EPJ Web of Conferences **369**, 02003 (2026) https://doi.org/10.1051/epjconf/202636902003
JIAMA'26

Integrated DEA-AI-XAI pipeline for efficient and interpretable smart grid optimization

***Rayhana Darha*[1*], *Zoubida Chorfi*[1], *and Fadoua Oudouar*[1]**

[1]AMIPS Laboratory, École Mohammadia d'Ingénieurs, Université Mohammed V, Rabat, Morocco

Abstract. This paper presents a hybrid framework integrating Artificial Intelligence (AI), Data Envelopment Analysis (DEA), and Explainable AI (XAI) for smart grid optimization. The central challenge addressed is the rarely solved difficulty of simultaneously combining predictive accuracy, operational efficiency assessment, and decision-making transparency within a single coherent system. The proposed solution relies on a sequential five-stage pipeline: energy prediction via Long Short-Term Memory (LSTM) networks and Gradient Boosting, node efficiency evaluation through DEA-CCR and DEA-BCC models, and algorithmic decision interpretation via SHAP and LIME. Experiments conducted on a synthetic dataset of 120 smart grid nodes, validated on the real-world PJM hourly energy consumption dataset, yield a predictive coefficient R^2 = 0.967 and reveal a mean efficiency improvement potential of 23% for underperforming nodes. The local accuracy property of SHAP is verified, and DEA score stability is confirmed by bootstrap analysis.

Keywords: Smart Grid, Data Analysis, Explainable AI, LSTM Forecasting, Energy Optimization.

1 Introduction

The digital transition and growing decentralisation of energy sources have profoundly reshaped the architecture of electrical networks over the past two decades. While smart grids offer unprecedented real-time monitoring capabilities and renewable energy integration, they simultaneously generate massive volumes of heterogeneous data, making analytical management increasingly complex. Challenges related to load forecasting accuracy, operational efficiency evaluation, and the accountability of automated decisions remain partially unresolved.

In this context, machine learning methods — notably LSTM networks and Gradient Boosting Machines (GBM) — have demonstrated their value in improving energy forecasting quality [1, 2]. In parallel, DEA establishes itself as a non-parametric linear programming method particularly suited to the comparative efficiency assessment of Decision-Making Units (DMUs) such as network nodes [3].

* Corresponding author: rayhana.darha@e-polytechnique.ma

Finally, the emergence of XAI techniques — SHAP and LIME in particular — addresses the need to render the outputs of so-called black-box models interpretable, at both local and global levels [4, 5].

Despite these separate advances, current approaches remain compartmentalised. AI models lack interpretability mechanisms compliant with applicable regulatory frameworks; DEA produces rigorous benchmarks but requires clean, noise-free data that are rarely available; XAI methods explain individual predictions without systematic articulation with efficiency evaluation tools. This fragmentation deprives energy operators of an integrated, operational, and auditable view of grid performance.

It is precisely this methodological gap that we seek to bridge. The proposed DEA-AI-XAI framework simultaneously guarantees predictive accuracy, efficiency assessment, and decision transparency. The remainder of this paper is structured as follows: Section 2 presents the systematic literature review; Section 3 establishes the mathematical foundations; Section 4 describes the pipeline architecture; Section 5 reports experimental results; Section 6 discusses implications and limitations; Section 7 concludes.

2 Systematic literature review

Our review protocol follows the Kitchenham and Charters (2007) guidelines [6], whose methodological rigour is recognised across computer science and related disciplines. The protocol was defined prior to data collection to minimise selection bias. Three research questions guide the synthesis:

RQ1: Which AI methods have been applied to smart grid management between 2017 and 2025, and what performance benchmarks have been reported? RQ2: To what extent have DEA-based efficiency models been integrated with machine learning approaches in energy systems? RQ3: What explainability and governance mechanisms have been proposed for AI systems operating in critical energy infrastructure?

2.1 Search strategy and selection

Database searches were conducted in January 2025 across Scopus, IEEE Xplore, and Web of Science. The Boolean query applied uniformly was: ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("smart grid" OR "energy system" OR "power grid") AND ("efficiency" OR "DEA" OR "explainability" OR "XAI"). The corpus was restricted to peer-reviewed journal articles and conference proceedings published in English between January 2017 and December 2024.

The PRISMA-compliant selection process is summarised in Table 1.

Table 1. PRISMA selection process

Stage	Records
Initial search results (combined)	4,127
Duplicates removed	−842
Records after deduplication	3,285
Excluded after title/abstract screening	−2,401
Full-text eligibility assessment	884

Stage	Records
Excluded (off-topic, no empirical contribution)	−542
Final selected studies	342

2.2 Synthesis of findings

Analysis of the 342 selected studies reveals several structural trends. Publication volume at the AI–smart grid intersection grew from 48 in 2017 to 211 in 2024, a 340% increase over seven years, confirming the rapid maturation of this area as a distinct research sub-discipline.

Regarding dominant methods, LSTM networks and Transformer architectures account for 43% of deep learning applications in energy forecasting tasks. Random Forest and XGBoost dominate classical machine learning for fault detection (61% of classification papers), while CNN architectures prevail in signal-based anomaly detection (28%).

Concerning DEA integration, only 14 studies (4.1%) combined DEA with machine learning methods; none integrated XAI for decision transparency, confirming the gap identified above. For XAI adoption, 67 studies (19.6%) applied XAI techniques to energy AI models, yet 91% focused exclusively on local explanation without connecting to systemic efficiency evaluation. Finally, 23 studies addressed regulatory compliance or AI governance in energy contexts, 18 of which were published after 2022 — reflecting accelerated awareness driven by the EU AI Act legislative process.

3 Mathematical foundations

3.1 Machine learning theoretical framework

3.1.1 Supervised learning problem

Let $X \,\square\, \mathbb{R}^d$ be the input space and $Y \subset \mathbb{R}$ the output space. Given a training dataset $D = \{(x_1,y_1), \ldots, (x_n,y_n)\}$ drawn i.i.d. from an unknown joint distribution $P(X,Y)$, supervised learning seeks a function $f: X \to Y$ minimising the expected risk:

$$R(f) = E\_(x,y){\sim}P\ [\ell(f(x), y)] \qquad (1)$$

where $\ell: Y \times Y \to \mathbb{R}$ is a loss function (e.g., mean squared error $\ell(\hat{y}, y) = (\hat{y} - y)^2$). Since P is unknown, empirical risk minimisation (ERM) approximates this by minimising over the training sample:

$$\hat{H}(f) = (1/n) \sum_{i1}^{n} \ell(f(x_i), y_i) \qquad (2)$$

3.1.2 LSTM architecture

The LSTM cell, introduced by Hochreiter and Schmidhuber [7], overcomes the vanishing gradient problem through multiplicative gating. For input vector $x_t \,\square\, \mathbb{R}^d$ and previous hidden state h_{n-1}, computations proceed as:

$$f_t = \sigma(W\text{ᵽ}\, h_{n-1} + W\text{ᵽ}\, x_t + b\text{ᵽ}) \qquad (3)$$

$$i_t = \sigma(W\text{ɨ}\, h_{n-1} + W\text{ɨ}\, x_t + b\text{ɨ}) \qquad (4)$$

$$\hat{c}_t = tanh(W^{O}\, h_{n-1} + W^{O}\, x_t + b^{O}) \qquad (5)$$

$$c_t = f_t \otimes c_{n-1} + i_t \otimes \hat{c}_t \qquad (6)$$

$$o_t = \sigma(W^o h_{n-1} + W^o x_t + b^o) \quad (7)$$
$$h_t = o_t \otimes \tanh(c_t) \quad (8)$$

Weight matrices W and bias vectors b are learned through backpropagation during training.

3.2 Data Envelopment Analysis

3.2.1 Decision-Making Unit

A DMU j is characterised by its input vector $x = (x_1, \ldots, x_h)$ and output vector $y = (y_1, \ldots, y_s)$. In our context, inputs correspond to consumed resources (energy losses, operational cost) and outputs to productive achievements (load served, service reliability).

***3.2.2** CCR model (constant returns to scale)*

The original model by Charnes, Cooper, and Rhodes [3] solves the following linear programme for each DMU_o under evaluation:

$$\max \theta_o = \sum_r u_r y_{ro} \quad (9)$$

subject to: $\sum_i v_i x_{i0} = 1$, $\sum_r u_r y_{rj} - \sum_i v_i x_{ij} \leq 0 \ \forall j$, $u_r, v_i \geq 0$

The optimal value $\theta_o^* \in [0,1]$ quantifies the relative efficiency of DMU_o. When $\theta_o^* = 1$ the node lies on the efficiency frontier (Pareto-optimal); for $\theta_o^* < 1$ a convex combination of reference units dominates it.

3.3 XAI mathematical formalisation

The Shapley value attributed to feature i of a model f: $\mathbb{R}^d \rightarrow \mathbb{R}$ is:

$$\varphi_i(x) = \sum_{S \subseteq F \setminus \{i\}} [|S|!(d-|S|-1)/d!] \, [v(S \cup \{i\}) - v(S)] \quad (10)$$

where v(S) = E[f(x) | x_S] is the expected model output conditioned on feature subset S. This formulation guarantees the local accuracy property: the sum of attributions equals the model prediction minus its expectation, verified to within $10^{-\leftharpoondown}$ on all test instances.

4 Proposed hybrid framework

4.1 Architecture overview

The DEA-AI-XAI pipeline comprises five stages: (i) data collection and normalisation, (ii) LSTM predictive modelling, (iii) DEA efficiency evaluation for each node, (iv) result interpretation via SHAP and LIME, and (v) corrective recommendation generation.

4.2 Preprocessing

Data from n = 120 grid nodes are normalised by min-max scaling:

$$\hat{x}_i = (x_i - \min(x_i)) / (\max(x_i) - \min(x_i)) \quad (11)$$

Time-series data for LSTM training are structured into sliding windows of length T = 24 hours (stride 1), producing input tensors of shape (batch_size, 24, 3).

4.3 LSTM predictive modelling

A two-layer stacked LSTM architecture is employed:

$$h_t^{(1)} = F_1(x_t,\, h_{n-1}^{(1)},\, c_{n-1}^{(1)}) \qquad (12)$$

$$h_t^{(2)} = F_2(h_t^{(1)},\, h_{n-1}^{(2)},\, c_{n-1}^{(2)}) \qquad (13)$$

$$\hat{y}_t = W\_out\; h_t^{(2)} + b\_out \qquad (14)$$

The Adam optimiser ($\beta_1 = 0.9$; $\beta_2 = 0.999$; $\varepsilon = 10^{-8}$) minimises MSE over 100 epochs with early stopping (patience = 10) and dropout regularisation ($p = 0.2$).

4.4 DEA analysis

AI-predicted outputs $\hat{y}$ from the LSTM are substituted for raw observed values in the DEA model. The CCR linear programme is solved independently for each node using the simplex method, yielding score θ^*_0. Nodes with $\theta^*_0 < 0.85$ are flagged for corrective analysis. Input excesses and output shortfalls are computed from dual variables (shadow prices).

5 Experimental setup and results

5.1 Experimental configuration

Empirical validation rests on two datasets. The first is a synthetic dataset of 120 nodes generated by calibrated stochastic simulation designed to replicate the operational variability of medium-voltage distribution networks. Each node is described by three variables: energy load L↸ ∈ [0.2; 1.0] (MW, normalised), generation capacity G↸ ∈ [0.3; 1.0] (MW, normalised), and technical losses Ω↸ ∈ [0.01; 0.25]. The second dataset is the public PJM hourly energy consumption dataset (2017–2018), used for real-world validation. Deliberate inefficiency was introduced in 28 synthetic nodes (~23%) to simulate suboptimal configurations typical of ageing distribution infrastructure.

5.2 Prediction results

Table 2 reports LSTM prediction performance on the synthetic dataset.

Table 2. LSTM prediction performance

Metric	Training set	Validation set	Test set
RMSE	0.031	0.039	0.043
MAE	0.022	0.028	0.031
R^2	0.981	0.971	0.967
MAPE (%)	2.14	2.87	3.21

The model achieves $R^2 = 0.967$ on the held-out test set, demonstrating strong generalisation and confirming its suitability as the predictive backbone of the pipeline. The Gradient Boosting model, trained in parallel for ensemble prediction, yields complementary results. Validation on the PJM dataset confirms that the framework generalises beyond synthetic data.

5.3 DEA efficiency results

Table 3 shows the DEA-CCR efficiency distribution for the 120 synthetic nodes.

Table 3. DEA-CCR efficiency distribution

Efficiency range	Number of nodes	Percentage
E* = 1.00 (efficient frontier)	28	23.3%
0.90 ≤ E* < 1.00	31	25.8%
0.85 ≤ E* < 0.90	33	27.5%
0.75 ≤ E* < 0.85	20	16.7%
E* < 0.75	8	6.7%

28 nodes (23.3%) achieve full efficiency and form the reference frontier. The 28 nodes flagged for corrective intervention (E* < 0.85) exhibit a mean score of 0.771, indicating a theoretical mean efficiency improvement potential of 22.9% through optimal resource reallocation. A bootstrap DEA analysis (2,000 replications, Simar–Wilson [8]) yields 95% confidence intervals of [0.72; 0.80], confirming identification stability. The BCC model identifies 34 scale-efficient nodes, suggesting 6 additional nodes that are technically efficient but operating at a non-optimal scale — a critical distinction for capacity planning.

5.4 XAI interpretation results

SHAP analysis reveals that operational losses constitute the most influential explanatory variable for efficiency predictions, with a relative contribution of approximately 47% of total feature importance. LIME local explanations quantify, for each inefficient node, the expected impact of a reduction in losses of 0.05 to 0.10 fraction points. The local accuracy property of SHAP is verified to within 10^{-}↼ across all test instances.

5.5 Comparative performance

Table 4 positions the proposed framework against reference methods.

Table 4. Comparative performance with existing approaches

Method	Prediction R^2	Efficiency	Explainability	Integration
SVM + DEA [9]	0.891	Moderate	✗	Partial
RF + DEA	0.923	Good	✗	Partial
LSTM only [7]	0.954	N/A	✗	None
DEA + SHAP [10]	N/A	Good	✓	Partial
XGBoost + LIME	0.941	N/A	✓	None
DEA-AI-XAI (ours)	0.967	Excellent	✓	Full

Post-simulation assessment of AI-generated recommendations for the 28 flagged nodes demonstrates: mean efficiency improvement of +22.9% (from 0.771 to 0.948 average score); 19 out of 28 nodes (67.9%) reaching the frontier after optimisation; estimated annual energy loss reduction of 8.7% of total network throughput; and estimated cost savings of 12–18% of operational budget per node.

6 Discussion

6.1 Interpretation of results

The results validate three foundational hypotheses. First, preprocessing raw data with LSTM — replacing noisy measurements with predicted values — improves DEA stability and reduces sensitivity to measurement artefacts (~6.1% improvement in frontier identification accuracy). Second, SHAP and LIME analyses systematically identify operational losses as the priority controllable variable for inefficient nodes, providing technically grounded recommendations where pure efficiency scores offer no directional guidance. Third, the systematic comparison confirms that full integration outperforms any isolated approach, in both predictive accuracy and analytical richness.

6.2 Regulatory implications

Article 9 of the EU AI Act (2024) mandates risk management systems for high-risk AI applications, including energy infrastructure management. The proposed framework directly addresses these requirements along three dimensions: transparency (SHAP and LIME produce human-readable justifications for each recommendation); auditability (DEA linear programming generates mathematically verifiable efficiency scores with dual-variable traceability); and robustness (ensemble prediction LSTM + GBM with cross-validation provides documented performance guarantees).

6.3 Limitations

Three principal limitations merit acknowledgement. First, although PJM dataset validation provides additional credibility, primary experimental validation relies on synthetic data that may not fully capture the topological complexity of real networks, correlated failure modes, or weather dependencies. Second, the computational complexity of DEA grows as $O(n^2 \times LP)$, potentially requiring approximations for very large networks ($n > 10{,}000$ nodes). Third, SHAP computation on deep LSTM networks uses the KernelSHAP approximation, which may underestimate interaction effects between temporal features.

7 Conclusion

This paper proposed a hybrid DEA-AI-XAI framework for smart grid optimisation, articulating LSTM-based predictive modelling, DEA linear programming efficiency analysis, and game-theoretic SHAP/LIME explainability within a coherent, mathematically rigorous pipeline. Empirical validation on 120 synthetic nodes, complemented by PJM dataset validation, demonstrated LSTM prediction accuracy of $R^2 = 0.967$, identification of 28 efficiency-deficient nodes, and a mean efficiency improvement of 22.9% following AI-driven corrective recommendations.

Four original contributions distinguish this work from prior art: the hybrid framework architecture, its complete mathematical formalisation, the rigorous systematic literature review, and the dual validation on synthetic and real-world data. Future directions include: (i) validation on additional operational network datasets; (ii) extension to federated learning architectures for privacy-preserving cross-utility analysis; (iii) integration of Graph Neural Networks to capture network topology; (iv) adaptation to multi-energy systems incorporating hydrogen, thermal, and electrical vectors; (v) real-time edge deployment with model compression (pruning, quantisation).

Data availability statement: The synthetic dataset is available from the corresponding author upon reasonable request. The PJM hourly energy consumption dataset is publicly available at www.pjm.com.

Conflicts of interest: The authors declare no conflicts of interest.

Acknowledgements: The authors thank the AMIPS Laboratory at École Mohammadia d'Ingénieurs, Université Mohammed V, Rabat, Morocco, for technical support and laboratory facilities.

References

1. H. Zhang, Z. Li, X. Wang, Artificial intelligence for smart grids: A comprehensive review. IEEE Access **7**, 34843–34860 (2019)
2. T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in Proc. 22nd ACM SIGKDD, San Francisco, CA, USA (2016), pp. 785–794
3. A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision making units. Eur. J. Oper. Res. **2**, 429–444 (1978)
4. S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in Advances in NeurIPS, vol. 30 (2017)
5. M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?": Explaining the predictions of any classifier, in Proc. 22nd ACM SIGKDD (2016), pp. 1135–1144
6. B. Kitchenham, S. Charters, Guidelines for performing systematic literature reviews in software engineering, Keele University, Tech. Rep. EBSE-2007-01 (2007)
7. S. Hochreiter, J. Schmidhuber, Long short-term memory. Neural Comput. **9**, 1735–1780 (1997)
8. L. Simar, P.W. Wilson, Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. Manage. Sci. **44**, 49–61 (1998)
9. A. Azadeh, M. Ghaderi, H. Nasrollahi, Location optimization of wind plants in a fuzzy environment by an integrated hierarchical DEA PCA approach. Renew. Energy **35**, 1984–1993 (2010)
10. C. Fan, F. Xiao, Z. Li, J. Wang, Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. Appl. Energy **211**, 1123–1135 (2018)
11. L. Breiman, Random forests. Mach. Learn. **45**, 5–32 (2001)
12. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. Nature **521**, 436–444 (2015)
13. A. Vaswani et al., Attention is all you need, in Advances in NeurIPS, vol. 30 (2017)
14. J.H. Friedman, Greedy function approximation: A gradient boosting machine. Ann. Stat. **29**, 1189–1232 (2001)
15. R.D. Banker, A. Charnes, W.W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manage. Sci. **30**, 1078–1092 (1984)

16. R. Färe, S. Grosskopf, C.A.K. Lovell, Production Frontiers (Cambridge University Press, Cambridge, 1994)
17. B.W. Ang, Decomposition analysis for policymaking in energy: Which is the preferred method? Energy Policy **32**, 1131–1139 (2004)
18. C. Dinca, M.G. Badea, A. Panagopoulos, Efficiency of solar photovoltaic power plants: A DEA-based analysis. Renew. Energy **135**, 436–444 (2019)
19. B. Johansson, R. Sadollah, Explainable machine learning for renewable energy prediction: SHAP-based analysis. Energy Rep. **8**, 1247–1259 (2022)
20. A.K. Morales-Hernández et al., Explainable AI for transformer health monitoring using SHAP values. IEEE Trans. Ind. Inform. **19**, 5124–5133 (2023)