

Systematic Conservation Planning for Sustainable Land-use Policies: A Constrained Partitioning Approach to Reserve Selection and Design.

Dimitri Justeau-Allaire, Philippe Vismara, Philippe Birnbaum, Xavier Lorca

▶ **To cite this version:**

Dimitri Justeau-Allaire, Philippe Vismara, Philippe Birnbaum, Xavier Lorca. Systematic Conservation Planning for Sustainable Land-use Policies: A Constrained Partitioning Approach to Reserve Selection and Design.. IJCAI 2019 - 28th International Joint Conference on Artificial Intelligence, Aug 2019, Macao, China. pp.5902-5908, 10.24963/ijcai.2019/818 . hal-02283036

HAL Id: hal-02283036

https://hal.umontpellier.fr/hal-02283036

Submitted on 10 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Systematic Conservation Planning for Sustainable Land-use Policies: A Constrained Partitioning Approach to Reserve Selection and Design

Dimitri Justeau-Allaire[1,2,3*], **Philippe Vismara**[4,5], **Philippe Birnbaum**[1,2,3] and **Xavier Lorca**[6]
[1]CIRAD, UMR AMAP, Montpellier, France
[2]Institut Agronomique néo-Calédonien (IAC), Noumea, New Caledonia
[3]AMAP, Univ Montpellier, CIRAD, CNRS, INRA, IRD, Montpellier, France
[4]MISTEA, Montpellier SupAgro, INRA, Univ Montpellier, Montpellier, France
[5]LIRMM, Univ Montpellier, CNRS, Montpellier, France
[6]ORKID, Centre de Génie Industriel, IMT Mines Albi, Albi, France
{dimitri.justeau-allaire, philippe.birnbaum}@cirad.fr,
philippe.vismara@supagro.fr, xavier.lorca@mines-albi.fr

Abstract

Faced with natural habitat degradation, fragmentation, and destruction, it is a major challenge for environmental managers to implement sustainable land use policies promoting socioeconomic development and natural habitat conservation in a balanced way. Relying on artificial intelligence and operational research, reserve selection and design models can be of assistance. This paper introduces a partitioning approach based on Constraint Programming (CP) for the reserve selection and design problem, dealing with both coverage and complex spatial constraints. Moreover, it introduces the first CP formulation of the buffer zone constraint, which can be reused to compose more complex spatial constraints. This approach has been evaluated in a real-world dataset addressing the problem of forest fragmentation in New Caledonia, a biodiversity hotspot where managers are gaining interest in integrating these methods into their decisional processes. Through several scenarios, it showed expressiveness, flexibility, and ability to quickly find solutions to complex questions.

1 Introduction

In the context of the global biodiversity crisis, it is urgent to strengthen the conservation of natural habitats through the establishment of nature reserves [Prendergast *et al.*, 1993; Haddad *et al.*, 2015]. Accordingly, two of the United Nations Sustainable goals have been focused on the conservation of marine and terrestrial habitats. However, to be efficient the selection and the design of reserves must be based on a systematic approach [Margules and Pressey, 2000], and sustainable land use policies must promote socioeconomic development and nature conservation in a balanced way. On top of that, the design of protected buffer zones surrounding sensitive areas is an important aspect, as it can mitigate negative edge-effects and contribute to reducing fragmentation, by fostering recolonization and habitat restoration [Harris, 1988; Fahrig, 2003]. Buffer zones are, for instance, an important element of UNESCO World Heritage Sites [Feilden and Jokilehto, 1998] and Man and the Biosphere reserves [Batisse, 1982]. More recently, the IUCN protected areas management categories provided a more comprehensive framework by promoting the partitioning of the space into several levels of protection, which can be nested [Dudley, 2008].

In conservation biology, this concern lies in the framework of systematic conservation planning and formalizes as the reserve selection and design problem, which also lies in the more recent framework of computational sustainability [Gomes, 2009]. The reserve selection and design problem aims at partitioning the geographical space into at least two regions: one dedicated to habitats and biodiversity conservation, the other for socioeconomic development. However, effective strategies usually involve more regions with several nested levels of protection. Each region is defined by a combination of coverage and spatial constraints, and some other constraints such as the buffer zone, that can involve several regions. Finally, optimization objectives can be defined, such as minimizing the cost of a region or maximizing the coverage of certain features. Figure 1 depicts an example with a grid partitioned into three regions.

Many reserve selection and design models have been devised, e.g. [ReVelle *et al.*, 2002; Williams *et al.*, 2005; Sarkar, 2012; Dilkina *et al.*, 2017]. Usually, these meth-

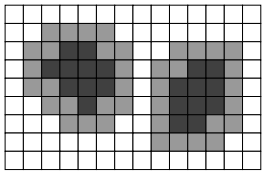


Figure 1: Grid partition with a core area (dark gray), a buffer (light gray), and an out-reserve area (white). The core and the buffer are both composed of two connected components.

*Contact Author

ods rely on ad-hoc heuristics, metaheuristics, or mixed-integer linear programming (MILP). More recently, CP models for and reserve selection and design has been published [Bessière *et al.*, 2015; Justeau-Allaire *et al.*, 2018]. While the first is focused on wildlife corridor design, the second is, to the best of our knowledge, the only generic CP model for this problem which combines both covering and spatial aspects. Sadly, in this model the search is focused on a single region, the reserve. As a consequence, socioeconomic constraints cannot be expressed on the out-reserve area. In addition, no more than two regions can be defined, and the buffer zone constraint is lacking. As a matter of fact, this constraint has, to our knowledge, only been modeled by some MILP approaches [Williams *et al.*, 2005; Billionnet, 2013] in a local fashion for three-regions configurations (core area, buffer zone and out-reserve area). This approach has some limitations since it does not account for the reciprocity between the regions (e.g. a buffer zone can exist without a core area).

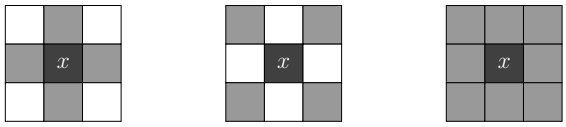
The current models are limited in their flexibility, as each address a specific subset of variants of the general problem. Because these subsets are usually different, it is also difficult to provide systematic model comparisons. For instance, to the best of our knowledge no existing MILP model is able to tackle problems with more than three regions. Some heuristics and metaheuristics are able to, e.g. Marxan with zones [Watts *et al.*, 2009], but they don't provide strict and explicit control over spatial attributes. Finally, no existing model implements a complete buffer zone constraint. However, it is of great interest for managers to have the possibility of seamlessly considering those aspects.

Although rarely used in this context, CP is a good candidate for devising a more generic model for reserve selection and design. In fact, CP provides both flexibility and expressiveness, as it allows seamless integration of complex and heterogeneous constraints into a single model. Moreover, CP is a complete approach that is able to provide satisfiability and optimality proofs. In this paper, we show how to encode a CP model that allows the definition of an arbitrary number of regions, on which any constraint can be seamlessly applied. On top of that, we provide a complete and generic formulation of the buffer zone constraint, which can be reused to compose more complex spatial constraints. Finally, we experiment our model on a real-world dataset addressing the problem of forest fragmentation in the south of New Caledonia, a biodiversity hotspot located in the South Pacific (this dataset was already used in [Justeau-Allaire *et al.*, 2018]). Through this use case, we show that the genericity provided by our model allows addressing problems that were not possible to tackle until now.

2 Problem Description

2.1 The Grid

The problem applies in a discretized geographical space. Several types of tessellation are possible (e.g. square grid, hexagonal grid, irregular grid). In accordance to available data, we only consider the regular square grid. However, the methods described in this paper can easily be transposed to other types of tessellation. In the context of reserve selection and design, a grid cell is called a *site* (the terms *parcel* and *planning unit* are also used in the literature).



(a) $\Gamma_4(x)$ (b) $\Gamma_D(x)$ (c) $\Gamma_8(x)$

Figure 2: Some neighborhood definitions (in light gray).

Definition 1 (2D regular square grid)**.** A $M \times N$ regular square grid $\mathcal{S}$ is a tessellation of the 2-dimensional space into $|\mathcal{S}| = M \times N$ unit squares (called *sites* in our context), M and N are respectively the number of rows and columns. A site is uniquely identified by its zero-based matrix coordinates x_i (row) and x_j (column) or by its flattened index $x = Nx_i + x_j$ (this indexing is independent of the tessellation).

Let $\mathcal{S}$ be a regular square grid.

Definition 2 (ω-connected neighborhood)**.** We denote by *ω-connected neighborhood* any function $\Gamma_\omega : \mathcal{S} \mapsto \mathcal{P}(\mathcal{S})$ (with $\mathcal{P}$ the power set) that associates to $x \in \mathcal{S}$ a set $\Gamma_\omega(x) \subseteq \mathcal{S}$ representing the neighbors of x, according to the label ω.

Given that, we derive some specific neighborhoods that will be useful in the rest of the paper (some of them are illustrated in Figure 2), $\forall x \in \mathcal{S}$, with $x \equiv (x_i, x_j)$:

$$\Gamma_H(x) = \{(x_i, x_j \pm 1)\} \cap \mathcal{S};$$
$$\Gamma_V(x) = \{(x_i \pm 1, x_j)\} \cap \mathcal{S};$$
$$\Gamma_D(x) = \{(x_i \pm 1, x_j \pm 1)\} \cap \mathcal{S};$$
$$\Gamma_4(x) = \Gamma_H(x) \cup \Gamma_V(x);$$
$$\Gamma_8(x) = \Gamma_4(x) \cup \Gamma_D(x).$$

Definition 3 (Connected component)**.** A set $cc \subseteq \mathcal{S}$ of connected sites (according to a neighborhood definition) is called a *connected component* (abbreviated CC). *Note:* This is an extension of the classical graph definition of connected component. If the neighborhood is not symmetric (e.g. representing environmental flows), the definition of *strongly connected component* must be used.

Definition 4 (Region)**.** A set $R \subseteq \mathcal{S}$ associated to a single land-use policy is called a *region*. A non-empty region is composed of one or several CCs. The set of CCs associated to a region R is denoted by $cc(R)$.

2.2 The Features

In the context of reserve selection and design, a feature corresponds to a characteristic of the geographical space that can be spatially represented with a numerical value for each site. A feature can represent biodiversity (e.g. species, habitats), or socioeconomic values (e.g. exploitable land, customary area). Three data types can describe a feature: binary data (e.g. presence of exploitable land), quantitative data (e.g. species abundance) and probabilistic data (e.g. species distribution model, SDM, representing either a probability of presence or a habitat suitability index). We denote the value associated to a feature f in the site x by $v_f(x)$.

2.3 Towards a Partitioning Formulation

In [Justeau-Allaire *et al.*, 2018], the problem is stated as follows: given a $M \times N$ grid $\mathcal{S}$, find $R \subseteq \mathcal{S}$ such that a set of constraints $\mathcal{C}$ are satisfied by R. In this formulation, the constraints are organized into two categories, coverage and spatial constraint. They can be formalized as follows.

Coverage Constraints

Let R be a region and $\mathcal{F}$ a set of features:

Constraint A (Covered features)**.** R is a cover of $\mathcal{F}$ if every feature of $\mathcal{F}$ is present in at least one site of R. In this context, a feature f is considered to be covered by a site x if and only if $v_f(x) \geq 1$ that is $\forall f \in \mathcal{F}, \exists x \in R, v_f(x) \geq 1$.

Constraint B (α-covered features)**.** The constraint holds if and only if every feature of $\mathcal{F}$ has a probability of at least α to lie in R: $\forall f \in \mathcal{F}, \prod_{x \in \mathcal{S}}(1 - v_f(x)) \leq 1 - \alpha$.

Constraint C (k-redundant features)**.** The constraint holds if and only if every feature of $\mathcal{F}$ is present in at least k site of R: $\forall f \in \mathcal{F}, \exists \mathcal{X} \subseteq R, |\mathcal{X}| \geq k \wedge \forall x \in \mathcal{X}, v_f(x) \geq 1$.

Spatial Constraints

Let R be a region:

Constraint D (Number of CCs, aka Number of reserves)**.** The constraint holds if and only if the number of CCs of R is bounded: $minNbCC \leq |cc(R)| \leq maxNbCC$.

Constraint E (Region size, aka Reserve System Area)**.** The constraint holds if and only if the size of the region is bounded: $minSize \leq |R| \leq maxSize$.

Constraint F (CCs size, aka Reserve areas)**.** The constraint holds if and only if the smallest (respectively largest) CC of R contains at least $minSizeCC$ (respectively $maxSizeCC$) sites: $\forall C \in cc(R), minSizeCC \leq |C| \leq maxSizeCC$.

We suggest here a more generic formulation of the problem: given a $M \times N$ grid $\mathcal{S}$, find a partitioning of $\mathcal{S}$ into n regions $\{R_0, ..., R_{n-1}\}$ such that each region R_u satisfies a given set of constraints $\mathcal{C}_u \subseteq \{A, \ldots, F\}$. Using this formulation, any constraint in the previous catalog can be seamlessly applied to any region.

3 A Generic CP Model

In this section we introduce a generic CP model associated with the partitioning formulation of the reserve selection and design problem.

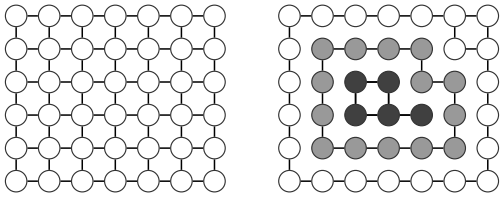
3.1 The Base Model

Decision Variables

To each site $x \in \mathcal{S}$ we associate an integer variable $\rho_x \in [0, n[$. If x lies in R_u then $\rho_x = u$. An instantiation of these variables defines de facto a partitioning of the grid into (at most) n regions: $\forall x \in \mathcal{S}, \rho_x \in [0, n[$.

Set Variables

Set variables are an abstraction providing an efficient, expressive and compact way to solve combinatorial problems through set-based modeling. The domain of a set variable X is a set interval $[\underline{X}, \overline{X}]$, with $\underline{X}$ and $\overline{X}$ two sets (respectively the lower and upper bounds). Given that, an instantiation of X is a subset of $\overline{X}$, such that $\underline{X}$ is a subset of X [Gervet, 1995]: $X \in [\underline{X}, \overline{X}] \Leftrightarrow \underline{X} \subseteq X \subseteq \overline{X}$.



(a) $G_{\mathcal{S}}$. (b) $\{G_0, G_1, G_2\}$

Figure 3: $G_{\mathcal{S}}$ associated with a 6×7 regular square grid (left) and $\{G_0, G_1, G_2\}$ associated with a 3-regions partitioning of a 6×7 regular square grid (right).

Each region is represented by a set variable $R_u \in [\varnothing, P(\mathcal{S})]$ that is channeled with the decision variables such that $\rho_x = u$ if and only if $x \in R_u$. This channeling ensures that the sets are all disjoint and that they form a partition of $\mathcal{S}$: $\forall u \in [0, n[, R_u \in [\varnothing, \mathcal{S}], \rho_x = u \Leftrightarrow x \in R_u$.

Graph Variables

Similarly to set variables, graph variables are an abstraction providing an efficient and expressive way to model combinatorial problems with graphs. A graph variable G is defined by a graph interval $[\underline{G}, \overline{G}]$ (with $\underline{G}$ and $\overline{G}$ two graphs, respectively the lower and upper bounds), such that an instantiation of G is a subgraph of $\overline{G}$ and $\underline{G}$ is a subgraph of G [Dooms, 2006; Fages, 2014]: $G \in [\underline{G}, \overline{G}] \Leftrightarrow \underline{G} \subseteq G \subseteq \overline{G}$.

First, we define the full spatial graph $G_{\mathcal{S}} = (\mathcal{S}, E_{\mathcal{S}})$, where a vertex is associated with each site of the grid $\mathcal{S}$, and such that there is an edge between two vertices if and only if they are adjacent in the grid, $E_{\mathcal{S}} = \{(x, y) \mid y \in \Gamma_4(x)\}$. In the scope of this paper we represent adjacency with the four-connected neighborhood Γ_4, but any other definition could be used. An illustration of $G_{\mathcal{S}}$ is provided in Figure 3. Then, similarly to the model defined in [Justeau-Allaire *et al.*, 2018], to each region R_u we associate a graph variable $G_u = (R_u, E_u)$. G_u is the subgraph of $G_{\mathcal{S}}$ induced by R_u. These variables will be used to define connectivity and size constraints on the regions and their CCs. Each graph G_u has the empty graph as lower bound, and the full spatial graph $G_{\mathcal{S}}$ as upper bound. Formally, for all $u \in [0, n[$, $G_u = (R_u, E_u) \subseteq G_{\mathcal{S}}$, such that $E_u = \{(x, y) \mid (x, y) \in {R_u}^2 \wedge y \in \Gamma_4(x)\}$. An illustration of $\{G_0, G_1, G_2\}$ for a 3-regions partitioning is provided in Figure 3.

User Constraints

Finally, any constraint defined in 2.3 can be seamlessly applied to any region. For more details on how to apply them, refer to [Justeau-Allaire *et al.*, 2018]. On top of that, the genericity provided by the partitioning perspective as well as the expressiveness provided by set and graph variables allows the modeling of more complex constraints. We detail, in the following, the buffer zone constraint.

3.2 The Buffer Zone Constraint

As mentioned in the introduction, the buffer zone constraint is of great interest for managers. Here, we provide a set-based

generic formulation of the constraint. A buffer zone is an area separating the periphery of two areas. We first introduce the notion of generalized neighborhood.

Definition 5 (Generalized neighborhood)**.** Let R be a region, and Γ_ω a neighborhood definition. The generalized neighborhood $\Gamma_\omega(R)$ of R is the union of the neighborhood of every site in R: $\Gamma_\omega(R) = \bigcup_{x \in R} \Gamma_\omega(x)$.

Constraint G (Buffer zone constraint)**.** Let Γ_ω be a neighborhood definition, R_u and R_v be two regions, and B a third region intended to be a buffer zone between R_u and R_v. The buffer zone constraint $\mathit{buffer}[\Gamma_\omega](R_u, R_v, B)$ holds if and only if:

$$\begin{aligned} &\Gamma_\omega(R_u) \cap R_v = \varnothing; \\ &R_u \cap \Gamma_\omega(R_v) = \varnothing; \\ &B = \Gamma_\omega(R_u) \cap \Gamma_\omega(R_v). \end{aligned} \tag{1}$$

Consistency of the Buffer Zone Constraint

We now study the consistency of the buffer zone constraint. To this end, we rely on the definitions and results introduced by [Walsh, 2003] on set/multiset constraints. In particular, we rely on the definition of bound consistency (BC) and on the following one: a constraint decomposition is a normal form if and only if decomposing constraints are at most ternary and of the form $X \subseteq Y, X = Y \cup Z, X = Y \cap Z, X = Y - Z, X \neq Y, |X| = I, \mathit{occ}(I, X) = m$ or $\mathit{occ}(m, X) = I$, where X, Y and Z are set/multiset variables, I an integer variable and m an integer. The constraint $\mathit{occ}(m, X) = I$ (respectively $\mathit{occ}(I, X) = m$) holds if and only if I (respectively m) equals the number of occurrences of m (respectively I) in X.

Proposition 1. *The constraint $N = \Gamma_\omega(R)$ (with N and R set variables) can be decomposed into a normal form.*

Proof. The decomposition is based on additional variables N_x^R, for each site x:

$$N = \Gamma_\omega(R) \Leftrightarrow N = \bigcup_{x \in S} N_x^R \text{ with } N_x^R = \begin{cases} \Gamma_\omega(x) & \text{if } x \in R \\ \varnothing & \text{otherwise} \end{cases} \tag{2}$$

These additional variables N_x^R are constrained by the following decomposition:

$$\begin{aligned} \forall x \in \mathcal{S}, \quad & N_x^R \in \{\varnothing, \Gamma_\omega(x)\} \\ & B_x = \mathit{occ}(x, R) \\ & I_x = |\Gamma_\omega(x)| B_x \\ & |N_x^R| = I_x \end{aligned} \tag{3}$$

Then, $\bigcup_{x \in \mathcal{S}} N_x^R$ can be decomposed into a normal form with $|\mathcal{S}| - 1$ ternary union constraints and $|\mathcal{S}| - 2$ intermediary set variables: $M_1 = N_0^R \cup N_1^R \wedge M_2 = M_1 \cup N_2^R \wedge ... \wedge N = M_{|\mathcal{S}|-2} \cup N_{|\mathcal{S}|-1}^R$. □

Corollary 1. *BC on $N = \Gamma_\omega(R)$ is equivalent to BC on the decomposed normal form.*

Proof. It is straightforward to decompose $N = \Gamma_\omega(R)$ into a normal form by Proposition 1. Moreover, $N = \Gamma_\omega(R)$ does not contain a repeated occurrence of variables, thus, BC on $N = \Gamma_\omega(R)$ is equivalent to BC on the decomposed normal form (see [Walsh, 2003]). □

Corollary 2. *The constraint $\mathit{buffer}[\Gamma_\omega](R_u, R_v, B)$ can be decomposed into a normal form. Consequently, BC on $\mathit{buffer}[\Gamma_\omega](R_u, R_v, B)$ is equivalent to BC on the decomposed normal form.*

Proof. The *buffer* constraint as defined in (1) can be decomposed as following:

$$\begin{aligned} &N_u = \Gamma_\omega(R_u);\ N_v = \Gamma_\omega(R_v); \\ &N_u \cap R_v = \varnothing;\ R_u \cap N_v = \varnothing; \\ &B = N_u \cap N_v. \end{aligned} \tag{4}$$

Then, Corollary 1's exact same reasoning applies. □

Time Complexity of the Buffer Zone Constraint

Finally, we study the time complexity of the buffer zone constraint filtering. Once more, we rely on [Walsh, 2003] which provides both filtering rules to enforce BC on the decomposed normal form of the constraint, and the worst-case time complexity associated with such a filtering: it is at most $O(enm^2)$ where e is the number of constraints, n the number of variables and m the maximum cardinality of the set variables.

Proposition 2. *Enforcing Bound Consistency on $\mathit{buffer}[\Gamma_\omega](R_u, R_v, B)$ can be done in $O(|S|^4)$.*

Proof. According to the proofs of Proposition 1 and Corollary 2, the number of constraints and the number of additional variables in the normal form are in $O(|S|)$. Since the cardinality of all the set variables is bounded by $|S|$, the complexity is in $O(|S|^4)$. □

3.3 Extending the Model

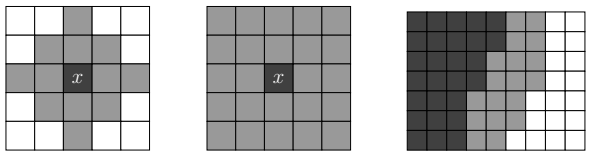
We now consider a set of operational scenarios to demonstrate how expressive our CP model is, and how it can be extended and adapted to complex requirements.

Width of the Buffer Zone

The *buffer* constraint allows great control over the spatial attributes of the buffer zone through the neighborhood definition. A good example consists in controlling the width of the buffer zone. To do so, we introduce an alternative version of any neighborhood which integrates the notion of width, as illustrated in Figure 4.

Definition 6 (k-wide neighborhood)**.** Let $\mathcal{S}$ be a regular square grid and Γ_ω a neighborhood definition. The k-wide neighborhood of a site $x \in \mathcal{S}$, denoted by $\Gamma_\omega^k(x)$, is defined by the following recursion:

$$\begin{aligned} &\Gamma_\omega^1(x) = \Gamma_\omega(x); \\ &\Gamma_\omega^k(x) = \left(\Gamma_\omega^{k-1}(x) \cup \Gamma_\omega\left(\Gamma_\omega^{k-1}(x)\right)\right) \setminus \{x\}. \end{aligned}$$



(a) $\Gamma_4^2(x)$ (b) $\Gamma_8^2(x)$ (c) 2-wide buffer zone

Figure 4: k-wide neighborhood examples (in light gray). Example of a 2-wide buffer zone.

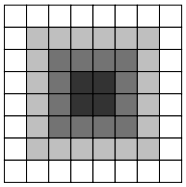


Figure 5: Nested regions example, 4 nested regions, $\Gamma_\alpha = \Gamma_8$.

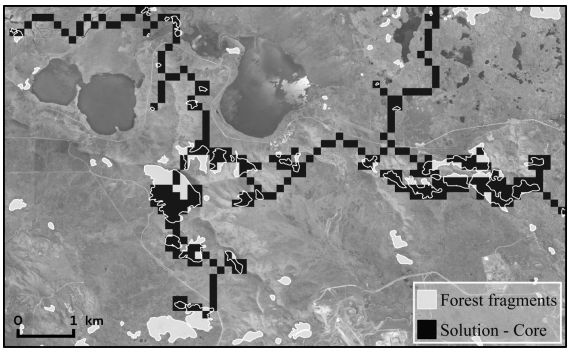
Nesting Several Protection Levels

Another interesting application of the buffer zone constraint is the spatial nesting of regions representing several levels of protection. Such a configuration, thereby, can be desired to design a landscape where the level of protection increases gradually as the habitat gets more sensitive, as illustrated in Figure 5. The IUCN protected area management categories provide good guidelines to define the policies associated with such configurations [Dudley, 2008]. In our CP model, expressing a nesting is easy. Let Γ_ω be a neighborhood definition, $\{R_0, ..., R_{n-1}\}$ a partition, and $\{R_0, ..., R_{m-1}\}$ be the regions to be nested. We assume that R_0 is the core region and R_{m-1} the periphery region. The nesting can be expressed as follows: $\forall i \in [0, m[, \mathit{buffer}[\Gamma_\omega](R_i, \mathcal{S} \setminus (R_i \cup R_{i+1}), R_{i+1})$. The term $\mathcal{S} \setminus (R_i \cup R_{i+1})$ represents the area outside R_i and R_{i+1}, the latter thus being a buffer separating R_i from this outside area, as Russian nested dolls.

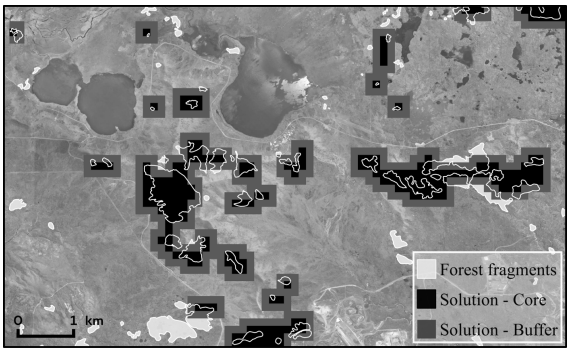
4 Use Case

New Caledonia is a large archipelago located in the South Pacific and the smallest biodiversity hotspot in the world. New Caledonian terrestrial flora, notably, is distinguished by a high rate of endemism (one of the highest in the world), and the presence of a large number of ancient lineages. However, New Caledonian forests are, as are most tropical forests, endangered with surface loss and fragmentation. Regarding this, Ibanez *et al.* [2017] conducted a study in a 60km^2 sensitive area located in the South of the main Island, "Grande Terre". In this section, we rely on the dataset from this study and the Species Distribution Models (SDMs) produced from [Pouteau *et al.*, 2015; Schmitt *et al.*, 2017]. The area is tessellated into a 46×75 square grid (a site is 1.7ha) were 5431 trees were identified among 97 communities, over 88 forest fragments. This area harbors 223 tree species for which 173 SDMs were produced. The 50 species without SDM were arbitrarily considered as endangered. According to the recent literature, this problem can be considered as a large one [Wang *et al.*, 2018]. Our model was implemented[1] and ran with Choco-solver and its Choco-graph extension (for graph variables) [Prud'homme *et al.*, 2017], on a Linux laptop with (Intel Core i5-5200U CPU 2.20GHz×4, 8GB RAM). We ran optimization scenarios under a time limit of 4h, focusing on the ability to find solutions for many constraint configurations rather than optimality proof. Numbers of solutions found and solving times for the first and best solutions found are provided in Table 1 (none were proven optimal).

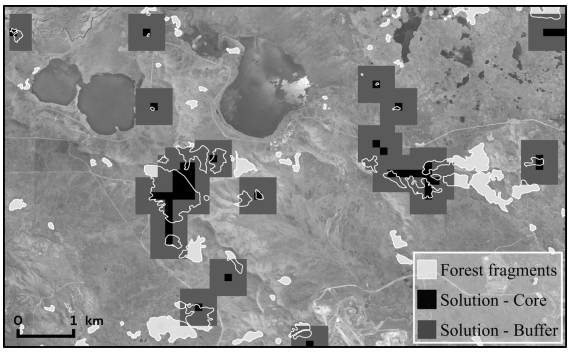
[1]The source code is available on GitHub as an open source project: https://github.com/dimitri-justeau/choco-reserve.



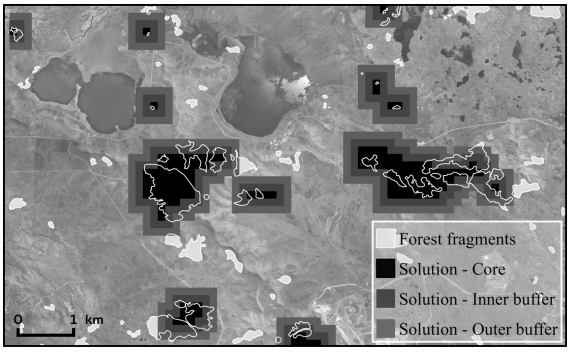
(a) Mapping of SC1's best solution.



(b) Mapping of SC2's best solution.



(c) Mapping of SC3's best solution.



(d) Mapping of SC4's best solution.

Figure 6: Use case scenarios best solutions mappings.

4.1 Original Scenario (SC1)

First, we show that our model is able to tackle the use case as defined by Justeau-Allaire *et al.*. This use case consists in highlighting a partitioning of the study area into two regions: R_0 the reserve and R_1 the out-reserve area. The reserve is the only constrained region: it must be composed of at most two CCs (Constraint D) with a surface area of at least 300ha (Constraint F). The total area of the reserve must not exceed 1000ha (Constraint E). Moreover, each endangered species must have at least one known occurrence in the reserve (Constraint A), every other species must be covered with a minimum probability of 0.8 (Constraint B), and at least 340ha of forest area must be covered (Constraint C). Finally, a set of sites that are lakes or mining areas cannot belong to the reserve. This constraint was added by restricting the domain of R_0. We tried to minimize the total size of the reserve (SC1, Figure 6a). The solving time of our approach is comparable to [Justeau-Allaire *et al.*, 2018], the best solution found within the time limit is even slightly better. If such a scenario provides useful insights, the produced delineation suffers from a major limitation: 99.3% of the selected sites are on the edge (i.e. adjacent to the out-reserve area).

4.2 Extended Scenarios (SC2 and SC3)

We show how the previous use case can be extended to a more realistic one including buffer zone constraints. To this end, we rely on Ibanez *et al.* conclusions on edge-effects and their conservation implications in the studied area: "[...] the surrounding vegetation including secondary forest at the edge and the vegetation matrix should also be protected to promote the long process of forest extension and subsequently reduce edge-effects [...]". Accordingly, we refined the use case by defining three regions: R_0 the core, R_1 the buffer zone and R_2 the out-reserve area, the protected area being $R_0 \cup R_1$. Pursuant with Ibanez *et al.* results which suggest that the spatial extent of the edge-effect applies within the first 100m to 300m, two scenarios were considered: an optimistic one (SC2, Figure 6b) with a 130–180 m wide buffer zone $\mathit{buffer}[\Gamma_8](R_0, R_2, R_1)$, and a pessimistic one (SC3, Figure 6c) with a 260–360 m wide buffer zone $\mathit{buffer}[\Gamma_8^2](R_0, R_2, R_1)$ [2]. In these scenarios, the region size constraint was applied to the protected area and became conflicting with the number of CCs constraint. Relying on conservation scientists' recommendations, we relaxed it and defined another optimization objective: minimize $|cc(R_1)|$, that is minimize the fragmentation.

In SC2, we could satisfy the constraints with 14 CCs and 37% of the protected area located in the core (R_0). In SC3, the constraints could not be satisfied, as the core cannot cover 200 sites of forest without exceeding the total area constraint. We thus relaxed the forest covering constraint to the protected area. This relaxation lead to a 10 CCs solution with only 10.6% of the protected area located in the core.

[2] Width is variable because of the square grid: diagonal is $\sqrt{2}$ larger than horizontal/vertical.

	SC1	SC2	SC3	SC4
Nb. solutions	4	3	1	47
Solving time - First found	16s	7s	9s	4s
Solving time - Best found	220s	20s	9s	1385s
Nb. sites core	277	216	51	135
Nb. sites buffer	-	370	429	192 R_1 + 262 R_2
Nb. sites total	277	586	480	589
Nb. CCs	1	14	10	9
Ratio core/total (%)	0.7%	37%	10.6%	22.6%

Table 1: Use case results characteristics.

4.3 Final Scenario (SC4)

Finally, we suggest a trade-off between SC2 and SC3 with SC4 (Figure 6d). In this final scenario we defined four regions: R_0 the core, R_1 the inner buffer, R_2 the outer buffer and R_3 the out-reserve area, the protected area being $R_0 \cup R_1 \cup R_2$. These three regions were defined as nested: $\mathit{buffer}[\Gamma_8](R_0, R_2 \cup R_3, R_1) \wedge \mathit{buffer}[\Gamma_8](R_1, R_0 \cup R_3, R_2)$. The species coverage constraints were still restricted to the core, however, the forest coverage constraint was only relaxed to the core and the inner buffer ($R_0 \cup R_1$). According to conservation scientists' feedbacks on SC3 results, we changed the optimization objective to: maximize the core area. The best solution found is composed of 9 CCs with 22.6% of the protected area located in the core, which is significantly better than SC3.

5 Conclusion

In this paper, we introduced a generic CP model that is able to tackle a high variety of reserve selection and design problems by providing high levels of flexibility and expressiveness. It is the first approach to allow the definition of an arbitrary number of regions on top of which any coverage or spatial constraint can be explicitly expressed. In addition, we provided the first CP formulation of the buffer zone constraint, which is compatible with any neighborhood definition in the tessellated geographical space and can be reused to compose more complex spatial constraints. Moreover, we provided insights on the consistency associated with the constraint, as well as on its worst-time complexity. Relying on a use case based on a real-world dataset, we showed how our model is able to support systematic conservation planning through a progressive and exploratory process. Through diverse scenarios, we highlighted useful insights for managers and conservation scientists. In particular, we showed how the buffer zone constraint can be composed to prospect more complex conservation scenarios. On top of that, our implementation showed its ability to quickly find solutions to the decision problem (cf. Table 1), demonstrating its potential for exploring many scenarios.

To conclude, our constrained partitioning approach for reserve selection and design provides the basis of a generic and exploratory decision support tool for systematic conservation planning and computational sustainability. It now remains to work closely with conservation scientists and managers to refine it and integrate it in decisional processes, in order to move towards more sustainable land-use policies. Providing proofs of optimality is, on the other hand, a prospect for technical future work.

References

[Batisse, 1982] Michel Batisse. The Biosphere Reserve: A Tool for Environmental Conservation and Management. *Environmental Conservation*, 9(2):101–111, 1982.

[Bessière *et al.*, 2015] Christian Bessière, Emmanuel Hébrard, George KATSIRELOS, and Toby Walsh. Reasoning about Connectivity Constraints. In *IJCAI: International Joint Conference on Artificial Intelligence*, pages 2568–2574, Buenos Aires, Argentina, July 2015. AAAI Press.

[Billionnet, 2013] Alain Billionnet. Mathematical optimization ideas for biodiversity conservation. *European Journal of Operational Research*, 231(3):514–534, December 2013.

[Dilkina *et al.*, 2017] Bistra Dilkina, Rachel Houtman, et al. Trade-offs and efficiencies in optimal budget-constrained multispecies corridor networks. *Conservation Biology*, 31(1):192–202, February 2017.

[Dooms, 2006] Grégoire Dooms. *The CP(Graph) Computation Domain in Constraint Programming*. PhD thesis, UCL - Université Catholique de Louvain, 2006.

[Dudley, 2008] Nigel Dudley. *Guidelines for Applying Protected Area Management Categories*. IUCN, 2008.

[Fages, 2014] Jean-Guillaume Fages. *Exploitation de structures de graphe en programmation par contraintes*. PhD thesis, Ecole des Mines de Nantes, October 2014.

[Fahrig, 2003] Lenore Fahrig. Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):487–515, 2003.

[Feilden and Jokilehto, 1998] Berbard M. Feilden and Jukka Jokilehto. *Management Guidelines for World Cultural Heritage Sites*. ICCROM. Second Edition. Rome, 1998.

[Gervet, 1995] Carmen Gervet. *Set Intervals in Constraint Logic Programming: Definition and Implementation of a Language*. PhD thesis, Université de Franche Comté Besançon, September 1995.

[Gomes, 2009] Carla P. Gomes. Challenges for Constraint Reasoning and Optimization in Computational Sustainability. In Ian P. Gent, editor, *Principles and Practice of Constraint Programming - CP 2009*, Lecture Notes in Computer Science, pages 2–4. Springer Berlin Heidelberg, 2009.

[Haddad *et al.*, 2015] Nick M. Haddad, Lars A. Brudvig, et al. Habitat fragmentation and its lasting impact on Earths ecosystems. *Science Advances*, 1(2):e1500052, March 2015.

[Harris, 1988] Larry D. Harris. Edge Effects and Conservation of Biotic Diversity. *Conservation Biology*, 2(4):330–332, 1988.

[Ibanez *et al.*, 2017] Thomas Ibanez, Vanessa Hequet, et al. How does forest fragmentation affect tree communities? A critical case study in the biodiversity hotspot of New Caledonia. *Landscape Ecology*, 32(8):1671–1687, August 2017.

[Justeau-Allaire *et al.*, 2018] Dimitri Justeau-Allaire, Philippe Birnbaum, and Xavier Lorca. Unifying Reserve Design Strategies with Graph Theory and Constraint Programming. In John Hooker, editor, *Principles and Practice of Constraint Programming*, Lecture Notes in Computer Science, pages 507–523. Springer International Publishing, 2018.

[Margules and Pressey, 2000] Chris R. Margules and Robert L. Pressey. Systematic conservation planning. *Nature*, 405(6783):243–253, May 2000.

[Pouteau *et al.*, 2015] Robin Pouteau, Élise Bayle, et al. Accounting for the indirect area effect in stacked species distribution models to map species richness in a montane biodiversity hotspot. *Diversity and Distributions*, 21(11):1329–1338, 2015.

[Prendergast *et al.*, 1993] John R. Prendergast, Rachel M. Quinn, et al. Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature*, 365(6444):335–337, September 1993.

[Prud'homme *et al.*, 2017] Charles Prud'homme, Jean-Guillaume Fages, and Xavier Lorca. *Choco Documentation*. 2017.

[ReVelle *et al.*, 2002] Charles S. ReVelle, Justin C. Williams, and John J. Boland. Counterpart Models in Facility Location Science and Reserve Selection Science. *Environmental Modeling & Assessment*, 7(2):71–80, June 2002.

[Sarkar, 2012] Sahotra Sarkar. Complementarity and the selection of nature reserves: Algorithms and the origins of conservation planning, 1980–1995. *Archive for History of Exact Sciences*, 66(4):397–426, July 2012.

[Schmitt *et al.*, 2017] Sylvain Schmitt, Robin Pouteau, et al. Ssdm: An r package to predict distribution of species richness and composition based on stacked species distribution models. *Methods in Ecology and Evolution*, 8(12):1795–1803, August 2017.

[Walsh, 2003] Toby Walsh. Consistency and Propagation with Multiset Constraints: A Formal Viewpoint. In Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, and Francesca Rossi, editors, *Principles and Practice of Constraint Programming – CP 2003*, volume 2833, pages 724–738. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[Wang *et al.*, 2018] Yicheng Wang, Hayri Önal, and Qiaoling Fang. How large spatially-explicit optimal reserve design models can we solve now? An exploration of current models' computational efficiency. *Nature Conservation*, 27:17–34, June 2018.

[Watts *et al.*, 2009] Matthew E. Watts, Ian R. Ball, et al. Marxan with Zones: Software for optimal conservation based land- and sea-use zoning. *Environmental Modelling & Software*, 24(12):1513–1521, December 2009.

[Williams *et al.*, 2005] Justin C. Williams, Charles S. ReVelle, and Simon A. Levin. Spatial attributes and reserve design models: A review. *Environmental Modeling & Assessment*, 10(3):163–181, September 2005.