Planification systématique de la conservation basée sur les contraintes, une approche générique et expressive : application à l'aide à la décision pour la conservation des forêts de Nouvelle-Calédonie

Dimitri Justeau-Allaire

To cite this version:

Dimitri Justeau-Allaire. Planification systématique de la conservation basée sur les contraintes, une approche générique et expressive : application à l'aide à la décision pour la conservation des forêts de Nouvelle-Calédonie. Sciences agricoles. Université Montpellier, 2020. Français. NNT : 2020MONTG058 . tel-03329679

HAL Id: tel-03329679
https://theses.hal.science/tel-03329679
Submitted on 31 Aug 2021

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En: Planification de la Conservation

École doctorale: GAIA

Unité de recherche: AMAP

Planification systématique de la conservation basée sur les contraintes, une approche générique et expressive

Application à l'aide à la décision pour la conservation des forêts de Nouvelle-Calédonie

Présentée par Dimitri Justeau-Allaire
Le 15 décembre 2020

Sous la direction de Philippe Birnbaum et Xavier Lorca

Devant le jury composé de

Philippe BIRNBAUM, Chercheur, Cirad	Directeur
Xavier LORCA, Directeur du centre de Génie Industriel, IMT Mines Albi	Co-directeur
Ana RODRIGUES, Directrice de recherche, CNRS	Présidente du jury
François MUNOZ, Professeur, Université Grenoble-Alpes	Rapporteur
Christine SOLNON, Professeure, INSA Lyon	Rapporteur
Jean-Jérôme CASSAN, Chef de service adjoint SIEC, DDEE Province Nord NC	Examinateur
Morgan MANGEAS, Directeur de recherche, IRD	Membre invité

UNIVERSITÉ DE MONTPELLIER

CONSTRAINT-BASED SYSTEMATIC CONSERVATION PLANNING, A GENERIC AND EXPRESSIVE APPROACH

Application to decision support in the conservation of New Caledonian forests

DIMITRI JUSTEAU-ALLAIRE

Under the supervision of Philippe Birnbaum and Xavier Lorca

AMAP

cirad
LA RECHERCHE AGRONOMIQUE
POUR LE DÉVELOPPEMENT

IAC
INSTITUT AGRONOMIQUE
NÉO-CALÉDONIEN

IMT Mines Albi-Carmaux
École Mines-Télécom

UNIVERSITÉ MONTPELLIER

Ecole Doctorale GAIA
Biodiversité
Agriculture
Alimentation
Environnement
Terre
Eau

Dimitri Justeau-Allaire: *Constraint-based systematic conservation planning, a generic and expressive approach. Application to decision support in the conservation of New Caledonian forests.* © October 2020

SUPERVISORS:
Philippe Birnbaum, Researcher at CIRAD/AMAP (thesis director).
Xavier Lorca, Head of Industrial Engineering Center at IMT Mines Albi (thesis co-director).

FUNDING:
Cirad – 50% of the doctoral grant + operating budget.
IAC – 50% of the doctoral grant + operating budget.
IMT Mines Atlantique – Travel costs for IJCAI 2020 conference.
CNRS – Accommodation costs for IJCAI 2020 conference.
UM (ED GAIA) – Accomodation costs for Island Biology 2019 conference.

LOCATIONS:
Montpellier (France) and Nouméa (New Caledonia).

TIME FRAME:
2017 – 2020.

“La fin justifie les moyens ? Cela est possible. Mais qui justifiera la fin ?”
— Albert Camus, *L'Homme révolté* (1951).

Cette thèse est dédicacée à tous les arbres de cette planète, tenez-bon !

FOREWORD

This PhD thesis is the result of three years of work carried out between the AMAP research unit (botAny and Modelling of Plant Architecture and vegetation) in Montpellier and Nouméa, and the SolVeG team of the IAC (New Caledonian Institute of Agronomy).

The French agricultural research and international cooperation organization (Cirad) and the IAC co-funded the doctoral grant and the operating budget of this PhD thesis. The IMT Mines Atlantique, the CNRS (The French National Center for Scientific Research), and the University of Montpellier (through the GAIA doctoral school) also financially supported this PhD thesis through travel and accommodation costs for conferences attendance.

This is a thesis by publication, all content of Part II (excepting Chapter 4) was published or accepted in international conferences proceedings and international scientific journals. Parts I and III are original works written for this PhD thesis.

The covers of this thesis are all photos taken in New Caledonia, by ©Philippe Birnbaum (Chapter 1), ©Nicolas Rinck (Chapters 7 and 8), and ©Dimitri Justeau-Allaire (Title page, Chapters 2, 3, 4, 5, 6, and 9).

AVANT PROPOS

Cette thèse de doctorat est le résultat de trois années de travail entre l'unité de recherche AMAP (botAnique et Modélisation de l'Architecture des Plantes et des végétations) de Montpellier et Nouméa, et l'équipe SolVeG de l'IAC (Institut Agronomique néo-Calédonien).

Le Cirad (Centre de coopération internationale en recherche agronomique pour le développement) et l'IAC ont cofinancé la bourse doctorale et le budget de fonctionnement de cette thèse. L'IMT Mines Atlantique, le CNRS (Centre National de la Recherche Scientifique), et l'Université de Montpellier (à travers l'école doctorale GAIA) ont également soutenu financièrement cette thèse de doctorat via des frais de voyage et de logement pour la participation à différentes conférences.

Il s'agit d'une thèse par publication, tout le contenu de la Partie II (à l'exception du Chapitre 4) a été publié ou accepté dans des actes de conférences internationales et des revues scientifiques internationales. Les Parties I et III sont des travaux originaux rédigés pour cette thèse de doctorat.

Les couvertures de cette thèse sont toutes des photos prises en Nouvelle-Calédonie, par ©Philippe Birnbaum (Chapitre 1), ©Nicolas Rinck (Chapitres 7 et 8) et ©Dimitri Justeau-Allaire (page de titre, Chapitres 2, 3, 4, 5, 6 et 9).

ABSTRACT

In the context of the global biodiversity crisis, human activities are the principal cause of natural habitat degradation, fragmentation, and destruction. Globally, the species extinction rate has reached an unprecedented level in human history and about one million species are nowadays threatened with extinction[1]. Conservation biology is a multidisciplinary research area which attempts to address the current biodiversity crisis challenges. The development of practical approaches to promote conservation and reduce the research-implementation gap is one of its objectives. Last decades, systematic conservation planning (SCP) emerged in this direction as a framework relying on optimization and computer science research. Its main target is to provide decision support in the planning of conservation actions through the integration of ecological targets along with socio-economical constraints.

In this PhD thesis, we introduced a formal approach for modelling and solving SCP problems based on constraint programming, a method from artificial intelligence based on automated reasoning. The main motivation of this approach was to provide more expressiveness into SCP (i.e. extend the breadth and variety of problems that users can represent and solve), notably through the integration of advanced spatial constraints and landscape indices. Formal approaches are often more demanding to implement and scale up than heuristic approaches. However, they provide satisfiability and optimality guarantees on the produced solutions. The insights offered by these guarantees can substantially improve the quality of decision support.

We evaluated the methods developed in this thesis on real data from New Caledonian forests. As the smallest biodiversity hotspot in the world, New Caledonia has to struggle with many conservation challenges. Moreover, the developed, insular and low populated New Caledonian context allows high proximity between conservation stakeholders, which makes it an appropriate field of study to experiment novel approaches. We illustrated this particularity through a real case study, conducted in close collaboration with the managers of the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" provincial park. In this study, we aimed to provide decision support in a reforestation project, with an emphasis on reducing fragmentation and improving structural connectivity. Overall, we demonstrated the genericity, flexibility, and expressiveness of the constraint-based approach to SCP. Our results also opened new perspectives for decision support in New Caledonia, systematic conservation planning, and constraint programming.

---

[1] According to the last Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) report.

RÉSUMÉ

A l'ère de la crise mondiale de la biodiversité, les activités humaines sont la principale cause de dégradation, de fragmentation et de destruction des habitats naturels. À l'échelle mondiale, le taux d'extinction des espèces a atteint un niveau sans précédent, et environ un million d'espèces sont aujourd'hui menacées d'extinction[2]. La biologie de la conservation est un domaine de recherche multidisciplinaire qui tente de relever les défis de cette crise. Le développement d'approches pratiques promouvant la conservation tout en réduisant le fossé entre la recherche et l'implémentation est l'un de ses objectifs. Ces dernières décennies, la planification systématique de la conservation (PSC) a émergé dans cette direction comme un cadre reposant sur l'optimisation et l'informatique. Son principal objectif est de fournir une aide à la décision dans la planification des actions de conservation en intégrant les objectifs écologiques avec les contraintes des gestionnaires.

Dans cette thèse de doctorat, nous avons introduit une approche formelle pour modéliser et résoudre des problèmes de PSC basée sur la programmation par contraintes, une méthode issue de l'intelligence artificielle et basée sur le raisonnement automatique. La motivation principale de cette approche était d'apporter plus d'expressivité à la PSC (i.e. d'accroître l'étendue et la variété des problèmes qui peuvent être représentés et résolus), notamment par l'intégration de contraintes spatiales avancées et d'indices du paysage. Les approches exactes sont souvent plus exigeantes à implémenter et à transposer à large échelle que les approches heuristiques. Cependant, elles fournissent des garanties de satisfaction et d'optimalité sur les solutions produites qui peuvent améliorer considérablement la qualité de l'aide à la décision.

Nous avons évalué les méthodes développées dans cette thèse sur des données réelles issues des forêts de Nouvelle-Calédonie. Point chaud de la biodiversité, la Nouvelle-Calédonie doit faire face à de nombreux défis pour la conservation de sa biodiversité. De plus, le contexte développé, insulaire et peu peuplé de cet archipel permet une grande proximité entre les différents acteurs de la conservation, ce qui en fait un terrain d'étude approprié pour expérimenter de nouvelles approches. Nous avons illustré cette particularité à travers un cas d'étude mené en étroite collaboration avec les gestionnaires du parc provincial de la "Côte Oubliée - Woen Vùù - Pwa Pereeù". Dans cette étude, nous avons fourni une aide à la décision dans un projet de reforestation, en mettant l'accent sur la réduction de la fragmentation et l'amélioration de la connectivité structurelle. Dans l'ensemble, nous avons démontré le caractère générique, la flexibilité et l'expressivité de l'approche basée sur les contraintes appliquée à la PSC. Nos résultats ont également ouvert de nouvelles perspectives pour l'aide à la décision en Nouvelle-Calédonie, la PSC, et la programmation par contraintes.

---

[2] Selon le dernier rapport de la plateforme intergouvernementale scientifique et politique sur la biodiversité et les services écosystémiques (IPBES).

PUBLICATIONS

This PhD thesis has led to the following publications:

International journal articles

- Dimitri Justeau-Allaire, Ghislain Vieilledent, Nicolas Rinck, Philippe Vismara, Xavier Lorca, and Philippe Birnbaum (2020). "Constrained Optimization of Landscape Indices in Conservation Planning to Support Ecological Restoration in New Caledonia". In: *Journal of Applied Ecology*. ISSN: 1365-2664. DOI: 10.1111/1365-2664.13803.

International conference articles

- Dimitri Justeau-Allaire, Philippe Birnbaum, and Xavier Lorca (2018). "Unifying Reserve Design Strategies with Graph Theory and Constraint Programming". en. In: *Principles and Practice of Constraint Programming*. Ed. by John Hooker. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 507–523. ISBN: 978-3-319-98334-9. DOI: 10.1007/978-3-319-98334-9_33.

- Dimitri Justeau-Allaire, Philippe Vismara, Philippe Birnbaum, and Xavier Lorca (2019a). "Systematic Conservation Planning for Sustainable Land-Use Policies: A Constrained Partitioning Approach to Reserve Selection and Design." en. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*. Macao, China: International Joint Conferences on Artificial Intelligence Organization, pp. 5902–5908. ISBN: 978-0-9992411-4-1. DOI: 10.24963/ijcai.2019/818.

International conference abstracts

- Dimitri Justeau-Allaire, N. Rink, Xavier Lorca, E. Couture, and Philippe Birnbaum (2019b). "Systematic Conservation Planning in New Caledonia: Supporting Sustainable Land-Use Policies with Reserve Selection Models". In: *Island Biology 2019*, Saint-Denis, Réunion: Dominique STRASBERG and Claudine AH-PENG.

National conference articles

- Dimitri Justeau-Allaire, Philippe Birnbaum, and Xavier Lorca (June 2019c). "Unifier Les Stratégies de Sélection de Réserve Avec La Programmation Par Contraintes et Les Graphes". In: *JFPC 2019 - Actes Des*

15es Journees Francophones de Programmation Par Contraintes. Albi, France: IMT Mines Albi, p. 27–29.

- Dimitri Justeau-Allaire, Philippe Vismara, Philippe Birnbaum, and Xavier Lorca (2020). “Partitionnement de l’espace Sous Contraintes: Un Modèle Générique et Expressif Pour La Planification de La Conservation.” In: *ROADEF 2020 - Actes du 21ème congrès annuel de la société Française de Recherche Opérationnelle et d’Aide à la Décision*.

REMERCIEMENTS

Le doctorat est une formidable aventure scientifique et surtout une incroyable expérience humaine. Elle s'inscrit dans une tranche de vie avec une intimité telle que les deux sont, à mon sens, difficilement dissociables. Ainsi, à travers ces remerciements, je tiens à faire honneur à toutes les âmes qui ont marqué ces trois années et assurément contribué à leur aboutissement.

Pour commencer, je tiens à remercier l'IAC (Institut Agronomique Néo-calédonien) et le CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement) qui ont financé ce projet. Derrière ces acronymes si chers au monde de la recherche, il y a des êtres humains, qui ont remué ciel et terre en amont de ce projet afin qu'il puisse se réaliser dans de bonnes conditions. Parmi eux je remercie tout particulièrement Philippe Birnbaum, Thierry Fourcaud, Laurent l'Huilier, Daniel Barthélémy, Maguy Rodier-Goud et Valérie Buisson. Je remercie également l'ensemble de l'équipe administrative de l'UMR AMAP, qui réalise un travail d'une efficacité remarquable dans un esprit de convivialité qui l'est tout autant.

Je tiens à exprimer toute ma gratitude à Philippe Birnbaum, cette fois-ci dans son rôle de directeur de thèse. Voilà maintenant près de six ans que nous nous connaissons, dans le travail, mais aussi dans l'amitié. Il m'a transmis sa fascination pour les plantes et les forêts, et son engouement pour la recherche. De cette dernière, il m'as ouvert les portes et à travers toute la confiance qu'il m'a accordé, j'ai construit celle qui m'a permis de me lancer et d'aller au bout de cette aventure. Un jour, un sage vanuatais a dit "un bon shell vaut mieux qu'un long discours", donc on se retrouvera au Nakamal dès que possible.

Je remercie également Xavier Lorca, mon co-directeur de thèse. Il y a quelques années, lorsqu'il était encore mon professeur à l'École des Mines de Nantes, il m'a dit "on trouve rarement le job de ses rêves du premier coup, l'objectif, c'est d'essayer d'y tendre progressivement". J'ai suivi son conseil, et je suis heureux qu'il m'ait amené à travailler avec lui. Son efficacité, son sens de l'organisation, sa confiance et son enthousiasme m'ont beaucoup aidé pendant ces trois années. Il a été présent à chaque fois que j'ai eu besoin de son aide, tout en me laissant l'autonomie dont j'avais besoin le reste du temps.

Je tiens également à remercier les membres de la team NCPIPPN, AMAP NC et associés, avec qui j'ai fait mes premiers pas dans les majestueuses forêts de Nouvelle-Calédonie. Beaucoup sont devenus des amis, et même si certains ne sont plus sur le caillou leur souvenir y demeure. Merci donc à Thomas Ibanez pour sa sagesse et sa simplicité, merci à Robin Pouteau pour son ouverture et sa décontraction tahitienne, merci à Élodie Blanchard pour sa joie et ses encouragements vers cette aventure doctorale, gracias a Santiago Trueba-Sanchez para todos estos momentos filosóficos al Nakamal, merci à Grégoire Blanchard pour toutes nos discussion et pour les coups de pêche dans le lagon, merci à Vanessa Hecquet pour son énergie rayonnante, merci à Hervé

Vandrot pour sa personnalité hors du commun, merci à Jérémy Girardi pour sa gentillesse, merci à Laure Barrabé, David Bruy, Sandrine Isnard, Gildas Gateblé, Céline Chambrey, Tanguy Jaffré, William Nigote, Jacqueline Fambart-Tinel, Yohann Pillon, et Jean-Jérôme Cassan. Je remercie également tous ceux dont je n'ai pas cité le nom, mais qui ont rythmé ma vie et mes séjours en Nouvelle-Calédonie avant et pendant cette thèse. Merci également aux différentes tribus et particuliers qui en nous ouvrant l'accès à leurs terres m'ont donné la chance de découvrir des environnements naturels exceptionnels.

Revenons en métropole, et notamment à Montpellier, où j'ai passé la majeure partie de ces trois dernières années. Je remercie tous les amis et collègues d'AMAP, une unité de recherche où la convivialité fait partie du quotidien. Notamment, je remercie tous les amis de l'équipe Pl@ntnet, qui m'ont accueilli à bras ouverts et avec qui j'ai passé d'innombrables moments de rigolade. Et oui, rigoler pendant sa thèse, ça aide. Je remercie également mes collègues de bureau, Christophe Botella et Sue Han Lee pour cette bonne ambiance quotidienne, ainsi que tous mes camarades de café (parfois houblonné) du PS1, Titouan Lorieul le torréfacteur perfectionniste, Antoine Affouard le vétéran, Alain Ibrahim le chasseur-cueilleur, Mathias Chouet le troll, Rémi Knaff, Claudia Huertas, Maïri Souza Oliveira, Mélaine Aubry-Kientz, Anthony Laybros, Adan Mora-Fallas, Pierre Bonnet, Jean-François Molino, Yohan Soubeyran, Hervé Goëau, Grégoire Vincent, et tous les autres !

Je remercie également tous les collègues de l'IMT Mines Atlantique, et notamment Charles Prud'homme et Jean-Guillaume Fages pour leur aide sur Choco, Giovanni Lo Blanco pour les absurdes soirées de conférences, Gilles Simonin et Philippe David pour toutes ces discussions enrichissantes.

Merci également à tous les membres de mon comité de thèse, de mon jury de thèse et à mes différents co-auteurs, Philippe Vismara, Ghislain Vieilledent et Nicolas Rinck, pour leur précieux soutien.

De tout mon cœur, je remercie ma famille et mes amis, de Nantes, de Montpellier, de Nouméa et d'ailleurs, pour tout l'amour et tout le soutien qu'ils m'ont apporté, peut-être sans le savoir. J'aimerais citer tous leurs noms, mais il me faudrait bien une dizaine de pages pour ça, alors j'espère qu'ils se reconnaîtront tous.

Pour finir, je tiens à remercier une personne en particulier, Camille, pour avoir ensoleillé ma dernière année de thèse, pour son amour, pour le soutien qu'elle m'a apporté pendant les derniers moments de cette aventure et pour ces belles perspectives de vie qui s'ouvrent.

Dimitri Justeau-Allaire.

CONTENTS

I GENERAL INTRODUCTION AND STATE OF THE ART

II A CONSTRAINT-BASED APPROACH TO SYSTEMATIC CONSERVATION PLANNING

LIST OF FIGURES

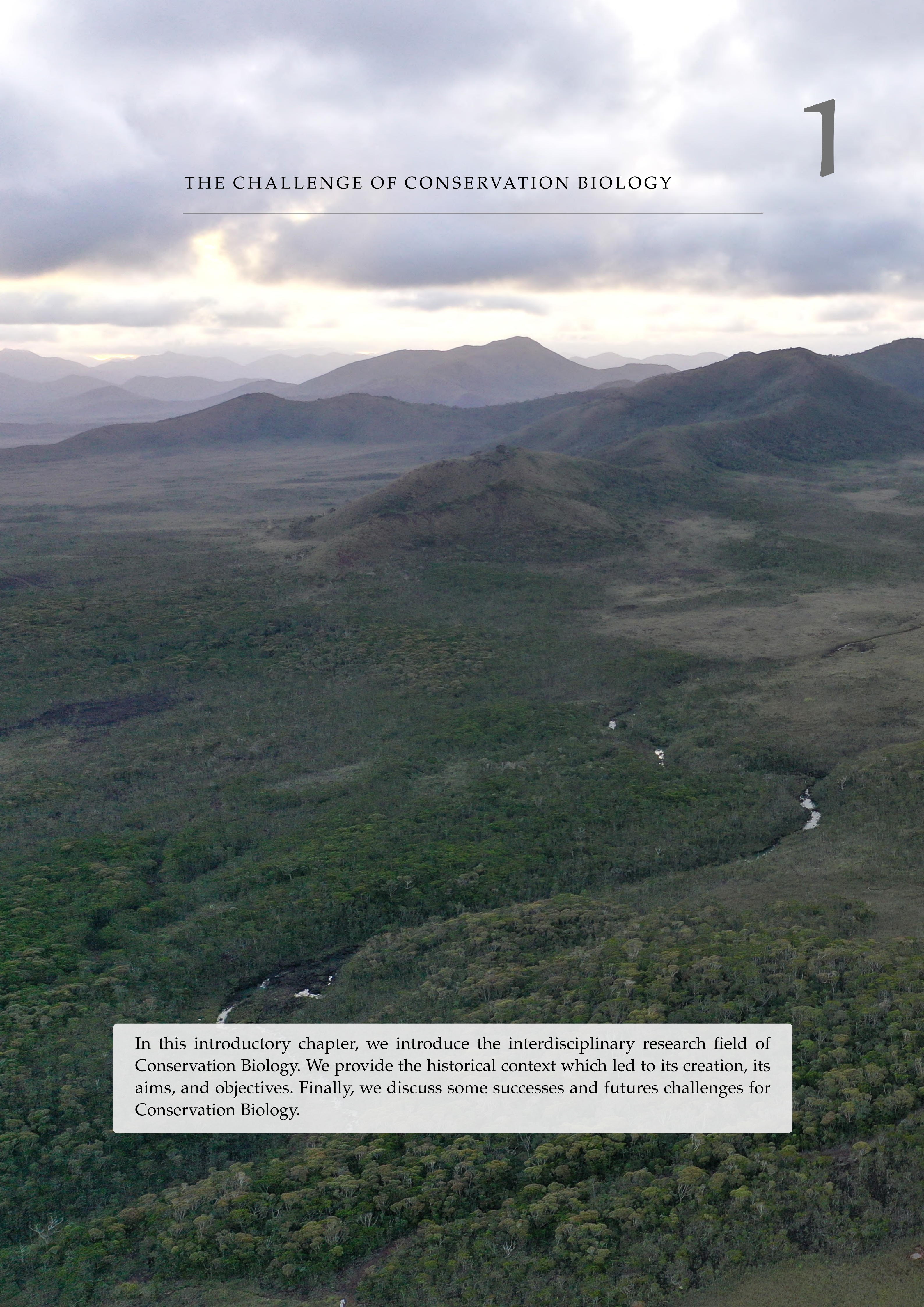
LIST OF TABLES

Part I

GENERAL INTRODUCTION AND STATE OF THE ART

1

THE CHALLENGE OF CONSERVATION BIOLOGY



In this introductory chapter, we introduce the interdisciplinary research field of Conservation Biology. We provide the historical context which led to its creation, its aims, and objectives. Finally, we discuss some successes and futures challenges for Conservation Biology.

1.1 THE GLOBAL NATURE CRISIS

The world is facing an unprecedented nature's crisis in human history, as ecosystems are declining along with biodiversity. Although accurate numbers are difficult to obtain, current species extinction rates are estimated 1000 times higher than the natural background rates, and about one million species are today threatened with extinction (Vos et al., 2015; Díaz et al., 2020). Anthropic activities are the principal cause of this crisis, which is known as the Holocene extinction (also referred to as the sixth mass extinction, or Anthropocene extinction). Land-use change is the most impacting of these activities, which includes agricultural expansion, urbanization, and natural resources exploitations such as mining or overfishing. Climate change, pollution, and invasive alien species are also major drivers of biodiversity and ecosystems decline. Many of these impacts on nature are already irreplaceable.

Though, in addition to its irrefutable intrinsic values, nature's provides resources and services that are vital to humans (Batavia and Nelson, 2017; Díaz et al., 2018). For instance, forests and woodlands are believed to host about 50% of the world's species, they produce just under half of the global terrestrial annual net primary production (amount of biomass produced per unit area and time, less plant respiration costs), and store about 50% of the world's terrestrial carbon stocks in their soils (Field et al., 1998; Groombridge and Jenkins, 2002). Moreover, forests provide many services. They protect watersheds, supply clean water, preserve soils from erosion, mitigate climate change, reduce air pollution, and provide many essential resources such as food, wood, or medicines. Despite, deforestation is one of the main consequences of land-use change. Between 1990 and 2015, the global forest area declined by 129 million ha. This loss mainly occurred in tropical forests, which are the richest and the most productive (Keenan et al., 2015).

As individuals of the species which has been causing so much damage to nature, we can feel concerned and willing to participate in the momentum of a society built around more respectful and sustainable models. This horizon is the leitmotif of nature conservation, which have led to the emergence of a dedicated science: conservation biology.

1.2 ORIGINS, AIM, AND OBJECTIVES OF CONSERVATION BIOLOGY

In the broadest sense, conservation defines as the preservation of something undamaged over time, or as the prevention of wasteful use of a resource. Nature conservation embraces both definitions, with two ethical motivations that are not mutually exclusive. In the first case, nature conservation is motivated by its perceived intrinsic value. In the second case, it is driven by the wish to provide sustainable access to the resources that nature offers to humans (from this point, conservation implicitly refers to nature conservation). It is possible to find conservation practices far back in the history of human societies. For instance, in medieval Europe, forests were named so from the low Latin word *foresta* which designates a prohibited land. Indeed, many forests were private areas whose enjoyment were reserved to nobles, for hunting, fishing and logging. Although based on a privilege now highly questionable, this practice can be seen as a conservation effort to preserve the resources provided by forests. In our recent history, the modern conservation movement was impulsed by naturalists in the mid-19th century. The first conservation organization, the Association for Protection of Sea Birds, was founded in England in 1868. A few years later, the first national park, Yellowstone, was created in the United States. It was the first example of a landscape-scale conservation effort, which is now at the core of most conservation policies through nature reserves.

Science always played an essential role in modern conservation. It has, however, only recently become a science itself, which emerged in 1978 with the First International Conference on Conservation Biology organized at the University of California by the biologist Michael E. Soulé. This meeting led to the publication of a foundational book, *Conservation Biology: An Evolutionary-Ecological Perspective* (Soulé and Wilcox, 1980). Five years later, the Society for Conservation Biology was the first scientific organization of this new discipline, which is now a well-established field, yet still burgeoning with many advances and innovative approaches[1]. The aim of conservation biology is to provide ethical and scientific responses to face the global nature crisis, protect species, preserve biodiversity and ecosystems. In this respect, it relies on three main objectives:

- Describing and understanding the diversity of species, communities, and ecosystems.
- Studying and quantifying the impacts of human activities on species, communities, and ecosystems.
- Developing practical, interdisciplinary, and integrated approaches to prevent species extinction, maintain genetic diversity among communities, preserve and restore biodiversity and ecosystems.

The first two objectives fall within the scope of fundamental research through a quest for knowledge and understanding. The third objective, on the other hand, also defines conservation biology as action and standard-setting science. In this regard, conservation biology is an applied scientific discipline based on the ethical values of nature conservation. Another specificity of this research field is its highly interdisciplinary nature (cf. Figure 1.1). If environmental and ecological sciences were at the core of conservation biology, it also relies on human and social sciences where some fields emerged along with the modern nature conservation movement. For example, environmental philosophy, which developed in the 1970s, is concerned by humans' relationship with their natural environment and by morals values that stem from it. More recently, mathematical and information sciences have been increasingly involved in conservation biology. Taking the example of computer science, bioinformatics established about fifty years ago but is mostly associated with

---

[1] For a more detailed history of conservation biology, interested readers can refer to the first chapter of *Conservation Biology: Foundations, Concepts, Applications* (Van Dyke, 2008).

genomics. Biodiversity informatics gained prominence at the beginning of the 21st century (Bisby, 2000) with a focus on taxonomic and biodiversity data and a broader scope of applications. In the same line, ecoinformatics appeared with an emphasis on ecological rather than taxonomical applications (Kareiva, 2001; Dengler et al., 2011). Lately, another term emerged from computer science communities: computational sustainability (Gomes, 2008). There is little doubt that these fields are overlapping. There is even less doubt that we are witnessing a major shift in conservation biology through the growing use of mathematical and information sciences.

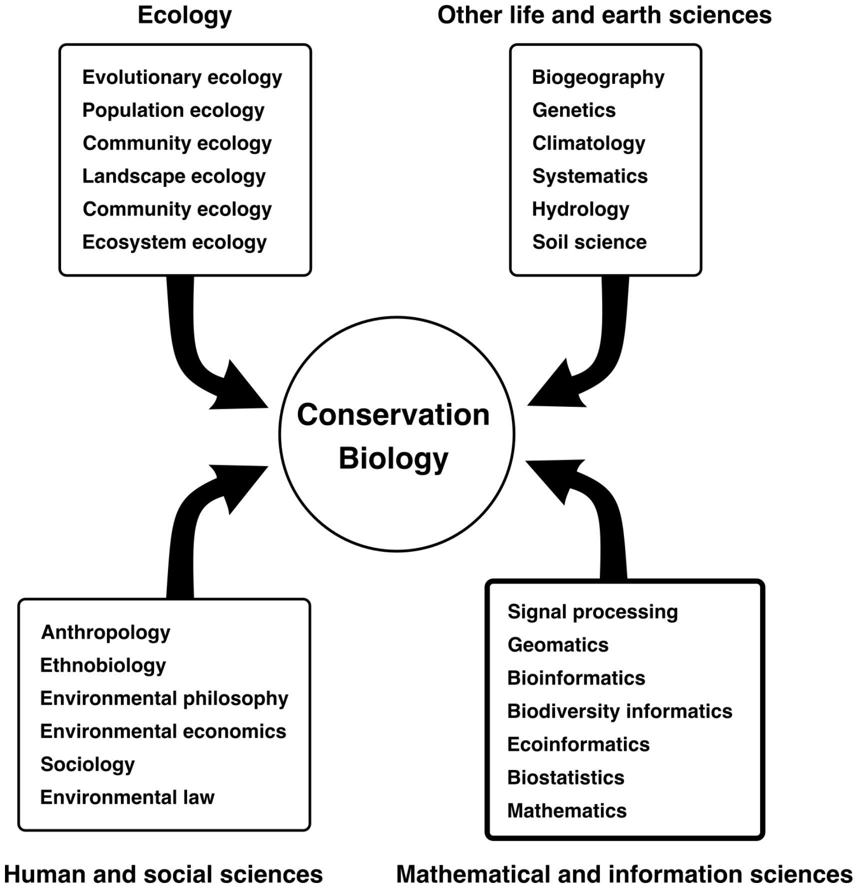


Figure 1.1: Scientific disciplines contributing to the interdisciplinary field of conservation biology. Adapted from Temple, 1991, with the addition of mathematical and information sciences fields.

Following the third objective of conservation biology, the union of these many disciplines aims to synthesize knowledge into new techniques and principles to improve the management of nature. Such transfers from research to implementation may involve identifying priority areas to protect, defining environmental legislative frameworks, developing new economic models, or supporting sustainable land-use planning (cf. Figure 1.2).

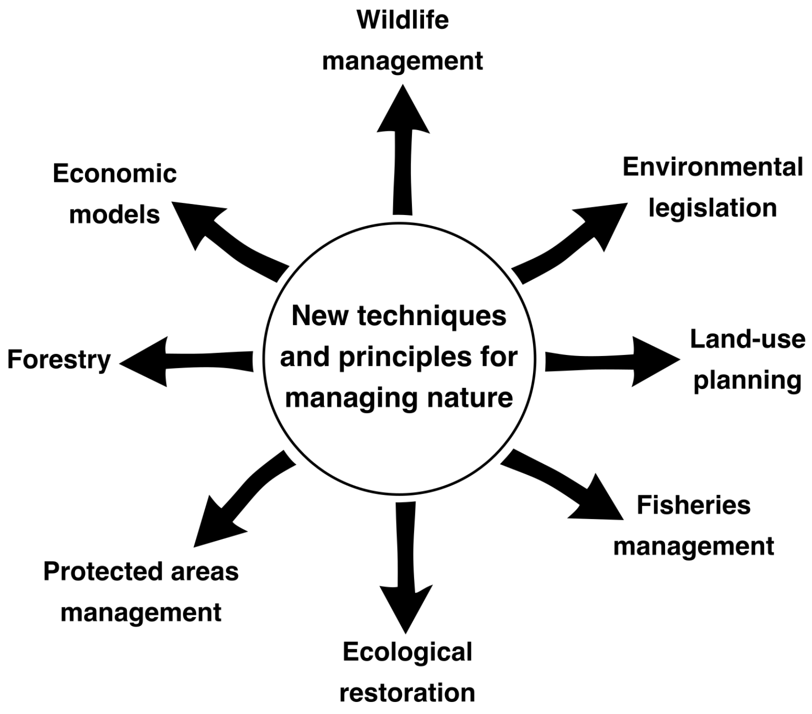


Figure 1.2: Transfers for conservation biology to nature management. Adapted from Temple, 1991.

1.3 SUCCESSES AND CHALLENGES OF CONSERVATION BIOLOGY

With more than forty years of existence, conservation biology established as a recognized and successful science from an academic perspective. Many universities teach conservation biology, and the number of publications has been steadily growing since the creation of the discipline. For instance, the number of papers submitted to the journal *Conservation Biology* from 1993 to 2006 has increased by 13.2% on average (Meffe, 2006). The field has also been successful in getting closer to the non-academic world: conservation biology has involved in the United Nations, national and regional agencies, non-governmental organizations (NGOs), citizen groups, etc.

But, is conservation biology achieving its goal? This question has intricate answers. What stands clear is that the global nature crisis has not been halted, as the latest Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) report indisputably shows, as have many other studies (Cardinale et al., 2012; Haddad et al., 2015; Keenan et al., 2015; Woinarski et al., 2015; Strona et al., 2018; Díaz et al., 2020). However, conservation biology significantly influenced the principles and practices of conservation and has several success stories to relate. One significant example is the International Union for Nature Conservation (IUCN) red list of threatened species (see Figure 1.3). Established by the IUCN (an international conservation NGO), it is now a recognized international standard which is influencing policies, resource allocation, and conservation planning. Assessments from conservation scientists contribute to this list, as well as the list guides conservation research (IUCN, 2020). More success stories are described in Boxes 1.1, 1.2, and 1.3.

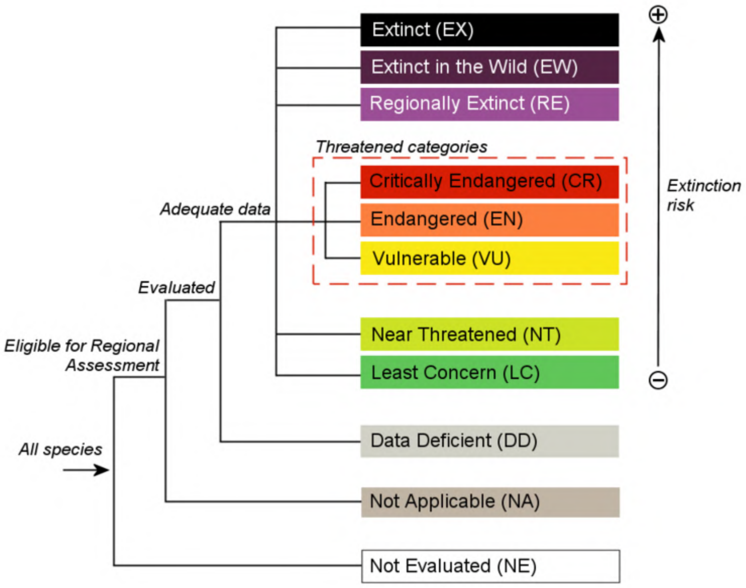


Figure 1.3: Categories of the International Union for Nature Conservation (IUCN) red list of threatened species. ©IUCN

Yet, despite substantial growth, numerous achievements, and clear recognition, conservation remains a weak social machine when compared to industries, finance, or armies. Powerful political and commercial organizations whose interests are conflicting with those of conservation constitute an obstacle, along with the consumerist way of life in many societies (Buckley, 2015). As a scientific field, conservation biology cannot solve the problem alone, but should nonetheless be involved in advocacy, following its third objective (see Section 1.2). In this sense, many conservation scientists depicted the *research-implementation gap* over the past few years (Robinson, 2006; Knight et al., 2008; Game et al., 2015; Wistbacka et al., 2018; Williams et al., 2020). Indeed, despite a burgeoning literature on conservation assessments and techniques for performing such assessments, research results are rarely translated into actions. Several trails could help to bridge this gap, such as engaging a closer dialogue between scientists and managers (Prendergast et al., 1999), working more closely with political and social sciences (Balmford and Cowling, 2006), directly involving as scientists in implementation (Arlettaz et al., 2010), or even getting more politically involved (Ellison, 2016).

In this PhD thesis, we focus on one particular aspect of conservation biology, conservation planning, which aims to provide decision support for land-use planning, wildlife and protected areas management, or even ecological restoration. Conservation planning directly falls within the third main objective of conservation biology (see Section 1.2) by adopting as a strategy an involvement in decision-making processes that are related to conservation. Through better decisions, conservation planning can offer solutions to reduce the negative impacts of land-use change and improve the effectiveness of conservation actions. Over its history, conservation planning have been increasingly relying on modelling and computer science, attracting more and more scientists from these fields. This modern and systematic approach of conservation planning is now referred as systematic conservation planning.

Box 1.1: Recovery of the Seychelles warbler after translocation.

The Seychelles warbler (*Acrocephalus sechellensis*) is a small passerine endemic to the granitic islands of Seychelles. Destruction of its habitat and introduction of invasive predators made it one of the rarest birds in the world in the mid 20th century, with a population of 26 individuals. Previously occurring in several islands, it only remained in the Cousin Island (Crook, 1965). Intensive conservation action started in 1968 with the purchase of the Cousin island by a consortium led by BirdLife International (at this time called the Internation Committee for Bird Protection). Habitat management led the population to recover and to remain stable, and the species was translocated to four other islands between 1988 and 2011 (Wright et al., 2014). The population of the Seychelles warbler now comprises more than 3000 mature individuals and is increasing. Its conservation status has changed from Threatened to Near Threatened between 1988 and today (IUCN, 2016).

Figure 1.4: The Seychelles warbler (*Acrocephalus sechellensis*). ©Remi Jouan, CC BY-SA 3.0.

Box 1.2: Ban of deep-sea trawling in European Union.

Deep-sea trawling is a fishing technique developed at the end of the 20th century as a response to the collapse of surface-water resources. It consists of dragging a fishing net along the seafloor. This method is not selective and extremely harmful to deep-sea slow-growing species and ecosystems. Many conservation scientists warned on the dangers of this practice (Gage et al., 2005; Pusceddu et al., 2014; Victorero et al., 2018). Their scientific results provided the basis for intensive advocacy campaigns, led by NGOs (such as Bloom and the Deep Sea Conservation Coalition) and citizen groups. In June 2016, the European Union finally banned deep-sea trawling below 800 meters in its waters.

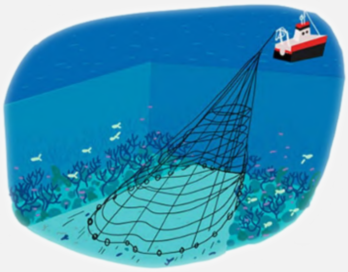


Figure 1.5: Illustration of deep-sea trawling ©Pénélope Bagieu.

Box 1.3: Protection of Tasmanian endemic and endangered plant species with the first reserve selection algorithm.

Tasmania is an Australian island located 240 km to the south of the mainland, which hosts a unique flora and fauna. For example, among 1530 native vascular plant species, 20% are endemic. However, logging and fire have been threatening several of these species. In 1983, the researcher Jamie Kirkpatrick developed and applied an iterative method to identify a set of areas that should be protected to preserve 25 endemic and endangered plant species. By the late 1990s, each of the seven new reserves recommended by Kirkpatrick was protected and preserved from logging. His method is now considered as the first reserve selection algorithm for conservation planning (Kirkpatrick, 1983; Pressey, 2002).

Figure 1.6: Douglas-Apsley National Park in Tasmania, extended in 1989 after Kirkpatrick's recommendations. ©VirtualWolf, CC BY-SA 2.0.

2

SYSTEMATIC CONSERVATION PLANNING

In this chapter, we first introduce systematic conservation planning, which is a subfield of conservation biology, and the research area of this PhD thesis. Then, we provide a review of the problems that were addressed by systematic conservation planning, along with existing methods and software packages. Finally, we discuss some trends and challenges of this research field.

2.1 WHAT IS SYSTEMATIC CONSERVATION PLANNING?

Landscape-scale conservation through reserves is a foundational and well-established practice in nature management. However, reserves have long been developed in an opportune manner and motivated by the protection of scenic, recreational, or inspirational places of little economic interest. By the end of the 20th century, conservation scientists started advocating a more systematic selection and delineation of reserves, with the primary motivation of effectively sustaining biological diversity (Myers, 1988; Pressey et al., 1993). In 2000, this concern came to be known as a sub-discipline of conservation biology called *systematic conservation planning* (SCP) (Margules and Pressey, 2000), which now encompasses other issues such as ecological corridor design or restoration planning, or conservation projects prioritization.

First concerns for SCP emerged from island biogeography theory: viewing nature reserves as islands surrounded by an ocean of altered habitat, a set of geometric principles (see Figure 2.1) were devised to produce efficient reserve delineations (MacArthur and Wilson, 1967; Diamond, 1975). However, those principles quickly became controversial, notably because areas surrounding reserves are not as inhospitable as the ocean is for terrestrial organisms and because they assume ecosystems to be at equilibrium state, which is very unlikely to be satisfied in anthropized and fragmented landscapes (Margules et al., 1982). Moreover, according to the context, it may be safer to establish several distant reserves to preserve the whole from natural catastrophes, forest fires, or diseases. Similarly, in heterogeneous landscapes, a single large reserve cannot guarantee the preservation of a large proportion of species (Higgs, 1981).

During the 1980s, many researchers tried to put emphasis on biodiversity features representation rather than geometrical properties of reserves. The first approaches consisted in multi-criteria scoring procedures (e.g. Margules and Usher, 1981; Smith and Theberge, 1986; Usher, 1986; Smith and Theberge, 1987). Nonetheless, several authors quickly showed that such methods are inefficient in their goal because they rely on a local reasoning which cannot spot the complementarity between sites in the representation of biodiversity features (Kirkpatrick, 1983; Pressey and Nicholls, 1989). Rather, they found that much better results can be achieved using iterative procedures, which were called reserve selection algorithms (see Box 1.3). Soon after, complementarity was recognized as key principle of reserve selection (Vane-Wright et al., 1991; Pressey et al., 1993) and many methods have been devised in this direction (e.g. Rebelo and Siegfried, 1992; Possingham et al., 1993; Underhill, 1994; Ball, 2000).

It is now widely recognized that, instead of geometric principles, we should regard spatial attributes contextually along with biodiversity features representation. Since its introduction, SCP has now well established, with many methodological and applied papers. The field has even attracted researchers from the mathematical optimization and computer science areas. Indeed, one interesting thing about SCP is that it involves constraint satisfaction and optimization problems that are theoretically and computationally challenging. As an example, identifying the minimum requirements to represent a complete set of endangered species in a reserve network, namely ensuring complementarity, equates the *set cover problem* in combinatorics, computer science, and operations research (Underhill, 1994; Camm et al., 1996). This problem lies in the NP-Complete computational complexity class. In simplified terms, solving this problem is challenging (see Boxes 2.1 and 2.2 for a more detailed explanation). Usually, SCP problem involves a combination of such hard problems and therefore addresses original and novel questions to modelling and computer science.

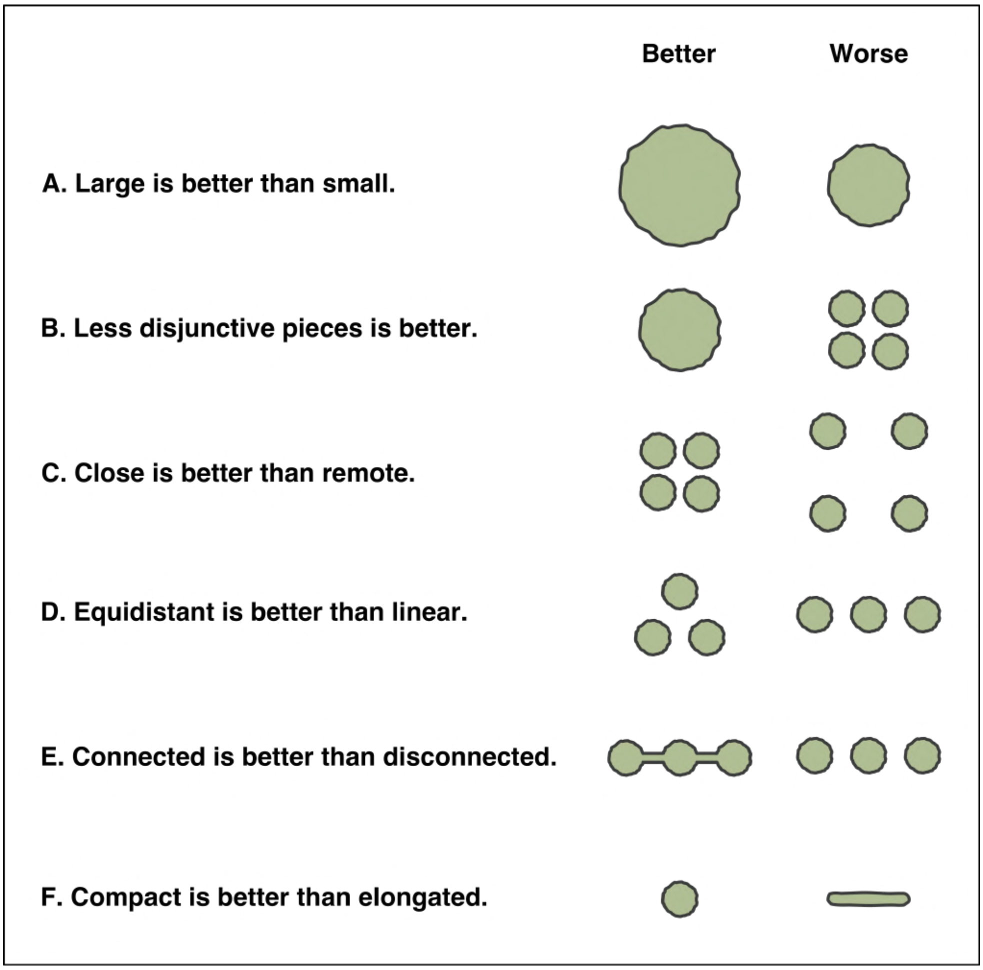


Figure 2.1: Diamond's geometric principles for reserve design, inspired from island biogeography theory. Adapted from Diamond, 1975.

Box 2.1: Computational problem.

In theoretical computer science, computational problems are those that analytical methods cannot solve, and for which we need algorithms. We distinguish four main types of such problems, all illustrated with the undirected graph G (also called network, see Section 4.1.3 for more details) represented in Figure 2.2:

- A *decision problem* is such that, given an input, the solution is either yes or no. For example: "Is there a path from a to f in the graph G ?". The answer is yes.
- A *search problem* is such that, given an input, a solution is an object satisfying the problem's statement. For example: "Find a path from a to f in the graph G". A solution is: a → c → d → f.
- An *optimization problem* is similar to a search problem, at the difference that the expected solution must be optimal according to one or more criteria. For example: "Find a shortest path from a to f in the graph G" (the length of a path being the number of links). A solution is a → b → f.
- A *counting problem* refers to a search or optimization problem for which is expected the number of solutions. For example: "Count the shortest paths from a to f in the graph G". The solution is 2.

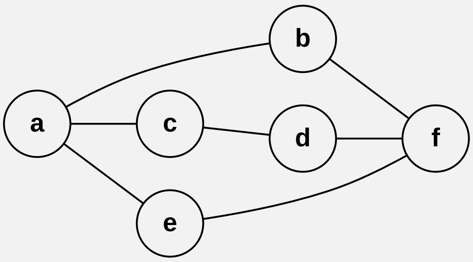


Figure 2.2: Undirected graph G.

Box 2.2: Computational complexity.

In theoretical computer science, computational complexity focuses on studying and classifying computational problems according to their intrinsic difficulty. It also investigates the relationships between classes. Computational complexity classes inform on the resources (time or memory) needed by an algorithm to solve a given problem. For example, the class **P** contains all decision problems that can be solved with a deterministic algorithm in polynomial time (i.e. number of operations needed) according to the input. The class **NP** contains all decision problems that can be solved with a non-deterministic algorithm in polynomial time. A problem is said **NP-hard** if any problem in **NP** can be reduced to it in polynomial time, and if such a problem belongs itself to **NP** it is also said **NP-Complete**. **NP-hard** problems are challenging because a non-deterministic computer is a theoretical construct with no physical implementation. We thus cannot solve **NP-hard** problems in polynomial time with deterministic computers (unless P=NP), such problem can quickly become intractable in practice. Computational complexity theory is a broad and complex research area, interested readers can refer to Arora and Barak, 2009 or Perifel, 2014.

2.2 STATE OF THE ART

Researchers devised many computational methods to solve SCP problems. These methods have, of course, much in common, but they also differ a lot in the techniques used and on the types of question they can address. In this section, we try to establish a concise state of the art, focusing on spatially-explicit methods, which are prevalent in this field (Margules and Pressey, 2000). We will complete it in chapter 4 with a comprehensive nomenclature of SCP problems. First, we provide a very generic formalism that is suitable for any spatially-explicit SCP problem: given a geographical space $\mathcal{S}$, delineate n regions $R_0, ..., R_{n-1}$ within $\mathcal{S}$ such that conservation goals are met. In the vast majority of cases, $\mathcal{S}$ is discretized into sites (e.g. regular square grid). For instance, if we need to delineate a reserve protecting a set of endangered species in a study area, we will try to delineate a single region such that each species occurs in it. Many criteria can be used to express conservation goals, they can be classified in two main categories: feature representation and spatial criteria.

2.2.1 *Feature representation criteria*

A feature designates a characteristic of the geographical space $\mathcal{S}$ that can be spatially represented with numerical values. Most of the time, biodiversity features (e.g. occurrences) are used to ensure properties such as complementarity. However, socio-economic values (e.g. exploitable land, customary area) can also be represented as features to express managers' constraints and quantify the trade-offs between different solutions. Features can be described with three data types: binary data (e.g. species occurrence), quantitative data (e.g. species abundance, land acquisition cost), and probabilistic data (e.g. species distribution models). Let $\mathcal{F}$ be a set of features and R_i a region, the feature representation criteria that have been used in the literature so far are the following:

Occurrence representation. The representation in R_i of at least one occurrence for each feature in $\mathcal{F}$, that is ensuring that R_i is a cover of $\mathcal{F}$. This criteria can also be used for optimization, by maximizing the number of occurring features from $\mathcal{F}$ in R_i for example. This criteria can be used when binary or quantitative data is available, and corresponds to the principle of complementarity (e.g. Vane-Wright et al., 1991; Church et al., 1996; ReVelle et al., 2002). See Figure 2.3 for an illustration.

Abundance representation. The representation in R_i of a minimum abundance for each feature in $\mathcal{F}$, when quantitative data is available (e.g. Margules et al., 1988; McDonnell et al., 2002; Watts et al., 2009). This criteria can also be used for optimization, by maximizing the abundance of a given feature for example.

Occurrence representation with backups. The representation in R_i of an occurrence for each feature in $\mathcal{F}$, in at least k distinct sites, when binary or quantitative data is available (e.g. Margules et al., 1988; ReVelle et al., 2002; Delmelle et al., 2017). This criteria can also be used for optimization, by maximizing k for example. See Figure 2.3 for an illustration.

Probabilistic representation. The representation in R_i of a minimum probability of presence for each feature in $\mathcal{F}$, when probabilistic data is available (e.g. Haight et al., 2000; ReVelle et al., 2002; Billionnet, 2011). This criteria can also be used for optimization, by maximizing the minimum probability of presence of features from $\mathcal{F}$ in R_i for example.

Phylogenetic diversity (PD). Some authors also considered the representation of a minimum phylogenetic diversity (PD) among the features represented in R_i (e.g. Moulton et al., 2007; Billionnet, 2017). This criteria can also be used for optimization, by maximizing PD. This criteria applies when $\mathcal{F}$ corresponds to binary or quantitative data of taxonomic entities, PD is thus computed from a phylogenetic tree where each entity in $\mathcal{F}$ must be represented.

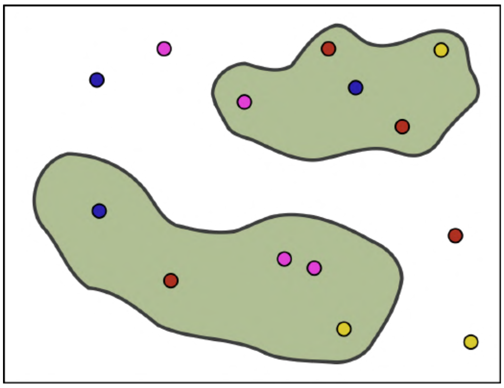


Figure 2.3: A reserve system ensuring complementarity (occurrence representation) for all species of interest, as well as occurrence representation with backups (each species is represented in at least two distinct sites). Each colour represents a species.

2.2.2 *Spatial criteria*

Spatial criteria are useful to control the spatial configuration of solutions. They can involve biological concerns such as the satisfaction of minimum requirements to ensure species persistence, but also management concerns such as ensuring that a reserve is accessible from an existing track network. Such criteria can involve a single region, or several regions that must relate to each other in a certain way (e.g. be adjacent).

Connectivity. Ensuring the connectivity of a region R_i is a frequent concern. Most of the time, connectivity is characterized through a graph, which can represent the adjacency between sites but also species dispersal capabilities (e.g. Sessions, 1992; Briers, 2002; Wang and Önal, 2011; Billionnet, 2016; Jafari et al., 2017). See Figure 2.4 for an illustration.

Number of connected components. Instead of ensuring full connectivity, the need for controlling the number of connected components of a region R_i was formalized by Williams et al. (2005). Hof and Flather (1996) devised a model allowing the control of the number of component, but only when regions are circular or rectangular, and Williams (2002) proposed a model allowing a maximum number of disconnections. For example, the reserve system represented in Figure 2.3 has two connected components.

Perimeter of a region. The perimeter (or boundary length) of a region R_i can also be considered. It can be used as a surrogate to approximate connectivity, or even be minimized to ensure the compactness of a reserve (e.g. McDonnell et al., 2002; Ball et al., 2009; Weerasena et al., 2014).

Size of a region. The size of a region R_i was spotted as an important spatial attribute which can be constrained, maximized or minimized (e.g. ReVelle et al., 2002; Williams et al., 2005). This criteria can, however, also be considered as a feature representation one, especially when the sites have different areas.

Size of connected components. Similarly, when a region R_i is not connected, it may be desirable to control the area of its connected components, to ensure populations persistence in a reserve network for example. This criteria has only been addressed when the candidate connected components are delineated a priori (Rothley, 1999; Williams et al., 2005).

Distance between regions. The distance between two regions R_i and R_j, or between the connected components of a region R_i, can be useful to control for facilitating species migration between reserves, or preserving a reserve from negative edge effects of a urbanized area for example (e.g. Rothley, 1999; Williams, 2008; Önal et al., 2016).

Buffer zone. Another frequent requirement is the design of a buffer zone R_b between a protected core area R_i and a non-protected area R_j. Such a buffer zone is useful to preserve a reserve from negative edge effects, or to nest several levels of protection (e.g. Williams and ReVelle, 1996; Billionnet, 2013; Cheng et al., 2015). See Figure 2.4 for an illustration.

Shape of a region. The shape of a region R_i can be important for ensuring species persistence or for management. Most of the time, we need a region to be compact, which can be ensured through its perimeter/area ratio, or by ensuring a maximum distance between every pair of sites within the region (e.g. Williams et al., 2005; Billionnet, 2016).

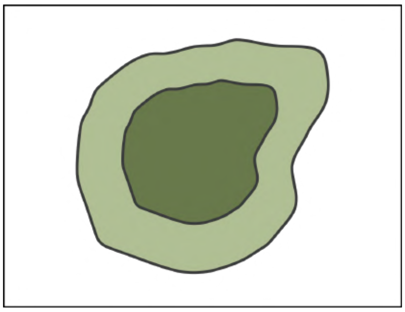


Figure 2.4: Example of a connected protected core area surrounded by a protected buffer zone.

2.2.3 *Techniques*

Different techniques can address and solve SCP problems. As the questions can widely vary in their structure, criteria, or application context, and because the underlying computational problems pose challenges, we cannot claim that one technique is dominating the others (to the exception of scoring techniques that have many limitations, as explained in Section 2.1). Instead, each has its strengths, weaknesses, and performs adequately for a specific class of problems. Current SCP solving approaches mostly rely on ad hoc heuristics, metaheuristics, and mixed-integer linear programming (Sarkar et al., 2006). In the following, we briefly describe these techniques. We also inform on their genericity, flexibility, and guarantees regarding runtime, satisfiability and optimality. The *genericity* of a technique refers to the extent to which it can be used to address different problems. Its *flexibility* refers to the ease with which it can be adapted to new types of problems. A technique offers *runtime guarantees* when it provides control over the time needed to produce a solution. It provides *satisfiability guarantees* when it can ensure whether the solutions satisfy the constraints. Finally, it offers *optimality guarantees* when it yields control over the quality of solutions (according to an optimization objective and relative to the optimal solution). These characteristics are summarized in Table 2.1.

Ad hoc heuristics. Problem-specific local search algorithms, they employ practical methods to quickly find solutions to a pre-formulated problem. Ad hoc heuristic are either based on a constructive (greedy) algorithm (e.g. Kirkpatrick, 1983; Rebelo and Siegfried, 1992; Nicholls and Margules, 1993), or destructive (stingy) algorithm (e.g. Zonation software; Moilanen et al., 2009a) Such algorithms are fast and can be adapted to prioritization problems, but lack genericity and cannot always provide guarantees such as constraint satisfaction and optimality, which is the necessary price for bypassing the combinatorial complexity. For overly large and complex problems, intractable by exact approaches, ad hoc heuristics are useful to find good solutions.

Metaheuristics. High-level and problem-independent optimization algorithms. They offer the same characteristics as ad hoc heuristic, with a higher level of genericity and flexibility, but still no systematic guarantees on constraint satisfaction and optimality. Popular metaheuristics include tabu search, genetic algorithms, ant colony optimization algorithms, or simulated annealing which is used in SCP by Marxan software (Ball et al., 2009), or in McDonnell et al. (2002).

Mixed-integer linear programming (MILP). Mathematical optimization approaches where the objective function and the constraints are stated as linear inequalities, with some or all the variables are integers. With less guarantees on the runtime for complex and large problems, MILP provides constraint satisfaction and optimality guarantees. Another advantage of MILP lies in its declarative nature, it offers a modelling language to represent a problem as well as a generic solving method, which confers it a high level of expressiveness, genericity, and flexibility (e.g. Rodrigues et al., 2000; Billionnet, 2013; Dilkina et al., 2017; oppr R package, Hanson et al., 2019a).

Table 2.1: Characteristics (genericity, flexibility, and guarantees) of the three main techniques currently used in systematic conservation planning: ad hoc heuristics, metaheuristics, and mixed-integer linear programming (MILP).

	Genericity	Flexibility	Guarantees
Ad hoc heuristics	Low	Low	Runtime
Metaheuristics	High	Middle	Runtime
MILP	High	High	Satisfiability, Optimality

2.2.4 *Software packages*

An important factor for dissemination and use of SCP approaches is their availability. In this respect, software packages are very useful, but only some approaches have been released in this form. We compiled a (non-exhaustive) list of currently available state-of-the art SCP software packages, and summarized it in Table 2.2.

Marxan (Ball et al., 2009). Marxan is a target-based SCP software based on simulated annealing, a metaheuristic optimization technique inspired from the annealing process in metallurgy (interested readers can refer to van Laarhoven and Aarts, 1987). Marxan addresses a problem which can be formulated as following: given a tessellated geographical space $\mathcal{S} = [0, n[$, a cost c_i for each site i, a set of features $\mathcal{F} = [0, k[$, a representation amount a_{ij} for each site i and each feature j, a target t_j for each feature j, and a connectivity cost cv_{i_1,i_2} for each pair of site $\{i_1, i_2\}$ (most of the time the connectivity cost corresponds to the boundary length), find a region $R \subseteq \mathcal{S}$ (most of the time a reserve network) minimizing the following objective function:

$$\sum_{i\ \in R} c_i + b \sum_{i_1 \in R} \sum_{i_2 \in \mathcal{S} \setminus R} cv_{i_1,i_2} + \sum_{j \in \mathcal{F}} FPF_j FR_j H(s) \left(\frac{s}{t_j} \right). \tag{2.1}$$

The first term corresponds to the total cost of the region R, as Marxan tries to achieve user targets at minimum cost. The second term corresponds to the weighted connectivity cost (e.g. boundary length) of R, where b is a penalty factor used to control the relative importance of connectivity in the objective function. The third term corresponds to the penalties for violating feature representation targets, FPF_j is a feature penalty factor for the feature j, and FR_j the cost for satisfying the target for feature j only. The term s corresponds to the amount of representation unsatisfied for the feature j, such that $s = t_j - \sum_{i \in R} a_{ij}$. $H(s)$ is a step function such that $H(s) = 0$ if $s \leqslant 0$ and $H(s) = 1$ otherwise. Because it relies on simulated annealing, Marxan needs to integrate the constraints in the objective function. Thus, their satisfaction cannot always be guaranteed, and the results are sensitive to penalty factors' values.

Marxan is distributed as a free software and offers a graphical user interface (GUI). Marxan with zones is an extension which allows the user to delineate several regions at the same time, with distinct targets (e.g. several levels of protection) (Watts et al., 2009). Zonae Cogito is a software which enables a direct interaction with Marxan and databases through a geographical information system (GIS) interface. More recently, Marxan Connect was released as an alternative GUI to facilitate the incorporation of connectivity issues in Marxan (Daigle et al., 2020).

Zonation (Moilanen et al., 2009b). Zonation is prioritization SCP software relying on an ad hoc local search heuristic. At the difference of most SCP approaches, Zonation does not rely on a target-based paradigm (although there is an option to use it for target-based planning, see Moilanen, 2007). Instead, it ranks every site from a tessellated geographical space $\mathcal{S} = [0, n[$ from the most to the least valuable, according to a set of features $\mathcal{F} = [0, k[$ given as input data. The user can define a number of regions to delineate as hierarchical zonations, for instance: the 10%, included in the top 20%, included in the top 30%, etc. The prioritization algorithm starts from the full landscape, then iteratively removes sets of sites from the remaining area. At each stage of the algorithm, the set to remove is selected such that is minimizes a loss function, which can be chosen among a predefined catalogue. For example, the basic core-area Zonation loss function is defined as follows for each site i:

$$\sigma_i = \max_{j \in \mathcal{F}} \frac{q_{ij} w_j}{c_i} \tag{2.2}$$

Where w_j is a weight associated to the feature j, c_i the cost associated to the site i, and q_{ij} the proportion of the feature j in the site i according to the remaining sites. Using this loss function, the algorithm will remove at each stage the site where the feature with the highest weighted proportion divided by the site cost is minimal. Besides, an aggregation method can be selected among a predefined catalogue to guide the heuristic towards less fragmented solutions (e.g. boundary length penalty). Such aggregation methods directly alter the loss function to include connectivity consideration. Zonation is distributed as a free software and offers a GUI.

C-Plan (Pressey et al., 2005; Pressey et al., 2009). C-Plan was designed as an integrated and interactive decision support system for SCP that can be used during negotiations between stakeholders. It links with GIS (ESRI ArcView 3) and relies on a statistical estimator to compute irreplaceability of sites and ensure complementarity (Ferrier et al., 2000). C-Plan can also be linked with Marxan. We did not find much more details on C-Plan's underlying mathematical model, but it seems to distinguish mainly by its emphasis on irreplaceability, interactivity, feature representation criteria, without spatial criteria. The project is apparently discontinued, as we could not find any active download link.

ConsNet (Ciarleglio et al., 2009, 2010). ConsNet is a target-based SCP software based on a metaheuristic optimization technique called tabu search (Glover, 1989, 1990; Ciarleglio, 2008). Given a tessellated geographical space and a set of features, ConsNet aims at delineating a region (most of the time a reserve network) either satisfying targets on feature representation at minimum cost or maximizing feature representation under cost constraints. The first problem is the same as Marxan's, but ConsNet offers more spatial criteria in both formulations:

- Compactness, through the perimeter/area ratio measure.
- Connectivity, through minimization of the number of connected components.
- Replication, which is the number of connected components criteria extended with feature representation with backups requirements.
- Alignment, which aims to coincide the delineated region with existing units (e.g. ecoregions, watersheds). As matter of fact, this criteria can also be considered as a feature representation criteria: the delineated region must cover a minimum area of existing units.

ConsNet also provides an interactive multi-criteria analysis to explore possible trade-offs between equivalent solutions, is distributed as a free software, and offers a GUI.

Conefor Sensinode (Saura and Torné, 2009). Conefor Sensinode cannot be exactly considered as a SCP software, it can however be useful for conservation planning. The aim of this software is to quantify the level of connectivity over a landscape through graph-based indices. A graph (also called network) is a mathematical object composed of vertices (or nodes) and arcs (also called edges or links) between the vertices (more details on graphs will be provided in Section 4.1, see Figure 2.2 for an example). Therefore, the user must provide a graph representation of a landscape, or a reserve network for example. Conefor Sensinode can compute several indices and quantify the importance of nodes. That is, how would a given index vary if a given node is added or removed from the network. Conefor is provided as a free software, with a GUI available only for Windows, and a command-line interface available for Windows, Mac, and Linux.

Prioritizr R package (Hanson et al., 2020). Prioritizr is an R package that was designed to solve a wide range of SCP problems with MILP. Prioritizr can solve similar problems as Marxan and Zonation, and also allows the user to delineate several regions, just as Marxan with Zones. Prioritizr integrates feature representation criteria in the form of constraints (e.g. occurrence representation) or optimization objectives. For example, Prioritizr allows the maximization of phylogenetic diversity under budget restrictions. The package also integrates spatial criteria, that can apply as penalties in the objective function (as Marxan does) or as hard constraints. Currently, Prioritizr allows control over the perimeter of regions (boundary length) as penalty in the objective function, over connectivity of regions as penalty or hard constraint, and over the minimum number of neighbours of selected sites as hard constraint. Because Prioritizr relies on MILP, it can provide optimality and satisfiability guarantees to the problems it formulates.

LQGraph (Fuller and Sarkar, 2006). LQGraph is a software package specifically designed to delineate corridors between existing protected areas. From quality scores (e.g. habitat suitability) defined for each site outside protected areas, LQGraph designs optimal corridors that can inform managers on the extension of protected areas. LQGraph can also be used to identify sites that can efficiently isolate protected areas to prevent the spread of pathogens or invasive species.

Linkage Mapper (McRae et al., 2012). Similarly to LQGraph, Linkage Mapper is specifically designed to identify optimal corridors between habitat patches. It relies on a raster-based resistance model where each raster cell is characterized by a value which represents the ability of focal species to migrate through. From such a model, Linkage Mapper computes minimal resistance corridors between habitat patches.

2.3 TRENDS AND CHALLENGES

Since its early days, SCP has already produced many methodological and practical advances. Many tools are available and frequently used by scientists and practitioners. Nonetheless, there remains many directions to explore to improve the efficiency and adequacy of SCP for providing decision support in conservation actions. Although the following list is certainly not exhaustive, we tried to synthesize most of the main trends and challenges in SCP research.

Table 2.2: Summary of currently available state-of-the-art systematic conservation planning software packages (non-exhaustive list). MILP: Mixed-Integer Linear Programming, GUI: Graphical User Interface, CLI: Command-Line Interface, GIS: Geographic Information System.

	Technique	Interface	Short description
Marxan (Ball et al., 2009)	Simulated annealing	GUI/CLI	Target-based conservation planning with boundary length penalty.
Zonation (Moilanen et al., 2009b)	Backward heuristic	GUI/CLI	Hierarchical prioritization for land-use planning.
C-Plan (Pressey et al., 2005)	Heuristic	GIS (ArcView)	Interactive decision support for conservation planning.
ConsNet (Ciarleglio et al., 2009)	Tabu search	GUI/CLI	Target-based conservation planning with several spatial spatial criteria.
Conefor Sensinode (Saura and Torné, 2009)	Graph-based	GUI/CLI	Quantification of inter-patch connectivity through graph-based indices.
Prioritizr (Hanson et al., 2020)	MILP	R package	Exact MILP target-based conservation planning in R.
LQGraph (Fuller and Sarkar, 2006)	Graph-based	GUI	Design of optimal corridors between existing protected areas.
Linkage Mapper (McRae et al., 2012)	Least-cost models	GIS (ArcGIS)	Optimal corridors between habitat patches based on a least-cost resistance model.

These share two main objectives: (i) bring SCP as close as possible to ecological processes and evidences, (ii) encourage and facilitate the dissemination and use of SCP among conservation biologists and managers.

2.3.1 *Formalization and structuration of SCP problems.*

The diversity of SCP problems directly reflects the diversity of conservation problems and context, which involve at the same time biological and socio-economical considerations. As a result, providing a general formalism for SCP problems which does not oversimplifies real conservation questions is a daunting and difficult task. There were some attempts to provide such a formalism. For example, ReVelle et al. (2002) proposed five classes of zero-one programming reserve selection models as counterparts to classical facility location problems. Williams et al. (2005) extended these models from a review of spatial attributes in reserve design. On the other hand, without proposing a general formalism, several reviews focused on cores concepts and principles of SCP (e.g. Pressey et al., 1993; Margules and Pressey, 2000; Sarkar et al., 2006; Kukkala and Moilanen, 2013). Nevertheless, there is still progress to be made towards a general formalism of SCP problems. For example, many terms are commonly used to designate very similar to identical concepts, as it was already depicted by Kukkala and Moilanen (2013). “Reserve selection”, “reserve design”, “conservation area network design”, “spatial conservation prioritization”, “systematic conservation planning”, “conservation planning”; these terms have different historical and geographical origins and can slightly differ in their intended meaning but often describe the same problem. The term “systematic conservation planning” was an attempt to provide an integrative term for any problem related to the planning of conservation actions, but is still not systematically used in articles keywords. Another example is the overlapping use of “sites”, “parcels”, and “planning units” to designate the granular decision units in spatially-explicit conservation planning problem. Maintaining this diversity in SCP terminology is important for expressing nuances, but having a standardized entry point such as a glossary and a general formalism to which any SCP problem could relate without losing specificity would undoubtedly be beneficial to the field. First, it would provide a structured and common entry point to students, scientists and practitioners to SCP. Secondly, it would help structuring the scientific community around common modelling concepts and jargon. Finally, it would provide a standardized basis to compare existing methods from a functional point of view (e.g. which classes of problems can solve a method?) as well as from a technical perspective (e.g. provide benchmark instances to compare methods that actually address the same problems).

2.3.2 *The importance of spatial configuration.*

Although early work based on island biogeography theory was focused on geometrical characteristics of nature reserves (Diamond, 1975), SCP method have long been focused on the conservation of biodiversity through complementarity in the representation of biodiversity features (Sarkar, 2012). Spatial configuration is therefore widely recognized as an important characteristic of spatially-explicit SCP. The first approaches that integrated spatial configuration along with feature representation focused on the closeness of selected sites (Nicholls and Margules, 1993). Additional spatial criteria, such as connectivity or boundary length, were then defined and integrated in several SCP methods (see Section 2.2.2). Recent work introduced several new perspective in the integration of spatial configuration into SCP procedures. For example, the recent release of Marxan Connect (Daigle et al., 2020) provided many options to preprocess Marxan’s inputs with ecological connectivity considerations. Many MILP approaches also provided solutions to ensure the connectivity and compactness of delineated regions (Billionnet, 2016; Wang and Önal, 2016), to design buffer zones (Williams et al., 2005; Billionnet, 2013), or to design optimal corridors between habitat patches (Dilkina et al., 2017). However, many spatial configuration criteria remain to be integrated into SCP procedures. For example, it would be useful to provide a strict control over the spatial attributes of delineated regions, such as controlling precisely the number and size of connected components. Such a control could particularly be useful to integrate managers’ spatial constraints such as accessibility and economies of scale. Another perspective would be the delineation of nested regions which could allow, for example, the design of nested protected areas with different levels of protection, in line with IUCN guidelines for applying protected areas management categories (Dudley, 2008). Finally, many fragmentation and connectivity indices were designed in the field of Landscape Ecology (e.g.

McGarigal, 2014; Pascual-Hortal and Saura, 2006; Saura and Pascual-Hortal, 2007). These indices are currently used to evaluate the level of fragmentation and connectivity within existing landscapes. In conservation planning, these indices are also used in a scenario analysis context to evaluate the potential impact of actions defined a priori on overall fragmentation and connectivity patterns (e.g. Bodin and Saura, 2010; Tambosi et al., 2014). A useful development would consist in the integration of these indices in SCP optimization procedures along with other spatial and feature representation requirements, as it would help identifying optimal solutions among all possible conservation actions, which is not possible by defining solutions a priori. Such a development would, for instance, help identifying optimal areas to restore connectivity and reduce the impacts of fragmentation through ecological restoration.

2.3.3 *Expressiveness as a key factor for successful applications.*

There is a continued debate on the notion of optimality in SCP methods, which mainly opposes local search approaches with MILP (Underhill, 1994; Pressey et al., 1996; Rodrigues and Gaston, 2002; Hanson et al., 2019a). However, there were much fewer discussions around the concept of expressiveness in SCP methods (Rodrigues et al., 2000; Moilanen, 2008). In modelling, expressiveness refers to the breadth and variety of problems that can be represented and solved by a modelling approach. Each conservation planning problem comes with contextual specificities, such as focal species ecological requirements or managers' constraint. Expressiveness directly informs on the ability of an SCP method to reflect this contextual diversity. In this respect, it is a prerequisite to ensure that for a given problem, the chosen method is expressive enough to model it accurately. The question of optimality should come afterwards. The main bottleneck in the evaluation of expressiveness directly lies in the issue depicted in the first paragraph of this section (*Formalization and structuration of SCP problems*). Indeed, without a general formalism or nomenclature of SCP problems, identifying the most suitable approach for a given problem implies a minimal knowledge of every method. Nonetheless, it is already possible to state that among the criteria depicted in Section 2.2, many combinations cannot be modelled with current approaches. Among existing techniques for addressing and solving SCP problems, generic constrained optimization methods such as MILP have the advantage of being constraint-based declarative paradigms, which is an important feature to provide expressiveness. In this context, declarative means that the modelling of a problem is decoupled from its solving method through an abstract modelling language. This feature allows to primarily focus on what must be solved rather than describing how to solve it.

2.3.4 *Planning under uncertainty in a dynamic world.*

Uncertainty is omnipresent in ecology and can considerably affect the outcomes of conservation planning approaches, as does the temporal dimension, to which uncertainty is often related (e.g. species probability of persistence; Cabeza and Moilanen, 2001). However, these are difficult parameters to take into account in constrained optimization. Several studies provided solutions to integrate the temporal dimension and uncertainty into SCP procedures. First, many authors addressed the biodiversity feature representation problem from a probabilistic perspective, using probabilities of presence from species distribution models instead of binary occurrences (e.g. Haight et al., 2000; Polasky et al., 2000; Billionnet, 2011). These approaches relied either on greedy heuristics or MILP. Other approaches focused on the species-specific probabilities of persistence over time (e.g. Burgman et al., 2005; Matisziw and Murray, 2006; Schapaugh and Tyre, 2013, 2014), robustness to climate and land-use change (e.g. Carvalho et al., 2011; Schlottfeldt et al., 2015; Albert et al., 2017; Alagador and Cerdeira, 2017; Lin et al., 2019), or even control of invasive species (e.g. Hof, 1998; Haider et al., 2018). Another interesting example is the study of (Xue et al., 2017), who embedded spatial capture-recapture information to optimize landscape connectivity for a given species. Most of these approaches relied on dynamic stochastic programming, robust optimization, or Markov decision processes. Although these studies already brought many advances, approaches combining uncertainty along with the temporal dimension were mainly evaluated on specific use case. A potential and useful perspective would be the integration of these developments into generic SCP approach to widen the possible application cases. This is a difficult perspective, as it would certainly imply hybridizing classical approaches with techniques from stochastic dynamic programming, robust optimization, and Markov decision processes.

3

STUDY LOCATION AND RESEARCH OBJECTIVES

This PhD thesis was co-funded as a collaboration between the Cirad (The French agricultural research and international cooperation organization) and the IAC (New Caledonian Institute of Agronomy). It was motivated by New Caledonian environmental managers needs for decision support in forest conservation. Thus this PhD thesis aimed to bridge some gaps in Systematic Conservation Planning, with an emphasis on the application context of New Caledonian forest conservation. In this chapter, we introduce our study location, New Caledonia, and our research objectives.

3.1 NEW CALEDONIA, A BIODIVERSITY HOTSPOT IN THE SOUTH PACIFIC

3.1.1 *Geography.*

New Caledonia is a tropical archipelago located in the South-West Pacific, 130 km north of the Tropic of Capricorn, about 1400 km east of Australia, and about 2000 km of New Zealand (see Figure 3.1). New Caledonia covers a land area of 18 575 km^2, about 3 400 km of coastline, and one of the largest lagoon in world (about 24 000 km^2). The main island, Grande Terre, stretches from north-west to south-east for about 400km long and 50 to 70 km wide. The Loyalty islands are three islands located at the west of the main island. Ouvéa, with a land area of 131 km^2 and a beach of 25 km, is the northernmost of these islands. Maré, with a land area of 641 km^2, is the southernmost. Finally, Lifou is the largest of Loyalty islands with a land area of 1207 km^2. The archipelago also comprises smaller islands, such as the Isle of Pines and Belep islands respectively located south and north of the main island (see the map of New Caledonia in Figure 3.2).

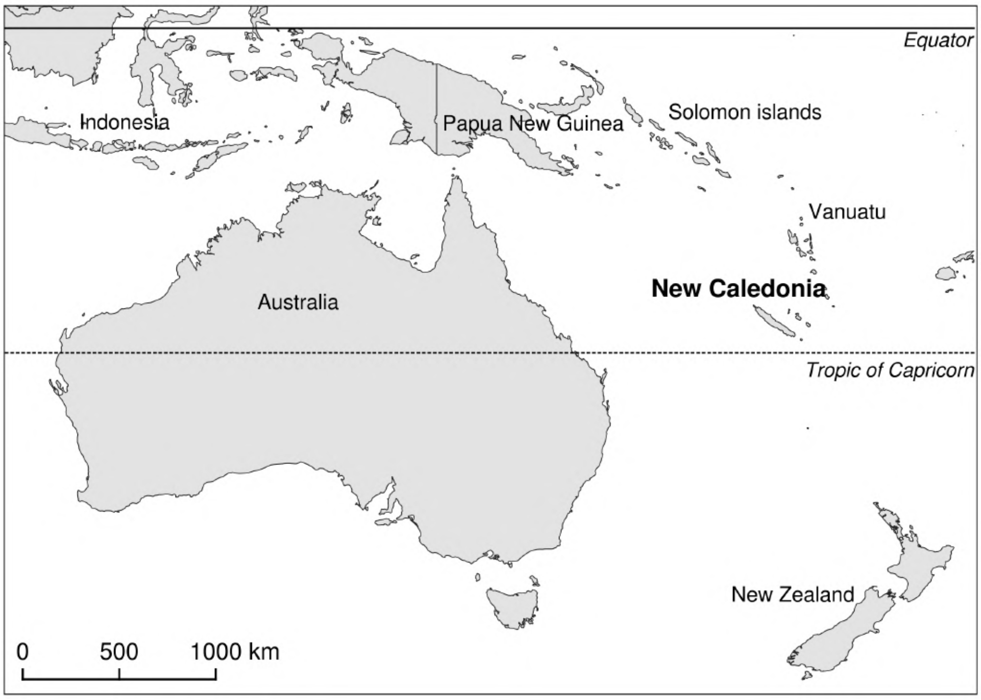


Figure 3.1: Location of New Caledonia.

3.1.2 *Orography and geology.*

The main island, Grande Terre, is traversed by a mountain range whose highest mounts are the Panié mount in the north and the Humboldt mount in the south (respectively 1628 m and 1618 m, see Figure 3.2). This mountain range forms a natural frontier between the large plains of the west coast, and the

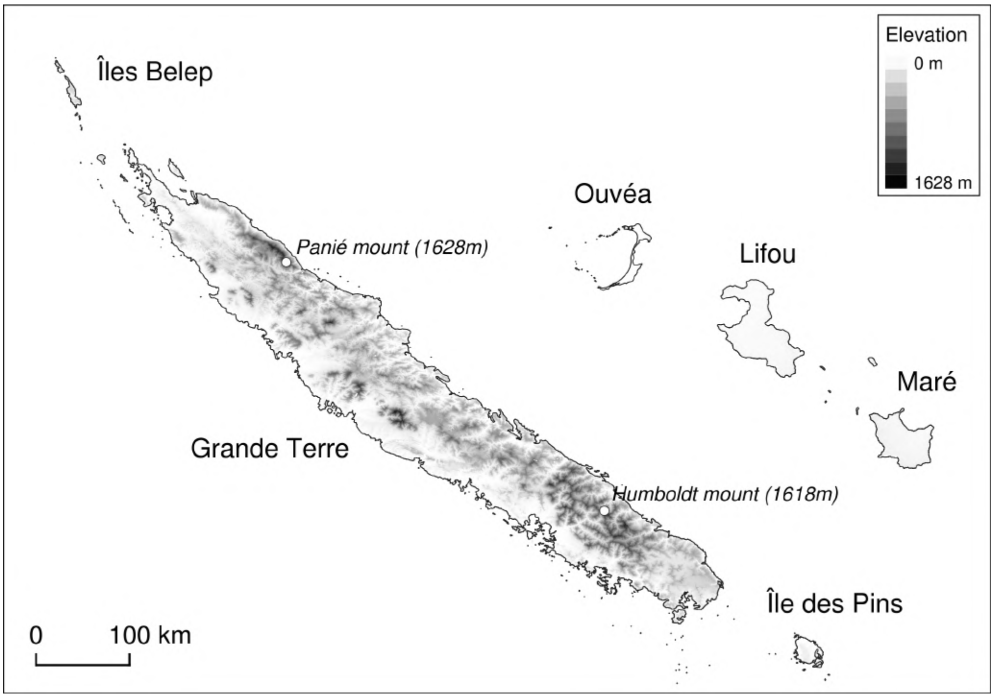


Figure 3.2: New Caledonia main islands, elevation, and highest mounts. Digital elevation model: ©Geomatics and Remote Sensing Service - DTSI - Government of New Caledonia, CC BY-NC-SA 4.0.

steep coastal landscape of the east coast. Other islands of the archipelago are much less mountainous. There are three main types of soils in New Caledonia: (i) Volcano-sedimentary soils, they cover about two third of the Grande Terre, mainly in the north. (ii) Ultramafic soils, they cover about one third of the Grande Terre, mainly in the south and in some isolated massifs of the north. They are also present in Belep islands and in the Isle of Pines. (iii) Calcareous soils, mainly present in Loyalty islands.

3.1.3 *Climate.*

New Caledonia has a tropical climate, with a hot and humid season from November to April (average maximum temperatures between 28°C and 32°C) and a cooler and dry season from May to September (average maximum temperatures between 24°C and 29°C) (Maitrepierre, 2012). The hot season is also called the cyclone season, due to the high frequency of tropical depressions and cyclones in the region during this period. The archipelago is subject to south-east trade winds which, according to the topography of the Grande Terre, result in a high variation of mean annual precipitations: from 800 mm per year in the western coastal plains to 4500 mm per year in the eastern coast.

3.1.4 *History and politics.*

Earliest traces of human occupation in New Caledonia were traced back about 3000 years in the past, with pottery fragments from the Lapita civilization. Until its discovery by the British explorer James Cook in 1774, the archipelago was only inhabited by the Kanaks, a Melanesian people. New Caledonia has been colonized by France from 1853, which established its capital city, Nouméa, in the south-west of the main island. New Caledonia was first a penal colony. After their imprisonment, deported prisoners were given land to settle and practice agriculture. These lands were taken from Kanak people, which until 1944 were subject to the "code de l'indigénat" and not considered as citizens. This unfair and unequal treatment led the Kanak people to revolt, and New Caledonian history has until very recently been marked by numerous violent confrontations. In 1988, as the territory was on the brink of civil war after a hostage taking in the island of Ouvéa, the Matignon-Oudinot Agreements were signed as a first step towards autonomy, independence, and better recognition of the Kanak culture and customary rights (Rocard, 1988). Ten years later, the Nouméa Accord was the beginning of a slow transition to autonomy, with a gradual competence transfer from the French government to the New Caledonian government (Jospin, 1998). The next step of this process is the self-determination referendum of the New Caledonian people for independence. The result of a first vote in 2018 was against independence, but another vote will occur in 2020, and in 2022 if the result is still against independence, as stated by the Nouméa Accord. Today, New Caledonia is divided into three provinces: the South Province, the North Province, and the Loyalty islands Province. The French Common Civil Code now coexists with the Customary Civil Code, and institutions such as the Customary Senate provide a political framework to the Kanak people for promoting their culture, traditions, and environment.

3.1.5 *Biodiversity, terrestrial flora and forests.*

As the smallest biodiversity hotspot in the world (Myers, 1988), New Caledonia hosts megadiverse marine and terrestrial ecosystems. Despite an ancient misconception according to which New Caledonian terrestrial fauna would be depauperate, the archipelago hosts a rich and unique fauna whose knowledge is still highly incomplete (Chazeau, 1993). Endemic species include chiroptera (e.g. the ornate flying fox, *Pteropus ornatus*), birds (e.g. the Cagou, *Rhynochetos jubatus*), geckos and skinks (e.g. the chameleon gecko, *Eurydactylodes vieillardi*, see Figure 3.3), freshwater fishes, invertebrates, insects (see Figure 3.4), etc.
On the other hand, New Caledonian terrestrial flora has long been recognized as exceptionally rich and unique, with good reasons. Indeed, New Caledonia flora distinguishes by one of the highest rates of endemism in the world, approximately 76% among more than 3400 vascular plant species (Myers et al., 2000; Morat et al., 2012), a high beta-diversity (Ibanez et al., 2014; Isnard et al., 2016), and the presence of relict taxa (Grandcolas et al., 2008; Pillon,

Figure 3.3: *Eurydactylodes vieillardi*, the chameleon gecko, endemic to New Caledonian rainforests. Picture taken at the Giant Fern Park during an excursion guided by Matthias Deuss, Farino, New Caledonia. ©Dimitri Justeau-Allaire.

Figure 3.4: New Caledonian phasma, unidentified species. Picture taken in the wilderness reserve of the Aoupinie massif in New Caledonia. ©Dimitri Justeau-Allaire.

2012). Angiosperms represent about 91% of New Caledonian plant species and gymnosperms are also very diversified, with 46 endemic species (Morat et al., 2012; Jaffré et al., 1994). This high level of endemism also reflects at higher taxonomic levels, as the archipelago harbours three endemic families (Morat et al., 2012) and between 62 and 91 endemic genera (Pillon et al., 2017). Some New Caledonian floristic curiosities are described in Boxes 3.1, 3.2, and 3.3. Several types of vegetation are present in New Caledonia: (i) climacic rainforests (see Figure 3.5), (ii) climacic sclerophyll (or dry) forests, (iii) low to high elevation shrublands that comprise both climacic and secondary formations, (iv) halophytic formations (mangrove and littoral vegetations), and (v) savanna and thickets that are both secondary formations (Jaffré et al., 1994; Jaffre et al., 1998). Rainforest is the richest vegetation type, with more than 2000 native vascular plant species and 82.4% of endemism for a total surface of about 4000 km^2, with one third located on ultramafic soils (Jaffré et al., 1994; Birnbaum et al., 2015). With the lowest surface (about 100 km^2), high altitude shrubland is therefore the vegetation type presenting the highest rate of endemism in New Caledonia (91%; Jaffré et al., 1994). Although it was shown that most rainforest tree species are ubiquist regarding substrate, rainfall and elevation (Birnbaum et al., 2015), many plant species are restricted to a single substrate (Ibanez et al., 2014), and some of them are narrow endemic to very small areas (Wulff et al., 2013). These specificities result in a mosaic of vegetal formations with diverse and unique floristic compositions, which has important implication for the conservation of New Caledonian forests and flora.

Figure 3.5: Fragmented montane rainforest in Arago, New Caledonia. ©Nicolas Petit.

3.2 CONTEXT: CONSERVATION OF NEW CALEDONIAN FORESTS

Forest conservation in New Caledonia poses great challenges, as human activities have profound and negative impacts on New Caledonian forests. The current main threats are nickel mining, bushfire, and invasive species. Because ultramafic substrates cover about 30% of the archipelago (against only 3% of world land area), nickel mining had quickly become the pillar of New Caledonian economy, at the price of a significant degradation and loss of forest. In addition, fires severely affect several types of vegetation in New Caledonia, where it is estimated that 99% of fires are triggered by humans and that every year around 27 000 ha of vegetation is burned (Hély-Alleaune, 2012). Finally, although invasive plant species have been restricted to disturbed areas, native vegetation is severely impacted by invasive herbivorous, and

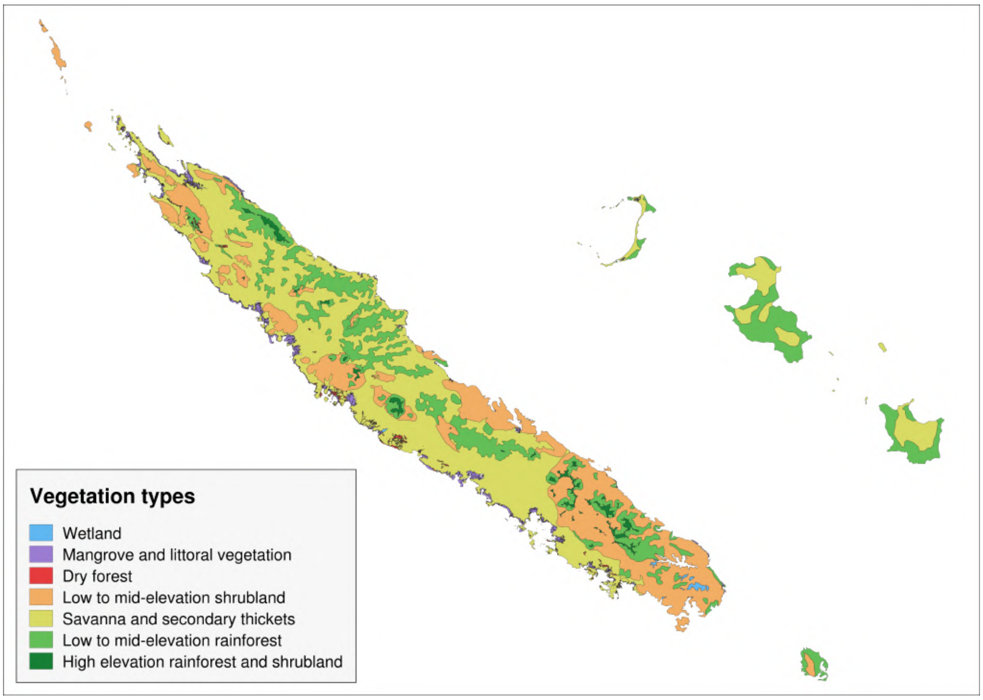


Figure 3.6: Distribution of vegetation types in New Caledonia. Data from the Atlas of New Caledonia (Jaffré et al., 2012).

mostly by the deer (De Garine-Wichatitsky et al., 2005). Fragmented forest areas in ultramafic substrates are particularly vulnerable to these threats that are amplified by negative edge effects (Ibanez et al., 2017).

In New Caledonia, conservation management is a provincial responsibility, to the exception of the Natural Park of the Coral Sea which is managed by the New Caledonian government. In this respect, each province has its own environmental code and protected areas categories. In North Province there are six of them, directly inspired by the IUCN protected areas categories (Dudley, 2008), from the most to the least protected: integral nature reserves (IUCN category Ia), wilderness reserves (Ib), provincial parks (II), natural reserves (IV), areas for the protection and valuation of natural and cultural heritage (V), and sustainable resources management areas (VI). In South Province, there are four protected areas categories, from the most to the least protected: integral nature reserves, natural reserves, sustainable resources management areas, and provincial parks. Finally in the Loyalty islands Province, which is mainly covered by customary lands, no protected areas were created. Instead, the Loyalty islands favoured a policy close to traditional and customary management methods. Indeed, the Loyalty islands environmental code claims that "The natural environment is inseparable from cultural practices and locally applicable customary rules. [...] The province of the Loyalty Islands takes into account the existence of customary modes of environmental management and integrates these modes of management into the regulations, while respecting the principle of subsidiarity." (Article 110-1; Citré et al., 2019). Although this provincial management favours diversity and local specificities, it can also lead to some confusions. For example, in South Province, a provincial park corresponds to a much lower level of protection than in North Province.

About twenty years ago, Jaffre et al. (1998) warned that the existing protected areas network in New Caledonia was inadequate and not sufficient to protect New Caledonian flora and forests. Notably, they have shown that 83% of the threatened plant species did not occur in a protected area, and that only 54% of protected area were under strict mining restrictions. Recently, Ibanez et al. (2018) conducted a study to evaluate the improvements and shortcomings of the protected areas network in New Caledonia, twenty years after Jaffre et al. (1998). They found that even if the terrestrial protected areas had increased by 35% in surface, they covered only 4% of New Caledonia, which was far below the Aichi target 11 of at least 17% of terrestrial protected areas by 2020 (CBD, 2010) (it should nonetheless be noted that since this study the South Province extended its protected areas network with the creation of the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" in April 2019 and reached the Aichi target 11). Moreover, 72% of the threatened plant species still did not occur in any protected area in 2017 (Endemia and NC, 2017). Conversely, the rate of nickel extraction has doubled between 1998 and 2018 (Ibanez et al., 2018).

In conclusion, there remain many challenges to improve the conservation of New Caledonian flora and forests. One obvious way to achieve this goal is to improve the adequacy of the protected areas network, but other conservation actions such as ecological restoration, ex-situ conservation, or reinforcement of management staff (Ibanez et al., 2018). New Caledonia has strong financial and administrative attributes, and the insular and low populated New Caledonian context allows high proximity between conservation stakeholders. These specificities makes it an appropriate field of study to experiment novel approaches. More particularly, several projects involving collaboration with research institutes and environmental managers (e.g. CORIFOR project, Birnbaum et al., 2016; COGEFOR project, Birnbaum et al., 2019) have highlighted the need to benefit from evidence-based decision support tools based on scientific knowledge and able to integrate conservation goals along with managers' constraints.

Box 3.1: *Parasitaxus usta*, the only parasitic conifer known in the world.

Parasitaxus usta is the sole species of the genus *Parasitaxus*, from the *Podocarpaceae* family. It is a rare woody shrub endemic to New Caledonian forests, which distinguishes by being the only know parasitic gymnosperm in the world. *Parasitaxus usta* lives in the undergrowth and can be found between 100 and 1000 m of elevation in undisturbed forests. It does not develop roots, but instead attaches to the roots or to the base of the trunk of its host species, *Falcatifolium taxoides*, another *Podocarpaceae* endemic to New Caledonia. *Parasitaxus usta* is classified as vulnerable (VU) in the IUCN Red List of threatened species (IUCN, 2009).

Figure 3.7: *Parasitaxus usta*, detail of stems with cones. Picture taken at mount Dzumac by ©Gildas Gateble.

Box 3.2: *Cerberiopsis candelabra*, the tree that blooms only once.

Cerberiopsis candelabra is a tree species in the family *Apocynaceae*, from the genus *Cerberiopsis* which is endemic to New Caledonia. It lives in low to mid-elevation rainforests and can reach 20-30m height. *Cerberiopsis candelabra* is a monocarpic tree, which means that it blooms only once in its life and dies. Monocarpy is a characteristic of several species among several families. However, the majority of these species are monocaulous (i.e. with a single unbranched trunk or stem). Monocarpy in long-lived ramified trees such as *Cerberiopsis candelabra* is a rare (it was observed in less than 10 species among only 3 genus in the world) and intriguing character that is still poorly understood, as it seems risky at first glance. Was it an adaptive strategy or an evolutionary coincidence which did not affect the long term fitness of the species? The question remains open (Read et al., 2008).

Figure 3.8: *Cerberiopsis candelabra*, blooming (left) and from the base (right). ©Camille Salmon and ©Patrick Heuret.

Box 3.3: *Pycnandra acuminata*, the nickel tree.

Pycnandra acuminata is tree species from the family *Sapotaceae* and the genus *Pycnandra*, which is the largest endemic angiosperms genus of New Caledonia. It is a rainforest species which grows in nickel-rich ultramafic soils. The main particularity of *Pycnandra acuminata* is its ability to accumulate high rates of nickels. Up to 25 % of nickel citrate as dry weight can be found in its latex, which makes it one of the most prolific hyperaccumulators known in the world. This high concentration of nickel gives a distinctive blue-green colour to its latex (see Figure 3.9), which earned it its vernacular name: "sève bleue" (blue sap). Hyperaccumulators such as *Pycnandra acuminata* are notably studied for their potential in agromining and mine rehabilitation projects (Erskine et al., 2018).

Figure 3.9: *Pycnandra acuminata*, stem with the famous blue-green latex, which can contain up to 25 % of nickel citrate as dry weight. Picture taken by ©Bernard Suprin.

3.3 RESEARCH OBJECTIVES

With a focus on New Caledonian needs in decision support tools for forest conservation planning, this PhD thesis organizes around three interlinked objectives.

3.3.1 *Improving the expressiveness of systematic conservation planning.*

Given the specific socio-economic context of New Caledonia, the heterogeneity of its flora's distribution, and the negative impact of fragmentation on its forests (Ibanez et al., 2017), conservation planning exercises in this area are heterogeneous and must take into account many constraints involving many stakeholders. In such a context, expressiveness of decision support approaches is more than ever a critical factor for successful applications. Accordingly, the first objective of this PhD thesis is to address the expressiveness challenge depicted in Section 2.3.3 and provide a practical, accountable, and expressive approach for modelling and solving conservation planning problems. This approach will be based on constraint programming, a formal constrained optimization technique that will be described in Chapter 4.

3.3.2 *Providing more control over the spatial configuration of solutions.*

Planning against forest fragmentation requires a precise control over the spatial configuration of produced solutions. Likewise, managers constraints in New Caledonia are not limited to budget and often require to take into account spatial aspects such as mining concessions, customary areas, accessibility in low urbanized and mountainous areas, etc. In this respect, a second objective addresses the challenge depicted in Section 2.3.2 to allow a precise control over spatial configuration in the expressive approach targeted by the previous objective, notably through fragmentation and connectivity indices.

3.3.3 *Getting involved in decision support in New Caledonia.*

As a baptism of fire, the third objective of this PhD thesis is to confront the approach resulting from the previous objectives to a real world conservation problem in New Caledonia. In this respect, we engaged a collaboration with the environmental managers of the South Province of New Caledonia in a current reforestation planning project in the Côte Oubliée – 'Woen Vùù – Pwa Pereeù provincial park, a symbolic area in New Caledonia with important conservation issues at stake.

Part II

A CONSTRAINT-BASED APPROACH TO SYSTEMATIC CONSERVATION PLANNING

4

INTRODUCTION TO CONSTRAINT PROGRAMMING

Constraint Programming is a formal paradigm for solving combinatorial problems which appeared in the 1970s and 1980s. Drawing on a wide range of techniques from artificial intelligence, operations research, digital analysis, and symbolic calculations, Constraint Programming heavily relies on concepts from mathematical logic, set theory and graph theory. In this chapter, we introduce Constraint Programming along with its mathematical prerequisites.

4.1 MATHEMATICAL PREREQUISITES

4.1.1 *Mathematical logic*

Mathematical logic is a branch of mathematics introduced at the end of the 19th century, to study mathematics as a formal language. This discipline played a critical role in the emergence of computer science, and both fields are still related. Among the pioneers of mathematical logic, we can cite Bertrand Russel and Alfred North Whitehead, who published a foundational book on the field (Whitehead and Russel, 1910). Interested readers can also read *Logicomix*, a playful and instructive comic book focused on Russel's quest for mathematics foundation (Doxiadis and Papadimitriou, 2015). Formulas are the fundamental objects of mathematical logic which represent mathematical statements with a formal language. *Propositions* are formulas that express mathematical facts (either true or false) and *predicates* are formulas expressing mathematical properties relying on the concept of *variable*. A variable is an object of a particular type (e.g. integer) whose possible values are defined by a *domain*. Predicates with a single type of variable are qualified as *first-order*, predicates with two types of variables as *second-order*, etc. The main elements of the logical formal language for propositions and predicates are respectively called *connectors* and *quantifiers*. A connector is an operator on propositions. It can imply one proposition (unary connector), two propositions (binary connector), or n propositions (n-ary connector). On the other hand, a quantifier is an operator on predicates, which imply variables from a given domain and a proposition. In the following, we introduce the main connectors and quantifiers.

Connector 1 (Negation $\neg$). Let P be a proposition, the negation of P, denoted by $\neg P$, is a proposition which is true if and only if P is false.

Connector 2 (Disjunction $\vee$). Let P and Q be two propositions, the disjunction of P and Q, denoted by $P \vee Q$, is a proposition which is true if and only if at least one of the two propositions is true.

Connector 3 (Conjunction $\wedge$). Let P and Q be two propositions, the conjunction of P and Q, denoted by $P \wedge Q$, is a proposition which is true if and only if P and Q are both true.

Connector 4 (Implication $\Rightarrow$). Let P and Q be two propositions, the implication of Q by P, denoted by $P \Rightarrow Q$, is a proposition which is true if and only if $(\neg P) \vee Q$ is true, namely if P is true then Q is true (but if Q is true, P is not necessarily true).

Connector 5 (Equivalence $\Leftrightarrow$). Let P and Q be two propositions, the equivalence of Q by P, denoted by $P \Leftrightarrow Q$, is a proposition which is true if and only $P \Rightarrow Q \wedge Q \Rightarrow P$ is true, that is P and Q necessarily have the same value.

Quantifier 1 (The existential quantifier $\exists$). Let x be a variable from a domain $\mathbb{X}$, and P a proposition. $\exists x \in \mathbb{X}, P$ must be read as "there exist at least one x in $\mathbb{X}$

such that P is true". For example, if x is a natural integer and $P = x > 3$, we can write $\exists x \in \mathbb{N}, x > 3$.

Quantifier 2 (The universal quantifier $\forall$). Let x be a variable from a domain $\mathbb{X}$, and P a proposition, $\forall x \in \mathbb{X}, P$ must be read "for every x in $\mathbb{X}$, P is true". For example, if x is a natural integer and $P = x \geqslant 0$, we can write $\forall x \in \mathbb{N}, x \geqslant 0$.

4.1.2 *Set theory*

Set theory is a branch of mathematical logic which developed at the end of the 19th century, with the impulse of the mathematician Georg Cantor. It focuses on the study of sets, that are collections of objects (e.g. integers, functions). Sets are heavily used in theoretical computer science, as they offer an abstract and expressive formal language to model combinatorial problems. Sets can be described in an *extensional* or *intentional* way. The extensional (or Roster) notation consists of enumerating all elements of the set between curly brackets. For example, the extensional notation of the set A containing all integers between 1 and 5 is $A = \{1, 2, 3, 4, 5\}$. On the other hand, the intentional (or set-builder) notation consists of defining the elements of the set with a predicate. For example, the intensional notation of the set A containing all integers between 1 and 5 is $A = \{n \mid n \in \mathbb{N} \land 1 \leqslant n \leqslant 5\ \}$. Note that for some object types such as numbers, it is also possible to use interval notations. For example, the set of all integers between 1 and 5 can be denoted by $[\![1, 5]\!]$. Finally, the unique set having no elements is called the *empty set* and is denoted by $\emptyset$. There are two fundamental relations on sets: *membership* and *inclusion*. Membership is a binary relation between an object x and a set A. If x is an element (or member) of the set A then the notation $x \in A$ is used. Otherwise, the notation $x \notin A$ is used. Inclusion is a binary relation between two sets. If all the elements of the set A are also elements of the set B, then A is a subset of B, denoted by $A \subseteq B$. If A is a subset of B but not equal to B, then the notation $A \subset B$ is used. Sets can be manipulated to construct new sets using specific operators. In the following, we describe the main set operators.

Set operator 1 (Union $\cup$). The set obtained with all the elements of the set A and the set B is the union of A and B, denoted by $A \cup B$. This operator can be generalized for n sets with the notation $\bigcup_{i \in [0,n[} A_i$, where each A_i is a set.

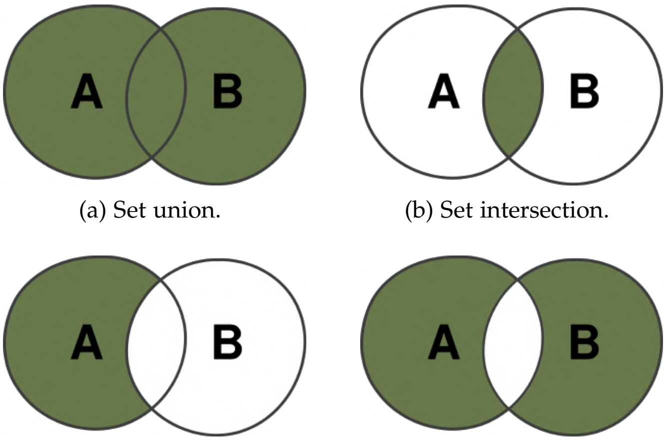
Set operator 2 (Intersection $\cap$). The set obtained by retaining only objects that are both elements of a set A and a set B is the intersection of A and B, denoted by $A \cap B$. This operator can also be generalized for n sets with the notation $\bigcap_{i \in [0,n[} A_i$.

Set operator 3 (Difference $\setminus$). The set obtained by removing every element of the set B from the set A is the difference of A and B, denoted by $A \setminus B$.

Set operator 4 (Symmetric difference $\triangle$). The set obtained by retaining elements of the set A that are not elements of the set B and elements of the set B that are not elements of the set A is the symmetric difference of A and B, denoted by $A \triangle B$.

Set operator 5 (Power set $\mathcal{P}$). The set of all possible subsets of the set A is the power set of A, denoted by $\mathcal{P}(A)$.

Set operator 6 (Cartesian product $\times$). The set of all possible ordered pairs between the set A and the set B is the Cartesian product of A and B, denoted by $A \times B$. For example, if $A = \{1, 2\}$ and $B = \{3, 4\}$, then $A \times B = \{(1, 3), (1, 4), (2, 3), (2, 4)\}$.



(a) Set union. (b) Set intersection.

(c) Set difference. (d) Set symmetric difference.

Figure 4.1: Illustration of set operators (union, intersection, difference and symmetric difference.

Additional definitions. We provide three additional definitions commonly used in set theory. First, the *cardinality* of a set A, denoted by $|A|$, is the number of elements of A. Secondly, two sets A and B are said *disjoints* if and only if their intersection is empty, that is $A \cap B = \emptyset$. Finally, a *partition* of a set A is a set of non-empty subsets of A that are all pairwise disjoint and such that the union of all these subsets is A.

4.1.3 *Graph theory*

Graph theory is a branch of discrete mathematics focused on the study of *graphs*, that are mathematical structures representing pairwise relations between objects. The notion of graphs has been introduced at the end of the 19th century. However, origins of graph theory can be traced back to 1741 with the work of Euler (1741) on the seven bridges of Königsberg problem. The city of Königsberg in the former Prussia (now Kaliningrad in Russia) was built around two islands in the Pregel river. The seven bridges of Königsberg problem consisted in determining whether it is possible to start a walk from one location and return to that location by crossing each of these bridges once and only once (see Figure 4.2). Euler has proved that this problem has no solutions. In the following, we introduce several concepts and definitions of graph theory, interested readers can go further by reading Berge (1973) or Diestel (1997).

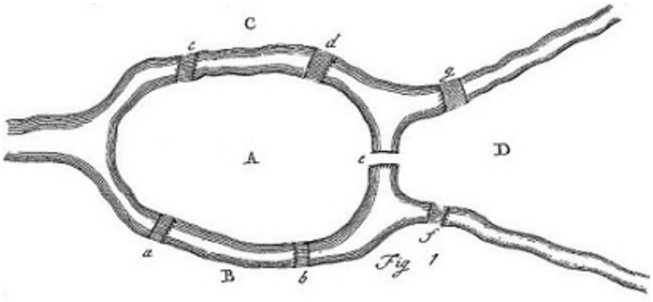
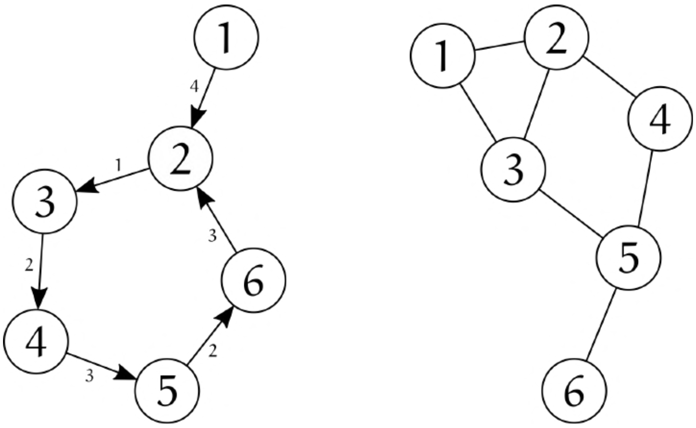


Figure 4.2: Original illustration of the seven bridges of Königsberg, from Euler (1741).

What is a graph? A graph (sometimes also called *network*) is a pair $G = (V, E)$ of sets, such that $E \subseteq V \times V$. The elements of V are the *vertices* (or *nodes*) and the elements of E are the *edges* (or *arcs*, or *links*) representing connections between vertices. A graph is said *directed* if its edges represent unilateral connections (i.e. there is a starting vertex and an ending vertex) and *undirected* if its edges represent bilateral connections. Note that the term *arcs* is more commonly used for directed graphs. Moreover, a graph is said *weighted* if its edges are labelled with a numerical value, *unweighted* otherwise. A graph can be graphically represented with lines and circles (see Figure 4.3), but it is important to remember that the structure of a graph is completely independent of its graphical representations.



(a) Directed weighted graph. (b) Undirected unweighted graph.

Figure 4.3: Illustration of a directed weighted graph and an undirected unweighted graph.

What are graphs useful for? Graph theory is a powerful modelling tool which is useful to address problems involving relationships between objects. The main strength of graph theory is its close relationship with computer science and operational research, which has led to the emergence of many efficient graph algorithms exploited in many real-world applications. Graphs have been intensively used in computer science to represent networks or computation processes, for example. They also have been used in linguistics, physics, chemistry, social sciences (to model and analyse social networks), biology (for DNA sequencing for example), etc.

In the following, we introduce some classical definitions from graph theory that will be useful in the next section and chapters of this PhD thesis.

Definition 4.1 (Subgraph)**.** Let $G = (V, E)$ be a graph. The graph $G' = (V', E')$ is a *subgraph* of G if and only if $V' \subseteq V$. This property is denoted by $G' \subseteq G$. If G' is such that E' contains all edges from E that have both endpoints in V', G' is called an induced subgraph of G, and is denoted by $G[V']$. Otherwise, G' is said to be a *non-induced subgraph* of G. If $V' = V \wedge E' \subset E$, G' is said to be a *spanning subgraph* of G.

Definition 4.2 (Path)**.** Let $G = (V, E)$ be a graph. A path is a sequence of edges from E joining a sequence of vertices from V. The number of edges of a path is its *length*. For example, the sequence $\{(1, 2), (2, 3), (3, 4), (4, 5), (5, 6)\}$ in the graph illustrated in Figure 4.3.a is a path.

Definition 4.3 (Cycle)**.** Let $G = (V, E)$ be a graph. A *cycle* is a non-empty path of G in which the first and last vertices are identical. If the only repeated vertices are the first and the last vertices, the cycle is *elementary*. For example, $\{(1, 2), (2, 3), (3, 1)\}$ is an elementary cycle in the graph illustrated in Figure 4.3.b.

Definition 4.4 (Connectivity)**.** Let $G = (V, E)$ be an undirected graph, G is *connected* if any pair of its vertices are linked by a path. Otherwise, it is *disconnected*. For example, the undirected graph illustrated in Figure 4.3.b is connected. If G is directed, G is *weakly connected* if replacing its arcs by undirected edges results in a connected graph and *strongly connected* if any pair of its vertices are linked by a directed path. For example, the directed graph illustrated in Figure 4.3. is weakly connected, as vertex 1 is not reachable from any other vertex.

Definition 4.5 (Connected components)**.** Let $G = (V, E)$ be a disconnected undirected graph, a maximal connected subgraph of G is a *connected component*. The set of connected components of G is denoted by $cc(G)$. If G is directed, a strongly (resp. weakly) connected subgraph of G is a strongly (reps. weakly) connected component.

Definition 4.6 (Tree and forest)**.** Let $G = (V, E)$ be a graph. If G is said to be a *tree* if it is connected and contains no cycle. If G is disconnected and contains no cycle, it is a *forest* and each of its connected components is a tree. See Figure 4.4 for an example.

Definition 4.7 (Planar graph)**.** A graph is *planar* if it can be drawn in a 2D plane such that no edges intersect. A drawing of a planar graph is called a *planar embedding* of the graph. For example, both graphs in Figure 4.3 are planar.

Definition 4.8 (Flow network)**.** A *flow network* is a directed graph where each edge is labelled by *capacity* value and optionally a *demand* value. The

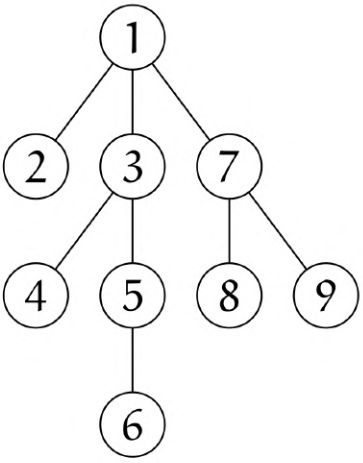


Figure 4.4: Illustration of a tree.

capacity (resp. demand) of an edge (x, y) is denoted by $u(x, y)$ (resp. $l(x, y)$). If $(x, y) \notin E$, then $u(x, y) = l(x, y) = 0$. Intuitively, such a graph can be seen as a network where a flow (e.g. water, electricity, data) can circulate between vertices according to capacity and demand values. Most often, flow networks contain two special vertices: the *source* where the flow starts and the *sink* where the flow ends. Note that it is also possible to have flow networks with several sources and sinks. See Figure 4.5 for an example.

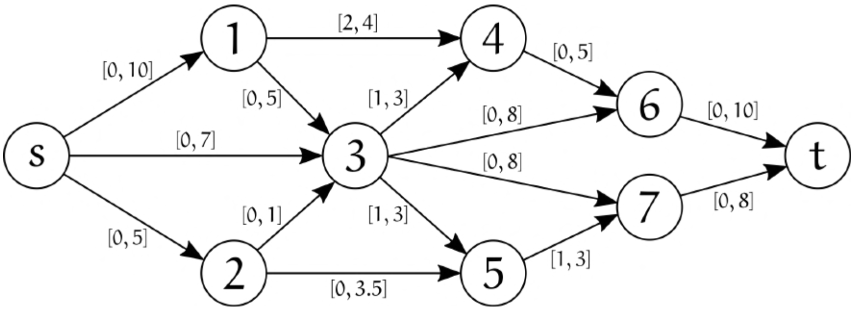


Figure 4.5: Flow network with edges demands and capacities represented as intervals. Source and sink vertices are respectively labelled with s and t.

Definition 4.9 (Flow)**.** Let $G = (V, E)$ be a flow network. A *flow* (or *feasible flow*) in G is a function $f : V \times V \mapsto \mathbb{R}$ such that:

(i) $\forall (x, y) \in V \times V$, $l(x, y) \leqslant f(x, y) \leqslant u(x, y)$, that is the flow satisfies demand and does not exceed capacity on each edge.

(ii) $\forall (x, y) \in V \times V$, $f(x, y) = -f(y, x)$, that is the flow circulating from a vertex x to a vertex y must be the opposite of the flow circulating from y to x (f is anti-symmetric).

(iii) $\forall x \in V \setminus \{s, t\}$, $\sum_{y \in V} f(x, y) = 0$ that is the net relative flow entering a vertex is zero (*flow conservation*, or *Kirchoff's law*).

Definition 4.10 (Bipartite graph)**.** A *bipartite graph* is a graph whose vertices can be partitioned into two sets U and V such that every edge has an end in U and an end in V. Formally, a bipartite graph G is denoted by $G = (U, V, E)$ and is such that $\forall (x, y) \in E$, $(x \in U \wedge y \in V) \vee (x \in V \wedge y \in U)$.

Definition 4.11 (Matching in a bipartite graph)**.** Let $G = (U, V, E)$ be a bipartite graph. A set of edges M such that no two edges share a same endpoint is a *matching* (edges in M are said *independent*). A *maximum cardinality matching* in G is such that there is no matching of strictly greater cardinality in G. See Figure 4.6 for an illustration.

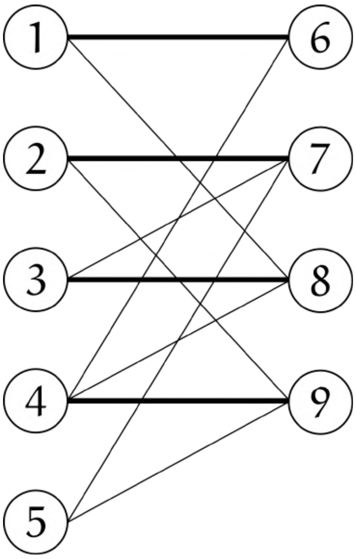


Figure 4.6: Illustration of a bipartite graph. The set of thick edges is a maximum cardinality matching.

4.2 CONSTRAINT PROGRAMMING

Constraint programming (CP) is a paradigm for modelling and solving constraint satisfaction and constrained optimization problems (respectively CSPs and COPs). The origins of CP can be first traced back to the work of logicians (see Section 4.1.1), and then to several works on constraint networks (Montanari, 1974), constraint consistency (Mackworth, 1977), symbolic representation of constraint satisfaction problems (Lauriere, 1978), and logic programming (Van Hentenryck et al., 1992). One particularity of CP is that it is a declarative paradigm, which means that the modelling of a problem is completely decoupled from its solving method, allowing the user to primarily focus on *what* must be solved rather than *how* to solve it. The central idea of CP is the declaration of *variables* and *constraints* representing logical relations between variables which must be satisfied by every solution of the problem modelled. In this sense, CP's paradigm is close to the Holy Grail of computer science: "the user states the problem, the computer solves it" (Freuder, 1996). In the following, we introduce the main concepts and principles of CP, interested reader can go further by reading the *Handbook of Constraint Programming* (Rossi et al., 2006). Before providing more details on variables, constraints, and the solving techniques used in CP, we introduce formal definitions of a CSP and a COP.

Definition 4.12 (Constraint satisfaction problem)**.** A *constraint satisfaction problem* (CSP) is a triplet $(\mathcal{V}, \mathcal{D}, \mathcal{C})$ such that:

- $\mathcal{V}$ is a set of n variables $\{v_1, ..., v_n\}$ that take their values in a finite set of objects from a certain type (e.g. integers). The term "variable" here and after refers the logical definition introduced in Section 4.1.1.
- $\mathcal{D}$ is the set of finite domains in which variables from $\mathcal{V}$ can take their values. The domain of a variable $v_i \in \mathcal{V}$ is denoted by $\mathcal{D}(v_i)$.
- $\mathcal{C}$ is a set of constraints, that are logical relations between subsets of $\mathcal{V}$.

A valid solution of a CSP is a tuple $(a_1, ..., a_n)$ of values such that $\forall i \in [1, n]$, $a_i \in \mathcal{D}(v_i)$ and such that $\forall c \in \mathcal{C}$, c is satisfied.

Definition 4.13 (Constrained optimization problem)**.** A *constrained optimization problem* (COP) is a CSP where one or several variables from $\mathcal{V}$ are defined as optimization objectives. If there is only one variable to maximize (resp. minimize) it is a *mono-objective* optimization problem whose solutions correspond to the subset of solutions of the associated CSP presenting the highest (resp. lowest) value for the objective variable. If there are several variables to maximize (resp. minimize) it is a *mult-objective* optimization problem whose solutions correspond to the subsect of solutions of the associated CSP such that none of the objective variables can be improved without degrading another. Such solutions are said *Pareto-optimal*.

4.2.1 *What is a variable in constraint programming?*

In the context of constraint programming, variables constitute the unknowns of the problem. The aim is to assign a value to each problem's variable such that all constraints are satisfied. The domain of a variable represents its possible values and is a finite set of objects. Domains can be defined extensively by individually enumerating all possible values or intensively by defining an interval of possible values. This abstract representation of variables has led to the definition of several types of variables, the main ones being integer variables, set variables, and graph variables.

Definition 4.14 (Integer variable)**.** An integer variable is a variable which takes its values in a finite set of integers. The interval domain of an integer variable v_i is an integer interval $[l_b(v_i), u_b(v_i)]$, where $l_b(v_i)$ is the *lower bound* and $u_b(v_i)$ the *upper bound* such that $l_b(v_i) \leqslant v_i \leqslant u_b(v_i)$.

Definition 4.15 (Set variable)**.** A set variable is a variable which takes its values in a finite set of sets (see Section 4.1.2). The domain of a set variable v_s is a set interval $[\underline{v_s}, \overline{v_s}]$, where $\underline{v_s}$ is the *kernel* and $\overline{v_s}$ the *envelope* such that an instantiation of v_s is a subset of the envelope and the kernel is a subset of v_s. Formally: $\underline{v_s} \subseteq v_s \subseteq \overline{v_s}$. The elements of the kernel are called *mandatory* values and the elements of the envelope are called *potential* values. See Figure 4.7 for an illustration.

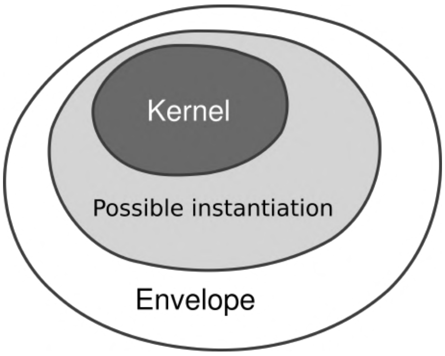


Figure 4.7: Illustration of a set variable's kernel (dark grey) and envelope (white), with a possible instantiation (union of light and dark grey).

Definition 4.16 (Graph variable)**.** A graph variable is a variable which takes its values in a finite set of graphs (see Section 4.1.3). Similarly to set variables, the domain of a graph variable v_g is a graph interval $[\underline{v_g}, \overline{v_g}]$ with $\underline{v_g}$ the kernel and $\overline{v_g}$ the envelope such that v_g is a subgraph of the envelope and the kernel is a subgraph of v_g. Formally, $\underline{v_g} \subseteq v_g \subseteq \overline{v_g}$. Similarly, the vertices (resp. edges) of the kernel are called the mandatory vertices (resp. edges) and the vertices (resp. edges) of the envelope are called the potential vertices (resp. edges). Graph variables can be either directed or undirected. See Figure 4.8 for an illustration.

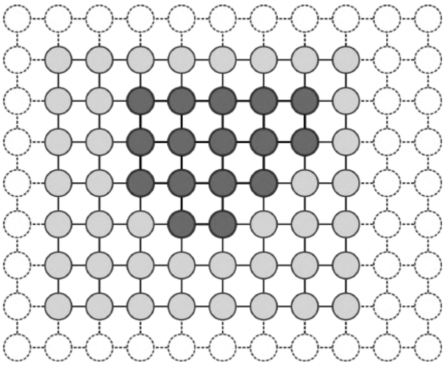


Figure 4.8: Illustration of an undirected graph variable's kernel (dark grey) and envelope (white and dashed lines), with a possible instantiation (union of light and dark grey).

This variety of variable types is one of CP's main strengths, as it allows to model problems with the most suitable mathematical objects as well as it permits to consider different types of variables in the same problem.

4.2.2 *What is a constraint in constraint programming?*

Conceptually, a constraint is a logical relation between a set of variables in a CSP or a COP. The number of variables implied in a constraint is the *arity* of the constraint. This arity can vary from one (*unary* constraint) to an arbitrary number of variables, in which case the constraint is *global*. Constraints with an arity of two are *binary* and constraints with an arity of three are *ternary*. Operationally, a constraint has two roles. First, it must embed a *constraint satisfaction checker* which can verify whether the constraint can be locally satisfied (i.e. independently of other constraints) by the variables' domains. This verification is done according to a necessary condition and optionally according to a sufficient condition. Secondly, it can embed a *filtering algorithm* which can identify and remove values in variables' domains that cannot lead to any solution satisfying the constraint, such values are said *locally inconsistent*. As an illustration, let us consider the constraint $c : v_x \neq v_y$ with v_x and v_y integer variables. The constraint satisfaction checker of c can identify the following necessary and sufficient conditions:

- *Necessary condition*: if $|\mathcal{D}(v_x) \cup \mathcal{D}(v_y)| \geqslant 2$ (and obviously both $\mathcal{D}(v_x)$ and $\mathcal{D}(v_y)$ are non-empty) then there exist $a_x \in \mathcal{D}(v_x)$ and $a_y \in \mathcal{D}(v_y)$ such that $a_x \neq a_y$ (i.e. c can be satisfied).
- *Sufficient condition*: if $\mathcal{D}(v_x) \cap \mathcal{D}(v_y) = \emptyset$ (and both $\mathcal{D}(v_x)$ and $\mathcal{D}(v_y)$ are non-empty) then c is satisfied for any instantiation of v_x and v_y.

The filtering algorithm of c consists in the following:

- *Filtering algorithm*: if $\mathcal{D}(v_x) = \{a\}$ (resp. $\mathcal{D}(v_y) = \{b\}$) then remove a (resp. b) from $\mathcal{D}(v_y)$ (resp. $\mathcal{D}(v_x)$).

The main strength of CP is the encapsulation of constraints logic into constraint satisfaction checkers and filtering algorithms, which both provide a generic interface completely independent of the constraints' nature. Especially, CP allows the expression of linear and non-linear constraints in a transparent way and a seamless integration of new constraints into a solver. This paradigm also allows embedding efficient and specialized algorithms into constraints, which has led to the conception of efficient global constraints that are key components in the efficiency of CP. The ALLDIFFERENT constraint is a historical and illustrative example of this strength (Régin, 1994; Van Hoeve, 2001). This constraint enforces n integer variables $\{v_1, ..., v_n\}$ to take distinct values. A simple way to model ALLDIFFERENT is to state $n(n-1)/2$ binary inequality constraints such that $v_1 \neq v_2 \wedge v_1 \neq v_3 \wedge ... \wedge v_2 \neq v_3 \wedge ... \wedge v_{n-1} \neq v_n$. However, this method relies on local reasoning that can only identify inconsistent values when a variable is instantiated, as illustrated in Table 4.1.

Table 4.1: Illustration of the local reasoning process when AllDifferent is modelled with binary inequality constraints. Three variables $\{v_1, v_2, v_3\}$ are considered such that initially $\mathcal{D}(v_1) = \{1,2,3\}$, $\mathcal{D}(v_2) = \{1,2\}$, and $\mathcal{D}(v_3) = \{1,2\}$. At each step, we describe the operation performed, each variable's domain after the operation, and the deductions that can be made on domains afterwards. When no deduction is possible, we assign a value to a variable to reactivate the deduction process. At each step, only the domains that were modified are displayed (except for the first and last steps).

Step	v_1	v_2	v_3	Deductions
#0 *Initialize*	$\{1,2,3\}$	$\{1,2\}$	$\{1,2\}$	$\emptyset$
#1 *Try* $v_1 = 1$	$\{1\}$	–	–	$v_2 \neq 1 \wedge v_3 \neq 1$
#2 *Remove 1 from* $\mathcal{D}(v_2)$ *and* $\mathcal{D}(v_3)$	–	$\{2\}$	$\{2\}$	$v_3 \neq 2 \wedge v_2 \neq 2$
#3 *Remove 2 from* $\mathcal{D}(v_2)$ *and* $\mathcal{D}(v_3)$	–	$\emptyset$	$\emptyset$	$v_1 \neq 1$
#4 *Backtrack to #0 and try* $v_1 = 2$	$\{2\}$	$\{1,2\}$	$\{1,2\}$	$v_2 \neq 2 \wedge v_3 \neq 2$
#5 *Remove 2 from* $\mathcal{D}(v_2)$ *and* $\mathcal{D}(v_3)$	–	$\{1\}$	$\{1\}$	$v_3 \neq 1 \wedge v_2 \neq 1$
#6 *Remove 1 from* $\mathcal{D}(v_2)$ *and* $\mathcal{D}(v_3)$	–	$\emptyset$	$\emptyset$	$v_1 \neq 2$
#7 *Backtrack to #0 and try* $v_1 = 3$	$\{3\}$	$\{1,2\}$	$\{1,2\}$	$\emptyset$
#8 *Try* $v_2 = 1$	–	$\{1\}$	–	$v_3 \neq 1$
#9 *Remove 1 from* $\mathcal{D}(v_3)$	$\{3\}$	$\{1\}$	$\{2\}$	$(3,1,2)$ is a solution

On the other hand, a global reasoning can be achieved by modelling AllDifferent with a bipartite graph $G = (U, V, E)$ such that $U = \{v_1, ..., v_n\}$ and $V = \bigcup_{i \in [1,n]} \mathcal{D}(v_i)$ and such that there is an edge (v_i, a) if $a \in \mathcal{D}(v_i)$. Then, inconsistent values can be detected by identifying edges that do not belong to any maximum cardinality matching in G. This global reasoning process is illustrated in Table 4.2 with the same example as Table 4.1 and the associated bipartite graph is illustrated in Figure 4.9. This global approach is unquestionably much more efficient than the local one as inconsistent values can be detected much earlier in the deduction process with efficient graph algorithms. The global filtering algorithm illustrated for the AllDifferent constraint even leads to the maximum possible level of filtering, as any remaining value in variables' domain belongs to an instantiation satisfying the constraint.

Table 4.2: Illustration of the global deduction process when AllDifferent is modelled with a bipartite graph. Three variables $\{v_1, v_2, v_3\}$ are considered such that initially $\mathcal{D}(v_1) = \{1,2,3\}$, $\mathcal{D}(v_2) = \{1,2\}$, and $\mathcal{D}(v_3) = \{1,2\}$. At each step, we describe the operation performed, each variable's domain after the operation, and the deductions that can be made on domains afterwards. When no deduction is possible, we assign a value to a variable to reactivate the deduction process. At each step, only the domains that were modified are displayed (except for the first and last steps).

Step	v_1	v_2	v_3	Deductions
#0 *Initialize*	$\{1,2,3\}$	$\{1,2\}$	$\{1,2\}$	$v_1 \neq 1 \wedge v_1 \neq 2$
#1 *Remove 1 and 2 from* $\mathcal{D}(v_1)$	$\{3\}$	–	–	$\emptyset$
#2 *Try* $v_2 = 1$	–	$\{1\}$	–	$v_3 \neq 1$
#3 *Remove 1 from* $\mathcal{D}(v_3)$	$\{3\}$	$\{1\}$	$\{2\}$	$(3,1,2)$ is a solution

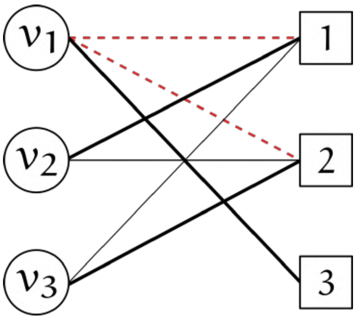


Figure 4.9: Illustration of the bipartite graph associated with the global reasoning of the AllDifferent constraint described in Table 4.2.

In CP, the efficiency of constraints is evaluated according to two concepts: the *consistency level* and the *worst-case algorithmic complexity* of its filtering algorithm. The latter refers to the classical worst-case algorithmic complexity definition (see Box 2.1) and is characterized with the *big O* notation.

Definition 4.17 (Big O notation)**.** In algorithmics, the *big O* notation is used to characterize the worst-case time or space complexity of an algorithm according to how time or space requirements increase with input size. If n is the input size of an algorithm, we denote by $f(n)$ its time or space requirement. We write $f(n) = O(g(n))(n \to +\infty)$ if there exist two constants N and C such that $\forall n \geqslant N$, $|f(n)| \leqslant C|g(n)|$, which intuitively means that f does not grow faster than g. In practice, we only write $O(g(n))$ to characterise the worst-case time or space complexity of an algorithm. For example, the worst-case time complexity of a linear algorithm is denoted as $O(n)$ and an algorithm with a logarithmic worst-case time complexity characterized by $O(\log(n))$.

The consistency level, on the other hand, is a characterization of the deductive power of a filtering algorithm. Several consistency levels have been defined (see Chapter 3 of the *Handbook of Constraint Programming*, Rossi et al., 2006), to illustrate the concept of consistency in CP we introduce two common consistency levels on integer variables: the *generalized arc-consistency* (GAC) and the *bound consistency* (BC). Most often, the choice of a filtering algorithm is a trade-off between the level of consistency that it can reach and its algorithmic complexity.

Definition 4.18 (Generalized arc-consistency)**.** Let c be a constraint defined on the integer variables $\{v_1, ..., v_n\}$. If for any pair (v_i, a) such that $a \in \mathcal{D}(v_i)$ there exist at least one tuple $(a_1, ..., a_n)$ such that $a_i = a$ satisfying c, then c is said to have reached *generalized arc-consistency* (GAC). A filtering algorithm thus reaches GAC when it is able to remove every value that cannot lead to a solution from variables' domains.

Definition 4.19 (Bound consistency)**.** Let c be a constraint defined on the integer variables $\{v_1, ..., v_n\}$. If for any pair $(v_i, l_b(v_i))$ and for any pair $(v_i, u_b(v_i))$ there exist at least one tuple $(a_1, ..., a_n)$ such that $a_i = l_b(v_i)$ satisfying c and at least one tuple $(b_1, ..., b_n)$ such that $b_i = u_b(v_i)$ satisfying c, then c is said to have reached *bound consistency* (BC). A filtering algorithm thus reaches

BC when it is able to guarantee that the lower (resp. upper) bound of every variable's domain belongs to at least one solution.

4.2.3 *How does a constraint programming solver work?*

We have presented the concepts of variables and constraints that are the building blocks of CP. It now remains to describe the mechanisms on which a CP solver relies to solve problems. CSPs and COPs can be seen as networks of variables linked by several constraints, where a variable is usually involved in several constraints. Thus, the modification of a variable's domain by a filtering algorithm can directly impact the deductions that can be made by the other constraints applied to this variable. Concretely, every constraint's filtering algorithm can reduce the search space through variables' domains, and any such reduction brings new pieces of information on the problem's structure that can be exploited by other constraints to keep reducing the search space. This mechanism is at the core of CP solvers and is called *constraint propagation*. A propagation algorithm repeatedly calls the filtering algorithms of the problem's constraints until it reaches a fixpoint. As the order in which it calls constraints' filtering algorithms can have an impact on the convergence speed, propagation algorithms have been subject to extensive research in CP. We provided an example of a basic propagation algorithm in Table 4.3 with a problem composed of three variables and three constraints.

Table 4.3: Illustration of a propagation algorithm for a CSP with three variables $\{v_1, v_2, v_3\}$ such that initially $\mathcal{D}(v_1) = \{1,2,3\}$, $\mathcal{D}(v_2) = \{2,3,4\}$, and $\mathcal{D}(v_3) = \{2,3\}$, and three constraints $c_1 : v_1 \geqslant v_2$, $c_2 : v_2 \neq v_3$, and $c_3 : \text{EVEN}(v_3)$. The constraints in the column *Constraints to propagate* are those whose consistency must be ensured through their filtering algorithm. The constraints in the column *Consistent* are those whose consistency is currently ensured. When a filtering algorithm modifies a variable's domain, every constraint in which this variable is involved is moved to the *Constraints to propagate* column. A fixed point is reached when the column *Constraints to propagate* is empty. At each step, only the domains that have been modified are displayed (except for the first and last steps).

Step	v_1	v_2	v_3	Constraints to propagate	Consistent
#0	$\{1,2,3\}$	$\{2,3,4\}$	$\{2,3\}$	$\mathbf{c_1}, c_2, c_3$	$\emptyset$
#1	–	$\{2,3\}$	–	$\mathbf{c_2}, c_3$	c_1
#2	–	–	–	$\mathbf{c_3}$	c_1, c_2
#3	–	–	$\{2\}$	$\mathbf{c_2}$	c_1, c_3
#4	–	$\{3\}$	–	$\mathbf{c_1}$	c_1, c_2
#5	$\{3\}$	$\{3\}$	$\{2\}$	$\emptyset$	c_1, c_2, c_3

In the example of Table 4.3, the propagation algorithm was sufficient to reach an instantiation. However, in most cases, this algorithm alone is not sufficient. In such situations, *backtracking search algorithms* come into play. As illustrated in Tables 4.1 and 4.2 with a little advance, the idea is, whenever the propagation algorithm reaches a fixpoint before complete instantiation, to select a variable and reduce its domain (such an action is called a *decision*) to reactivate the propagation algorithm. Most of the time reduction corresponds to a value assignment. This process is the principle of *search*. When a decision leads to a *contradiction* (i.e. one variable's domain becomes empty), it is necessary to go back in time before the last decision and try another decision. This process is the principle of *backtracking*. Search and backtracking can be represented by a binary search tree where each vertex corresponds to a decision and leaves correspond either to a contradiction or a solution. We illustrated the binary search tree corresponding to the process depicted in Table 4.1 in Figure 4.10.

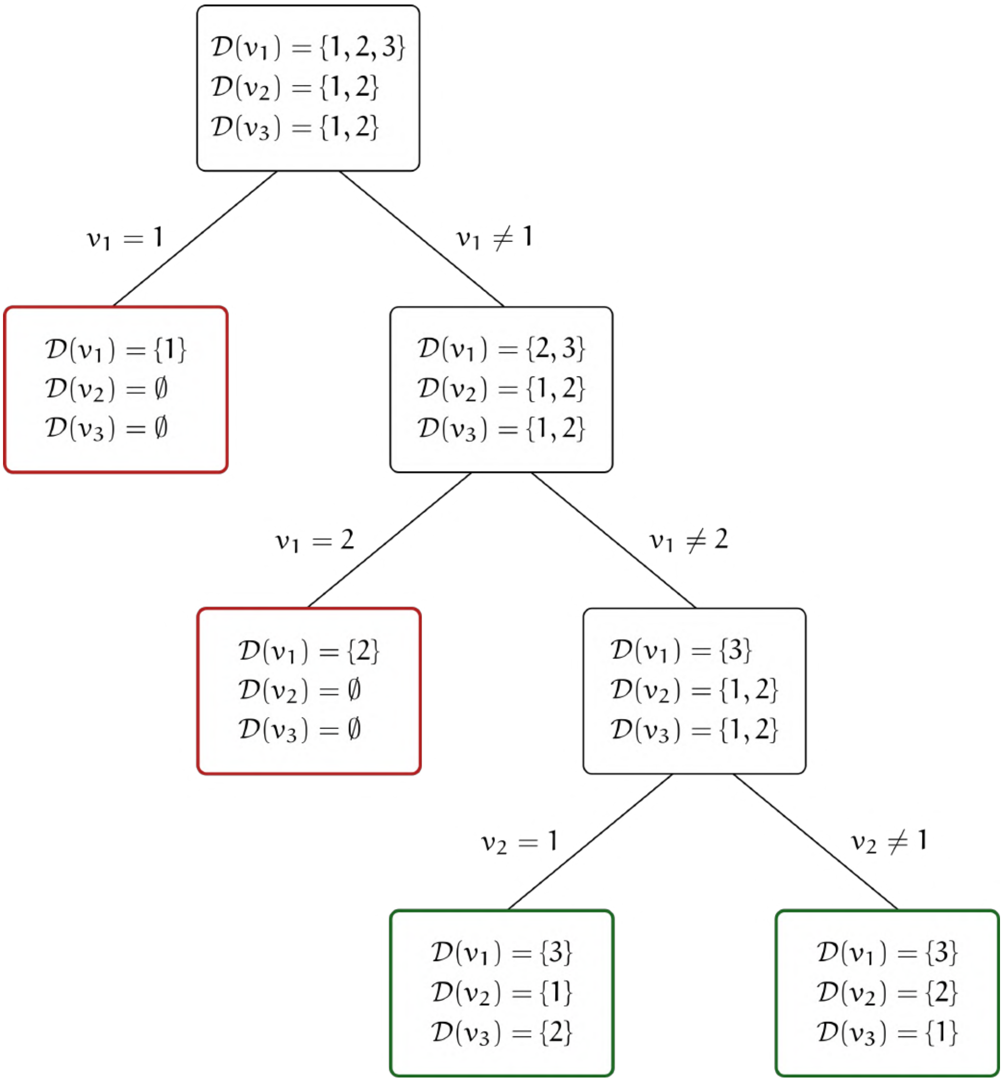


Figure 4.10: Illustration of the binary search tree associated with the deduction process depicted in Table 4.1. The root of the tree corresponds to the initial state, and each edge corresponds to a decision whose impacts on domains are indicated on the next level vertex. Red leaves correspond to contradictions and green leaves to solutions. This search tree is complete (i.e. all possible solutions are enumerated).

4.2.4 *Applications, state-of-the-art solvers and additional resources*

After nearly half a century of existence, constraint programming is well-established within the fields of operational research (OR) and artificial intelligence (AI). The discipline gathers a worldwide community of academic and industrial researchers and shapes around a dedicated international journal (*Constraints*) and international conferences in OR and AI, such as *The International Conference on Principles and Practice of Constraint Programming* (CP), *The International Joint Conferences on Artificial Intelligence* (IJCAI), *The Association for the Advancement of Artificial Intelligence conference* (AAAI), *The International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research* (CPAIOR), etc.

In terms of applications, CP has historically been successfully employed to solve scheduling, planning, routing, or packing problems with industrial (e.g. production optimization), military (e.g. aircraft scheduling), or medical (e.g. nurse scheduling) applications (Wallace, 1996). However, the diversity of possible applications of CP is constantly growing, with new applications in music (e.g. Hooker, 2016), cryptography (e.g. Gerault et al., 2016), testing and verification (e.g. Bardin et al., 2009), bioinformatics (e.g. Barahona et al., 2011; Kerdprasop and Kerdprasop, 2015), or even systematic conservation planning (e.g. Bessière et al., 2015, this PhD thesis).

The historical link of CP with both the academic and industrial world has undoubtedly been a determining factor in the success of CP, notably through the emergence of many state-of-the-art CP solvers developed with industry standards in an open-sourced academic world. Indeed, the majority of the most commonly used CP solvers are open-source, and most of these originate from academical initiatives. In the following, we provide a (non-exhaustive) list of state-of-the-art CP solvers.

- *MiniZinc/FlatZinc* (`https://www.minizinc.org/`; Nethercote et al., 2007; Stuckey et al., 2014). Minizinc is a free and open-source constraint modelling language, which can be used to model and solve CSPs and COPs with any compatible solver. A MiniZinc model is compiled into *FlatZinc*, a language which is accepted as input by a wide range of solver. MiniZinc also made possible the organisation of challenges to compare the performances of CP solvers.

- *SICStus Prolog* (`https://sicstus.sics.se/`; Carlsson, 2001). SICStus Prolog is a development system based on the Prolog logic programming language. It integrates a finite domain constraint programming module with global constraints and a MiniZinc/FlatZinc interface. SICStus is non-free proprietary software.

- *Gecode* (`https://www.gecode.org/`; Schulte et al., 2006, 2010). Gecode is an open-source C++ constraint programming toolkit. Gecode was initially developed at KTH University (Sweden) and is now mainly maintained at Monash University (Australia). Gecode integrates constraints (including global constraints) over booleans, integers, sets, graphs, and float, and provides a MiniZinc/FlatZinc interface.

- *Google OR-Tools* (`https://developers.google.com/optimization/`; Perron and Furnon, 2019). OR-Tools is an open-source software suite developed by Google to solve optimization problems. It includes dedicated tools to solve vehicles routing, flows, mixed-integer linear programming, and constraint-programming problems. OR-Tools provides a MiniZinc/FlatZinc interface, and programming interfaces for Java, C++, C#, and Python.

- *JaCoP* (`http://www.jacop.eu/`; Kuchcinski and Szymanek, 2013). JaCoP is a Java-based open-source constraint programming solver developed and maintained at the University of Lund (Sweden), and the Crossing-Tech company in Lausanne (Switzerland). JaCoP supports integer and sets of integers finite domains, floating-point domains, and offers a wide range of state-of-the-art constraints. It is used for research, teaching, industrial and commercial purposes, and provides a MiniZinc/FlatZinc interface.

- *Mistral-2.0* (`https://github.com/ehebrard/Mistral-2.0`; Hebrard and Siala, 2016). Mistral-2.0 is a C++ open-source constraint programming library. It provides a modelling API but can also be interfaced with MiniZinc/FlatZinc, and XCSP3 which is a XML-based format for representing CSPs and COPs. Mistral-2.0 is also fully interfaced with Numberjack a Python library for combinatorial optimization (`https://github.com/eomahony/Numberjack`).

- *Chuffed* (`https://github.com/chuffed/chuffed`; Chu et al., 2016). Chuffed is a C++ open-source constraint programming solver developed at the University of Melbourne (Australia) and CSIRO (Australia). Chuffed is a state-of-the-art lazy clause generation solver, which is a hybrid approach between finite domain propagation and Boolean satisfiability (SAT, see Biere et al. (2009) for more details on such problems). Chuffed provides a MiniZinc/FlatZinc interface.

- *PicatSAT* (`http://picat-lang.org/`; Zhou and Kjellerstrand, 2017). PicatSAT is a solver developed with the logic-based Picat (http://picat-lang.org/) language. PicatSAT solves CSPs and COPs by first traducing them into SAT problems and then by relying on a SAT solver.

- *Yuck* (`https://github.com/informarte/yuck`; Hentenryck and Michel, 2009; Björdal et al., 2015). Yuck is a constraint-solver based on local-search techniques. It implements Boolean, integer, and integer set variables. It includes global constraints and can be used as a FlatZinc interpreter.

- *OscaR* (`https://oscarlib.bitbucket.io/`; De Landtsheer and Ponsard, 2013). OscaR is an open-source solver developed at the Catholic University of Louvain (Belgium). It is based on the Scala language and provides techniques for constraint programming but also for constrained-based local search and derivative free optimization. OscaR provides an interface for MiniZinc/FlatZinc.

- *Sunny-CP* (`https://github.com/CP-Unibo/sunny-cp`; Amadini et al., 2015). Sunny-CP is a parallel constraint programming portfolio solver that takes as input MiniZinc models. Portfolio solvers as Sunny-CP are built on top of existing solvers that are launched in parallel on the same problem. During the solving process, when a solver finds a solution or detects domain inconsistencies, the portfolio solver shares the information with the other solvers so that they can take advantage of it.

- *Choco* (`https://choco-solver.org/`; Prud'homme et al., 2017; Fages et al., 2018). Choco is an open-source Java constraint programming solver, developed and maintained at the IMT Atlantique (Nantes, France) and by the Cosling company (Nantes, France). Choco supports Boolean, integer, integer set, graph, and real variables. It integrates a wide range of constraints including state-of-the-art global constraints, state-of-the-art search strategies, an explanation-based engine, and can be used for multi-objective Pareto optimization. Choco has been used for more than twenty years for teaching, academic, industrial, and commercial projects.

Although there are performance differences between existing solvers, the main bottleneck with regards to efficiency in CP is in most cases modelling. Indeed, there are several manners to model a problem as a CSP or as COP, but some are more efficient than others. For example, relying on global constraints instead of local constraints usually allow solvers to perform better reasoning and to solve problems faster (see the comparison between the global constraint AllDifferent and local inequality constraints in Section 4.2.2). Good modelling practices in CP come with experience in problem-solving, but also with theoretical and technical knowledge on algorithms and solvers. To go further, interested readers can refer to the comprehensive *Handbook of Constraint Programming* (Rossi et al., 2006). In addition, many courses are freely available on the Internet (e.g. `http://kti.ms.mff.cuni.cz/~bartak/constraints/`, `https://www.coursera.org/learn/discrete-optimization`). Most CP solvers' documentations also provide tutorials and example that are useful to improve modelling and technical CP skills. For example, Choco provides an excellent series of tutorials to get started with CP using Choco (`https://choco-solver.org/tutos/`). In this PhD thesis, we relied on the Choco solver for three main reasons: (i) it is, to our knowledge, the most comprehensive solver with regards to set and graph variables and constraints. (ii) Xavier Lorca, the co-supervisor of this PhD thesis, is one of the developers of Choco. Technical support was thus easily accessible using this solver. (iii) Choco is a state-of-the-art solver, awarded with numerous prizes at MiniZinc challenges (`https://www.minizinc.org/challenge.html`; Stuckey et al., 2010) and XCSP3 competitions (`http://www.xcsp.org/competition`; Lecoutre and Roussel, 2018). It is actively maintained by a dedicated team and used in several academic and industrial projects.

5

UNIFYING RESERVE DESIGN STRATEGIES WITH GRAPH THEORY AND CONSTRAINT PROGRAMMING

This chapter is a literal transcription of an article **published** in the proceeding of the 24th International Conference on Principles and Practices of Constraint Programming (CP 2018), which was held in Lille (France) between the 27th and the 30th of August 2018. This work was also subject to a talk at this conference. Dimitri Justeau-Allaire, Philippe Birnbaum, and Xavier Lorca (2018). "Unifying Reserve Design Strategies with Graph Theory and Constraint Programming". en. In: *Principles and Practice of Constraint Programming*. Ed. by John Hooker. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 507–523. ISBN: 978-3-319-98334-9. DOI: 10.1007/978-3-319-98334-9_33.

Abstract

The delineation of areas of high ecological or biodiversity value is a priority of any conservation program. However, the selection of optimal areas to be preserved necessarily results from a compromise between the complexity of ecological processes and managers' constraints. Current reserve design models usually focus on few criteria, which often leads to an oversimplification of the underlying conservation issues. This paper shows that Constraint Programming (CP) can be the basis of a more unified, flexible and extensible framework. First, the reserve design problem is formalized. Secondly, the problem is modeled from two different angles by using two graph-based models. Then CP is used to aggregate those models through a unique Constraint Satisfaction Problem. Our model is finally evaluated on a real use case addressing the problem of rainforest fragmentation in New Caledonia, a biodiversity hotspot. Results are promising and highlight challenging perspectives to overtake in future work.

5.1 INTRODUCTION

Human activities are exerting pressure on natural habitats, which generally results in a loss of surface and an increase of fragmentation. As a consequence, many species depending on those habitats are threatened, sometimes with extinction. In this context, it is essential to devote an important part of conservation efforts in the protection of natural habitats through the establishment of nature reserves (Beier et al., 2011; Baguette et al., 2013; Haddad et al., 2015; Prendergast et al., 1993). Designing a reserve system is a difficult process involving a trade-off between the conservation targets and the socioeconomic constraints. This problem is known as the *reserve design problem*. The associated questions are at the crossroad between conservation biology, geography, mathematics, computer science, decision theory and environmental philosophy (Sarkar, 2013). In this paper, we focus on the mathematical modeling and the computational solving of the reserve design problem. From this point of view, it is a decision and/or optimization problem. In almost all cases, the combinatorial complexity justifies the need of a systematic approach based on mathematical modeling and computational tools.

In the literature, two major aspects of the reserve design problem usually stand out: the *feature covering* and the *spatial configuration*. The first is often referred as the reserve (or site) selection problem (Pressey et al., 1993; ReVelle et al., 2002; Billionnet, 2011; Watts et al., 2009). In extension, we refer to the reserve design problem when spatial attributes are considered (Diamond, 1975; Williams et al., 2005; Billionnet, 2016; Dilkina et al., 2017; Jafari et al., 2017). Current models usually focus on a few aspects of the problem because: (1) they provide an ad-hoc solution to a specific instance of the problem, or (2) they are limited by the modeling paradigm. However, there is a need for a more unified and flexible framework (Rodrigues et al., 2000) which, in our opinion and based on our experience in New Caledonia, could help to reduce the gap between computer scientists, conservation scientists, and practitioners.

In this paper, we show how the combination of graph-based models with CP can be the basis of such a framework. After a detailed description of the reserve design problem (Section 5.2), we presents two graph-based models (Section **??**). One model is dedicated to the constraint representation of the features covering issues (Section 5.3.1) and the other one is dedicated to the constraint representation of the spatial issues (Section 5.3.2). We then unify the models throughout a single CP model based on the Choco constraint solver (Prud'homme et al., 2017) (Section 5.4). Finally, a realistic operational use case on the problem of rainforest fragmentation in New Caledonia is depicted and first results are discussed (Section 5.5).

5.2 DESCRIPTION OF THE PROBLEM

The reserve design is a decision and/or optimization problem in the discretized geographical space. Given a set of geographical features (e.g., Fig. 5.2), we are looking for a reserve system satisfying several criteria, in accordance to conservation targets. In this section, we describe and formalize the problem precisely. We start by defining the characteristics of the problem and then define a set of criteria that can be required for a reserve system.

5.2.1 *Characteristics of the Problem – Input Data*

The Discretized Geographical Space. The geographical space is tessellated into n granular parcels, which are the decision variables of the problem. Several tessellation methods are possible (Sahr et al., 2003; Birch et al., 2007). The most commonly used is the regular square grid (illustrated in Fig. 5.1). We choose to this method in this paper.

We denote the number of rows by r, the number of columns by c and the set of parcels by $\mathcal{P}$. We identify a single parcel with the letter i, and index the parcels with integers from 0 to $n-1$: $\mathcal{P} = \{\, i \mid i \in [\![0, n[\![\, \}$. While this indexing is not the most convenient for a grid, it has the advantage to be independent of the tessellation method and thus offers extensibility for future work. Finally, we use the 8-connected (cf. Fig. 5.1) neighborhood to define the adjacency between the parcels, in opposition to the 4-connected neighborhood.

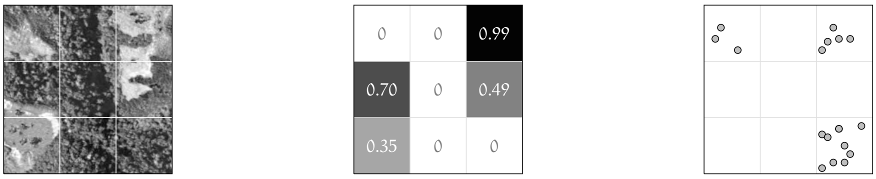
The Environmental Features. The geographical space is characterized by a set of m environmental features. A feature can be anything that can be spatially represented (e.g. the presence of a species, a certain type of habitat, human constructions). We denote by $\mathcal{F}$ the set of environmental features and use the letter j to identify the features: $\mathcal{F} = \{\, j \mid j \in [\![0, m[\![\, \}$.

The Values of the Features. To each feature j is associated a set $\mathcal{V}_j$, representing the available data about j among the parcels: $\mathcal{V}_j = \{\, v_{ji} \in \mathbb{R}^+ \mid i \in \mathcal{P} \}$. Each $v_{ji} \in \mathcal{V}_j$ corresponds to a value describing the feature j in the parcel i. Three types of data are possible: the presence-absence data, the abundance data and

(a) Square grid tesselation

(b) 8-connected neighborhood (gray) of the parcel 4 (black)

Figure 5.1: Square grid tessellation and 8-connected neighborhood illustrations.



(a) Feature 0: Dominance of dense rainforest habitat (presence-absence)

(b) Feature 1: SDM of an endemic ant species (probability of presence)

(c) Feature 2: Exhaustive inventory of a rare fern species (abundance)

Figure 5.2: Three feature examples.

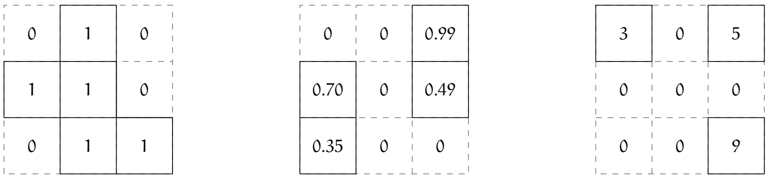
the probability of presence data. An example for each data type is given in Fig. 5.2, and below is a short description for each of them:

- *Presence-absence:* if j is present in the parcel i, $v_{ji} = 1$, else $v_{ji} = 0$. For each $(j, i) \in \mathcal{F} \times \mathcal{P}$ we then have $v_{ji} \in \{0, 1\}$. The presence-absence data is often used to describe the occurrence distribution of a species or a particular characteristic of the landscape (e.g. forest, savanna, fields, roads, cities).

- *Abundance:* in this case, v_{ji} represents a quantitative value about the feature j in the parcel i (e.g. density of trees per parcel, average annual rainfall). For each $(j, i) \in \mathcal{F} \times \mathcal{P}$ we then have $v_{ji} \in [0, +\infty[$.

- *Probability of presence:* it can be possible to evaluate the probability of presence of a feature j for every parcel in $\mathcal{P}$. The most common situation is the use of Species Distribution Models (SDMs), that are able to combine observations of a species with environmental data to predict its spatial distribution (Guisan and Zimmermann, 2000; Elith and Leathwick, 2009). With probability of presence data, for each $(j, i) \in \mathcal{F} \times \mathcal{P}$ we have $v_{ji} \in [0, 1]$.

The Domains/Anti-Domains of the Features. As illustrated in Fig. 5.3, to each feature, is associated a set $\mathcal{D}_j$ and its complement $\overline{\mathcal{D}_j}$. $\mathcal{D}_j$ represents the domain of j, that is, the parcels where j is present or where the probability of presence of j is not null: $\mathcal{D}_j = \{ i \in \mathcal{P} \mid v_{ji} > 0 \}$. Conversely, $\overline{\mathcal{D}_j}$ represents the anti-domain of j, that is, the parcels where j is not present or where the probability of presence of j is null: $\overline{\mathcal{D}_j} = \{ i \in \mathcal{P} \mid v_{ji} = 0 \}$.



(a) $\mathcal{V}_0$, $\mathcal{D}_0$ and $\overline{\mathcal{D}}_0$ (b) $\mathcal{V}_1$, $\mathcal{D}_1$ and $\overline{\mathcal{D}}_1$ (c) $\mathcal{V}_2$, $\mathcal{D}_2$ and $\overline{\mathcal{D}}_2$

Figure 5.3: The values, domains and anti-domains associated with the features in Fig. 5.2. The domains are represented with solid lines and the anti-domains with dashed lines.

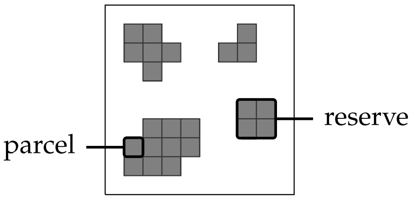


Figure 5.4: Illustration of a reserve system, composed of four reserves, themselves made of several adjacent parcels.

5.2.2 *The Reserve System – Solution of the Problem*

In the first place, we define the terms "*parcel*" (sometimes called "*site*" in the literature), "*reserve*" and "*reserve system*". As defined in the previous subsection, a parcel is a granular selection unit of the discretized geographical space. On top of that, a reserve is a set of spatially continuous selected parcels (note that a single selected isolated parcel is a reserve). Finally a reserve system is a set of spatially disjoint reserves (note that a reserve system can be composed of a single reserve). We illustrated the previous definitions in Fig. 5.4.

Given that, a solution to our problem is a reserve system whose attributes are satisfying a set of criteria, themselves depending on the conservation question. We denote such a reserve system by $\mathcal{S}$, its number of reserves by n_r and its k^{th} reserve by $\mathcal{X}_k$: $\mathcal{S} = \{\mathcal{X}_k \subseteq \mathcal{P} \mid k \in [\![0, n_r[\![\}$ (where $\mathcal{P}$ is the set of parcels).

5.2.3 *Required Criteria for a Reserve System*

According to the underlying conservation questions, several criteria can be required for a reserve system. We distinguish between the feature covering criteria and the spatial criteria.

Feature Covering Criteria. By providing one of the first formalization of the reserve selection problems, ReVelle et al. (2002) introduced three fundamental feature covering criteria.

- *Covered Features.* Among the features that are covered (with certainty) by the reserve system $\mathcal{S}$, we want a set of mandatory features $\mathcal{F}'$ to be represented (e.g. rare or endangered species).
- *α-Covered Features.* Assuming that the v_{ji}'s are pairwise independent, we want a set of features $\mathcal{F}'$ to be covered by $\mathcal{S}$ with a probability of at least α. This criterion is helpful when probability of presence data is available.
- *k-Redundant Features.* A feature j is k-redundant in the reserve system $\mathcal{S}$ if and only if it is covered (with certainty) by at least k distinct parcels. We want to enforce this property for a set of features $\mathcal{F}'$ (e.g. for increasing the chances of persistence of vulnerable species).

Spatial Criteria. A list of six geometric principles had been defined by Diamond (1975) and Williams et al. (2005) summarized them into six spatial attributes to take into account when designing a reserve system: the *number of reserves*, the *reserve areas* (by extension we define the *reserve system area*), the *reserve proximity*, the *reserve connectivity*, the *reserve shape* and *core areas and buffer zones*. Here we consider three of those spatial attributes (expressed as criteria) and keep the remaining ones for future work.

- *Number of reserves.* Determining if the best suited is a "single large or several small reserves" (SLOSS), or a "few large or many small reserves" (FLOMS) is a well known debate in ecology (Diamond, 1975; Etienne and Heesterbeek, 2000). The conclusion is that the answer strongly depends on the context, and that flexibility is needed. We therefore want to set a minimum value N_{min} and/or a maximum value N_{max} for the number of reserves.
- *Reserve Areas.* Following the previous criterion, it is also essential to provide control on the reserve areas by setting a minimum area A_{min} and a maximum area A_{max}.
- *Reserve System Area.* It should also be possible to express this criterion on the whole reserve system area, by setting a minimum total area $A_{T_{min}}$ and a maximum total area $A_{T_{max}}$.

5.3 THE GRAPH-BASED MODELS

In this section, we present two graph-based models. The first one is a resource allocation model that will enable us to express the feature covering criteria in the form of constraints. In the same way, the second one is a spatial model that will enable us to control the spatial criteria. In both models, each parcel is represented by a vertex. This common characteristic is essential since it is the one that makes the aggregation of the models possible, through a set of appropriate channeling constraints.

5.3.1 *The Resource Allocation Graph*

We consider parcels as resources that can be allocated to the conservation of features, then considered as tasks, and thus define the directed graph $G_r = (V_r, A_r)$, also called the *resource allocation graph*. The vertices of G_r are partitioned into three disjoint sets F_r, P_r and $\{s,t\}$. F_r represents the *feature* (or *task*) vertices, P_r represents the *parcel* (or *resource*) vertices, s is the source vertex and t the sink vertex.

$$\begin{aligned} V_r &= F_r \cup P_r \cup \{s,t\}; \\ F_r &= \{\, f_j \mid j \in \mathcal{F} \}; \\ P_r &= \{\, p_i \mid i \in \mathcal{P} \}. \end{aligned} \tag{5.1}$$

Furthermore, using $A_r(X,Y)$ as the notation for the set of all $X-Y$ arcs, we define the arcs of G_r in the following way:

$$A_r = A_r(s, F_r) \,\cup\, A_r(F_r, P_r) \,\cup\, A_r(P_r, t). \tag{5.2}$$

$A_r(s, F_r)$ and $A_r(P_r, t)$ are defined such that there is an arc from s to each feature vertex and an arc from each parcel vertex to t:

$$\begin{aligned} A_r(s, F_r) &= \{\, (s, f_j) \mid f_j \in F_r \,\}; \\ A_r(P_r, t) &= \{\, (p_i, t) \mid p_i \in P_r \,\}. \end{aligned} \tag{5.3}$$

Moreover, $A_r(F_r, P_r)$ represent the possible allocations between F_r and P_r. More precisely, there is an arc from a feature vertex f_j to a parcel vertex p_i if and only if the feature j is represented in the parcel i, that is $i \in \mathcal{D}_j$. We then have:

$$A_r(F_r, P_r) = \bigcup_{j \in \mathcal{F}} \{\, (f_j, p_i) \mid i \in \mathcal{D}_j \,\}. \tag{5.4}$$

On the arcs of G_r, we define a lower bound (or demand) function $l : A_r \mapsto \mathbb{R}^+_\infty$ and an upper bound (or capacity) function $u : A_r \mapsto \mathbb{R}^+_\infty$ such that if f is a flow in G_r:

$$\forall a \in A_r,\ l(a) \leqslant f(a) \leqslant u(a). \tag{5.5}$$

Finally, we define $H_r : \mathcal{P}(P_r) \mapsto \mathcal{P}(V_r) \times \mathcal{P}(A_r)$ that associates to a set $X \subseteq P_r$ the subgraph of G_r induced by $\{s,t\} \cup F_r \cup X$, that is, the resource allocation graph obtained when only considering a subset of parcels. We denote by $V_r[H_r(X)]$ the vertices of $H_r(X)$ and by $A_r[H_r(X)]$ the arcs of $H_r(X)$. An example is provided in Fig. 5.5.

$$H_r(X) = G_r[\{s,t\} \cup F_r \cup X]. \tag{5.6}$$

Expressing Feature Covering Criteria as Constraints. From this point, to each feature covering criterion (as defined in the previous section) we associate a constraint that can be applied on the resource allocation model. More precisely, if $\mathcal{S}$ is a reserve system and X_s its associated set of parcel vertices, a criterion is satisfied by $\mathcal{S}$ if and only if its associated constraint if satisfied by $H_r(X_s)$. We

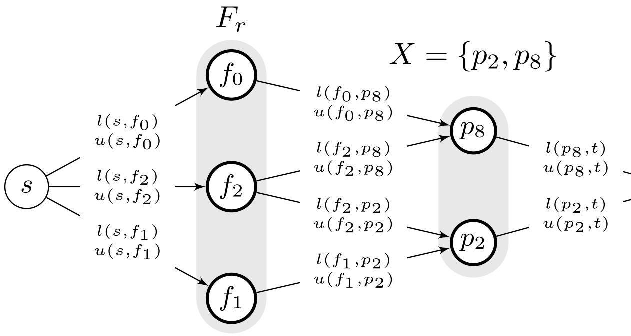


Figure 5.5: $H_r(\{p_2, p_8\})$ associated with the example in Fig. 5.2. Lower and upper bounds are represented on the arcs.

express these constraints as flow constraints by defining the value of l and u on certain arcs. When the value of l is not explicitly defined, it is unconstrained and then set to 0. Similarly, u is set to $+\infty$ when its value is not explicitly defined.

Covered Features. In our resource allocation model, we can easily express this criterion as a flow constraint on $H_r(X_s)$.

Constraint 1: Covered Features.

Input parameter(s): A set of features $\mathcal{F}' \subseteq \mathcal{F}$.

The set of features $\mathcal{F}'$ is covered by $\mathcal{S}$ if and only if $H_r(X_s)$ admits a feasible flow f verifying (5.5) when:

$$\begin{cases} l(s, f_j) = 1, & \forall f_j \in F'_r; \\ l(f_j, p_i) = 1, & \forall (f_j, p_i) \in A_r(F_r, P_r); \\ u(f_j, p_i) = 1, & \forall (f_j, p_i) \in A_r(F_r, P_r) \text{ such that } \nu_{ji} \geqslant 1; \\ u(f_j, p_i) = 0, & \forall (f_j, p_i) \in A_r(F_r, P_r) \text{ such that } \nu_{ji} < 1. \end{cases} \tag{5.7}$$

α-Covered Features. To express this criterion, we assume that the probabilities of presence ν_{ji} are pairwise independent. We then rely on the probability of absence $q_{ji} = (1 - \nu_{ji})$ and express the constraint as:

$$\forall j \in \mathcal{F}', \quad \prod_{i \in \mathcal{S}} q_{ji} \leqslant 1 - \alpha. \tag{5.8}$$

We then express the α-presence constraint in the following way:

Constraint 2: α-Covered Features.

Input parameter(s): A set of features $\mathcal{F}'$ and a real $\alpha \in [0, 1]$.

The set of features $\mathcal{F}'$ is covered by $\mathcal{S}$ with a probability of at least α if and only if $H_r(X_s)$ admits a feasible flow f verifying (5.5) when:

$$\begin{cases} l(s, f_j) = -\log(1 - \alpha), & \forall f_j \in F'_r; \\ u(f_j, p_i) = -\log(q_{ji}), & \forall (f_j, p_i) \in A_r(F_r, P_r) \text{ such that } \nu_{ji} < 1. \end{cases} \tag{5.9}$$

k-Redundant Features. Since the k-redundancy is actually a generalization of the covering features criterion, we can also express it as a flow constraint on $H_r(X_s)$.

Constraint 3: k-Redundant Features.

Input parameter(s): A set of features $\mathcal{F}'$ and a positive integer k.

The k-redundancy of the set of features $\mathcal{F}'$ in the reserve $\mathcal{S}$ is satisfied if and only if $H_r(X_s)$ admits a feasible flow f verifying (5.5) when:

$$\begin{cases} l(s, f_j) = k, & \forall f_j \in F'_r; \\ l(f_j, p_i) = 1, & \forall (f_j, p_i) \in A_r(F_r, P_r); \\ u(f_j, p_i) = 1, & \forall (f_j, p_i) \in A_r(F_r, P_r) \text{ such that } v_{ji} \geqslant 1; \\ u(f_j, p_i) = 0, & \forall (f_j, p_i) \in A_r(F_r, P_r) \text{ such that } v_{ji} < 1. \end{cases} \tag{5.10}$$

5.3.2 *The Spatial Graph*

We now define the undirected graph $G_s = (V_s, E_s)$, the *spatial graph*, which is a representation of the discretized geographical space $\mathcal{P}$ (a $r \times c$ regular square grid in our case). Once again, to each parcel i of $\mathcal{P}$, we associate a vertex p_i, we then have:

$$V_s = \{\, p_i \mid i \in \mathcal{P} \,\}. \tag{5.11}$$

Moreover, the edges of G_s are defined such that if p_u and p_v are two vertices, there is an edge between p_u to p_v if and only if the parcels u and v are spatially adjacent. The edges of G_s can be partitioned into four disjoint sets: the horizontal edges (E_H), the vertical edges (E_V), the north-west to south-east diagonal edges ($E_{N_W S_E}$) and the north-east to south-west diagonal edges ($E_{N_E S_W}$).

$$\begin{aligned} E_s &= E_H \cup E_V \cup E_{N_W S_E} \cup E_{N_E S_W}; \\ E_H &= \{\, (p_i, p_{i+1}) \mid i \in \mathcal{P} \land \neg\, (i+1) \equiv 0\, (c) \,\}; \\ E_V &= \{\, (p_i, p_{i+c}) \mid i \in \mathcal{P} \land i < c(r-1) \,\}; \\ E_{N_W S_E} &= \{\, (p_i, p_{i+c+1}) \mid i \in \mathcal{P} \land i < c(r-1) \land \neg\, (i+1) \equiv 0\, (c) \,\}; \\ E_{N_E S_W} &= \{\, (p_i, p_{i+c-1}) \mid i \in \mathcal{P} \land i < c(r-1) \land \neg\, i \equiv 0\, (c) \,\}. \end{aligned} \tag{5.12}$$

See Fig. 5.6 for an illustration of the above equation. Also note that it takes into account the extremal positions of the grid. In fact, the parcels located in the first column are the one whose index is a multiple of c, that is $i \equiv 0\, (c)$. Moreover, the parcels located in the last column are the ones preceding those that are located in the first column, that is $(i+1) \equiv 0\, (c)$. Finally, the parcels located in the last line are the ones satisfying $i < c(r-1)$.

Expressing Spatial Criteria as Constraints. Similarly to what had been defined for the resource allocation graph, to a solution $\mathcal{S}$ of the problem we associate $X_s \subseteq V_s$. Moreover, to each reserve $\mathcal{X}_k \in \mathcal{S}$ we associate $X_s(k) \in X_s$, the

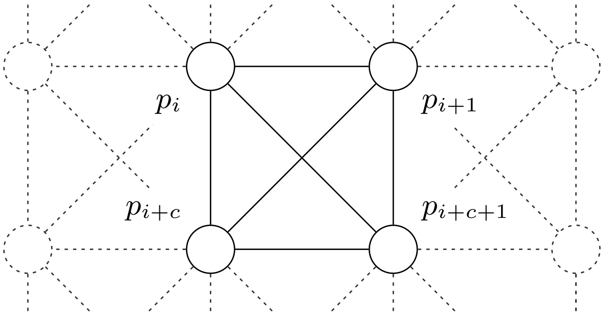


Figure 5.6: Illustration of a portion of a spatial graph G_s associated with a $r \times 4$ square grid, using the 8-connectivity neighborhood definition.

vertices associated to the parcels of $\mathcal{X}_k$. We now express each spatial criterion as a constraint that can be applied on $G_s[X_s]$.

Number of Reserves. We easily express this criterion by bounding the number of connected components (`NCC`, Dooms, 2006; Beldiceanu et al., 2005, 2006) in $G_s[X_s]$.

Constraint 4: Number of Reserves.

Input parameter(s): Two positive integer N_{min} and N_{max}.

Ensuring that the number of reserves in $\mathcal{S}$ is bounded by N_{min} and N_{max} is equivalent to bounding the `NCC` of $G_s[X_s]$ with N_{min} and N_{max}.

$$N_{min} \leqslant \texttt{NCC}(G_s[X_s]) \leqslant N_{max}. \tag{5.13}$$

Reserve Areas. We express this criterion as a constraint on the number of vertices of the smallest connected component of $G_s[X_s]$ (`MIN_NCC`, Beldiceanu et al., 2012, 2005, 2006) and on the number of vertices of the largest connected component of $G_s[X_s]$ (`MAX_NCC`, Beldiceanu et al., 2012, 2005, 2006).

Constraint 5: Reserve Areas.

Input parameter(s): Two positive integer A_{min} and A_{max}.

Ensuring that the area of every reserve $\mathcal{X}_k \in \mathcal{S}$ is bounded by A_{min} and A_{max} is equivalent to constraining the lower bound of $\texttt{MIN\_NCC}(G_s[X_s])$ to A_{min} and the upper bound of $\texttt{MAX\_NCC}(G_s[X_s])$ to A_{max}.

$$\forall k \in [\![0, n_r[\![, \quad \begin{aligned} &\texttt{MIN\_NCC}(G_s[X_s]) \geqslant A_{min}; \\ &\texttt{MAX\_NCC}(G_s[X_s]) \leqslant A_{max}. \end{aligned} \tag{5.14}$$

Reserve System Area. In the current case of a regular tessellation method, we can control the whole reserve system's area by bounding the norm of X_s.

Constraint 6: Reserve System Area.

Input parameter(s): Two positive integer $A_{T_{min}}$ and $A_{T_{max}}$.

Ensuring that the total area of the reserve system is bounded by $A_{T_{min}}$ and $A_{T_{max}}$ is equivalent to bounding $|X_s|$.

$$A_{T_{min}} \leqslant |X_s| \leqslant A_{T_{max}}. \tag{5.15}$$

5.4 THE CP MODEL

In this section we present our CP model for the reserve design problem. For its implementation, we rely on the solver Choco (Prud'homme et al., 2017) and its extension Choco-graph (Fages et al., 2018), which provides graph variables and constraints.

The Decision Variables. We naturally model the parcels with a boolean variable array, named `parcels`. If the parcel i is selected in the reserve system, `parcels[i]` $= 1$, else `parcels[i]` $= 0$.

```
BoolVar[] parcels = model.booVarArray("parcels", n);
```

These decision variables are the cornerstone of our CP model because they allow us to aggregate the two models we introduced in the previous section.

The Feature Covering Constraints. Given the particular configuration of the resource allocation graph, we are able to express each feature covering constraint with several local flow conservation inequalities, one for each feature involved in the constraint. Note that we would certainly benefit from the filtering of a global flow constraint (Bockmayr et al., 2001). However, there is no such constraint implemented in Choco at the time we are writing this paper. We thus keep this idea for future work.
Constraint A, Covered Features (5.7): with local flow conservation inequalities, (5.7) becomes:

$$\forall j \in \mathcal{F}', \quad \sum_{i=0}^{n-1} b_i \times (v_{ji} \geqslant 1) \geqslant 1.$$

Below is the implementation with Choco 4, using the `scalar` constraint.

```
for (int j : featuresToCover) {
  int[] coeffs = Arrays.stream(V[j])
    .mapToInt(v -> (v >= 1) ? 1 : 0)
    .toArray();
  model.scalar(parcels, coeffs, ">=", 1).post();
}
```

Constraint 2, α-Covered Features (5.9): the coefficients in the `scalar` constraint must be integers. We then retain only two digits of precision for the probabilities of presence. If $\alpha \in [0, 0.99]$ then $-\log(1-\alpha) \in [0, 2]$, moreover, with this precision the order of the smallest variation between two values ($\alpha = 0$

and $\alpha = 0.01$) is 10^{-3}, we thus multiply our local flow inequality by 10^3 in order to stay in the integer domain. If $v_{ji} \geqslant 1$, we set the flow upper bound to $-10^3 \log(1 - 0.999) = 3000$ as a replacement for $+\infty$ .Consequently, we reduce (5.9) to:

$$\forall j \in \mathcal{F}', \quad \sum_{i=0}^{n-1} b_i \times \min(-10^3 \log(1 - v_{ji}), 3000) \geqslant -10^3 \log(1 - \alpha).$$

Below is the implementation with Choco 4.

```
for (int j : featuresToCover) {
  int[] coeffs = Arrays.stream(V[j])
    .mapToInt(
      v -> (v >= 1) ? 3000 : (int) (-1000 * Math.log10(1 - v)))
    .toArray();
  int scaled = (int) (-1000 * Math.log10(1 - alpha));
  model.scalar(parcels, coeffs, ">=", scaled).post();
}
```

Constraint 3, k-Redundant Features (5.10): similarly, we reduce (5.10) to:

$$\forall j \in \mathcal{F}', \quad \sum_{i=0}^{n-1} b_i \times (v_{ji} \geqslant 1) \geqslant k.$$

And implement it the following way with Choco 4:

```
for (int j : featuresToCover) {
  int[] coeffs = Arrays.stream(V[j])
    .mapToInt(v -> (v >= 1) ? 1 : 0)
    .toArray();
  model.scalar(parcels, coeffs, ">=", k).post();
}
```

The Spatial Constraints. We rely on Choco-graph to express the spatial constraints in our CP Model. First, we use a graph variable g to model the reserve system. Its kernel is the empty graph (`GLB` in the code), and its envelope is G_s (`GUB` in the code).

```
UndirectedGraph GLB = new UndirectedGraph(model, n, BIPARTITESET, false);
UndirectedGraph GUB = new UndirectedGraph(model, n, BIPARTITESET, false);
for (int i = 0; i < n; i++) {
  GUB.addNode(i);
  for (int ii : getNeighbors(i)) {
    GUB.addEdge(i, ii);
  }
}
UndirectedGraphVar g = model.graphVar("g", GLB, GUB);
```

Then, we link the graph variable g with the boolean variables `parcels` using the `nodesChanneling` constraint.

```
model.nodesChanneling(g, parcels).post();
```

We also force the existence of an edge between two selected adjacent parcels through an `edgeChanneling` constraint with a reified `and` constraint between each pair $(\texttt{i1}, \texttt{i2})$ of adjacent parcels. Doing so, we ensure that every existing edges between two selected vertices are also present in our graph variable.

```
BoolVar forceEdge = model.and(parcels[i1], parcels[i2]).reify();
model.edgeChanneling(g, forceEdge, i1, i2).post();
```

Constraint 4, Number of Reserves (5.13): we use the `nbConnectedComponents` and the `arithm` constraints.

```
IntVar nbCC = model.intVar("nbCC", Nmin, Nmax);
model.nbConnectedComponents(g, nbCC).post();
```

Constraint 5, Reserve Areas (5.14): at the time we are writing this paper, there is no constraint in Choco-graph for controlling the `MIN_NCC` and `MAX_NCC` graph properties. We thus implemented the `sizeConnectedComponents`[1] constraint, which allows us to bound `MIN_NCC` and `MAX_NCC`.

```
model.sizeConnectedComponents(g, Amin, Amax).post();
```

Constraint 6, Reserve System Area (5.15): we can control the number of vertices of G_s (that is, the number of parcels) through the `nbNodes` graph constraint, or through the `sum` constraint over the decision variables.

```
IntVar nbParcels = model.intVar(Atmin, Atmax);
model.nbNodes(g, nbParcels).post(); // Option 1
model.sum(parcels, "=", nbParcels).post(); // Option 2
```

5.5 USE CASE: RAINFOREST FRAGMENTATION IN NEW CALEDONIA

New Caledonia is biodiversity hotspot located in the South Pacific, slightly north of the tropic of the Capricorn. The flora of this large archipelago is distinguished by an exceptionally high rate of endemism. Like most of the world's remaining natural forests, New Caledonian rainforests are endangered with surface loss and fragmentation. A case study had been conducted in the south of New Caledonia in order to highlight "how does forest fragmentation affect tree communities" (Ibanez et al., 2017). We relied on this case study and its associated dataset (up to date) for our use case, and considered the following fictive but realistic operational scenario:

---

[1] https://gist.github.com/dimitri-justeau/8098af35824bbf8d52ef21282291e621

> *"We want to establish a reserve system in which a pool of endangered species must be present. In addition, most of the other species known in the area must have a high probability to occur, or a high habitat suitability. The reserve system must be mostly covering rainforest areas. Its area and its number of reserves must be limited because of budget limitation. Moreover, each reserve must be large enough to ensure the persistence of the species."*

Note. In this scenario, the objective is to protect both existing and potential rainforest areas. To do so, we relied on SDM layers that were generated with presence-only data and thus produce a score of habitat suitability rather than a standardized probability of presence. A high habitat suitability in a non-rainforest zone can then be interpreted as an adequate zone for recolonization.

5.5.1 *Input Data, Constraints and Parameters*

The original dataset consists of the mapping of a 60 km^2 landscape where 97 tree communities had been sampled in 88 digitized rainforest fragments (forest/non-forest). The dataset gathers 5431 identified trees belonging to 223 species. Moreover, an SDM raster layer was available for 173 of the species (Pouteau et al., 2015; Schmitt et al., 2017). Arbitrarily, we considered the 50 species without SDM as the endangered ones. We then prepared this dataset by tessellating the study area into a 46×75 regular square grid and by rasterizing the dataset according to this grid. Each parcel then has an area of about 1.7 ha. Note that we also defined a set of forbidden parcels corresponding to lakes and mining sites.

From this point, we defined a feature for each observed species in the area. When available, we relied on the SDM layer for the feature data (probability of presence data). We forced the values to 1 for the parcels where an observation is available. When no SDM was available, we only relied on the occurrence dataset (presence-absence). We represented the rainforest coverage as a presence-absence feature.

We then applied the *Covered Features* constraint for the set of endangered species, and the α-*Covered Features* constraint for the other species with $\alpha = 0.8$. In order to ensure a minimum rainforest area of 340 ha in the reserve system, we applied the *k-Redundant Features* constraint for the rainforest coverage, with $k = 200$ parcels. Moreover, we enforced the forbidden parcels on the envelope of the graph variable g. We then set the minimum area of the reserves to $A_{min} = 176$ parcels (about 300 ha) with the *Reserve Areas* constraint. In addition, we limited the reserve system area using the *Reserve System Area* constraint, with $A_{t_{max}} = 589$ parcels (about 1000 ha). According to those restrictive parameters, we allowed the number of reserve to be at most two, using the *Number of Reserves* constraint, with $N_{min} = 1$ and $N_{max} = 2$.

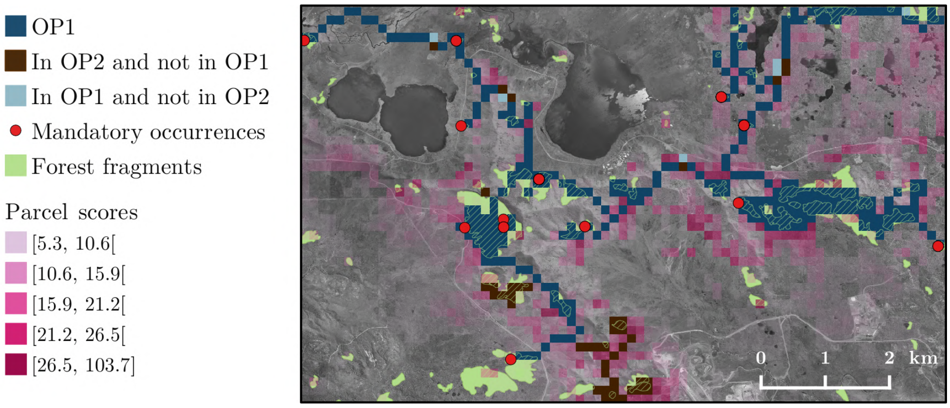


Figure 5.7: Mapping of the use case best solutions. The parcel scores correspond to the heuristic score and the mandatory occurrences to rare species observed only once in the study area (they must then be covered by any solution).

Table 5.1: Use case results: resolution times and solution characteristics. All experiments were run on an Intel Core i5-5200U CPU (2.20GHz×4), with 7.7GB of RAM.

	DP	OP1	OP2
Resolution time	28s	3h24m	1h5m
Number of solutions found	1	8	25
Number of reserves	1	1	1
Number of parcels	318	292	328
Number of rainforest parcels	200	200	224

5.5.2 *Questioning and Results*

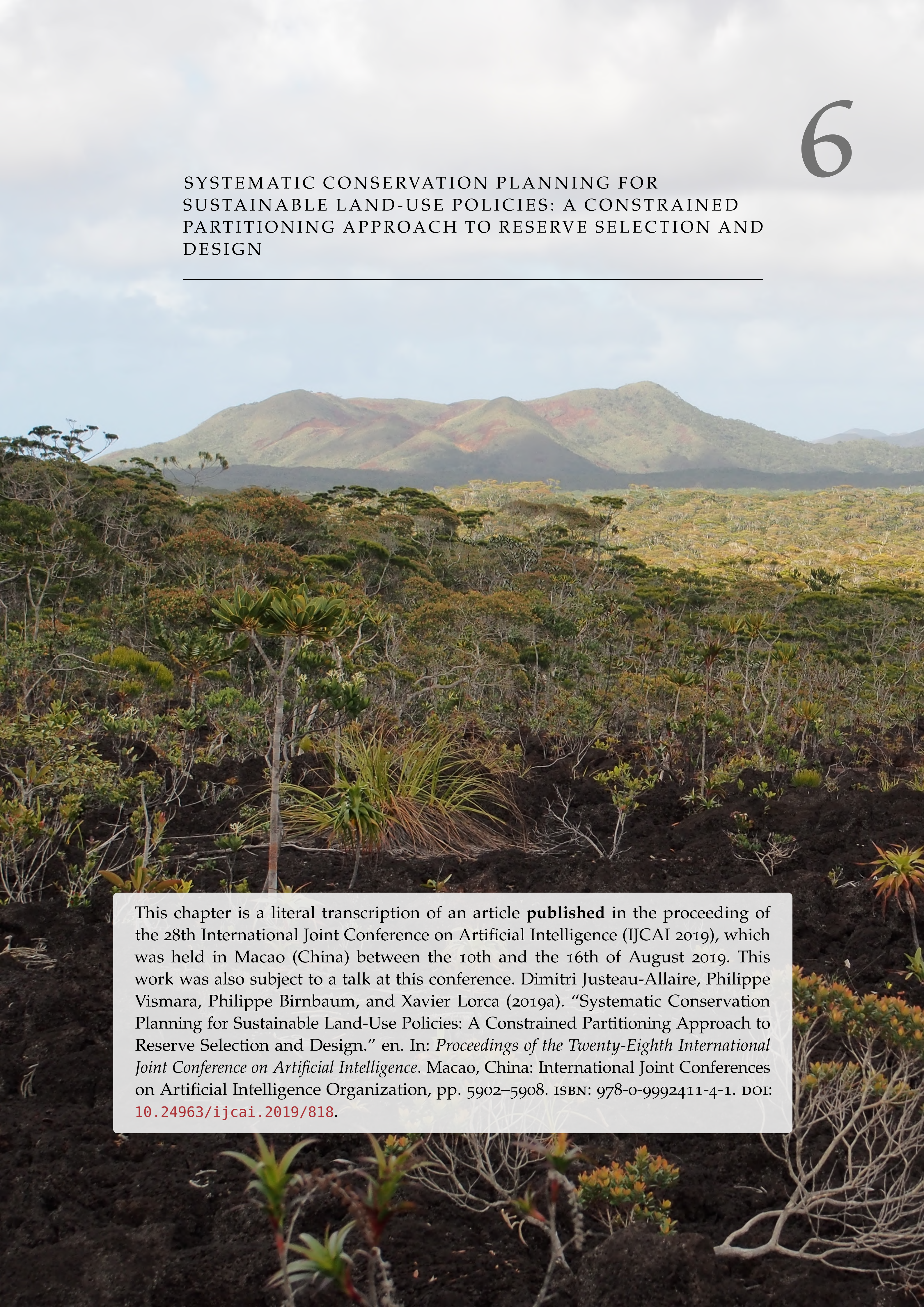
In the first place, the number of reserves and the number of parcels are critical parameters of our use case: the less the better. This is why we started by implementing a search strategy that starts by branching on the lower bound of the `nbCC` variable and continues by selecting the lower bound of the `parcels` variables, sorted in descending order by a score corresponding to the number of features with a value greater than 0.6 (cf. 5.7). The solver quickly found a solution to the decision problem (DP), as summarized in Table 5.1. Given that, we ran a first optimization problem (OP1) where we tried to minimize the total area of the reserve system, that is the `nbParcels` variable. We limited the computation time to 4 hours and retrieved the best solution found, which reduced the total area by 8% in comparison to DP (cf. Table 5.1). In order to cover more rainforest parcels, we ran a second optimization problem (OP2) in which we forced the `nbParcels` variable to be within 15% of the best value found in OP1, and tried to maximize k (the number of forest parcels), thus defined as an integer variable. After a limited run of 4 hours, we could increase the area of rainforest by 12% (cf. Table 5.1). A mapping of our results is provided in Fig. 5.7.

5.6 CONCLUSION AND CHALLENGES

To the best of our knowledge this paper tackles, for the very first time, the reserve design problem from a constraint programming point of view. It is also the first time that a reserve design model integrates such a diversity of constraints, simultaneously involving decisions based on occurrences, SDMs and spatial attributes with an exact approach. Although performance enhancements are needed, the combination of graph-based modeling and constraint programming reveals as a powerful and promising framework for dealing with the reserve design problem.

Based on a challenging use case, our model highlighted a solution compatible with the conservation strategy, namely a trend to link isolated forest patches in order to enhance the functioning of tree communities. However, in this use case we restrained to a binary landscape only composed of forest/non-forest while it is often assumed that a reserve system must include an assemblage of several landscape types. In such a mosaic, an important challenge lays in weighting and balancing the reserve system characteristics and shape in order to maintain (or restore) the functional connectivity inside and between the reserves. In fact, the functional isolation of an habitat leads to a reduction of biological flows, which tends to amplify its spatial isolation. Moreover, since the underlying processes are dynamic, robust solutions must rely both on the current state and future scenarios. It also remains to model the impacts of a reserve system on the off-reserve area, such as the creation of boundaries or enclosed areas.

These elements lead us to identify several lacks and challenges. First of all, the main lack is that Choco solver does not offer an implementation of the flow constraint (Steiger et al., 2011; Downing et al., 2012). We will focus on its implementation in future work. Next, a bottleneck in the constraint propagation is the interaction between the constraint on the number and the size of the connected components (Beldiceanu et al., 2012). We actually treat each one independently but we think that there is a possible enhancement of the filtering by dealing with their interaction. A first challenge for future work concerns our capacity to model constraints on the shape of the reserves by using graph properties, such as graph diameter, in order to design reserves that are compatible with the long-term persistence of species. A second challenge is more oriented to decision making aspects such as identifying key areas that have to be present in any solution. Finally, a last challenge is related to our capacity to take a dual point of view: is it possible to take into account managers' needs on the off-reserve area by adding constraints on the same graph representation?



6

SYSTEMATIC CONSERVATION PLANNING FOR SUSTAINABLE LAND-USE POLICIES: A CONSTRAINED PARTITIONING APPROACH TO RESERVE SELECTION AND DESIGN

This chapter is a literal transcription of an article **published** in the proceeding of the 28th International Joint Conference on Artificial Intelligence (IJCAI 2019), which was held in Macao (China) between the 10th and the 16th of August 2019. This work was also subject to a talk at this conference. Dimitri Justeau-Allaire, Philippe Vismara, Philippe Birnbaum, and Xavier Lorca (2019a). "Systematic Conservation Planning for Sustainable Land-Use Policies: A Constrained Partitioning Approach to Reserve Selection and Design." en. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*. Macao, China: International Joint Conferences on Artificial Intelligence Organization, pp. 5902–5908. ISBN: 978-0-9992411-4-1. DOI: 10.24963/ijcai.2019/818.

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.

6.1 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 6.1 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 methods 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

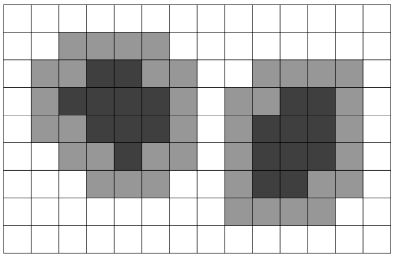


Figure 6.1: Grid partition with a core area (dark gray), a buffer (light gray), and an out-reserve area (white). The core and the buffer are both composed of two connected components.

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.

6.2 PROBLEM DESCRIPTION

6.2.1 *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).

Definition 6.1 (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).

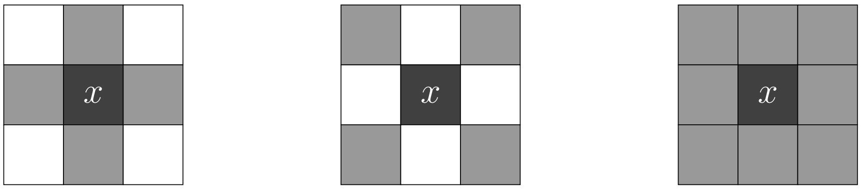
Let $\mathcal{S}$ be a regular square grid.

Definition 6.2 (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 6.2), $\forall x \in \mathcal{S}$, with $x \equiv (x_i, x_j)$:

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

Definition 6.3 (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.



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

Figure 6.2: Some neighborhood definitions (in light gray).

Definition 6.4 (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 $cc(R)$.

6.2.2 *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 $\nu_f(x)$.

6.2.3 *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.

6.2.3.1 *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 $\nu_f(x) \geqslant 1$ that is $\forall f \in \mathcal{F}, \exists x \in R, \nu_f(x) \geqslant 1$.

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: $\forall f \in \mathcal{F}, \prod_{x \in \mathcal{S}}(1 - \nu_f(x)) \leqslant 1 - \alpha$.

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: $\forall f \in \mathcal{F}, \exists \mathcal{X} \subseteq R, |\mathcal{X}| \geqslant k \wedge \forall x \in \mathcal{X}, \nu_f(x) \geqslant 1$.

6.2.3.2 *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: $minNbCC \leqslant |cc(R)| \leqslant maxNbCC$.

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

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: $\forall C \in cc(R), minSizeCC \leqslant |C| \leqslant maxSizeCC$.

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 $\mathcal{C}_u \subseteq \{A, \dots, F\}$. Using this formulation, any constraint in the previous catalog can be seamlessly applied to any region.

6.3 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.

6.3.1 *The Base Model*

6.3.1.1 *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: $\forall x \in \mathcal{S}, \rho_x \in [0, n[$.

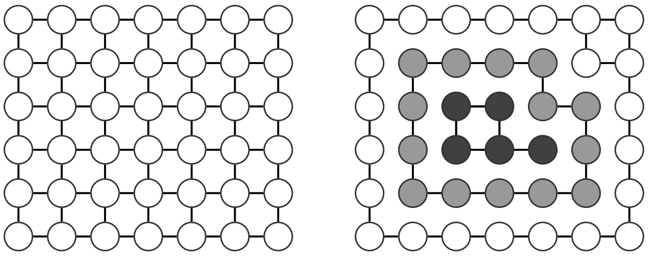
6.3.1.2 *Set Variables*

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): $X \in [\underline{X}, \overline{X}] \Leftrightarrow \underline{X} \subseteq X \subseteq \overline{X}$.

Each region is represented by a set variable $R_u \in [\emptyset, P(\mathcal{S})]$ 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 $\mathcal{S}$: $\forall u \in [0, n[, R_u \in [\emptyset, \mathcal{S}]$, $\rho_x = u \Leftrightarrow x \in R_u$.

6.3.1.3 *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



(a) $G_{\mathcal{S}}$.

(b) $\{G_0, G_1, G_2\}$

Figure 6.3: $G_{\mathcal{S}}$ 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).

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 6.3. 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 $E_u = \{(x, y) \mid (x, y) \in {R_u}^2 \land y \in \Gamma_4(x)\}$. An illustration of $\{G_0, G_1, G_2\}$ for a 3-regions partitioning is provided in Figure 6.3.

6.3.1.4 *User Constraints*

Finally, any constraint defined in 6.2.3 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.

6.3.2 *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 6.5 (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 *buffer*$[\Gamma_\omega](R_u, R_v, B)$ holds if and only if:

$$\begin{aligned} &\Gamma_\omega(R_u) \cap R_v = \emptyset; \\ &R_u \cap \Gamma_\omega(R_v) = \emptyset; \\ &B = \Gamma_\omega(R_u) \cap \Gamma_\omega(R_v). \end{aligned} \tag{6.1}$$

6.3.2.1 *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 $X \subseteq Y, X = Y \cup Z, X = Y \cap Z, X = Y - Z, X \neq Y, |X| = I, occ(I, X) = m$ or $occ(m, X) = I$, where X, Y and Z are set/multiset variables, I an integer variable and m an integer. The constraint $occ(m, X) = I$ (respectively $occ(I, X) = m$) 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:

$$N = \Gamma_\omega(R) \Leftrightarrow N = \bigcup_{x \in S} N_x^R \text{ with } N_x^R = \begin{cases} \Gamma_\omega(x) & \text{if } x \in R \\ \emptyset & \text{otherwise} \end{cases} \tag{6.2}$$

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

$$\begin{aligned} \forall x \in \mathcal{S}, \quad & N_x^R \in \{\emptyset, \Gamma_\omega(x)\} \\ & B_x = occ(x, R) \\ & I_x = |\Gamma_\omega(x)| B_x \\ & |N_x^R| = I_x \end{aligned} \tag{6.3}$$

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

Proof. The *buffer* constraint as defined in (6.1) can be decomposed as following:

$$\begin{aligned} &N_u = \Gamma_\omega(R_u);\ N_v = \Gamma_\omega(R_v); \\ &N_u \cap R_v = \emptyset;\ R_u \cap N_v = \emptyset; \\ &B = N_u \cap N_v. \end{aligned} \tag{6.4}$$

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

6.3.2.2 *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 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 $O(|S|^4)$. ☐

6.3.3 *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.

6.3.3.1 *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 6.4.

Definition 6.6 (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 $\Gamma^k_\omega(x)$, is defined by the following recursion:

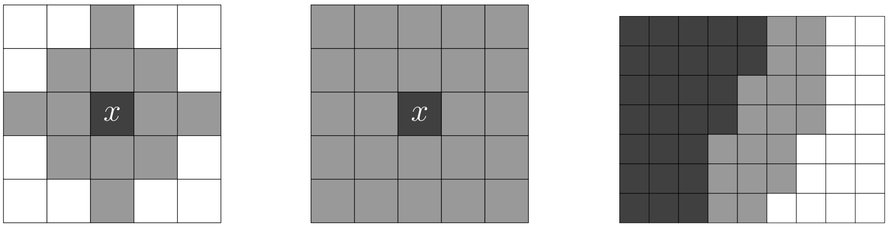
$$\Gamma^1_\omega(x) = \Gamma_\omega(x);$$
$$\Gamma^k_\omega(x) = \left(\Gamma^{k-1}_\omega(x) \cup \Gamma_\omega\left(\Gamma^{k-1}_\omega(x)\right)\right) \setminus \{x\}.$$

6.3.3.2 *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 6.5. 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: $\forall i \in [0, m[, \textit{buffer}[\Gamma_\omega](R_i, \mathcal{S} \setminus (R_i \cup R_{i+1}), R_{i+1})$. 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.

6.4 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) and Schmitt et al. (2017). The area is tessellated into a 46×75 square grid



(a) $\Gamma^2_4(x)$ (b) $\Gamma^2_8(x)$ (c) 2-wide buffer zone

Figure 6.4: k-wide neighborhood examples (in light gray). Example of a 2-wide buffer zone.

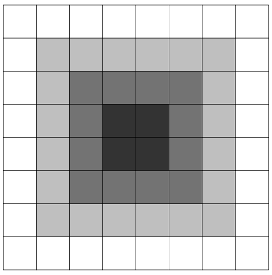


Figure 6.5: Nested regions example, 4 nested regions, $\Gamma_\alpha = \Gamma_8$.

(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 6.1 (none were proven optimal).

6.4.1 *Original Scenario (SC1)*

First, we show that our model is able to tackle the use case as defined by Justeau-Allaire et al. (2018). 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 6.6a). 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).

---

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

6.4.2 *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. (2017) 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. (2017) 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 6.6b) with a 130–180 m wide buffer zone $\mathit{buffer}[\Gamma_8](R_0, R_2, R_1)$, and a pessimistic one (SC3, Figure 6.6c) 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.

6.4.3 *Final Scenario (SC4)*

Finally, we suggest a trade-off between SC2 and SC3 with SC4 (Figure 6.6d). 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.

6.5 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

---

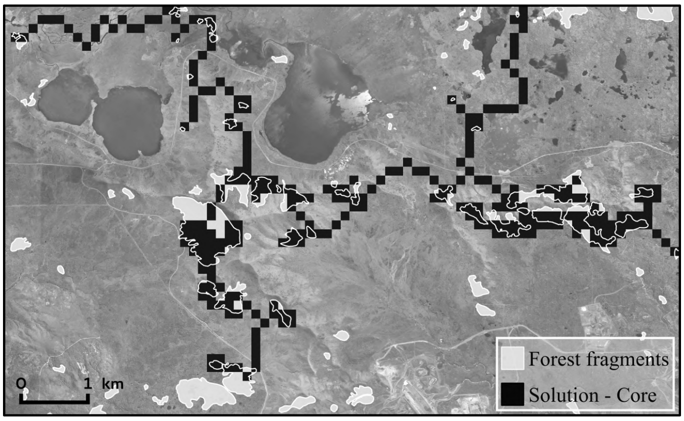
[2] Width is variable because of the square grid: diagonal is $\sqrt{2}$ larger than horizontal/vertical.

	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%

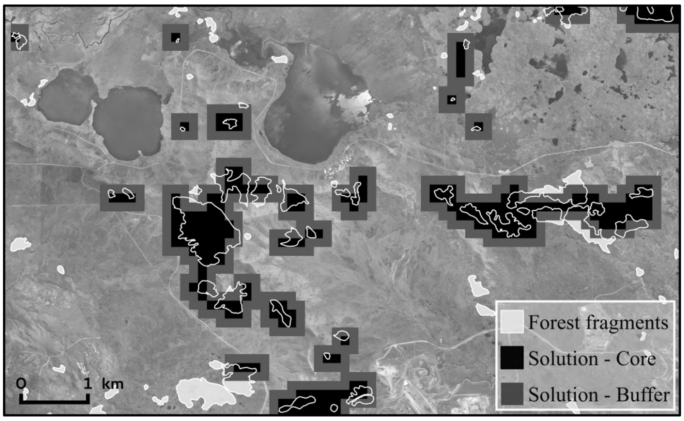
Table 6.1: Use case results characteristics.

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 6.1), demonstrating its potential for exploring many scenarios.

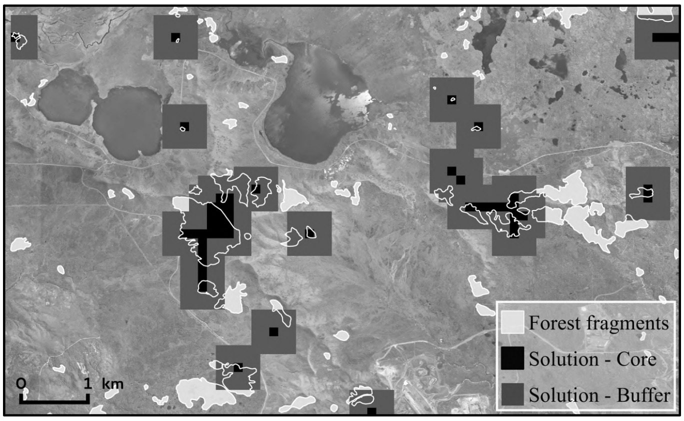
To conclude, our constrained partitioning approach for reserve selection and design provides the basis of a generic and exploratory decision support tool for systematic conservation planning and computational sustainability. It now remains to work closely with conservation scientists and managers to 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.



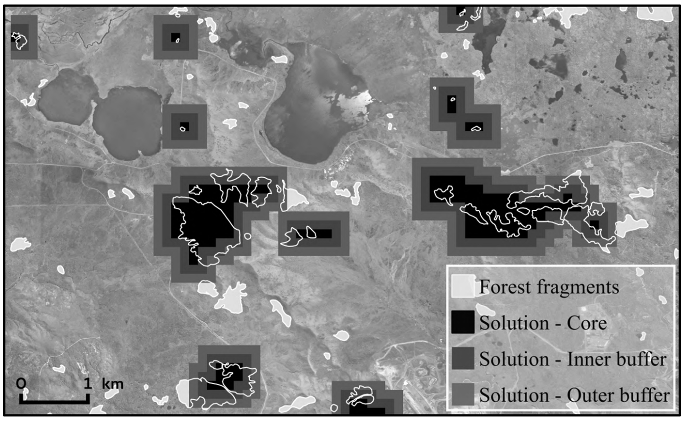
(a) Mapping of SC1's best solution.



(b) Mapping of SC2's best solution.

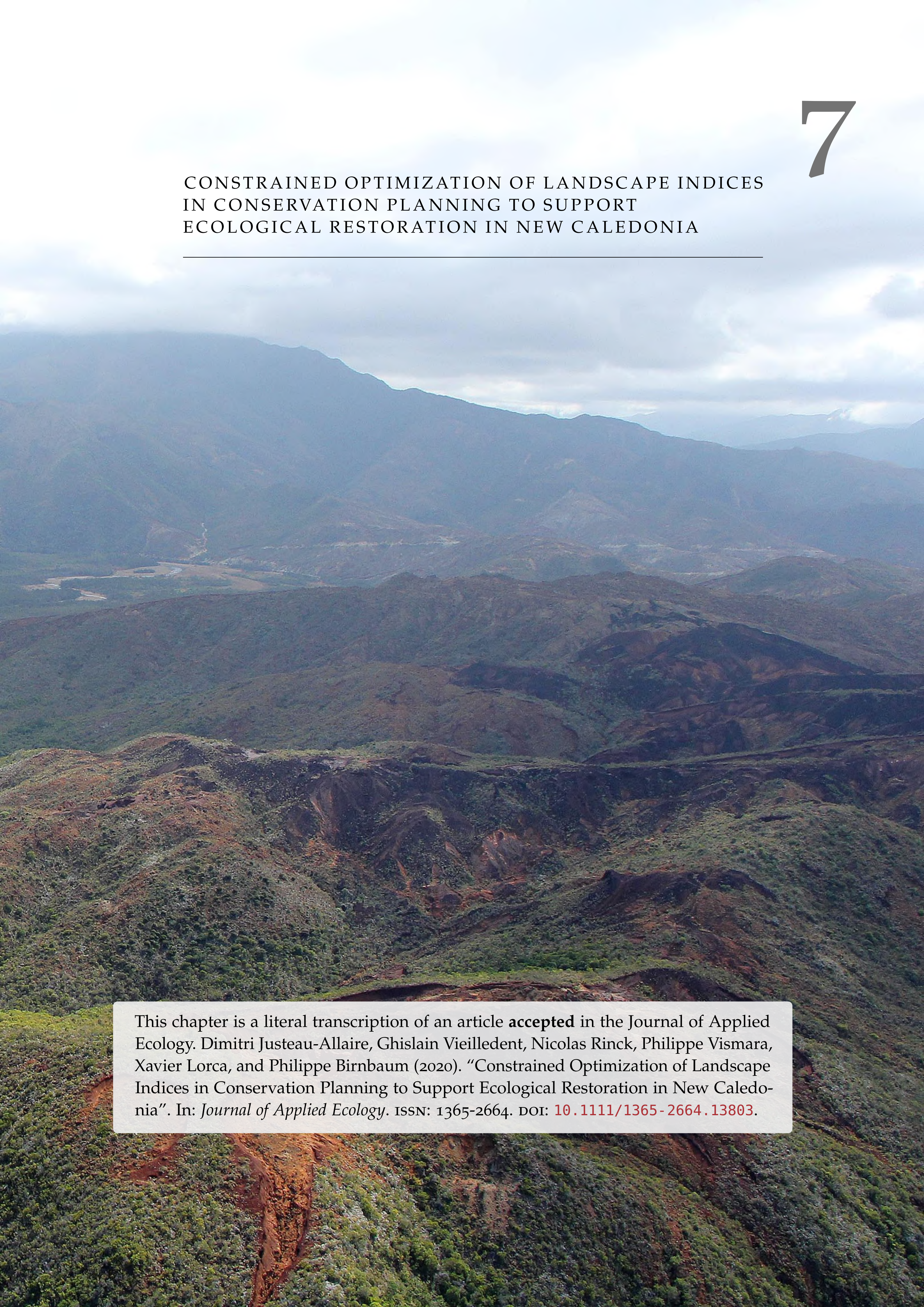


(c) Mapping of SC3's best solution.



(d) Mapping of SC4's best solution.

Figure 6.6: Use case scenarios best solutions mappings.



7

CONSTRAINED OPTIMIZATION OF LANDSCAPE INDICES IN CONSERVATION PLANNING TO SUPPORT ECOLOGICAL RESTORATION IN NEW CALEDONIA

This chapter is a literal transcription of an article **accepted** in the Journal of Applied Ecology. Dimitri Justeau-Allaire, Ghislain Vieilledent, Nicolas Rinck, Philippe Vismara, Xavier Lorca, and Philippe Birnbaum (2020). "Constrained Optimization of Landscape Indices in Conservation Planning to Support Ecological Restoration in New Caledonia". In: *Journal of Applied Ecology*. ISSN: 1365-2664. DOI: 10.1111/1365-2664.13803.

Abstract

1. Curbing habitat loss, reducing fragmentation, and restoring connectivity are frequent concerns of conservation planning. In this respect, the incorporation of spatial constraints, fragmentation, and connectivity indices into optimization procedures is an important challenge for improving decision support.

2. Here we present a novel optimization approach developed to accurately represent a broad range of conservation planning questions with spatial constraints and landscape indices. Relying on constraint programming, a technique from artificial intelligence based on automatic reasoning, this approach provides both constraint satisfaction and optimality guarantees.

3. We applied this approach in a real case study to support managers of the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" provincial park project, in the biodiversity hotspot of New Caledonia. Under budget, accessibility, and equitable allocation constraints, we identified restorable areas optimal for reducing forest fragmentation and improving inter-patch structural connectivity, respectively measured with the effective mesh size and the integral index of connectivity.

4. *Synthesis and applications.* Our work contributes to more effective and policy-relevant conservation planning by providing a spatially-explicit and problem-focused optimization approach. By allowing an exact representation of spatial constraints and landscape indices, it can address new questions and ensure whether the solutions will be socio-economically feasible, through optimality and satisfiability guarantees. Our approach is generic and flexible, thus applicable to a wide range of conservation planning problems such as reserve or corridor design.

Keywords: conservation planning, constraint programming, New Caledonia, structural connectivity, forest-fragmentation, ecological restoration, restoration planning, fragmentation.

7.1 INTRODUCTION

As the Earth has entered the Anthropocene, human impacts on the environment have led to the current global biodiversity crisis. Habitat loss and degradation due to land-use change are the leading causes of ecosystem collapse and biodiversity decline (Haddad et al., 2015). Landscape configuration can also have profound impacts on ecological processes such as dispersal, gene flow, or fire resistance (Taylor et al., 1993; Fahrig, 2003). These impacts are often assessed through habitat fragmentation metrics and inter-patch connectivity measures (Uuemaa et al., 2013). Fragmentation refers to the spatial patterns of habitat distribution (Fahrig, 2003) and inter-patch connectivity to the potential ability of species to migrate or disperse between habitat patches (Taylor et al., 1993).

Restoration and conservation planning can help to curb habitat loss and promote suitable landscape configurations, as well as helping to identify trade-offs between conservation targets and managers' objectives (Rodrigues et al., 2000; Knight et al., 2008). Efficient decision support processes must rely on spatially-explicit, systematic, and reproducible approaches (Pressey et al., 1993). Over the last few decades, many such approaches have been devised, from geometric principles derived from biogeography theory (Diamond, 1975) to the principle of complementarity in the representation of biodiversity features (Vane-Wright et al., 1991). Systematic conservation planning (SCP) is now an active field of conservation biology. There is also a consensus on the mutual importance of spatial configuration and the representation of biodiversity features in the planning of conservation actions, to express managers' constraints as much as ecological requirements (Margules and Pressey, 2000; Williams et al., 2005).

Many optimization methods for SCP have been proposed, mostly relying on *ad hoc* heuristics, metaheuristics, or mixed-integer linear programs (MILP). *Ad hoc* heuristics are problem-specific local search algorithms either based on a forward (greedy) e.g. Kirkpatrick, 1983; Nicholls and Margules, 1993 or backward (stingy) procedure e.g. Zonation software, Moilanen et al., 2014. In constructive heuristics (resp. destructive), solutions are obtained by iteratively adding (resp. removing) the planning unit which offers the highest gain (resp. loss) according to an objective function to maximize (resp. minimize). Metaheuristics are high-level and problem-independent stochastic search heuristics, such as simulated annealing e.g. Marxan software, Ball et al., 2009 or tabu search e.g. ConsNet software, Ciarleglio et al., 2010. The main advantage of heuristics is that they are often straightforward to understand and implement, but produce solutions of unknown quality relative to optimality. Finally, MILP is a constrained mathematical optimization approach where the objective function and the constraints are stated as linear equations, with some or all the variables being integers (Billionnet 2013; Dilkina et al. 2017; oppr R package, Hanson et al. 2019a). Exact approaches such as MILP can require more time to generate solutions than heuristics, however they offer guarantees relative to optimality and constraint satisfaction. Indeed, even though heuristics can reach constraint satisfaction for loosely constrained problems (e.g. species set covering problem, ReVelle et al. 2002), they can fail to provide this guarantee for highly constrained problems (e.g. Billionnet, 2013). Constraint satisfaction problems on a finite domain are indeed in general NP-Complete (Dechter, 2003). Although less widely used, dynamic programming approaches e.g. Meir et al., 2004 and Markov decision processes e.g. Schapaugh and Tyre, 2012 have also brought substantial advances in SCP but are limited to smaller problem sizes than the approaches described above.

Recent work has introduced several perspectives towards the integration of landscape spatial configuration in SCP optimization procedures. For instance, Marxan software uses a boundary length penalty in its objective function to influence the spatial configuration of the solutions. Additionally, Marxan Connect (Daigle et al., 2020) provides many options to include structural or functional connectivity data in Marxan's input. Similarly, Zonation provides eight different methods to integrate connectivity in its prioritization process (Moilanen et al., 2014). In MILP approaches, several options are available to ensure spatial requirements, such as strictly guaranteeing the connectivity and compactness of delineated areas, or designing buffer zones (Billionnet, 2013; Wang and Önal, 2016). Other approaches such as LQGraph (Fuller and Sarkar, 2006) or Linkage Mapper (McRae et al., 2012) specifically aim to identify optimal corridors between core areas or existing protected areas. On the other hand, landscape ecologists have devised many indices to evaluate the level of fragmentation (McGarigal, 2014) and connectivity (Pascual-Hortal and Saura, 2006; Saura and Pascual-Hortal, 2007) within a landscape. Except Xue et al., 2017 and to the best of our knowledge, such connectivity and fragmentation indices were mainly used in scenario analysis contexts e.g. Bodin and Saura, 2010. Integrating such indices into constrained optimization approaches is difficult due to their non-linearity and the curse of dimensionality. Nonetheless, it would improve the value of decision support by taking into account more powerful and ecologically relevant metrics in SCP.

Recently, we introduced a novel and generic SCP framework based on constraint programming (Justeau-Allaire et al., 2019a), an exact constrained optimization technique based on automated reasoning. In this article, we have extended this framework with landscape indices and applied it in a current reforestation project in the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" provincial park in the New Caledonia biodiversity hotspot. We worked in close collaboration with New Caledonian environmental managers to provide spatially-explicit decision support focused on reducing forest fragmentation and isolation, which are known to have adverse effects on tree communities in this region (Ibanez et al., 2017). Under budget, land accessibility and equitable allocation constraints, we computed optimal solutions for two landscape indices: the effective mesh size (MESH; Jaeger, 2000) and the integral index of connectivity (IIC; Pascual-Hortal and Saura, 2006) applied to structural connectivity. MESH is a measure of landscape fragmentation which is based on the probability that two randomly chosen points are located in the same patch. Maximizing it in the context of reforestation favours fewer and larger forest patches. On the other hand, IIC is a graph-based inter-patch connectivity index based on a binary connection model. Its maximization in the context of reforestation favours restoring structural connectivity between large patches. Our results demonstrated the flexibility of this approach and how its expressiveness (i.e. the breadth and variety of problems that it can represent and solve) facilitates the representation of the inherent diversity of real-world conservation problems, offering new perspectives for designing decision support tools in ecological restoration and more broadly in conservation planning (e.g. for reserve or corridor design).

7.2 MATERIAL AND METHODS

7.2.1 *Case study: reforestation planning in the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" provincial park, New Caledonia*

New Caledonia is a tropical archipelago located in the South Pacific (see Figure 7.1.a). As the smallest biodiversity hotspot in the world, it hosts megadiverse marine and terrestrial ecosystems. Notably, New Caledonian flora is distinguished by one of the highest rates of endemism in the world – approximately 76% (Myers et al., 2000; Morat et al., 2012), a high beta-diversity (Ibanez et al., 2014), and the presence of relict taxa (Grandcolas et al., 2008; Pillon, 2012). However, New Caledonian forests are under threat and the remaining forest is highly fragmented, as the result of anthropic activities such as bushfires, logging, urbanization, and nickel mining. New Caledonia is an overseas French collectivity which was first populated by the Kanak people. In this territory, the French Common Civil Code coexists with the Customary Civil Code, and institutions such as the Customary Senate provide a political framework to the Kanak people for promoting their culture, traditions, and environment. In this respect, customary authorities of the "Côte Oubliée 'Woen Vùù – Pwa Pereeù", a large area in the Southeast of the main island of New Caledonia, "Grande Terre" (see Figure 7.1.b), established a moratorium on nickel mining activity between 2014 and 2016. They called for a cessation on any road, mining or infrastructure project, in response to the erosion of many areas, due to bushfires and mining activity. This moratorium was renewed for ten years (from 2018 to 2028) and led to the creation in April 2019 of the "Côte Oubliée 'Woen Vùù – Pwa Pereeù" Provincial Park by the South Province of New Caledonia. With 93000 ha of terrestrial and 27000 ha of marine protected area, the provincial park blocked 102 mining concessions, includes three existing natural reserves and is adjacent to four existing natural reserves (see Figure 7.1.c). It now remains for the managers of the South Province's Sustainable Development Department for the Territories (SDDT) to establish the management plan of the park, with a strong emphasis on reducing forest fragmentation.

In this study, we focus on a reforestation project that must be planned by the SDDT. One of its objectives is expected to be the zoning of two suitable areas for reforestation, one in each of the two customary districts of the Côte Oubliée, respectively Borendy and Unia, so as to involve both communities in the project. Since the Côte Oubliée is a low urbanized and mountainous area, most locations are difficult to access. Accordingly, to be accessible reforestation areas must be compact (within an enclosing circle whose maximum diameter is 1500 m) and close to existing tracks (at a maximal distance of 1000 m). In this study, we considered a realistic cost corresponding to 200 ha to reforest, equitably divided between Borendy and Unia (100 ha ± 10% in each district). Under these constraints, the aim was to optimize the potential contribution of the reforested areas to reduce forest fragmentation and improve forest structural connectivity in the provincial park.

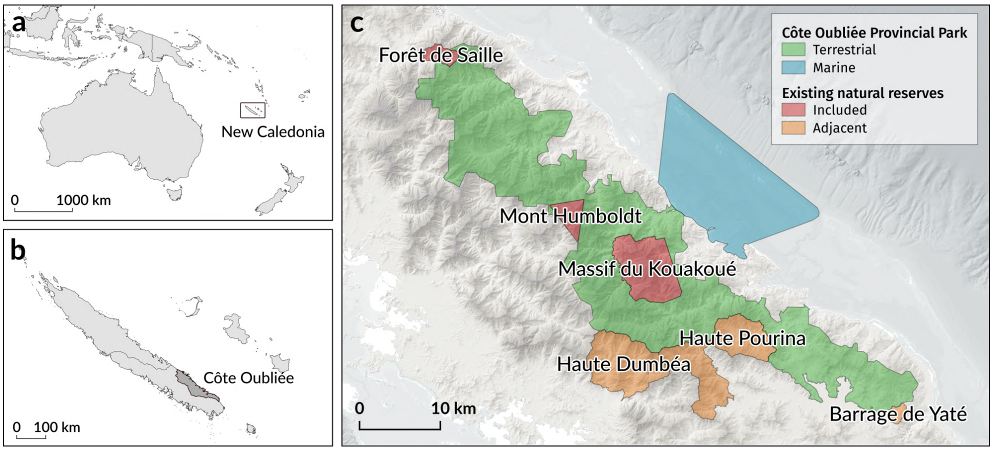


Figure 7.1: (**a**) Location of New Caledonia. (**b**) Location of the “Côte Oubliée” area. (**c**) Map of the “Côte Oubliée – Woen Vùù – Pwa Pereeù” provincial park, with included and adjacent existing natural reserves.

7.2.2 *Data*

The Côte Oubliée is a poorly studied area, and we still have little knowledge about the dispersal of New Caledonian animal and plant forest species see the last biological knowledge synthesis on the Côte Oubliée: Guillemot et al., 2016. Although species occurrences are useful to guide planning, the region is insufficiently sampled to ensure an unbiased selection. Species distribution models (SDMs) of tree species could also help to identify adequate reforestation areas. However, it would be necessary to have more occurrences in this region to obtain reliable predictions, due to the heterogeneity of tree community compositions which is still not well understood (Pouteau et al., 2019). In this respect, we adopted a forest-cover approach using remote sensing data (the dominant forest type in this area is dense rainforest). In this respect, we relied on a 2019 30 m binary forest-cover raster (cf. Figure 7.2.a), based on the historical analysis of temporal series from Landsat data (1982 to 2018) (Vancutsem et al., 2020). We focused on the extent of the Côte Oubliée Provincial Park (55.68 km height and 81.6 km width) and resampled the forest-cover raster to a resolution of 480 m (16×16 30 m cells) as a compromise between conservation planning and computational solving (480 m × 480 m ≈ 23 ha). We obtained a 116×170 raster map where each 480 m cell is characterized by a forest-cover proportion, according to the number of covered 30 m forest pixels. A 480 m cell was considered as degraded if its forest-cover proportion was smaller than 70% (Fahrig, 2013; Vieilledent et al., 2018). As reforestation must occur in the provincial park, we retained the cells within the boundaries of the provincial park to which we included parts of forest patches extending outside the park to avoid the boundary problem (Moser et al., 2007). The resulting raster map contained 3629 forest cells and 2715 non-forest cells, as illustrated in Figure 7.2.b. Consequently, we quantified the area to be reforested in each 480 m cell as the area needed to reach a forest-cover proportion of 70% (cf. Figure 7.2.d). Finally, we identified accessible areas for reforestation as a 1000 m buffer around tracks using the tracks vector data provided by the SDDT, classified according to the two customary districts covered by the provincial park, Borendy and Unia (see Figure 7.2.c).

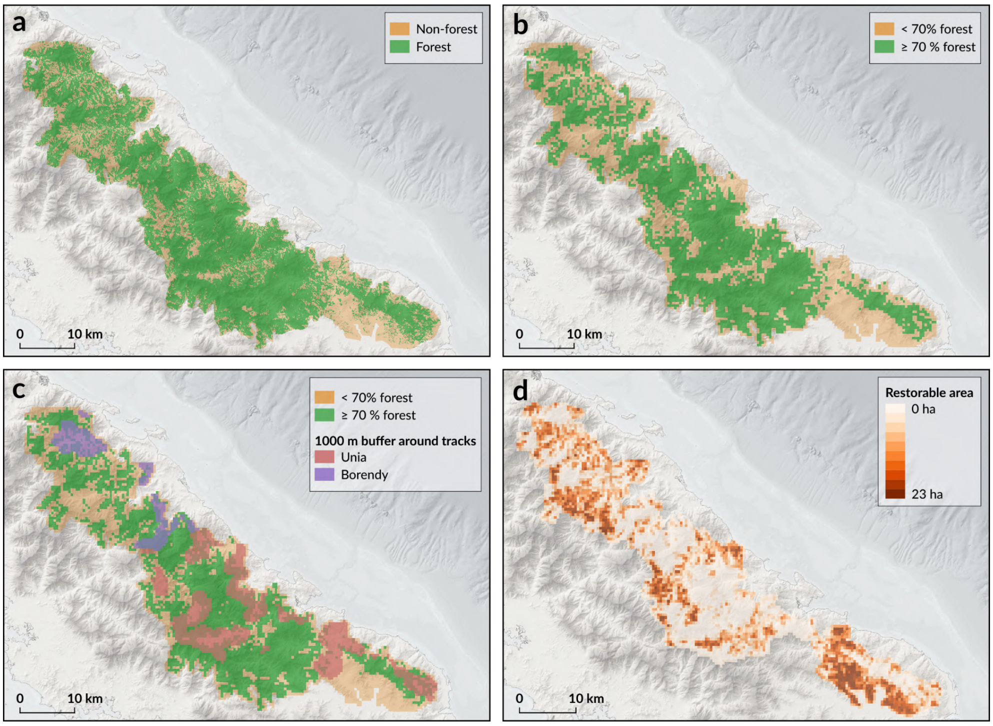


Figure 7.2: Input data maps. (**a**) 2019 30 m binary forest map produced from Landsat historical data analysis. (**b**) Upscaled 480 m binary forest map. A 480 m cell was considered as forest if its forest-cover proportion at 30 m was at least 70%. (**c**) 480 m accessible areas (1000 m buffer around tracks) map, classified by customary districts. (**d**) 480 m restorable area map, that is the non-forest area for each cell.

7.2.3 *Mathematical formulation*

7.2.3.1 *Base problem: variables and managers' constraints*

To each cell of the input raster grid we associate a planning unit (PU) that can be selected for reforestation, these are the decision variables of our base model. Let $\mathcal{S}$ be the set of PUs in the study area, we define the following subsets of $\mathcal{S}$ according to the data:

$$\begin{aligned}
&\mathcal{U}, \text{the set of accessible PUs located in the Unia district;}\\
&\mathcal{B}, \text{the set of accessible PUs located in the Borendy district;}\\
&\mathcal{F}_{\geqslant 70\%}, \text{the set of PUs with forest-cover proportion} \geqslant 70\%;\\
&\mathcal{F}_{<70\%}, \text{the set of PUs with forest-cover proportion} < 70\%.
\end{aligned} \tag{7.1}$$

Let $R_u \subseteq \mathcal{F}_{<70\%}$ and $R_b \subseteq \mathcal{F}_{<70\%}$ be the sets of PUs to reforest respectively in Unia and Borendy, that is sets of PUs available for restoration. The sets $R_u, R_b, \mathcal{F}_{\geqslant 70\%}$, and $\mathcal{F}_{<70\%} \setminus (R_u \cup R_b)$ must form a partition of $\mathcal{S}$, and $R_u \cup R_b \cup$

$\mathcal{F}_{\geqslant 70\%}$ corresponds to the potential forest-cover resulting from reforestation. To each of these sets is associated a grid graph. For a given set, each PU in the set is a node and two nodes are connected if and only if the corresponding PUs are adjacent according to the four-connected neighbourhood definition in the regular square grid. We now introduce the following constraints:

Constraint H (CONNECTED)**.** Let $R \subseteq \mathcal{S}$ be a region, $\textsc{Connected}(R)$ holds if and only if the region R is connected according to its associated graph.

Constraint I (RESTORABLE)**.** Let $R \subseteq \mathcal{S}$ be a region, a an integer variable, and $p \in [0, 1]$. $\textsc{Restorable}(R, a, p)$ holds if and only if each PU in R can be restored to a forest-cover proportion of p by reforesting at least a ha. In any solution satisfying this constraint, the value of a thus corresponds to the minimum area to restore to reach a forest-cover proportion of p. Formally, let v_x^p be the minimum area to reforest to restore the PU x to p, then: $\textsc{Restorable}(R, a, p) \Leftrightarrow a = \sum\limits_{x \in R} v_x^p$.

Constraint J (RADIUS)**.** Let $R \subseteq \mathcal{S}$ be a region and ρ a real variable. $\textsc{Radius}(R, \rho)$ holds if and only if the radius of the smallest enclosing circle containing R equals ρ (in meters).

Given two regions $R_u \subseteq \mathcal{S}$ and $R_b \subseteq \mathcal{S}$, the budget, accessibility, and equitable allocation requirements are satisfied if and only if all the following constraints are satisfied:

$$R_u \subseteq \mathcal{U} \cap \mathcal{F}_{<70\%} \land R_b \subseteq \mathcal{B} \cap \mathcal{F}_{<70\%}; \tag{7.2}$$

$$\textsc{Connected}(R_u) \land \textsc{Connected}(R_b); \tag{7.3}$$

$$a_u \in 0.5 \cdot A_{max} \pm 10\% \land \textsc{Restorable}(R_u, a_u, 70\%); \tag{7.4}$$

$$a_b \in 0.5 \cdot A_{max} \pm 10\% \land \textsc{Restorable}(R_b, a_b, 70\%); \tag{7.5}$$

$$a_u + a_b \leqslant A_{max}; \tag{7.6}$$

$$a_{max} \in [0, +\infty] \land \textsc{Restorable}(R_u \cup R_b, a_{max}, 100\%); \tag{7.7}$$

$$a_{max} \geqslant A_{max}; \tag{7.8}$$

$$\rho_u \in [0, P_{max}] \land \textsc{Radius}(R_u, \rho_u); \tag{7.9}$$

$$\rho_b \in [0, P_{max}] \land \textsc{Radius}(R_b, \rho_b). \tag{7.10}$$

With A_{max} the total area to reforest (200 ha) and P_{max} the maximum radius of the smallest circle enclosing reforested areas (1500 m). Constraint (9.3) ensures that the reforested regions are located in accessible and degraded areas respectively in Unia and Borendy. Constraint (9.4) ensures that each reforested region is connected. Constraints (9.5) and (9.6) ensure that the budget is equitably allocated between Unia and Borendy, with a_u and a_b two integer variables representing the minimum areas to restore respectively in Unia and Borendy. Constraint (9.7) ensures that the minimum area to restore in Unia and Borendy together does not exceed A_{max}. Constraint (9.8) ensures that the integer variable a_{max} equals the total area that can be reforested in Unia and Borendy together. Constraint (9.9) ensures that the totality of the budget can be invested in the selected areas. Finally, Constraints (9.10) and (9.11) ensure that each selected region is compact.

7.2.3.2 *Constrained optimization of fragmentation indices*

From the base problem described in the previous section, we defined two optimization problems, respectively associated with the maximization of MESH and IIC. We computed the value of each index in the current landscape, then we found every optimal solution and retained the index optimal value, the improvement brought by the optimal value compared to the current one, the number of optimal solutions, and the solving times for reaching the optimal value and then enumerate all optimal solutions. In the following, we denote the set of patches of a region R by $P(R)$. These patches are directly derived from the raster representation of the landscape by extracting the connected components of the grid graph associated to the raster grid, as illustrated in Figure 7.3.

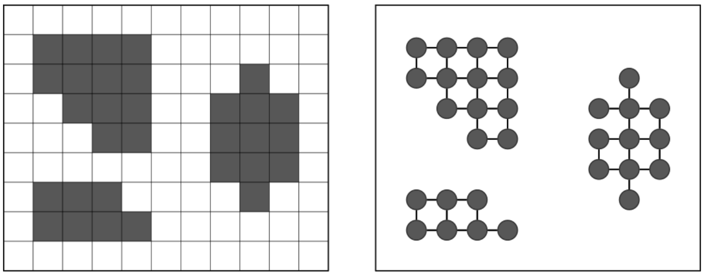


Figure 7.3: Raster representation of the landscape (left) and the associated grid graph (right). In this example, there are three connected components, thus three patches.

Maximization of MESH. MESH is a fragmentation index based on habitat patch sizes distribution within the landscape. It expresses an area unit and corresponds to the area of patches when the investigated region is divided into equally sized patches such that the probability that two randomly chosen points are in the same patch remains the same (Jaeger, 2000). For a region R, it is given by:

$$\mathrm{MESH}(R) = \frac{1}{A_L} \sum_{k \in P(R)} A_k^2. \tag{7.11}$$

With A_k the area of patch k, and A_L the total landscape area. The constrained optimization of MESH associated with our case study is given by:

$$\begin{aligned} &\underset{(R_u, R_b) \subseteq \mathcal{S}^2}{\text{maximize}} && \mathrm{MESH}(R_u \cup R_b \cup \mathcal{F}_{\geqslant 70\%}); \\ &\text{subject to:} && (9.3) \wedge (9.4) \wedge (9.5) \wedge (9.6) \wedge (9.7) \wedge (9.10) \wedge (9.11) \end{aligned} \tag{7.12}$$

Maximization of IIC. IIC is a graph-based inter-patch connectivity index introduced by Pascual-Hortal and Saura, 2006. It focuses on groups of patches (components) that are structurally or functionally connected and evaluates their sizes distribution along with the topological complexity of these components (i.e. the potential ability to move from one patch to another within a component). It ranges from 0 (no habitat in the landscape) to 1 (all the landscape is occupied by habitat). For a region R, it is given by:

$$\mathrm{IIC}(R) = \frac{1}{A_L^2} \sum_{k \in P(R)} \sum_{l \in P(R)} \frac{A_k \cdot A_l}{1 + d_{kl}} \tag{7.13}$$

Where A_k is the area of the patch k, A_L the total landscape area and d_{kl} the topological distance (i.e. shortest path length) between k and l in the landscape graph. Due to the lack of knowledge on species dispersal in the Côte Oubliée area, we used IIC as a structural connectivity index. To determine whether two forest patches are structurally connected, which is required to calculate IIC (see Pascual-Hortal and Saura, 2006), we used the smallest possible edge-to-edge distance threshold of at most one non-forest cell. This distance threshold can be represented by the two-wide-four-connected neighbourhood (Justeau-Allaire et al., 2019a). Two examples illustrating the construction of the landscape graph from a raster representation are provided in Figure 7.4 and Figure 7.5. The constrained optimization of IIC associated with our case study is given by:

$$\begin{aligned} \underset{(R_u, R_b) \subseteq \mathcal{S}^2}{\text{maximize}} \quad & \mathrm{IIC}(R_u \cup R_b \cup \mathcal{F}_{\geqslant 70\%}); \\ \text{subject to:} \quad & (9.3) \wedge (9.4) \wedge (9.5) \wedge (9.6) \wedge (9.7) \wedge (9.10) \wedge (9.11) \end{aligned} \tag{7.14}$$

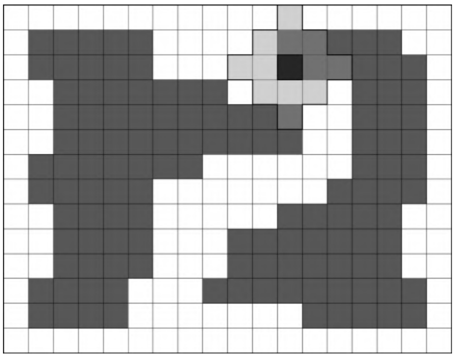


Figure 7.4: Illustration of the two-wide-four-connected neighbourhood distance threshold used to construct the landscape graph needed to compute IIC. The left patch intersects with the two-wide-four-connected neighbourhood of the black pixel located in the right patch. The patches are thus considered structurally connected.

7.2.4 *Solving method: The constraint-based systematic conservation planning framework*

To solve this problem, we used the constraint-based systematic conservation planning (SCP) framework briefly presented in the introduction (Justeau-Allaire et al., 2019a). As this framework relies on constraint programming (CP), we have provided a quick description of this technique's fundamental principles in Box 1. In this constraint-based SCP framework, any problem states as follows: given a tessellated geographical space $\mathcal{S}$, find a partitioning

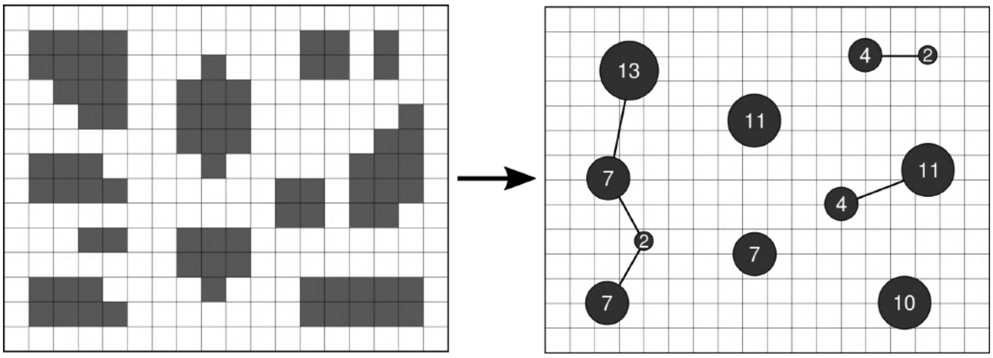


Figure 7.5: Construction of the forest landscape graph from a raster-based representation, using the two-wide-four-connected neighbourhood distance threshold.

of $\mathcal{S}$ into n regions $\{R_0, ..., R_{n-1}\}$ satisfying a set of constraints C, available from a constraint catalogue. The CP model associated with this formulation relies on three representations of the space: integer variables (one for each PU), set variables (one for each region), and graph variables (one for each region), and each user constraint applies to the most relevant space representation. This formulation allows the modelling of regions' expected properties through constraints. This framework was implemented upon the java open-source CP solver Choco (Prud'homme et al., 2017), and its source code is available on GitHub[1]. Most of the constraints needed by the case study were already available in the framework, we however extended it with the RADIUS constraint, implemented with a linear-time filtering algorithm based on the best-known algorithm for the smallest enclosing circle problem (Welzl, 1991), the MESH constraint, and the IIC constraint, implemented with a two-stage algorithm which first constructs the landscape graph from the raster representation and then computes all-pairs shortest paths by performing a breadth-first search from each node of the landscape graph. We ran all optimization problems described in the previous section on a Linux server (Intel Xeon E5-2620 CPU 2.40GHz × 12, 64GB RAM). The case study source code is available on GitHub[2] and we packaged a executable command-line jar to reproduce the single-region version of the problem (installation and usage instruction are available on the GitHub page).

---

[1] https://github.com/dimitri-justeau/choco-reserve

[2] https://github.com/dimitri-justeau/cote-oubliee-choco-reserve-code

Box 7.1: Constraint programming in a nutshell.

Constraint programming (CP) is a declarative paradigm for modelling and solving constraint satisfaction and constrained optimization problems. In this context declarative means that the modelling of a problem is decoupled from its solving process, which allows the primary focus to be on *what* must be solved rather than describing *how* to solve it. CP is a subfield of artificial intelligence which relies on automated reasoning, constraint propagation and search heuristics. As an exact approach, CP can provide constraint satisfaction and optimality guarantees, as well as enumerating every solution of a problem. In CP, the modeller represents a problem by declaring *variables* whose possible values belong to a specified finite *domain*, by stating *constraints* (mainly logical relations between variables), and eventually by defining an objective function to minimize or maximize. A solution to the problem is an instantiation of every variable such that every constraint is satisfied. As opposed to mixed-integer linear programming, constraints can be non-linear and variables of several types (e.g. integer, real, set, graph). A CP solver then handles the solving process relying on an automated reasoning method alternating a constraint propagation algorithm (deduction process on values within domains that does not lead to any solution) and a backtracking search algorithm. In a nutshell, more than satisfiability, each constraint embeds a filtering algorithm able to detect inconsistent values in variables domains. At each step of the backtracking search algorithm, the solver calls the constraint propagation algorithm that repeatedly applies these algorithms until a fix point is reached. When it is proven that a part of the search tree contains no solution, the solver rolls back to a previous state and explores another part of the search tree: this is backtracking. Note that most CP solvers are also able to handle Pareto multi-objective optimization. Interested readers can go further by reading the Handbook of Constraint Programming (Rossi et al., 2006).

7.3 RESULTS

We summarized the results of the constrained optimization of MESH and IIC in Table 7.1 and mapped optimal solutions in Figures 7.6 and 7.7. First, the solver found the optimal value for MESH in about 30 minutes and quickly enumerated all optimal solutions. Conversely, the solver took several hours to reach the optimal solution for IIC and about 20 minutes to enumerate all optimal solutions. Moreover, although several optimal solutions were found, for a given index they were all located in the same zone and reconnected the same patches.

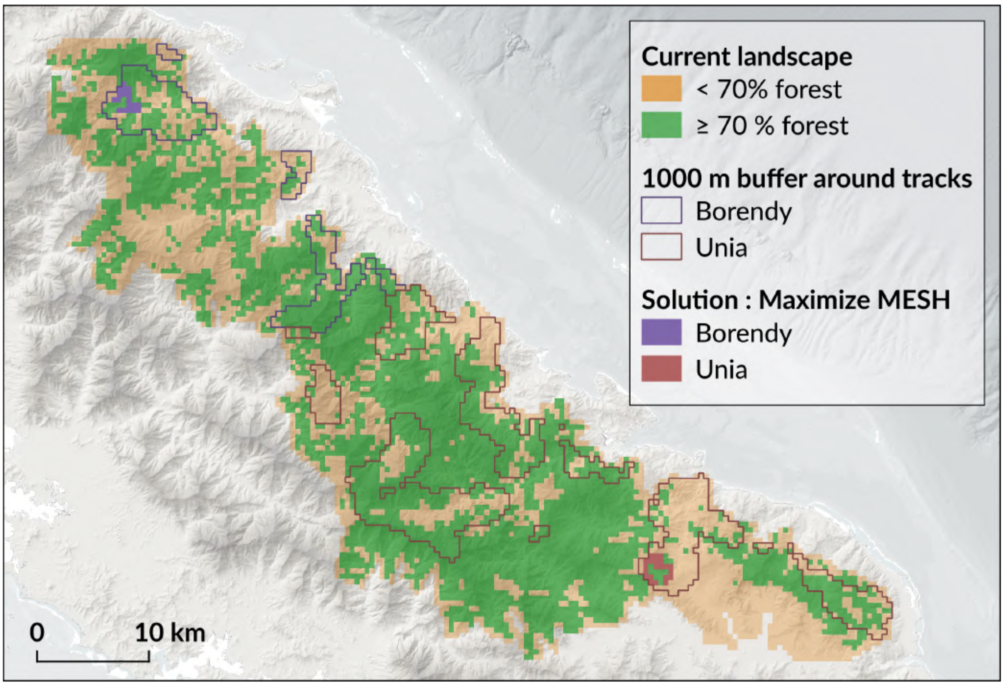


Figure 7.6: Mapping of a solution maximizing the effective mesh size (MESH).

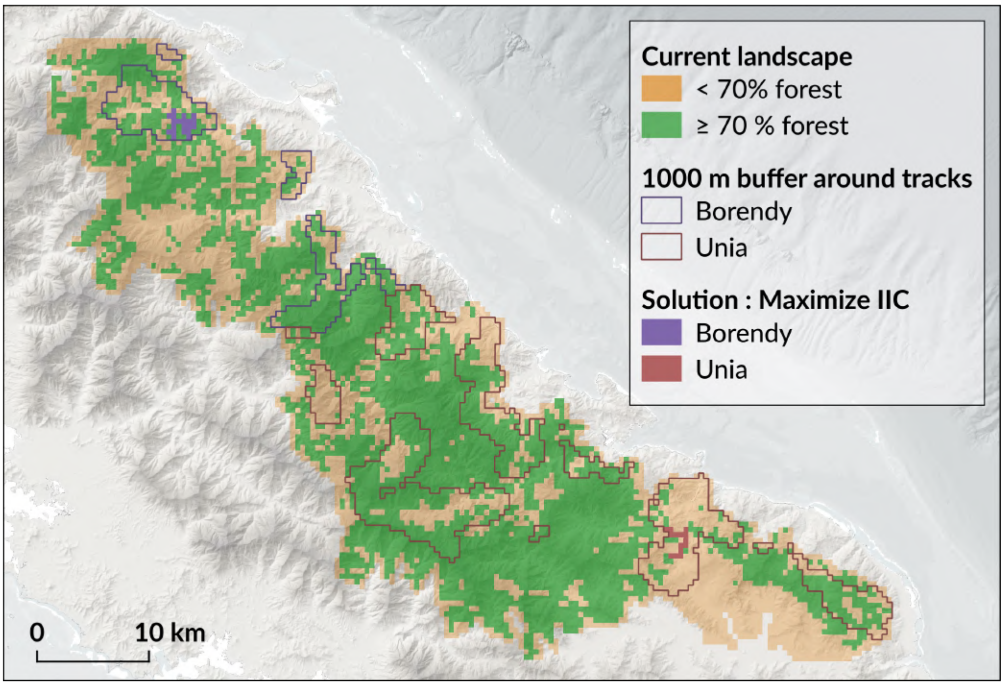


Figure 7.7: Mapping of a solution maximizing the integral index of connectivity (IIC).

Objective	maximize MESH	maximize IIC
Current value	24 542 ha	0.20691
Optimal value	25 502 ha	0.22986
Improvement	3.91%	11.09%
No. optimal solutions	7	3
Solving time (optimize)	14.7 min	5.8 h
Solving time (enumerate)	18 s	19.7 min

Table 7.1: Results characteristics: for each index, its value in the current landscape, its optimal value, the improvement after optimization, the number of optimal solutions and solving times. MESH: effective mesh size, IIC: integral index of connectivity.

7.4 DISCUSSION

7.4.1 *Contribution to decision support in the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" reforestation project*

Under budget, accessibility and equitable allocation constraints, we computed all optimal solutions for a fragmentation index (MESH) and an inter-patch connectivity index (IIC) within relatively short amounts of time. There was a considerable computing time difference between MESH and IIC, due to the combinatorial complexity involved by the construction of the patch-based landscape graph from a raster landscape representation. Optimal areas for MESH and IIC were not overlapping and offered two reforestation scenarios for managers. MESH did not assume any possible link between physically disconnected forest patches, thus highlighted areas favouring the physical connection of large patches together. In Borendy, it connected medium-sized patches into a large patch. In Unia, it merged two small patches with a large patch. On the other hand, IIC assumed possible links between physically disconnected but close patches, thus did not consider the medium-sized patch in Borendy as disconnected and favoured merging several small patches to reduce the topological complexity of the forest component. In Unia, it reconnected the southernmost forest component with the main forest component of the provincial park.

These results contributed to decision support by providing two scenarios that are optimal according to their respective index. In this regard, they provided a spatially-explicit and problem-focused baseline for discussions between stakeholders of the project, as well as specific areas presenting particular landscape-scale properties, thus potential candidates to prospection for local-scale assessments. Such results, along with the proposed methods, were well received and considered useful by the stakeholders of the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù". Most importantly, they were enthusiastic to see that the solver guarantees that every constraint will be satisfied by the solutions and that it will inform the user when no solution exists that satisfies all the constraints.

7.4.2 *On the use of landscape indices in systematic conservation planning*

These results illustrated the potential for integrating more complex and ecologically meaningful landscape indices into conservation planning to reduce fragmentation and improve connectivity. Fragmentation is known to have adverse effects on forest tree communities in New Caledonia (Ibanez et al., 2017) and there is strong evidence on the importance of structural connectivity for facilitating species dispersal, persistence, and gene flow between communities (Taylor et al., 1993). Optimizing such indices in systematic conservation planning (SCP) is thus useful to inform on the potential benefits of conservation actions on landscape fragmentation and connectivity. Being able to take into account the benefits of conservation projects over several indices is also an important step for providing holistic management recommendations. The main advantage of constrained optimization over prioritization and scenario analysis approaches is that the solutions are produced considering every possible combination of planning units satisfying user-defined constraints. This characteristics assures decisions makers that no feasible or better (according to an optimization objective) opportunity has been missed.

7.4.3 *Advantages of the constraint-based approach for systematic conservation planning*

Our constraint-based SCP framework demonstrated its ability to address and solve real-world SCP problems with satisfiability and optimality guarantees. By emphasizing a spatially-explicit and problem-focused approach, it presents several strengths. First, its expressiveness (i.e. the breadth and variety of problems that it can represent and solve) allows an accurate representation of the various constraints that stakeholders need to take into account for implementing conservation actions. Combined with a satisfiability guarantee, we can ensure that the proposed solutions will satisfy every managers' constraint and thus be socio-economically feasible, which is a requirement for policy-relevant conservation science (Game et al., 2015; Williams et al., 2020). Moreover, the flexibility of our approach makes it relevant to a wide range of conservation planning questions, as constraints and objectives can be seamlessly modified, added, or removed from the model without affecting the solving process. For instance, it can help to design optimal corridors, protected areas, fire-protected zones, or even provide insight for maintaining and restoring connectivity for migratory species. Note that although our use case was focused on forest cover, our constraint-based approach is also suited to include several biodiversity features and can handle multiple management zones. We believe that, besides being a useful methodological tool, such an approach can contribute to narrowing the "research-implementation gap" (Knight et al., 2008). With a modelling tool expressive enough to represent accurately conservation scientists' aims along with managers' constraints, it becomes possible to design conservation actions that are realistic for managers, as well as offering an integrative and evidence-based tool for scientists.

7.4.4 *Current limitations and perspectives for systematic conservation planning*

A lot of effort is still required to invest in development to provide a wide-audience software package, as our framework in its current state still requires knowledge of constraint programming (CP) to be used correctly. Moreover, as CP is an exact optimization approach, computation of optimal solutions can take time for large problems, and it is difficult to predict this time as it depends on the problem's structure (e.g. problem size, number and nature of the constraints). In its current implementation, we can however assert that exercises involving 50000 planning units (which is Marxan's limit in most cases; Ardron et al., 2008) would likely exceed the memory capacity of a standard desktop computer or not complete within a feasible amount of time. Another limitation directly relates to the regular square grid representation, which involves a trade-off between the spatial resolution and the sophistication of the model. In our case study, this spatial resolution limited the distance threshold needed to compute IIC to at least 480 m, which can be too large for some species. A promising perspective to overcome this limitation would consist of using an irregular grid representation to locally increase the spatial resolution without increasing the number of planning units.

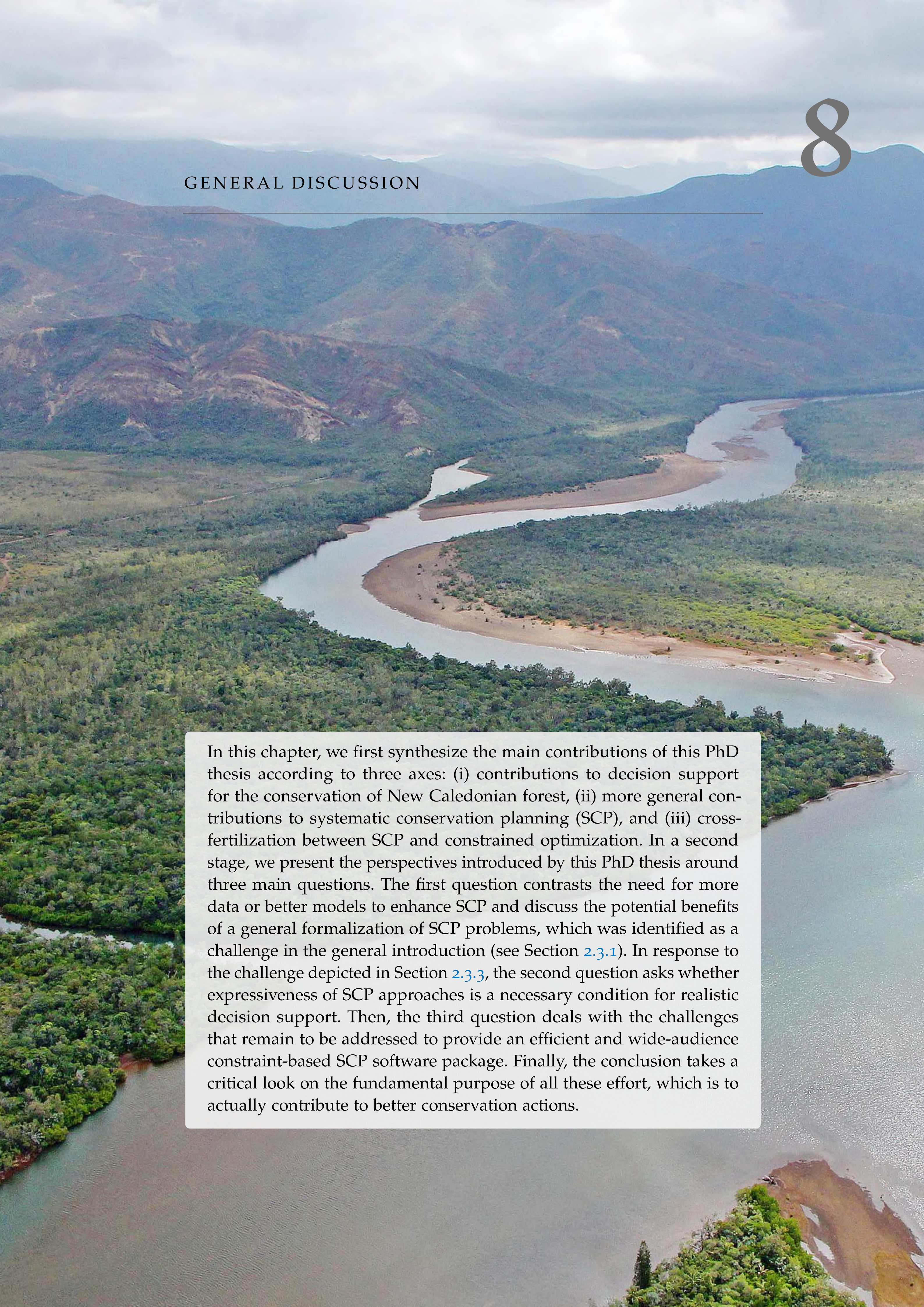
Nevertheless, we have shown that there is good potential for formulating and solving SCP problems using CP. There is a continued debate on the importance of optimality in SCP methods, which mainly contrasts local search approaches with MILP (Underhill, 1994; Pressey et al., 1996; Rodrigues and Gaston, 2002; Hanson et al., 2019a). However, optimality should not be the only consideration. We even argue that expressiveness is a prerequisite to optimality (Rodrigues et al., 2000; Moilanen, 2008). To conclude, recent years have seen substantial advances in artificial intelligence. We believe that, as illustrated by this study, such advances are providing new opportunities for formulating and solving conservation planning problems.

Part III

GENERAL DISCUSSION AND CONCLUSION

8

GENERAL DISCUSSION



In this chapter, we first synthesize the main contributions of this PhD thesis according to three axes: (i) contributions to decision support for the conservation of New Caledonian forest, (ii) more general contributions to systematic conservation planning (SCP), and (iii) cross-fertilization between SCP and constrained optimization. In a second stage, we present the perspectives introduced by this PhD thesis around three main questions. The first question contrasts the need for more data or better models to enhance SCP and discuss the potential benefits of a general formalization of SCP problems, which was identified as a challenge in the general introduction (see Section 2.3.1). In response to the challenge depicted in Section 2.3.3, the second question asks whether expressiveness of SCP approaches is a necessary condition for realistic decision support. Then, the third question deals with the challenges that remain to be addressed to provide an efficient and wide-audience constraint-based SCP software package. Finally, the conclusion takes a critical look on the fundamental purpose of all these effort, which is to actually contribute to better conservation actions.

8.1 SYNTHESIS OF THE CONTRIBUTIONS

8.1.1 *How did we contribute to decision support in the conservation of New Caledonian forests?*

As a collaboration between the Cirad through the AMAP lab in France and the IAC through the SolVeg team in New Caledonia, this PhD thesis was co-funded in response to recurring needs for decision support in forest conservation from New Caledonian environmental managers (see the third objective of this PhD thesis, Section 3.3.3). This need was mainly identified by Philippe Birnbaum, the director of this PhD thesis, through several partnership projects with New Caledonian managers (e.g. CORIFOR project, Birnbaum et al., 2016; COGEFOR project, Birnbaum et al., 2019). Indeed, as explained in Section 3.1.5, New Caledonia is a biodiversity hotspot recognized for its unique, rich, and heterogeneous flora. Sadly, we have also seen in Section 3.2 that the conservation of this exceptional heritage poses considerable challenges since New Caledonia has to face the threats of bushfires, invasive species, and is carrying the environmental weight of an economy almost entirely based on nickel extraction. Regarding this last aspect, the current situation is intricate as any arbitration must take into account ecological, economic, social, and cultural considerations in an unsettled political context. New Caledonia is, indeed, a territory undergoing a period of political transition ruled by the Nouméa Accord (Jospin, 1998), which outcome depends on the result of the self-determination referendum process which should achieve by 2022 at the latest. In this context where the stakes intertwine, it was shown that the current protected areas network, as well as previous conservation actions, are insufficient to efficiently protect New Caledonian forests (Jaffre et al., 1998; Ibanez et al., 2018). Despite these seemingly impossible challenges, there is a real willingness from New Caledonian environmental managers to improve the efficiency of their conservation policies. One of their most recurring concern is the ability to evaluate the potential benefits and trade-offs of conservation actions with an accurate and evidence-based decision support process. Although this issue neatly falls within the scope of systematic conservation planning (SCP), New Caledonian managers and conservationists did not adopt existing tools (see Section 2.2.4) as they did not provide their expected level of decision support. As introduced in Section 3.3.1, one of our objectives was to contribute to more expressive SCP approaches. Our initiative was based on the assumption that, with a model that can be precisely tailored to New Caledonian managers' questions, we would encourage them to rely on SCP while at the same time establishing a richer dialogue around this approach.

In Chapter 5, we have introduced the first generic constraint programming (CP) approach to SCP. Focused on the delineation of two regions (the protected area and the non-protected area), this approach allowed the combination of formal feature representation constraints (occurrence representation with or without backups and probabilistic representation, see Sections 2.2.1 and

5.3) along with formal spatial constraints (number and size of connected components, and total area of the protected area, see Sections 2.2.2 and 5.3). We evaluated this early generic CP formulation with a use case based on a fragmented area located in the south of New Caledonia, the CORIFOR's project area (see Figure 8.1; Birnbaum et al., 2016), which was subject to a study on the impacts of fragmentation on forest tree communities by Ibanez et al. (2017).

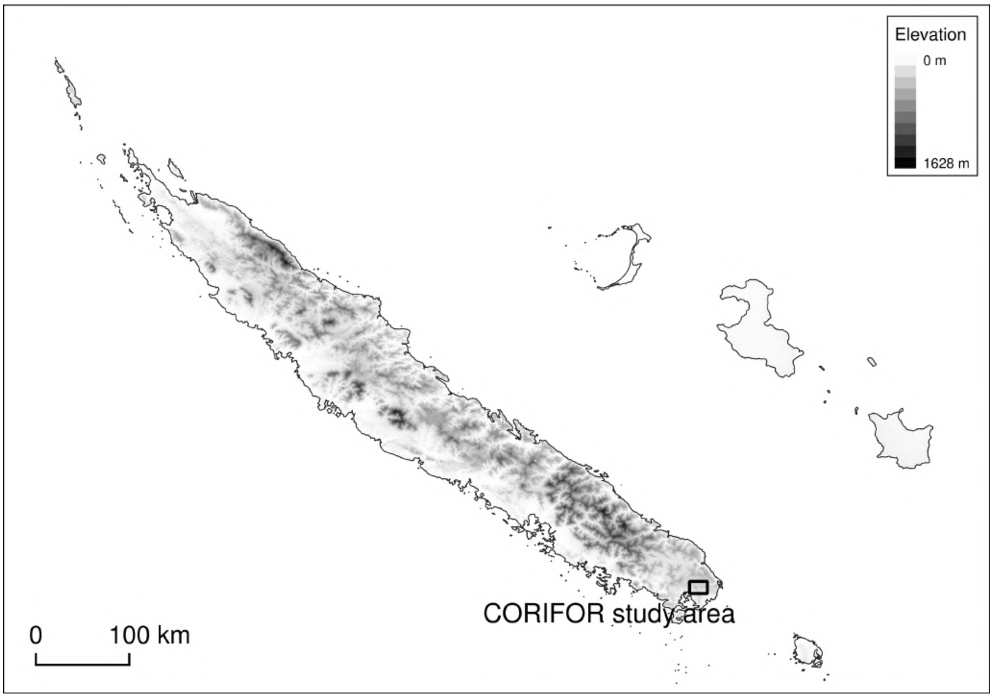
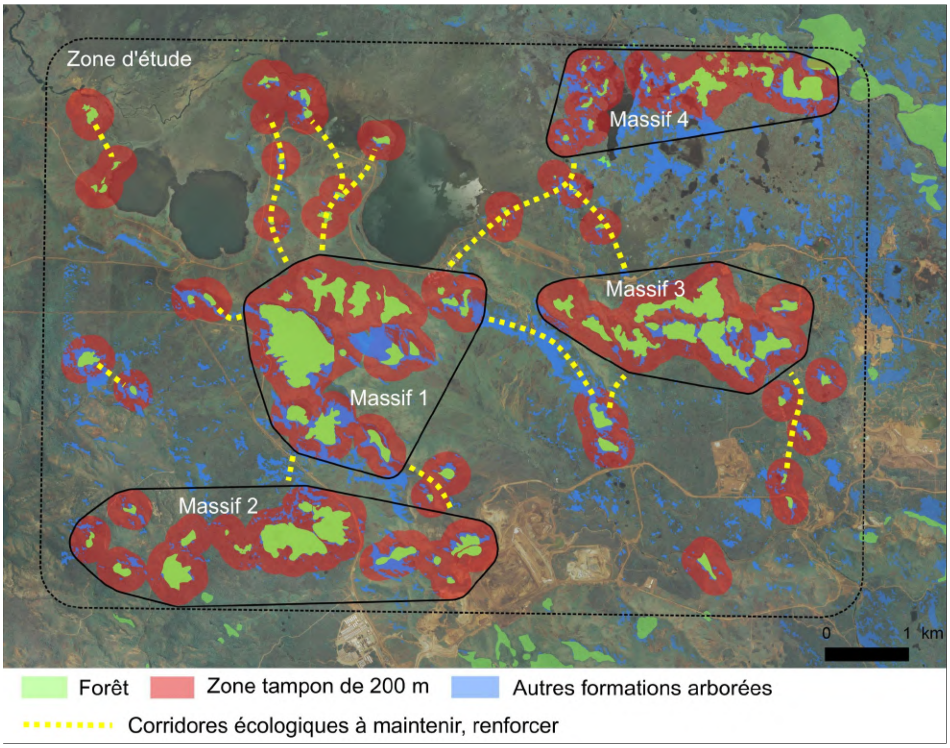


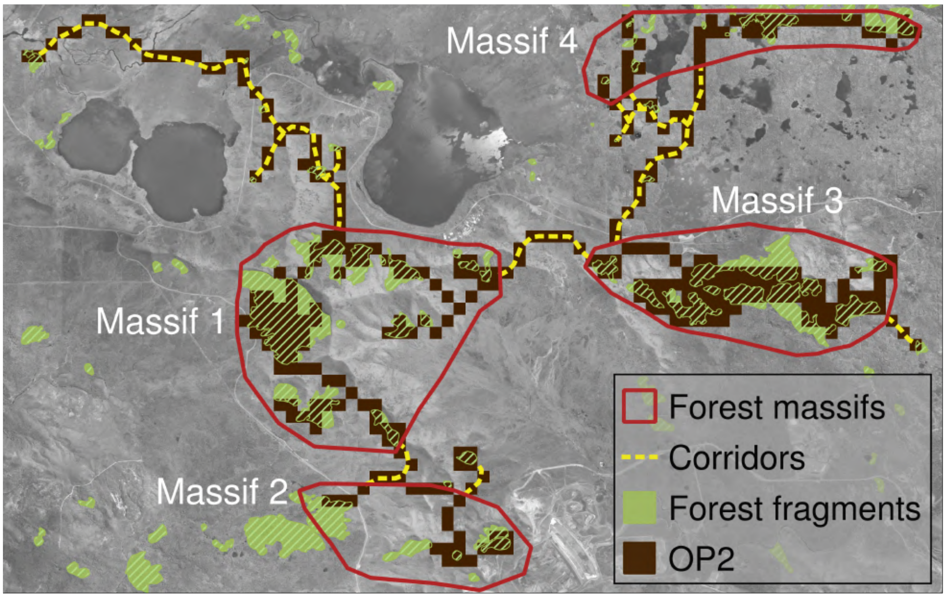
Figure 8.1: Location of the CORIFOR study area (Birnbaum et al., 2016; Ibanez et al., 2017). Digital elevation model: ©Geomatics and Remote Sensing Service - DTSI - Government of New Caledonia, CC BY-NC-SA 4.0.

Although this use case was fictive, it was elaborated with the help of the ecologists of the CORIFOR project, and conducted with real data including occurrence data for 223 species, species distribution models (SDMs) for 173 species, and forest cover data from expert digitization. Not only the results of this use case have shown that our CP model was able to address conservation planning questions with heterogeneous and real data in New Caledonia, but they also highlighted the complementarity of expert-based and automated SCP approaches. Indeed, with an expert-based approach, the authors of the CORIFOR projet's report (Birnbaum et al., 2016) highlighted four forest massifs and a set of corridors between these massifs and smaller forest patches. Our SCP approach identified the same massifs, but the inclusion of occurrence and SDM data in the analysis revealed slightly different results.

As we can see on Figure 8.2, only one part of Massif 2 was selected by our approach, as the rest of this massif was not necessary to ensure the complementarity of species representation from occurrence and SDM data. Moreover, we highlighted different corridors that take into account habitat suitability for



(a) Forest massifs and corridors identified by Birnbaum et al. (2016) with an expert-based spatial approach. Figure extracted from the original publication.



(b) Forest massifs and corridors identified by the SCP approach in Justeau-Allaire et al. (2018). OP2 is the solution obtained for the second optimization scenario (see Chapter 5).

Figure 8.2: Forest massifs and corridors identified in the CORIFOR study area by Birnbaum et al. (2016) with an expert-based spatial approach compared to forest massifs and corridors identified by the SCP approach introduced in Justeau-Allaire et al. (2018).

the considered tree species. However, the identification of these massifs and corridors needed a manual interpretation from the output of our model. Indeed, the connectivity constraint combined with the low density of occurrences for the tree species concerned by the binary feature representation constraint forced the solution into an elongated configuration that was not suitable as-is for management. Besides, our model was restricted to the delineation of only two regions, and constraints could only be expressed in the protected area.

In response to these limitations, we extended the model introduced in Chapter 5 into a more generic constrained partitioning model in Chapter 6. This new approach allowed the delineation of an arbitrary number of regions and introduced a buffer zone constraint which can be composed into more complex spatial constraints. Taking another look at the CORIFOR use case, we were able to bring additional insights compared to the initial results. Pursuant with Ibanez et al. (2017) results on negative edge-effects in the CORIFOR area, we included a protected buffer zone in our fictive conservation scenario. Mostly, we could compare the outcomes of an optimistic and a pessimistic scenario, to finally provide a trade-off scenario in which we delineated two buffer zones (inner and outer) with distinct protection levels.

Because these results related to a fictive use case, they did not inform managers on practical and planned conservation actions nor lead to a concrete implementation in the field. However, we used them to illustrate the potential benefits of using a constraint-based SCP approach to support conservation decisions in New Caledonia. In addition to the talks given at CP 2018[1] and IJCAI 2019[2] conferences, we organised two public seminars in New Caledonia (October 2018 and September 2019) in which we invited the environmental managers of the different provinces, local conservation scientists, and conservationists from NGOs based in New Caledonia. Their feedbacks were positive, and managers were enthusiastic to see that our approach can guarantee that every stated constraint will be satisfied by the solutions and that it will inform when no solution exists satisfying all the constraints.

Most importantly, these exchanges with New Caledonian managers opened an opportunity that we did not expect. By the end of 2018, environmental managers of the South Province of New Caledonia were working on the creation of the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" Provincial Park that would enable them to reach the Aichi target 11 (CBD, 2010). Although this Provincial Park was not officially created until April 2019, we started a collaboration with the South Province's Sustainable Development Department for the Territories (SDDT), in charge of this project, right after our first public seminary in New Caledonia. We aimed to use our constraint-based SCP approach to provide decision support in the establishment of the management plan for this Provincial Park. Our first study consisted in identifying potential areas within the Provincial Park suitable to be upgraded nature reserves or integral nature reserves (see Section 3.2 for more details on protected areas

---

[1] *International Conference on Principles and Practices of Constraint Programming* – `https://cp2018.a4cp.org`

[2] *International Joint Conference on Artificial Intelligence* – `https://www.ijcai19.org`

categories in the South Province of New Caledonia). We presented the results of this study in the Island Biology 2019 conference[3] (Justeau-Allaire et al., 2019b). However, due to the socio-economical constraints, the SDDT had to change its priorities and preferred to focus on a reforestation project for which they had a budget coming from the ecological compensation measures imposed on actors in the mining sector. In this respect, we focused on the identification of two suitable areas for reforestation, one in each customary district of the Côte Oubliée, to involve both communities in the project. We aimed to optimize the potential contribution of the reforested areas to the overall forest structural connectivity in the region using advanced fragmentation and connectivity indices. This study led to the article corresponding to Chapter 7, which was accepted in *Journal of Applied Ecology*. We cannot yet claim that this study will have an impact on implementation, as the official reporting of these results to the SDDT had to be postponed to 2021 because of the COVID crisis. What is nonetheless certain is that this collaboration established a close dialogue between managers, biologists, and computer scientists. Regarding the approach, managers of the SDDT were enthusiastic, and this first real-world conservation planning exercise has opened perspectives to improve the relevance of our decision support approach. Indeed, we are currently planning a postdoctoral project in collaboration with the environmental managers of the North Province of New Caledonia.

8.1.2 *How did we contribute to systematic conservation planning?*

More generally, two objectives of this PhD thesis were focused on: (i) contribute to more expressive systematic conservation planning (SCP) through the use of constraint programming (CP), and (ii) provide precise control over the spatial configuration of SCP's answers through the integration of formal spatial constraints, notably fragmentation and connectivity indices (see Sections 3.3.1 and 3.3.2). In this respect, articles corresponding to Chapters 5, 6, and 7 brought successive and gradual contributions with regard to both objectives.

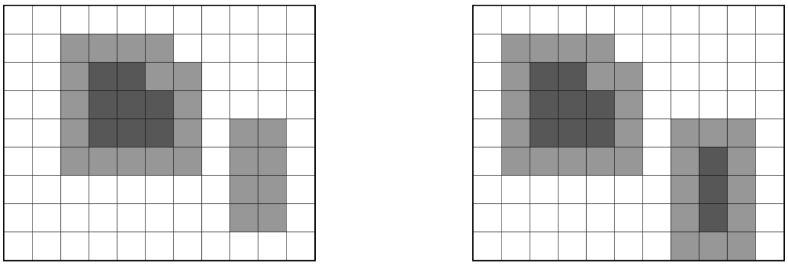
In Chapter 5, we introduced the first generic CP model to solve SCP problems. We have shown that this technique can formally address spatial constraints that were until now, and to the best of our knowledge, never addressed before, despite Williams et al. (2005) they were already identified as important contributions. Precisely, these spatial constraints are the control of the connected components of a delineated region as well as their sizes (see Sections 2.2.2 and 5.3). Besides, this first formulation benefited from the high levels of genericity, flexibility, and expressiveness that are provided by CP: constraints can be seamlessly added, removed or modified, and any variable of the model can be defined as an optimization objective. Finally, the solving process can be scaled up through the use of local search techniques.

In Chapter 6, we extended the previous model into an even more generic model. First, we turned the two-regions approach of Chapter 5 into a con-

---

[3] *International Conference on Island ecology, evolution, and conservation* – `https://ib2019.sciencesconf.org`

strained partitioning approach which allowed the delineation of an arbitrary number of regions. This feature was already provided by heuristic approaches such as Marxan with Zones (Watts et al., 2009) or Zonation (Moilanen et al., 2014) but to our knowledge, the only other formal approach offering this feature is the *Prioritizr* R package (Hanson et al., 2020). The main advantage of the constrained partitioning approach is that the policies associated with delineated regions are not defined *a priori*, but instead user-defined through the application of any constraint available from a catalogue. Also, this chapter introduced the first CP formulation of the buffer zone constraint (see Section 6.3.2). Although this constraint was already introduced in formal MILP models (e.g. Williams and ReVelle, 1996; Billionnet, 2013), our formulation was actually the first *complete* formulation of this constraint. Indeed, previous MILP formulations considered the buffer zone constraint in a local-fashion were the reciprocity between the buffered regions is not taken into account. The correct definition of a buffer zone is a region which directly separates two other regions. In this sense, the separated regions must not touch, but they both must be adjacent to the buffer zone such that the passage from one to the other necessarily occurs through the buffer zone. However, MILP formulations allow the existence of buffer zones without core area. Therefore, it would be more accurate to designate such formulations as *edge area*. In contrast, the formulation introduced in Section 6.3.2 does not assume that one buffered region is the protected core area and the other the non-protected area. Instead, it ensures that the region designated as the buffer zone will be a proper buffer reciprocally separating two given regions (see Figure 8.3).



(a) Edge area constraint with a protected core (dark gray), a non-protected area (white), and a protected edge area (light gray). Protected edge patches can exist without protected core.

(b) Buffer zone constraint. Light gray region is a proper buffer zone between the dark gray region and the white region. Buffer patches cannot exist without the two other regions.

Figure 8.3: Difference between the *edge area* constraint (a) and the *buffer zone* constraint (b). Current MILP formulations of the buffer zone constraint are actually edge area constraints.

As shown in Section 6.3.3, a substantial advantage of this complete formulation is that it can be easily extended to control the width and shape of the buffer zone and to express more complex spatial constraints by composition, such as nesting several protection levels, in line with the IUCN protected areas management categories guidelines (Dudley, 2008).

Finally, we have shown in Chapter 7 how the constrained partitioning approach introduced in Chapter 6 can be extended to optimize advanced landscape indices. Precisely, we implemented the MESH and the IIC constraints, which respectively allow controlling the effective mesh size (Jaeger, 2000) and the integral index of connectivity (Pascual-Hortal and Saura, 2006). The first is a measure of landscape fragmentation based on the cumulative patch sizes distribution and the second is a graph-based connectivity measure based on a binary connection model. To our knowledge, these indices had not been integrated into SCP optimization procedures. Indeed, most SCP approaches integrate fragmentation and connectivity through simple indices (e.g. boundary length modifier in Marxan; Ball et al., 2009). The main challenge related to the integration of indices such as MESH or IIC is to move beyond simple proxies of ecological processes and integrate metrics that directly derive from research in theoretical ecology (Saura and Rubio, 2010). Such indices are, however, hard to implement into optimization procedures because of their non-linearity and the curse of dimensionality. In Chapter 7, we have shown that CP is perfectly adapted to seamlessly optimize such indices along with other constraints through specialized filtering algorithms.

In conclusion, the three articles published in this PhD thesis demonstrated that many methodological advances can be achieved in SCP by drawing from advanced techniques in artificial intelligence. As shown in Section 2.2, SCP approaches until today mainly relied on heuristics, metaheuristics, and mixed-integer linear programming (MILP). The first two are usually straightforward to implement but produce solutions of unknown quality relative to optimality and satisfiability. On the other hand, MILP approaches can provide such guarantees but are limited in their expressiveness when it comes to integrating non-linear constraints and optimization objectives. Relying on an alternative paradigm (CP), we have shown that it was possible to catch up the state-of-the-art and integrate new advanced features to reach an unrivalled level of expressiveness and control over the spatial configuration of the solutions. Far be it from us to assert that CP dominates other approaches in all aspects (e.g. it is still not adapted to various large problems), but we hope that these contributions will stimulate the experimentation of new techniques in SCP.

8.1.3 *Did we contribute to cross-fertilization between systematic conservation planning and constrained optimization?*

Although the contributions brought by constraint programming (CP) and more generally constrained optimization with regards to decision support in New Caledonia and systematic conservation planning (SCP) are rather noticeable, it also seems legitimate to consider the question in the other direction. In this respect, the first thing to notice is how "exotic" the application context offered by SCP is to constrained optimization. Indeed, constrained optimization applications are mainly related to industrial, transportation, logistics, or military applications. As an example, in the two computer science and artificial intelligence international conferences in which we participated (CP 2018[4] and IJCAI 2019[5]), our articles were the only ones related to conservation biology (over 114 accepted articles for CP 2018 and 850 accepted articles for IJCAI 2019). In both conferences, feedbacks from other participants showed a strong interest in our application field. Especially, we noticed a strong enthusiasm for the ethical motivations of our work, which aims at contributing to better decisions in the conservation of nature. Accordingly, we believe that our first contribution from SCP to constrained optimization was to participate in the dissemination of an unusual but high-potential application field in constrained optimization scientific communities.

Furthermore, our work also showed that SCP can stimulate theoretical and technical developments in constrained optimization. For example, the work presented in Chapter 5 introduced bounds and filtering schemes in the general case for the constraints MIN_NCC and MAX_NCC, which both apply to a graph variable G and an integer variable ν and hold if the size of the smallest (respectively largest) connected component of G equals ν. These constraints were already defined in the global constraint catalogue (Beldiceanu et al., 2012), but bounds and filtering schemes were only devised for the specific `path_with_loops` graph class (Beldiceanu et al., 2006). We implemented these constraints in the Choco-graph extension of the Choco solver (Prud'homme et al., 2017; Fages et al., 2018) and suggested them to the developers of the solver, who integrated them in its next release. Moreover, in Chapter 6 we introduced theoretical results on the worst-case time complexity needed to enforce bound consistency (as defined for set variables by Walsh, 2003) on the buffer zone constraint. Although this constraint was devised specifically for our SCP application, it can be reused for other types of spatially-explicit problems applying on a tessellated geographical space, such as risk mitigation in humanitarian spatial planning. In Chapter 7 we introduced the effective mesh size and integral index of connectivity constraint (see Section 7.2.3.2) which, even if they are very specific to landscape ecology, have required challenging algorithmic design to develop efficient filtering algorithms. In contrast, although we devised the RADIUS constraint for Chapter 7's application, it was based on the SMALLESTENCLOSINGCIRCLE constraint that we priorly designed to be completely generic. Many application cases can potentially reuse this constraint. The SMALLESTENCLOSINGCIRCLE constraint applies on a set of boolean variables $\{b_0, ..., b_n\}$ to which are associated Cartesian coordinates $\{(x_0, y_0), ..., (x_n, y_n)\}$ and to a triplet of real variables (r, c_x, c_y). The boolean variables represent points in the Cartesian plane that can either be selected or not, and the real variables represent the smallest circle enclosing these points (r is its radius, c_x and c_y the Cartesian coordinates of its centre). SMALLESTENCLOSINGCIRCLE holds if the circle (r, c_x, c_y) is the smallest possible circle enclosing the set of points defined by $\{(x_i, y_i) \mid b_i = 1\}$. This constraint has two noticeable particularities. First, it is a hybrid constraint involving finite

---

[4] *International Conference on Principles and Practices of Constraint Programming* – `https://cp2018.a4cp.org`

[5] *International Joint Conference on Artificial Intelligence* – `https://www.ijcai19.org`

and continuous domain variables, which can be of great interest as some finite and continuous domain solvers can be combined to handle hybrid constrained optimization (e.g. Choco and Ibex, Fages et al., 2013). Secondly, the SMALLESTENCLOSINGCIRCLE filtering algorithm is an example of cross-fertilization between two fields related to computer science, as it is based on the best known worst-case time complexity algorithm (Welzl, 1991) which was introduced in the field of computational geometry. The formalization of this constraint in a short note is a perspective for future work.

Finally, as the implementation of the models devised during this PhD thesis relied on the Choco solver and its Choco-graph extension, we established a frequent and close communication with the developers of the solver. Our intensive use of graph and set variables along with our need for "exotic" constraints led us to identify several possible improvements for the solver. These improvements were mainly bug fixes and design enhancements that could increase the performance and modularity of the code. These technical perspectives can potentially benefit both our constraint-based SCP approach and the users of the solver.

8.2 PERSPECTIVES

8.2.1 *Do we need more data or better models to enhance systematic conservation planning?*

At the era of big data and machine learning, the emphasis is often on data. Artificial intelligence is even sometimes confounded with deep learning, whilst the latter is one form of machine learning, which is itself a subfield of artificial intelligence. Such data-driven approaches to problem-solving have for a long time been opposed to knowledge-driven ones (e.g. Dubois et al., 2000), among which we find constraint programming (CP) and mixed-integer linear programming (MILP) for instance. A current trend in artificial intelligence is to mix data-driven and knowledge-driven approaches to overcome their respective limitations, which in buzzwords could express as designing "learning and reasoning machines" (e.g. Lombardi et al., 2017; Adadi and Berrada, 2018; Ignatiev et al., 2018; Sukor et al., 2019). In our opinion, the choice of a method should not be a matter of trend, but rather the result of a rigorous definition and analysis of the problem to solve. In what refers to systematic conservation planning (SCP), the two main characteristics to outline are that (i) it involves complex combinatorial problems that (ii) can vary considerably in their structure from an instance to another. From these two characteristics, we can already argue that pure machine learning approaches are unlikely suitable for solving SCP problems. Indeed, although machine learning was successfully employed to solve combinatorial problems, this has been done relying on recurrent problem structures (Khalil et al., 2017). Moreover, realistic training datasets of sufficient size are, to date, not available and unlikely to be generated without knowledge-based approaches. Concerning this first conclusion, we can provide a first insight into the question posed by this section.

Before data, SCP requires knowledge-based models. Naturally, such models would be useless without the sufficient amount of data needed to apply them. However, data alone is not enough to deduce a model using pure data-driven approaches.

A question that remains is whether we need more models to improve the efficiency of SCP. In response to the latter and before designing more models, we argue in favour of the definition of a general modelling formalism able to encompass the vast majority of SCP problems that conservation research has addressed until now. The purpose of this formalism would not directly be to solve SCP problems, but instead, express them with a common formal modelling language. Although some studies attempted to provide such a formalism (e.g ReVelle et al., 2002; Williams et al., 2005; Billionnet, 2013), they were strongly correlated with the solving method (MILP) and restricted to some classical SCP problems. For example, it would not be possible to use such formalisms to address the case study depicted in Chapter 7. We, nonetheless, draw attention to the fact that devising a completely general SCP modelling formalism is a tedious task due to the high heterogeneity of problems structures. Still, we see many potential benefits motivating this effort. First, such a formalism would provide a standardized entry point to SCP problems, to which researchers and practitioners could refer to formulate questions and identify suitable methods to address them. Secondly, it would provide a systematic tool to compare the functionalities and performances of existing solving approaches, which would both benefit to end-users and stimulate research and development around a shared basis. For example, one inspiring perspective that such a formalism could offer would be the possibility to build open problem instances datasets and organize annual challenges to stimulate and federate SCP researchers in virtuous competition. Such challenges are common in artificial intelligence communities, and they have shown their ability to encourage research (e.g. MiniZinc Challenge – Stuckey et al., 2010; General Game Playing Competition – Genesereth et al., 2005; Angry Birds Artificial Intelligence Competition – Renz, 2015). Such challenges also gained popularity in biology and stimulated research in automatic plant species identification (PlantCLEF challenge; Goëau et al., 2016) or species distribution models (GeoLifeCLEF challenge; Botella et al., 2019), for example.

Although there remain many efforts to develop a generic SCP modelling formalism, we believe that the constrained partitioning approach introduced in Chapter 6 provided some elements that can be reused in this direction. For example, the set-based definitions of the geographical space and the regions already offer a high level of genericity. Indeed, it is suited to represent problems with an arbitrary number of areas to delineate and allows to associate land-use policies to groups of regions through set unions. Moreover, the graph representation that can be associated with the regions through abstract neighbourhood relationships has shown genericity and flexibility in both Chapters 6 and 7. Finally, the mathematical representation of constraints over sets and graphs objects offers an ideal level of abstraction as it allows to express several criteria independently from each other. However, although we can reuse such elements, the resulting formalism will need to be decorrelated from CP and any solving approach to provide the expected benefits.

Finally, despite our advocacy for the development of more generic SCP models, data remains an essential element not only for models' inputs but also for fundamental research, whose lack involves bottlenecks in the formulation of conservation planning questions. Indeed, data is a prerequisite to ecological knowledge, which is itself a prerequisite to formulating comprehensive conservation planning questions. This bottleneck was illustrated with the use case in the Côte Oubliée region, which was insufficiently studied to include species-specific considerations in our problem (see Chapter 7). We, nonetheless, believe that the availability of advanced SCP models has the potential to settle iterative decisions processes which can help to identify data and knowledge gaps. With such a configuration, decision support can become the cornerstone of a virtuous feedback loop between research and management.

8.2.2 *Are we expressive enough to provide realistic decision support?*

In Section 8.1.2 we have synthesized the contributions of this PhD thesis to systematic conservation planning (SCP), notably to expressiveness in the formulation and solving of problems. One important question remains to know whether we can improve this expressiveness, and how much it matters to provide realistic decision support. First, the more knowledge we can integrate into a decision support model, the closer the results are to real-world problems. Similarly, the more expressive this decision support model is, the more knowledge it can faithfully integrate. Then, we can already affirm that providing more expressiveness into SCP is a way to improve the effectiveness of decision support. However, even the most expressive decision support model remains useless if not adopted by stakeholders of real conservation projects. Accordingly, the relevance of decision support directly depends on the real-world benefits related to the use of a model. Thus, a simplistic model can have enormous benefits if it is used appropriately in a real-world context with stakeholders. The example of the first reserve selection algorithm (Kirkpatrick, 1983; see Box 1.3) is an excellent illustration of this fact, as despite being very simple and applied without the help of computers, it led to the creation of seven new protected areas in Tasmania. Another interesting example is the Marxan software (Ball et al., 2009), which is currently the most widely used SCP software. Despite not being the most sophisticated approach to date, its creators successfully disseminated the tool as the result of numerous communication efforts aimed at a non-academic audience (e.g. Marxan good practices handbook; Ardron et al., 2008). These two examples are good illustrations of the third objective of conservation biology, which is to develop practical, interdisciplinary, and integrated approaches to prevent species extinction, maintain genetic diversity among communities, preserve and restore biodiversity and ecosystems (see Section 1.2). Most importantly, as many conservation scientists, we believe that this connection between the academic and the non-academic world is a key leverage to reduce the research-implementation gap (e.g. Knight et al., 2008; Game et al., 2015; Ellison, 2016; Williams et al., 2020). In this respect, our position is that SCP should be ambitious in both improving the expressiveness of its models and committing to the dissemination of its advances in the non-academic world. Indeed, SCP has demonstrate its ability to improve decision support when researchers truly involve with stakeholders of real conservation projects. If this commitment comes along with intensive efforts to improve the expressiveness of models, we believe that decision support in conservation planning can be greatly improved. Finally, let us introduce three possible perspectives to improve the expressiveness of decision support in SCP: (i) integrate more indices resulting from research in ecology, in the continuity of the work presented in Chapter 7), (ii) devise automated methods to characterize solutions from problems' input data, which would be useful to help stakeholders discriminating solutions when a problem has several optimal solutions, and (iii) investigate the potential for user-friendly explanations in this particular application field (e.g. Jussien and Ouis, 2001), which would help understanding why some problems are not satisfiable and what it does imply for conservation.

8.2.3 *Is there room for a wide audience constraint-based systematic conservation planning software package?*

As said in Section 2.2.4, the availability of new methods in the form of software packages is a key factor for their success and dissemination. Besides, distribution as free and open-source software packages is an additional asset to facilitate access to a wide audience. These rules also applies to systematic conservation planning (SCP), as evidenced by the high number of citations reached by Marxan (Ball et al., 2009; 885 citations to date), Zonation (Moilanen et al., 2009b; 138 citations to date), or C-Plan (Pressey et al., 2009; 133 citation to date) free software. The three previous software are heuristics and metaheuristics approaches to SCP and are provided through graphical user interfaces (GUI). On the other hand, although many mixed-integer linear programming have been devised over time, SCP software packages relying on this technique were not available until recently (e.g. Hanson et al., 2017, 2019b, Raptr and Prioritizr R packages). However, even though the two previous examples are provided as free and open-source software packages, they are recommended to be used along with commercial solvers (e.g. Gurobi, IBM CPLEX) to be efficient. It should nonetheless be noted that, although state-of-the-art MILP solvers are mostly commercial, they usually offer free licences for academic users.

Before answering the question posed by this section, let us look at one significant aspect of SCP software packages: their availability through a GUI or an application programming interface (API), that is through a programming language (e.g. R, Python, Java), as it determines the definition of "wide audience". In the first case (GUI), the audience can be very large as few technical knowledge is required to use the software. Because it is straightforward to get familiar with such software, their advantages for decision support are quick to perceive for users. However, this simplicity of use also presents a high risk of misuses, as users can obtain results without going to the stage of understanding the issues and pitfalls to be avoided to use the tool properly and get reliable results. On the other hand, a software offering an API is more difficult to handle at first glance as it requires both skills with the concerned programming language and the API itself. However, once the steep part of the learning curve past, users are more likely to have a sufficient understanding of the approach to avoid misuses. Researchers in conservation biology usually have skills in using programming languages (at least R), thus using an API should not be a blocking element for this audience. However, users outside the academic world might not be used to handle programming languages. Thus, SCP software packages proposed through an API might be less accessible to this audience. Yet, we believe that this characteristic is not necessarily a weakness, as it can foster collaborations between stakeholders from and outside the academic world, and with different scientific and technical skills, as it was the case in Chapter 7's study.

Finally, to the question posed by this section we argue that there is room for an API-based constraint-based SCP software package. Indeed, as shown along with this PhD thesis, our constraint-based approach brought new perspectives to tackle SCP problems, and its availability to a wide audience is an important factor to ensure the sustainability of this work. Because flexibility and genericity mean a large variety of possible usages, we believe that a GUI software would necessarily imply a reduction of these possibilities. Our approach can be seen as a language based on constraints to model SCP problems, in this sense, an API is the best approach to make this language accessible without losing its power. Note that such an API can also be used to build more specialized GUI tools. Ideally, this API should be accessible through languages that are commonly used by conservation scientists (e.g. R, Python) to provide them with a toolbox to address SCP problems. At this stage, there remain many efforts to invest in development to provide such an API from our work. However, all the methods developed during this PhD thesis were implemented with this perspective in mind. We designed the *choco-reserve* framework (`https://github.com/dimitri-justeau/choco-reserve`) as a generic constraint-based SCP framework based on the Choco solver and its Choco-graph extension (Prud'homme et al., 2017; Fages et al., 2018). This framework is not yet ready for a wide-audience release, but can already be used by curious users as it is free and open-source. This framework is developed with the Java language, and relies on industry-standard tools for ensuring software quality (unit tests, automatic code coverage and quality review, and continuous integration). The *choco-reserve* framework was designed with an abstract and declarative object model: the user declares a tessellated geographical space (which can be regular or irregular), a set of regions to delineate within this space, applies constraints on or between these regions, and finally uses the Choco solver to solve the CP model that was automatically built from the SCP problem description. A future perspective is to enhance the design and performances of this framework, provide an comprehensive documentation with usage examples, and develop APIs to use it from R and Python, that are languages frequently used by conservation scientists.

8.3 CONCLUSION: ARE WE REALLY CONTRIBUTING TO BETTER CONSERVATION ACTIONS?

In other words, is the investment in research and development for systematic conservation planning (SCP) worthwhile, or would it not be a pretext to address challenging research questions? Clearly and as synthesised in Section 8.1.2, we have enough hindsight to affirm that SCP has the potential to support better decisions in what refers to conservation. However, let us not delude ourselves, SCP remains first and foremost a decision support tool. First, its results can be relevant only if the upstream fundamental research provides enough knowledge. Secondly, its influence on conservation actions depends on both the implication of researchers with stakeholders of conservation projects and the willingness of decision-makers to consider conservation on an equal footing with social and (mostly) economic aspects. In this respect, we mainly see in SCP an opportunity to synthesize results from research in biology along with managers' constraints. Incidentally, it is also an opportunity to involve researchers in computer science into problems that are at the same time technically challenging and ethically motivated. To conclude, we argue that SCP is one possibility among others to strengthen conservation in the face of powerful social machines such as industrial, financial, political, or commercial organizations (Buckley, 2015). Likely, there are many situations where its benefits cannot meet our expectations, and where other tools such as militancy or environmental law are much more adapted. However, the more tools we have at our disposal to support conservation, the more chances we have to improve conservation actions. In conclusion, this PhD thesis has contributed to improving the diversity and quality of the decision support that can offer SCP, and it opened research and development perspective in SCP as well as in constrained optimization. Thus, without going so far as to say that through this PhD thesis contributed to better conservation actions, it has at least proposed tools that can help that purpose. We sincerely hope that it will be the case.

Part IV

RÉSUMÉ EN FRANÇAIS

9

RÉSUMÉ EN FRANÇAIS

Ce chapitre est un résumé en français de la thèse de doctorat. Il est structuré de la même manière que la thèse avec un résumé de l'introduction générale, trois chapitres qui correspondent à des résumés en français des articles qui constituent la thèse (les deux premiers ont été publiés dans les actes de conférences francophones), et un résumé de la discussion générale.

9.1 INTRODUCTION GÉNÉRALE ET ÉTAT DE L'ART

9.1.1 *La Conservation de la Biologie et ses défis*

La crise mondiale de la biodiversité

Nous sommes face à une crise de la biodiversité sans précédent dans l'histoire de l'humanité. Si des chiffres précis sont difficiles à obtenir, le taux d'extinction des espèces à l'échelle mondiale est estimé à environ 1000 fois le taux naturel d'extinction. De plus, environ un million d'espèces sont aujourd'hui menacées d'extinction (Vos et al., 2015; Díaz et al., 2020). Les activités humaines sont la principale cause de cette crise, qu'on nomme la sixième extinction de masse (ou extinction de l'Holocène). Parmi celles-ci, le changement d'utilisation des terres (e.g. agriculture, urbanisation, exploitation des ressources) est la plus impactante. Le changement climatique, la pollution, et les espèces invasives introduites sont également des causes majeures du déclin de la biodiversité et des écosystèmes. Malheureusement, beaucoup de ces impacts sur la nature sont irremplaçables.

Pourtant, en plus de sa valeur intrinsèque, la nature fournit des ressources et des services qui sont vitaux pour les humains (Batavia and Nelson, 2017). Par exemple, les forêts protègent les sols de l'érosion, fournissent de l'eau propre, tamponnent les changements climatiques, sont une source de nourriture, de bois, ou de substances médicinales. Malgré cela, la déforestation est une des principales conséquences du changement d'utilisation des terres. En effet, entre 1990 et 2015, 129 millions d'hectares ont été perdus. Cette perte a principalement eu lieu dans des forêts tropicales, qui sont pourtant les plus riches et les plus productives (Keenan et al., 2015).

Origines et objectifs de la biologie de la conservation

En réponse aux dommages, qu'en tant qu'espèce, nous infligeons à la nature, le mouvement moderne de la conservation de la nature a été impulsé par différents naturalistes au milieu du XIXe siècle. Par exemple, la première association de conservation a été fondée en 1868 par des naturalistes anglais (l'association pour la protection des oiseaux marins). Quelques années plus tard, le premier parc national au monde, Yellowstone, fut créé aux États-Unis, devenant le premier exemple de conservation à l'échelle paysagère.

Si la science a toujours joué un rôle fondamental dans le mouvement de conservation moderne, la conservation n'est elle-même devenue une science qu'en 1978 avec la première conférence internationale sur la biologie de la conservation. Cette rencontre a abouti à la publication d'un livre qui a posé les fondations de cette nouvelle discipline scientifique : *Conservation Biology: An Evolutionary-Ecological Perspective* (Soulé and Wilcox, 1980). Aujourd'hui, la biologie de la conservation est une discipline scientifique bien établie, qui offre de très nombreuses perspectives. L'objectif de cette discipline est d'apporter des réponses éthiques et scientifiques pour faire face à la crise mondiale de la nature, protéger les espèces, préserver la biodiversité et les écosystèmes.

Pour atteindre cet objective, la biologie de la conservation s'appuie sur trois objectifs principaux :

- Décrire et comprendre la diversité des espèces, des communautés et des écosystèmes.
- Étudier et quantifier les impacts des activités humaines sur les espèces, les communautés et les écosystèmes.
- Développer des approches pratiques intégrées et interdisciplinaires pour prévenir l'extinction des espèces, maintenir la diversité génétiques dans les communautés, préserver et restaurer la biodiversité et les écosystèmes.

Tandis que les deux premiers de ces objectives correspondent à une recherche fondamentale, le troisième objectif définit également la Biologie de la Conservation comme une discipline normative et d'action. De ce fait, c'est une discipline scientifique appliquée et basée sur les valeurs éthiques de la conservation de la nature. Une autre spécificité de la biologie de la conservation est son interdisciplinarité. Si les sciences de l'environnement et l'écologie sont au cœur de la Biologie de la Conservation, celle-ci s'appuie également sur les sciences humaines et sociales, ou sur les mathématiques et les sciences de l'information. Il est important de noter que la Biologie de la Conservation contribue également au développement des disciplines sur lesquelles elle s'appuie, comme en témoigne la naissance de disciplines telles que la philosophie environnementale, la bio-informatique, ou l'éco-informatique.

Succès et défis de la conservation de la biologie

Après plus de quarante ans d'existence, la Biologie de la Conservation est une discipline reconnue et bien établie d'un point de vue académique. Cependant, est-ce cette discipline remplit ses objectifs ? Il est difficile de répondre clairement à cette question. Une chose est cependant sûre, la crise mondiale de la nature n'a pas été stoppée, comme en atteste le dernier rapport de la plateforme intergouvernementale scientifique et politique sur la biodiversité et les services écosystémiques (IPBES) et de nombreuses autres études (Cardinale et al., 2012; Haddad et al., 2015; Keenan et al., 2015; Woinarski et al., 2015; Strona et al., 2018; Díaz et al., 2020). Néanmoins, la Conservation de la Biologie a substantiellement influencé les principes et les pratiques de conservation et a de nombreux succès à raconter.

Malgré tous ces efforts, la conservation reste faible face au pouvoir des industries, de la finance, et d'organisations politiques et commerciales dont les intérêts sont souvent conflictuels avec ceux de la conservation. En tant que discipline scientifique, la Conservation de la Biologie ne peut pas stopper la crise de nature seule, mais peut cependant continuer à fournir des efforts pour la préservation de la nature, en accord avec son troisième objectif. Sur ce point, plusieurs biologistes de la conservation déplorent le fossé qui existe parfois entre la recherche et l'implémentation (Robinson, 2006; Knight et al., 2008;

Game et al., 2015; Wistbacka et al., 2018; Williams et al., 2020). Plusieurs pistes sont possibles pour réduire ce fossé, comme par exemple mettre en place un dialogue plus étroit entre les scientifiques et les gestionnaires (Prendergast et al., 1999), établir plus de collaboration avec les sciences sociales et politiques (Balmford and Cowling, 2006), s'impliquer directement dans l'implémentation en tant que scientifiques (Arlettaz et al., 2010), ou s'engager politiquement (Ellison, 2016).

Dans cette thèse de doctorat, nous nous intéressons à un aspect particulier de la Biologie de la Conservation, la Planification de la Conservation, qui vise à fournir une aide à la décision pour la planification de l'utilisation des sols, la gestion de la nature et des aires protégées, ou la restauration écologique. La planification de la conservation s'inscrit directement dans le troisième objectif de la Biologie de la Conservation en s'impliquant dans les processus décisionnels liés à la conservation. Depuis quelques années, la planification de la conservation s'appuie de plus en plus sur la modélisation et l'informatique. Cette approche moderne s'appelle la Planification Systématique de la Conservation.

9.1.2 *Planification Systématique de la Conservation*

Qu'est ce que la Planification Systématique de la Conservation ?

La conservation à l'échelle paysagère via les aires protégées est une pratique bien établie dans la gestion de la nature. Cependant, les aires protégées ont longtemps été identifiées de manière opportune et motivées par la protection d'espaces scéniques et réactionnels, présentant en général peu d'intérêts économiques. À la fin du XXe siècle, les scientifiques de la conservation ont commencer à plaider pour sélection et une délimitation plus systématique des aires protégées. L'objectif étant bien sûr de mettre l'accent sur leur efficacité dans leur rôle de protection de la biodiversité et des écosystèmes (Myers, 1988; Prendergast et al., 1993). En 2000, cette préoccupation a donné naissance à une nouvelle sous-discipline de la biologie de la conservation nommée Planification Systématique de la Conservation (PSC) (Margules and Pressey, 2000). Cette discipline ne se rétreint aujourd'hui plus à la délimitation des aires protégées et est également adaptée pour la délimitation de corridors écologiques, la planification de la restauration écologique, ou même la priorisation des projets de conservation.

Les premiers concepts de la PSC ont été inspirés par la théorie de la biogéographie insulaire : en considérant les aires protégées comme des îles entourées par un océan d'habitat altéré, un ensemble de principes géométriques ont été conçus pour la délimitation des aires protégées (MacArthur and Wilson, 1967; Diamond, 1975). Ces principes ont toutefois été rapidement critiqués, puisque les zones autour des aires protégées ne sont pas nécessairement inhospitalières, et parce qu'ils s'appuient sur une hypothèse d'écosystèmes en état d'équilibre qui est rarement satisfaite dans des paysages anthropisés et fragmentés (Margules et al., 1982). Par la suite, de nombreux chercheurs ont cherché à mettre l'accent sur la représentation des caractéristiques de la biodiversité (e.g. espèces représentées) plutôt que sur la configuration spatiale des aires protégées. Les premières approches étaient basées sur des procédures de notation multi-critères e.g Margules and Usher, 1981; Smith and Theberge, 1986; Usher, 1986; Smith and Theberge, 1987. Plusieurs auteurs ont rapidement démontré l'inefficacité de ces approches, elles s'appuient en effet sur un raisonnement local qui ne peut pas mettre en évidence la complémentarité entre les unités de planification dans la représentation des caractéristiques de la biodiversité (Kirkpatrick, 1983; Pressey and Nicholls, 1989). Ces même auteurs ont montré qu'en s'appuyant sur des approches itératives il est possible d'obtenir des résultats bien meilleurs. Après ces travaux, la complémentarité a été reconnue comme un des principes fondamentaux de la PSC (Vane-Wright et al., 1991; Pressey et al., 1993). De nombreuses méthodes ont par la suite été développées dans cette direction (e.g. Rebelo and Siegfried, 1992; Possingham et al., 1993; Underhill, 1994; Ball, 2000).

Aujourd'hui, il est largement admis que plutôt que de s'appuyer sur des principes géométriques universels, la configuration spatiale doit être considérée de manière contextuelle avec la représentation des caractéristiques de la biodiversité. La PSC est aujourd'hui une discipline bien établie qui compte de nombreuses production scientifiques méthodologiques et appliquées. Cette discipline a même attiré des chercheurs issus de l'optimisation mathématique et de l'informatique. En effet, les problèmes sous-jacents à la PSC posent de nombreux défis en terme de modélisation et de résolution. Une des originalités de ces problèmes est leur hétérogénéité et leur tendance à être composés de sous-problèmes difficiles qui doivent être résolus simultanément.

État de l'art

De nombreuses approches ont été développées pour modéliser et résoudre des problèmes de PSC. Si ces méthodes ont de nombreux points communs, elles diffèrent également dans les techniques qu'elles utilisent et dans les types de questions auxquelles elles peuvent répondre. De manière générale, on peut voir un problème de SCP de la manière suivante : étant donné un espace géographique $\mathcal{S}$, délimiter n régions $R_0, R_1, \ldots, R_{n-1}$ dans $\mathcal{S}$ de manière à satisfaire un ensemble de critères. En général, $\mathcal{S}$ est discrétisé en *unités de planifications* (ou sites, ou parcelles), et les critères correspondent à des contraintes socioéconomiques et à des objectifs de conservation. On peut distinguer deux types de critères : les critères de représentation de caractéristiques spatiales (biodiversité ou autre) et les critères de configuration spatiale. Dans la première catégorie, on retrouve les critères suivants :

- *Représentation d'un ensemble de caractéristiques spatiales*. Étant donné une région R et un ensemble de caractéristiques spatiales $\mathcal{F}$ (représentées par des données d'occurrence), chacune des caractéristiques spatiales de $\mathcal{F}$ doit être présente dans au moins une unité de planification de la région R. Ce critère peut également correspondre à un objectif d'optimisation (e.g. maximiser le nombre d'éléments de $\mathcal{F}$ représentés dans R).

- *Représentation de l'abondance d'un ensemble de caractéristiques spatiales*. Étant donné une région R et un ensemble de caractéristiques spatiales $\mathcal{F}$ (représentées par des données d'abondance), chacune des caractéristiques spatiales de $\mathcal{F}$ doit être représentée avec une abondance minimale définie par l'utilisateur. Ce critère peut également correspondre à un objectif d'optimisation (e.g. maximiser l'abondance minimale).

- *Représentation multiple d'un ensemble de caractéristiques spatiales*. Ce critère est similaire au premier critère, à la différence que chacune des caractéristiques spatiales de $\mathcal{F}$ doit être présente dans au moins k unités de planification distinctes. Ce critère peut également correspondre à un objectif d'optimisation (e.g. maximiser k).

- *Représentation probabiliste d'un ensemble de caractéristiques spatiales*. Étant donné une région R et un ensemble de caractéristiques spatiales $\mathcal{F}$ (représentées par des données probabilistes), chacune des caractéristiques spatiales de $\mathcal{F}$ doit être représentée dans R avec une probabilité minimale. Ce critère peut également correspondre à un objectif d'optimisation (e.g. maximiser la probabilité minimale).

- *Représentation de la diversité phylogénétique*. Certains auteurs ont également considéré la représentation d'une diversité phylogénétique minimale parmi les caractéristiques représentées dans R (e.g. Moulton et al., 2007; Billionnet, 2017). Ce critère peut également être utilisé comme objectif d'optimisation (e.g. maximiser la diversité phylogénétique). Ce critère s'applique uniquement lorsque $\mathcal{F}$ correspond à des entités taxonomiques.

Dans la seconde catégorie (configuration spatiale), on retrouve les caractéristiques suivantes :

- *Connectivité d'une région*. Ce critère consiste à assurer la connectivité au sein d'une région, selon un graphe qui caractérise la relation d'adjacence entre les unités de planification.

- *Nombre de composantes connexes d'une région*. Lorsque la connectivité stricte n'est pas souhaitée, on peut préférer contrôler le nombre de composantes connexes (sous-ensembles connectés) d'une région.

- *Périmètre d'une région*. Ce critère consiste à contrôler le périmètre d'une région. Il peut être utile pour approximer la connectivité, assurer la compacité, ou minimiser les effets de bords dans une région.

- *Taille d'une région*. La taille (ou surface) d'une région peut être importante à contrôler, notamment lorsqu'il s'agit de la minimiser ou de la maximiser.

- *Taille des composantes connexes d'une région*. Lorsqu'une région n'est pas connectée, il peut être souhaitable de contrôler la taille de ses composantes connexes, par exemple en leur imposant une taille minimale.

- *Distance entre deux régions*. La distance entre deux régions (ou entre les composantes connexes d'une région) est un critère utile pour faciliter la dispersion des espèces entre différentes aires protégées, ou pour préserver une aire protégée des effets de bords liées à une zone urbanisée par exemple.
- *Zone tampon entre deux régions*. La délimitation de zones tampon est une préoccupation fréquente qui permet part exemple de mitiger les effets de bords entre une aire protégée et une zone non protégée.
- *Forme d'une région*. Le contrôle de la forme d'une région peut permettre de favoriser la persistance des espèces au sein d'une aire protégée, ou bien faciliter sa gestion.

Différentes techniques ont été mises en place pour modéliser et résoudre des problèmes de PSC. Du fait de la grande variabilité des problèmes qui peuvent être adressées (e.g. structure, critères considérés, contexte d'application) et des défis posés par la résolution de ces problèmes, il est difficile d'affirmer qu'une approche domine les autres (à l'exception des approches de notation multicritères qui présentent de nombreuses limites). En revanche, chaque technique a ses avantages, ses faiblesses, et est adaptée pour résoudre certaines classes de problèmes. A ce jour, la majorité des approches s'appuient sur des heuristiques ad hoc, sur des méta-heuristiques et sur la programmation linéaire mixte en nombres entiers (PLNE).

- *Heuristiques ad hoc*. Algorithmes de recherche locale adaptés à des problèmes spécifiques. Ils permettent de trouver rapidement des solutions à des problèmes prédéfinis, sans fournir de garanties sur la satisfaction des contraintes ni sur la qualité des solutions pour les problèmes d'optimisation (relativement à la valeur optimale).
- *Méta-heuristiques*. Algorithmes d'optimisation haut-niveau basés sur une recherche locale. Ils offrent les mêmes caractéristiques que les heuristiques ad hoc tout en offrant un plus haut niveau de généricité et de flexibilité.
- *Programmation linéaire mixte en nombres entiers (PLNE)*. Méthode formelle d'optimisation mathématique basée sur des équations linéaires. Tout en offrant moins de garanties sur le temps d'exécution, cette approche permet néanmoins d'obtenir des garanties sur la satisfaction des contraintes et sur la qualité des solutions.

Un facteur important pour la diffusion de la PSC est la disponibilité des méthodes sous la forme de logiciels. En effet, si de nombreuses méthodes ont été développées et décrites dans la littérature scientifique, toutes n'ont pas fait l'objet d'une valorisation sous la forme d'un logiciel. Nous avons donc compilé une liste (qui ne se veut pas exhaustive) des logiciels de PSC disponibles à ce jour.

- *Marxan (Ball et al., 2009)*. Marxan est un logiciel de PSC basé sur une méta-heuristique qui s'appelle le recuit simulé. Marxan met l'accent sur la représentation des caractéristiques spatiales et permet également un contrôle sur la configuration spatiale des solutions via une pénalité sur le périmètre des régions délimitées. Marxan est un logiciel gratuit qui offre une interface graphique et plusieurs extensions ont été proposées (e.g. Marxan with zones qui permet de délimiter plusieurs régions, correspondant à différentes actions de conservation, Watts et al., 2009).

- *Zonation (Moilanen et al., 2009b)*. Zonation est un logiciel de priorisation basé sur une heuristique ad hoc. À la différence de la plupart des outils de PSC, Zonation n'est pas basé sur des objectifs de représentation des caractéristiques spatiale à atteindre par les régions délimitées (il existe cependant une option pour tendre vers cette approche). À la place, Zonation produit une classification hiérarchique de l'ensemble des unités de planification de l'espace géographique, de la plus à la moins importante pour la conservation. Cette classification est basée sur les caractéristiques spatiales définies par l'utilisateur, mais peut également prendre en compte la configuration spatiale via une des huit méthodes d'agrégation spatiale prédéfinies. Zonation est un logiciel gratuit qui propose une interface graphique.

- *C-Plan (Pressey et al., 2005; Pressey et al., 2009)*. C-Plan est un outil intégré et interactif d'aider à la décision pour la PSC qui peut être utilisé pendant les négociations entre les différents acteurs concernés par un projet. L'outil s'intègre avec un système d'information géographique (ESRI ArcView 3) et s'appuie sur un estimateur statistique pour calculer l'importance des unités de planification, ou sur Marxan.

- *ConsNet (Ciarleglio et al., 2009, 2010)*. ConsNet est un outil de PSC basé sur une méta-heuristique qui s'appelle la recherche tabou. ConsNet propose de résoudre un problème similaire à celui que résout Marxan tout en offrant plus d'options pour contrôler la configuration spatiale des solutions. ConsNet est un logiciel gratuit qui propose une interface graphique.

- *Conefor Sensinode (Saura and Torné, 2009)*. Sans être exactement un outil de PSC, Conefor Sensinode peut toutefois être utile pour la planification des actions de conservation. L'objectif de cet outil est de quantifier le niveau de connectivité inter-patch d'un paysage à travers un ensemble d'indices basés sur les graphes (objet mathématique composé de nœuds et d'arêtes, représentant un réseau ou des relations entre différents objets). Cet outil permet également de quantifier l'importance relative de chacun des patchs d'habitats dans la connectivité globale du paysage. Conefor Sensinode est un logiciel gratuit qui propose une interface graphique sur Windows et une interface en ligne de commande sur Linux, Mac et Windows.

- *Prioritizr R package (Hanson et al., 2020)*. Prioritizr est un package R qui permet de modéliser et résoudre un large panel de problèmes de PSC avec la PLNE. Cet outil permet de traiter des problèmes similaires à ceux que traitent Marxan et Zonation et il permet la délimitation de plusieurs régions (à la manière de Marxan with zones). Prioritizr propose plusieurs critères de représentation de caractéristiques spatiales (e.g. représentation d'un ensemble de caractéristiques, diversité phylogénétique) et plusieurs critères spatiaux qui peuvent être appliqués de manière stricte ou en tant que pénalité dans une fonction objectif.

- *LQGraph (Fuller and Sarkar, 2006)*. LQGraph est un outil qui permet d'identifier des corridors écologiques entre des aires protégées existantes. L'outil se base sur des scores de qualité de l'habitat pour délimiter des corridors optimaux. LQGraph peut également être utilisé pour identifier des unités de planification qui peuvent être utilisées pour isoler des aires protégées en cas de propagation de pathogènes ou d'invasion d'espèces exotiques.

- *Linkage Mapper (McRae et al., 2012)*. Dans le même esprit que LQGraph, Linkage Mapper permet d'identifier des corridors optimaux entre des patchs d'habitats. L'outil est basé sur un modèle de résistance à la dispersion à partir duquel il identifie des chemins de moindre coût sur une grille raster.

Tendances et défis de la planification systématique de la conservation

Depuis son apparition, la planification systématique de la conservation (PSC) a produit de nombreuses avancées méthodologiques et pratiques, et de nombreux outils qui sont utilisés par les biologistes de la conservation et les gestionnaires. Cependant, il reste de nombreuses voies à explorer pour améliorer l'efficacité et la pertinence de la PSC dans son rôle d'aide à la décision. Sans chercher à être exhaustifs, nous avons synthétisé la plupart des tendances et défis actuels pour la recherche en PSC. Ces derniers partagent deux objectifs principaux : (i) rapprocher autant que possible la PSC des processus écologiques et de la réalité du terrain, (ii) promouvoir et faciliter la diffusion de la PSC en biologie de la conservation et auprès des gestionnaires.

Premièrement, la diversité des problèmes de SCP reflète directement celle des contextes et des problématiques de conservation. Si cette particularité fait de la PSC une discipline particulièrement riche et stimulante, elle pose également un défi majeur : comment proposer un formalisme général qui englobe toute la diversité des problèmes de PSC ? S'il y a eu quelques travaux dans ce sens (e.g. Pressey et al., 1993; Margules and Pressey, 2000; ReVelle et al., 2002; Williams et al., 2005; Sarkar et al., 2006; Kukkala and Moilanen, 2013), il reste de nombreux progrès à faire pour structurer la diversité inhérente à la PSC autour d'un formalisme générique et unifiant.

Deuxièmement, beaucoup de progrès ont été faits pour contrôler la configuration spatiale des solutions produites par la PSC. En effet, la plupart des outils actuels proposent différentes options pour contrôler cet aspect en plus des critères de représentation de caractéristiques spatiales. Cependant, de nombreux critères spatiaux ont été définis dans la littérature sans pour autant être intégrés dans des approches pratiques. Par exemple, à notre connaissance aucune approche ne permet de contrôler exactement le nombre et la taille des composantes connexes d'une région délimitée. D'autre part, de nombreux indices de fragmentation et de connectivité ont été développés en écologie du paysage (e.g. McGarigal, 2014; Pascual-Hortal and Saura, 2006; Saura and Pascual-Hortal, 2007), cependant, ces indices n'ont pas été intégrés dans les procédures d'optimisation en PSC, qui s'appuient en général sur des métriques simples (e.g. pénalité sur le périmètre d'une région).

Finalement, si la notion d'optimalité a été sujette à de nombreux débats dans la littérature (Underhill, 1994; Pressey et al., 1996; Rodrigues and Gaston, 2002; Hanson et al., 2019a), la notion d'expressivité a été beaucoup moins discutée. En modélisation, l'expressivité fait référence à l'étendue et la variété des problèmes qui peuvent être représentés et résolus par une approche donnée. Le niveau d'expressivité d'une approche informe donc sur sa capacité à s'adapter aux spécificités d'un problème donné. Du fait de la grande diversité des problèmes qui peuvent se présenter en PSC, l'expressivité se présente donc comme un facteur clef pour la pertinence de l'aide à la décision. Nous pensons même que cet aspect est plus important encore que la notion d'optimalité.

9.1.3 *Zone d'étude et objectifs de recherche*

La Nouvelle-Calédonie, un point chaud de la biodiversité dans le Pacifique Sud

La Nouvelle-Calédonie est un archipel tropical situé dans le Pacifique Sud-ouest à 130 km au nord du Tropique du Capricorne, à environ 1 400 km à l'est de l'Australie et à environ 2 000 km au nord de la Nouvelle-Zélande. L'archipel couvre une surface terrestre de 18 575 km, se distingue par le plus grand lagon du monde (environ 24 000 km). L'île principale, la Grande Terre, s'étend du nord-ouest au sud-est sur environ 400 km de longueur et entre 50 et 70 km de largeur. La Grande Terre est traversée par une chaîne de montagnes qui forme une barrière naturelle entre les plaines de la côte ouest et les paysages escarpés de la côte est (voir Figure 9.1). Environ un tiers de la surface de la Grande Terre est couverte par un substrat ultramafique riche en métaux lourds, les deux tiers restants sont principalement recouverts par un substrat volcano-sédimentaire. Les îles loyautés situées à l'est de la Grande Terre, Lifou, Maré et Ouvéa, présentent quant à elles un relief beaucoup plus plat et sont principalement couverte par un substrat calcaire. L'île des Pins et les îles Belep, respectivement situées au sud et au nord de la Grande Terre, sont les deux autres îles principales de l'archipel. La Nouvelle-Calédonie jouit d'un climat tropical, avec une saison chaude et humide de novembre à avril (températures maximales moyennes : entre 28 °C et 32 °C) et une saison fraîche de mai à septembre (températures maximales moyennes : entre 24 °C et 29 °C). La saison chaude est également appelée saison cyclonique en raison de la fréquence élevée de dépressions tropicales et cyclone pendant cette période. L'archipel est sujet à l'alizé qui dans cette région souffle de sud-est. Du fait de la topographie de la Grande Terre, ce régime de vent conduit à une forte variation des précipitations annuelles moyennes : de 800 mm par an sur les plaines de la côte ouest jusqu'à 4 500 mm sur la côte est.

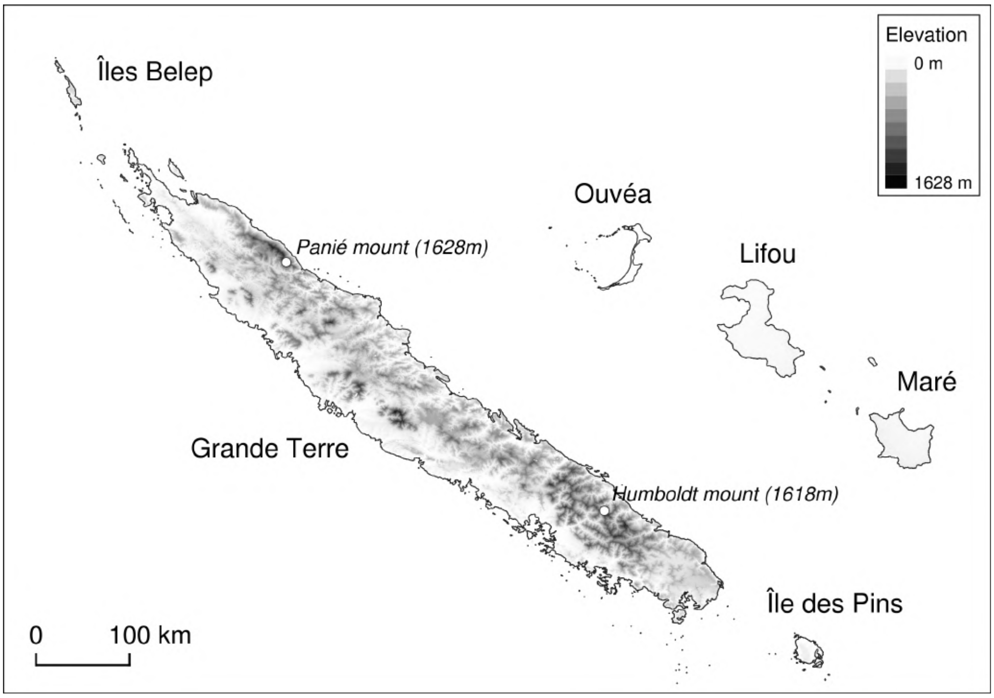


Figure 9.1: Nouvelle-Calédonie: îles, altitude et points culminants. Modèle numérique de terrain: ©Service de la géomatique et de la télédetection - DTSI - Gouvernement de Nouvelle-Calédonie, CC BY-NC-SA 4.0.

En tant que plus petit point chaud de la biodiversité au monde (Myers, 1988), la Nouvelle-Calédonie abrite des écosystèmes marins et terrestres extrêmement riches. Par exemple, la flore terrestre de Nouvelle-Calédonie est reconnue pour sa richesse et son caractère unique au monde. En effet, la flore néo-calédonienne se distingue par : (i) un des taux d'endémisme les plus élevés au monde, environ 76% parmi plus de 3 400 espèces de plantes vasculaires (Myers et al., 2000; Morat et al., 2012), (ii) par une diversité bêta élevée (Ibanez et al., 2014; Isnard et al., 2016) et (iii) la présence d'espèces reliques (Grandcolas et al., 2008; Pillon, 2012). Les angiospermes représentent environ 91% des espèces de plantes en Nouvelle-Calédonie, même si les gymnospermes y sont également très diversifiés avec 46 espèces endémiques (Jaffré et al., 1994; Morat et al., 2012). La diversité de la flore néo-calédonienne se reflète également à des niveaux taxonomiques plus hauts, l'archipel recense en effet trois familles endémiques et entre 62 et 91 genres endémiques (Pillon et al., 2017). Plusieurs types de formations végétales sont présentes en Nouvelle-Calédonie: (i) forêts humides climaciques, (ii) forêts sclérophylles (ou sèches) climaciques, (iii) maquis de basse à haute altitude, comprenant à la fois des formations climaciques et secondaires, (iv) formations hélophytique (mangroves et formations littorales) et (v) savanes et fourrés, toutes deux secondaires (Jaffré et al., 1994; Jaffre et al., 1998). Les forêts humides constituent en Nouvelle-Calédonie le type de formation végétale le plus riche avec plus de 2 000 espèces natives de plantes vasculaires et 82,4% d'endémisme sur une surface d'environ 4 000 km^2, dont un tiers sur substrat ultramafique (Jaffré et al., 1994; Birnbaum et al., 2015). Avec la surface la plus faible (environ 100 km2), les maquis de haute altitude présentent le plus haut taux d'endémisme (91%; Jaffré et al., 1994). Finalement, s'il a été montré que la plupart des espèces d'arbres sont ubiquistes en ce qui concerne le substrat, la pluviométrie et l'altitude (Birnbaum et al., 2015), de nombreuses espèces sont inféodées à un type de substrat et certaines d'entre elles sont même micro-endémiques (i.e. restreintes à un territoire très petit) (Wulff et al., 2013; Ibanez et al., 2014). Toutes ces spécificités font de la Nouvelle-Calédonie une mosaïque de formations végétale présentant des compositions floristiques uniques et diversifiées, ce qui a des implications importantes pour la conservation des forêts et de la flore en Nouvelle-Calédonie.

Contexte: conservation des forêts de Nouvelle-Calédonie

En Nouvelle-Calédonie, les activités humaines ont des impacts négatifs et importants sur les forêts. Actuellement, les menaces principales sont l'extraction de nickel, les feux et les espèces invasives introduites. Comme nous l'avons vu précédemment, les substrats ultramafiques couvrent environ 30% du territoire néo-calédonien, contre environ 3% à l'échelle mondiale. Cette particularité fait de l'extraction du nickel un des piliers de l'économie néo-calédonienne, au prix d'une importante perte et dégradation des forêts. Par ailleurs, les feux (pour 99% d'origine humaine) font chaque année des ravages en Nouvelle-Calédonie, avec environ 27 000 ha de végétation qui part en fumée chaque année. Finalement, si en Nouvelle-Calédonie les plantes invasives sont principalement restreintes aux zones dégradées, la végétation native est fortement impactée par les espèces animales herbivores invasives (en particulier, le cerf rusa).

En Nouvelle-Calédonie, la gestion de la conservation est une responsabilité provinciale, à l'exception du Parc Naturel de la mer de Corail, géré par le gouvernement de Nouvelle-Calédonie. Il y a trois provinces en Nouvelle-Calédonie, la Province Sud, la Province Nord et la Province des îles Loyautés. Chaque province dispose de son propre code de l'environnement et de ses propres catégories d'aires protégées. En Province Nord il y en six, inspirée des catégories d'aire protégées de l'UICN (Dudley, 2008), de la plus à la moins protégée : réserve naturelle intégrale (catégorie UICN Ia), réserve de nature sauvage (Ib), parc provincial (II), réserve naturelle (IV), aire de protection et de valorisation du patrimoine naturel et culturel (V) et aire de gestion durable des ressources (VI). En Province Sud, il existe quatre catégories d'aire protégées, de la plus à la moins protégée : réserve naturelle intégrale, réserve naturelle, aire de gestion durable des ressources et parc provincial. Finalement, la Province des îles Loyautés, principalement couverte par des aires coutumières, n'a mis en place aucune aire protégée. En revanche, le code de l'environnement de la Province des îles Loyautés promeut une politique plus proche des modes de gestion traditionnels coutumiers et affirme que "l'environnement naturel est indissociable des pratiques culturelles et des règles coutumières localement applicables. [...] La Province des îles Loyauté prend en compte l'existence de modes de gestion coutumière de l'environnement et intègre ces modes de gestion dans la réglementation, dans le respect du principe de subsidiarité." (Article 110-1; Citré et al., 2019). Si cette gestion provinciale a l'avantage de promouvoir la diversité et les spécificités locales, elle peut aussi être source de confusion. Par exemple, en Province Sud, un parc provincial correspond à un niveau de protection bien plus faible qu'en province nord.

Il y a environ vingt ans, Jaffre et al. (1998) ont alerté sur les insuffisances du réseau d'aires protégées en Nouvelle-Calédonie pour protéger efficacement la flore et les forêts. Ils ont notamment montré que 83% des espèces de plantes menacées n'étaient présentes dans aucune aire protégée et que seulement 54% des aires protégées étaient soumises à des restrictions minières strictes. Récemment, Ibanez et al. (2018) ont conduit une étude afin d'évaluer les évolutions du réseau d'aires protégées depuis Jaffre et al. (1998). Ils ont montré que, même si la surface totale des aires protégées terrestres a augmenté de 35%, elles ne couvraient en 2017 que 4% du territoire néo-calédonien. Ce chiffre est bien en deçà de l'objectif d'Aichi 11 qui visait au moins 17% de la surface terrestre protégée pour 2020 (CBD, 2010). Il est cependant important de souligner la création en avril 2019 (donc postérieure à cette étude) du parc provincial de la Côte Oubliée – 'Woen Vùù – Pwa Pereeù en Province Sud, qui a permis à cette province d'atteindre l'objectif d'Aichi 11. Par ailleurs, Ibanez et al. (2018) a également montré que 72% des espèces de plantes protégées n'étaient toujours pas couvertes par des aires protégées en 2017. En revanche, le taux d'extraction du nickel en Nouvelle-Calédonie a doublé entre 1998 et 2018.

Pour conclure, il reste de nombreux défis à relever pour conserver la flore et les forêts de Nouvelle-Calédonie. Une façon évidente pour aller dans ce sens est d'agrandir et d'améliorer le réseau d'aires protégées. D'autres actions de conservation sont également possibles, comme la restauration écologique, la conservation ex-situ ou alors le renforcement du personnel de gestion des aires protégées (Ibanez et al., 2018). La Nouvelle-Calédonie dispose d'une administration et d'une économie solides, et le contexte insulaire et peu peuplé de l'archipel permet une grande proximité entre les différents acteurs de la conservation. Ces spécificités font de ce territoire un candidat idéal pour expérimenter de nouvelles approches de conservation. En particulier, les différents projets menés ces dernières années entre des instituts de recherche et les gestionnaires environnementaux néo-calédoniens (e.g. projet CORIFOR, Birnbaum et al., 2016; projet COGEFOR, Birnbaum et al., 2019) ont mis en évidence le besoin d'avoir accès à des outils d'aide à la décision précis, basés sur des données scientifiques, et capables de prendre en compte les contraintes spécifiques de gestionnaires.

Objectifs de recherche

En mettant l'accent sur les besoins de la Nouvelle-Calédonie en matière d'aide à la décision pour la planification de la conservation des forêts, cette thèse de doctorat s'articule autour de trois objectifs interdépendants.

1. *Améliorer l'expressivité de la planification systématique de la conservation.* Compte tenu du contexte socioéconomique de la Nouvelle-Calédonie, de l'hétérogénéité de la répartition de sa flore et de l'impact négatif de la fragmentation sur ses forêts (Ibanez et al., 2017), les problématiques de planification de la conservation dans cette zone sont hétérogènes et doivent prendre en compte de nombreuses contraintes impliquant différents acteurs. Dans un tel contexte, l'expressivité est une caractéristique critique pour l'aide à la décision. Par conséquent, le premier objectif de cette thèse de doctorat est de proposer une approche expressive pour la modélisation et la résolution des problèmes de planification de la conservation. Cette approche sera basée sur la programmation par contraintes, une technique formelle d'optimisation sous contraintes.
2. *Offrir plus de contrôle sur la configuration spatiale des solutions.* Pour lutter contre la fragmentation des forêts, il est nécessaire d'avoir un contrôle précis sur la configuration spatiale des solutions produites en planification de la conservation. De plus, les contraintes des gestionnaires en Nouvelle-Calédonie ne se limitent pas au budget et nécessitent souvent la prise en compte de critères spatiaux divers et variés (e.g. concessions minières, zones coutumières, accessibilité). À cet égard, le deuxième objectif de cette thèse est de permettre un contrôle précis sur la configuration spatiale des solutions produites par l'approche proposée en réponse au premier objectif.
3. *S'impliquer dans l'aide à la décision en Nouvelle-Calédonie.* En guise de baptême du feu, le troisième objectif de cette thèse consiste à confronter l'approche résultant des objectifs précédents à un problème réel de conservation en Nouvelle-Calédonie. Pour cela, nous avons engagé une collaboration avec les gestionnaires de l'environnement de la Province Sud de la Nouvelle-Calédonie dans un projet de planification de la restauration écologique dans le parc provincial de la Côte Oubliée - 'Woen Vùù - Pwa Pereeù, une zone symbolique de la Nouvelle-Calédonie avec des enjeux de conservation importants.

9.2 UNIFIER LES STRATÉGIES DE SÉLECTION DE RÉSERVE AVEC LA PROGRAMMATION PAR CONTRAINTES ET LES GRAPHES

Cette section est une retranscription littérale d'un résumé de l'article Justeau-Allaire et al. (2018) publié dans les actes des 15èmes journées francophones de la programmation par contraintes (JFPC 2019). Ce résumé a également été l'objet d'une présentation lors de ces journées. Dimitri Justeau-Allaire, Philippe Birnbaum, and Xavier Lorca (June 2019c). "Unifier Les Stratégies de Sélection de Réserve Avec La Programmation Par Contraintes et Les Graphes". In: *JFPC 2019 - Actes Des 15es Journees Francophones de Programmation Par Contraintes*. Albi, France: IMT Mines Albi, p. 27–29.

9.2.1 *Le problème de sélection de réserve*

Le problème de *sélection de réserve*[1] a été introduit entre le milieu des années 1970 et le début des années 1980 (cf. Sarkar, 2012 pour une revue historique sur le sujet). C'est un problème combinatoire qui s'applique dans un espace géographique maillé $\mathcal{P}$, généralement selon une grille carrée. L'objectif est de mettre en évidence une zone qui répond à un ensemble de contraintes écologiques et socioéconomiques, avec ou sans optimisation. Cette zone est composée de réserves, chacune composée d'un ensemble de cellules connectées selon un voisinage défini dans la maille. Un exemple est fourni dans la Figure 9.2 et deux voisinages dans une grille carrée sont illustrés dans la Figure 9.3.

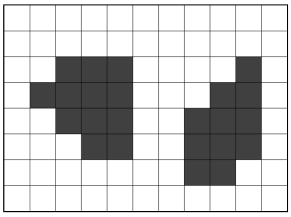


Figure 9.2: Exemple de solution au problème de sélection de réserve: deux réserves.

On distingue deux catégories de contraintes: les contraintes de couverture et les contraintes spatiales.

Contraintes de couverture L'espace géographique $\mathcal{P}$ est caractérisé par un ensemble de caractéristiques spatiales $\mathcal{F}$. Pour chaque caractéristique $j \in \mathcal{F}$, une valeur numérique v_{ji} (binaire, quantitative ou probabiliste) est définie pour chaque cellule $i \in \mathcal{P}$. Les caractéristiques spatiales peuvent représenter l'occupation du sol (e.g. écosystèmes), la biodiversité (e.g. espèces observées), ou des valeurs socioéconomiques (e.g. zones exploitables). Les contraintes de couverture s'appliquent sur ces caractéristiques et permettent de définir les seuils de couverture minimaux et/ou maximaux que la zone doit couvrir pour chacune d'elles.

---

[1] Aussi connu comme *conception de réserve*, ou *planification systématique de la conservation* Margules and Pressey, 2000.

Contraintes spatiales Les contraintes spatiales permettent de contrôler la configuration spatiale de la zone à délimiter, selon les attributs définis par Williams et al. Williams et al., 2005: *nombre de réserves* (i.e. de zones connectées), *surface des réserves*, *surface totale*, distance entre les réserves, connectivité dans les réserves, forme des réserves et zones tampon autour des réserves. Dans le cadre de ce travail nous nous sommes limités aux contraintes représentées en italique, les autres feront l'objet de travaux futurs.

(a) Voisinage 4-connecté. (b) Voisinage 8-connecté.

Figure 9.3: Deux voisinages dans une grille carrée (en gris: voisinage de x).

9.2.2 *Un modèle CP unifié*

Nous avons proposé un modèle CP permettant d'intégrer les contraintes de couverture et les contraintes spatiales de manière transparente. Dans ce modèle, chaque cellule i de la maille est représentée par une variable booléenne b_i. La cellule i est dans la zone à délimiter si et seulement si $b_i = 1$. Ces variables sont les variables de décision du modèle et il est possible d'y exprimer directement les contraintes de couverture via des contraintes de flot ou des contraintes de couverture par ensembles. Afin d'exprimer les contraintes spatiales, nous avons introduit $G = (\mathcal{P}, E)$, un graphe spatial non dirigé qui représente la maille. Un sommet est associé à chaque cellule $i \in \mathcal{P}$ et E est défini tel qu'il existe une arrête (u, v) si et seulement si u et v sont adjacents selon le voisinage choisi. A partir de G nous avons défini la variable graphe $g = (v, e) \in [\emptyset, G]$, chaînée avec les variables de décision de manière à ce que $b_i = 1 \Leftrightarrow i \in v \wedge (i, j) \in e \Leftrightarrow b_i = b_j = 1$. La variable g permet alors d'exprimer les contraintes de connectivité et celles sur la taille des réserves via ses composantes connexes. Ce modèle a été implémenté avec le solveur `Choco` (Prud'homme et al., 2017) et son extension `Choco-graph` (Fages et al., 2018) qui permet l'utilisation de variables graphes.

9.2.3 *Cas d'utilisation: fragmentation forestière au sud de la Nouvelle-Calédonie*

La Nouvelle-Calédonie est un archipel tropical situé dans la région du Pacifique Sud. C'est un point chaud de la biodiversité (Myers et al., 2000) dont la flore se caractérise par un taux d'endémisme exceptionnellement élevé. Malheureusement, les forêts néo-calédoniennes ne sont pas épargnées par la déforestation et elles sont menacées par la fragmentation, qui fragilise les communautés végétales et animales. Pour notre cas d'utilisation, nous nous sommes basés sur une étude d'Ibanez et al. conduite dans le sud de la

Nouvelle-Calédonie (Ibanez et al., 2017). Celle-ci visait à mieux comprendre l'impact de la fragmentation sur les communautés d'arbres dans une zone où la biodiversité est menacée par l'extraction du nickel, un des pilier de l'économie néo-calédonienne.

En concertation avec les écologues à l'origine de cette étude, nous avons considéré un scénario fictif basé sur des données réelles. Dans celui-ci, il est question de protéger une zone qui couvre 223 espèces d'arbres, au moins 340ha de forêt, qui est formée d'au plus deux réserves d'au moins 300ha et dont la surface totale n'excède pas 1000ha. Afin de minimiser les coûts d'établissement et de gestion, nous avons cherché à minimiser la surface totale de la zone à protéger. La meilleure solution obtenue (sans preuve d'optimalité) est illustrée dans la Figure 9.4.

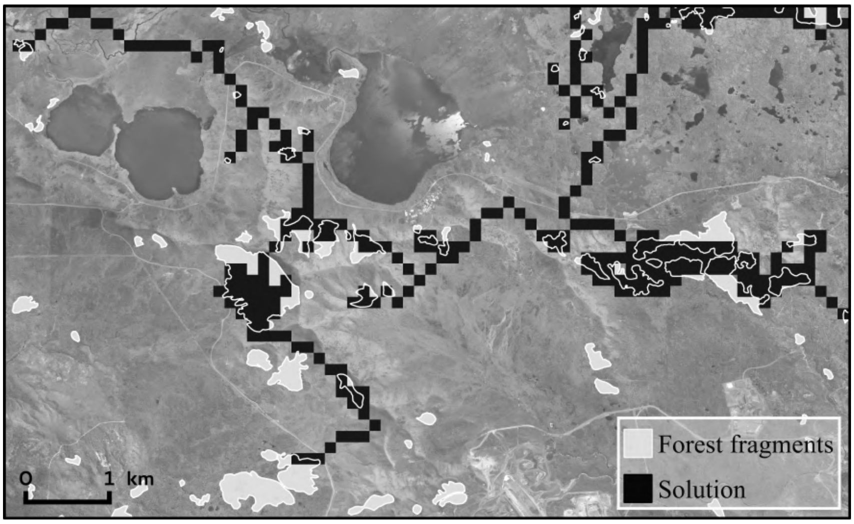


Figure 9.4: Meilleure solution obtenue dans le cas d'utilisation: environ 496ha, 1 réserve.

9.2.4 *Conclusions et perspectives*

Dans cet article, nous avons introduit le premier modèle CP permettant d'aborder le problème de sélection de réserve en combinant des contraintes spatiales et de couverture. Nous avons montré que cette technique permet d'unifier les différents aspects du problème dans un modèle à partir duquel nous avons pu traiter un cas d'étude basé sur des données réelles. Ce travail a été présenté aux autorités néo-calédoniennes et ouvre différentes perspectives en matière d'aide à la décision pour la conservation de la nature. Parmi celles-ci, les priorités pour nos travaux futurs concernent la modélisation de contraintes spatiales complexes (e.g. zone tampon) et la généralisation du modèle en un problème de partitionnement. Cette approche permettrait ainsi d'exprimer des contraintes sur des zones correspondant à différents niveaux de protection, en accord avec les lignes conductrices de l'IUCN pour la gestion des aires protégées (Dudley, 2008).

9.3 PARTITIONNEMENT DE L'ESPACE SOUS CONTRAINTES: UN MODÈLE GÉNÉRIQUE ET EXPRESSIF POUR LA PLANIFICATION DE LA CONSERVATION

Cette section est une retranscription littérale d'un résumé de l'article (Justeau-Allaire et al., 2019a) publié dans les actes du 21ème congrès annuel de la société française de recherche opérationnelle et d'aide à la décision (ROADEF 2020). Ce résumé a également été l'objet d'une présentation lors de ce congrès. Dimitri Justeau-Allaire, Philippe Vismara, Philippe Birnbaum, and Xavier Lorca (2020). "Partitionnement de l'espace Sous Contraintes: Un Modèle Générique et Expressif Pour La Planification de La Conservation." In: *ROADEF 2020 - Actes du 21ème congrès annuel de la société Française de Recherche Opérationnelle et d'Aide à la Décision*.

9.3.1 *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 apparait 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 (Pressey et al., 1993). En biologie de la conservation, on parle de planification de la conservation (Margules and Pressey, 2000). Concrètement, il s'agit de partitionner une zone géographique donnée en un ensemble de régions (cf. Figure 9.5 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 naturew (e.g. réserve naturelle), d'autre 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 des approches existantes (Justeau-Allaire et al., 2019a). 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

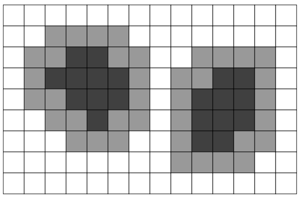


Figure 9.5: Un exemple de partitionnement sur une grille carrée, avec une zone de cœur de réserve (gris foncé), une zone tampon (gris clair), et une zone non protégée (blanc).

données réelles: la fragmentation des forêts sur substrat ultramafique dans le sud de la Nouvelle Calédonie.

9.3.2 *Un modèle CP générique et expressif pour la planification de la conservation*

9.3.2.1 *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 cette 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 9.6. 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 $\{R_0, ..., R_{n-1}\}$ 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 (Justeau-Allaire et al., 2018), qui s'organise selon deux catégories: contraintes de couverture et contraintes spatiales. Ce catalogue qui peut facilement être étendu.

9.3.3 *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 $\nu_f(x) \geqslant 1$, soit $\forall f \in \mathcal{F}, \exists x \in R, \nu_f(x) \geqslant 1$.

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 $\forall f \in \mathcal{F}, \prod_{x \in \mathcal{S}}(1 - \nu_f(x)) \leqslant 1 - \alpha$.

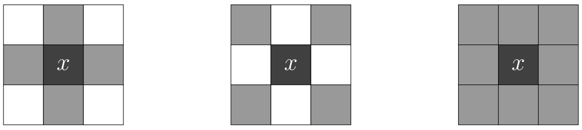
Contrainte C (Caractéristiques k-redondantes) Cette contrainte est satisfaite si et seulement si chaque caractéristique de $\mathcal{F}$ est couverte par au moins k sites distincts de R, soit $\forall f \in \mathcal{F}, \exists \mathcal{X} \subseteq R, |\mathcal{X}| \geqslant k \wedge \forall x \in \mathcal{X},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

9.3.4 *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.*



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

Figure 9.6: Quelques exemples de voisinage (en gris clair): le voisinage 4-connecté, le voisinage D-connecté (les cellules voisines diagonales) et le voisinage 8-connecté.

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.*

9.3.5 *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 (Gervet, 1995). 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 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 (Dooms, 2006; Fages, 2014). Étant donné un voisinage Γ_ω, l'espace géographique $\mathcal{S}$ peut être représenté par un graphe $G_\mathcal{S} = (\mathcal{S}, E_\mathcal{S})$, avec $E_\mathcal{S} = \{(x, y) \mid y \in \Gamma_\omega(x)\}$. Une illustration de $G_\mathcal{S}$ est proposée dans la figure Figure 9.7, avec $\Gamma_\omega = \Gamma_4$, le voisinage 4-connecté dans une grille carrée. A chaque région est associée une variable graphe $G_u = (R_u, E_u) \in [\emptyset, G_\mathcal{S}]$, qui est contrainte de manière à être le sous-graphe de $G_\mathcal{S}$ induit par R_u. Une illustration de $\{G_0, G_1, G_2\}$ est proposée en Figure 9.7 pour un partitionnement en 3 régions d'une grille carrée.

Contraintes utilisateur. Toutes les contraintes définies dans les sections 9.3.3 et 9.3.4 peuvent être appliquées de manière transparente à toutes les régions.

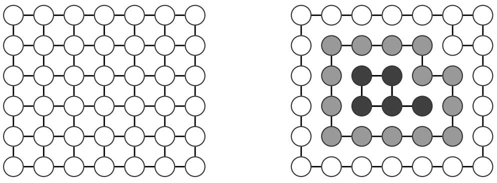

(a) G_S. (b) $\{G_0, G_1, G_2\}$

Figure 9.7: 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).

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 Justeau-Allaire et al. (2018, 2019a).

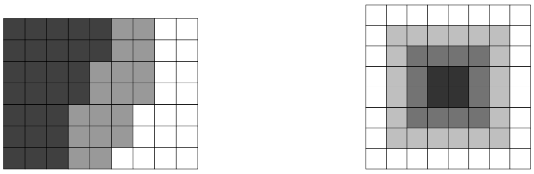
9.3.6 *Un modèle extensible: illustration avec la contrainte de zone tampon*

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 (Harris, 1988; Fahrig, 2003). 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.

Constraint K (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:

$$\begin{aligned} &\Gamma_\omega(R_u) \cap R_v = \emptyset; \\ &R_u \cap \Gamma_\omega(R_v) = \emptyset; \\ &B = \Gamma_\omega(R_u) \cap \Gamma_\omega(R_v). \end{aligned} \tag{9.1}$$

Exprimée ainsi, il est possible d'atteindre la consistance aux bornes (telle que définie sur les variables ensemblistes dans Walsh (2003)) en $O(|S|^4)$ (Justeau-Allaire et al., 2019a). 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 9.8 pour une illustration).



(a) Zone tampon de largeur 2 (gris clair)

(b) 4 régions imbriquées, $\Gamma_\omega = \Gamma_8$

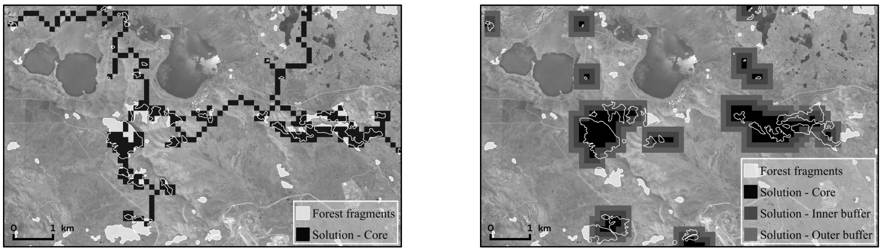
Figure 9.8: Exemples d'utilisation de la contrainte de zone tampon.

9.3.7 *Cas d'étude: Fragmentation forestière en Nouvelle-Calédonie*

Nous avons implémenté le modèle présenté précédemment avec le solver Choco et son extension Chocho-graph (Prud'homme et al., 2017), 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. (Ibanez et al., 2017) dans une zone de 60km^2 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. (Pouteau et al., 2015) et Schmitt et al. (Schmitt et al., 2017). A partir de ce jeu 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 (Justeau-Allaire et al., 2019a). Deux des solutions obtenues sont illustrées en Figure 9.9.

9.3.8 *Conclusion et perspectives*

Dans cet article, nous avons résumé les résultats introduits dans Justeau-Allaire et al. (2019a), 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 questionnement 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.

Figure 9.9: Résultats obtenus pour deux des scénarios du cas d'étude.

9.4 OPTIMISATION SOUS CONTRAINTES D'INDICES DU PAYSAGE POUR LA RESTAURATION ÉCOLOGIQUE EN NOUVELLE CALÉDONIE

Cette section est un résumé en français de l'article Justeau-Allaire et al., 2020.

9.4.1 *Introduction*

Dans le contexte de la crise mondiale de la nature, la dégradation et perte des habitats naturels dus aux changements d'utilisation des terres sont les causes majeures de l'effondrement des écosystèmes et du déclin de la biodiversité (Haddad et al., 2015). La configuration du paysage peut avoir des impacts importants sur les processus écologiques tels que la dispersion, la résistance au feu, ou les flux génétiques (Taylor et al., 1993; Fahrig, 2003). Ces impacts sont en général évalués via des indices de fragmentation et de connectivité entre les patches d'habitats (Uuemaa et al., 2013). La restauration écologique et la planification de la conservation peuvent contribuer à réduire la perte d'habitat et à favoriser des configurations paysagères favorables à la persistance des communautés d'espèces. La planification systématique de la conservation (PSC) est une approche qui s'appuie sur différentes méthodes d'optimisation (e.g. heuristiques ad hoc, métaheuristiques, programmation linéaire en nombres entiers mixtes – MILP). Les travaux récents dans ce domaine ont introduit de nombreuses perspectives pour prendre en compte la configuration du paysage dans les procédures d'optimisation pour la PSC. En parallèle, la recherche en écologie du paysage a fourni de nombreux indices pour évaluer le niveau de fragmentation (McGarigal, 2014) et de connectivité (Pascual-Hortal and Saura, 2006; Saura and Pascual-Hortal, 2007) à l'échelle du paysage. Jusqu'à présent, ces indices ont principalement été utilisés dans des contextes d'analyse de scénarios, mais très peu dans des procédures d'optimisation pour la PSC.

Récemment, nous avons introduit une approche générique et expressive pour la PSC de la conservation basée sur la programmation par contraintes (PPC) (Justeau-Allaire et al., 2019a). La PPC est une approche exacte d'optimisation sous contraintes basée sur le raisonnement automatique. Dans l'article (Justeau-Allaire et al., 2020) ici résumé, nous avons étendu cette approche avec des indices du paysage et nous l'avons appliqué pour un projet de reforestation situé dans le parc provincial de la "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" dans le sud-est de la Nouvelle-Calédonie. En collaboration avec les gestionnaires en charge du projet, nous avons fourni une aide à la décision pour identifier des zones de reforestation respectant les contraintes des gestionnaires et optimales pour réduire la fragmentation et l'isolation forestière, connus pour avoir des effets négatifs sur les communautés d'arbres dans cette région (Ibanez et al., 2017). Tout en respectant des contraintes de budget, d'accessibilité et d'allocation équitable des ressources, nous avons identifié des solutions optimales pour deux indices du paysage : la taille effective des mailles (MESH; Jaeger, 2000) et l'indice intégral de connectivité (IIC; Pascual-Hortal and Saura, 2006). Dans un contexte de reforestation, MESH favorise la fusion physique de gros patches entre eux tandis qu'IIC favorise la restauration de la connectivité structurelle entre des grands groupes de patches. Nos résultats ont démontré la flexibilité de cette approche et la manière avec laquelle son expressivité (i.e. l'étendue et la variété des problèmes qui peuvent être représentés et résolus) facilite la représentation des problèmes de PSC dans toute leur diversité. Cette approche offre de nouvelles perspectives d'aide à la décision pour la restauration écologique et de manière générale pour la planification de la conservation (e.g. délimitation de réserves naturelles, de corridors écologiques).

9.4.2 *Matériel et méthodes*

Cas d'étude: planification de la reforestation dans le parc provincial de la "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" en Nouvelle-Calédonie

La Nouvelle-Calédonie est un archipel tropical situé dans le Pacifique Sud. En tant que plus petit point chaud de la biodiversité du monde, l'archipel héberge des écosystèmes terrestres et marins très riches. Notamment, la flore néo-calédonienne se distingue par l'un des plus forts taux d'endémisme dans le monde (76%), une diversité bêta élevée, et la présence d'espèces reliques (Ibanez et al., 2014; Grandcolas et al., 2008; Pillon, 2012). Les forêts de Nouvelle-Calédonie sous menacées par les activités humaines (e.g. extraction minière, feu, espèces invasives). Dans le sud-est de la Grande Terre (île principale de l'archipel), les autorités coutumières de la zone de la Côte Oubliée ont établi, entre 2014 et 2016, un moratoire sur la construction d'infrastructures et sur l'exploitation minière suite à l'érosion de nombreuse zones. Ce moratoire a été renouvelé pour dix ans (de 2018 à 2028) et a conduit la Province Sud de Nouvelle-Calédonie a créer le parc provincial de la "Côte Oubliée 'Woen Vùù – Pwa Pereeù" en 2019, qui contient 93 000 ha de zone terrestre et 27 000 ha de zone marine. Les gestionnaires de la direction du développement durable des territoires (DDDT) de la Province Sud de Nouvelle-Calédonie doivent maintenant établir un plan de gestion pour le parc, avec notamment un accent mis sur la réduction de la fragmentation forestière.

Dans cette étude, nous nous intéressons à un projet de reforestation qui doit être planifié par la DDDT. Un de ses objectifs consiste à identifier deux zones pour la reforestation, une dans chacun des districts coutumiers de la Côte Oubliée (Borendy et Unia). Le budget disponible correspond à environ 200 ha de reforestation. Pour des raisons d'accessibilité, les zones identifiées doivent être connectées, compactes et accessibles (proche des pistes). Sous ces contraintes, l'objectif est d'optimiser la contribution potentielle des zone reforestées à la réduction de la fragmentation forestière et à l'amélioration de la connectivité structurelle.

Données

La zone de la Côte Oubliée a été peu prospectée et peu de connaissances sont disponibles sur la dispersion des espèces animales et végétales des forêts néo-calédoniennes. Pour ces raisons, nous avons opté pour une approche basée sur la couverture forestière à partir de données issues de la télédétection. En s'appuyant sur un raster de couverture forestière en 2019 avec une résolution de 30 mètres (Vancutsem et al., 2020), nous avons produit (voir Figure 9.10):

- Un raster binaire avec une résolution de 480 mètres qui représente les cellules dégradées ($\leqslant 70\%$ de couverture forestière) et non dégradées ($\geqslant 70\%$ de couverture forestière).
- Un raster avec une résolution de 480 mètres avec pour chaque cellule la surface reforestable.
- Un raster avec une résolution de 480 mètres représentant les cellules accessible dans les districts de Borendy et d'Unia.

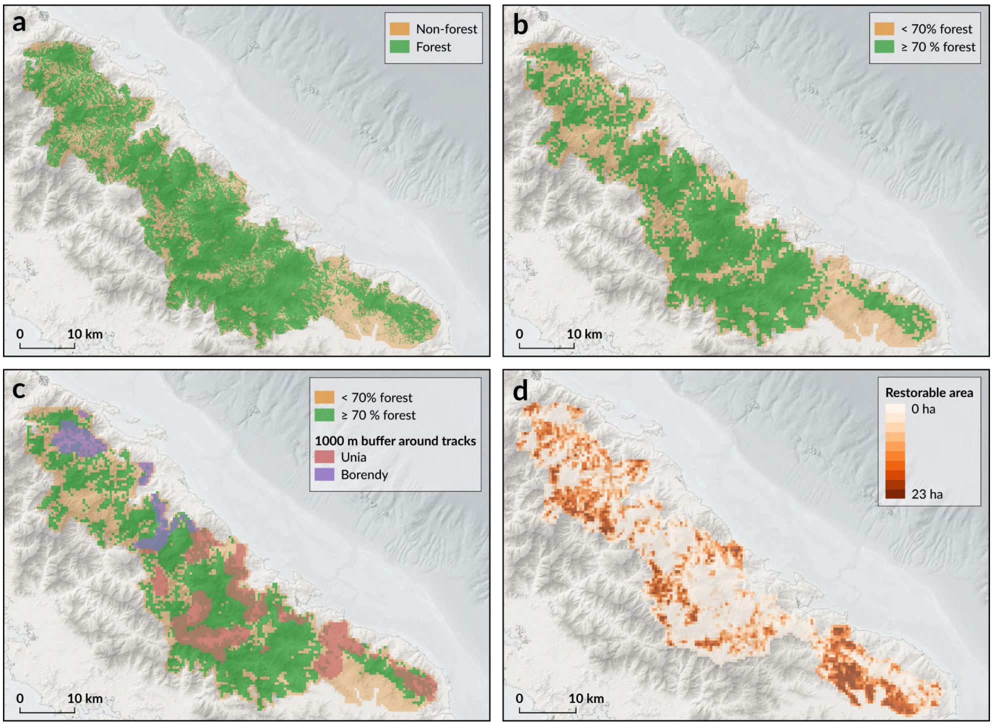


Figure 9.10: Données. (**a**) Raster binaire de couverture forestière en 2019 avec une résolution de 30 mètres. (**b**) Raster de couverture forestière avec une résolution de 480 mètres, les cellules avec au moins 70% de couverture forestières sont considérées comme non-dégradées, les autres dégradées. (**c**) Raster représentant les zones accessibles dans les districts de Borendy et d'Unia, avec une résolution de 480 mètres. (**d**) Raster représentant la surface de zone reforestable avec une résolution de 480 mètres.

Formulation mathématique du problème

À chaque cellule de 480 mètres, est associée une unité de planification (UP) qui peut être sélectionnée pour la reforestation. Ces UPs constituent l'ensemble $\mathcal{S}$ et sont les variables de décision du modèle. Aux données d'entrées, sont associés les ensembles suivants :

$$\begin{aligned}&\mathcal{U}, \text{l'ensemble des UPs accessibles dans le district d'Unia;}\\&\mathcal{B}, \text{l'ensemble des UPs accessibles dans le district de Borendy;}\\&\mathcal{F}_{\geqslant 70\%}, \text{l'ensemble des UPs avec une couverture forestière} \geqslant 70\%;\\&\mathcal{F}_{<70\%}, \text{l'ensemble des UPs avec une couverture forestière} < 70\%.\end{aligned} \tag{9.2}$$

Soient $R_u \subseteq \mathcal{F}_{<70\%}$ et $R_b \subseteq \mathcal{F}_{<70\%}$ les ensembles d'UPs à reforester respectivement dans les districts d'Unia et de Borendy. L'ensemble formé par l'union de R_u, R_b, $\mathcal{F}_{\geqslant 70\%}$ et $\mathcal{F}_{<70\%}$ doit alors former une partition de l'ensemble $\mathcal{S}$ et l'ensemble $R_u \cap R_b \cap \mathcal{F}_{\geqslant 70\%}$ correspond à la couverture forestière potentielle après reforestation. À chacun de ces ensembles est associé un graphe grille construite à partir de la 4-connectivité dans une grille. Étant donné une région $R \subseteq \mathcal{S}$, nous définissons les contraintes suivantes :

- CONNECTED(R), qui est satisfaite si le graphe associé à la région R est connecté.
- RESTORABLE(R, a, p), qui est satisfaite si toutes les UPs dans la région R peuvent être restaurées pour atteindre une couverture forestière d'au moins p en investissant au total a (en unité de surface).
- RADIUS(R, ρ), qui est satisfaite si le rayon du plus petit cercle englobant l'ensemble des UPs de R est égal à ρ.

Un couple (R_u, R_b) satisfait les contraintes de budget, d'accessibilité et d'allocation équitable si et seulement si l'ensemble des contraintes suivantes sont satisfaites, avec A_{max} la surface totale à reforester (200 ha) et P_{max} le rayon maximum du plus petit cercle englobant chacune des zones reforestées:

$$R_u \subseteq \mathcal{U} \cap \mathcal{F}_{<70\%} \wedge R_b \subseteq \mathcal{B} \cap \mathcal{F}_{<70\%}; \tag{9.3}$$

$$\textsc{Connected}(R_u) \wedge \textsc{Connected}(R_b); \tag{9.4}$$

$$a_u \in 0.5 \cdot A_{\mathrm{max}} \pm 10\% \wedge \textsc{Restorable}(R_u, a_u, 70\%); \tag{9.5}$$

$$a_b \in 0.5 \cdot A_{\mathrm{max}} \pm 10\% \wedge \textsc{Restorable}(R_b, a_b, 70\%); \tag{9.6}$$

$$a_u + a_b \leqslant A_{\mathrm{max}}; \tag{9.7}$$

$$a_{\mathrm{max}} \in [0, +\infty] \wedge \textsc{Restorable}(R_u \cup R_b, a_{\mathrm{max}}, 100\%); \tag{9.8}$$

$$a_{\mathrm{max}} \geqslant A_{\mathrm{max}}; \tag{9.9}$$

$$\rho_u \in [0, P_{\mathrm{max}}] \wedge \textsc{Radius}(R_u, \rho_u); \tag{9.10}$$

$$\rho_b \in [0, P_{\mathrm{max}}] \wedge \textsc{Radius}(R_b, \rho_b). \tag{9.11}$$

À partir de ce problème de base, qui permet de garantir la satisfaction des contraintes formulées par les gestionnaires de la DDDT, nous avons défini deux problèmes d'optimisation sous contraintes, respectivement associés avec la maximisation de l'indice MESH et de l'indice IIC. Ces deux indices du paysage s'intéressent à une région (ou classe de paysage, e.g. forêt) R, dont l'ensemble des patches est donné par $P(R)$.

- MESH est un indice de fragmentation basé sur la distribution cumulative des surfaces de patches (dans notre cas de forêt) dans un paysage. Il s'exprime en unités de surface (dans notre cas en ha) et correspond à la surface des patches lorsque le paysage est transformé en un ensemble de patches de même taille tout en conservant la probabilité que deux points sélectionnés au hasard sont dans le même patch (Jaeger, 2000). Pour une région R, avec A_k la surface du patch k et A_L la surface totale du paysage, MESH est donné par :

$$\mathrm{MESH}(R) = \frac{1}{A_L} \sum_{k \in P(R)} A_k^2. \tag{9.12}$$

- IIC est un indice de connectivité inter-patch basé sur une représentation du paysage sous la forme d'un graphe, où chaque patch est représenté par un nœud et où deux patches sont connectés selon un critère structurel (e.g. distance maximale) ou fonctionnel (e.g. trajectoires de migration) (Pascual-Hortal and Saura, 2006). Dans notre cas d'étude, nous utilisons le critère structurel d'au plus une cellule vide entre deux patches. Pour une région R, avec A_k la surface du patch k, A_L la surface totale du paysage et d_{kl} la longueur du plus court chemin (distance topologique) entre le patch k et le patch l, IIC est donné par :

$$\mathrm{IIC}(R) = \frac{1}{A_L^2} \sum_{k \in P(R)} \sum_{l \in P(R)} \frac{A_k \cdot A_l}{1 + d_{kl}} \tag{9.13}$$

Méthode de résolution: l'approche de planification systématique de la conservation basée sur les contraintes

Afin de résoudre le problème introduit dans la section précédente, nous nous sommes appuyés sur l'approche de PSC basée sur les contraintes rapidement brièvement présentée dans l'introduction (Justeau-Allaire et al., 2019a), dont le code source est disponible sur GitHub[2]. Cette approche repose sur la programmation par contraintes (PPC), une méthode d'optimisation sous contraintes basée sur des techniques de raisonnement automatique. Cette approche est particulièrement adaptée pour résoudre des problèmes combinatoires fortement contraints, et ses principaux avantages sont sa flexibilité, son expressivité, et son caractère formel qui permet d'obtenir des garanties sur la satisfaisabilité et l'optimalité des problèmes. Dans le cadre de cette étude, nous avons étendu l'approche avec la contrainte RADIUS, la contrainte MESH et la contrainte IIC. Nous avons résolu les deux problèmes d'optimisation décrits dans la section précédente sur un serveur Linux (Intel Xeon E5-2620 CPU 2.40GHz × 12, 64 GB RAM). Le code source du cas d'étude est également disponible sur GitHub[3].

---

[2] https://github.com/dimitri-justeau/choco-reserve

[3] https://github.com/dimitri-justeau/cote-oubliee-choco-reserve-code

9.4.3 *Résultats*

Les résultats sont résumés dans la Table 9.1, et une des solutions optimale du problème de maximisation de MESH (respectivement IIC) est représentée en Figure 9.11 (respectivement Figure 9.12). Le solveur a pu trouver l'ensemble des solutions optimales pour les deux problèmes, avec un temps de calcul plus long pour l'indice IIC. Plusieurs solutions optimales existent pour les deux problèmes, mais elles sont situées dans les mêmes zones et reconnectent les mêmes patches de forêt.

Objectif	maximisation de MESH	maximisation d'IIC
Valeur dans le paysage actuel	24 542 ha	0.20691
Valeur optimale après reforestation	25 502 ha	0.22986
Amélioration potentielle	3.91%	11.09%
Nombre de solutions optimales	7	3
Temps de résolution (optimisation)	14.7 min	5.8 h
Temps de résolution (énumération)	18 s	19.7 min

Table 9.1: Résultats: pour chaque indice, sa valeur dans le paysage actuel, sa valeur optimale après reforestation, l'amélioration potentielle, le nombre de solutions optimales et les temps de résolution. MESH: taille effective des mailles, IIC: indice intégral de connectivité.

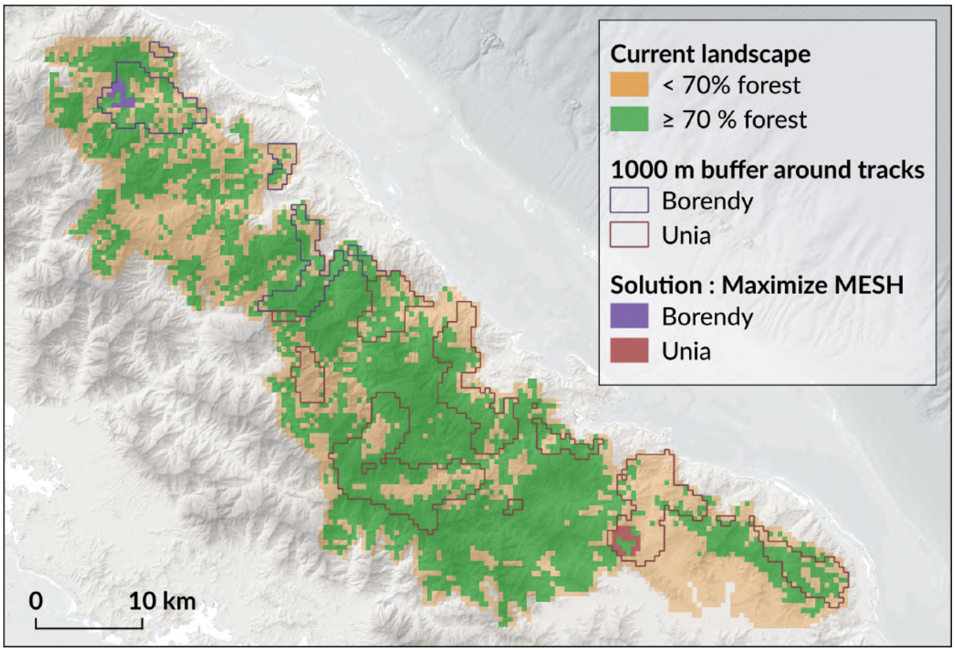


Figure 9.11: Une des solutions maximisant la taille effective des mailles (MESH).

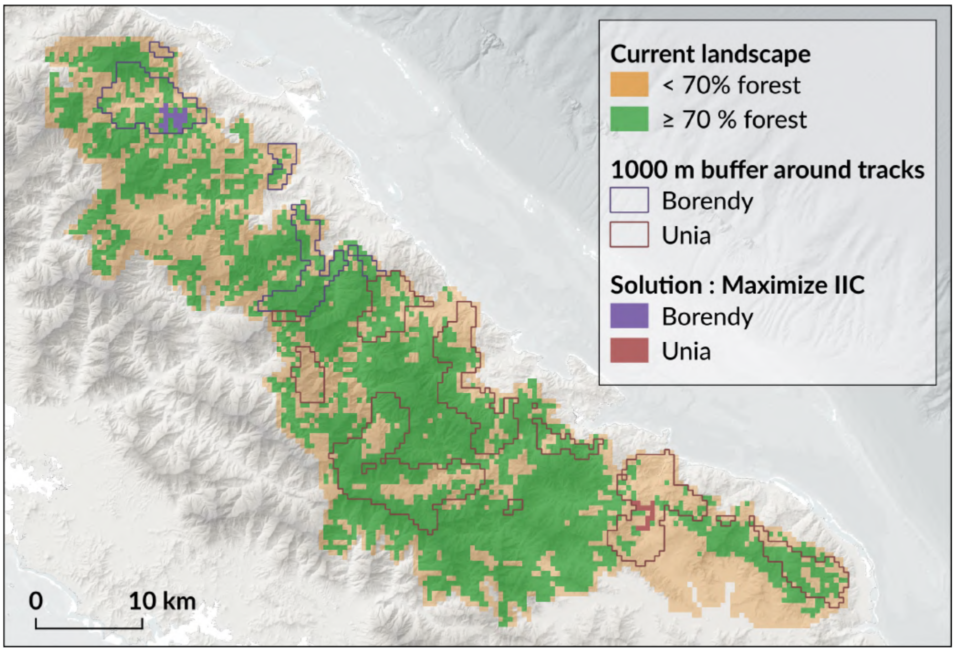


Figure 9.12: Une des solutions maximisant l'indice intégral de connectivité (IIC).

9.4.4 *Discussion*

Contribution à l'aide à la décision dans le projet de reforestation du parc provincial de la "Côte Oubliée – 'Woen Vùù – Pwa Pereeù"

En respectant des contraintes de budget, d'accessibilité et d'allocation équitable, nous avons pu identifier l'ensemble des solutions optimales pour un indice de fragmentation (MESH) et un indice de connectivité structurelle (IIC) avec des temps de calcul relativement courts. Ces résultats ont contribué à l'aide à la décision en fournissant à la DDDT deux scénarios optimaux selon l'indice considéré. À cet égard, ces résultats ont constitué une base spatialement explicite pour enrichir les discussions entre les parties prenantes du projet, tout en proposant un ensemble de zones spécifiques à considérer pour une étude plus fine sur le terrain. Ces résultats ont été bien reçus par les gestionnaires de la DDDT, qui étaient notamment enthousiastes de voir qu'il est possible de garantir que toutes leurs contraintes sont respectées par les solutions et que le solveur est capable d'informer l'utilisateur s'il n'est pas possible de toutes les satisfaire.

À propos de l'utilisation d'indices du paysage en planification systématique de la conservation

Les résultats de cette étude illustrent le potentiel offert par l'intégration d'indices avancés issus de l'écologie du paysage pour réduire la fragmentation et améliorer la connectivité. En Nouvelle-Calédonie, la fragmentation forestière est connue pour avoir des effets néfastes sur les communautés d'arbres (Ibanez et al., 2017) et la connectivité structurelle est un élément important pour la dispersion et la persistance des espèces, ainsi que les flux génétiques entre les communautés (Taylor et al., 1993). L'optimisation de ces indices en planification systématique de la conservation (PSC) permet d'informer sur les bénéfices potentiels des actions de conservation sur la fragmentation et sur la connectivité à l'échelle du paysage. L'avantage principal de l'optimisation combinatoire sur l'analyse de scénarios prédéfinis réside dans le fait que les solutions sont produites en considérant toutes les combinaisons possibles d'unités de planification satisfaisant les contraintes définies par les utilisateurs. Cette particularité offre la garantie aux gestionnaires qu'aucune opportunité n'a été manquée.

Avantages de l'approche basée sur les contraintes pour la planification systématique de la conservation

Dans cette étude, notre approche de PSC basée sur les contraintes a montré sa capacité à exprimer et résoudre des problèmes réels de PSC tout en apportant des garanties de satisfaisabilité et d'optimalité. Son expressivité permet une représentation précise des contraintes que les gestionnaires doivent prendre en compte pour la mise en place des actions de conservation. Combinée avec la garantie de satisfaisabilité, cette caractéristique permet d'assurer que les solutions proposées seront réalistes d'un point de vue socio-économique (Game et al., 2015; Williams et al., 2020). De plus, cette approche est suffisamment flexible pour être utilisée dans d'autres types de problèmes de PSC (e.g. délimitation de réserves ou de corridors écologiques) car les contraintes peuvent être modifiées, ajoutées ou retirées d'un problème sans impact sur la méthode de résolution.

Limitations actuelles et perspectives

Il est encore nécessaire d'investir beaucoup d'efforts dans le développement logiciel pour proposer un outil accessible à une large audience. En effet, l'approche requiert actuellement une expérience en programmation par contraintes (PPC) pour être utilisée correctement. De plus, la PPC étant une approche exacte, le temps de résolution peut être long pour les problèmes de taille élevée. Une autre limite vient de la grille carrée régulière utilisée pour ce cas d'étude, qui implique un compromis entre la résolution spatiale et la sophistication du modèle. Une perspective pour repousser cette limite pourrait consister en l'utilisation de grilles irrégulières afin d'augmenter (ou de diminuer) localement la résolution spatiale du problème sans augmenter le nombre total d'unités de planification. Malgré ces limites, nous avons montré qu'il y a du potentiel dans l'utilisation de la PPC pour formuler et résoudre des problèmes de PSC. Notamment, l'expressivité offerte par cette approche permet d'aller plus loin dans la finesse de représentation des problèmes. La notion d'optimalité est sujette à de nombreux débats en PSC (Underhill, 1994; Pressey et al., 1996; Rodrigues and Gaston, 2002; Hanson et al., 2019a), nous soutenons cependant le fait que l'expressivité est un attribut plus important, qui devrait être considéré avant l'optimalité (Rodrigues et al., 2000; Moilanen, 2008).

9.5 DISCUSSION GÉNÉRALE ET CONCLUSION

9.5.1 *Synthèse des contributions*

Co-financée à travers une collaboration entre le Cirad et l'IAC, cette thèse de doctorat avait pour objectif de répondre à des besoins récurrents de la part des gestionnaires en terme d'aide à la décision pour la conservation des forêts de Nouvelle-Calédonie. Comme nous l'avons vu dans la Section 9.1.3, la Nouvelle-Calédonie est un point chaud de la biodiversité qui doit faire face à de nombreux défis pour la conservation de sa flore, exceptionnellement riche. Ces enjeux de conservation interviennent dans un contexte socioéconomique complexe puisque l'économie de l'archipel repose majoritairement sur l'extraction minière qui est une des trois principales menaces sur la flore néo-calédonienne. De plus, la Nouvelle-Calédonie est engagée dans un processus d'autodétermination depuis les accords de Nouméa (Jospin, 1998), à l'issue duquel le peuple néo-calédonien décidera de son indépendance ou non vis-à-vis de la France. Ce contexte de transition politique implique un certain nombre de contraintes sur lesquelles les décisions relatives à la conservation ne peuvent faire l'impasse. Malgré ce contexte particulier, les gestionnaires environnementaux de Nouvelle-Calédonie font preuve d'une réelle volonté d'améliorer leurs politiques de gestion de la nature à travers des collaborations avec le monde scientifique et la société civile. Si les problématiques qu'ils rencontrent s'inscrivent clairement dans le cadre de la planification de la conservation, ils n'ont jusqu'à présent pas adopté les outils existants qui, selon notre analyse, n'offrent pas un niveau d'expressivité suffisant pour s'adapter à leurs besoins.

Quelles contributions pour la planification systématique de la conservation?

De manière plus générale, deux objectifs de cette thèse étaient : (i) de permettre plus d'expressivité dans la planification systématique de la conservation (PSC) et de (ii) offrir plus de contrôle sur la configuration spatiale des solutions produites par la PSC, notamment via l'intégration d'indices du paysage avancés. À ce titre, nous avons pu voir que les articles résumés dans les Sections 9.2, 9.3 et 9.4 ont tous contribué à ces deux objectifs. Les deux premiers articles ont introduit le premier modèle de programmation par contraintes (PPC) générique pour la SCP et ont montré que cette technique permet l'intégration formelle de contraintes spatiales qui n'était jusque-là pas disponible dans les approches basées sur la programmation linéaire en entiers mixtes, l'approche formelle la plus utilisée en SCP. Finalement, le troisième article a montré qu'il est possible d'intégrer des indices du paysage avancés de manière formelle dans cette approche basée sur la PPC et de les optimiser pour évaluer l'impact potentiel des actions de conservation sur la configuration du paysage. À notre connaissance, les indices considérés dans cette étude n'avaient jamais été intégrés dans des procédures d'optimisation en PSC. En conclusion, les trois articles publiés dans le cadre de cette thèse de doctorat ont montré qu'il est possible de réaliser des progrès méthodologiques substantiels en SCP via l'utilisation de techniques issues de l'intelligence artificielle. Nous avons montré que la PPC est une technique adaptée pour la PSC et qu'elle permet d'apporter de nouveaux éléments à la modélisation et à la résolution des problèmes associés. Nous espérons que ces résultats stimuleront l'exploration de nouvelles approches en SCP.

Quelles contributions pour la fertilisation croisée entre la planification systématique de la conservation et l'optimisation sous contraintes?

Si les contributions apportées par la PPC (et plus généralement l'optimisation sous contraintes) à la PSC et à l'aide à la décision en Nouvelle-Calédonie sont assez visibles, il est intéressant de se poser la question dans l'autre sens. La première chose à remarquer est "l'exotisme" du cas d'application offert par la PSC pour l'optimisation sous contraintes, traditionnellement utilisée dans des contextes industriels, logistiques ou militaires. Cet exotisme se remarque notamment dans les premières publications de cette thèse de doctorat (cf. Sections 9.2 et 9.3), publiées dans les actes de deux conférences internationales de référence en PPC et en intelligence artificielle. En effet, dans ces deux conférences, nos travaux étaient les seuls à s'intéresser à la PSC (sur 114 articles acceptés à CP 2018 et 850 articles acceptés à IJCAI 2019) et nous avons perçu un fort enthousiasme lors de nos échanges avec les participants de ces conférences sur les motivations éthique de la PSC. En outre, nos travaux ont également montré que la PSC peut stimuler les développements techniques et théoriques en optimisation sous contraintes, à travers le développement de contraintes originales sur les graphes et les ensembles. De plus, certaines de ces contraintes peuvent être réutilisées dans d'autres contextes applicatifs, tels que l'atténuation des risques ou la planification spatiale humanitaire.

9.5.2 *Perspectives*

Avons-nous besoin de plus de données ou de meilleurs modèles pour améliorer la planification systématique de la conservation?

À l'ère du big data et du machine learning, l'accent est régulièrement mis sur les données. Cependant, l'approche que nous avons utilisée dans cette thèse est principalement basée sur le problème. En intelligence artificielle, on distingue souvent les approches basées sur les données (data-driven) et les approches basées sur le problème (knowledge-driven). Un axe de recherche en intelligence artificielle consiste à combiner ces approches pour en repousser les limites respectives. Le choix d'une de ces deux approches doit se faire après une analyse et une définition rigoureuse du problème à résoudre. Dans le cas de la planification systématique de la conservation (PSC), il est important de remarquer que les problèmes impliqués sont : (i) fortement combinatoires et (ii) peuvent varier considérablement dans leur structure d'une instance à l'autre. À partir de ces deux caractéristiques, on peut anticiper le fait que les approche purement basées sur les données seront probablement peu adaptées. Par conséquent, la PSC a avant tout besoin de modèles basés les problèmes. Naturellement, ces modèles ont besoin d'un volume de données suffisant et d'une qualité suffisante pour être utilisés. Nous pensons que l'amélioration de l'efficacité de la PSC passe avant tout par la définition d'un formalisme générique permettant d'exprimer la diversité des problèmes qui peuvent être rencontrés. L'objectif d'un tel formalisme serait principalement de fournir un outil commun pour définir les problèmes et ses avantages seraient multiples. Premièrement, il permettrait aux utilisateurs de mieux définir leurs problèmes et donc de mieux identifier les méthodes les plus adaptées pour le résoudre. Un formalisme de ce type permettrait également aux différentes approches de se comparer en terme de fonctionnalités et de performances de manière systématique. En conclusion, malgré notre prise de position pour le développement de meilleurs modèles pour la PSC, les données restent un élément essentiel pour alimenter les modèles et pour la recherche fondamentale en écologie, elle-même essentielle pour formuler des questions pertinentes en PSC.

Avons-nous atteint un niveau d'expressivité suffisant pour fournir une aide à la décision réaliste?

Dans cette thèse de doctorat, nous avons montré qu'il est possible d'apporter plus d'expressivité dans la formulation des problèmes de PSC. Est-il encore possible d'améliorer cette expressivité et dans quelle mesure est-ce important pour une aide à la décision réaliste ? Premièrement, plus il est possible d'intégrer de la connaissance dans un modèle d'aide à la décision, plus les résultats que ce modèle fournira seront proche du problème "réel". De même, plus un modèle est expressif, plus il possible d'y intégrer une grande variété de connaissances. Nous pouvons donc d'ores et déjà affirmer que l'expressivité est un attribut de qualité qui permet d'améliorer la pertinence de l'aide à la décision. Cependant, l'adoption d'un outil d'aide à la décision par les parties prenantes d'un projet et son usage adéquat sont des pré-requis indispensables. De fait, un modèle très simple peut avoir des bénéfices énormes s'il est utilisé à bon escient dans un contexte réel. C'est le cas par exemple du premier algorithme de sélection de réserve (Kirkpatrick, 1983) qui, malgré sa simplicité et le fait qu'il fut appliqué sans l'aide d'ordinateurs, a conduit à la création de sept nouvelles aires protégées en Tasmanie. Un autre exemple intéressant est le logiciel Marxan (Ball et al., 2009) qui est l'approche la plus utilisée à ce jour. Sans être l'approche la plus sophistiquée du marché, les créateurs de Marxan ont réussi à diffuser largement leur outil grâce à de nombreux efforts de communication auprès d'un public non-académique (e.g. guide des bonnes pratiques avec Marxan; Ardron et al., 2008). Ces efforts vont dans une direction qui est défendue par de nombreux biologistes de la conservation : réduire le fossé entre la recherche et l'implémentation sur le terrain en créant de connections entre le monde académique et le monde non-académique (Knight et al., 2008; Game et al., 2015; Ellison, 2016; Williams et al., 2020). En conclusion, l'expressivité est un attribut qui présente de nombreux bénéfices potentiels pour l'aide à la décision, mais ces bénéfices ne pourront exister qu'à travers un fort engagement pour la diffusion des méthodes expressives auprès des gestionnaires et plus généralement du monde non-académique.

9.5.3 *Conclusion: contribuons-nous à de meilleures actions de conservation?*

En d'autres termes, l'investissement en recherche et développement pour la planification systématique de la conservation (PSC) est utile pour la conservation de la nature, ou est-ce un prétexte pour la recherche ? Comme nous l'avons montré dans les Sections **??** et 9.5.1, nous avons suffisamment de recul pour affirmer que la PSC a le potentiel nécessaire pour soutenir de meilleures actions de conservation. Cependant, la PSC reste avant tout un outil d'aide à la décision qui ne peut être pertinent que si (i) la recherche fondamentale en amont lui fournit suffisamment de connaissances et (ii) si elle est appliquée dans des contextes réels avec les différentes parties prenantes des projets de conservation. De plus, ses bénéfices reposent directement sur la volonté des décideurs de considérer la conservation de la nature sur un pied d'égalité avec les aspects socioéconomiques. À cet égard, nous voyons avant tout la PSC comme une opportunité pour synthétiser des résultats issus de la biologie avec les contraintes socioéconomiques auxquelles les gestionnaires doivent faire face. Par ailleurs, la PSC offre également une opportunité pour impliquer des chercheurs en informatique dans des problèmes qui sont à la fois stimulants et basés sur l'éthique de la conservation de la nature. En conclusion, cette thèse de doctorat a contribué à enrichir la diversité et améliorer la qualité de l'aide à la décision que peut offrir la PSC, tout en ouvrant différentes perspectives de recherche et de développement. Sans aller jusqu'à dire que nous avons contribué à de meilleures actions de conservation, nous avons proposé un ensemble d'outils qui peuvent aider à atteindre cet objectif.

BIBLIOGRAPHY

Adadi, Amina and Mohammed Berrada (2018). “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)”. In: *IEEE Access* 6, pp. 52138–52160. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2018.2870052.

Alagador, Diogo and Jorge Orestes Cerdeira (2017). “Meeting Species Persistence Targets under Climate Change: A Spatially Explicit Conservation Planning Model”. en. In: *Diversity and Distributions* 23.6, pp. 703–713. ISSN: 1472-4642. DOI: 10.1111/ddi.12562.

Albert, Cécile H., Bronwyn Rayfield, Maria Dumitru, and Andrew Gonzalez (Dec. 2017). “Applying Network Theory to Prioritize Multispecies Habitat Networks That Are Robust to Climate and Land-Use Change”. en. In: *Conservation Biology* 31.6, pp. 1383–1396. ISSN: 1523-1739. DOI: 10.1111/cobi.12943.

Amadini, Roberto, Maurizio Gabbrielli, and Jacopo Mauro (2015). “SUNNY-CP: A Sequential CP Portfolio Solver”. In: *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pp. 1861–1867.

Ardron, Jeff A, Hugh P Possingham, and Carissa J Klein (2008). “Marxan Good Practices Handbook”. In: *Pacific Marine Analysis and Research Association, Vancouver*, p. 149.

Arlettaz, Raphaël, Michael Schaub, Jérôme Fournier, Thomas S. Reichlin, Antoine Sierro, James E. M. Watson, and Veronika Braunisch (Nov. 2010). “From Publications to Public Actions: When Conservation Biologists Bridge the Gap between Research and Implementation”. en. In: *BioScience* 60.10, pp. 835–842. ISSN: 0006-3568. DOI: 10.1525/bio.2010.60.10.10.

Arora, Sanjeev and Boaz Barak (Apr. 2009). *Computational Complexity: A Modern Approach*. en. Cambridge University Press. ISBN: 978-1-139-47736-9.

Baguette, Michel, Simon Blanchet, Delphine Legrand, Virginie M. Stevens, and Camille Turlure (May 2013). “Individual Dispersal, Landscape Connectivity and Ecological Networks”. en. In: *Biological Reviews* 88.2, pp. 310–326. ISSN: 1469-185X. DOI: 10.1111/brv.12000.

Ball, Ian R, Hugh P Possingham, and Matthew E Watts (2009). “Marxan and Relatives: Software for Spatial Conservation Prioritization”. en. In: p. 12.

Ball, Ian Randall (2000). “Mathematical Applications for Conservation Ecology : The Dynamics of Tree Hollows and the Design of Nature Reserves / Ian R. Ball.” en. Thesis.

Balmford, Andrew and Richard M. Cowling (June 2006). “Fusion or Failure? The Future of Conservation Biology”. en. In: *Conservation Biology* 20.3, pp. 692–695. ISSN: 0888-8892, 1523-1739. DOI: 10.1111/j.1523-1739.2006.00434.x.

Barahona, Pedro, Ludwig Krippahl, and Olivier Perriquet (2011). “Bioinformatics: A Challenge to Constraint Programming”. In: *Hybrid Optimization*. Springer, pp. 463–487.

Bardin, Sebastian, Bernard Botella, Frédéric Dadeau, Florence Charreteur, Arnaud Gotlieb, Bruno Marre, Claude Michel, Michel Rueher, and Nicky Williams (2009). “Constraint-Based Software Testing”. In: *Journée du GDR-GPL* 9, p. 1.

Batavia, Chelsea and Michael Paul Nelson (May 2017). “For Goodness Sake! What Is Intrinsic Value and Why Should We Care?” In: *Biological Conservation* 209, pp. 366–376. ISSN: 0006-3207. DOI: `10.1016/j.biocon.2017.03.003`.

Batisse, Michel (1982). “The Biosphere Reserve: A Tool for Environmental Conservation and Management”. en. In: *Environmental Conservation* 9.2, pp. 101–111. ISSN: 1469-4387, 0376-8929. DOI: `10.1017/S0376892900019937`.

Beier, Paul, Wayne Spencer, Robert F. Baldwin, and Brad H. McRAE (Oct. 2011). “Toward Best Practices for Developing Regional Connectivity Maps”. en. In: *Conservation Biology* 25.5, pp. 879–892. ISSN: 1523-1739. DOI: `10.1111/j.1523-1739.2011.01716.x`.

Beldiceanu, Nicolas, Mats Carlsson, Jean-Xavier Rampon, and Charlotte Truchet (2005). *Graph Invariants as Necessary Conditions for Global Constraints*. eng. Swedish Institute of Computer Science.

Beldiceanu, Nicolas, Mats Carlsson, Sophie Demassey, and Thierry Petit (Sept. 2006). “Graph Properties Based Filtering”. en. In: *Principles and Practice of Constraint Programming - CP 2006*. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 59–74. ISBN: 978-3-540-46267-5 978-3-540-46268-2. DOI: `10.1007/11889205_7`.

Beldiceanu, Nicolas, Mats Carlsson, and Jean-Xavier Rampon (2012). *Global Constraint Catalog, 2nd Edition (Revision a)*. eng. Swedish Institute of Computer Science.

Berge, Claude (1973). *Graphs and Hypergraphs*. English.

Bessière, Christian, Emmanuel Hébrard, George KATSIRELOS, and Toby Walsh (July 2015). “Reasoning about Connectivity Constraints”. In: *IJCAI: International Joint Conference on Artificial Intelligence*. Buenos Aires, Argentina: AAAI Press, pp. 2568–2574.

Biere, Armin, Marijn Heule, Hans van Maaren, and Toby Walsh, eds. (2009). *Handbook of Satisfiability*. Vol. 185. Frontiers in Artificial Intelligence and Applications. IOS Press. ISBN: 978-1-58603-929-5.

Billionnet, Alain (2011). “Solving the Probabilistic Reserve Selection Problem”. In: *Ecological modelling*. 222, pp. 546–554.

Billionnet, Alain (Dec. 2013). “Mathematical Optimization Ideas for Biodiversity Conservation”. en. In: *European Journal of Operational Research* 231.3, pp. 514–534. ISSN: 03772217. DOI: `10.1016/j.ejor.2013.03.025`.

Billionnet, Alain (Apr. 2016). “Designing Connected and Compact Nature Reserves”. en. In: *Environmental Modeling & Assessment* 21.2, pp. 211–219. ISSN: 1420-2026, 1573-2967. DOI: `10.1007/s10666-015-9465-3`.

Billionnet, Alain (Dec. 2017). “How to Take into Account Uncertainty in Species Extinction Probabilities for Phylogenetic Conservation Prioritization”. en. In: *Environmental Modeling & Assessment* 22.6, pp. 535–548. ISSN: 1420-2026, 1573-2967. DOI: `10.1007/s10666-017-9561-7`.

Birch, Colin P. D., Sander P. Oom, and Jonathan A. Beecham (Aug. 2007). "Rectangular and Hexagonal Grids Used for Observation, Experiment and Simulation in Ecology". In: *Ecological Modelling* 206.3, pp. 347–359. ISSN: 0304-3800. DOI: `10.1016/j.ecolmodel.2007.03.041`.

Birnbaum, Philippe, Thomas Ibanez, Robin Pouteau, Hervé Vandrot, Vanessa Hequet, Elodie Blanchard, and Tanguy Jaffré (Jan. 2015). "Environmental Correlates for Tree Occurrences, Species Distribution and Richness on a High-Elevation Tropical Island". en. In: *AoB PLANTS* 7. DOI: `10.1093/aobpla/plv075`.

Birnbaum, Philippe et al. (2016). *Corifor, Caractérisation de Connectivités Structurelles et Fonctionnelles Des Paysages Fragmentés Sur Sols Ultramafiques*. Tech. rep.

Birnbaum, Philippe, Vanessa Hequet, Thomas Ibanez, Jérémy Girardi, Grégoire Blanchard, Robin Pouteau, Hervé Vandrot, Dimitri Justeau-Allaire, and Tristan Mangeard (2019). *Connaissances et Outils et Pour La Gestion Conservatoire Des Forêts Naturelles En Province Nord (Cogefor)*. Tech. rep.

Bisby, F. A. (Sept. 2000). "The Quiet Revolution: Biodiversity Informatics and the Internet". en. In: *Science* 289.5488, pp. 2309–2312. ISSN: 00368075, 10959203. DOI: `10.1126/science.289.5488.2309`.

Björdal, Gustav, Jean-Noël Monette, Pierre Flener, and Justin Pearson (July 2015). "A Constraint-Based Local Search Backend for MiniZinc". en. In: *Constraints* 20.3, pp. 325–345. ISSN: 1572-9354. DOI: `10.1007/s10601-015-9184-z`.

Bockmayr, Alexander, Nicolai Pisaruk, and Abderrahmane Aggoun (Nov. 2001). "Network Flow Problems in Constraint Programming". en. In: *Principles and Practice of Constraint Programming — CP 2001*. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 196–210. ISBN: 978-3-540-42863-3 978-3-540-45578-3. DOI: `10.1007/3-540-45578-7_14`.

Bodin, Örjan and Santiago Saura (Sept. 2010). "Ranking Individual Habitat Patches as Connectivity Providers: Integrating Network Analysis and Patch Removal Experiments". en. In: *Ecological Modelling* 221.19, pp. 2393–2405. ISSN: 0304-3800. DOI: `10.1016/j.ecolmodel.2010.06.017`.

Botella, Christophe, Maximilien Servajean, Pierre Bonnet, and Alexis Joly (Sept. 2019). "Overview of GeoLifeCLEF 2019: Plant Species Prediction Using Environment and Animal Occurrences". In: *Working Notes of CLEF 2019 - Conference and Labs of the Evaluation Forum*. Vol. CEUR Workshop Proceedings. Lugano, Switzerland.

Briers, Robert A. (Jan. 2002). "Incorporating Connectivity into Reserve Selection Procedures". In: *Biological Conservation* 103.1, pp. 77–83. ISSN: 0006-3207. DOI: `10.1016/S0006-3207(01)00123-9`.

Buckley, Ralf C. (2015). "Grand Challenges in Conservation Research". English. In: *Frontiers in Ecology and Evolution* 3. ISSN: 2296-701X. DOI: `10.3389/fevo.2015.00128`.

Burgman, M. A., D. B. Lindenmayer, and J. Elith (2005). "Managing Landscapes for Conservation Under Uncertainty". en. In: *Ecology* 86.8, pp. 2007–2017. ISSN: 1939-9170. DOI: `10.1890/04-0906`.

CBD, Secretariat (2010). “The Strategic Plan for Biodiversity 2011-2020 and the Aichi Biodiversity Targets”. In: *Secretariat of the Convention on Biological Diversity: Nagoya, Japan*.

Cabeza, M. and A. Moilanen (May 2001). “Design of Reserve Networks and the Persistence of Biodiversity”. eng. In: *Trends in Ecology & Evolution* 16.5, pp. 242–248. ISSN: 1872-8383.

Camm, Jeffrey D., Stephen Polasky, Andrew Solow, and Blair Csuti (Dec. 1996). “A Note on Optimal Algorithms for Reserve Site Selection”. In: *Biological Conservation* 78.3, pp. 353–355. ISSN: 0006-3207. DOI: `10.1016/0006-3207(95)00132-8`.

Cardinale, Bradley J. et al. (June 2012). “Biodiversity Loss and Its Impact on Humanity”. en. In: *Nature* 486.7401, pp. 59–67. ISSN: 0028-0836, 1476-4687. DOI: `10.1038/nature11148`.

Carlsson, Mats (2001). “Finite Domain Constraints in Sicstus Prolog”. In:

Carvalho, Sílvia B., José C. Brito, Eduardo G. Crespo, Matthew E. Watts, and Hugh P. Possingham (July 2011). “Conservation Planning under Climate Change: Toward Accounting for Uncertainty in Predicted Species Distributions to Increase Confidence in Conservation Investments in Space and Time”. en. In: *Biological Conservation* 144.7, pp. 2020–2030. ISSN: 00063207. DOI: `10.1016/j.biocon.2011.04.024`.

Chazeau, Jean (1993). “Research on New Caledonian Terrestrial Fauna: Achievements and Prospects”. In: *Biodiversity letters*, pp. 123–129.

Cheng, Hung-Chih, Pierre-Alexandre Château, and Yang-Chi Chang (May 2015). “Spatial Zoning Design for Marine Protected Areas through Multi-Objective Decision-Making”. In: *Ocean & Coastal Management*. Estuaries and Coastal Areas in Times of Intense Change 108, pp. 158–165. ISSN: 0964-5691. DOI: `10.1016/j.ocecoaman.2014.08.018`.

Chu, Geoffrey, PJ Stuckey, A Schutt, T Ehlers, G Gange, and K Francis (2016). “Chuffed, a Lazy Clause Generation Solver”. In:

Church, Richard L., David M. Stoms, and Frank W. Davis (Jan. 1996). “Reserve Selection as a Maximal Covering Location Problem”. In: *Biological Conservation* 76.2, pp. 105–112. ISSN: 0006-3207. DOI: `10.1016/0006-3207(95)00102-6`.

Ciarleglio, Michael Ian (May 2008). “Modular Abstract Self-Learning Tabu Search (MASTS) : Metaheuristic Search Theory and Practice”. eng. In:

Ciarleglio, Michael, J. Wesley Barnes, and Sahotra Sarkar (2009). “ConsNet: New Software for the Selection of Conservation Area Networks with Spatial and Multi-Criteria Analyses”. en. In: *Ecography* 32.2, pp. 205–209. ISSN: 1600-0587. DOI: `10.1111/j.1600-0587.2008.05721.x`.

Ciarleglio, Michael, J. Wesley Barnes, and Sahotra Sarkar (Aug. 2010). “ConsNet—A Tabu Search Approach to the Spatially Coherent Conservation Area Network Design Problem”. en. In: *Journal of Heuristics* 16.4, pp. 537–557. ISSN: 1381-1231, 1572-9397. DOI: `10.1007/s10732-008-9098-7`.

Citré, B. et al. (2019). *Code de l’environnement de La Province Des Îles Loyauté*. Province des îles Loyauté.

Crook, John Hurrell (1965). *The Present Status of Certain Rare Land Birds of the Seychelles Islands*.

Daigle, Rémi M., Anna Metaxas, Arieanna C. Balbar, Jennifer McGowan, Eric A. Treml, Caitlin D. Kuempel, Hugh P. Possingham, and Maria Beger (2020). "Operationalizing Ecological Connectivity in Spatial Conservation Planning with Marxan Connect". en. In: *Methods in Ecology and Evolution* 11.4, pp. 570–579. ISSN: 2041-210X. DOI: `10.1111/2041-210X.13349`.

De Garine-Wichatitsky, Michel, Yohann Soubeyran, Daniel Maillard, and Patrick Duncan (2005). "The Diets of Introduced Rusa Deer (Cervus Timorensis Russa) in a Native Sclerophyll Forest and a Native Rainforest of New Caledonia". eng. In: *New Zealand Journal of Zoology*. ISSN: 0301-4223.

De Landtsheer, Renaud and Christophe Ponsard (2013). "OscaR. Cbls: An Open Source Framework for Constraint-Based Local Search". In: *Proceedings of ORBEL* 27.

Dechter, Rina et al. (2003). *Constraint Processing*. Morgan Kaufmann.

Delmelle, Eric, Michael R. Desjardins, and Jing Deng (Dec. 2017). "Designing Spatially Cohesive Nature Reserves with Backup Coverage". In: *International Journal of Geographical Information Science* 31.12, pp. 2505–2523. ISSN: 1365-8816. DOI: `10.1080/13658816.2017.1357820`.

Dengler, Jürgen, Jörg Ewald, Ingolf Kühn, and Robert K. Peet (2011). "Ecoinformatics and Global Change – an Overdue Liaison". en. In: *Journal of Vegetation Science* 22.4, pp. 577–581. ISSN: 1654-1103. DOI: `10.1111/j.1654-1103.2011.01313.x`.

Diamond, Jared M. (Feb. 1975). "The Island Dilemma: Lessons of Modern Biogeographic Studies for the Design of Natural Reserves". In: *Biological Conservation* 7.2, pp. 129–146. ISSN: 0006-3207. DOI: `10.1016/0006-3207(75)90052-X`.

Díaz, S. et al. (May 2020). "Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services". en. In:

Díaz, Sandra et al. (Jan. 2018). "Assessing Nature's Contributions to People". en. In: *Science* 359.6373, pp. 270–272. ISSN: 0036-8075, 1095-9203. DOI: `10.1126/science.aap8826`.

Diestel, Reinhard (1997). *Graph Theory*. Two thousand, fifth.

Dilkina, Bistra, Rachel Houtman, Carla P. Gomes, Claire A. Montgomery, Kevin S. McKelvey, Katherine Kendall, Tabitha A. Graves, Richard Bernstein, and Michael K. Schwartz (Feb. 2017). "Trade-offs and efficiencies in optimal budget-constrained multispecies corridor networks". es. In: *Conservation Biology* 31.1, pp. 192–202. ISSN: 1523-1739. DOI: `10.1111/cobi.12814`.

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

Downing, Nicholas, Thibaut Feydy, and Peter J. Stuckey (May 2012). "Explaining Flow-Based Propagation". en. In: *Integration of AI and OR Techniques in Contraint Programming for Combinatorial Optimzation Problems*. Lecture

Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 146–162. ISBN: 978-3-642-29827-1 978-3-642-29828-8. DOI: 10.1007/978-3-642-29828-8_10.

Doxiadis, Apostolos and Christos Papadimitriou (July 2015). *Logicomix: An Epic Search for Truth*. en. Bloomsbury Publishing USA. ISBN: 978-1-63286-480-2.

Dubois, Didier, Petr Hájek, and Henri Prade (Jan. 2000). "Knowledge-Driven versus Data-Driven Logics". en. In: *Journal of Logic, Language and Information* 9.1, pp. 65–89. ISSN: 1572-9583. DOI: 10.1023/A:1008370109997.

Dudley, Nigel (2008). *Guidelines for Applying Protected Area Management Categories*. en. IUCN. ISBN: 978-2-8317-1086-0. DOI: 10.2305/IUCN.CH.2008.PAPS.2.en.

Elith, Jane and John R. Leathwick (2009). "Species Distribution Models: Ecological Explanation and Prediction Across Space and Time". In: *Annual Review of Ecology, Evolution, and Systematics* 40.1, pp. 677–697. DOI: 10.1146/annurev.ecolsys.110308.120159.

Ellison, Aaron M. (Oct. 2016). "It's Time to Get Real about Conservation". en. In: *Nature News* 538.7624, p. 141. DOI: 10.1038/538141a.

Endemia and RLA Flore NC (2017). *La Liste Rouge de La Fore Menacée de Nouvelle-Calédonie, Synthèse Octobre 2017*. Tech. rep.

Erskine, Peter D., Gavin Lee, Bruno Fogliani, Laurent L'Huillier, and Stéphane McCoy (2018). "Incorporating Hyperaccumulator Plants into Mine Rehabilitation in the Asia-Pacific Region". en. In: *Agromining: Farming for Metals: Extracting Unconventional Resources Using Plants*. Ed. by Antony Van der Ent, Guillaume Echevarria, Alan J.M. Baker, and Jean Louis Morel. Mineral Resource Reviews. Cham: Springer International Publishing, pp. 189–204. ISBN: 978-3-319-61899-9. DOI: 10.1007/978-3-319-61899-9_10.

Etienne, R. S. and J. A. Heesterbeek (Mar. 2000). "On Optimal Size and Number of Reserves for Metapopulation Persistence". eng. In: *Journal of Theoretical Biology* 203.1, pp. 33–50. ISSN: 0022-5193. DOI: 10.1006/jtbi.1999.1060.

Euler, Leonhard (1741). "Solutio Problematis Ad Geometriam Situs Pertinentis". In: *Commentarii academiae scientiarum Petropolitanae*, pp. 128–140.

Fages, Jean-Guillaume (Oct. 2014). "Exploitation de structures de graphe en programmation par contraintes". fr. PhD thesis. Ecole des Mines de Nantes.

Fages, Jean-Guillaume, Gilles Chabert, and Charles Prud'Homme (Sept. 2013). "Combining Finite and Continuous Solvers Towards a Simpler Solver Maintenance". In: *The 19th International Conference on Principles and Practice of Constraint Programming*. Uppsala, Sweden, TRICS'13 Workshop: Techniques foR Implementing Constraint programming Systems.

Fages, Jean-Guillaume, Charles Prud'homme, and Xavier Lorca (Feb. 2018). *Choco Graph Documentation*.

Fahrig, Lenore (2003). "Effects of Habitat Fragmentation on Biodiversity". In: *Annual Review of Ecology, Evolution, and Systematics* 34.1, pp. 487–515. DOI: 10.1146/annurev.ecolsys.34.011802.132419.

Fahrig, Lenore (2013). "Rethinking Patch Size and Isolation Effects: The Habitat Amount Hypothesis". en. In: *Journal of Biogeography* 40.9, pp. 1649–1663. ISSN: 1365-2699. DOI: 10.1111/jbi.12130.

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

Ferrier, Simon, Robert L. Pressey, and Thomas W. Barrett (May 2000). “A New Predictor of the Irreplaceability of Areas for Achieving a Conservation Goal, Its Application to Real-World Planning, and a Research Agenda for Further Refinement”. en. In: *Biological Conservation* 93.3, pp. 303–325. ISSN: 0006-3207. DOI: `10.1016/S0006-3207(99)00149-4`.

Field, null, null Behrenfeld, null Randerson, and null Falkowski (July 1998). “Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components”. eng. In: *Science (New York, N.Y.)* 281.5374, pp. 237–240. ISSN: 1095-9203. DOI: `10.1126/science.281.5374.237`.

Freuder, Eugene (1996). “In Pursuit of the Holy Grail”. In: *ACM Computing Surveys (CSUR)* 28.4es, 63–es.

Fuller, Trevon and Sahotra Sarkar (May 2006). “LQGraph: A Software Package for Optimizing Connectivity in Conservation Planning”. en. In: *Environmental Modelling & Software* 21.5, pp. 750–755. ISSN: 1364-8152. DOI: `10.1016/j.envsoft.2006.01.005`.

Gage, John D., J. Murray Roberts, John P. Hartley, and John D. Humphery (Sept. 2005). “Potential Impacts of Deep-Sea Trawling on the Benthic Ecosystem along the Northern European Continental Margin: A Review”. English. In: *Benthic Habitats and the Effects of Fishing*. American Fisheries Society, pp. 503–517.

Game, Edward T., Mark W. Schwartz, and Andrew T. Knight (2015). “Policy Relevant Conservation Science”. en. In: *Conservation Letters* 8.5, pp. 309–311. ISSN: 1755-263X. DOI: `10.1111/conl.12207`.

Genesereth, Michael, Nathaniel Love, and Barney Pell (June 2005). “General Game Playing: Overview of the AAAI Competition”. en. In: *AI Magazine* 26.2, pp. 62–62. ISSN: 2371-9621. DOI: `10.1609/aimag.v26i2.1813`.

Gerault, David, Marine Minier, and Christine Solnon (2016). “Constraint Programming Models for Chosen Key Differential Cryptanalysis”. In: *International Conference on Principles and Practice of Constraint Programming*. Springer, pp. 584–601.

Gervet, Carmen (Sept. 1995). “Set Intervals in Constraint Logic Programming: Definition and Implementation of a Language”. en. PhD thesis. Université de Franche Comté Besançon.

Glover, Fred (Aug. 1989). “Tabu Search—Part I”. In: *ORSA Journal on Computing* 1.3, pp. 190–206. ISSN: 0899-1499. DOI: `10.1287/ijoc.1.3.190`.

Glover, Fred (Feb. 1990). “Tabu Search—Part II”. In: *ORSA Journal on Computing* 2.1, pp. 4–32. ISSN: 0899-1499. DOI: `10.1287/ijoc.2.1.4`.

Goëau, Hervé, Pierre Bonnet, and Alexis Joly (Sept. 2016). “Plant Identification in an Open-World (LifeCLEF 2016)”. In: *CLEF: Conference and Labs of the Evaluation Forum*. Vol. CEUR Workshop Proceedings. Évora, Portugal, pp. 428–439.

Gomes, Carla P. (2008). “Computational Sustainability : Computational Methods for a Sustainable Environment , Economy , and Society Optimal Forest

Fire Fuel Management and Timber Harvest In The Face Of Endogenous Spatial Risk”. In:

Gomes, Carla P. (2009). “Challenges for Constraint Reasoning and Optimization in Computational Sustainability”. en. In: *Principles and Practice of Constraint Programming - CP 2009*. Ed. by Ian P. Gent. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 2–4. ISBN: 978-3-642-04244-7.

Grandcolas, Philippe, Jérôme Murienne, Tony Robillard, Laure Desutter-Grandcolas, Hervé Jourdan, Eric Guilbert, and Louis Deharveng (Oct. 2008). “New Caledonia: A Very Old Darwinian Island?” In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 363.1508, pp. 3309–3317. ISSN: 0962-8436. DOI: 10.1098/rstb.2008.0122.

Groombridge, B. and M. D. Jenkins (2002). “World Atlas of Biodiversity: Earth’s Living Resources in the 21st Century.” English. In: *World atlas of biodiversity: earth’s living resources in the 21st century.*

Guillemot, N., T. Gaillard, and A. Lagrange (2016). *Biodiversité et Environnement Dans La Région de La Côte Oubliée « Woen Vùù » : État Des Connaissances et Identification Des Intérêts Écologiques | Oeil Nouvelle-Calédonie*. Tech. rep. OEIL.

Guisan, Antoine and Niklaus E. Zimmermann (Dec. 2000). “Predictive Habitat Distribution Models in Ecology”. In: *Ecological Modelling* 135.2, pp. 147–186. ISSN: 0304-3800. DOI: 10.1016/S0304-3800(00)00354-9.

Haddad, Nick M. et al. (Mar. 2015). “Habitat Fragmentation and Its Lasting Impact on Earths Ecosystems”. en. In: *Science Advances* 1.2, e1500052. ISSN: 2375-2548. DOI: 10.1126/sciadv.1500052.

Haider, Zulqarnain, Hadi Charkhgard, and Changhyun Kwon (Jan. 2018). “A Robust Optimization Approach for Solving Problems in Conservation Planning”. In: *Ecological Modelling* 368, pp. 288–297. ISSN: 0304-3800. DOI: 10.1016/j.ecolmodel.2017.12.006.

Haight, Robert G., Charles S. Revelle, and Stephanie A. Snyder (Oct. 2000). “An Integer Optimization Approach to a Probabilistic Reserve Site Selection Problem”. In: *Operations Research* 48.5, pp. 697–708. ISSN: 0030-364X. DOI: 10.1287/opre.48.5.697.12411.

Hanson, Jeffrey O., Jonathan R. Rhodes, Hugh P. Possingham, and Richard A. Fuller (2017). “Raptr: Representative and Adequate Prioritization Toolkit in R”. en. In: *Methods in Ecology and Evolution* 9.2, pp. 320–330. ISSN: 2041-210X. DOI: 10.1111/2041-210X.12862.

Hanson, Jeffrey O., Richard Schuster, Matthew Strimas-Mackey, and Joseph R. Bennett (2019a). “Optimality in Prioritizing Conservation Projects”. en. In: *Methods in Ecology and Evolution* 10.10, pp. 1655–1663. ISSN: 2041-210X. DOI: 10.1111/2041-210X.13264.

Hanson, Jeffrey O., Richard Schuster, Nina Morrell, Matthew Strimas-Mackey, Matthew E. Watts, Peter Arcese, Joseph Bennett, and Hugh P. Possingham (Oct. 2019b). *Prioritizr: Systematic Conservation Prioritization in R.*

Hanson, Jeffrey O., Richard Schuster, Nina Morrell, Matthew Strimas-Mackey, Matthew E. Watts, Peter Arcese, Joseph Bennett, and Hugh P. Possingham (May 2020). *Prioritizr: Systematic Conservation Prioritization in R.*

Harris, Larry D. (1988). “Edge Effects and Conservation of Biotic Diversity”. en. In: *Conservation Biology* 2.4, pp. 330–332. ISSN: 1523-1739. DOI: `10.1111/j.1523-1739.1988.tb00196.x`.

Hebrard, E and M Siala (2016). “Mistral 2.0”. In:

Hély-Alleaune, Christell (2012). *Incendies et Biodiversité Des Écosystèmes En Nouvelle-Calédonie - Compte Rendu de Fin de Projet | Oeil Nouvelle-Calédonie*. Tech. rep.

Hentenryck, Pascal Van and Laurent Michel (2009). *Constraint-Based Local Search*. The MIT press.

Higgs, A. J. (1981). “Island Biogeography Theory and Nature Reserve Design”. In: *Journal of Biogeography* 8.2, pp. 117–124. ISSN: 0305-0270. DOI: `10.2307/2844554`.

Hof, John (1998). “Optimizing Spatial and Dynamic Population-Based Control Strategies for Invading Forest Pests”. en. In: *Natural Resource Modeling* 11.3, pp. 197–216. ISSN: 1939-7445. DOI: `10.1111/j.1939-7445.1998.tb00308.x`.

Hof, John and Curtis H. Flather (1996). “Accounting for Connectivity and Spatial Correlation in the Optimal Placement of Wildlife Habitat”. en. In: *Ecological Modelling. 88(1-3): 143-155*.

Hooker, John N (2016). “Finding Alternative Musical Scales”. In: *International Conference on Principles and Practice of Constraint Programming*. Springer, pp. 753–768.

IUCN (Nov. 2009). *Parasitaxus Usta: Thomas, P.: The IUCN Red List of Threatened Species 2010: E.T31002A9597883*. en. DOI: `10.2305/IUCN.UK.2010-3.RLTS.T31002A9597883.en`.

IUCN (Oct. 2016). *Acrocephalus Sechellensis: BirdLife International: The IUCN Red List of Threatened Species 2016: E.T22714882A94431883*. en. DOI: `10.2305/IUCN.UK.2016-3.RLTS.T22714882A94431883.en`.

IUCN (2020). *The IUCN Red List of Threatened Species. Version 2020-1. Downloaded on 19 March 2020.*

Ibanez, Thomas, Jérôme Munzinger, Gilles Dagostini, Vanessa Hequet, Frédéric Rigault, Tanguy Jaffré, and Philippe Birnbaum (2014). “Structural and Floristic Diversity of Mixed Tropical Rain Forest in New Caledonia: New Data from the New Caledonian Plant Inventory and Permanent Plot Network (NC-PIPPN)”. en. In: *Applied Vegetation Science* 17.3, pp. 386–397. ISSN: 1654-109X. DOI: `10.1111/avsc.12070`.

Ibanez, Thomas, Vanessa Hequet, Céline Chambrey, Tanguy Jaffré, and Philippe Birnbaum (Aug. 2017). “How Does Forest Fragmentation Affect Tree Communities? A Critical Case Study in the Biodiversity Hotspot of New Caledonia”. en. In: *Landscape Ecology* 32.8, pp. 1671–1687. ISSN: 0921-2973, 1572-9761. DOI: `10.1007/s10980-017-0534-7`.

Ibanez, Thomas, Philippe Birnbaum, Gildas Gâteblé, Vanessa Hequet, Sandrine Isnard, Jérôme Munzinger, Yohan Pillon, Robin Pouteau, Hervé Vandrot, and Tanguy Jaffré (Nov. 2018). “Twenty Years after Jaffré et al. (1998), Is the System of Protected Areas Now Adequate in New Caledonia?” en. In: *Biodiversity and Conservation*, pp. 1–10. ISSN: 0960-3115, 1572-9710. DOI: `10.1007/s10531-018-1659-y`.

Ignatiev, Alexey, Filipe Pereira, Nina Narodytska, and Joao Marques-Silva (2018). "A SAT-Based Approach to Learn Explainable Decision Sets". en. In: *Automated Reasoning*. Ed. by Didier Galmiche, Stephan Schulz, and Roberto Sebastiani. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 627–645. ISBN: 978-3-319-94205-6. DOI: 10.1007/978-3-319-94205-6_41.

Isnard, Sandrine, Laurent L'huillier, Frédéric Rigault, and Tanguy Jaffré (June 2016). "How Did the Ultramafic Soils Shape the Flora of the New Caledonian Hotspot?" en. In: *Plant and Soil* 403.1, pp. 53–76. ISSN: 1573-5036. DOI: 10.1007/s11104-016-2910-5.

Jaeger, Jochen A.G. (Feb. 2000). "Landscape Division, Splitting Index, and Effective Mesh Size: New Measures of Landscape Fragmentation". en. In: *Landscape Ecology* 15.2, pp. 115–130. ISSN: 1572-9761. DOI: 10.1023/A:1008129329289.

Jafari, Nahid, Bryan L. Nuse, Clinton T. Moore, Bistra Dilkina, and Jeffrey Hepinstall-Cymerman (May 2017). "Achieving Full Connectivity of Sites in the Multiperiod Reserve Network Design Problem". In: *Computers & Operations Research* 81, pp. 119–127. ISSN: 0305-0548. DOI: 10.1016/j.cor.2016.12.017.

Jaffré, T, P Morat, and JM Veillon (1994). "La Flore de Nouvelle-Calédonie. Caractéristiques et Composition Floristique Des Principales Formations Végétales". In: *Bois et Forêts des Tropiques* 242, pp. 8–28.

Jaffre, TANGUY, PHILIPPE Bouchet, and JEAN-MARIE Veillon (Jan. 1998). "Threatened Plants of New Caledonia: Is the System of Protected Areas Adequate?" en. In: *Biodiversity and Conservation* 7.1, pp. 109–135. ISSN: 1572-9710. DOI: 10.1023/A:1008815930865.

Jaffré, Tanguy, Frédéric Rigault, and Jérôme Muzinger (2012). "La végétation". fre. In: *Atlas de la Nouvelle Calédonie*. Ed. by Jacques Bonvallot, Jean-Christophe Gay, and Elisabeth Habert. Marseille: IRD, pp. 77–80.

Jospin, Lionel (1998). *Déclaration de M. Lionel Jospin, Premier Ministre, Sur Le Bilan Des Accords de Matignon, La Négociation Sur l'avenir de La Nouvelle-Calédonie, La Signature de l'accord de Nouméa Du 21 Avril Sur Les Principes de l'organisation Politique et Sociale de l'île, et Le Suivi de l'application de l'accord, Nouméa Le 5 Mai 1998.*

Jussien, Narendra and Samir Ouis (Nov. 2001). "User-Friendly Explanations for Constraint Programming". In: *arXiv:cs/0111037*. arXiv: cs/0111037.

Justeau-Allaire, Dimitri, Philippe Birnbaum, and Xavier Lorca (2018). "Unifying Reserve Design Strategies with Graph Theory and Constraint Programming". en. In: *Principles and Practice of Constraint Programming*. Ed. by John Hooker. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 507–523. ISBN: 978-3-319-98334-9. DOI: 10.1007/978-3-319-98334-9_33.

Justeau-Allaire, Dimitri, Philippe Vismara, Philippe Birnbaum, and Xavier Lorca (2019a). "Systematic Conservation Planning for Sustainable Land-Use Policies: A Constrained Partitioning Approach to Reserve Selection and Design." en. In: *Proceedings of the Twenty-Eighth International Joint Conference* *on Artificial Intelligence*. Macao, China: International Joint Conferences on Artificial Intelligence Organization, pp. 5902–5908. ISBN: 978-0-9992411-4-1. DOI: `10.24963/ijcai.2019/818`.

Justeau-Allaire, Dimitri, N. Rink, Xavier Lorca, E. Couture, and Philippe Birnbaum (2019b). “Systematic Conservation Planning in New Caledonia: Supporting Sustainable Land-Use Policies with Reserve Selection Models”. In: *Island Biology 2019,* Saint-Denis, Réunion: Dominique STRASBERG and Claudine AH-PENG.

Justeau-Allaire, Dimitri, Philippe Birnbaum, and Xavier Lorca (June 2019c). “Unifier Les Stratégies de Sélection de Réserve Avec La Programmation Par Contraintes et Les Graphes”. In: *JFPC 2019 - Actes Des 15es Journees Francophones de Programmation Par Contraintes*. Albi, France: IMT Mines Albi, p. 27–29.

Justeau-Allaire, Dimitri, Ghislain Vieilledent, Nicolas Rinck, Philippe Vismara, Xavier Lorca, and Philippe Birnbaum (2020). “Constrained Optimization of Landscape Indices in Conservation Planning to Support Ecological Restoration in New Caledonia”. In: *Journal of Applied Ecology*. ISSN: 1365-2664. DOI: `10.1111/1365-2664.13803`.

Justeau-Allaire, Dimitri, Philippe Vismara, Philippe Birnbaum, and Xavier Lorca (2020). “Partitionnement de l’espace Sous Contraintes: Un Modèle Générique et Expressif Pour La Planification de La Conservation.” In: *ROADEF 2020 - Actes du 21ème congrès annuel de la société Française de Recherche Opérationnelle et d’Aide à la Décision*.

Kareiva, Peter (May 2001). “Ecoinformatics: Facilitating Access to Existing Data Sets”. English. In: *Trends in Ecology & Evolution* 16.5, p. 226. ISSN: 0169-5347. DOI: `10.1016/S0169-5347(01)02176-0`.

Keenan, Rodney J., Gregory A. Reams, Frédéric Achard, Joberto V. de Freitas, Alan Grainger, and Erik Lindquist (Sept. 2015). “Dynamics of Global Forest Area: Results from the FAO Global Forest Resources Assessment 2015”. en. In: *Forest Ecology and Management*. Changes in Global Forest Resources from 1990 to 2015 352, pp. 9–20. ISSN: 0378-1127. DOI: `10.1016/j.foreco.2015.06.014`.

Kerdprasop, Nittaya and Kittisak Kerdprasop (2015). “Constraint-Based System for Genomic Analysis”. In: *International Journal of Information and Education Technology* 5.2, p. 119.

Khalil, Elias, Hanjun Dai, Yuyu Zhang, Bistra Dilkina, and Le Song (2017). “Learning Combinatorial Optimization Algorithms over Graphs”. In: *Advances in Neural Information Processing Systems*, pp. 6348–6358.

Kirkpatrick, J. B. (Feb. 1983). “An Iterative Method for Establishing Priorities for the Selection of Nature Reserves: An Example from Tasmania”. In: *Biological Conservation* 25.2, pp. 127–134. ISSN: 0006-3207. DOI: `10.1016/0006-3207(83)90056-3`.

Knight, Andrew T., Richard M. Cowling, Mathieu Rouget, Andrew Balmford, Amanda T. Lombard, and Bruce M. Campbell (2008). “Knowing But Not Doing: Selecting Priority Conservation Areas and the Research–Implementation

Gap". en. In: *Conservation Biology* 22.3, pp. 610–617. ISSN: 1523-1739. DOI: 10.1111/j.1523-1739.2008.00914.x.

Kuchcinski, Krzysztof and Radoslaw Szymanek (2013). "JaCoP - Java Constraint Programming Solver". English. In: *CP Solvers: Modeling, Applications, Integration, and Standardization, Co-Located with the 19th International Conference on Principles and Practice of Constraint Programming*.

Kukkala, Aija S. and Atte Moilanen (May 2013). "Core Concepts of Spatial Prioritisation in Systematic Conservation Planning". eng. In: *Biological Reviews of the Cambridge Philosophical Society* 88.2, pp. 443–464. ISSN: 1469-185X. DOI: 10.1111/brv.12008.

Lauriere, Jena-Lonis (Feb. 1978). "A Language and a Program for Stating and Solving Combinatorial Problems". en. In: *Artificial Intelligence* 10.1, pp. 29–127. ISSN: 0004-3702. DOI: 10.1016/0004-3702(78)90029-2.

Lecoutre, Christophe and Olivier Roussel (Dec. 2018). "Proceedings of the 2018 XCSP3 Competition". In: *arXiv:1901.01830 [cs]*. arXiv: 1901.01830 [cs].

Lin, Yu-Pin, Chi-Ju Chen, Wan-Yu Lien, Wen-Hao Chang, Joy R. Petway, and Li-Chi Chiang (Jan. 2019). "Landscape Conservation Planning to Sustain Ecosystem Services under Climate Change". en. In: *Sustainability* 11.5, p. 1393. DOI: 10.3390/su11051393.

Lombardi, Michele, Michela Milano, and Andrea Bartolini (Mar. 2017). "Empirical Decision Model Learning". en. In: *Artificial Intelligence*. Combining Constraint Solving with Mining and Learning 244, pp. 343–367. ISSN: 0004-3702. DOI: 10.1016/j.artint.2016.01.005.

MacArthur, Robert H. and Edward O. Wilson (1967). *The Theory of Island Biogeography*. REV - Revised. Princeton University Press. ISBN: 978-0-691-08836-5.

Mackworth, Alan K. (Feb. 1977). "Consistency in Networks of Relations". en. In: *Artificial Intelligence* 8.1, pp. 99–118. ISSN: 0004-3702. DOI: 10.1016/0004-3702(77)90007-8.

Maitrepierre, L (2012). "Les Types de Temps et Les Cyclones, Les Éléments Du Climat". In: *Atlas de la Nouvelle-Calédonie. Marseille-Nouméa: IRD-Congrès de la Nouvelle-Calédonie*, pp. 53–60.

Margules, C. R., A. O. Nicholls, and R. L. Pressey (Jan. 1988). "Selecting Networks of Reserves to Maximise Biological Diversity". In: *Biological Conservation* 43.1, pp. 63–76. ISSN: 0006-3207. DOI: 10.1016/0006-3207(88)90078-X.

Margules, C. and M.B. Usher (Oct. 1981). "Criteria Used in Assessing Wildlife Conservation Potential: A Review". en. In: *Biological Conservation* 21.2, pp. 79–109. ISSN: 00063207. DOI: 10.1016/0006-3207(81)90073-2.

Margules, C., A. J. Higgs, and R. W. Rafe (Oct. 1982). "Modern Biogeographic Theory: Are There Any Lessons for Nature Reserve Design?" In: *Biological Conservation* 24.2, pp. 115–128. ISSN: 0006-3207. DOI: 10.1016/0006-3207(82)90063-5.

Margules, Chris R. and Robert L. Pressey (May 2000). "Systematic Conservation Planning". eng. In: *Nature* 405.6783, pp. 243–253. ISSN: 0028-0836. DOI: 10.1038/35012251.

Matisziw, Timothy C. and Alan T. Murray (Sept. 2006). "Promoting Species Persistence through Spatial Association Optimization in Nature Reserve Design". en. In: *Journal of Geographical Systems* 8.3, pp. 289–305. ISSN: 1435-5949. DOI: `10.1007/s10109-006-0020-2`.

McDonnell, Mark D., Hugh P. Possingham, Ian R. Ball, and Elizabeth A. Cousins (June 2002). "Mathematical Methods for Spatially Cohesive Reserve Design". en. In: *Environmental Modeling & Assessment* 7.2, pp. 107–114. ISSN: 1420-2026, 1573-2967. DOI: `10.1023/A:1015649716111`.

McGarigal, Kevin (2014). "Landscape Pattern Metrics". en. In: *Wiley StatsRef: Statistics Reference Online*. American Cancer Society. ISBN: 978-1-118-44511-2. DOI: `10.1002/9781118445112.stat07723`.

McRae, Brad H., Sonia A. Hall, Paul Beier, and David M. Theobald (Dec. 2012). "Where to Restore Ecological Connectivity? Detecting Barriers and Quantifying Restoration Benefits". en. In: *PLOS ONE* 7.12, e52604. ISSN: 1932-6203. DOI: `10.1371/journal.pone.0052604`.

Meffe, Gary K. (2006). "The Success-and Challenges-of "Conservation Biology"". In: *Conservation Biology* 20.4, pp. 931–933. ISSN: 0888-8892.

Meir, Eli, Sandy Andelman, and Hugh P. Possingham (Aug. 2004). "Does Conservation Planning Matter in a Dynamic and Uncertain World?" en. In: *Ecology Letters* 7.8, pp. 615–622. ISSN: 1461-023X, 1461-0248. DOI: `10.1111/j.1461-0248.2004.00624.x`.

Moilanen, A, FM Pouzols, L Meller, V Veach, A Arponen, J Leppänen, and H Kujala (2014). "Zonation–Spatial Conservation Planning Methods and Software. Version 4. User Manual". In: *University of Helsinki: Finland*.

Moilanen, Atte (Feb. 2007). "Landscape Zonation, Benefit Functions and Target-Based Planning: Unifying Reserve Selection Strategies". In: *Biological Conservation* 134.4, pp. 571–579. ISSN: 0006-3207. DOI: `10.1016/j.biocon.2006.09.008`.

Moilanen, Atte (July 2008). "Two Paths to a Suboptimal Solution – Once More about Optimality in Reserve Selection". en. In: *Biological Conservation* 141.7, pp. 1919–1923. ISSN: 0006-3207. DOI: `10.1016/j.biocon.2008.04.018`.

Moilanen, Atte, Kerrie A. Wilson, and Hugh Possingham, eds. (May 2009a). *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools*. Oxford, New York: Oxford University Press. ISBN: 978-0-19-954776-0.

Moilanen, Atte, Heini Kujala, and John R Leathwick (2009b). "The Zonation Framework and Software for Conservation Prioritization". In: *Spatial conservation prioritization* 135, pp. 196–210.

Montanari, Ugo (Jan. 1974). "Networks of Constraints: Fundamental Properties and Applications to Picture Processing". en. In: *Information Sciences* 7, pp. 95–132. ISSN: 0020-0255. DOI: `10.1016/0020-0255(74)90008-5`.

Morat, P. et al. (2012). "The Taxonomic Reference Base Florical and Characteristics of the Native Vascular Flora of New Caledonia." en. In: *Adansonia* 34.2, pp. 179–221. ISSN: 1280-8571.

Moser, Brigitte, Jochen A. G. Jaeger, Ulrike Tappeiner, Erich Tasser, and Beatrice Eiselt (Mar. 2007). "Modification of the Effective Mesh Size for Measuring Landscape Fragmentation to Solve the Boundary Problem". en. In:

Landscape Ecology 22.3, pp. 447–459. ISSN: 1572-9761. DOI: 10.1007/s10980-006-9023-0.

Moulton, Vincent, Charles Semple, and Mike Steel (May 2007). "Optimizing Phylogenetic Diversity under Constraints". eng. In: *Journal of Theoretical Biology* 246.1, pp. 186–194. ISSN: 0022-5193. DOI: 10.1016/j.jtbi.2006.12.021.

Myers, N. (1988). "Threatened Biotas: "Hot Spots" in Tropical Forests". eng. In: *The Environmentalist* 8.3, pp. 187–208. ISSN: 0251-1088. DOI: 10.1007/BF02240252.

Myers, Norman, Russell A. Mittermeier, Cristina G. Mittermeier, Gustavo A. B. da Fonseca, and Jennifer Kent (Feb. 2000). "Biodiversity Hotspots for Conservation Priorities". en. In: *Nature* 403.6772, pp. 853–858. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/35002501.

Nethercote, Nicholas, Peter J. Stuckey, Ralph Becket, Sebastian Brand, Gregory J. Duck, and Guido Tack (2007). "MiniZinc: Towards a Standard CP Modelling Language". en. In: *Principles and Practice of Constraint Programming – CP 2007*. Ed. by Christian Bessière. Vol. 4741. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 529–543. ISBN: 978-3-540-74969-1. DOI: 10.1007/978-3-540-74970-7_38.

Nicholls, A. O. and C. R. Margules (Jan. 1993). "An Upgraded Reserve Selection Algorithm". In: *Biological Conservation* 64.2, pp. 165–169. ISSN: 0006-3207. DOI: 10.1016/0006-3207(93)90654-J.

Önal, Hayri, Yicheng Wang, Sahan T. M. Dissanayake, and James D. Westervelt (June 2016). "Optimal Design of Compact and Functionally Contiguous Conservation Management Areas". In: *European Journal of Operational Research* 251.3, pp. 957–968. ISSN: 0377-2217. DOI: 10.1016/j.ejor.2015.12.005.

Pascual-Hortal, Lucía and Santiago Saura (Oct. 2006). "Comparison and Development of New Graph-Based Landscape Connectivity Indices: Towards the Priorization of Habitat Patches and Corridors for Conservation". en. In: *Landscape Ecology* 21.7, pp. 959–967. ISSN: 1572-9761. DOI: 10.1007/s10980-006-0013-z.

Perifel, Sylvain (2014). *Complexité Algorithmique*.

Perron, Laurent and Vincent Furnon (July 2019). *OR-Tools*. Google.

Pillon, Yohan (2012). "Time and Tempo of Diversification in the Flora of New Caledonia". en. In: *Botanical Journal of the Linnean Society* 170.3, pp. 288–298. ISSN: 1095-8339. DOI: 10.1111/j.1095-8339.2012.01274.x.

Pillon, Yohan, Laure Barrabé, and Sven Buerki (Feb. 2017). "How Many Genera of Vascular Plants Are Endemic to New Caledonia? A Critical Review Based on Phylogenetic Evidence". en. In: *Botanical Journal of the Linnean Society* 183.2, pp. 177–198. ISSN: 0024-4074. DOI: 10.1093/botlinnean/bow001.

Polasky, Stephen, Jeffrey D. Camm, Andrew R. Solow, Blair Csuti, Denis White, and Rugang Ding (June 2000). "Choosing Reserve Networks with Incomplete Species Information". In: *Biological Conservation* 94.1, pp. 1–10. ISSN: 0006-3207. DOI: 10.1016/S0006-3207(99)00171-8.

Possingham, Hugh, J R. Day, M Goldfinch, and F Salzborn (Jan. 1993). “The Mathematics of Designing a Network of Protected Ares for Conservation”. In:

Pouteau, Robin, Élise Bayle, Élodie Blanchard, Philippe Birnbaum, Jean-Jérôme Cassan, Vanessa Hequet, Thomas Ibanez, and Hervé Vandrot (2015). “Accounting for the Indirect Area Effect in Stacked Species Distribution Models to Map Species Richness in a Montane Biodiversity Hotspot”. In: *Diversity and Distributions* 21.11, pp. 1329–1338.

Pouteau, Robin, François Munoz, and Philippe Birnbaum (Apr. 2019). “Disentangling the Processes Driving Tree Community Assembly in a Tropical Biodiversity Hotspot (New Caledonia)”. en. In: *Journal of Biogeography* 46.4, pp. 796–806. ISSN: 0305-0270, 1365-2699. DOI: `10.1111/jbi.13535`.

Prendergast, John R., Rachel M. Quinn, J. H. Lawton, B. C. Eversham, and D. W. Gibbons (Sept. 1993). “Rare Species, the Coincidence of Diversity Hotspots and Conservation Strategies”. en. In: *Nature* 365.6444, pp. 335–337. ISSN: 1476-4687. DOI: `10.1038/365335a0`.

Prendergast, John R., Rachel M. Quinn, and John H. Lawton (June 1999). “The Gaps between Theory and Practice in Selecting Nature Reserves”. en. In: *Conservation Biology* 13.3, pp. 484–492. ISSN: 1523-1739. DOI: `10.1046/j.1523-1739.1999.97428.x`.

Pressey, R. L. and A. O. Nicholls (Jan. 1989). “Efficiency in Conservation Evaluation: Scoring versus Iterative Approaches”. In: *Biological Conservation*. Australian Developments in Conservation Evaluation 50.1, pp. 199–218. ISSN: 0006-3207. DOI: `10.1016/0006-3207(89)90010-4`.

Pressey, R. L., C. J. Humphries, C. R. Margules, R. I. Vane-Wright, and P. H. Williams (Apr. 1993). “Beyond Opportunism: Key Principles for Systematic Reserve Selection”. In: *Trends in Ecology & Evolution* 8.4, pp. 124–128. ISSN: 0169-5347. DOI: `10.1016/0169-5347(93)90023-I`.

Pressey, R. L., H. P. Possingham, and C. R. Margules (Jan. 1996). “Optimality in Reserve Selection Algorithms: When Does It Matter and How Much?” In: *Biological Conservation* 76.3, pp. 259–267. ISSN: 0006-3207. DOI: `10.1016/0006-3207(95)00120-4`.

Pressey, RL, M Watts, M Ridges, and T Barrett (2005). “C-Plan Conservation Planning Software”. In: *User Manual. NSW Department of Environment and Conservation*.

Pressey, Robert L, Matthew E Watts, Thomas W Barrett, and Malcolm J Ridges (2009). “The C-Plan Conservation Planning System: Origins, Applications, and Possible Futures”. In: *Spatial conservation prioritization: quantitative methods and computational tools*, pp. 211–34.

Pressey, Robert (Sept. 2002). “The First Reserve Selection Algorithm - A Retrospective on Jamie Kirkpatrick’s 1983 Paper”. In: *Progress in Physical Geography - PROG PHYS GEOG* 26, pp. 434–441. DOI: `10.1191/0309133302pp347xx`.

Prud’homme, Charles, Jean-Guillaume Fages, and Xavier Lorca (2017). *Choco Documentation*.

Pusceddu, A., S. Bianchelli, J. Martin, P. Puig, A. Palanques, P. Masque, and R. Danovaro (June 2014). “Chronic and Intensive Bottom Trawling Impairs

Deep-Sea Biodiversity and Ecosystem Functioning". en. In: *Proceedings of the National Academy of Sciences* 111.24, pp. 8861–8866. ISSN: 0027-8424, 1091-6490. DOI: `10.1073/pnas.1405454111`.

ReVelle, Charles S., Justin C. Williams, and John J. Boland (June 2002). "Counterpart Models in Facility Location Science and Reserve Selection Science". en. In: *Environmental Modeling & Assessment* 7.2, pp. 71–80. ISSN: 1420-2026, 1573-2967. DOI: `10.1023/A:1015641514293`.

Read, Jennifer, Gordon D. Sanson, Martin Burd, and Tanguy Jaffré (2008). "Mass Flowering and Parental Death in the Regeneration of Cerberiopsis Candelabra (Apocynaceae), a Long-Lived Monocarpic Tree in New Caledonia1". en. In: *American Journal of Botany* 95.5, pp. 558–567. ISSN: 1537-2197. DOI: `10.3732/ajb.2007194`.

Rebelo, A. G. and W. R. Siegfried (1992). "Where Should Nature Reserves Be Located in the Cape Floristic Region, South Africa? Models for the Spatial Configuration of a Reserve Network Aimed at Maximizing the Protection of Floral Diversity". en. In: *Conservation Biology* 6.2, pp. 243–252. ISSN: 1523-1739. DOI: `10.1046/j.1523-1739.1992.620243.x`.

Régin, Jean-Charles (1994). "A Filtering Algorithm for Constraints of Difference in CSPs". In: *AAAI*. Vol. 94, pp. 362–367.

Renz, Jochen (2015). "AIBIRDS: The Angry Birds Artificial Intelligence Competition." In: *AAAI*, pp. 4326–4327.

Robinson, John G. (June 2006). "Conservation Biology and Real-World Conservation". en. In: *Conservation Biology* 20.3, pp. 658–669. ISSN: 0888-8892, 1523-1739. DOI: `10.1111/j.1523-1739.2006.00469.x`.

Rocard, Michel (1988). *Accords de Matignon-Oudinot*.

Rodrigues, Ana S. L. and Kevin J. Gaston (Sept. 2002). "Optimisation in Reserve Selection Procedures—Why Not?" In: *Biological Conservation* 107.1, pp. 123–129. ISSN: 0006-3207. DOI: `10.1016/S0006-3207(02)00042-3`.

Rodrigues, Ana S., J. Orestes Cerdeira, and Kevin J. Gaston (Oct. 2000). "Flexibility, Efficiency, and Accountability: Adapting Reserve Selection Algorithms to More Complex Conservation Problems". en. In: *Ecography* 23.5, pp. 565–574. ISSN: 1600-0587. DOI: `10.1111/j.1600-0587.2000.tb00175.x`.

Rossi, Edited F, P van Beek, and T Walsh (2006). "Handbook of Constraint Programming". en. In: p. 969.

Rothley, K. D. (Aug. 1999). "Designing Bioreserve Networks to Satisfy Multiple, Conflicting Demands". en. In: *Ecological Applications* 9.3, pp. 741–750. ISSN: 1939-5582. DOI: `10.1890/1051-0761(1999)009[0741:DBNTSM]2.0.CO;2`.

Sahr, Kevin, Denis White, and A. Jon Kimerling (Jan. 2003). "Geodesic Discrete Global Grid Systems". In: *Cartography and Geographic Information Science* 30.2, pp. 121–134. ISSN: 1523-0406. DOI: `10.1559/152304003100011090`.

Sarkar, Sahotra (July 2012). "Complementarity and the Selection of Nature Reserves: Algorithms and the Origins of Conservation Planning, 1980–1995". en. In: *Archive for History of Exact Sciences* 66.4, pp. 397–426. ISSN: 0003-9519, 1432-0657. DOI: `10.1007/s00407-012-0097-6`.

Sarkar, Sahotra (Jan. 2013). "Environmental Philosophy: From Theory to Practice". In: *Studies in History and Philosophy of Science Part C: Studies in History* *and Philosophy of Biological and Biomedical Sciences* 45. DOI: 10.1016/j.shpsc.2013.10.010.

Sarkar, Sahotra et al. (2006). "Biodiversity Conservation Planning Tools: Present Status and Challenges for the Future". In: *Annual Review of Environment and Resources* 31.1, pp. 123–159. DOI: 10.1146/annurev.energy.31.042606.085844.

Saura, Santiago and Lucía Pascual-Hortal (Nov. 2007). "A New Habitat Availability Index to Integrate Connectivity in Landscape Conservation Planning: Comparison with Existing Indices and Application to a Case Study". en. In: *Landscape and Urban Planning* 83.2, pp. 91–103. ISSN: 0169-2046. DOI: 10.1016/j.landurbplan.2007.03.005.

Saura, Santiago and Lidón Rubio (2010). "A Common Currency for the Different Ways in Which Patches and Links Can Contribute to Habitat Availability and Connectivity in the Landscape". en. In: *Ecography* 33.3, pp. 523–537. ISSN: 1600-0587. DOI: 10.1111/j.1600-0587.2009.05760.x.

Saura, Santiago and Josep Torné (Jan. 2009). "Conefor Sensinode 2.2: A Software Package for Quantifying the Importance of Habitat Patches for Landscape Connectivity". en. In: *Environmental Modelling & Software* 24.1, pp. 135–139. ISSN: 1364-8152. DOI: 10.1016/j.envsoft.2008.05.005.

Schapaugh, Adam W. and Andrew J. Tyre (2012). "A Simple Method for Dealing with Large State Spaces". en. In: *Methods in Ecology and Evolution* 3.6, pp. 949–957. ISSN: 2041-210X. DOI: 10.1111/j.2041-210X.2012.00242.x.

Schapaugh, Adam W. and Andrew J. Tyre (Apr. 2013). "Accounting for Parametric Uncertainty in Markov Decision Processes". In: *Ecological Modelling* 254, pp. 15–21. ISSN: 0304-3800. DOI: 10.1016/j.ecolmodel.2013.01.003.

Schapaugh, Adam W. and Andrew J. Tyre (June 2014). "Maximizing a New Quantity in Sequential Reserve Selection". en. In: *Environmental Conservation* 41.2, pp. 198–205. ISSN: 0376-8929, 1469-4387. DOI: 10.1017/S0376892913000544.

Schlottfeldt, Shana, Jon Timmis, Maria Walter, Andre de Carvalho, Lorena Mendes Simon, Rafael Loyola, and Jose Alexandre Diniz-Filho (Mar. 2015). "A Multi-Objective Optimization Approach Associated to Climate Change Analysis to Improve Systematic Conservation Planning". In: pp. 458–472. DOI: 10.1007/978-3-319-15892-1_31.

Schmitt, Sylvain, Robin Pouteau, Dimitri Justeau, Florian Boissieu, and Philippe Birnbaum (Aug. 2017). "Ssdm: An r Package to Predict Distribution of Species Richness and Composition Based on Stacked Species Distribution Models". In: *Methods in Ecology and Evolution* 8.12, pp. 1795–1803. ISSN: 2041-210X. DOI: 10.1111/2041-210X.12841.

Schulte, Christian, Mikael Lagerkvist, and Guido Tack (2006). "Gecode". In: *Software download and online material at the website: http://www. gecode. org*, pp. 11–13.

Schulte, Christian, Guido Tack, and Mikael Z Lagerkvist (2010). "Modeling and Programming with Gecode". In: *Schulte, Christian and Tack, Guido and Lagerkvist, Mikael* 1.

Sessions, John (Feb. 1992). “Solving for Habitat Connections as a Steiner Network Problem”. en. In: *Forest Science* 38.1, pp. 203–207. ISSN: 0015-749X. DOI: `10.1093/forestscience/38.1.203`.

Smith, Paul G. R. and John B. Theberge (Nov. 1986). “A Review of Criteria for Evaluating Natural Areas”. en. In: *Environmental Management* 10.6, pp. 715–734. ISSN: 0364-152X, 1432-1009. DOI: `10.1007/BF01867726`.

Smith, Paul G. R. and John B. Theberge (Aug. 1987). “Evaluating Natural Areas Using Multiple Criteria: Theory and Practice”. en. In: *Environmental Management* 11.4, pp. 447–460. ISSN: 0364-152X, 1432-1009. DOI: `10.1007/BF01867653`.

Soulé, Michael E. and Bruce A. Wilcox (1980). *Conservation Biology : An Evolutionary-Ecological Perspective*. en. ISBN: 978-0-87893-800-1.

Steiger, Robin, Willem-Jan van Hoeve, and Rados\law Szymanek (2011). “An Efficient Generic Network Flow Constraint”. In: *Proceedings of the 2011 ACM Symposium on Applied Computing*. SAC ’11. New York, NY, USA: ACM, pp. 893–900. ISBN: 978-1-4503-0113-8. DOI: `10.1145/1982185.1982381`.

Strona, Giovanni, Simon D. Stringer, Ghislain Vieilledent, Zoltan Szantoi, John Garcia-Ulloa, and Serge A. Wich (Aug. 2018). “Small Room for Compromise between Oil Palm Cultivation and Primate Conservation in Africa”. en. In: *Proceedings of the National Academy of Sciences* 115.35, pp. 8811–8816. ISSN: 0027-8424, 1091-6490. DOI: `10.1073/pnas.1804775115`.

Stuckey, Peter J, Ralph Becket, and Julien Fischer (2010). “Philosophy of the M Ini Z Inc Challenge”. In: *Constraints* 15.3, pp. 307–316.

Stuckey, Peter J., Thibaut Feydy, Andreas Schutt, Guido Tack, and Julien Fischer (June 2014). “The MiniZinc Challenge 2008–2013”. en. In: *AI Magazine* 35.2, pp. 55–60. ISSN: 2371-9621. DOI: `10.1609/aimag.v35i2.2539`.

Sukor, Abdul Syafiq Abdull, Ammar Zakaria, Norasmadi Abdul Rahim, Latifah Munirah Kamarudin, Rossi Setchi, and Hiromitsu Nishizaki (Jan. 2019). “A Hybrid Approach of Knowledge-Driven and Data-Driven Reasoning for Activity Recognition in Smart Homes”. en. In: *Journal of Intelligent & Fuzzy Systems* 36.5, pp. 4177–4188. ISSN: 1064-1246. DOI: `10.3233/JIFS-169976`.

Tambosi, Leandro R., Alexandre C. Martensen, Milton C. Ribeiro, and Jean P. Metzger (2014). “A Framework to Optimize Biodiversity Restoration Efforts Based on Habitat Amount and Landscape Connectivity”. en. In: *Restoration Ecology* 22.2, pp. 169–177. ISSN: 1526-100X. DOI: `10.1111/rec.12049`.

Taylor, Philip D., Lenore Fahrig, Kringen Henein, and Gray Merriam (Dec. 1993). “Connectivity Is a Vital Element of Landscape Structure”. en. In: *Oikos* 68.3, p. 571. ISSN: 00301299. DOI: `10.2307/3544927`.

Temple, Stanley A (1991). “Conservation Biology: New Goals and New Partners for Managers of Biological Resources”. In: *Decker, DJ; Krasny, ME; Goff, GR; Smith, CR*, pp. 45–54.

Underhill, L. G. (Jan. 1994). “Optimal and Suboptimal Reserve Selection Algorithms”. In: *Biological Conservation* 70.1, pp. 85–87. ISSN: 0006-3207. DOI: `10.1016/0006-3207(94)90302-6`.

Usher, Michael B. (1986). “Wildlife Conservation Evaluation: Attributes, Criteria and Values”. en. In: *Wildlife Conservation Evaluation*. Ed. by Michael B.

Usher. Dordrecht: Springer Netherlands, pp. 3–44. ISBN: 978-94-009-4091-8. DOI: 10.1007/978-94-009-4091-8_1.

Uuemaa, Evelyn, Ülo Mander, and Riho Marja (May 2013). “Trends in the Use of Landscape Spatial Metrics as Landscape Indicators: A Review”. en. In: *Ecological Indicators*. 10 Years Ecological Indicators 28, pp. 100–106. ISSN: 1470-160X. DOI: 10.1016/j.ecolind.2012.07.018.

“The History and Distinctions of Conservation Biology” (2008). en. In: *Conservation Biology: Foundations, Concepts, Applications*. Ed. by Fred Van Dyke. Dordrecht: Springer Netherlands, pp. 1–27. ISBN: 978-1-4020-6891-1. DOI: 10.1007/978-1-4020-6891-1_1.

Van Hentenryck, Pascal, Helmut Simonis, and Mehmet Dincbas (1992). “Constraint Satisfaction Using Constraint Logic Programming”. In: *Artificial intelligence* 58.1-3, pp. 113–159.

Van Hoeve, Willem-Jan (2001). “The Alldifferent Constraint: A Survey”. In: *arXiv preprint cs/0105015*. arXiv: cs/0105015.

Vancutsem, C., F. Achard, J.-F. Pekel, G. Vieilledent, S. Carboni, D. Simonetti, J. Gallego, L. Aragao, and R. Nasi (Jan. 2020). “Long-Term (1990-2019) Monitoring of Tropical Moist Forests Dynamics”. In: *bioRxiv*, p. 2020.09.17.295774. DOI: 10.1101/2020.09.17.295774.

Vane-Wright, R. I., C. J. Humphries, and P. H. Williams (Jan. 1991). “What to Protect?—Systematics and the Agony of Choice”. In: *Biological Conservation* 55.3, pp. 235–254. ISSN: 0006-3207. DOI: 10.1016/0006-3207(91)90030-D.

Victorero, Lissette, Les Watling, Maria L. Deng Palomares, and Claire Nouvian (2018). “Out of Sight, But Within Reach: A Global History of Bottom-Trawled Deep-Sea Fisheries From >400 m Depth”. English. In: *Frontiers in Marine Science* 5. ISSN: 2296-7745. DOI: 10.3389/fmars.2018.00098.

Vieilledent, Ghislain, Clovis Grinand, Fety A. Rakotomalala, Rija Ranaivosoa, Jean-Roger Rakotoarijaona, Thomas F. Allnutt, and Frédéric Achard (June 2018). “Combining Global Tree Cover Loss Data with Historical National Forest Cover Maps to Look at Six Decades of Deforestation and Forest Fragmentation in Madagascar”. en. In: *Biological Conservation* 222, pp. 189–197. ISSN: 0006-3207. DOI: 10.1016/j.biocon.2018.04.008.

Vos, Jurriaan M. De, Lucas N. Joppa, John L. Gittleman, Patrick R. Stephens, and Stuart L. Pimm (2015). “Estimating the normal background rate of species extinction”. es. In: *Conservation Biology* 29.2, pp. 452–462. ISSN: 1523-1739. DOI: 10.1111/cobi.12380.

Wallace, Mark (Sept. 1996). “Practical Applications of Constraint Programming”. en. In: *Constraints* 1.1-2, pp. 139–168. ISSN: 1383-7133, 1572-9354. DOI: 10.1007/BF00143881.

Walsh, Toby (2003). “Consistency and Propagation with Multiset Constraints: A Formal Viewpoint”. en. In: *Principles and Practice of Constraint Programming – CP 2003*. Ed. by Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, and Francesca Rossi. Vol. 2833. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 724–738. ISBN: 978-3-540-20202-8 978-3-540-45193-8. DOI: 10.1007/978-3-540-45193-8_49.

Wang, Yi-cheng and Hayri Önal (Oct. 2011). "Designing Connected Nature Reserve Networks Using a Graph Theory Approach". In: *Acta Ecologica Sinica* 31.5, pp. 235–240. ISSN: 1872-2032. DOI: `10.1016/j.chnaes.2011.06.001`.

Wang, Yicheng and Hayri Önal (Apr. 2016). "Optimal Design of Compact and Connected Nature Reserves for Multiple Species: Reserve Networks for Multiple Species". en. In: *Conservation Biology* 30.2, pp. 413–424. ISSN: 08888892. DOI: `10.1111/cobi.12629`.

Wang, Yicheng, Hayri Önal, and Qiaoling Fang (June 2018). "How Large Spatially-Explicit Optimal Reserve Design Models Can We Solve Now? An Exploration of Current Models' Computational Efficiency". en. In: *Nature Conservation* 27, pp. 17–34. ISSN: 1314-3301. DOI: `10.3897/natureconservation.27.21642`.

Watts, Matthew E., Ian R. Ball, Romola S. Stewart, Carissa J. Klein, Kerrie Wilson, Charles Steinback, Reinaldo Lourival, Lindsay Kircher, and Hugh P. Possingham (Dec. 2009). "Marxan with Zones: Software for Optimal Conservation Based Land- and Sea-Use Zoning". In: *Environmental Modelling & Software*. Special Issue on Simulation and Modelling in the Asia-Pacific Region 24.12, pp. 1513–1521. ISSN: 1364-8152. DOI: `10.1016/j.envsoft.2009.06.005`.

Weerasena, Lakmali, Douglas Shier, and David Tonkyn (Oct. 2014). "A Hierarchical Approach to Designing Compact Ecological Reserve Systems". en. In: *Environmental Modeling & Assessment* 19.5, pp. 437–449. ISSN: 1420-2026, 1573-2967. DOI: `10.1007/s10666-013-9393-z`.

Welzl, Emo (1991). "Smallest Enclosing Disks (Balls and Ellipsoids)". en. In: *New Results and New Trends in Computer Science*. Ed. by Hermann Maurer. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 359–370. ISBN: 978-3-540-46457-0. DOI: `10.1007/BFb0038202`.

Whitehead, Alfred North and Bertrand Russel (1910). *Principia Mathematica*. Cambridge.

Williams, David R., Andrew Balmford, and David S. Wilcove (2020). "The Past and Future Role of Conservation Science in Saving Biodiversity". en. In: *Conservation Letters* n/a.n/a, e12720. ISSN: 1755-263X. DOI: `10.1111/conl.12720`.

Williams, J C and C S ReVelle (1996). "A 0 - 1 Programming Approach to Delineating Protected Reserves". en. In: *Environment and Planning B: Planning and Design* 23.5, pp. 607–624. ISSN: 0265-8135, 1472-3417. DOI: `10.1068/b230607`.

Williams, Justin C. (Oct. 2002). "A Zero-One Programming Model for Contiguous Land Acquisition". en. In: *Geographical Analysis* 34.4, pp. 330–349. ISSN: 1538-4632. DOI: `10.1111/j.1538-4632.2002.tb01093.x`.

Williams, Justin C. (Feb. 2008). "Optimal Reserve Site Selection with Distance Requirements". In: *Computers & Operations Research*. Part Special Issue: Location Modeling Dedicated to the Memory of Charles S. ReVelle 35.2, pp. 488–498. ISSN: 0305-0548. DOI: `10.1016/j.cor.2006.03.012`.

Williams, Justin C., Charles S. ReVelle, and Simon A. Levin (Sept. 2005). "Spatial Attributes and Reserve Design Models: A Review". en. In: *Environmental*

Modeling & Assessment 10.3, pp. 163–181. ISSN: 1420-2026, 1573-2967. DOI: `10.1007/s10666-005-9007-5`.

Wistbacka, Ralf, Markku Orell, and Andrea Santangeli (Aug. 2018). “The Tragedy of the Science-Policy Gap – Revised Legislation Fails to Protect an Endangered Species in a Managed Boreal Landscape”. In: *Forest Ecology and Management* 422, pp. 172–178. ISSN: 0378-1127. DOI: `10.1016/j.foreco.2018.04.017`.

Woinarski, John C. Z., Andrew A. Burbidge, and Peter L. Harrison (Apr. 2015). “Ongoing Unraveling of a Continental Fauna: Decline and Extinction of Australian Mammals since European Settlement”. en. In: *Proceedings of the National Academy of Sciences* 112.15, pp. 4531–4540. ISSN: 0027-8424, 1091-6490. DOI: `10.1073/pnas.1417301112`.

Wright, David J, Nirmal J Shah, and David S Richardson (2014). “Translocation of the Seychelles Warbler Acrocephalus Sechellensis to Establish a New Population on Frégate Island, Seychelles”. In: *Conserv Evid* 11, pp. 20–24.

Wulff, Adrien S., Peter M. Hollingsworth, Antje Ahrends, Tanguy Jaffré, Jean-Marie Veillon, Laurent L’Huillier, and Bruno Fogliani (Sept. 2013). “Conservation Priorities in a Biodiversity Hotspot: Analysis of Narrow Endemic Plant Species in New Caledonia”. In: *PLoS ONE* 8.9. ISSN: 1932-6203. DOI: `10.1371/journal.pone.0073371`.

Xue, Yexiang, Xiaojian Wu, Dana Morin, Bistra Dilkina, Angela Fuller, Jeffrey Andy Royle, and Carla P Gomes (2017). “Dynamic Optimization of Landscape Connectivity Embedding Spatial-Capture-Recapture Information”. en. In: p. 7.

Zhou, Neng-Fa and Håkan Kjellerstrand (2017). “Optimizing SAT Encodings for Arithmetic Constraints”. In: *International Conference on Principles and Practice of Constraint Programming*. Springer, pp. 671–686.

van Laarhoven, Peter J. M. and Emile H. L. Aarts (1987). *Simulated Annealing: Theory and Applications*. en. Dordrecht: Springer Netherlands. ISBN: 978-90-481-8438-5 978-94-015-7744-1. DOI: `10.1007/978-94-015-7744-1`.

ABSTRACT

In the context of the global biodiversity crisis, human activities are the principal cause of natural habitat degradation, fragmentation, and destruction. Humanity's global impacts have considerably increased species extinction rates and about one million species are nowadays threatened with extinction, according to the last Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) report. Conservation biology is a multidisciplinary research area which attempts to address the current biodiversity crisis challenges. The development of practical approaches to promote conservation by reducing the research-implementation gap is one of its objectives. Last decades, systematic conservation planning (SCP) emerged as a framework relying on optimization and computer science research. Its main target is to provide efficient decision support in the planning of conservation actions. SCP offers an opportunity to bridge the research-implementation gap by providing accountable decision support tools able to integrate ecological targets along with socio-economical constraints.

In this PhD thesis, we introduce a unifying approach for modelling and solving SCP problems based on constraint programming, a method from artificial intelligence based on automated reasoning. The motivations of this approach are to provide more expressiveness into SCP (i.e. the breadth and variety of problems that can be represented and solved), notably through the integration of advanced spatial constraints and landscape indices within a formal solving approach. Formal approaches are often more difficult to implement and scale up to large problems than heuristic approach. However, they present the advantage of providing satisfiability and optimality guarantees on solutions.

The methods developed in this thesis are evaluated on real data from New Caledonian forests. As the smallest biodiversity hotspot in the world, New Caledonia has to struggle with many conservation challenges. Moreover, the developed, insular and low populated New Caledonian context allows high proximity between conservation stakeholders, which makes it an appropriate field of study to experiment novel approaches. This PhD thesis illustrates this particularity through a real case study, conducted in close collaboration with New Caledonian managers of the "Côte Oubliée – 'Woen Vùù – Pwa Pereeù" provincial park to provide decision support in a reforestation project. The constraint-based paradigm proposed in this thesis has proven its genericity, its flexibility, and its expressiveness, offering new perspectives for systematic conservation planning.