Pre-Draft Questions

**Literature Review Outline for EV Charging Optimization (RL, Cost/Time/Hybrid)**

**1. EV Charging Infrastructure & Challenges**

**Guiding Questions:**

* What are the main **challenges in urban EV charging** (availability, pricing fragmentation, range anxiety, peak demand)?
* How is EV charging infrastructure being **developed in London/UK**, and what policy/regulation influences optimization?
* What constraints have been used in EV charging (pricing, connector limits, charging curves)?

**2. Traditional Optimization Approaches for EV Charging**

**Guiding Questions:**

* What has been done to optimize **charging costs and time** using traditional methods?
* What optimization methods (MILP, shortest-path algorithms, heuristics) have been applied to EV charging?
* What are the **strengths and limitations** of these approaches (e.g., scalability, adaptability)?

**3. Charging Curves, Pricing, and User Constraints**

**Guiding Questions:**

* How do **charging curves (SoC vs kW)** impact charging optimization?
* How has **pricing variability** (operator tariffs, time-of-day rates, idle fees) been handled in prior work?
* What approaches have been taken to balance **user trade-offs between cost and time**?
* What limitations exist in current pricing/charging curve modeling?

**4. Reinforcement Learning in EV & Energy Systems**

**Guiding Questions:**

* How has **RL been used in EV charging optimization** before?
* What **reward functions** have been designed (time, cost, battery life, hybrid)?
* What factors affect the performance of RL algorithms like PPO in EV charging optimization?
* How do you measure the results of PPO (or RL methods generally)?

**5. Algorithms & Frameworks (PPO, DQN, Gym, SUMO)**

**Guiding Questions:**

* What is PPO, and why use PPO instead of other RL algorithms for this problem?
* What is DQN, and how does it compare to PPO for EV optimization tasks?
* Why use OpenAI Gym instead of other RL frameworks?
* Why use SUMO as a traffic simulator, and how has it been used in similar contexts?
* Has SUMO been used for EV charging optimization before?

**6. Multi-Objective Optimization (Cost, Time, Hybrid)**

**Guiding Questions:**

* What research has addressed **cost vs time trade-offs** in EV charging?
* How has **hybrid optimization** (time + cost) been approached in prior studies?
* What techniques exist for **reward shaping** in multi-objective RL?
* What makes your approach different from existing hybrid optimization research?

**7. Positioning of This Project**

**Guiding Questions:**

* Why did we choose reinforcement learning instead of other optimization methods?
* Why did we choose PPO and Gym/SUMO specifically?
* Why did we choose this project instead of something else?
* What is the importance of EV charging optimization using cost and time constraints?
* What problem are we solving with EV charging optimization in London?
* What are the limitations faced (pricing data, model constraints, scaling)?
* What makes our approach different from existing research?

Literature Review first draft

**1 Introduction**

The rapid adoption of electric vehicles (EVs) is reshaping the transportation and energy landscape. With growing commitments to decarbonization and stricter air quality policies, cities such as London are actively expanding their EV infrastructure to support the anticipated surge in demand. However, this transition introduces significant challenges for both users and operators. From the user’s perspective, the key concerns revolve around minimizing charging costs, ensuring route feasibility, and reducing delays due to charging availability. On the system side, the heterogeneous charging behavior of EVs contributes to uncertainties in demand, congestion, and pricing dynamics [1, 2].

Traditional optimization approaches have primarily focused on **grid-oriented challenges** such as voltage stability, renewable integration, and load balancing [3, 4]. While these methods are valuable in stabilizing distribution networks, they often neglect the **user-centric dimension** of EV charging, including travel times, tariff variability, and driver satisfaction [5]. This gap has motivated a shift towards methods that integrate both **mobility constraints** and **charging economics**, particularly in urban environments with fragmented charging networks.

Recent advances in reinforcement learning (RL) provide promising pathways for addressing these challenges. RL methods have demonstrated adaptability in high-dimensional, uncertain environments, making them well-suited for modeling EV charging optimization where decisions depend on dynamic pricing, state of charge (SoC), and traffic conditions [6, 7]. Algorithms such as Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO) have been applied in related contexts, showing strong performance in balancing multiple objectives such as cost, satisfaction, and grid stability [8, 9]. However, existing studies often emphasize **grid coordination** rather than **driver-level optimization** in real-world city contexts.

This study contributes to bridging this gap by applying PPO-based reinforcement learning to optimize **cost-based, time-based, and hybrid EV charging decisions** in **Inner London**. Unlike previous work, our approach explicitly incorporates real-world charging tariffs, operator heterogeneity, and charging curves to design policies that are both realistic and scalable.

**2 EV Charging Infrastructure and Challenges**

The effectiveness of EV adoption depends heavily on the availability, accessibility, and affordability of charging infrastructure. Literature typically frames the EV charging problem along several dimensions: (a) **where to site and size chargers**, (b) **how to manage station operations**, (c) **how drivers behave in real-world contexts**, and (d) **how pricing and market mechanisms influence charging choices** [10].

**2.1 Urban Infrastructure Constraints**

In dense metropolitan regions such as London, challenges include limited space for charging station deployment, uneven distribution of chargers, and operator fragmentation. UK-focused infrastructure studies highlight significant “infrastructure gaps,” where charging provision lags behind adoption projections [11]. These gaps are especially pronounced in Inner London, where demand clustering and limited land availability complicate the rollout of new stations {need a paper referencing Inner London infrastructure gaps specifically}.

**2.2 User Behavior and Demand Variability**

User heterogeneity adds another layer of complexity. Empirical UK charging data shows strong clustering in charging times, with peak plug-in events occurring in the evenings [12]. Behavioral differences in charging willingness, range anxiety, and flexibility further affect the load profile [13]. Without demand-aware scheduling, urban charging stations face risks of congestion and long waiting times, directly impacting user satisfaction.

**2.3 Pricing and Tariff Complexity**

Pricing structures vary across operators, including per-kWh rates, session fees, idle fees, and time-of-day tariffs. Some works model station competition through game-theoretic frameworks, showing how strategic pricing influences station selection and congestion [14]. However, most existing literature examines **static or simplified pricing models**, which do not reflect the real complexity of London’s multi-operator environment {need a paper referencing operator-specific London tariffs}.

**2.4 Implications for Optimization**

Taken together, these challenges highlight the limitations of grid-only optimization. Any realistic optimization framework must consider not only **system-level efficiency** but also **driver-level objectives** such as minimizing charging time, reducing costs, and balancing trade-offs between the two. Reinforcement learning offers a promising approach for navigating these constraints in real time, making it a suitable candidate for EV charging optimization in London’s dynamic context.

**3 Traditional Optimization Approaches for EV Charging**

Before reinforcement learning gained traction, optimization of EV charging was primarily addressed using deterministic and heuristic methods. These approaches sought to improve system efficiency, minimize costs, and reduce grid impact, but often relied on static assumptions that limit their scalability in real-world urban contexts.

**3.1 Mathematical Programming and Deterministic Models**

Classical formulations such as **Mixed-Integer Linear Programming (MILP)** and **Linear Programming (LP)** have been widely applied to charging station scheduling and siting problems. These models are valued for their transparency and guaranteed optimality under fixed assumptions [1]. For example, MILP has been used to determine optimal siting and sizing of charging infrastructure, balancing costs and voltage stability constraints in power systems [2]. However, such formulations are computationally expensive and often unsuitable for real-time decision-making in dynamic environments.

**3.2 Meta-Heuristic Approaches**

To address computational bottlenecks, meta-heuristic algorithms such as **Genetic Algorithms (GA)**, **Particle Swarm Optimization (PSO)**, and **Ant Colony Optimization (ACO)** have been proposed. These methods provide near-optimal solutions with reduced complexity, making them more practical for large-scale systems [3]. Multi-criteria decision-making (MCDM) frameworks have also been used for charger siting, incorporating factors such as accessibility, grid constraints, and service quality [4]. While flexible, these heuristic approaches lack adaptability when confronted with dynamic pricing and user mobility patterns.

**3.3 EV Routing with Charging Constraints**

The **Electric Vehicle Routing Problem (EVRP)** extends classical vehicle routing formulations by introducing battery capacity, charging time, and service constraints. Early EVRP studies focused on deterministic travel and charging assumptions, treating charging stations as fixed service nodes [5]. However, real-world applications involve nonlinear charging curves, waiting times, and congestion at charging stations, which classical EVRP models fail to capture adequately.

**3.4 Limitations of Traditional Approaches**

While deterministic and heuristic models provide valuable baselines, they exhibit several limitations when applied to urban EV charging optimization:

* **Static assumptions**: Models often ignore real-time fluctuations in traffic, demand, and pricing.
* **Limited scalability**: Solvers become intractable as the number of EVs and charging stations increases.
* **Neglect of user heterogeneity**: Cost sensitivity, charging preferences, and behavioral patterns are rarely included.
* **Weak integration with tariffs**: Most models assume uniform or static prices, ignoring operator-specific pricing structures and idle fees {need a paper referencing idle fee integration}.

These limitations highlight the need for **adaptive, data-driven approaches** capable of handling the dynamic and stochastic nature of EV charging. Reinforcement learning, with its ability to learn policies from interaction with uncertain environments, has emerged as a promising alternative to overcome these shortcomings.

**4 Charging Curves, Tariffs, and User Constraints**

The optimization of EV charging in real-world contexts must account for the complex interplay between vehicle capabilities, charging station characteristics, and pricing structures. Unlike simplified models that assume linear charging and uniform costs, actual charging dynamics are nonlinear and highly variable across users, vehicles, and operators.

**4.1 Charging Curves and Nonlinear Power Delivery**

EV charging is inherently nonlinear due to battery characteristics and management systems. Charging curves show that power delivery is highest at low-to-mid states of charge (SoC) and tapers significantly as SoC approaches full capacity [1]. This tapering effect means that marginal charging times increase nonlinearly, complicating scheduling and route planning. Optimizing routes without incorporating charging curves risks overestimating available power delivery and underestimating charging times. Some studies have begun to integrate nonlinear charging into optimization models [2], but most continue to rely on simplified constant-power assumptions, limiting realism {need a paper explicitly modeling EV charging curves in routing}.

**4.2 Tariffs and Operator-Specific Pricing**

Another key challenge is the heterogeneity of charging tariffs. Operators employ a wide range of cost structures, including per-kWh rates, session fees, membership discounts, idle fees, and time-of-day tariffs [3]. For example, idle fees penalize vehicles that occupy chargers after completing a session, while dynamic pricing reflects fluctuations in electricity demand. Existing optimization models frequently assume uniform or static tariffs, which does not reflect fragmented operator landscapes such as London’s. Game-theoretic frameworks have been used to model station competition, showing how pricing strategies affect user decisions and congestion [4]. However, these models typically operate at a strategic market level rather than capturing the fine-grained tariffs that individual users face.

**4.3 User Behavior and Demand Variability**

User heterogeneity adds an additional dimension to charging optimization. Empirical studies of UK charging demand reveal clustering of charging sessions in the evenings, with strong dependence on household and workplace charging access [5]. Household charging behavior is influenced by daily routines, time-of-use tariffs, and range anxiety, all of which affect public charging demand [6]. Without behavior-aware modeling, optimization frameworks risk overestimating flexibility and underestimating congestion at urban public chargers.

**4.4 Implications for Cost-Time Optimization**

The combined effects of nonlinear charging curves, tariff complexity, and behavioral variability underscore the difficulty of balancing **cost and time objectives**. A driver may minimize cost by charging at lower per-kWh stations, but this may increase travel time or risk long queues. Conversely, prioritizing time may push users toward high-power ultra-rapid chargers with higher tariffs. Few existing models attempt to explicitly address this **multi-objective trade-off**, particularly in real-world multi-operator contexts like London {need a reference on multi-objective EV charging cost vs time}.

**5 Reinforcement Learning for EV Charging Optimization**

Reinforcement learning (RL) has recently emerged as a promising paradigm for addressing the limitations of deterministic and heuristic EV charging optimization approaches. Unlike static methods, RL enables agents to learn adaptive policies through interaction with dynamic environments, making it well-suited for modeling uncertainties in **traffic, charging demand, and tariff variability**.

**5.1 Reinforcement Learning in Energy and Charging Systems**

RL has been widely applied in power and energy domains, including demand response, microgrid scheduling, and smart charging [1]. In EV contexts, RL treats charging as a **sequential decision-making problem**, where the agent learns when, where, and how much to charge based on system states such as SoC, travel time, station availability, and tariffs. Several studies have demonstrated RL’s effectiveness in reducing costs, flattening grid loads, and improving user satisfaction compared to traditional optimization methods [2, 3].

**5.2 Value-Based Methods: DQN and Extensions**

Deep Q-Networks (DQN) extend classical Q-learning by using neural networks to approximate value functions in high-dimensional spaces. DQN has been applied to EV charging scheduling problems, modeling the environment as a Markov Decision Process (MDP) with states representing SoC, tariff levels, and demand [4]. Extensions such as Double DQN and Dueling DQN further improve learning stability and efficiency [5]. While DQN performs well in **discrete action spaces**, such as choosing between predefined charging stations or time slots, it struggles when continuous control is required—for example, determining optimal charging power or duration.

**5.3 Policy Gradient and Actor–Critic Methods: PPO and Beyond**

Policy gradient methods overcome DQN’s limitations by directly optimizing policies for continuous or large action spaces. **Proximal Policy Optimization (PPO)** has gained particular traction due to its stability, sample efficiency, and ability to handle multi-objective optimization [6]. PPO has been successfully applied in energy management and EV fleet scheduling, demonstrating superior adaptability compared to value-based methods [7]. Actor–critic variants such as Deep Deterministic Policy Gradient (DDPG) and Soft Actor–Critic (SAC) have also been explored, with SAC showing promise in minimizing battery degradation costs [8]. Despite these advances, relatively few studies have applied PPO to **urban EV charging optimization at the driver level**, where decisions must balance cost, time, and route feasibility in real-world networks.

**5.4 Reward Function Design for EV Optimization**

Reward functions are central to RL’s performance. In EV charging, rewards have been designed to capture diverse objectives, including:

* **Cost minimization**: per-kWh tariffs, session fees, idle fees [9].
* **Time minimization**: charging time, travel time, and waiting penalties [10].
* **Hybrid objectives**: weighted functions balancing cost, time, user satisfaction, and grid stability [11].
* **Battery health**: degradation linked to depth of discharge and temperature [8].

Most prior works focus on single objectives (cost **or** time), with limited exploration of **explicit hybrid reward functions**. This presents a key research opportunity: developing RL frameworks that can adapt reward weights dynamically to reflect user or operator priorities.

**5.5 Gaps in Existing RL Applications**

While RL has shown clear advantages, several gaps remain:

* **Focus on grid coordination**: Many studies optimize for grid stability rather than user-centric routing.
* **Limited integration of real-world tariffs**: Few models incorporate operator-specific pricing schemes with session and idle fees.
* **Weak use of charging curves**: Most environments assume constant charging rates, ignoring nonlinear SoC dynamics.
* **Sparse urban deployment studies**: RL applications often test on synthetic datasets or single-station setups, not dense multi-operator networks like London.

These gaps underline the novelty of this project, which leverages PPO to optimize **cost-based, time-based, and hybrid objectives** in London, integrating **real tariffs, charging curves, and operator heterogeneity** into the RL environment.

**6 Algorithms and Frameworks for EV Charging Optimization**

The choice of algorithm and framework is central to developing a reinforcement learning environment for EV charging optimization. While the literature has explored a range of RL algorithms, including value-based methods such as Deep Q-Networks (DQN), this project adopts **Proximal Policy Optimization (PPO)** as the primary method due to its stability and suitability for continuous control.

**6.1 Value-Based Methods in the Literature (DQN and Variants)**

DQN and its extensions (Double DQN, Dueling DQN) are frequently applied in EV charging optimization because of their ability to approximate value functions in complex environments [1]. Studies have demonstrated their effectiveness in reducing costs and improving user satisfaction [2]. However, DQN’s reliance on **discrete action spaces** makes it less suitable for urban charging problems that require continuous decisions such as charging duration and power level. Thus, while DQN serves as a useful reference in prior work, this project does not implement it directly.

**6.2 Proximal Policy Optimization (PPO) for This Study**

This project adopts **PPO** as the primary algorithm for EV charging optimization. PPO combines the strengths of policy-gradient methods with mechanisms that ensure training stability by preventing large policy updates [3]. Its advantages include:

* **Continuous control**: PPO can handle continuous action spaces, making it suitable for modeling charging duration and energy delivery.
* **Multi-objective adaptability**: PPO is well-suited to problems where agents must balance competing objectives, such as minimizing charging cost and time.
* **Sample efficiency and robustness**: PPO has demonstrated strong convergence properties across diverse RL applications, making it reliable for large, dynamic environments.

In this project, PPO is applied to optimize EV charging in **Inner London**, where agents must navigate heterogeneous tariffs, charging curves, and traffic conditions to minimize either cost, time, or a hybrid objective.

**6.3 Simulation Environment: OpenAI Gym**

OpenAI Gym provides the structural backbone for this RL environment. Gym standardizes the definition of states, actions, and rewards, ensuring that PPO can be trained in a modular and reproducible framework [4]. This allows seamless integration of EV-specific dynamics such as charging curves, tariffs, and state of charge into the environment without compromising flexibility.

**6.4 Mobility Simulation: SUMO**

The **Simulation of Urban Mobility (SUMO)** adds traffic and spatial realism to EV charging optimization by modeling vehicle movement and congestion patterns [5]. SUMO has been widely used to analyze charging demand and urban traffic interactions, and its ability to integrate with external RL controllers makes it an ideal complement to the Gym environment [6]. While some prior works have used SUMO for infrastructure planning, this project employs it to capture **mobility constraints** relevant to real-time charging route optimization.

**6.5 Implications for This Project**

By focusing on PPO, this project aligns with recent trends in applying policy-gradient methods to energy optimization while addressing the limitations of value-based approaches. Gym provides the modular RL interface, and SUMO introduces realistic urban traffic conditions, together creating a comprehensive environment for testing **cost-, time-, and hybrid-based EV charging strategies in London**.

**7 Multi-Objective Optimization: Cost, Time, and Hybrid**

EV charging optimization involves inherently competing objectives: drivers seek to minimize both **charging cost** and **charging time**, yet these two goals often conflict. Public charging stations vary in power delivery rates, tariffs, and congestion levels, which means that minimizing cost may increase time, and vice versa. A practical optimization framework must therefore address both objectives, either separately or in combination.

**7.1 Cost-Oriented Optimization**

Several studies have framed EV charging as a cost-minimization problem. Optimization models based on dynamic pricing and demand response aim to reduce user bills while flattening grid demand [1]. User-centric models such as those developed by Bian et al. (2019) demonstrated how drivers adapt charging decisions to minimize expenditure under variable tariffs [2]. Reinforcement learning approaches have further improved cost savings by adaptively selecting charging times and stations in response to real-time price signals [3]. However, cost-focused models often neglect user time constraints, which are critical in urban mobility.

**7.2 Time-Oriented Optimization**

Time minimization focuses on reducing **charging session duration**, **waiting times**, and **travel delays**. Approaches in this category often favor ultra-rapid charging stations to minimize overall travel time, even at higher cost [4]. EV routing problems that integrate charging constraints highlight how charging station availability and power ratings directly influence route feasibility [5]. Reinforcement learning methods have also been applied to minimize queueing delays, though most studies prioritize grid efficiency rather than end-user travel time [6].

**7.3 Hybrid and Multi-Objective Approaches**

Hybrid optimization frameworks explicitly balance cost and time by designing reward functions or objective functions that combine both metrics. Some works use weighted linear combinations, while others employ Pareto-optimal approaches to generate trade-off solutions [7]. Actor–critic RL algorithms such as PPO and SAC are particularly well-suited to this setting because of their ability to learn from continuous trade-offs [8]. For example, hybrid RL models have been shown to minimize both charging costs and battery degradation, demonstrating the flexibility of multi-objective reward design [9].

Despite these advances, **very few studies apply hybrid optimization to real-world urban charging networks** where tariffs are operator-specific, charging curves are nonlinear, and congestion is location-dependent. This gap is particularly evident in London, where fragmented operators and diverse tariff structures create highly variable cost–time trade-offs {need a reference explicitly addressing hybrid EV optimization in urban environments}.

**7.4 Positioning of This Project**

This project directly addresses the hybrid optimization gap by training PPO agents to minimize **cost, time, and hybrid objectives** in **Inner London**. Unlike prior studies that simplify tariffs and charging rates, this work integrates:

* **Real-world operator tariffs** (per-kWh rates, session fees, idle fees, and membership rules).
* **Nonlinear charging curves** that capture tapering effects at high SoC.
* **Traffic and station congestion dynamics** via SUMO integration.

By combining these elements, the framework provides a more realistic and user-focused optimization of charging strategies, moving beyond grid-centric or single-objective approaches.

📌 **References for Section 1 & 2:**  
[1] Bian et al. (2019) – user-centric MDP model.  
[2] Xydas (2016) – UK charging demand data.  
[3] Kiani et al. (2021) – robust grid optimization.  
[4] Weckx et al. (2014) – reactive power management.  
[5] Draft 2 framing (grid vs user trade-offs).  
[6] Liu et al. (2024) – DQN for charging scheduling.  
[7] Bertolini (2022) – DRL in charging hubs.  
[8] Yalçın & Herdem (2024) – hybrid RL models.  
[9] Gautam (2023) – review of DRL in power systems.  
[10] EV\_Literature\_Review doc – Section 1 synthesis.  
[11] Nicholas & Lutsey (2020) – UK infrastructure gap.  
[12] Xydas (2016) – UK empirical charging demand.  
[13] Almaghrebi et al. (2024) – household charging behavior.  
[14] Gerding (2021) – game-theoretic station competition.

📌 **References for Section 3**  
[1] Schmidt, M. et al. (2025). Multiple-Criteria-Based EV Charging Infrastructure Design Problem. *Energies*.  
[2] (Author truncated) (2024). Optimal Siting and Sizing of EV Charging Station Using Stochastic Power Flow Analysis for Voltage Stability. *IEEE Transactions on Transportation Electrification*.  
[3] Nicholas, M. (2012). Estimating EV Charging Infrastructure Costs Across Major U.S. Metropolitan Areas. *Transportation Research Record*.  
[4] Anagnostopoulou, A. (2014). Electric Vehicle Routing Problem with Industry Constraints: Trends and Insights for Future Research. *Transportation Research Procedia*.  
[5] {need a strong survey paper on EVRP variants with charging constraints}

📌 **References for Section 4**  
[1] Yalçın, S., & Herdem, M. S. (2024). Optimizing EV Battery Management: Advanced Hybrid RL Models for Efficient Charging and Discharging. *Energies*.  
[2] Bertolini, A. (2022). Power Output Optimization of EV Smart Charging Hubs Using DRL. *Expert Systems with Applications*.  
[3] Lin, R. et al. (2024). Charging Management and Pricing Strategy of EV Stations Based on Mean-Field Game Theory. *Asian Journal of Control*.  
[4] Gerding, E. H. (2021). A Comprehensive Game-Theoretic Model for EV Charging Station Competition. *IEEE Transactions on Intelligent Transportation Systems*.  
[5] Xydas, E. (2016). A Data-Driven Approach for Characterising EV Charging Demand: A UK Case Study. *Applied Energy*.  
[6] Almaghrebi, A. et al. (2024). Insights into Household EV Charging Behavior: Analysis and Predictive Modeling. *Energies*.

📌 **References for Section 5**  
[1] Gautam, M. (2023). Deep Reinforcement Learning for Resilient Power and Energy Systems: Progress, Prospects, and Future Avenues. *Electricity*.  
[2] Bian, J. et al. (2019). Optimizing EV Charging: A Customer’s Perspective. *IEEE Transactions on Smart Grid*.  
[3] Liu, L. et al. (2024). Research on Reinforcement Learning-Based Optimization Algorithm for EV Charging Scheduling. *POWERCON*.  
[4] Li, S. et al. (2022). EV Charging Management Based on Deep Reinforcement Learning. *Sensors*.  
[5] Bertolini, A. (2022). Power Output Optimization of EV Smart Charging Hubs Using DRL. *Expert Systems with Applications*.  
[6] Schulman, J. et al. (2017). Proximal Policy Optimization Algorithms. *arXiv preprint arXiv:1707.06347*.  
[7] (Author not indicated). Reinforcement Learning for EV Charging Optimization: A Holistic Perspective for Commercial Fleets. *Thesis/Monograph*.  
[8] Yalçın, S., & Herdem, M. S. (2024). Optimizing EV Battery Management: Advanced Hybrid RL Models. *Energies*.  
[9] Lin, R. et al. (2024). Charging Management and Pricing Strategy of EV Stations Based on Mean-Field Game Theory. *Asian Journal of Control*.  
[10] Gerding, E. H. (2021). EV Charging Station Competition: A Game-Theoretic Model. *IEEE Transactions on Intelligent Transportation Systems*.  
[11] Kiani, H. et al. (2021). Adaptive Robust Operation of Active Distribution Networks with Flexible Sources. *Sustainable Energy, Grids and Networks*.

📌 **References for Section 6**   
[1] Li, S. et al. (2022). EV Charging Management Based on Deep Reinforcement Learning. *Sensors*.  
[2] Liu, L. et al. (2024). RL-Based Optimization for EV Charging Scheduling. *POWERCON*.  
[3] Schulman, J. et al. (2017). Proximal Policy Optimization Algorithms. *arXiv preprint arXiv:1707.06347*.  
[4] Brockman, G. et al. (2016). OpenAI Gym. *arXiv preprint arXiv:1606.01540*.  
[5] Behrisch, M. et al. (2011). SUMO – Simulation of Urban Mobility: An Overview. *SIMUL 2011*.  
[6] {need a paper explicitly showing SUMO + RL integration for EV routing}.

📌 **References for Section 7**  
[1] Kiani, H. et al. (2021). Adaptive Robust Operation of Active Distribution Networks with Flexible Sources. *Sustainable Energy, Grids and Networks*.  
[2] Bian, J. et al. (2019). Optimizing EV Charging: A Customer’s Perspective. *IEEE Transactions on Smart Grid*.  
[3] Liu, L. et al. (2024). RL-Based Optimization for EV Charging Scheduling. *POWERCON*.  
[4] Bertolini, A. (2022). Power Output Optimization of EV Smart Charging Hubs Using DRL. *Expert Systems with Applications*.  
[5] Anagnostopoulou, A. (2014). Electric Vehicle Routing Problem with Industry Constraints. *Transportation Research Procedia*.  
[6] Ding, L. et al. (2023). DRL for EV Scheduling with Voltage Stability Considerations. *Electricity*.  
[7] Nicholas, M. & Lutsey, N. (2020). Quantifying the EV Charging Infrastructure Gap in the UK. *ICCT Report*.  
[8] Schulman, J. et al. (2017). Proximal Policy Optimization Algorithms. *arXiv:1707.06347*.  
[9] Yalçın, S. & Herdem, M. S. (2024). Hybrid RL Models for Efficient Charging and Discharging. *Energies*.