Biological Conservation 282 (2023) 110066

Contents lists available at ScienceDirect

Biological Conservation

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

Cumulative effects on habitat networks: How greedy should we be?

François Hamonic [a,b,*], Basile Couëtoux [a], Yann Vaxès [a], Cécile H. Albert [b]

[a] Aix-Marseille Univ, CNRS, Université de Toulon, LIS, Marseille, France
[b] Aix-Marseille Univ, CNRS, Univ Avignon, IRD, IMBE, Marseille, France

ARTICLE INFO

Keywords:
Biodiversity
Landscape connectivity
Connectivity conservation
Spatial graph
Landscape change
Biological restoration

ABSTRACT

Connectivity conservation and restoration are key strategies to preserve biodiversity under the pressure of habitat loss and fragmentation. Numerous quantitative approaches have been developed to help conservation practitioners take the most effective and cost-efficient actions to conserve, restore or create habitat patches or corridors that preserve or enhance landscape connectivity. This problem is often solved by ranking habitat patches (or corridors) with suboptimal algorithms that do not account for the cumulative effects that may occur in the case of simultaneous or sequential losses (or enhancements) of multiple habitat patches. Accounting for these cumulative effects is one of the current challenges for connectivity conservation. Here, we quantify these cumulative effects and explore the trade-off between solution quality and computational time when optimizing the selection of conservation/restoration actions that maximize landscape connectivity under a budget constraint; connectivity is here measured by the frequently-used Equivalent Connected Area. We compare the solutions obtained with a new optimization pipeline with solutions obtained with four simpler algorithms used in most connectivity conservation studies. Comparison is performed for four case studies covering a wide range of possible applications for conservation practitioners. We show that the simpler algorithms can provide suboptimal solutions, when conservation/restoration actions have a strong impact on least-cost paths between the habitat patches, and that optimal resolution should be considered whenever possible.

1. Introduction

Connectivity conservation has been identified as an essential lever for biodiversity conservation in the face of natural habitat loss and fragmentation (Crooks and Sanjayan, 2006; Lawler, 2009). Implementing connectivity conservation requires that conservation practitioners identify areas where habitat conservation or restoration will be most effective and cost-efficient for landscape connectivity (Beier et al., 2011). A multitude of frameworks have been developed to address this question for a variety of contexts and biological systems (Magris et al., 2016; Albert et al., 2017; Tarabon et al., 2019). However, they largely struggle to account for the cumulative effects that can occur when multiple impacts or small-scale decisions accumulate over large habitat networks (Foley et al., 2017; Whitehead et al., 2017).

One of the most typical workflows, widely applied by conservation practitioners, is to rank each of the individual habitat patches (or corridors) of the study area according to their importance in maintaining overall landscape connectivity. This ranking thus guides decision-making for conservation measures, with high-ranked patches (or corridors) targeted for conservation or restoration actions within a given budget and low-ranking ones allocated to other human uses (Yemshanov et al., 2019). Quantifying the contribution of habitat patches (or corridors) to connectivity can be conducted through a variety of methods and metrics, including graph-based approaches (Urban and Keitt, 2001). Measures of connectivity which relate to habitat availability, such as the Equivalent Connected Area (hereafter ECA, Saura et al. (2011)) or the Probability of Connectivity (hereafter PC, Saura and Pascual-Hortal (2007)), that are currently widely used for this purpose (Schivo et al., 2020) as they have received encouraging experimental support (M. Pereira et al., 2011; Awade et al., 2012). The ranking of habitat patches (or corridors) is then estimated by a patch removal experiment, i.e. by calculating the relative decrease in connectivity obtained when patches are hypothetically removed in turn (Bodin and Saura, 2010). Similarly, such virtual experiments can be made to identify the best candidates for restoration or re-creation of habitat patches or corridors by calculating the relative increase in connectivity (Blazquez-Cabrera et al., 2019).

* Corresponding author at: LIS UMR 7020, Campus universitaire de Luminy Bat. TPR2, 163 avenue de Luminy, 13288 Marseille cedex 09, France.
E-mail addresses: francois.hamonic@lis-lab.fr (F. Hamonic), basile.couetoux@lis-lab.fr (B. Couëtoux), yann.vaxes@lis-lab.fr (Y. Vaxès), cecile.albert@imbe.fr (C.H. Albert).

https://doi.org/10.1016/j.biocon.2023.110066
Received 23 January 2023; Received in revised form 31 March 2023; Accepted 7 April 2023
Available online 26 April 2023

Despite its widespread use, this approach still largely fails to capture the cumulative effects that may occur in the case of simultaneous or sequential losses (or enhancements) of multiple patches (or corridors) through time (Foley et al., 2017). Indeed, a large majority of studies use this framework with single-node (nodes being habitat patches) or single-link (links being corridors in a graph-based approach) removal/addition (Table A.1). This means that they assess the potential effects of conserving (or restoring) each patch (or corridor) independently and therefore do not take into account the potential interactions of multiple changes, i.e. the cumulative effects. These interactions can be: synergies when potential changes are in series and only their combination might improve the connectivity, redundancies when potential changes are in parallel and choosing only one might be sufficient, or antagonisms when a potential change might make no sense without additional actions (Fig. 1). A few years ago, Rubio et al. (2015) have warned that single-node approaches might lead to suboptimal conservation solutions. For instance, single-node approaches will overlook the potential benefit of connecting large parks through a series of stepping stones (Fig. 1). Rubio et al. (2015) also pointed out that identifying the N most important patches in the study area on the basis of an exhaustive search through all combinations of deletions of patches is computationally demanding and can quickly become impractical for networks with more than 20 nodes.

Since then, a few studies have used greedy algorithms that build solutions iteratively (Hodgson et al., 2016; Tarabon et al., 2019). At each step, simulations are performed taking into account the decisions made in the previous steps and the best element is added to the solution. Various algorithms have also been developed for multi-node selection based on network centrality or habitat availability (An and Liu, 2016; J. Pereira et al., 2017) for landscapes with up to 150 nodes, or in the specific case of dendritic networks (Wu et al., 2014; Sethi et al., 2017). In particular, Xue et al. (2017) have proposed a new algorithm to obtain the exact optimal solution when optimizing PC on a graph, but it remains applicable to few nodes only due to computational constraints (<30). Consequently, no study has yet compared optimal and approximated solutions for larger graphs. This means that we still do not know, both ecologically speaking and in terms of decision-making, how important it is to consider cumulative effects in connectivity conservation.

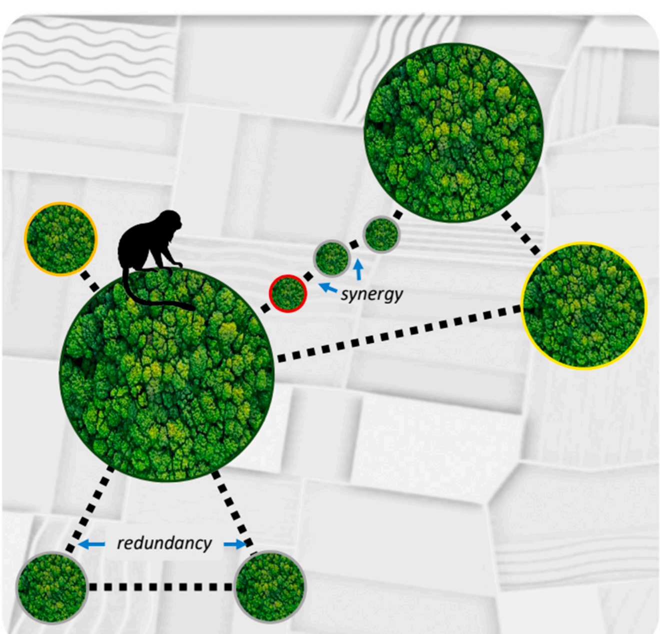

Fig. 1. Hypothetical case study of the restoration of corridors between protected areas, for the conservation of a monkey species. The acquisition of a dash costs one budget unit. A single-link approach will primarily lead to the reconnection of red and/or orange parks to larger ones. It will overlook the potential benefit of linking the two larger parks through a series of stepping stones. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Here, we propose to quantify these cumulative effects on larger graphs (up to hundreds of habitat patches) and to explore the trade-off between computational time and solution quality for real case studies. To do this, we place ourselves in the case of the framework described above, namely the identification of the N most important elements (nodes or links) to be targeted for conservation or restoration actions when maximizing the ECA under a budget constraint. We compare the optimal solution — which takes cumulative effects into account — with solutions obtained with simplified approaches classically used by ecologists that do not account for cumulative effects (single node/link selection) or partially account for them (greedy selection). To obtain the optimal solution, we use the new pipeline we have developed (Hamonic et al., 2023) which combines a preprocessing algorithm and a mixed integer formulation and allows to solve larger problems than previous methods while accounting for all the cumulative effects. We compare the different algorithms both in terms of solution quality and computational time on four contrasted case studies that cover a wide range of possible applications for conservation and restoration practitioners. We address the following questions: 1) Should we account for the synergies and redundancies among options when running connectivity conservation/restoration prioritization analyses? 2) Are the greedy algorithms sufficient to address these synergies/redundancies or do we need an optimal solution?

2. Methods

2.1. Model and problem formulation

The landscape is modeled by a graph $G = (V, A, w, \pi)$ where each node $u \in V$ represents a habitat patch, and is associated with a weight w_u that represents its ecological quality (e.g. habitat area for a given species). Each directed link $(u, v) \in A$ represents a connection that individuals of this species can use to move from node u to node v and π_{uv} is the probability for an individual to succeed in this move.

The ECA indicator (Saura et al., 2011) is a combined measure of habitat amount and connectivity derived from the Probability of Connectivity index (Saura and Pascual-Hortal, 2007). While PC is defined as *the probability that two animals randomly placed within the landscape fall into habitat areas that are reachable from each other*, ECA is *the area of a single habitat patch that would have the same PC value as the landscape*. ECA is computed as:

$$ECA(G = (V, A, w, \pi)) = \sqrt{\sum_{s,t \in V} w_s \cdot w_t \cdot \Pi_{st}}$$

where Π_{st} is the probability of the most probable path from s to t where the probability of a path is the product of the probabilities of its links.

We define the *Max-ECA* problem as follows. Let $G = (V, A, w, \pi)$ be a landscape and let Φ be a set of options. Each option $i \in \Phi$ has a cost c_i, a weight $w_u^i \geq 0$ for each node $u \in V$ and a probability $\pi_{uv}^i \geq \pi_{uv}$ for each link $(u, v) \in A$. The *Max-ECA* problem consists in finding the set of options $S \subseteq \Phi$ whose total cost remains below a budget B and which maximizes the ECA value of the landscape improved by these options, i.e. $ECA(G' = (V, A, w', \pi'))$. The improved weight of a patch u is calculated as $w_u' = w_u + \sum_{i \in S} w_u^i$, i.e. the improvements of different options on the same patch accumulate, and the improved probability of a link (u, v) is calculated as $\pi_{uv}' = max(\pi_{uv}, max_{i \in S} \pi_{uv}^i)$, i.e. only the best allowed probability of the link is taken. This problem captures both restoration and conservation problems. In the case of restoration, G is the current landscape, Φ is the set of restoration options, w_u^i represents the amount by which the habitat area of the node u is increased and π_{uv}^i represents the increased probability of the link (u, v) if the option i is chosen. In the case of conservation, G represents the landscape as it would be if no conservation action is taken, Φ is the set of conservation options, w_u^i represents the amount of habitat area of the node u that is protected by the option i and π_{uv}^i is a lower bound on the probability of the link (u, v) guaranteed by the choice of i. In both cases, the *Max-ECA* problem is to find a set of options S of total cost at most B that maximizes the ECA of the restored (resp. protected) landscape G'.

2.2. Algorithms

For the *Max-ECA* problem, we compare the optimal solution, obtained with the mixed integer program of Hamonic et al. (2023), to the solutions obtained with four simpler algorithms: the *incremental local*, the *decremental local*, the *incremental greedy* and the *decremental greedy*. Local algorithms correspond to single-node approaches while greedy algorithms take decisions step by step, taking into account their past decisions at each step. Incremental selection means we start with the empty solution and add iteratively the most beneficial improvements while decremental selection means we start with all the improvements and progressively remove the less interesting ones.

The incremental local algorithm starts by calculating the potential gain in ECA, Δ_i^+, for each option $i \in \Phi$, i.e. the difference between the ECA value of the graph without enhanced elements and the ECA value of the graph where only the elements of the option i are enhanced. It then selects, in descending order, the options with the greatest ECA gain/cost ratio, i.e. Δ_i^+/c_i, that fit within the budget. The incremental local algorithm does not take cumulative effects into account as options are chosen independently of each other. To overcome this problem, the incremental greedy algorithm iteratively selects the option with the best ECA gain/cost ratio and recalculates these ratios at each step based on the graph updated with the options that were taken in the previous steps. In this way, the incremental greedy algorithm partly takes into account the cumulative effects between the options taken, but at the cost of additional calculations. Because it does not question past decisions, it can easily miss synergies that exist in certain combinations of options.

The decremental local algorithm starts from the graph with all potential improvements made and computes the potential loss in ECA Δ_i^-, for each option $i \in \Phi$, i.e. the difference between the ECA value of the graph with all improvements and the ECA value of the graph with all improvements except those of the option i. Then it iteratively removes the option with the smallest ECA loss/cost ratio, i.e. Δ_i^-/c_i until the remaining options fit within the budget. The decremental greedy does the same, but recomputes potential losses in ECA after each option removal.

2.3. Case studies

We have selected four case studies that cover a wide range of potential applications for conservation practitioners (Fig. 2), including conservation and restoration problems focused on habitat patches or corridors for different types of organisms (mammal, fish, bird, amphibian) in different types of ecosystems (terrestrial or freshwater). These case studies also reflect different types of graphs or topologies that are classically used in connectivity conservation studies, such as the tree graph (which contains only the minimum number of links to connect all nodes, typically used for river systems), the complete graph (which contains all possible links to connect all the pairs of nodes), the minimum planar graph (Fall et al., 2007) (which contains only links that connect two adjacent nodes with no crossings allowed) or the lattice graph (which forms a regular tiling of the study area). Note also that case studies 2 and 3 account for the resistance to movement that animals can encounter when moving among patches in the landscape matrix while case studies 1 and 4 don't. These case studies relate to real conservation or restoration problems, but the results presented here and the parameterizing of the models aim to be illustrative. In each case, the parameters selected (dispersal distance, matrix resistance) are intended to

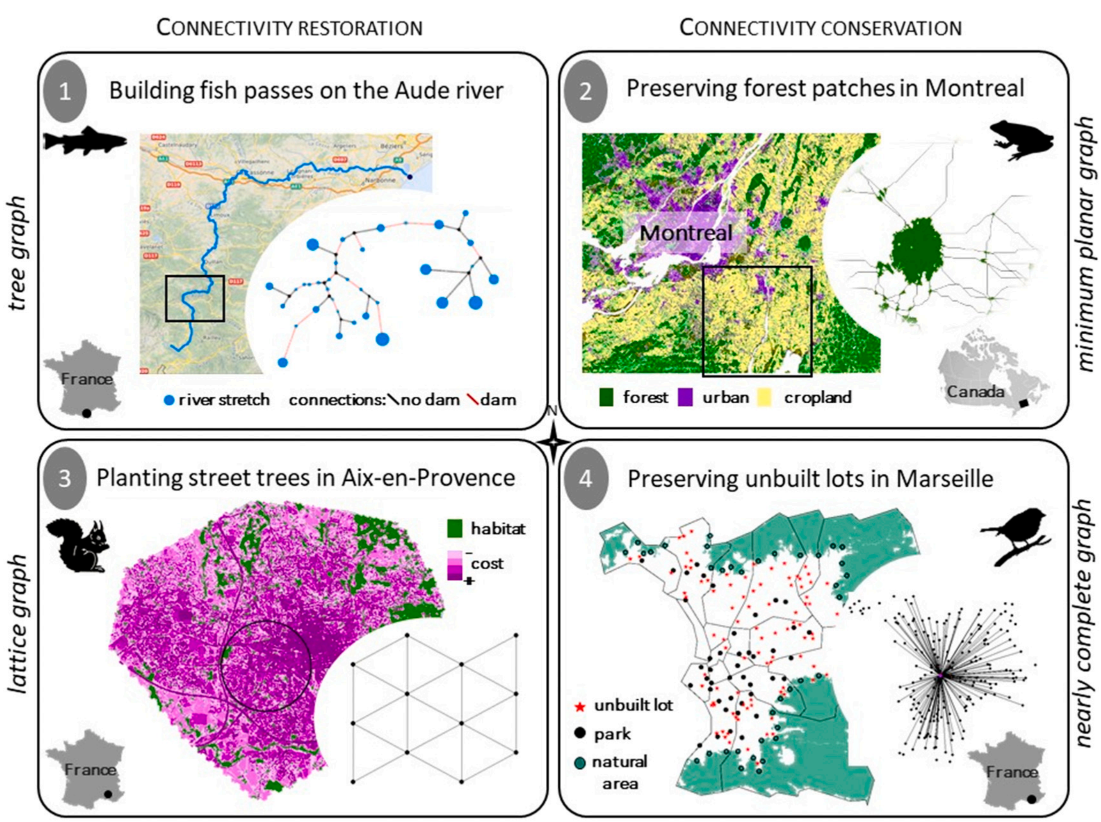

Fig. 2. Case studies.

reflect the biology of the study species and the parameters related to the improvement of nodes or links in the graph have been arbitrarily set to highlight the effects of potential conservation or restoration actions. Application-oriented results would require the involvement of local stakeholders, which was beyond the scope of this study. We implemented our model as well as the preprocessing and greedy algorithms in C++ using Gurobi Optimizer (Gurobi Optimization, 2022).

Case study 1 consists in identifying among a set of 15 dams present on the Aude river (France) those that need to be equipped with fish passes in order to restore the river connectivity for trout (Saint-Pé, 2019). The river is modeled by a graph whose 34 nodes represent the river stretches obtained by cutting the river tributaries at each confluence point and each dam (Erös and Lowe, 2019; Segurado et al., 2013). The weight of each stretch is its length, as an approximation for its area. Two adjacent stretches u and v are connected by reciprocal links (u,v) and (v,u) which represents the ability for an individual to move from the center of one stretch to the center of the other in upstream and downstream directions. Each link (u,v) is associated with a probability π_{uv} representing the feasibility for an individual to make the corresponding movement. If the stretches u and v are separated by a dam we fix this probability to zero, i.e. $\pi_{uv} = 0$, otherwise we compute it according to the negative exponential model, considering a median dispersal range of 10 km (Crook, 2004), i.e. $\pi_{uv} = exp\left(-\frac{log(0.5)}{10000} \cdot d(u,v)\right)$ where $d(u,v)$ is the distance in meters between the centers of the stretches u and v along the river course. The installation of a fish pass on a dam is modeled by increasing the probability of the corresponding links from 0 to 80 % of the probability computed with the negative exponential model. We assume that all fish passes have the same construction cost.

Case study 2 consists in identifying the remnant forest patches that need to be preserved from deforestation in the Montreal vicinity (Canada) to guarantee habitat connectivity for the wood frog (Albert et al., 2017). Here the 518-nodes graph is a minimal planar graph (Fall et al., 2007) whose 909 links are least-cost paths among habitat patches. Each link is weighted with a probability of movement which is a negative exponential function of the least-cost path length among patches, with a median dispersal of 300 m. For each node u, we have an estimate of the area it could lose by 2050 to agriculture or urbanisation (‘business as usual’ scenario, Albert et al., 2017). A total of 260 nodes could be reduced in size ($w_u^i > 0$), of which 80 could disappear if not protected ($w_u^i = w_u$). While for partially-threatened nodes we consider that individuals can continue to move across adjacent links with unchanged probabilities, for each fully-threatened node u we also set the probability of each incident link $\{u,v\}$ to 0, so that they no longer contribute to any most-probable path. The disappearance of these nodes thus also affects the nodes to which they were connected. We assume that the protection cost of each node is proportional to its potential area loss by 2050.

Case study 3 consists in identifying street sections in which planting trees can improve the connectivity of the urban canopy for the European red squirrel in the city of Aix-en-Provence (France). The landscape is here modeled with 6186 hexagon grid cells of 187 m^2 each. A hexagon is associated with a quality weight of 1 if it contains mostly treed areas and 0 otherwise. Each hexagon u is also associated with a probability of connection μ_u (probability that an individual succeeds when moving through it) that depends on the underlying land cover: 1 for habitat, 0.97 for partially treed areas, 0.7 for low vegetated areas, 0.45 for non vegetated areas, 0.25 for roads and 0 for buildings. These values correspond to a median dispersal distance of about 2 km when matrix is suitable and lower otherwise (Wauters et al., 2010). The probability of each link (u,v) is computed as $\pi_{uv} = (\mu_u \cdot \mu_v)^{\frac{1}{2}}$. As a substitute for empirical data, we assume that planting trees along a street section allows squirrels to travel 6 times further: the probability of using the link uv increases, $\pi_{uv}^i = (\pi_{uv})^{\frac{1}{6}}$; the cost of these actions is proportional to the number of crossed hexagons. For illustration, we chose here a set of 47 candidate street sections for which we implemented ECA optimization.

Case study 4 consists in identifying the unbuilt lots (usually biodiversity rich grasslands or shrublands) in the city of Marseille (France) that need to be preserved from development to maintain connectivity among urban parks and the surrounding natural areas, both of which provide habitat for songbirds (e.g. Eurasian blackcap). These unbuilt lots, that are mainly present in the city periphery, can indeed act like stepping stones between natural areas and urban parks. The baseline graph is composed of 196 nodes of which 42 model the frontier of natural areas (20 ha each), 43 represents smaller parks (1 ha), 11 larger parks (5 ha), and 100 unbuilt lots (0.1 ha). The graph is nearly complete, and each link is weighted with a probability of movement which is a negative exponential function of the border to border distance among patches, with median dispersal distance of 3000 m (Paradis et al., 1998). We have arbitrarily chosen to remove links with probabilities less than 0.135 in order to reduce the graph size and to speed up the calculations. When an unbuilt lot is not selected for conservation, its area becomes zero and it no longer contributes to any shortest path (probability of moving along adjacent links set to zero). When selected, its attributes and those of its adjacent links do not change.

3. Results

3.1. Solution quality

By comparing the quality of the solutions obtained by the different algorithms, we have confirmed experimentally that greedy and local algorithms can return - at least for some values of the budget - suboptimal solutions (Fig. 3). In all cases, the four simpler algorithms converge towards the optimal solutions as the budget increases, rapidly up to 10–20 % of the maximum budget (all options are chosen), then more slowly. As expected, the greedy algorithms lead on average to solutions closer to the optimum than the local algorithms (Fig. A.1). The decremental greedy algorithm leads on average to slightly better solutions than the incremental greedy algorithm, while the decremental local algorithm leads on average to worse solutions than the incremental local algorithm. Although solutions can quickly approach the optimal solution as the budget increases, we also observe sharp drops in solution quality for the four simpler algorithms. In the case of Montreal, the incremental local suddenly deviates from the optimum around 11 % of the maximum budget. In the case of Marseille, the decremental local deviates from the optimum around 4 % of the maximum budget. In the case of Aude, the incremental local and incremental greedy show smaller but frequent deviations from the optimum up to 80 % of the maximum budget. The observed deviations from the optimum mean that at some point previous decisions were not the best given the new decisions that can be made as the budget increases. Our results also show some discrepancies between the cases, with the solutions being on average better for Marseille and Montreal cases than for Aix and Aude cases (Fig. 3).

3.2. Execution time

Execution time varies considerably between case studies and algorithms, ranging from a few milliseconds to a few minutes. It thus varies by several orders of magnitude for a given case study processed by the different algorithms (Fig. 4). It is not easy to disentangle the different sources of time distribution.

The incremental local algorithm is the fastest for all case studies, followed by the decremental local algorithm (Fig. 4). For both local algorithms the computation time depends mainly on the graph’s size, because the change in ECA is calculated once for each option and the time needed to calculate ECA depends directly on this.

The greedy algorithms have a longer computation time than the local algorithms, and can be faster (Aix and Aude cases) or equivalent to the optimal algorithm (Montreal and Marseille cases, Fig. 4). Their computation times depend on the graph’s size but also on the value of the budget (Fig. A.2). Indeed, the incremental (resp. decremental)

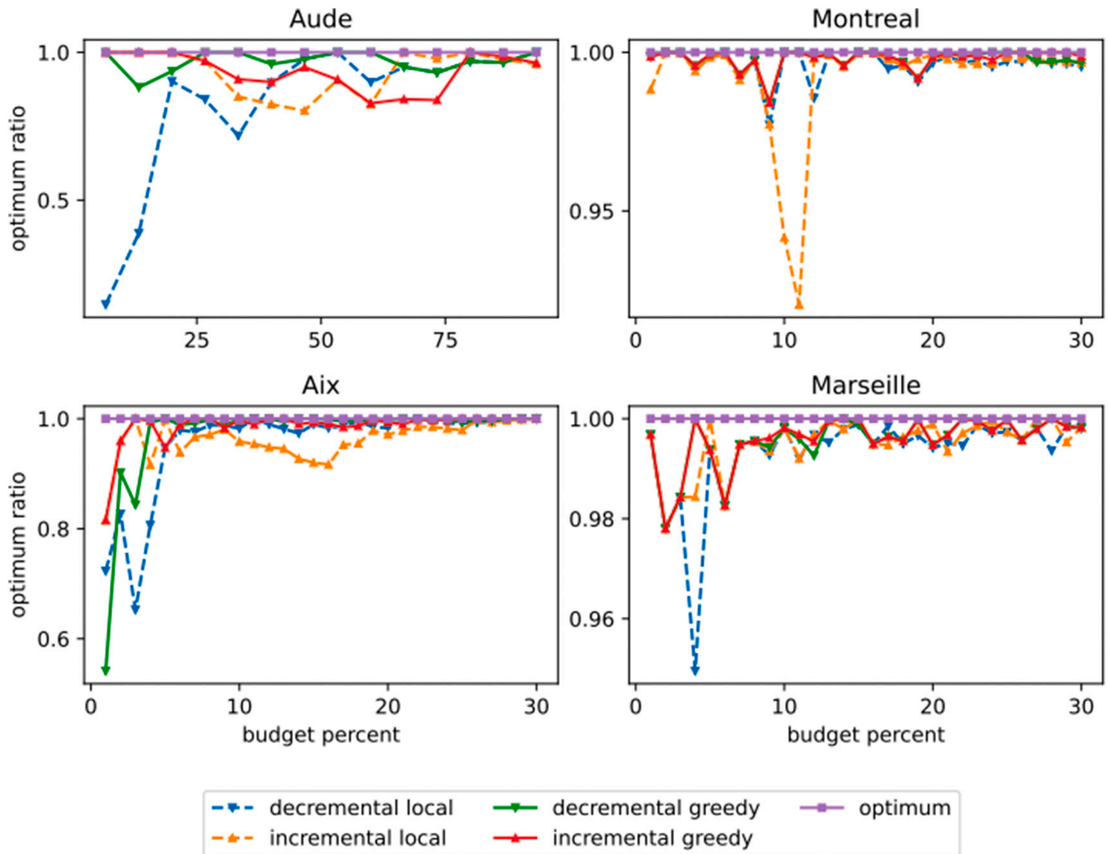

Fig. 3. Percentage gain in ECA achieved by the solutions of the different algorithms compared to the optimal solution (Hamonic et al., 2023) as a function of the budget.

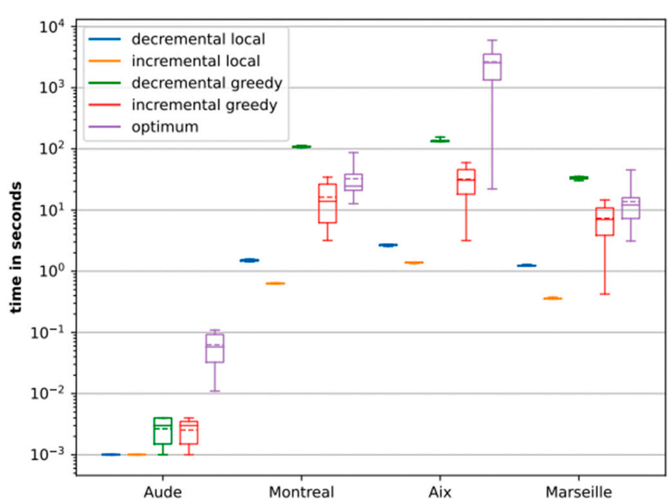

Fig. 4. Box plot synthesizing the computation times achieved by the different algorithms for each case study with different budget values. The whiskers represent the min and max values, the box corresponds to the values between the 25th and 75th percentiles, the horizontal line is the median and the dashed line is the mean.

greedy algorithm needs to compute ECA after each option selection (resp. removal) a larger (resp. smaller) value of the budget may increase the number of steps of the algorithm and thus the number of times ECA has to be computed. Therefore, the two greedy algorithms are equivalent in terms of computation time at an intermediate point of the budget.

The computation time of the optimal algorithm is higher than that of the other algorithms in the cases of Aix and Aude, but is between the computation time of the local algorithms and the greedy algorithms in the cases of Montreal and Marseille (Fig. 4). It depends on the space that needs to be explored on the decision tree of options. The search time is at most a function of 2^N (N being the number of options), but the optimization process allows a large reduction of this time by cutting some branches of the decision tree that cannot lead to an optimal solution.

Note that here the preprocessing time is included in the computation time of the optimal algorithm (it corresponds to a constant part of this computation time as it does not depend on the budget and could be performed only once if several budgets were to be tested as different scenarios of a same project), it allows to reduce the computation time by a factor of 6 to 60, depending on the case study and the density of the graph (Hamonic et al., 2023).

4. Discussion

To better integrate biodiversity issues into land use planning, it is now necessary to determine whether the cumulative effects of decisions should be taken into account (Rubio et al., 2015). Here, we use a new optimal algorithm developed to solve the *Max-ECA* problem (Hamonic et al., 2023) on several full-size instances whose optimal solutions were out of reach by brute force algorithms. We compare this new method with four simpler algorithms that are currently widely used in conservation (Table A.1) but can lead to suboptimal solutions Rubio et al. (2015). Below, we discuss the potential limits of the different algorithms. Based on four contrasted case studies, we establish some guidelines to help decide.

4.1. Quality of the solutions for the different algorithms

As expected, the mixed integer optimization algorithm always finds the best solution for a given budget (higher ECA values), while the local and greedy algorithms mostly lead to under-optimal solutions. The greedy algorithms lead on average to solutions that are closer to the optimum than the local algorithms, because they partly account for cumulative effects. All algorithms, whether local or greedy, incremental or decremental, deviate from the optimal solution for some budget values. There are two possible and non-exclusive explanations to these deviations.

On the one hand, the deviations could be due to the suboptimality of the greedy and local algorithms for the knapsack problem, a problem in combinatorial optimization. The knapsack problem seeks to determine, given a set of objects with a weight and a value, which objects should be chosen to fill a knapsack so that the total weight is less than or equal to its load limit and the total value is as large as possible (Nemhauser and Wolsey, 1988). It is similar to our problem of resource allocation, where the decision-makers have to choose from a set of non-divisible options under a fixed budget. Thus, even if there are no cumulative effects between the chosen options, local and greedy algorithms do not guarantee an optimal budget allocation (Dilkina et al., 2011). However, this suboptimality issue only arises when the costs of the options are not all the same. Otherwise, i.e. when the options all have equal costs, the knapsack can easily be filled optimally by choosing the options with the highest values first. Unlike the other three cases, in the Aude case, all options have a similar cost, so we know that the observed deviations cannot be explained by the suboptimality of the greedy algorithm for the knapsack problem.

On the other hand, these deviations could be due to synergies or redundancies of certain combinations of options, i.e., cumulative effects. The fact that the four simpler algorithms perform poorly in some arbitrary and unpredictable cases is due to the way they work sequentially, and their inability to question previous decisions based on novel ones. As stated earlier, we know that in the Aude case the observed deviations in quality between the simpler algorithms and the optimal algorithm are not related to the suboptimality of greedy algorithms for the knapsack problem; we thus assume these deviations are due to interactions among the potential options and the order in which they are chosen. Sequential selections may overlook the potential effects of selecting multiple options along the same branch if each individual option brings only small improvements in ECA (e.g. Fig. 1). In the other three cases, we assume that the observed deviations may be due to both explanations; the size and density of the graphs make it difficult to understand the effect of each option or budget level on the overall graph and its ECA value.

4.2. Computation time of the different algorithms

Computation time varies by several orders of magnitude between case studies and between algorithms. The computation time of the different algorithms depends on several elements, including: the search procedure, the size of the graph, the number of options, the budget, and the complexity of the problem (existence of synergies/redundancies among options), all being not necessarily independent. For both local algorithms, the computation time depends only on the number of possible options and the time needed to calculate ECA (so the graph size). That is why they are faster to compute, and that also explains why they are currently largely used in applied conservation problems (Table A.1). Here, decremental (less used in practice) is a bit longer than incremental due to implementation procedure which implies more calculation steps. For both greedy algorithms, the computation time depends on the number of options, but also largely on the size of the graph, as the time required to compute ECA directly depends on it, and ECA needs to be recalculated $N - k$ times for each kth additional option. The computation time of the incremental greedy algorithm increases with the value of the budget, this is due to the fact that a larger budget allows for the inclusion of more options and therefore requires the addition of the computation of the effects of these additional options on top of those already chosen.

For the optimal algorithm, the computation time depends on the space that needs to be explored on the decision tree of options, which is at most a function of 2^N (N being the number of options), but the optimization process allows a large reduction of this time by cutting branches of the decision tree that cannot lead to an optimal solution. For this reason, the computation time of the optimal algorithm is shorter for smaller problems (smaller graphs with fewer options) but also when the problem is simpler (fewer complex interactions between the potential solutions) because more branches can be pruned earlier. This can also occur when the maximal ECA value has been reached and additional options only marginally (or do not) improve this value; in the case of Aix, for instance, this arises around 30 % of the budget (or 80 % in the case of Aude). The computation time of the optimal algorithm is overall less predictable.

We have chosen to illustrate this paper by demonstrating the method on cases on which the four algorithms run easily and thus quantify the differences in the orders of magnitude required by the different algorithms. Obviously, no one is afraid to run a calculation for a few minutes, hours, or even days if they want to arrive at an efficient and cost-effective solution to their conservation problem. However, months or years quickly become problematic. The experiments presented in Hamonic et al. (2023) show that the computation time needed to solve the mixed integer program grows almost exponentially with the number of options (see Fig. 5). This means that by adding options to the problem, one can quickly switch from a calculation time of a few hours to a time of a few days. Combining and comparing different conservation scenarios with different budgets, or different sets of possible options, would also only multiply the computation time by the number of different scenarios chosen.

4.3. Importance of the cumulative effects for connectivity conservation

Although the optimal algorithm always finds the best solution for a given budget, in our case studies, local and greedy algorithms provide solutions that are often close to the optimal solution (5 % lower on average, Fig. A.1). This suggests that in our case studies, cumulative effects have a relatively small impact. However, we also observe some discrepancies between the cases, the suboptimal solutions being on average better for Marseille and Montreal cases than for Aix and Aude cases (Fig. A.1). These discrepancies could be explained by the density of these graphs (how close to complete they are); low-density graphs like trees being more sensitive to potential synergies among options. More generally, this sensitivity is related to the impact of conservation/restoration options. When the options have strong impacts on the least cost paths in the graph, we can expect more cumulative effects and thus a higher risk that suboptimal algorithms produce poor quality solutions.

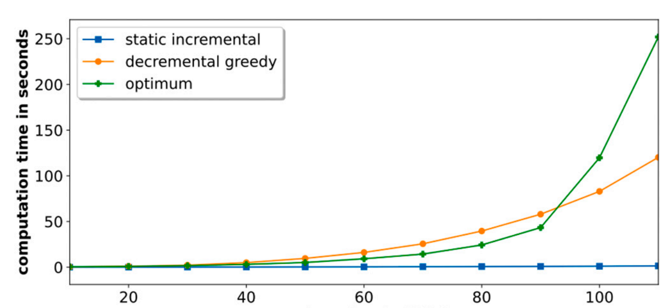

Fig. 5. The average computation time to solve the mixed integer (optimum) program in the Marseille case grows exponentially with the number of unbuilt lots considered (approximately doubles every 10 unbuilt lots) while the computation time of the suboptimal algorithms increases only polynomially.

Indeed, if some options have a large impact on the least-cost paths between nodes, the selection of these options will significantly alter the way other options can increase the ECA value. When a large proportion of the options have a significant impact on the least-cost paths between nodes, larger cumulative effects can be expected, as the effects of decisions are not independent of each other (as illustrated by the stepping stones in Fig. 1). This also explains why the different algorithms tend to converge towards the optimal solution as the budget increases. The maximum value of ECA is reached when all options that maximize the connection probabilities between habitat patches have been chosen. After this point, additional options are only redundant. This emerging property actually corresponds to something we believe is a shortcoming of the ECA indicator which only takes into account the shortest path distance among pairs of nodes and not the number of short paths: redundancies taken into account by circuit connectivity measures can lead to more robust networks (McRae et al., 2008).

In practice, it also seems that cumulative effects appear more strongly when the problem is related to the restoration/conservation of links (corridors) as in the cases of Aude and Aix rather than nodes (patches) as in the cases of Marseille and Montreal. As we have illustrated our work with four very contrasted and complementary case studies, we believe that our results are rather robust and are not related to the modelling of the landscape (graph types) nor to the type of problem (conservation or restoration, node or link).

4.4. Practical guidelines

As we have seen, both local and greedy algorithms can lead to suboptimal solutions, especially when the budget is small, when the network is weakly connected, or when conservation/restoration options strongly impact the least-cost paths in the graph. We thus conclude that the optimal resolution is to be preferred whenever possible. However, the computation time of the optimal algorithm depends on a combination of factors and is not easy to predict accurately. To help conservation practitioners decide which algorithm is best to use in their case study, we therefore propose a rule of thumb as follows.

If the number of variables in the *Max-ECA* problem (i.e. the number of columns in the mixed integer program matrix) reaches one million, we believe that the execution of the optimal algorithm may become complicated. As an order of magnitude, the number of variables can be computed as the number of links multiplied by the number of nodes in the graph; to reach one million variables, the order of magnitude would be about 100 nodes for a complete graph and about 500 nodes for a planar graph. This cutoff is not firm however given that the preprocessing step can greatly reduce the number of variables. Trying to run the optimal solution on a subset of the problem first might therefore be the most reasonable approach to more precisely estimate the cutoff for the case study under investigation.

If the number of variables is slightly above one million, we suggest reducing the number of variables to fall within the previous case with a sequential approach. One can first use a decremental greedy algorithm under a budget constraint that is larger than the actual constraint (e.g. twice) to pre-select a subset of options. One can then run the optimal algorithm with the real budget on this subset of options (the set of solutions is therefore smaller, resulting in a shorter computation time); thanks to the preprocessing step, reducing the number of options will in this case also reduce the size of the graph used to run the optimal algorithm.

If the number of variables is largely above one million, we recommend using the decremental greedy algorithm, that can provide reasonable solutions. To ensure the quality of the solution, one can also run both greedy algorithms and keep the solution corresponding to the highest ECA value.

As a last resort, if the greedy algorithms do not run in a reasonable time, one could use both local algorithms and take the best solution. Note that these suboptimal algorithms could already be improved by solving exactly the knapsack instance, explained above, instead of selecting the options in descending order of their ECA gain/cost ratio.

The alternative solutions we propose here (sequential approach or decremental/incremental combination) deserve further investigation to better quantify the trade-offs between solution quality and computation time for a variety of conservation problems.

CRediT authorship contribution statement

All authors conceived the analyses. F.H. ran the experiments. F.H and C.H.A wrote the first draft of the manuscript. All authors contributed to finalize the manuscript and approved its final submission.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

The research on this paper was supported by Région Sud Provence-Alpes-Côte d'Azur and Natural Solutions. The Aix case study belongs to the Baum program (Biodiversity Urban Development Morphology) supported by the PUCA, the OFB and the DGALN. For stimulating exchanges on the case studies, we also thank: Patrick Bayle, Simon Blanchet, Aurélie Coulon, Andrew Gonzalez, Maria Dumitru, Jérôme Prunier, Bronwyn Rayfield, Benoit Romeyer and Keoni Saint-Pé.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biocon.2023.110066.

References

Albert, C.H., Rayfield, B., Dumitru, M., Gonzalez, A., 2017. Applying network theory to prioritize multispecies habitat networks that are robust to climate and land-use change. Conserv. Biol. 31, 1383–1396.

An, W., Liu, Y.-H., 2016. Keyplayer: an r package for locating key players in social networks. R J. 8 (1).

Awade, M., Boscolo, D., Metzger, J.P., 2012. Using binary and probabilistic habitat availability indices derived from graph theory to model bird occurrence in fragmented forests. Landsc. Ecol. 27, 185–198.

Beier, P., Spencer, W., Baldwin, R.F., McRAE, B.H., 2011. Toward best practices for developing regional connectivity maps. Conserv. Biol. 25 (5), 879–892.

Blazquez-Cabrera, S., Ciudad, C., Gastón, A., Simón, M., Saura, S., 2019. Identification of strategic corridors for restoring landscape connectivity: application to the Iberian lynx. Anim. Conserv. 22 (3), 210–219.

Bodin, Ö., Saura, S., 2010. Ranking individual habitat patches as connectivity providers: integrating network analysis and patch removal experiments. Ecol. Model. 221 (19), 2393–2405.

Gurobi Optimization, L., 2022. Gurobi Optimizer Reference Manual.

Crook, D.A., 2004. Is the home range concept compatible with the movements of two species of lowland river fish? J. Anim. Ecol. 73 (2), 353–366.

Crooks, K.R., Sanjayan, M., 2006. Connectivity Conservation, volume 14. Cambridge University Press.

Dilkina, B., Lai, K.J., Gomes, C.P., 2011. Upgrading shortest paths in networks. In: International Conference on AI and OR Techniques in Constriant Programming for Combinatorial Optimization Problems. Springer, pp. 76–91.

Erös, T., Lowe, W.H., 2019. The landscape ecology of rivers: from patch-based to spatial network analyses. Curr. Landsc. Ecol. Rep. 4 (4), 103–112.

Fall, A., Fortin, M.-J., Manseau, M., O'Brien, D., 2007. Spatial graphs: principles and applications for habitat connectivity. Ecosystems 10 (3), 448–461.

Foley, M.M., Mease, L.A., Martone, R.G., Prahler, E.E., Morrison, T.H., Murray, C.C., Wojcik, D., 2017. The challenges and opportunities in cumulative effects assessment. Environ. Impact Assess. Rev. 62, 122–134.

Hamonic, F., Albert, C., Couëtoux, B., Vaxès, Y., 2023. Optimizing the ecological connectivity of landscapes. Networks 81 (2), 278–293.

Hodgson, J.A., Wallis, D.W., Krishna, R., Cornell, S.J., 2016. How to manipulate landscapes to improve the potential for range expansion. Methods Ecol. Evol. 7 (12), 1558–1566.

Lawler, J.J., 2009. Climate change adaptation strategies for resource management and conservation planning. Ann. N. Y. Acad. Sci. 1162 (1), 79–98.

Magris, R.A., Treml, E.A., Pressey, R.L., Weeks, R., 2016. Integrating multiple species connectivity and habitat quality into conservation planning for coral reefs. Ecography 39 (7), 649–664.

McRae, B.H., Dickson, B.G., Keitt, T.H., Shah, V.B., 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89, 2712–2724.

Nemhauser, G., Wolsey, L., 1988. In: The Scope of Integer and Combinatorial Optimization, Chapter I.1. John Wiley & Sons, Ltd, pp. 1–26.

Paradis, E., Baillie, S.R., Sutherland, W.J., Gregory, R.D., 1998. Patterns of natal and breeding dispersal in birds. J. Anim. Ecol. 67 (4), 518–536.

Pereira, M., Segurado, P., Neves, N., 2011. Using spatial network structure in landscape management and planning: a case study with pond turtles. Landsc. Urban Plan. 100, 67–76.

Pereira, J., Saura, S., Jordán, F., 2017. Single-node vs. Multi-node centrality in landscape graph analysis: key habitat patches and their protection for 20 bird species in ne Spain. Methods Ecol. Evol. 8, 1458–1467.

Rubio, L., Bodin, O., Brotons, L., Saura, S., 2015. Connectivity conservation priorities for individual patches evaluated in the present landscape: how durable and effective are they in the long term? Ecography 38, 782–791.

Saint-Pé, K., 2019. In Situ Quantification of Brown Trout Movements. Université Paul Sabatier-Toulouse III. PhD thesis.

Saura, S., Pascual-Hortal, L., 2007. A new habitat availability index to integrate connectivity in landscape conservation planning: comparison with existing indices and application to a case study. Landsc. Urban Plan. 83, 91–103.

Saura, S., Estreguil, C., Mouton, C., Rodríguez-Freire, M., 2011. Network analysis to assess landscape connectivity trends: application to European forests (1990–2000). Ecol. Indic. 11, 407–416.

Schivo, F., Mateo-Sánchez, M.C., Bauni, V., Quintana, R.D., 2020. Influence of land-use/land-cover change on landscape connectivity for an endemic threatened amphibian (argenteohyla siemersi pederseni, anura: Hylidae). Landsc. Ecol. 35 (6), 1481–1494.

Segurado, P., Branco, P., Ferreira, M.T., 2013. Prioritizing restoration of structural connectivity in rivers: a graph based approach. Landsc. Ecol. 28 (7), 1231–1238.

Sethi, S.A., O'Hanley, J.R., Gerken, J., Ashline, J., Bradley, C., 2017. High value of ecological information for river connectivity restoration. Landsc. Ecol. 32 (12), 2327–2336.

Tarabon, S., Bergès, L., Dutoit, T., Isselin-Nondedeu, F., 2019. Maximizing habitat connectivity in the mitigation hierarchy. A case study on three terrestrial mammals in an urban environment. J. Environ. Manag. 243, 340–349.

Urban, D., Keitt, T., 2001. Landscape connectivity: a graph-theoretic perspective. Ecology 82, 1205–1218.

Wauters, L.A., Verbeylen, G., Preatoni, D., Martinoli, A., Matthysen, E., 2010. Dispersal and habitat cuing of eurasian red squirrels in fragmented habitats. Popul. Ecol. 52 (4), 527–536.

Whitehead, A.L., Kujala, H., Wintle, B.A., 2017. Dealing with cumulative biodiversity impacts in strategic environmental assessment: a new frontier for conservation planning. Conserv. Lett. 10 (2), 195–204.

Wu, X., Sheldon, D.R., Zilberstein, S., 2014. Stochastic network design in bidirected trees. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems, volume 27. Curran Associates Inc, pp. 882–890.

Xue, Y., Wu, X., Morin, D., Dilkina, B., Fuller, A., Royle, J.A., Gomes, C.P., 2017. Dynamic optimization of landscape connectivity embedding spatial-capture-recapture information. In: 31st AAAI Conference on Artificial Intelligence, 31, pp. 4552–4558.

Yemshanov, D., Haight, R.G., Koch, F.H., Parisien, M.-A., Swystun, T., Barber, Q., Burton, A.C., Choudhury, S., Liu, N., 2019. Prioritizing restoration of fragmented landscapes for wildlife conservation: a graph-theoretic approach. Biol. Conserv. 232, 173–186.