Pesquisa Florestal Brasileira
Brazilian Journal of Forestry Research
http://pfb.cnpf.embrapa.br/pfb/
e-ISSN: 1983-2605
Article

Definition of tourism itineraries in a Brazilian conservation unit using artificial intelligence

Carlos Alberto Araújo Júnior[1]*, Helio Garcia Leite[2], João Batista Mendes[3]

[1]Universidade Federal de Minas Gerais, Instituto de Ciências Agrárias, Avenida Universitária, 1.000, Bairro Universitário, CEP 39404-547, Montes Claros, MG, Brasil

[2]Universidade Federal de Viçosa, Departamento de Engenharia Florestal, Avenida Purdue, s/nº, Campus Universitário, CEP 36570-900, Viçosa, MG, Brasil

[3]Universidade Estadual de Montes Claros, Departamento de Ciências da Computação, Avenida Rui Braga, S/Nº – Vila Mauricéia, CEP 39401-089, Montes Claros, MG, Brasil

***Corresponding author:**
carlosaraujo@ufmg.br

Index terms:
Geographical information systems
Forest conservation
Operations research

Termos para indexação:
Sistemas de informação geográfica
Conservação florestal
Pesquisa operacional

Received in 21/04/2024
Technical approval in 21/11/2024
Final approval in 10/01/2025
Published in 20/02/2025

Abstract - The Pandeiros River Environmental Protection Area is an important Brazilian conservation unit used for ecotourism. However, there is a lack of research guiding decision-making regarding tourist movements. The objective of this study was to evaluate the use of a simplified version of the clonal selection metaheuristic for optimizing tourist itineraries. Thirty-one tourist sites were considered, with routes starting from three origins. A mathematical model based on the vehicle routing problem is proposed. This problem was solved using the branch and bound, clonal selection, and simulated annealing algorithms, and the proposed simplification for the clonal selection metaheuristic. Random solutions were evaluated to simulate tourist behaviour. Random solutions yield the worst results. The proposed simplification produced better results for itineraries starting from two origins. It provided an average reduction of 42% in the total distance of tourist itineraries and a 17% reduction in the use of available road networks.

Definição de roteiros turísticos em uma unidade de conservação brasileira, utilizando inteligência artificial

Resumo - A Área de Proteção Ambiental do Rio Pandeiros é uma unidade de conservação brasileira utilizada para ecoturismo. No entanto, faltam pesquisas que orientem a tomada de decisões em relação aos movimentos de turistas. Objetivou-se avaliar a utilização de uma versão simplificada da seleção clonal metaheurística para a otimização de roteiros turísticos. Foram considerados trinta e um pontos turísticos, com roteiros partindo de três origens. Um modelo matemático baseado no problema de roteamento de veículos foi proposto e resolvido utilizando os algoritmos *branch and bound*, *seleção clonal* e *análise de recozimento*, bem como a simplificação proposta para a seleção clonal metaheurística. Foram avaliadas soluções aleatórias para simular o comportamento do turista. Soluções aleatórias produziram os piores resultados. A simplificação proposta produziu melhores resultados para itinerários partindo de duas origens. Proporcionou uma redução média de 42% na distância total dos roteiros turísticos e uma redução de 17% na utilização das redes rodoviárias disponíveis.

https://doi.org/10.4336/2025.pfb.45e202402303

Introduction

The Pandeiros River basin is located in the northern region of Minas Gerais State, Brazil, on the left bank of the São Francisco River. It includes the municipalities of Januária, Cônego Marinho, and Bonito de Minas (Santos et al., 2020). It is part of the Water Resources Planning and Management Unit of Pandeiros River (UPGRH-SF9) under IGAM (Minas Gerais Water Management Institute) and consists of an Environmental Protection Area (APA Pandeiros) and a Wildlife Refuge (REVS Rio Pandeiros).

Pandeiros River is approximately 145 km in length and is the only river with a swamp area in the state of Minas Gerais. It is located in a region that serves as a transition between the Cerrado (Brazilian savanna) and Caatinga biomes (Rezende et al., 2012). According to Souza et al. (2008) and Nunes et al. (2009), the river is considered a natural fish nursery, responsible for almost 70% of São Francisco River's fish procreation. This mainly occurs in the swamp area, which covers approximately 3,000 ha and is composed of interconnected lakes that expand during rainy periods.

One possible way to conserve the environment and the scenic aspects of this important area is by promoting ecotourism. It can also serve as a complementary source of income for local communities. This activity involves the use of natural resources but is designed to minimize negative environmental impacts, as it occurs in areas with native vegetation, allowing visitors to observe animals and plants in their natural environment (Freitas & Portuguez, 2014).

However, many tourist routes are often designed based on previous experiences and intuition. Therefore, it is overly optimistic to believe that an optimal solution or a solution close to the best can be consistently obtained by a single individual every time such a plan is created (Karagul & Gungor, 2014). Consequently, extensive research has advocated the use of mathematical programming models for determining the best routes in tourist visitation problems, as mentioned in Brach & Górski (2014) and Gavalas et al. (2014).

In most cases, these studies are conducted in the context of the tourist trip design problem (TTDP), which falls within the category of planning problems aimed at identifying the optimal route for tourists interested in visiting multiple destinations (Gavalas et al., 2014). The TTDP can be understood as a variation of the vehicle routing problem (VRP). VRP is one of the most extensively studied problems in combinatorial optimization and its variations, often belonging to the nondeterministic polynomial time (NP-hard class), with clear practical applications (Silva Júnior & Lopes, 2011; Kantawong & Pravesjit, 2020).

Numerous papers have been dedicated to developing alternatives for solving vehicle routing problems, including those by Shukla & Jharkharia (2013), Karagul & Gungor (2014), Shui et al. (2015), Anggodo et al. (2016), Ogiolda (2017), Hosseinabadi et al. (2018), Silva et al. (2018), Kantawong & Pravesjit (2020), and Rozidi et al. (2021). One common characteristic of these papers is the use of metaheuristics, a tool from the field of artificial intelligence that has proven effective in addressing complex problems. One of its advantages is the ability to achieve near-optimal solutions without a complete understanding of the problem's internal dependencies (Mrówczyńska et al., 2019).

One aspect of mathematical modelling for optimization is its complexity. In addition to the inherent complexity of the problem itself, challenges arise from defining parameter values and dealing with uncertainty, which significantly influence decision-making. Complex models, while striving to represent reality, such as the model proposed by Liao & Zheng (2018), may require extensive processing time and rely on data that is difficult to acquire. On the other hand, overly simplistic models may overlook critical aspects of the problem. The reality is that there is a need to simplify the inputs of the models and obtaining outputs quickly.

Therefore, our objective is to propose the use of a simplified version of the clonal selection algorithm, a metaheuristic, to address the problem of route planning for tourist visits. This problem involves minimizing the total distance required to visit all tourist destinations within a conservation unit while considering time restrictions.

Material and methods

Study area

This study was conducted in the Environmental Protection Area of Rio Pandeiros, located in the northern region of Minas Gerais State. Its territory covers 393,866 ha, making it the largest sustainable use conservation unit in the state. The climate in the region is classified as Aw (Tropical Savanna) according to the Köppen classification. The average annual temperature ranges from 18.4 °C to 31.9 °C, and the average annual precipitation is 876 mm, based on data from the meteorological station in the city of Januária, MG for the years 1988 to 2018, obtained from the National Institute of Meteorology (INMET).

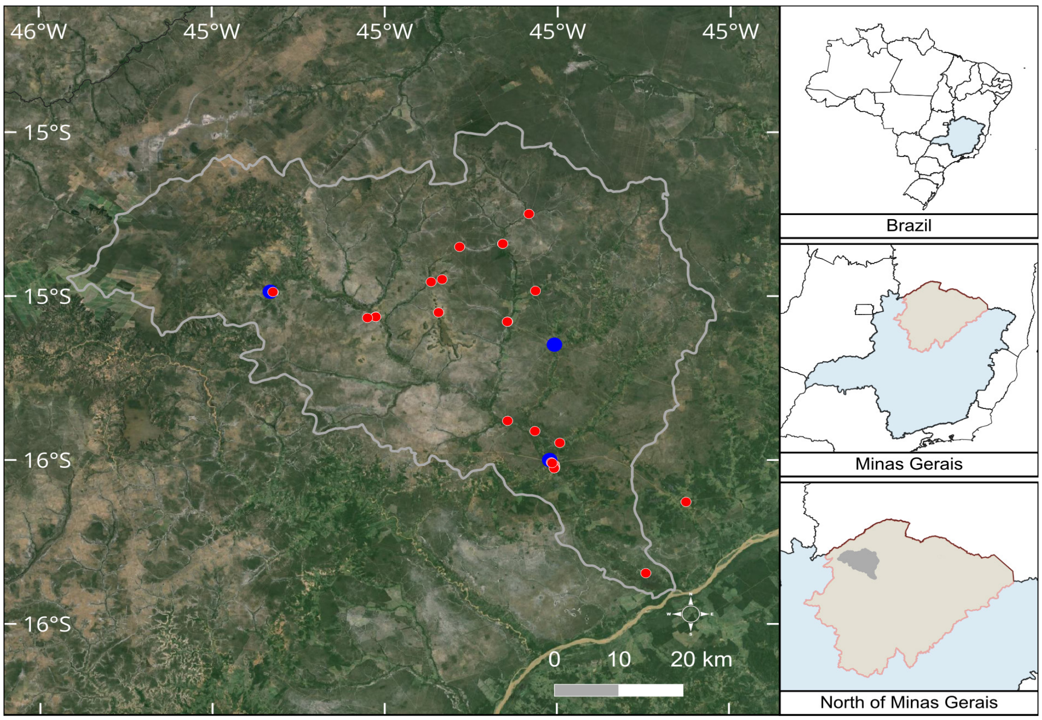

Figure 1. Location map of Environmental Protection Area of Rio Pandeiros. Red dots are places of tourist interest and blue dots are departing places for tourist itineraries.

Places of tourist interest

After conducting a literature review, consulting government websites, and conducting field visits, 31 tourist attractions were identified in the State Environmental Protection Area of Rio Pandeiros. These include five bathing areas, three waterfalls, one town, three villages, and seventeen other locations with tourist visitation potential. The highest number of attractions is located in the district of Pandeiros, where the Pandeiros River bathing area and the three waterfalls are situated.

A survey with 21 interviewees was conducted to estimate the time spent at each tourist attraction. An image of each location, along with its description, was presented to the analyst, who was asked to indicate the amount of time a person would like to spend at that location during a tourist visit. The visitation time for each attraction was then calculated by averaging the values provided.

Three alternatives for the tourists' starting point upon arrival in the Environmental Protection Area (APA) were evaluated. These locations were chosen on the basis of the existing infrastructure, which allows tourists to be accommodated more comfortably in terms of the possible services available, such as restaurants, hotels, and supermarkets. Therefore, the assessment of itineraries departing from Pandeiros, Bonito de Minas, and Várzea Bonita was chosen.

Vehicle routing problem

The vehicle routing problem (VRP) is a classic optimization challenge that has been widely studied due to its complexity and practical applications of the generated solutions. The problem at hand involves determining which tourist attractions should be included in a given route, similar to the problem addressed by Rozidi et al. (2021). It is assumed that the tourist will be accommodated at a pre-defined location (usually a hotel or inn), will leave in the morning to explore other places, and must return to the starting point before exceeding the daily time limit. Attractions not visited in one day should be included in another route, to be taken on a different day.

Other considerations assumed for modelling the routing problem were as follows (Equations 1-8):

1. There is a specified number of available days for conducting tourist visitation routes, and as many routes as necessary can be generated. Each route represents a day of visit.
2. The vehicle's capacity constraint is related to the available time for its use in a single day.
3. Each tourist attraction must be visited only once and included in a single route.
4. When visiting a particular location, the vehicle will remain stationary for the time required for the tourist to engage in recreational activities. This time can be understood as the location's demand and is subtracted from the vehicle's capacity.
5. The time spent traveling between locations is also subtracted from the vehicle's capacity.
6. The vehicle must return to the starting point before the available time expires.

Thus, the mathematical model for the VRP in this study can be described as follows:

$$Z = \text{Min} \sum_{i=0}^{n} \sum_{j=0}^{n} X_{ij} C_{ij} \quad (1)$$

$$\sum_{i=0}^{n} X_{ik} = 1 \quad \forall\, k = 1, \dots, n \quad (2)$$

$$\sum_{i=0}^{n} X_{ik} = 1 \quad \forall\, k = 1, \dots, n \quad (3)$$

$$\sum_{j=1}^{n} X_{0j} = \sum_{j=1}^{n} X_{i0} \quad (4)$$

$$q_k = T_k + \sum_{i=0}^{n} T_{ik} X_{ik} + T_{k0} \quad \forall\, k = 1, \dots, n \quad (5)$$

$$\sum_{i=0}^{n} F_{ik} - \sum_{j=0}^{n} F_{kj} = q_k \quad \forall\, k = 1, \dots, n \quad (6)$$

$$F_{ik} \leq Q X_{ij} \quad \forall\, i, j = 0, 1, \dots, n \quad (7)$$

$$X_{ij} \in \{0,1\} \; \forall\, i, j \quad (8)$$

Where: X_{ij} is the binary decision variable, equal to 1 when the arc (i, j) is traversed by the vehicle and 0 otherwise; C_{ij} is the distance between locations i and j (in km); the index k represents the current location, i corresponds to the location before location k, and the index j corresponds to the location after location k; the depot (or the vehicle's starting point) is represented by the value 0 for the indices i and j; n is the total number of locations or vertices in the graph; T_k represents the time required for staying at location k (in min); q represents the time demand for visiting location k, considering round-trip transit time and stay time (in min); T_{ik} is the travel time from the previous location to the current location (in min); Q is the vehicle's capacity, corresponding to the available time between departure and return to the route's point of origin (in min).

The considered model minimizes the total distance (in km) travelled to visit all points of interest (Equation 1). The first and second constraints (Equations 2 and 3) ensure that each point of visitation, different from the origin, is assigned to a single route (or visitation day). The third constraint (Equation 4) ensures that the routes start at the depot and end at the same location. Equation 5 calculates the time required for visiting any location, which is composed of the time needed for the tourist's stay, the time of the route from the origin to location k, and the return time to the depot. The constraint represented by Equation 6 ensures that the time available to the tourist is allocated for visiting each location. Equation 7 indicates that the available time must be greater than the time that will be spent until the next location, and equation 8 guarantees that the decision variable is binary.

Random solution generation (RND)

To simulate the behaviour of tourists when planning a trip, various random solutions for visiting all points of interest in the Environmental Protection Area were generated. The construction of these solutions involved the random allocation of locations along a route that began at the defined starting point. To ensure the feasibility of each solution, the time required for travel between the route's origin and the current visiting point (Li), added to the time spent at that location, was calculated and combined with the time needed to reach the next location (Lj), stay there, and return immediately to the origin of the route. If the total calculated time was less than the available daily visitation time, the inclusion of location Li in the route was accepted; otherwise, that location would be the first to be visited after departing from the origin in the following route. So, 500 random solutions were generated for each starting location and compared with the solutions obtained from optimization algorithms.

Clonal selection algorithm (CSA)

The metaheuristic clonal selection algorithm was proposed by Castro & Von Zuben (2002) and has since been used to seek high-level solutions for combinatorial optimization problems. Its operation is inspired by the immune system, abstracting the concepts of the relationship between antigens and antibodies.

In the context of an optimization problem, solutions can be treated as antibodies. A population of solutions is randomly created. The best solutions, those with the highest values for the evaluation function, are retained and cloned. During the evolution of the optimization system, random mutations occur in the solutions, leading to the generation of individuals different from those that exist. Some solutions may become degenerate in this step, meaning that not all problem constraints are satisfied. In such cases, the fitness function can be penalized, reducing the chances of these solutions remaining in the population of the next generation. The processes of selection, cloning, and mutation is repeated until a stopping criterion is met.

In the case of the proposed VRP, each solution represents a travel itinerary for visiting all tourist points. Itineraries with lower total distance values are kept in the population and evolve to find increasingly better solutions. Constraints on the time spent at locations and the total daily route time are considered when developing initial solutions to ensure a completely viable initial population. When the hyper mutation process occurs, solutions that do not meet the constraints are penalized.

The algorithm's execution considered a population of antibodies with a size (N) of 30, the selection of the top 3 solutions, a cloning coefficient (β) of 0.10, and a stopping condition of 500 generations. These values were determined using preliminary tests. In each generation, the number of clones (N_c) for each of the three selected antibodies was defined using Equation 9 (Ogiolda, 2017), where i represents the index of the antibody in a descending list based on adaptation level.

$$N_c = round((\beta . N)/i) \tag{9}$$

The mutation rate (αc) was calculated for each clone (Equation 11), considering a ratio (ρ) equal to 4 and the significance of each antibody (f_c) from Equation 10. The number of mutations (N_{mc}) was determined as equation 12, where L is the number of positions in the location vector.

$$f_c = \frac{fit_c}{\sum_{c=1}^{N} fit_c} \tag{10}$$

$$\alpha_c = \left(\frac{1}{\rho}\right) e^{-f_c} \tag{11}$$

$$N_{mc} = int(\alpha_c . L) \tag{12}$$

Therefore, the new population is composed of the best individuals and clones after mutation. If the number of antibodies is less than the population size defined, the antibodies not selected in the previous population were ranked in terms of fitness and chosen without cloning or mutation.

Simplified clonal selection algorithm (SCS)

The proposed simplification involves considering fixed values for the number of clones per antibody and the number of mutations per clone in every generation. Furthermore, the next population is composed only of the selected individuals and the generated clones, completely discarding the antibodies from the previous generation that were not among the best adapted for solving the problem.

Thus, an initial population of 500 antibodies was randomly generated. From these, three were selected for the next generation, with each of them being cloned nine times (fixed cloning rate, independent of the antibody's adaptation level). Hyper mutation was applied to each of the generated clones, involving only a random exchange of positions between two locations on different days of visitation (fixed mutation rate).

The new population was then defined as having thirty antibodies, of which the top three were selected, and so on until the algorithm completed 500 iterations. These values were chosen to maintain compatibility with the algorithm described earlier. The best solution obtained at the end of processing was considered for analysis. For each starting point, fifty repetitions of the proposed algorithm were executed.

Simulated annealing (SA)

We chose to construct an algorithm based on the simulated annealing metaheuristic, similar to what was done in Shukla & Jharkharia (2013). The algorithm begins by obtaining an initial random and feasible solution (S_i). This solution is used to create a new neighbouring solution by modifying a segment of the route (S_j). The neighbouring solution is compared to the initial solution and becomes the current solution if it has a fitness value lower than that of S_i. If the fitness value of S_j is greater than S_i, a random value (Rnd_j) is generated, and the acceptance probability (P_{ac}) of this solution, considered inferior, is calculated using the Equation 13, where T is the current system temperature.

$$P_{ac} = e^{\frac{S_i - S_j}{T}} \quad (13)$$

If Rnd_j is less than P_{ac}, S_j becomes the basis for generating a new neighbouring solution in the next iteration; otherwise, the solution S_j is discarded.

Based on preliminary tests, the initial system temperature was set to 10^7 °C, the cooling rate was 5% per iteration, and each repetition had 15,500 iterations to obtain a quantity of analyzed solutions close to that considered for the clonal selection algorithm.

Performing calculations

A time limit of 600 min was considered for each day of visitation, meaning that the route starts at 8:00 AM and must finish by 6:00 PM on the same day. The same time window was adopted for all tourist points because, in practice, there are no restrictions on the timing of visiting the locations throughout the day.

The Lingo software was used to execute the branch & bound algorithm to evaluate the possibility of obtaining an optimal solution for each instance. A runtime of 20 min was arbitrarily defined, and the best solution found within this time frame was stored, even if it was not the optimal solution.

The execution of the metaheuristics considered a penalty criterion for infeasible solutions. In this case, for each minute beyond the available time for the daily route, a penalty factor of a thousand units was applied. For example, if the route consumed 602 min, the total distance was penalized by 2000 km. The penalty should be reasonably high to ensure that infeasible solutions do not persist for many iterations.

The SCS, CSA, and SA algorithms were executed fifty times using a computer with an Intel(R) Core (TM) i7-8565U CPU @ 1.80GHz 1.99 GHz processor. The codes were implemented in custom programs written in the Java language.

Evaluation of the results

Because the optimal solution for the problem at hand cannot be obtained using classical and deterministic algorithms, we chose to compare the best solution obtained at the end of the processing with the one obtained after the random generation process. In this case, the latter can be considered as a possible itinerary that a tourist would plan without the use of an optimization system, essentially through trial and error.

Evaluations were conducted in terms of differences in the total lengths of the generated itineraries, road network utilization, and the behaviour of the solutions obtained by the metaheuristics. The Wilcoxon test was conducted to evaluate whether there were significant differences between the medians of the total distances obtained from each execution of the algorithms.

Results

The mean coefficients of variation for the different instances were 8.2% for randomly obtained solutions (RND), 6.2% for simplified clonal selection algorithm (SCS), 5.9% for clonal selection algorithm (CSA), and 6.6% for simulated annealing (SA). In all cases, the randomly obtained solutions performed worse than those generated by the evaluated algorithms (Table 1), with the best results achieved by SCS, except for Várzea Bonita, where the CSA algorithm provided the best solution.

The best solutions for SCS were obtained after 222 iterations when Bonito de Minas was considered as the starting point, 171 iterations for Pandeiros, and 243 iterations for Várzea Bonita (Figure 2). The best solutions for SA were obtained after 2,646 iterations when Bonito de Minas was considered as the starting point, 2,537 iterations for Pandeiros, and 2,301 iterations for Várzea Bonita. For CSA, these values were 345, 406, and 361, respectively. The high fitness values for the SA algorithm in the initial iterations indicate the application of penalties to the fitness function. The average time for each repetition was 99 milliseconds for SCS, 116 milliseconds for CSA, and 266 milliseconds for SA.

Table 1. Results obtained (in terms of total distance) using each algorithm in the proposed instances.

Algorithm	Origins								
	Bonito de Minas			Pandeiros			Várzea Bonita		
	Min (km)	Avg (km)	Max (km)	Min (km)	Avg (km)	Max (km)	Min (km)	Avg (km)	Max (km)
RND	855	1,210	1,496	882	1,192	1,481	1,245	1,611	2,026
SCS	496	557	655	460	546	624	808	899	970
CSA	510	586	667	479	562	638	778	917	1,050
SA	527	600	678	478	597	684	869	979	1,090

RND = random solutions, CSA = clonal selection algorithm, SCS = simplified clonal selection algorithm, SA = simulated annealing, Min = minimum, Avg = average, Max = maximum.

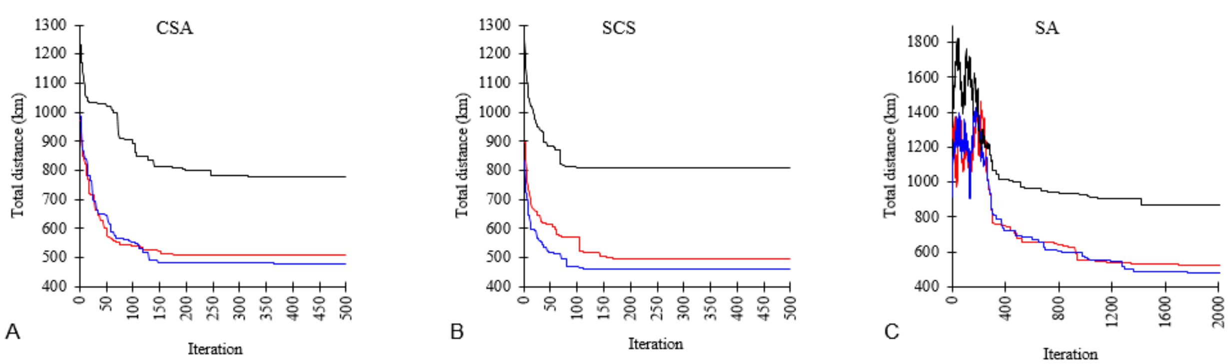

Figure 2. Evolution of the best solutions found at each iteration of the metaheuristics for each evaluated starting point (black line: Várzea Bonita; blue line: Pandeiros; red line: Bonito de Minas). a) Clonal selection algorithm; b) simplified clonal selection algorithm; c) simulated annealing.

For all starting points considered, the results obtained after applying the SCS algorithm showed an average reduction of 42% in the total distance and 16% in the total time used for visiting the tourist attractions (Table 2). The average gains provided by CSA were 41% and 16%. On the other hand, SA showed reductions in the order of 38% and 15% for the same parameters. The greatest gains in terms of total distance travelled occurred for routes that started in Pandeiros, with the smallest gain for those starting in Várzea Bonita.

The longer distance required to complete the proposed routes starting from Várzea Bonita resulted in a longer time for tourist visits, necessitating an increase in the number of days for visitation compared with that proposed for departures from Pandeiros and Bonito de Minas. Using the location of Pandeiros as the starting point proved to be the best alternative, both in terms of the number of days required and the total distance travelled.

The solutions found by the metaheuristics showed a reduction in the road network used, with an average savings of 17% for SCS, 28% for CSA, and 27% for SA (Figure 3). The results obtained by the branch and bound algorithm indicate the use of a road network that is 23% smaller (on average) than the best solutions obtained randomly.

Despite having a relatively low coefficient of variation for the set of solutions found by each algorithm, there is a considerable distinction between the longest and shortest distances obtained in repeated executions. The attempts that yielded the best results showed an average reduction of 26.4% in the total distance travelled compared with the worst solutions for the same algorithm and starting location, representing an average decrease of 263 km in the total itinerary required for tourist visits.

The results of the Wilcoxon test indicate that the SCS algorithm outperformed the other metaheuristics for all three departures (Table 3). There was a significant difference between the SCS and CSA algorithms only for the Bonito de Minas departure. The medians of SCS were statistically different from those of SA for all departures. Additionally, there were significant differences between the CSA and SA algorithms for the Pandeiros and Várzea Bonita departures.

Table 2. Results of the best solutions obtained for each algorithm used in the study.

Departure	Criteria	RND	CSA	SCS	SA	B&B
Bonito de Minas	Number of routes	7	7	7	8	7
	Total distance (km)	855	510	496	527	560
	Length of the road network (km)	351	228	323	295	276
Pandeiros	Number of routes	8	7	7	7	6
	Total distance (km)	882	479	460	478	493
	Length of the road network (km)	379	318	326	237	287
Várzea Bonita	Number of routes	8	8	8	9	-
	Total distance (km)	1,245	778	808	869	-
	Length of the road network (km)	465	314	328	339	-

RND = random solutions, CSA = clonal selection algorithm, SCS = simplified clonal selection algorithm, SA = simulated annealing, B&B = branch and bound algorithm.

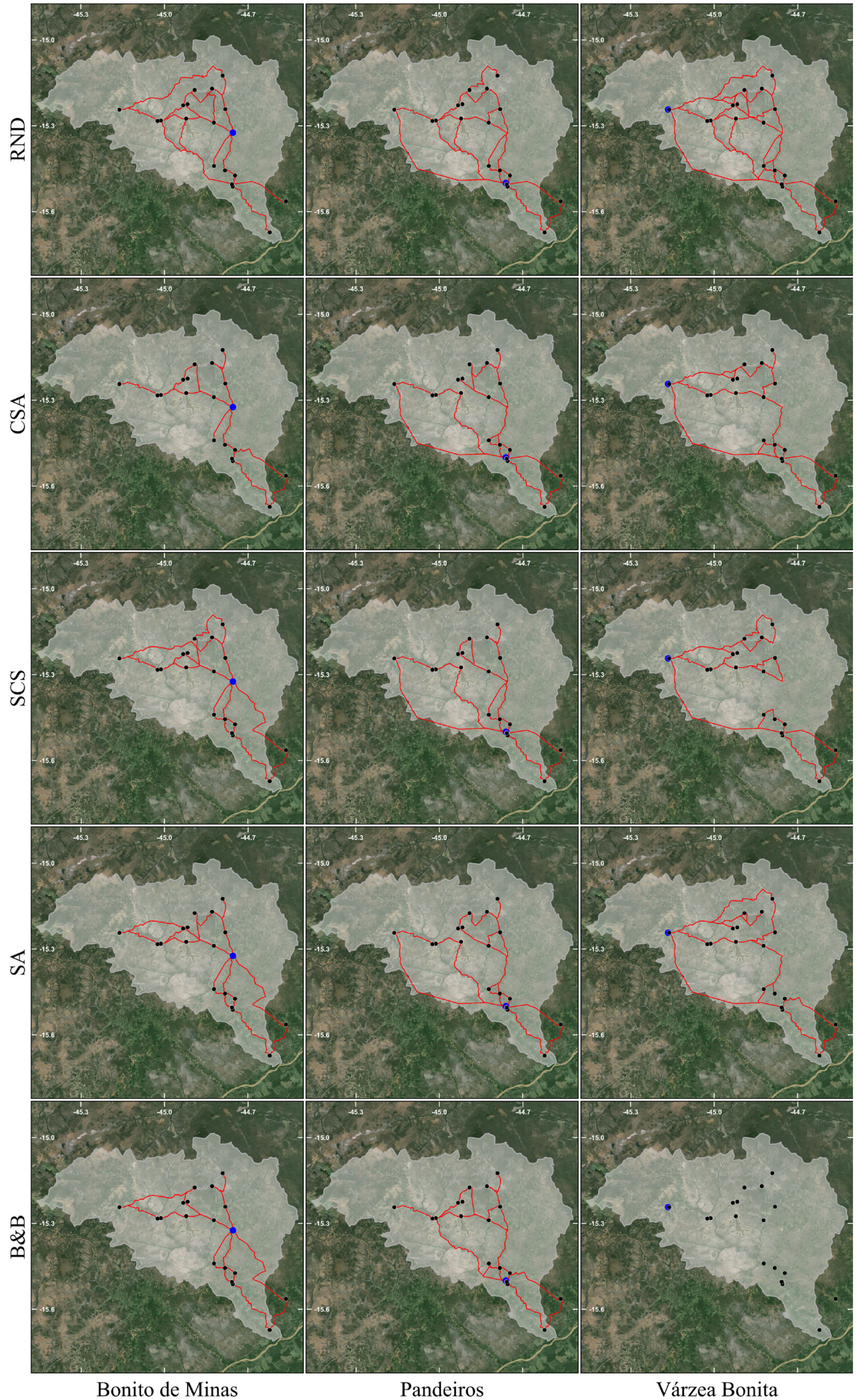

Figure 3. Routes obtained with the best solutions for each algorithm. Red lines are the roads used in each itinerary. Black dots are places for tourism visiting and blue dots are departure places.

Table 3. P-values from the Wilcoxon test comparing the results obtained for each algorithm across the different departure locations.

Departure	Algorithm	Median (km)	Algorithm			
			RND	CSA	SCS	SA
Bonito de Minas	RND	1,208.65	1.0000			
	CSA	579.86	<2e-16	1.0000		
	SCS	547.58	<2e-16	0.0151	1.0000	
	SA	604.08	<2e-16	1.0000	0.0001	1.0000
Pandeiros	RND	1,185.85	1.0000			
	CSA	563.42	<2e-16	1.0000		
	SCS	547.88	<2e-16	1.0000	1.0000	
	SA	604.23	<2e-16	0.0014	<4.1e-06	1.0000
Várzea Bonita	RND	1,613.11	1.0000			
	CSA	923.55	<2e-16	1.0000		
	SCS	902.90	<2e-16	1.0000	1.0000	
	SA	984.92	<2e-16	0.0002	<5.4e-08	1.0000

RND = random solutions, CSA = clonal selection algorithm, SCS = simplified clonal selection algorithm, SA = simulated annealing.

Discussion

Despite having a relatively low coefficient of variation for the set of solutions found by each algorithm, there is a considerable distinction between the longest and shortest distances obtained in repeated executions. The attempts that yielded the best results showed an average reduction of 26.4% in the total distance travelled compared with the worst solutions for the same algorithm and starting location, representing an average decrease of 263 km in the total itinerary required for tourist visits.

Low coefficients of variation indicate that the algorithms used have stable behaviours in terms of converging to good solutions (Mrówczyńska et al., 2019), or at least to similar solutions. This is highly desirable when it's not possible to run hundreds or thousands of repetitions. However, when considering the practical use of these algorithms for creating tourist itineraries, it's reasonable to believe that a sufficient number of repetitions should be performed before the final solution is presented to the user. Otherwise, there is a risk that the result of a new execution requested by the tourist may be considerably different from that obtained in the previous execution. Results with differences of 263 km can influence the tourist's decision regarding the use of a route planning system or not.

Although there is such a difference between the best and worst routes provided by different algorithm executions, the results presented are still better than those obtained from the set of solutions generated randomly. The average difference in this case is 673 km. This suggests the difficulty that a tourist may encounter when trying to create an itinerary randomly, which could even make it impossible to visit all the tourist spots due to the traveller's limited time. Karagul & Gungor (2014) mention that mathematical models produce better solutions in less time than those based on human experience and insights that people may have when trying to solve a problem intuitively.

One of the aspects that introduce variability into many existing metaheuristics is the generation of the initial population. Since these algorithms often aim to solve combinatorial problems, providing good initial solutions can enhance their performance, a fact already discussed by Araújo Júnior et al. (2021). This is particularly important for the clonal selection algorithm, as it benefits from its memorization capability (Hassen et al., 2019), retaining the best solutions as part of the adaptation process for the next population. In preliminary tests, it was observed that some individuals in the initial population led to worse final results, all other parameters being constant, justifying the choice of an initial population of five hundred individuals.

Simulated annealing is not affected by this decision because there is a high probability that inferior solutions will be accepted in the initial iterations of the algorithm. In this case, even if the process of generating the initial solution is well-crafted to provide an excellent starting point for the search in the solution space, the effort may be lost in the first evaluations of the neighbourhood.

The difference between the considered algorithms leads to the need to evaluate the number of iterations required in each case. It was observed that the two versions of clonal selection (simplified clonal selection algorithm - SCS and clonal selection algorithm - CSA) showed much faster convergence than simulated annealing (Figure 1) in all three cases studied. While the former tends to stabilize with just over 150 iterations, this only occurred after 2,300 iterations for the latter. Shukla & Jharkharia (2013) mentioned that artificial immune system algorithms have demonstrated superiority in terms of convergence when compared with other algorithms, including simulated annealing and genetic algorithm.

In each iteration, an algorithm analyzes a different number of solutions, with thirty solutions for SCS and CSA and only one for simulated annealing (SA). To compare them fairly, it is reasonable to consider the same number of solutions evaluated in each run (Mrówczyńska et al., 2019). Even though this was done, the practical result was a faster execution for SCS, with a processing time 63% shorter than that of SA, followed by CSA with a 56% reduction in average processing time compared to SA. Indeed, algorithms in the field of artificial immune systems have shown excellent results for the vehicle routing problem, as evidenced in the works of Yu & Lau (2013), Mrówczyńska et al. (2019), Hassen et al. (2019), Ogiolda (2017), and Shukla & Jharkharia (2013).

The time required to obtain a solution is a crucial factor in combinatorial problems, and in such cases, waiting for the optimal solution provided by deterministic algorithms may not be the best option. In fact, the solution presented by the branch and bound algorithm (non-optimal) for each instance analyzed in this study was inferior to the solutions provided by the metaheuristics. Moreover, the time required to find a feasible solution was much longer, with the noteworthy fact that no feasible solution was even presented when the problem involved creating tourist routes starting from the location of Várzea Bonita. Because of these challenges, research in the field of heuristics and metaheuristics remains highly active today (Armas & Melian-Batista, 2015).

The worst solutions found by SCS showed an average improvement of 25% compared with the best solutions obtained randomly. This improvement was 22% for CSA and 19% for SA. In other words, considering 50 repetitions, the metaheuristics significantly contribute to reducing the distances required for tourist visitation in the area considered, even when taking the worst routes into account.

The gains presented are even greater when comparing the best solutions of each algorithm with the best solution obtained randomly. SCS provided an average reduction of 42%, CSA reduced it by 41%, and SA by 38% in the total distance travelled. Indeed, there is no question about the use of optimization models to reduce costs or distances, as is the case discussed here. We emphasize how important it is for the knowledge and mathematical models researched to be applied in society, minimizing waste, and improving the quality of visits to natural environments.

The comparison between the metaheuristics used shows that SA presented solutions, on average, 5.5% worse than SCS and 4.5% worse than CSA. Mrówczyńska et al. (2019) presented results where SA was up to 23% worse than CSA for a vehicle routing problem with time windows.

In addition to the reduction in the total distance travelled, optimization indicates that the available road network can be reduced. Compared with the roads used by the best randomly generated solutions, the SCS algorithm provides an average reduction of 17%, CSA of 28%, and SA of 27%. This reduction is related to the repetition of road segments over the visits. From the perspective of managing the conservation area, it is better for fewer roads to be used, allowing for better maintenance of the area, reducing the impacts related to vehicle traffic, and favouring tourist control measures. Other benefits associated with optimization of routing problems in tourism activities are presented by Karagul & Gungor (2014), such as fuel savings, reduced emissions of polluting gasses, and decreased damage to natural and historical beauty.

The village of Pandeiros can be considered the best starting point for tourist visits. This can be explained by the fact that the location is in the central region of the Pandeiros River Environmental Protection Area and encompasses most of the tourist attractions of greatest interest, such as the waterfalls in the Pandeiros resort. This finding is important for directing public policies aimed at improving the conditions for receiving tourists in the conservation unit.

Conclusions

The simplified version of the clonal selection algorithm produced superior results compared to the original version for the problem under investigation.

Utilizing optimal itineraries can minimize the number of required roads in the conservation unit, particularly when the emphasis is on tourism. This approach can mitigate the environmental impacts associated with such activities.

Identifying the appropriate departure location for tourism is crucial when planning visits to the Pandeiros River Environmental Protection Area.

Acknowledgments and funding source

This work is the result of a research project funded by the Minas Gerais Research Foundation (FAPEMIG), process APQ-04498-16.

Conflict of interest

The authors have no conflict of interest to declare.

Authors' Contributions

Carlos Alberto Araújo Júnior: Conceptualization, investigation, methodology, and writing – original draft; **Helio Garcia Leite**: Formal analysis, supervision, and writing – review & editing; **João Batista Mendes**: Formal analysis, methodology, supervision, and writing – review & editing.

References

Anggodo, Y. P. et al. Optimization of multi-trip vehicle routing problem with time windows using genetic algorithm. **Journal of Environmental Engineering & Sustainable Technology**, v. 3, n. 2, p. 92-97, 2016. http://dx.doi.org/10.21776/ub.jeest.2017.003.02.4.

Araújo Júnior, C. A. et al. Can linear programming assist metaheuristics in forest production planning problem? **Floresta**, v. 51, n. 1, p. 751-759, 2021. http://dx.doi.org/10.5380/rf.v51i3.72612.

Armas, J. & Melian-Batista, B. Constrained dynamic vehicle routing problems with time windows. **Soft Computing**, v. 19, n. 1, p. 2481-2498, 2015. http://dx.doi.org/10.1007/s00500-014-1574-4.

Brach, M. & Górski, D. Application of network analysis for development and promotion of sustainable tourism in public forests. **Folia Forestalia**, v. 56, n. 2, p. 105-112, 2014. https://doi.org/10.2478/ffp-2014-0010.

Castro, L. N. de & Von Zuben, F. J. Learning and optimization using the clonal selection principle. **IEEE Transactions on Evolutionary Computation**, v. 6, n. 3, p. 239-251, 2002. https://doi.org/10.1109/TEVC.2002.1011539.

Freitas, B. & Portuguez, A. P. Uso, ocupação do espaço e perspectivas de desenvolvimento do turismo ecorrural na bacia hidrográfica do ribeirão São Vicente, Ituiutaba, MG. **Campo-Território**: Revista de Geografia Agrária, v. 9, n. 17, p. 330-361, 2014.

Gavalas, D. et al. A survey on algorithmic approaches for solving tourist trip design problems. **Journal of Heuristics**, v. 20, n. 1, p. 291-328, 2014. https://doi.org/10.1007/s10732-014-9242-5.

Hassen, H. B. et al. An artificial immune algorithm for HHC planning based on multi-agent system. **Procedia Computer Science**, v. 164, n. 1, p. 251-256, 2019. https://doi.org/10.1016/j.procs.2019.12.180.

Hosseinabadi, A. A. R. et al. OVRP_GELS: solving open vehicle routing problem using the gravitational emulation local search algorithm. **Neural Computing and Applications**, v. 29, n. 1, p. 955-968, 2018. https://doi.org/10.1007/s00521-016-2608-x.

Kantawong, K. & Pravesjit, S. An enhanced ABC algorithm to solve the vehicle routing problem with time windows. **ECTI Transactions on Computer and Information Technology**, v. 14, n. 1, p. 46-52, 2020. https://doi.org/10.37936/ecti-cit.2020141.200016.

Karagul, K. & Gungor, I. A case study of heterogeneous fleet vehicle routing problem: touristic distribution application in Alanya. **An International Journal of Optimization and Control**: Theories & Applications, v. 4, n. 2, p. 67-76, 2014. https://doi.org/10.11121/ijocta.01.2014.00185.

Liao, Z. & Zheng, W. Using a heuristic algorithm to design a personalized day tour route in a time-dependent stochastic environment. **Tourism Management**, v. 68, n. 1, p. 284-300, 2018. https://doi.org/10.1016/j.tourman.2018.03.012.

Mrówczyńska, B. et al. Artificial immune system in planning deliveries in a short time. **Bulletin of the Polish Academy of Sciences**, v. 67, n. 5, p. 969-980, 2019. https://doi.org/10.24425/bpas.2019.126630.

Nunes, Y. R. F. et al. Pandeiros: o Pantanal Mineiro. **MG Biota**, v. 2, n. 2, p. 4-17, 2009.

Ogiolda, M. The use of clonal selection algorithm for the vehicle routing problem with time windows. **Symposium for Young Scientists in Technology, Engineering and Mathematics**, v. 1, n. 1, p. 68-74, 2017.

Rezende, R. S. et al. Avaliação ambiental do rio Pandeiros utilizando macroinvertebrados como indicadores de qualidade da água. **Ecología Austral**, v. 22, n. 1, p. 159-169, 2012.

Rozidi, A. et al. Determination of the tourist route in Malang Raya by using ant colony optimization. **Journal of Physics**: Conference Series, v. 1872, n. 1, p. 1-9, 2021. https://doi.org/10.1088/1742-6596/1872/1/012003.

Santos, G. L. et al. Anthropogenic and climatic influences in the swamp environment of the Pandeiros River basin, Minas Gerais-Brazil. **Environmental Monitoring and Assessment**, v. 192, n. 219, p. 218-219, 2020. https://doi.org/10.1007/s10661-020-8192-7.

Silva, A. A. et al. Optimization approaches to support the planning and analysis of travel itineraries. **Expert Systems with Applications**, v. 112, n. 1, p. 321-330, 2018. https://doi.org/10.1016/j.eswa.2018.06.045.

Silva Júnior, O. S. & Lopes, L. A. S. A free geographic information system as a tool for multi-depot vehicle routing. **Brazilian Journal of Operations & Production Management**, v. 8, n. 1, p. 103-120, 2011. https://doi.org/10.4322/bjopm.2011.006.

Shui, X. et al. A clonal selection algorithm for urban bus vehicle scheduling. **Applied Soft Computing**, v. 36, n. 1, p. 36-44, 2015. http://dx.doi.org/10.1016/j.asoc.2015.07.001.

Shukla, M. & Jharkharia, S. Artificial immune system-based algorithm for vehicle routing problem with time window constraint for the delivery of agri-fresh produce. **Journal of Decision Systems**, v. 22, n. 3, p. 224-247, 2013. http://dx.doi.org/10.1080/12460125.2013.810859.

Souza, A. M. S. et al. **Proposta de instituição do comitê da bacia hidrográfica afluentes mineiros do médio São Francisco** (UPGRH-S9): diagnóstico sócio-econômico-ambiental apresentado ao Conselho Estadual de Recursos Hídricos, como pré-requisito de aprovação do comitê. Comissão Pró-Comitê, 2008. 32 p.

Yu, C. & Lau, H. Y. K. AIS-based algorithm for solving Vehicle Routing Problem with Simultaneous Pick-up and Delivery (VRP-SPD). **Journal of Traffic and Logistics Engineering**, v. 1, n. 2, p. 174-178, 2013. http://dx.doi.org/10.12720/jtle.1.2.174-178.