sustainability

MDPI

Article

Reinforcement Learning-Based Energy Management for Sustainable Electrified Urban Transportation with Renewable Energy Integration: A Case Study of Alexandria, Egypt

Amany El-Zonkoly

Department of Electrical & Control Engineering, Arab Academy for Science, Technology & Maritime Transport, Alexandria 21937, Egypt; amanyelz@aast.edu

Abstract

To enhance access to efficient and low-carbon public transportation, the city of Alexandria, Egypt, has introduced a fleet of electric buses. Additionally, an ongoing project aims to upgrade and electrify the existing urban railway system, which is expected to alleviate traffic congestion in this densely populated city. The implementation of electric vehicle (EV) parking facilities is also under consideration. This paper investigates the integration of photovoltaic (PV) systems and green hydrogen-powered gas turbines as components of the integrated energy system (IES). An optimal energy management strategy is proposed to maximize the benefits of incorporating renewable energy sources into the urban transportation system (UTS). The proposed energy management algorithm incorporates demand-side management (DSM) for UTS loads and EVs, increasing the complexity of the decision-making process due to the high uncertainty of decision variables. To address this challenge, a modified multi-agent reinforcement learning (MRL) approach is employed, in which uncertainty is incorporated through stochastic environment sampling. Simulation results demonstrate the economic potential of integrating renewable and sustainable energy resources into the IES of the electrified urban transportation system, achieving a 40.2% reduction in the average daily energy consumption cost.

Keywords: sustainable mobility; urban transport planning; electrified urban transportation system (UTS); renewable energy resources (RES); Integrated energy management; modified reinforcement learning (MRL)

Check for updates

Received: 20 January 2026
Revised: 20 February 2026
Accepted: 24 February 2026
Published: 28 February 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

As the transportation sector is a major energy consumer and a significant contributor to greenhouse gas emissions, many countries are transitioning toward its electrification. Policies have been adopted to electrify urban transportation systems through the integration of electric vehicles (including cars and buses) and the modernization of urban railway networks.

1.1. Motivation

The urban railway system (URS) serves as a fundamental element of urban transportation, playing a key role in alleviating traffic congestion in metropolitan areas. It offers multiple benefits, such as high operational speed, substantial passenger capacity, and reliable scheduling [1]. To minimize energy consumption and curb greenhouse gas emissions associated with the URS, significant efforts have been made to enhance energy efficiency and incorporate renewable energy sources [2–4].

Additionally, railway stations contribute to nearly 50% of the total energy demand in electrified railway networks. As a result, efficient energy management (EM) at these stations is essential for improving the overall energy performance of the URS [1]. To promote sustainability and optimize power supply efficiency, clean electrification solutions have been implemented. These include the integration of photovoltaic (PV) systems in elevated railway stations [4] and the utilization of green hydrogen-powered gas turbines [5].

1.2. Literature Review

The energy consumption of electrified urban railway systems (URS) has been extensively analyzed from multiple perspectives. A primary research focus has been on characterizing energy usage patterns in railway stations. Authors in [6,7] developed a generalized framework that outlines the key characteristics of energy consumption in metro stations, covering both traction and non-traction loads. Similarly, in [8], Guan et al. conducted an in-depth study on the hourly energy profile of a metro station. Their research utilized a multi-objective mixed-integer linear programming (MILP) approach to enhance renewable energy utilization through on-site photovoltaic (PV) generation and energy storage systems (ESSs). However, this study primarily aimed at increasing renewable energy penetration and minimizing the payback period, without accounting for uncertainties in PV generation.

Another crucial research direction involves forecasting railway station energy consumption to optimize energy management strategies and infrastructure design. Authors in [9] designed forecasting models using machine learning (ML) algorithms to predict the energy consumption and carbon footprint of urban rail transit systems, but their study did not incorporate energy management strategies. Furthermore, their model did not consider additional station-related energy demands, such as those from electric vehicle (EV) parking or energy storage systems. Ref. [10] presented analysis of energy consumption in the Italian railway system, utilizing a bottom-up deterministic modeling approach to estimate energy demand for railway traction. However, their model did not account for the uncertainties in some station loads. In contrast, authors in [11] introduced binary nonlinear fitting regression (BNFR) and support vector regression (SVR) models to predict the electricity consumption of individual load types in subway stations. Nevertheless, this study exclusively focused on station loads without addressing available energy resources.

Given that railway stations consume nearly 50% of the total energy in electrified railway systems, their energy management (EM) has drawn significant research attention. Studies have explored the optimization of heating, ventilation, and air conditioning (HVAC) systems [12] and the implementation of energy-efficient lighting solutions [13]. Authors in [14] employed a probabilistic clustering-based framework for the optimal operation of a railway station. Their model incorporated regenerative braking energy (RBE), ESSs, and PV generation while addressing uncertainties related to PV generation. However, other uncertainties, such as variations in station loads and electricity prices, were not considered. Additionally, authors in [2,15,16] proposed an optimized EM strategy for smart traction systems to improve energy efficiency. Their approach focused on demand-side management (DSM) of railway station loads. Furthermore, refs. [2,16] integrated PV systems into station power networks while managing the energy consumption of EV parking garages. Their method accounted for RBE, ESSs, and rooftop PV units; however, the consideration of uncertainties in EM models remained limited.

In this paper, Reinforcement learning (RL) is employed to solve the proposed EM–DSM problem because it can learn optimal policies for sequential decision-making under uncertainty. Unlike traditional optimization or metaheuristic methods that require repeated re-optimization, RL enables real-time control in dynamic and stochastic environments. This

makes RL particularly suitable for energy management problems involving renewable integration, EVs, and time-varying loads [17,18].

In contrast, RL learns a control policy that maps system states to optimal actions, allowing real-time decision-making under uncertainty and without repeated offline optimization. This approach is particularly effective when the system is subject to mixed uncertainties in renewable generation, load demand, and energy prices.

Electric buses have also emerged as a vital component of electrified urban transportation, significantly reducing greenhouse gas emissions. However, their increased adoption places additional pressure on the electricity grid. Authors in [19] developed an MILP-based algorithm to optimize e-bus charging schedules, incorporating PV generation and ESSs. In a related study, Liu et al. introduced a surrogate-based optimization method to determine the most efficient PV and ESS configurations to minimize e-bus battery charging costs [20].

Electric buses have also emerged as a vital component of electrified urban transportation, significantly reducing greenhouse gas emissions; however, their large and time-constrained charging demand can impose substantial stress on the electricity grid. Several studies have addressed this challenge using forecasting-based optimization approaches. For example, authors in [19] developed an MILP-based algorithm to optimize e-bus charging schedules by explicitly incorporating PV generation and energy storage systems, while Liu et al. [20] employed a surrogate-based optimization framework to determine optimal PV and ESS sizing for minimizing charging costs. Similarly, Jarvis et al. [21] integrate renewable energy forecasting with optimization-based charging schedules to align e-bus charging with predicted clean-energy availability while maintaining service quality. In contrast to these forecasting-driven, fleet-level scheduling methods, this paper embeds e-bus charging within a broader integrated urban transportation energy system managed using reinforcement learning. Rather than relying on explicit renewable forecasts or predefined clean-energy windows, the proposed RL-based framework learns charging and energy exchange policies through interaction with stochastic system dynamics, coordinating e-bus stations jointly with metro station loads, EV parking facilities, PV generation, hydrogen production, and grid transactions under a unified decision-making strategy. This distinction in scope and control philosophy differentiates the proposed approach from existing e-bus scheduling studies and highlights its contribution to system-level energy coordination under uncertainty.

Regarding the EVs' charging/discharging behavior, recent studies have highlighted the importance of modeling user-centric uncertainty. For example, Wu et al. [22] provide a detailed characterization of EV owner decision-making by explicitly considering multiple uncertain attributes such as arrival time, departure time, energy demand, and charging willingness. Their work demonstrates that ignoring behavioral uncertainty can significantly overestimate the flexibility potential of charging stations and lead to overly optimistic scheduling results.

Similarly, authors in [23,24] emphasize that EV participation in grid-support services is strongly influenced by user preferences, perceived risk, and situational constraints. These works clearly show that deterministic or fully compliant charging assumptions limit the realism and transferability of optimization-based scheduling approaches.

Building on these insights, it can be shown that many existing transportation electrification studies rely on forecasting-based or scenario-specific optimization frameworks that require explicit assumptions regarding user behavior, compliance, and charging willingness. While these approaches provide valuable analytical insights, their performance depends heavily on the accuracy of assumed behavioral models and predefined uncertainty distributions.

In contrast, in this paper, the proposed reinforcement learning–based framework does not impose fixed charging commitments or explicit behavioral utility functions. Instead, user voluntariness and scenario variability are implicitly embedded in the environment through stochastic availability, SOC constraints, and state-dependent feasible action sets. The RL agent learns adaptive policies through repeated interaction with this uncertain environment, enabling it to respond to heterogeneous and time-varying participation patterns without requiring explicit ex-ante behavioral modeling.

To enhance the energy self-sufficiency of railway stations and support sustainable electrification, researchers have explored the integration of PV systems in elevated railway infrastructure. Authors in [4] examined the feasibility of installing PV units along the trackside of elevated metro lines to improve energy efficiency and power quality. Authors in [3] investigated the techno-economic potential of deploying PV systems on station rooftops and platform canopies. Additionally, authors in [1,25] formulated mixed-integer programming algorithms to optimize PV-battery systems for energy savings and economic benefits. However, these studies primarily focused on off-grid PV systems, leading to energy surplus wastage due to the absence of effective utilization mechanisms.

Despite the growing integration of renewable energy into urban railway systems, large-scale power generation for these networks remains heavily dependent on gas turbines (GTs), which primarily use diesel or natural gas, thereby contributing to greenhouse gas emissions. Hydrogen-fueled GTs have been identified as a more environmentally sustainable alternative [26]. To enhance sustainability, researchers have investigated the integration of renewable energy sources with hydrogen production [5,27]. In this study, surplus PV-generated energy is repurposed for hydrogen production, further advancing the sustainability of railway power systems.

To clarify the contributions and motivations of this research, a comparative analysis with previous studies is provided in Table 1. While existing studies typically focus on forecasting-based optimization of isolated components (e.g., EV or e-bus charging), the present work contributes a learning-based coordination framework that jointly manages metro station loads, EV parking, E-bus stations, PV generation, hydrogen production, and grid interaction under uncertainty. Additionally, although uncertainties in PV generation and energy prices have been considered in earlier research, the uncertainty associated with station load variations has largely been overlooked. Moreover, this study incorporates the impact of e-bus availability patterns, which have not been widely examined in the previous literature. The novelty lies not in introducing individual technologies, but in their integrated control through a unified RL-based energy management strategy.

Table 1. Comparison of this paper with previous research papers.

Ref.	PV System	EVs	E-Buses	Green-Hydrogen	EM	Uncertainties			Elevated Railway Stations
						RES	Energy Price	Loads	
[1]	√	•	•	•	•	•	•	•	√
[2]	√	√	•	•	√	√	√	•	√
[3]	√	•	•	•	•	•	•	•	√
[4]	√	•	•	•	√	•	•	•	√
[8]	√	•	•	•	√	•	•	•	√
[14]	√	•	•	•	√	√	•	•	•
[15]	•	•	•	•	√	•	•	•	•
[16]	√	√	•	•	√	√	•	•	√
[25]	√	•	•	•	√	•	•	•	√
This paper	√	√	√	√	√	√	√	√	√

√ topic addressed; • topic not addressed.

1.3. Contributions

This paper investigates the integrated energy system (IES) of an urban transportation system (UTS) in Alexandria, consisting of an electrified metro network with 20 elevated stations and an electric bus (E-bus) station accommodating up to 100 buses. The metro and bus stations are equipped with rooftop photovoltaic (PV) installations. In addition, the UTS includes a green hydrogen production facility supplying ten hydrogen-fueled gas turbines (HGTs). Energy exchange occurs with the utility grid as well as with an electric vehicle (EV) parking facility, enabling internal energy exchange across multiple subsystems.

The primary objective of this study is to optimally manage UTS energy resources through coordinated demand-side management of station loads and scheduling of EV and E-bus charging and discharging. This problem is challenging due to the stochastic nature of renewable generation, load demand, regenerative braking, and electricity prices. To address these challenges, a reinforcement learning (RL)–based energy management framework is proposed, enabling adaptive decision-making under uncertainty.

The main contributions of this work are summarized as follows:

- A comprehensive urban transportation energy system model is developed that integrates metro station loads, E-bus station, EV parking, PV generation, hydrogen production, and grid interaction within a unified operational framework.
- A renewable-aware energy utilization strategy is proposed, enabling surplus PV power to be effectively allocated to internal loads, EV and E-bus charging, or hydrogen production that fuels a set of gas turbines, thereby reducing energy wastage and grid dependence.
- The energy management problem is formulated as a factored Markov decision process (FMDP), capturing the interactions among heterogeneous subsystems while maintaining computational tractability.
- A multi-agent reinforcement learning–based coordination strategy is developed to jointly manage demand-side flexibility and distributed energy resources under stochastic operating conditions.

2. Methodology

This paper proposes an optimal energy management approach for Alexandria's metro system to reduce reliance on the electricity grid while enhancing passenger satisfaction across various stations along the metro line. As depicted in Figure 1, the metro's integrated energy system (IES) consists of photovoltaic (PV) generation units, a green hydrogen production facility, hydrogen-fueled gas turbines (HGTs), electric vehicle (EV) parking areas for both cars and buses, and regenerative braking (RB) as a supplementary energy source.

2.1. System Model

2.1.1. PV System

Photovoltaic (PV) units are planned for deployment on the line's way-sides and on the rooftops of station buildings and platform canopies within Alexandria's metro system, as demonstrated in [2] for one of the interchange stations along the line. Furthermore, the canopies of the electric bus station are also considered for PV installation. The deterministic model for PV power generation is presented in Equation (1) [27].

$$P_t^{PV} = \eta^{PV} P_{max}^{PV} \frac{SR_t}{SR_{STC}} \left[1 + tc\left(T_t^C - T_{STC}^{PV}\right)\right] \quad (1)$$

$$T_t^C = T_t^{amb} + 0.0256\, SR_t \quad (2)$$

where η^{PV} is the efficiency of the PV array, P_{max}^{PV} is the maximum PV power, SR_t is the solar radiation at time t, SR_{STC} is the solar radiation under standard testing conditions (STC). t_c is the temperature coefficient of the maximum output power, T_t^C is the temperature of the PV cell at time t, T_{STC}^{PV} is the temperature of the PV cell under STC, and T_t^{amb} is the ambient temperature at time t.

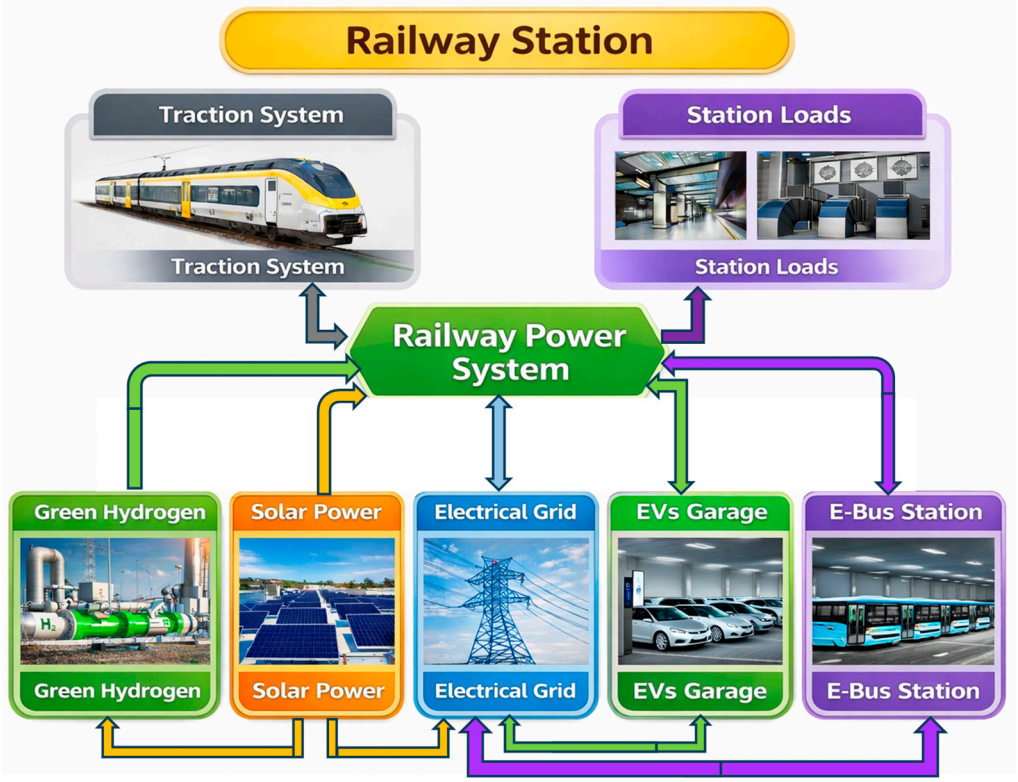


Figure 1. Structure of the energy management system of the UTS.

2.1.2. Electrolyzer

Since the surplus power generated by PV units is allocated for green hydrogen production, the electrical energy stored in the produced hydrogen can be determined using Equation (3).

$$\begin{cases} P_t^{H2} = \eta_{EL} \times P_t^{EL-PV} \\ P_{min}^{EL} \leq P_t^{EL-PV} \leq P_{max}^{EL} \end{cases} \quad (3)$$

where P_t^{H2} is the electrolyzer's output power at time t, P_t^{EL-PV} is the PV power input to the electrolyzer (EL) at time t, and η_{EL} is the conversion efficiency of the EL. This efficiency represents the portion of the electrical energy contained in the produced hydrogen gas to the electrical energy supplied to the EL.

2.1.3. Electric Vehicles Model

This study examines two categories of electric vehicles (EVs): private cars and public buses, each exhibiting distinct charging and discharging behaviors. The charging and discharging patterns of electric cars are influenced by their arrival and departure times, which are modeled stochastically using probability distribution functions with varying means and standard deviations. Moreover, private car owners can decide whether to share their battery charge with the network, a choice determined by a willingness factor (Wf) assigned to each vehicle. In addition, the minimum SOC of each vehicle is assumed to depend on its next journey information given by the vehicle's owner. For that reason, a factor J_{EV}^i is given to each vehicle to determine its allowable minimum SOC as given in (5).

Conversely, electric buses operate on fixed arrival and departure schedules. These factors introduce significant uncertainty in the power exchange between the metro's power system and EVs. Another source of uncertainty is the initial state of charge (SOC) of each vehicle.

The update of stored energy in EV batteries is described in Equations (4)–(8), while SOC constraints and EV-specific limitations on charging/discharging rates and intervals are detailed in Equation (9).

$$B_{EV} = \begin{cases} 1 & \text{if the E}V\text{ is charging} \\ -1 & \text{if the EV is discharging} \\ 0 & \text{if the EV is idle} \end{cases} \tag{4}$$

$$SOC_{min}^{EV-i} = SOC_{min}^{EV} . J_{EV}^{i} \tag{5}$$

$$SOC_{t+1}^{EV} \cdot E_{max}^{EV} = SOC_{t}^{EV} \cdot E_{max}^{EV} + \eta_{EV} \left[B_{EV} \left(P_{t+1}^{ch} . \frac{B_{EV}+1}{2} + P_{t+1}^{dis} . Wf_{EV} . \frac{1-B_{EV}}{2} \right) \right] \Delta t \tag{6}$$

where

$$J_{EV}^{i} \geq 1 \text{ and } Wf_{EV} \in [0,1] \tag{7}$$

$$\eta_{EV} = \begin{cases} \eta_{EV}^{ch} & \text{if the EV is charging} \\ \frac{1}{\eta_{EV}^{dis}} & \text{if the EV is discharging} \end{cases} \tag{8}$$

$$\begin{cases} SOC_{min}^{EV-i} \leq SOC_{t}^{EV} \leq SOC_{max}^{EV} \\ P_{t}^{ch} \leq CR^{EV} \\ P_{t}^{dis} \leq DR^{EV} \\ T_{arv} \leq t_{ch}^{start} \leq T_{dep} - 1 \\ t_{ch}^{start} + 1 \leq t_{ch}^{stop} \leq T_{dep} \\ t_{ch}^{stop} + 1 \leq t_{dis}^{stop} \leq T_{dep} - 1 \\ t_{dis}^{start} + 1 \leq t_{dis}^{stop} \leq T_{dep} \end{cases} \tag{9}$$

2.1.4. Hydrogen Fueled Gas Turbines

The energy conversion relationship and operational constraints of the HGT are given in (10).

$$\begin{cases} P_{t}^{HGT} = \eta_{HGT} \cdot P_{t}^{H2} \\ P_{min}^{H2-HGT} \leq P_{t}^{H2} \leq P_{max}^{H2-HGT} \\ P_{min}^{HGT} \leq P_{t}^{HGT} \leq P_{max}^{HGT} \end{cases} \tag{10}$$

2.2. Problem Formulation

This study proposes a reinforcement learning (RL) optimization approach for efficiently managing energy resources in the urban transportation system (UTS), aiming to minimize total energy consumption while ensuring minimal user dissatisfaction. The energy management (EM) strategy considers the optimal demand-side management (DSM) of flexible loads. Additionally, the RL algorithm is enhanced to incorporate various system uncertainties.

The problem is structured as a finite Markov decision process (FMDP) to align with the RL framework for decision-making. The objective is to minimize a total cost function, as defined in Equation (11), subject to multiple constraints.

$$\min \sum_{t=1}^{T} [Cost_{t}] \tag{11}$$

2.2.1. Cost Functions

The objective is to minimize the total cost function, which comprises both electricity consumption costs and user dissatisfaction costs, as outlined in Equation (12).

$$Cost_t = [P_t^{E-buy} \cdot C_t^{E-buy} - P_t^{E-sell} \cdot C_t^{E-sell}] \cdot \Delta t + \sum_{n=1}^{N} C_t^{D,n} \quad (12)$$

The dissatisfaction cost $\left(C_t^{D,n}\right)$ reflects the deviation of controllable station load levels from their predetermined set points while accounting for a tolerance factor ε, as specified in Equation (13).

$$C_t^{D,n} = \begin{cases} 0 & \text{if } \left(P_t^{Set,n} - \varepsilon_n\right) \le P_t^n \le \left(P_t^{Set,n} + \epsilon_n\right) \\ C_t^D \cdot df_D \cdot \Delta t \sqrt{\left(P_t^n - \left(P_t^{Set,n} - \varepsilon_n\right)\right)\left(P_t^n - \left(P_t^{Set,n} + \varepsilon_n\right)\right)} & \text{else} \end{cases} \quad (13)$$

2.2.2. Operational Constraints

The amount of power bought and sold is supplied by different resources and is determined as given in (14).

$$\begin{cases} P_t^{E-buy} = P_t^{G-buy} + P_t^{EV-dis} \\ P_t^{E-sell} = P_t^{G-sell} + P_t^{EV-ch} \end{cases} \quad (14)$$

The interaction between multiple energy resources, the railway power system, and the utility grid (UG) is regulated by a power balance condition, as specified in Equation (15).

$$P_t^{PV} + P_t^{HGT} + P_t^{EV-dis} + B_G \cdot P_t^{G-buy} + P_t^{RB} = P_t^{load} + P_t^{EL-PV} + (1 - B_G) \cdot P_t^{G-sell} + P_t^{EV-ch} \quad (15)$$

where $B_G = \begin{cases} 1 \text{ if buying from the grid} \\ 0 \text{ if selling to the grid} \end{cases}$

Electric vehicles (EVs) within the parking garage and E-bus station are accounted for in Equation (15) as energy sources when discharging surplus power (P_t^{EV-dis}) and as loads when drawing additional charging power from the system (P_t^{EV-ch}). However, power is first exchanged internally between vehicles (V2V) before interacting with the railway power system.

Beyond the power balance equality constraint, a set of inequality constraints defines the operational boundaries of flexible loads, limits on power exchange with the utility grid, restrictions on power transfer with the railway power system, and power supply constraints. These constraints ensure that the transportation system's energy requirements are met, as specified in Equation (16).

$$\begin{cases} P_{max}^G + P_{max}^{PV} + P_{max}^{HGT} + P_{max}^{EV-dis} + P_{max}^{RB} \ge P_{max}^{e-load} \\ P_t^G \le P_{max}^G \\ P_t^{TPS} \le P_{max}^{TPS} \\ P_{min}^{light} \le P_t^{light} \le P_{max}^{light} \\ P_{min}^{HVAC} \le P_t^{HVAC} \le P_{max}^{HVAC} \end{cases} \quad (16)$$

As outlined in Equation (16), the controllable loads in railway stations managed by the demand-side management (DSM) algorithm include lighting and HVAC systems. However, vertical transportation loads, such as escalators and elevators, remain fixed and cannot be reduced or rescheduled. Furthermore, the total electrical load of the transportation system (P_t^{e-load}) encompasses the railway stations' loads, the power supplied to the electrolyzers at the hydrogen production facility, and the energy required for EV charging.

2.2.3. FMDP Formulation

The energy management and demand-side management (EM-DSM) problem is modeled as a multi-agent finite Markov decision process (FMDP). This formulation involves three fundamental components: states, actions, and rewards.

The selected RL approach is motivated by the sequential, stochastic, and decentralized nature of the urban transportation energy management problem, where system states evolve in response to both exogenous uncertainties and endogenous control actions. The problem is formulated as a factored Markov decision process (FMDP), which enables decomposition of the global decision space into interacting subproblems associated with metro stations, EVs, E-buses, and hydrogen assets. This factorization significantly reduces the effective state–action space and allows the use of multi-agent Q-learning without requiring global function approximation.

- States

At time $\boldsymbol{t}$, the state vector $\boldsymbol{st}$ consists of several sub-vectors that represent the power generated by photovoltaics (PV), vertical transportation loads, regenerative braking power, and the prices for buying and selling energy, as defined in (17).

The state vector in (17) is designed to capture all exogenous variables that directly influence energy management decisions but are not controllable by the agent. The selection of these state variables follows the Markov property, ensuring that the current state contains sufficient information to determine future system evolution without requiring historical data. Each state component has a direct impact on the system cost and constraints.

$$s_t = \left[sP_t^{PV}, sP_t^{V-load}, sP_t^{RB}, sp_t^B, \ sp_t^S \right] \tag{17}$$

Due to varying accuracies in forecasting PV generation, energy prices, regenerative braking, and vertical transportation load consumption, the EM-DSM problem involves mixed uncertainty conditions. These uncertainties stem from the unpredictable nature of PV generation, load consumption, regenerative braking power, and energy buying/selling prices (P_t^{PV}, P_t^{V-load}, P_t^{RB}, p_t^B, p_t^S), which impacts the problem states. To address these uncertainties, at the beginning of each episode, realizations of PV generation, loads, regenerative braking power, and energy prices are sampled from their respective probability distributions, generating a single trajectory of state transitions for the agent. By augmenting the state representation with stochastic environment sampling, the proposed FMDP formulation allows the agent to learn policies that are robust under mixed uncertainty conditions rather than relying on deterministic forecasts.

Uncertainty is now handled exclusively through stochastic environment dynamics, consistent with canonical RL formulations. Upon the realizations of the system's states, the reward function is defined based on the realized operating cost, and robustness is achieved implicitly through repeated exposure to diverse uncertainty realizations across episodes. This revision removes any ambiguity with stochastic programming and ensures internal consistency between state transitions, rewards, and learning updates.

- Actions

At time t, the action vector a_t, as defined in (18), consists of several sub-vectors that represent the controllable load levels (P_t^{Load}) of Sn metro stations and the charging/discharging power of electric vehicles (P_t^{EV} and P_t^{Ebus}).

$$a_t = \left[aP_t^{Load}, \ aP_t^{EV-ch}, \ aP_t^{EV-dis} \right] \tag{18}$$

where

$$\begin{cases} aP_t^{Load} = \left[aP_t^{S1}, \ldots, aP_t^{Si}, \ldots, aP_t^{Sn}\right] & \\ aP_t^{Si} = \left[aP_t^{Light-i}, aP_t^{HVAC-i}\right] & \\ aP_t^{EV} = \left[aP_t^{ch-EV}, aP_t^{dis-EV}, aP_t^{ch-Ebus}, aP_t^{dis-Ebus}\right] & \\ aP_t^{Light-i} = \left[P_t^{L1}, \ldots, P_t^{Lm}, \ldots\right] & Lm \in A_{Light} \\ aP_t^{HVAC-i} = \left[P_t^{H1}, \ldots, P_t^{Hm}, \ldots\right] & Hm \in A_{HVAC} \\ aP_t^{ch-EV} = \left[P_t^{cEV1}, \ldots, P_t^{cEVm}, \ldots\right] & cEVm \in A_{ch-EV} \\ aP_t^{dis-EV} = \left[P_t^{dEV1}, \ldots, P_t^{dEVm}, \ldots\right] & dEVm \in A_{dis-EV} \\ aP_t^{ch-Ebus} = \left[P_t^{cEbus1}, \ldots, P_t^{cEbusm}, \ldots\right] & cEbusm \in A_{ch-Ebus} \\ aP_t^{dis-Ebus} = \left[P_t^{dEbus1}, \ldots, P_t^{dEbusm}, \ldots\right] & dEbusm \in A_{dis-Ebus} \end{cases} \quad (19)$$

The action vector defined in (18)–(19) represents all decision variables that can be directly controlled by the energy management system at each time step. These include: adjustable load levels of metro station subsystems (lighting and HVAC), and charging and discharging powers of electric vehicles and electric buses.

The action space is structured hierarchically to reflect the physical and operational structure of the system: station-level decisions are decomposed into subsystem-level actions, while EV-related actions are separated into charging and discharging modes. Each action is constrained within predefined feasible sets (A_{Light}, A_{HVAC}, A_{ch}, A_{dis}), derived from technical limits, comfort constraints, and battery operating boundaries.

This design ensures that all actions selected by the learning agent are physically realizable, avoids infeasible exploration during training, and improves convergence stability of the multi-agent reinforcement learning process.

- Reward

The reward function ***rt*** represents the benefit obtained at time ***t*** for a given state-action pair (action ***at***—state ***st***). It is defined as the negative of the cost function, with the total reward over the simulation period R formulated in (20) and (21), respectively, which aligns the reinforcement learning objective with the original optimization goal of cost minimization.

Additional uncertainties arise from the forecasting of various loads, regenerative braking (RB), PV generation, and energy prices. These uncertainties are modeled through the stochastic dynamics of the environment, consistent with the finite Markov decision process (FMDP) formulation. Each source of uncertainty is characterized by a probability distribution.

At the beginning of each training episode, realizations of the uncertain variables are sampled from their respective distributions, and the environment evolves according to these sampled trajectories over the episode horizon. At each time step ***t***, the agent observes a single realized system state ***st***, selects an action ***at***, and transitions to the next state ***st+1*** according to the stochastic environment dynamics. This preserves the Markov property and avoids explicit enumeration of multiple scenarios at a given decision step.

The instantaneous reward ***rt*** is computed based on the realized operational cost associated with the observed state–action pair. Over repeated interactions, the agent experiences a diverse set of state transitions induced by stochastic realizations of renewable generation, loads, regenerative braking, and energy prices. In this manner, the learned policy converges toward minimizing the expected long-term cost across the underlying uncertainty distributions.

It is emphasized that scenario probabilities are not used to optimize over parallel futures, as in stochastic programming. Instead, they govern the frequency with which

different operating conditions are encountered across episodes, allowing the reinforcement learning agent to implicitly learn robust decision policies through experience. The learning process, therefore, remains fully consistent with standard Q-learning and FMDP theory. The reward associated with each state-action pair is then computed using the modified reward function defined in (20).

$$r_t = -Cost_t(w_t) \tag{20}$$

$$R = \sum_{t=1}^{T} r_t \tag{21}$$

where w_t denotes the realized uncertainty sample at time t.

2.2.4. Multi-Agent RL Algorithm

In the reinforcement learning (RL) decision-making framework, the Q-learning algorithm is utilized to optimize the expected cumulative rewards. The Bellman equation, as presented in (22), calculates the Q-values for each state-action pair $Q(s_t,a_t)$, offering a precise reward estimation and updating based on the learning parameter γ.

$$Q(s_t, a_t) = r_t + \gamma \cdot \max[Q(s_{t+1}, a_{t+1})] \tag{22}$$

Throughout the learning process, the Q-table is iteratively updated during each training cycle. This continuous update process ensures the selection of the optimal action with the highest Q-value for each state, based on the corresponding reward [28].

In this study, the action vector consists of 44 sub-vectors, as defined in (19), representing the loads of 22 agents. These agents correspond to the controllable lighting and HVAC loads of 20 metro stations, along with an electric vehicle (EV) parking garage and an electric bus (E-bus) station.

2.3. Methodological Scope and Limitations

This study adopts a decentralized, tabular Q-learning framework to address the energy management and demand-side management (EM–DSM) problem under mixed uncertainty conditions. The methodological scope is intentionally focused on discretized state and action representations, derived from physical operating limits, tariff structures, and demand categories, which enable stable and interpretable learning within a simulation-based environment.

The reinforcement learning formulation is designed to capture system-level coordination rather than algorithmic novelty. Each agent operates over a constrained local decision space, and no centralized joint state–action table is constructed. As a result, tabular Q-learning remains computationally tractable and suitable for the considered problem scale, while avoiding the training instability and high data requirements commonly associated with deep reinforcement learning methods.

Nevertheless, the proposed approach has several limitations. First, discretization inevitably reduces resolution and may smooth fine-grained dynamics present in continuous control systems. Second, tabular learning does not scale efficiently to fully continuous state–action spaces, and therefore may not be appropriate for systems with significantly higher dimensionality or unbounded operating ranges. Third, the learned policies are optimized for the statistical characteristics of the simulated scenarios and may require retraining if system configurations or uncertainty distributions change substantially.

Despite these limitations, the proposed framework offers a transparent, computationally efficient, and extensible foundation for coordinated EM–DSM. Future work will investigate the integration of function approximation, deep reinforcement learning, and

hybrid learning–optimization methods to address larger-scale systems and continuous control settings.

3. Case Study

3.1. System Description

The urban transit system (UTS) analyzed in this study is an electrified metro network in Alexandria, linking the densely populated northeastern town of Abou-Qir to downtown Alexandria. Based on the technical design, 22 electric trains, each comprising at least nine cars, are required to meet the targeted service level. The metro line extends over 21.7 km and includes 20 modern elevated stations, four of which function as interchange stations, connecting the metro to other railway lines and transportation systems [29].

Among these transportation systems is a proposed electric bus (E-bus) station at Abou-Qir, designed to accommodate up to 100 buses. Additionally, an EV parking garage with a capacity of 400 electric vehicles (EVs) is planned at the Sidi-Gaber station [2]. A schematic representation of the UTS is provided in Figure 2.



Figure 2. The UTS under study.

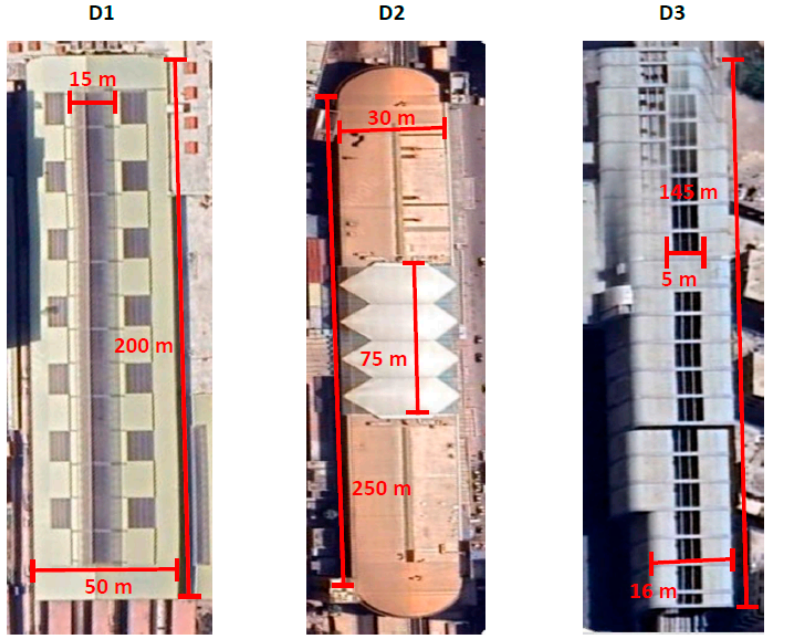
To promote the integration of sustainable energy sources, rooftop photovoltaic (PV) units will be installed at selected elevated metro stations, the way-sides of metro lines and on the canopies of the E-bus station. Moreover, to optimize the utilization of surplus PV-generated power, a green hydrogen production facility is proposed as an addition to the system, as illustrated in Figure 2. The produced green hydrogen will be used to fuel hydrogen gas turbines (HGTs) to contribute to the system's energy demand.

The optimal scheduling of energy selling, purchasing, and storage is determined by the energy management and demand-side management (EM-DSM) algorithm, which adapts to fluctuations in energy prices.

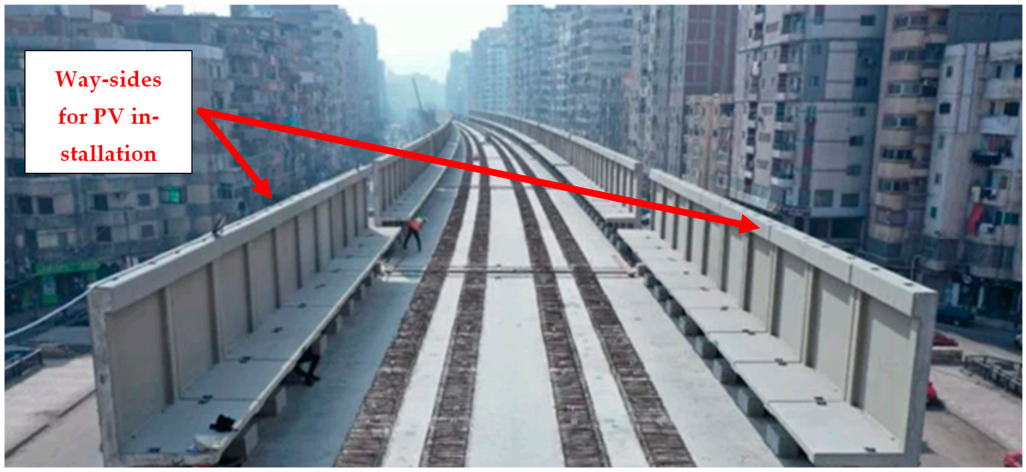
3.1.1. Metro Line Stations and Way-Sides

Concerning the PV installation areas in metro stations, three distinct station building designs are considered, as depicted in Figure 3a. As previously stated, the network includes four interchange stations. The first interchange station, Masr Station, follows a unique

architectural design (D1) due to its historical significance. While upgrades are possible, structural modifications are restricted.



(**a**) Available area for PV installation in different metro stations



(**b**) Available area for PV installation along metro line's way-sides

Figure 3. Available area for PV installation in different metro stations and way-sides.

The other three interchange stations feature a more contemporary design (D2), similar to that of the second interchange station, Sidi-Gaber. Meanwhile, the remaining 16 elevated stations will adopt a design corresponding to (D3).

The PV areas available for installation, as shown in Figure 3, represent the remaining rooftop spaces of the stations after excluding skylight-covered sections. Considering the PV installation capacity at each station, the total available area across all 20 stations amounts to approximately 48,700 m^2. As for the available PV installation area along the way-sides, according to the direction of the way-side, an area of 14 km long by 3 m height is considered, with an approximate area of 42,000 m^2. The specifications of the PV arrays installed on the metro station rooftops are detailed in Table 2. The annual distribution of average solar radiation is depicted in Figure 4. As illustrated, the highest peak radiation of 879 W/m^2 occurs in May, while the lowest peak radiation of 278 W/m^2 is recorded in January.

Table 2. The data of the PV array used on rooftops of metro stations and E-bus station canopies.

η^{PV}	95%
P_{max}^{PV}	200 W/1.6 m^2
SR_{STC}	1000 W/m^2
tc	−0.41%/°C
T_{STC}^{PV}	25 °C

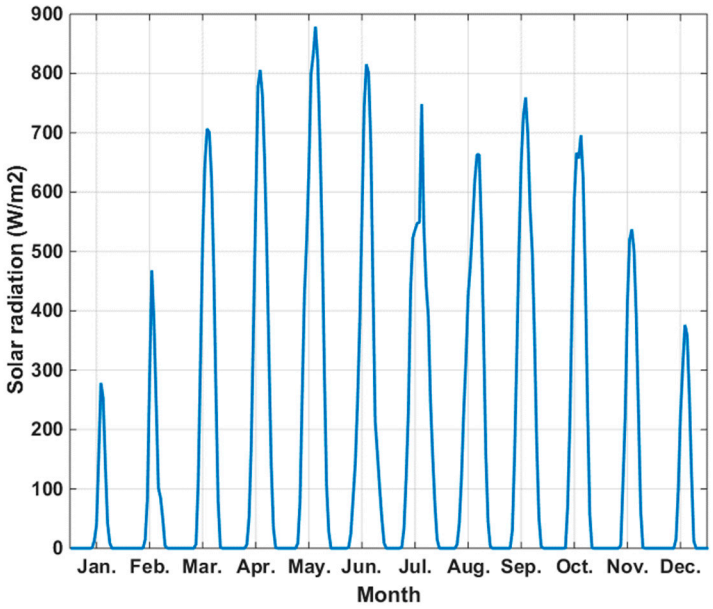


Figure 4. The average solar radiation during a year.

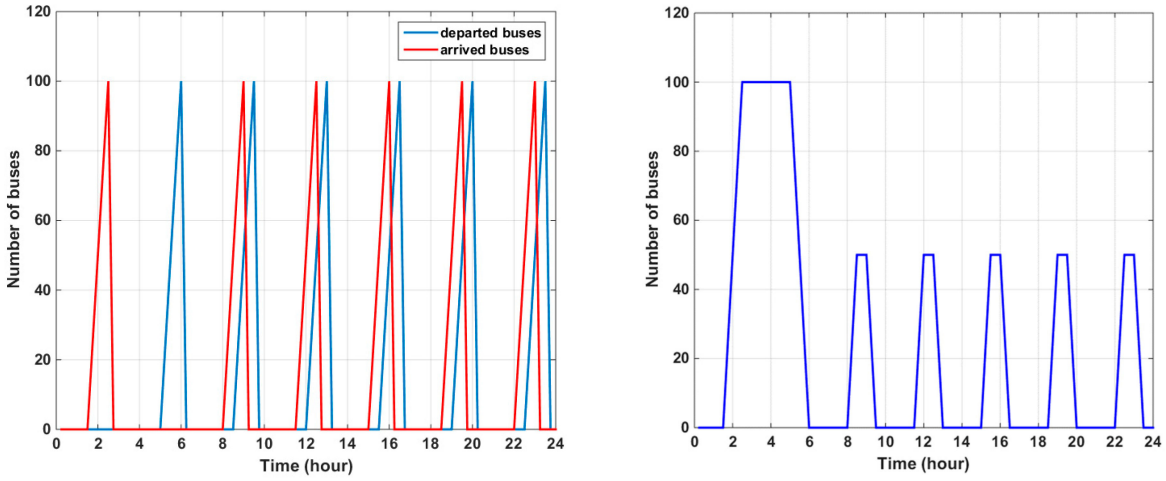
3.1.2. E-Bus Station

The UTS will feature an E-bus charging station, similar to the one illustrated in Figure 5, designed to simultaneously charge up to 100 buses using dispensers with a maximum output power of 150 kW. Each bus is equipped with a 462 kWh battery and has an energy consumption rate of 2.9 kWh/km. The permissible state of charge (SOC) ranges from a minimum of 20% to a maximum of 95% of the battery's full capacity [30]. The charging station will also include five canopies covering a total area of 780 m^2, allocated for PV unit installation.



Figure 5. An E-bus charging station.

The charging and discharging processes of the E-bus batteries follow the EV model outlined in Equations (4)–(8). The arrival and departure schedules of the fleet are depicted in Figure 6a, showing a consistent operational pattern. Departures begin at 5:00 AM, with 25 buses leaving the station every 15 min. The duration required to complete a route varies between 2 and 4 h, after which the buses return to the station for a 30 min dwell time before their next departure. The resulting occupancy levels of the E-bus station are presented in Figure 6b.



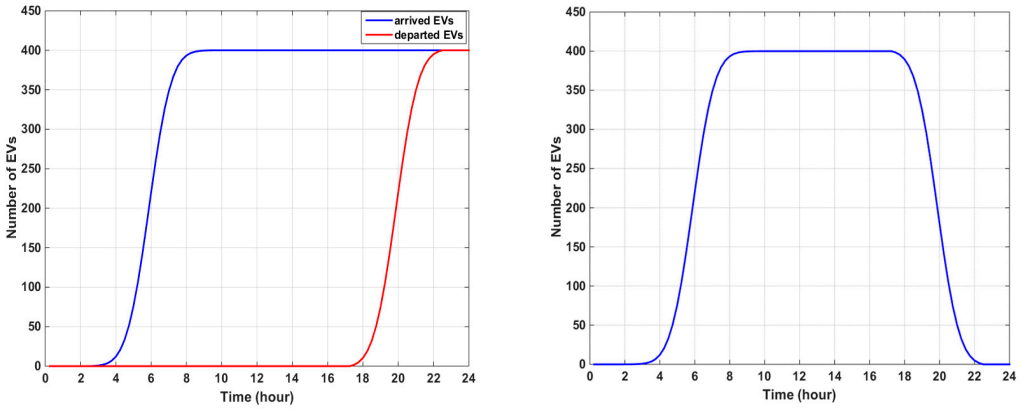
(**a**) (**b**)

Figure 6. (**a**) Total number of buses arrived/departed during 24 h, (**b**) Occupation level of the E-bus station.

During their stay at the station, E-buses have the capability to exchange power with both the railway's power system and the electricity grid. However, this energy exchange is contingent upon their SOC, which exhibits considerable uncertainty.

3.1.3. EVs Parking Garage

In contrast to E-buses, the arrival and departure patterns of EVs are modeled using normal distribution functions with distinct means and standard deviations, as shown in Figure 7a. The mean arrival time is 6:00 AM with a standard deviation of 60 min, while the mean departure time is 8:00 PM, also with a standard deviation of 60 min. Based on these patterns, the occupancy level of the EV parking garage is illustrated in Figure 7b.



(**a**) (**b**)

Figure 7. (**a**) Total number of EVs arrived/departed during 24 h; (**b**) Occupation level of the EVs parking garage.

This study incorporates a variety of EVs with different battery capacities, as well as varying charging and discharging rates. The distribution of battery capacities among EVs is assumed to follow a normal distribution. Additionally, the initial state of charge (SOC) is considered to follow a uniform distribution within the range of 20% to 50% of the battery's total capacity. The specifications of the EVs are detailed in Table 3 [2].

Table 3. EVs' Data [2].

Batteries' Rated Capacity (kWh)	CR^{EV}/DR^{EV} (kW)	J^{i}_{EV}	SOC^{max}_{EV} (%)	SOC^{min}_{EV} (%)	Mean		Standard Deviation	
					Arrival	Departure	Arrival	Departure
8	1.6	1–3	95	20	6 a.m.	8 p.m.	60 min	60 min
17	3.4							
18	3.6							
48	9.6							

3.1.4. Green-Hydrogen Generating Station & HGT

The green hydrogen station will utilize excess PV energy that remains unconsumed by the railway stations. It is capable of producing up to 1300 Nm^3/h of hydrogen gas, with an average energy consumption of 4.5 kWh/Nm^3. As previously stated, the produced hydrogen will be used to power hydrogen-fueled microturbines. Consequently, the station is equipped with 10 hydrogen-fueled microturbines, each with a capacity of 300 kW and an efficiency of 40.3%.

3.2. PV System Integration and Optimal Energy Management

Figure 8 illustrates the available regenerative braking energy along with the internal load consumption across the 20 metro stations, while Figure 9 presents the PV power generation from metro stations, line's way-sides and E-bus station.

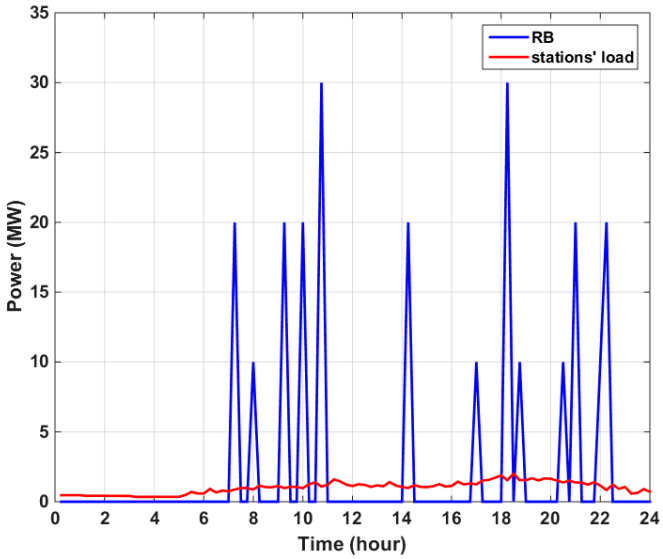


Figure 8. The available regenerative braking and the stations' internal load consumption of the 20 metro stations.

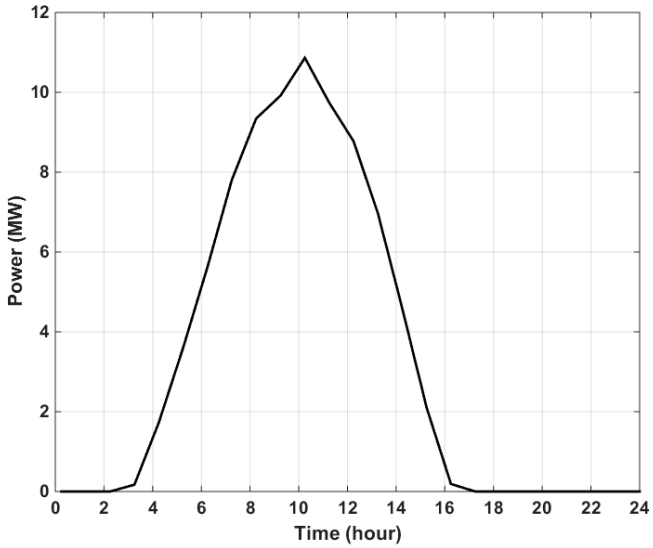


Figure 9. The available PV-generated power of metro stations, line's way-sides and E-bus station.

Furthermore, the energy trading prices with the utility grid are presented in Figure 10. As shown, the selling price of energy to the utility grid is approximately 66% of the purchasing price. However, since the energy management algorithm accounts for internal energy trading among system components, the associated prices differ from those of transactions with the utility grid. In this study, the price of energy exchanged between EVs in a parking lot is set at 90% of the utility grid's selling price, while the price for energy traded between the UTS and EVs is assumed to be 80% of the utility grid's selling price.

As integral components of the UTS, the E-bus station, green hydrogen production facility, hydrogen-fueled turbines (HGTs), and PV units engage in energy trading with metro stations at no cost.

This study considers 22 agents within the RL algorithm, as detailed in (19). To define the action sets for demand-side management (DSM) of lighting and HVAC loads across the 20 metro stations, acceptable load levels are specified in Table 4. For the two agents representing the E-bus station and the EV parking garage, the action sets—corresponding to vehicle charging and discharging—are determined as a percentage of maximum capacity while ensuring compliance with SOC limits, as indicated in Table 4.

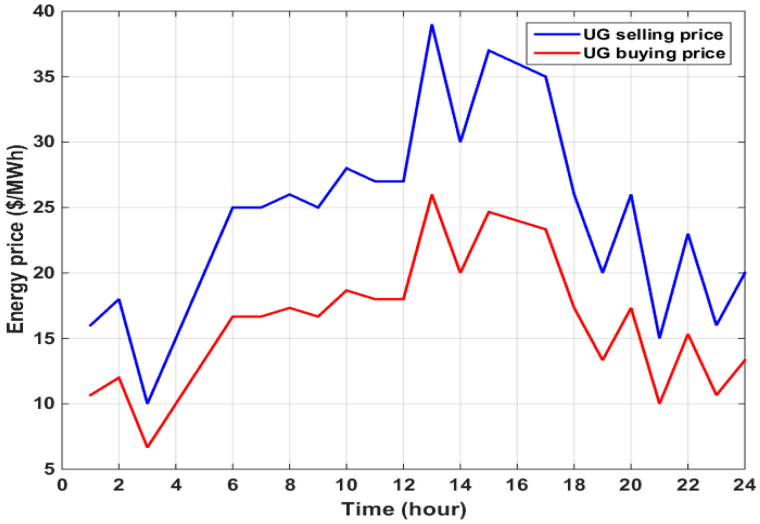


Figure 10. The energy trading prices with the utility grid.

Table 4. The action sets of RL agents.

Agent ID	Action Set
$aP_t^{Light-i}$	[0.75, 0.85, 0.95, 1] of P_{max}^{light}
aP_t^{HVAC-i}	[0.7, 0.8, 0.9, 1] of P_{max}^{light}
aP_t^{ch-EV}	[0.6, 0.7, 0.8, 0.9] of P_{max}^{EV}
aP_t^{dis-EV}	[0.25, 0.35, 0.45, 0.55] of P_{max}^{EV}
$aP_t^{ch-Ebus}$	[0.6, 0.7, 0.8, 0.9] of P_{max}^{E-bus}
$aP_t^{dis-Ebus}$	[0.25, 0.35, 0.45, 0.55] of P_{max}^{E-bus}

Although the urban transportation energy system includes 44 controllable sub-actions across 22 agents, a centralized reinforcement learning formulation would lead to a prohibitively large joint action space. Based on the action discretization summarized in Table 4, each controllable variable is discretized into four levels. Consequently, each metro station agent—controlling lighting and HVAC loads—has $4 \times 4 = 16$ feasible actions, while the EV parking and E-bus station agents each have $4 \times 4 = 16$ actions associated with charging and discharging decisions. A centralized controller would therefore face a theoretical joint action space of 16^{22}, which is computationally intractable for tabular learning.

Therefore, to overcome this challenge, the problem is formulated as a factored Markov decision process (FMDP) and solved using a decentralized multi-agent Q-learning framework. In this formulation, each agent learns an independent Q-function over its own local state and action space, eliminating the need to enumerate or store joint state–action pairs. As a result, the learning complexity scales linearly with the number of agents rather than exponentially with the number of controllable variables.

With the adopted discretization, each agent's action space is limited to 16 actions, and the corresponding Q-table size is given by $|Qi| = |Si| \times 16$, where $|Si|$ denotes the number of discrete local states observed by agent $\boldsymbol{i}$. For the considered system, this results in Q-tables of manageable size and enables stable convergence within a practical training time. Therefore, despite the large theoretical joint action space, the proposed RL framework remains computationally feasible for the studied system scale. Extensions to higher-resolution control or larger transportation networks would require function approximation techniques, which are left for future work.

To evaluate the impact of PV integration on the transportation power system and optimized energy management, three scenarios are examined. In the first scenario, which is the base scenario, all energy transactions within the transportation system occur exclusively through the utility grid (UG). Under this condition, metro station internal loads are supplied by the UG without DSM, and any available regenerative braking (RB) energy is also sold to the UG. Additionally, the charging and discharging of EVs in the parking lot and E-bus station are conducted solely via the UG, with no direct energy exchange between vehicles. Moreover, this scenario does not include green hydrogen production, meaning that HGTs do not contribute energy to the transportation system.

In the second scenario, the RL-based energy management (RLEM) algorithm is applied, incorporating both PV generation and hydrogen-fueled turbines (HGTs) into the system. This scenario allows internal energy trading among EVs; however, no direct energy exchange takes place between EVs and the UTS. The remaining energy required to meet the demands of the UTS, including metro stations and E-buses, is supplied by the utility grid (UG), while any surplus energy generated within the system is sold back to the UG.

In the third scenario, the RL-based energy management algorithm is implemented with an added priority for energy trading between the UTS and the EV parking lot before

engaging in transactions with the UG. While energy exchange among E-buses remains free of charge, it reduces reliance on the UG, thereby lowering the cost of purchased energy. Furthermore, this pricing strategy ensures that internal energy trading between EVs and among E-buses takes precedence before transactions with other system components. Likewise, energy exchange between the UTS and EVs is prioritized over trading with the UG.

Figure 11 presents the internal power consumption of metro stations before and after the implementation of demand-side management (DSM) within the RL-based energy management (RLEM) algorithm. As observed in Figure 11, shifting loads from peak price periods to off-peak intervals has resulted in approximately a 3% reduction in energy costs when relying on the utility grid (UG) for power supply.

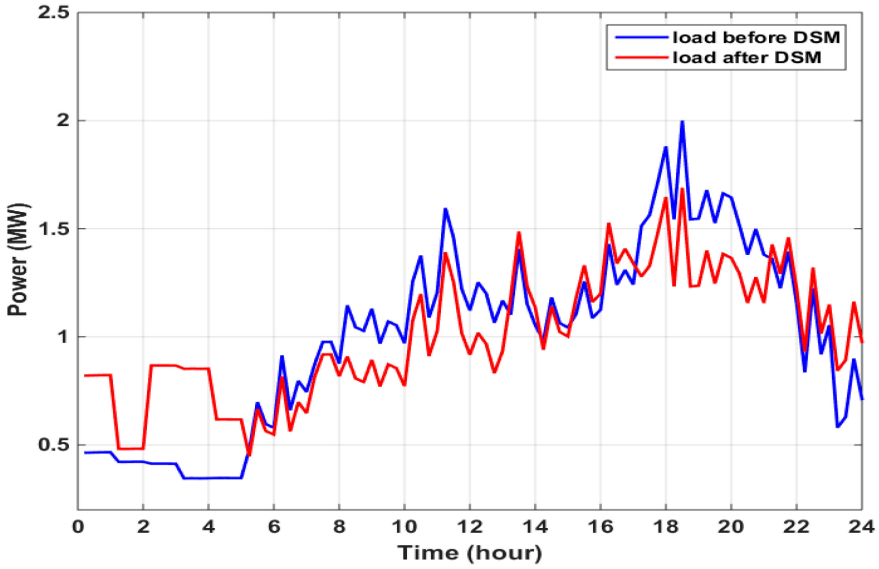


Figure 11. The internal power consumption of the metro stations before and after applying DSM.

Table 5 summarizes the energy trading costs among the UTS, EVs, and UG across different scenarios. As indicated in Table 5, with the integration of RLEM, PV units, green hydrogen production, and hydrogen-fueled turbines (HGTs) into the UTS, the revenue from energy sold to the UG in Scenario 2 increased by 62.5% compared to Scenario 1. This rise is attributed to the incorporation of renewable energy sources within the UTS and internal energy trading between E-buses. In Scenario 3, where energy trading between the UTS and EVs is also introduced, the revenue from UTS energy sales increased by 58.9% relative to Scenario 1 but decreased by 2.17% compared to Scenario 2 because of the energy sold to EVs.

Likewise, the cost of energy purchased by the UTS from the UG in Scenario 2 decreased by 40.2% compared to Scenario 1. Additionally, in Scenario 3, with energy trading between the UTS and the EV parking lot taken into account, the total cost of energy acquired by the UTS from both the UG and discharging EVs amounted to $3440, representing a 40.2% reduction relative to Scenario 1.

Figure 12 presents the power consumption of the UTS loads alongside the power supplied by its various energy sources. As illustrated, there are periods throughout the day when the energy demand of the UTS surpasses the available supply from internal resources. To address this shortfall, additional power is drawn from the utility grid (UG) and the EV parking lot, as detailed in Table 5, ensuring compliance with the power balance constraint specified in Equation (14). Conversely, at other times, the power generated by UTS energy sources exceeds the system's load demand. During these intervals, the excess

energy is traded with both the UG and the EV parking lot, maintaining adherence to the power balance constraint in Equation (14).

Table 5. The cost of energy traded between the UTS, EVs and UG.

Energy Traded		Energy Cost ($)		
		Scenario 1	**Scenario 2**	**Scenario 3**
UTS and UG	Sold	4.7533×10^3	7.724×10^3	7.5557×10^3
	bought	5.7568×10^3	3.4427×10^3	3.4297×10^3
EVs and UG	Sold	136.5771	11.8405	3.1739
	bought	450.8924	263.7876	11.3364
EVs and UTS	Sold	X	X	10.4
	bought	X	X	201.961
EVs internal trading		X	168.3943	168.3943

X: not applicable.

With regard to the influence of RLEM on EV energy trading costs, a comparison between Scenario 2 and Scenario 1 is provided in Table 5. The internal trading among EVs within the parking lot resulted in a 41.5% reduction in the cost of energy purchased from the UG and a 91.3% decrease in the cost of energy sold to the UG. However, the overall revenue generated from energy sold by discharging EVs—both to the UG and to other EVs for charging—increased by 31.9%, while the total cost of energy procured from both the UG and discharging EVs decreased by 4.1%.

In Scenario 3, incorporating both internal energy exchanges among EVs and energy transactions with the UTS resulted in an additional 15.3% reduction in the total cost of purchased energy and a 33.2% increase in revenue from sold energy compared to Scenario 1.

Figure 13a shows the power needed to charge EVs and E-buses, the power available from discharging EVs and E-buses, and the power traded internally in the EVs parking lot and E-bus station, while Figure 13b shows the power traded externally with the system.

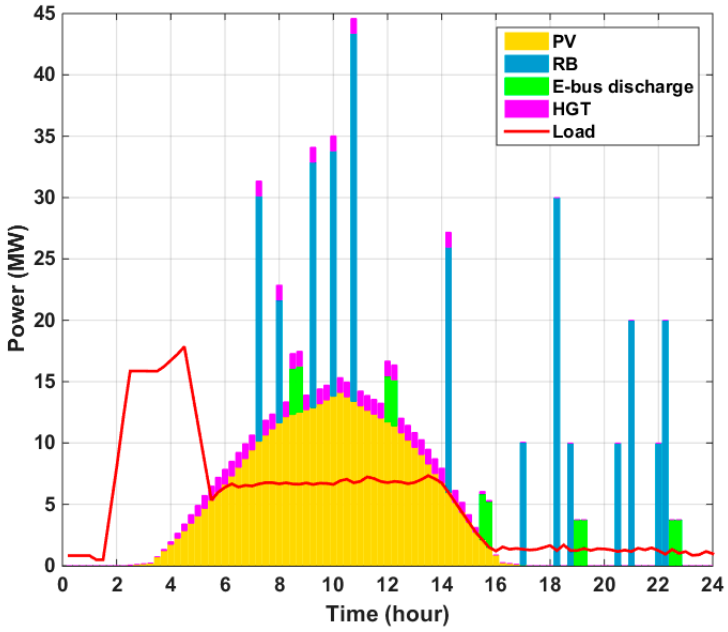
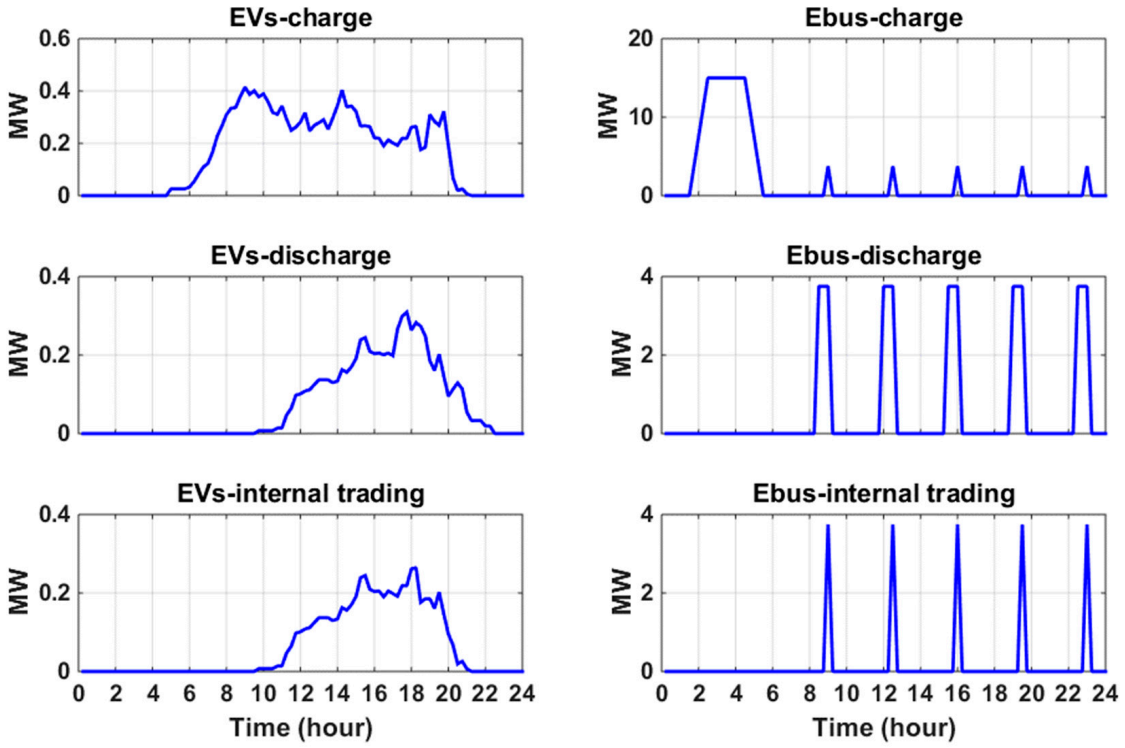
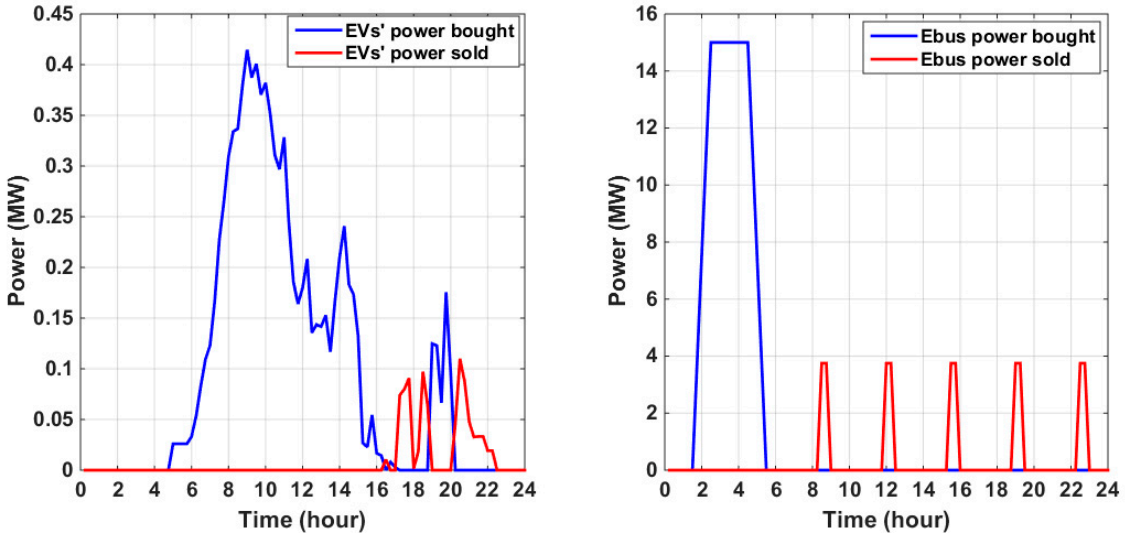


Figure 12. The power consumed by the UTS's loads and the power supplied to it by different energy resources.



(**a**)



(**b**)

Figure 13. (**a**) The power needed to charge EVs and E-buses, the power available from discharging EVs and E-buses, and the power traded internally in the EVs parking lot and E-bus station; (**b**) The power traded externally with the system.

Figure 14 depicts the convergence of energy costs during the training process of the RL algorithm for both scenarios 2 and 3. The training is conducted over 10,000 iterations with a learning parameter (γ) of 0.8. As shown in Figure 14, the cost value stabilizes at a minimum, as anticipated. Furthermore, Table 5 reveals that the revenue from energy sold by the UTS surpasses the cost of energy purchased, resulting in the convergence of the objective value to a negative value, as illustrated in Figure 14.

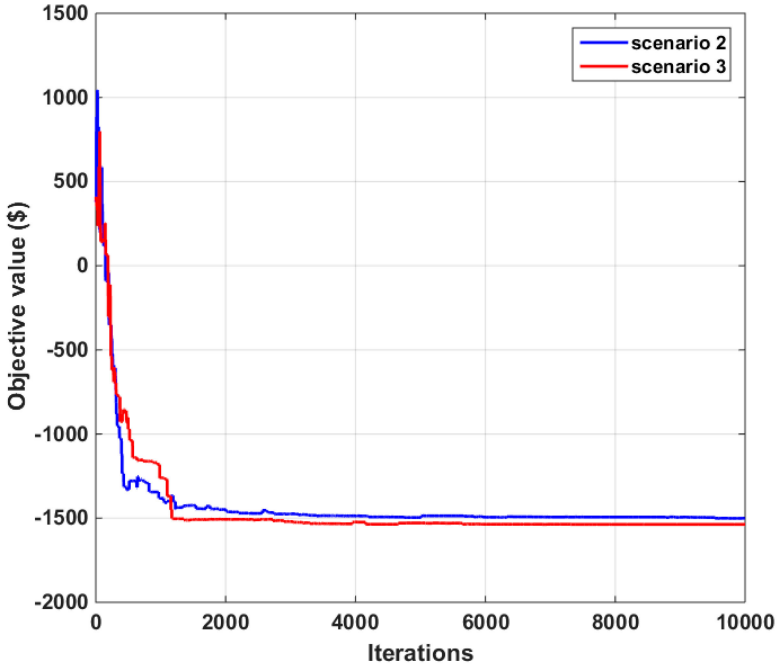


Figure 14. The convergence of the objective value through the training process of the RL algorithm both in scenario 2 and 3.

4. Conclusions

This paper investigated the electrified transportation system in Alexandria to evaluate the impact of integrating renewable energy sources (RES) into an urban transit system (UTS). To enhance the system's efficiency and sustainability, rooftop photovoltaic (PV) units and green-hydrogen production for hydrogen-powered gas turbines were incorporated as clean and renewable energy solutions. This integration ensured that any surplus PV-generated power exceeding the system's demand is effectively utilized rather than wasted. The examined UTS consists of an electrified metro line, a fleet of electric buses, and designated parking lots for electric vehicles (EVs).

To maximize the benefits of RES integration, an optimal management strategy for the integrated energy system (IES) of the UTS was developed. The proposed energy management approach, which included demand-side management (DSM) for UTS loads and EVs, introduced complexity in decision-making due to the uncertainty of decision variables and the large search space. To address these challenges, a modified multi-agent reinforcement learning (MRL) algorithm was implemented. The optimization problem was structured as a finite Markov decision process, incorporating mixed uncertainties as additional states and action scenarios within the MRL framework.

The simulation results highlighted the economic advantages of integrating renewable and sustainable energy sources into the IES of the electrified UTS, demonstrating a potential reduction of 40.2% in the average daily energy consumption cost.

Despite these promising results, this study has several limitations. First, the reinforcement learning model relies on discretized state and action spaces and simulation-based training data, which may limit scalability to larger systems with highly continuous dynamics. Second, uncertainties in load demand, renewable generation, and energy prices are modeled using stochastic dynamics, which may not fully capture extreme or rare events. Third, the analysis is based on a specific case study in Alexandria, and system performance may vary under different climatic, market, or regulatory conditions, in addition to simplified modeling of some components of the system.

Future work will focus on addressing these limitations by extending the proposed framework to continuous-state reinforcement learning methods, incorporating more detailed component models, and validating the approach using real-time operational data. Additionally, expanding the framework to multi-regional transit networks and investigating the interaction with emerging electricity markets and carbon pricing mechanisms represent promising directions for further research.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data from this study are available from the corresponding author upon reasonable request.

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

Nomenclature

η^{PV}	the efficiency of the PV array
P_{max}^{PV}	the maximum PV power (MW)
SR_t	the solar radiation at time t (W/m^2)
SR_{STC}	the solar radiation under standard testing conditions (STC) (W/m^2)
t_c	the temperature coefficient of the maximum output power (1/K)
T_t^C	the temperature of the PV cell at time t (K)
T_{STC}^{PV}	the temperature of the PV cell under STC (K)
T_t^{amb}	the ambient temperature at time t (K)
P_t^{H2}	the electrolyzer's output power (MW)
P_t^{EL-PV}	the PV power input to the electrolyzer (EL) at time t (MW)
P_{max}^{EL}, P_{min}^{EL}	the maximum and minimum power input to the EL (MW)
η_{EL}	the conversion efficiency of the EL
Wf_{EV}	the EV's factor of willingness
J_{EV}^i	a factor representing the next journey of the ith EV
E_{max}^{EV}	the maximum capacity of an EV's battery (MWh)
SOC_t^{EV}	the state of charge (SOC) of an EV at time t as percentage of its battery's maximum capacity
SOC_{max}^{EV}, SOC_{min}^{EV}	the percentage maximum and minimum SOC of an EV at time t, respectively
η_{EV}^{ch}, η_{EV}^{dis}	the charging and discharging efficiency of an EV, respectively
P_t^{ch}, P_t^{dis}	the charging and discharging power of an EV (MW), respectively
$\boldsymbol{CR^{EV}}$, $\boldsymbol{DR^{EV}}$	the charging and discharging rates of an EV (MW/h), respectively
$\boldsymbol{T_{arv}}$, $\boldsymbol{T_{dep}}$	the EV's arrival and departure time, respectively
t_{ch}^{start}, t_{ch}^{stop}	the starting and stopping time of EV's charging, respectively
t_{dis}^{start}, t_{dis}^{stop}	the starting and stopping time of EV's discharging, respectively
η_{HGT}	the efficiency of the hydrogen-fueled gas turbine (HGT)
P_t^{HGT}	the output power of the HGT (MW)
P_{max}^{H2-HGT}, P_{min}^{H2-HGT}	the maximum and minimum power input to a HGT (MW), respectively
P_{min}^{HGT}, P_{max}^{HGT}	the minimum and maximum power output of a HGT (MW), respectively
P_t^{E-buy}, P_t^{E-sell}	the electric power bought and sold (MW), respectively
C_t^{E-buy}, C_t^{E-sell}	the energy buying and selling prices ($/MWh)
$C_t^{D,n}$	the dissatisfaction cost of nth load at time t ($)
C_t^D	the dissatisfaction energy price at time t ($/MWh)
df_D	the dissatisfaction coefficient
P_t^n	the consumed power of the nth load at time t (MW)
$P_t^{Set,n}$	the rated power of the nth load at time t (MW)

P_t^{G-buy}, P_t^{G-sell}	the power bought from and sold to the electricity grid at time t (MW)
P_t^{EV-ch}, P_t^{EV-dis}	the power traded with EVs (MW)
P_t^{RB}	the regenerative braking power (MW)
P_t^{load}	the loads power at time t (MW)
P_t^{TPS}	the power traded with the transportation system at time t (MW)
P_t^{light}	the power of the lighting loads at time t (MW)
P_t^{HVAC}	the power of the HVAC loads at time t (MW)

References

1. Bowen, G.; Yang, H.; Zhang, T.; Liu, X.; Wang, X. Technoeconomic analysis of rooftop PV system in elevated metro station for cost-effective operation and clean electrification. *Renew. Energy* **2024**, *226*, 120305. [CrossRef]
2. El-Zonkoly, A. Optimal P2P based energy trading of flexible smart inter-city electric traction system and a wayside network: A case study in Alexandria, Egypt. *Electr. Power Syst. Res.* **2023**, *223*, 109708. [CrossRef]
3. Li, X.; Zhao, Y.; Zhang, W.; Wang, F.; Yin, W.; Liu, K. Photovoltaic potential prediction and techno-economic analysis of China railway stations. *Energy Rep.* **2023**, *10*, 3696–3710. [CrossRef]
4. Shen, X.; Wei, H.; Wei, L. Study of trackside photovoltaic power integration into the traction power system of suburban elevated urban rail transit line. *Appl. Energy* **2020**, *260*, 114177. [CrossRef]
5. Laimon, M.; Yusaf, T. Towards energy freedom: Exploring sustainable solutions for energy independence and self-sufficiency using integrated renewable energy-driven hydrogen system. *Renew. Energy* **2024**, *222*, 119948. [CrossRef]
6. Guan, B.; Liu, X.; Zhang, T.; Wang, X. Hourly energy consumption characteristics of metro rail transit: Train traction versus station operation. *Energy Built Environ.* **2023**, *4*, 568–575. [CrossRef]
7. Yu, Y.; You, S.; Ye, T.; Wang, Y.; Guo, X.; Chen, W.; Liu, T.; Wei, S.; Na, Y. Characteristics and assessment of the electricity consumption of metro systems: A case study of Tianjin, China. *Energy Sci. Eng.* **2023**, *11*, 4408–4420. [CrossRef]
8. Guan, B.; Yang, H.; Li, H.; Gao, H.; Zhang, T.; Liu, X. Energy consumption characteristics and rooftop photovoltaic potential assessment of elevated metro station. *Sustain. Cities Soc.* **2023**, *99*, 104928. [CrossRef]
9. Savaş, S.; Külahcı, K. Machine Learning-Based Energy Consumption and Carbon Footprint Forecasting in Urban Rail Transit Systems. *Appl. Sci.* **2026**, *16*, 1369. [CrossRef]
10. Gad, K.; Tonini, F.; Agati, G.; Borello, D.; Colombo, E. Energy demand prediction and scenario analysis for vehicle traction: A bottom-up approach applied to the Italian railway system. *Energy Rep.* **2025**, *13*, 4196–4208. [CrossRef]
11. Tang, Z.; Yin, H.; Yang, C.; Yu, J.; Guo, H. Predicting the electricity consumption of urban rail transit based on binary nonlinear fitting regression and support vector regression. *Sustain. Cities Soc.* **2021**, *66*, 102690. [CrossRef]
12. He, D.; Teng, X.; Chen, Y.; Liu, B.; Wu, J. Piston wind and energy saving based on the analysis of fresh air in the subway system. *Sustain. Energy Technol. Assess.* **2022**, *50*, 101805. [CrossRef]
13. Cheng, Y.; Fang, C.; Yuan, J.; Zhu, L. Design and application of a smart lighting system based on distributed wireless sensor networks. *Appl. Sci.* **2020**, *10*, 8545. [CrossRef]
14. Akbari, S.; Hashemi-Dezaki, H.; Fazel, S. Optimal clustering-based operation of smart railway stations considering uncertainties of renewable energy sources and regenerative braking energies. *Electr. Power Syst. Res.* **2022**, *213*, 108744. [CrossRef]
15. Gabaldón, A.; García-Garre, A.; Ruiz-Abellón, M.C.; Guillamón, A.; Molina, R.; Medina, J. Management of railway power system peaks with demand-side resources: An application to periodic timetables. *Sustainability* **2023**, *15*, 2746. [CrossRef]
16. Mohajer, B.K.; Mousavi, S.M.G. Energy management optimization in smart railway stations with the ability to charge plug-in hybrid electric vehicles. *J. Energy Storage* **2023**, *70*, 107867. [CrossRef]
17. Vamvakas, D.; Michailidis, P.; Korkas, C.; Kosmatopoulos, E. Review and Evaluation of Reinforcement Learning Frameworks on Smart Grid Applications. *Energies* **2023**, *16*, 5326. [CrossRef]
18. Li, H.; Dai, X.; Goldrick, S.; Kotter, R.; Aslam, N.; Ali, S. Reinforcement Learning for EV Fleet Smart Charging with On-Site Renewable Energy Sources. *Energies* **2024**, *17*, 5442. [CrossRef]
19. Liu, X.; Yeh, S.; Plötz, P.; Ma, W.; Li, F.; Ma, X. Electric bus charging scheduling problem considering charging infrastructure integrated with solar photovoltaic and energy storage systems. *Transp. Res. Part E* **2024**, *187*, 103572. [CrossRef]
20. Liu, X.; Liu, X.C.; Xie, C.; Ma, X. Impacts of photovoltaic and energy storage system adoption on public transport: A simulation-based optimization approach. *Renew. Sustain. Energy Rev.* **2023**, *181*, 113319. [CrossRef]
21. Jarvis, P.; Climent, L.; Arbelaez, A. Smart and sustainable scheduling of charging events for electric buses. *TOP* **2024**, *32*, 22–56. [CrossRef]
22. Wu, F.; Yang, J.; Li, B.; Crisostomi, E.; Rafiq, H.; Rashed, G.I. Uncertain scheduling potential of charging stations under multi-attribute uncertain charging decisions of electric vehicles. *Appl. Energy* **2024**, *374*, 124036. [CrossRef]

23. Fan, P.; Bu, S.; Li, S.; Fang, S.; Zhang, C.; Ke, S. Resilience enhancement strategy for multi-microgrids with electric vehicles under extreme weather conditions. *CSEE J. Power Energy Syst.* **2026**, 1–12. [CrossRef]
24. Ke, S.; Zhang, K.; Mai, W.; Guo, R.; He, S.; Tian, J.; Chung, C.Y. Maximizing intra-day V2G feasible capacity of EVs: A cross-disciplinary approach with traffic flow and prospect theory. *IEEE Trans. Transp. Electr.* **2026**. [CrossRef]
25. Guan, B.; Li, H.; Yang, H.; Zhang, T.; Liu, X.; Wang, X. Leveraging cost-effectiveness of photovoltaic-battery system in metro station under time-of-use pricing tariff. *J. Clean. Prod.* **2024**, *434*, 140268. [CrossRef]
26. Taneja, S.; Jain, A.; Bhadoriya, Y. Green Hydrogen as a Clean Energy Resource and Its Applications as an Engine Fuel. *Eng. Proc.* **2023**, *59*, 159. [CrossRef]
27. Yang, Y.; Xu, X.; Luo, Y.; Liu, J.; Hu, W. Distributionally robust planning method for expressway hydrogen refueling station powered by a wind–PV system. *Renew. Energy* **2024**, *225*, 120210. [CrossRef]
28. Ahammed, T.; Khan, I. Ensuring power quality and demand-side management through IoT-based smart meters in a developing country. *Energy* **2022**, *250*, 123747. [CrossRef]
29. Egypt: Alexandria–Abou Qir Metro Line. Available online: https://www.aiib.org/en/projects/details/2022/approved/Egypt-Alexandria-Abou-Qir-Metro-Line.html (accessed on 31 December 2025).
30. Mwasalat Misr Company. Available online: https://mwasalatmisr.com/ (accessed on 31 December 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.