The original pricing information was gathered directly from public energy provider websites (e.g., SureCharge, Believe, Ubitricity, BP Pulse) [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17].

A screenshot of a computer

AI-generated content may be incorrect.

Figure 1: SureCharge – Raw Tariff Information (Website Extract)

A screenshot of a computer

AI-generated content may be incorrect.

Figure 2: Believ – Raw Tariff Information (Website Extract)

A close-up of a message

AI-generated content may be incorrect.

Figure 3: Ubitricity – Raw Tariff Information (Website Extract)

A screenshot of a website

AI-generated content may be incorrect.

Figure 4: Bppulse – Raw Tariff Information (Website Extract)

The raw pricing data included inconsistent formats, varying currencies (pence/kWh vs £/kWh), and heterogeneous fee structures (energy rates, idle fees, connection fees). To ensure comparability, we standardised units, harmonised VAT inclusion, and separated the data into three structured datasets: pricing\_core (base tariffs), pricing\_conditions (time-of-use, idle rules), and pricing\_by\_charger\_type (connector/power-based overrides). These structured datasets were integrated into the charging cost estimation module, enabling the reinforcement learning environment to compute realistic, provider-specific charging session costs.

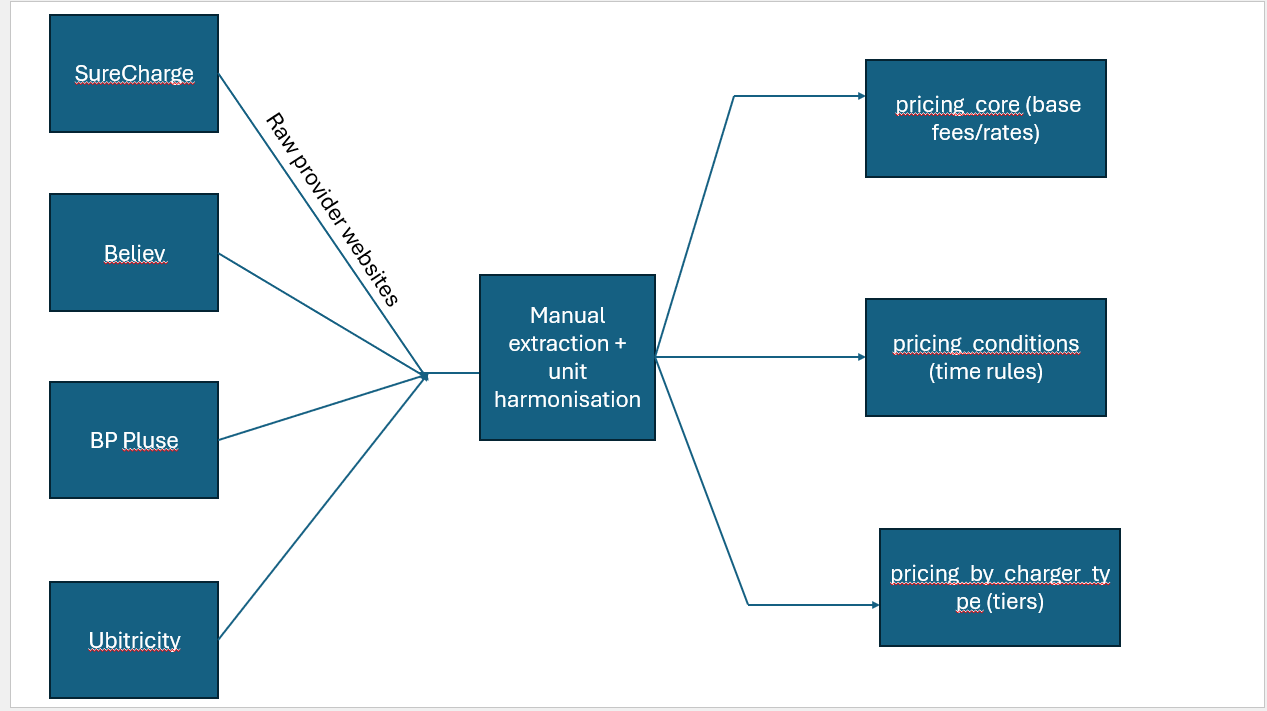


Figure 5: Pricing Data Collection and Processing Pipeline

The original charging station dataset was obtained from the National Chargepoint Registry (NCR), archived by the UK Department for Transport [18]. We cleaned duplicates, standardised geolocations, and separated connector details from station metadata, producing two datasets: *charging\_station\_connectors* and *charging\_station\_metadata*. These cleaned datasets support station feature generation and action feasibility within the RL environment.

A diagram of a system

AI-generated content may be incorrect.

Figure 6: Charging Station Data Cleaning and Splitting Process

EV specifications and charging curves were sourced from Open EV Data v2 [19]. We parsed JSON shards to extract usable\_battery\_size, AC/DC ports/power, energy\_consumption and charging\_curve points. Cleaning: standardised plug names, converted consumption to kWh/km, interpolated 0–100% curves, capped powers. Used to build EV\_Metadata and EV\_Charging\_Curve\_Data powering SOC and charging-time.

A diagram of a diagram

AI-generated content may be incorrect.

Figure 7: EV Data Extraction and Normalisation Pipeline

The datasets underlying the simulation were first pre-processed into structured and harmonised bundles. Tariffs were collected directly from operator websites, where heterogeneous formats included per-kWh rates, session fees, idle charges, and membership discounts. These were normalised and split into three structured datasets: base tariffs, conditional rules, and connector overrides. Station data were obtained from the UK National Chargepoint Registry (NCR), cleaned for duplicates and coordinate inconsistencies, and divided into station-level metadata and connector-level data. Vehicle specifications and charging curves were sourced from Open EV Data v2, where nonlinear charging curves were interpolated from 0–100% state-of-charge (SoC). Together, these pre-processed datasets were unified into a single bundle that provided tariffs, station features, and charging curves in a consistent format (see Fig. 1). This avoided runtime CSV parsing and ensured reproducibility across experiments.

The charging decision process was formalised as a Markov decision process (MDP), defined as M=(S,A,P,R,γ)\mathcal{M} = (S, A, P, R, \gamma)M=(S,A,P,R,γ). The state vector consisted of SoC, vehicle position (latitude, longitude), distance to destination, and candidate charging stations. The action space was a discrete choice among up to kkk charging station candidates, where k=5k=5k=5 by default, padded if fewer stations were feasible. The transition function PPP updated SoC according to energy consumption, increased elapsed time using congestion-adjusted travel speeds, and replenished SoC during charging using interpolated nonlinear curves constrained by station power ratings. The reward function RRR varied by experiment. In cost-oriented runs, the reward was the negative charging expenditure:

rtcost=−Ct,r\_t^{cost} = -C\_t ,rtcost​=−Ct​,

where CtC\_tCt​ is the session cost in GBP. In time-oriented runs, the reward penalised both detour and charging duration:

rttime=−(Tdetour,t+Tcharge,t).r\_t^{time} = -(T\_{detour,t} + T\_{charge,t}) .rttime​=−(Tdetour,t​+Tcharge,t​).

For hybrid optimisation, the reward combined both elements via a normalised weighted sum:

rthybrid=−(α⋅Tt30+(1−α)⋅Ct10),r\_t^{hybrid} = -\left(\alpha \cdot \frac{T\_t}{30} + (1-\alpha)\cdot \frac{C\_t}{10}\right),rthybrid​=−(α⋅30Tt​​+(1−α)⋅10Ct​​),

where TtT\_tTt​ is journey time in minutes, CtC\_tCt​ is cost in GBP, and denominators represent scaling constants to balance magnitudes. Figure 2 illustrates this state–action–reward loop (see Fig. 2).

To ensure that training episodes provided non-trivial learning signals, several deliberate environment design choices were made. Initial SoC was sampled uniformly from 10–30%, forcing the agent to encounter charging events during most episodes. Without this, many trips could be completed without recharging, leading to sparse reward signals. Trip distances were sampled uniformly from 12–25 km, sufficiently long to require at least one charging stop. A fixed overhead of three minutes was added to every charging session, representing plug-in and authentication delays, and discouraging the agent from exploiting unrealistic micro-charging strategies. Traffic congestion was modelled using scalar multipliers applied to travel times: ×1.6 during morning peaks, ×1.5 during evening peaks, and ×1.0 during off-peak hours. These values approximate Inner London congestion patterns without incurring the runtime cost of full SUMO simulations. Table 1 summarises the differences between the prototype environment and the PPO training environment (see Table 1).

Reward shaping was introduced to stabilise training. Beyond the base objectives, the environment applied a success bonus of +50 upon reaching the destination, a stranding penalty of –200 when the vehicle depleted its battery, and an invalid action penalty of –2 when infeasible chargers were selected. The magnitude of the stranding penalty was chosen to be roughly twenty times the cost of a typical charging session, ensuring that feasibility always dominated over small cost savings. In hybrid mode, time and cost were compared directly by introducing a value-of-time of £0.05/minute, a standard figure in transport studies {with reference needed here}. This enabled the agent to trade between financial and temporal objectives on a consistent scale. The reward trade-off surface is visualised in Fig. 3, showing how the weight parameter α\alphaα shifts optimisation between time and cost (see Fig. 3).

The PPO algorithm was employed in this study via the Stable-Baselines3 library {with reference needed here}, without modifications to its implementation. PPO was chosen for its robustness in continuous decision-making and its sample efficiency relative to older policy gradient methods. Its clipped surrogate objective, as proposed by Schulman et al. {with reference needed here}, is given by:

LCLIP(θ)=E^t[min⁡(rt(θ)A^t, clip(rt(θ),1−ϵ,1+ϵ)A^t)],L^{CLIP}(\theta) = \hat{\mathbb{E}}\_t \Big[ \min \big( r\_t(\theta) \hat{A}\_t, \, \text{clip}(r\_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}\_t \big) \Big],LCLIP(θ)=E^t​[min(rt​(θ)A^t​,clip(rt​(θ),1−ϵ,1+ϵ)A^t​)],

where rt(θ)=πθ(at∣st)πθold(at∣st)r\_t(\theta) = \frac{\pi\_\theta(a\_t|s\_t)}{\pi\_{\theta\_{old}}(a\_t|s\_t)}rt​(θ)=πθold​​(at​∣st​)πθ​(at​∣st​)​ and A^t\hat{A}\_tA^t​ is the advantage estimate. Advantages were computed using Generalised Advantage Estimation (GAE):

A^t=∑l=0∞(γλ)lδt+l,\hat{A}\_t = \sum\_{l=0}^\infty (\gamma \lambda)^l \delta\_{t+l} ,A^t​=l=0∑∞​(γλ)lδt+l​,

where δt\delta\_tδt​ is the temporal-difference error, γ\gammaγ is the discount factor, and λ\lambdaλ trades off bias and variance. While these equations form the theoretical basis of PPO, their implementation in this study relied entirely on the Stable-Baselines3 framework. Our contribution therefore lies in the environment design, reward shaping, and evaluation framework, rather than in modifications to PPO itself. Figure 4 illustrates the PPO training loop, where the agent interacts with the Gym environment, which in turn queries tariff and charging modules, applies congestion multipliers, and returns rewards to the agent (see Fig. 4).

The experimental setup was designed to evaluate PPO across cost, time, and hybrid objectives. Each trained policy was tested on a held-out dataset of trips, distinct from those used for training. Metrics included mean charging cost per journey, mean journey time (including detours and charging), hybrid weighted score, and success rate, defined as the percentage of trips completed without battery depletion. Baseline strategies were implemented for comparison: a greedy nearest-station heuristic, a greedy cheapest-station heuristic, and shortest-path routing with deterministic charging. These baselines provided interpretable benchmarks; the nearest-station heuristic typically reduced detours but ignored tariffs, while the cheapest-station heuristic minimised expenditure at the expense of long detours or slow charging. PPO was expected to learn more adaptive strategies, particularly under hybrid objectives. Figure 5 shows a typical convergence plot of PPO training, with episodic rewards stabilising after sufficient timesteps (see Fig. 5).

Reproducibility was prioritised throughout the methodology. All preprocessing pipelines were deterministic and version-controlled, ensuring identical structured datasets could be reproduced from raw inputs. The Gym environment was modular, with clear configuration options specifying SoC distributions, trip ranges, and reward objectives. Random seeds were fixed in PPO training runs to control variance, and evaluation was performed on disjoint datasets to avoid leakage. While SUMO was available for high-fidelity traffic simulation, training predominantly used lightweight congestion multipliers to balance realism and computational efficiency. Together, these steps ensured the methodology provided both reproducibility and generalisability to other urban contexts.

References

1. FM Conway. (n.d.) SureCharge pricing. Available from: [https://www.fmconway.co.uk/our-services/electric-vehicle-charging/faqs#:~:text=How%20much%20does%20SureCharge%20cost,midnight%20Friday%20to%20midnight%20Sunday](https://www.fmconway.co.uk/our-services/electric-vehicle-charging/faqs?utm_source=chatgpt.com#:~:text=How%20much%20does%20SureCharge%20cost,midnight%20Friday%20to%20midnight%20Sunday) [Accessed 29 August 2025].
2. Char.gy. (n.d.) Pricing. Available from: [https://char.gy/pricing](https://char.gy/pricing?utm_source=chatgpt.com) [Accessed 29 August 2025].
3. Zest. (n.d.) Charge with us. Available from: [https://www.zest.uk.com/charge-with-us](https://www.zest.uk.com/charge-with-us?utm_source=chatgpt.com) [Accessed 29 August 2025].
4. Zap-Map. (n.d.) Zest network overview. Available from: [https://www.zap-map.com/ev-guides/public-charging-point-networks/zest](https://www.zap-map.com/ev-guides/public-charging-point-networks/zest?utm_source=chatgpt.com) [Accessed 29 August 2025].
5. BP Pulse. (n.d.) Public EV charging pricing. Available from: [https://www.bppulse.com/en-gb/public-ev-charging/pricing](https://www.bppulse.com/en-gb/public-ev-charging/pricing?utm_source=chatgpt.com) [Accessed 29 August 2025].
6. Driving Electric. (n.d.) Complete guide to Source London charging network. Available from: [https://www.drivingelectric.com/your-questions-answered/1191/complete-guide-source-london-charging-network](https://www.drivingelectric.com/your-questions-answered/1191/complete-guide-source-london-charging-network?utm_source=chatgpt.com) [Accessed 29 August 2025].
7. Ubitricity. (n.d.) Driver pricing. Available from: [https://ubitricity.com/en/driver/pricing/](https://ubitricity.com/en/driver/pricing/?utm_source=chatgpt.com) [Accessed 29 August 2025].
8. Qwello. (n.d.) Pricing. Available from: [https://qwello.uk/en](https://qwello.uk/en?utm_source=chatgpt.com) [Accessed 29 August 2025].
9. Believ. (n.d.) Driver pricing. Available from: [https://www.believ.com/drivers/pricing/](https://www.believ.com/drivers/pricing/?utm_source=chatgpt.com) [Accessed 29 August 2025].
10. CityEV. (n.d.) Street charging information. Available from: [https://cityev.net/street-how-to/](https://cityev.net/street-how-to/?utm_source=chatgpt.com) [Accessed 29 August 2025].
11. WhatCar. (n.d.) The true cost of using public charging points. Available from: [https://www.whatcar.com/news/the-true-cost-of-using-public-charging-points/n21009](https://www.whatcar.com/news/the-true-cost-of-using-public-charging-points/n21009?utm_source=chatgpt.com) [Accessed 29 August 2025].
12. ESB Energy. (n.d.) EV charging pricing (general membership). Available from: [https://www.esbenergy.co.uk/ev/pricing](https://www.esbenergy.co.uk/ev/pricing?utm_source=chatgpt.com) [Accessed 29 August 2025].
13. Connected Kerb. (n.d.) How much does it cost to charge an electric car? Available from: [https://connectedkerb.com/stories-reports-and-events/how-much-does-it-cost-to-charge-an-electric-car/](https://connectedkerb.com/stories-reports-and-events/how-much-does-it-cost-to-charge-an-electric-car/?utm_source=chatgpt.com) [Accessed 29 August 2025].
14. Shell Recharge. (n.d.) Public charging pricing. Available from: [https://www.shell.co.uk/electric-vehicle-charging/public-charging-with-shell-recharge/shell-recharge-pricing.html](https://www.shell.co.uk/electric-vehicle-charging/public-charging-with-shell-recharge/shell-recharge-pricing.html?utm_source=chatgpt.com) [Accessed 29 August 2025].
15. GeniePoint. (n.d.) Rates and tariffs. Available from: [https://www.equans.co.uk/drivers-geniepoint/rates-tariffs](https://www.equans.co.uk/drivers-geniepoint/rates-tariffs?utm_source=chatgpt.com) [Accessed 29 August 2025].
16. Leccy.net. (n.d.) Blink charging (7kW fast). Available from: [https://leccy.net/charging/public/blink-charging/7kw-fast#google\_vignette](https://leccy.net/charging/public/blink-charging/7kw-fast?utm_source=chatgpt.com#google_vignette) [Accessed 29 August 2025].
17. Smart Charge. (n.d.) Pricing. Available from: [https://smartcharge.co.uk/](https://smartcharge.co.uk/?utm_source=chatgpt.com) [Accessed 29 August 2025].
18. UK Department for Transport. (2025) *National Chargepoint Registry (NCR) dataset [archived version, supplied on request]*. Provided to author via email communication, March 2025.
19. KilowattApp. (n.d.) *Open EV Data v2.* GitHub. Available from: [https://github.com/KilowattApp/open-ev-data/tree/master/data/v2](https://github.com/KilowattApp/open-ev-data/tree/master/data/v2?utm_source=chatgpt.com)