

Article

A Hybrid Memetic Algorithm for Asymmetric Vehicle Routing with Topographic Constraints: Quantifying the Orographic Gap in Mountain Urban Networks

Alejandra María Restrepo-Franco ^{1,*} , Orlando Valencia-Rodriguez ² , Eliana Mirledy Toro-Ocampo ³ 
and Omar Danilo Castrillón-Gómez ² 

¹ Departamento de Física y Matemáticas, Universidad Autónoma de Manizales, Manizales 170001, Caldas, Colombia

² Departamento de Ingeniería Industrial, Organizaciones y Logística, Universidad Nacional de Colombia, Manizales 170001, Caldas, Colombia; ovalenciaro@unal.edu.co (O.V.-R.); odcastrillong@unal.edu.co (O.D.C.-G.)

³ Facultad de Ciencias Económicas y Administrativas, Universidad Tecnológica de Pereira, Pereira 660002, Risaralda, Colombia; elianam@utp.edu.co

* Correspondence: alejandra.restrepof@autonoma.edu.co

Abstract

Classical vehicle routing models assume flat, symmetric road networks, yet mountain cities exhibit gravitational asymmetry and steep gradients that invalidate two-dimensional cost estimates and may produce mechanically infeasible routes. This study formalizes the Asymmetric Capacitated Vehicle Routing Problem with Topographic Constraints (ACVRP-TC) and introduces a generalized cost function (GCF) that linearizes direction-dependent energy consumption into an impedance metric within a mixed-integer linear programming (MILP) formulation. A hybrid memetic algorithm coupled with stochastic large neighborhood search (H-MA-LNS) is proposed to solve this NP-hard variant, combining evolutionary global exploration with structured local intensification. Benchmark validation yields a mean improvement of 9.4% over the state of the art on flat reference instances. An extensive computational evaluation on 24 real geospatial instances from Seoul (Republic of Korea), enriched with satellite elevation data, reveals a topographic gap of 14.1% for this urban network (Wilcoxon: $p = 0.002$), quantifying the cost underestimation that orography imposes on theoretical planning in the studied setting. Furthermore, the feasibility-based arc pruning reduces the search space by 51%, accelerating convergence and inducing a shift toward radial route structures in high-density environments.

Keywords: capacitated vehicle routing problem; topographic constraints; generalized cost function; memetic algorithm; large neighborhood search; digital elevation model; slope-dependent fuel consumption; urban logistics; asymmetric cost matrix



Academic Editors: Ming Wei and Bo Sun

Received: 4 May 2026

Revised: 13 July 2026

Accepted: 16 July 2026

Published: 22 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Road freight transport is a major component of the logistics sector; it currently faces critical challenges, including high operational costs, low efficiency, and significant environmental impact [1]. Within the field of operations research, the vehicle routing problem (VRP) has been extensively studied as a means to optimize fleet efficiency [2]. However, a persistent gap exists between theoretical routing models and actual operating conditions in geographically complex environments [3].

Classical VRP formulations typically assume that routes are executed over flat, horizontal terrain—an assumption that becomes inadequate when elevation varies as a nonlinear function of the distance traveled [4]. This simplification is particularly problematic in “mountain cities”, characterized by abrupt orography, such as those found in the Andean region or East Asia [5]. The physics of heavy-duty transport introduces gravitational asymmetry: the energy required for ascending significantly exceeds that for descending, where inertial forces alter the consumption dynamics. Neglecting this vertical component leads to fuel estimation errors exceeding 20% [6] and generates infeasible routes that violate traction capabilities. Simultaneously, urban infrastructure introduces structural asymmetry through one-way restrictions, with differences of up to 45.5% in total distance between symmetric and asymmetric formulations [7].

Despite the recognized need to integrate environmental variables into urban logistics [8], the current literature exhibits critical limitations in terms of scalability and physical fidelity. Recent research, exemplified by [9], addresses the Pollution-Routing Problem with Time Windows and Speed Optimization (PRP-SO) through mixed-integer nonlinear programming (MINLP) models. Although such approaches enable continuous speed profile optimization, their computational complexity limits scalability when confronted with large-scale instances. Furthermore, a significant gap remains in the realism of traditional benchmarks. While Lee et al. [10] proposed asymmetric instances based on real geospatial data from Seoul and Busan, these models do not incorporate topographic gradients. By omitting terrain slopes, the impact of engine effort on fuel consumption and emissions is systematically underestimated, thereby limiting the applicability of the results in urban environments with complex orography.

The central objective of this research is to develop an optimization framework capable of capturing the three-dimensionality and asymmetry of urban road networks without compromising computational tractability. To this end, the Asymmetric Capacitated Vehicle Routing Problem with Topographic Constraints (ACVRP-TC) is formally addressed, demonstrating that the linearization of topographic gradients enables high-quality solutions within reasonable computational times for large-scale instances. A generalized cost function (GCF) is introduced to translate slope-sensitive energy consumption into an equivalent cost metric within a mixed-integer linear programming (MILP) structure, and a hybrid memetic algorithm coupled with stochastic large neighborhood search (H-MA-LNS) is proposed, combining the global exploration of a genetic algorithm with structured local intensification.

The main contributions of this work are organized into three levels:

At the level of problem formulation and genuinely novel modeling, we propose:

1. A piecewise-linearized generalized cost function that internalizes gravitational asymmetry and mechanical-feasibility constraints within MILP, without the complexity of dynamic physical simulations.
2. The formalization and quantification of the topographic gap through realistic geospatial instances, providing critical evidence for sustainable urban logistics.

At the level of problem-specific algorithmic adaptation, we tailor established operators to direction-dependent cost ($c_{ij} \neq c_{ji}$): a directional swap-intra operator that replaces 2-opt to preserve arc directionality, an asymmetry-biased Clarke–Wright initialization, and a feasibility-based arc-pruning scheme derived from the GCF.

At the level of algorithmic integration, we combine these elements with genetic search, large neighborhood search, simulated annealing, and a polishing phase into a single H-MA-LNS framework. This is validated against the state of the art, with a mean improvement of 9.4%.

We emphasize that the novelty of this work lies in the modeling contributions (1)–(2) and the problem-specific adaptations, rather than in a fundamentally new metaheuristic template.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 details the mathematical formulation and the proposed algorithmic methodology. Section 4 presents the computational experiments and results. Section 5 discusses the findings and their implications. Finally, Section 6 offers conclusions and future research directions.

2. Related Work

The VRP literature has evolved from simplified theoretical models toward robust formulations that capture the complexity of real-world operations. This section reviews the relevant contributions along three axes: environmental and topographic factors, structural asymmetry, and hybrid metaheuristics.

2.1. Environmental and Topographic Routing

Recent studies demonstrate that terrain geography and vehicle physics invalidate the traditional shortest-path premise. Suzuki [11] and Schröder and Cabral [12] showed that incorporating sequenced payload and digital elevation models (DEMs) reduces fuel consumption and emissions by up to 20%.

In complex geographies, neglecting topography leads to markedly suboptimal decisions. Suarez-Chilma et al. [6] achieved emission reductions of 12.2% by optimizing over elevation profiles in Andean topography. More recently, Lai et al. [9] formalized the Pollution-Routing Problem with Speed Optimization (PRP-SO), demonstrating that avoiding steep gradients with loaded vehicles can yield energy savings of up to 53%, thereby underscoring that omitting the vertical component (z) compromises the validity of estimated costs.

2.2. Structural Asymmetry and Fleet Heterogeneity

Urban operational reality imposes constraints that break cost symmetry ($c_{ij} \neq c_{ji}$). Fischetti et al. [13] laid the foundations of the asymmetric VRP (ACVRP), and subsequent extensions such as that of Herrero et al. [14] demonstrated that combining asymmetry with heterogeneous fleets can reduce operational costs by up to 10% through constructive heuristics.

However, the validity of these models strictly depends on the quality of input data. Lee and Chae [10] found that Euclidean distances underestimate urban reality by a factor of $1.57\times$, with round-trip cost differences reaching 30%. Consequently, recent studies [15,16] have recommended the use of real cost matrices obtained via mapping APIs to capture both structural and temporal asymmetry, thereby avoiding bias in logistics planning.

2.3. Hybrid Metaheuristics and Memetic Algorithms

Given the NP-hard complexity of the ACVRP, exact methods become impractical at large scale. Memetic algorithms (MAs) constitute one of the most effective approaches, balancing the global exploration of a genetic algorithm (GA) with local search intensification [17]. This hybrid architecture has demonstrated strong practical performance; for instance, Köroğlu [18] integrated advanced operators (Lin–Kernighan and 2-opt) to mitigate delays in dynamic traffic scenarios.

Recent developments in hybridization can be classified into three categories. Constructive hybridization combines randomized heuristics (Biased-CWS) with multi-start strategies for asymmetric costs [14]. Metaheuristic–exact hybridization couples discrete optimization (Tabu Search for sequencing) with exact sub-algorithms for continuous variables such as speed [9]. Multi-criteria hybridization dynamically selects genetic operators—for example, integrating TOPSIS within a GA to improve adaptability [19].

2.4. Summary and Research Gap

Table 1 summarizes the transition from traditional approaches (CVRP) toward models that integrate transport physics, topography, and asymmetry in dense environments, predominantly solved through hybrid architectures.

Table 1. Related work: VRP variants, key factors, and solution methodologies.

Reference	Problem	Key Factors	Solution Methodology	Data/Context
[9]	PRP-SO	Slope, payload, speed, CMEM	Hybrid: Tabu Search + Exact	Hong Kong (real)
[4]	Steep-road VRP	Hard slope limits, truck dynamics	MILP + route heuristics	Valparaíso (real)
[12]	Green 3D-VRP	DEM integration, CO ₂ estimation	GIS-based 3D routing	Germany (real)
[6]	Mountain VRP	Andean topography, emissions	Multi-objective metaheuristic	Manizales (real)
[11]	Green VRP	Payload sequencing, idling	Enumeration heuristic	Simulations (USA)
[10]	ACVRP	Asymmetry, real vs. aerial dist.	Spatial analysis (T-map API)	Seoul/Busan (real)
[7]	ACVRP	One-way asymmetry, real road networks	Iterated greedy heuristic	Madrid (real)
[14]	AHVRP	Asymmetry, heterogeneous fleet	Randomized CWS + local search	Villalobos instances
[18]	VRPTW	Dynamic traffic, variable time	Hybrid: GA + Lin-Kernighan	Istanbul (real)
[19]	CVRP	Dynamic operator selection	Hybrid: GA + TOPSIS	Classical benchmarks

Beyond these axes, the alternative approaches to gravitational asymmetry deserve explicit comparison. Full-MINLP energy models with speed optimization capture the physics most faithfully but scale poorly [9,20]; GIS-based 3D routing operates directly on elevation-enriched graphs yet addresses single-path rather than fleet-level optimization [12,21]; steep-road VRP formulations impose hard slope limits without direction-dependent costs [4]; and data-driven asymmetric matrices obtained from mapping APIs capture structural but not gravitational asymmetry [10,15,16]. Each family thus treats the vertical dimension, computational tractability, or asymmetry in isolation, which motivates the linearized GCF adopted in this work.

The foregoing analysis reveals that, although substantial research has been conducted on topography and asymmetry in isolation, their joint integration within a computationally tractable and scalable model remains an open problem. The present work addresses this gap by proposing an H-MA-LNS framework that internalizes heavy-duty transport physics within asymmetric topographies through a linearized generalized cost function.

3. Materials and Methods

This section presents the mathematical formulation of the ACVRP-TC (Sections 3.1 and 3.2) and the proposed hybrid metaheuristic framework (Sections 3.4–3.7).

3.1. Problem Formulation and Topographic Impedance Model

Following the formalization of routing problems with irregular topography [9], the problem is defined over a directed graph $G = (\mathcal{V}, \mathcal{A})$, where $\mathcal{V} = \{0, 1, \dots, n\}$ is the node set, with node 0 representing the depot and $\mathcal{V}_c = \mathcal{V} \setminus \{0\}$ the set of customers. The arc set $\mathcal{A}$ contains the feasible connections (i, j) , with $i \neq j$. The objective is to determine minimum-cost routes that satisfy all demands and operational constraints while accounting for the impact of topography.

In contrast to traditional models based on idealized flat topologies, this work incorporates the vertical dimension z (elevation), extracted from digital elevation models (DEMs) [21]. For each directed arc $(i, j) \in \mathcal{A}$, the gradient is defined as follows:

$$\theta_{ij} = \frac{z_j - z_i}{d_{ij}} \quad (1)$$

where d_{ij} denotes the horizontal projected distance between nodes i and j .

To internalize the terrain impact without incurring the computational complexity of dynamic physical models, a generalized cost function (GCF) is introduced. The adjusted operational cost c_{ij} is defined as an equivalent distance weighted by physical effort:

$$c_{ij} = d_{ij} \cdot \Omega(\theta_{ij}) \quad (2)$$

The dimensionless impedance factor $\Omega(\theta_{ij})$ models the consumption dynamics, penalizing traction effort and adjusting for inertial recovery [22]:

$$\Omega(\theta_{ij}) = \begin{cases} 1 + \alpha \cdot \theta_{ij} & \text{if } \theta_{ij} \geq 0 \\ \max(\kappa, 1 + \beta \cdot \theta_{ij}) & \text{if } \theta_{ij} < 0 \end{cases} \quad (3)$$

where α and β are sensitivity coefficients and $\kappa = 0.6$ establishes a lower safety bound to prevent spurious negative costs.

It should be emphasized that the GCF constitutes a deliberately simplified impedance approximation designed for tractability within an MILP structure, rather than a fully calibrated vehicle energy model such as CMEM [10].

Its purpose is to preserve the two physically essential properties of slope-dependent consumption—gravitational asymmetry ($\alpha > \beta$) and a strictly positive consumption floor (κ)—while enabling constant-time arc-cost evaluation. Field calibration of (α, β, κ) against measurements from specific vehicle types remains an avenue for future work (Section 5.6).

3.1.1. Physical Parameter Validation

The GCF parameter configuration follows a validation criterion grounded in heavy-duty transport physics [20] and road design standards [23]. Table 2 details the technical justification for each coefficient.

Table 2. Technical justification and validation of the physical model parameters.

Parameter	Value	Physical–Normative Basis	Sensitivity Analysis
$\theta_{\max}$	25%	Mechanical traction limit based on mountain road design standards.	Defines the arc pruning rate; excessive reduction compromises connectivity.
α	10	Represents the power increase (P_{grade}) required to overcome gravitational resistance on ascent.	Dominant driver of the topographic gap, with sensitivity $\Delta\text{Gap}/\Delta\alpha \approx 2.83$ pp per unit (Section 3.1.3).
β	5	Coefficient of inertial recovery and regenerative braking efficiency on descent.	Lower sensitivity ($ \Delta\text{Gap}/\Delta\beta \approx 1.31$ pp per unit); stabilizes gravitational asymmetry ($\alpha > \beta$).
κ	0.6	Irreducible base consumption integrating rolling resistance and engine idling losses.	Establishes the floor of $\Omega(\theta)$ on descent, guaranteeing $c_{ij} > 0 \forall (i, j) \in \mathcal{A}$ and preventing negative-cost cycles in the MILP formulation.

3.1.2. Physical Asymmetry Property

A fundamental characteristic of the GCF is its capture of gravitational asymmetry, consistent with the grade-resistance term of comprehensive fuel-consumption models [20] and with the direction-dependent energy formulations reported for routing on uneven topography [4,9]. Since $\alpha > \beta$, it follows that $\Omega(\theta) \neq \Omega(-\theta)$. This property induces a strictly asymmetric cost matrix ($c_{ij} \neq c_{ji}$), thereby justifying the formulation as an asymmetric CVRP (ACVRP).

3.1.3. Parametric Sensitivity Analysis

To validate the robustness of the reference configuration ($\alpha = 10, \beta = 5$) and to quantify its influence on the topographic gap, a factorial sensitivity analysis was conducted over the

actual slope distribution of Seoul, following standard factorial screening practice for parametric models [24]. A total of 36 combinations $(\alpha, \beta) \in \{3, 5, 7, 8, 10, 12, 15\} \times \{2, 3, 4, 5, 6, 7, 8\}$ were evaluated, subject to the physical constraint $\alpha > \beta$ (ascending effort always exceeds descending recovery). For each configuration, the topographic gap was computed as the percentage divergence between the generalized cost $\sum c_{ij}$ and the flat cost $\sum d_{ij}$, evaluated over feasible arcs ($|\theta_{ij}| \leq \theta_{\max}$). Table 3 summarizes the representative configurations.

The analysis reveals a marked asymmetry in parametric sensitivity. With $\beta = 5$ fixed, the Gap exhibits a quasi-linear relationship with α , at a sensitivity of $\Delta \text{Gap} / \Delta \alpha \approx 2.83$ pp per unit, ranging from 3.6% ($\alpha = 5$) to 30.7% ($\alpha = 15$). In contrast, with $\alpha = 10$ fixed and β varying from 2 to 8, the sensitivity is significantly lower ($|\Delta \text{Gap} / \Delta \beta| \approx 1.31$ pp per unit), with a range of 22.1% to 14.3%. This confirms that α is the dominant driver of the topographic gap, consistent with the underlying physics: the power required to overcome gravitational resistance on ascent ($P_{\text{grade}} \propto \alpha \cdot \theta$) dominates the energy balance relative to the partial inertial recovery on descent.

The reference configuration ($\alpha = 10, \beta = 5$) produces a simulated Gap of 16.5%, consistent with the 14.1% experimental value (Section 4). The ~ 2 pp discrepancy is attributable to the experimental Gap incorporating the effect of algorithmic optimization, which preferentially selects lower-impedance arcs. The gravitational asymmetry is quantified by the ratio $\bar{\Omega}_{\uparrow} / \bar{\Omega}_{\downarrow} = 2.33$, indicating that the average cost of an ascending arc is 2.33 times that of its symmetric descending counterpart. Perturbations of ± 2 units in α (from 8 to 12) or β (from 3 to 7) produce bounded Gap variations (10.9–22.2% and 14.8–19.6%, respectively), ensuring that the study conclusions are robust to reasonable uncertainties in the estimation of vehicle physical parameters. The resulting asymmetry ratios (1.6–3.0, Table 3) are moreover of the same order as the ascent/descent consumption ratios reported in empirical studies of slope-dependent fuel use [9,20].

Table 3. GCF sensitivity analysis. Simulated topographic gap over the Seoul slope distribution for representative (α, β) configurations, with $\alpha > \beta$. $\bar{\Omega}_{\uparrow}$ and $\bar{\Omega}_{\downarrow}$: mean impedance on ascent and descent, respectively. The reference configuration is highlighted in bold.

α	β	α/β	Topo. Gap (%)	$\bar{\Omega}_{\uparrow}$	$\bar{\Omega}_{\downarrow}$	Asymmetry Ratio
5	2	2.50	7.96	1.348	0.855	1.576
5	3	1.67	5.45	1.348	0.798	1.690
7	3	2.33	11.11	1.487	0.798	1.865
7	5	1.40	8.02	1.487	0.727	2.045
8	4	2.00	12.12	1.557	0.756	2.059
8	5	1.60	10.86	1.557	0.727	2.141
10	3	3.33	19.61	1.696	0.798	2.126
10	5	2.00	16.52	1.696	0.727	2.332
10	7	1.43	14.85	1.696	0.690	2.459
12	5	2.40	22.18	1.835	0.727	2.523
12	7	1.71	20.51	1.835	0.690	2.660
15	5	3.00	30.67	2.044	0.727	2.811
15	8	1.88	28.42	2.044	0.677	3.020

3.1.4. Feasibility Constraint

A hard navigability constraint is imposed based on the mechanical limitations of heavy-duty vehicles. The subset of infeasible arcs $\mathcal{A}_{\text{infeas}} \subset \mathcal{A}$ is defined such that $(i, j) \in \mathcal{A}_{\text{infeas}}$ if $|\theta_{ij}| > \theta_{\max}$, with a critical threshold $\theta_{\max} = 25\%$.

3.2. MILP Formulation

The pre-computation of generalized costs c_{ij} enables the problem to be formulated as a mixed-integer linear programming (MILP) problem, significantly reducing computational

complexity relative to MINLP formulations [25]. The choice of an MILP structure—rather than MINLP—is justified by the fact that the GCF transfers all nonlinear physical computations to a preprocessing phase, yielding constant-time arc-cost evaluation during optimization.

The decision variables are as follows: $x_{ij} \in \{0, 1\}$, a binary variable taking value 1 if the vehicle traverses arc (i, j) and 0 otherwise; and $u_i \in \mathbb{R}^+$, a continuous auxiliary variable representing the accumulated load upon visiting node i . The parameters are as follows: q_i , the demand of node i ($q_0 = 0$); Q , the maximum capacity of the homogeneous vehicle; and K , the available fleet size.

The objective function minimizes the total generalized cost,

$$\min Z = \sum_{(i,j) \in \mathcal{A}} c_{ij} x_{ij} \quad (4)$$

subject to flow conservation and assignment constraints, ensuring each customer is visited exactly once:

$$\sum_{j \in \mathcal{V}, j \neq i} x_{ij} = 1 \quad \forall i \in \mathcal{V}_c \quad (5)$$

$$\sum_{i \in \mathcal{V}, i \neq j} x_{ij} = 1 \quad \forall j \in \mathcal{V}_c \quad (6)$$

The fleet constraints limiting the number of active routes are as follows:

$$\sum_{j \in \mathcal{V}_c} x_{0j} \leq K, \quad \sum_{i \in \mathcal{V}_c} x_{i0} \leq K \quad (7)$$

The Miller–Tucker–Zemlin (MTZ) constraints for subtour elimination and capacity enforcement are as follows:

$$u_j - u_i \geq q_j - Q(1 - x_{ij}) \quad \forall i, j \in \mathcal{V}_c, i \neq j \quad (8)$$

$$q_i \leq u_i \leq Q \quad \forall i \in \mathcal{V}_c \quad (9)$$

These inequalities serve a dual purpose: they eliminate subtours disconnected from the depot and ensure that the accumulated demand does not exceed capacity Q at any point along a route. The MTZ formulation was selected over alternatives (e.g., Dantzig–Fulkerson–Johnson) for its compact size— $O(n^2)$ constraints versus exponentially many—which is advantageous for the metaheuristic framework where the MILP structure guides the solution encoding rather than being solved to optimality by a general-purpose solver.

Finally, the topological domain constraints are as follows:

$$x_{ij} = 0 \quad \forall (i, j) \in \mathcal{A}_{\text{infeas}} \quad (10)$$

$$x_{ij} \in \{0, 1\} \quad \forall (i, j) \in \mathcal{A} \setminus \mathcal{A}_{\text{infeas}} \quad (11)$$

3.3. Model Verification

The consistency of the formulation (4)–(11) was verified analytically and computationally. Analytically, the constraint system is non-contradictory: the MTZ inequalities (8) and (9) are jointly satisfiable whenever the demand of each route respects Q , since the assignment $u_i = \sum_{k \leq i} q_{c_k}$ along any capacity-feasible route satisfies both (8) and (9) simultaneously; and the domain restriction (10) cannot induce infeasibility as long as the pruned graph $G' = (V, \mathcal{A} \setminus \mathcal{A}_{\text{infeas}})$ remains strongly connected—a condition verified computationally on all Seoul instances after $\theta_{\max}$ pruning.

Computationally, the complete MILP was solved to proven optimality with the CBC exact solver on ten subsampled Seoul instances derived from SLAS100 ($n \in \{10, 12, 15, 18, 20\}$

customers under two demand configurations: DV1M5, where capacity is slack; and DV1M20, where capacity binds and the optimum requires 2–3 routes with loads of up to 99.7% of Q), using the same GCF cost matrices of Equations (1)–(3), and the same demands and capacities as the metaheuristic. The formulation was strengthened with the standard bin-packing valid inequality $\sum_j x_{0j} \geq \lceil \sum_i q_i / Q \rceil$, which does not alter the feasible set but tightens the linear relaxation. Table 4 summarizes the results: in all cases the exact optimum satisfies every constraint simultaneously—assignment (5) and (6), fleet bound (7), capacity and subtour elimination (8) and (9), and exclusion of pruned arcs (10)—as confirmed by an independent post hoc verification routine; and the H-MA-LNS solutions match the exact optima in all ten instances, confirming both model correctness and decoder consistency. The verification script is available in the public repository (Section 5.7).

Table 4. Exact MILP verification on subsampled Seoul instances (SLAS100, CBC solver). R : number of routes in the exact optimum. In all ten instances the pruned graph is strongly connected and the exact optimum satisfies all constraints (5)–(10), as confirmed by the independent post hoc verification routine.

Instance	Config.	n	R	MILP Opt.	Time (s)	H-MA-LNS	Gap
SLAS100-n10	DV1M5	10	1	21,149.8	1.05	21,149.8	0.0%
SLAS100-n12	DV1M5	12	1	28,155.8	2.50	28,155.8	0.0%
SLAS100-n15	DV1M5	15	1	40,037.8	3.86	40,037.8	0.0%
SLAS100-n18	DV1M5	18	1	43,099.8	3.52	43,099.8	0.0%
SLAS100-n20	DV1M5	20	1	45,374.0	14.51	45,374.0	0.0%
SLAS100-n10	DV1M20	10	2	23,513.8	0.11	23,513.8	0.0%
SLAS100-n12	DV1M20	12	2	30,916.0	2.36	30,916.0	0.0%
SLAS100-n15	DV1M20	15	2	42,348.0	3.92	42,348.0	0.0%
SLAS100-n18	DV1M20	18	3	47,510.8	1.94	47,510.8	0.0%
SLAS100-n20	DV1M20	20	3	50,226.8	9.01	50,226.8	0.0%

3.4. Hybrid Memetic Algorithm

Since the problem formulated in Sections 3.1 and 3.2 is NP-hard, a metaheuristic resolution strategy is proposed, formally defined as a hybrid memetic algorithm with large neighborhood search (H-MA-LNS). It should be noted that H-MA-LNS integrates well-established metaheuristic components; its contribution at the algorithmic level is therefore the problem-specific adaptation of these components to the asymmetric, topographically constrained setting (Sections 3.1.4, 3.4.2 and 3.5.2), rather than a new metaheuristic paradigm. The choice of a memetic architecture over alternative metaheuristics (e.g., PSO or ACO coupled with LNS) is dictated by the structure of the problem rather than by preference. First, the solution space consists of customer permutations decoded by an optimal DP split; GA operators (OX, inversion) and LNS destroy-and-repair act natively on permutations, whereas PSO is defined over continuous vector spaces and its discrete adaptations (random keys, swap sequences) rely on velocity-based moves that implicitly assume metric symmetry—an assumption violated by the strongly asymmetric cost matrix ($c_{ij} \neq c_{ji}$) and the pruned-arc set of the ACVRP-TC. Second, the literature consistently identifies memetic frameworks combining population-based exploration with intensive local search as the state of the art for capacitated and asymmetric VRP variants [13,17,26,27]. Third, the valley-corridor structure induced by orography produces spatially correlated sub-optimal assignments, which the relatedness-removal operator of LNS is explicitly designed to dismantle and rebuild—a mechanism absent from swarm-based velocity updates. The incremental development protocol of Section 4.2 validates this design choice empirically: each component was retained only when it measurably improved solution quality.

Let $\mathcal{S}$ denote the space of all feasible customer permutations $\pi = (\pi_1, \dots, \pi_n)$ that can be decoded into feasible routes. The algorithm seeks $s^* \in \mathcal{S}$ such that $Z(s^*) \leq Z(s)$, $\forall s \in \mathcal{S}$.

3.4.1. General Architecture

The algorithm is modeled as an iterative stochastic process composed of four sequential operators: initialization ($\mathcal{I}$), memetic evolution ($\mathcal{E}$), refinement ($\mathcal{R}$), and final polishing ($\mathcal{P}$). The final solution is obtained through the composition:

$$\mathcal{H} = \mathcal{P} \circ \mathcal{R} \circ \mathcal{E} \circ \mathcal{I} \quad (12)$$

3.4.2. Stratified Population Initialization

Let P_0 be the initial population with cardinality $N_{\text{pop}} = |P_0| = 150$, corresponding to the final Optuna-tuned value reported in Table 5. To ensure structural diversity and quality, P_0 is constructed as the union of four subsets:

$$P_0 = P_{\text{CW}} \cup P_{\text{CW}}^{\text{pert}} \cup P_{\text{elite}} \cup P_{\text{RND}} \quad (13)$$

The subset P_{CW} (one individual) is generated via the Clarke and Wright savings heuristic [28], adapted for asymmetric matrices with a biased weighting: $w = 0.6 \cdot \max(s_{ij}, s_{ji}) + 0.4 \cdot \min(s_{ij}, s_{ji})$. From this base solution, 20 perturbed variants $P_{\text{CW}}^{\text{pert}}$ are generated by applying between 2 and $\lfloor n/10 \rfloor$ random swaps, creating a solution cloud in the neighborhood of the constructive heuristic. The subset P_{elite} (one individual, if available) injects the best permutation from the previous run, transferring accumulated knowledge without compromising diversity. The remainder P_{RND} is filled with uniform random permutations $\pi \sim U(\mathcal{S})$.

3.4.3. Memetic Evolutionary Search

Let P_t denote the population at generation t . The state transition $P_t \rightarrow P_{t+1}$ is governed by standard genetic operators (OX and inversion mutation), with the calibrated hyperparameters shown in Table 5.

Offspring generation is defined through the order crossover (OX) and inversion mutation (Inv) operators [29]. To ensure robust and reproducible performance, hyperparameter tuning was articulated in three stages: (i) a systematic literature review to delimit initial ranges [19,30,31]; (ii) a Design of Experiments based on a Taguchi L_{18} orthogonal array [32]; and (iii) fine-tuning via Bayesian optimization with the Optuna framework [33]. The operators are formally defined as follows:

$$(o_1, o_2) = \text{OX}(p_1, p_2), \quad \text{with probability } p_c = 0.85 \quad (14)$$

$$o' = \text{Inv}(o), \quad \text{with probability } p_m = 0.35 \quad (15)$$

where OX preserves relative adjacency subsequences and $\text{Inv}(\cdot)$ reverses a random subsequence $(\pi_i, \dots, \pi_j)$, introducing the variability needed to escape local optima. Table 5 details the evolution of key factors from initial ranges to final values.

Table 5. Systematic hyperparameter calibration of the GA.

Factor	Initial Ranges	DoE Values (L_{18})	Final (Optuna)
Population size	40, 80, 128, 150, 200	128	150
Number of generations	200, 500, 1000	500	300
Random savings count	3, 6, 10	3	3
Elite size	2, 4, 8	8	8
Crossover prob. (p_c)	0.7, 0.85, 0.9	0.9	0.85
Mutation prob. (p_m)	0.1, 0.2, 0.35	0.1	0.35
Selection	Tournament; RW	Tournament; RW	Tournament; RW
Crossover	OX, SX, ERX, Alt-Edges	OX, SX, ERX	OX

3.4.4. Baldwinian Memetic Improvement

Following Baldwinian evolutionary theory, where learned behaviors influence genotype–phenotype relationships by modifying individual fitness [34], the engine implements a two-level learning scheme with differentiated frequencies and intensities.

Level 1—Frequent and light learning (every $\lambda_1 = 10$ generations): An elite subset $\mathcal{E}_t \subset P_t$ is defined such that $|\mathcal{E}_t| = 0.10 \cdot N_{\text{pop}}$ (top 10%). For each elite individual, the permutation is decoded into routes via optimal DP split (Section 3.5.1) and a composite local search operator $\mathcal{LS}_{\text{light}}$ is applied (Section 3.5.4):

$$\forall s \in \mathcal{E}_t, \quad s \leftarrow \mathcal{LS}_{\text{light}}(\text{SplitDP}(s)) \quad (16)$$

where $\mathcal{LS}_{\text{light}}$ sequentially applies four operators (swap-intra, or-opt, inter-route relocate, inter-route swap) until a local optimum is reached, refining the phenotype without permanently altering the genotype structure.

Level 2—Sporadic and intensive learning (every $\lambda_2 = 60$ generations): An LNS-SA refinement of $I_{\text{mini}} = 20$ iterations is applied exclusively to the best individual in the population:

$$s_{\text{best}} \leftarrow \text{LNS-SA}(s_{\text{best}}, I_{\text{mini}} = 20) \quad (17)$$

This level explores structurally distinct neighborhoods through destruction and reconstruction, generating adaptive jumps that reconfigure client-to-route assignments. By operating at low frequency on a single individual, deep perturbations are prevented from dominating the computational budget or inducing premature convergence.

In Level 1, each elite individual is directly overwritten by its locally improved encoding; in Level 2, the LNS-SA-refined individual replaces its original copy in the offspring population (elitist reinsertion). In both cases the refined solution is inherited by the next generation.

Algorithm 1 describes the complete memetic evolutionary engine.

3.5. Local Search Operators

A set of five operators was specifically designed to respect the asymmetry of the cost matrix ($c_{ij} \neq c_{ji}$), organized into two intensity levels.

3.5.1. Optimal DP Split (Bellman/Prins)

Given a customer permutation $\pi = (c_1, c_2, \dots, c_n)$, the minimum-cost partition into routes is formulated as a shortest-path problem on an auxiliary graph with $n + 1$ nodes. Let $dp[i]$ denote the minimum cost to serve the first i customers,

$$dp[j+1] = \min_{0 \leq i \leq j} \{dp[i] + \text{RouteCost}(c_i, \dots, c_j)\} \quad \text{s.t.} \quad \sum_{k=i}^j q_{c_k} \leq Q \quad (18)$$

where the route cost is computed incrementally:

$$\text{RouteCost}(c_i, \dots, c_j) = \begin{cases} c_{0,c_i} + c_{c_i,0} & \text{if } j = i \\ \text{RouteCost}(c_i, \dots, c_{j-1}) - c_{c_{j-1},0} + c_{c_{j-1},c_j} + c_{c_j,0} & \text{if } j > i \end{cases} \quad (19)$$

This runs in $O(n^2)$ worst-case time, with practical acceleration via early termination when the accumulated load exceeds Q . The DP split is the component with the highest individual contribution to algorithmic performance: the proportion of optimal solutions found triples upon its incorporation.

Algorithm 1: Hybrid Memetic Evolutionary Search (H-MA).

Input: Population P , Generations $G_{\max}$, $p_c = 0.85$, $p_m = 0.35$, $\lambda_1 = 10$, $\lambda_2 = 60$
Output: Best solution S_{best}

```

1  $S_{\text{best}} \leftarrow \text{GetBest}(P)$ ;
2 for  $gen \leftarrow 1$  to  $G_{\max}$  do
3    $P_{\text{new}} \leftarrow \emptyset$ ;
4   while  $|P_{\text{new}}| < |P|$  do
5      $p_1, p_2 \leftarrow \text{TournamentSelection}(P, k = 2)$ ;
6     if  $\text{Rand}() < p_c$  then
7        $o_1, o_2 \leftarrow \text{OrderCrossover}(p_1, p_2)$ ;
8     else
9        $o_1, o_2 \leftarrow p_1, p_2$ ;
10    if  $\text{Rand}() < p_m$  then  $o_1 \leftarrow \text{InversionMutation}(o_1)$ ;
11    if  $\text{Rand}() < p_m$  then  $o_2 \leftarrow \text{InversionMutation}(o_2)$ ;
12     $P_{\text{new}} \leftarrow P_{\text{new}} \cup \{o_1, o_2\}$ ;
    // Level 1: Light LS on elite (every  $\lambda_1$  gen)
13    if  $gen \bmod \lambda_1 = 0$  then
14       $P_{\text{elite}} \leftarrow \text{SelectTopK}(P_{\text{new}}, 10\%)$ ;
15      foreach  $S \in P_{\text{elite}}$  do
16         $R \leftarrow \text{SplitDP}(S)$ ;
17         $R \leftarrow \mathcal{LS}_{\text{light}}(R)$ ;
18         $S \leftarrow \text{Encode}(R)$ ;
    // Level 2: Mini LNS on best (every  $\lambda_2$  gen)
19    if  $gen \bmod \lambda_2 = 0$  then
20       $S_{\text{top}} \leftarrow \text{GetBest}(P_{\text{new}})$ ;
21       $S_{\text{top}} \leftarrow \text{LNS-SA}(S_{\text{top}}, I_{\text{mini}} = 20)$ ;
22       $P_{\text{new}} \leftarrow \text{Replace}(P_{\text{new}}, S_{\text{top}})$ ;
23     $P \leftarrow P_{\text{new}}$ ;
24    if  $\text{Cost}(\text{GetBest}(P)) < \text{Cost}(S_{\text{best}})$  then
25       $S_{\text{best}} \leftarrow \text{GetBest}(P)$ ;
26 return  $S_{\text{best}}$ ;

```

3.5.2. Intra-Route Operators

Swap-Intra: This exchanges the positions of two nodes c_i and c_j within a route without reversing the intermediate subsequence, thereby preserving arc directionality in asymmetric matrices. The delta cost is evaluated in $O(1)$. For adjacent nodes ($j = i + 1$):

$$\Delta = [c_{p_i, c_j} + c_{c_j, c_i} + c_{c_i, n_j}] - [c_{p_i, c_i} + c_{c_i, c_j} + c_{c_j, n_j}] \quad (20)$$

For nonadjacent nodes ($j > i + 1$):

$$\Delta = [c_{p_i, c_j} + c_{c_j, n_i} + c_{p_j, c_i} + c_{c_i, n_j}] - [c_{p_i, c_i} + c_{c_i, n_i} + c_{p_j, c_j} + c_{c_j, n_j}] \quad (21)$$

where p_i, n_i denote the predecessor and successor of position i in the route. The move is accepted if $\Delta < -10^{-6}$.

Directional Or-Opt: This relocates a segment of length $\ell \in \{1, 2, 3\}$ from its current position to another position within the same route, without reversing the segment. The cost decomposes into two independent deltas:

$$\Delta_{\text{total}} = \underbrace{(c_{b,a} - c_{b,s_s} - c_{s_e,a})}_{\Delta_{\text{remove}}} + \underbrace{(c_{p_j,s_s} + c_{s_e,n_j} - c_{p_j,n_j})}_{\Delta_{\text{insert}}} \quad (22)$$

where s_s, s_e are the first and last nodes of the segment; b, a are the nodes adjacent to the original segment position; and p_j, n_j flank the insertion position.

3.5.3. Inter-Route Operators

These operators modify client assignments between distinct routes, subject to the capacity constraint Q .

Inter-Route Relocate: This extracts a client c from route R_i and inserts it at the best position in route R_j , subject to $\text{load}(R_j) + q_c \leq Q$, with per-route load caching for efficiency.

Inter-Route Swap (1×1): This exchanges a client $c_i \in R_i$ with a client $c_j \in R_j$, verifying capacity feasibility in both routes.

Cross-Exchange (segments 1–3): This exchanges segments of length $\ell_i, \ell_j \in \{1, 2, 3\}$ between two routes. The cost is evaluated in $O(1)$ by considering only the four arcs at the cut points:

$$\Delta = [c_{b_i,s_j^s} + c_{s_j^e,a_i} + c_{b_j,s_i^s} + c_{s_i^e,a_j}] - [c_{b_i,s_i^s} + c_{s_i^e,a_i} + c_{b_j,s_j^s} + c_{s_j^e,a_j}] \quad (23)$$

where s_i^s, s_i^e denote the first and last nodes of segment i , and b_i, a_i the nodes adjacent to the cut point. Segments are inserted without reversal, ensuring correctness under asymmetric matrices. All combinations are explored and the best-improvement move is executed. The restriction to segments of length $\ell \leq 3$ balances exploration breadth against computational cost, as preliminary experiments showed diminishing returns for $\ell \geq 4$.

3.5.4. Intensity-Level Organization

The operators are organized into two computational intensity levels:

$$\mathcal{LS}_{\text{light}}(S) = \text{SwapInter} \circ \text{Relocate} \circ \text{OrOpt} \circ \text{SwapIntra}(S) \quad (24)$$

$$\mathcal{LS}_{\text{full}}(S) = \text{CrossExchange} \circ \mathcal{LS}_{\text{light}}(S) \quad (25)$$

This separation allows $\mathcal{LS}_{\text{light}}$ to be applied frequently (every LNS iteration) while reserving $\mathcal{LS}_{\text{full}}$ (which includes the costly cross-exchange) for periodic interventions (every 5 iterations in LNS, every 10 during polishing), thereby optimizing the computational budget.

3.6. LNS Refinement and Final Polishing

3.6.1. Stochastic LNS Refinement

The stochastic refinement phase iteratively subjects a solution to controlled degradation followed by optimized reconstruction [26]; it is distinguished from standard local search by the size of the explored neighborhood: LNS can remove and reinsert between 10% and 40% of the nodes, enabling significant jumps in the fitness landscape [27].

Let s_{best} denote the incumbent solution after the evolutionary phase. The refinement phase applies stochastic LNS for k_{max} iterations (150–400, calibrated by instance size). At each step k ,

$$s' = \mathcal{LS}(\mathcal{R}(\mathcal{D}(s_k, q_k))) \quad (26)$$

Adaptive Destruction: This removes q nodes with an adaptive fraction that grows with stagnation:

$$\gamma_k = 0.25 + 0.15 \cdot \frac{k_{\text{no_improve}}}{k_{\text{max}}}, \quad \gamma_k \leq 0.45 \quad (27)$$

$$q_k = \max(4, \lfloor |\text{clients}| \cdot \gamma_k \rfloor) \quad (28)$$

The operator alternates stochastically (50%/50%) between two complementary strategies. Worst removal eliminates nodes with the highest marginal cost $\Delta_c = c_{\text{prev},c} + c_{c,\text{next}} - c_{\text{prev},\text{next}}$ to intensify the search. Relatedness removal starts from a seed node and removes spatially correlated nodes according to $R_{\text{ref},c} = c_{\text{ref},c}$ with biased selection ($\text{idx} = \lfloor |\text{cand}| \cdot r^6 \rfloor$, $r \sim U(0,1)$), creating geographically coherent gaps that facilitate reconstruction and diversify exploration.

Repair: This reinserts the q removed nodes via the k -regret heuristic [27],

$$\Delta f_{\text{regret}}(i) = \sum_{h=1}^k (f_{\text{ins}}(i, r_h) - f_{\text{ins}}(i, r_1)) \quad (29)$$

prioritizing clients with a single good insertion option. To diversify reconstruction, an adaptive probability $p_{\text{stoch}} = \min(0.20, 0.20 \cdot k_{\text{no_improve}}/k_{\text{max}})$ of selecting suboptimal insertion positions is introduced, growing linearly from 0% (deterministic at start) to 20% as the algorithm stagnates.

Post-Repair Local Search: After reconstruction, local search is applied in a stratified manner:

$$\mathcal{LS}_k = \begin{cases} \mathcal{LS}_{\text{full}}(s') & \text{if } k \bmod 5 = 0 \text{ or } k_{\text{no_improve}} = 0 \\ \mathcal{LS}_{\text{light}}(s') & \text{otherwise} \end{cases} \quad (30)$$

This alternation reduces computational cost by approximately 40% without degrading solution quality, reserving $\mathcal{LS}_{\text{full}}$ for moments of the greatest improvement potential.

Stochastic Acceptance Criterion: The transition $s_k \rightarrow s_{k+1}$ follows a simulated annealing-inspired criterion,

$$s_{k+1} = \begin{cases} s' & \text{if } \Delta Z < 0 \\ s' & \text{if } \Delta Z \geq 0 \text{ and } \text{rand}(0,1) < \exp(-\Delta Z/T_k) \\ s_k & \text{otherwise} \end{cases} \quad (31)$$

with a geometric cooling schedule:

$$T_0 = 0.02 \cdot Z(s_{\text{best}}), \quad T_{\text{min}} = 0.0005 \cdot Z(s_{\text{best}}), \quad T_{k+1} = T_k \cdot \alpha_c \quad (32)$$

where the cooling factor is computed as $\alpha_c = (T_{\text{min}}/T_0)^{1/k_{\text{max}}}$. At the start ($T_0 = 2\%$ of cost), solutions up to 2% worse are accepted with significant probability, facilitating the exploration of new basins of attraction. As the temperature decreases geometrically, the criterion converges toward strictly greedy acceptance. The value $T_0 = 0.02 \cdot Z$ was determined experimentally (Section 4): lower values ($0.005 \cdot Z$) produce premature convergence (+0.4 pp in GAP), while higher values ($0.05 \cdot Z$) degrade the final quality by 0.3–0.5 pp, consistent with recommendations for SA schemes coupled with LNS [27].

Table 6 consolidates the LNS/SA hyperparameters, their tested ranges, and the retained values, complementing the evolutionary calibration of Table 5.

Algorithm 2 presents the complete stochastic LNS refinement procedure.

Table 6. Calibration of the LNS/SA hyperparameters.

Parameter	Tested/Range	Retained Value and Rationale
Destruction fraction, γ_k	0.10–0.45	Adaptive: $0.25 + 0.15 k_{\text{no_improve}}/k_{\text{max}}$, capped at 0.45 (Equation (27)); base 0.25 balances jump size and repair cost.
LNS iterations, k_{max}	100–500	150–400, calibrated by instance size.
Initial temperature, T_0	$\{0.005, 0.02, 0.05\} \cdot Z$	$0.02 \cdot Z(s_{\text{best}})$; lower values cause premature convergence (+0.4 pp GAP), higher values degrade quality by 0.3–0.5 pp (Section 3.6.1).
Final temperature, T_{min}	—	$0.0005 \cdot Z(s_{\text{best}})$, with geometric cooling $\alpha_c = (T_{\text{min}}/T_0)^{1/k_{\text{max}}}$.
Stochastic insertion, p_{stoch}	0–0.30	Adaptive cap at 0.20, growing linearly with stagnation (Section 3.6.1).
Destruction operator mix	30/70–70/30	50%/50% worst removal/relatedness removal.
Light-LS frequency, λ_1	$\{5, 10, 20\}$ gen.	10 generations (elite top 10%).
Deep-LNS frequency, λ_2	$\{30, 60, 100\}$ gen.	60 generations, $I_{\text{mini}} = 20$ iterations on the best individual.
LS_{full} cadence	Every $\{3, 5, 10\}$ iter.	Every 5 iterations in LNS, every 10 in polishing ($\approx 40\%$ cost reduction, Section 3.5.4).
Polishing iterations, I_{polish}	100–800	200–600, calibrated by instance size; strict acceptance, $p_{\text{stoch}} = 0$.

Algorithm 2: Stochastic Large Neighborhood Search with SA (LNS-SA).

Input: Initial solution S_{init} , iterations I_{LNS}
Output: Refined solution S_{global}

```

1  $S_{\text{curr}} \leftarrow \mathcal{LS}_{\text{full}}(\text{SplitDP}(S_{\text{init}}));$ 
2  $S_{\text{global}} \leftarrow S_{\text{curr}};$ 
3  $T \leftarrow 0.02 \cdot \text{Cost}(S_{\text{curr}}); T_{\text{min}} \leftarrow 0.0005 \cdot \text{Cost}(S_{\text{curr}}); \alpha_c \leftarrow (T_{\text{min}}/T)^{1/I_{\text{LNS}}}; k_{\text{stag}} \leftarrow 0;$ 
4 for  $k \leftarrow 1$  to  $I_{\text{LNS}}$  do
5    $\gamma \leftarrow \min(0.45, 0.25 + 0.15 \cdot k_{\text{stag}}/I_{\text{LNS}}); q \leftarrow \max(4, \lfloor |S_{\text{curr}}| \cdot \gamma \rfloor);$ 
6   if  $\text{rand}() < 0.5$  then
7      $S_{\text{partial}}, R \leftarrow \text{WorstRemoval}(S_{\text{curr}}, q);$ 
8   else
9      $S_{\text{partial}}, R \leftarrow \text{RelatednessRemoval}(S_{\text{curr}}, q);$ 
10   $p_{\text{stoch}} \leftarrow \min(0.20, 0.20 \cdot k_{\text{stag}}/I_{\text{LNS}});$ 
11   $S_{\text{new}} \leftarrow \text{RegretInsertion}(S_{\text{partial}}, R, p_{\text{stoch}});$ 
12  if  $k \bmod 5 = 0$  or  $k_{\text{stag}} = 0$  then
13     $S_{\text{new}} \leftarrow \mathcal{LS}_{\text{full}}(S_{\text{new}});$ 
14  else
15     $S_{\text{new}} \leftarrow \mathcal{LS}_{\text{light}}(S_{\text{new}});$ 
16   $\Delta \leftarrow \text{Cost}(S_{\text{new}}) - \text{Cost}(S_{\text{curr}});$ 
17  if  $\Delta < 0$  then
18     $S_{\text{curr}} \leftarrow S_{\text{new}};$ 
19    if  $\text{Cost}(S_{\text{new}}) < \text{Cost}(S_{\text{global}})$  then
20       $S_{\text{global}} \leftarrow S_{\text{new}};$ 
21       $k_{\text{stag}} \leftarrow 0;$ 
22  else if  $T > 0$  and  $\text{rand}() < \exp(-\Delta/T)$  then
23     $S_{\text{curr}} \leftarrow S_{\text{new}};$ 
24     $k_{\text{stag}} \leftarrow k_{\text{stag}} + 1;$ 
25  else
26     $k_{\text{stag}} \leftarrow k_{\text{stag}} + 1;$ 
27   $T \leftarrow T \cdot \alpha_c;$ 
28 return  $S_{\text{global}};$ 

```

3.6.2. Final Polishing Phase

After completing 10 independent runs, an intensive polishing phase is applied to the best global solution found. This procedure executes 200–600 LNS iterations (calibrated by instance size), with three differences from the main LNS: (1) no SA acceptance criterion—only strict improvements are accepted, maximizing intensification; (2) deterministic insertion—pure regret without stochastic component ($p_{\text{stoch}} = 0$); and (3) $\mathcal{LS}_{\text{full}}$ applied every 10 iterations, with $\mathcal{LS}_{\text{light}}$ on the remaining ones:

$$s_{\text{final}} = \text{Polish}(s_{\text{best}}^{\text{global}}, I_{\text{polish}}) \quad (33)$$

Our experiments (Section 4) demonstrated that polishing improves the best solution from the 10 runs in 60–80% of instances, with cost reductions of up to 2%.

3.7. Computational Complexity Analysis

To assess the theoretical scalability of the proposed H-MA-LNS, an asymptotic complexity analysis of its main components is performed. Let n be the number of customers, K the number of active routes, P the population size, q the number of nodes removed during the LNS destruction phase, and G the maximum number of generations.

Table 7 provides the asymptotic complexity of each individual component, including the five local search operators and both chain configurations ($\mathcal{LS}_{\text{light}}$ and $\mathcal{LS}_{\text{full}}$).

Table 7. Asymptotic complexity of H-MA-LNS components.

Component	Complexity	Observation
Split DP (Bellman/Prins)	$O(n^2)$	Early termination by capacity; replaces greedy split $O(n)$.
Swap-intra	$O(n^2)$ per route	Delta evaluation $O(1)$ per move.
Directional or-opt	$O(n^2)$ per route	Segments of length $\ell \in \{1, 2, 3\}$.
Inter-route relocate	$O(n^2 \cdot K)$	Per-route load caching.
Inter-route swap	$O(n^2 \cdot K^2)$	Worst case over all route pairs.
Cross-exchange	$O(n^2 \cdot K^2 \cdot \ell^2)$	$\ell \leq 3$; only in $\mathcal{LS}_{\text{full}}$.
$\mathcal{LS}_{\text{light}}$ (1 pass)	$O(n^2 \cdot K^2)$	4 operators (without cross-exchange).
$\mathcal{LS}_{\text{full}}$ (1 pass)	$O(n^2 \cdot K^2 \cdot \ell^2)$	5 operators (with cross-exchange).
LNS-SA (1 iteration)	$O(q^2Kn + \mathcal{LS})$	Destruction + Repair + LS.
GA (1 evaluation)	$O(n^2)$	Split DP dominant.
Polish (1 iteration)	$O(q^2Kn + \mathcal{LS})$	Identical to LNS without SA.
Total per run	$O(GPn^2 + I_{\text{ins}}(q^2Kn + n^2K^2))$	Polynomial.

Table 8 compares the complexity order between the proposed framework and the MINLP approaches prevalent in the green VRP literature.

The competitive advantage of H-MA-LNS lies in transferring physical complexity to a preprocessing phase via the GCF. While conventional MINLP models must solve energy consumption equations at each iteration, the proposed framework reduces this operation to a constant-time memory lookup $O(1)$. Additionally, the mechanical feasibility constraint $\theta_{\text{max}} = 25\%$ acts as a topological filter that drastically reduces the cardinality of the arc set $\mathcal{A}$, mathematically justifying the competitive convergence times in scenarios with complex orography. A third efficiency factor is the stratified alternation between $\mathcal{LS}_{\text{light}}$ and $\mathcal{LS}_{\text{full}}$: by reserving the costly cross-exchange operator ($O(n^2K^2\ell^2)$) for periodic interventions, the average computational cost per iteration is reduced by approximately 40% relative to systematic application of the full chain, without significantly degrading the solution quality. Formal derivations of the bounds reported in Tables 7 and 8 are provided in Appendix A.

Table 8. Asymptotic complexity and scalability comparison.

Component	Traditional MINLP	Proposed H-MA-LNS	Scalability Implication
Arc cost computation	$O(f_{NL})$ (numerical integration or nonlinear solvers)	$O(1)$ (pre-computed matrix via GCF)	High efficiency in solution evaluation.
Local search	$O(I \cdot f_{NL} \cdot n^2)$	$O(I_{lms} \cdot q^2 \cdot K \cdot n)$	Superior scalability in dense instances ($q \ll n$).
Search-space pruning	None (explores full speed spectrum)	$O(\mathcal{A} \setminus \mathcal{A}_{infeas})$ (exclusion for $\theta > 25\%$)	Reduces branching factor by 51%.
Total per iteration	Exponential / non-deterministic	$O(Pn^2) + O(I_{lms}(q^2Kn + n^2K^2))$	Polynomial-time convergence.

3.8. Convergence Properties

Beyond the complexity bounds, H-MA-LNS satisfies the standard sufficient conditions for convergence in probability to the global optimum of elitist stochastic search. First, *elitism*: the incumbent S_{best} is never lost—it is updated only upon strict improvement in the evolutionary engine (Algorithm 1) and in the refinement phase (Algorithm 2), so the sequence $Z(S_{best,t})$ is non-increasing. Second, *reachability*: inversion mutation acts on permutations with strictly positive probability $p_m > 0$ and can generate any permutation from any other in finitely many steps; jointly with LNS destruction of up to 45% of the customers followed by regret insertion with stochastic component $p_{stoch} > 0$ and the SA acceptance rule of Equation (31), the resulting transition kernel assigns strictly positive probability to reaching any feasible solution from any state. Under these two conditions, the classical results for elitist evolutionary algorithms [35] and simulated-annealing acceptance schemes [36] imply $\Pr(Z(S_{best,t}) \rightarrow Z^*) \rightarrow 1$ as $t \rightarrow \infty$. We emphasize that this is an asymptotic guarantee: as for all metaheuristics, it does not bound finite-time performance, which is why the empirical protocol of Section 4 (10 independent seeds, coefficient of variation, and nonparametric hypothesis testing) remains the operative evidence of solution quality.

4. Results

This section evaluates the performance of the proposed H-MA-LNS through a sequential experimental protocol structured in three phases: (i) characterization of the algorithmic refinement process, (ii) comparative validation on standard and asymmetric benchmarks, and (iii) evaluation on real topographic instances.

4.1. Experimental Setup

All experiments were conducted on Microsoft Windows 10 Pro, equipped with an Intel Xeon CPU E5-2698 v3 @ 2.30 GHz (16 physical cores) and 16 GB RAM. The algorithms were implemented entirely in Python 3.11.7, selected for its availability of specialized scientific libraries and its open-source nature, which facilitates replicability. The technological stack integrates Pymoo [37] as the base framework for evolutionary operators, complemented with custom local search operators in NumPy; OSMnx, NetworkX, and Rasterio [38] for geospatial processing of road graphs and DEM raster data; and Joblib for parallelized computation of the asymmetric impedance cost matrix (`n_jobs=-1`).

Each instance was executed over 10 independent runs with distinct random seeds. The reported metrics include the best cost found, the percentage GAP relative to the best known solution (BKS), the average computation time, and the coefficient of variation (CV) as a stability indicator.

4.2. Phase 1: Algorithmic Engine Evolution

A central methodological aspect of this research was the incremental development of the H-MA-LNS algorithm, which evolved through six successive versions—from a

basic GA with post-evolutionary LNS (V1) to the final configuration with computational optimization (V5-fast)—progressively incorporating: population diversification and elite transfer between runs (V2); adaptive destruction with worst/relatedness removal and stochastic insertion (V3); operator correction for asymmetry by replacing 2-opt with directional swap-intra (V4); and the structural improvements of optimal DP split, SA criterion, cross-exchange, and final polishing (V5). Version V5-fast adds $O(1)$ delta evaluation across all operators and stratified $\mathcal{LS}_{\text{light}}/\mathcal{LS}_{\text{full}}$ alternation without algorithmic change.

To measure the contribution of each component, the versions were executed on the Augerat Set A [39] (27 symmetric instances) and the Fischetti instances [13] (8 asymmetric instances), with 10 independent runs each. Table 9 presents the aggregated results.

Table 9. H-MA-LNS engine performance evolution on symmetric (Augerat, 27 inst.) and asymmetric (Fischetti, 8 inst.) instances, 10 runs each. Time: mean wall-clock time per instance on the symmetric set; n.r.: not recorded.

Version	Symmetric [39]			Asymmetric [13]		Key Improvement
	Avg. GAP	% Optimal	Time (s)	Avg. GAP	% Optimal	
V1	1.3%	7%	7.9	2.9%	0%	Base engine
V2	0.9%	22%	n.r.	—	—	Elites + callback
V3	0.6%	22%	25.1	1.2%	38%	Adaptive destruction
V4	—	—	—	1.0%	38%	Swap-intra (asymmetry fix)
V5	0.2%	67%	108.1	1.6%	25%	Split DP + SA + cross-exch.
V5-fast	0.3%	52%	36.0	1.6%	25%	Delta $O(1)$

The evolution exhibits a pattern of diminishing returns with domain-differentiated contributions. In symmetric instances, the most significant jump comes from the optimal DP split and SA criterion (V3→V5: -0.4 pp), which triple the percentage of optimal solutions from 22% to 67%, confirming that the Bellman/Prins partition is the component with the single highest impact. Version V5-fast sacrifices only 0.1 pp in GAP in exchange for a 66% reduction in mean execution time (from 108.1 s to 36.0 s per instance; Table 9), making it the preferred variant for large-scale instances. The speed-up stems from the $O(1)$ delta evaluation and the stratified $\mathcal{LS}_{\text{light}}/\mathcal{LS}_{\text{full}}$ alternation (Section 3.5.4), and grows with instance size—for A-n80-k10 the mean time drops from 514.3 s to 103.0 s (80%).

In the asymmetric domain, versions V3–V4 achieve the best performance (GAP: 1.0–1.2%), while V5 increases the GAP to 1.6%. This is explained by the reduced size of Fischetti instances (34–71 nodes), where the overhead of DP split reduces effective GA evaluations without the partition improvement compensating for the loss of exploratory diversity. For the Seoul topographic instances (250+ nodes, strongly asymmetric), V5-fast was adopted for its balance between quality and time.

Exploration–Exploitation Phase Decomposition

Following the structure of the composition $H = P \circ R \circ E \circ I$ (Equation (12)), the separate contribution of global exploration and structured local search was quantified through instrumented re-runs (seeds $s \in \{0, \dots, 9\}$) on five Augerat instances, logging the incumbent cost at the end of every phase and at every generation. Table 10 reports the step-by-step results and Figure 1 visualizes the trajectories on two representative instances.

The decomposition shows that the two mechanisms are complementary and their contributions separable. The Clarke–Wright initialization (I) starts 4.1–7.4% above the proven optimum. The memetic evolutionary phase (E), responsible for global exploration, performs the bulk of the descent, reaching gaps of 0.0–5.7%; in A-n33-k5 and A-n61-k9 it already attains the best value found. The stochastic LNS-SA refinement (R), the structured local-search

component, then contributes up to an additional 3.0% improvement, precisely on the instances where exploration has not yet converged (A-n32-k5: 829 $\rightarrow$ 804; A-n64-k9: 1456 $\rightarrow$ 1446), while requiring only 6–19% of the run time. The elite-transfer mechanism of the multi-run architecture (Section 3.4.2) propagates and further improves the incumbent across runs (e.g., 1456 $\rightarrow$ 1431 in A-n64-k9), and the final polishing phase (P) yields no further improvement in any of the five instances, acting as a local-optimality verification. Together with the incremental component analysis in Table 9, these results isolate the role of each phase of the hybrid framework.

Table 10. Phase-by-phase decomposition of H-MA-LNS on five Augerat instances (instrumented re-runs with seeds $s \in \{0, \dots, 9\}$, reported separately from the benchmark protocol of Table 11). I: Clarke–Wright initialization; E: end of the memetic evolutionary phase (cold-start run); R: end of the LNS-SA refinement (cold-start run); best-of-10 after elite transfer across runs; P: final polishing. $\bar{t}_E/\bar{t}_R$: mean phase times.

Phase/Instance	A-n32-k5	A-n33-k5	A-n34-k5	A-n61-k9	A-n64-k9
BKS	784	661	778	1034	1401
I: Initialization (CW)	832	696	810	1110	1462
E: After memetic evolution	829	661	791	1035	1456
R: After LNS-SA (Δ vs. E)	804 (−3.0%)	661 (0%)	786 (−0.6%)	1035 (0%)	1446 (−0.7%)
Best of 10 runs (elite transfer)	784	661	778	1035	1431
P: Polishing	=	=	=	=	=
Final gap to BKS	0.0%	0.0%	0.0%	0.1%	2.1%
$\bar{t}_E/\bar{t}_R$ (s)	13.9/0.9	13.1/1.1	13.4/1.1	40.5/8.1	45.1/10.4

“=” denotes no further improvement (local-optimality verification).

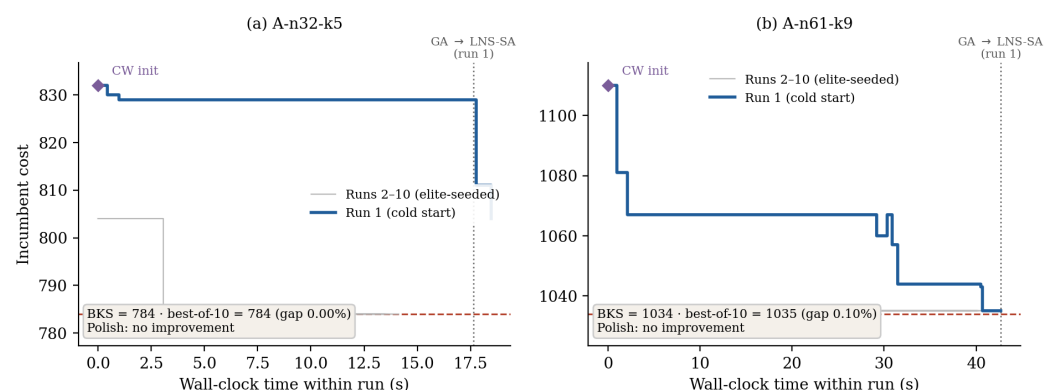


Figure 1. Incumbent convergence on two of the instances of Table 10. (a) A-n32-k5; (b) A-n61-k9. Blue: cold-start run (Clarke–Wright initialization, memetic evolution, LNS-SA refinement; the dotted line marks the phase boundary). Gray: subsequent runs, which start from the injected elite solution (Section 3.4.2) and therefore collapse onto the incumbent. Dashed red line: proven optimum.

4.3. Phase 2: Benchmark Validation

Before introducing topographic complexity, the robustness of the solution method is validated on both standard capacitated routing problems and asymmetric variants. The validation is structured through three levels of strategic comparison: exact methods as ground truth [13,39], structural asymmetry references [14], and real-network validation [10].

4.3.1. Symmetric Instances

The Augerat Set A [39] was used to evaluate performance on the classical CVRP ($c_{ij} = c_{ji}$). As presented in Table 11, the algorithm achieves a mean deviation (GAP) of 0.2%, reaching the exact optimum (BKS) in 67% of instances (18 of 27). The average CV

over 10 independent runs is 0.87%, guaranteeing algorithmic stability, with an average computation time of 108.1 s per instance.

Table 11. H-MA-LNS on the Augerat Set A instances [39] compared against published per-instance results of LNS-family metaheuristics and a savings-based matheuristic taken from Akpinar [40] (best percentage deviation from the BKS over 30 independent runs): LNS with improving-only acceptance (LNS_i), LNS with a simulated-annealing-type acceptance criterion in the style of the ALNS framework of Ropke and Pisinger [27] (LNS_a), the hybrid LNS-ACO of Akpinar [40], and the set-covering-based extended savings algorithm (SC-ESA) of Stanojević et al. [41]. H-MA-LNS gaps are best-of-10 runs; baseline gaps are best-of-30 runs, which makes the comparison conservative.

Instance	BKS	H-MA-LNS	GAP	LNS-ACO	LNS _a	LNS _i	SC-ESA
A-n32-k5	784	784	0.00%	0.00%	0.00%	0.00%	0.00%
A-n33-k5	661	661	0.00%	0.00%	0.00%	0.00%	0.00%
A-n33-k6	742	742	0.00%	0.00%	0.00%	0.00%	0.00%
A-n34-k5	778	778	0.00%	0.00%	0.00%	0.77%	0.00%
A-n36-k5	799	799	0.00%	0.00%	0.00%	0.75%	0.00%
A-n37-k5	669	669	0.00%	0.00%	0.00%	1.20%	0.00%
A-n37-k6	949	949	0.00%	0.00%	0.00%	0.42%	0.00%
A-n38-k5	730	730	0.00%	0.00%	0.00%	1.37%	0.00%
A-n39-k5	822	822	0.00%	0.00%	0.00%	1.70%	0.00%
A-n39-k6	831	833	0.24%	0.00%	0.00%	1.44%	0.00%
A-n44-k6	937	937	0.00%	0.00%	0.32%	1.49%	0.00%
A-n45-k6	944	944	0.00%	1.48%	0.74%	3.07%	0.00%
A-n45-k7	1146	1146	0.00%	0.00%	0.35%	2.36%	0.00%
A-n46-k7	914	914	0.00%	0.00%	0.44%	1.75%	0.00%
A-n48-k7	1073	1073	0.00%	1.03%	1.58%	2.24%	1.03%
A-n53-k7	1010	1010	0.00%	0.00%	0.50%	2.38%	0.10%
A-n54-k7	1167	1167	0.00%	0.00%	0.51%	1.71%	0.09%
A-n55-k9	1073	1073	0.00%	0.00%	0.37%	2.05%	0.00%
A-n60-k9	1354	1354	0.00%	0.00%	0.44%	1.92%	0.07%
A-n61-k9	1034	1037	0.29%	3.19%	3.68%	4.93%	0.00%
A-n62-k8	1288	1293	0.39%	1.55%	2.17%	3.34%	0.78%
A-n63-k9	1616	1630	0.87%	2.04%	2.23%	2.85%	0.50%
A-n63-k10	1314	1321	0.53%	1.14%	1.52%	3.65%	0.08%
A-n64-k9	1401	1421	1.43%	1.00%	1.36%	3.00%	0.57%
A-n65-k9	1174	1178	0.34%	0.94%	1.36%	3.24%	0.34%
A-n69-k9	1159	1163	0.35%	0.95%	1.81%	3.02%	0.00%
A-n80-k10	1763	1768	0.28%	2.95%	3.40%	4.71%	0.74%
Average	—	—	0.17%	0.60%	0.84%	2.05%	0.16%
Optima	—	—	18/27	17/27	10/27	3/27	17/27

To position H-MA-LNS against a broader range of metaheuristics, Table 11 also reports the published per-instance best deviations of four methods evaluated on the same 27 instances under identical BKS references, taken from the computational study of Akpinar [40]: two LNS variants—LNS_i (improving-only acceptance) and LNS_a, which employs a simulated-annealing-type acceptance criterion in the style of the ALNS framework [27]—the hybrid LNS-ACO of the same study, and the savings-based matheuristic SC-ESA [41]. Following common practice, the comparison relies on published results rather than reimplementations. H-MA-LNS attains the largest number of proven optima (18/27) and a mean best gap of 0.17%, and a paired Wilcoxon signed-rank test over the 27 instances confirms that it significantly outperforms all three LNS variants (vs. LNS-ACO: $W = 3$, $p = 0.008$; vs. LNS_a: $W = 3$, $p = 0.0003$; vs. LNS_i: $W = 0$, $p < 0.0001$), while being statistically indistinguishable from SC-ESA ($p = 0.81$; mean gaps: 0.17% vs. 0.16%). We finally note that all Set A optima are proven and that highly specialized CVRP solvers, such as HGS [42], reproduce them with essentially zero gap; the purpose of this benchmark is therefore not to claim superiority over dedicated symmetric-CVRP solvers, but to verify

that the proposed framework—designed for the asymmetric, topographically constrained setting—remains competitive on the classical symmetric domain.

4.3.2. Asymmetric Instances

To validate directed-cost handling ($c_{ij} \neq c_{ji}$), the Fischetti instances [13] were used. Table 12 contrasts the V3 engine (best configuration for this set) with the state of the art. The H-MA-LNS converges in an average of 65.0 s, compared to 530.9 s for the exact branch-and-bound method—an 87.8% reduction in execution time while maintaining a competitive GAP of 1.2%.

It must be emphasized that this comparison is indicative rather than strictly fair: branch-and-bound certifies optimality, whereas H-MA-LNS provides high-quality solutions without an optimality guarantee. The reported time reduction should therefore be interpreted as evidence of the practical efficiency of the heuristic in operational settings where certificates of optimality are not required, and not as superiority over exact methods.

Table 12. H-MA-LNS results on Fischetti asymmetric instances [13].

Instance	BKS	BKS Time (s)	H-MA-LNS	Time (s)	GAP
A034-02f	1406	64.0	1424	26.43	1.3%
A036-03f	1644	209.7	1681	19.31	2.25%
A039-03f	1654	10.8	1654	20.58	0.0%
A045-03f	1740	40.6	1740	34.81	0.0%
A048-03f	1891	35.3	1891	58.43	0.0%
A056-03f	1739	1939.9	1759	68.10	1.15%
A065-03f	1974	1332.7	2034	105.71	3.04%
A071-03f	2054	613.9	2087	185.86	1.61%
Average	—	530.9	—	65.0	1.2%

4.3.3. Complex Asymmetric Instances

To assess generalization and scalability in high-complexity environments, the instances of Rodríguez and Ruiz [7] were selected, which exhibit superior difficulty due to their structural asymmetry and clustered topology derived from real urban road networks with high one-way restriction density in Madrid. The results are validated against two literature references [7,14]. Table 13 presents the detailed comparison.

The algorithm achieves a mean GAP of 2.0% relative to [7] and 7.8% relative to [14], surpassing the BKS of [7] in 50% of instances (05xx–10xx). Notably, instance G-C-CAA0505CW reaches a cost of 331.42, 19.8% lower than reported by [7]. Instance G-C-CAA1502CW achieves a cost of 850.52, 1.9% below the reported BKS (866.90), constituting a new best known solution. The average CV of 1.0% across 10 runs confirms exceptional stability despite the combinatorial difficulty.

4.4. Phase 3: Topographic Instance Evaluation

The final validation is structured as a factorial experimental design using real geospatial instances from Seoul, South Korea, enriched with high-precision elevation data (SRTM 1 Arc-Second, USGS EarthExplorer [43]). The design crosses two topological dimensions: network scale and radius (100 nodes at 2.5 km vs. 250 nodes at 5 km) and client typology (residential type A vs. restaurant type F), producing four base groups (SLAS100, SLFS100, SLAM250, SLFM250). Each group was subjected to six fleet and demand configurations combining two vehicle types (V1: 1-ton/5 m³; V2: 2.5-ton/12.5 m³) with three demand multipliers (M5, M10, M20), yielding 24 test instances and a corpus of 240 executions.

Table 13. Performance on complex asymmetric instances of Rodríguez and Ruiz [7] and Herrero et al. [14].

Instance	BKS [7] (1)	BKS [14] (2)	Time (s) Herrero	H-MA-LNS (3)	Time (s) H-MA-LNS	GAP (1)–(3)	GAP (2)–(3)
G-C-CAA0501CW	396.954	370.26	17.83	368.31	63.92	−7.2%	−0.5%
G-C-CAA0502CW	383.354	372.48	13.30	411.37	53.19	+7.3%	+10.4%
G-C-CAA0503CW	418.768	404.30	10.48	441.48	43.89	+5.4%	+9.2%
G-C-CAA0504CW	376.837	361.74	17.95	358.09	59.13	−5.0%	−1.0%
G-C-CAA0505CW	413.310	386.73	15.59	331.42	57.70	−19.8%	−14.3%
G-C-CAA1001CW	631.386	596.54	43.03	651.18	248.70	+3.1%	+9.2%
G-C-CAA1002CW	636.504	578.15	39.36	621.60	232.46	−2.3%	+7.5%
G-C-CAA1003CW	584.511	561.08	42.23	624.76	253.03	+6.9%	+11.4%
G-C-CAA1004CW	689.650	660.81	34.84	681.42	205.96	−1.2%	+3.1%
G-C-CAA1005CW	703.048	652.08	29.03	818.88	212.45	+16.5%	+25.6%
G-C-CAA1501CW	903.766	—	—	925.52	464.73	+2.4%	—
G-C-CAA1502CW	866.901	—	—	850.52	398.04	−1.9%	—
G-C-CAA1503CW	825.169	—	—	861.05	440.68	+4.3%	—
G-C-CAA1504CW	740.339	—	—	779.46	332.94	+5.3%	—
G-C-CAA1505CW	787.306	—	—	845.30	500.09	+7.4%	—
Average	—	—	26.36	—	143.0	2.0%	7.8%

The comparison is structured across three scenarios: (1) the literature reference [10]; (2) H-MA-LNS without topographic constraints (flat 2D space, $\Omega = 1$); and (3) H-MA-LNS with the full topographic model ($\alpha = 10$, $\beta = 5$, $\theta_{\max} = 25\%$). This structure decomposes the total improvement into two independent components: algorithmic efficiency (GAP 1–2) and the cost of topographic reality (GAP 2–3).

4.4.1. Geospatial Characterization

Figure 2 illustrates the Seoul spatial network, where altimetry is explicitly integrated into the arc cost function, generating a strictly asymmetric structure dependent on direction and gradient. Demand concentrates in low- and mid-elevation valleys, while surrounding highlands act as topographic barriers that constrain the feasible routing space and induce strongly directional displacement patterns. Seoul’s mean slope of 7.9 degrees (8.57% in urban basins) places it in the “mountain city” category, alongside Hong Kong and Chongqing [23].

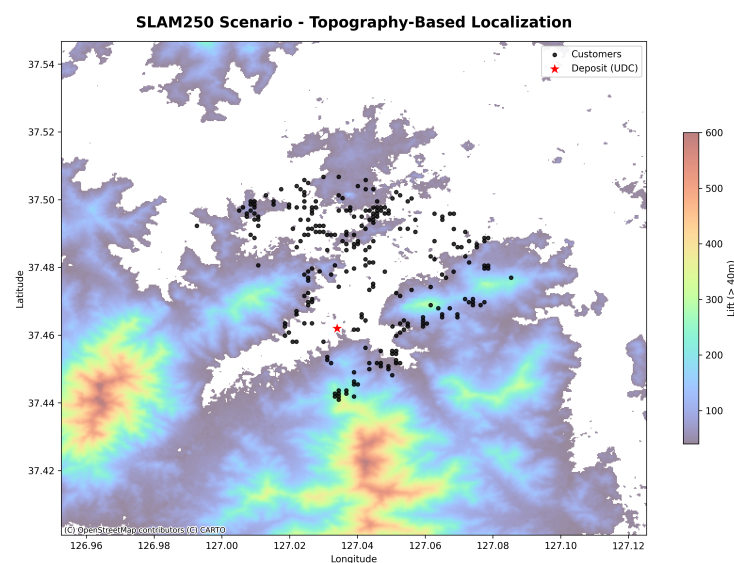


Figure 2. Routing scenario with topographic complexity (SLAM250—Seoul). Demand concentrates in valleys while surrounding elevated areas act as barriers constraining the feasible routing space.

To ground the slope constraints ($\theta_{\max}$) implemented in the model, Table 14 summarizes the normative slope limits for mountain terrain road design. The analysis reveals that a 16–17% gradient constitutes the soft transitivity limit; the model sets a technical infeasibility threshold at 25%, capturing extreme last-mile scenarios while excluding arcs that violate mechanical traction capacity.

Table 14. Normative slope limits for mountain terrain road design.

Road Category	Design Speed	Max. Absolute	Operational Implication
Highway	100 km/h	6%	Strict limit to prevent brake fade.
Principal arterial	60–80 km/h	9%	Main logistics flow corridors.
Local road	20–40 km/h	16–17%	Critical threshold defining standard transitivity.

4.4.2. Computational Results on Asymmetric Urban Instances

Table 15 presents the computational results across the 24 Seoul instances, comparing the three scenarios to isolate the impact of orography.

Gap definitions. All gaps in Table 15 share the literature cost C_1 as a common denominator: $\text{Gap (1-2)} = (C_2 - C_1)/C_1$, $\text{Gap (1-3)} = (C_3 - C_1)/C_1$, and $\text{Gap (2-3)} = (C_3 - C_2)/C_1$, where C_1 , C_2 , and C_3 denote the best cost over 10 runs under scenarios (1)–(3), respectively. The common denominator makes the decomposition exactly additive, $\text{Gap (1-3)} = \text{Gap (1-2)} + \text{Gap (2-3)}$, so the net deviation from the literature splits precisely into algorithmic efficiency and topographic surcharge. The global topographic gap of 14.1% is the arithmetic mean of the Gap (2–3) values over the 23 instances of Table 15, excluding the anomalous instance SLFS100_DV2M5.

To assess the statistical significance of the algorithmic improvement (Gap 1–2), a Wilcoxon signed-rank test was applied to the 23 paired instance costs of scenarios (1) and (2), excluding the anomalous instance SLFS100_DV2M5. The test rejects the null hypothesis of zero median difference ($W = 15$, $p < 0.0001$), with a median improvement of 7.3% and a matched-pairs rank-biserial correlation of $r = 0.89$, indicating a large and systematic effect. The global average of -9.4% is therefore not attributable to a few favorable instances.

This aggregate behavior is not uniform across the instance set. Table 16 disaggregates the three gaps by network scale and typology, showing that the algorithmic advantage over the literature is marginal on the 100-node instances but reaches -17.9% (SLAM) and -14.3% (SLFM) at 250 nodes, while the topographic surcharge remains substantial at every scale. Table 16 also reports the topographic-to-flat runtime ratio, indicating that topological pruning reduces computation time by 38–51%.

4.4.3. Statistical Validation

To provide inferential rigor, 10 independent replications were analyzed across all 24 experimental instances. As detailed in Table 17, the H-MA-LNS exhibits notable stability: the CV averages 1.52% for the flat model and 3.08% under orographic constraints, both strictly below the 5% critical threshold required in metaheuristic design.

Note that Table 17 reports average paired gaps computed per seed over the 10 runs of each instance—the paired structure on which the Wilcoxon signed-rank test operates. Since these per-seed gaps are normalized by the flat cost of the same run rather than by C_1 (Section 4.4.2), the mean of the Avg. Gap column (17.8%) differs from the 14.1% headline figure both in run basis (average vs. best) and in normalization.

To assess the structural divergence of the topographic gap, despite the Shapiro–Wilk test suggesting normality ($p > 0.05$), a conservative approach was adopted, using the non-parametric Wilcoxon signed-rank test. The test yielded a systematic value of $p = 0.00195$

across all scenarios, categorically rejecting the null hypothesis at the 95% confidence level. It is worth noting that with $n = 10$ paired samples, $p = 2/1024 \approx 0.00195$ is the minimum attainable p -value when all differences share the same sign, which does not invalidate the result but should be acknowledged for rigor. The 95% confidence intervals consolidate the finding: the omission of gravitational asymmetry induces a systematic underestimation of operational costs in urban logistics planning, with a global topographic gap of 14.1%.

Table 15. Computational results: cost comparison and normalized gaps relative to the literature.

Instance	(1) Literature [10]	(2) H-MA-LNS Flat	(3) H-MA-LNS Topographic	Gap (1)–(2) Efficiency	Gap (1)–(3) Reality	Gap (2)–(3) Topography
SLAS100_DV1M5	76,358	72,727	79,640	−4.8%	+4.3%	+9.05%
SLAS100_DV1M10	83,938	79,228	85,830	−5.6%	+2.3%	+7.87%
SLAS100_DV1M20	108,139	104,124	110,320	−3.7%	+2.0%	+5.73%
SLAS100_DV2M5	74,887	71,893	77,740	−4.0%	+3.8%	+7.81%
SLAS100_DV2M10	74,837	69,986	76,830	−6.5%	+2.7%	+9.14%
SLAS100_DV2M20	80,057	75,460	82,380	−5.7%	+2.9%	+8.64%
SLFS100_DV1M5	39,488	39,900	48,710	+1.0%	+23.4%	+22.31%
SLFS100_DV1M10	48,164	49,570	57,770	+2.9%	+19.9%	+17.03%
SLFS100_DV1M20	66,487	68,160	76,880	+2.5%	+15.6%	+13.12%
SLFS100_DV2M5 †	80,984	33,770	41,557	−58.3%	−48.7%	+9.62%
SLFS100_DV2M10	36,134	36,180	44,360	+0.1%	+22.8%	+22.64%
SLFS100_DV2M20	43,654	44,330	52,420	+1.5%	+20.1%	+18.53%
SLAM250_DV1M5	182,120	142,384	174,290	−21.8%	−4.3%	+17.52%
SLAM250_DV1M10	217,286	183,426	217,696	−15.6%	+0.2%	+15.77%
SLAM250_DV1M20	287,510	259,731	303,907	−9.7%	+5.7%	+15.36%
SLAM250_DV2M5	162,042	123,398	150,980	−23.8%	−6.8%	+17.02%
SLAM250_DV2M10	175,111	138,488	173,780	−20.9%	−0.8%	+20.15%
SLAM250_DV2M20	206,558	174,164	204,260	−15.7%	−1.1%	+14.57%
SLFM250_DV1M5	135,313	114,550	134,210	−15.3%	−0.8%	+14.53%
SLFM250_DV1M10	183,650	161,660	191,560	−12.0%	+4.3%	+16.28%
SLFM250_DV1M20	295,320	273,820	309,410	−7.3%	+4.8%	+12.05%
SLFM250_DV2M5	107,839	86,300	100,870	−20.0%	−6.5%	+13.51%
SLFM250_DV2M10	123,386	99,660	114,560	−19.2%	−7.2%	+12.08%
SLFM250_DV2M20	161,406	142,330	162,520	−11.8%	+0.7%	+12.51%
Global average	—	—	—	−9.4%	+4.7%	+14.1%

† Instance SLFS100_DV2M5 exhibited extreme divergence from the literature (reported base resolution time of 0 s and atypical cost). Its gaps are reported for completeness but are excluded from all global averages in this table, as the reference value is unreliable for this specific topology.

Table 16. Efficiency comparison by scale and typology.

Metric	100 Nodes	250 Nodes	Interpretation
GAP (1–2) Algorithmic	−5.1% (SLAS)/+1.6% (SLFS)	−17.9% (SLAM)/−14.3% (SLFM)	Algorithmic advantage grows with network scale.
GAP (2–3) Topographic	+8.0% (SLAS)/+18.7% (SLFS)	+16.7% (SLAM)/+13.5% (SLFM)	Surcharge present in every group (8–19%), governed by client typology as much as by network size.
GAP (1–3) Net	+3.0% (SLAS)/+20.4% (SLFS)	−1.2% (SLAM)/−0.8% (SLFM)	At 250 nodes the algorithmic gain compensates the surcharge.
Time (topo/flat)	0.62×	0.49×	Topological pruning reduces computation time by 38–51%.

Entries are arithmetic means of the per-instance gaps reported in Table 15 (best cost over 10 runs) across the six fleet–demand configurations of each group. Instance SLFS100_DV2M5 is excluded from the SLFS column ($n = 5$), consistent with Table 15. Rows are additive by construction, $\text{GAP}(1-3) = \text{GAP}(1-2) + \text{GAP}(2-3)$, up to 0.1 pp of rounding. Weighting each group by its number of instances reproduces the global averages of Table 15 (−9.4%, +14.1% and +4.7%).

Table 17. Statistical inference analysis: stability and significance of the topographic gap (10 runs per instance).

Instance	CV Flat	CV Topo.	Avg. Gap	Shapiro–Wilk (<i>p</i>)	Wilcoxon (<i>p</i>)	95% CI (Gap)
SLAS100_DV1M5	2.32%	1.88%	8.41%	0.8943	0.00195 *	[6.64%, 10.17%]
SLAS100_DV1M10	1.92%	2.42%	9.31%	0.5950	0.00195 *	[6.75%, 11.88%]
SLAS100_DV1M20	0.84%	3.95%	11.27%	0.6931	0.00195 *	[7.72%, 14.82%]
SLAS100_DV2M5	1.53%	1.31%	8.38%	0.9182	0.00195 *	[6.15%, 10.60%]
SLAS100_DV2M10	1.34%	1.60%	10.10%	0.6284	0.00195 *	[8.23%, 11.97%]
SLAS100_DV2M20	1.42%	2.56%	10.47%	0.0423	0.00195 *	[7.72%, 13.23%]
SLAM250_DV1M5	1.90%	4.74%	29.79%	0.1384	0.00195 *	[24.13%, 35.46%]
SLAM250_DV1M10	2.18%	3.38%	22.26%	0.4409	0.00195 *	[19.54%, 24.98%]
SLAM250_DV1M20	1.70%	2.12%	17.96%	0.3875	0.00195 *	[15.09%, 20.83%]
SLAM250_DV2M5	0.74%	2.75%	25.04%	0.4319	0.00195 *	[22.72%, 27.37%]
SLAM250_DV2M10	1.29%	3.51%	30.17%	0.7287	0.00195 *	[26.67%, 33.67%]
SLAM250_DV2M20	1.61%	5.54%	24.13%	0.9123	0.00195 *	[19.92%, 28.35%]
SLFS100_DV1M5	1.75%	0.73%	21.02%	0.3530	0.00195 *	[19.51%, 22.52%]
SLFS100_DV1M10	0.55%	1.32%	16.86%	0.2300	0.00195 *	[15.46%, 18.25%]
SLFS100_DV1M20	2.58%	2.28%	13.20%	0.4304	0.00195 *	[10.13%, 16.28%]
SLFS100_DV2M5	2.12%	0.66%	17.72%	0.5310	0.00195 *	[15.76%, 19.69%]
SLFS100_DV2M10	1.64%	1.72%	22.69%	0.6307	0.00195 *	[20.41%, 24.98%]
SLFS100_DV2M20	0.92%	1.70%	18.92%	0.5390	0.00195 *	[17.24%, 20.60%]
SLFM250_DV1M5	1.46%	3.05%	18.49%	0.3505	0.00195 *	[15.84%, 21.14%]
SLFM250_DV1M10	1.61%	1.84%	17.97%	0.5654	0.00195 *	[16.39%, 19.55%]
SLFM250_DV1M20	1.27%	2.63%	17.76%	0.5548	0.00195 *	[15.92%, 19.59%]
SLFM250_DV2M5	0.74%	2.38%	18.53%	0.8720	0.00195 *	[16.85%, 20.21%]
SLFM250_DV2M10	2.20%	2.77%	18.61%	0.1304	0.00195 *	[15.45%, 21.77%]
SLFM250_DV2M20	0.81%	3.65%	17.30%	0.2497	0.00195 *	[14.14%, 20.46%]
Average	1.52%	3.08%	17.8%	—	<0.05	Significant

* Statistically significant difference at $\alpha = 0.05$. Note: Column mean = 17.8%; the headline global topographic gap of 14.1% is defined from the best-run gaps of Table 15 (see gap definitions, Section 4.4.2).

4.4.4. Spatial Reconfiguration Patterns

Analysis of the eight routing configurations (comparative pairs between the flat Euclidean model and the orography-constrained model for high-demand instances DV1M20), combined with planimetric measurement of the resulting routes, reveals five consistent structural patterns.

Fleet dimensioning invariance: Across all four high-demand instances, the route cardinality remains identical between flat and topographic spaces (SLAS100: 12/12; SLFS100: 10/10; SLAM250: 23/23; SLFM250: 25/25), as evidenced in Figures 3–6. Topographic constraints modify node assignment and visit sequences but do not perturb the strategic fleet optimum.

Scale-dependent spatial deformation: Topography redistributes planimetric route length rather than merely increasing it, and the direction of the redistribution depends on network scale. Measuring, for each customer, the length of the route serving it, and grouping customers into three concentric rings by distance from the depot, the deformation at 100 nodes is peripheral: the outermost ring lengthens (+2.0% in SLAS100, +12.8% in SLFS100) while the core compacts or grows less (−12.5% and +7.0%, respectively), so peripheral trajectories extend along enveloping arcs while the depot neighbourhood remains geometrically similar to the flat model. At 250 nodes the pattern reverses: the innermost ring absorbs the deformation (+18.1% in SLAM250, +6.2% in SLFM250) whereas the outermost ring is essentially unchanged (+0.1% and +4.0%). Under flat costs the solution stratifies by distance, with compact local routes near the depot; at high density the orographic constraint dissolves this stratification and near-depot customers are absorbed

into long radial corridors spanning the full radial range. This is the same spatial saturation effect that drives the petal-to-star transition described below.

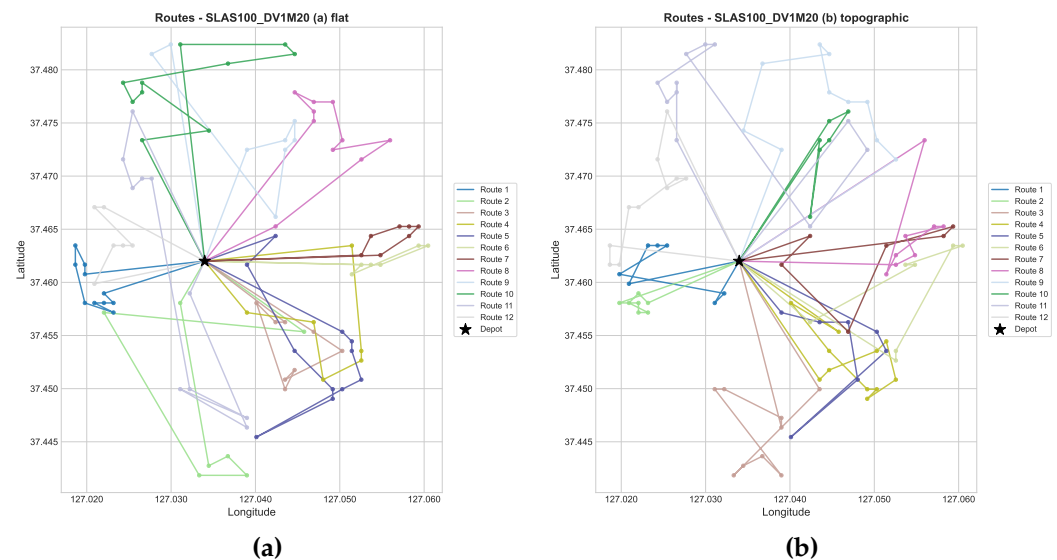


Figure 3. Routing comparison for the 100-node residential instance (SLAS100_DV1M20; 12 routes in both scenarios). (a) Flat Euclidean model; (b) orography-constrained model. The star marks the depot. Each route is identified in the legend by a unique colour; route indices are assigned independently in each panel and are therefore not comparable across them. Wide, compact “petal” cluster morphology is observed due to low spatial scale. The deformation follows the expected pattern at this scale: the innermost ring compacts (-12.5%) while peripheral routes lengthen ($+2.0\%$), with a slight net reduction in total planimetric length (-2.0%).

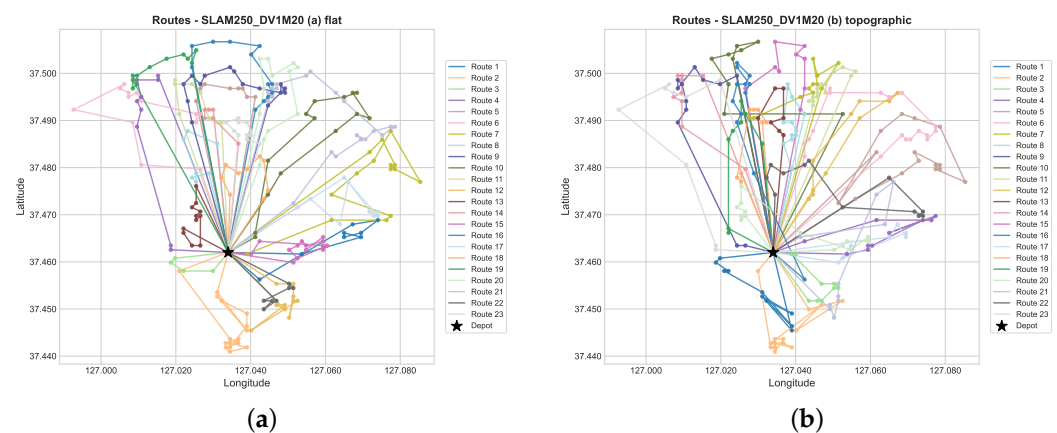


Figure 4. Evidence of radial route structure and core deformation in the large-scale residential instance (SLAM250_DV1M20; 23 routes in both scenarios). (a) Flat 2D space; (b) topographic model. The star marks the depot. Each route is identified in the legend by a unique colour; route indices are assigned independently in each panel and are therefore not comparable across them. Topographic constraints increase total planimetric length by 10.4% at constant fleet size, with the increase concentrated in routes serving near-depot customers ($+18.1\%$) rather than peripheral ones ($+0.1\%$).

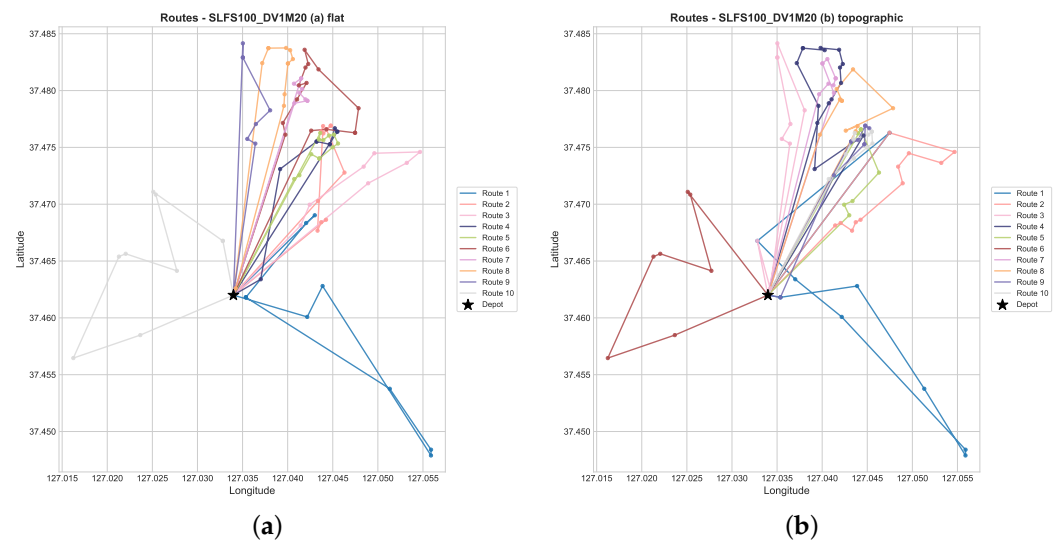


Figure 5. Spatial reconfiguration under commercial asymmetry for the 100-node restaurant network (SLFS100_DV1M20; 10 routes in both scenarios). (a) Flat model; (b) orographic model. The star marks the depot. Each route is identified in the legend by a unique colour; route indices are assigned independently in each panel and are not comparable across them. The bimodal distribution (dense valley core versus isolated peripheral hillside customers) generates the most aggressive disruptions observed at this scale: total planimetric length grows by 7.7% and the longest route by 23.3%, with peripheral routes lengthening nearly twice as much as the core (+12.8% versus +7.0%).

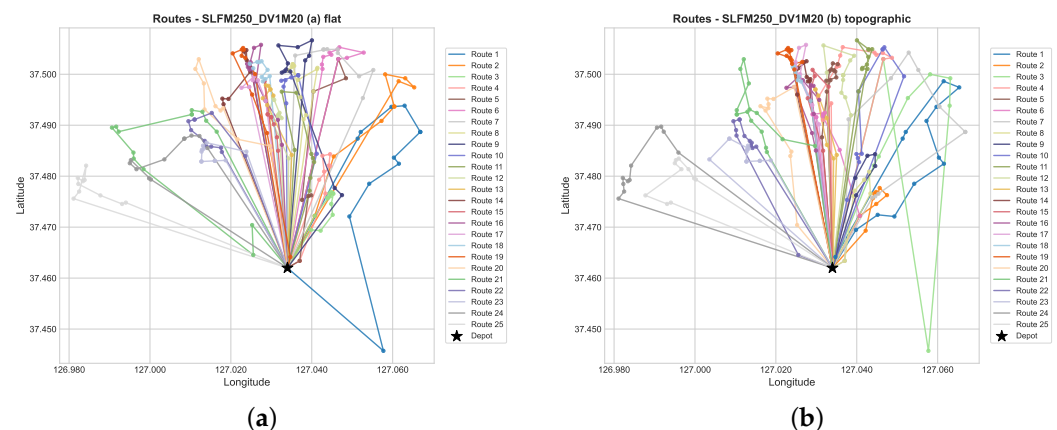


Figure 6. Topological homogenization due to high density in the 250-node commercial instance (SLFM250_DV1M20; 25 routes in both scenarios). (a) Flat 2D space; (b) topographic model. The star marks the depot. Each route is identified in the legend by a unique colour; route indices are assigned independently in each panel and are not comparable across them. At spatial saturation the morphology converges toward the same radial-stellar pattern observed in residential networks, and the deformation concentrates in the core (+6.2% for the innermost ring) rather than in the periphery (+4.0%), with a modest net increase in total planimetric length (+1.7%).

5. Discussion

This section interprets the experimental results in the context of the existing literature, identifies the practical implications, and acknowledges the limitations of the present study.

5.1. Algorithmic Efficiency

The H-MA-LNS demonstrates consistent algorithmic exploration superiority over the literature reference [10] when evaluated on a flat Euclidean space. The magnitude of this advantage exhibits a pronounced scale dependence: at 100 nodes it ranges between -5.1% (SLAS100) and $+1.6\%$ (SLFS100), while at 250 nodes the divergence amplifies to -17.9%

(SLAM250) and -14.3% (SLFM250). This confirms that the global exploration of the proposed memetic architecture (GA + LNS) exploits the expansion of the search space more effectively than the sweep and simulated annealing methods prevalent in the state of the art.

A topological asymmetry is also detected at 100 nodes: dispersed residential networks (SLAS100) reveal systematic improvements (-3.7% to -6.5%), whereas compact restaurant clusters (SLFS100) show marginal variations ($+0.1\%$ to $+2.9\%$). This confirms that the evolutionary mechanism provides critical value in topologies with high geographic dispersion, overcoming the limited reach of classical constructive heuristics.

5.2. Quantification of the Topographic Gap

A central result of this study is the quantification of the topographic gap, defined as the percentage cost difference between 2D planning and execution over real orography. With a global mean difference of 14.1% (Wilcoxon $p = 0.00195$), this work demonstrates that traditional symmetric metrics systematically underestimate the physical effort of logistics transport. The gap depends on three factors.

Scale effect: Expanded networks (250 nodes) suffer moderately greater penalties (16.7% in SLAM250) than compact ones (8.0% in SLAS100), although the ordering is not monotonic across typologies. Extending the operational radius from 2.5 km to 5 km increases the probability of traversing zones with steep gradients.

Demand topology: At 100 nodes, the bimodal distribution of restaurants (dense valley center and isolated peripheral hillsides) generates altimetric contrasts that amplify the gap (18.7% vs. 8.0% residential). When scaling to 250 nodes, spatial saturation homogenizes this bimodality, inducing an impact inversion, where the residential distribution assimilates a higher penalty (16.7% vs. 13.5%).

Load consolidation: High-demand-per-node scenarios (DV1M20) exhibit a lower relative topographic gap. Higher volume demands shorter intra-route consolidations, reducing sustained vehicle exposure to prolonged gradients—a penalty fully absorbed by the extended routes of low-density scenarios (DV1M5).

5.3. Net Efficiency Trade-Off

The GAP 1–3 indicator quantifies the net trade-off between algorithmic optimization and topographic friction penalty, revealing a structural bifurcation dependent on combinatorial complexity.

At 100 nodes (topographic dominance), gravitational physics overcomes the algorithmic gain, yielding positive net gaps ($+3.0\%$ in SLAS100 and $+20.4\%$ in SLFS100). This increase does not reflect an optimization deficiency but rather quantifies the systematic underestimation error in the literature, which reports theoretical global minima that prove mechanically unattainable in real operations.

At 250 nodes (algorithmic compensation), the exploration power of H-MA-LNS fully compensates the orographic penalty. The model generates physically feasible, terrain-adapted solutions that achieve lower costs (-1.2% and -0.8%) than the flat and infeasible routes of the state of the art. This finding indicates that prior methods produce suboptimal solutions in high-density urban problems.

5.4. Computational Acceleration Through Topological Pruning

A counterintuitive finding is that incorporating topographic constraints reduces computation times rather than increasing them. The mechanical feasibility constraint ($\theta_{\max} = 25\%$) eliminates a priori all arcs whose gradients exceed the traction capacity of heavy vehicles. In the Seoul instances, this reduces the feasible-arc set cardinality by 51% , proportionally contracting the stochastic LNS search space. The practical consequence is that integrating transport physics not only improves model fidelity but paradoxically en-

hances computational tractability—a result that contrasts with the conventional assumption that incorporating greater realism increases computational cost.

5.5. Implications for Logistics Practice

The results support four implications for urban logistics in cities with complex topography. First, budgeting logistics operations using projected Euclidean distances induces structural estimation errors between 8.0% and 18.7% (systematic cost underestimation). Second, planimetric models may generate routes with gradients exceeding the traction capacity of heavy vehicles (slopes $> 25\%$); the implemented topological constraint demonstrably guarantees strictly navigable trajectories (mechanical feasibility). Third, bidirectional route design (typical of symmetric models) causes critical inefficiencies by omitting the asymmetric gravitational penalty on ascending versus descending slopes (operational irreversibility). Fourth, the algorithmic integration of orography reduces computation times by up to 51% through search-space pruning via infeasible-arc exclusion (induced computational acceleration).

5.6. Limitations and Future Work

While the ACVRP-TC model substantially reduces the gap between mathematical planning and physical operation, several limitations should be acknowledged. The current model assumes constant speed and does not incorporate time windows, which limits its direct applicability to time-critical delivery scenarios. The GCF parameters (α , β , κ) were validated through sensitivity analysis but not calibrated against field measurements from specific vehicle types, which would further strengthen the physical fidelity of the model. The topographic gap of 14.1% was measured on Seoul instances; generalization to other mountain cities (e.g., Andean or Alpine) requires validation with local DEM data and road network characteristics.

Three methodological extensions emerge as priority future directions: (i) transition toward a multi-objective framework that explicitly exposes the trade-off between operational cost and pollutant emissions through Pareto frontiers; (ii) incorporation of environmental stochasticity and dynamic traffic, adapting the impedance function $\Omega(\theta_{ij})$ to climatic uncertainty and congestion windows; and (iii) deployment of heterogeneous fleets that relax the $\theta_{\max} = 25\%$ constraint through light or electric vehicles capable of penetrating extreme hillside clusters.

5.7. Reproducibility and Data Availability

To ensure result replicability, the following provisions are adopted. The complete H-MA-LNS implementation—encompassing the local search operators, the LNS-SA module, and the final polishing phase—is publicly available at <https://github.com/amrestrepof/ACVRP-Topographic-Results> (accessed on 4 May 2026), under the MIT license. While the 24 Seoul-based instances were originally proposed by [10], our repository provides the necessary SRTM elevation mosaics and the processed road graphs (OSMnx) required to replicate the topographic impedance.

The repository also hosts the exact MILP verification script (`milp_verification.py`) and its complete results log (`milp_verification_results.csv`), which reproduce the model verification of Section 3.3: the construction of the node-level GCF cost matrices from the SLAS100 road distances and the SRTM mosaics, the CBC exact resolution of the ten subsampled instances, the independent post hoc constraint checks, and the H-MA-LNS comparison runs.

All experimental outputs, comprising 240 individual runs derived from executing each instance with sequential seeds $s \in \{0, 1, \dots, 9\}$ (initialized via `numpy.random.seed(s)`), are comprehensively documented across 24 per-instance result files hosted directly in the

GitHub repository. These individual logs are further consolidated into instance-specific summary files, such as `resultados_vrp_SLFM250_DV1M5.xlsx`. Execution times correspond to wall-clock time measured with `time.perf_counter()`. It is important to note that the evolutionary algorithm was executed sequentially to establish a standardized performance baseline; parallelization (Joblib, `n_jobs=-1`) was strictly limited to the impedance matrix preprocessing phase, and its duration is rigorously excluded from the reported metaheuristic efficiency metrics. The computational environment is fully documented via a `requirements.txt` file, specifying Python 3.11.7, NumPy 1.26.2, Pymoo 0.6.1, OSMnx 1.7.1, NetworkX 3.2.1, Rasterio 1.3.9, and Joblib 1.3.2.

6. Conclusions

This study addressed the challenge of logistics planning in mountain cities through the formulation and solution of the Asymmetric Capacitated Vehicle Routing Problem with Topographic Constraints (ACVRP-TC). Unlike traditional approaches that simplify the operational environment as an idealized two-dimensional Euclidean space, this research integrated heavy-duty transport physics and the structural asymmetry of urban road networks into a unified algorithmic framework.

Regarding computational performance and scalability, the proposed H-MA-LNS algorithm obtained competitive results against the state of the art across three validation sets. On the Augerat symmetric instances, it achieved a mean GAP of 0.2%, with 67% exact optima. On the Fischetti asymmetric instances, it converged with a GAP of 1.2% in 87.8% less time than the exact branch-and-bound method (65 s vs. 530.9 s), bearing in mind that exact methods pursue optimality certificates while the proposed heuristic targets high-quality solutions efficiently. On the complex asymmetric instances of Rodríguez and Ruiz (Madrid), it obtained a GAP of 2.0%, surpassing the best known solution in 50% of cases. Globally, the evaluation across 24 geospatial Seoul instances demonstrated a mean algorithmic improvement of 9.4% over the literature reference (GAP 1–2), with a scale-dependent advantage that amplifies from −5.1% at 100 nodes to −17.9% at 250 nodes.

With respect to algorithmic acceleration through topological pruning, it was demonstrated that the physical complexity of the terrain, far from increasing computational cost, optimizes problem tractability. The mechanical feasibility constraint ($\theta_{\max} = 25\%$) acts as a topological filter that reduces the cardinality of the feasible-arc set by 51%, proportionally contracting the stochastic LNS search space. This pruning effect induced acceleration of 38% to 51% in convergence times (topo/flat ratio of $0.62\times$ at 100 nodes and $0.49\times$ at 250 nodes), validating that the linearized generalized cost function (GCF) constitutes a computationally superior alternative to conventional MINLP approaches.

Concerning the structural quantification of the topographic gap, the integration of digital elevation models (DEMs) and real cartography enabled the measurement of a global topographic gap of 14.1% (Wilcoxon $p = 0.00195$), quantifying the divergence in operational costs between theoretical 2D planning and physical execution over real orography. This figure is specific to the Seoul instances; its extrapolation to other mountain cities requires validation with local elevation data and road network characteristics (Section 5.6). This gap exhibits a structural dependence on three factors: network scale (group means rising from 12.9% at 100 nodes to 15.1% at 250 nodes), demand typology (bimodal distributions amplify the penalty), and load consolidation (high demand per node mitigates the relative gap). At 250 nodes, the algorithmic advantage of H-MA-LNS fully compensated the orographic penalty, generating physically feasible, terrain-adapted solutions that proved more economical (−1.2% in SLAM250, −0.8% in SLFM250) than the flat and infeasible routes of the state of the art.

Finally, regarding inferential robustness and industrial applicability, the statistical analysis across the full experimental corpus yielded coefficients of variation averaging 1.52% under the flat model and 3.08% under topographic constraints, with all instances but one below the 5% threshold conventionally required in metaheuristic design, and 95% confidence intervals that exclude zero for the topographic gap in every instance. The algorithmic stability and the integration of real geospatial data suggest that the model is applicable to logistics planning in urban environments with complex orography.

Author Contributions: Conceptualization, A.M.R.-F. and E.M.T.-O.; methodology, A.M.R.-F. and O.V.-R.; software, A.M.R.-F.; validation, A.M.R.-F., O.V.-R. and E.M.T.-O.; formal analysis, A.M.R.-F.; investigation, A.M.R.-F.; resources, E.M.T.-O. and O.D.C.-G.; data curation, A.M.R.-F.; writing—original draft preparation, A.M.R.-F.; writing—review and editing, O.V.-R., E.M.T.-O. and O.D.C.-G.; visualization, A.M.R.-F.; supervision, O.V.-R. and O.D.C.-G.; project administration, E.M.T.-O.; funding acquisition, A.M.R.-F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Colombian Ministry of Science, Technology and Innovation (Minciencias) through the project “Formación de capital humano de alto nivel para atender las demandas territoriales de la región Eje Cafetero—Caldas, Antioquia, Quindío, Risaralda” [BPIN 2022000100055].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data, source code supporting the findings of this study are openly available in the GitHub repository at <https://github.com/amrestrepof/ACVRP-Topographic-Results> (accessed on 4 May 2026). The repository provides the complete H-MA-LNS implementation under the MIT license, along with the SRTM elevation mosaics, processed OSMnx road graphs, and the resulting asymmetric cost matrices for the 24 Seoul-based instances. Furthermore, it contains the comprehensive execution logs for all 240 computational runs (10 sequential seeds $\times$ 24 instances) and interactive Folium-based route visualizations. The classical benchmark instances proposed by Augerat [39], Fischetti [13], and Rodríguez and Ruiz [7] are publicly accessible via CVRPLib and their respective original cited sources.

Acknowledgments: Special recognition is extended to the Universidad Nacional de Colombia, Manizales Campus, for providing the facilities and resources during the research stay in which this work was consolidated.

Conflicts of Interest: The authors declare no conflicts of interest. The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ACVRP-TC	Asymmetric Capacitated Vehicle Routing Problem with Topographic Constraints
BKS	Best known solution
CVRP	Capacitated vehicle routing problem
CV	Coefficient of variation
DEM	Digital elevation model
GCF	Generalized cost function
H-MA-LNS	Hybrid memetic algorithm with large neighborhood search
LNS	Large neighborhood search
MILP	Mixed-integer linear programming
MINLP	Mixed-integer nonlinear programming
MTZ	Miller–Tucker–Zemlin

SA	Simulated annealing
SRTM	Shuttle Radar Topography Mission
VRP	Vehicle routing problem

Appendix A. Complexity Derivations for Tables 7 and 8

This appendix provides brief formal derivations for the asymptotic bounds stated in Tables 7 and 8. Throughout, n denotes the number of customers, K the number of active routes (with $K \leq n$ and, in practice, $K = O(n/\bar{q})$ for mean demand $\bar{q}$), q the number of nodes removed in the LNS destruction phase, $\ell \leq 3$ the maximum segment length, P the population size, and G the number of generations.

DP split: The recursion of Equation (18) evaluates, for each end position $j \in \{0, \dots, n-1\}$, all start positions $i \leq j$, giving at most $n(n+1)/2$ state transitions. By Equation (19), each transition is computed incrementally from its predecessor in $O(1)$, so the total is $O(n^2)$. Early termination when $\sum_{k=i}^j q_{c_k} > Q$ prunes the inner loop to the maximal capacity-feasible window, which tightens the practical bound to $O(n \cdot w)$, with w the largest number of consecutive customers fitting in one vehicle ($w \ll n$ for high-demand instances).

Intra-route operators: Swap-intra enumerates all position pairs (i, j) within a route of length r , i.e., $O(r^2)$ moves; Equations (20) and (21) evaluate each delta with at most eight arc lookups; hence, $O(1)$ per move and $O(r^2)$ per route, which summed over routes is bounded by $O(n^2)$ since $\sum_r r \leq n$. Directional or-opt enumerates $O(r)$ segments of each length $\ell \in \{1, 2, 3\}$ and $O(r)$ insertion positions, with the $O(1)$ two-delta evaluation of Equation (22) giving $O(\ell r^2) = O(r^2)$ per route.

Inter-route operators: Relocate tries each of the n customers in each position of each other route; with per-route load caching the feasibility check is $O(1)$, so the bound is $O(n \cdot n \cdot K) = O(n^2 K)$ in the worst case, where positions are distributed over K routes. Inter-route swap enumerates customer pairs across route pairs; $O(n^2)$ pairs over $O(K^2)$ route combinations in the worst case, i.e., $O(n^2 K^2)$. Cross-exchange additionally enumerates $\ell_i, \ell_j \in \{1, 2, 3\}$ segment lengths at each cut pair; Equation (23) evaluates each exchange via the four cut-point arcs in $O(1)$, yielding $O(n^2 K^2 \ell^2)$.

Composite chains: LS_{light} applies the four cheaper operators sequentially; its single-pass cost is dominated by inter-route swap, $O(n^2 K^2)$. LS_{full} adds cross-exchange and is dominated by it: $O(n^2 K^2 \ell^2)$.

LNS-SA iteration: Worst removal sorts marginal costs, $O(n \log n)$; relatedness removal is $O(qn)$. Regret- k insertion evaluates, for each of the q removed customers, its best insertion in each route, $O(qKn)$ per reinsertion round, repeated q times: $O(q^2 Kn)$. Adding the post-repair local search gives $O(q^2 Kn + LS)$ per iteration, where $LS \in \{O(n^2 K^2), O(n^2 K^2 \ell^2)\}$ according to the cadence of Equation (30).

Totals: One GA generation performs $O(P)$ fitness evaluations dominated by the DP split, $O(Pn^2)$, over G generations, $O(GPn^2)$. Adding I_{Ins} refinement iterations yields the per-run bound $O(GPn^2 + I_{\text{Ins}}(q^2 Kn + n^2 K^2))$ of Table 7, polynomial in all inputs. The comparison of Table 8 follows because the GCF replaces per-evaluation nonlinear cost computation $O(f_{NL})$ with a precomputed matrix lookup $O(1)$, and the pruning of Equation (10) restricts enumeration to $|A \setminus A_{\text{infeas}}|$ arcs.

References

1. Li, J.; Zheng, Y.; Liu, B.; Chen, Y.; Zhong, Z.; Dong, C.; Wang, C. The Synergistic Relationship between Low-Carbon Development of Road Freight Transport and Its Economic Efficiency—A Case Study of Wuhan, China. *Sustainability* **2024**, *16*, 2636. [CrossRef]
2. Pustilnik, M.; Borrelli, F. Clustering heuristics for Stochastic Energy Capacitated Vehicle Routing Problem (ECVRP). *Int. J. Syst. Sci. Oper. Logist.* **2025**, *12*, 2451764. [CrossRef]
3. Aghaabdollahian, B.; Javadi, B.; Aazami, A. An overview of vehicle routing problems: A comprehensive classification and description of types. *Transp. Res. Interdiscip. Perspect.* **2026**, *36*, 101808. [CrossRef]

4. Brunner, C.; Giesen, R.; Klapp, M.A.; Flórez-Calderón, L. Vehicle routing problem with steep roads. *Transp. Res. Part A Policy Pract.* **2021**, *151*, 1–17. [\[CrossRef\]](#)
5. Bonilla-Bedoya, S.; López-Ulloa, M.; Mora-Garcés, A.; Macedo- Pezzopane, J.E.; Salazar, L.; Ángel Herrera, M. Urban soils as a spatial indicator of quality for urban socio-ecological systems. *J. Environ. Manag.* **2021**, *300*, 113556. [\[CrossRef\]](#)
6. Suarez-Chilma, V.F.; Sarache, W.A.; Costa, Y.J. Una Solución al Enrutamiento de Vehículos en Ciudades Montañasas Considerando Aspectos Ambientales y Económicos. *Inf. Tecnol.* **2018**, *29*, 3–14. [\[CrossRef\]](#)
7. Rodríguez, A.; Ruiz, R. A study on the effect of the asymmetry on real capacitated vehicle routing problems. *Comput. Oper. Res.* **2012**, *39*, 2142–2151. [\[CrossRef\]](#)
8. Restrepo-Franco, A.M.; Valencia-Rodríguez, O.; Toro-Ocampo, E.M. The vehicle routing problem as applied to residential solid waste collection operations: Systematic literature review. *Int. J. Ind. Eng. Comput.* **2025**, *16*, 197–220. [\[CrossRef\]](#)
9. Lai, D.; Costa, Y.; Demir, E.; Florio, A.M.; Van Woensel, T. The pollution-routing problem with speed optimization and uneven topography. *Comput. Oper. Res.* **2024**, *164*, 106557. [\[CrossRef\]](#)
10. Lee, K.; Chae, J. A Proposal and Analysis of New Realistic Sets of Benchmark Instances for Vehicle Routing Problems with Asymmetric Costs. *Appl. Sci.* **2021**, *11*, 4790. [\[CrossRef\]](#)
11. Suzuki, Y. A new truck-routing approach for reducing fuel consumption and pollutants emission. *Transp. Res. Part D Transp. Environ.* **2011**, *16*, 73–77. [\[CrossRef\]](#)
12. Schröder, M.; Cabral, P. Eco-friendly 3D-Routing: A GIS based 3D-Routing-Model to estimate and reduce CO₂-emissions of distribution transports. *Comput. Environ. Urban Syst.* **2019**, *73*, 40–55. [\[CrossRef\]](#)
13. Fischetti, M.; Toth, P.; Vigo, D. A Branch-and-Bound Algorithm for the Capacitated Vehicle Routing Problem on Directed Graphs. *Oper. Res.* **1994**, *42*, 846–859. [\[CrossRef\]](#)
14. Herrero, R.; Rodríguez, A.; Cáceres-Cruz, J.; Juan, A.A. Solving vehicle routing problems with asymmetric costs and heterogeneous fleets. *Int. J. Adv. Oper. Manag.* **2014**, *6*, 58–80. [\[CrossRef\]](#)
15. Liu, Y.; Yuan, Y.; Shen, J.; Gao, W. Emergency response facility location in transportation networks: A literature review. *J. Traffic Transp. Eng. (Engl. Ed.)* **2021**, *8*, 153–169. [\[CrossRef\]](#)
16. Fan, H.; Zhang, Y.; Tian, P.; Lv, Y.; Fan, H. Time-dependent multi-depot green vehicle routing problem with time windows considering temporal-spatial distance. *Comput. Oper. Res.* **2021**, *129*, 105211. [\[CrossRef\]](#)
17. Nagata, Y. Edge assembly crossover for the capacitated vehicle routing problem. In *Proceedings of the European Conference on Evolutionary Computation in Combinatorial Optimization (EvoCOP)*, Valencia, Spain, 11–13 April 2007; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4446, pp. 142–153. [\[CrossRef\]](#)
18. Köroğlu, C. The Vehicle Routing Problem with Traffic Conditions. Master's Thesis, Dokuz Eylül University, Graduate School of Natural and Applied Sciences, İzmir, Turkey, 2018.
19. Meniz, B.; Tiryaki, F. Genetic Algorithm Optimization with Selection Operator Decider. *Arab. J. Sci. Eng.* **2025**, *50*, 6931–6941. [\[CrossRef\]](#)
20. Demir, E.; Bektaş, T.; Laporte, G. A review of recent research on green road freight transportation. *Eur. J. Oper. Res.* **2014**, *237*, 775–793. [\[CrossRef\]](#)
21. Zhao, M.; Zhao, N. Method for the Construction of Urban Road Digital Elevation Models Integrated with Semantic Information. *Appl. Sci.* **2023**, *13*, 4210. [\[CrossRef\]](#)
22. Toro, E.M.; Franco, J.F.; Echeverri, M.G.; Guimarães, F.G. A multi-objective model for the green capacitated location-routing problem considering environmental impact. *Comput. Ind. Eng.* **2017**, *110*, 114–125. [\[CrossRef\]](#)
23. Jeong, E.; Cho, M.; Cho, H.; Cho, B.; Han, S. Characteristics of forest road cut slopes affecting the movement of mammals in South Korea. *For. Sci. Technol.* **2021**, *17*, 155–161. [\[CrossRef\]](#)
24. Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. *Global Sensitivity Analysis. The Primer*; John Wiley & Sons: Chichester, UK, 2008. [\[CrossRef\]](#)
25. Flores-Luyo, L.; Agra, A.; Figueiredo, R.; Ocaña, E. Mixed integer formulations for a routing problem with information collection in wireless networks. *Eur. J. Oper. Res.* **2020**, *280*, 621–638. [\[CrossRef\]](#)
26. Pisinger, D.; Ropke, S. A unified heuristic for a large class of Vehicle Routing Problems. *Eur. J. Oper. Res.* **2006**, *171*, 750–775. [\[CrossRef\]](#)
27. Ropke, S.; Pisinger, D. An Adaptive Large Neighborhood Search Heuristic for the Pickup and Delivery Problem with Time Windows. *Transp. Sci.* **2006**, *40*, 455–472. [\[CrossRef\]](#)
28. Segerstedt, A. A simple heuristic for vehicle routing—A variant of Clarke and Wright's saving method. *Int. J. Prod. Econ.* **2014**, *157*, 74–79. [\[CrossRef\]](#)
29. Arifuddin, A.; Utamima, A.; Mahananto, F.; Vinarti, R.A.; Fernanda, N. Optimizing the Capacitated Vehicle Routing Problem at PQR Company: A Genetic Algorithm and Grey Wolf Optimizer Approach. *Procedia Comput. Sci.* **2024**, *234*, 420–427. [\[CrossRef\]](#)
30. Majumdar, J.; Bhunia, A. Genetic algorithm for asymmetric traveling salesman problem with imprecise travel times. *J. Comput. Appl. Math.* **2011**, *235*, 3063–3078. [\[CrossRef\]](#)

31. Ochelska-Mierzejewska, J.; Poniszewska-Maraña, A.; Maraña, W. Selected Genetic Algorithms for Vehicle Routing Problem Solving. *Electronics* **2021**, *10*, 3147. [\[CrossRef\]](#)
32. Romero-Ocampo, A.D.; Cosío-León, M.; Valenzuela-Alcaraz, V.M.; Avilés-Rodríguez, G.J.; Martínez-Vargas, A. Efecto de la calibración de parámetros mediante un diseño Taguchi L_93^4 en el algoritmo GRASP resolviendo el problema de rutas de vehículos con restricciones de tiempo. *Comput. Sist.* **2018**, *22*, 657–673. [\[CrossRef\]](#)
33. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-Generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage, AK, USA, 4–8 August 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 2623–2631. [\[CrossRef\]](#)
34. Jiaojiao, M.; Gulyás, L.; Botzheim, J. Comparing Lamarckian and Baldwinian Approaches in Memetic Optimization. In *Advances in Computational Collective Intelligence (ICCCI 2023)*, Budapest, Hungary, 27–29 September 2023; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2023; pp. 521–533. [\[CrossRef\]](#)
35. Rudolph, G. Convergence Analysis of Canonical Genetic Algorithms. *IEEE Trans. Neural Netw.* **1994**, *5*, 96–101. [\[CrossRef\]](#)
36. Hajek, B. Cooling Schedules for Optimal Annealing. *Math. Oper. Res.* **1988**, *13*, 311–329. [\[CrossRef\]](#)
37. Blank, J.; Deb, K. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* **2020**, *8*, 89497–89509. [\[CrossRef\]](#)
38. Boeing, G. Modeling and Analyzing Urban Networks and Amenities with OSMnx. *Geogr. Anal.* **2025**, *57*, 567–577. [\[CrossRef\]](#)
39. Augerat, P.; Naddef, D.; Belenguer, J.M.; Benavent, E.; Corberan, A.; Rinaldi, G. Computational Results with a Branch and Cut Code for the Capacitated Vehicle Routing Problem; Technical Report RR949-M; IMAG: Grenoble, France, 1995. Available online: <https://www.osti.gov/etdeweb/biblio/289002> (accessed on 18 February 2026).
40. Akpinar, S. Hybrid large neighbourhood search algorithm for capacitated vehicle routing problem. *Expert Syst. Appl.* **2016**, *61*, 28–38. [\[CrossRef\]](#)
41. Stanojević, M.; Stanojević, B.; Vujošević, M. Enhanced savings calculation and its applications for solving capacitated vehicle routing problem. *Appl. Math. Comput.* **2013**, *219*, 10302–10312. [\[CrossRef\]](#)
42. Vidal, T. Hybrid genetic search for the CVRP: Open-source implementation and SWAP* neighborhood. *Comput. Oper. Res.* **2022**, *140*, 105643. [\[CrossRef\]](#)
43. U.S. Geological Survey. EarthExplorer. 2024. Available online: <https://earthexplorer.usgs.gov/> (accessed on 22 May 2024).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.