LLM-Enhanced Scenario-Aware Vehicle-to-Home Operation under Grid Resiliency

Guang Chen
School of Electrical & Automation Engineering Nanjing Normal University
Nanjing, China
241812150@nnu.edu.cn

Difei Tang, *IEEE Member*
School of Electrical & Automation Engineering Nanjing Normal University
Nanjing, China
dftang@njnu.edu.cn

Ziqi Zhang, *IEEE Member*
School of Electrical & Automation Engineering Nanjing Normal University
Nanjing, China
zhangziqi@nnu.edu.cn

Qi Wang, *IEEE Senior Member*
School of Electrical & Automation Engineering Nanjing Normal University
Nanjing, China
wangqi@njnu.edu.cn

Yuanyi Chen, *IEEE Member*
School of Electrical & Automation Engineering Nanjing Normal University
Nanjing, China
chenyuanyi@njnu.edu.cn

Peng Wang, *IEEE Fellow*
School of Electrical & Automation Engineering Nanjing Normal University
Nanjing, China
epwang@njnu.edu.cn

***Abstract*—Long-duration power outages, driven by extreme weather, threaten residential energy resilience. The effectiveness of Vehicle-to-Home (V2H) operation heavily relies on the accurate characterization of load profiles. Existing methods are dependent on fixed net-load estimations, fail to adapt to dynamic crisis scenarios. This paper proposes a Large Language Model (LLM)-based scenario-aware V2H operation for household resilience enhancement. At the offline planning layer, the LLM generates scenario-specific net-load profiles by interpreting unstructured information such as weather and household composition, providing realistic inputs for Mixed-Integer Quadratic Programming (MIQCP) optimization. At the online control layer, the scenario-aware information generated by the LLM is integrated into the real-time dispatch via a weighted fusion and hybrid trigger mechanism. Simulations demonstrate that the proposed framework provides more informed energy management by effectively capturing the intensity of outage scenarios and diverse human behaviors. Even in cases where high-demand scenarios necessitate earlier resource depletion, the approach bridges the gap between qualitative scenario descriptions and quantitative numerical optimization, offering a more realistic and robust assessment of household energy resilience compared to traditional methods relying on optimistic historical averages.**

***Index Terms*--Vehicle-to-Home, Residential Energy Resilience, Large Language Model, Scenario Awareness.**

I. Introduction

Long-duration power grid outages triggered by extreme weather events have become increasingly frequent and prolonged, severely threatening residential energy resilience. Vehicle-to-Home (V2H) supports critical household loads by utilizing the battery and engine-generator (EG) of plug-in hybrid electric vehicles (PHEVs) and has demonstrated potential to enhance household energy resilience and yield operational and economic benefits across multiple scenarios [1].

Existing V2H operations have limitations. Offline optimization can theoretically provide near-optimal strategies, but its results are highly dependent on the input load curves [2]. If the input is a smooth curve based on historical averages, it often loses its reference value in extreme or behavior-driven outage scenarios [1]. Online heuristic control is simple for implementation and low computational burden. It typically uses static net load estimations to determine the optimal operating point of the engine, which may lead to resource misalignment in crisis scenarios where behaviors change significantly [3].

On the other hand, human behavior during crisis scenarios, such as power outages, exhibits significant context dependency. Residents' electricity consumption decisions are influenced by panic reactions, prioritization of power supply for critical equipment, and even judgments based on ambient temperature and information. Such qualitative behaviors and their impacts on electricity usage patterns have been discussed in behavioral science and emergency management literature [4], suggesting that relying solely on historical statistics cannot fully capture the load evolution characteristics during a crisis period.

Large Language Models (LLMs) have advantages in natural language understanding and common-sense reasoning. Recent work indicates that language models can be utilized for load forecasting, scenario generation, and are capable of mapping unstructured scenario descriptions [5], such as "persistent high temperatures," into quantitative judgments on load trends, thereby providing scenario-aware inputs or constraints for traditional optimization and control. Employing LLMs as a scenario reasoning layer integrated with mathematical optimization and heuristic control has emerged as a promising approach to bridge the gap where quantitative models struggle to incorporate qualitative scenario information.

This paper proposes a two-layer LLM-enhanced V2H control framework. In the offline planning layer, an LLM is utilized to generate scenario-aware net load profiles to enhance the input quality for the MIQCP-based optimization.

This work is supported by the National Natural Science Foundation of China under Grant 52277105. *Corresponding author: Difei Tang (dftang@njnu.edu.cn).*

979-8-3315-8138-1

Subsequently, in the online control layer, a hybrid trigger and dynamic weighted fusion mechanism is designed to integrate these pre-generated semantic insights into the real-time heuristic process, enabling key operation parameters to be adaptively updated based on the identified scenario-persona context.

The main contributions of this paper are: 1) A hierarchical LLM-enhanced V2H framework that integrates semantic scenario intelligence into both offline strategic planning and online tactical control. 2) A scenario-aware profile generation method that transforms unstructured text into quantitative load inputs, addressing the limitations of traditional optimization relying on static averages. 3) A robust online integration strategy featuring a hybrid triggering and weighted fusion mechanism, which adaptively incorporates pre-generated LLM insights into real-time heuristic dispatch to enhance energy resilience.

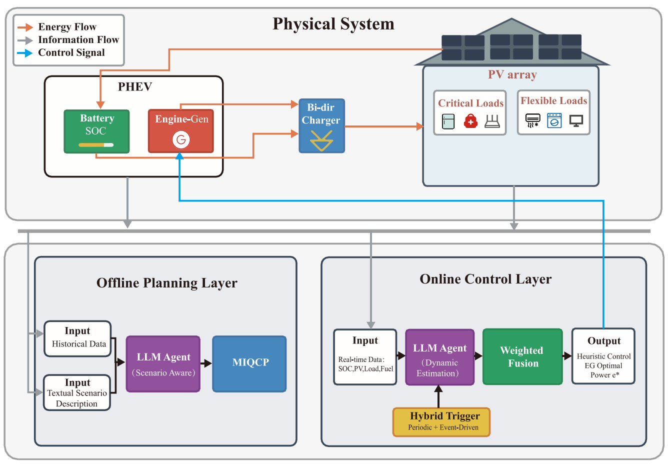

Figure 1. Proposed V2H System.

II. LLM-Enhanced V2H Systems

A. Physical Configuration and Energy Flow

As illustrated in Fig. 1, the V2H system functions as a cyber-physical platform to ensure household energy resilience [1]. The physical layer integrates a PHEV, a rooftop PV array, and household loads. Within the PHEV, the traction battery serves as the primary energy buffer for balancing power fluctuations, while the EG acts as a fuel-consuming secondary power source. Notably, the EG's operation is not merely a reactive response to battery depletion; instead, it is strategically triggered by the control logic to maintain energy security. The PV array provides clean energy that can either directly supply the household loads or charge the battery when surplus power is available. To maximize the survival duration during prolonged outages, household loads are prioritized and managed based on their criticality.

B. Control Architecture and Information Interaction

On the information layer, sensors continuously collect real-time data including the battery SoC, EG operational status, PV output, and household load consumption. These real-time measurements are transmitted to a fusion and decision module, where they are integrated with the pre-generated scenario-aware insights from the LLM agent. By combining the LLM's quantitative net load estimation (l') and qualitative reasoning with the historical average (l_{hist}), the module dynamically generates optimal operating parameters for the EG, ensuring the control strategy is both data-driven and scenario-aware.

III. Methodology

A. Offline Planning Layer

As illustrated in Fig. 1, the core value of the LLM-enhanced MIQCP model is to provide a theoretically optimal dispatch under the assumption of perfect foresight. The effectiveness of the original model is heavily contingent upon the accuracy of the input load profile $l(t)$. Traditional models typically rely on a static average load derived from historical data, which fails to capture the behavioral shifts and environmental volatility inherent in outage scenarios.

To address this, the offline layer leverages the LLM's contextual awareness to transform qualitative scenario descriptions into high-fidelity quantitative load inputs. This process refines the critical parameters within the physical constraints, ensuring that the MIQCP model operates on a scenario-aware baseline that closely aligns with the dynamic realities of a crisis.

To translate qualitative scenario descriptions into high-fidelity quantitative inputs, the offline layer implements a multi-factor driven synthesis protocol. This protocol bridges the gap between LLM semantic reasoning and the physical operational boundaries of power systems through a four-stage process. First, a feature-oriented schema encodes environmental constraints (E_{env}), household personas (P_{user}), and outage specifics (S_{out}) into a structured prompt. Acting as a domain expert, the LLM maps these features into an initial 24-point net-load curve l_{llm} while providing an internal confidence score. Second, to rigorously determine the influence of the LLM's insights, an adaptive weight $w \in [0,1]$ is formulated by synthesizing three complementary factors: the confidence factor C, the scenario severity S, and the expected deviation intensity D. The deviation intensity D is defined as the normalized absolute mismatch between the LLM's estimation and the historical baseline $\tilde{l}_{hist}$

$$w = \lambda_1 \cdot C + \lambda_2 \cdot D + \lambda_3 \cdot S$$

$$D = \min\left(\frac{\left|l_{llm} - \tilde{l}_{hist}\right|}{\tilde{l}_{hist}}, 1\right)$$

Third, the initial profile l_{llm} is fused with an expert-rule-based extreme profile l_{ext} using the derived weight w to ensure the result remains within established physical boundaries:

$$l' = w \cdot l_{llm} + (1 - w) \cdot l_{ext}$$

Finally, a robustness guardrail is applied to filter potential hallucinations by clipping l' within a $\pm 30\%$ range of the historical average, yielding the final scenario-aware benchmark l_{final}:

$$l_{final} = \mathrm{clip}\left(l', \left[0.7 \cdot \tilde{l}_{hist}, 1.3 \cdot \tilde{l}_{hist}\right]\right)$$

The resulting l_{final} is resampled to the optimization horizon T and stored in a shared knowledge base to serve as the consistent reference for both MIQCP planning and online heuristic execution.

979-8-3315-8138-1

Table I. Household Persona Definitions

Category	Type	Key Characteristics & Contextual Information
Environmental Scenarios	Winter Storm	Extreme snow, low PV; morning/night heating peaks; peak storm at noon
	Summer Heatwave	Extreme heat, high PV; afternoon AC cooling peak; load subsides at night
User Profile	Elderly care	High morning care and hot water demand; low afternoon rest load
	Night owl	Late-night activity; late wake-up with delayed morning peaks
	Shift worker	Low daytime load for sleeping; concentrated nighttime energy use
	Work from home	Continuous daytime office load; brief afternoon dip due to outings
	Young child	Early laundry and cooking peaks; early sleep leads to early load drop

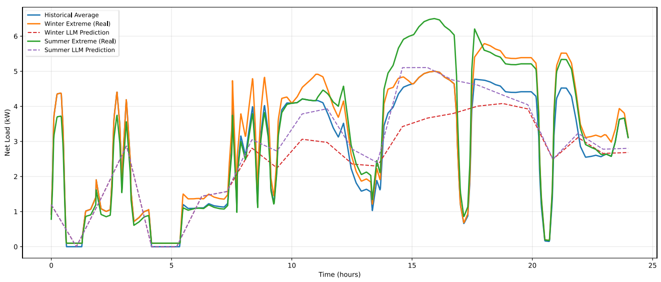

Figure 2. Scenario-Aware Net Load Comparison of night owl.

B. *Online Control Layer*

The online control layer prioritizes real-time responsiveness by leveraging traditional heuristic algorithms. While these algorithms are computationally efficient, their performance is often constrained by the reliance on static historical average loads ($\tilde{l}_{hist}$) for calculating the optimal EG output e^*. To address this, the proposed framework enhances the heuristic logic through a three-stage adaptive process: 1) Contextual Profile Retrieval, 2) Hybrid Triggering, and 3) Adaptive Weighted Fusion. This process replaces the static input with a context-aware profile $\tilde{l}_{final}$, which is sampled from the scenario-aware baseline pre-generated in the offline layer. This ensures that the EG output e^* can dynamically adapt to the volatile net load fluctuations characteristic of outage scenarios without incurring the computational latency of real-time LLM inference.

To maintain an optimal balance between control stability and situational sensitivity, a hybrid triggering mechanism encompassing both periodic and event-driven components is implemented. The periodic component operates at a fixed interval, such as every 60 minutes, ensuring that the online controller consistently aligns the dispatch strategy with the macro-scale trends predicted by the offline LLM agent. Complementing this, the event-driven component enables an immediate re-evaluation of the fusion strategy when the system detects significant environmental volatility. Specifically, a trigger is activated when sudden fluctuations in net power exceed a predefined threshold ϵ, or when the real-time measured net load l_{meas} deviates from the historical average $\tilde{l}_{hist}$ by more than a threshold σ. Such a mismatch indicates that the static baseline is no longer representative of the current operational state. This integrated approach ensures that the scenario-aware intelligence is prioritized only during critical intervals, thereby effectively minimizing unnecessary computational overhead while enhancing overall energy resilience.

Building upon the hybrid triggering framework, the online layer translates the retrieved scenario-aware profile $l_{final}(t)$ into actionable control commands by dynamically updating the optimal generator power setpoint e^*. Unlike conventional heuristic methods that reactively compensate for instantaneous power deficits, the proposed strategy integrates contextual expectations into the efficiency-oriented dispatch logic. First, the fuel-efficiency optimal operating point g_{opt} is determined based on the generator's quadratic fuel consumption characteristic:

$$g_{opt} = -q_2/(2q_1)$$

where q_1 and q_2 are the specific fuel consumption coefficients. To balance operational economy with supply reliability, the target setpoint e^* is derived through a multi-constrained maximization operator:

$$e^* = \max(\gamma \cdot g_{opt}, \tilde{l}_{final}(t), P_{min})$$

In this formulation, $\gamma = 0.88$ is an efficiency-prioritized coefficient designed to keep the generator within its high-performance window, while P_{min} serves as the lower stability limit to prevent inefficient low-load operation. Crucially, by explicitly incorporating $\tilde{l}_{final}(t)$ into the decision-making process, the generator output is proactively scaled to cover the anticipated demand peaks identified by the LLM, effectively preempting potential supply shortfalls. Finally, the output is capped by the physical maximum capacity P_{max} to ensure operational safety:

$$e^*_{final} = \min(e^*, P_{max})$$

This refined logic ensures that the generator not only operates near its peak efficiency but also remains highly responsive to the specific behavioral shifts of the current outage scenario.

IV. Case Study

A. *Experimental Configuration and Scenario Description*

The simulation data for this study are sourced from [1], including the load characteristics and PV generation patterns of a typical household.

To validate the effectiveness of the proposed framework, a 24-hour simulation is conducted under two extreme disaster scenarios: A Winter Storm and a Summer Heatwave. The specific environmental parameters and user profiles are

Table II. Statistical Performance Matrix. Data entries are presented in the format of (Backup Duration, Unmet Load). For LLM-enhanced methods, values are reported as mean $\pm$ standard deviation ($\mu \pm \sigma$). Units for backup duration and unmet load are hours (h) and kWh, respectively.

User Profile	Season	MIQCP-Static	MIQCP-LLM (avg±std)	Heuristic-Static	Heuristic-LLM (avg±std)
Elderly care	Winter	16.17, 59.29	15.86±3.35, 63.58±0.44	8.03, 44.89	11.42±0.54, 48.41±0.60
Elderly care	Summer	16.17, 59.33	21.46±1.10, 64.37±0.25	8.03, 48.39	12.41±1.22, 48.21±0.48
Night owl	Winter	16.17, 60.07	19.51±0.52, 65.19±0.51	8.03, 46.81	11.60±0.09, 49.77±0.03
Night owl	Summer	16.17, 60.08	19.25±0.44, 64.82±0.46	8.03, 50.12	12.96±0.29, 48.61±0.10
Shift worker	Winter	16.17, 60.78	16.26±3.91, 65.36±0.12	8.03, 47.57	10.70±0.27, 50.84±0.18
Shift worker	Summer	16.17, 60.98	14.62±2.96, 65.52±0.62	8.03, 50.90	12.32±1.44, 50.42±1.77
Work from home	Winter	16.17, 61.25	18.45±2.00, 65.98±0.49	8.03, 46.79	11.48±0.40, 50.76±0.41
Work from home	Summer	16.17, 61.23	16.49±4.08, 66.19±0.45	8.03, 50.14	11.92±1.58, 51.00±1.36
Young child	Winter	16.17, 60.37	19.53±0.50, 65.47±0.25	8.03, 46.40	11.75±0.44, 49.56±0.45
Young child	Summer	16.17, 60.65	15.51±3.72, 65.06±0.60	8.03, 49.71	12.00±1.59, 50.33±1.29

formalized in Table I, which serves as the primary knowledge source for guiding the generative process.

The Winter Storm is characterized by extreme cold and negligible solar power, whereas the Summer Heatwave is driven by high cooling demand and strong solar generation. To ensure a diverse evaluation of system resilience, five household profiles are modeled: Elderly Care, Young Child, Night Owl, Work from Home, and Shift Worker.

The following four operation strategies are compared to isolate the performance gains:

1. *MIQCP-Static:* Offline optimization utilizing historical average load data.
2. *MIQCP-LLM:* Offline optimization utilizing the profiles synthesized by the LLM.
3. *Heuristic-Static:* Real time rule based control employing fixed operation thresholds.
4. *Heuristic-LLM:* Real time control where the generator output is updated via the pre-generated LLM profiles.

To account for variability in load estimation caused by different models, these scenario aware profiles are synthesized using three advanced LLM backends known as GLM 4.7, DeepSeek V3.2, and MiniMax M2.1. The statistical performance metrics, consisting of backup duration and unmet load, are reported in Table II as mean plus or minus standard deviation ($\mu \pm \sigma$). This setup establishes a rigorous foundation for evaluating system resilience as visualized through the SOC and power output trajectories in Fig. 3 and Fig. 4.

B. Analysis of Scenario Aware Load Profile Generation

This section examines the performance of the LLM backends in translating complex qualitative descriptions into quantitative demand profiles. As illustrated in Fig. 2, which features the Night Owl persona during a Winter Storm synthesized by DeepSeek V3.2, the generated l_{final} profile exhibits significant contextual intelligence compared to the static historical average $\tilde{l}_{hist}$. It is observed that while the LLM derived profile successfully captures the nocturnal consumption surge characteristic of the Night Owl, it does not perfectly overlap with the rule based extreme load profile. This divergence is intentional and stems from the integration protocol described in Section III A.

The rule based extreme load serves as a synthetic upper bound for stress testing, whereas the LLM aims to generate a plausible and interpretable scenario. Furthermore, the application of fusion and clipping mechanisms ensures that the resulting l_{final} remains within a realistic operating envelope, leading to a smoother and more conservative estimation than the raw extreme rules. However, the primary value of the LLM integration lies in its ability to provide situational awareness regarding the timing and trend of demand shifts. By anticipating the late night load peak identified in the case study, the system gains critical foresight that allows for proactive strategy adjustments. This behavioral alignment, rather than simple numerical curve fitting, is what effectively reduces the mismatch between energy supply and demand during the most vulnerable hours of the outage.

C. Statistical Performance and Robustness Assessment

The aggregate performance of the proposed framework is quantitatively evaluated through the statistical metrics summarized in Table II, revealing distinct impacts of Large Language Model integration on system resilience. Across the evaluated scenarios, the LLM enhanced strategies demonstrate a primary strength in extending the system backup duration. Specifically, the Heuristic LLM approach achieves an average increase of 3.83 hours in backup duration compared to the Heuristic Static baseline, with individual gains spanning a wide range from 2.67 to 4.93 hours. In comparison, the MIQCP LLM strategy yields a more variable average improvement of 1.52 hours, where the performance fluctuates between a slight decrease of 1.55 hours and a substantial gain of 5.29 hours. This overall enhancement is largely attributed to the information gain provided by the scenario aware profiles which allow the controller to proactively align energy supply with context specific demands such as nocturnal peaks.

However, the impact on the unmet load metric remains inconsistent across different operation strategies. Within the MIQCP LLM framework, the system exhibits an average increase in unmet load of 4.75 kilowatt hours compared to its

979-8-3315-8138-1

static counterpart, with values distributed between 4.29 and 5.12 kilowatt hours. This performance degradation indicates a magnification effect inherent to offline global optimization. Although the synthesis protocol ensures physical plausibility, any residual prediction errors in the 24 hour load profile can be amplified during the optimization process, leading to sub-optimal resource allocation during peak intervals. Conversely, the Heuristic LLM strategy demonstrates superior robustness and a more stable safety margin. The unmet load for this approach shows an average change of 1.62 kilowatt hours within a span from a reduction of 1.51 to an increase of 3.97 kilowatt hours. Such stability is further evidenced by the lower standard deviation across the different model backends known as GLM 4.7, DeepSeek V3.2, and MiniMax M2.1. These results suggest that while LLM insights significantly bolster long term energy availability, the sensitivity of offline optimization to prediction accuracy remains a critical trade off.

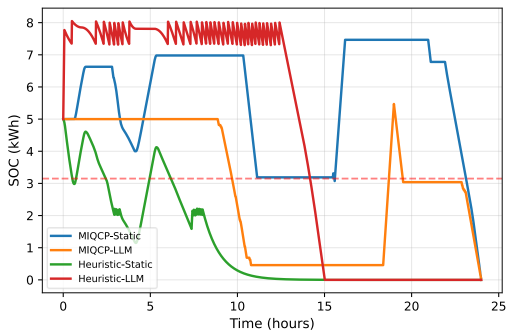

Figure 3. SOC Trajectory.

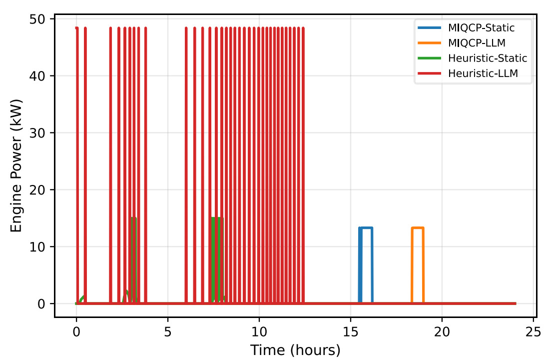

Figure 4. Generator Output.

D. Dynamic Operational Behavior and Strategy Insights

The operational dynamics of the V2H system are further analyzed through the state of charge trajectories and generator power outputs depicted in Fig. 3 and Fig. 4, respectively. This specific case focuses on the Night Owl persona during a Winter Storm as managed by the DeepSeek V3 backend. In Fig. 3, the state of charge profiles reveals a distinct advantage for the Heuristic LLM strategy over its static counterpart. Unlike the Heuristic Static baseline which reacts only when the battery is near exhaustion, the LLM enhanced controller anticipates the nocturnal load surge characteristic of the Night Owl profile. By retrieving the pre-generated l_{final} profile from the shared knowledge base, the system maintains a more robust energy buffer during the early morning hours. This proactive energy management prevents the rapid depletion of the state of charge observed in static methods, where the battery level often drops below critical thresholds during unpredicted demand peaks.

The corresponding generator behavior in Fig. 4 further elucidates the efficacy of the adaptive e^* logic discussed in Section III B. For the Heuristic-LLM strategy, the generator setpoint is dynamically adjusted to align with the expected demand peaks identified in the scenario aware profile. Specifically, the generator ramps up its output in anticipation of the night peak, ensuring that the power supply covers the estimated l_{final} while adhering to the efficiency prioritized operating window. This behavior demonstrates the practical utility of the information gain where the system effectively trades a small amount of fuel for early energy maintenance. In contrast, the MIQCP-LLM strategy, while theoretically optimal, shows a higher sensitivity to the load structure. In some intervals, the magnification of prediction errors in the global optimization leads to delayed generator firing, which explains the previously noted fluctuations in unmet load. Ultimately, the dynamic adaptation observed in the Heuristic-LLM curves confirms that the integration of LLM insights transforms the V2H system from a reactive backup into a proactive resilience asset.

V. CONCLUSIONS

This paper presents an LLM enhanced V2H framework that bolsters grid resilience through situational awareness. The integration of Large Language Models effectively transforms the heuristic dispatch logic from a reactive discharge mode into a proactive control asset that maintains a robust state of charge via strategic generator scheduling. While the impact on the MIQCP strategy remains primarily limited to adjustments in operational timing, the benefits of the proactive approach are particularly pronounced during winter outages with low solar availability. In summer scenarios, the LLM enhanced version ensures superior energy management during the critical hours after photovoltaic generation subsides. This research establishes a viable pathway for utilizing situational intelligence to safeguard residential energy stability under extreme uncertainty.

REFERENCES

[1] H. Shin and R. Baldick, "Plug-in electric vehicle to home (V2H) operation under a grid outage," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2254-2263, Sept. 2017.

[2] B. Zafar, G. A. S. Sidhu, M. S. J. M. S. Khan, R. M. Asif, and A. Ali, "PV-EV integrated home energy management using vehicle-to-home (V2H) technology and household occupant behaviors," *Applied Energy*, vol. 307, p. 118228, Feb. 2022.

[3] N. V. Pham, T. T. T. B. Tran, and T. T. H. Nguyen, "An MIQP model for microgrid unit commitment incorporating grid-connected/islanded operating modes and power network," *Mathematical and Computational Applications*, vol. 27, no. 1, p. 13, Feb. 2022.

[4] A. Rubin and R. Rogers, "Behavioural and psychological responses of the public during a major power outage: A literature review," *International Journal of Disaster Risk Reduction*, vol. 42, p. 101341, Jan. 2020.

[5] G. Spiliopoulos, P. Spanakis, and V. K. G. Spiliopoulos, "Utilizing language models for energy load forecasting," in *Proc. 14th ACM International Conference on Future Energy Systems (e-Energy '23)*, June 2023, pp. 586-590.

[6] Pecan Street Inc. (2023). *Dataport - Residential electricity and PV/EV dataset*. [Online]. Available: https://www.pecanstreet.org/dataport/

979-8-3315-8138-1