\frac{5}{6}$, 1, and $\frac{7}{6}$. So the reserve shown in Fig. 1(c) would be classified as the most compact configuration of the three, which is intuitively reasonable. As a further example, suppose that the selected reserve sites form a rectangle with R rows and C columns. Then the density is calculated to be $\delta = 2 - [(R + C)/RC]$. This quantity is maximized by making R and C as equal as possible, conforming to the idea that a square-shaped reserve achieves the most compact rectangular shape. A square-shaped reserve with R rows and R columns then has density $\delta = 2 - \frac{2}{R}$, approaching the limiting value $\delta = 2$. As discussed in Section 2.2, an upper bound of 2 on reserve graph density holds for rectangular grids and is important in developing the solution technique discussed there.

2.2. Models for creating compact reserves

In this section, we develop an optimization framework for creating compact reserves based on maximizing the graph density $\delta = m/n$. We suppose that the system of possible reserve sites is defined by a rectangular grid $\mathcal{G}$ of uniformly-sized cells, each represented by an ordered pair $(i, j) \in \mathcal{G}$. If two cells (i, j) and (k, l) share a common side, we say they are adjacent. The set $N(i, j)$ indicates the neighborhood of (i, j), namely those grid cells adjacent to cell (i, j).

The optimization models developed here and subsequently identify a compact set of sites constituting a possible reserve, taking into account the desired representation of species and a budgetary constraint. These models require the following inputs:

S = number of conservation species to be protected
A_s = set of sites (i, j) inhabited by species of type s, where $s = 1, 2, \ldots, S$
B = total budget available for the reserve system
b_{ij} = budgetary cost of purchasing, conserving or maintaining site (i, j).

We define decision variables x_{ij} to indicate which grid sites (i, j) are selected for inclusion in the reserve system:

$$x_{ij} = \begin{cases} 1 & \text{if site } (i, j) \in \mathcal{G} \text{ is selected} \\ 0 & \text{otherwise.} \end{cases}$$

We use this notation to define the graph $G = (N, E)$ associated with a stipulated reserve. Namely, $N = \{(i, j) : x_{ij} = 1\}$ and $E = \{[(i, j), (k, l)] : x_{ij} = x_{kl} = 1, (k, l) \in N(i, j)\}$. Then the graph density can be expressed as the ratio

$$\delta = \frac{\frac{1}{2} \sum_{(i,j)\in\mathcal{G}} \sum_{(k,l)\in N(i,j)} x_{ij} x_{kl}}{\sum_{(i,j)\in\mathcal{G}} x_{ij}} \tag{1}$$

The denominator of (1) counts the total number of sites in the reserve system G. The numerator of (1) counts the number of adjacent pairs of sites in G and is divided by 2 to avoid double counting the number of actual edges in G. To obtain a feasible configuration of sites with maximum density, we can solve the following 0–1 nonlinear optimization model (P_1):

$$(P_1) \quad \text{maximize} \quad \frac{\frac{1}{2} \sum_{(i,j)\in\mathcal{G}} \sum_{(k,l)\in N(i,j)} x_{ij} x_{kl}}{\sum_{(i,j)\in\mathcal{G}} x_{ij}} \tag{2}$$

$$\text{subject to} \quad \sum_{(i,j)\in A_s} x_{ij} \ge n_s, \quad \text{for all } s = 1, 2, \ldots, S \tag{3}$$

$$\sum_{(i,j)\in\mathcal{G}} b_{ij} x_{ij} \le B \tag{4}$$

$$x_{ij} \in \{0, 1\}, \quad \text{for all } (i, j) \in \mathcal{G} \tag{5}$$

Constraint (3) enforces the requirement that to protect species of type s adequately, we must select at least n_s sites in which species s is present. Constraint (4) ensures that the total cost of selected sites does not exceed the conservation budget B.

The objective function represented in (2) is a ratio of two functions $f(x)$ and $g(x)$, with the numerator being a quadratic in the 0–1 decision variables x. Let Ω be the feasible region defined by (3)–(5). Using concepts from fractional programming (Borrero et al., 2017), we instead

KOP	AIJLNP	HMOP	HIP	IJMNP	DKNOP	ABILMP	IKLMOP	NOP	CKNP
LMN	GHOP	ABOP	NOP	AELMNP	CKNOP	AFMOP	EIL	MNP	AJKLMNOP
KMO	BNO	FKNOP	O	IKNOP	OP	GO	DJKNP	MOP	LNO
P	FKOP	NOP	FIP	BJKO	IKNO	ILOP	MOP	DEFGJOP	O
DNP	EO	JMNOP	IJMNP	MNOP	ACGHKNP	KMOP	CMOP	BKP	M
BNOP	MOP	CDGHKMOP	ABCGLMP	DNO	EMOP	OP	NOP	NOP	DHMNO
JKLMOP	LMNO	EHMO	ACEGMO	NOP	JNOP	DNP	JMNOP	BKLNP	MOP
IKLNP	NP	DOP	HN	DMNP	AGHLOP	LP	IOP	NOP	MNOP
KMNOP	GHJMOP	NOP	AFKMNP	LOP	ACEGOP	IMNP	HMOP	DHLMOP	ADIJKLNOP
IMNOP	GMNO	MOP	KMNOP	MOP	DNO	MNOP	KNOP	DLMOP	IMN

Fig. 2. Distribution of species A–P over a 10 × 10 grid; see Pimm and Lawton (1998).

solve the following problem: $\max\{f(x) - \lambda g(x) : x \in \Omega\}$. The value λ is to be selected so that this new objective function value achieves the maximum value 0. This can be accomplished by applying Dinkelbach's algorithm (Dinkelbach, 1967) in order to identify the optimal $\lambda^\star$, which will then equal the maximum value of the ratio $\frac{f(x)}{g(x)}$ over $x \in \Omega$. For our problem, we have some natural bounds on the optimal value $\lambda^\star$. From Section 2.1, we know $4n = p + 2m > 2m$ so that $f(x)/g(x) = m/n < 2$. Also we expect that $m > n$ holds for any reasonable configuration, so that $m/n > 1$. Consequently, the optimal $\lambda^\star$ can be found by searching the interval $(1, 2)$, say starting with the initial value $\lambda = 1.5$.

Note that in the optimization problem $\max\{f(x) - \lambda g(x) : x \in \Omega\}$, the function $f(x)$ involves quadratic terms $x_{ij}x_{kl}$. Using a standard transformation, this optimization problem can be converted into an equivalent one involving 0–1 decision variables only in a linear fashion. Namely, define the new binary variables $y_{ijkl} = x_{ij}x_{kl}$:

$$y_{ijkl} = \begin{cases} 1 & \text{if both sites } (i,j) \text{ and } (k,l) \text{ are selected} \\ 0 & \text{otherwise.} \end{cases}$$

The following constraints then ensure that y_{ijkl} equals 1 if and only if sites (i,j) and (k,l) are both selected:

$$y_{ijkl} \le x_{ij}, \quad \text{for all } (i,j),(k,l) \in \mathcal{G} \tag{6}$$

$$y_{ijkl} \le x_{kl}, \quad \text{for all } (i,j),(k,l) \in \mathcal{G} \tag{7}$$

$$x_{ij} + x_{kl} - y_{ijkl} \le 1, \quad \text{for all } (i,j),(k,l) \in \mathcal{G} \tag{8}$$

$$y_{ijkl} \ge 0, \quad \text{for all } (i,j),(k,l) \in \mathcal{G} \tag{9}$$

Specifically, constraints (6)–(7) ensure that y_{ijkl} must equal to 0 unless both x_{ij} and x_{kl} equal 1, while constraint (8) ensures that if both (i,j) and (k,l) are selected for conservation, then y_{ijkl} must equal 1. Consequently, $y_{ijkl} = x_{ij}x_{kl}$ always holds. Moreover, it can be shown that these y_{ijkl} variables can be relaxed to be continuous rather than binary.

Thus the following linear mixed-integer model P_2 can be iteratively solved using Dinkelbach's algorithm to identify a reserve configuration with maximum density $\lambda^\star$:

$$(P_2) \text{ maximize } \frac{1}{2} \sum_{(i,j)\in\mathcal{G}} \sum_{(k,l)\in N(i,j)} y_{ijkl} - \lambda \sum_{(i,j)\in\mathcal{G}} x_{ij} \tag{10}$$

$$\text{subject to} \quad (3)\text{–}(9)$$

As an illustration of this approach, consider the problem posed by Pimm and Lawton (1998), in which there are 16 species A–P distributed across a 10 × 10 grid; see Fig. 2. Suppose that each species s should appear in at least two sites, so that $n_s = 2$ for all s. In addition, suppose that at most $B = 15$ sites can be chosen, achieved by setting all $b_{ij} = 1$.

We used MATLAB version R2020a (MATLAB, 2020), CPLEX Optimization Studio version 12.9.0 (IBM, 2020), Python 3.7 and a MacBook Pro equipped with an Intel i7 processor and 16 GB memory to solve our optimization models. Dinkelbach's algorithm converged to an optimal solution $x^\star = (x^\star_{ij})$ after just two iterations (in less than 10 s) with $\lambda^\star = \delta = \frac{21}{15} = 1.4$. This solution is given in Fig. 3, which displays the 0–1 values of $x^\star$. The shaded squares highlight the 15 sites selected for the reserve.

0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	1	1	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0	0	0
0	1	1	1	0	0	0	0	0	0
0	1	1	1	0	0	0	0	0	0
0	1	1	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Fig. 3. Solution obtained using Algorithm 1 for the Pimm and Lawton (1998) data set ($B = 15, n_s = 2$ for all species s). Selected sites are shown shaded.

0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	0	0	0	0	0
0	0	1	1	1	0	0	0	0	0
0	0	1	1	1	0	0	0	0	0
0	0	1	1	1	0	0	0	0	0
0	0	1	1	1	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

Fig. 4. Alternative solution to model P_2 for the Pimm and Lawton (1998) data set ($B = 15, n_s = 2$ for all species s). Selected sites are shown shaded.

2.3. Alternative solutions

The maximum density solution to the Pimm and Lawton problem shown in Fig. 3 is just one optimal solution to the model P_2. An alternative solution, also with $\delta = \frac{m}{n} = \frac{21}{15} = 1.4$, is displayed in Fig. 4. Having alternative optimal solutions is quite valuable, as it provides decision makers the opportunity to apply additional criteria to select a configuration from among those alternative optima.

A family of optimal solutions to the original problem can be found by adding suitable constraints to the original formulation. Let $x^\star = (x^\star_{ij})$ denote an optimal solution found by solving model P_2, let I_0 be the set of indices ij for which $x^\star_{ij} = 0$, and let I_1 be the set of indices ij for which $x^\star_{ij} = 1$. Then add to the original model P_2 the following constraint

$$\sum\{x_{ij} : ij \in I_1\} - \sum\{x_{ij} : ij \in I_0\} \le |I_1| - 1 \tag{11}$$

and solve model P_2 with this added constraint. Constraint (11) cuts off the current optimal solution $x^\star$. Moreover it can be shown that this added constraint does not preclude any other alternative optimal solution (Balas and Jeroslow, 1972). If solution of model P_2 with this added constraint produces the same objective function value, we have now identified an alternative optimal solution.

a	bc	abc	a	ac	abc	c	ab		ac		c		b
a	ab	bc	abc	ac	bc	bc	c	a		ac	a	b	bc
abc	bc	abc	abc	c	bc	c	ab	ab	abc	ab	bc	a	a
ab		c	a	abc	a		b	b		c	ac		c
abc	bc	c			b			ab	bc	b	abc	a	
	abc	ab	b	c	c	a	a	b	b	abc	ab	b	bc
	ac	ac		c		b				a	abc	a	b
abc	abc	a	c	b	a		c		c	b	b	c	
bc	b	a	ac	b		bc		a	bc	a	ac	c	a
c	c			bc	ab	a	b	abc	b	ac	a		
b		a	ac	b	abc	abc	ac	ab	abc	b	c	a	a
		ac	ac	b	ac	ab	abc	bc	abc	ac	a	c	c
b	c		ac	ac	a	abc	abc	abc	a	abc			b
a		ab	c	b			a	ab	a	b	bc		a

Fig. 5. Distribution of species a, b, c over a 14×14 grid; see Weeraseena et al. (2014).

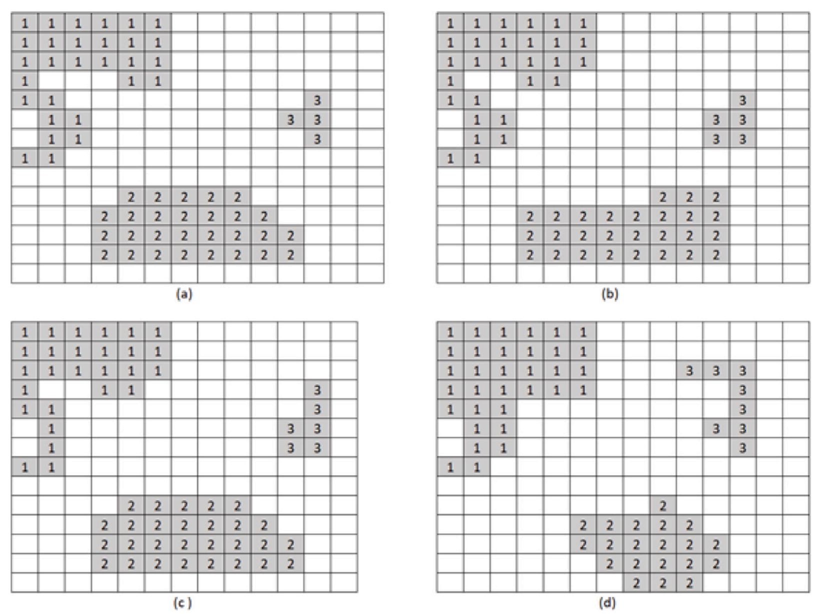

Fig. 6. Some alternative optimal solutions for the 14×14 problem of Weeraseena et al. (2014). Species requirements are $n_a = 49, n_b = 41, n_c = 44$ with an upper bound $B = 61$. Clusters obtained are labeled 1, 2, 3, etc.

This process continues by adding one new constraint of the form (11) each time, based on the 0/1 nature of the optimal $x^\star$ just found. Once we obtain a smaller value of the objective function, the process can stop. This procedure will generate all alternative optimal configurations. We illustrate this generation process by solving a problem previously introduced in Weerasena et al. (2014). This problem involves a 14×14 grid of sites in which there are three species a, b, c with $n_a = 49, n_b = 41, n_c = 44$ and an upper bound $B = 61$ on the number of selected sites. The distribution of species on these 196 sites is depicted in Fig. 5.

Solving this problem using model P_2 results in the solution shown in Fig. 6(a). The optimal density was therefore achieved not by creating a single cluster but was identified by the model as consisting of three separate compact clusters of sites. By applying the procedure discussed here, we generated a number of alternative optimal solutions, several of which are displayed in Fig. 6(b)–(d). Notice that the models developed here do not ensure connectivity of the entire reserve system.

To explore this further, we consider the 14×14 problem just introduced, now with more general site costs b_{ij} generated using a uniform distribution over the interval $[1, 10]$. Fig. 7 displays the cost values for the individual reserve sites. We now apply model P_2 in order to generate compact clusters using the upper bound $B = 290$.

Two of the resulting optimal solutions are shown in Fig. 8; the first contains four clusters and the second contains three clusters. Since our objective is to generate a reserve system that is connected, the following section will discuss an efficient approach for judiciously adding corridors to connect such clusters. It should be noted that solution of model P_2 never required more than 10 CPU seconds for any of the results shown in Figs. 6 and 8.

2.4. Addition of corridors

As seen in Section 2.3, model P_2 can produce compact disjoint clusters of core areas to be protected. We now develop a heuristic for adding corridors to connect these clusters in a manner that uses relatively few corridor sites. Since the added corridor sites provide habitat for species, a secondary objective will be to bias the selection of corridor sites in favor of those that are species rich.

The addition of corridors to connect wildlife reserves has been approached using a variety of optimization techniques. Önal and Briers (2006) develop an integer programming formulation to create a fully connected reserve that satisfies species representation constraints. However there is no assurance that the resulting configuration achieves compactness. By contrast a number of researchers have developed models for connecting existing clusters of wildlife reserves (possibly

1.23	4.92	3.97	6.57	3.40	5.76	5.62	8.07	5.45	1.72	1.59	1.87	6.37	1.96
4.15	2.82	5.35	4.48	6.22	7.31	5.50	4.07	4.85	7.99	9.58	1.74	8.66	1.24
1.60	9.74	6.42	2.52	5.72	1.41	4.97	3.91	4.48	7.63	1.12	3.42	1.23	4.49
4.73	5.96	2.02	1.38	6.92	4.75	6.95	8.93	2.21	7.74	5.89	9.27	4.12	9.23
5.86	8.44	2.59	5.40	7.30	2.68	3.47	6.66	2.69	3.55	3.57	5.95	9.90	3.09
2.45	2.96	6.88	1.28	7.34	1.28	6.31	1.59	4.56	2.78	4.89	3.61	9.58	2.86
5.00	5.73	2.48	7.37	7.99	9.61	5.33	8.72	8.64	7.44	1.23	1.92	3.55	4.50
6.13	7.32	5.55	1.04	6.07	7.10	9.34	8.34	4.12	8.81	4.14	8.25	2.85	8.23
1.96	2.99	7.48	7.93	3.06	4.73	6.78	8.70	2.60	5.25	7.23	3.22	7.86	8.55
2.13	3.67	1.71	4.87	1.79	1.92	7.42	9.85	5.30	6.03	5.63	8.50	9.90	8.44
7.60	9.56	6.77	6.61	2.48	8.42	5.17	3.35	4.93	9.04	4.36	2.77	4.34	2.14
8.93	3.59	7.01	4.40	9.57	5.81	3.23	6.01	1.09	3.29	4.10	1.57	5.47	9.42
5.18	3.62	2.90	2.05	6.80	8.92	6.50	3.93	8.63	9.59	5.43	2.21	5.55	3.99
7.98	8.07	6.66	1.89	8.40	2.75	3.78	8.32	2.98	9.68	2.46	2.26	8.23	6.37

Fig. 7. General site costs for the 14 × 14 problem of Weeraseena et al. (2014).

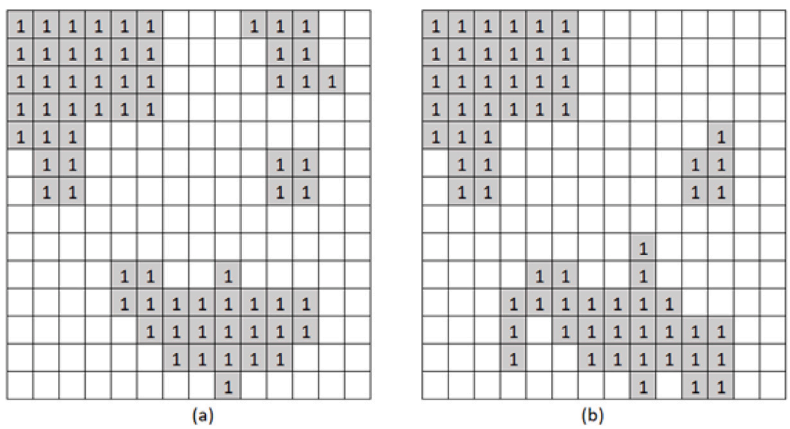

Fig. 8. Some alternative optimal solutions for the 14 × 14 problem of Weeraseena et al. (2014) with nonuniform cost values.

compact) in an optimal way, by treating the problem as a Network Steiner Tree problem (Dilkina et al., 2017; Hamaide et al., 2022; Williams and ReVelle, 1998; Williams and Snyder, 2005) or as a connected subgraph problem (Conrad et al., 2012; Dilkina and Gomes, 2010). Our approach takes as given a compact set of wildlife reserves and seeks to connect them using a one parcel wide corridor system in a computationally effective manner. Possibly our approach can be extended to develop a corridor system that takes into account corridor geometry (Matisziw et al., 2015; St John et al., 2018; Wang et al., 2022; Yemshanov et al., 2022); this is clearly an avenue for further exploration.

In order to connect clusters using relatively few corridor sites in an efficient way, we developed a heuristic that *sequentially* joins "nearest" clusters until all are connected, at which point we can identify efficient corridors joining the given clusters. Initially we identify two existing clusters that are closest to one another (measured by the minimum number of corridor sites needed to join them). Since there will generally be multiple ways to minimally connect these two clusters, we select a minimum length corridor whose constituent sites contain the largest number of species. The original two clusters plus this identified corridor are then coalesced into a new (pseudo) cluster, and the process is repeated. Note that such corridors need not necessarily directly join two original clusters; they may attach to a previously added corridor (in this way producing a heuristically obtained Steiner tree of corridors joining the original clusters).

More specifically, consider disjoint clusters $C_1, C_2, \ldots, C_k$ produced by applying model P_2. We expand in turn clusters $C_1, C_2, \ldots, C_k$ by one layer of adjacent sites until two expanded sets connect. This can be accomplished by keeping the current collection of k sets in a queue, adding to the set at the front of the queue all (new) sites that are adjacent to that set, and then placing this expanded set at the end of the queue. Eventually, two of the original clusters, say C_i and C_j, will become connected. Since there are multiple ways in which C_i and C_j can be connected with the fewest number of corridors sites, we use a suitably adapted longest path algorithm to identify a *minimum* number of connecting sites that contain the *maximum* number of species, giving rise to a new cluster to replace C_i and C_j. Then this process is repeated, now with a smaller number of (expanded) clusters, to eventually identify a set of corridors that connect all of the original clusters $C_1, C_2, \ldots, C_k$. This heuristic runs very quickly, especially in comparison to the exact approach discussed in Section 3.1; our computational experience shows that the heuristic solutions produced are either optimal or near optimal.

There are several ways in which we can (if desired) generate alternative configurations that join the original clusters $C_1, C_2, \ldots, C_k$ in a near-optimal way. First of all, it is straightforward to modify our longest path algorithm to generate alternative corridor designs that join C_i and C_j with a minimum length corridor containing the maximum number of species. If additional alternative configurations are desired, then the original ordering of $C_1, C_2, \ldots, C_k$ in the queue data structure can be permuted and the entire process can be repeated. Since our heuristic is reasonably efficient, such multiple invocations can produce a rich set of alternative optimal solutions.

We illustrate in Fig. 9 the result of applying this approach to the configuration displayed in Fig. 6(a), which contains three compact

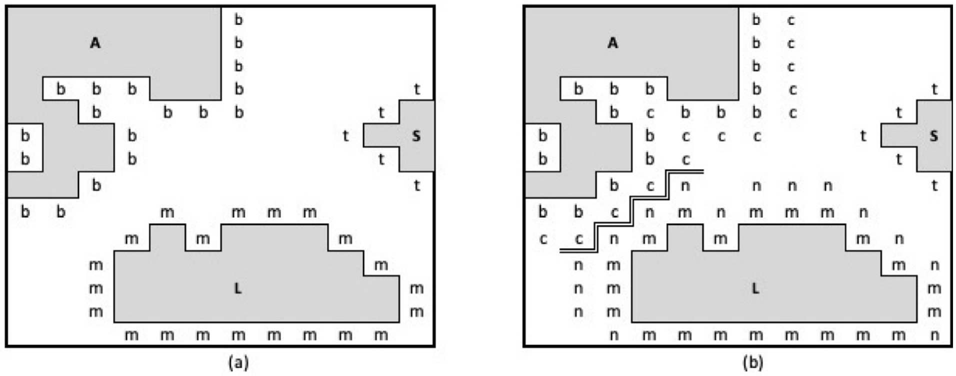

Fig. 9. Expansion process obtained from initial compact clusters in Fig. 6(a). Part (a) shows the expansion of A, L, S by one level to $A1, L1, S1$. Part (b) shows the expansion of $A1, L1$ by one level to $A2, L2$.

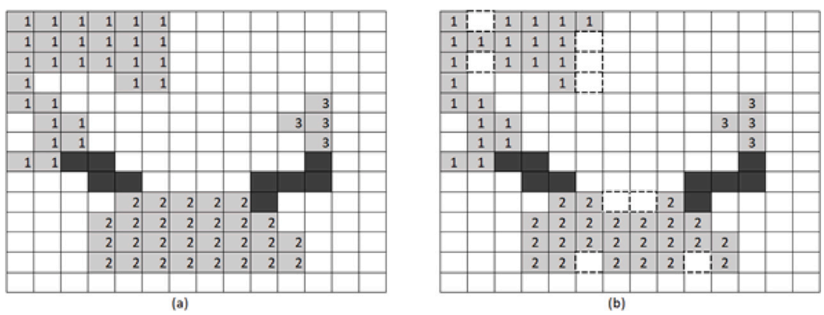

Fig. 10. The result of adding corridors (shown dark) and removing boundary sites (shown with dotted line) using alternative solution (a) of Fig. 6 for the 14 × 14 problem of Weeraseena et al. (2014). The solution in (a) adds nine corridor sites, while the solution in (b) then removes nine boundary sites.

clusters. For ease of illustration, we rename these clusters as A, L, S. Then we expand each cluster in turn by adding one layer at a time. The first layer added to cluster A contains adjacent sites labeled b; the first layer added to cluster L contains adjacent sites labeled m; and the first layer added to cluster S contains adjacent sites labeled t. The result of these first-level expansions is shown in Fig. 9(a); let us denote these expanded clusters as $A1, L1, S1$. Since these expanded clusters are still disjoint, the expansion process is continued. Namely, a second layer of adjacent sites (labeled c) is added to $A1$, and then a second layer of adjacent sites (labeled n) is added to $L1$, resulting in the enlarged clusters $A2$ and $L2$ shown in Fig. 9(b). At this point, we observe that clusters $A2$ and $L2$ meet along the indicated double-lines shown in Fig. 9(b), meaning that the two original clusters A and L can be merged using a corridor involving sites b, c, n, m. Then among the relevant possibilities, we use a modified longest path algorithm to select the $b-c-n-m$ corridor with maximum species representation. This process is then repeated with the two (expanded) clusters now available until they merge. The added corridors that result are shown dark in Fig. 10(a). This heuristic added nine extra sites connecting the three original clusters shown in Fig. 6(a). These added corridor sites cover an additional $[4, 4, 4]$ of species a, b, c, respectively.

Having identified a set of (say M) promising corridor sites to connect the original clusters C_i, we can consider two options. The first is to identify for decision makers the proposed set of corridors — which involves adding the fewest number of unit width corridor sites that have the most impact in terms of increasing species numbers. They can then weight the benefit of adding these particular sites to the overall reserve, versus the added cost. Alternatively, we can seek to identify certain "boundary" sites whose removal will still result in a connected system that respects the species and budgetary constraints (3)–(4). Removing a relatively small number of these boundary sites is expected to have small impact on the overall compactness of the original reserve clusters. The following discussion indicates how this can be approached by developing an appropriate 0–1 integer programming model.

For species s, let $E(s)$ denote the number of sites (in the current clusters) containing s in excess of n_s, plus the number of corridor sites containing s. So $E(s)$ represents the excess number of sites containing s present in the current configuration of clusters and corridors. In addition, certain sites appearing on the boundary of the original clusters C_i can be removed without disconnecting its containing cluster; let K denote the set of such boundary sites. Such (non-disconnecting) boundary sites can be found efficiently, with the computation time growing linearly with the number of nodes and edges in the graph (Tarjan, 1972).

We would like to determine a set of boundary sites $(i, j) \in K$ to remove to maintain feasibility of the system with the added corridors and removed boundary sites. Consequently, we solve the following optimization problem:

$$(P_3) \text{ minimize } \sum_{(i,j)\in K} x_{ij} \tag{12}$$

$$\text{subject to } \sum_{(i,j)\in(A_s\cap K)} x_{ij} \le E(s), \text{ for all } s \in S \tag{13}$$

$$\sum_{(i,j)\in K} x_{ij} \ge M \tag{14}$$

$$x_{ij} \in \{0,1\}, \text{ for all } (i,j) \in K \tag{15}$$

Constraint (13) ensures that the removed boundary sites maintain feasibility of the species requirement (3) after adding the corridor sites and removing the boundary sites K. Constraint (14) ensures that enough boundary sites are removed to account for the M added corridor sites, thus maintaining the original budgetary requirement (4). If model P_3 provides a feasible solution, then we are able to incorporate all the extra sites used to create the corridors. If this model is infeasible, then we can solve again with an alternative set of corridors.

To illustrate this process, we return to the situation shown in Fig. 6(a), in which the species and cost requirements are given by $[n_a, n_b, n_c] = [49, 41, 44]$ and $B = 61$. For the solution shown in

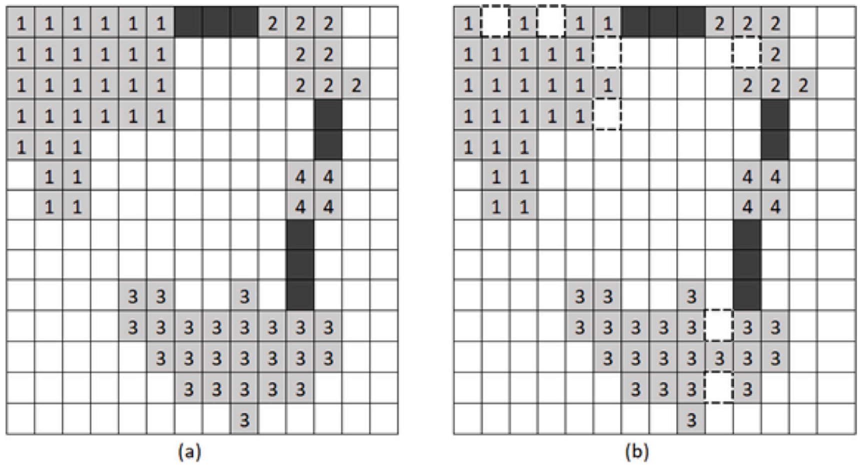

Fig. 11. The result of adding corridors (shown dark) to alternative solution (a) of Fig. 8 for the 14 × 14 problem of Weeraseena et al. (2014). The solution in (a) adds eight corridor sites, while the solution in (b) then removes seven corridor sites.

Fig. 10(a), we obtain species coverage $[49, 42, 44]$ for the three clusters involving 61 sites. The corridors shown add nine corridor sites with species coverage $[4, 4, 4]$. Model P_3 produces a feasible solution with nine boundary sites, represented in Fig. 10(b) as cells with dotted sides. These nine boundary sites remove species coverage of $[4, 5, 4]$, giving a net species coverage (original + corridor — boundary) of $[49, 42, 44] + [4, 4, 4] - [4, 5, 4] = [49, 41, 44]$, a connected feasible solution involving $B = 61$ sites.

This approach can also be applied to instances with general site costs b_{ij}. To illustrate, consider the solution with four clusters in Fig. 8(a) that was generated by applying model P_2. Using our previously described method for adding species-rich corridors, we identified the eight corridor sites shown in Fig. 11(a). We now discuss in general how to determine a set of boundary sites to remove to compensate for the addition of corridor sites.

Suppose that M corridor sites have been added to a configuration of disjoint clusters. We solve the following optimization problem (P_4), which generalizes the model (P_3) employed in the uniform cost case. Here B' is the additional cost (in excess of B) incurred when these M corridor sites are added.

$$(P_4) \text{ minimize } \sum_{(i,j)\in K} x_{ij} \tag{16}$$

$$\text{subject to } \sum_{(i,j)\in(A_s\cap K)} x_{ij} \leq E(s), \text{ for all } s \in S \tag{17}$$

$$\sum_{(i,j)\in K} b_{ij} x_{ij} \geq B' \tag{18}$$

$$x_{ij} \in \{0, 1\}, \text{ for all } (i, j) \in K \tag{19}$$

A solution of P_4 identifies a set of boundary sites that can be removed from the configuration to restore feasibility: that is, a solution that satisfies constraints (3)–(4) in model P_1. For the example with corridors shown in Fig. 11(a), application of model P_4 identifies seven sites to remove, giving the final configuration shown in Fig. 11(b). Here the total cost is $288.18 + 51.36 - 51.76 = 287.78 \leq 290$. The species requirements are also satisfied.

3. Model evaluation and real-world data sets

This section assesses the quality and the effectiveness of our overall approach. We evaluate our methodology using test cases previously discussed as well as a more extensive real-world data set. First, we compare the heuristic approach for corridor generation discussed in Section 2.4 with an exact approach. Next, we study the performance of our proposed models on data sets derived from real-world examples. All computations reported were performed using the computational environment specified in Section 2.2.

3.1. An exact model for corridor generation

We used the optimization model proposed in Billionnet (2021) to evaluate the quality of the solutions produced by the heuristic algorithm described in Section 2.4. This flow-based optimization model is similar to those developed for various Steiner tree problems (Ljubić, 2021).

Specifically, the Billionnet approach seeks to connect up at minimum total cost a given set of clusters $C_1, C_2, \ldots, C_k$ by using a set of intermediate sites (i, j), where there is a cost k_{ij} associated with each site (i, j). The model assumes that if a site (i, j) is already protected or included in a cluster, the cost associated with that site is zero. Recall that the heuristic approach developed in Section 2.4 is designed to identify a fewest number of corridors that connect the given clusters, and that have the largest impact in terms of increasing species number. Thus, to have a reasonable comparison of our heuristic algorithm with the Billionnet flow model, we modified the objective function of the Billionnet model as follows. For each site (i, j) that is not included in one of the given clusters, we define the site cost $k_{ij} = M - s_{ij}$ where M is a sufficiently large positive number and s_{ij} is the total number of species contained in site (i, j). As a result, identifying a set of corridor sites with the minimum total cost k_{ij} is equivalent to identifying a smallest set of corridor sites that overall contain the maximum number of species.

When we applied this exact model to the two examples illustrated in Figs. 6(a) and 8(a), the exact model and the heuristic algorithm produced corridors with the same number of corridor sites and the same number of additional species. Specifically, the exact and heuristic solutions were identical for the instance shown in Fig. 8(a). In contrast, the exact model produced an alternative set of corridors for the instance shown in Fig. 6(a); both solutions are displayed in Fig. 12. The heuristic approach took 0.0015 and 0.0020 s, while the exact model took 41 and 45 s, to produce corridors for the test cases given in Figs. 6(a) and 8(a), respectively.

3.2. Results using a real-world data set

Next we discuss the quality and effectiveness of our overall approach using a data set obtained from the Atlas of the Vascular Flora of the Iberian Peninsula (AFLIBER) (Ramos-Gutiérrez et al., 2021). Thirty species were selected at random from this source, which contains distribution data for thousands of plant species in Spain and Portugal. Distribution data for the 30 species was then filtered down to a 200 km× 200 km block and analyzed at the 10 km × 10 km resolution. The region of Portugal and Spain outlined in Fig. 13(a) was selected because the

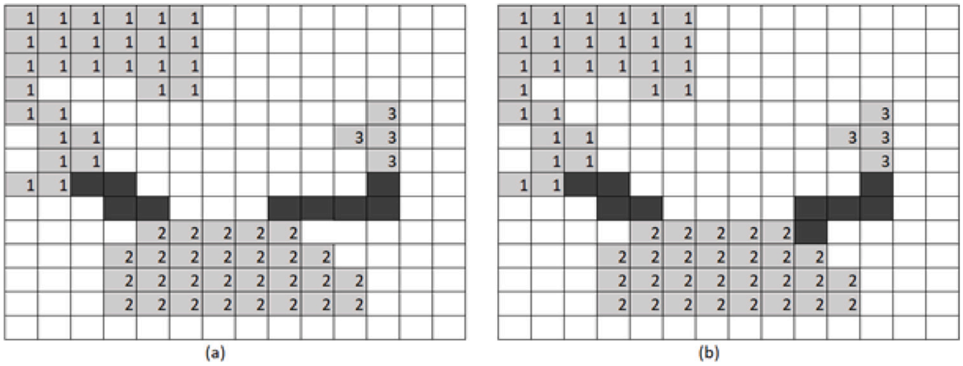

Fig. 12. Corridors produced by (a) the exact model and by (b) the heuristic algorithm for the three clusters shown in Fig. 6(a).

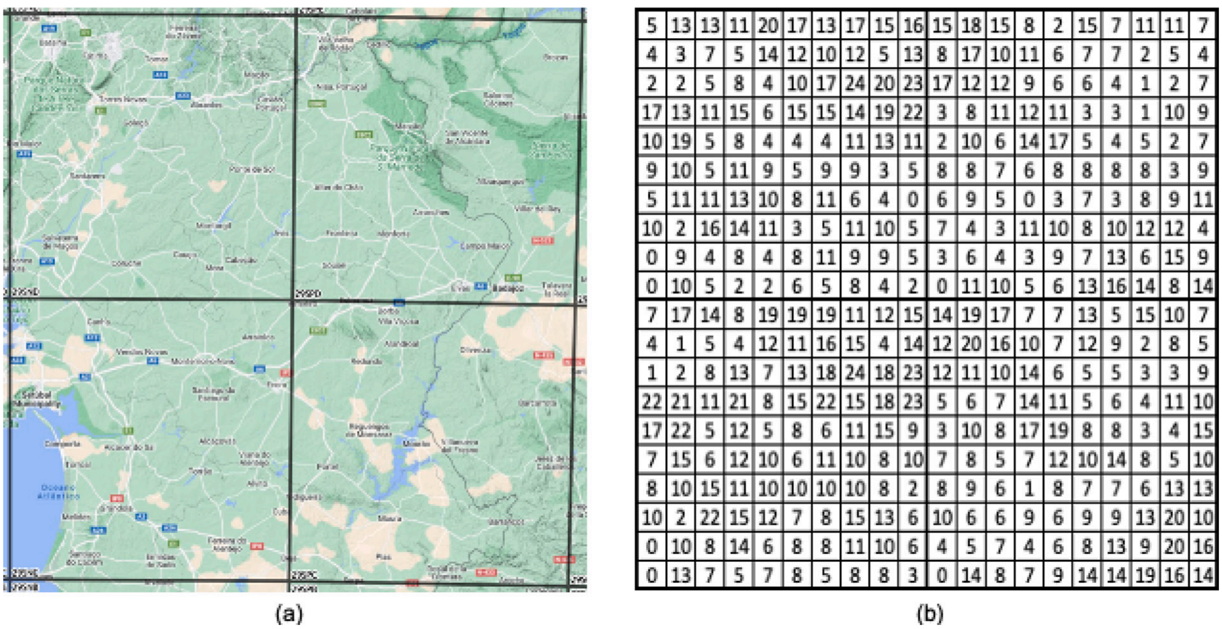

Fig. 13. The 40,000 km^2 area of Portugal and Spain from which the real species data is taken is shown in Fig. 13(a), while Fig. 13(b) gives the corresponding species distribution map for the 30 selected species.

level of completeness of distribution records was relatively high and the species richness was relatively low. The corresponding species numbers s_{ij} are tabulated in Fig. 13(b).

For our next set of analyses, we assume uniform site costs $b_{ij} = 1$. The first set of computational tests investigated the 10×10 data set from the top-left block of Fig. 13 and considered separately the cases $n_s = 1$ and $n_s = 2$ for all species s.

For each of these two cases, the overall cost budget B was set equal to 10, 20, 30, 40, and 50. The results of running model P_2 are shown in Fig. 14, where it is seen that in all cases we obtained a single, nearly square cluster. Table 1 shows the optimal value $\lambda^{\star}$ and the CPU time (in seconds) required by the Dinkelbach algorithm to solve model P_2. In each case, convergence was achieved in at most two iterations of the algorithm.

In order to explore larger problems and the addition of corridors, we considered the entire 20×20 system displayed in Fig. 13(b). First, we applied model P_2 to five separate instances, whose characteristics are given in Table 2. For each species s, the number of required species n_s is a random integer uniformly drawn from the interval $[1, 20]$. The results obtained from model P_2 for these five instances are shown in Fig. 15, where it is seen that in all cases we obtained multiple clusters. Table 3 displays the number of clusters produced, the optimal graph density value $\lambda^{\star}$, and the computation time, respectively. For example,

Table 1
Solution of model P_2 for the 10×10 data set.

$n_s = 1$			$n_s = 2$		
B	Optimal λ	Time (s)	B	Optimal λ	Time (s)
10	1.3333	9.65	10	1.3333	6.81
20	1.5500	12.60	20	1.5500	12.49
30	1.6333	10.25	30	1.6333	3.84
40	1.6750	6.14	40	1.6750	8.62
50	1.7143	5.59	50	1.7143	7.80

Table 2
Characteristics of the 20×20 test problems.

Test cases	B	n_s
Problem 1	30	$n_s \in [1, 20]$
Problem 2	30	$n_s \in [1, 20]$
Problem 3	35	$n_s \in [1, 20]$
Problem 4	35	$n_s \in [1, 20]$
Problem 5	40	$n_s \in [1, 20]$

model P_2 produced four clusters with graph density 1.1333 for Problem 1; it took 19.49 s to obtain this result.

Second, we applied the exact and heuristic methods to add corridors to these clustered systems, resulting in the corridors shown dark in

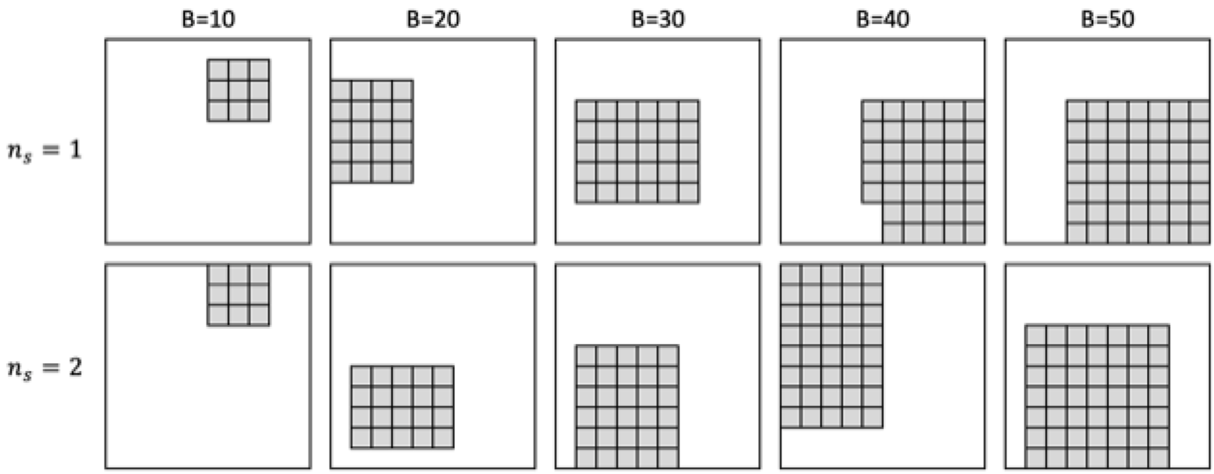

Fig. 14. Clusters obtained by applying model P_2 to the 10 × 10 data set, for various species coverage requirements n_s and budget constraints B.

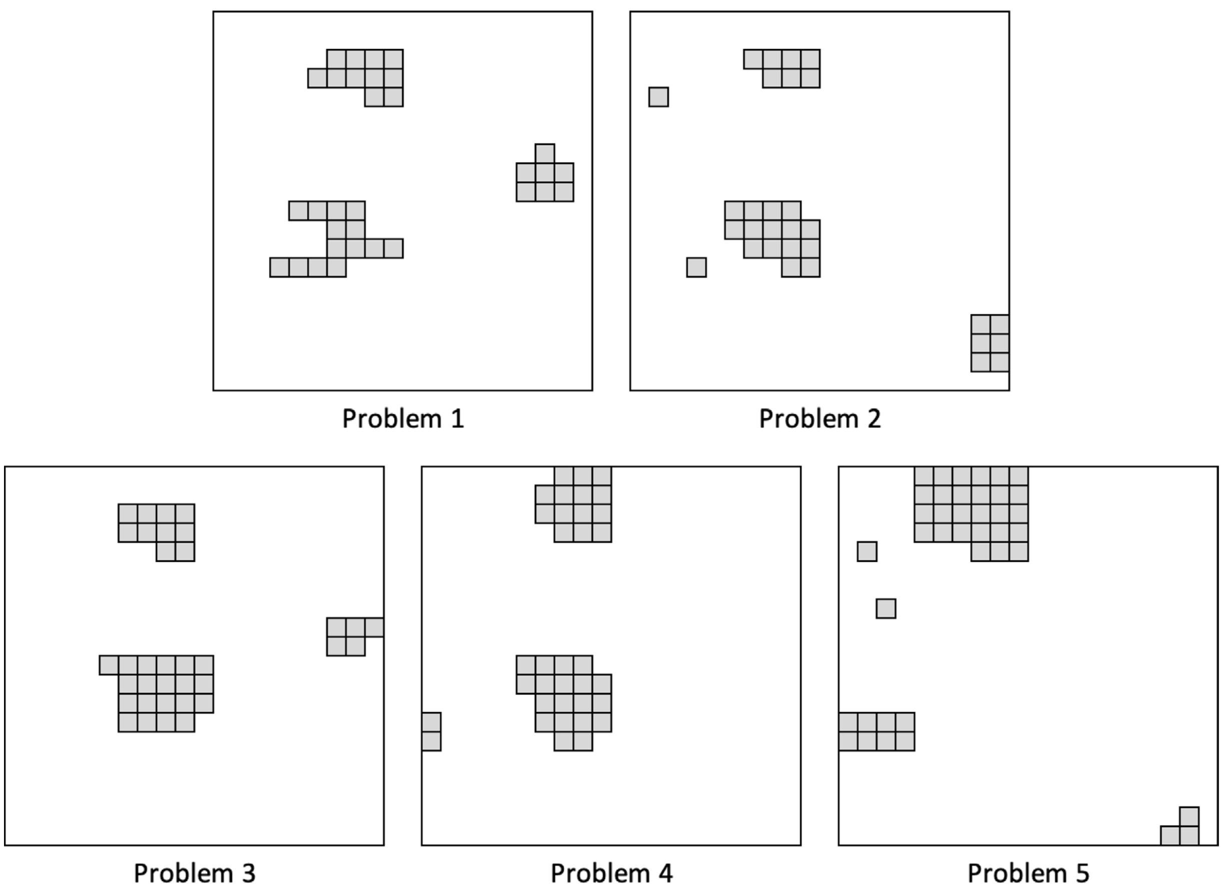

Fig. 15. Clusters produced by model P_2 for the 20 × 20 data set.

Fig. 16. The leftmost column of this figure displays the corridors produced by the exact method while the rightmost column displays the corridors produced by our heuristic algorithm. Numerical results associated with this analysis are provided in Table 4. For Problems 3 and 4, the exact and heuristic approaches used the same set of corridor sites to connect all clusters. For Problems 1, 2 and 5, the heuristic method identified corridors containing one extra site compared to the exact method. Notice that with one extra site, we achieve a significantly larger number of species; this reflects the advantage of prioritizing species-rich sites in our models for corridor selection. Table 4 shows that the computation time of the heuristic method is less than 0.3 s for all test problems, which is some 600 times faster than the exact flow-based model (Billionnet, 2021).

To further validate the effectiveness of our heuristic, we applied model P_2 to a set of test instances with nonuniform site costs b_{ij}, randomly generated from the interval $[1, 10]$. Three different random scenarios were considered, corresponding to three independent sets of randomly selected b_{ij} with associated sum of all site costs $Btotal = 2190, 2176, 2159$. For each such $Btotal$, the overall budget value B was varied as 5%, 10%, and 15% of $Btotal$. As in the prior analyses, the number of required species n_s is a random integer uniformly drawn from the interval $[1, 20]$. We considered ten separate test instances with

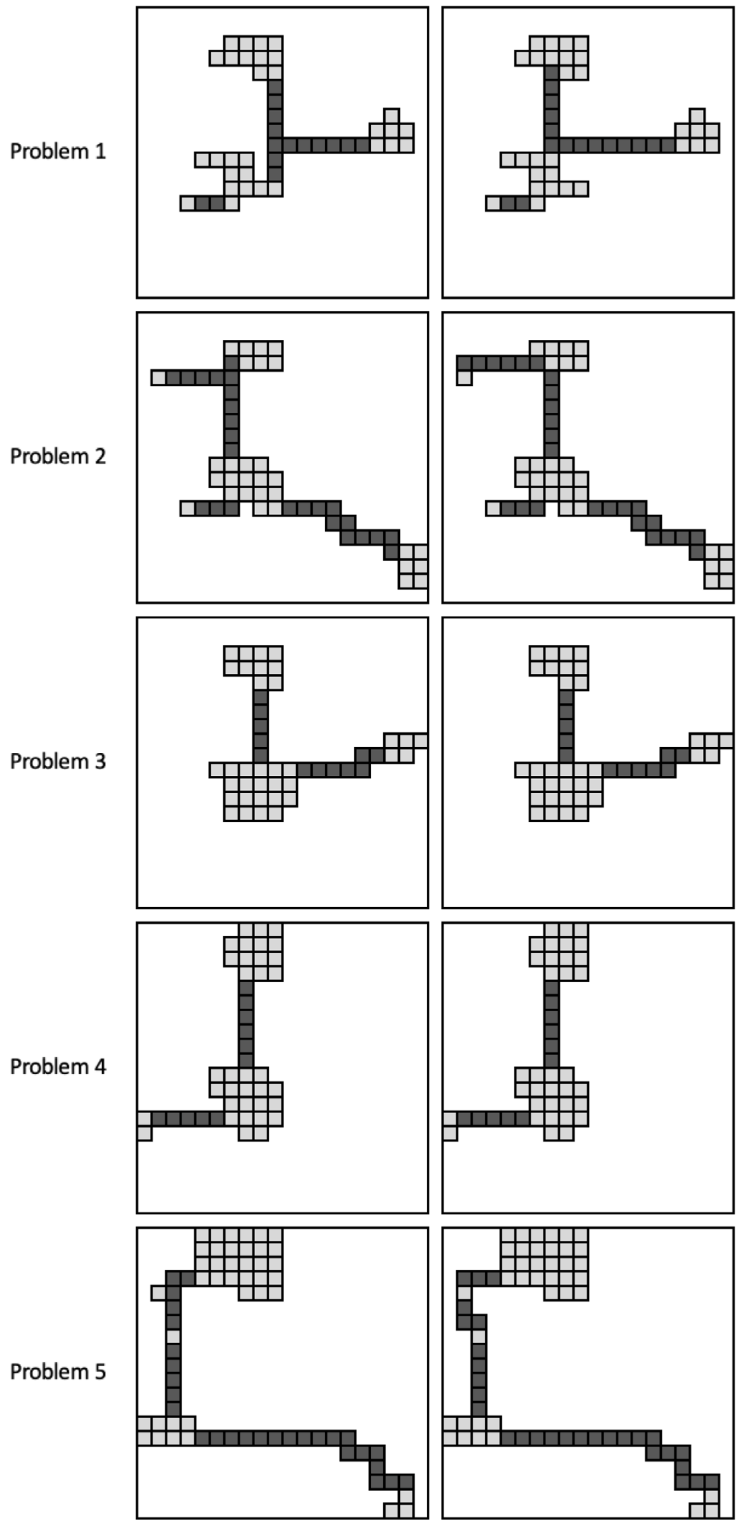

Fig. 16. Corridors (shown dark) added for the 20 × 20 test problems. The leftmost column shows corridors added using the exact method while the rightmost column shows corridors added using the heuristic method.

nonuniform site costs, whose characteristics are given in Table 5. Model P_2 was then applied to each test instance. Table 6 displays the number of clusters produced, the optimal graph density λ^*, and the computation time, respectively.

We then applied the exact and heuristic methods to add corridors to these clustered systems with nonuniform site costs. The numerical results associated with this analysis are provided in Table 7. Only for Problem 14 did the heuristic method identify more corridors to connect all clusters, compared to the exact method; in that case the heuristic added just one extra site and in the process added five to the number of covered species. Furthermore, as seen in Table 7, the computation time of the heuristic method is less than 0.8 s for all test problems, some two orders of magnitude less than that for the exact method. This analysis

Table 3
Solution of model P_2 for the 20 × 20 data set.

Test cases	B	# Clusters	Optimal λ	Time (s)
Problem 1	30	4	1.1333	19.49
Problem 2	30	5	1.2000	21.81
Problem 3	35	3	1.3714	23.54
Problem 4	35	3	1.4000	40.45
Problem 5	40	5	1.3750	19.31

Table 4
Summary of adding corridors for the 20 × 20 data set.

Test cases	# Corridor sites		Species count		Time (s)	
	Exact	Heuristic	Exact	Heuristic	Exact	Heuristic
Problem 1	15	16	114	128	189.12	0.2991
Problem 2	25	26	244	292	181.44	0.0514
Problem 3	12	12	122	122	190.02	0.2141
Problem 4	11	11	130	130	187.26	0.0663
Problem 5	28	29	279	303	182.11	0.0112

Table 5
Characteristics of the 20 × 20 test problems with nonuniform site costs.

Test cases	$Btotal$	B	n_s
Problem 6	2190	10%	$n_s \in [1, 20]$
Problem 7	2190	10%	$n_s \in [1, 20]$
Problem 8	2176	5%	$n_s \in [1, 20]$
Problem 9	2176	5%	$n_s \in [1, 20]$
Problem 10	2176	10%	$n_s \in [1, 20]$
Problem 11	2159	10%	$n_s \in [1, 20]$
Problem 12	2159	10%	$n_s \in [1, 20]$
Problem 19	2159	15%	$n_s \in [1, 20]$
Problem 14	2159	10%	$n_s \in [1, 20]$
Problem 15	2159	15%	$n_s \in [1, 20]$

again highlights the efficiency of our heuristic method in comparison to the exact model, when applied to problems with nonuniform site costs.

4. Discussion

Protected areas are generally viewed as the primary mechanism by which to preserve biological diversity, for example in the Aichi Biodiversity Target 11 of protecting at least 17% of terrestrial lands with “ecologically represented and well connected systems of protected areas” (COBD, 2010). There has been significant progress toward meeting these global goals of coverage and connectivity, but challenges remain and the goals themselves are now viewed as inadequate. We have described a method to delineate a system of connected reserves that is efficient in protecting species. In the next sections, we summarize the contributions (and limitations) of our approach, address some caveats, and outline some further extensions of this work.

4.1. Contributions of our approach

We have described a suite of algorithms to aid in designing efficient systems of protected areas that contain extensive, compact cores along with their connecting corridors. Specifically, we first identify reserve systems comprised of a single or several compact clusters; we discuss how in fact a set of alternative compact areas can be generated by our model, important in pursuing the investigation of alternative designs. Using a heuristic approach, we then connect any resulting disjoint clusters with corridors chosen for their conservation values. Our computational study suggests that this heuristic is reasonably efficient and typically near optimal in performance. This approach addresses our four goals of designing efficient reserve systems that provide species coverage, and are extensive, compact, and connected. A third optimization model can be applied, if desired, to remove certain boundary sites of the compact cores to compensate for the addition of the identified corridor sites. It is important to note that these three optimization models can be applied sequentially, or individually, to aid decision makers.

Our approach shares features with many others in systematic reserve planning. First, it is implemented for presence/absence data on species (or habitat types, communities, or other features of conservation interest) on a rectangular array. Second, it can take existing protected areas as fixed and build strategically on them for greater coverage and connectivity. Third, our model can be used to justify the inclusion of particular sites in a reserve and identify alternatives if those sites are not chosen.

There are ways in which our approach departs from others in the literature. Notably, we employ a somewhat novel measure of compactness to guide the optimization models. While maximizing the graph density δ is mathematically equivalent to certain other measures proposed in the literature, our measure satisfies $1 \le \delta \le 2$, making it easy to initialize the Dinkelbach algorithm using $\lambda = 1.5$. (For hexagonal grids we have the bounds $1 \le \delta \le 3$ and so could initialize with $\lambda = 2$.) Even for the largest problems analyzed here, this algorithm converged in at most four iterations. Moreover our density measure allows a comparison between candidate reserves possessing different sizes since δ must have a numerical value within the interval $[1, 2]$. We also indicate how it is possible to generate alternative compact clusters, all having maximum density, by systematically adding a new constraint (11) at each step to problem P_2. To illustrate the potential for such exploration, we found over 40 alternative solutions when investigating the 14 × 14 test problem of Weerasena et al. (2014) by sequentially adding constraints of the form (11) to model P_2. This is an important consideration since it enables the analyst to provide a rich set of alternative configurations, which can then be evaluated by decision makers. Additionally, we explicitly consider exchanging core for corridor sites if they can better serve the twin goals of coverage and connectivity. This option is a natural result of seeking to maximize the conservation return on investments (Kareiva, 2010); indeed, Fuller et al. (2010) suggests that strategically trading the 1% least cost-effective of Australia’s nearly 7000 protected areas could dramatically increase the conservation value of the protected area system.

Our approach can also be used to quantify the costs and benefits of connecting versus enlarging reserves (Adams et al., 2019; Falcy and Estades, 2007), critical when the funds and political and social will for the protection of lands are limited (Seidl et al., 2021). Corridors will often be through developed landscapes and may require restoration, making them expensive competitors for limited funds with core reserves. Indeed, this was one of Simberloff et al.’s (1992) primary concerns about proposed corridor systems in Florida.

One other feature of our modeling approach is worth emphasizing. The optimization model P_1 strictly segregates the objective function from the species representation and budget constraints, unlike approaches like Marxan (Ball et al., 2009; Daigle et al., 2020), which incorporate cost, boundary length, and possibly conservation targets into a single objective function. The latter approach requires a judicious selection of weighting factors since these components have quite different units. Moreover, if the decision maker wishes to systematically vary these weighting factors to obtain a set of alternative reserve designs, the binary nature of the variables makes it impossible in general to generate all “non-inferior (Pareto optimal)” solutions (Williams, 1998). By contrast, we optimize density (a single objective) subject to constraints on total cost (in $ units) and species representation (in numbers); the decision maker will have a good sense of the total cost budget and also the species representation requirements and (if desired) can parametrically vary these values.

4.2. Caveats and extensions

There are a number of caveats in this work. First, the distributions of many species, even among relatively well-studied birds (Ocampo-Peñuela et al., 2016), may not be known with certainty, and there will

Table 6
Solution of model P_2 for the 20 × 20 test problems with nonuniform site costs.

Test cases	# Clusters	Optimal λ	Time (s)	Test cases	# Clusters	Optimal λ	Time (s)
Problem 6	2	1.5122	272.88	Problem 11	2	1.5417	33.14
Problem 7	3	1.3958	52.18	Problem 12	2	1.5294	220.96
Problem 8	4	1.1081	14.59	Problem 13	2	1.6944	168.08
Problem 9	4	1.1471	20.65	Problem 14	3	1.4255	44.05
Problem 10	3	1.3750	5304	Problem 15	2	1.6620	174.80

Table 7
Summary of adding corridors for the 20 × 20 test problems with nonuniform site costs.

Test cases	# Corridor sites		Species count		Time (s)	
	Exact	Heuristic	Exact	Heuristic	Exact	Heuristic
Problem 6	18	18	222	222	49.9750	0.751
Problem 7	25	25	261	261	61.5339	0.2084
Problem 8	21	21	290	290	82.6125	0.7403
Problem 9	20	20	255	255	82.7988	0.1567
Problem 10	11	11	144	144	60.9956	0.0107
Problem 11	8	8	70	70	46.626	0.0062
Problem 12	6	6	40	40	47.5666	0.0056
Problem 13	7	7	51	51	42.9829	0.0059
Problem 14	14	15	181	186	59.8559	0.0141
Problem 15	10	10	150	150	43.8985	0.0066

be cases in which species are inaccurately omitted from or assigned to sites. These errors are of particular concern in tropical forests which are rich in biodiversity but not in survey data (Amano and Sutherland, 2013; Donaldson et al., 2016; Meyer et al., 2016; Rodrigues, 2011; Tydecks et al., 2018). Methods for reserve planning have addressed these uncertainties in various ways (reviewed in Arthur et al., 2002). The threshold approach considers a species to be present in a site if its expectation of occurring there exceeds some threshold, say 90%, and then maximizes the coverage of species that exceed this threshold (Haight et al., 2000). This ignores sites for which species just miss the threshold but uses standard maximum coverage algorithms. In contrast, the expected coverage approach maximizes the expected number of species covered, using all the data but in a nonlinear 0–1 optimization problem of greater difficulty (Polasky et al., 2000). Watts et al. (2020) modified Marxan to incorporate four kinds of uncertainty: that a species or some other feature occurs in a site, that it occurs but might disappear, that it occurs but might be degraded and no longer contribute to conservation goals, and that it does not occur but might appear through ecological succession or other mechanisms. Haider et al. (2018) presented a robust optimization approach that protects against worst-case scenarios, while Rosing et al. (2002) defended heuristic methods, including their own, when the underlying data are too soft to support "optimal" solutions. After all, Csuti et al. (1997) found that 17 of 18 heuristic approaches covered at least 422 species in Oregon with 23 sites, just shy of the optimum of 426. Additional errors can arise simply from the choice of grid size and origin (Dunn, 2010; Witte et al., 2008), and all can lead to the inefficiencies or failures in a reserve design.

We view this investigation as a first step in creating an integrated approach to reserve design. Encouraged by the efficiency and quality of results for the 20 × 20 real-world test problems, we plan to expand our set of test data to encompass larger examples. In order to accommodate additional data sources, we also plan to generalize our approach to incorporate sites defined on hexagonal grids or more generally defined on an arbitrary graph. Another avenue for future investigation is expanding the width of added corridors (currently one parcel wide) to promote a more accommodating habitat for transit.

5. Closing comments

The 2011 Aichi Accords set a lofty goal of protecting 17% of terrestrial lands worldwide, which helped focus attention on the biodiversity crisis and launch an enormous increase in protected lands. Unfortunately, this growth has been far from optimal in terms of protecting biodiversity, often favoring lands of limited agricultural or other value rather than ones satisfying specific conservation goals (Maxwell et al., 2020; Stokstad, 2020). The resulting biases and shortfalls represent a lost opportunity (Visconti et al., 2019), which must be countered in future conservation efforts. In addition, there has been a continued downgrading and degazetting of protected areas around the world (Mascia and Pailler, 2011; Ruaro et al., 2020; Symes et al., 2016), conservatively affecting an area equal to that of Mexico (Kroner et al., 2019), and undermining all these conservation efforts. Finally, there is growing concern that the goal of protecting 17% of terrestrial lands will not suffice to protect the earth's biodiversity — even 40% may not be enough (Leclère et al., 2020). There are now calls to protect far more lands, from 30% (Dinerstein et al., 2019) to 50% (Wilson, 2016), and with more specific objectives (Baillie and Zhang, 2018; Noss et al., 2012). Proposed objectives include protecting key biodiversity areas (Visconti et al., 2019), climate refugia and connectivity (Carrasco et al., 2021; Carroll and Noss, 2021), wilderness (Watson et al., 2018), and ecosystem functions (Maxwell et al., 2020), achieving no net loss of natural ecosystems (Maron et al., 2020), limiting further climate change (Arneth et al., 2020; Dinerstein et al., 2020) and extinctions (Rounsevell et al., 2020), and accommodating other conservation and development goals (Di Marco et al., 2016; Sala et al., 2021; Tilman et al., 2017). These conservation goals will require significant additional growth in the protected area networks on earth that is strategic in minimizing costs and conflicts with other needs. Our integrated approach begins to address this challenge by providing for the joint design of large, compact reserves with their connecting corridors.

Software availability section

We used MATLAB version R2020a (MATLAB, 2020), CPLEX Optimization Studio version 12.9.0 (IBM, 2020) to solve our optimization models. We used Python 3.7 to implement the heuristic algorithm.

CRediT authorship contribution statement

Lakmali Weerasena: Conceive the ideas and methodology, Developed and implemented the mathematical methods and analyses. **Douglas Shier:** Conceive the ideas and methodology, Developed and implemented the mathematical methods and analyses. **David Tonkyn:** Conceive the ideas and methodology, Wrote the biological sections. **Mark McFeaters:** Computation analysis and implementations. **Christopher Collins:** Computation analysis and implementations.

Data availability

All empirical data used in this study can be found in the MODERN repository in GitHub https://github.com/lweeras/MODERN.git.

Acknowledgment

All contributed to the final draft and gave final approval for submission.

Funding

This work was supported by the National Science Foundation, United States under Grant #2137622

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Adams, V.M., Iacona, G.D., Possingham, H.P., 2019. Weighing the benefits of expanding protected areas versus managing existing ones. Nat. Sustain. 2 (5), 404–411.

Alagador, D., Trivino, M., Cerdeira, J.O., Brás, R., Cabeza, M., Araújo, M.B., 2012. Linking like with like: optimising connectivity between environmentally-similar habitats. Landsc. Ecol. 27 (2), 291–301.

Amano, T., Sutherland, W.J., 2013. Four barriers to the global understanding of biodiversity conservation: wealth, language, geographical location and security. Proc. R. Soc. B: [Biol] 280 (1756), 20122649.

Arneth, A., Shin, Y.-J., Leadley, P., Rondinini, C., Bukvareva, E., Kolb, M., Midgley, G.F., Oberdorff, T., Palomo, I., Saito, O., 2020. Post-2020 biodiversity targets need to embrace climate change. Proc. Natl. Acad. Sci. 117 (49), 30882–30891.

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. Environ. Model. Assess. 7 (2), 81–89.

Baillie, J., Zhang, Y.-P., 2018. Space for nature. Science 361 (6407), 1051.

Balas, E., Jeroslow, R., 1972. Canonical cuts on the unit hypercube. SIAM J. Appl. Math. 23 (1), 61–69.

Ball, I.R., Possingham, H.P., Watts, M.E., 2009. Marxan and relatives: software for spatial conservation prioritization. In: Moilanen, A., Wilson, K., H.P., P. (Eds.), Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools. Oxford University Press, pp. 185–195.

Beier, P., 2012. Conceptualizing and designing corridors for climate change. Ecol. Restor. 30 (4), 312–319.

Beier, P., Loe, S., 1992. In my experience: a checklist for evaluating impacts to wildlife movement corridors. Wildl. Soc. Bull. 20 (4), 434–440.

Bengtsson, J., Angelstam, P., Elmqvist, T., Emanuelsson, U., Folke, C., Ihse, M., Moberg, F., Nyström, M., 2003. Reserves, resilience and dynamic landscapes. AMBIO: A J. Human Environ. 32 (6), 389–396.

Beyer, H.L., Dujardin, Y., Watts, M.E., Possingham, H.P., 2016. Solving conservation planning problems with integer linear programming. Ecol. Model. 328, 14–22.

Billionnet, A., 2012. Designing an optimal connected nature reserve. Appl. Math. Model. 36 (5), 2213–2223.

Billionnet, A., 2013. Mathematical optimization ideas for biodiversity conservation. European J. Oper. Res. 231 (3), 514–534.

Billionnet, A., 2016. Designing connected and compact nature reserves. Environ. Model. Assess. 21 (2), 211–219.

Billionnet, A., 2021. Designing Protected Area Networks: A Mathematical Programming Approach. EDP Sciences-Current Natural Sciences.

Borrero, J., Gillen, C., Prokopyev, O., 2017. Fractional 0-1 programming: applications and algorithms. J. Global Optim. 69, 255–282.

Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D., Rodrigues, A.S., 2006. Global biodiversity conservation priorities. Science 313 (5783), 58–61.

Bruner, A.G., Gullison, R.E., Balmford, A., 2004. Financial costs and shortfalls of managing and expanding protected-area systems in developing countries. BioScience 54 (12), 1119–1126.

Butler, E.P., Bliss-Ketchum, L.L., de Rivera, C.E., Dissanayake, S., Hardy, C.L., Horn, D.A., Huffine, B., Temple, A.M., Vermeulen, M.E., Wallace, H., 2022. Habitat, geophysical, and eco-social connectivity: benefits of resilient socio–ecological landscapes. Landsc. Ecol. 37 (1), 1–29.

Cabeza, M., Moilanen, A., 2001. Design of reserve networks and the persistence of biodiversity. Trends Ecol. Evol. 16 (5), 242–248.

Carrasco, L., Papeş, M., Sheldon, K.S., Giam, X., 2021. Global progress in incorporating climate adaptation into land protection for biodiversity since Aichi targets. Global Change Biol. 27 (9), 1788–1801.

Carroll, C., Noss, R.F., 2021. Rewilding in the face of climate change. Conserv. Biol. 35 (1), 155–167.

Chartrand, G., Zhang, P., 2005. Introduction To Graph Theory. McGraw-Hill Education.

Church, R.L., Stoms, D.M., Davis, F.W., 1996. Reserve selection as a maximal covering location problem. Biol. Cons. 76 (2), 105–112.

COBD. Convention on Biological Diversity, Nagoya, Japan, 18 to 29 October 2010, 2010. https://www.cbd.int/sp/targets/.

Conrad, J.M., Gomes, C.P., van Hoeve, W.-J., Sabharwal, A., Suter, J.F., 2012. Wildlife corridors as a connected subgraph problem. J. Environ. Econ. Manag. 63 (1), 1–18.

Cova, T.J., Church, R.L., 2000. Contiguity constraints for single-region site search problems. Geogr. Anal. 32 (4), 306–329.

Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L., Camm, J.D., Kershaw, M., Kiester, A.R., Downs, B., Hamilton, R., Sahr, K., 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. Biol. Cons. 80 (1), 83–97.

Daigle, R.M., Metaxas, A., Balbar, A.C., McGowan, J., Treml, E.A., Kuempel, C.D., Possingham, H.P., Beger, M., 2020. Operationalizing ecological connectivity in spatial conservation planning with Marxan Connect. Methods in Ecol. Evol. 11 (4), 570–579.

Di Marco, M., Butchart, S.H., Visconti, P., Buchanan, G.M., Ficetola, G.F., Rondinini, C., 2016. Synergies and trade-offs in achieving global biodiversity targets. Conserv. Biol. 30 (1), 189–195.

Diamond, J.M., 1975. The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. Biol. Cons. 7 (2), 129–146.

Dilkina, B., Gomes, C.P., 2010. Solving connected subgraph problems in wildlife conservation. In: International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming. Springer, pp. 102–116.

Dilkina, B., Houtman, R., Gomes, C.P., Montgomery, C.A., McKelvey, K.S., Kendall, K., Graves, T.A., Bernstein, R., Schwartz, M.K., 2017. Trade-offs and efficiencies in optimal budget-constrained multispecies corridor networks. Conserv. Biol. 31 (1), 192–202.

Dinerstein, E., Joshi, A.R., Vynne, C., Lee, A.T.L., Pharand-Deschênes, F., França, M., Fernando, S., Birch, T., Burkart, K., Asner, G., Olson, D., 2020. A "Global Safety Net" to reverse biodiversity loss and stabilize Earth's climate. Sci. Adv. 6 (36), eabb2824.

Dinerstein, E., Vynne, C., Sala, E., Joshi, A.R., Fernando, S., Lovejoy, T.E., Mayorga, J., Olson, D., Asner, G.P., Baillie, J.E., et al., 2019. A global deal for nature: guiding principles, milestones, and targets. Sci. Adv. 5 (4), eaaw2869.

Dinkelbach, W., 1967. On nonlinear fractional programming. Manage. Sci. 13 (7), 492–498.

Dissanayake, S.T., Önal, H., Westervelt, J.D., Balbach, H.E., 2012. Incorporating species relocation in reserve design models: an example from Ft. Benning GA. Ecol. Model. 224 (1), 65–75.

Donaldson, M.R., Burnett, N.J., Braun, D.C., Suski, C.D., Hinch, S.G., Cooke, S.J., Kerr, J.T., 2016. Taxonomic bias and international biodiversity conservation research. Facets 1, 105–213.

Dunn, A.G., 2010. Grid-induced biases in connectivity metric implementations that use regular grids. Ecography 33 (3), 627–631.

Falcy, M.R., Estades, C.F., 2007. Effectiveness of corridors relative to enlargement of habitat patches. Conserv. Biol. 21 (5), 1341–1346.

Fischer, D.T., Church, R.L., 2003. Clustering and compactness in reserve site selection: an extension of the biodiversity management area selection model. For. Sci. 49 (4), 555–565.

Fischman, R.L., 2005. The significance of national wildlife refuges in the development of US conservation policy. J. Land Use & Environ. Law 21 (1), 1–22.

Frankham, R., 2005. Genetics and extinction. Biol. Cons. 126 (2), 131–140.

Fuller, R.A., McDonald-Madden, E., Wilson, K.A., Carwardine, J., Grantham, H.S., Watson, J.E., Klein, C.J., Green, D.C., Possingham, H.P., 2010. Replacing underperforming protected areas achieves better conservation outcomes. Nature 466 (7304), 365–367.

Fuller, T., Sarkar, S., 2006. LQGraph: a software package for optimizing connectivity in conservation planning. Environ. Model. Softw. 21 (5), 750–755.

Gomes, C.P., 2011. Computational sustainability. In: IDA 2011 COnference, Lecture Notes in Computer Science. Vol. 7014, Springer, Berlin, Heidelberg.

Gupta, A., Dilkina, B., Morin, D.J., Fuller, A.K., Royle, J.A., Sutherland, C., Gomes, C.P., 2019. Reserve design to optimize functional connectivity and animal density. Conserv. Biol. 33 (5), 1023–1034.

Haider, Z., Charkhgard, H., Kwon, C., 2018. A robust optimization approach for solving problems in conservation planning. Ecol. Model. 368, 288–297.

Haight, R.G., Revelle, C.S., Snyder, S.A., 2000. An integer optimization approach to a probabilistic reserve site selection problem. Oper. Res. 48 (5), 697–708.

Hall, K.R., Anantharaman, R., Landau, V.A., Clark, M., Dickson, B.G., Jones, A., Platt, J., Edelman, A., Shah, V.B., 2021. Circuitscape in Julia: empowering dynamic approaches to connectivity assessment. Land 10 (3), 301.

Hamaide, V., Hamaide, B., Williams, J.C., 2022. Nature reserve optimization with buffer zones and wildlife corridors for rare species. Sustain. Anal. Model. 2, 100003.

Hannah, L., 2011. Climate change, connectivity, and conservation success. Conserv. Biol. 25 (6), 1139–1142.

Hess, G.R., Fischer, R.A., 2001. Communicating clearly about conservation corridors. Landsc. Urban Plan. 55 (3), 195–208.

Hilty, J., Worboys, G.L., Keeley, A., Woodley, S., Lausche, B., Locke, H., Carr, M., Pulsford, I., Pittock, J., White, J.W., et al., 2020. Guidelines for conserving connectivity through ecological networks and corridors. Best Practice Protected Area GUidelines Series NO. 30. Gland, Switzerland: IUCN. IUCN, Gland, Switzerland.

IBM ILOG CPLEX Optimization Studio, 2020. Version (12.9.0). Armonk, NewYork: IBM. https://www.ibm.com/docs/en/icos/12.9.0?topic=v1290-introducing-ilog-cplex-optimization-studio.

Jenkins, C.N., Joppa, L., 2009. Expansion of the global terrestrial protected area system. Biol. Cons. 142 (10), 2166–2174.

Kareiva, P., 2010. Trade-in to trade-up. Nature 466 (7304), 322–323.

Keeley, A.T., Beier, P., Jenness, J.S., 2021. Connectivity metrics for conservation planning and monitoring. Biol. Cons. 255, 109008.

Kiester, A.R., Scott, J.M., Csuti, B., Noss, R.F., Butterfield, B., Sahr, K., White, D., 1996. Conservation prioritization using GAP data. Conserv. Biol. 10 (5), 1332–1342.

Kingsland, S.E., 2002. Creating a science of nature reserve design: perspectives from history. Environ. Model. Assess. 7 (2), 61–69.
Kroner, R.E.G., Qin, S., Cook, C.N., Krithivasan, R., Pack, S.M., Bonilla, O.D., Cort-Kansinally, K.A., Coutinho, B., Feng, M., Garcia, M.I.M., et al., 2019. The uncertain future of protected lands and waters. Science 364 (6443), 881–886.
Krosby, M., Tewksbury, J., Haddad, N.M., Hoekstra, J., 2010. Ecological connectivity for a changing climate. Conserv. Biol. 24 (6), 1686–1689.
Lande, R., 1988. Genetics and demography in biological conservation. Science 241 (4872), 1455–1460.
Leclère, D., Obersteiner, M., Barrett, M., Butchart, S.H., Chaudhary, A., De Palma, A., DeClerck, F.A., Di Marco, M., Doelman, J.C., Dürauer, M., et al., 2020. Bending the curve of terrestrial biodiversity needs an integrated strategy. Nature 585 (7826), 551–556.
Ljubić, I., 2021. Solving steiner trees: recent advances, challenges, and perspectives. Networks 77 (2), 177–204.
Margules, C.R., Pressey, R.L., 2000. Systematic conservation planning. Nature 405 (6783), 243–253.
Maron, M., Simmonds, J.S., Watson, J.E., Sonter, L.J., Bennun, L., Griffiths, V.F., Quétier, F., von Hase, A., Edwards, S., Rainey, H., et al., 2020. Global no net loss of natural ecosystems. Nat. Ecol. Evol. 4 (1), 46–49.
Mascia, M.B., Pailler, S., 2011. Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications. Conserv. Lett. 4 (1), 9–20.
Matisziw, T.C., Alam, M., Trauth, K.M., Inniss, E.C., Semlitsch, R.D., McIntosh, S., Horton, J., 2015. A vector approach for modeling landscape corridors and habitat connectivity. Environ. Model. Assess. 20 (1), 1–16.
MATLAB, 2020. Version (R2020a). Natick, Massachusetts: The MathWorks inc.. https://www.mathworks.com/.
Maxwell, S.L., Cazalis, V., Dudley, N., Hoffmann, M., Rodrigues, A.S., Stolton, S., Visconti, P., Woodley, S., Kingston, N., Lewis, E., et al., 2020. Area-based conservation in the twenty-first century. Nature 586 (7828), 217–227.
McCullough, D.R., 1996. Metapopulations and Wildlife Conservation. Island Press.
McDonnell, M.D., Possingham, H.P., Ball, I.R., Cousins, E.A., 2002. Mathematical methods for spatially cohesive reserve design. Environ. Model. Assess. 7 (2), 107–114.
Meyer, C., Weigelt, P., Kreft, H., 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. Ecol. Lett. 19 (8), 992–1006.
Murcia, C., 1995. Edge effects in fragmented forests: implications for conservation. Trends Ecol. Evol. 10 (2), 58–62.
Nalle, D.J., Arthur, J.L., Montgomery, C.A., Sessions, J., 2002a. Economic and spatial impacts of an existing reserve network on future augmentation. Environ. Model. Assess. 7 (2), 99–105.
Nalle, D.J., Arthur, J.L., Sessions, J., 2002b. Designing compact and contiguous reserve networks with a hybrid heuristic algorithm. For. Sci. 48 (1), 59–68.
Noss, R.F., Dobson, A.P., Baldwin, R., Beier, P., Davis, C.R., Dellasala, D.A., Francis, J., Locke, H., Nowak, K., Lopez, R., et al., 2012. Bolder thinking for conservation. Conserv. Biol. 26 (1), 1–4.
Ocampo-Peñuela, N., Jenkins, C.N., Vijay, V., Li, B.V., Pimm, S.L., 2016. Incorporating explicit geospatial data shows more species at risk of extinction than the current Red List. Sci. Adv. 2 (11), e1601367.
O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W., Frankham, R., 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biol. Cons. 133 (1), 42–51.
Öhman, K., Lämås, T., 2005. Reducing forest fragmentation in long-term forest planning by using the shape index. Forest Ecol. Manag. 212 (1–3), 346–357.
Olson, D.M., Dinerstein, E., 2002. The Global 200: priority ecoregions for global conservation. Ann. Missouri Botanical Gard. 89 (2), 199–224.
Önal, H., Briers, R.A., 2002. Incorporating spatial criteria in optimum reserve network selection. Proc. R. Soc. B 269 (1508), 2437–2441.
Önal, H., Briers, R.A., 2003. Selection of a minimum-boundary reserve network using integer programming. Proc. R. Soc. B 270 (1523), 1487–1491.
Önal, H., Briers, R.A., 2005. Designing a conservation reserve network with minimal fragmentation: a linear integer programming approach. Environm. Model. Assess. 10 (3), 193–202.
Önal, H., Briers, R.A., 2006. Optimal selection of a connected reserve network. Oper. Res. 54 (2), 379–388.
Önal, H., Wang, Y., Dissanayake, S.T., Westervelt, J.D., 2016. Optimal design of compact and functionally contiguous conservation management areas. European J. Oper. Res. 251 (3), 957–968.
Pascual-Hortal, L., Saura, S., 2006. Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. Landsc. Ecol. 21 (7), 959–967.
Pimm, S.L., Lawton, J., 1998. Planning for biodiversity. Science 279, 2068–2069.
Polasky, S., Camm, J.D., Solow, A.R., Csuti, B., White, D., Ding, R., 2000. Choosing reserve networks with incomplete species information. Biol. Cons. 94 (1), 1–10.
Possingham, H., Ball, I., Andelman, S., 2000. Mathematical methods for identifying representative reserve networks. In: Ferson, S., Burgman, M. (Eds.), Quant. Methods for Conserv. Biol. 291–306.
Pressey, R.L., Humphries, C.J., Margules, C.R., Vane-Wright, R.I., Williams, P.H., 1993. Beyond opportunism: key principles for systematic reserve selection. Trends Ecol. Evol. 8 (4), 124–128.
Ramos-Gutiérrez, I., Lima, H., Pajarón, S., Romero-Zarco, C., Sáez, L., Pataro, L., Molina-Venegas, R., Rodríguez, M., Moreno-Saiz, J.C., 2021. Atlas of the vascular flora of the Iberian Peninsula biodiversity hotspot (AFLIBER). Global Ecol. Biogeogr. 30 (10), 1951–1957.
Rodrigues, A.S., 2011. Improving coarse species distribution data for conservation planning in biodiversity-rich, data-poor, regions: no easy shortcuts. Animal Conserv. 14 (2), 108–110.
Rodrigues, A.S., Akcakaya, H.R., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Chanson, J.S., Fishpool, L.D., Da Fonseca, G.A., Gaston, K.J., et al., 2004a. Global gap analysis: priority regions for expanding the global protected-area network. BioScience 54 (12), 1092–1100.
Rodrigues, A.S., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Cowling, R.M., Fishpool, L.D., Da Fonseca, G.A., Gaston, K.J., Hoffmann, M., et al., 2004b. Effectiveness of the global protected area network in representing species diversity. Nature 428 (6983), 640–643.
Rosing, K.E., ReVelle, C.S., Williams, J.C., 2002. Maximizing species representation under limited resources: a new and efficient heuristic. Environ. Model. Assess. 7 (2), 91–98.
Rounsevell, M.D., Harfoot, M., Harrison, P.A., Newbold, T., Gregory, R.D., Mace, G.M., 2020. A biodiversity target based on species extinctions. Science 368 (6496), 1193–1195.
Ruaro, R., Laurance, W., Mormul, R.P., 2020. Brazilian national parks at risk. Science 367 (6481), 990.
Sala, E., Mayorga, J., Bradley, D., Cabral, R.B., Atwood, T.B., Auber, A., Cheung, W., Costello, C., Ferretti, F., Friedlander, A.M., et al., 2021. Protecting the global ocean for biodiversity, food and climate. Nature 592 (7854), 397–402.
Scott, J.M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., D'Erchia, F., Edwards Jr., T.C., et al., 1993. Gap analysis: a geographic approach to protection of biological diversity. Wildlife Monographs 123, 3–41.
Seidl, A., Mulungu, K., Arlaud, M., van den Heuvel, O., Riva, M., 2021. The effectiveness of national biodiversity investments to protect the wealth of nature. Nat. Ecol. Evol. 5 (4), 530–539.
Sessions, J., 1992. Solving for habitat connections as a steiner network problem. For. Sci. 38 (1), 203–207.
Simberloff, D., Farr, J.A., Cox, J., Mehlman, D.W., 1992. Movement corridors: conservation bargains or poor investments? Conserv. Biol. 6 (4), 493–504.
St John, R., Tóth, S.F., Zabinsky, Z.B., 2018. Optimizing the geometry of wildlife corridors in conservation reserve design. Oper. Res. 66 (6), 1471–1485.
Stokstad, E., 2020. Global efforts to protect biodiversity fall short. Science 369 (6510), 1418.
Suter, J.F., Conrad, J.M., Gomes, C.P., Van Hoeve, W.J., Sabharwal, A., 2008. Optimal corridor design for grizzly bear in the U.S. Northern Rockies. Technical report.
Symes, W.S., Rao, M., Mascia, M.B., Carrasco, L.R., 2016. Why do we lose protected areas? Factors influencing protected area downgrading, downsizing and degazettement in the tropics and subtropics. Global Change Biol. 22 (2), 656–665.
Tarjan, R., 1972. Depth-first search and linear graph algorithms. SIAM J. Comput. 1 (2), 146–160.
Tilman, D., Clark, M., Williams, D.R., Kimmel, K., Polasky, S., Packer, C., 2017. Future threats to biodiversity and pathways to their prevention. Nature 546 (7656), 73–81.
Tydecks, L., Jeschke, J.M., Wolf, M., Singer, G., Tockner, K., 2018. Spatial and topical imbalances in biodiversity research. PLoS One 13 (7), e0199327.
UNEP-WCMC. World database on protected areas, 2021. https://www.unep-wcmc.org/resources-and-data/wdpa (Accessed: Jan 2021).
Visconti, P., Butchart, S.H., Brooks, T.M., Langhammer, P.F., Marnewick, D., Vergara, S., Yanosky, A., Watson, J.E., 2019. Protected area targets post-2020. Science 364 (6437), 239–241.
Wang, Y., Önal, H., 2016. Optimal design of compact and connected nature reserves for multiple species. Conserv. Biol. 30 (2), 413–424.
Wang, Y., Qin, P., Önal, H., 2022. An optimisation approach for designing wildlife corridors with ecological and spatial considerations. Methods in Ecol. Evol. 13 (5), 1042–1051.
Ward, M., Saura, S., Williams, B., Ramírez-Delgado, J.P., Arafeh-Dalmau, N., Allan, J.R., Venter, O., Dubois, G., Watson, J.E., 2020. Just ten percent of the global terrestrial protected area network is structurally connected via intact land. Nature Commun. 11, 4563.
Watson, J.E., Venter, O., Lee, J., Jones, K.R., Robinson, J.G., Possingham, H.P., Allan, J.R., 2018. Protect the last of the wild. Nature 563 (7729), 27–30.
Watts, M., Klein, C.J., Tulloch, V.J.D., Carvalho, S.B., Possingham, H.P., 2020. Software for prioritizing conservation actions based on probabilistic information. Conserv. Biol. 35 (4), 1299–1308.
Weerasena, L., Shier, D., Tonkyn, D., 2014. A hierarchical approach to designing compact ecological reserve systems. Environ. Model. Assess. 19 (5), 437–449.
Williams, J.C., 1998. Delineating protected wildlife corridors with multi-objective programming. Environ. Model. Assess. 3 (1–2), 77–86.
Williams, P., Margules, C.R., Hilbert, D.W., 2002. Data requirements and data sources for biodiversity priority area selection. J. Biosci. 27 (4), 327–338.
Williams, J.C., ReVelle, C.S., 1996. A 0–1 programming approach to delineating protected reserves. Environ. Plan. B: Plann. Des. 23 (5), 607–624.
Williams, J.C., ReVelle, C.S., 1998. Reserve assemblage of critical areas: a zero-one programming approach. European J. Oper. Res. 104 (3), 497–509.

Williams, J.C., ReVelle, C.S., Levin, S.A., 2005. Spatial attributes and reserve design models: a review. Environ. Model. Assess. 10 (3), 163–181.
Williams, J.C., Snyder, S.A., 2005. Restoring habitat corridors in fragmented landscapes using optimization and percolation models. Environ. Model. Assess. 10 (3), 239–250.
Wilson, E.O., 2016. Half-Earth: Our Planet's Fight for Life. WW Norton & Company.
Wilson, E.O., Willis, E.O., 1975. Applied biogeography. In: Cody, M.L., Diamond, J.M. (Eds.), Ecol. Evol. Communities 522–534.
Witte, J.-P.M., He, F., Groen, C.L., 2008. Grid origin affects scaling of species across spatial scales. Global Ecol. Biogeogr. 17 (3), 448–456.
Wright, J., Revelle, C., Cohon, J., 1983. A multiobjective integer programming model for the land acquisition problem. Reg. Sci. Urban Econ. 13 (1), 31–53.
Yemshanov, D., Haight, R.G., Liu, N., Rempel, R., Koch, F.H., Rodgers, A., 2022. Exploring the tradeoffs among forest planning, roads and wildlife corridors: a new approach. Optim. Lett. 16 (3), 747–788.