# Federated Learning-Based Sustainable Electric Vehicle Routing to Alleviate Urban Grid Pressure

Lucas Hartman[1], Ethan Pigou[2], Santiago Gomez-Rosero[3], Miriam A. M. Capretz[4]
Department of Electrical and Computer Engineering
Western University, London, Ontario, Canada, N6A 5B9
Email: [1]lhartma8@uwo.ca, [2]epigou@uwo.ca, [3]sgomezro@uwo.ca, [4]mcapretz@uwo.ca

*Github page:*
*https://github.com/LucasHartmanWestern/rl-for-vrp-csp*

***Abstract*—This paper presents a federated learning-based framework to tackle sustainable and dynamic electric vehicle routing, integrating multiple decision-making (DM) paradigms within a realistic simulation environment. Our environment models real-world conditions, including variable traffic patterns, seasonal temperature effects on battery performance, and authentic charging station distributions and constraints. In this setting, we compare the performance and environmental impact of reinforcement learning and transformer-based decision-makers, analyzing their ability to minimize travel distance, reduce peak charging station loads, and conserve energy. We use a federated learning approach which allows fine-grained per-vehicle DM, while aggregating experience at city, zone, and car model levels to accelerate learning and enhance scalability. By sharing policies among agents, the system adapts more rapidly to new conditions and achieves robust solutions that improve upon centralized methods. We evaluate the training overhead and estimate $CO_2$ emissions, demonstrating that federated learning can reduce environmental costs and training durations. Our findings offer actionable insights for electric vehicle fleet operators, urban planners, and policymakers, illustrating the promise of federated, data-driven DM methods in promoting environmentally responsible and grid-friendly electric vehicle routing.**

***Index Terms*—Electric vehicles, Reinforcement learning, Decision transformers, Federated learning, AI carbon footprint, Decision-making**

## I. Introduction

Sustainable artificial intelligence (AI) has emerged as an important research area of focus, as the environmental costs of training and deploying advanced decision-making (DM) models become increasingly evident. Recent studies have shown that state-of-the-art models can incur significant carbon footprints due to their intensive computational demands [1]. While several papers have explored methods for assessing AI's sustainability, there remains a gap in frameworks dedicated to evaluating the environmental impact of DM systems, particularly in realistic and dynamic settings. The dynamic setting explored in this work is the problem of routing fleets of electric vehicles (EVs) to charging stations along the way to their destinations, and evenly distributing the power demand incurred across the charging station network.

Electrical vehicle adoption continues to face obstacles that impede widespread acceptance. Battery performance degrades in cold weather [2], leading to shorter ranges and more frequent charging stops, and charging infrastructure remains inadequate [3]. Existing research often lacks models that integrate both transportation and power systems or account for the actual patterns of charging behavior. Overcoming these challenges is necessary for improving the accessibility and practicality of EVs, especially in colder climates and areas with sparse charging networks [2], [3].

Addressing the challenges of EV adoption requires the integration of strategic charging infrastructure planning with advanced vehicle routing methods to optimize both energy use and charger availability. The electric vehicle routing problem (EVRP) provides a structured framework for incorporating battery constraints, multiple charging options, and station placement into route planning. Efficient EV routing not only reduces travel and charging wait times [4], thereby improving user experience, but also plays a key role in balancing the load on the electrical grid, which is often less robust than that supporting internal combustion vehicles [5]. However, traditional routing algorithms typically focus on minimizing travel time without accounting for congestion at charging stations or the impact on grid stability [6], making them less effective in dynamic environments characterized by fluctuating traffic, variable charger availability, and shifting energy demands [7]. While advanced DM models can handle such complexities, their high computational cost introduces new sustainability concerns that remain underexplored [7].

Federated learning (FL) is an effective decentralized approach for training AI models with a privacy focus, especially in scenarios where data is distributed across multiple agents or devices [8]. FL allows individual EVs to locally compute routing decisions based on private or locally generated data without sharing raw data directly. By leveraging FL, the proposed framework enhances scalability, facilitates faster learning convergence, and supports adaptive DM tailored to individual vehicle characteristics and local environment conditions.

While numerous approaches have been proposed for EV routing and charge scheduling [6], none have measured the carbon emissions associated with different DM strategies in dynamic and realistic settings. To the best of our knowledge, no prior research has conducted a comparative analysis of carbon emissions resulting from well-known DM paradigms, including reinforcement learning (RL) and transformer-based approaches for EV routing problems.

This paper introduces a new framework for sustainable urban routing evaluation of decision-makers (SURE-DM). The framework is designed to measure the carbon emissions of DM models trained to tackle the EVRP within a simulation environment that reflects real-world conditions. The environment considers variable traffic patterns, weather conditions and their effect on battery degradation, charging station availability, and capacity. It also uses real-world data for the charging station locations and weather patterns. In this paper, two primary DM paradigms are applied to the same problem, and their effectiveness and carbon emissions are compared. This enables the selection of a sustainable approach that does not compromise performance.

By focusing on a realistic and dynamic system, unlike simplified or static models often used in previous studies [9], [10], this research advances the understanding of how different DM models affect carbon emissions for EV routing. The contributions of this work are as follows:

- A novel framework for sustainable urban routing evaluation of decision-makers (SURE-DM) is presented, which evaluates how DM models route EVs under realistic conditions. It focuses on their ability to adapt to changing traffic and charging station demands.
- To the best of our knowledge, SURE-DM is the first framework to measure and compare the computational demand of different DM models on a realistic and dynamic environment.
- SURE-DM integrates the EVRP and the charge scheduling problem (CSP) as one problem to solve. This allows for optimized route planning, ensuring efficiency.
- In addition, the SURE-DM framework avoids clustering EVs at centralized charging stations. This minimizes delays for EVs at charging stations, improves overall route efficiency and increases electrical grid stability.

The rest of this paper is organized as follows: Section II examines similar existing approaches. Section III describes the SURE-DM framework and its components. Section IV presents the experiments and results obtained while section V discusses their implications. Finally, Section VI concludes the paper.

## II. Related Work

Existing research on the EVRP often overlooks key factors such as the current load on a charging station's power supply and the traffic at charging stations, limiting its practical applicability [11]. Most studies addressing the EVRP focus primarily on minimizing the time an EV owner spends driving or reducing the distance traveled [11], without accounting for broader system-level impacts. Similarly, research addressing the CSP with EVs fails to explore strategies for balancing power loads across stations [12], such as routing vehicles to alternative locations to alleviate congestion. As mentioned by Abid *et al*. [6], the challenges of routing and charging EVs are deeply connected, but existing approaches treat them separately, within a manageable scope. The routing of vehicles is abstracted to the traveling salesmen problem [13], and the scheduling of EV charging is treated as a mixed-integer linear programming problem [14]. Machine learning techniques have been used to solve the EVRP and CSP separately [15], [16], [17], [18], but have not been applied to the combination. SURE-DM is a framework that evaluates the performance of DM models on the EVRP and CSP as it considers both the time spent traveling and the load placed on the power grid.

Park and Moon [16] used RL to deal with the scheduling and allocation of EV charging, but did not consider the routing of the EVs to the allotted charging stations. Adetunji *et al.* [15] proposed a deep deterministic policy gradient approach to distribute charging demand by adjusting incentives for EV drivers to encourage them to charge at less busy charging stations. The main downside of this approach is that it does not consider the increased distance or the time it takes to route the drivers to the desired charging stations. Wang *et al.* [19] used RL combined with FL to create an adaptive pricing strategy for power at charging stations, intending to alleviate traffic congestion and the load on the power grid. However, the simulation environment used by Wang *et al.* does not account for seasonal effects on demand, and the work also does not consider update-to-date approaches such as transformer-based models.

Federated reinforcement learning approaches have recently been applied to address EV charging control while preserving privacy. Qian *et al.* [20] proposed an algorithm to optimize EV charging strategies, however, their work focuses on optimizing grid stability and reducing the impact of EV charging on the distribution network, without considering factors such as the travel time to charging stations or the total route performance. Khatua *et al.* [21] used evolutionary computation with FL to route vehicles across cities but did not consider factors such as temperature, different EV models, and how the battery depletion rates may affect the decisions made. The proposed framework differs as it evaluates the DM models in a highly dynamic and realistic environment, specifically applying them to the EVRP and CSP.

There is a notable lack of comparison between the different types of DM models in prior research for handling the EVRP and CSP. Existing works typically focus on a single paradigm in isolation. Several studies apply RL to optimize EV routing or charging policies [12], [11], and transformer-based architectures have also been explored in this domain [22]. However, no prior work directly compares these different approaches on the same task. This work systematically compares multiple types of DM models (RL and transformer-based) within a unified experimental setting, evaluating their relative strengths and weaknesses under identical conditions.

While EV routing systems aim to reduce vehicular emissions, training advanced models can be computationally intensive and thus carbon-intensive. Training state-of-the-art AI models has been shown to consume large amounts of energy and produce significant $CO_2$ emissions [23]. Recent work has advocated for more environmentally sustainable AI and that new algorithms should be evaluated not only on performance, but also on their carbon footprint [24]. However, none of the existing EV routing or charging optimization studies report or compare the carbon emissions resulting from training their models. The SURE-DM framework provides an emissions-aware evaluation of the proposed DM methods. It measures and compares the energy usage and $CO_2$-equivalent emissions

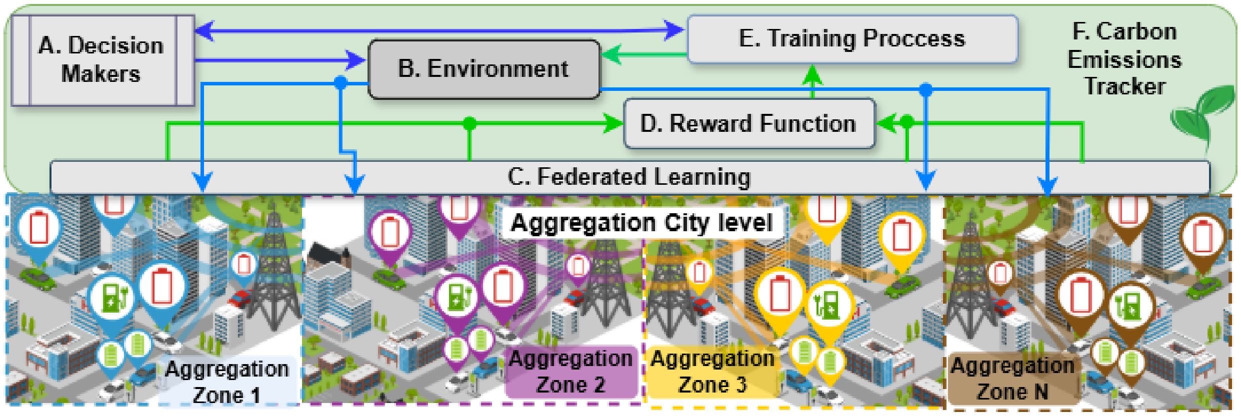



Fig. 1. SURE-DM framework illustrating the setup, where the nearest communication tower assigns a decision-maker tailored to each EV's route to compute the optimal path.

of training each DM model, highlighting the sustainability implications of deploying different EV routing strategies.

## III. Methodology

This section presents the sustainable urban routing evaluation of decision-makers (SURE-DM) framework, designed to assess the carbon footprint and computational performance of different AI-based DM methods in the routing and charging optimization of EVs. The framework operates within a realistic and dynamic simulation environment that considers key factors such as traffic conditions at charging stations (which increase waiting duration), travel distance, EV energy consumption, and scalability across different EV fleet sizes. For this work, EV fleets are considered a group of EVs with different characteristics sharing the same zone location.

To evaluate AI-driven DM methods, SURE-DM incorporates multiple methods, including deep Q-networks (DQN), REINFORCE, and online decision transformers (ODT). These methods are compared in terms of their computational requirements, efficiency in finding optimal charging routes, and impact on sustainability metrics.

The SURE-DM framework operates within a simulation environment designed to address the EVRP and CSP, incorporating several novel features to improve simulation fidelity. These improvements include seasonal battery depletion modeling, diverse EV energy consumption patterns, and computational complexity optimizations that enhance the performance of large-scale simulations. Another key enhancement in SURE-DM is the integration of federated learning (FL), which enables distributed agents to train DM models while preserving data privacy collaboratively. This approach accelerates the learning process, reduces the carbon footprint during training, and ensures secure data handling.

SURE-DM is structured into six main processes to systematically analyze the framework and its components, as shown in Figure 1. The DM methods are discussed in Subsection III-A. The simulation environment is described in Subsection III-B, detailing the real-world conditions modeled in the study. Subsection III-C presents how federated learning is integrated into the framework. Subsection III-D defines the objective function used for evaluation. Subsection III-E outlines the training process for DM agents, explaining how they interact with the environment. Finally, Subsection III-F explains the carbon emissions tracker.

### A. Decision-makers

The SURE-DM framework evaluates DM methods across two major paradigms: reinforcement learning (RL) and transformer-based models. These paradigms employ distinct learning and DM strategies, enabling a comparative analysis of their effectiveness in optimizing EVs charging-routing problems while monitoring computational expense.

To illustrate these paradigms, three representative methods were selected for implementation in the framework: Deep Q-Networks (DQN), REINFORCE, and Online Decision Transformers (ODT). DQN and REINFORCE are well-established RL-based approaches that exemplify value-based and policy-based learning, respectively. ODT represents the transformer-based paradigm, offering strong integration with RL principles by leveraging trajectory data from offline environments. These methods were chosen for their clarity in representing the underlying learning paradigms and their relevance to sequential DM tasks in EV routing contexts.

*a) Deep Q-networks (DQN) [25]:* is an RL method that uses deep learning to approximate Q-values, enabling agents to select actions that maximize expected rewards over time.

*b) Policy gradient techniques [26]:* are implemented using the REINFORCE algorithm, which is an effective method, particularly suitable for continuous action spaces.

*c) Online Decision Transformers (ODT) [27]:* uses self-attention to model temporal dependencies and predict future actions based on past trajectories, eliminating the need for explicit reward signals.

The SURE-DM framework is designed to accommodate a wide variety of DM methods. While this study focuses on three representative methods to demonstrate the framework's capabilities, it remains extensible and can incorporate additional approaches as needed. The next subsection outlines the training process used to optimize these techniques within the simulation environment.

### B. The Environment

The SURE-DM environment builds upon the existing multi-factor edge-weighting with reinforcement learning [10] approach, incorporating several enhancements to improve its realism and accuracy in modeling EV routing and charging optimization. These enhancements include the integration of temperature data to model seasonal battery behavior, the introduction of diverse EV models and trims with distinct energy consumption profiles, and a scalable architecture optimized for GPU processing.

The methodology models charging stations, route starting points, and destinations as a graph structure, where groups of EVs traverse from their origins to destinations, stopping at charging stations as needed. Each EV is assigned to a zone, simulating real-world geographical constraints based on its start and end coordinates. The environment dynamically transforms an unweighted graph into a weighted one by computing edge weights based on multiple factors, including traffic congestion, charging station availability, and energy performance. These weights influence the routing decisions made by DM agents, which are specific instances of the DM models that optimize travel paths using a path-finding algorithm.

To enhance realism, SURE-DM introduces three key improvements:

*a) Temperature-Based Seasonal Battery Behavior:* Temperature has a significant impact on the performance of EV batteries, influencing depletion rates and route feasibility. To model this effect, SURE-DM integrates historical temperature data from the nearest weather stations to each zone. This data is used to dynamically adjust battery performance based on real-world observations of temperature-dependent range fluctuations, as studied by Senol *et al.* [2]. Colder temperatures lead to faster battery depletion, making some routes viable in one season but inefficient in another. The battery performance adjustment is performed using Eq. 1, ensuring that EV performance accurately reflects temperature variations.

$$\text{eff}(temp) = 0.75 + 0.25 \tanh\left(\frac{temp - 12.5}{6.25}\right) \tag{1}$$

*b) Diverse EV Models and Trims:* Due to differences in battery capacity, seasonal performance, and aerodynamics, different EV models exhibit varying energy consumption patterns. SURE-DM incorporates a range of EV models with distinct energy consumption profiles, allowing for a more realistic simulation of fleet-wide energy usage. By considering vehicle-specific consumption characteristics, the framework provides a more granular evaluation of routing strategies, helping assess the impact of DM methods across various vehicle types.

*c) Scalability with GPU Optimization:* To handle large-scale EV routing scenarios efficiently, SURE-DM restructures the simulation to run on a graphical processing unit (GPU). The environment's state and operations are represented using matrices, allowing for parallel computations and simultaneous EV movements. This optimization significantly reduces computational complexity, improving simulation performance. Specifically, movement operations now achieve O(1) time complexity, compared to the O(n) complexity in the original environment, enabling the efficient simulation of thousands of EVs.

### C. Federated Learning

FL is introduced to improve training time while preserving data privacy. The learning process is accelerated by grouping agents based on shared attributes, such as geographical zone or EV model, and periodically aggregating their policies. Aggregation occurs at different levels: global (all agents in the environment), city (agents in the same city), zone (agents in the same zone), and EV model (agents managing the same type of electric vehicle). The updated local policy is computed as:

$$\pi_i = \sigma_c \pi_i^c + \sigma_z \pi_i^z + \sigma_m \pi_i^m, \tag{2}$$

where $\pi_i$ is the updated local policy, $\pi_i^c$, $\pi_i^z$, and $\pi_i^m$ represent the city-level, zone-level, and EV-model-level policies relevant to agent $i$, respectively. The weights $\sigma_c$, $\sigma_z$ and $\sigma_m$ determine the relative importance of each policy and $\sigma_c + \sigma_z + \sigma_m = 1$.

When a new agent joins the network, it initializes with a base policy from global and semi-global policies, reducing adaptation time and minimizing retraining. This approach improves scalability and coordination, ensuring agents learn strategies that enhance the system's performance.

### D. The Objective Function

The SURE-DM framework trains DM agents using an objective function that minimizes travel distance, traffic congestion at charging stations, and energy consumption. At each time step, a reward is computed based on these factors and returned to the agents. The accumulated reward calculated per episode $ep$ guides agents toward optimizing their routing decisions over time.

The reward function $r^z_{j,t}$, used to train each vehicle $v^z_j$ in zone $z$ with $N^z$ vehicles at time step $t$, is defined as:

$$r^z_{j,t} = \alpha_d d^z_{j,t} + \alpha_T \max_c T^z_{c,t} + \alpha_e e^z_{j,t}, \tag{3}$$

where:

- $d^z_{j,t} = \left|\left|\psi(v^z_{j,t}) - \psi(v^z_{j,t-1})\right|\right|_2$ is the Euclidean distance traveled by $v^z_j$,
- $e^z_{j,t} = \left|\beta(v^z_{j,t}) - \beta(v^z_{j,t-1})\right|$ is the energy consumed by $v^z_j$,
- $T_{c,t}$ represents the traffic level at charging station $c$,

$\alpha_d$, $\alpha_T$, and $\alpha_e$ are normalization weights that balance the influence of each factor. $\psi(v^z_{j,t})$ represents the position of $v^z_j$, and $\beta(v^z_{j,t})$ denotes the battery level at time step $t$.

The reward function aligns training objectives with real-world goals, ensuring that agents minimize travel distance, avoid peak traffic at charging stations, and optimize energy usage.

**Algorithm 1:** SURE-DM Training Process

**Output:** Aggregated Policies $(\pi^c, \pi^z, \pi^m)$, $CO_2 \leftarrow$ Emission Logs
**Input:** $E \leftarrow max_{episodes}$, $S \leftarrow max_{steps}$, $AGG \leftarrow max_{aggregations}$, $EV \leftarrow numEV$
**Training:** TRAIN($max_{episodes}$, $max_{steps}$, $max_{aggregations}$, $numEV$)

`/* Step 1) Initialize decision makers */`
**for** $i \leftarrow 1$ **to** $AGG$ **do**
  **for** $j \leftarrow 1$ **to** $EV$ **do**
    $DM_j \leftarrow getLocalPolicy(\pi^c, \pi^z, \pi^m)$
  `/* Step 2) Train over episodes */`
  **for** $j \leftarrow 1$ **to** $E$ **do**
    $state[1..EV] \leftarrow$ `ResetEnv()`
    **for** $k \leftarrow 1$ **to** $S$ **do**
      **for** $l \leftarrow 1$ **to** $EV$ **do**
        $action \leftarrow DM_l(state)$
        $[state', reward] \leftarrow$ `StepEnv`($action[1..EV]$)
        $DM_l$.`learn`($state_l$, $action_l$, $reward_l$, $state'_l$)
        $state \leftarrow state'$
      **if** *all EVs at destination* **then**
        **break**
    $CO_2 \leftarrow$ collect emission data
  `/* Step 3) Aggregate updated policies */`
  $\pi^c, \pi^z, \pi^m \leftarrow aggregatePolicies(DM)$
**return** $\{\pi^c, \pi^z, \pi^m, CO_2\}$

### *E. The Training Process*

The SURE-DM framework trains DM agents either by zone (transformer-based) or per-car (RL-based) basis by iteratively adjusting weights on an initially unweighted graph representing the environment. Transformer-based approaches were configured to operate with a single agent controlling all cars within a zone, as loading individual transformer agents for each car in a zone during training is too computationally demanding to be feasible. The agents simulate EV movement using the updated weighted graph and refine their strategies based on the rewards obtained during each simulation episode. Training occurs in episodes, where all EVs travel from their unique starting locations to their destinations. Each episode consists of multiple time steps, and agents can update their previously assigned weights at each time step. Each time step represents a fixed interval in the environment, during which EVs may be in motion, charging at a station, or both.

Algorithm 1 outlines the high-level training process used by the SURE-DM framework. It starts with the DM agent generating an action based on the current state of the environment. This action modifies the environment, influencing EV routing decisions. The environment then evaluates the resulting routes and provides feedback as a reward, which helps refine the agent's DM process. Carbon emissions are tracked throughout the training process by measuring the per-episode power consumption of the hardware used for training and running the simulation. This tracking allows for a dual evaluation, assessing the effectiveness of each DM method at maximizing the reward function and the environmental impact of its computational consumption.

### *F. Carbon Emissions Tracking*

The SURE-DM framework evaluates the sustainability of each DM method by monitoring the duration of training and the power consumption of the hardware used. The power drawn by the GPU and CPU is tracked at regular intervals, and this data is used to estimate the $CO_2$ emissions generated during training. By incorporating power consumption into the evaluation, the framework enables a direct comparison of DM methods in terms of performance and their environmental impact. The total $CO_2$ emissions are estimated using:

$$CO_2 = \left( \sum_{i=1}^{K} (P_{\text{CPU},i} + P_{\text{GPU},i}) \Delta h \right) CI, \tag{4}$$

where $K$ represents the total number of time intervals in the training process, $P_{\text{CPU},i}$ and $P_{\text{GPU},i}$ are the power consumption of the CPU and GPU in Watts at the $i$-th interval, $\Delta h$ is the duration of each interval in hours, and $CI$ is the carbon intensity of the energy source in grams of $CO_2$ per watt-hour.

This tracking method allows for a dual evaluation of DM methods by assessing their performance alongside their energy consumption and emissions. The carbon footprint of each method is influenced by factors such as training complexity, convergence speed, and hardware efficiency. Faster training processes reduce total power usage, resulting in lower emissions, whereas methods that require extensive computation may have higher environmental costs.

By integrating carbon footprint tracking into the evaluation, the SURE-DM framework provides insights into the trade-offs between computational performance and sustainability, supporting the development of more environmentally friendly DM methods.

## IV. Experiments and Results

This section presents the experiments conducted to evaluate various aspects of the proposed SURE-DM framework and their corresponding results. Subsection IV-A outlines the experimental setup, including the configuration and parameters used for testing. Subections IV-B-IV-D outline the various experiments that were ran to evaluate the SURE-DM framework. The experiments focus on analyzing the performance and behavior of DM agents under different scenarios and configurations.

### *A. Experimental Setup*

The experiments are structured as follows: The first experiment evaluates how aggregation steps in FL affect agent performance, optimizing the balance between episodes and aggregation frequency. The second experiment examines DM performance under various seasonal conditions. Finally, the third experiment examines carbon emissions during training for the various DM methods.

Each experiment was a combination of DM model, seed, and season. For reproducibility purposes, nine seeds were used for all experiments. Additionally, all experiments were conducted across four different seasons. The seasons influenced the temperature of the environment, and historical data

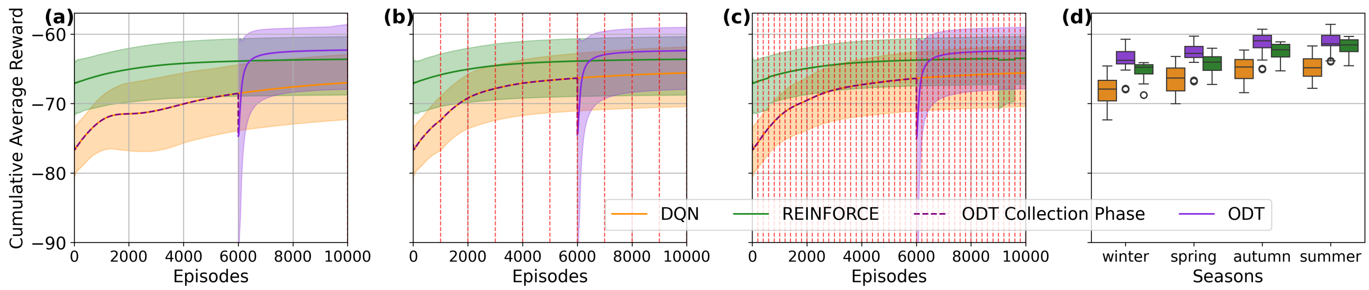



Fig. 2. Training and rewards per decision-makers method over aggregation variation. (a) No aggregations, (b) 1000 episodes per aggregation, (c) 200 episodes per aggregation, (d) seasonality for the last 100 episodes as steady behavior.

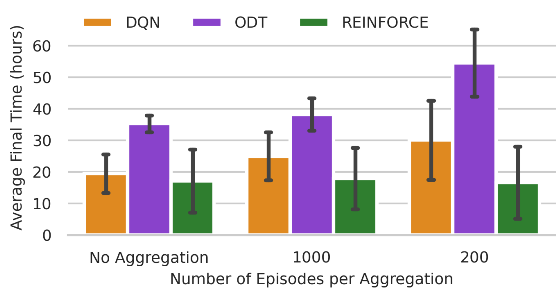



Fig. 3. Average time to train for each aggregation frequency.

TABLE I

SIMULATION ENVIRONMENT MEAN METRICS BY DECISION-MAKER AND SEASON

| Algorithm | Season | Distance [km] | Peak Traffic [No. of EVs] | Energy [kWh] |
|---|---|---|---|---|
| DQN | Winter | 37.14 ± 1.55 | 26 | 15.83 ± 3.29 |
| | Autumn | 37.11 ± 1.52 | 26 | 21.33 ± 3.10 |
| | Spring | 37.12 ± 1.52 | 26 | 17.57 ± 3.16 |
| | Summer | 37.11 ± 1.52 | 26 | 22.79 ± 3.08 |
| ODT | Winter | 36.90 ± 1.23 | 26 | **15.29 ± 2.99** |
| | Autumn | 36.88 ± 1.19 | 25 | 21.20 ± 3.01 |
| | Spring | 36.88 ± 1.19 | 26 | 18.52 ± 3.70 |
| | Summer | **36.88 ± 1.18** | 25 | 22.54 ± 2.93 |
| REINFORCE | Winter | 36.73 ± 0.99 | 24 | 15.37 ± 3.03 |
| | Autumn | 36.70 ± 0.92 | 25 | 21.20 ± 3.01 |
| | Spring | 36.71 ± 0.93 | 25 | 17.57 ± 3.15 |
| | Summer | 36.70 ± 0.91 | **23** | 22.54 ± 2.93 |

from weatherstats.ca [28] were used to determine the average seasonal temperature of the location. Real data from the National Renewable Energy Laboratory [29] was also used to place the location of each charging station and set their output rates. There are three types of EV models considered in these experiments, which were most popular models from the EV brands with the highest global sales [30] (Tesla and BYD). The type of EV model affects the depletion rates of the car battery, as well as the maximum capacity of the battery, which were both set in the environment based on the manufacturers' estimates.

### *B. Experiment 1. Convergence and Episodes per Aggregation*

This experiment was selected to analyze how an aggregation policy in FL can accelerate the plateau searching, and with that, also save computational resources for the different DM methods.

Experiment 1 evaluates the training performance of the DM methods under varying aggregation frequencies. Three configurations were tested: no aggregation (Fig. 2.a), aggregation every 1000 episodes (Fig. 2.b), and aggregation every 200 episodes (Fig. 2.c). Figure 2 presents the range (minimum, maximum) and average reward values across nine seeds for each DM model configuration. To visualize the behavior of the DM methods across different seasons, Fig. 2.d presents the seasonality patterns over the final 100 episodes, providing insight into how each DM model responds to recurring environmental changes during the later stages of training. The corresponding training durations for each aggregation frequency are shown in Fig. 3. Results indicate that aggregation intervals of 1000 and 200 episodes both facilitate faster convergence (i.e., earlier reward plateauing) compared to no aggregation. However, the 200-episode configuration incurs a higher training time. Consequently, an aggregation interval of 1000 episodes is adopted in the subsequent experiments to balance performance and training time.

### *C. Experiment 2. Decision-Maker Evaluation*

Experiment 2 evaluates every combination of DM model, seed, and season. Table I compares the reward, distance traveled, traffic level, and energy consumed in the simulation environment during training. The distance metric measures the distance traveled by the EVs in the simulation environment, the traffic metric measures the number of EVs at charging stations in the simulation environment, and the energy metric measures the energy consumed by the EVs in the simulation environment. In Table I, the ± symbol indicates an standard deviation of the mean.

These results show that REINFORCE was the best at minimizing peaks in traffic levels, but ODT was the overall best performing model as it was only marginally worse than REINFORCE at minimizing traffic peaks but better at minimizing distance traveled and energy used.

### *D. Experiment 3. Computational Resources*

This experiment was selected to compare the computational resources used by each DM agent. This effectively examines the environmental sustainability of training each DM agent which is an important factor to consider when deploying any machine learning approach at scale.

TABLE II
TRAINING TIME, CARBON EMISSIONS, AND PERFORMANCE BY ALGORITHM

| **Algorithm** | **Duration** [h] | **Emissions** [g of $CO_2$] | **Reward** |
|---|---|---|---|
| DQN | 24.17 ± 7.41 | 0.44 ± 0.13 | -64.13 ± 6.32 |
| ODT | 36.05 ± 0.94 | 2.24 ± 1.26 | **-62.22 ± 4.80** |
| REINFORCE | **20.32 ± 5.78** | **0.35 ± 0.12** | -63.35 ± 4.04 |

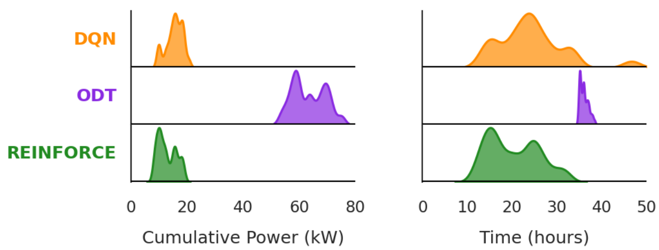



Fig. 4. Power and training duration by DM method.

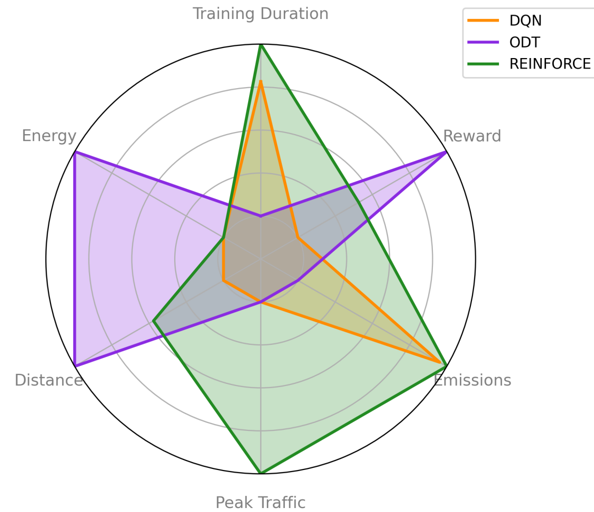



Fig. 5. Performance profiles of decision-makers across key attributes. Values closer to the outer edge of the chart indicate better performance for the corresponding attribute.

Experiment 3 measures the $CO_2$ emissions from the simulations done in Experiment 2. Table II shows how long it took to train each algorithm (duration), the estimated carbon emissions during the training of each algorithm (emissions), and the average reward on the last episode. Figure 4 shows the cumulative power used over time during the training for each DM agent.

This experiment demonstrates ODT's large computational resource demand compared to the leaner DQN and REINFORCE DM models.

## V. DISCUSSION

The experimental results demonstrated that when using FL, finding the right balance of aggregation frequency can significantly reduce the time it takes for the model to plateau without decreasing the reward level to which the model plateaus. Tuning the aggregation frequency also reduced the average time to train the model by up to 30-40% when using ODT and DQN. This reduction in training time directly reduces computation demand, as the hardware resources spend less time training the models, resulting in a more sustainable implementation.

Finally, these results show a significant difference in computational sustainability between the different DM methods. Transformer-based models require more extensive computational resources and generate higher carbon emissions compared to the leaner methods like DQN and REINFORCE. While ODT was the most computationally intensive, it also performed the best, with DQN being the worst performer, and REINFORCE being the most sustainable approach.

Figure 5 compares the relative performance of the DM agents across the 6 primary attributes. This figure shows that while ODT reaches the highest reward, REINFORCE achieves the best balance across all 3 DM methods overall.

## VI. CONCLUSION

This paper presented a FL-based framework for dynamic electric vehicle routing that aims to reduce stress on urban grids and promote more sustainable transportation. By modeling realistic conditions such as varying traffic patterns, weather effects on batteries, and actual charging station locations, this framework supports decision-makers in guiding vehicles toward routes that save energy and minimize peak loads at charging stations.

Our experiments showed that FL helps improve performance while reducing the time and computational power needed for training. This leads to better routing decisions and lower environmental costs associated with model development. The results suggest that the approach can adapt to new conditions, making it suitable for large-scale use in changing urban environments.

In the future, this work can be extended to include more factors, like fluctuating electricity prices and evolving grid conditions. These additions could help decision-makers create even more robust and sustainable routing strategies. The framework offers a promising pathway for electric vehicle fleet operators, city planners, and policymakers to collaborate in building greener transportation systems.

## VII. ACKNOWLEDGMENTS

This research has been funded by the NSERC Alliance Missions grant: ALLRP 577133-2022 and supported by the Vector Scholarship in Artificial Intelligence, provided through the Vector Institute. Additionally, this work is supported by the Canadian Graduate Scholarship (Masters), awarded by the Government of Canada. S.G.-R. thanks the SENESCYT from Ecuador for the support received during his doctoral studies in Canada.

## REFERENCES

[1] R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, "Green AI," *Communications of the ACM*, vol. 63, no. 12, pp. 54–63, 11 2020.

[2] M. Senol, I. S. Bayram, Y. Naderi, and S. Galloway, "Electric Vehicles Under Low Temperatures: A Review on Battery Performance, Charging Needs, and Power Grid Impacts," *IEEE Access*, vol. 11, pp. 39 879–39 912, 2023.

[3] T. Unterluggauer, J. Rich, P. B. Andersen, and S. Hashemi, "Electric vehicle charging infrastructure planning for integrated transportation and power distribution networks: A review," *eTransportation*, vol. 12, p. 100163, 5 2022.

[4] C. Lv, Y. Liu, X. Hu, H. Guo, D. Cao, and F. Y. Wang, "Simultaneous Observation of Hybrid States for Cyber-Physical Systems: A Case Study of Electric Vehicle Powertrain," *IEEE Transactions on Cybernetics*, vol. 48, no. 8, pp. 2357–2367, 8 2018.

[5] F. Razi and I. Dincer, "A review of the current state, challenges, opportunities and future directions for implementation of sustainable electric vehicle infrastructure in Canada," *Journal of Energy Storage*, vol. 56, p. 106048, 12 2022.

[6] M. Abid, M. Tabaa, A. Chakir, and H. Hachimi, "Routing and charging of electric vehicles: Literature review," *Energy Reports*, vol. 8, pp. 556–578, 11 2022.

[7] S. Padakandla, "A Survey of Reinforcement Learning Algorithms for Dynamically Varying Environments," *ACM Computing Surveys*, vol. 54, no. 6, pp. 1–25, 7 2022.

[8] D. Namakshenas, A. Yazdinejad, A. Dehghantanha, R. M. Parizi, and G. Srivastava, "IP2FL: Interpretation-Based Privacy-Preserving Federated Learning for Industrial Cyber-Physical Systems," *IEEE Transactions on Industrial Cyber-Physical Systems*, vol. 2, pp. 321–330, 7 2024.

[9] B. Lin, B. Ghaddar, and J. Nathwani, "Deep Reinforcement Learning for the Electric Vehicle Routing Problem With Time Windows," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 11 528–11 538, 8 2022.

[10] L. Hartman, S. Gomez-Rosero, P. Adjei, and M. A. M. Capretz, "Multi-Factor Edge-Weighting with Reinforcement Learning for Load Balancing of Electric Vehicle Charging Stations," in *2023 International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 12 2023, pp. 580–587.

[11] X. A. Martin, M. Escoto, A. Guerrero, and A. A. Juan, "Battery Management in Electric Vehicle Routing Problems: A Review," *Energies*, vol. 17, no. 5, p. 1141, 2 2024.

[12] Z. Zhao, C. K. Lee, X. Yan, and H. Wang, "Reinforcement learning for electric vehicle charging scheduling: A systematic review," *Transportation Research Part E: Logistics and Transportation Review*, vol. 190, p. 103698, 10 2024.

[13] J. D. C. Little, K. G. Murty, D. W. Sweeney, and C. Karel, "An Algorithm for the Traveling Salesman Problem," *Operations Research*, vol. 11, no. 6, pp. 972–989, 12 1963.

[14] J. P. Vielma, "Mixed Integer Linear Programming Formulation Techniques," *SIAM Review*, vol. 57, no. 1, pp. 3–57, 1 2015.

[15] K. E. Adetunji, I. W. Hofsajer, A. M. Abu-Mahfouz, and L. Cheng, "A Two-Tailed Pricing Scheme for Optimal EV Charging Scheduling Using Multiobjective Reinforcement Learning," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 3361–3370, 3 2024.

[16] K. Park and I. Moon, "Multi-agent deep reinforcement learning approach for EV charging scheduling in a smart grid," *Applied Energy*, vol. 328, p. 120111, 12 2022.

[17] M. Wang, Y. Wei, X. Huang, and S. Gao, "An End-to-End Deep Reinforcement Learning Framework for Electric Vehicle Routing Problem," *IEEE Internet of Things Journal*, vol. 11, no. 20, pp. 33 671–33 682, 10 2024.

[18] C. Wang, M. Cao, H. Jiang, X. Xiang, and X. Zhang, "A Deep Reinforcement Learning-Based Adaptive Large Neighborhood Search for Capacitated Electric Vehicle Routing Problems," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 9, no. 1, pp. 131–144, 2 2025.

[19] Y. Wang, Q. Guo, and Y. Yu, "Cooperative Pricing for Vehicle-Traffic-Power Nexus to Support City Decarbonization: A Federated Deep Reinforcement Learning Framework," *IEEE Transactions on Industrial Cyber-Physical Systems*, vol. 1, pp. 182–191, 9 2023.

[20] J. Qian, Y. Jiang, X. Liu, Q. Wang, T. Wang, Y. Shi, and W. Chen, "Federated Reinforcement Learning for Electric Vehicles Charging Control on Distribution Networks," *IEEE Internet of Things Journal*, vol. 11, no. 3, pp. 5511–5525, 2 2024.

[21] S. Khatua, D. De, S. Maji, S. Maity, and I. E. Nielsen, "A federated learning model for integrating sustainable routing with the Internet of Vehicular Things using genetic algorithm," *Decision Analytics Journal*, vol. 11, p. 100486, 6 2024.

[22] Z. Feng, J. Zhang, H. Jiang, X. Yao, Y. Qian, and H. Zhang, "Energy consumption prediction strategy for electric vehicle based on LSTM-transformer framework," *Energy*, vol. 302, p. 131780, 9 2024.

[23] P. Delanoë, D. Tchuente, and G. Colin, "Method and evaluations of the effective gain of artificial intelligence models for reducing CO2 emissions," *Journal of Environmental Management*, vol. 331, p. 117261, 4 2023.

[24] V. Bolón-Canedo, L. Morán-Fernández, B. Cancela, and A. Alonso-Betanzos, "A review of green artificial intelligence: Towards a more sustainable future," *Neurocomputing*, vol. 599, p. 128096, 9 2024.

[25] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2 2015.

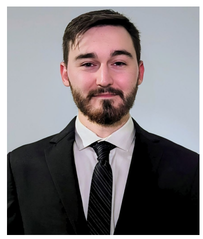
[26] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Machine Learning*, vol. 8, no. 3-4, pp. 229–256, 5 1992.

[27] Q. Zheng, A. Zhang, and A. Grover, "Online Decision Transformer," pp. 27 042–27 059, 6 2022.

[28] WeatherStats, "Weather Data for Canada." [Online]. Available: https://www.weatherstats.ca

[29] NREL: Developer Network, "Alternative Fuel Stations API," 2023. [Online]. Available: https://developer.nrel.gov/docs/transportation/alt-fuel-stations-v1/all/

[30] G. A. Samawi, O. M. Bwaliez, M. Jreissat, and A. Kandas, "Advancing Sustainable Development in Jordan: A Business and Economic Analysis of Electric Vehicle Adoption in the Transportation Sector," *World Electric Vehicle Journal*, vol. 16, no. 1, p. 45, 1 2025.

**Lucas Hartman** is a graduate student in the Department of Electrical and Computer Engineering at Western University, Canada, where he also received his B.E.Sc. in 2024. Before graduate school, he worked as an AI Engineer at Optimy.ai, and during graduate school he works as a research intern at the Vector Institute for Artificial Intelligence. His experience is in software engineering and applied AI. His research interests include reinforcement learning applications, computer vision, and generative AI.

**Ethan Pigou** is a graduate student in the Department of Electrical and Computer Engineering at Western University, Canada, where he received his B.E.Sc in 2024. Prior to graduate school, Ethan worked within Bell Canada's leadership program where he explored machine learning applications for improved customer retention rates. His experience resides in software engineering and applied AI. His research interests include reinforcement learning applications, transformer architectures, and efficient AI.

**Santiago Gomez-Rosero** is currently working as a postdoctoral associate at the Department of Electrical and Computer Engineering at Western University, Canada. He obtained his Ph.D. at Western University in 2022. Before joining Western University, he taught at the International University SEK, Ecuador. He received his B.Sc. from the Escuela de las Fuerzas Armadas (ESPE), Ecuador, and his M.Sc. from the Federal University of Rio de Janeiro, Brazil. His experience is in electronic and software engineering. His research interests include machine learning applications, neural architecture search, neuroevolution, and IoT.

**Miriam A. M. Capretz** is a professor in the Department of Electrical and Computer Engineering at Western University, Canada. Before joining Western University, she was with the University of Aizu, Japan. She received her B.Sc. and M.E.Sc. degrees from UNICAMP, Brazil, and her Ph.D. from the University of Durham, UK. She has been working in the software engineering area for more than 35 years and has been involved with the organization of workshops and symposia and has been serving on program committees at international conferences. Her current research interests include machine learning, big data analytics, IoT, blockchain and cloud computing.