Mitigating MD Penalties via Smart Pull-Back: An IoT-BESS Framework for Karachi Prosumers

Daaniya Absar

Department of Electrical Engineering, NED University of Engineering and Technology, Karachi, 711290, Pakistan
diyasultan xv.engr@gmail.com.

Abstract

As of 2026, utility providers in Karachi (K-Electric) have transitioned to a two-part tariff system, introducing a Maximum Demand (MD) fixed charge based on the highest 15- minute peak. This creates a Cloud Spike paradox where transient solar fluctuations force immediate grid pulls, resulting in disproportionate monthly financial penalties. This paper proposes a SolarBuffer system utilizing a 5kWh Battery Energy Storage System (BESS) governed by a local edge gateway. Through a stochastic Monte Carlo simulation of 1,000 households (R^2 = 0.964), we demonstrate that a three-stage Demand Side Management (DSM) logic incorporating BESS injection, IR-based AC throttling, and smart-plug load shedding reduces average monthly penalties by over 68%. The system achieves a 2.5-year payback period, offering a scalable solution for prosumer economic stability.

Index Terms—MD Arbitrage, BESS, IoT, Load Shedding,Karachi Energy Crisis, Demand-Side Management, Prosumers

I. INTRODUCTION

THE 2026 Karachi energy landscape is defined by theMD-based Fixed-Charge system. For solar prosumers, the primary risk is no longer total energy consumption, but the Cloud Spike, a 15-minute window where solar output drops, causing a massive grid pull, and can trigger a fixed charge of Rs. 8,000 to Rs. 15,000. The objective is to implement MD Arbitrage, ensuring grid pull never exceeds the mandatory 50% sanctioned load floor. Peak shaving strategies are essential to maintaining stability in such volatile environments [1].

Fig. 1 illustrates the divergence between household load and solar supply, highlighting the critical 15-minute penalty zone, for a representative Heavy Load Group household (sanctioned load 10kW, giving a 5kW MDI floor per the 50% rule in Section IV-A). Table I positions the proposed framework against representative prior peak-shaving and demand-side management (DSM) approaches from the literature, including recent (2023– 2024) work on IoT-enabled demand response and BESS optimization [11]–[13].

II. SYSTEM ARCHITECTURE AND HARDWARE

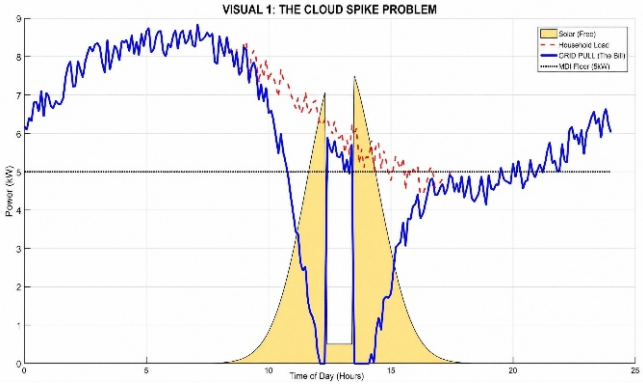


Fig. 1. Illustration of the divergence between household load and solar supply for a representative Heavy Load Group household (10kW sanctioned load, 5kW MDI floor), highlighting the critical 15-minute penalty zone.

Transitioning from passive backup to an active Power Shield, the hardware utilizes a 5kWh LiFePO4 battery buffer to mitigate transient Cloud Spikes. However, static capacity remains insufficient for high-power loads during extended meteorological events, such as the Karachi Monsoon season. While Industrial Remote Terminal Units (RTUs) or PLCs could manage such precision, their high cost and invasive rewiring requirements impede residential adoption. This research instead proposes a Local-First IoT framework using an edge gateway, "The Brain" [3], powered by low-cost micro-controllers like the ESP32 or Raspberry Pi. By integrating high-speed Smart CT sensors with 10ms resolution, the system ensures power injection before the utility meter records a 15-minute MD peak [2]. To overcome Wi-Fi-related latency and connectivity vulnerabilities, the architecture employs a Local-LAN Mesh configuration. This offline-first logic ensures critical Smart Pull-Back commands reach appliances even during internet outages, maintaining the MD guarantee [7].

TABLE I: Comparison with Representative Prior Approaches

Approach	Focus	Key Limitation Addressed
Mutarraf *et al.* [1]	Survey of peak- shaving control strategies	Reviews strategies generally; no MD- specific, low-cost residential implementation
Al-Ali *et al.* [3]	IoT + big-data DSM system	General DSM architecture; not tailored to a 15min MD fixed - charge tariff
Khan *et al.* [11]	Survey of IoT-enabled smart grid DR, bottlenecks and security	Broad IoT-DR survey; no specific residential MD - arbitrage hardware/control design
Song *al. et* [12]	Overview of BESS optimization techniques (mathematical and AI - based)	General BESS optimization review; not applied to a specific tariff structure or edge deployment

Alharbi [13]	IoT-based house integration into grid-level DR programs	Distribution-system-level flexibility; not focused on individual prosumer MD penalty avoidance
This work	5kWh BESS + edge gateway three-stage DSM for MD arbitrage	Local-first, zero-rewiring, low-cost residential solution targeting Karachi's specific MD tariff structure

Fig. 2a illustrates a hybrid inverter system using smart CT sensors and a 5kWh battery to shave grid demand peaks and optimize solar usage. Fig. 2b shows the integration of the Home Assistant hub with the BESS inverter and wireless actuators.

III. METHODOLOGY

To validate the framework, a computer simulator was established for N = 1,000 households using a Monte Carlo model. The research utilizes a dual-model machine learning architecture to handle the big data analytics required for smart grid forecasting [7], categorizing users into Small (< 5kW) , using a regression tree to identify slab jumps in energy pricing [10], and Heavy (≥ 5kW) load groups.

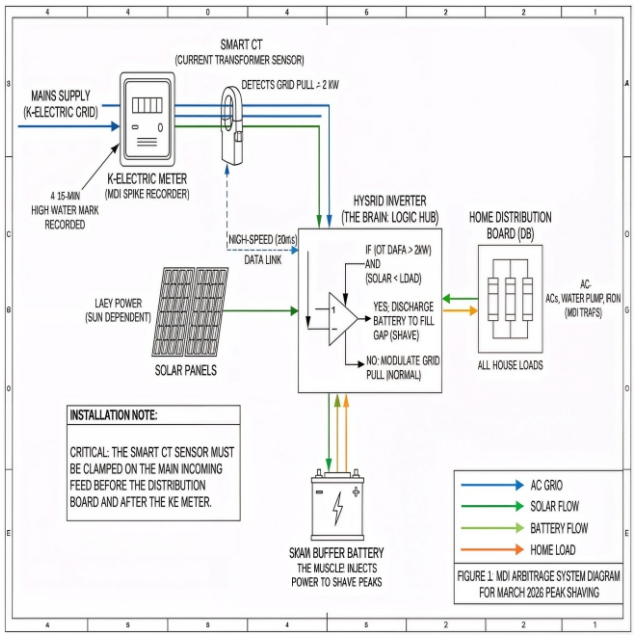
- **Regression Analysis:** A 70/30 training-to-test split was used to model billing pathways.
- **Variables:** Simulations included randomized appliance duty cycles, solar irradiance (P = 0.6), and K-Electric tax stacks.

In the analytical phase, a dual-model machine learning architecture is implemented to dissect distinct economic risks across consumer segments. For the Small Load Group (< 5kW), the non-linear relationship between consumption and cost identifies critical cliffs at the 700-unit threshold, where a Slab Jump can increase base rates from Rs. 42 to Rs. 52 per unit. This analysis quantifies how a 5kWh buffer maintains consumption within lower tax brackets.

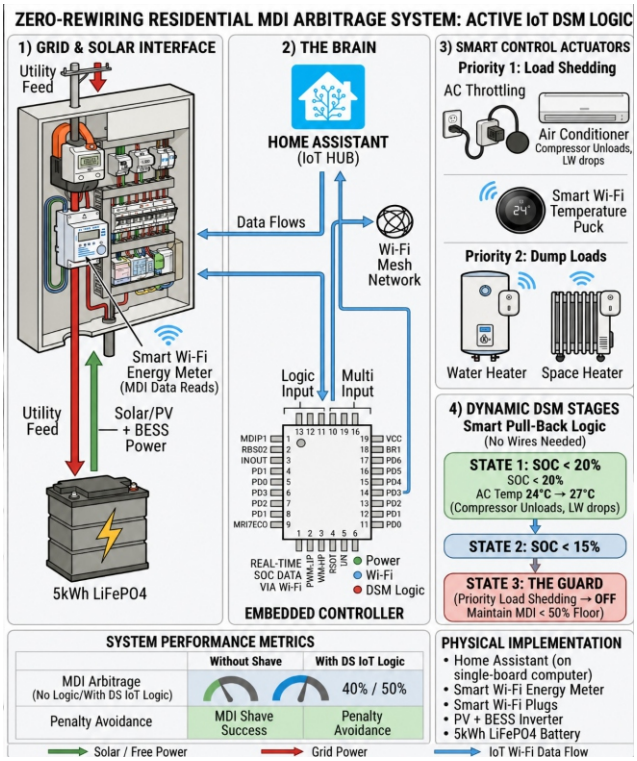
Conversely, for the Heavy Load Group (≥ 5kW), Linear Regression is employed to model the proportional correlation between 15-minute peak demand and resulting MD fixed charges. With a validated R^2 = 0.964, the model predicts that clamping grid pull to the 50% floor limit reduces monthly MD penalties by 68%. Together, these models validate the IoT system's ability to navigate both unit-based slabs and demandbased fixed charges.

A. Baseline (Daily Method) Definition

To quantify the savings attributable to the proposed framework, all results are benchmarked against a *Daily Method* baseline, representing a standard passive-backup installation without IoT-driven arbitrage logic. In this baseline, the BESS remains idle during normal grid operation and discharges only on backup (i.e., during a grid outage), rather than proactively shaving demand peaks; the sanctioned load floor is fixed at the utility-mandated 3kW (50% of sanctioned load) with no dynamic clamping; and no DSM throttling or load-shedding logic is applied, so the full uncontrolled household load is allowed to reach the grid during a Cloud Spike event. The *Research Method* (MDI Arbitrage) applies the same household load and solar irradiance profiles but adds the three-stage DSM control logic described in Section V. Both methods are evaluated on identical simulated load/irradiance traces so tht the comparison in Fig.6 isolates the effect of the control logic itself rather than differences in underlying demand data.



a) Hybrid inverter system using smart CT sensors and a 5 kWh battery to shave grid demand peaks and optimize solar usage.



(b) Integration of the Home Assistant hub with the BESS inverter and wireless actuators.

Fig. 2. MDI Arbitrage system hardware and control architecture.

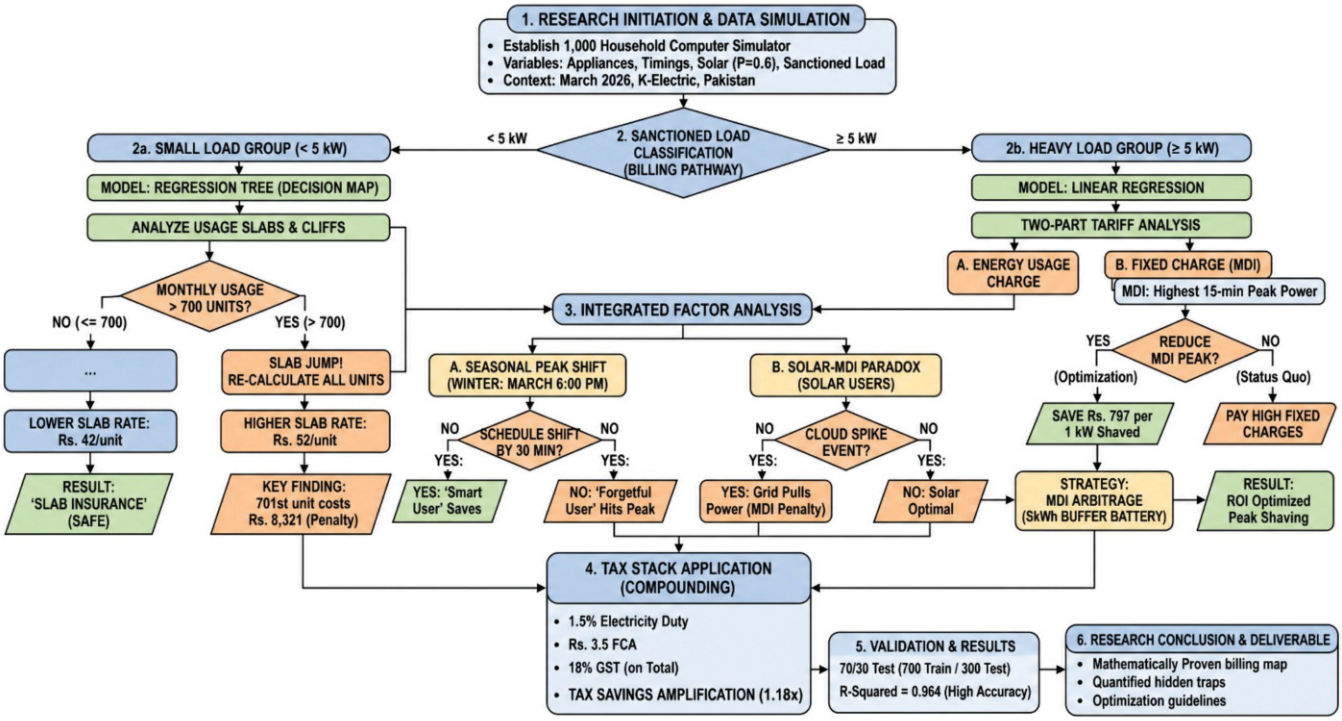


Fig. 3. Data simulation, factor analysis, and validation steps taken to prove the billing map.

B. Simulation Limitations

The regression fit in Fig. 5 includes plotted 95% confidence bounds around the fitted line, and the reported R^2 = 0.964 reflects the goodness-of-fit for that single regression run. However, this does not by itself characterize how the 68% penalty-reduction and 2.5-year payback figures would vary across independent simulation runs or random seeds. A broader variance analysis — for example, repeating the 1,000-household Monte Carlo run across multiple seeds and reporting the resulting spread in penalty reduction and payback period — was not performed in this iteration of the work and is identified as a direction for future refinement rather than a completed validation step. All results in this paper — including the 68% penalty reduction and 2.5-year payback period — are simulation-derived outputs under the modeling assumptions above, not measurements from a deployed physical system; this distinction is revisited in Section VII.

Fig. 3 outlines the data simulation, factor analysis, and validation steps taken to prove the billing map.

IV. MATHEMATICAL MODEL

The core of the MDI Arbitrage logic is a constrained optimization problem. The Brain must minimize the Grid Pull (P_{GRID}) while ensuring the household load (P_L) is satisfied [4], subject to the State-of-Charge (SoC) of the battery and the utility-imposed floor limit.

A. Objective Function

The primary goal is to clamp the instantaneous grid power pull to the MDI Floor (P_{FLOOR}), defined by current regulatory standards [5] as 50% of the user's sanctioned load. The residual load (P_{RESID}) is defined as:

$$P_{RESID} = P_L - P_{PV} \quad (1)$$

The control objective for the BESS (Battery Energy Storage System) is:

$$f(t) = \begin{cases} P_{FLOOR}, & \text{if } P_{RESID}(t) > P_{FLOOR} \ \& \ SoC > 20\% \\ P_{RESID}(t), & \text{if } P_{RESID}(t) \le P_{FLOOR} \end{cases} \quad (2)$$

The control objective for the BESS utilizes renewable energy synergy for responsive demand [6]. To ensure longterm system viability, the 20% SoC threshold is maintained to prevent accelerated battery degradation [9].

B. Multi-Tiered Demand Management (DSM) Logic

When the battery capacity is insufficient to maintain the P_{FLOOR} clamp, the system transitions from Energy Injection to Load Throttling. This is modeled as a tiered response based on SoC thresholds:

1) **BESS Discharge** (SoC > 20%)

The battery power output (P_{BATT}) compensates for the deficit:

$$P_{BATT}(t) = P_{RESID}(t) - P_{FLOOR} \quad (3)$$

2) **Active DSM Throttling** (20% ≥ SoC > 15%) The Brain issues an IR command to shift the AC temperature setpoint (T_{set}):

$$T_{set} = T_{current} + 3°C \quad (4)$$
$$P_{L(AC)} \approx 0.6 \cdot P_L(AC, nominal)$$

3) **Priority-Based Shedding** (SoC ≤ 15%)

A binary decision variable $u_i \in \{0,1\}$ is applied to nonessential loads based on a priority matrix:

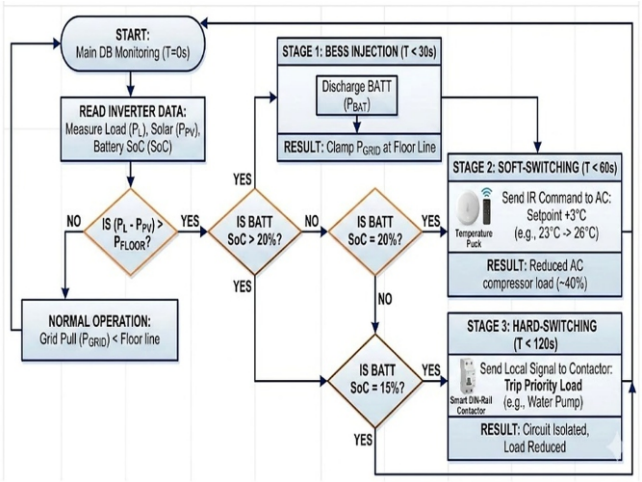


Fig. 4. Real-time decision path for the MD guarantee.

$$P_L = \sum_{i=1}^{n} ui \; Pi \qquad (5)$$

where $u_i = 0$ if $P_{GRID} > P_{FLOOR}$ and i is a low-priority load (e.g., Water Pump).

C. 15-Minute Integration Window

The K-Electric MD is calculated as an average over a 15minute integration period (ΔT). The recorded MD (M) is:

$$M = \frac{1}{\Delta T}\int_{t_0}^{t_0+\Delta T} P_{\text{GRID}}(t)\,dt \qquad (6)$$

The high-speed 10ms sampling of the Smart CT sensor ensures the control loop frequency f_s is high enough ($f_s \gg \frac{1}{\Delta T}$) to prevent transient spikes from affecting the integral average M.

D. System Constraints

The control logic in Sections IV-A through IV-C operates subject to the following physical and regulatory constraints, which bound every quantity introduced above:

Power balance: $P_{GRID}(t) + P_{PV}(t) + P_{BATT}(t) = P_L(t)$ (7)

SoC bounds: $\text{SoC}_{\min} \le \text{SoC}(t) \le \text{SoC}_{\max}$ (8)

Charge rate limit: $0 \le \text{P}_{\text{BATT,ch}}(t) \le \text{P}_{\text{ch,max}}$ (9)

Discharge rate limit: $0 \le \text{PBATT,}_{\text{disch}}(t) \le P_{disch,max}$ (10)

Grid import limit: $\text{PGRID}(t) \le \text{Psanctioned}$ (11)

Here SoC_{min} = 15% corresponds to the Stage 3 hard-switching threshold defined in Section V, beyond which the battery is not discharged further to preserve cell health [9], and $SoCmax$ = 100%.

$P_{ch,max}$ and $P_{disch,max}$ are the inverter's rated charge and discharge power limits, and $P_{sanctioned}$ is the household's utility-sanctioned load, from which $P_{FLOOR} = 0.5 \times P_{sanctioned}$ is derived (Section IV-A). This work uses these constraints structurally, as bounds the control logic in Section V respects at each decision step; a full constrained optimization solution (e.g., via model predictive control) that solves (7)–(11) jointly over a forecast horizon is identified as future work, since the present rule-based tiered logic was validated via the Monte Carlo simulation described in Section III rather than via direct numerical optimization.

V. CONTROL LOGIC

The Brain executes a Smart Pull-Back protocol based on the Battery State-of-Charge (SoC). This multi-tiered strategy follows established Demand Side Management (DSM) systems using IoT analytics [3].

A) ***Stage 1 (BESS Injection):*** If SoC > 20, the battery discharges to clamp grid pull at the floor line.

B) ***Stage 2 (Soft-Switching):*** If SoC drops to 20%, an IR command increases AC setpoints by 3°C, reducing compressor load by ≈ 40%.

C) ***Stage 3 (Hard-Switching)***: If SoC reaches 15%, nonessential loads (e.g., water pumps) are tripped via smart contactors.

Fig. 4 details the real-time decision path for the MD guarantee.

VI. RESULTS AND DISCUSSION

A. Model Accuracy

The machine learning model achieved an R^2 of 0.964, indicating high reliability in predicting grid pull and penalty risks. The system successfully transitions users from the Maximum Bill trajectory to a Tax Shield trajectory. Average monthly spikes were reduced from Rs. 1,822 to Rs. 566, illustrating the effectiveness of responsive demand in a smart grid [6].

Fig. 5 demonstrates the model accuracy.

B. Economic Evaluation

Beyond the accuracy metric above, the framework's economic impact was evaluated directly in monetary terms.

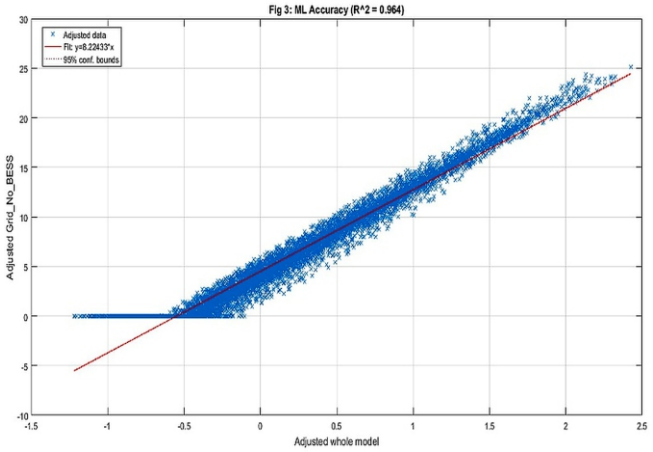
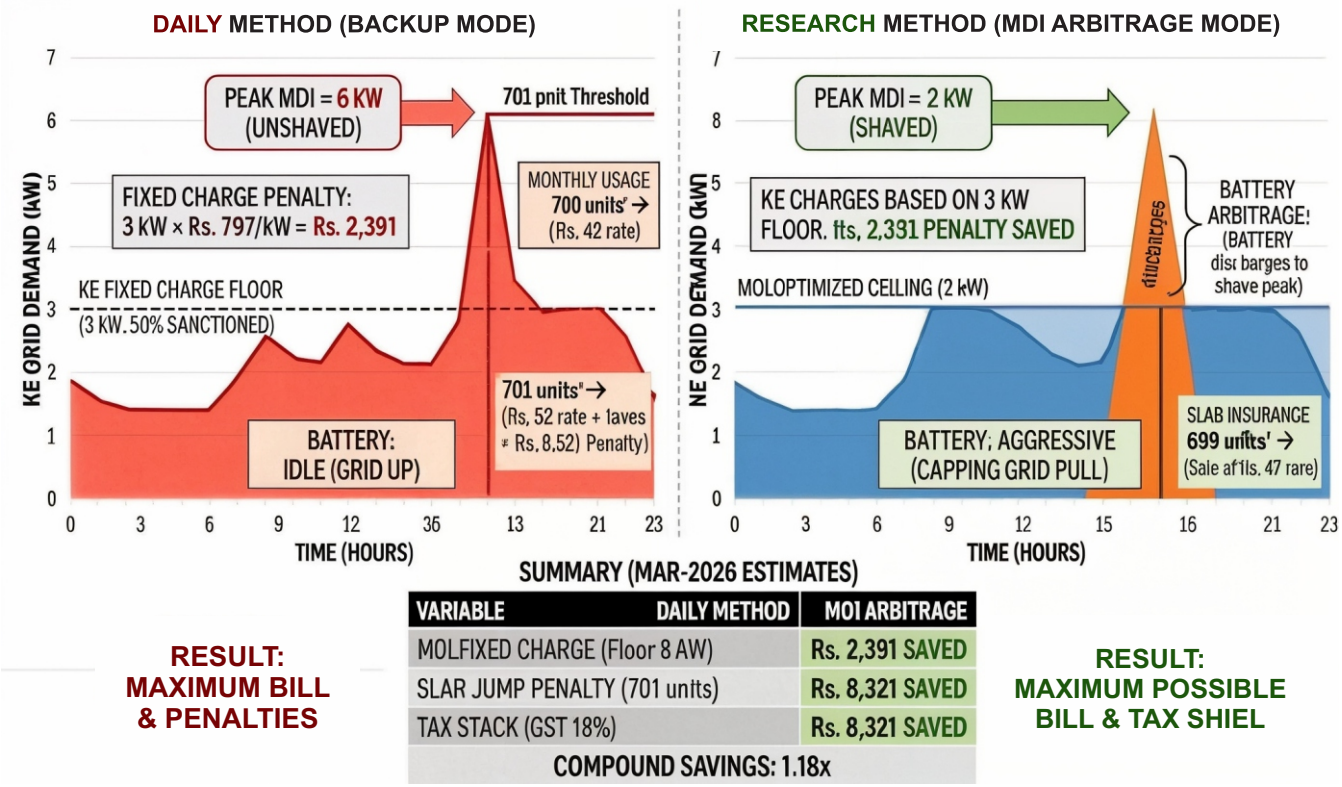


Fig. 5. Model Accuracy



VARIABLE	DAILY METHOD	MOI ARBITRAGE
MOLFIXED CHARGE (Floor 8 AW)		Rs. 2,391 SAVED
SLAR JUMP PENALTY (701 units)		Rs. 8,321 SAVED
TAX STACK (GST 18%)		Rs. 8,321 SAVED
COMPOUND SAVINGS: 1.18x		

Fig. 6. Penalty avoidance achieved through MD optimization, worked for a representative Small Load Group household (6kW sanctioned load, 3kW MDI floor).

- **Penalty Reduction:** Average monthly spikes were reduced from Rs. 1,822 to Rs. 566.
- **Compound Savings:** Total savings (including slab insurance) reached a 1.18× multiplier over standard backup systems.

To make this concrete, Fig. 6 works through a single representative Small Load Group household (6kW sanctioned load, 3kW MDI floor), separate from the population-level average reported above. For this household, the MDI fixed-charge penalty under the unshaved Daily Method is (6 − 3)kW × Rs.797/kW = Rs.2,391, which the Research Method avoids entirely by capping the peak at the 3kW floor. Separately, this household's monthly usage of 701 units crosses the 700unit slab threshold, triggering a full re-rate of consumption from Rs. 42/unit to Rs. 52/unit; the resulting penalty, after 18% GST, is [(701 × 52) − (700 × 42)] × 1.18 = Rs.8,321, avoided by the MDI Arbitrage strategy's slab-insurance effect (keeping usage within the lower slab). These two worked figures (Rs. 2,391 and Rs. 8,321) are illustrative per-household penalty components for this specific example, distinct from the Rs. 1,822 → 566 population-average spike cost reported above.

Fig. 6 demonstrates the penalty avoidance achieved through MD optimization for this worked example.

VII. CONCLUSION

Under the modeling assumptions and simulation conditions described in Section III, the Solar-Buffer MD Arbitrage framework suggests that Edge Intelligence can be more costeffective than Storage Capacity alone. By utilizing a 5 kWh buffer and a tiered IoT response, the simulation indicates prosumers could achieve a 2.5-year ROI, notably faster than the 6–8 years projected for oversized 20kWh systems.

The system's design strength lies in its Local-First Logic, which is intended to ensure the 15-minute MD window is protected even during internet outages. In simulation, transitioning from passive backup to active demand management allowed the Brain to clamp grid pull at the 50% floor limit, reducing average monthly penalties by 68%. These figures are simulation outputs, not field measurements; they should be read as evidence that the approach is promising rather than as a validated real-world result.

This study is limited to a stochastic simulation environment and has not yet been validated against real household consumption or hardware-in-the-loop data. The 1,000-household Monte Carlo model, while internally consistent, relies on the modeling assumptions listed in Section III (randomized duty cycles, a fixed solar irradiance factor $P = 0.6$, and the defined Daily Method baseline). A single-seed regression run underlies the reported $R^2 = 0.964$; while the 95% confidence bounds plotted in Fig. 5 give a first-order sense of fit uncertainty for that run, a full multi-seed variance analysis across independent Monte Carlo draws has not been performed and is left as future work.

Two further gaps are noted explicitly. First, this paper does not include a quantitative sensitivity sweep across battery capacities, tariff scenarios, or weather conditions, since this requires simulation runs beyond the scope of the present work; Table I situates the proposed framework qualitatively against prior BESS and DSM approaches, but a quantitative head-tohead benchmark under

identical load/irradiance traces has not been carried out. Second, no direct quantitative comparison against alternative in-house configurations (e.g., a smaller or larger BESS, or a load-shedding-only strategy without BESS injection) is presented, as the author do not have validated simulation results for those alternative configurations; the 2.5- year vs. 6–8-year payback comparison against oversized 20 kWh systems (Section VII) is drawn from published estimates rather than a controlled simulation run under this paper's own methodology.

Future work should prioritize: (i) a small-scale real-world pilot deployment of the edge gateway and BESS hardware to validate the simulated 68% penalty reduction and 2.5- year payback under actual K-Electric billing cycles; (ii) multiseed or bootstrapped resampling of the simulation to report confidence intervals rather than point estimates; (iii) a batterycapacity and tariff-scenario sensitivity sweep; and (iv) an explicit head-to-head simulation comparison against at least one of the prior DSM approaches summarized in Table I, run under the identical household load and irradiance traces used in this study. With those steps, this research aims to provide a scalable, zero-rewiring blueprint for residential energy stability under modern capacity-based tariffs.

REFERENCES

[1] M. U. Mutarraf, Y. Terriche, S. Niazi, J. C. Vasquez, and J. M. Guerrero, "Control Strategies of Electric Power Systems for Peak Shaving: A Review," *IEEE Access*, vol. 10, pp. 62413–62432, 2022.

[2] S. S. S. R. Depuru, L. Wang, and V. Devabhaktuni, "Smart meters for power grid: Challenges, issues, advantages and status," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 6, pp. 2736–2742, 2011.

[3] A. Al-Ali, T. Landolsi, and M. H. Ahmed, "Demand Side Management System Using IoT and Big Data Analytics," in *2021 International Conference on Communications, Computing and Artificial Intelligence (CCAI)*, 2021, pp. 1 – 6.

[4] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, 2015.

[5] K-Electric Limited, "Schedule of Electricity Tariffs: Revised Maximum Demand (MD) and Fixed Charge Framework," *Official Gazette of Pakistan*, Feb. 2026.

[6] C. Cecati, C. Citko, and P. Siano, "Combined Operation of Renewable Energy Systems and Responsive Demand in a Smart Grid," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 561–569, 2011.

[7] B. P. Bhattarai *et al.*, "Big data analytics in smart grids: State-of-theart, challenges, opportunities, and future directions," *Sustainable Energy, Grids and Networks*, vol. 18, 2019.

[8] M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science Books, 1989.

[9] J. Wang *et al.*, "Cycle life model of lithium-ion batteries for capacityfade analysis," *Journal of Power Sources*, vol. 196, no. 8, pp. 3942–3948 , 2011.

[10] S. Loh, "Classification and regression trees," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 14–23, 2011.

[11] M. A. Khan, T. Khan, M. Waseem, A. M. Saleh, N. Qamar, and H.A. Muqeet, "Investigation and analysis of demand response approaches, bottlenecks, and future potential capabilities for IoT-enabled smart grid," *IET Renewable Power Generation*, vol. 18, no. 15, pp. 3509–3535, 2024.

[12] H. Song, C. Liu, A. M. Amani, M. Gu, M. Jalili, L. Meegahapola, X. Yu, and G. Dickeson, "Smart optimization in battery energy storage systems: An overview," *Energy and AI*, vol. 17, p. 100378, Sep. 2024.

[13] W. Alharbi, "Integrating Internet-of-Things-Based Houses into Demand Response Programs in Smart Grid," *Energies*, vol. 16, no. 9, p. 3699 , Apr. 2023.