Optimization of the Ecological Restoration Process of the Desert Scrub in the Mexican Plateau

Asgard Mendoza-Flores[1*], Luis Alcázar-Díaz[1], Rogelio Coria-López[1], Fernando Elizalde-Ramírez[2]

[1]Undergraduate, Data Science and Mathematics Engineering, Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Ave. Eugenio Garza Sada 2501, Monterrey, N.L., México, 64849.
[2]Professor, Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Ave. Eugenio Garza Sada 2501, Monterrey, N.L., México, 64849.

Contributing authors: asgard.mendoza@gmail.com; alcazar.dev@gmail.com; rogecl02@gmail.com; fer_elizalde@tec.mx;

Abstract

Addressing and mitigating the impact of human development on native ecosystems has recently become a critical and widely adopted strategy in Mexico. The scope and success of such projects strongly relies on the possibility to accurately predict the amount of time and resources needed to complete each restoration, as any unforeseen activities translate directly into economical loses that could jeopardize current and future funding for such projects.
This paper presents three mathematical and computational algorithms that were developed to optimize the plant distribution on a restoration project in the Mexican plateau by formulating it as a CVRP (Capacitated Vehicle Routing Problem). The first method uses linear programming and a Mixed Integer Programming (MIP) solver, while the second and third methods use meta-heuristics. The three algorithms showed optimal and computationally efficient results for networks with less than 16 nodes, where each node represents a quadrant to be restored.
In the second part of our work, the same algorithms were evaluated in problems with up to 26 nodes, finding that the exact mathematical formulations grow exponentially in execution time. However, the formulation based on meta-heuristics (OR-Tools and KNN) kept the execution time to just over a second for all the cases considered.

Keywords: CVRP, MIP, Meta-Heuristics, Ecological Restoration, Algorithm

1 Introduction

Ecological restoration projects have gained momentum in Mexico as a measure aimed at reducing the impact of deforestation, industrial development and problems related to the degradation of endemic ecosystems such as: droughts, forest fires, soil erosion, as well as the poverty and marginalization generated by the loss of the resources that provide and sustain native communities. The Mexican Institute for competitiveness (IMCO)[1] highlighted the doubling of droughts in Mexico by 2050, as warned by the UN, which will have a greater impact on the country's economic drivers (the northern and central regions). Considering the importance of ecological restoration as a comprehensive step in the solution of environmental and social problems in the central-northern region of the country three tools have been proposed and exhaustively analyzed for the optimization such projects with the objective of reducing the time and costs associated. The execution time, memory usage and the percentage error with respect to the optimal solution were the metrics of performance used to compare the linear programming algorithm against both meta-heuristics. The methodology proposed for the creation of an optimized plant distribution strategy came down to formulate it as a Capacitated Vehicle Routing Problem (CVRP), with the main objective being to find a strategy to transfer every specimen needed to complete the restoration with the least amount of trips made by the trucks that distribute the plants across each quadrant of the project. In the formulation the trucks are represented by the capacitated vehicles, and a graph is created to represent the space of the restoration. In the graph created each quadrant to be restored is represented as a node and the arcs represent the cost in time for a truck to travel from one quadrant to another. The CVRP formulation allows for the creation of an optimal schedule for the trucks to deliver all the specimens to each quadrant in the least amount of time possible.

While exploring Konstantinos (2022)[2] partitioning formulation for CVRP problems he remarks that it will require an exponentially increasing number of binary variables each associated with all feasible circuits. The mathematical modeling reviewed and proposed by Konstantinos was heavily based on P.Thoth, D, Vigo (2002)[3] two index vehicle flow formulation without upper bound which becomes a computationally expensive formulation and forces the amount of routes created to be the same as the specified number of vehicles even if the number of vehicles could be smaller while maintaining feasibility. Another relevant approach came by seeking to solve the pollution-routing problem, when Bektas and Laporte (2011)[4] delved into branch and cut algorithms to face the possible complexities of CVRP problems. Their research emphasizes the use of efficient cutting planes and inequalities to reduce the solution space and improving computational performance. Similar to MIP, this approach crumbles to the complexity of large-sized networks, however, they introduce heuristics to guide the search process and escape local optima.

2 Mathematical modeling

In general terms the Vehicle Routing Problem (VRP) aims to determine the most efficient routes for a fleet of vehicles to service a set of customers, the CVRP adds a capacity restriction per vehicle to the VRP as well as a demand of a given product for each client. This restriction must guarantee that the capacity of the transport vehicle is not exceeded for any trip made while ensuring that each client receives the needed number of products. Here we present a possible formulation of the capacity restriction proposed by Borcinová (2017)[5]:

$$\sum_{i=1} \sum_{j=1, i\neq j} B_i \cdot x_{ijk} \leq A \quad \forall k,\ i,j \neq 0$$

where B_i represents the demand of the i-th client, A represents the vehicle capacity and x_{ijk} it's a binary variable that indicates if the vehicle k transits the arc (i, j). For the modeling of the problem, the following assumptions were made:

1. The restoration area has a defined set of quadrants and each quadrant has a defined dimension.
2. Each vehicle has a 524 plant capacity which is also the exact amount needed per hectare.
3. The time for loading and unloading plants to a vehicle is 15 minutes for each case totaling 30 minutes when all 524 plants are loaded.
4. The distance between each quadrant is determined by the Euclidean distance between the centers of each quadrant.
5. A fixed speed for all trucks was set to 10 km/h.
6. One quadrant must have the supply warehouse.
7. All quadrants must be restored.

The element 2 is highlighted because when considering the vehicle's capacity identical to the demand for plants in a hectare, a direct relationship is created between the dimension of a quadrant and the quantity of plants necessary for its restoration.

$$Required\ vehicles = B_i/524$$

where B_i is equivalent to the demand for plants in the quadrant. In this way, the demand used in the formulation of the problem (B_i) is changed to be equivalent to the number of hectares that the quadrant covers. So a demand $B_1 = 1$ now represents that the full vehicle load of 524 plants. Hence, the vehicle routing problem is divided in two parts:

- Supply routing of entire hectares: let $\bar{B}_i \subset B_i$ be a subset representing the integer part of the demand value for each node i. A solution is proposed that seeks to satisfy this demand.
- Fractional demands: for this formulation the decimal part of the demand of each node is used: $B_i - \bar{B}_i$.

By creating this division in the problem, the optimization process is simplified since the supply of entire hectares can be modeled as an assignment problem leaving only the second part as CVRP, allowing to save a considerable amount of time for the mathematical formulation.

2.1 Optimization of the supply of entire hectares with Mixed Integer Programming (MIP)

Since the original problem has been divided into two parts: the allocation problem in whole hectares and the allocation problem in fractional hectares, it was decided to solve the problem of whole hectares by taking advantage of the relation between the vehicle capacity and demand by quadrant in which a trip supplies a single hectare until all the quadrant has been restored. The model was solved using a mixed integer programming (MIP) solver [6] the problem was solved using the software General Algebraic Modeling System (GAMS).

Parameters:

- $i = \{depot, K1, K2, ..., Kn\}$: Quadrants.
- $k = \{T1, T2, ..., Tn\}$: Vehicles.
- O_i: Offer available at the i-th quadrant.
- $\bar{B}_j$: Demand expressed in complete loads of the $j - th$ quadrant.
- C_{ij}: Time (cost) in minutes needed to travel from i to j.

Variables:

- w_{ij} : Number of trips made from node i to node j.
- z : Time needed to complete the restoration of whole hectares.

Model:

$$\text{minimize: } z = \sum_i \sum_j 2C_{ij} w_{ij} + 30 w_{ij} \tag{1}$$

Subject to:

$$\sum_j w_{ij} \leq O_i, \forall i \tag{2}$$

$$\sum_i w_{ij} \geq \bar{B}_j, \forall j \tag{3}$$

$$w_{ij} \geq 0, \quad \forall i, j \tag{4}$$

Where the objective function 1 considers the time spend for each trip of each vehicle plus the time for loading and unloading. The restrictions 2 and 3 guarantee that the supply is never exceeded and the demand always reached. Restriction 4 maintains w_{ij} as a positive variable. The solution to this part of the problem is computationally efficient, so this was the only algorithm developed for the logistics of entire hectares restoration.

2.2 Optimization of hectare fractions

To solve the restoration problem in the remaining spaces, the capacitated vehicle routing problem paradigm (CVRP) was used. [3]. As previously stated, this paradigm allows us to model the allocation of the remaining resources to each quadrant in the most efficient way possible, since the vehicle is now allowed to visit more than one quadrant after re-supplying in the depot quadrant. To solve this problem, we based our formulation on Liza's (2022)[7] model for vehicle routing with capacity demand and sub-tour elimination, and an additional section of the objective function was added to consider loading and unloading time.

Parameters:

- $i = \{depot, K1, K2, ..., Kn\}$: Quadrants.
- $k = \{T1, T2, ..., Tn\}$: Vehicles.
- $A_k = \{T1, T2, ..., Tn\}$: Capacity of the k-th vehicle.
- B_i: Remaining demand of the i-th quadrant after having supplied all entire hectares within it.
- C_{ij}: Time of traveling from i to j.

Variables:

- x_{ijk} : Binary variable that indicates whether the k-th transport travels from node i to node j.
- y_{ik} : Binary variable that represents whether the demand for hectare i has been fully met by the k-th vehicle.
- u_{ik}: Variables used for sub-tour elimination.
- z: Time needed to complete the restoration of the hectare fractions.

Model:

$$\text{Minimize: } z = \sum_i \sum_j \sum_k C_{i,j} * x_{i,j,k} + \sum_i B_i / A_1 \cdot 30$$

Subject to:

$$\sum_i B_i \cdot y_{ik} \leq A_k, \forall k \tag{5}$$

$$\sum_k y_{ik} = |k|, \forall i = 1 \tag{6}$$

$$\sum_k y_{ik} = 1, \forall i \neq 1 \tag{7}$$

$$\sum_j x_{ijk} = y_{ik}, \forall j, k \tag{8}$$

$$\sum_i x_{ijk} = y_{jk}, \forall j, k \tag{9}$$

$$u_{ik} - u_{jk} + A_k * x_{ijk} \leq A_k - B_j \forall i, j, k \tag{10}$$

$$u_{ik} \geq B_i, \forall i, k \tag{11}$$

$$A_k \geq u_{ik}, \forall i, k \tag{12}$$

Where:

- The objective function adds up the time for all vehicle routes as well as the load and unload time with respect to the amount of specimens transported on each trip.
- The first constraint 5 ensures that the capacity constraint of any vehicle is never surpassed.
- The second constraint 6 guarantees that each route starts and ends at the deposit node.
- The third constraint 7 ensures that each quadrant is visited by one vehicle
- Constraints 8 and 9 ensure continuity between the nodes visited by the same vehicle on a route.
- Constraints 10, 11, and 12 are responsible for sub-tour elimination.

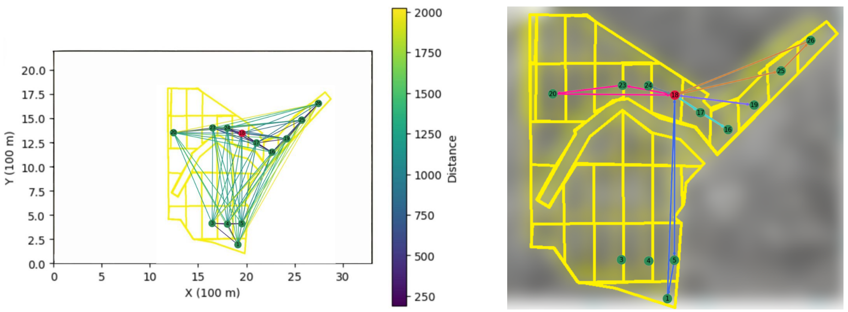
This formulation guarantees finding an optimal solution; however, it is a model that does not scale well due to the amount of variables and iterations required to solve the problem using the MIP solver for mixed integer programming.

2.2.1 Formulations based on meta-heuristics

Due to the computational cost required to find a result through mathematical modeling, a second and third algorithm were tested. The first method belongs to the open-license software developed by Google: OR-Tools[8]. This implemented algorithm follows the selection criterion of finding the node with the shortest travel time but implements a meta-heuristic algorithm that allows it to select other paths and escape local optima based on penalizing previous routes. The base documentation for CVRP modeling using OR-Tools can be reviewed on their official site[8]. The third algorithm implemented was K-Nearest Neighbors based on Zayyan A.[9]. The algorithm looks for the lowest cost vehicle routes that satisfy the CVRP demands until the demand on each node has been met, further depth into the algorithm can be found in the supplementary information of this paper.

3 Results

On this table a figure of the results obtained for the solution of each case can be found. It is worth mentioning that on the cases with more than 16 nodes the linear programming algorithm was unable to converge to the optimal value on a reasonable amount of time, hence the execution had to be terminated. The first figure shows a real life abstraction of a restoration project into a graph as previously detailed on the introduction of this paper, whereas the next figure recreates the most efficient routing to distribute the specimens needed to satisfy the demand of the hectare fractions on the same case.

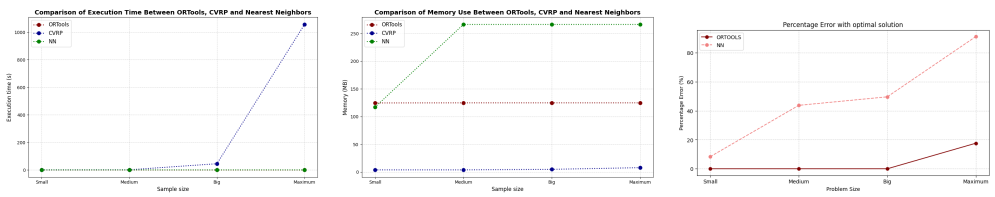


(a) Graph configuration (b) Optimal routing for hectare fractions.

Fig. 1: Summary of graph configuration and optimal routing on an applied case. The yellow lines represent the borders of the quadrants to be restored.

Batch size	**Solver**	**Execution Time (s)**	**Memory (MB)**	**Solution (s)**
Small: 11 Nodes	OR-Tools	1.01	123.85	12,244.2
	GAMS	0.16	4	12,244.2
	KNN	0.011	117.12	13,306.2
Medium: 16 Nodes	OR-Tools	1.01	124.87	12,259.2
	GAMS	0.75	4	12,259.2
	KNN	0.0298	266.25	17,624.4
Big: 21 Nodes	OR-Tools	1.01	124.87	19,027.8
	GAMS	44.672	5	19,027.8
	KNN	0.072	266.25	28,485.6
Maximum: 26 Nodes	OR-Tools	1.01	124.87	18,609.6
	GAMS	1056.98	8	15,821.4
	KNN	0.0144	266.25	30,249.6

Table 1: Comparison of execution times and memory usage for different batch sizes between OR-Tools and GAMS



(a) Comparative graph of execution times for the two algorithms (b) Comparative graph of memory usage for the two algorithms (c) Comparative graph of solution times for the two algorithms

Fig. 2: Accuracy and performance evaluations for all algorithms

Route	Total Time (s)	Total area (hectares)
1	1,680	0.75
2	2,225.4	0.8
3	1,299.6	0.64
4	2,694.6	0.96
5	2,084.4	0.92
6	2,295	0.91
Total	**12,278.16**	**4.98**

Table 2: Mathematical model: Summary of times and areas per Route

Route	Total Time (s)	Total area (hectares)
1	2,698.8	0.96
2	2,224.8	0.8
3	2,204.4	0.92
4	1,680	0.75
5	2,294.4	0.91
6	1,299.9	0.64
Total	**12,282**	**4.98**

Table 3: Summary of times and areas per Route (OR-Tools Algorithm)

4 Conclusions

In this article three different algorithms were developed and evaluated to optimize plant distribution in ecological restoration projects on the Mexican plateau by modeling them as Capacitated Vehicle Routing Problems (CVRP).

For the analysis of small graphs, the implemented methods showed efficiencies with negligible differences in terms of the accuracy and execution time as shown on figure 2 and tables 1, 2 and 3 producing optimal and functional results for the ecological restoration agencies. However, if the goal is to scale the implementation of any of these models, it is crucial to consider the specific details of each one and decide whether a completely rigorous solution from a mathematical perspective is more important or one that provides an approximate solution with a fraction of the computational resources required to achieve the optimal solution.

In the context studied, understanding the parameters and capabilities of the organization was crucial, as this context determines how easily a model can be modified, scaled, and interpreted. This makes a significant difference in the ability to save time and resources satisfactorily for ecological restoration projects and facilitates the implementation of computationally efficient algorithms such as OR-Tools meta-heuristics to improve the logistics in restoration projects.

Based on this, it has been concluded that to solve this type of problem optimally, the mathematical formulation developed can be used if the case can be represented with few nodes. On the other hand, if the problem contains multiple nodes, the solution with meta-heuristics, specially the OR-Tools algorithm should provide a satisfactory result at a significantly lower computational cost.

Supplementary information. GitHub repository where you can find the code for each algorithm implemented, as well as a pseudo-code: GitHub Repository

Next-steps. For a future edition we are seeking to extend comparison and analysis of optimization methods for ecological restoration processes to cases where there is more than one deposit and the assumption of Euclidean distances between sectors will be replaced for the algorithms to consider real paths only. It is expected that these contributions will enrich and contribute to optimal resource management and greater efficiency in ongoing restoration projects.

Acknowledgements. We sincerely thank Fernando Elizalde Ramírez for providing his expertise, orientation and feedback throughout the development of this project.

We thank Liza for her publications on the official GAMS Forum regarding her development of the mathematical CVRP model adapted for this project.

References

[1] López, S.: Escasez de agua y sequía en México: crisis actual (2023). https://imco.org.mx/escasez-de-agua-y-sequia-en-mexico-crisis-actual/

[2] Gkiotsalitis, K.: Public Transport Optimization, (2022). https://doi.org/10.1007/978-3-031-12444-0 . https://doi.org/10.1007/978-3-031-12444-0

[3] Toth, P., Vigo, D.: The Vehicle Routing Problem, (2002). https://doi.org/10.1137/1.9780898718515 . https://doi.org/10.1137/1.9780898718515

[4] Bektaş, T., Laporte, G.: The pollution-routing problem. Transportation Research Part B: Methodological **45**(8), 1232–1250 (2011) https://doi.org/10.1016/j.trb.2011.02.004 . Supply chain disruption and risk management

[5] Borcinova, Z.: Two models of the capacitated vehicle routing problem. Croatian Operational Research Review **8**, 463–469 (2017) https://doi.org/10.17535/crorr.2017.0029

[6] Bazaraa, M.S., Jarvis, J.J., Sherali, H.D.: Linear Programming and Network Flows, (1990)

[7] Liza: Capacitated Vehicle Routing Problem (2022). https://forum.gamsworld.org/viewtopic.php?t=11125

[8] Google: Capacity constraints. https://developers.google.com/optimization/routing/cvrp

[9] Firdaus, A.Z.: Solving Vehicle Routing Problems with Python Heuristics Algorithm (2023)