

Decentralized AI-Driven flexibility for transformer-constrained industrial energy communities

Merkebu Z. Degefa^{1*}, Bjørn Grønning² and John Arild Arnø³

¹ Department of Electrical Engineering and Computer Science, University of Stavanger, Stavanger, Norway

² Flexpartner.Energy, Stavanger, Norway

³ IT & Teknologi, Fagne AS, Haugesund, Norway

*E-mail: merkebu.z.degefa@uis.no

Abstract. This study investigates whether decentralised, AI-driven flexibility agents can collectively reduce transformer peak load in an industrial park, even when each agent has no real-time knowledge of aggregate transformer loading or co-located consumers' actions. A two-stage simulation framework is used. First, a centralised day-ahead linear programme (LP) solves the joint scheduling problem for four industrial consumers over a synthetic training year, minimising peak transformer load subject to a $\pm 10\%$ per-hour flexibility envelope and daily energy neutrality. Its solutions serve as labelled training data, not an operational schedule. Second, one Random Forest model is trained per industry to map local inputs (own load profile, hour-of-day, calendar features) to the LP-optimal flexibility response; no transformer measurements or other industries' data are used. On a held-out synthetic year never seen during training, the four independently operating models reduce annual peak transformer load by 4.1%, compared to 5.7% for the LP benchmark. They eliminate all transformer overloading hours present in the unoptimised baseline, improve the Peak-to-Average Ratio, and retain 95.8% of the LP cost saving. These results offer preliminary simulation evidence that centralised optimisation data could transfer implicit system-level knowledge to independently acting agents, enabling coordinated grid impact without real-time communication infrastructure. Scalability to larger consumer populations and robustness to real industrial load variability still require further investigation before deployment conclusions can be drawn.

1. Introduction

The accelerating integration of renewable energy sources and the electrification of industrial processes are placing growing stress on distribution networks. Industrial parks are clusters of manufacturing facilities sharing a common grid connection point, typically a medium-voltage transformer. This represent both a significant source of peak demand and an underutilised reservoir of controllable flexibility. Activating this flexibility offers Distribution System Operators (DSOs) a cost-effective alternative to capital-intensive network reinforcement, extracting additional capacity without physical infrastructure upgrades.

Demand response (DR) is the adjustment of electricity consumption in response to system conditions or price signals has matured as a cornerstone of demand-side management (DSM). Within the industrial sector, DR programmes shift load from high-demand to lower-cost intervals. Industrial flexibility is primarily categorised into load shedding, load shifting, and load substitution. Load shifting refers to the rescheduling of manufacturing processes to periods of lower grid demand. This approach is particularly relevant to industrial parks, as it preserves production targets while altering temporal consumption patterns. Integrating industrial flexibility into electricity markets can convert formerly rigid operational costs into dynamic sources of value through participation in energy and capacity markets [1,2]. Research on multi-type flexible loads in industrial parks has proposed multi-time-scale dynamic optimisation for integrated energy systems, demonstrating that DR strategies can shift demand across multiple energy carriers while reducing operational costs [3].

From an operational perspective, industrial processes exhibit diverse flexibility characteristics. A flexibility score framework [4] quantifies the degrees of freedom of individual energy flexibility measures (EFMs), providing a DSM potential indicator requiring only limited data. Multi-agent reinforcement learning (MARL) frameworks have emerged as particularly promising for industrial DR: a MARL-based energy management system modelling energy-intensive processes as Partially Observable Markov Decision Processes [5] incorporates buffer constraints and a flexible reward function balancing production efficiency against cost reduction which then naturally extending to the multi-facility industrial park context where each facility responds to shared price signals without explicit coordination.

This paper addresses flexibility management for industrial consumers connected to a common transformer, with the dual objective of shifting aggregate peak load at the transformer level while allowing each industrial actor to optimise its own operations. As a representative to individual operational objectives, the market price signal is utilized. The study adopts a two-stage imitation learning framework, also known as behavioural cloning. In the first stage, a centralised day-ahead linear programme solves the joint peak-shaving problem for four industrial consumers sharing a distribution transformer. The resulting optimal flexibility activations are then used as supervised labels to train one Random Forest regressor per consumer. At inference, each agent operates solely on local inputs (own load profile, hour-of-day, and calendar features), reproducing the coordination that the LP achieved jointly, without any real-time inter-agent communication.

2. Experiences on industrial flexibility resources activation

The distribution transformer serves as the critical interface between medium-voltage network infrastructure and downstream consumers. Transformer overloading is one of the primary triggers for costly network reinforcement. It is driven by coincident industrial demand peaks, electrification of heat and transport, and intermittent generation. Activating demand-side flexibility to manage transformer loading offers a market-based and operationally reversible alternative to asset upgrades. The economic case is well supported: peak load reduction translates directly into avoided capacity costs, reduced generation reserves, and deferred infrastructure investment [6–8]. Thermal modelling studies [9] have quantified overloading potential and associated ageing costs, demonstrating that price-sensitive industrial loads can cost-effectively postpone asset upgrades.

Flexibility services for congestion management have been studied extensively in the context of the European CoordiNet project, which demonstrated a DSO flexibility platform based on optimal power flow for congestion resolution [10]. The aggregator-centric model conveys congestion information through dynamic cost signals, enabling aggregators to reschedule consumer loads. While effective, this approach requires near-real-time data exchange between the DSO, aggregator, and end consumers; a limitation this paper explicitly seeks to relax through pre-trained decentralised models.

Price-based demand response, exposing consumers to time-varying prices reflecting system conditions, represents the most scalable and communication-light approach to flexibility activation. Unlike explicit DR, which requires bilateral communication of control signals, price-based DR relies on the autonomous response of consumers to market signals. Dynamic pricing has been shown to effectively stimulate demand-side flexibility in industrial contexts, with time-of-use and real-time pricing tariffs incentivising load shifts to off-peak periods without direct load control.

The design of optimal price signals has been formulated as a learning problem in several recent studies. Reinforcement learning (RL) approaches cast dynamic pricing as a Markov decision process, enabling adaptive retail price setting by learning from consumer responses. Offline RL for price-based DR design [11] demonstrates that offline pre-training using simulated demand response data can significantly reduce data costs and accelerate deployment convergence. Implicit flexibility approaches,

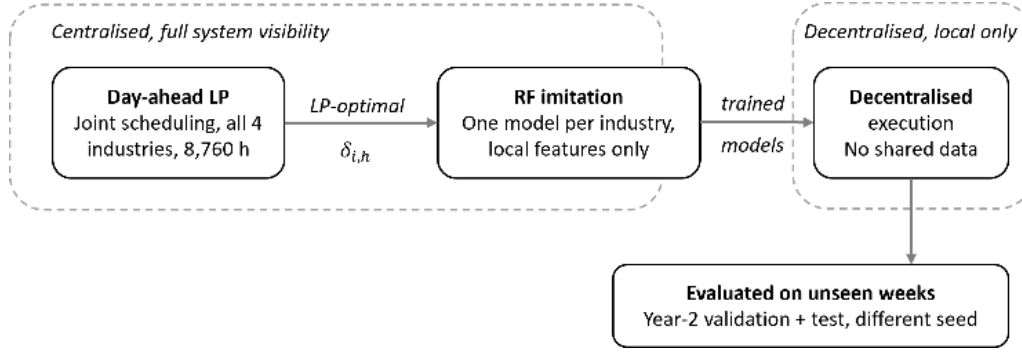


Figure 1. Two stage framework for training and executing decentralized flexibility resources

where machine learning determines optimal price signals to achieve congestion avoidance, have been demonstrated in real-world trials using consumer price sensitivity models built from historical data.

Privacy constraints are a defining challenge of industrial flexibility management. Federated learning (FL) has emerged as a leading paradigm for privacy-preserving coordination, enabling collaborative model training without sharing raw data [12]. Privacy-preserving federated reinforcement learning has been applied to virtual power plant scheduling [13,14], allowing multiple agents to collaboratively learn coordination policies without exposing individual data. These approaches demonstrate that coordination performance comparable to centralised solutions is achievable without privacy compromise.

The concept of treating flexibility activation as a training problem, where consumer response models are learned from historical data and deployed autonomously, represents a rapidly growing research direction. The LearnAMR framework in [15] combines adaptive model predictive control with RL, embedding trained facility models as MILP constraints within an optimisation framework to enable fast real-time energy optimisation without a continuous feedback loop with a central coordinator. Decentralised DR frameworks [10] and cooperative flexibility exchange architectures further demonstrate that decentralised models responding to common price signals can achieve coordination outcomes comparable to centralised solutions, validating the core architectural premise of the proposed approach.

Collectively, these experiences reveal four critical gaps that motivate this paper. First, existing frameworks address individual facilities or heterogeneous DER portfolios rather than the specific topology of an industrial park bound by a shared transformer capacity constraint. Second, most implementations require some form of real-time data exchange. A fully offline-trained architecture has yet to be formalised, in which agents are trained on centralised optimisation data with aggregate peak load reduction as the primary objective and market energy price retained as a secondary, facility-level cost objective. Third, transformer-level peak shifting has not previously been framed as a decomposed, facility-level training problem within a unified system. Fourth, federated and blockchain approaches avoid disclosing raw data during training, but the proposed framework eliminates data exchange entirely during operational activation. This is a qualitatively stronger privacy guarantee, and one more compatible with industrial operational realities.

3. The two-stage framework

The simulation framework was structured as a two-stage framework. In the first stage, a centralised day-ahead linear programme solved the joint scheduling problem for all four industries simultaneously over a full synthetic year (8,760 hours). The LP had full visibility of all baseline load profiles and the shared inflexible background load, and minimised the peak transformer load as its primary objective, with energy cost minimisation as a subordinate secondary goal. Each industry was constrained to shift at most $\pm 10\%$ of its hourly load, with daily energy neutrality enforced, total daily consumption was preserved exactly.

The LP solutions, specifically, the hourly flexibility activation decisions (delta values) for each industry, formed the labelled training dataset for the second stage. One Random Forest regression model was trained per industry, mapping local inputs (own load forecast, hour-of-day, day-of-year, day-of-week) to the LP-optimal flexibility response. Crucially, no transformer-level information or data from other industries was included as a model feature. At deployment, the four AI models operated fully

independently, each deciding its own flexibility activation based solely on local information. The methodology is illustrated in Figure 1.

Random Forest suits this task because the LP-optimal flexibility delta is a bounded, piecewise-smooth function shaped by the LP's own discrete constraints, which tree ensembles capture well without assuming a functional form. It also handles the mixed continuous and cyclic features with minimal preprocessing and tuning, making it a reasonable, computationally lightweight default, though not one tested here against simpler or more expressive alternatives.

3.1. Synthetic Data Generation

Hourly base load time series are generated for all 8 760 hours of the reference year 2024 (a leap year). Each industry's load is characterised by a bell-shaped daily profile, a seasonal modulation factor, and Gaussian noise:

$$P(h) = base + (1 - base) \cdot \exp \left[-\frac{1}{2} \left(\frac{h - peak}{\sigma} \right)^2 \right] \times seasonal(doy) \times noise \quad (1)$$

where:

- *base*: base load which is assumed to be 0.6 MW
- *peak*: the hour of day at which the profile peaks. Gaussian bell curve centred at the industry's characteristic peak hour.
- σ : the standard deviation in hours; controls how wide/narrow the peak is (larger $\sigma \rightarrow$ broader peak)
- $seasonal(doy) = 0.85 + 0.15 \cdot \left(\frac{2\pi(doy-15)}{365} \right)$ winter peak, summer trough. *doy* = day of year (1–365)
- $noise \sim N(0, 0.03)$: Gaussian noise simulating stochastic process variability.

Day-ahead prices are generated with two daily peak windows (morning 08:00–10:00 and evening 17:00–20:00) receiving the peak price of 150 €/MWh, a seasonal premium of $20 \cdot (1 + \cos(2\pi \cdot doy/365))$ €/MWh representing winter scarcity, and Gaussian noise $\sigma = 5$ €/MWh. Prices are clipped to [10, 300] €/MWh.

$$price(h) = price_{base}(h) + seasonal_{premium}(doy) + noise \quad (2)$$

3.2. Stage 1: Centralised Day-Ahead LP Training

For each of the 365 days in the year, a linear programme is solved jointly over all four flexible industries. The decision variable $\delta[i, h]$ (MW) represents the flexibility activation in hour *h* for industry *i*. A positive value indicates load shedding (demand reduction); a negative value indicates a rebound (additional consumption to compensate for an earlier shed).

Decision Variables

- $\delta_{i,h} \in \mathbb{R} \quad \forall i, h$ (flexibility shift, MW – positive = load shed)
- $M \in \mathbb{R}_{\geq 0}$ (daily peak transformer load, MW)

Indices and sets

- $i \in \{1,2,3,4\}$ – flexible industries
- $h \in \{0,1, \dots, 23\}$ – hours of the day

Dual Objective Function

The objective simultaneously minimises transformer peak load (primary, weight $w_{peak} = 1.0$) and total energy cost across all industries (secondary, weight $w_{cost} = 0.001$):

$$\min_{\delta, M} W_{peak} \cdot M - W_{cost} \cdot \sum_i \sum_h p_h \cdot \delta_{i,h} \quad (3)$$

The secondary term rewards shedding ($\delta > 0$) in expensive hours, aligning cost reduction with peak reduction. The weight ratio 1.0 : 0.001 ensures the peak objective always dominates.

Constraints

Peak load definition: M must be at least as large as the transformer load in every hour:

$$\sum_i (b_{i,h} - \delta_{i,h}) + f_h \leq M \quad \forall h \quad (4)$$

Energy neutrality: each industry must consume the same total daily MWh as its unscheduled baseline:

$$\sum_i \delta_{i,h} = 0 \quad \forall i \quad (5)$$

Flexibility bounds: shifts are bounded by $\pm 10\%$ of the baseline load each hour:

$$-\alpha \cdot b_{i,h} \leq \delta_{i,h} \leq +\alpha \cdot b_{i,h} \quad \forall i, h \quad (6)$$

Non-negativity of peak variable:

$$M \geq 0 \quad (7)$$

Solution

The LP is formulated with $n \cdot T + 1 = 97$ decision variables (4 industries $\times$ 24 hours + 1 peak variable M), 24 inequality constraints (one per hour), 4 equality constraints (one per industry), and 97 bound constraints. The HiGHS solver (via `scipy.optimize.linprog`) produces the globally optimal $\delta^*[i, h]$ for each day. Solutions across all 365 days generate the complete annual label dataset for Stage 2.

Once the LP is solved, the actual scheduled load for industry i at hour h is:

$$\hat{b}_{i,h} = b_{i,h} - \delta_{i,h} \quad (8)$$

And the total transformer load at hour h :

$$T_h = \sum_i \hat{b}_{i,h} + f_h = \sum_i b_{i,h} - \sum_i \delta_{i,h} + f_h \leq M \quad (9)$$

3.3. Stage 2: Decentralised Activator Training

One Random Forest regressor is trained per industry using only local features. Features include price, local load forecast, hour-of-day, and day-of-year signals. The target is the LP-derived hourly flexibility delta for each industry. This creates a decentralised policy that does not require direct exchange of raw industry data. All features are computable by the industry locally, using only its own metered load history and the broadcast day-ahead price signal (See table 1).

Table 1. Feature Vector (per industry, per hour)

Feature	Description
Market price [EUR/MWh]	The single coordination signal, raw market energy price, no augmentation
Local load forecast [MW]	Day-ahead load forecast for that specific industry
$\sin/\cos(2\pi \cdot h/24)$	Cyclic hour-of-day encoding (captures intra-day patterns)
$\sin/\cos(2\pi \cdot \text{doy}/365)$	Cyclic day-of-year encoding (captures seasonal patterns)
Day-of-week (0–6)	Distinguishes weekday vs. weekend behaviour

The Random Forest Regressor model consists of 250 decision trees, each trained on a bootstrap sample of the data and averaged at prediction time to reduce variance. A minimum leaf size of 2 samples provides mild regularisation by preventing leaves from fitting to single observations, offering a slight bias–variance trade-off over the default fully-grown setting. All available CPU cores are used in parallel to accelerate training, while a fixed random seed ensures full reproducibility of both bootstrap sampling and feature subsampling across runs.

Table 2. Synthetic Data Generation

	Peak hour	Capacity	Time spread (σ)
FlexInd#1	10	3 MW	3 h
FlexInd#2	14	2.5 MW	4 h
FlexInd#3	19	2 MW	3 h
FlexInd#4	22	1.5 MW	5 h
Inflexible Background	18	5 MW	5 h

3.4. Decentralised Inference and Energy Neutrality

At deployment time, each industry runs its trained local model hour-by-hour through the scheduling day. The procedure is iterative: the model predicts $\hat{\delta}[h]$ at each hour h using the current feature vector, which includes the cumulative energy balance accumulated from predictions in hours 0 through $h-1$ and applies the result before moving to hour $h+1$.

The inference loop for each industry proceeds as follows:

1. Receive broadcast day-ahead price vector $\text{price}[0..23]$.
2. Compute own base load forecast from historical meter data.
3. Initialise cumulative balance: $\text{cum} = 0$.
4. For each hour $h = 0..23$: assemble feature vector $X[h]$, predict raw activation, clip to $\pm 10\%$ bounds, apply neutrality correction, update $\text{cum} += \hat{\delta}[h]$.
5. Submit scheduled load profile: $\text{sched}[h] = \text{base}[h] - \hat{\delta}[h]$.

To preserve the energy neutrality constraint (total daily consumption unchanged) without access to a central solver, each industry's inference loop applies a soft neutrality correction. When the cumulative balance is drifting away from zero and insufficient hours remain to correct it, the raw prediction is steered toward the compensating direction:

$$\hat{\delta}[h] \leftarrow \text{clip}\left(\hat{\delta}[h] - \frac{1}{24} \sum_{h'=0}^{23} \hat{\delta}_i[h'], -\text{Flex}_{\text{down}}, -\text{Flex}_{\text{up}}\right) \quad (9)$$

At the final hour ($h = 23$), a hard correction forces neutrality:

$$\hat{\delta}[23] \leftarrow \text{clip}\left(\frac{1}{24} \sum_{h'=0}^{23} \hat{\delta}_i[h'], -\text{Flex}_{\text{down}}, -\text{Flex}_{\text{up}}\right) \quad (10)$$

4. Case study

An industrial park consisting of multiple industries connected to a common secondary substation transformer is considered in this study (see Figure 2). Four of the industries are considered to have flexibility potentials while the rest of the industries are aggregated as inflexible loads. The flexibility management plan is implemented for the four industries. For synthetic data generation the parameter listed in table 2 are utilized.

To generate price signal the following parameters are considered:

- Base: 50 EUR/MWh
- Morning/Evening peaks: 150 EUR/MWh
- Seasonal premium + noise

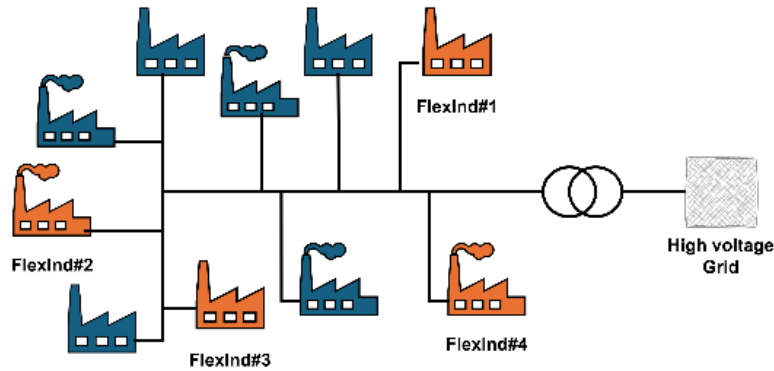


Figure 2. Industries in industrial park sharing transformer. Part of the industries subscribed to flexibility provisioning.

5. Results and discussion

The centralised LP represents the best achievable outcome given the $\pm 10\%$ flexibility envelope and full system observability, and serves here as the benchmark against which the decentralised AI is judged on a held-out synthetic year never used for training. On this unseen year, the LP reduced the annual peak transformer load from 12.426 MW to 11.714 MW, a 5.7% reduction, and eliminated all 49 hours per year during which the transformer exceeded its 12 MVA rated capacity. The decentralised AI, despite having no access to transformer-level measurements or other industries' data, reduced the same annual peak to 11.915 MW, a 4.1% reduction, and matched the LP in eliminating every hour of overload. The Peak-to-Average Ratio improved from 1.537 to 1.449 under the LP and to 1.474 under the AI, both representing a measurably flatter annual load profile. Total energy costs for the flexible industries fell by 3.20% under the LP and by 3.07% under the AI, meaning the decentralised agents retained 95.8% of the LP's cost saving while using only local information.

These results establish the LP as the upper-bound reference against which the decentralised AI must be evaluated. The LP achieves this outcome through perfect foresight and full system visibility; conditions that are not replicable in real-time decentralised deployment.

The core finding of this study is that the decentralised AI models, operating with no knowledge of the transformer state and no communication with each other, replicated the LP outcome with high fidelity. The KPI's from the simulation results are presented in table 3.

Table 3. LP Centralised vs AI Decentralised evaluation with out-of-Sample data

KPI	Baseline	LP Centralised	AI Decentralised
Annual Peak (MW)	12.43	11.71 (+5.7%)	11.92 (+4.1%)
Avg Daily Peak (MW)	10.30	9.70	9.87
Annual Energy Cost	4.750	4.598 (saving 152 kEUR)	4.604 (saving 146 kEUR)
Hours Over 12 MVA	49 h	0 h	0 h
Peak-to-Average Ratio	1.5372	1.4491	1.4737
Price Signal Used	-	Market EUR/MWh only	Market EUR/MWh only
Data Exchange	-	Centralised (all industries)	None (fully local)

Figure 3 shows the decentralized AI model rescheduled load profiles of the flexible industries in a typical week. More importantly, Figure 4 shows the comparison of LP-based and decentralized AI-based

total transformer load profiles for a typical week. These results provide strong evidence in support of the central hypothesis. The AI models, trained entirely on LP-generated labels, acquired an implicit understanding of how each industry's individual flexibility activation contributes to the collective transformer load reduction. This knowledge was encoded during the offline training phase and was reproduced faithfully at deployment, without any real-time system-level information.

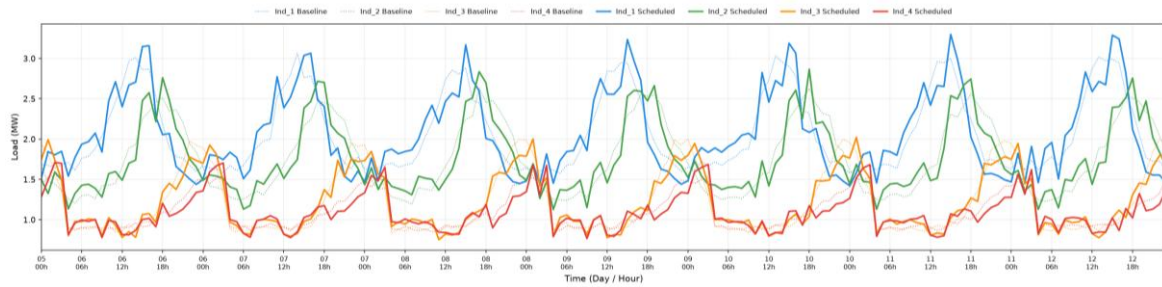


Figure 3. Baseline vs AI-scheduled load profile per industry for a typical week

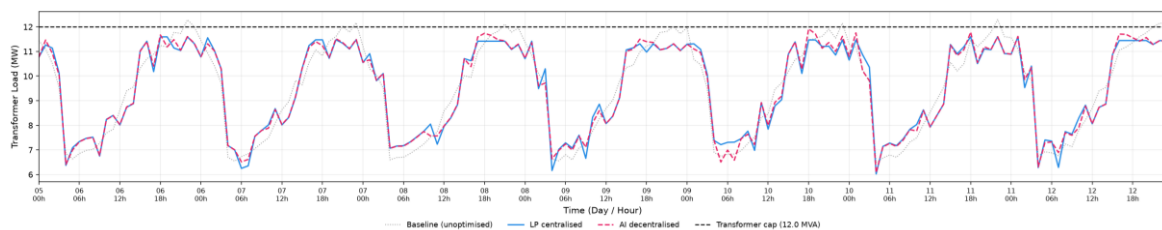


Figure 4. Transformer Aggregate Load- LP vs AI (Typical Winter Week, Jan 15-21)

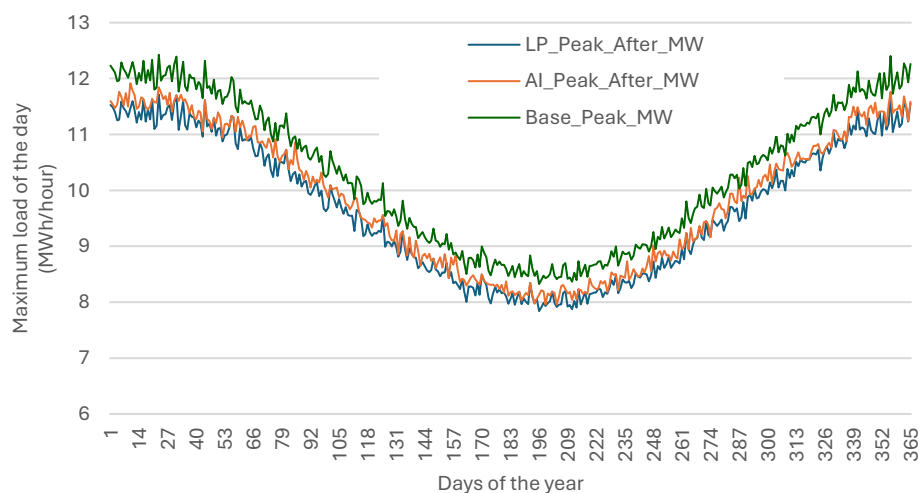


Figure 5. Comparison of Transformer level daily peak load among the baseline profile, the LP optimized and the AI model output profiles

The near-zero gap between LP and AI on the overload-hours metric is particularly significant: this is the operationally critical outcome from the perspective of transformer protection and grid reliability (see Figure). The AI deployment met this objective as completely as the theoretical optimum.

The following limitations should be considered when interpreting these results:

- Load profiles were generated synthetically from parameterised Gaussian shapes. Real industrial loads may exhibit stronger non-stationarity, process coupling, or equipment-level constraints not captured here.
- The $\pm 10\%$ flexibility envelope is applied uniformly and continuously. In practice, flexibility may be time-restricted, asymmetric, or subject to ramp-rate limits specific to each industry's processes.
- The AI models were trained and evaluated on the same synthetic distribution. Robustness to distribution shift, e.g. new seasonal patterns, demand anomalies, or structural load changes, was not assessed.
- The study assumes that LP solutions from historical or simulated data are sufficiently representative of future operating conditions for effective knowledge transfer. Periodic retraining may be required as load patterns evolve.
- The simulation does not model network power flow, voltage constraints, or reactive power, which would be relevant in a real distribution network context.

Future extensions could explore reinforcement learning agents, adaptive retraining pipelines triggered by load pattern drift, multi-energy flexibility (thermal storage, compressed air), and validation against real industrial metering data.

6. Conclusion

The central question of this study is whether individual industrial consumers can activate demand-side flexibility in a way that reduces the aggregate load on a shared distribution transformer, without access to real-time transformer measurements, and without direct communication with each other. The hypothesis is that a decentralised AI model, trained exclusively on the output of a centralised LP optimisation, inherits sufficient system-level knowledge to behave in a grid-beneficial manner at deployment time, even though it operates in complete informational isolation.

The results support the hypothesis:

- The decentralised AI agents eliminated all transformer overload events, matching the LP benchmark on the most critical grid reliability metric
- The annual peak load was reduced by 4.1% on unseen data (vs 5.7% for LP), with a gap of 0.20 MW between the two approaches
- 95.8% of the LP energy cost saving was retained on unseen data, demonstrating that the system-level objective generalises beyond the training year, through the knowledge-transfer training mechanism
- Each AI model operated in complete informational isolation at inference. No transformer measurements, no inter-industry communication, no real-time coordination signal

These findings suggest the proposed framework, centralised optimisation as an offline knowledge source, and decentralised AI as the operational delivery mechanism, is a viable approach, though further validation is needed. The approach could offer a practical pathway to scalable, privacy-preserving, and infrastructure-light demand-side flexibility management in shared-grid industrial settings.

Acknowledgment

The authors acknowledge the contribution of partners in FlexLEC project and the funding from Ulla-Førrefondet.

Generative AI was used to assist with editing and refining the language of this manuscript for clarity and conciseness; all technical content, analysis, and conclusions are the authors' own. The authors reviewed and take full responsibility for the final text.

References

- [1] Nolzen N, Ganter A, Baumgärtner N, Leenders L and Bardow A 2022 Where to Market Flexibility? Optimal Participation of Industrial Energy Systems in Balancing-Power, Day-Ahead, and Continuous Intraday Electricity Markets *arXiv preprint arXiv:2212.12507*.
- [2] Garfield D, Jadun P, Hwang S, O'Malley M and Ruth M 2021 Opportunities for Industry to Provide Flexibility While Increasing Profitability. No. NREL/TP-6A50-75784. National Renewable Energy Laboratory (NREL), Golden, CO (United States)
- [3] Cui J, Gao M, Zhou X, Li Y, Liu W, Tian J and Zhang X 2023 Demand response method considering multiple types of flexible loads in industrial parks *Engineering Applications of Artificial Intelligence* **122** 106060
- [4] Rusche S, Weissflog J, Wenninger S and Häckel B 2023 How flexible are energy flexibilities? Developing a flexibility score for revenue and risk analysis in industrial demand-side management *Applied Energy* **345** 121351
- [5] Bashyal A, Boroukhan T, Veerachanchai P, Naransukh M and Wicaksono H 2025 Multi-agent deep reinforcement learning based demand response and energy management for heavy industries with discrete manufacturing systems *Applied Energy* **392** 125990
- [6] Holweger J, Pena-Bello A, Jeannin N, Ballif C and Wyrsh N 2023 Distributed flexibility as a cost-effective alternative to grid reinforcement *Sustainable Energy, Grids and Networks* **34** 101041
- [7] Nolden C, Banks N, Irwin J, Wallom D and Parrish B 2025 The economics of flexibility service contracting in local energy markets: A review *Renewable and Sustainable Energy Reviews* **215** 115549
- [8] Brylle L, Andersen N and Madsen H 2026 Scaling Demand-Side Flexibility Through Dynamic Tariffs
- [9] Cheilas D, Bindner H W and Weckesser T 2025 Investigating the benefit of flexibility services in distribution grids under uncertainty *Sustainable Energy, Grids and Networks* **41** 101582
- [10] García-Muñoz F, Ivanova A, Farré Montané J, Serrano M and Corchero C 2023 A DSO flexibility platform based on an optimal congestion management model for the European CoordiNet project *Electric Power Systems Research* **221** 109386
- [11] Jang D, Spangher L, Srivistava T, Khattar M, Agwan U, Nadarajah S and Spanos C 2021 Offline-online reinforcement learning for energy pricing in office demand response: lowering energy and data costs *Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation BuildSys '21: The 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (Coimbra Portugal: ACM) pp 131–9
- [12] Dong Y, Chen G and Qin J 2025 Federated Aggregation of Demand Flexibility *arXiv preprint arXiv:2509.19612*
- [13] Yang T, Feng X, Cai S, Niu Y and Pen H 2025 A Privacy-Preserving Federated Reinforcement Learning Method for Multiple Virtual Power Plants Scheduling *IEEE Trans. Circuits Syst. I* **72** 1939–50
- [14] Feng B, Liu Z, Huang G and Guo C 2023 Robust federated deep reinforcement learning for optimal control in multiple virtual power plants with electric vehicles *Applied Energy* **349** 121615
- [15] Liu J, Wu B, Meng X, Wu J and Ma Z 2025 LearnAMR: Learning-based adaptive model predictive control enhanced by reinforcement learning for optimizing energy flexibility in building energy systems incorporating demand-side management *Applied Energy* **401** 126707