Literature Review

The rapid adoption of electric vehicles (EVs) is transforming the transport and energy sectors, driven by ambitious decarbonisation targets and air quality policies. Cities such as London are expanding charging networks to meet rising demand, yet urban charging remains characterised by fragmented operators, uneven charger distribution, and complex tariffs that combine per-kWh fees, session costs, idle penalties, and membership discounts [9]. For users, the immediate concerns are minimising charging costs, reducing delays, and ensuring route feasibility, while operators and regulators focus on infrastructure adequacy and grid stability. Traditional optimisation approaches have primarily addressed charging from a grid perspective, using techniques such as adaptive robust optimisation [2] and reactive power management [3] to stabilise distribution networks. Although effective for system reliability, these studies do not reflect user-side constraints, including travel times, heterogeneous tariffs, and individual charging preferences [4].

This gap has motivated the exploration of methods that jointly consider mobility and cost in dynamic settings. Reinforcement learning (RL) has shown promise in such contexts, offering adaptability to uncertainty and the ability to learn policies through interaction with complex environments [5, 6]. Deep Q-Networks (DQN) have been applied to charging scheduling [5], but their discrete action space limits applicability. Proximal Policy Optimisation (PPO), by contrast, is well suited to continuous decisions and hybrid objectives [7]. Building on this potential, the present study applies PPO to optimise cost-based, time-based, and hybrid charging strategies in Inner London, explicitly incorporating real-world tariffs, nonlinear charging curves, and operator heterogeneity.

The effectiveness of EV adoption depends not only on technology but also on accessibility, affordability, and the spatial distribution of charging infrastructure. Research frames the charging problem across several dimensions: siting and sizing of chargers, operational management, user behaviour, and tariff structures [8]. In the UK, infrastructure deployment has been uneven, with significant "infrastructure gaps" between adoption projections and available charging capacity [9]. These gaps are most acute in metropolitan areas such as London, where high demand clustering and limited space for station expansion exacerbate congestion [10]. User behaviour introduces further uncertainty. Empirical studies show clustering of UK charging sessions in evening hours, reflecting commuting patterns and household dependence [4]. Household charging is shaped by daily schedules and time-of-use tariffs, while public charging behaviour reflects range anxiety and limited access to off-street charging [11]. Tariff structures add complexity: operators impose layered cost schemes including session fees, idle penalties, and membership discounts, with some introducing time-of-day variability to incentivise off-peak charging [12]. Comparative studies also highlight how tariff design affects EV adoption and charging flexibility [13]. UK field experiments demonstrate strong price elasticity: for example, a 40% discount

increased public charging volumes by 117% during single sessions [14]. Game-theoretic and mean-field models capture how pricing affects congestion and station selection [15, 16, 27], yet most fail to model the fine-grained operator diversity of London. For Inner London specifically, tariff fragmentation is pronounced, with multiple operators (e.g., Source London, bp pulse, Shell Recharge) applying divergent pricing schemes [10].

Technical constraints from vehicle charging curves further complicate optimisation. Battery charging is nonlinear: power is high at lower states of charge (SoC) but tapers significantly as SoC approaches full capacity, meaning marginal charging times increase disproportionately. Some studies have incorporated this tapering into optimisation models [17, 18], yet many still assume constant charging rates, leading to underestimation of charging duration. Combined with tariff heterogeneity, this creates cost–time trade-offs: users may minimise cost by choosing slower, cheaper stations at the expense of longer travel or waiting times, while those prioritising time gravitate towards ultra-rapid but more expensive chargers. These challenges illustrate why optimisation cannot focus solely on grid stability; any practical framework must reflect user behaviour, tariff fragmentation, and charging dynamics in dense urban settings such as London.

Before reinforcement learning gained prominence, EV charging optimisation was addressed through deterministic and heuristic approaches. Mixed-integer linear programming (MILP) and linear programming (LP) have been widely applied to siting, scheduling, and grid integration. These formulations are valued for transparency and their ability to guarantee optimal solutions under fixed assumptions [19, 20]. However, their computational burden scales poorly with the number of vehicles and stations, making them impractical for real-time urban applications. To address scalability, metaheuristic techniques such as genetic algorithms (GA), particle swarm optimisation (PSO), and ant colony optimisation (ACO) have been employed, as well as multi-criteria decision-making (MCDM) frameworks that balance grid stability, accessibility, and cost [21]. While these methods are more flexible, they lack adaptability to dynamic environments where tariffs, demand, and mobility patterns fluctuate rapidly. A related strand of research focuses on the electric vehicle routing problem (EVRP), which extends traditional vehicle routing by incorporating charging constraints. Early EVRP formulations treated stations as deterministic service nodes, overlooking congestion, waiting times, and nonlinear charging curves [18]. These limitations have motivated the move towards adaptive, data-driven approaches, with reinforcement learning offering a compelling alternative.

EV charging optimisation must therefore account for the combined influence of vehicle dynamics, tariffs, and user behaviour. The nonlinearity of charging curves complicates cost and route forecasting [17]. Tariff heterogeneity adds further complexity, with operators applying per-kWh rates, session fees, idle penalties, and membership discounts, sometimes coupled with time-of-day variability [12]. Game-theoretic and mean-field approaches highlight how pricing affects congestion and user choice [15, 16]. Meanwhile, behavioural studies show clustering of demand in evening hours [4], alongside tariff sensitivity and

range anxiety [11]. Together, these factors make balancing cost and time objectives difficult: slower, cheaper charging often increases waiting or travel times, while prioritising speed incurs higher costs. Few frameworks explicitly model this trade-off in real-world, multi-operator urban contexts.

RL offers significant advantages in this setting. Unlike deterministic or heuristic methods, RL allows agents to learn adaptive policies by interacting with dynamic environments. Applications in energy systems demonstrate RL's ability to reduce costs, flatten grid loads, and improve user satisfaction [22, 23]. Early EV charging studies used DQN, which showed success in cost reduction and scheduling efficiency [5, 6], though discrete action spaces limited their suitability for continuous charging problems. Policy gradient methods, such as PPO, address this limitation by directly optimising policies for continuous action spaces [7]. PPO has gained traction for its stability, sample efficiency, and suitability for multi-objective problems, and has outperformed value-based methods in scalability and convergence [24]. Extensions such as DDPG and SAC have also been explored for EV charging, offering further flexibility in continuous domains [25]. Despite these advances, few studies apply PPO at the driver level in dense cities, where tariffs, nonlinear charging curves, and congestion interact.

Reward function design is central to RL's success. Studies typically target cost minimisation, time reduction, or hybrid combinations, but most fix reward priorities instead of dynamically balancing them. This limits adaptability to heterogeneous user needs and system constraints. Addressing this, the present study applies PPO to optimise cost, time, and hybrid objectives in London, explicitly integrating operator tariffs, charging curves, and traffic congestion into the RL environment.

The effectiveness of RL also depends on training frameworks and simulation tools. Value-based methods such as DQN have been widely used for scheduling [5, 6], but their discrete action space makes them unsuitable for continuous decision problems. PPO offers robustness and flexibility for multi-objective optimisation [7]. For this project, PPO is implemented in OpenAI Gym, which provides a modular and reproducible structure for defining states, actions, and rewards [26]. To capture mobility constraints, the Simulation of Urban Mobility (SUMO) is incorporated, providing spatial and temporal modelling of traffic and congestion [27], which has been applied to London congestion studies. Together, Gym and SUMO provide a realistic framework for evaluating PPO in dense urban EV charging scenarios.

Finally, EV charging inevitably involves trade-offs between cost and time. Cost-oriented approaches minimise user expenditure by leveraging dynamic pricing or demand response, often reducing bills while supporting grid stability [28, 29]. However, these generally overlook time constraints, which are critical in urban mobility. Time-focused approaches instead reduce charging duration, waiting, or travel delays, often directing users to ultra-rapid but more expensive stations [30, 18]. Hybrid frameworks explicitly address this

challenge, employing weighted sums or Pareto methods to capture trade-offs between cost and time [31]. RL, particularly PPO, is well suited to this setting, adapting policies to continuous spaces and balancing competing rewards [7]. Hybrid RL models have shown promise in jointly minimising cost, time, and even secondary objectives such as battery degradation [32]. Yet hybrid optimisation remains underexplored in real-world urban contexts where tariffs are operator-specific, charging curves nonlinear, and congestion highly variable. This gap is especially pronounced in London, where multiple operators impose diverse cost structures and clustered user demand creates congestion. The present study therefore applies PPO to design cost-based, time-based, and hybrid charging strategies in Inner London, explicitly integrating real-world tariffs, nonlinear charging curves, and traffic constraints into a unified optimisation framework.

References

1. Kiani, H., Hesami, K., Azarhooshang, A., Pirouzi, S., & Safaee, S. (2021). Adaptive robust operation of the active distribution network including renewable and flexible sources. Sustainable Energy, Grids and Networks, 26, 100476.

2. Weckx, S., & Driesen, J. (2015). Load balancing with EV chargers and PV inverters in unbalanced distribution grids. IEEE Transactions on Sustainable Energy.

3. Xydas, E., Marmaras, C., & Cipcigan, L. (2016). A data-driven approach for characterising the charging demand of electric vehicles: A UK case study. Applied Energy, 162, 763–771.

4. Qian, T., Shao, C., Wang, X., & Shahidehpour, M. (2019). Deep reinforcement learning for EV charging navigation by coordinating smart grid and intelligent transportation system. IEEE Transactions on Smart Grid, 11(2), 1714–1723.

5. Han, Y., Li, T., & Wang, Q. (2024). A DQN-based approach for large-scale EVs charging scheduling. Complex & Intelligent Systems, 10, 8319–8339.

6. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimisation algorithms. arXiv preprint arXiv:1707.06347.

7. Sovacool, B. K., Noel, L., Axsen, J., & Kempton, W. (2018). The neglected social dimensions to a vehicle-to-grid (V2G) transition: a critical and systematic review. Environmental Research Letters, 13(1), 013001.

8. Nicholas, M., & Lutsey, N. (2020). Quantifying the EV charging infrastructure gap in the UK. ICCT Report.

9. Transport for London. (2021). London's 2030 Electric Vehicle Infrastructure Strategy. TfL.

10. Robinson, C., Blythe, P., Bell, M., Hübner, Y., & Hill, G. (2013). Analysis of electric vehicle driver recharging demand profiles and subsequent impacts on the electricity distribution network. Energy Policy, 61, 337–348.

11. Figenbaum, E., & Kolbenstvedt, M. (2016). Learning from Norwegian battery electric and plug-in hybrid vehicle users: Results from a survey of vehicle owners. Institute of Transport Economics Report.

12. Hardman, S., Jenn, A., Tal, G., Axsen, J., Beard, G., Daina, N., ... & Turrentine, T. (2018). A review of consumer preferences of and interactions with electric vehicle charging infrastructure. Transportation Research Part D, 62, 508–523.

13. Zavvos, E., Gerding, E. H., & Brede, M. (2021). A comprehensive game-theoretic model for EV charging station competition. IEEE Transactions on Intelligent Transportation Systems.

14. Lin, R., Chu, H., Gao, J., & Chen, H. (2024). Charging management and pricing strategy of electric vehicle charging station based on mean field game theory. Asian Journal of Control, 26(2), 803-813.

15. Montoya, A., Guéret, C., Mendoza, J. E., & Villegas, J. G. (2017). The electric vehicle routing problem with nonlinear charging function. Transportation Research Part B, 103, 87–110.

16. Froger, A., Mendoza, J. E., Jabali, O., & Laporte, G. (2019). Improved formulations and algorithmic components for the EV routing problem with nonlinear charging functions. Computers & Operations Research, 104, 256–294.

17. Luo, C., Huang, Y. F., & Gupta, V. (2019). Stochastic dynamic programming for EV charging station placement in urban areas. IEEE Transactions on Smart Grid, 10(2), 2272–2282.

18. He, F., Wu, D., Yin, Y., & Guan, Y. (2013). Optimal deployment of public charging stations for plug-in hybrid electric vehicles. Transportation Research Part B: Methodological, 47, 87–101.

19. García-Álvarez, J., González, M. A., & Vela, C. R. (2018). Metaheuristics for solving a real-world electric vehicle charging scheduling problem. Applied Soft Computing, 65, 292–306.

20. Gautam, M. (2023). Deep reinforcement learning for resilient power and energy systems: Progress, prospects, and future avenues. Electricity, 4(4), 336–380.

21. Perera, A. T. D., & Kamalaruban, P. (2021). Applications of reinforcement learning in energy systems. Renewable and Sustainable Energy Reviews, 137, 110618.

22. Xu, B., Rubenis, A., & Long, C. (2024, October). Reinforcement learning based smart charging for electric vehicle fleet. In International Symposium on Intelligent Technology for Future Transportation (pp. 375–383). Cham: Springer Nature.

23. Zhou, Z., Wen, X., Wang, Y., & Zhang, C. (2021). Battery-aware continuous charging control using Soft Actor–Critic. Applied Energy, 298, 117180.

24. Brockman, G., et al. (2016). OpenAI Gym. arXiv preprint arXiv:1606.01540.

25. Behrisch, M., Bieker, L., Erdmann, J., & Krajzewicz, D. (2011). SUMO – Simulation of Urban Mobility: An Overview. Proceedings of SIMUL 2011, 55–60.

26. Tan, J., & Wang, L. (2014). Real-time charging navigation of electric vehicles to fast charging stations: A game theoretic approach. IEEE Transactions on Smart Grid.

27. Sassi, S., et al. (2020). Fast charging station scheduling for minimising EV waiting time. Applied Energy.

28. Guo, Y., et al. (2021). Multi-objective optimisation of EV charging: Cost and time trade-offs. Energy, 214, 118891.

29. Cui, F., Lin, X., Zhang, R., & Yang, Q. (2022). Multi-objective optimal scheduling of charging stations based on deep reinforcement learning. Frontiers in Energy Research, 10, 1042882.