Reducing Feature and Temporal Complexity in High-Frequency NILM for Urban IoT Energy Monitoring

Gabriel Arquelau Pimenta Rodrigues[1], **André Luiz Marques Serrano**[1], **Geraldo Pereira Rocha Filho**[1,2], **Rodolfo Ipolito Meneguette**[3], **Vinícius Pereira Gonçalves**[1]

[1]University of Brasilia, Brasília, DF, Brazil
[2] State University of Southwest Bahia, Vitória da Conquista, BA, Brazil
[3] University of São Paulo, São Carlos, SP, Brazil

gabriel.arquelau@redes.unb.br, {andrelms,vpgvinicius}@unb.br, geraldo.rocha@uesb.edu.br, meneguette@icmc.usp.br

Abstract. *High-frequency non-intrusive load monitoring (NILM) has achieved improved estimation accuracy with the use of harmonic electrical features in smart homes. However, the increased feature space produces significant computational, storage, and communication overheads, especially for scalable urban sensing deployments. Hence, this work investigates the reduction of feature dimensionality and of temporal complexity in high-frequency NILM for urban IoT energy monitoring while preserving predictive performance. A feature importance analysis is conducted using a Random Forest model, resulting in a reduced subset of 19 features, comprising those that together represent 90% of the aggregated importance for all appliances. The results show that models trained with the selected features achieve performance comparable to models trained with the complete feature set, despite a 48.65% dimensionality reduction, which reduces training and inference times and supports scalable deployment in urban sensing infrastructures. Assuming 32-bit floating-point representation, this reduction decreases the input feature vector from 148 bytes to 76 bytes per sample, reducing computational demands. Additionally, a temporal relevance analysis based on mutual information indicates that instantaneous measurements provide the most predictive information. Thus, these findings suggest that accurate high-frequency NILM can be achieved using a reduced feature set and minimal temporal context, promoting efficient deployment on IoT platforms to support scalable urban energy monitoring and more sustainable energy use.*

1. Introduction

The digitalization of electrical infrastructure and the evolution of smart grids have intensified demand for mechanisms to monitor energy consumption in distributed environments [Kamenska et al. 2022], including smart homes, buildings, and urban environments. In this scenario, non-intrusive load monitoring (NILM) is an important technique for disaggregating overall electrical consumption into appliance estimates, thus providing a scalable monitoring solution that addresses this demand [Athanasiadis et al. 2021].

Load disaggregation may rely on low-frequency electrical measurements, which are data sampled at a rate slower than the AC waveform's frequency [Huber et al. 2021].

This approach, however, may not adequately characterize each signal due to insufficient information [Papageorgiou et al. 2025]. Consequently, low-frequency NILM models may perform poorly on devices with overlapping power signatures.

It has been demonstrated that the extraction of harmonic features improves the accuracy of NILM models [Li et al. 2020]. However, it also leads to significant system overhead, as processing more features increases computational complexity. This additional cost may be prohibitive for deployments with limited resources. For example, Internet of Things (IoT) sensors in smart grids may struggle to process and transmit large volumes of data in real time [Goudarzi et al. 2022].

It is therefore essential to reduce the dimensionality to retain only the most diagnostically meaningful high-frequency features. This process is necessary to balance model accuracy and system feasibility. Removing non-discriminatory information reduces the computational burden on IoT devices, conserves bandwidth for transmission, and minimizes storage requirements, which are desirable characteristics for scalable urban sensing infrastructures. These advantages can be achieved while preserving the model's ability to distinguish among complex appliance signatures. However, it is essential to identify which specific high-frequency features provide the greatest discriminative power for NILM and how much temporal context is necessary.

To address these questions, a high-frequency dataset has been published with detailed electrical measurements, including 32 harmonic components, Root Mean Square (RMS) voltage and current, power factor, and active and apparent power, measured with a distributed sensing architecture based on the ATM90E36A metering IC [Dinar et al. 2025]. This dataset allows the exploration of feature relevance in high-frequency NILM.

In this paper, we investigate this dataset to (i) identify which electrical features are most relevant for NILM; and (ii) evaluate the usefulness of temporal context, determining whether past and future observations improve regression. In addition to NILM, these questions are relevant for practical deployment in IoT and smart city environments, as well as in distributed monitoring infrastructures, where reducing feature dimensionality results in lower communication overhead, improved latency, and greater energy efficiency.

To achieve this, we analyze the importance of the features using a Random Forest NILM model, selected due to its good performance on the dataset [Dinar et al. 2025] and its support for feature ranking. Comparing models trained with complete and reduced feature sets quantifies the relationship between accuracy and computational cost. The results clarify how NILM can be efficiently integrated into distributed smart-grid platforms and sensor networks.

1.1. Contributions and limitations

This paper contributes to the field of high-frequency NILM with a feature-relevance study to quantify the discriminative power of harmonic components. Additionally, we evaluate the temporal context requirements for accurate regression. These analyses result in an analysis of the relationship between predictive performance and input dimensionality, with implications for computational and communication cost, which may be used to assist the deployment of efficient models on IoT platforms for smart cities.

As a limitation, this study uses one model and a single dataset, which is sufficient to investigate whether the reduction of the feature space is feasible in the evaluated scenario and to quantify its effect on predictive performance under a controlled protocol. However, the findings may not be generalized for all devices, households, or datasets.

1.2. Structure of the work

Section 2 reviews related literature, whilst Section 3 presents the methodology used in this work. Section 4 discusses the results obtained, and Section 5 concludes the paper.

2. Literature review

Accurate monitoring of energy consumption is fundamental to the development of smart cities and the modernization of smart grids, supporting future urban energy independence and environmental goals. Advanced technological integration helps to build resilient, efficient, and self-optimizing energy networks [AT et al. 2024]. This aligns with the Internet of Energy (IoE) concept, wherein energy systems are enhanced through pervasive sensing, connectivity, and intelligent analytics [Mishra and Singh 2023].

The introduction of MELISSA exemplifies this evolution [Rodrigues et al. 2025], as a framework that uses large language models to transform raw data from IoE devices into natural-language personalized recommendations. The work promotes intelligent human-computer interaction, which enables users to understand consumption patterns, receive cost-saving advice, and identify potential appliance anomalies through simple conversation.

Solutions such as MELISSA, however, require the deployment of an individual smart sensor on each monitored appliance in the home to generate the granular consumption data needed for the analysis. The associated costs of these devices can be substantial, particularly for larger households and neighborhoods, emphasizing the scalability challenges of per-device sensing. A more scalable approach is to use NILM [Hart 1992], which infers appliance consumption from aggregate electrical measurements acquired at a single metering point [Schirmer and Mporas 2022].

The development of NILM techniques has been supported by publicly accessible datasets that capture electrical consumption at different temporal resolutions and levels of granularity. Early datasets, such as REDD [Kolter and Johnson 2011] and UK-DALE [Kelly and Knottenbelt 2015], provide low-frequency aggregate and sub-metered measurements suitable for long-term disaggregation studies. High-frequency and event-oriented datasets, including [Filip et al. 2011], PLAID [Gao et al. 2014], and COOLL [Picon et al. 2016], provide detailed voltage and current waveforms, improving appliance discrimination. Our work uses a high-frequency dataset that includes harmonic features, suitable for advanced NILM analysis [Dinar et al. 2025].

In addition to these datasets, a high-frequency NILM framework that generates labeled aggregate traces from collections of isolated appliance signals has been proposed. It synthesizes large sets of single-appliance waveforms from a comparatively small pool of real measurements [Nour et al. 2023]. This approach enables the extraction of high-frequency signatures that are typically absent from low-frequency datasets.

A high-frequency NILM framework that balances disaggregation accuracy with the computational demands of real-time applications has been proposed

[Papageorgiou et al. 2023]. The work proposes a feature extraction method that transforms one second of aggregated current data into two-dimensional current harmonic distortion images, created by plotting the 3rd-order harmonic distortion against the 5th-order harmonic distortion. These images serve as input to a Convolutional Neural Network (CNN) for appliance classification. The method demonstrates robust performance for single and combined appliance operation.

An information-theoretic analysis of NILM, framing load disaggregation as a coding-decoding process, indicates that certain appliances, such as hair dryer and electric water heater, are more distinguishable than others from the aggregate signal, and also that some appliance signatures, such as laptop charger, may mask others [Rodrigues et al. 2026].

However, using too many features may be computationally heavy for hardware with limited resource [da Silva et al. 2020]. To address this, a lightweight NILM approach designed for edge computing deployment has been proposed [Dong et al. 2025]. The method uses a Time & Voltage Dual-Reference-Based Time-Domain Window Subtraction (TVDR-TDWS) technique, which isolates the transient current waveform of a single appliance activation from the aggregated signal. These transient signatures are then used to train a low-parameter model. In contrast, our work uses feature-importance analysis to discard non-informative features from the data, reducing model complexity. This process creates optimized training sets, which we then use to compare the accuracy of different NILM models.

The reviewed literature indicates that although high-frequency features improve NILM accuracy, methods commonly use signal-processing feature representations [Papageorgiou et al. 2023] or deep learning models with significant computational demands [Dong et al. 2025]. In contrast, fewer studies examine the relative importance of individual electrical features and harmonic components, particularly their impact on model accuracy and deployment feasibility. This limits the practical adoption of high-frequency NILM in urban IoT environments. Therefore, this work focuses on identifying which existing high-frequency features are meaningful, advancing model simplification without degrading performance.

3. Materials and Methods

This work uses a high-frequency NILM dataset that contains harmonic-level electrical features [Dinar et al. 2025]. The dataset was acquired using a measurement system based on the ATM90E36A energy metering integrated circuit (IC), which simultaneously monitors aggregate household consumption and individual appliance consumption. The measured appliances are a hair dryer, an electric water heater, a hair straightener, a fridge, an iron, a screen, a laptop charger, and a lamp. The dataset is divided into measurement sessions, collected over multiple recording periods.

This sensing configuration represents a general smart-home energy monitoring scenario that can be integrated into urban IoT infrastructures. In such environments, reducing feature dimensionality is important for model efficiency and for lowering computational costs.

Each sensing node records RMS voltage and current, active and apparent power, power factor, and the magnitudes of current harmonics from the 1st to the 32nd order, computed using the IC's onboard discrete Fourier transform engine applied to an eight kHz-sampled signal. Measurements are logged at two-second intervals.

A representation of the dataset and of the tools used in this study is given in Figure 1. Python version 3.12.13 is used to conduct the experiment. All experiments were performed on a machine equipped with an Intel Xeon CPU running at 2.20 GHz (2 vCPUs) and approximately 13 GB of RAM.

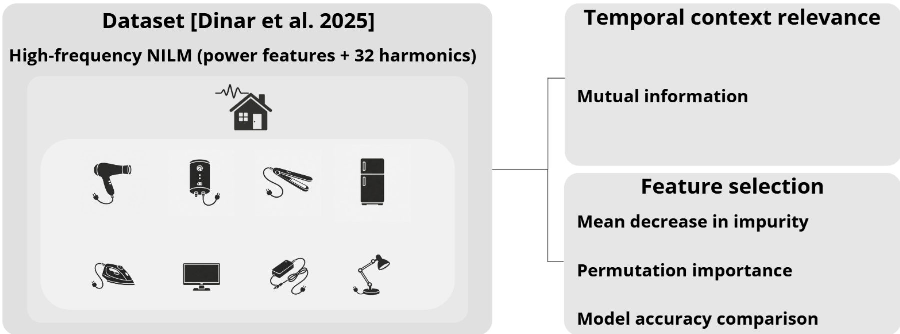


Figure 1. Overview of the dataset and analysis conducted in this work.

To assess the temporal relevance of aggregate electrical features for predicting each appliance's energy consumption, we compute mutual information (MI) over different temporal lags using the nonparametric mutual information estimator for continuous targets from `scikit-learn` (version 1.6.1). For each appliance-session pair and each lag $k \in [-10, 10]$, MI was estimated between the target appliance active power at time t and the aggregate feature values at time $t + k$. At each lag, MI was computed independently for each aggregate feature and then averaged across features to obtain a single score. Thus, lag zero corresponds to temporally aligned measurements, whereas nonzero lags quantify the predictive relevance of shifted aggregate observations. For comparability, each lagwise MI is min-max normalized.

For the NILM tasks, we use a Random Forest regressor, as it has been shown to achieve the best performance compared with alternative machine learning approaches on the dataset [Dinar et al. 2025].

To identify the most informative features and evaluate opportunities for model simplification, we compute the mean decrease in impurity (MDI) from the trained Random Forest model, which measures the average reduction in the regression impurity criterion contributed by each feature across the decision trees. Then, we compute permutation importance by randomly shuffling each feature's values in the test set and measuring the resulting reduction in regression accuracy, quantifying each feature's contribution to predictive performance independent of the model's internal structure.

The features are then ranked according to these importance metrics to identify candidates for removal. For the feature ablation experiment, one Random Forest regressor was trained independently for each target appliance using the aggregate measurements as input features and the target appliance active power as regression output. The implementation used `RandomForestRegressor` from `scikit-learn` with 100 trees, with a fixed random seed (set to 0). All other parameters were kept at their library defaults. For each appliance, three input configurations were evaluated: (i) only the selected relevant features; (ii) all aggregate features except the selected subset; and (iii) the full aggregate feature set.

The aggregate and appliance files were aligned by integer timestamp within each session, and the resulting data were partitioned by session into disjoint training and test sets. This protocol prevents temporal leakage across splits, since samples from the same session were not shared between training and evaluation. Performance is quantified with the Mean Absolute Error (MAE).

4. Results and Discussion

This section presents the experimental results obtained from the proposed analyses.

4.1. Temporal context relevance

A NILM strategy may use sequence-based models to learn temporal dependencies in electrical signals, but it requires higher computational complexity. For deployments with limited resources, it is essential to determine whether appliance energy consumption can be accurately inferred from instantaneous measurements or whether incorporating past and future context provides significant additional information.

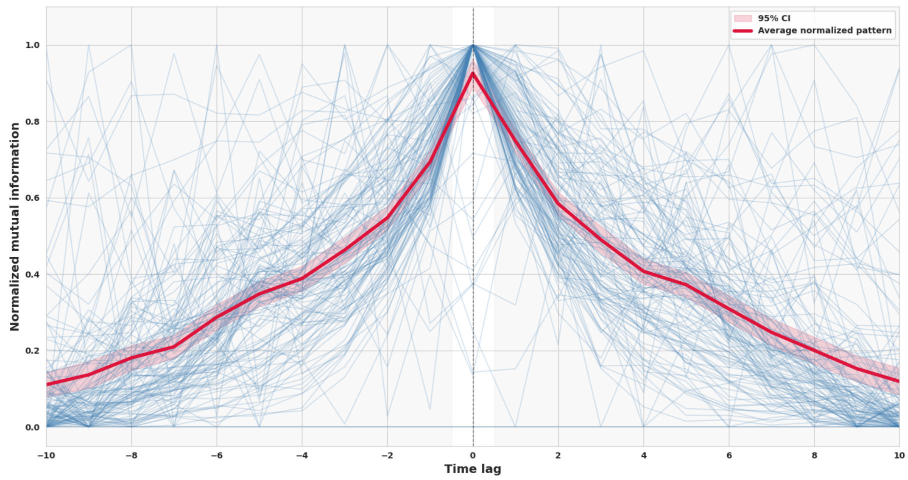


Figure 2. Normalized mutual information between features from aggregated power data and the target appliance's active power. Each line represents an appliance in a session. The shaded region represents the 95% confidence interval (CI).

To assess the relevance of the temporal context, Figure 2 shows the normalized mutual information between the aggregated signal's electrical features and the target appliance's active power as a function of temporal lag. For each temporal lag, mutual information was computed between the lagged aggregate input features and the instantaneous active power of the target appliance.

A peak is observed at zero lag, indicating that the strongest predictive relationship occurs when features and target are temporally aligned. Deviation from this alignment in either direction results in an approximately symmetric decline in MI, with an average reduction of 20.47% per one-time-step shift. This decay suggests that most of the discriminative information is concentrated in instantaneous measurements, whereas past and future observations contribute substantially less.

Despite some appliance-session pairs deviating from a peak at zero lag, the dominant pattern observed suggests that temporal dependencies appear limited in this dataset and evaluation protocol. As a result, complex sequence-based models that explicitly exploit long temporal contexts, such as recurrent neural networks or attention-based architectures, are likely to achieve progressively minor performance improvements as model complexity, computational requirements, and memory costs increase. This observation aligns with previous findings on the same dataset, which showed that convolutional models that explore temporal structure did not outperform a Random Forest model operating on independent samples [Dinar et al. 2025].

Hence, these results indicate that instantaneous high-frequency features capture most of the relevant information for appliance power regression, supporting the use of simpler models for deployment on resource-constrained IoT platforms in urban energy monitoring scenarios.

4.2. Feature selection and evaluation

The high-frequency dataset used in this study provides several electrical measurements, including harmonic components and power metrics, which can significantly improve appliance discrimination. However, this also results in feature spaces that increase computational complexity, memory usage, and data transmission requirements. In practical deployment scenarios, processing and storing large feature vectors can limit applicability, especially in distributed energy monitoring infrastructures.

4.2.1. Feature relevance analysis

Hence, to identify the most informative electrical features and reduce model complexity, we use mean decrease in impurity and permutation importance. Since these two metrics are based on different principles, we evaluate the concordance between their rankings to assess the consistency of their measures. The agreement between the rankings produced by the two methods is quantified using Spearman's rank correlation coefficient and Kendall's tau for each appliance.

Table 1. Concordance between MDI and permutation feature importance rankings.

Appliance	Spearman ρ	Kendall τ
Hair dryer	0.961	0.841
Electric water heater	0.887	0.769
Hair straightener	0.904	0.754
Fridge	0.969	0.856
Iron	0.736	0.571
Screen	0.924	0.766
Laptop charger	0.961	0.832
Lamp	0.907	0.733

The strong agreement observed between mean decrease in impurity and permutation-based feature importance, as shown in Table 1, indicates that feature relevance is a stable property of the data. This consistency suggests that the most informative electrical features are discriminative with different evaluation perspectives, reinforcing confidence in the resulting feature rankings. This agreement reduces the risk of bias introduced by the limitations of individual importance measures.

Considering this concordance, MDI and permutation importance are aggregated by averaging their normalized values. The resulting score is subsequently used for dimensionality reduction.

Figure 3 indicates the most critical features from the aggregated signal to distinguish each appliance. Features were selected to together represent at least 90% of the total importance for each appliance, consistent with prior work on importance-based feature selection [Duan 2025, Awal et al. 2021]. This feature set is unchanged for importance thresholds between 86.28% and 91.77%, indicating low sensitivity to the 90% cutoff.

This results in the selection of 19 features out of the 37 total, including the harmonics h_1, h_2, h_3, h_4, h_5, h_7, h_9, h_{11}, h_{13}, h_{15}, h_{17}, h_{19}, h_{22}, and h_{27}; along with RMS current and voltage (irms and vrms), active and apparent power (p_active and p_apparente), and power factor.

It is noteworthy that most of the selected harmonics are odd. This is because odd harmonics are a characteristic outcome of nonlinear loads commonly found in household appliances [Ewald F. Fuchs 2023]. This is corroborated by Table 2, which indicates that the odd harmonics magnitude contribution, given by the sum of the mean absolute values of the harmonic components, is significantly greater than that of even harmonics.

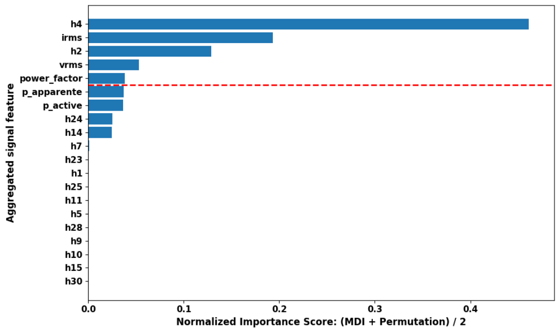
Table 2. Odd versus even harmonic magnitude contribution per appliance.

Appliance	Odd harmonics	Even harmonics	Odd ratio
Hair dryer	0.134	0.025	0.844
Electric water heater	0.163	0.000	0.997
Hair straightener	0.023	0.001	0.969
Fridge	0.602	0.026	0.958
Iron	0.043	0.000	0.995
Screen	0.093	0.008	0.924
Laptop charger	0.144	0.004	0.976
Lamp	0.056	0.001	0.986

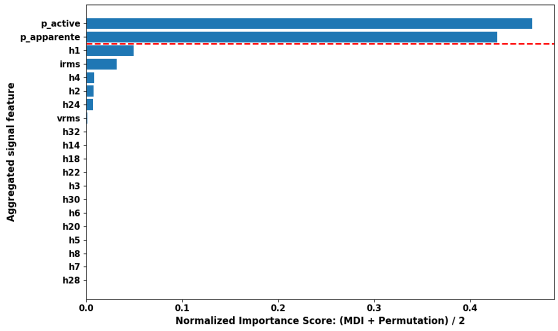
Consequently, odd harmonics have been preferentially used in NILM tasks [Kang et al. 2020, Fu et al. 2024, Mylona and Bouhouras 2025]. Even harmonics, however, may also be meaningful for these models, as depicted in Figure 3(a), which indicates a significant importance of the h_4 and h_2 harmonics for estimating the hair dryer consumption, possibly due to asymmetries in its waveform. This is also evident in Table 2, which shows a lower odd-to-even ratio for this appliance.

4.2.2. Impact of feature selection on model performance and computational cost

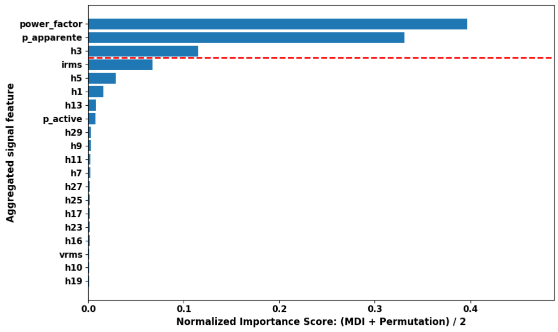
Considering the 19 features that represent 90% of the aggregated importance for all appliances, the performance of Random Forest NILM models trained with the complete feature set is compared with models trained using only the selected features and with models



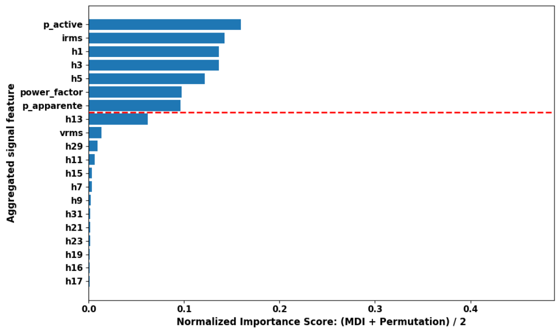
(a) Hair dryer



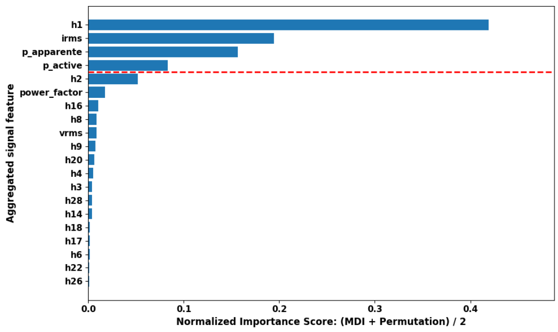
(b) Electric water heater



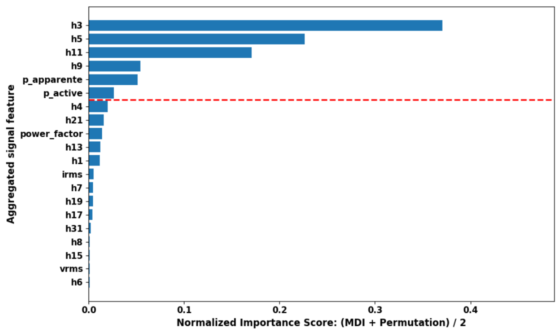
(c) Hair straightener



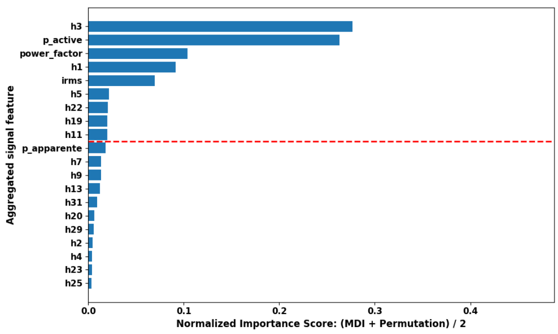
(d) Fridge



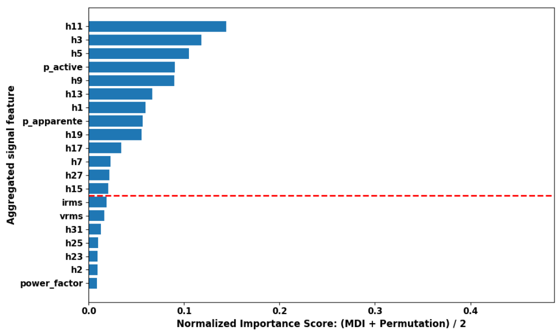
(e) Iron



(f) Screen



(g) Laptop charger



(h) Lamp

Figure 3. Feature importance per appliance for the 20 highest ranked features. The dashed horizontal line indicates the cumulative 90% importance threshold.

trained after removing these features. This comparison assesses whether the selected subset preserves the predictive information of the original data and demonstrates that, when considered in isolation, the remaining features are insufficient for accurate NILM.

The results, presented in Figure 4, show that the point-to-point Random Forest disaggregation model trained with the selected features achieves a performance comparable to that obtained using all available features. This indicates that the selected subset retains the dominant information required for accurate appliance estimation, despite a substantial reduction in dimensionality. Conversely, when the model is trained using only the features that were not selected, a significant increase in the MAE is observed across all appliances, indicating that these features, when considered in isolation, are insufficient to represent the discriminative structure of the data.

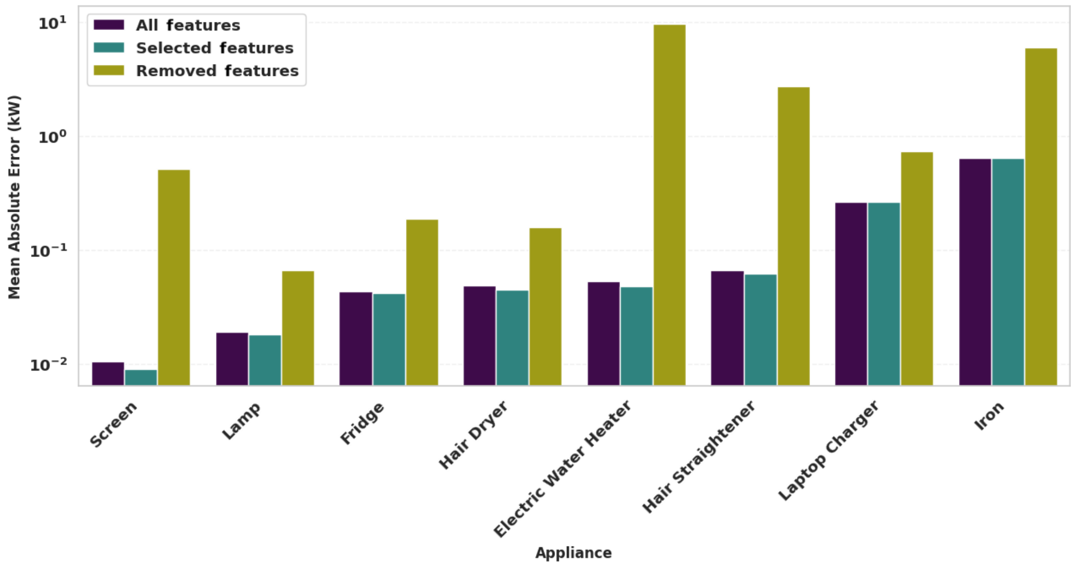
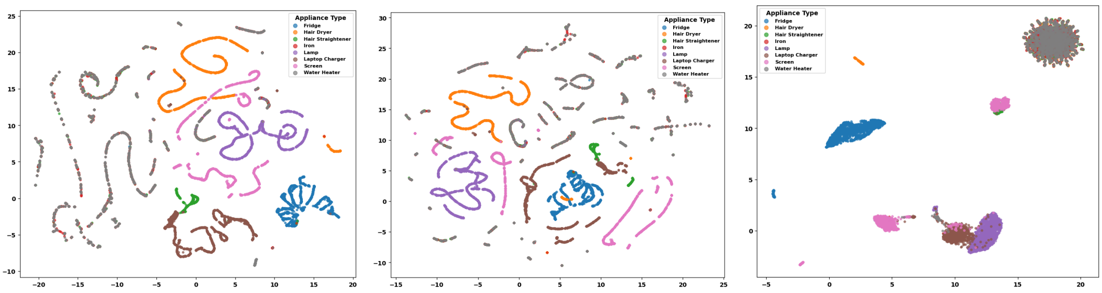


Figure 4. Mean Absolute Errors between the RF model trained with all features, the selected features, and the removed features.

Using the selected feature subset improved model performance, reducing the MAE for most appliances. The median improvement was 5.5%, with variations ranging between a reduction of 13.6% and an increase in 0.2%. This indicates that the selected features improve estimation accuracy with the removal of non-informative data. A possible explanation for this improvement is the removal of redundant and noisy features, which may reduce overfitting and improve the model's ability to generalize from the available data.

This finding is also supported by the Uniform Manifold Approximation and Projection (UMAP) plots shown in Figure 5, which provide a view of how the different feature configurations affect class separability. The UMAP is configured with `n_neighbors=30`, `min_dist=0.1`, Euclidean distance, and `random_state=42`.

When all features are used, the appliances form distinguishable classes, indicating that the high-dimensional feature space contains sufficient information to separate appliance signatures. A similar structure is preserved when only the selected features are considered. This observation confirms that the reduction process retains the essential



(a) All features **(b) Selected features** **(c) Removed features**

Figure 5. UMAP projection of features for all appliances.

structure of the data, preserving inter-class separability despite removing nearly half of the original features.

In contrast, the UMAP projection obtained using only the removed features exhibits a loss of structure, with appliances overlapping across classes. This degradation in separability suggests that the non-selected features do not encode the dominant characteristics required to distinguish appliance behavior.

Table 3. Average computational cost over appliances for the reduced and full representations.

	No. of features	**Train time (s)**	**Infer. time (s)**	**Bytes/sample (32 b)**
Selected feat.	19	158.35 ± 78.78	0.77 ± 0.65	76 ± 0
All feat.	37	292.58 ± 141.65	0.89 ± 0.46	148 ± 0

As shown in Table 3, the proposed feature selection also reduces the computational burden associated with NILM deployment. In addition to a 48.65% reduction in dimensionality, with 18 of the 37 features removed, the reduced representation lowers the average training time and the average inference time across appliances. This reduction also proportionally decreases the memory required to store feature vectors and the volume of data to be transmitted in distributed sensing scenarios.

For example, assuming 32-bit floating-point representation, the input vector size decreases from 148 bytes to 76 bytes per sample, proportionally reducing storage and communication demands. These gains support a simpler and more efficient deployment, with lower storage requirements, reduced transmission costs, and decreased processing demands, which are desirable for scalable urban sensing and IoT energy monitoring systems.

5. Conclusions and Future Works

The proposed strategy resulted in the selection of 19 features from 37, reducing dimensionality and bytes per sample by 48.65% while preserving predictive performance. Experimental results showed that models trained with the selected features achieved higher accuracy than those trained with the complete feature set. In contrast, models trained using only the non-selected features exhibited a significant performance degradation. Complementary UMAP projections also indicated that the selected features preserve the structure and separability of appliance signatures, whilst the removed features do not maintain meaningful class organization.

The analysis also revealed that odd-order harmonics dominate the selected feature set, consistent with the physical behavior of nonlinear household appliances and corroborated by the observed predominance of odd-harmonic contributions across appliances. Nevertheless, the results indicate that certain even harmonics may still provide relevant information for specific devices, reinforcing the importance of feature selection rather than relying solely on generalized assumptions.

Additionally, the temporal context analysis showed that instantaneous high-frequency measurements provide most of the predictive information required for appliance estimation. This finding supports the use of memory-efficient NILM models and strengthens the suitability of the proposed approach for edge and IoT deployments.

These findings also indicate that high-frequency NILM can be simplified without significant loss of accuracy, supporting its practical adoption in urban IoT energy monitoring applications and smart-city sensing infrastructures.

Future work may explore extending this analysis to event-based NILM scenarios to detect the precise turn-on and turn-off times of individual devices within aggregated power data. Additionally, a comparison of learning models, including classical and deep learning approaches, could be conducted to assess their relative performance and generalization across NILM datasets. Additionally, it is proposed to investigate privacy-preserving learning paradigms, such as federated learning and differential privacy, to support collaborative NILM model training over distributed households and to prevent the disclosure of sensitive consumption data. Future work may also investigate the integration of the proposed reduced NILM approach into urban IoT energy monitoring scenarios involving multiple distributed households.

References

AT, M. R., B, B., RR, S. A. P., Naidu, R. C., M, R. K., Ramachandran, P., Rajkumar, S., Kumar, V. N., Aggarwal, G., and Siddiqui, A. M. (2024). Intelligent energy management across smart grids deploying 6G IoT, AI, and blockchain in sustainable smart cities. *IoT*, 5(3):560–591.

Athanasiadis, C. L., Papadopoulos, T. A., and Doukas, D. I. (2021). Real-time non-intrusive load monitoring: A light-weight and scalable approach. *Energy and Buildings*, 253:111523.

Awal, M. A., Hossain, M. S., Debjit, K., Ahmed, N., Nath, R. D., Habib, G. M., Khan, M. S., Islam, M. A., and Mahmud, M. P. (2021). An early detection of asthma using BOMLA detector. *IEEE Access*, 9:58403–58420.

da Silva, M. V., Silva, G. G., and Gomes, R. L. (2020). Minimização do consumo de energia de dispositivos sem fio em ambientes industriais. In *Workshop de Computação Urbana (CoUrb)*, pages 43–56. SBC.

Dinar, F., Paris, S., and Busvelle, E. (2025). Capturing high-frequency harmonic signatures for NILM: Building a dataset for load disaggregation. *Sensors*, 25(15).

Dong, Z., Zhang, X., Cai, S., Yang, Y., Jiang, W., Song, Y., Zhao, W., and Zhang, D. (2025). Real-time NILM: A lightweight and low power approach. *Energy*, page 138100.

Duan, J. (2025). Dynamic feature selection for silicon content prediction in blast furnace using BOSVRRFE. *Scientific Reports*, 15(1):20555.

Ewald F. Fuchs, M. A. S. M. (2023). *Power Quality in Power Systems, Electrical Machines, and Power-Electronic Drives*. Academic Press, 3 edition.

Filip, A. et al. (2011). BLUED: A fully labeled public dataset for event-based nonintrusive load monitoring research. In *2nd workshop on data mining applications in sustainability (SustKDD)*, volume 2012, page 5.

Fu, Y., Chen, S., Zhao, F., Shi, Y., Ye, S., Jiang, J., Wang, L., and Yang, X. (2024). Identification of electrical equipment based on VI trajectory, odd harmonic currents. In *2024 8th International Conference on Smart Grid and Smart Cities (ICSGSC)*, pages 164–169. IEEE.

Gao, J., Giri, S., Kara, E. C., and Bergés, M. (2014). Plaid: a public dataset of high-resoultion electrical appliance measurements for load identification research: demo abstract. In *proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, pages 198–199.

Goudarzi, A., Ghayoor, F., Waseem, M., Fahad, S., and Traore, I. (2022). A survey on IoT-enabled smart grids: emerging, applications, challenges, and outlook. *Energies*, 15(19):6984.

Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891.

Huber, P., Calatroni, A., Rumsch, A., and Paice, A. (2021). Review on deep neural networks applied to low-frequency NILM. *Energies*, 14(9):2390.

Kamenska, L., Moiseenko, V., Shendryk, V., Kamenskyi, S., and Shendryk, S. (2022). Implementation of distributed information systems in solving problems of energy consumption monitoring. In *International Conference "New Technologies, Development and Applications"*, pages 758–767. Springer.

Kang, H., Kim, H., et al. (2020). Household appliance classification using lower odd-numbered harmonics and the bagging decision tree. *IEEE Access*, 8:55937–55952.

Kelly, J. and Knottenbelt, W. (2015). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes. *Scientific data*, 2(1):1–14.

Kolter, J. Z. and Johnson, M. J. (2011). REDD: A public data set for energy disaggregation research. In *Workshop on data mining applications in sustainability (SIGKDD), San Diego, CA*, volume 25, pages 59–62.

Li, K., Yin, B., Du, Z., and Sun, Y. (2020). A nonintrusive load identification model based on time-frequency features fusion. *IEEE Access*, 9:1376–1387.

Mishra, P. and Singh, G. (2023). Energy management systems in sustainable smart cities based on the internet of energy: A technical review. *Energies*, 16(19):6903.

Mylona, D. N. and Bouhouras, A. S. (2025). A digital twin-based framework for load identification using odd harmonic current plots. *Applied Intelligence*, 55(7):1–15.

Nour, M., Le Bunetel, J.-C., Ravier, P., and Raingeaud, Y. (2023). Data augmentation strategies for high-frequency NILM datasets. *IEEE Transactions on Instrumentation and Measurement*, 72:1–9.

Papageorgiou, P., Mylona, D., Stergiou, K., and Bouhouras, A. S. (2023). A time-driven deep learning NILM framework based on novel current harmonic distortion images. *Sustainability*, 15(17):12957.

Papageorgiou, P. G., Christoforidis, G. C., and Bouhouras, A. S. (2025). NILM in high frequency domain: A critical review on recent trends and practical challenges. *Renewable and Sustainable Energy Reviews*, 213:115497.

Picon, T., Meziane, M. N., Ravier, P., Lamarque, G., Novello, C., Bunetel, J.-C. L., and Raingeaud, Y. (2016). COOLL: Controlled on/off loads library, a public dataset of high-sampled electrical signals for appliance identification. *arXiv preprint arXiv:1611.05803*.

Rodrigues, G. A. P., de Oliveira, M. N., Serrano, A. L. M., Rocha Filho, G. P., Vergara, G. F., Mosquéra, L. R., and Gonçalves, V. P. (2025). MELISSA: An LLM-powered smart home energy consumption monitoring framework. In *Simpósio Brasileiro de Computação Ubíqua e Pervasiva (SBCUP)*, pages 11–20. SBC.

Rodrigues, G. A. P., Serrano, A. L. M., Filho, G. P. R., Gonçalves, V. P., and Meneguette, R. I. (2026). An information-theoretic analysis of high-frequency load disaggregation. *Entropy*, 28(3):334.

Schirmer, P. A. and Mporas, I. (2022). Non-intrusive load monitoring: A review. *IEEE Transactions on Smart Grid*, 14(1):769–784.