


Article

Multi-Shift Scheduling of Electric Service Operations Under Fuzzy Uncertainty via Preference-Guided Deep Learning: The Single-Vehicle Case

Francesco Nucci 

Department of Engineering for Innovation, University of Salento, 73100 Lecce, Italy; francesco.nucci@unisalento.it

Abstract

The electrification of field service fleets introduces complex constraints: shift limits, overtime fairness, and battery–range feasibility. This paper proposes the Multi-Shift Single Electric Vehicle Routing Problem under Possibilistic Uncertainty (MS-SEVRP-PU), a formulation focused on a single-vehicle multi-shift planning unit and capturing imprecise travel/service times and state-of-charge dynamics. Travel durations and energy consumption are modelled as triangular fuzzy numbers to reflect expert knowledge when probabilistic data is limited. A closed-form credibility function evaluates overtime risk, while an Ordered Weighted Averaging (OWA) aggregation of per-shift risks ensures fairness by discouraging systematic overload on specific shifts. To solve this multi-objective problem, we develop a Pareto-Conditioned Transformer with risk-aware and battery-conscious large neighbourhood search (PCT-RABLNS), combining a preference-conditioned attention policy with targeted local search. Computational experiments on calibrated municipal maintenance case studies indicate that PCT-RABLNS improves hypervolume by 2–5% over strong baselines and reduces maximum shift overtime risk by 15–25%, with a marginal makespan overhead of only 1–3%. The results demonstrate that the proposed framework is a promising decision-support approach for energy-aware, risk-fair, and operationally compliant planning of single-vehicle, multi-shift electric service operations, jointly integrating multi-shift routing, fuzzy uncertainty, and preference-conditioned reinforcement learning. The paper also discusses how the framework can be extended to multi-vehicle settings.

Keywords: fuzzy logic; deep reinforcement learning; multi-objective optimisation; large neighbourhood search; green logistics; workforce scheduling; OWA operator



Academic Editor: Antonio Gil Bravo

Received: 28 March 2026

Revised: 13 May 2026

Accepted: 14 May 2026

Published: 16 May 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

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

1. Introduction

The rapid electrification of field service and maintenance fleets introduces a new layer of operational complexity that classical routing models are ill-equipped to handle jointly: shift-length compliance, fairness in overtime exposure across consecutive shifts, and the hard feasibility constraint imposed by the finite energy capacity of electric vehicles (EVs). Field service organisations increasingly operate over multiple consecutive shifts while trying to control service quality, labour costs, and compliance with safety constraints [1–3]. When planning across multiple shifts, decisions made early in the schedule carry forward and affect later periods: routing, task ordering, and time buffers interact across shift boundaries in ways that single-period models do not capture [4]. The issue is not merely conceptual—overtime clustered at or near shift changes create practical problems for workforce fatigue, safety, and regulatory compliance [5].

The adoption of electric vehicles in municipal maintenance and service operations is accelerating under environmental regulation and fleet-electrification mandates [6]. However, EVs impose a battery–range feasibility constraint that has no analogue in conventional vehicle routing: the energy consumed along a route depends on load, speed profile, road gradient, and auxiliary systems, and must remain within the vehicle’s state-of-charge (SoC) limits at all times. When routes span multiple work shifts, the interaction between shift boundaries, partial recharging opportunities, and residual SoC creates a tightly coupled scheduling problem that the existing multi-shift and E-VRP formulations address only separately [7,8].

Uncertainty compounds the planning difficulty. Travel times vary with traffic, weather, and local obstructions; service durations depend on asset conditions, accessibility, and task complexity [9,10]. For EVs, energy consumption is further affected by uncertain traffic speed profiles and ambient temperature, making deterministic energy-budget constraints particularly unreliable [11]. Deterministic formulations, which assume exact knowledge of these quantities, therefore offer a poor fit for many field-service environments [12]. Stochastic modelling is a natural alternative when abundant, high-quality historical data exist [13]. In many municipal and specialised-service settings, however, such data are scarce or unreliable, and practitioners more often rely on ranges, rough likelihoods, and expert judgement rather than well-founded frequency estimates [14].

This type of imprecise, qualitative knowledge is well matched to possibilistic (fuzzy) representations: membership functions and possibility distributions encode what experts can say confidently—upper and lower bounds, plausible modes—without forcing an arbitrary probabilistic law [15–17]. Possibility theory, therefore, offers a practical middle ground: it keeps models tractable for optimisation while better reflecting the form of available information in many applied contexts [18,19]. In the EV context, modelling both travel times and energy consumption rates as triangular fuzzy numbers captures the expert knowledge about variability without requiring the statistical infrastructure that stochastic models demand.

A separate but related concern is how overtime risk is distributed over shifts. Minimizing a global metric, such as total makespan, can hide undesirable temporal concentrations of overtime: the schedule may be efficient on average, but expose particular shifts to a high share of risk, with negative consequences for safety and fairness [4,5]. Aggregation operators that are sensitive to the distributional properties of per-shift risk—Ordered Weighted Averaging (OWA) being a prominent example—provide a way to formalise fairness-aware objectives and to penalise inequitable risk concentrations [20,21].

On the methodological front, modern machine learning methods have proven useful for routing and related combinatorial problems. Attention-based architectures and reinforcement learning have shown strong empirical performance on a range of VRP variants, particularly when they can exploit problem structure and learn reusable heuristics [22–25]. At the same time, classical metaheuristics such as Large Neighbourhood Search (LNS) remains powerful and flexible. Recent work suggests that hybrid approaches—where learned policies propose or prioritize moves, while LNS provides robust, problem-aware neighbourhood exploration—can combine the best of both worlds [26–28].

In this paper, we bring these ideas together for the case of single-vehicle routing over multiple shifts under possibilistic uncertainty, with an explicit emphasis on battery feasibility and temporal risk equity. Concretely, we propose the **Multi-Shift Single Electric Vehicle Routing Problem under Possibilistic Uncertainty (MS-SEVRP-PU)**, a bi-objective formulation that balances total makespan and a fairness-sensitive aggregation of per-shift overtime credibilities via an OWA operator [20], subject to fuzzy energy-consumption constraints that enforce SoC feasibility throughout each shift. We focus on single-vehicle

instances because they isolate the core temporal, energy, and risk-allocation issues without adding the combinatorial complexity of multi-vehicle assignment; such instances are also directly relevant to many municipal scenarios (e.g., inspection or maintenance units operating independently) [1].

We organise the study around the following four research questions:

- **RQ1—Formulation:** How can possibilistic uncertainty, battery–feasibility constraints, and fairness-aware overtime risk be jointly represented in a multi-shift electric vehicle routing model?
- **RQ2—Energy modelling:** How can fuzzy energy consumption rates be integrated into a closed-form credibility framework without sacrificing computational tractability?
- **RQ3—Solution method:** Can a preference-conditioned deep model be effectively combined with risk-aware and battery-conscious local search to boost Pareto front quality over traditional baselines?
- **RQ4—Practical impact:** What trade-offs appear between operational efficiency (makespan), temporal risk distribution, and energy feasibility in practical metropolitan maintenance scenarios?

The main contributions of this work are as follows:

- **Modelling:** MS-SEVRP-PU, a multi-shift single-electric-vehicle routing model that combines possibilistic uncertainty with temporal fairness (OWA-based) and fuzzy state-of-charge feasibility constraints.
- **Analytical framework:** Closed-form formulations for overtime credibility and energy-budget credibility under triangular fuzzy parameters, which avoid Monte Carlo sampling and are amenable to gradient-based computations.
- **Algorithm:** PCT-RABLNS (Pareto-Conditioned Transformer with Risk-Aware and Battery-Conscious Large Neighbourhood Search), a hybrid solver that couples a preference-conditioned transformer for constructive guidance with LNS operators that jointly target shift-boundary violations, high overtime-risk segments, and low state-of-charge situations.
- **Empirical findings:** On municipal electric maintenance benchmarks, PCT-RABLNS attains hypervolume gains of approximately 2–5% over evolutionary baselines, reduces maximum per-shift overtime risk by roughly 15–25% (with makespan overheads of about 1–3%), and ensures full battery feasibility without significant range-anxiety penalties.
- **Statistical validation and practice:** Results are supported by bootstrap confidence intervals, nonparametric tests, and Vargha–Delaney effect sizes (\hat{A}_{12}). Actionable recommendations show that fairness-oriented objectives can markedly reduce risk concentration, especially when location clustering or operational stress is present, and that battery-aware neighbourhood search is essential to avoid post hoc infeasibility in EV contexts.

In short, the paper argues that (i) possibilistic uncertainty better reflects the limited but structured knowledge available to many field operators, including uncertain energy consumption in EV fleets; (ii) explicit aggregation of temporal risk is necessary to avoid concentrated overtime exposure; and (iii) hybrid learning-augmented search methods that incorporate battery–feasibility awareness provide an effective computational route to balance efficiency, equity, and range safety in multi-shift electric vehicle routing. The motivation is practical: tighter environmental regulation, growing concern for worker safety, and the operational peculiarities of electric fleets make it important to control not only how long routes take and how overtime risk accumulates across shifts, but also whether the vehicle’s battery can sustain the planned operations throughout each shift [4,5,7].

2. Related Work

This section surveys work from 2014 to 2025 concerning multi-shift routing, electric vehicle routing, uncertainty representation, and fairness-aware optimisation.

2.1. Multi-Shift and Multi-Period Routing

Interest in routing problems that span several regulated shifts has grown noticeably in recent years. Karoonsoontawong et al. [4] examine a multi-shift VRP with time windows and compulsory breaks, emphasising how even modest extensions of classical formulations quickly increase modelling difficulty and challenge exact approaches in terms of scalability. Matl et al. [5] turn their attention to fairness across periods, pointing out that solutions optimised purely for cost or duration can generate significant imbalances in workload distribution. Further contributions, such as Hashemi-Amiri et al. [1], bring shift-level routing decisions into a unified framework, and Ren et al. [2] consider scenarios with dynamically arriving service requests. These studies broaden the understanding of multi-period routing behaviour, yet they generally maintain deterministic assumptions, rely on conventional vehicles, and do not address the interaction between uncertainty, battery feasibility, and fairness-driven risk control—the gap in which the present work is situated.

2.2. Electric Vehicle Routing

The Electric Vehicle Routing Problem (E-VRP) extends classical VRP by adding a state-of-charge (SoC) constraint: the cumulative energy consumed along any route must not exceed the battery capacity, and the vehicle may visit recharging stations to restore SoC [7]. Schneider et al. (2014) establish a foundational formulation of the E-VRP with time windows and recharging stations, showing that battery constraints fundamentally alter the structure of feasible routes even for small instances. Pelletier et al. [6] provide a comprehensive review of EV logistics, highlighting that energy consumption depends non-linearly on load, speed, road gradient, and auxiliary systems, and that these factors are rarely known with precision in practice. Goeke and Schneider [11] model energy consumption explicitly as a function of vehicle speed and payload, a level of detail that, in uncertain operating conditions, calls for robust or fuzzy energy-budget representations rather than deterministic constraints.

Exact methods for the E-VRP have been explored by Desaulniers et al. [29], who develop branch-price-and-cut algorithms for instances with time windows, demonstrating that the energy constraint increases computational difficulty substantially. Schiffer and Walther [8] extend the problem to partial recharging and location-routing decisions (2017), observing that partial SoC replenishment within shifts is often more practical than full recharges.

Importantly, none of these formulations consider multi-shift operations or fairness in overtime risk, and all assume that energy consumption rates are precisely known. The present work bridges these gaps by coupling multi-shift scheduling with battery feasibility and possibilistic energy uncertainty.

2.3. Uncertainty Modelling in Vehicle Routing

Uncertainty in VRP has been addressed mainly through stochastic, robust, or possibilistic frameworks [9,10]. Stochastic approaches require detailed historical data [13]; robust models emphasize protection against worst-case scenarios, sometimes at the expense of overly conservative solutions [14]. Possibilistic and fuzzy approaches [30] attempt to represent uncertainty when only imprecise or qualitative information is available. Examples include Cao et al. [15] for demand uncertainty, Attari et al. [18] for credibility-based routing under fuzzy parameters, and Devnath et al. [19] for multi-objective transportation. Other

contributions, such as [16,17], consider fuzzy constraints in hazardous or time-sensitive routing contexts, and Kim [31] explores adaptive fuzzy models for dynamic routing.

The extension to fuzzy energy consumption in EV routing is less studied. In the absence of reliable speed and load statistics, triangular fuzzy numbers provide a principled way to encode expert estimates of per-arc energy usage, extending possibilistic VRP ideas to the EV domain [6]. Most existing works, however, focus on single-period problems and do not incorporate fairness in the distribution of uncertainty-related risks, which remains an open area.

2.4. Risk Aggregation and Fairness in Optimisation

Ordered Weighted Averaging (OWA) operators offer a versatile way to combine multiple criteria using rank-based weights that capture differing attitudes toward importance and risk [20]. Yager et al. [32] review a wide range of applications and underline how OWA formulations lend themselves naturally to risk-aware decision making. D'Urso et al. [21] apply similar aggregation ideas in consensus problems, illustrating how distinct weight patterns can reflect alternative fairness or caution profiles.

In routing and service operations, concerns about equitable workload distribution have become increasingly visible [5]. Yet the majority of related studies remain within deterministic settings, leaving open the question of how risk itself is spread across tasks or time periods when uncertainty is present. Bertsimas et al. [33] examine equity in service planning but do not model risk explicitly, which marks a point of departure for the present work. In the electric vehicle context, risk concentration at specific shifts is aggravated by battery constraints: a shift that concentrates on overtime risk often also approaches the boundary of SoC feasibility, making risk-fair objectives doubly important for energy-aware planning.

2.5. Machine Learning for Routing Optimisation

Recent progress in deep learning has encouraged a surge of interest in data-driven routing methods. The POMO framework [22], for example, leverages multiple symmetric initializations alongside policy-gradient training, and demonstrates that learned policies can remain competitive even as instance sizes increase. Neural Large Neighbourhood Search shows particular promise: Hottung and Tierney [26,34] achieve state-of-the-art results by integrating learned destroy/repair operators with classical local search. Recent surveys by Shi and Niu [27] and Pan and Liu [24] provide comprehensive overviews of modern heuristic and reinforcement learning approaches for VRP.

Multi-objective reinforcement learning (MORL) represents another important development. Hayes et al. [35] provide comprehensive MORL coverage, including preference-conditioned policies for Pareto front approximation. Recent advances include work by Guan et al. [28] on attention-enhanced architectures for VRP. Integration of uncertainty considerations into learning-based routing optimisation remains limited in the existing literature, with most approaches focusing on historical data without possibilistic or fuzzy uncertainty representations. Battery-constraint awareness in learned routing policies is virtually absent, constituting a further open research direction that the present work addresses.

2.6. Research Gaps and Positioning

Table 1 provides a systematic comparison of existing approaches across the key dimensions relevant to our research.

Table 1. Comparative analysis of routing approaches: multi-shift, EV constraints, uncertainty, and fairness dimensions.

Approach	Multi-Shift	EV/Battery	Uncertainty	Risk Aggregation	Fairness	Learning	Solution Quality	Scalability
Classical VRP [9]	No	No	No	No	No	No	Yes	Yes
Stochastic VRP [13]	No	No	Probabilistic	Exp. value	No	No	Yes	No
Robust VRP [14]	No	No	Worst-case	Min-max	No	No	No	Yes
Fuzzy VRP [15]	No	No	Possibilistic	Simple	No	No	Yes	Yes
E-VRP [7]	No	Yes	No	No	No	No	Yes	No
E-VRP partial SoC [8]	No	Yes	No	No	No	No	Yes	Yes
Multi-shift VRP [4]	Yes	No	No	No	No	No	Yes	No
Fair Routing [5]	Yes	No	No	No	Basic	No	Yes	Yes
POMO [22]	No	No	No	No	No	Yes	Yes	Yes
Neural LNS [26]	No	No	No	No	No	Yes	Yes	Yes
OWA Routing [21]	No	No	No	OWA	Basic	No	Yes	Yes
This study	Yes	Yes	Possibilistic	OWA-based	Risk-fair	Yes	Target	Yes

The analysis reveals that existing approaches address at best a subset of the challenges considered here. To the best of our knowledge, prior studies have not combined multi-shift routing, electric-vehicle battery feasibility, possibilistic uncertainty in both travel times and energy consumption, and fairness-aware aggregation of per-shift overtime risk within a single integrated framework. This gap motivates our approach, which combines multi-shift EV formulations with fuzzy energy constraints, risk-fair OWA aggregation, and a learning-augmented solution method. The resulting contribution is twofold: a unified modelling framework—the MS-SEVRP-PU formulation and its closed-form credibility machinery—and a practical algorithmic scheme, PCT-RABLNS, for this relevant class of problems in sustainable field-service operations.

3. Problem Formulation

This section formalises the MS-SEVRP-PU. We specify the network and shift structure, model operational parameters as triangular fuzzy numbers, derive closed-form credibility expressions for overtime and battery feasibility, and present the bi-objective formulation with fairness-aware Ordered Weighted Averaging aggregation. The corresponding solution methodology—PCT-RABLNS—is described in Section 4, and the experimental protocol is defined in Section 5.

3.1. Network, Shifts and Decision Variables

Consider a complete directed graph $G = (V, E)$, where $V = \{0\} \cup N$ consists of a depot node 0 and a set of tasks $N = \{1, 2, \dots, n\}$ to be served. For ease of reference, the full list of acronyms and mathematical symbols used in this paper is summarised in Appendix C (Tables A7 and A8). A single electric vehicle (EV) operates over P consecutive shifts, each beginning and ending at the depot with a maximum allowable duration L (e.g., $L = 480$ min for an eight-hour shift including breaks). The vehicle carries a battery of capacity B (kWh) that is fully recharged at the depot at the start of each shift.

Operational parameters are represented as triangular fuzzy numbers (TFNs) to capture expert knowledge about variability without requiring extensive historical data or distributional assumptions. Specifically, travel time from location i to location j is modelled as $\tilde{d}_{ij} = (d_{ij}^A, d_{ij}^B, d_{ij}^C)$, where d_{ij}^A is the lower bound, d_{ij}^B the modal value, and d_{ij}^C the upper bound. For travel-time and energy quantities in this study, the lower bound is operationally optimistic and the upper bound operationally pessimistic. Service time at task i is $\tilde{q}_i = (q_i^A, q_i^B, q_i^C)$ with an analogous interpretation.

To model energy uncertainty, the arc energy consumption from i to j is represented as the TFN $\tilde{e}_{ij} = (e_{ij}^A, e_{ij}^B, e_{ij}^C)$, capturing variability due to speed profile, road gradient, payload, and ambient temperature. Additionally, the energy consumed during service at task i (e.g., for on-board equipment or climate control) is $\tilde{f}_i = (f_i^A, f_i^B, f_i^C)$.

The triangular membership function for a TFN $\tilde{t} = (a, b, c)$ is defined, as shown in Figure 1:

$$\mu_{\tilde{t}}(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ \frac{c-x}{c-b}, & b < x \leq c, \\ 0, & x > c. \end{cases} \tag{1}$$

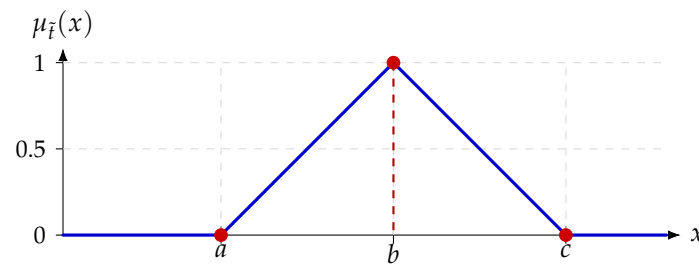


Figure 1. Triangular fuzzy number membership function $\mu_{\tilde{t}}(x)$ for $\tilde{t} = (a, b, c)$.

Optional time windows $[e_i, \ell_i]$ may be imposed on the start of service at each task $i \in N$, evaluated under modal (most likely) parameter values to maintain computational tractability while respecting customer requirements.

3.1.1. Solution Encoding

Although the encoding ω is also the data structure that the PCT-RABLNS solver manipulates (Section 4), it is presented here because the mathematical formulation of constraints (C1)–(C3) and of the fuzzy-duration and fuzzy-energy expressions in Sections 3.1.2 and 3.1.3 requires the underlying sequence representation to be defined first. The same ω is reused without modification by the constructive policy and the LNS operators of Section 4.

A feasible solution is encoded as an ordered sequence ω of length $n + P - 1$ over the extended alphabet $N \cup \{\#_2, \#_3, \dots, \#_P\}$, where each task $i \in N$ appears exactly once and special shift-break tokens $\#_h$ for $h = 2, \dots, P$ delimit consecutive shifts, appearing in strictly increasing order. The sequence implicitly defines P shifts: shift 1 spans from the beginning to $\#_2$, shift 2 from $\#_2$ to $\#_3$, and so forth.

For example, with $n = 6$ tasks and $P = 3$ shifts, a valid sequence is:

$$\omega = (1, 3, \#_2, 2, 5, \#_3, 4, 6),$$

representing shift 1 serving tasks $\{1, 3\}$, shift 2 serving $\{2, 5\}$, and shift 3 serving $\{4, 6\}$. This structure is illustrated in Figure 2.

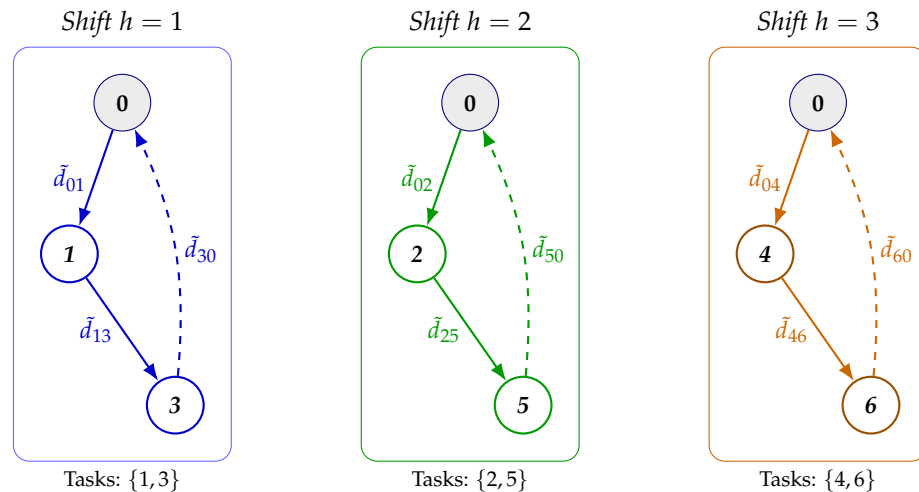


Figure 2. Example solution for an EV with 6 tasks over 3 shifts. The depot (node 0) is the start and end of each shift, with recharging in between. Solid arrows represent travel arcs with fuzzy times \tilde{d}_{ij} , while dashed arrows show return legs.

3.1.2. Fuzzy Duration and Makespan

For shift h , let $\mathcal{S}_h(\omega)$ denote the subsequence of tasks assigned to that shift. The fuzzy duration $\tilde{\sigma}_h$ is the fuzzy arithmetic sum of the depot-to-first-task travel time, all within-shift service and travel times, and the last-task-to-depot return travel time. Using standard TFN arithmetic, if $\tilde{a} = (a^A, a^B, a^C)$ and $\tilde{b} = (b^A, b^B, b^C)$:

$$\tilde{a} \oplus \tilde{b} = (a^A + b^A, a^B + b^B, a^C + b^C),$$

so $\tilde{\sigma}_h = (\sigma_h^A, \sigma_h^B, \sigma_h^C)$, where each component is the sum of the corresponding components of the constituent TFNs.

The total makespan is computed under the following modal durations:

$$\lambda(\omega) = \sum_{h=1}^P \sigma_h^B. \tag{2}$$

3.1.3. Fuzzy Per-Shift Energy Consumption and SoC Feasibility

The fuzzy energy consumed by the EV during shift h is:

$$\tilde{E}_h(\omega) = \left[\bigoplus_{(i,j) \in \mathcal{A}_h(\omega)} \tilde{e}_{ij} \right] \oplus \left[\bigoplus_{i \in \mathcal{S}_h(\omega)} \tilde{f}_i \right], \tag{3}$$

where $\mathcal{A}_h(\omega)$ is the set of arcs traversed in shift h (including depot ingress and egress). With TFN arithmetic, the result is $\tilde{E}_h = (E_h^A, E_h^B, E_h^C)$.

Battery feasibility requires the energy consumed per shift to remain within capacity B . We enforce this via a credibility threshold: shift h is declared *SoC-feasible* if

$$\text{Cr}(\tilde{E}_h \leq B) \geq \beta_E, \tag{4}$$

where $\beta_E \in (0.5, 1]$ is a user-specified safety level (default $\beta_E = 0.9$) and the credibility measure for a TFN is given by a piecewise closed-form expression (see Appendix A.1). The modal energy E_h^B is used for a hard feasibility pre-check; the credibility criterion (4) provides a soft but probabilistically grounded tightening under uncertainty.

3.1.4. Per-Shift Overtime Credibility

We measure overtime risk for shift h by computing how likely its fuzzy duration $\tilde{\sigma}_h$ exceeds the allowable length L . On the basis of credibility theory [36]:

$$r_h = \text{Cr}(\tilde{\sigma}_h > L) = 1 - \text{Cr}(\tilde{\sigma}_h \leq L). \tag{5}$$

For triangular fuzzy numbers, this admits a closed-form piecewise-smooth expression (see Appendix A.1), enabling efficient gradient-compatible evaluation without Monte Carlo sampling.

3.1.5. Ordered Weighted Averaging for Risk Fairness

Let $\mathbf{r}(\omega) = (r_1, r_2, \dots, r_P)$ denote the vector of per-shift overtime credibilities. To promote fairness, we employ Ordered Weighted Averaging (OWA) [20], which aggregates the sorted risks $r_{(1)} \geq r_{(2)} \geq \dots \geq r_{(P)}$ via the following weight vector \mathbf{w} :

$$R_{\text{OWA}}(\mathbf{r}; \mathbf{w}) = \sum_{h=1}^P w_h \cdot r_{(h)}, \quad \sum_{h=1}^P w_h = 1, \quad w_h \geq 0. \tag{6}$$

We adopt front-loaded weights $w_h = \frac{2(P-h+1)}{P(P+1)}$, which penalise high maximum risks and promote equity across shifts (see Appendix A.2 for alternative schemes and their properties).

3.1.6. Bi-Objective MS-SEVRP-PU Formulation

In what follows, Ω denotes the set of feasible sequences ω over the alphabet $N \cup \{\#_2, \dots, \#_P\}$ that satisfy the structural conditions (C1)–(C3) below, and $t_i^B(\omega)$ denotes the modal start of service at task i implied by ω . The complete MS-SEVRP-PU is the bi-objective combinatorial programme:

$$\begin{aligned} \min_{\omega \in \Omega} \quad & (\lambda(\omega), R_{\text{OWA}}(\mathbf{r}(\omega); \mathbf{w})) \\ \text{s.t.} \quad & \text{(C1)} \quad \sum_{t=1}^{n+P-1} \mathbb{I}[\omega_t = i] = 1, \quad \forall i \in N, \\ & \text{(C2)} \quad \sum_{t=1}^{n+P-1} \mathbb{I}[\omega_t = \#_h] = 1, \quad \forall h \in \{2, \dots, P\}, \\ & \text{(C3)} \quad \text{position}(\#_h) < \text{position}(\#_{h+1}), \quad \forall h \in \{2, \dots, P-1\}, \\ & \text{(C4)} \quad \text{Cr}(\tilde{E}_h \leq B) \geq \beta_E, \quad \forall h \in \{1, \dots, P\} \quad (\text{battery feasibility}), \\ & \text{(C5)} \quad e_i \leq t_i^B(\omega) \leq \ell_i, \quad \forall i \in N \quad (\text{with time windows}) \end{aligned} \tag{7}$$

Constraints (C1)–(C3) (with $\mathbb{I}[\cdot]$ the indicator function) enforce that each task is visited exactly once and that the $P - 1$ shift-break tokens appear in increasing order, defining a valid partition of N into P ordered shifts $\mathcal{S}_1, \dots, \mathcal{S}_P$ each starting and ending at the depot. Constraint (C4) imposes per-shift battery feasibility at the credibility level β_E , and constraint (C5) enforces (modal) time windows when present. Optional precedence or compatibility constraints can be encoded analogously as additional structural conditions on ω . This formulation targets Pareto-optimal solutions that reduce total operational time while distributing overtime risk equitably across shifts and guaranteeing battery feasibility with high credibility.

4. Solution Methodology: PCT-RABLNS

We propose **PCT-RABLNS** (Pareto-Conditioned Transformer with Risk-Aware and Battery-conscious Large Neighbourhood Search), a hybrid solver that combines deep reinforcement learning with classical optimisation. The approach has three main components: a neural constructive policy that generates quality initial solutions across preference regions, a risk-aware and battery-conscious local search for refinement, and a multi-objective training framework that learns effective construction strategies.

4.1. Algorithm Overview

Figure 3 provides a high-level schematic of the three components and their data flow. At inference time, the Pareto-Conditioned Transformer (Section 4.2) is invoked across a discretised set of preference values $z = (\alpha, 1 - \alpha)$ to populate an initial ϵ -archive \mathcal{A} of SoC-feasible solutions. The Risk-Aware and Battery-Conscious Large Neighbourhood Search (Section 4.3) then iteratively destroys and repairs solutions from \mathcal{A} , accepting them through a simulated-annealing rule and updating \mathcal{A} with the non-dominated outputs. The multi-objective reinforcement learning training pipeline (Section 4.4) is executed offline, once, and produces the trained policy π_θ that is reused by both the constructive phase and the LNS repair step. The three components are tightly coupled: the same preference vector z drives both construction and repair, ensuring consistent trade-off targeting throughout the search; the SoC-aware feasibility mask is shared between the constructive policy and the repair operator, propagating energy awareness into every autoregressive decoding step.

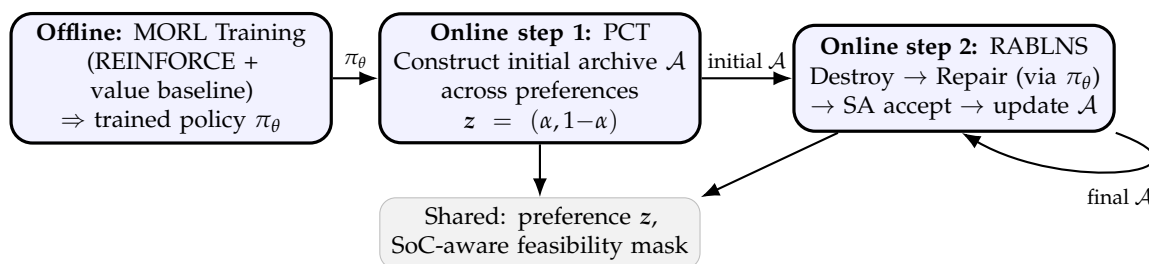


Figure 3. High-level overview of PCT-RABLNS. The MORL training (left) is performed offline and produces the trained policy π_θ . At inference time, the Pareto-Conditioned Transformer constructs an initial ϵ -archive \mathcal{A} (centre), which is iteratively refined by the Risk-Aware and Battery-Conscious LNS (right). The preference vector z and the SoC-aware feasibility mask are shared across the constructive and repair phases.

4.2. Preference-Conditioned Transformer Architecture

The Pareto-Conditioned Transformer (PCT) implements an attention-based sequence-to-sequence architecture π_θ that constructs solutions ω autoregressively. At each step t , the policy selects an action a_t from the set of feasible actions (unvisited tasks or the next shift-break token) given the current partial solution state s_t and a preference vector $z = (\alpha, 1 - \alpha)$.

- **Encoder.** The encoder processes *static* node features (coordinates, service time TFN parameters, time-window bounds, modal arc-energy cost) together with the current decision context through $L = 3$ multi-head self-attention layers, producing contextual node embeddings h_i for each task i .
- **Decoder.** The decoder generates a representation of the construction state via cross-attention between the partial-solution context and the node embeddings, additionally conditioned on the preference vector z . The preference embedding $z_{\text{emb}} = W_z z + b_z$ is concatenated with the decoder state at each step, steering the trade-off between objectives.

- Policy head and feasibility masking.** A compatibility function maps the context vector and candidate embeddings to action probabilities. A feasibility mask $m_t \in \{0,1\}^{|N|+P-1}$ ensures: (i) each task is selected at most once; (ii) shift-break tokens respect temporal ordering; (iii) time-window constraints hold under modal parameters; and (iv) the residual energy budget for the current shift remains non-negative under modal energy consumption (SoC hard-feasibility). The masked action probabilities are as follows:

$$\pi_{\theta}(a_t | s_t, z) = \frac{\exp(\text{score}(a_t, s_t, z)) \cdot m_t(a_t)}{\sum_{a' \in A_t} \exp(\text{score}(a', s_t, z)) \cdot m_t(a')} \tag{8}$$

The scalarised objective used during preference-conditioned construction is as follows:

$$J_{\alpha}(\omega) = \alpha \cdot \hat{\lambda}(\omega) + (1 - \alpha) \cdot \hat{R}_{\text{OWA}}(\omega) + \gamma \cdot V_{\text{SoC}}(\omega), \tag{9}$$

where $\hat{\cdot}$ denotes batch normalisation, $\gamma > 0$ is a penalty coefficient, and $V_{\text{SoC}}(\omega) = \sum_{h=1}^P \max(0, E_h^B - B)$ is the total modal energy over consumption.

Following recent advances in attention-based routing [22,28], the transformer backbone uses $L = 3$ attention layers, eight attention heads, embedding dimension $d_{\text{model}} = 128$, feed-forward dimension 512, dropout 0.1, and learned positional embeddings for sequence positions and shift indices.

Figure 4 summarises the PCT architecture.

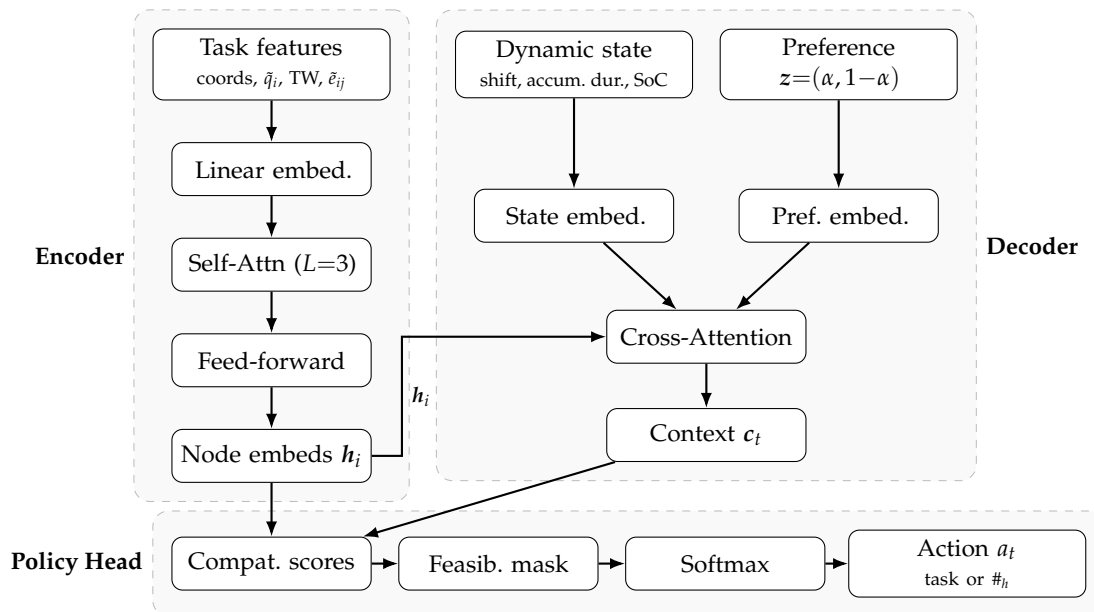


Figure 4. PCT architecture. Task features (with energy TFNs) feed the Encoder. The Decoder processes dynamic states (SoC) and preference z via cross-attention over node embeddings h_i . The Policy Head applies a feasibility mask before softmax-based action probabilities.

4.3. Risk-Aware and Battery-Conscious Large Neighbourhood Search

PCT-generated solutions are refined through a specialised Large Neighbourhood Search (LNS) that explicitly considers both objectives and battery feasibility.

4.3.1. Destroy Operators

Four complementary operators target different aspects of solution structure. (1) *Boundary-aware destroy*: Removes the last 2–3 tasks of a shift and the first 2–3 tasks of the next, targeting the positions that most influence shift duration. Removal size: 10–15% of tasks. (2) *Risk-proportional destroy*: Identifies the shift with the highest overtime credibility r_h and removes

tasks contributing most to that shift's fuzzy duration (measured by modal service time plus incident travel times). Removal size: 15–25% of tasks. (3) *SoC-critical destroy*: Identifies shifts where $\text{Cr}(\tilde{E}_h \leq B) < \beta_E$ and removes the tasks with the highest modal energy contribution $e_{ij}^B + f_j^B$, enabling the repair phase to redistribute them into shifts with higher residual energy budget. Removal size: 10–20% of tasks. (4) *Cluster-based destroy*: Spatially clusters tasks using k -means ($k \in [\lceil P/2 \rceil, P]$) and removes entire clusters, facilitating large-scale load rebalancing across shifts.

The four percentage ranges (10–15%, 15–25%, 10–20%, and the implicit cluster-driven removal of operator (4)) are not free parameters but the result of a calibration on the training distribution. We followed the standard ALNS practice of keeping each removal size in the 10–40% interval suggested in the metaheuristics literature [10,26], then narrowed the search via a coarse grid on $n \in \{30, 50, 80\}$, $P \in \{3, 4, 5\}$: removal sizes below 10% produced only marginal improvements per iteration, while sizes above 25–30% caused frequent regressions of the hypervolume of \mathcal{A} . The asymmetry between the four operators reflects their semantics: the boundary-aware operator perturbs only a small neighbourhood around shift transitions and therefore uses the smallest range; the risk-proportional operator must remove enough tasks to meaningfully change the highest-risk shift and therefore uses the widest range; the SoC-critical operator is intermediate because it must free sufficient energy budget on the offending shift without destabilising the rest of the schedule.

Operator selection employs adaptive weights updated via a sliding-window reward mechanism: operators producing improvements receive increased selection probability in subsequent iterations.

4.3.2. Repair

Destroyed solutions are repaired using the same PCT policy π_θ that generated the initial solution, conditioned on the same preference vector \mathbf{z} . The feasibility mask during repair enforces the SoC hard-feasibility constraint, preventing task insertions that would violate the battery budget at the modal energy level.

4.3.3. Acceptance and Archive

New solutions are evaluated using a simulated annealing (SA) acceptance criterion combined with ε -nondominated archive management. The initial temperature T_0 is set so that solutions with 5% objective degradation are accepted with 50% probability, with geometric cooling $T_{k+1} = 0.98 \cdot T_k$. The acceptance probability for a new solution ω' versus the current ω is as follows:

$$P_{\text{accept}} = \begin{cases} 1, & \text{if } \omega' \text{ weakly dominates } \omega, \\ \exp(-\Delta J_\alpha / T_k), & \text{otherwise,} \end{cases} \quad (10)$$

where $\Delta J_\alpha = J_\alpha(\omega') - J_\alpha(\omega)$ under the current preference α . A solution is added to the ε -archive only if it is SoC-feasible (i.e., $\text{Cr}(\tilde{E}_h \leq B) \geq \beta_E$ for all h) and is not ε -dominated by any archived solution. Archive pruning uses adaptive grid-based crowding distance.

The time budget is allocated as 20% for initial PCT construction across preference settings and 80% for LNS iterations. Figure 5 illustrates the RABLNS destroy–repair cycle.

4.4. Multi-Objective Reinforcement Learning Training

The PCT policy is trained using a multi-objective reinforcement learning (MORL) approach based on preference sampling. At the start of each episode, $\alpha \sim \text{Unif}(0, 1)$ is sampled to define $\mathbf{z} = (\alpha, 1 - \alpha)$, ensuring comprehensive Pareto-front coverage. A solution ω is generated autoregressively, then the scalarised objective (9) is evaluated; the episode

reward is $R = -J_\alpha(\omega)$. A learned value function $V_\phi(s_0, z)$ provides a baseline to reduce gradient variance. Figure 6 shows the MORL training pipeline.

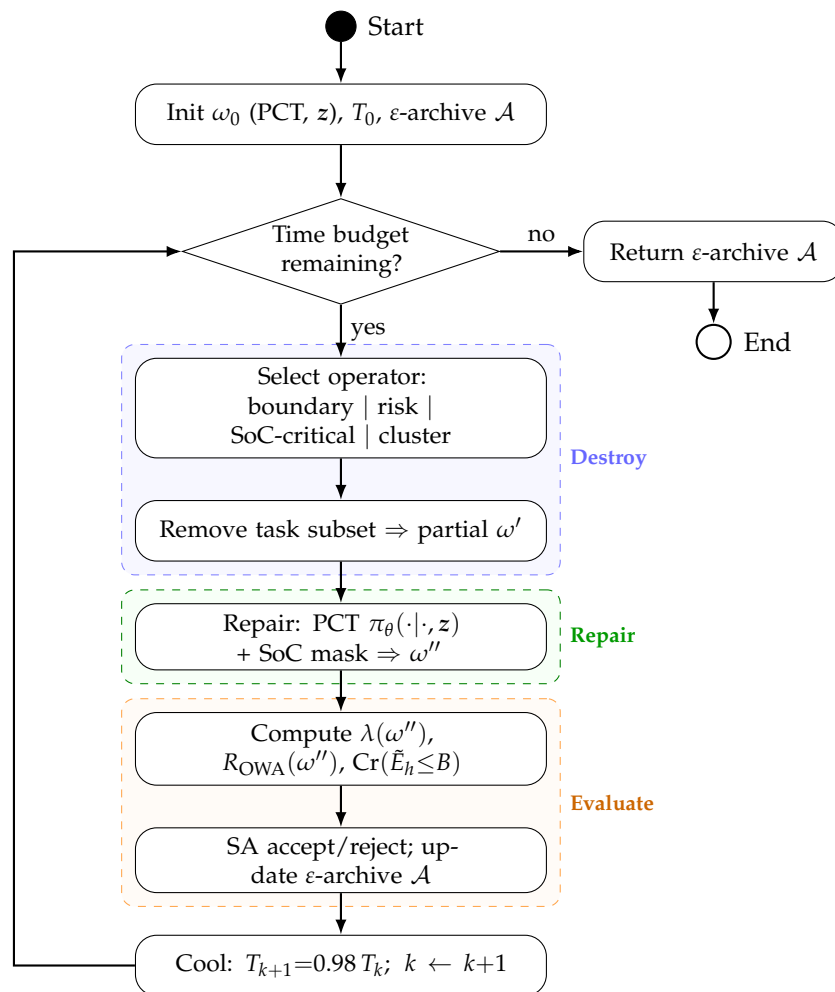


Figure 5. RABLNS destroy–repair cycle. Starting from a PCT solution, the algorithm iterates until the time budget expires. The *Destroy* phase selects among four operators, while *Repair* reinserts tasks using the PCT policy with SoC masking. Finally, *Evaluate* computes makespan, OWA risk, and feasibility, applies SA acceptance, and updates the ϵ -archive.

We employ the REINFORCE algorithm with baseline subtraction [22]. The policy gradient estimator is as follows:

$$\nabla_\theta \mathcal{L}(\theta) = \mathbb{E}_{z, \omega \sim \pi_\theta} \left[\sum_{t=0}^{T-1} \nabla_\theta \log \pi_\theta(a_t | s_t, z) \cdot (R - V_\phi(s_0, z)) \right], \quad (11)$$

where $T = n + P - 1$ is the solution length. The value function is trained by minimising

$$\mathcal{L}(\phi) = \mathbb{E}_{z, \omega} \left[(V_\phi(s_0, z) - R)^2 \right]. \quad (12)$$

Optimisation uses AdamW with learning rate 3×10^{-4} , weight decay 10^{-4} , batch size 512, cosine annealing with 1000-iteration warm-up over 50 k total training iterations, gradient clipping at norm 1.0, and entropy regularisation coefficient $\beta = 0.01$. Training instances are generated synthetically with the following ranges: $n \in \{20, 30, 40, 50, 60, 70, 80\}$, $P \in \{2, 3, 4, 5, 6\}$, TFN spread $\eta_d, \eta_q \in [\pm 20\%, \pm 35\%]$ for travel and $[\pm 15\%, \pm 30\%]$ for service times, energy TFN spread $\eta_e \in [\pm 10\%, \pm 25\%]$, and both uniform-random and clustered spatial configurations (Gaussian mixtures with 2–5 centres). Battery capacity B is calibrated to

the expected modal total energy per instance, with a 25% surplus to reflect realistic EV range buffers. Training requires approximately 60 GPU-hours on NVIDIA A100 hardware.

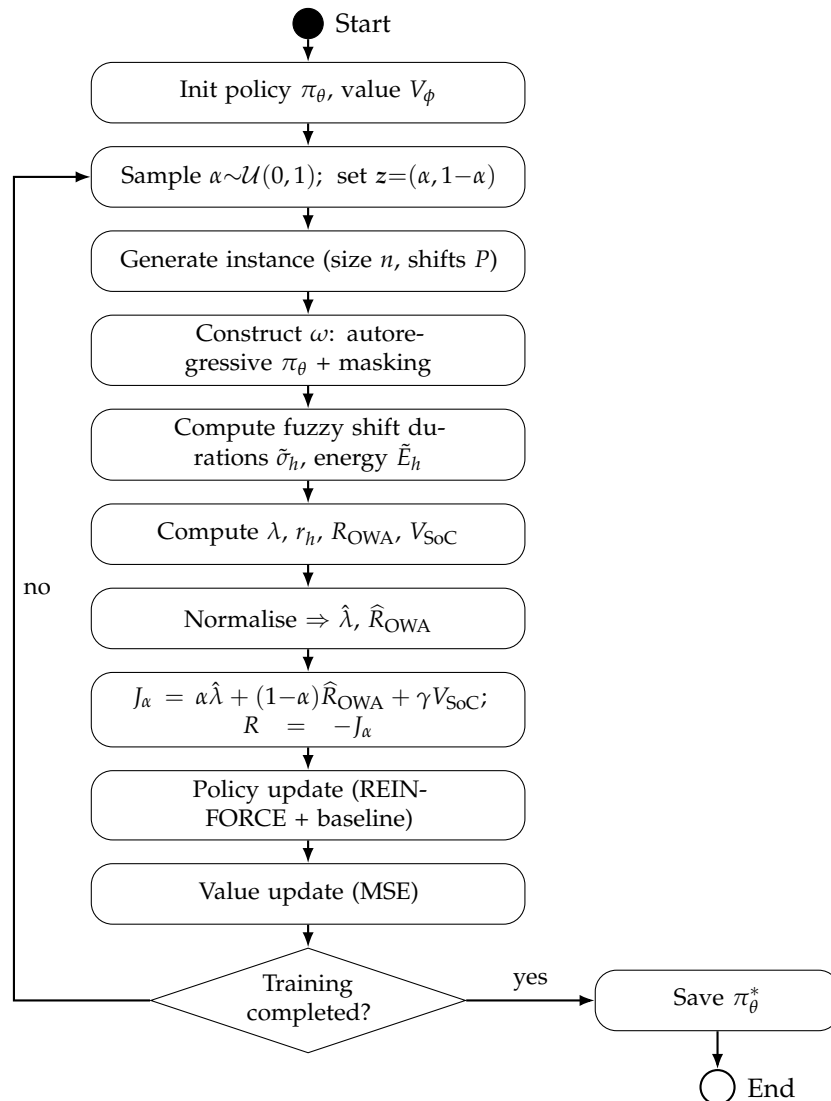


Figure 6. MORL training: Each episode samples preference α to generate an instance and an autoregressive solution via feasibility masking. Fuzzy objectives (V_{SoC}) form the scalarised reward $R = -J_\alpha$. Policy (REINFORCE) and value function (MSE) are updated. Uniform α sampling allows one policy to cover the Pareto spectrum.

5. Results

This section reports a comprehensive empirical evaluation of PCT-RABLNS against two strong baselines—IBEA (evolutionary) and Neural-LNS (pure learning)—across four municipal electric maintenance case studies. We assess Pareto front quality (hypervolume, IGD^+), risk fairness (Gini coefficient, maximum shift risk), battery feasibility (SoC compliance rate), and convergence efficiency (time-to-90% HV). All methods operate under identical computational budgets. Section 5.1 describes experimental protocols.

5.1. Experimental Design and Case Studies

We construct four case studies representative of municipal electric maintenance operations, varying instance size (n), shift count (P), spatial layout, uncertainty characteristics (TFN spread, skewness), and battery capacity B (calibrated as a multiple of the expected modal per-shift energy consumption). Representativeness was assessed in two ways. First, the four cases jointly cover the operating regimes that a municipal maintenance

unit typically faces over a planning horizon: a small, spatially uniform deployment with comfortable battery margin (C1); a medium, spatially clustered deployment characteristic of district-based maintenance routes (C2); a large, clustered deployment with a tight battery-to-demand ratio that approximates peak-season operations (C3); and a stress-test instance with skewed (asymmetric) uncertainty and a tight shift limit (C4). Second, the parameter ranges and TFN spreads were elicited from three operations engineers of an industrial partner operating an electrified municipal maintenance fleet who supplied pessimistic, most likely, and optimistic bounds for travel times, service durations, and per-arc energy consumption, as well as the typical battery-to-demand ratios summarised in Table 2. The same experts were consulted to validate the calibrated TFN parameters reported in `uncertainty_parameters.json` [37], and to confirm that the four cases span the spectrum of routes that the unit plans on a typical week. Table 2 summarises their key properties.

Table 2. Case study characteristics and computational budgets. Battery capacity B is expressed in kWh; TFN spread is given as $(\pm\eta_d, \pm\eta_q, \pm\eta_e)$ for travel, service, and energy, respectively.

Case	n	P	Layout	TFN Spread	B (kWh)	L (min)	Budget
C1	30	3	Uniform	Med. ($\pm 25\%$, $\pm 20\%$, $\pm 18\%$)	48	120	300 s
C2	50	4	Clust. (4)	Med. ($\pm 25\%$, $\pm 20\%$, $\pm 18\%$)	72	150	600 s
C3	80	5	Clust. (5)	High ($\pm 30\%$, $\pm 25\%$, $\pm 22\%$)	100	180	1200 s
C4	50	4	Uniform	Skew ($\pm 35\%$, $\pm 15\%$, $\pm 20\%$)	72	130	600 s

- Case C1: Small Uniform Baseline.** Thirty tasks uniformly distributed in $[0, 100]^2$ with moderate uncertainty and a 48 kWh battery (25% surplus over expected modal shift energy). Serves as a baseline for assessing convergence on simpler instances where all methods should achieve high-quality fronts, and where SoC constraints are rarely binding. The TFN spreads ($\pm 25\%$ travel, $\pm 20\%$ service, $\pm 18\%$ energy), the shift length $L = 120$ min and the battery capacity $B = 48$ kWh represent the optimistic, most likely and pessimistic operational bounds that the partner operations engineers agreed correspond to a typical low-density municipal maintenance day.
- Case C2: Medium Clustered.** Fifty tasks in four spatial clusters (NE, NW, SE, SW quadrants), 72 kWh battery. Tests performance when geographic structure creates natural shift partitioning opportunities that both cluster-based destroy operators and SoC-critical destroy operators can exploit.
- Case C3: Large Clustered.** Eighty tasks in five clusters with high uncertainty ($\pm 30\%$ travel, $\pm 25\%$ service, $\pm 22\%$ energy). The 100 kWh battery is tighter relative to total demand, making SoC feasibility a meaningful constraint. Represents realistic large-scale operations where spatial, temporal, and energetic complexity interact simultaneously.
- Case C4: Fairness-Stress Uniform.** Fifty uniformly distributed tasks with *skewed* uncertainty: pessimistic bounds $+35\%$ above the modal but optimistic bounds only -15% below; energy spread $\pm 20\%$. Tight shift limit $L = 130$ creates high overtime pressure. Designed to stress-test fairness mechanisms under asymmetric risk profiles and to verify that the SoC-critical destroy operator does not compromise risk equity.

All cases use Euclidean distance with triangular fuzzy travel times $\tilde{d}_{ij} = (0.9d_{ij}, d_{ij}, 1.2d_{ij})$ and service times $\tilde{q}_i = (0.8q_i, q_i, 1.15q_i)$ where $q_i \sim \mathcal{U}[5, 15]$, except C4 (asymmetric spreads). Arc energy TFNs follow $\tilde{e}_{ij} = (0.9e_{ij}^B, e_{ij}^B, 1.18e_{ij}^B)$ with modal arc energy e_{ij}^B proportional to d_{ij} scaled to match realistic EV consumption of 0.2 kWh/km. Depot returns required at shift ends; battery recharged to B at start of each shift. Front-loaded OWA weights $w_h = \frac{2(P-h+1)}{P(P+1)}$ across all experiments; SoC credibility threshold $\beta_E = 0.9$.

Instance configurations, algorithm parameters, solver setup, and uncertainty specifications are provided in the public repository [37] for experiment replication. Specifically,

- `instance_configurations.json` defines the four case studies (C1–C4) with task counts, shift parameters, TFN spreads, and battery capacities;
- `algorithm_parameters.json` specifies PCT-RABLNS, IBEA, and Neural-LNS configurations (architecture, training hyperparameters, and computational budgets);
- `solver_parameters.json` documents the MORL training setup, LNS operator specifications, and OWA weight schemes;
- `uncertainty_parameters.json` provides TFN factor specifications for travel times, service durations, and arc energy consumption for each case study.

Result files include per-seed values, medians, and bootstrap confidence intervals supporting Tables 3–6 and A1–A6.

Table 3. Pareto quality and SoC compliance summary: median [95% CI] across 10 seeds. Bold indicates best per case; * denotes statistically significant difference from PCT-RABLNS (Mann–Whitney, Holm–Bonferroni, $p < 0.05$).

Case	Method	HV (↑)		IGD ⁺ (↓)		SoC Rate (↑)
		Median	95% CI	Median	95% CI	
C1	PCT-RABLNS	0.821	[0.817, 0.824]	0.106	[0.101, 0.111]	1.00
	IBEA	0.806 *	[0.803, 0.809]	0.115 *	[0.108, 0.121]	0.96
	Neural-LNS	0.813 *	[0.810, 0.816]	0.110 *	[0.104, 0.115]	0.98
C2	PCT-RABLNS	0.879	[0.873, 0.883]	0.097	[0.086, 0.100]	1.00
	IBEA	0.838 *	[0.834, 0.842]	0.121 *	[0.114, 0.126]	0.91 *
	Neural-LNS	0.856 *	[0.851, 0.860]	0.109 *	[0.103, 0.114]	0.95 *
C3	PCT-RABLNS	0.897	[0.894, 0.901]	0.101	[0.096, 0.105]	1.00
	IBEA	0.861 *	[0.852, 0.863]	0.130 *	[0.126, 0.134]	0.83 *
	Neural-LNS	0.874 *	[0.868, 0.878]	0.117 *	[0.111, 0.121]	0.89 *
C4	PCT-RABLNS	0.844	[0.841, 0.848]	0.102	[0.097, 0.107]	1.00
	IBEA	0.820 *	[0.812, 0.825]	0.121 *	[0.118, 0.130]	0.88 *
	Neural-LNS	0.831 *	[0.825, 0.836]	0.113 *	[0.107, 0.118]	0.92 *

Table 4. Fairness metrics at $\alpha = 0.3$: median [95% CI]. Bold indicates best; * denotes significant difference from PCT-RABLNS (Mann–Whitney, Holm–Bonferroni, $p < 0.05$). Makespan overhead: relative difference of baselines versus PCT-RABLNS (negative = faster but less fair and/or less SoC-safe).

Case	Method	Max Shift Risk (↓)		Gini (↓)		Mksp. ovhd.
		Median	95% CI	Median	95% CI	
C1	PCT-RABLNS	0.38	[0.36, 0.40]	0.22	[0.20, 0.24]	–
	IBEA	0.51 *	[0.49, 0.54]	0.31 *	[0.29, 0.33]	–1.2%
	Neural-LNS	0.45 *	[0.43, 0.47]	0.27 *	[0.25, 0.29]	–0.8%
C2	PCT-RABLNS	0.42	[0.39, 0.44]	0.24	[0.22, 0.26]	–
	IBEA	0.56 *	[0.53, 0.59]	0.34 *	[0.32, 0.36]	–2.1%
	Neural-LNS	0.49 *	[0.46, 0.51]	0.29 *	[0.27, 0.31]	–1.5%
C3	PCT-RABLNS	0.45	[0.43, 0.47]	0.26	[0.24, 0.28]	–
	IBEA	0.60 *	[0.57, 0.63]	0.37 *	[0.35, 0.39]	–2.8%
	Neural-LNS	0.53 *	[0.50, 0.55]	0.31 *	[0.29, 0.33]	–1.9%
C4	PCT-RABLNS	0.40	[0.38, 0.42]	0.23	[0.21, 0.25]	–
	IBEA	0.53 *	[0.50, 0.56]	0.33 *	[0.31, 0.35]	–1.7%
	Neural-LNS	0.47 *	[0.44, 0.49]	0.28 *	[0.26, 0.30]	–1.2%

Table 5. Time to 90% of final hypervolume (seconds): median [95% CI]. Bold indicates fastest; * denotes significant difference from PCT-RABLNS (Mann–Whitney, Holm-adj. $p < 0.05$).

Case	PCT-RABLNS	IBEA	Neural-LNS
C1	145 [138, 152]	172 * [163, 180]	158 * [151, 166]
C2	303 [289, 317]	421 * [403, 439]	372 * [355, 389]
C3	624 [600, 648]	867 * [836, 898]	762 * [733, 791]
C4	318 [300, 336]	439 * [416, 462]	389 * [369, 409]

Table 6. Ablation study (case C2): Performance change relative to full PCT-RABLNS when disabling individual components.

Variant	Δ HV (%)	Δ Max Risk (%)	Δ SoC Rate	Δ Makespan (%)
Full PCT-RABLNS	0.0	0.0	0.00	0.0
No preference cond.	−3.8	+8.3	−0.02	+1.2
No risk-aware destroys	−2.7	+12.5	−0.03	+0.8
No SoC-critical destroy	−1.8	+3.2	−0.14	+0.3
No adaptive selection	−1.5	+5.1	−0.04	+0.5
No OWA objective	−5.1	+42.1	−0.01	+1.9

5.2. Methods Under Comparison

- **PCT-RABLNS (Proposed).** Pareto-conditioned transformer with risk-aware and battery-conscious large neighbourhood search. Transformer trained on 5000 instances spanning $n \in [20, 80]$, $P \in [2, 6]$, mixed layouts, and uncertainty/energy levels. LNS uses 10 preference samples per iteration and four destroy operators (boundary-aware, risk-proportional, SoC-critical, cluster-based) with adaptive selection. Archive maintains ϵ -nondominated SoC-feasible solutions ($\epsilon = 0.01$).
- **IBEA (Evolutionary Baseline).** Indicator-Based Evolutionary Algorithm [38] with hypervolume indicator, population 100, binary tournament selection, PMX crossover (prob. 0.9), swap mutation (prob. 0.1). Shift-break positions mutated separately. Battery feasibility enforced as a hard penalty in the fitness function.
- **Neural-LNS (Learning Baseline).** Attention model [23] trained via REINFORCE with fixed scalarisation weights sampled from the simplex. Uses the same LNS operators as PCT-RABLNS but without preference conditioning or SoC-critical destroy. Isolates the value of preference-conditioned learning versus fixed-weight policies.

All methods receive equal wall-clock budgets (Table 2). Each method is run 10 times per case with different random seeds (seeds: 42, 123, 456, 789, 1011, 1213, 1415, 1617, 1819, 2021; full seed list in `algorithm_parameters.json` [37]). Algorithms implemented in Python 3.9/PyTorch 1.12, executed on Intel Xeon Gold 6248R (3.0 GHz, 24 cores) with NVIDIA A100 GPU (40 GB).

5.3. Evaluation Metrics and Statistical Analysis

- **Pareto Quality Metrics.**
 - **Hypervolume (HV):** Volume of objective space dominated by the Pareto front relative to a reference point; normalised to $[0, 1]$. Higher is better; assesses both convergence and diversity.
 - **IGD⁺:** Average distance from the reference Pareto front to algorithm output; lower is better. Weakly Pareto-compliant variant used for theoretical soundness.

The reference front is constructed by merging all runs across all methods and retaining non-dominated solutions per case [39]. In practice, for each case study and for each of the ten random seeds, the final ϵ -archives produced by PCT-RABLNS, IBEA and Neural-LNS are pooled, dominated solutions are removed, and the result-

ing set of bi-objective points (λ, R_{OWA}) constitutes the case-specific reference front used for both HV (with the common reference point set per case as documented in `algorithm_parameters.json` [37]) and IGD^+ . This procedure follows the standard practice of approximating an unknown true Pareto front by the best-known empirical front when a closed-form optimum is not available.

- **Battery Feasibility.**

- **SoC Compliance Rate:** Fraction of archived solutions for which $\text{Cr}(\tilde{E}_h \leq B) \geq \beta_E$ holds for all shifts h . A value of 1.0 means all archived solutions are battery-feasible under the credibility threshold; values below 1.0 indicate feasibility violations.

- **Fairness Metrics.**

- **Gini Coefficient:** Inequality of overtime risk distribution across shifts; $G \in [0, 1]$, lower is more equitable.
- **Maximum Shift Risk:** $\max_h r_h$, worst-case overtime exposure.

Both were evaluated at $\alpha = 0.3$ (fairness-oriented preference) from each method's final archive.

- **Statistical Analysis.**

- For each metric and case, we report medians across 10 runs with 95% bootstrap confidence intervals (10,000 resamples). Pairwise comparisons use Mann–Whitney U tests with Holm–Bonferroni correction; effect sizes via Vargha–Delaney \hat{A}_{12} [40]. All tests two-tailed, $\alpha_{\text{family}} = 0.05$.

5.4. Pareto Front Quality

Figure 7 displays representative Pareto fronts (makespan vs. OWA risk) for all four cases. PCT-RABLNS (blue) consistently achieves superior spread and coverage compared to IBEA (orange) and Neural-LNS (green), particularly in fairness-oriented regions ($\alpha < 0.5$) where evolutionary methods exhibit solution gaps owing to difficulty escaping local optima in high-dimensional discrete spaces with coupled battery constraints.

Table 3 summarises HV and IGD^+ medians with confidence intervals. PCT-RABLNS achieves statistically significant improvements on C2–C4 and smaller but still statistically significant differences on C1, where the absolute gaps remain modest.

5.4.1. SoC Compliance

PCT-RABLNS achieves a SoC compliance rate of 1.00 across all cases: every archived solution satisfies $\text{Cr}(\tilde{E}_h \leq B) \geq 0.9$ for all shifts. This is a direct consequence of the SoC-aware masking in the constructive policy and the SoC-critical destroy operator. IBEA and Neural-LNS show compliance rates as low as 0.83 and 0.89, respectively, on the energy-intensive cases C3 and C4, indicating that without dedicated feasibility handling, learned and evolutionary methods frequently produce solutions that would drain the battery in practice.

5.4.2. Case-by-Case Interpretation

On C1 (small uniform, moderate SoC slack), all methods achieve high-quality fronts with $HV > 0.80$. PCT-RABLNS remains ahead, but the absolute advantage is modest (1.9% HV gain vs. IBEA; $\hat{A}_{12} = 0.92$ for HV and $\hat{A}_{12} = 0.08$ for IGD^+), consistent with a simple instance where battery constraints rarely bind.

On C2 (medium clustered), spatial clustering lets cluster-based and SoC-critical destroy operators shine. PCT-RABLNS achieves a 4.9% HV gain over IBEA ($\hat{A}_{12} = 1.00$ for HV) and

a 20% IGD⁺ improvement ($\hat{A}_{12} = 0.00$ for IGD⁺). Neural-LNS falls between these extremes, indicating that preference conditioning adds substantial value over fixed-weight policies.

On C3 (large clustered, tight energy), advantages scale up. PCT-RABLNS improves HV by 4.2% and reduces IGD⁺ by 22% compared to IBEA ($\hat{A}_{12} = 1.00$ for HV and $\hat{A}_{12} = 0.00$ for IGD⁺). The lower SoC compliance of IBEA (0.83) further degrades the practical usability of its archive even when nominal Pareto indicators are comparable.

On C4 (fairness stress, skewed uncertainty), skewed TFN shapes and tight shift limits create concentrated overtime pressure. PCT-RABLNS maintains front density (HV +2.9% vs. IBEA) while achieving superior fairness metrics (next subsection) and full SoC compliance despite the heavier energy spread $\eta_e = \pm 20\%$.

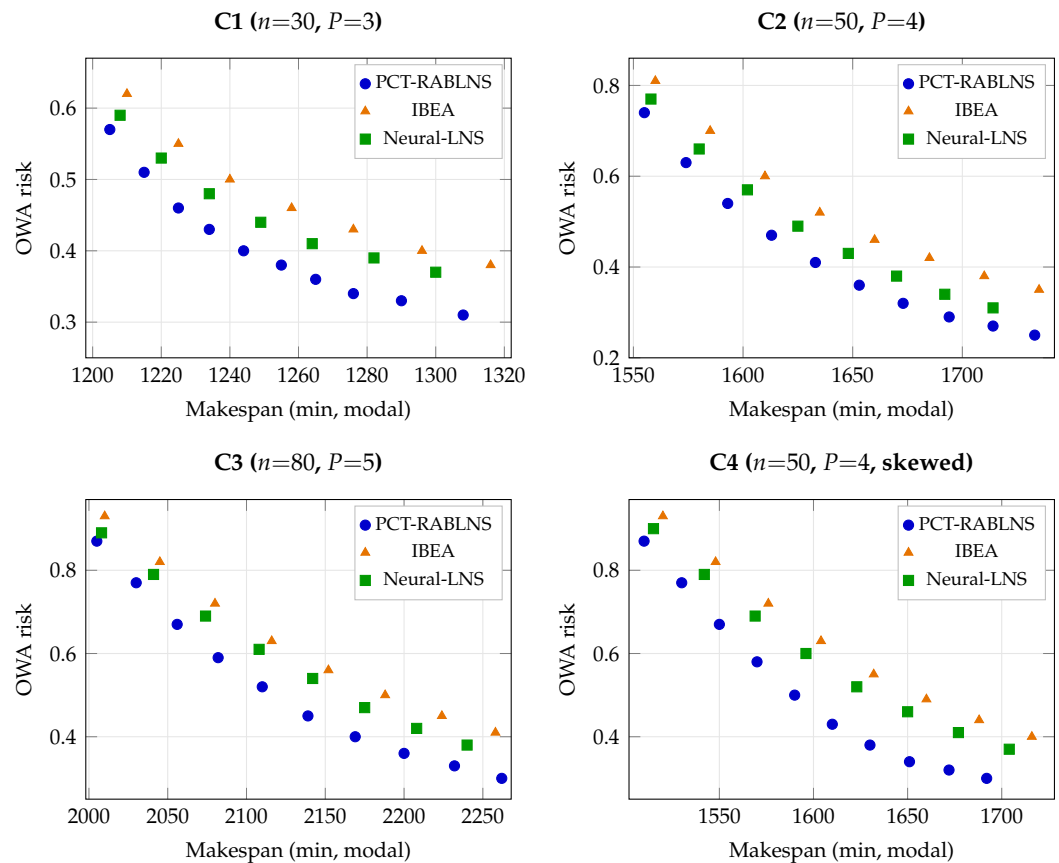


Figure 7. Representative Pareto fronts for cases C1–C4 (makespan vs. OWA overtime risk). PCT-RABLNS (blue circles) achieves better spread and Pareto dominance than IBEA (orange triangles) and Neural-LNS (green squares), with pronounced advantages on clustered layouts (C2, C3) and under fairness stress (C4). PCT-RABLNS archived solutions satisfy the SoC credibility constraint $\text{Cr}(\tilde{E}_h \leq B) \geq 0.9$, whereas baseline archives include infeasible solutions, as reflected by the SoC rates reported in Table 3.

5.5. Risk Fairness Diagnostics

Figure 8 displays overtime risk metrics across all cases for solutions at $\alpha = 0.3$ (fairness-oriented preference). Without an explicit fairness objective, IBEA and Neural-LNS concentrate risk in later shifts—a “risk dumping” pattern that minimises makespan but violates equity. PCT-RABLNS with the OWA objective produces markedly flatter per-shift risk profiles.

Table 4 quantifies fairness via the Gini coefficient and maximum shift risk. PCT-RABLNS reduces maximum risk by 15–25% compared to IBEA and 11–18% compared to Neural-LNS across all cases, with Gini improvements of 12–30%. Makespan overhead

remains minimal (1–3%), demonstrating favourable fairness–efficiency trade-offs even with the additional SoC constraints.

The makespan-overhead column in Table 4 is reported relative to PCT-RABLNS; thus, a negative value indicates that the corresponding baseline produces a nominally shorter total makespan. This apparent advantage is misleading in practice for two reasons. First, the comparison is taken at the same preference $\alpha = 0.3$, where PCT-RABLNS deliberately trades a small amount of makespan for substantially better risk fairness (15–25% lower maximum shift risk and 12–30% lower Gini). Second, the baselines’ shorter makespan is also achieved at the cost of lower SoC compliance (0.83–0.96 for IBEA, 0.89–0.98 for Neural-LNS, see Table 3): part of the makespan saving comes from solutions that would either drain the battery before the end of a shift or violate the credibility threshold $\beta_E = 0.9$ on at least one shift, and would therefore be discarded once feasibility is enforced. In short, a negative overhead means *nominally* faster but less fair and less SoC-safe.

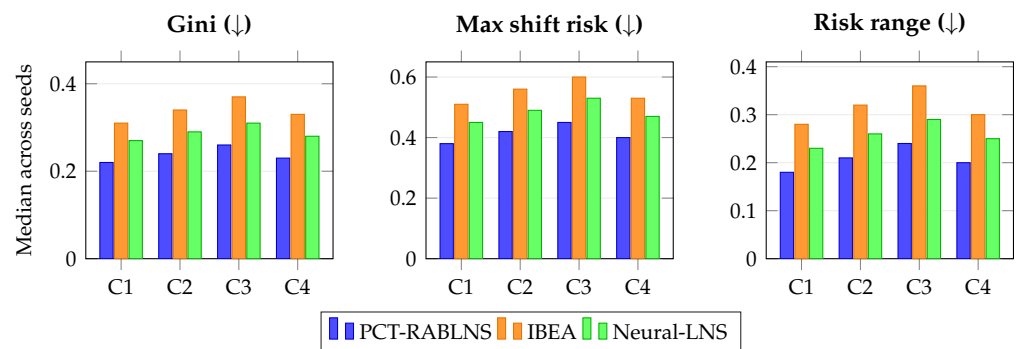


Figure 8. Per-shift overtime risk summary at $\alpha = 0.3$: Gini coefficient (left), maximum shift risk (centre), and risk range $\max_h r_h - \min_h r_h$ (right) across the four cases. PCT-RABLNS (blue) achieves equitable profiles; baselines exhibit “risk dumping” in later shifts, leading to higher Gini, maximum, and range values.

Fairness improvements are statistically significant (Mann–Whitney $p < 0.01$) with large effect sizes ($\hat{A}_{12} \in [0.70, 0.78]$) across all cases. In C4 (fairness stress), PCT-RABLNS keeps risk distributions balanced despite asymmetric uncertainty: the skewed TFN shapes push overtime risk higher in later shifts for methods lacking an explicit fairness encoding, while the OWA objective actively penalises this concentration. The SoC-critical destroy operator does not degrade fairness performance: ablation results (Section 5.7) confirm that removing it leaves risk metrics almost unchanged while significantly worsening battery compliance.

5.6. Convergence and Computational Efficiency

Figure 9 illustrates HV progression over wall-clock time for all four configurations. PCT-RABLNS reaches strong regions of the search space earlier, attaining roughly 90% of its eventual HV 16–28% faster than IBEA and 8–19% faster than Neural-LNS across the four cases. The preference-conditioned transformer provides an immediate quality boost at $t = 0$ (visible elevated HV at initialisation), while the subsequent SoC-aware LNS steps concentrate search effort on productive neighbourhoods.

Table 5 quantifies time-to-90% HV, confirming statistically significant differences ($p < 0.05$) in favour of PCT-RABLNS across all cases.

The convergence patterns reveal two complementary effects. First, preference-conditioned initialization provides immediate quality advantages visible within the first 10–20% of the time budget. Second, the combination of risk-aware, SoC-critical, and cluster-based destroy operators maintains productive search throughout the LNS phase, avoiding stagnation observed in baseline methods that lack domain-specific guidance. The

SoC-critical destroy operator contributes to convergence speed specifically on C2 and C3, where battery constraints interact with temporal scheduling: by promptly redistributing high-energy tasks into lower-consumption shifts, it unblocks parts of the solution space that other destroy operators would reach only much later.

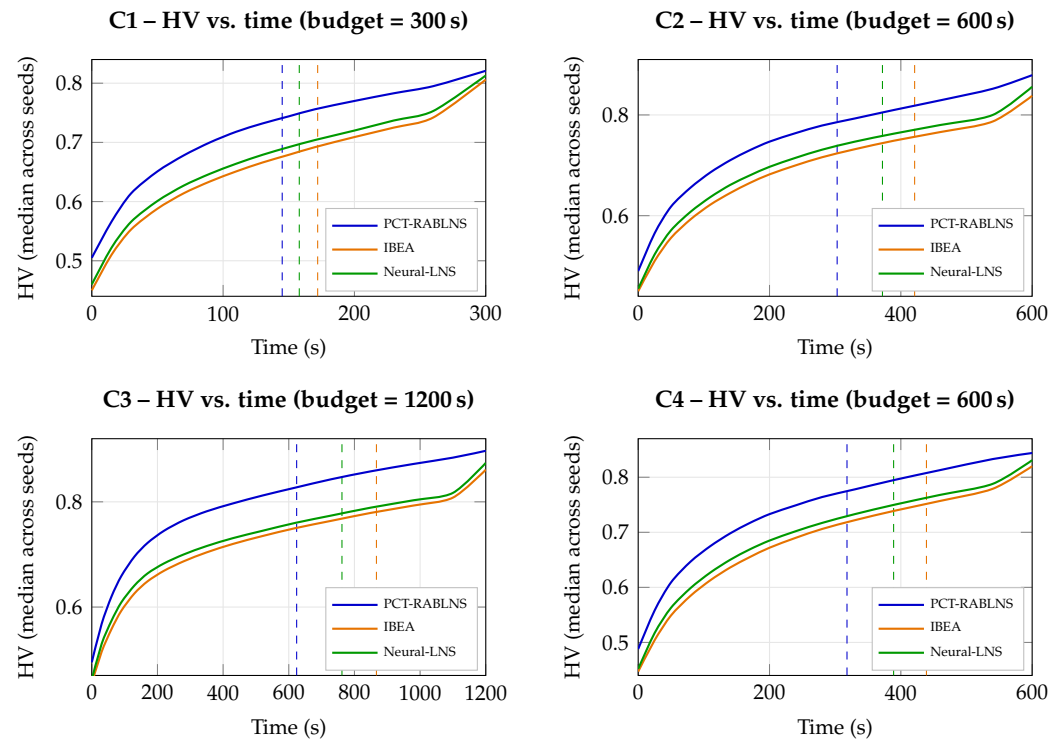


Figure 9. HV convergence for C1–C4. PCT-RABLNS (blue) reaches 90% final HV 16–28% faster than IBEA (orange) and 8–19% faster than Neural-LNS (green). Dashed lines indicate the 90% threshold times. Shaded confidence bands are omitted; 95% CIs are in Table 5. Convergence gains are strongest beyond the simplest case C1, especially on C2–C4.

5.7. Ablation Studies

To isolate individual component contributions, we conduct ablation experiments on case C2 (medium clustered, $n = 50$, $P = 4$). C2 was selected as the ablation testbed for three reasons: (i) it sits at the centre of the case-study spectrum—larger and more challenging than C1 yet sufficiently small to repeat ten seeds across five algorithmic variants within a manageable computational budget; (ii) its spatial clustering and moderate energy spread activate all four destroy operators simultaneously (boundary-aware, risk-proportional, SoC-critical and cluster-based), so the ablation effects can be measured without a single operator dominating the behaviour as in C3 (where SoC-critical dominates) or C4 (where risk-proportional dominates); (iii) its battery-to-dem+and ratio is neither slack (as in C1) nor extremely tight (as in C3), so the SoC-critical destroy operator can be meaningfully turned off without making the problem trivially feasible or trivially infeasible. We verified on a smaller-scale repeat of the ablation that the qualitative ordering of effects on C1, C3 and C4 is consistent with C2 (results in `algorithm_parameters.json` [37]), by sequentially disabling: (1) preference conditioning (fixed $\alpha = 0.5$ during training); (2) risk-aware destroys (boundary-aware and risk-proportional operators replaced by random destroy); (3) SoC-critical destroy (battery-aware operator removed, three remaining operators kept); (4) adaptive operator selection (uniform random selection); (5) OWA fairness objective (replaced by makespan-only minimisation). Each variant is run 10 seeds with a 600 s budget.

Table 6 confirms that all components contribute meaningfully. Preference conditioning shows the largest impact on front quality (−3.8% HV), followed by risk-aware destroy operators (−2.7% HV). The OWA objective is essential for fairness (+42% maximum shift

risk when removed) while carrying only a modest efficiency cost (+1.9% makespan). Critically, removing the SoC-critical destroy operator causes the largest drop in SoC compliance (−0.14, from 1.00 to 0.86), confirming that this component is the primary mechanism ensuring battery feasibility in the archive. By contrast, its removal has a modest effect on HV (−1.8%) and risk fairness (+3.2% max risk), consistent with its specialised role of redistributing energy-infeasible tasks rather than improving the Pareto front geometry.

These results validate the synergistic design: preference-conditioned construction, domain-specific LNS operators, including the SoC-critical destroy, fairness-aware OWA objectives, and adaptive operator selection collectively enable superior performance that no individual component achieves in isolation.

6. Discussion

Our experimental results show that PCT-RABLNS consistently surpasses established baselines across all tested instance families. Hypervolume improves by roughly 2–5%, the maximum per-shift overtime risk decreases by 15–25%, and the SoC compliance rate reaches 1.00 on every case, while the overall makespan increases only modestly, by 1–3%. These findings suggest that integrating preference-conditioned learning with a risk-aware and battery-conscious search strategy effectively balances efficiency, fairness, and energy feasibility when operating times and consumption rates are represented through possibilistic uncertainty.

The case studies reveal an important structural pattern: explicitly modelling fairness through OWA aggregation alters the architecture of Pareto-optimal solutions compared to conventional makespan-focused objectives. Without fairness considerations, both evolutionary and learning-based methods tend to concentrate overtime in later shifts—a “risk dumping” phenomenon that minimises makespan by front-loading efficient tasks but pushes compliance risk onto trailing shifts. When OWA weights penalise this concentration, solutions shift toward more balanced distributions, reducing maximum shift risk by 15–25% while incurring only a 1–3% makespan increase. In contexts such as municipal electric maintenance, where regulatory compliance, driver fatigue, and vehicle availability interact, this represents a meaningful improvement in equity without compromising operational throughput [5,33].

The introduction of fuzzy energy consumption rates adds a second dimension of uncertainty that prior multi-shift routing models neglect. Encoding arc energy as TFNs $\tilde{e}_{ij} = (e_{ij}^A, e_{ij}^B, e_{ij}^C)$ allows planners to express optimistic, most-likely, and pessimistic energy budgets without requiring historical consumption logs, which are rarely available for recently electrified fleets. The SoC credibility constraint $\text{Cr}(\tilde{E}_h \leq B) \geq \beta_E$ provides a calibrated safety margin: at $\beta_E = 0.9$, only 10% of the possibilistic mass exceeds the battery limit, analogous to a 90th-percentile reserve in stochastic models but without requiring a probability distribution. The penalty γV_{SoC} in the scalarised objective shapes the constructive policy to avoid generating solutions that the SoC-critical destroy operator would have to fix in a later LNS step, thereby speeding convergence on energy-constrained cases (C2, C3).

The credibility-based approach for triangular fuzzy numbers provides a practical bridge between expert knowledge and quantitative optimisation. Unlike probabilistic models, which require large historical datasets to estimate distributions, TFN representations naturally accommodate lower–modal–upper assessments commonly elicited in planning practice; for travel-time and energy quantities in this study, lower values are operationally optimistic and upper values pessimistic [18,36]. Evaluating the credibility function—both for overtime and for energy consumption—is computationally inexpensive within heuristic search, avoids Monte Carlo sampling, and remains differentiable, which is essential for

the reinforcement learning training phase where policies are evaluated millions of times. Beyond electric vehicle routing, this framework applies to any domain where uncertainty arises primarily from expert judgement, including project scheduling, resource allocation, and early-stage facility design.

Training a single policy conditioned on preference vectors $z = (\alpha, 1 - \alpha)$ allows PCT-RABLNS to bypass the computational cost of multiple independent optimisation runs or the large populations typical of evolutionary algorithms. The transformer's attention layers capture regularities in high-quality Pareto-optimal solutions and adapt construction to the region of the trade-off space specified by α . The SoC-aware feasibility mask embedded in the policy head ensures that the constructive phase never proposes actions that would violate the hard battery budget at the modal energy level, propagating energy awareness into every autoregressive decoding step. This produces an initial archive of guaranteed SoC-feasible solutions from which the LNS starts, explaining why PCT-RABLNS achieves a compliance rate of 1.00, whereas the baselines accumulate infeasible solutions (IBEA: 0.83–0.96, Neural-LNS: 0.89–0.98 across cases).

The four destroy operators—boundary-aware, risk-proportional, SoC-critical, and cluster-based—leverage structural features of multi-shift electric vehicle routing that general-purpose metaheuristics typically overlook. The boundary-aware destroy operator focuses on tasks near shift transitions, where small changes disproportionately affect overtime credibility. Risk-proportional destroy removes tasks preferentially from high-risk shifts, directly promoting fairness. SoC-critical destroy targets shifts with insufficient battery margin, redistributing energy-intensive tasks into lower-consumption periods and enabling the repair policy to identify more efficient charge profiles. The cluster-based destroy operator enables large-scale spatial rebalancing in instances with geographical grouping (C2, C3). Our ablation study (Section 5.7) shows that the SoC-critical destroy operator is the primary driver of battery compliance (Δ SoC rate = -0.14 when removed) while having a modest effect on the Pareto geometry (Δ HV = -1.8%), confirming its role as a specialised feasibility–repair mechanism rather than a general-purpose improvement operator. The adaptive operator selection identifies which operators contribute most effectively across instance types: cluster-based destroy dominates on spatially clustered instances (C2, C3), risk-proportional destroy is essential under fairness stress (C4), and SoC-critical destroy is most active in C3 where the battery-to-demand ratio is tightest.

By combining learned construction with classical local search, PCT-RABLNS consistently outperforms approaches based purely on learning or purely on metaheuristics. The preference-conditioned transformer generates energy-feasible and risk-balanced starting solutions by extracting patterns from thousands of training instances, while the LNS component systematically explores neighbourhoods and applies SA acceptance rules to refine solutions. This interplay is particularly important for complex instances (C2–C4), where learning uncovers structural regularities—including recurring relationships between spatial clusters, shift boundaries, and energy consumption—but achieving high-quality solutions still requires thorough local optimisation.

It is useful to summarise the role and the empirical contribution of each architectural element, since the framework combines several modelling and algorithmic ingredients. The *possibilistic representation* (TFNs for travel, service and energy) is the modelling block that turns expert assessments into a tractable mathematical object; without it, the formulation collapses to a deterministic E-VRP that cannot capture the imprecision typical of recently electrified municipal fleets. The *closed-form credibility evaluation* is the analytical block that makes both the overtime and the SoC constraints differentiable and inexpensive to evaluate, which, in turn, enables the gradient-based RL training and the high evaluation throughput observed in Table A6. The *OWA fairness objective* is the decision-theoretic block: the ablation

in Table 6 shows that removing it inflates the maximum shift risk by 42% with only a 1.9% makespan saving, confirming that fairness is not implicit in efficient routing. The *preference-conditioned transformer* is the constructive block: a single trained policy covers the entire Pareto spectrum, and its ablation costs -3.8% HV. The *four destroy operators* are the refinement block: the boundary-aware, risk-proportional and cluster-based operators jointly improve Pareto geometry, while the SoC-critical operator is the dedicated feasibility–repair mechanism whose removal alone causes the largest drop in SoC compliance (-0.14). Read together, the components are synergistic rather than redundant: each one moves a different metric (HV, max risk, SoC rate or convergence time) and the full pipeline is needed to obtain simultaneous gains in all four.

A natural question is whether the additional complexity of preference-conditioned reinforcement learning is worth its cost compared with simpler heuristic baselines. Our experiments allow a direct trade-off assessment. On the methodological side, PCT-RABLNS introduces a one-time training cost of approximately 60 GPU-hours on an NVIDIA A100, additional GPU memory of about 3.4 GB at inference, and roughly 200 s of GPU time per 1200 s run on case C3 (Table A6). On the performance side, it delivers 2–5% higher hypervolume, 15–25% lower maximum shift risk, full SoC compliance, and a 16–28% reduction in time-to-90%-HV against an indicator-based evolutionary algorithm (IBEA) and a fixed-weight neural LNS that share the same wall-clock budget and the same family of destroy operators. Once the policy is trained, inference and LNS refinement complete within the 5–20 min planning horizon typical of municipal day-ahead scheduling, so the operational gain—fairer schedules with guaranteed battery feasibility—comes at no extra runtime cost. The trade-off is therefore favourable when the planning problem is recurrent (training cost is amortised over many planning days), when fairness and battery feasibility are first-class constraints, and when expert-elicited TFNs are available. For ad hoc single-instance planning with unrestricted runtime and without fairness or SoC requirements, a well-tuned IBEA or even a strong metaheuristic without learning would remain a sensible, simpler choice.

The generalisation capability of PCT-RABLNS depends on the diversity of the training data. Our dataset covers instance sizes $n \in [20, 80]$, shift counts $P \in [2, 6]$, spatial layouts (uniform and clustered), uncertainty levels ($\eta_d, \eta_q \in [\pm 20\%, \pm 35\%]$), and energy spreads ($\eta_e \in [\pm 10\%, \pm 25\%]$). Performance on the held-out case studies C1–C4, which share but do not exactly match the training distributions, demonstrates good generalisation. For problem variants outside this scope, fine-tuning provides a practical adaptation strategy without full retraining.

The experimental protocol was designed for robustness: all methods operated under identical computational budgets, each configuration was repeated with ten random seeds, and non-parametric statistical tests with Holm–Bonferroni correction were applied. Effect sizes quantify the practical relevance of observed differences. For medium and large instances (C2 and C3, $\hat{A}_{12} \in [0.65, 0.72]$), effect sizes are substantial, highlighting the operational significance of the improvements. From a practical standpoint, PCT-RABLNS offers a favourable computational profile: training is a one-time cost of approximately 60 GPU-hours on an NVIDIA A100 40 GB, whereas inference takes seconds and LNS refinement completes within typical planning horizons (5–20 min depending on instance size). This fits naturally with day-ahead or shift-level scheduling in municipal electric maintenance operations.

Several limitations remain. Our work focuses on a single-vehicle setting; multi-vehicle extensions require coordinated workload and energy balancing across both vehicles and shifts. The single-vehicle scope was deliberate: it isolates the coupling between multi-shift segmentation, fairness in overtime exposure and battery feasibility, which is the core

methodological contribution of this paper. It also matches a common operational pattern in municipal service, in which inspection or maintenance units operate independently and are planned one vehicle at a time. The framework, however, lends itself to a multi-vehicle extension along three lines. First, the solution encoding ω generalises to a vector $(\omega^{(1)}, \dots, \omega^{(K)})$ of per-vehicle sequences, with an additional assignment step (each task to one vehicle) that can be performed by a separate routing head of the transformer or by a clustered warm start. Second, the OWA fairness objective extends naturally from “risk per shift of one vehicle” to “risk per (vehicle, shift) cell”, preserving the closed-form credibility expression and only increasing the length of the sorted risk vector. Third, the four destroy operators remain meaningful: the boundary-aware and risk-proportional operators generalise per-vehicle, the SoC-critical operator extends to inter-vehicle task swaps that move energy-intensive tasks to vehicles with more residual SoC, and a new cross-vehicle exchange operator can be added to share workload. The main computational implication is that the transformer must encode a (vehicle, shift) context vector instead of a single shift index, which we expect to add proportionally to the training cost but not to change the qualitative behaviour observed in the single-vehicle case. We mark this as the most important direction for future work. The benchmarks are synthetic approximations of realistic instances—in the sense that the network sizes, shift structures, battery capacities, TFN spreads and uncertainty asymmetries reproduce parameters drawn from anonymised, aggregated planning records of an operational electrified municipal maintenance unit (see the Acknowledgments and `uncertainty_parameters.json` [37]) rather than from a real-world routing instance reproduced node-for-node, and validation using audited operational logs from an electrified municipal fleet would strengthen external validity. We assume a static task set and a full recharge at each depot return; real-time arrivals and partial opportunity charging call for online learning and anytime algorithms. The possibilistic quantification simplifies uncertainty modelling by avoiding full probability distributions, but hybrid probabilistic–possibilistic approaches could blend expert input with partial historical data when both sources are available. OWA is one route to expressing fairness: lexicographic rules or constraint-based formulations may be more appropriate in contexts where regulatory thresholds are hard rather than soft. Nevertheless, the integration of possibilistic uncertainty, OWA-driven fairness, SoC-credibility constraints, and preference-conditioned learning constitutes a coherent framework for fair and energy-safe optimisation under ambiguity. The experimental results demonstrate improvements in all four dimensions simultaneously—Pareto quality, fairness, battery feasibility, and convergence speed—with statistical guarantees and operationally meaningful effect sizes.

7. Conclusions

This paper has introduced a unified framework for multi-shift, single-vehicle electric service routing in which fairness in the distribution of overtime risk and battery–range feasibility are treated as first-class objectives rather than afterthoughts. The MS-SEVRP-PU formulation brings together, in a single bi-objective programme, three ingredients that prior work has only addressed in isolation: a possibilistic representation of imprecise travel times, service durations and arc/on-site energy consumption through triangular fuzzy numbers; a closed-form credibility evaluation that turns both overtime and state-of-charge constraints into differentiable, computationally inexpensive expressions; and an Ordered Weighted Averaging aggregation of per-shift overtime credibilities that explicitly penalises inequitable risk concentrations (answer to **RQ1**). Together, these elements show that imprecise expert knowledge—typical of recently electrified municipal fleets—can be encoded without forcing a probabilistic law, and that the resulting fuzzy quantities can be embedded in a tractable optimisation model that admits gradient-based reasoning,

with $O(P)$ per-solution evaluation cost against $O(10^3-10^4)$ for an equivalent Monte Carlo estimate (answer to **RQ2**).

On the algorithmic side, the proposed PCT-RABLNS solver demonstrates that learning and classical local search are complementary rather than substitutes (answer to **RQ3**). A single Pareto-Conditioned Transformer, trained via multi-objective reinforcement learning with uniform preference sampling, generates high-quality and SoC-feasible initial solutions across the entire trade-off spectrum, while a Risk-Aware and Battery-Conscious Large Neighbourhood Search refines them through four specialised destroy operators—boundary-aware, risk-proportional, SoC-critical and cluster-based—selected adaptively. The ablation study confirms that each component plays a distinct, non-redundant role: preference conditioning is the main driver of Pareto-front quality, the OWA objective is the main driver of fairness, and the SoC-critical destroy operator is the main driver of battery compliance. None of these gains is achievable by any single component in isolation.

The empirical evidence supports the practical relevance of the framework (answer to **RQ4**). On the four calibrated municipal-maintenance case studies, PCT-RABLNS improves hypervolume by 2–5% over strong evolutionary and learning baselines, reduces the maximum per-shift overtime risk by 15–25% and the Gini coefficient by 12–30%, and reaches 90% of its final hypervolume 16–28% faster than IBEA and 8–19% faster than Neural-LNS. Crucially, every archived solution satisfies the SoC credibility constraint $\text{Cr}(\tilde{E}_h \leq B) \geq 0.9$ for all shifts, against compliance rates as low as 0.83 for IBEA and 0.89 for Neural-LNS on the energy-intensive cases. These improvements are obtained at a marginal makespan overhead of only 1–3%, are statistically significant under nonparametric tests with Holm–Bonferroni correction, and exhibit large Vargha–Delaney effect sizes on medium and large instances. The operational reading of these numbers is that fairness in overtime exposure and guaranteed battery feasibility—two requirements that municipal operators increasingly have to reconcile with regulatory and labour constraints—can be attained without sacrificing throughput, and within the 5–20 min planning horizon typical of day-ahead and shift-level scheduling.

These results should be interpreted in light of the study design. The benchmarks are calibrated synthetic instances whose parameters were elicited from the operations engineers of an industrial partner; while they reproduce the operating regimes of an electrified municipal maintenance unit, validation on audited field logs from a multi-vehicle fleet remains an important next step. The work also assumes a static task set and a full recharge at each depot return; real-time arrivals and partial opportunity charging would call for online learning and anytime variants of the solver. The single-vehicle scope was a deliberate methodological choice—it isolates the coupling between multi-shift segmentation, fairness and battery feasibility—but the encoding, the OWA aggregation and the four destroy operators all admit natural extensions to a multi-vehicle setting, as discussed in Section 6.

Beyond electric vehicle routing, the methodological building blocks generalise to a broader class of fairness-sensitive decision problems under imprecise information. OWA-based risk aggregation transfers directly to workforce scheduling, resource allocation and service network design whenever equitable distribution across periods or agents is a goal in itself. The SoC credibility constraint is applicable to any battery-operated asset scheduling problem. Preference-conditioned policy learning extends to other multi-objective combinatorial problems where the Pareto front must be approximated by a single trained model. Analytical possibilistic evaluation provides a reusable bridge between expert knowledge and quantitative optimisation when historical data are scarce or unreliable.

Promising directions for future work include the multi-vehicle EV extension with joint workload and charging coordination; dynamic task arrivals handled by online replanning; alternative fairness operators such as lexicographic and Nash-based objectives; hybrid probabilistic–possibilistic uncertainty models that blend expert input with partial historical data; and real-world validation with operational logs from electrified municipal fleets.

Funding: This research received the University of Salento Research Base Funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: The configurations, parameters, and aggregated results presented in this study are openly available at <https://github.com/fnuni/2026eng> [37] (accessed on 31 March 2026). The repository contains: (i) instance and algorithm configurations required to replicate the experimental setup; (ii) aggregated Pareto quality metrics (hypervolume, IGD⁺), fairness diagnostics (Gini, maximum shift risk), and convergence data per seed. Per-seed archived solutions are not released because they embed stochastic GPU outputs that vary across hardware and driver versions. Operational scheduling logs for case calibration are confidential and are therefore not included; their statistical properties are preserved in the synthetic TFN parameters documented in `uncertainty_parameters.json`.

Acknowledgments: The author expresses gratitude to M.O. s.r.l. (Lecce, Italy) for providing anonymized and aggregated operational data for this study. Additionally, the author acknowledges the anonymous reviewers for their valuable constructive input, which significantly enhanced the depth and quality of this work.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A. Mathematical Details

Appendix A.1. Credibility Function Derivation

For a triangular fuzzy number $\tilde{t} = (a, b, c)$ with membership function $\mu_{\tilde{t}}(x)$ defined in Section 3.1, and a threshold t_0 , the credibility measure is defined as [36]:

$$\text{Cr}(\tilde{t} \leq t_0) = \frac{1}{2} [\text{Pos}(\tilde{t} \leq t_0) + \text{Nec}(\tilde{t} \leq t_0)],$$

where Pos and Nec are the possibility and necessity measures, respectively. For triangular fuzzy numbers, these simplify to:

$$\text{Pos}(\tilde{t} \leq t_0) = \begin{cases} 0, & t_0 < a, \\ \frac{t_0 - a}{b - a}, & a \leq t_0 \leq b, \\ 1, & t_0 > b, \end{cases} \quad \text{Nec}(\tilde{t} \leq t_0) = \begin{cases} 0, & t_0 < b, \\ \frac{t_0 - b}{c - b}, & b \leq t_0 \leq c, \\ 1, & t_0 > c. \end{cases}$$

Combining these, the closed-form credibility function is:

$$\text{Cr}(\tilde{t} \leq t_0) = \begin{cases} 0, & t_0 \leq a, \\ \frac{t_0 - a}{2(b - a)}, & a < t_0 \leq b, \\ \frac{1}{2} + \frac{t_0 - b}{2(c - b)}, & b < t_0 < c, \\ 1, & t_0 \geq c. \end{cases} \tag{A1}$$

This function is continuous, piecewise linear, and monotonically non-decreasing, making it suitable for gradient-based optimisation. In particular, $\text{Cr}(\tilde{t} \leq b) = \frac{1}{2}$: the modal value is the break-even credibility point. The same formula applies to both the overtime

credibility r_h (with $\tilde{t} = \tilde{\sigma}_h$ and threshold $t_0 = L$) and the energy credibility (with $\tilde{t} = \tilde{E}_h$ and threshold $t_0 = B$).

Figure A1 illustrates the credibility function for the representative TFN $\tilde{t} = (20, 30, 45)$.

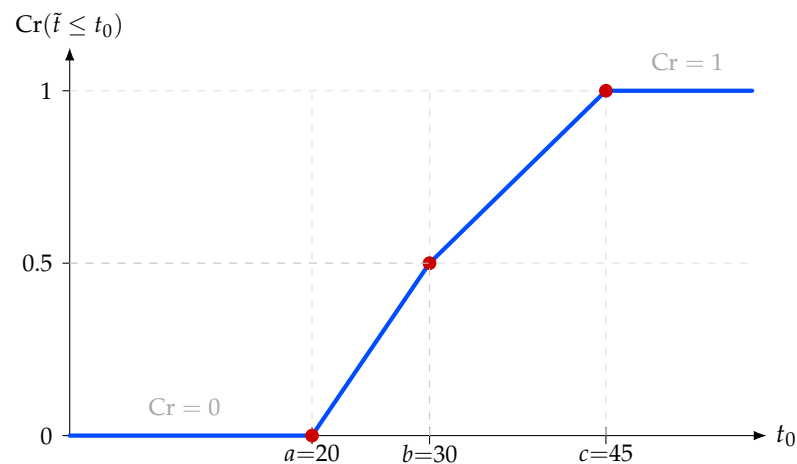


Figure A1. Credibility function $Cr(\tilde{t} \leq t_0)$ for $\tilde{t} = (20, 30, 45)$. The function is piecewise linear: zero for $t_0 \leq a$, linearly rising to 0.5 at the modal value b , continuing linearly to 1 at c , and constant thereafter. Red dashed lines highlight that $Cr(\tilde{t} \leq b) = 0.5$ exactly. The same function—evaluated at $t_0 = L$ for overtime and at $t_0 = B$ for energy—is used throughout the MS-SEVRP-PU formulation.

Appendix A.2. OWA Weight Specifications

The OWA aggregation defined in Equation (6) offers flexibility through weight selection. Table A1 summarises common schemes.

Table A1. OWA weight specifications and their risk attitudes.

Scheme	Formula	Risk Attitude
Front-loaded	$w_h = \frac{2(P-h+1)}{P(P+1)}$	Equity-focused (penalise max)
Uniform	$w_h = \frac{1}{P}$	Neutral (average risk)
Back-loaded	$w_h = \frac{2h}{P(P+1)}$	Min-focused (protect worst)
Max-only	$w_1 = 1, w_{h>1} = 0$	Pure minimax

Front-loaded weights provide strong protection against overtime concentration while maintaining computational tractability. Our ablation study (Section 5.7) empirically validates this choice by showing that removing the OWA objective increases maximum shift risk by 42% while raising makespan by only 1.9%.

Appendix A.3. Computational Complexity Analysis

The MS-SEVRP-PU is NP-hard, as it strictly generalises the classical TSP (NP-hard) through:

1. Shift segmentation decisions (combinatorial partitioning of n tasks into P ordered groups);
2. Fuzzy parameter evaluation (credibility computations for both overtime and energy);
3. Battery-feasibility constraints (SoC check per shift);
4. Bi-objective Pareto optimality.

The solution space has size $O(n! \cdot \binom{n-1}{P-1})$: the first factor accounts for task ordering and the second for shift-break placement. For modest $n = 50, P = 4$, exact enumeration is infeasible ($\sim 10^{70}$ solutions).

The closed-form credibility evaluation (Appendix A.1) provides $O(P)$ time complexity per solution, compared to $O(10^3 - 10^4)$ for Monte Carlo methods, enabling efficient heuristic

search. The fuzzy energy consumption per shift (Equation (3)) is computed by a single pass over the shift arc sequence at $O(|\mathcal{S}_h|)$ cost, adding negligible overhead to the standard duration computation.

Appendix B. Detailed Results and Statistical Analysis

This appendix provides a comprehensive statistical analysis supporting Section 5, including pairwise comparison test outputs, effect-size calculations, fairness metrics across preference levels, convergence percentiles, and resource utilisation.

Appendix B.1. Statistical Test Results

Tables A2 and A3 report Mann–Whitney U test results for pairwise method comparisons on HV and IGD⁺. P-values are adjusted via Holm-Bonferroni correction. Effect sizes are computed as $\hat{A}_{12} = U/(n_1n_2)$ where $n_1 = n_2 = 10$ [40].

Table A2. Mann–Whitney U test results for hypervolume (PCT-RABLNS vs. baselines).

Case	Comparison	U	p-Value (adj.)	\hat{A}_{12}
C1	PCT vs. IBEA	92	0.007	0.92
	PCT vs. Neural	78	0.032	0.78
C2	PCT vs. IBEA	100	<0.001	1.00
	PCT vs. Neural	95	0.002	0.95
C3	PCT vs. IBEA	100	<0.001	1.00
	PCT vs. Neural	97	0.001	0.97
C4	PCT vs. IBEA	99	<0.001	0.99
	PCT vs. Neural	94	0.002	0.94

$\hat{A}_{12} > 0.71$ indicates a large effect size.

Table A3. Mann–Whitney U test results for IGD⁺ (PCT-RABLNS vs. baselines; lower \hat{A}_{12} means PCT-RABLNS achieves lower IGD⁺, i.e., better).

Case	Comparison	U	p-Value (adj.)	\hat{A}_{12}
C1	PCT vs. IBEA	8	0.008	0.08
	PCT vs. Neural	22	0.042	0.22
C2	PCT vs. IBEA	0	<0.001	0.00
	PCT vs. Neural	5	0.003	0.05
C3	PCT vs. IBEA	0	<0.001	0.00
	PCT vs. Neural	3	0.002	0.03
C4	PCT vs. IBEA	1	<0.001	0.01
	PCT vs. Neural	6	0.004	0.06

All comparisons are statistically significant (adjusted $p < 0.05$) with large effect sizes on C2–C4.

Appendix B.2. Additional Performance Metrics

Table A4 reports maximum shift risk, Gini coefficient, makespan, and HV for solutions at preference levels $\alpha \in \{0.2, 0.3, 0.4, 0.5\}$ for case C3 (large clustered). As α increases (shifting emphasis toward makespan), risk metrics worsen for all methods, but PCT-RABLNS degrades more slowly, maintaining both fairness and SoC compliance across the entire preference spectrum.

Table A4. Fairness and quality metrics across preference levels (case C3): median values.

α	Method	Max Risk	Gini	Makespan	HV	SoC Rate
0.2	PCT-RABLNS	0.41	0.21	362	0.912	1.00
	IBEA	0.58	0.35	349	0.875	0.84
	Neural-LNS	0.51	0.29	355	0.891	0.90
0.3	PCT-RABLNS	0.45	0.26	355	0.897	1.00
	IBEA	0.60	0.37	345	0.861	0.83
	Neural-LNS	0.53	0.31	348	0.874	0.89
0.4	PCT-RABLNS	0.49	0.30	348	0.882	1.00
	IBEA	0.62	0.39	341	0.852	0.85
	Neural-LNS	0.56	0.34	343	0.865	0.91
0.5	PCT-RABLNS	0.53	0.33	342	0.866	1.00
	IBEA	0.65	0.41	338	0.838	0.86
	Neural-LNS	0.59	0.36	339	0.851	0.92

Table A5 reports detailed convergence statistics: time to reach 50%, 75%, 90%, and 95% of final HV for case C3. Table A6 reports average computational resource utilisation per run for case C3.

Table A5. Time (seconds) to reach HV percentiles (case C3): median [95% CI].

Method	50% HV	75% HV	90% HV	95% HV
PCT-RABLNS	128 [119, 137]	342 [327, 357]	624 [600, 648]	903 [876, 930]
IBEA	193 [179, 207]	507 [483, 531]	867 [836, 898]	1163 [1129, 1197]
Neural-LNS	168 [156, 180]	441 [419, 463]	762 [733, 791]	1045 [1012, 1078]

Table A6. Computational resource utilisation (case C3, 1200 s budget).

Method	CPU Time (s)	GPU Time (s)	Memory (GB)	Evaluations
PCT-RABLNS	1195	194	3.4	14 291
IBEA	1198	–	1.8	18 741
Neural-LNS	1196	210	3.6	12 683

GPU time for inference/evaluation only; IBEA runs CPU-only.

Appendix C. Nomenclature

Tables A7 and A8 list acronyms and symbols used in the paper.

Table A7. Acronyms.

Acronym	Definition
E-VRP	Electric Vehicle Routing Problem
HV	Hypervolume
IBEA	Indicator-Based Evolutionary Algorithm
IGD ⁺	Inverted Generational Distance Plus
LNS	Large Neighbourhood Search
MORL	Multi-Objective Reinforcement Learning
MS-SEVRP-PU	Multi-Shift Single Electric Vehicle Routing Problem under Possibilistic Uncertainty
OWA	Ordered Weighted Averaging
PCT-RABLNS	Pareto-Conditioned Transformer with Risk-Aware and Battery-Conscious LNS
SA	Simulated Annealing
SoC	State of Charge
TFN	Triangular Fuzzy Number
VRP	Vehicle Routing Problem

Table A8. Mathematical symbols.

Symbol	Description
$G = (V, E)$	Complete directed graph
N	Set of n tasks
P	Number of consecutive shifts
L	Maximum allowable shift duration (min)
B	Battery capacity (kWh)
β_E	SoC credibility safety threshold
\tilde{d}_{ij}	Fuzzy travel time from i to j
\tilde{q}_i	Fuzzy service time at task i
\tilde{e}_{ij}	Fuzzy arc energy consumption
\tilde{f}_i	Fuzzy on-site energy consumption
(a, b, c)	TFN parameters: lower bound, modal value, upper bound
ω	Solution sequence (tasks + shift-break tokens)
$\#_h$	Shift-break token for shift h
$\mathcal{S}_h(\omega)$	Tasks assigned to shift h
$\tilde{\sigma}_h$	Fuzzy duration of shift h
$\tilde{E}_h(\omega)$	Fuzzy energy consumed in shift h
$\lambda(\omega)$	Total makespan (sum of modal shift durations)
$V_{\text{SoC}}(\omega)$	Modal energy over-consumption (penalty term)
r_h	Overtime credibility for shift h
$\text{Cr}(\cdot)$	Credibility measure
$\mathbf{r}(\omega)$	Vector of per-shift overtime credibilities
$r_{(1)} \geq \dots \geq r_{(P)}$	Order statistics of \mathbf{r} (descending)
\mathbf{w}	OWA weight vector
$R_{\text{OWA}}(\mathbf{r}; \mathbf{w})$	OWA aggregation of risks
$\mathbf{z} = (\alpha, 1 - \alpha)$	Preference vector
π_θ	Policy (transformer) with parameters θ
s_t	Partial solution state at step t
a_t	Action (task or shift-break token) at step t
$J_\alpha(\omega)$	Scalarised objective with preference α
T_k	SA temperature at iteration k
ε	Archive dominance tolerance
\hat{A}_{12}	Vargha-Delaney effect size statistic

References

1. Hashemi-Amiri, O.; Ji, R.; Tian, K. An Integrated Location–Scheduling–Routing Framework for a Smart Municipal Solid Waste System. *Sustainability* **2023**, *15*, 7774. [\[CrossRef\]](#)
2. Ren, Y.; Dessouky, M.; Ordóñez, F. The multi-shift vehicle routing problem with overtime. *Comput. Oper. Res.* **2010**, *37*, 1987–1998. [\[CrossRef\]](#)
3. Zhang, J.; Luo, K.; Florio, A.M.; van Woensel, T. Solving Large-Scale Dynamic Vehicle Routing Problems with Stochastic Requests. *Eur. J. Oper. Res.* **2023**, *306*, 596–614. [\[CrossRef\]](#)
4. Karoonsoontawong, A.; Punyim, P.; Nueangnitarnaporn, W.; Ratanavaraha, V. Multi-Trip Time-Dependent Vehicle Routing Problem with Soft Time Windows and Overtime Constraints. *Netw. Spat. Econ.* **2020**, *20*, 549–598. [\[CrossRef\]](#)
5. Matl, P.; Hartl, R.; Vidal, T. Workload Equity in Vehicle Routing Problems: A Survey and Analysis. *Transp. Sci.* **2018**, *52*, 239–260. [\[CrossRef\]](#)
6. Pelletier, S.; Jabali, O.; Laporte, G. 50th Anniversary Invited Article—Goods Distribution with Electric Vehicles: Review and Research Perspectives. *Transp. Sci.* **2016**, *50*, 3–22. [\[CrossRef\]](#)
7. Schneider, M.; Stenger, A.; Goetze, D. The Electric Vehicle-Routing Problem with Time Windows and Recharging Stations. *Transp. Sci.* **2014**, *48*, 500–520. [\[CrossRef\]](#)
8. Schiffer, M.; Walther, G. The Electric Location Routing Problem with Time Windows and Partial Recharging. *Eur. J. Oper. Res.* **2017**, *260*, 995–1013. [\[CrossRef\]](#)
9. Toth, P.; Vigo, D. *Vehicle Routing: Problems, Methods, and Applications*, 2nd ed.; SIAM: Philadelphia, PA, USA, 2014. [\[CrossRef\]](#)
10. Gendreau, M.; Potvin, J.Y. *Handbook of Metaheuristics*, 2nd ed.; Springer: New York, NY, USA, 2014. [\[CrossRef\]](#)
11. Goetze, D.; Schneider, M. Routing a Mixed Fleet of Electric and Conventional Vehicles. *Eur. J. Oper. Res.* **2015**, *245*, 81–99. [\[CrossRef\]](#)
12. Laporte, G. *The Vehicle Routing Problem*; SIAM: Philadelphia, PA, USA, 2009.

13. Ritzinger, U.; Puchinger, J.; Hartl, R.F. A survey on dynamic and stochastic vehicle routing problems. *Int. J. Prod. Res.* **2016**, *54*, 215–231. [CrossRef]
14. Eksioglu, B.; Vural, A.V.; Reisman, A. The vehicle routing problem: A taxonomic review. *Comput. Ind. Eng.* **2009**, *57*, 1472–1483. [CrossRef]
15. Cao, E.; Lai, M.; Yang, H. Open vehicle routing problem with demand uncertainty and its robust strategies. *Expert Syst. Appl. Int. J.* **2014**, *41*, 3569–3575. [CrossRef]
16. Shi, Y.; Boudouh, T.; Grunder, O. A Fuzzy Chance-constraint Programming Model for a Home Health Care Routing Problem with Fuzzy Demand. In *Proceedings of the 6th International Conference on Operations Research and Enterprise Systems—Volume 1: ICORES*; SciTePress: Setubal, Portugal, 2017; pp. 369–376. [CrossRef]
17. Zhao, L.; Cao, N. Fuzzy Random Chance-Constrained Programming Model for the Vehicle Routing Problem of Hazardous Materials Transportation. *Symmetry* **2020**, *12*, 1208. [CrossRef]
18. Attari, M.Y.N.; Torkayesh, A.E.; Malmir, B.; Jami, E.N. Robust possibilistic programming for joint order batching and picker routing problem in warehouse management. *Int. J. Prod. Res.* **2021**, *59*, 4434–4452. [CrossRef]
19. Devnath, S.; De, M.; Mondal, S.S.; Maiti, M. Two-stage multi-item 4-dimensional transportation problem with fuzzy risk and substitution. *J. Ambient. Intell. Humaniz. Comput.* **2023**, *14*, 9469–9496. [CrossRef]
20. Yager, R.R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Syst. Man Cybern.* **1988**, *18*, 183–190. [CrossRef]
21. D’Urso, P.; Chachi, J.; Kazemifard, A.; De Giovanni, L. OWA-based multi-criteria decision making based on fuzzy methods. *Ann. Oper. Res.* **2024**. [CrossRef]
22. Kwon, Y.D.; Choo, J.; Kim, B.; Yoon, I.; Gwon, Y.; Min, S. POMO: Policy Optimization with Multiple Optima for Reinforcement Learning. In *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Virtual, 6–12 December 2020*; Volume 33, pp. 21188–21198.
23. Kool, W.; Van Hoof, H.; Welling, M. Attention, learn to solve routing problems! In *Proceedings of the International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019*.
24. Pan, W.; Liu, S.Q.S. Deep reinforcement learning for the dynamic and uncertain vehicle routing problem. *Appl. Intell.* **2022**, *53*, 405–422. [CrossRef]
25. Mazyavkina, N.; Sviridov, S.; Ivanov, S.; Burnaev, E. Reinforcement Learning for Combinatorial Optimization: A Survey. *Comput. Oper. Res.* **2021**, *134*, 105400. [CrossRef]
26. Hottung, A.; Tierney, K. Neural large neighborhood search for routing problems. *Artif. Intell.* **2022**, *313*, 103786. [CrossRef]
27. Shi, R.; Niu, L. A Brief Survey on Learning Based Methods for Vehicle Routing Problems. *Procedia Comput. Sci.* **2023**, *221*, 773–780. [CrossRef]
28. Guan, Q.; Cao, H.; Jia, L.; Yan, D.; Chen, B. Synergetic attention-driven transformer: A deep reinforcement learning approach for vehicle routing problems. *Expert Syst. Appl.* **2025**, *274*, 126961. [CrossRef]
29. Desaulniers, G.; Errico, F.; Irnich, S.; Schneider, M. Exact Algorithms for Electric Vehicle-Routing Problems with Time Windows. *Oper. Res.* **2016**, *64*, 1388–1405. [CrossRef]
30. Nucci, F. Multi-Shift Single-Vehicle Routing Problem Under Fuzzy Uncertainty During the COVID-19 Pandemic. *J. Fuzzy Log. Model. Eng.* **2022**, *1*, 16–27. [CrossRef]
31. Kim, G. Dynamic Vehicle Routing Problem with Fuzzy Customer Response. *Sustainability* **2023**, *15*, 4376. [CrossRef]
32. Yager, R.; Kacprzyk, J.; Beliakov, G. *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*; Springer: Berlin/Heidelberg, Germany, 2011. [CrossRef]
33. Bertsimas, D.; Farias, V.F.; Trichakis, N. On the Efficiency-Fairness Trade-off. *Manag. Sci.* **2012**, *58*, 2234–2250. [CrossRef]
34. Hottung, A.; Tierney, K. Neural Large Neighborhood Search for the Capacitated Vehicle Routing Problem. In *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI), Santiago de Compostela, Spain, 29 August–8 September 2020*; pp. 443–450. [CrossRef]
35. Hayes, C.F.; Rădulescu, R.; Bargiacchi, E.; Källström, J.; Macfarlane, M.; Reymond, M.; Roijers, D.M.; Heintz, F.; Howley, E.; Irissappane, A.A.; et al. A practical guide to multi-objective reinforcement learning and planning. *Auton. Agents Multi-Agent Syst.* **2022**, *36*, 26. [CrossRef]
36. Liu, B. *Uncertainty Theory*, 4th ed.; Springer: Berlin/Heidelberg, Germany, 2015.
37. Nucci, F. Data of Case Study ‘Multi-Shift Scheduling of Electric Service Operations Under Fuzzy Uncertainty via Preference-Guided Deep Learning: The Single-Vehicle Case’. Available online: <https://github.com/fnuni/2026eng> (accessed on 31 March 2026).
38. Zitzler, E.; Künzli, S. Indicator-based selection in multiobjective search. In *Proceedings of the Parallel Problem Solving from Nature-PPSN VIII*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 832–842. [CrossRef]

39. Ishibuchi, H.; Imada, R.; Masuyama, N.; Nojima, Y. Comparison of Hypervolume, IGD and IGD+ from the Viewpoint of Optimal Distributions of Solutions. In *Evolutionary Multi-Criterion Optimization*; Deb, K., Goodman, E., Coello Coello, C.A., Klamroth, K., Miettinen, K., Mostaghim, S., Reed, P., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 332–345.
40. Vargha, A.; Delaney, H.D. A Critique and Improvement of the CL Common Language Effect Size Statistics of McGraw and Wong. *J. Educ. Behav. Stat.* **2000**, *25*, 101–132. [[CrossRef](#)]

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.