

ORIGINAL RESEARCH

Open access

A Methodological Framework for Breaking Down and Predicting Demand Flexibility in Large-Scale Consumers Using Non-Intrusive Load Monitoring Techniques

Nikolai Ivanov^{1*}, Sergey Volkov¹, Elena Morozova²

Abstract

Advancements in technology, continuous innovation, and the emergence of the Industry 4.0 paradigm are driving Distribution System Operators toward increasingly competitive environments where flexible demand response mechanisms are essential. In many developing nations, however, the absence of standardized measurement frameworks across industrial process chains hinders the effective deployment of demand-side management models, leading to inefficiencies in data analysis and limiting the accurate assessment of large-scale consumers' contributions to flexibility. To address this gap, the present study utilizes load profile data—specifically energy and power measurements obtained from utility meters—as input variables to generate disaggregated forecasts at 15-minute intervals across four defined time segments. The proposed approach is validated through performance metrics, with outcomes assessed using the Root Mean Square Error (RMSE) to quantify total methodological error. The framework integrates Machine Learning and Big Data techniques implemented in Python, employing Combinatorial Disaggregation Optimization alongside Factorial Hidden Markov Models to enhance prediction accuracy.

Keywords Machine learning, Non-intrusive load monitoring, Big data, Combinatorial optimization, Factorial hidden markov model, Time-of-use tariffs

*Correspondence:

Nikolai Ivanov
nikolai.ivanov@nsu.ru

¹ Department of Smart Grid Control Systems, Faculty of Engineering, Novosibirsk State University, Novosibirsk, Russia

² Department of Energy Analytics, Faculty of Engineering, Tomsk State University, Tomsk, Russia

Introduction

Limited understanding of consumption flexibility within electricity demand management initiatives, combined with insufficient standardization and data acquisition techniques for industrial energy use, constrains both operational efficiency and economic performance, ultimately reducing industries' capacity to optimize resource utilization [1, 2]. The industrial sector accounts for a major share of total electricity demand in most countries; however, in developing economies, the adoption of advanced monitoring and innovation systems to track end-use energy consumption remains limited, negatively affecting global competitiveness [3, 4].

Implementing modern energy management strategies and technologies can significantly enhance productivity and competitive positioning, particularly in emerging markets. Increasingly, data-driven methodologies—especially those leveraging Machine Learning (ML)—are being adopted in research to process extensive datasets characteristic of Big Data (BD) and to address complex nonlinear challenges [5]. Segmenting and analyzing end-use energy consumption in industrial contexts enables a deeper understanding of inefficiencies and supports the evaluation of energy-saving

measures before and after their implementation. Proper identification of classification features is therefore critical when examining energy usage across production systems [6].

Non-Intrusive Load Monitoring (NILM), also known as energy disaggregation, aims to determine the distribution of energy consumption across end uses. This approach opens new avenues for intelligent energy management by enabling behavioral changes in consumption patterns, while also assisting distribution utilities in forecasting aggregated demand and in designing pricing schemes, such as Time-of-Use (TOU) tariffs [7, 8].

Establishing standardized methods for collecting and managing industrial electricity consumption data is crucial. A lack of visibility into energy consumption across production stages leads to inefficiencies in both operational and financial management, thereby limiting the potential for optimization [9]. Continuous monitoring of electricity demand plays a key role in production planning, and disaggregating energy use at the process level provides actionable insights for more effective energy management [10, 11].

Industry 4.0 represents a global shift toward modernization, driven by advancements in information and communication technologies. This transformation involves integrating new communication protocols, enhanced cybersecurity measures, interconnected devices, and flexible user interfaces [12].

Within this framework, the Internet of Things (IoT) and Cyber-Physical Systems (CPS) are central to enabling predictive and intelligent production environments. By applying mathematical models to process-monitoring data, these systems identify behavioral patterns that can be utilized for forecasting, classification, and anomaly detection, which are core challenges of Industry 4.0 [12, 13]. In developed economies, industrial facilities are typically highly automated, incorporating advanced machinery and robotic systems. Data exchange is facilitated through cloud-based infrastructures, where each component performs specialized functions [3, 4].

Conversely, many Latin American countries face limitations in advancing Research and Development (R&D) due to structural characteristics of their production sectors and delayed innovation policies. Although awareness of Industry 4.0 challenges has increased in recent years, particularly regarding digital transformation and infrastructure optimization, significant gaps remain in technological adoption and human capital development [3, 14]. Differences across countries are largely influenced by the scale and effectiveness of governmental intervention, as well as institutional capacity to implement policies that support sustainable industrial value chains. In many emerging economies, clear strategic direction and policy consistency are still lacking [14]. Financial constraints further restrict organizations' ability to invest in advanced equipment, infrastructure, and energy-efficient innovations, posing a major barrier to Industry 4.0 adoption [14, 15].

Globally, both governments and industries are placing growing emphasis on energy efficiency and the development of forward-looking indicators to guide policy-making in energy-intensive sectors [6, 16].

In developed countries, industries often rely on Supervisory Control and Data Acquisition (SCADA) systems to enhance energy awareness and operational efficiency. These systems enable real-time monitoring and control of industrial processes by integrating sensors, communication networks, and software interfaces that present data in user-friendly dashboards [17, 18].

Accurate measurement of electrical parameters is fundamental for effective load management and fault detection. Real-time communication systems facilitate the monitoring and control of electrical loads while integrating energy and production data at the operational level, providing detailed insights that support informed decision-making in energy management [17, 19].

Earlier research has attempted to categorize industrial processes by energy end use, often using the concept of unit processes. These are typically divided into production and support functions. While this classification has been widely

adopted, it does not fully capture the diversity and complexity of industrial operations, as many processes do not fit neatly into this framework and lack hierarchical differentiation based on energy-use complexity [19].

Therefore, it is necessary to establish a more detailed stratification of industrial processes by identifying relevant classification criteria, including supply characteristics and production methodologies that influence energy consumption [19].

In Aslam et al. [19], a hierarchical framework for categorizing energy end-use processes is proposed, distinguishing different levels based on complexity, technological characteristics, and operational methods.

In Gan et al. [20], an IoT-based monitoring architecture is introduced, consisting of three layers: a data-acquisition layer for measuring energy-related variables, a transmission layer for sending data to cloud-based servers, and an application layer for real-time visualization and data analysis to support energy management.

In Czosnyka et al. [21], a structured methodology is presented to reduce energy consumption without compromising production efficiency. This approach involves three stages: first, deploying advanced energy management controllers that eliminate the need for additional measurement devices; second, equipping production lines with supplementary systems for measurement, communication, control, and energy recovery; and third, optimizing financial returns through improved operational efficiency. This work emphasizes the role of technological upgrades in enhancing manufacturing performance. Table 1 summarizes the key contributions in this research area.

Table 1. NILM state-of-the-art operational research summary.

Chapter	Topics	Subtopic	Bibliography
Introduction			[1-21]
Disaggregation of the Load Profile in the Industry and Application of Real Data			[22]
	Supervised Learning		[23-25]
	Unsupervised and semi-supervised learning		[23, 25, 26]
	Machine Learning algorithms for disaggregation		[6, 19]
		Combinatorial optimization (CO)	[23]
		Hidden Markov Models (HMM)	[21, 23, 27, 28]
		Factorial Hidden Markov models (FHMM)	[23, 27, 28]
	Evaluation Metrics		[8, 25, 29]
Methodology for acquisition, structure, and analysis of information	NILMTK: Open Source Toolkit for Non-Intrusive Load Monitoring		[8, 30, 31]
	Information Acquisition	Information Structure	[24, 32]

The purpose of this research is to develop a structured approach for evaluating the flexibility of energy and power demand that large-scale consumers can provide to Distribution System Operators. This is achieved by combining Non-Intrusive Load Monitoring (NILM) techniques with Machine Learning (ML) methods to decompose overall energy consumption into its constituent components. The proposed framework offers a practical solution for industries that lack advanced, process-level metering systems, enabling them to extract detailed information from aggregated load profiles, analyze internal consumption patterns, and ultimately contribute controllable energy blocks to the grid.

The principal contributions of this study are outlined below:

- To design a methodological framework capable of disaggregating electricity demand using data obtained from utility meters, specifically through 15-minute interval load profiles of large or significant consumers, thereby enabling correlation of demand response without disrupting production activities.
- To determine the extent of flexibility in demand-side management that industrial consumers can offer by analyzing disaggregated load data and evaluating the impact of adjustments in production processes.
- To assess the proposed methodology through advanced data analytics and ML techniques, with performance evaluated using error-based metrics.

Disaggregation of the Load Profile in Large Consumers-Application to Industry

In industrial environments, load profile disaggregation relies on Machine Learning algorithms to estimate the distribution of energy consumption according to end-use processes. The NILMTK framework incorporates multiple disaggregation approaches, including classification methods based on trained models. These approaches can be categorized into three primary learning paradigms: supervised, unsupervised, and semi-supervised learning.

In Buzachis *et al.* [22], an approach based on the Fast Fourier Transform (FFT) is applied within a residential context to analyze current waveforms. This method focuses on identifying and characterizing electrical loads and devices by analyzing spectral variations, aiming to detect anomalies and prevent operational faults in household electrical systems.

Supervised learning

Supervised learning techniques are trained using a combination of aggregated load measurements at the point of delivery and detailed consumption data corresponding to individual industrial processes. These methods rely on labelled datasets during the training phase [23]. Two main strategies are typically employed: optimization-based approaches and pattern recognition techniques. Examples include clustering methods and Bayesian frameworks used to infer the probable operational states of processes [24, 25].

Unsupervised and semi-supervised learning

Unsupervised learning models are developed using only aggregated load profile data from the delivery point, without requiring labelled information during training [23, 25, 26]. In contrast, semi-supervised learning integrates both labelled and unlabeled datasets. Within NILM applications, semi-supervised approaches are the most widely adopted, as they tend to deliver superior performance. In conventional ML settings, training and testing datasets are usually drawn from the same labelled source. However, in NILM applications, training may use either labelled or unlabeled data, whereas testing is performed strictly on labelled datasets derived from the activity-specific dataset [23].

Machine learning algorithms for disaggregation

Machine Learning algorithms enable the decomposition of aggregated load profiles into individual end-use consumption patterns at the delivery point. These methods offer several key advantages [6]:

- They significantly reduce the cost associated with deploying large-scale metering infrastructure [19].
- They allow for the utilization of existing electrical infrastructure without requiring extensive modifications [19].
- They facilitate the implementation of demand-side management strategies and operational control mechanisms [19].

Combinatorial optimization (CO)

Combinatorial Optimization (CO), originally introduced by George Hart, is a disaggregation algorithm that minimizes the discrepancy between the estimated sum of individual process loads and the total consumption measured at the main electricity meter. The method identifies the optimal combination of process states by aggregating their respective energy contributions and comparing the result with the observed total demand.

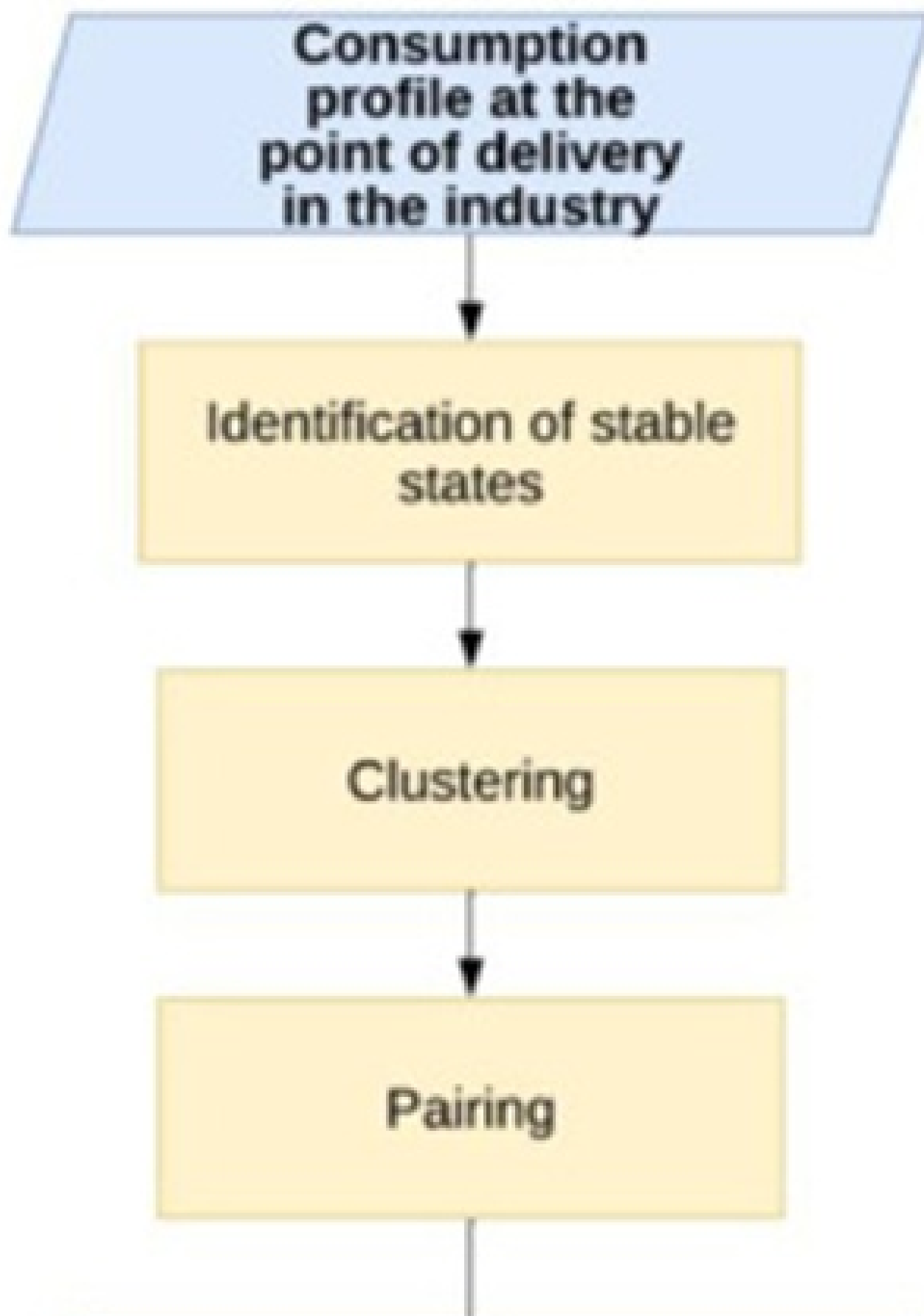
The formulation of this problem, as expressed in Eq. 1, considers the load profile at the delivery point as a time series $Y = \{y_1, y_2, \dots, y_k\}$, where k represents the number of samples. Each process is described by a binary state vector $A = \{a_1, a_2, \dots, a_n\}$, corresponding to n processes [23].

$$y_t = \sum_{i=1}^n a_i p_i + e_t \quad (1)$$

At each time step t , the solution involves identifying the state a_i that minimizes the error e_t , which accounts for measurement noise.

CO has the following stages

The CO algorithm proceeds through several steps. First, it identifies steady-state conditions by detecting variations in the industrial load profile. Subsequently, clustering techniques are applied to locate transitions in the space defined by active power and time. Clusters exhibiting similar magnitudes but opposite signs are then paired, while unmatched clusters are associated with existing or newly formed groups using a maximum likelihood approach. Finally, the detected events—represented as disaggregated load curves—are mapped to the corresponding process labels in the validation dataset, as illustrated in Figure 1.



Disaggregated curves

This Markov process can be visualized as a directed graph, where each node represents a possible state $S = \{1, 2, \dots, n\}$, and the edges indicate transitions between these states, as illustrated in [Figure 2](#).

Within this framework, a random variable X is defined over the set of states SSS , with X_t denoting the system state at time t . Since X_t is not directly observable, an output variable Y is instead measured, where Y_t represents the observed value at time t [\[27\]](#).

Typically, the relationship between these variables is expressed as $Y_t = f(X_t)$, where f may represent a probability distribution, a constant mapping, or a more complex functional relationship. [Figure 2](#) illustrates the temporal evolution of this Markov process: the upper sequence corresponds to the hidden system states over time, while the lower sequence reflects the observable outputs associated with each state [\[30\]](#).

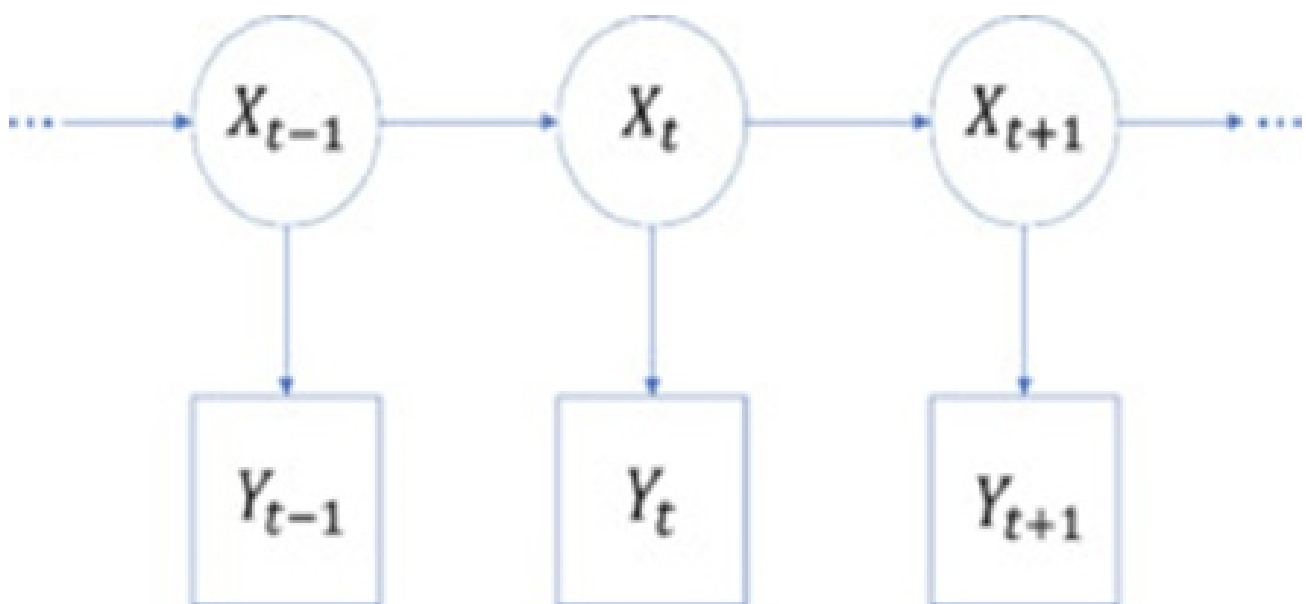


Figure 2. Representation of the evolution of a hidden Markov model.

In [Figure 2](#), the states X_{t-1} , X_t , and X_{t+1} are produced by a Markov chain. Observers can only see the corresponding observation sequences Y_{t-1} , Y_t , $Y_{t+1}, \dots$. It should be noted that the observations also obey the Markov property with respect to the underlying hidden states.

Factorial Hidden Markov Models (FHMM)

A generalization of the standard HMM is the Factorial Hidden Markov Model (FHMM). This model employs multiple variables to describe the system state. Specifically, an FHMM consists of k separate state variables, $X(0)$, $X(1)$, $\dots$, $X(k)$, corresponding to k independent HMMs operating in parallel over time.

The main goal of FHMMs is to analyze individual load patterns in industrial settings. The approach relies on several Markov chains running concurrently, with the overall observation defined as the sum of their individual outputs. In this setup, every chain models a distinct process, and the combined output reflects the total consumption recorded at the delivery point [23, 28].

An FHMM is essentially a collection of k HMMs that progress simultaneously and independently across time steps. For each individual HMM i , the state is denoted by $X(i)$. Consequently, at every time t , there exists a full sequence of states for that chain: $X_0(i)$, $X_1(i)$, ..., $X_t(i)$. Each of these component HMMs can have its own number of states and its own emission distributions, indicated as $f_{ij}(i)$ for the emission function of the j -th state in the i -th HMM [27].

Figure 3 represents the FHMM equivalent of **Figure 2**, where the top three rows illustrate three independent Markov chains whose states evolve separately. These states collectively determine the final observation of the FHMM, shown in the bottom row as Y_t [27].

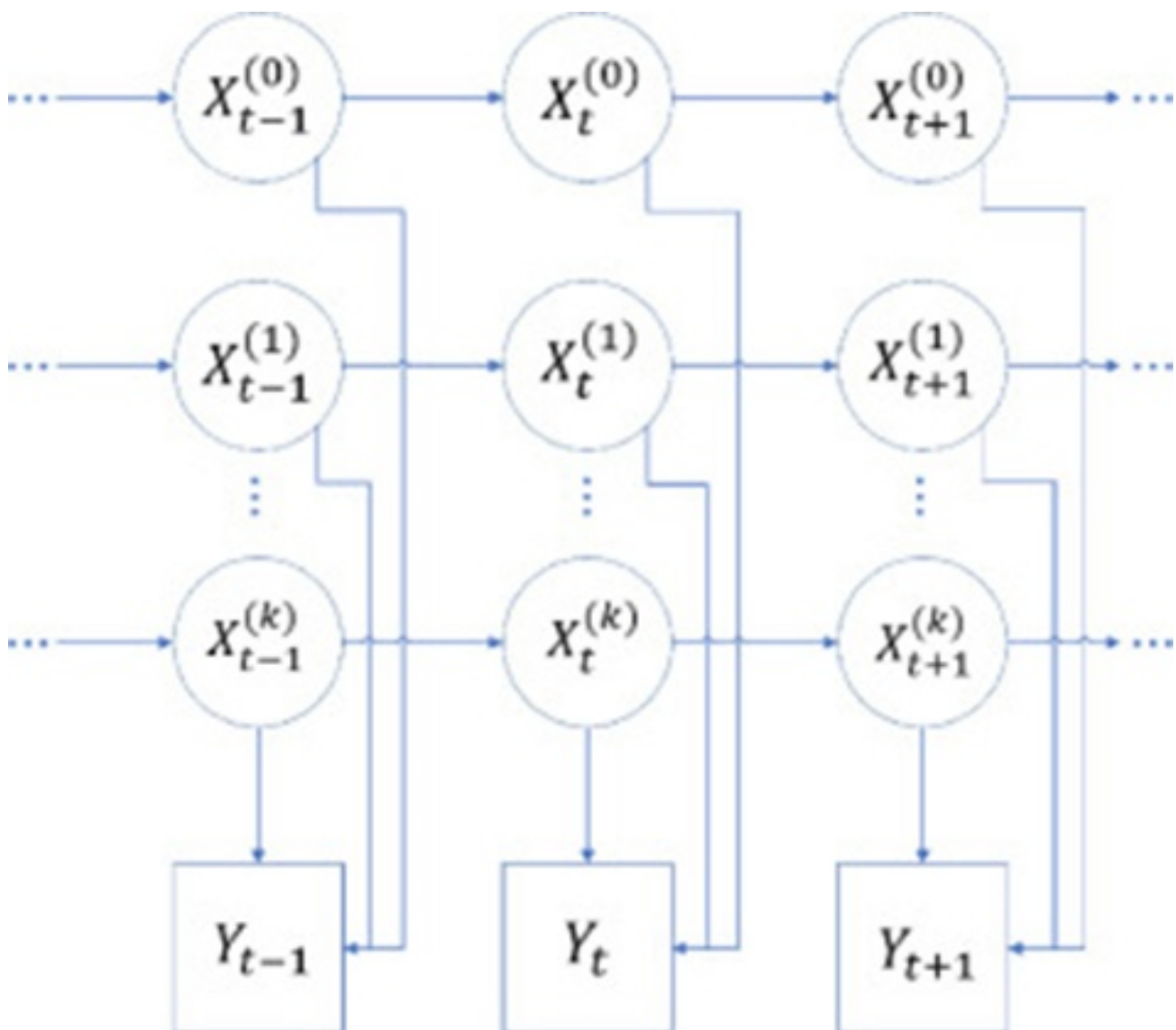


Figure 3. Representation of an FHMM.

Evaluation metrics

The progress in algorithm development and the rapid growth of NILM have created a strong need for reliable ways to assess model performance. Evaluation metrics compare the disaggregated outputs (predictions) produced by trained models with ground-truth data in the validation set, which contains precise measurements of individual processes. NILMTK supports this assessment by offering built-in metric calculations through the Meter Group object when testing results on the validation dataset. The key metrics considered are the following:

Total assigned energy error (EAE)

This metric quantifies the difference between the total energy predicted by the model and the actual total energy used by the n industrial processes, expressed in kWh [8, 29].

Normalised error in the assigned power (MNEAP)

It measures the deviation between the estimated and true power for each process at every time step. The accumulated power deviations are then divided by the process's total energy consumption to obtain a normalised value [8].

RMS error in rated power (RMSE)

This is the root-mean-square error calculated from the differences between the predicted and actual power values for each industrial process [8, 25, 29].

F-Score (F1)

The F-Score evaluates how accurate and dependable the trained model is by computing the harmonic mean of Precision and Recall [8].

Methodology For Acquisition, Structure, and Analysis of Information

This section explains the functionalities available in NILM-TK for organizing and processing the dataset. It presents a systematic procedure for gathering data from multiple industries and large-scale consumers. The approach includes categorizing production and auxiliary processes based on an initial evaluation of their demand-response flexibility, which serves as the basis for further data examination.

Utilities are required to provide electrical power continuously, reliably, and securely through the distribution network operator. In vertically integrated markets, energy sales are managed by the DSO. These operations rely on digital platforms that gather technical and commercial data. Using disaggregation algorithms, such systems estimate the level of flexibility large consumers can provide to the DSO. The overall methodology used for building and consolidating the database in this research is presented in [Figure 4](#).

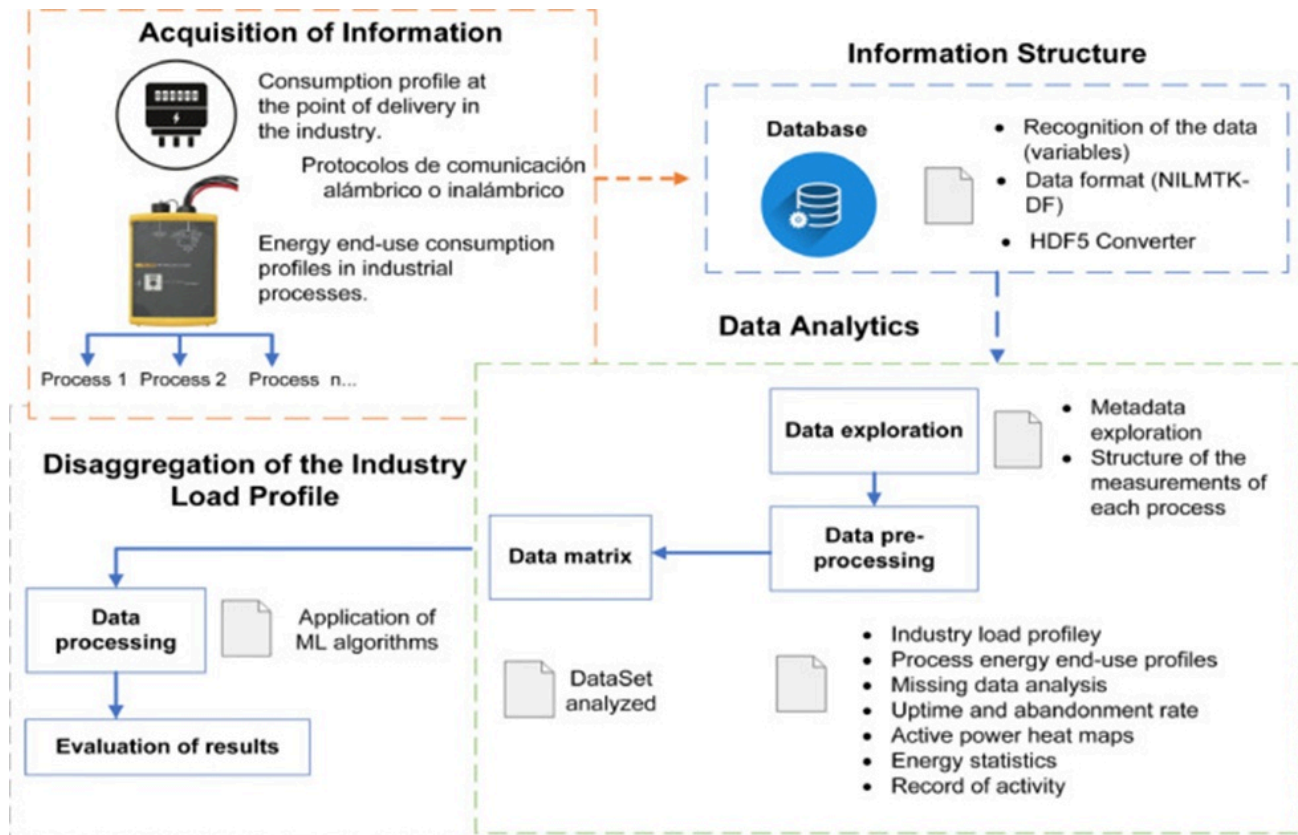


Figure 4. The methodology proposed for acquiring and structuring information.

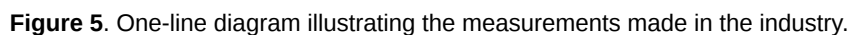
Figure 4 presents the designed approach used to collect raw data and systematically arrange it into a structured format suitable for analysis.

NILMTK: open-source toolkit for non-intrusive load monitoring

Non-intrusive load monitoring (NILM), also described as energy disaggregation, seeks to separate total electrical demand into the individual contributions of appliances or electrical circuits. The NILM Toolkit (NILMTK), an open-source Python-based framework, is specifically designed to support the evaluation and comparison of machine learning techniques for this task. It delivers a comprehensive workflow that spans from data processing and analysis to the application of performance metrics. To ensure accurate model training, the dataset must first be preprocessed to remove corrupted, missing, or implausible values, thereby ensuring data integrity throughout the learning phase [8, 30, 31].

Information acquisition

For this investigation, data are sourced from large-scale energy consumers using a combination of acquisition strategies [24]. These include initial evaluations to estimate demand response flexibility, as well as the integration of advanced metering systems positioned at critical points across industrial environments, as depicted in **Figure 5**.



Through an initial evaluation focused on demand response flexibility, the study characterized how load profiles vary across different industrial sectors. This assessment incorporated factors such as production levels and seasonal consumption patterns observed across industries, including ceramics, plastics, and food processing. These variables enabled a detailed breakdown of how energy is ultimately utilized within companies, distinguishing between consumption assigned to core production activities and that used for auxiliary operations, as presented in [Figure 6](#).



PRE-EVALUACIÓN DE LA FLEXIBILIDAD DE PROVEEDORES DE RESPUESTA DE LA DEMANDA
HERRAMIENTA DE VISUALIZACIÓN





DRIP

Empresas Rival CIA. LTDA.
Sector el Tablón - Código Postal Cuenca (Azuay) - España
Ricourte

A.I. GENERAL

A.III. USO DE LA ENERGÍA Y PROCESOS

A.III.1. ¿Qué tecnologías transversales tiene en su planta, qué proporción de la energía anual consumen y cuáles son sus fuentes de suministro energético?

	Proporción del consumo anual, %		Uso de la energía	
	Electricidad	Otras fuentes	Horas al día	Días por semana
☐ Producción y distribución de vapor	0 %	0 %	0	0
☒ Agua caliente para procesos	40 %	0 %	24	5

A.III.4. ¿Cuáles son sus principales procesos productivos, cuál es su proporción en el consumo eléctrico anual y qué uso hace de la energía?

	Descripción	Proporción del consumo anual	Turnos/día	Horas/turno	Días/semana
☒ Proceso 1	Materia Prima	10 %	2	8	5
☒ Proceso 2	Mezcladores	10 %	2	8	5
☒ Proceso 3	Producción	80 %	3	8	5
☐ Proceso 4		0 %	0	0	0
☐ Proceso 5		0 %	0	0	0
☐ Proceso 6		0 %	0	0	0
☐ Proceso 7		0 %	0	0	0
☐ Proceso 8		0 %	0	0	0
Tecnologías transversales		100%			
Otros procesos		-100%			
CONSUMO ELÉCTRICO ANUAL		100%			

Others must be positive

Figure 6. Initial evaluation of demand response flexibility.

Figure 6 presents the findings from the preliminary evaluation, which examined the detailed load behavior of sectors including ceramics, plastics, and food processing, with a visual summary provided in Figure 7.

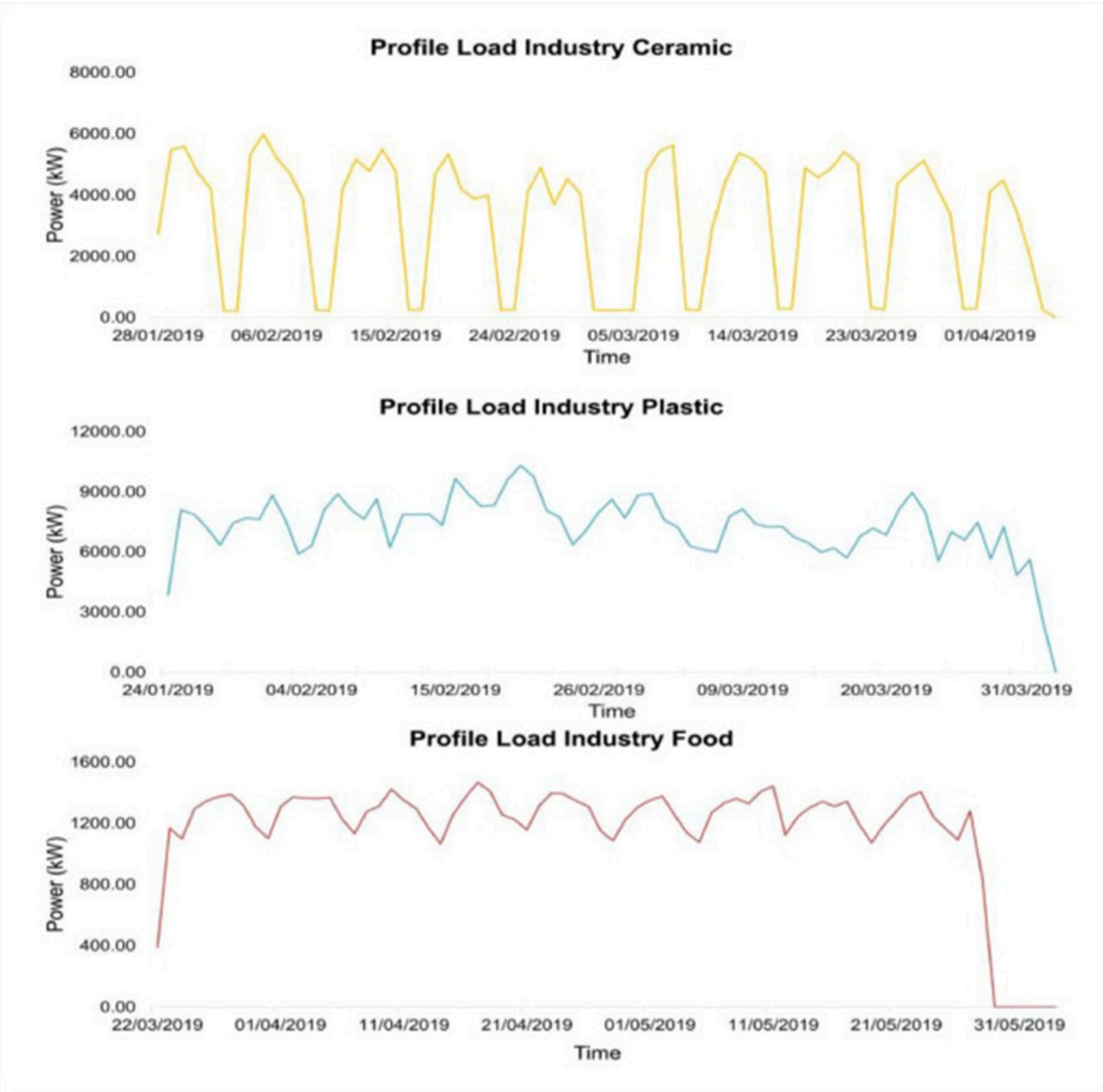


Figure 7. Load trends of various industrial sectors.

The electricity consumption patterns for different industries were analyzed for the initial five months of 2019, representing a pre-COVID-19 period, which facilitated examining industrial behavior during the preliminary assessment and comparing it with measurements recorded before and throughout the pandemic, as illustrated in [Figure 8](#).

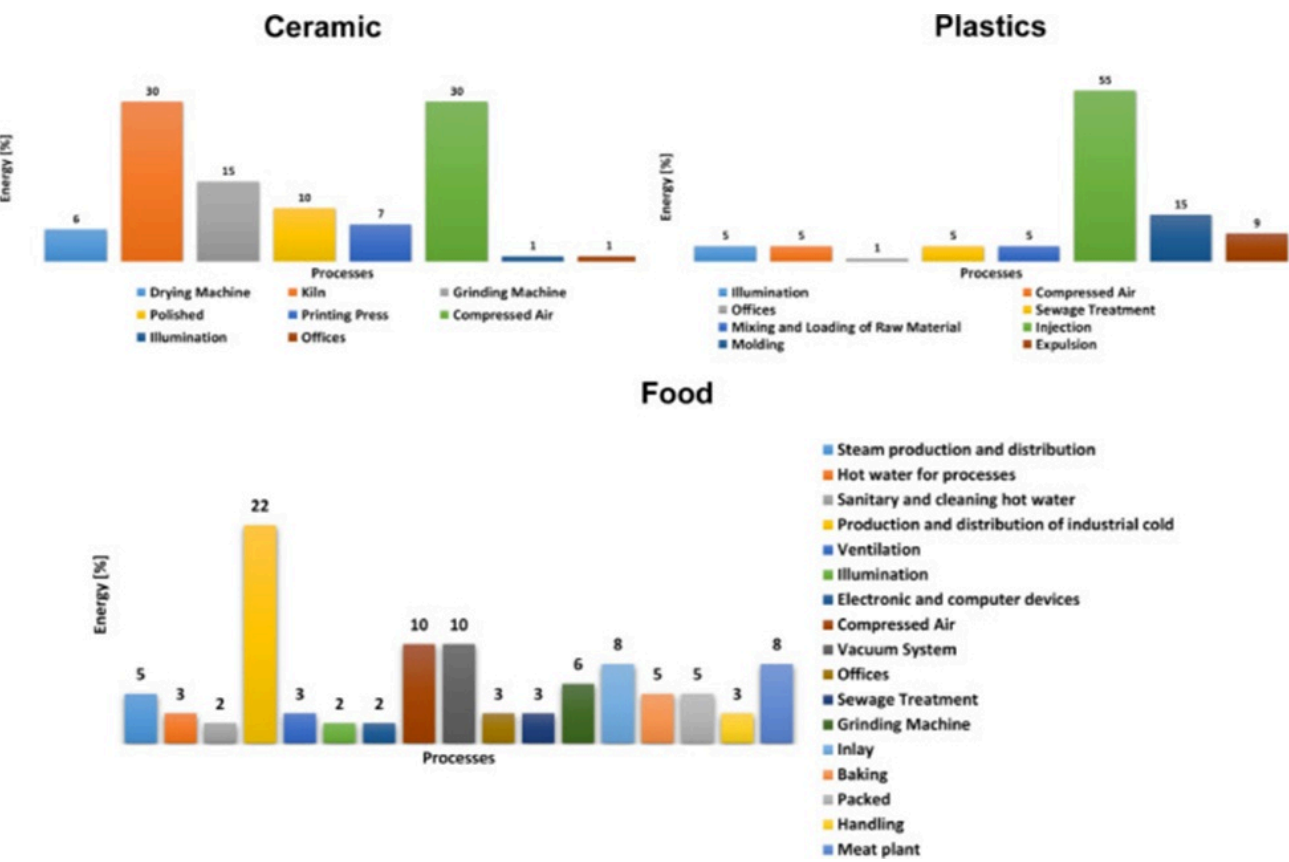


Figure 8. Outcomes of the preliminary evaluation of industrial demand response.

Information structure

To ensure reliable model training for energy disaggregation, it is crucial to obtain consumption data at strategic points within industrial operations. The smart meters collected information including date, time, phase currents (R, S, T in A), line voltages (RS, ST, TR in V), active power (kW), and power factor (pf).

For the disaggregation process, only the date, time, and active power (kW) of each process are utilized after appropriate data preprocessing. All input variables must be numeric, with dates and times converted to UNIX timestamp format as required by NILMTK [32]. The dataset is structured into two columns: the first represents the timestamp in 10-digit Unix Epoch format, and the second contains the active power of the production process in NILMTK-DF format. After reformatting, the industrial measurements are organized into channels based on the REDD database layout [32], as shown in Figure 9.

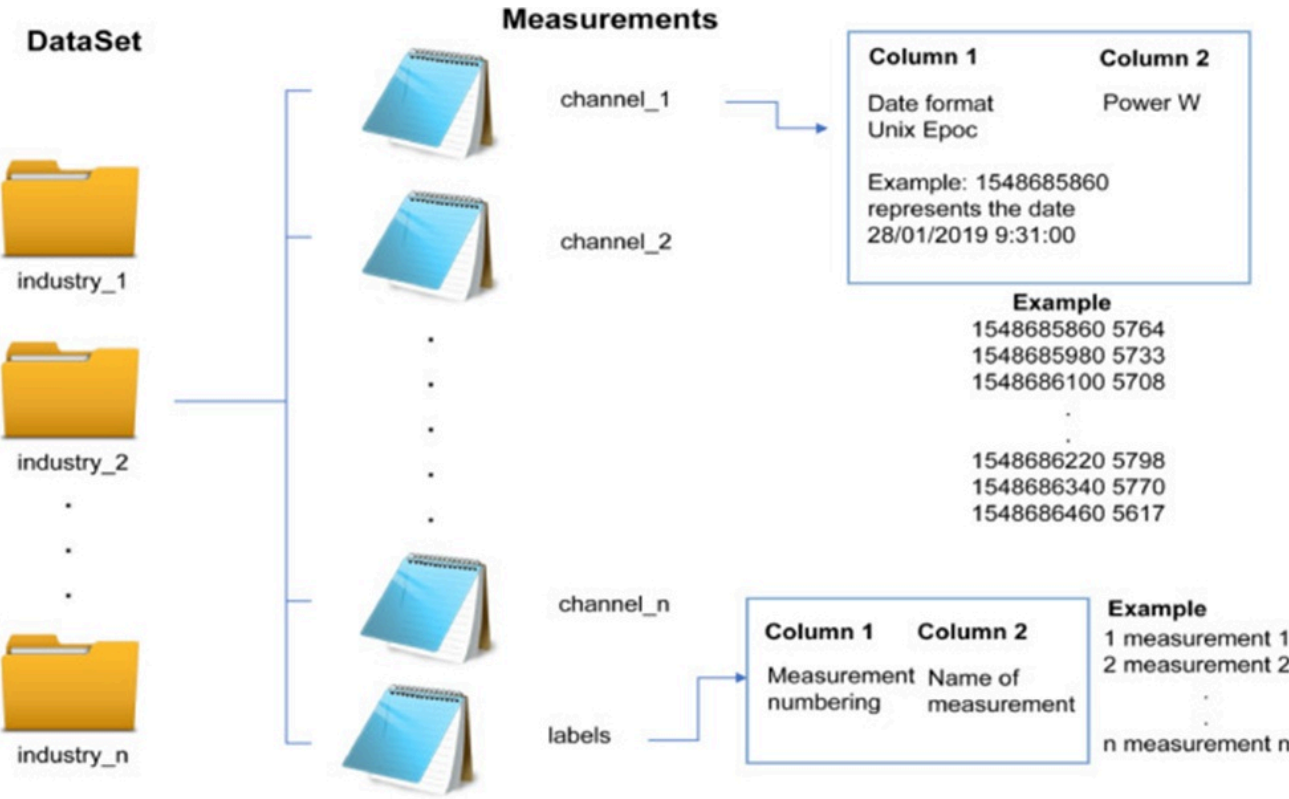


Figure 9. Configuration of the ecdiac dataset.

Data analytics

The ecdiac dataset consists of several fields: the first column records the date and time as a timestamp, and the remaining columns capture power consumption across various industrial process measurement channels. Table 2 provides a detailed overview of this dataset's structure.

Table 2. Structure of the ecdiac dataset measurements.

Empty Cell	Delivery Point	Clinker M2 1	Clinker M2 2	Clinker P1	Graiman 1–2	Graiman pox	Tugal Vand
2019–01–28 09:30:00	5764.0	0.0	110.0	848.0	3211.0	1287.0	363.0
2019–01–28 09:32:00	5733.0	0.0	183.0	856.0	3087.0	1346.0	363.0
2019–01–28 09:34:00	5708.0	0.0	109.0	830.0	3052.0	1357.0	345.0
2019–01–28 09:36:00	5798.0	0.0	184.0	846.0	3142.0	1301.0	351.0
2019–01–28 09:38:00	5770.0	0.0	110.0	844.0	3221.0	1322.0	348.0
...	...	...	...	...	...	...	...

2020-11-01 13:20:00	5185.0	0.0	57.0	119.0	3744.0	918.0	344.0
2020-11-01 13:22:00	5099.0	0.0	58.0	116.0	3556.0	987.0	333.0
2020-11-01 13:24:00	4955.0	0.0	62.0	79.0	3501.0	981.0	336.0
2020-11-01 13:26:00	5205.0	0.0	62.0	79.0	3699.0	948.0	461.0
2020-11-01 13:28:00	5023.0	0.0	59.0	110.0	3428.0	961.0	467.0

Load profile at the industry delivery point

Analyzing the load at the delivery point reveals the industry's operational characteristics, highlighting both power consumption levels and the specific time intervals of the recordings. **Figure 10** displays the load pattern observed at the industry's point of delivery.

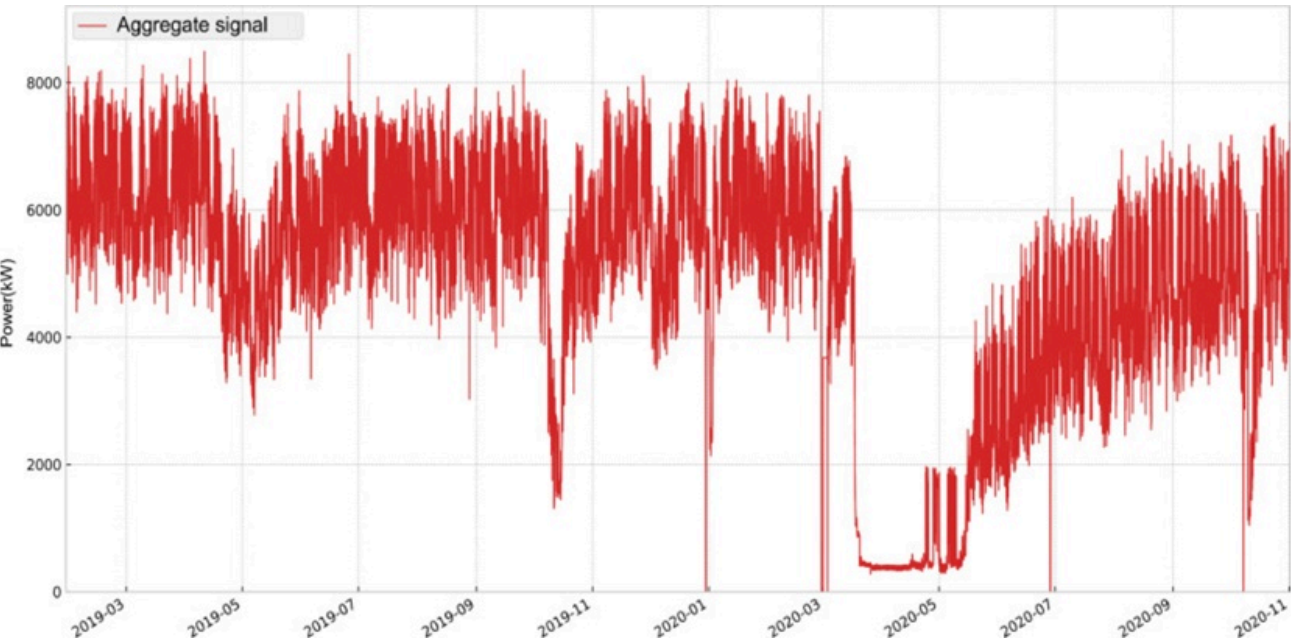


Figure 10. Load profile at the industry delivery point.

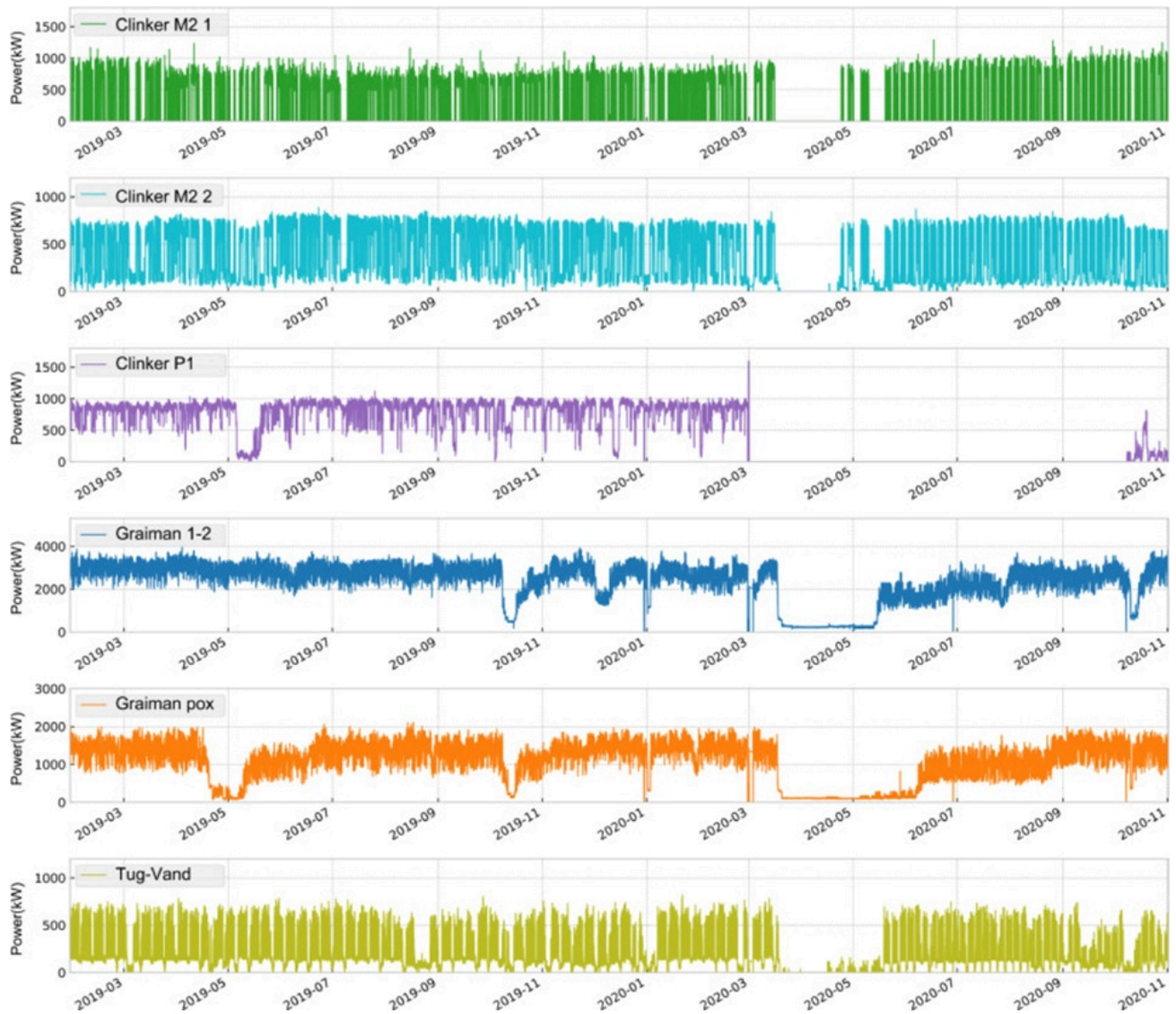


Figure 11. Consumption profiles of final energy use in industrial processes.

Consumption profiles of the final use of energy in the processes of the industry

By examining the final energy consumption profiles across industrial processes, it becomes possible to obtain a clear breakdown of the overall load. This detailed view reveals the typical operating patterns and power-demand ranges for each process. **Figure 11** displays these final-use energy consumption profiles for the different industrial activities.

Data preprocessing

Preprocessing of the dataset was performed using the NILMTK toolkit, while considering several key factors:

Analysis of the missing data of the dataset

Gaps and missing values in the records are detected using metadata associated with each meter, primarily by adjusting the maximum sample period setting. This parameter specifies the maximum allowable interval between readings before the meter is considered inactive. The examination shows several time intervals where the metering equipment did not capture data as expected.

Uptime and abandonment rate

Uptime is calculated based on the maximum sample period defined in each meter’s metadata. Any period exceeding this limit without recorded data is interpreted as the meter being turned off. The abandonment (dropout) rate, however, refers exclusively to unintended data losses caused by problems in the data acquisition system, even when measurements should have been taken within the allowed time window.

Active power spectrogram of the meter at the point of delivery

The active power spectrogram provides valuable visual information to identify periods of increased operational activity within the industrial facility.

Energy statistics

These statistics provide an overview of the facility’s total energy usage, along with average power values calculated daily, weekly, and monthly across the full dataset.

Activity register

The activity register records time windows during which recorded power exceeds a specific threshold. This threshold is set based on the average consumption level observed for each meter.

Data matrix

After preprocessing, the time window used to construct the data matrix spans February 2019 to February 2020. This interval includes a disruption caused by the pandemic, making it necessary to handle atypical data points carefully to avoid distorting statistical results.

The monthly average energy consumption during this period was analyzed to identify the most appropriate segments for model training and validation.

Table 3 summarizes the total energy for the selected 13-month period. The data was divided into 70% for model training and 30% for validation. Training was performed on two separate intervals: February–April 2019 and June–September 2019. May and October were excluded from training due to their below-average energy consumption, which could negatively affect model performance. Validation was conducted using data from November 2019 to February 2020 (a total of 4 months).

Table 3. The monthly average energy of industrial processes from actual measurements.

Empty cell	kWh Feb 2019	kWh Mar 2019	kWh Abr 2019	kWh May 2019	kWh Jun 2019	kWh Jul 2019	kWh Aug 2019	kWh Sep 2019	kWh Oct 2019	kWh Nov 2019	kWh Dec 2019	kWh Jan 2020	kWh Feb 2020
Delivery point	4621	4535	4235	3695	4363	4517	4428	4475	3455	4420	4123	4320	4230
Clinker M2 1	348	272	284	226	300	305	284	313	335	292	285	296	297
Clinker M2 2	311	310	349	276	414	400	376	398	376	340	320	337	327
Clinker P1	576	591	589	355	613	597	606	578	572	585	538	600	567
Graiman 1–2	2185	2159	2152	2090	1977	2025	2052	2048	1359	2084	1773	1921	1862

Graiman prox	987	1031	679	537	875	996	992	977	658	956	1061	982	1013
Tugal vandervil	239	197	206	227	209	220	142	186	174	187	176	209	215

NILMTK applied to industry delivery point load profile disaggregation

Data processing

In this phase, the focus shifts to building and utilizing the disaggregation models. The CO and FHMM algorithms are trained on the dataset, after which the trained models are exported for later use. The actual disaggregation step is performed by loading these saved models, which allow the system to produce the separated load component results. This complete procedure is shown in [Figure 12](#).

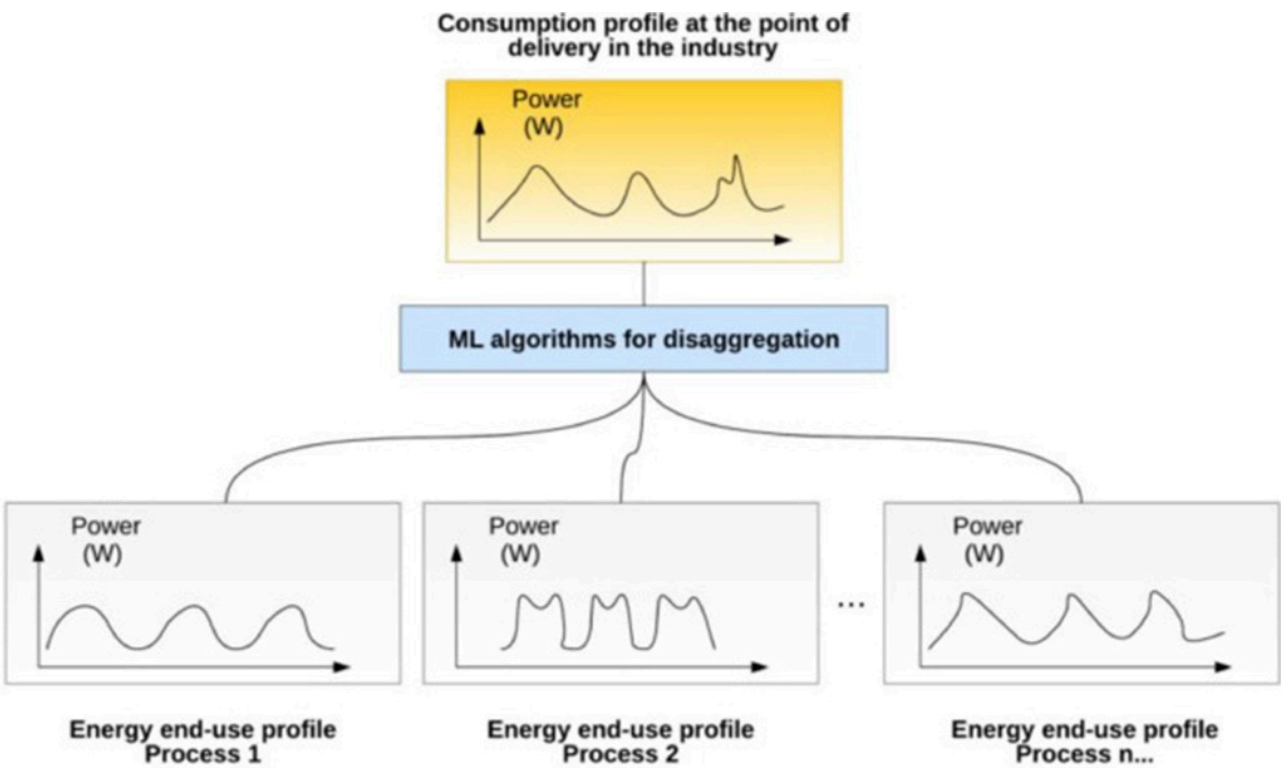


Figure 12. Disaggregation model.

Training and export of disaggregation models

The creation of disaggregation models was based on the final energy end-use profiles from the various industrial processes included in the dataset. The available data was split into 70% for training and 30% for performance validation using the chosen evaluation metrics.

Models were trained at five sampling rates: 2, 5, 10, 15, and 30 minutes. To determine the most effective setups, three resampling approaches were tested: using the first valid value, calculating the mean, and calculating the median.

Since mean and median resampling involve recomputing the original signal, training with these methods was limited to the 5-minute, 10-minute, 15-minute, and 30-minute intervals.

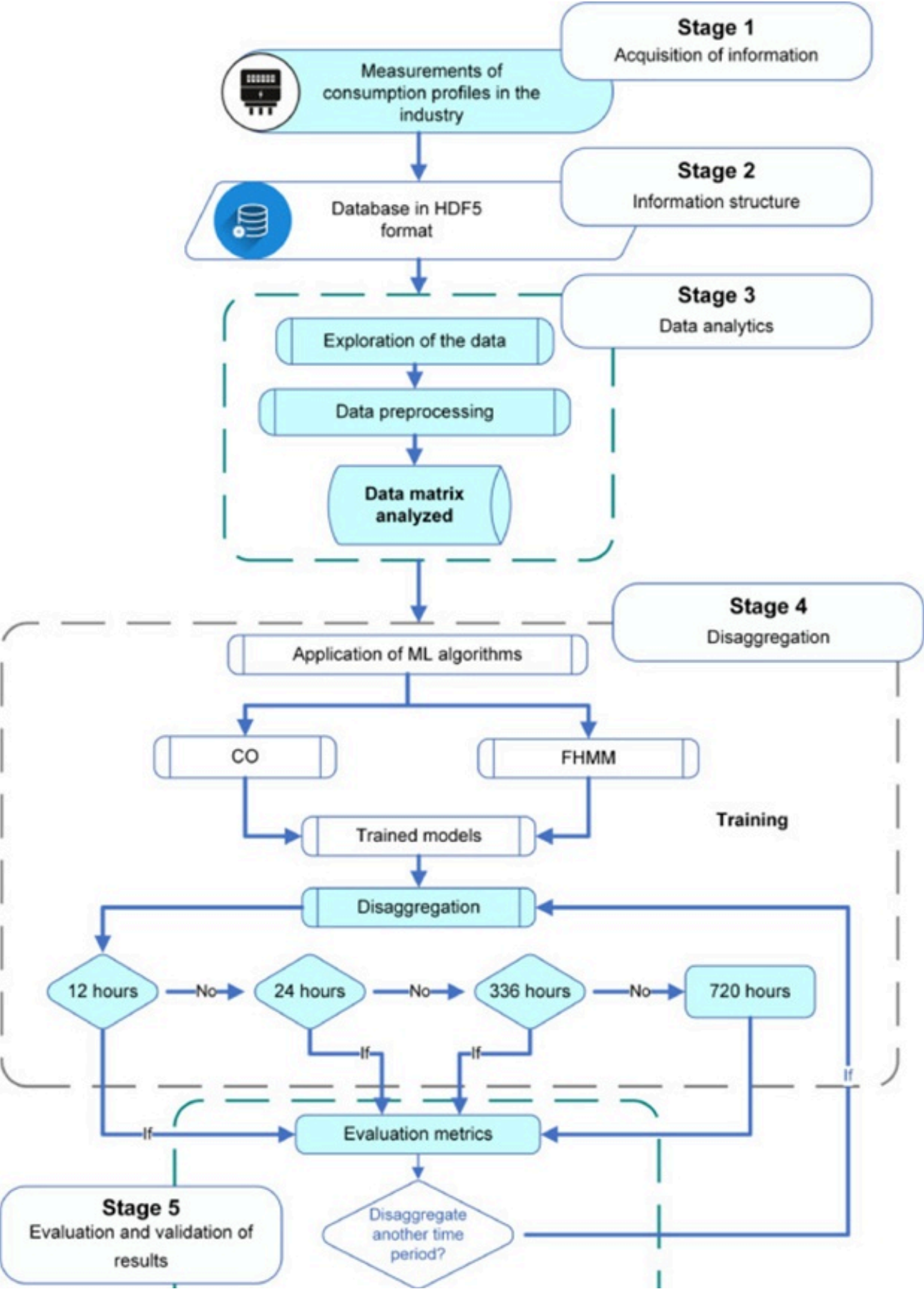
Import and disaggregation of trained models

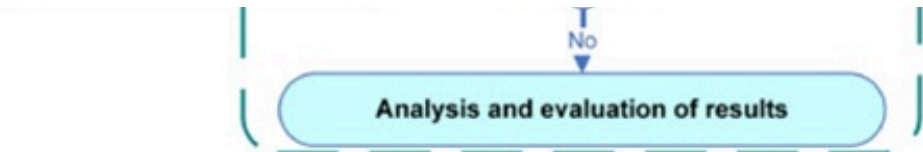
Once trained, the models were loaded and applied to the overall load profile recorded at the industry's point of delivery. This step utilized the Meter Group called `valid_Main`, which stores measurements from the facility's primary electrical meter.

The first attempt at disaggregation covered the entire four-month validation period, yet the resulting metrics were unsatisfactory. Consequently, the process was repeated across several alternative time windows to improve accuracy. The additional windows tested were 12 hours, 24 hours (one full day), 336 hours (14 days), and 720 hours (30 days).

Results and Discussion

The diagram in **Figure 13** illustrates the full methodology developed for disaggregating industrial load profiles. This approach is organized into five key stages. The fourth stage proved to be the most influential, as it incorporates the evaluation of disaggregation across varying time windows. Based on the metrics' performance, the system determines whether to repeat the disaggregation using a different time interval.





fundamental requirement. These promises provide the necessary data foundation to ensure the disaggregation algorithms can be applied with confidence.

Information structure

All collected measurements must be stored in HDF5 format. The database needs to properly organize all measurement channels related to the overall industry load, as well as the separate channels from each specific process.

Data analytics

During this stage, the dataset is examined in depth, including metadata and raw measurements. Preprocessing steps are applied using the NILMTK tools described earlier. The final output of this phase is a clean, structured data matrix that is ready for the disaggregation stage.

Disintegration

Here, the selected machine learning algorithms — CO and FHMM — are applied. They are responsible for both training the models on the prepared data and executing the load separation process.

Analysis of the results

The highest-performing model is chosen and rigorously checked against the evaluation metrics. This step enables a comprehensive interpretation and assessment of how well the disaggregation performed.

Evaluation and validation of ML algorithms for disaggregation

This stage consists of applying the performance metrics to the outputs generated by the disaggregation models. The aim is to determine which approach delivers the strongest results, confirm that the total of all predicted loads equals the measured aggregate load at the delivery point, and calculate the energy deviation percentage between predicted and actual values for every individual process.

Table 4 displays the models that achieved the strongest performance for each tested time window. The disaggregation performed on the first time window stood out as the most successful according to the evaluation metrics, particularly in its ability to accurately estimate each process.

Table 4. Models with better results for the time windows.

Time Windows			
12 h	24 h (1 day)	336 h (14 days)	720 h (30 days – 1 month)
CO-median 15 min.	FHMM-median 10 min.	FHMM-median 30 min.	FHMM-median 30 min.

The evaluation metrics from the CO-median model at 15-minute resolution provide a clear picture of the model's overall effectiveness. Among all processes, Clinker P1 achieved the strongest performance, recording an RMSE of 84,353 kW and a notably low EAE of 0.311 kWh. In contrast, the Tug-Vand process yielded considerably weaker results, with an RMSE of 200.64 kW and an EAE of 1,142 kWh. These large discrepancies stem primarily from the frequent and abrupt power fluctuations characteristic of the Tug-Vand process, which also accounts for a substantial portion of total energy use. Further details are presented in **Table 5**.

Table 5. CO-median model evaluation metrics 15 min.

Empty Cell	Clinker M2 1	Clinker M2 2	Clinker P1	Graiman 1–2	Graiman pox	Tugal-vanderbilt
F1 (%)	0.907	0.968	1000	1000	1000	0.979
EAE (kWh)	0.373	0.949	0.311	0.083	0.439	1142
MNEAP	0.193	0.192	0.052	0.031	0.050	0.351
RMSE (kW)	158,797	145,486	84,353	139,859	111,287	200,640

Figure 14 shows the disaggregation outcomes produced by the CO-median model with 15-minute sampling over a 12-hour time window. These results are directly compared with the measured consumption for each process in the validation dataset.

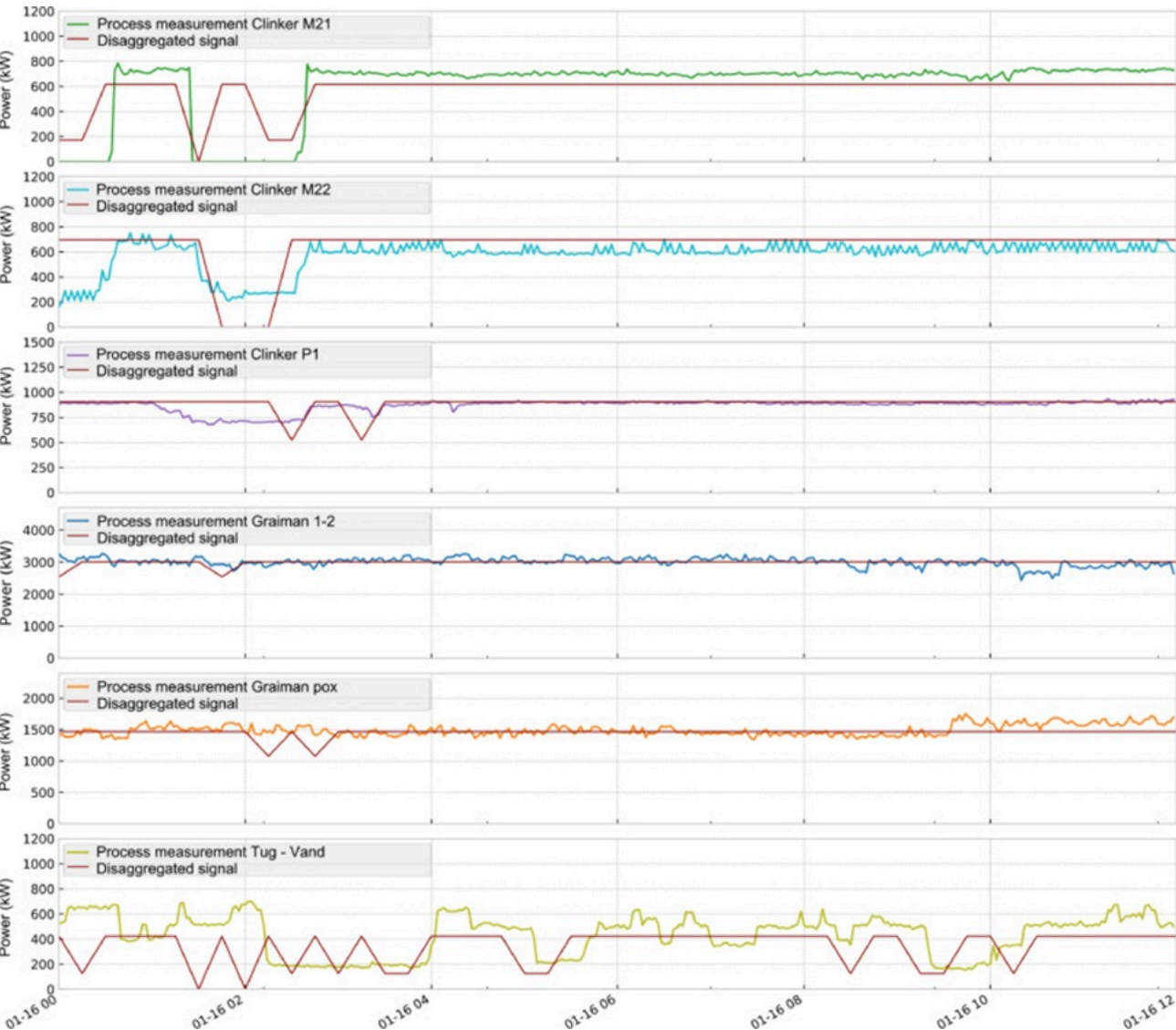


Figure 14. Outcomes of the 12-hour energy disaggregation using the CO-median model at 15-minute intervals.

The aggregate of all process-level predictions generated by the CO-median model (using 15-minute sampling over the 12-hour window) was compared with the actual total load profile recorded at the industry’s delivery point. This comparison revealed a small energy difference of 0.388 kWh, corresponding to an error percentage of just 0.467%. Further visualization of this comparison is provided in **Figure 15**. **Table 5** details the energy differences and percentage errors between the predicted and measured values (**Table 6**).

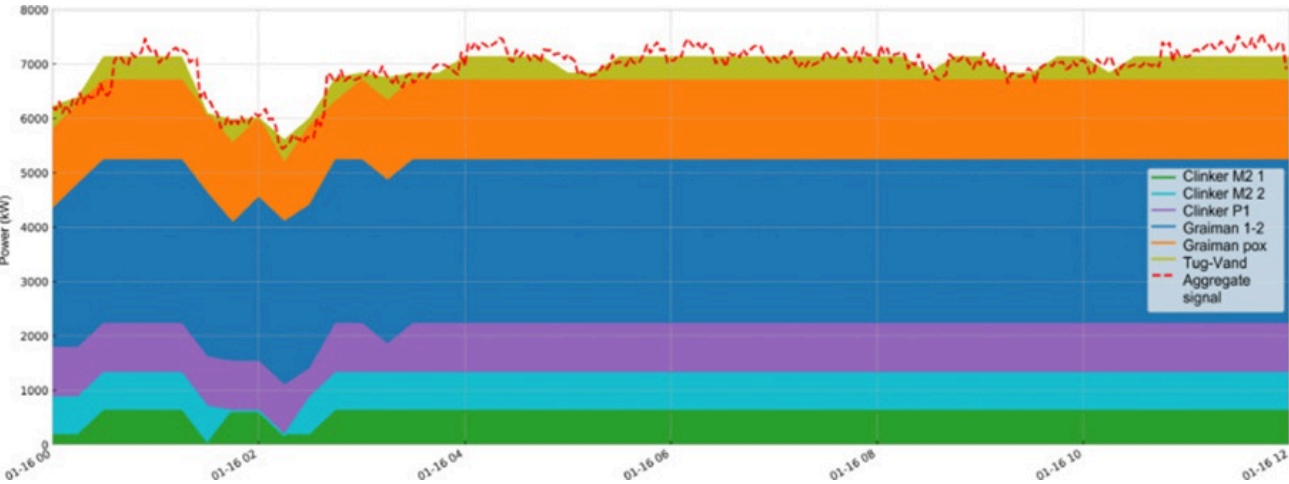


Figure 15. Aggregate disaggregation results of the OC model across processes at 15-minute intervals.

Table 6. Percentage error comparing the total disaggregation with the industry load profile using the CO-median model at 15-minute intervals.

Description	kWh
Energy sum predictions	82,699
Energy actual load profile	83,086,967
Difference	0.388
Mistake%	0.467

Figure 16 illustrates the proportion of energy accurately allocated across the various processes, showing that, for certain methods, the discrepancy between predicted and measured energy does not exceed 1%.

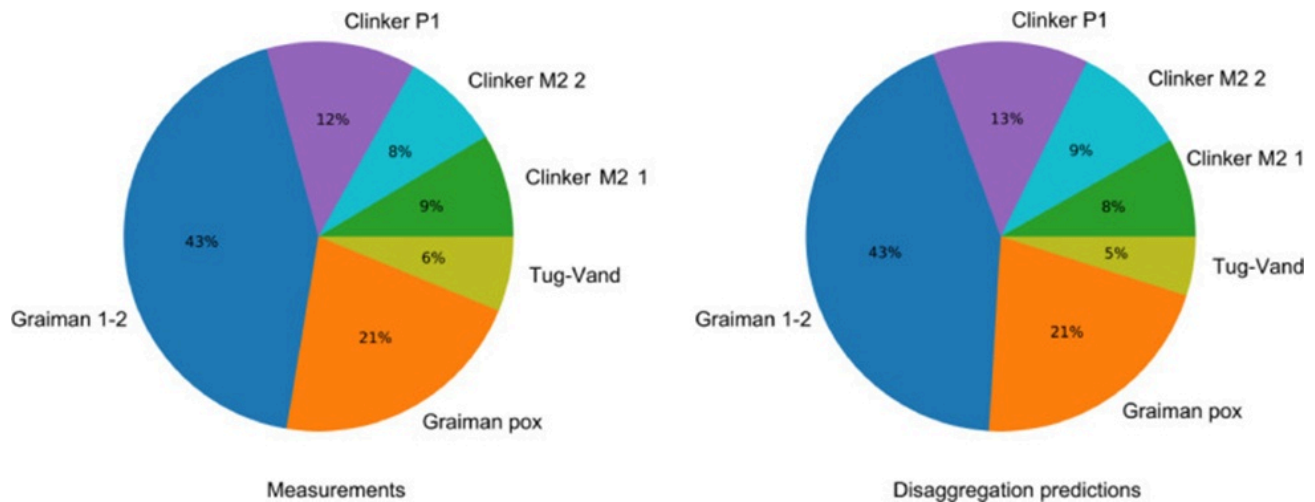


Figure 16. Proportion of total energy accurately allocated by the CO-median 15 min model.

In most studies, Non-Intrusive Load Monitoring (NILM) has been primarily applied to residential consumption, typically leveraging public datasets such as REDD and UKDALE, which contain power usage data for common household devices such as televisions, kettles, and microwaves. In contrast, this work applies NILM to an industrial setting. Training was performed using the CO and FHMM algorithms, which operate through multiple stages, including detecting states within the load profile and clustering data to match similar magnitude patterns.

Analysis across various time windows revealed that disaggregation accuracy improved when conducted over a shorter 12-hour period with a 15-minute sampling interval, employing median resampling. Evaluation metrics indicated superior performance for processes exhibiting stable, patterned measurements and minimal abrupt state changes. Additionally, using mean and median resampling, FHMM-trained models across the last three time windows with 10–30 minute sampling intervals outperformed those trained using the CO algorithm.

Conclusion

The industrial sector is a key component of the Fourth Industrial Revolution, driving economic development and enhancing the quality of life. In developing nations, the expansion of R&D systems has lagged due to production-focused policies and delayed prioritization of innovation frameworks.

Industry 4.0 technologies, such as SCADA systems and advanced communication protocols, facilitate energy monitoring and data acquisition. Implementing these technologies and strategies to analyze final energy usage supports energy efficiency and optimizes process performance. The integration of IoT, Big Data storage, and machine learning enables industries to collect detailed information across the supply chain, anticipate events, and make informed decisions about production activities.

Machine learning techniques have become essential in tackling nonlinear industrial problems, with NILM providing a tool to optimize energy usage and reduce waste by analyzing detailed process-level consumption and overall industrial load profiles. NILM, or energy disaggregation, allows precise monitoring of each production process, supporting intelligent energy management, behavioral load adjustments, and the development of informed policies, such as Time-of-Use (TOU) pricing for utility companies.

Acronyms

BD = Big Data

CO = Combinatorial Optimization

CPS = Cyber-Physical Systems

DCS = Distributed Control System

DSM = Demand Side Management System

EAE = Error in Total Energy Assigned

FHMM = Factorial Hidden Markov Model

FFT = Fast Fourier Transform

ML = Machine Learning

NILM = Non-Intrusive Load Monitoring

NILM-TK = NILM Toolkit

PLC = Programmable Logic Controller

R&D = Research and Development

RMSE = Root Mean Square Error in Assigned Power

SCADA = Supervisory Control and Data Acquisition

TOU = Time-of-Use Tariffs

Acknowledgements

This work was developed under the auspices of the Polytechnic University of Valencia (Spain), the Polytechnic University Salesian (Cuenca, Ecuador), and the Catholic University of Cuenca (Ecuador).

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Universitat Politècnica de Valencia provided Marco Toledo-Orozco reports. Marco Toledo-Orozco reports a relationship with Universitat Politècnica de Valencia that includes non-financial support. Marco Toledo-Orozco hasn't patented the Methodology for the Disaggregation and Forecast of Demand Flexibility in Large Consumers with the Application of Non-Intrusive Load Monitoring Techniques. The author is a doctoral candidate in the Doctoral Program in Engineering and Industrial Production at the Polytechnic University of Valencia; his areas of interest are optimization methods, aggregator models, electrical markets, and electromobility. He belongs to the research group of the Energy Engineering Department at the Polytechnic University of Valencia.

Financial support

None

Ethics statement

None

Received: 27 Mar 2025

Revised: 08 May 2025

Accepted: 20 Jun 2025

Published online: 20 September 2025

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Buele Nugra IG, Cuesta Astudillo FL, Chillogalli Peña CF. Metricas para diagnosticar os balanços financeiros das companhias do setor industrial manufatureiro EM Cuenca, Equador. *Innovar*. 2020;30(76):63-75.
- Glas J, Alvarado V, Leon S, Parra J. Política Industrial del Ecuador 2016-2025. 2016 [cited 2026 Jun 28]. Available from: <https://www.industrias.gob.ec/wp-content/uploads/2017/01/politicaIndustrialweb-16-dic-16-baja.pdf>.
- Riaga CO, Cante MAA. Factores determinantes de los procesos de innovacion: una mirada a la situacion en Latinoamerica. *Rev Esc Adm Neg*. 2012;(72):148-63.
- Stock T, Seliger G. Opportunities of sustainable manufacturing in Industry 4.0. *Procedia CIRP*. 2016;40:536-41.
- Rastegar S. Multi parametric intelligent identification and robust control methodologies for industrial processes [dissertation]. [place unknown]; 2017.
- Kanchiralla F, Jalo N, Johnsson S, Thollander P, Andersson M. Energy end-use categorization and performance indicators for energy management in the engineering industry. *Energies*. 2020;13(2):369.
- Arvind A, Bourne D. Architecture for industry 4.0-based manufacturing systems [master's thesis]. Pittsburgh (PA): School of Computer Science, Carnegie Mellon University; 2016.
- Batra N, Kelly J, Parson O, Dutta H, Knottenbelt W, Rogers A, et al. NILMTK. In: Proceedings of the 5th International Conference on Future Energy Systems; 2014 Jun; Cambridge, UK. New York: ACM; 2014. p. 265-76.
- Paucar Uyaguari PE, Yubi Quito LA. Pronostico de perfiles de consumo de electricidad a traves de modelos estadísticos y desarrollo de una metodología comparativa [dissertation]. [place unknown]; 2020.
- Cimoli M, Castillo M, Porcile G, Stumpo G. Políticas industriales y tecnologicas en America Latina. Santiago: CEPAL; 2017.
- Cugliari J, Goude Y, Poggi J. Disaggregated electricity forecasting using wavelet-based clustering of individual consumers. In: 2016 IEEE International Energy Conference (ENERGYCON); 2016 Apr 4-8; Leuven, Belgium. IEEE; 2016. p. 1-6.

Diez-Olivan A, del Ser J, Galar D, Sierra B. Data fusion and machine learning for industrial prognosis: trends and perspectives towards industry 4.0. *Inf Fusion*. 2019;50:92-111.

Peralta AG. Planning of FiWi networks to support communications infrastructure of SG and SC. In: *Broadband Communications Networks - Recent Advances and Lessons from Practice*. [place unknown]: IntechOpen; 2018.

Luthra S, Mangla SK. Evaluating challenges to Industry 4.0 initiatives for supply chain sustainability in emerging economies. *Process Saf Environ Prot*. 2018;117:168-79.

Medojevic M, Diaz Villar P, Cosic I, Rikalovic A, Sremcevic N, Lazarevic M. Energy management in industry 4.0 ecosystem: a review on possibilities and concerns. In: *Proceedings of the 29th DAAAM International Symposium*; 2018; Vienna, Austria. DAAAM; 2018. p. 0674-80.

Carrillo-Rojas GJ, Andrade-Rodas J, Barragan-Escandón A, Astudillo-Aleman A. Impact of electrical energy efficiency programs, case study: food processing companies in Cuenca, Ecuador. *Dyna (Medellin)*. 2014;81(184):41.

Shrouf F, Gong B, Ordieres-Mere J. Multi-level awareness of energy used in production processes. *J Clean Prod*. 2017;142:2570-85.

Perez-Lopez E. Los sistemas SCADA en la automatizacion industrial. *Rev Tecnol Marcha*. 2015;28(4):3-14.

Aslam W, Soban M, Akhtar F, Zaffar NA. Smart meters for industrial energy conservation and efficiency optimisation in Pakistan: scope, technology and applications. *Renew Sustain Energy Rev*. 2015;44:933-43.

Gan S, Li K, Wang Y, Cameron C. IoT based energy consumption monitoring platform for industrial processes. In: *2018 UKACC 12th International Conference on Control (CONTROL)*; 2018 Sep 5-7; Sheffield, UK. IEEE; 2018. p. 236-40.

Czosnyka M, Wnukowska B, Karbowa K. Energy efficiency as a base of the energy economy in a modern industrial plants. In: *2018 Progress in Applied Electrical Engineering (PAEE)*; 2018 Jun 18-22; Koscielisko, Poland. IEEE; 2018.

Buzachis A, Galletta A, Celesti A, Fazio M, Villari M. Development of a smart metering microservice based on fast fourier transform (FFT) for edge/internet of things environments. In: *2019 IEEE 3rd International Conference on Fog and Edge Computing (ICFEC)*; 2019 May 14-17; Larnaca, Cyprus. IEEE; 2019. p. 1-6.

Gonzalez DA. En galego: desagregacion do consumo enerxetico nun fogar inteli [dissertation]. [place unknown]; 2016.

Reuelta Herrero J, Pérez Alonso A, Armentia Sáez JA, Lozano Ruiz Á, García Berjano P, Álvarez Utrera A, et al. Non intrusive load monitoring (NILM): a state of the art. In: *Proceedings of the 15th International Conference on Soft Computing Models in Industrial and Environmental Applications*. Springer; 2018. p. 125-38.

Tien PW, Wei S, Darkwa J, Wood C, Calautit JK. Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality—a review. *Energy AI*. 2022;100198.

Orzaez JMA. Non-intrusive load monitoring techniques for the activity of daily living recognition [dissertation]. 2017. Available from: <http://hdl.handle.net/10017/41536>.

Fiol Arguimbau A. Algorithms for energy disaggregation [master's thesis]. 2016.

Osathanunkul K, Osathanunkul K. Simple Load Disaggregation Library based on NILMTK. In: *2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*; 2020; [place unknown]. IEEE; 2020. p. 141-4.

Schwegmann S, Faulhaber J, Pfaffel S, Yu Z, Dörenkämper M, Kersting K, et al. Enabling Virtual Met Masts for wind energy applications through machine learning methods. *Energy AI*. 2023;11:100209.

Kelly J, Knottenbelt W. Metadata for Energy Disaggregation. In: 2014 IEEE 38th International Computer Software and Applications Conference Workshops; 2014 Jul 21-25; Västerås, Sweden. IEEE; 2014. p. 578-83.

Kelly D. Disaggregation of domestic smart meter energy data [dissertation]. [place unknown]; 2016.

Kolter JZ, Johnson MJ. REDD: a public data set for energy disaggregation research. In: Workshop on Data Mining Applications in Sustainability (SIGKDD); 2011; San Diego, CA. [place unknown: publisher unknown]; 2011. p. 59-62.

Zoha A, Gluhak A, Imran M, Rajasegarar S. Non-intrusive load monitoring approaches for disaggregated energy sensing: a survey. *Sensors*. 2012;12(12):16838-66.