

Partitionnement de l'espace sous contraintes : un modèle générique et expressif pour la planification de la conservation.

Dimitri Justeau-Allaire^{1,2,3}, Philippe Vismara^{4,5}, Philippe Birnbaum^{1,2,3}, Xavier Lorca⁶

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
{dimitri.justeau-allaire, philippe.birnbaum}@cirad.fr

⁴ MISTEA, Montpellier SupAgro, INRA, Univ Montpellier, Montpellier, France

⁵ LIRMM, Univ Montpellier, CNRS, Montpellier, France
philippe.vismara@supagro.fr

⁶ ORKID, Centre de Génie Industriel, IMT Mines Albi, Albi, France
xavier.lorca@mines-albi.fr

Mots-clés : *planification de la conservation, programmation par contraintes, intelligence artificielle, écologie, Nouvelle-Calédonie.*

1 Introduction

À l'ère de l'anthropocène et dans un contexte de crise globale de la biodiversité, de nombreuses espèces souffrent de la dégradation, de la fragmentation et de la disparition de leurs habitats naturels. Alors que les activités humaines constituent la cause principale de ce phénomène, la mise en place de politiques d'utilisation des sols durables apparaît comme un des enjeux majeurs de notre époque. Afin d'être efficaces, ces politiques doivent promouvoir le développement socioéconomique de manière équilibrée avec la conservation des habitats naturels, et être construites à partir d'une approche rationnelle et systématique [11]. En biologie de la conservation, on parle de planification de la conservation [9]. Concrètement, il s'agit de partitionner une zone géographique donnée en un ensemble de régions (cf. Figure 1 pour un exemple de partitionnement). Chaque région correspond à une politique d'utilisation des sols : tandis que certaines sont dédiées à la conservation de la nature (e.g. réserve naturelle), d'autres répondent à des besoins socioéconomiques (e.g. zone agricole). Avec ce modèle, il est possible de traduire une politique d'utilisation des sols (i.e. une région) en un ensemble de contraintes. Ces contraintes peuvent concerner les caractéristiques couvertes par une région (e.g. présence d'un ensemble d'espèces, abondance d'une communauté), ou la configuration spatiale d'une région ou d'un ensemble de régions entre elles (e.g. la région doit former un ensemble connecté). Dans un article publié dans le cadre de la conférence IJCAI 2019 (*International Joint Conference on Artificial Intelligence*), nous avons proposé un modèle de programmation par contraintes (CP) qui met l'accent sur la généricité et l'expressivité, en complément aux approches existantes [8]. Ce modèle permet de définir un nombre arbitraire de régions, sur lesquelles un nombre arbitraire de contraintes peuvent être appliquées depuis un catalogue. Nous avons également proposé une formulation complète et générique de la contrainte de zone tampon, qui peut être utilisée pour composer des contraintes spatiales plus complexes. Dans cet article, nous résumons ce modèle et les résultats obtenus sur un cas d'étude basé sur des données réelles : la fragmentation des forêts sur substrat ultramafique dans le sud de la Nouvelle Calédonie.

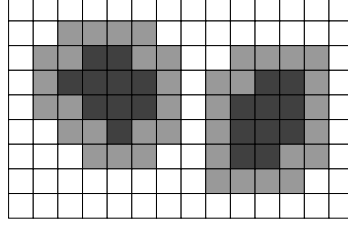


FIG. 1 – Un exemple de partitionnement sur une grille carrée, avec une zone de cœur de réserve (gris foncé), une zone tampon (gris clair), et une zone non protégée (blanc).

2 Un modèle CP générique et expressif pour la planification de la conservation

2.1 Description du problème

L'espace géographique discrétisé \mathcal{S} . Le problème s'applique sur un espace géographique discrétisé, selon une maille de *sites* qui peut prendre différentes formes (e.g. maille carrée régulière, maille hexagonale, maille irrégulière). On note cet espace géographique \mathcal{S} et on appelle *région* tout sous-ensemble $R \subseteq \mathcal{S}$. Les exemples illustrés dans cet article le sont via une maille régulière carrée, mais toutes les méthodes décrites peuvent être transposées à n'importe quel type de maille.

Voisinages dans \mathcal{S} et composantes connexes. Il est possible de définir différentes fonctions de voisinage dans la maille, qu'on note $\Gamma_\omega : \mathcal{S} \mapsto \mathcal{P}(\mathcal{S})$ (avec $\mathcal{P}(\mathcal{S})$ l'ensemble des parties de \mathcal{S}), quelques exemples sont illustrés dans la Figure 2. Par extension de la définition classique des composantes connexes pour un graphe, on définit par composante ω -connexe d'une région $R \subseteq \mathcal{S}$ tout ensemble $cc_\omega \subseteq R$ de sites connectés selon la fonction de voisinage Γ_ω . Une fonction de voisinage peut être naturellement généralisée à une région : $\Gamma_\omega(R) = \bigcup_{x \in R} \Gamma_\omega(x)$.

Caractéristiques spatiales. Une caractéristique spatiale correspond à une fonction qui peut être représentée avec une valeur numérique pour tous les sites de l'espace géographique considéré. Une caractéristique peut représenter des valeurs environnementales (e.g. altitude, substrat), de biodiversité (e.g. occurrence d'espèce, habitat naturel) ou socioéconomiques (e.g. zone urbanisée). On note $v_f(x)$ la valeur associée à la caractéristique f pour le site x .

Le problème. En utilisant les définitions précédentes, un problème générique de planification de la conservation peut s'exprimer de la manière suivante : *étant donné un espace géographique maillé \mathcal{S} , on cherche un partitionnement de \mathcal{S} en n régions $\{R_0, \dots, R_{n-1}\}$ tel que chaque région R_u satisfait un ensemble de contraintes \mathcal{C}_u .* Avec cette formulation, il devient possible de définir un nombre arbitraire de régions, sur lesquelles toutes les contraintes d'un catalogue peuvent s'appliquer. Dans la section suivante, nous présentons un catalogue de contraintes issu d'un article précédent [7], qui s'organise selon deux catégories : contraintes de couverture et contraintes spatiales. Ce catalogue qui peut facilement être étendu.

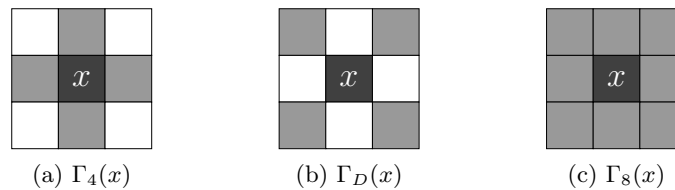


FIG. 2 – Quelques exemples de voisinage (en gris clair) : le voisinage 4-connecté, le voisinage D-connecté (les cellules voisines diagonales) et le voisinage 8-connecté.

2.2 Contraintes de couverture

Soit R une région et \mathcal{F} un ensemble de caractéristiques spatiales :

Contrainte A (Caractéristiques couvertes) R est une couverture de \mathcal{F} si chaque caractéristique de \mathcal{F} est couverte par au moins un site de R . Dans ce contexte, une caractéristique f est couverte par un site x si et seulement si $v_f(x) \geq 1$, soit $\forall f \in \mathcal{F}, \exists x \in R, v_f(x) \geq 1$.

Contrainte B (Caractéristiques α -couvertes) Cette contrainte est satisfaite si et seulement si chaque caractéristique de \mathcal{F} a une probabilité d’au moins α d’être couverte par R , soit $\forall f \in \mathcal{F}, \prod_{x \in \mathcal{S}} (1 - v_f(x)) \leq 1 - \alpha$.

Contrainte C (Caractéristiques k -redondantes) Cette contrainte est satisfaite si et seulement si chaque caractéristique de \mathcal{F} est couverte par au moins k sites distincts de R , soit $\forall f \in \mathcal{F}, \exists \mathcal{X} \subseteq R, |\mathcal{X}| \geq k \wedge \forall x \in \mathcal{X}, v_f(x) \geq 1$.

2.3 Contraintes spatiales

Soit R une région et Γ_ω un voisinage :

Contrainte D (Nombre de composantes ω -connexes) Cette contrainte est satisfaite si et seulement si le nombre de composantes ω -connexes de R est borné par $minNbCC$ et $maxNbCC$. *e.g. On cherche à délimiter une réserve naturelle connectée, ou composée au plus de deux zones connectées.*

Contrainte E (Taille de la région) Cette contrainte est satisfaite si et seulement si la taille de la région R est bornée par $minSize$ et $maxSize$. *Note : Ici la taille correspond au nombre de sites, elle peut être convertie en surface si la maille est régulière. e.g. On souhaite identifier entre 1000ha et 2000ha pour faire de la restauration écologique.*

Contrainte F (Taille des composantes ω -connexes) Cette contrainte est satisfaite si et seulement si la plus petite (respectivement plus grande) composante ω -connexe de la région R contient au moins $minSizeCC$ (respectivement $maxSizeCC$) sites. *Note : Ici la taille correspond au nombre de sites, elle peut être convertie en surface si la maille est régulière. e.g. Chaque zone connectée de la réserve à délimiter doit faire au moins 300ha afin d’assurer la pérennité des communautés qu’elle protège.*

2.4 Description du modèle CP

Le modèle s’articule autour de trois représentations de l’espace géographique : en variables entières (variables de décision), en variables ensemblistes et en variables graphe.

Variables entières de décision. À chaque site $x \in \mathcal{S}$ est associée une variable entière $\rho_x \in [0, n[$. Si x est alloué à la région R_u alors $\rho_x = u$. Une instanciation complète de ces variables définit de facto un partitionnement de l’espace en (au plus) n régions : $\forall x \in \mathcal{S}, \rho_x \in [0, n[$.

Variables ensemblistes. Les variables ensemblistes sont une abstraction offerte par CP qui permet d’exprimer de manière compacte et naturelle des problèmes combinatoires avec des ensembles. Le domaine d’une variable ensembliste X est un intervalle d’ensembles $[\underline{X}, \overline{X}]$, avec \underline{X} et \overline{X} deux ensembles (respectivement la borne inférieure et la borne supérieure). Une instanciation de X est un sous-ensemble de \overline{X} , tel que \underline{X} est un sous-ensemble de X [4]. Chaque région à délimiter est représentée par une variable $R_u \in [\emptyset, P(\mathcal{S})]$ et chaînée avec les variables de décision de manière à ce que $\rho_x = u$ si et seulement si $x \in R_u$. Cette contrainte de chaînage garantit que tous les ensembles sont disjoints et forment une partition of \mathcal{S} .

Variables graphe. Dans un esprit similaire aux variables ensemblistes, les variables graphe sont une abstraction qui permet de modéliser naturellement des problèmes avec des graphes en CP. Une variable graphe G est définie par un intervalle de graphes $[\underline{G}, \overline{G}]$ (avec \underline{G} et \overline{G} deux

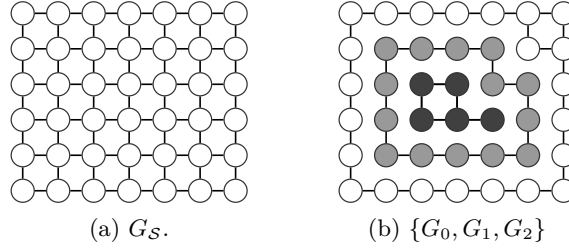


FIG. 3 – G_S associé à une grille carrée régulière 6×7 (gauche) et $\{G_0, G_1, G_2\}$ associé avec un partitionnement en 3 régions de la même grille (droite).

graphes, respectivement la borne inférieure et la borne supérieure), tel qu’une instanciation de G est un sous-graphe de \overline{G} et \underline{G} un sous-graphe de G [1, 2]. Étant donné un voisinage Γ_ω , l’espace géographique \mathcal{S} peut être représenté par un graphe $G_S = (\mathcal{S}, E_S)$, avec $E_S = \{(x, y) \mid y \in \Gamma_\omega(x)\}$. Une illustration de G_S est proposée dans la figure Figure 3, avec $\Gamma_\omega = \Gamma_4$, le voisinage 4-connecté dans une grille carrée. A chaque région est associée une variable graphe $G_u = (R_u, E_u) \in [\emptyset, G_S]$, qui est contrainte de manière à être le sous-graphe de G_S induit par R_u . Une illustration de $\{G_0, G_1, G_2\}$ est proposée en Figure 3 pour un partitionnement en 3 régions d’une grille carrée.

Contraintes utilisateur. Toutes les contraintes définies dans les sections 2.2 et 2.3 peuvent être appliquées de manière transparente à toutes les régions. Certaines contraintes s’expriment naturellement sur les variables ensemblistes (e.g. les contraintes de couverture, ou la contrainte de zone tampon qui est présentée dans la section suivante), d’autres sur les variables graphe (e.g. nombre de composantes connexes). Plus de détails sur leur implémentation sont disponibles dans les articles [7, 8].

2.5 Un modèle extensible : illustration avec la contrainte de zone tampon

Le modèle présenté ci-dessus profite d’un des principaux avantages de la CP : l’extensibilité. Nous avons illustré ces attributs en enrichissant le catalogue de contraintes utilisateur avec une contrainte d’intérêt majeur pour les gestionnaires : la contrainte de zone tampon. En effet, la délimitation de zones tampon protégées autour des habitats naturels sensibles permet de mitiger les effets de bord négatifs et d’encourager la régénération naturelle et la restauration des zones dégradées [5, 3]. En définissant une zone tampon comme l’intersection de la périphérie de deux zones non adjacentes, on peut exprimer naturellement la contrainte de zone tampon en utilisant les variables ensemblistes du modèle.

Contrainte A (Zone tampon) Soit Γ_ω un voisinage, R_u et R_v deux régions, et B une troisième région qui doit faire tampon entre R_u et R_v . La contrainte de zone tampon est satisfaite si et seulement si :

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

Exprimée ainsi, il est possible d’atteindre la consistance aux bornes (telle que définie sur les variables ensemblistes dans [14]) en $O(|S|^4)$ [8]. Nous avons également montré comment à partir de cette contrainte il est possible de contrôler la taille et la forme de la zone tampon, et comment composer celle-ci pour former la contrainte de zones imbriquées (cf. Figure 4 pour une illustration).

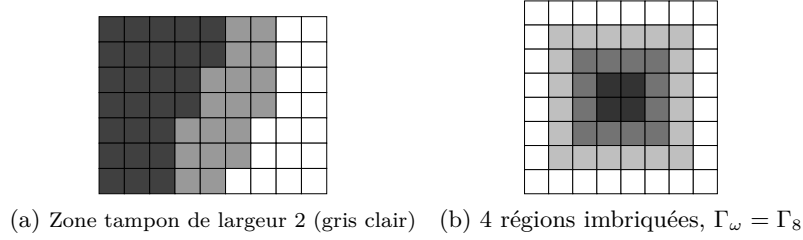


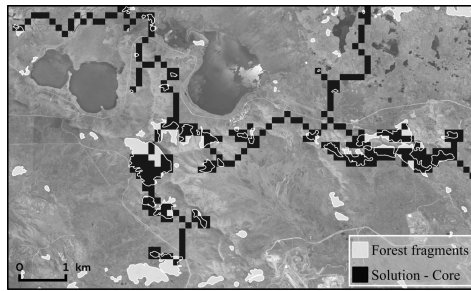
FIG. 4 – Exemples d'utilisation de la contrainte de zone tampon.

3 Cas d'étude : Fragmentation forestière en Nouvelle-Calédonie

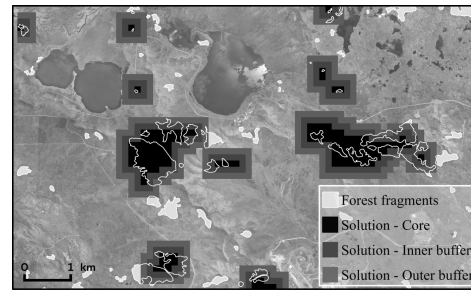
Nous avons implémenté le modèle présenté précédemment avec le solveur Choco et son extension Chocho-graph [12], afin de le valider à travers un cas d'étude basé sur des données réelles en Nouvelle-Calédonie. La Nouvelle-Calédonie est un archipel tropical situé dans le Pacifique Sud et qui se distingue par une biodiversité exceptionnelle, mais menacée : c'est un point chaud de la biodiversité. En particulier, les forêts de Nouvelle-Calédonie sont menacées par la fragmentation, dont les principales causes sont d'origine anthropique : le feu et l'extraction du nickel dans des mines à ciel ouvert. Nous nous sommes intéressés à une étude conduite par Ibanez et al. [6] dans une zone de 60km² située dans le sud de l'île principale de la Nouvelle-Calédonie, la Grande Terre. Nous avons accès à un jeu de données qui regroupe 5431 occurrences d'arbres parmi 223 espèces, identifiés via 97 inventaires répartis sur les 88 patchs forestiers de la zone. Pour 173 de ces espèces, nous avons pu disposer de modèles de distribution d'espèces (SDMs) produits par Pouteau et al. [10] et Schmitt et al. [13]. A partir de ce jeu de données, nous avons maillé la zone en une grille carrée régulière de dimension 46 × 75, puis considéré plusieurs questions de conservation fictives avec des écologues et des gestionnaires. A travers quatre scénarios, nous avons pu illustrer la capacité de notre modèle à exprimer des questions diverses, et sa capacité à y répondre rapidement. Le détail de ces scénarios et les résultats sont disponibles dans l'article original [8]. Deux des solutions obtenues sont illustrées en Figure 5.

4 Conclusion et perspectives

Dans cet article, nous avons résumés les résultats introduits dans [8], à savoir un modèle CP générique capable d'exprimer et de résoudre une grande variété de problèmes de planification de la conservation. C'est à notre connaissance la première approche permettant la définition d'un nombre abstrait de régions sur lesquelles peuvent être explicitement appliquées n'importe quelles contraintes de couverture ou spatiales. C'est aussi la première approche à intégrer une formulation CP complète de la contrainte de zone tampon, qui est compatible avec différents voisinages et peut être réutilisée pour composer des contraintes spatiales complexes. A partir d'un cas d'étude basé sur un jeu de données réel, nous avons pu illustrer le potentiel de ce modèle pour l'aide à la décision en planification de la conservation à travers un questionnaire progressif et itératif. En conclusion, nous avons montré qu'une approche basée sur le partitionnement sous contraintes de l'espace peut fournir les bases d'un outil d'aide à la décision générique pour la planification de la conservation. En mettant l'accent sur l'expressivité et la diversité des questions qui peuvent être posées, nous pensons qu'un tel outil peut faciliter la collaboration entre les modélisateurs, les écologues et les gestionnaires et contribuer à la construction de politiques d'utilisation des sols plus durables.



(a) Cartographie de la meilleure solution obtenue dans le premier scénario.



(b) Cartographie de la meilleure solution obtenue dans le dernier scénario.

FIG. 5 – Résultats obtenus pour deux des scénarios du cas d'étude.

Références

- [1] Grégoire Dooks. *The CP(Graph) Computation Domain in Constraint Programming*. PhD thesis, UCL - Université Catholique de Louvain, 2006.
- [2] Jean-Guillaume Fages. *Exploitation de structures de graphe en programmation par contraintes*. PhD thesis, Ecole des Mines de Nantes, October 2014.
- [3] Lenore Fahrig. Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 34(1) :487–515, 2003.
- [4] Carmen Gervet. *Set Intervals in Constraint Logic Programming : Definition and Implementation of a Language*. PhD thesis, Université de Franche Comté Besançon, September 1995.
- [5] Larry D. Harris. Edge Effects and Conservation of Biotic Diversity. *Conservation Biology*, 2(4) :330–332, 1988.
- [6] Thomas Ibanez, Vanessa Hequet, Céline Chambrey, Tanguy Jaffré, and Philippe Birnbaum. 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.
- [7] 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.
- [8] Dimitri Justeau-Allaire, Philippe Vismara, Philippe Birnbaum, and Xavier Lorca. Systematic Conservation Planning for Sustainable Land-use Policies : A Constrained Partitioning Approach to Reserve Selection and Design. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, pages 5902–5908, Macao, China, August 2019. International Joint Conferences on Artificial Intelligence Organization.
- [9] Chris R. Margules and Robert L. Pressey. Systematic conservation planning. *Nature*, 405(6783) :243–253, May 2000.
- [10] Robin Pouteau, Élise Bayle, Élodie Blanchard, Philippe Birnbaum, Jean-Jérôme Cassan, Vanessa Hequet, Thomas Ibanez, and Hervé Vandrot. 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.
- [11] R. L. Pressey, C. J. Humphries, C. R. Margules, R. I. Vane-Wright, and P. H. Williams. Beyond opportunism : Key principles for systematic reserve selection. *Trends in Ecology & Evolution*, 8(4) :124–128, April 1993.
- [12] Charles Prud'homme, Jean-Guillaume Fages, and Xavier Lorca. *Choco Documentation*. 2017.

- [13] Sylvain Schmitt, Robin Pouteau, Dimitri Justeau, Florian Boissieu, and Philippe Birnbaum. 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.
- [14] 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.