A Systematic Review on Low-Resolution NILM: Datasets, Algorithms, and Challenges

Deepika R. Chavan and Dagadu S. More

Abstract The active large-scale deployment of electrical smart meters throughout the world offers opportunities to analyze smart meter data to generate numerous innovative applications, Non-Intrusive Load Monitoring (NILM) is one such application that goes beyond remote and precise billing. The NILM has been a popular and growing methodology for monitoring the energy profile of a household building and disaggregating overall power consumption into individual appliance usage. The device-level energy consumption information would assist users to understand their device usage behavior and take required actions to reduce energy consumption. This paper systematically reviews the NILM approaches exclusively for low-resolution smart meter data. This review highlighted the low-resolution energy datasets and their feature measurements, the state-of-the-art algorithms explored and developed for low-resolution NILM systems. Furthermore, this study discussed the challenges related to the low-resolution NILM model performance, data scarcity, three-phase data, etc. Finally, the existing research gaps as well as potential research directions in the Indian context are described in detail.

Keywords Energy disaggregation · Low-resolution NILM · Machine learning · Deep learning

1 Introduction

India is the world's third-largest producer and consumer of electricity, as well as the fourth-largest emitter of CO_2 [1]. India's energy sector accounted for 68.7% of greenhouse gas emissions. The residential sector's energy consumption is 24.01% and is increasing year after year [2]. Electric production and consumption are major sources of CO_2 emissions. Rapid urbanization causes high energy demand, ultimately burdened to the limited energy resources, so it is highly essential to manage the energy

D. R. Chavan (✉) · D. S. More
Department of Electrical Engineering, Walchand College of Engineering, Sangli, India

D. S. More
e-mail: dagadu.more@walchandsangli.ac.in

P. K. Mallick et al. (eds.), *Electronic Systems and Intelligent Computing*, Lecture Notes in Electrical Engineering 860, https://doi.org/10.1007/978-981-16-9488-2_9

expenses. One of the solutions is monitoring the load behavior and load energy consumption patterns for efficient and effective energy utilization. The problem of effective energy consumption monitoring has attracted a lot of researchers. The prior work reported in [3] divided the load monitoring system into two categories.

1. Intrusive Appliance Load Monitoring (ILM)
2. Non-Intrusive Load Monitoring (NILM)

The ILM technique utilizes a single measurement device connected to each home appliance, leading to increased costs and complexity, however in NILM, only one measuring tool is required to find the individual information about the device [3]. Due to its non-intrusive nature, NILM has a major benefit that it does not require to modify the existing building infrastructure. NILM finds a better solution to disaggregate the total power into the individual device level power consumption. Energy Disaggregation is generally worked on Software as a Service (SaaS) or sensor-based solution to classify fine energy consumption data from whole aggregated data. With the help of customer electricity usage patterns logged by the NILM system, recent electronics companies such as Samsung Electronics, ABB, LG, Apples, and others projected that efficient management of electricity demand saves $6 million. Hence, appliance-centric electricity management and monitoring have equal importance [4].

The concept behind NILM was first introduced by G.W. Hart in 1992, based on steady-state active power feature. Many researchers have been proposed several NILM techniques that are based on Factorial Hidden Markov Model (FHMM), V-I trajectory, Wavelet Transform, Graph Signal Processing, Neural Network, Machine Learning, genetic algorithms [5]–[9], additional methodologies can be found in [10]–[12]. Large availability of energy datasets like REDD, AMPds, etc. many of these are listed in [13] encouraging to the researchers to adopt NILM in various applications like scheduling appliances to reduce peak hour demand [14], smart home energy management including ambient parameters [15], anomaly detection in appliances [16], developing household characteristics [17], non-technical losses reduction[18].

According to the recent development in NILM, a very few review articles on NILM have been published in the last 15 years. Most of the reviews [12, 19, 20] briefly elaborated with specific approaches for appliance classification, supervised/unsupervised learning, event/non-event feature detection/extraction, and performance measure metrics. Furthermore, articles like [21, 22] reviewed different degrees of freedom of NILM technique. However, a complete overview of low-resolution NILM is missing up to now, based on datasets, input features, disaggregation approaches as well as research challenges particularly in machine learning and deep learning area of research. Thus, this paper contributes a complete overview of state-of-the-art NILM techniques based on low-resolution smart meter data. The main contributions of this work in each section are specified as follow:

1. This paper summarizes the existing status of NILM research in Sect. 1.1, providing essential details.
2. Sect. 1.2 illustrates a structured review of data acquisition and low-resolution energy datasets with their specifications and future research directions.

3. This study describes the feature extraction and detection in terms of macroscopic features and it is presented in Sect. 1.3
4. An overview of state-of-the-art algorithms utilized in the low-resolution NILM with machine learning and deep learning approaches is discussed in Sect. 1.4.
5. Finally, this paper summarizes the research gap regarding NILM performance comparison, multiple input features, data scarcity, and three-phase datasets are discussed in Sect. 1.6. Furthermore, this study also points out the future direction of research in Indian context.

We hope our contributions will inspire future researchers and lead to new achievements.

1.1 Smart Meter Rollout and Low-Resolution NILM

India expected to rise in electricity by 79% in next decade, with this energy production enhancement, nation needs to cut down Aggregate Technical and Commercial (AT&C) losses below 10% by the year 2027. To achieve this aim, India forms Advanced Metering Infrastructure (AMI) along with the new range of smart meters. Under Smart Meter National Programme (SMNP), Government of India has been working on to replace 250 millions of conventional meters by new Smart meters [23]. The active large-scale roll-out of electrical smart meters throughout the world offers opportunities to analyze smart meter data to generate numerous innovative applications; one of such applications that go beyond precise and remote billing is Non-Intrusive Load Monitoring (NILM).

It is vital that some usual NILM nomenclature be clarified and how that applies to the data being utilized. Internally, smart electric meters sample voltage and current signals at different frequencies. These frequencies categorize as low- and high-frequency ranges. The raw data can be produced directly, or the averaged value can be calculated and produced such as, the root mean square (RMS) value of voltage and current. For better understanding of smart meter data categorization, we cited low-resolution data as low frequency data.

The high- and low-frequency smart meter data with NILM are discussed briefly in literature [24]. The low-frequency sampling approaches are the ones that utilize information generated at rates below the AC Fundamental frequency (50 Hz in India). However, the high-frequency sampling approach utilizes data generated higher than AC fundamental frequency usually up to few KHz. The benefits of using high frequency data should be pretty evident as this preserves all the signals and allows to extract the greatest amount of information. This is revealed in [25], However, obtaining high-frequency data is costly in terms of both hardware and installation time. On the other hand, the loss of information at low frequencies can be compensated without the need for additional hardware installation. In practice, the resolution of a smart electric meter maintained lower than 1–60 s due to limitations in data storage, data handling, and privacy protection. This leads to motivate researchers to investigate NILM with low-resolution smart energy meter data with existing building infrastructure. For more information about NILM with high- and low-frequency characteristics, researchers can refer [26].

1.2 Fundamentals of NILM

The primary purpose of NILM is to break down or disaggregate the overall amount of power drawn into its component. The resulting power in a residential building is the total power consumption of each electrical device. Therefore, the goal is to determine how much electricity is consumed by each appliance. The aggregated power of N devices with respective time T is specified in Eq. (1)

$$P(t) = P_{noise}(t) + \sum_{i=1}^{N} P_i(t) t \epsilon \{1, T\} \tag{1}$$

where;

P_i= Power of each appliance.

P_{noise}= power of unwanted signal.

In order to solve the problem of power disaggregation, many different ways have been developed, the most common is to calculate P_i for $i = 1, 2, 3, \ldots, N$, from $P(t)$. According to Eq. (1), variations of power disaggregation expression are described in [24, 27]. When an issue is solved using machine learning, particularly deep learning, it is referred to as a regression problem. Although most publications employ only the active power component, the aggregation signal may also be solved by other information such as apparent power, reactive power, and current.

1.3 NILM Framework

From the recent literature review [13, 15], NILM has the following working stages:

1. Data Acquisition: Electrical signals (current, voltage, Active Power, reactive Power, Harmonic contents) are collected from measuring meter (Smart meter or by using specific Hardware) at the low or high sampling rate.
2. Feature Extractions and Event Detection: Individual appliance has its own load signature or feature pattern that leads to differentiate one appliance from another. The Event is nothing but change in electrical signals with respect to time. This transition includes appliance ON/OFF, operational mode change, and speed variations.

3. Load Classification and Energy estimation: By using features extracted from the above stage, identify which appliance is operated at a given time with power consumption. This stage includes inference and learning of models.

2 Data Acquisition and Low-Resolution Energy Datasets

The very first stage of NILM system is data acquisition or data collection. This stage has a significant role in developing NILM algorithms for a specific application. Data acquisition is associated to electronic measuring devices. Typically, a NILM system is equipped with a voltage and current sensor module that is connected to the main power line. Depending upon the data acquisition framework, the communication devices have the task of transferring measured data over a communication network [4]. Currently, NILM data have been transferred via different wireless communication protocols and stored on the server/cloud. [28].

The market offers a variety of measuring meters with different sample rates [29]. The selection of measuring meters depends on the requirement of application. The data acquisition is discussed here in terms of the sampling rate of the measuring equipment. Sampling frequency in Hz is referred as low, whereas sampling frequency in KHz and above is high. Commercial smart meters are capable of capturing low-frequency energy signals, while high-frequency signals are acquired with special acquisition boards and equipment, high-frequency data are costlier in terms of hardware and software and required more communication bandwidth to transmit the data. [25].

Companies like Neurio Technology [30], Smappee [31], ENTERTALK [32], etc. brought a straightforward solution for data acquisition with plug-in devices. These provide basic functionality of data acquisition with some considerable drawbacks regarding sampling rates, flexibility, and cost [29].

Table 1 lists the NILM energy datasets explicitly for low-frequency sampling rate. From Table 1, datasets such as RAE, I-BLEND collect data from entire domestic buildings, which are referred to as aggregated data, and Tracebace, Dataport are such datasets that gather data at both aggregate and appliance levels. On the other hand, just a few datasets from the business sector are available as shown in Table 1. Due to higher energy consumption in commercial sector, the implementation of NILM techniques will save more money than in the residential sector. The survey discovered that the majority of datasets focused on home appliances, with only a few datasets (such as COMBED) contributed to dataset of office appliances. Table 1 depicts the differences in electrical features, to the difference in disaggregation results.

Each dataset has a different recording length, ranging from 1 week to several years. To adopt a universal cost-cutting strategy, researchers can conduct a comparative study of different countries' usage patterns, by recording data in a consistent manner. Comparing and testing NILM algorithms might be difficult due to differences in sample frequency between datasets. From Table 1, it is cleared that a very small

Table 1 Low-resolution datasets with data captured location, data stored duration, different sampling rates, measurements, and published year

Serial number	Dataset	Location	Duration	Number of houses	Sampling rate	Purpose	Year	Measurements
1	REDD	USA	1 month	6	1 Hz (aggregate), 1/3 Hz (appliance)	Residential	2011	I, V, P
2	Tracebace [33]	Germany	1 day	–	1 Hz (appliance)	Residential	2012	P
3	Smart [34]	USA	3 months	3	1 Hz (aggregated and appliance)	Residential	2012	V, P, S, F
4	HES [35]	UK	1 and 12 months	251	2 min and 10 min (aggregated and appliance)	Residential	2012	P
5	Dataport [36]	USA	4 + years	1200 +	1 Hz to 1 min (aggregated and appliance)	Residential, commercial	2013	P, S
6	AMPds [37]	Canada	1 year	1	1 min (aggregated and appliance)	Residential	2013	f, V, pf, I, Q, S, P
7	iAWE [19]	India	73 days	1	1 Hz (aggregate), 1 Hz or 6 s (appliance)	Residential	2013	E, V, pf, I, Q, S, P
8	IHEPCDS	France	4 years	1	1 min (aggregated and appliance)	Residential	2013	I, V, P, Q
9	ACS-Fx	Switzerland	1 h	–	10 s (appliance)	Residential	2013	I, P, Q, pf
10	BERDS	USA	1 year	1	20 s (aggregated and appliance)	Commercial	2013	P, Q, S
11	ECO[38]	Switzerland	8 months	6	1 Hz (aggregated and appliance)	Residential	2014	I, V, P, pf

(continued)

Table 1 (continued)

Serial number	Dataset	Location	Duration	Number of houses	Sampling rate	Purpose	Year	Measurements
12	GREEND [39]	Austria/Italy	1 year	9	1 Hz (aggregated and appliance)	Residential	2014	P
13	RBSA [40]	USA	27 months	101	15 min aggregated	Residential	2014	V, P, Q, S, E
14	COMBED [41]	India	1 month	6	30 s (aggregated and appliance)	University Building	2014	I, P
15	DRED [42]	Holland	6 months	1	1 Hz (aggregated and appliance)	Residential	2015	P
16	REFIT	UK	2 years	20	8 s (aggregated and appliance)	Residential	2015	P
17	OPLD	Singapore	-	-	1 Hz (aggregated and appliance)	Commercial	2016	P, S, I
18	EEUD [43]	Canada	1 year	23	1 min (aggregated)	Residential	2017	P
19	ESHL	Germany	4 years	–	0.5–1 Hz (aggregated)	Lab equipment	2017	I, V, P
20	RAE [44]	Canada	72 days	1	1 Hz (aggregated)	Residential	2018	I, V, P, Q, S
21	ENTERTALK [45]	Korea	29–122 days	22	15 Hz (aggregated and appliance)	Residential	2019	P, Q
22	I-BLEND [46]	India	52 months	7	1 min (aggregated)	Commercial	2019	V, I f, pf, P
23	IDEAL [47]	UK	–	255	1 Hz	Residential	2020	E
24	CU-BEMS[48]	Thailand	18 months	1	1 Hz (appliance)	Commercial	2020	P, E

subset of countries (like UK, USA, Canada, Germany) contributed to energy datasets, thus it is necessary to develop country or region-specific energy datasets.

3 Appliance Feature Detection and Extraction

Individual appliance has its own load signature or feature pattern that leads to differentiate one appliance from another. The load identification in NILM is highly subject to the feature uniqueness of the appliances. So, the feature extraction methods have major role in the NILM system. The feature extraction process involved the extraction of important information from voltage and current signals through the signal processing techniques. The unique features are highly dependent on the sampling frequency of the data; this data rate is nothing but the output by measuring instrument.

The data rate separated into two groups depending upon the sampling rate, these are macroscopic and microscopic, and these are also called as low frequency and high frequency, respectively [49]. This paper reviews the feature extraction of low-frequency data rate, further these are divided into very low, low, and medium, whereas high-frequency data are categorized into high, very high, extremely high ranges. Table 2 shows the respective sampling rates with utilized features.

Most of the features employed power variables with respect to time. These power variables are voltage (V), current (I), active power (P), reactive power (Q), apparent power (S), power factor (PF), phase angle, and total harmonic distortion (THD). Most employed feature is active power and is widely used in [3, 50–54].

3.1 *Macroscopic Features*

Feature extracted from aggregated low-frequency data (from medium to very low range) is called macrolevel or macroscopic feature [55]. Generally, the macroscopic feature includes real power and reactive power variants. The actual power consumed during operation by an electric appliance is called real energy, However, unused

Table 2 Macroscopic and microscopic data rate

Parameter	Data rate
Very low	Slower than 1 min
Low	1 min to 1 s
Medium	Faster than 1 Hz to fundamental frequency (50 Hz in India)
High	Fundamental frequency to 2 kHz
Very high	2 kHz to 40 kHz
Extremely high	Faster than 40 kHz

power produced by capacitive and inductive components is reactive power, which gives further information to simplify appliance identification process [10].

Initially, the macrolevel features were examined by EPRI and MIT institutes [3, 56]. From these investigations, it is possible to detect the occurrence of an appliance being turned on or off by measuring actual and reactive power in relation to time and the accompanying positive and negative changes. Later, the MIT researchers expanded their work to apply to an industrial building's aggregate load [51]. After filtering out the sudden peaks, their research found that the appliances would have a lengthy transient period and low reactive power. As a result, ref.[11] discovered that employing transitory events as additional signatures can improve appliance detection. From recent studies, it is observed that all high-frequency data is analyzed with event-based feature extraction whereas low-frequency data is analyzed by event as well as non-event-based approach.

Feature detection in low-frequency NILM complicates the disaggregation process due to low sampling; however, the key benefit is that low-frequency data may be easily accessed without the need for any additional hardware. Considering the low sampling rate, appliance feature generation using eigenvector and to match the features during testing time pattern recognition methods has been proposed in [57]. Study [58] disaggregates total domestic electricity usage into five different categories of load. Here, evaluation made between various sparse coding algorithms. Furthermore, accuracy of a Support Vector Machine (SVM) classifier based on features is also suggested but not demonstrated. In [59], authors considered power levels and ON/OFF duration as a feature to identify appliances, both features computed with normal distribution and Weibull distribution, respectively. This work proposed maximum likelihood classifier and subtractive clustering technique, an event-based approach improvised result by exhibiting temporal relations among appliances features. Various feature extraction methods have been projected over the period of time, the related literature can be found in [10, 15].

4 Energy Disaggregation Algorithms

Energy disaggregation finds an effective and efficient solution for extracting appliance-level data from an aggregate data with an appropriate set of algorithms. In order to identify individual load data from aggregated consumption data, various disaggregation algorithms have been developed. The categorization of disaggregation algorithms is based on system learning approach and can be classifies into two major categories, one is supervised and other is unsupervised. Appliances are well labeled in supervised learning whereas unsupervised learning does not require. The energy disaggregation algorithms used in this section are heavily influenced by machine learning and deep learning area in NILM with low-resolution data.

4.1 Preprocessing

Before using the disaggregation algorithms, the raw data are transformed. The following section discusses the pre-processing steps.

(1) Resampling

Since datasets have missing values because of the failure in measurement or transmission equipment, resampling technique is utilized to get sampled data uniformly. In literature [60, 61], the original dataset has been up-sampled to the higher frequency. For on/off classification of appliances like TV, washing machine, and rice cooker, data have been down-sampled to 0.03 Hz from 10 Hz [32]. Furthermore, the study concluded that to avoid performance degradation, the sampling rate for classification task should be at least 1 Hz whereas and for regression task, it should be 3 Hz. The effect of resampling on disaggregation is carried out in numerous studies [32, 62–64].

(2) Normalization

To normalize the data, a variety of approaches have been used, the majority of methods compute the mean over the entire training set in order to normalize the training data. To reduce the statistical sensitivity of the data to outliers, prior to normalization, arcsinh employed to transforms the data [44]. Study [34] carried instance normalization whereas [146] revealed that batch normalization produced better results than instance normalization. The studies [40,147] found that L2 normalization yielded the best results.

4.2 Post-processing

Post-processing is a strategy for addressing the validity of disaggregation results in order to improve NILM further. The article [65] presented an optimization-based strategy to ensure the summing of disaggregated loads is as near to the genuine aggregate consumption as possible. The authors of [66] discover that partial activation of neural networks impacts appliance power. As a result of this, the mean also affects the ground truth. To overcome this, Ref. [66] proposed to use median, which is relatively unaffected. To improve disaggregation results, Ahmed et al. [67] use Generative Adversarial Networks (GANs) technique.

4.3 Machine Learning-Based Approach

1. **Hidden Markov Model (HMM):**

   HMM is an example of unsupervised learning model and is widely used for the disaggregation of load having low-frequency data resolution. HMM has been well discovered in the literature [22]. The Hidden Markov Model is constructed with data preprocessing stage, then it calculates the hidden events and observed events using k-means clustering [54]. Hidden events identify appliance on/off state while the observed events associate with the energy consumption of each load. The transition matrix is used to identify the state transitions of the appliance. The various versions of HMM explored in NILM application are, Factorial HMM (FHMM), Conditional (FHMM), Conditional Factorial Hidden Semi-Markov Model (CFHSMM), Factorial Hidden Semi-Markov Model (FHSMM) [68].

Numerous algorithms have been developed that make use of various kinds of HMMs and have proven outstanding outcomes. The fundamental constraints of classical Markov models, on the other hand, have remained unsolved. Though, the fundamental constraints of classic Markov models remain unaddressed. Though several researches have been conducted employing HMM and its variants, it is observed that as the appliances on the power line increase, the time complexity exponentially increases. A limitation in the ability to classify multi-state appliances is a result of the fact that many Markov models are based on first-order Markov chains [69].

2. **Support Vector Machine (SVM)**

When it comes to machine learning, SVM has been one of the most powerful algorithms. Data extraction and classification based on identified patterns are advantages of this method [9]. To separate the samples, SVM uses either a linear kernel (which uses the features in original feature space) or a non-linear kernel (which uses features in higher-dimensional feature space) [70]. Since the data in this work are from the four CFL lamps in 16 potential electrical network topologies, the author constructed this as a 16-class problem.

In [71], the SVM learns about a specific electrical appliance's features. With good accuracy, the trained network identifies the specified electrical item and calculates the total household power used.

3. **Sparse Viterbi Algorithms**

Reference [72] proposed a novel algorithm that addresses the Viterbi algorithm's efficiency issue. The author demonstrates a strategy created on super-state Hidden Markov Model (SSHMM) with the Viterbi algorithm variation. In SSHMM, a super-state represents the power status of appliances, which can be either on or off. Each combination of appliances has its own super-state, which results in the disaggregation of appliances with complex multi-state power features.

4. **Decision Tree (DT):**

Decision tree-based NILM is a supervised technique with a modest level of complexity that can be trained with a small amount of labeled data. Decision trees are rule-based models that are simple and easy to visualize once they have been constructed. The difference between two successive active power measurements, referred to as ΔP, used as a training feature [73], furthermore, the system performance improved using active power (P) as additional feature[74].

5. **K-Nearest Neighbor (KNN):**

   Another type of supervised learning is KNN, in order to employ KNN dataset must be a labeled dataset. The value of K is determined based on the validation set, which contains 60% of the labeled data. In article [75], KNN demonstrated on the AMPds2 dataset and shows that KNN has potential to disaggregate appliances like, dishwasher and clothes dryer. The study achieved a classification accuracy of 95% by considering active and reactive power as an input feature, whereas the accuracy degraded to 73% by taking only active power as a feature. Study [73] describes a number of various strategies for pre-processing data in order to reduce the effects of noisy data. The KNN tested on two datasets, namely, REFIT and REDD.

4.4 Deep Learning Based

The majority of NILM systems are based on hand-engineered features taken from the aggregated power stream. Deep learning algorithms have demonstrated their capacity to solve numerous complicated problems in a variety of applications in recent years, including speech recognition, computer vision, and asset status monitoring, among others. Recently, researchers have been investigating deep learning methods such as recurrent neural network (RNN), convolution neural network (CNN), and autoencoder (AE) in the NILM problem to better classify appliances and disaggregate energy [70, 76]–[78]. It has been established that deep neural networks (DNN) can be used as a multi-class classifier for discriminating between different appliances using deep learning techniques [79].

1. **Recurrent Neural Network (RNN):**

   A neural network implementation that permits connections between neurons of the same layer is known as a recurrent neural network. Sequential data such as the readings of power usage in NILM are ideally suited for RNNs [78]. Article [80] proposed RNN in NILM to disaggregate the appliances. Author compares Combinatorial Optimization (CO). The results show that CO lags behind RNN in those cases, while RNN may operate well in unknown cases. However, significant progress must be made in order to improve the RNN performance for multi-state appliances in the near future.

2. **Convolution Neural Network (CNN):**

   For machine vision, image processing, and natural language processing, CNN has proven to be a very successful modeling system. Motivated by recent advancements of CNN, ref.[76] used the model to disaggregate appliances from total household data. When it comes to training the model, CNN has a significant advantage over other methods because it does not require hand-crafted features.

A supervised (CNN)-based approach was adopted in [81], which is trained on small subset of aggregated data. To enhance the disaggregation performance of CNN, author considered time of the day as an additional feature.

3. **Long short-term memory (LSTM):**

   LSTMs have been successfully applied to a number of sequence applications, such as automatic speech recognition and machine translation. To overcome vanishing gradient issue rises in RNN, ref.[76] adopted LSTM architecture, which employs a 'memory cell' with all gated input, output, and feedback loops. Multiple feature with four-layered bidirectional LSTM is adopted in [60]. The performance evaluation in this work showed that the MFS-LSTM method is more computationally efficient, scalable, and accurate in a noisy environment, as well as generalized to unforeseen loads, when compared to standard algorithms.

5 Performance Metrics

The efficacy of NILM algorithms is based on the outcome of the performance evaluation metrics. Numerous assessment metrics have been employed to assess the performance of event detection/classification and energy estimation, as well as to compare the findings. For the NILM system, first performance evaluation performed by G. W. Hart [3], in which a fraction of correct event identifications and a fraction of total energy consumed employed. The effectiveness of energy disaggregation is assessed by calculating the difference between the estimated and actual consumed energy. For low resolution energy disaggregation systems, many metrics related to estimated error, such as standard deviation of error (SDE), root mean square error (RMSE), average error (AE), energy error, and R-squared are commonly employed in [27]. As the details of performance measurements are outside the scope of this article, they can be found in [26] with mathematical expressions.

6 Discussion and Challenges

6.1 Performance Comparison

The NILM algorithm's performance is evaluated in a variety of ways. In the studied literature, MAE and F-score were the most commonly used metrics to measure predicted energy use and appliance on/off condition for low-resolution NILM. The results have been acquired by a variety of algorithms that are completely distinct. The following are the opinions associated with model performance comparisons:

- Approaches that are published should provide a set of standard metrics, set of assumptions and set of constraints.

- For model cross-validation, there should be a standardized evolution procedure for defining training and testing conditions.
- Authors should make their code publicly available; this will simplify the retraining of models for comparison with new ways.
- Trained models may be published too, as the computer vision community does in [82]. Only trained models from [83] have been made publicly available.

6.2 Multiple Features

Numerous authors made advantage of a variety of different input features. The literature [60, 84] present the findings of a comparison of multiple input features. The research article [85] specifically makes use of reactive power (Q). According to the authors, Q has been found to have an impact on the F1-score in both the AMPds and the UK-DALE datasets. They discover approximately 12.5%, a significant improvement in the seen evaluation situation. Furthermore, an improvement of approximately 8% in the unseen evaluation scenario across all of the investigated appliances.

The observed improvement is minor or negative for pure resistive loads, such as a kettle or an electric oven, which is unusual. Therefore, hypothesize made in such instances, reactive power does not give any information, but rather is only background noise. The features such as P, Q, I, S versus P, using distinct performance metrics, such as mean absolute error (MAE), normalized root mean square error (NRMSE) and the root mean square error (RMSE), examined in [86]. The benefits with the extra features are substantially larger in this work: roughly 40%–50% for all measures. Depending on the outcomes, conclusion made those additional features other than P can help disaggregation better. No judgments can be drawn about the amount of improvement due to the wide range of outcomes. It may be worthwhile to investigate what aspects, e.g., architectures, can best utilize information from attributes other than P.

Except for [85], all outcomes are from observed evaluation scenarios. This implies extra features help to estimate an appliance's power usage. The amount of accuracy they can provide to disaggregate type of appliance (Type I, II, III for details refer [12]) is unknown. It would be fascinating to look into a bigger feature set.

6.3 Data Scarcity

The biggest difficulty with applying NILM is the scarcity of labeled data. It is possible to adapt semi-supervised deep learning to low-frequency NILM, overcoming the data scarcity problem in practical applications.

In the Netherlands, Net2Grid is a company that assists power utilities with NILM compliance and management. In a demonstration, they emphasized the fact that high-quality data are needed for greater accuracy of NILM system [87]. They further pointed out that appliance with different program or different settings has different load patterns, therefore necessitating numerous observed cycles. The authors of [32] analyze low-frequency NILM approach with the implementation of deep neural network (DNN), in which disaggregation error depends on different households utilized for training. They investigate that disaggregation error falls continuously in proportion to the number of houses added to the training dataset till 40 houses. Therefore, both literatures show that complicated machines require a high diversity of training data to generalize fresh data successfully. Both studies conclude that sophisticated machines require a high degree of diversity in their training data in order to successfully generalize to previously unseen information. This finding is, at the fundamental level, known as 'data scarcity'.

6.4 Data from Three-Phase Appliances

The European countries like Switzerland, three-phase power supply arrives at main distribution board then it separates into single phases. However, multi-phase electric appliances like pool pumps, electrical storage heaters, heat pumps, and electric vehicles charging stations provide a significant challenge to energy conservation measures. In order to disaggregate information from all three phases, it is necessary to use the NILM algorithm. Datasets that contributed three-phase appliance data are: ECO[38] and iAWE [19]. NILM algorithm is required for these three-phase appliances in order to disaggregate information from all three stages. One of the most difficult challenges to overcome when developing a method that should work in every household is the fact that multi-phase equipment can be connected in any number of different ways. In order to be invariant to these permutations, the outcome of the low-frequency NILM technique must be consistent.

6.5 Prospect of NILM

Taking future speculation, one can imagine a variety of scenarios and possibilities for the NILM industry. Considering rapid growth of the Internet of Things (IOT), in the future, appliances may be programmed to be conscious of their own energy usage and able to communicate information to the outside world through their own communication interface. In order to have this, it is necessary to create a business case for appliance manufacturers and provide the groundwork for interfaces and protocols. The exponential growth in computing power of edge devices (gateway devices) will soon enable NILM close to the meter without transmitting data to a cloud service. In this circumstance, NILM algorithms can learn and improve on local data. In order to be success in NILM, the learning problem must first be phrased in such a way that the data from the meter may be used to create future improvements.

6.6 Indian Outlook Toward NILM

We are seeing changes in power generation, regulatory measures, and consumption patterns in India's power sector, which is undergoing rapid transformation. This environment allows Discoms (Distribution Company) to transition from being just an electricity supplier to being an energy service provider by lowering costs, engaging with customers to help them save energy, and improving the overall customer experience. The country's smart metering initiatives are perfectly timed to facilitate this transition. Discoms should look beyond metering, billing, and collection efficiencies to get the most out of their smart meter investments with utilizing NILM technology.

Utility companies can use granular data on household electricity usage to assess and design appropriate mechanisms to manage rising demand at the consumer level. Many households have insufficient information about their electricity usage because they tend to over- or under-estimate their appliance usage. Discoms could encourage consumers to use electricity carefully by providing daily or weekly information about their usage, as well as assist them in making the best appliance purchase decisions. For example, discoms could provide individualized advice to households about the potential savings from switching to a more efficient air conditioner. Routine feedback to consumers on their own consumption via their electricity bills or via mobile communication could also help to reduce consumption.

6.7 Conclusion

To summarize, this paper gives an overview of the literature on NILM with low-resolution smart meter data. NILM is being investigated because it has the potential to benefit a wide range of applications. This is coupled with the realization that low-resolution data will almost probably become readily accessible on a large scale in the near future. This study involved articles that use machine learning and deep learning approaches to separate appliances from aggregated low-frequency data. The study reviewed many degrees of flexibility offered by these approaches. The fundamental study of energy disaggregation is presented and is followed by lists of low-resolution datasets, shown in Table 1, which provides details of datasets in terms of location, recorded data at aggregated level as well as appliance level, duration of recorded data, number of houses, measurement parameters, etc.

Numerous difficulties were identified, related to data scarcity, model performance comparison, outlook toward NILM in Indian context, three-phase energy datasets for NILM and many more, these opinions are resulting in valuable conclusions and recommendations for future studies. Comparing multiple NILM systems is still time-consuming, although there are fresh methodologies and mathematical tools that have not yet been implemented. Although it is still missing in the current literature, this contribution may prove useful.

Acknowledgements The research work has been carried out under AICTE Doctoral fellowship (ADF) scheme, authors are thankful to the All-India Council for Technical Education (AICTE) for providing fellowship.

References

1. "Each Country's Share of CO2 Emissions | Union of Concerned Scientists." https://www.ucsusa.org/resources/each-countrys-share-co2-emissions (accessed May 30, 2021).
2. GROWTH OF ELECTRICITY SECTOR IN INDIA FROM 1947–2020, GOVERNMENT OF INDIA , MINISTRY OF POWER, CENTRAL ELECTRICITY AUTHORITY NEW DELHI, OCTOBER 2020
3. Hart GW (1992) Nonintrusive Appliance Load Monitoring. Proc IEEE 80(12):1870–1891
4. Abubakar I, Khalid SN, Mustafa MW, Shareef H, Mustapha M (2017) Application of load monitoring in appliances' energy management – A review. Renew Sustain Energy Rev 67:235–245
5. Kolter JZ, Jaakkola T (2012) Approximate inference in additive factorial HMMs with application to energy disaggregation. J Mach Learn Res 22:1472–1482
6. L. De Baets, C. Develder, T. Dhaene, and D. Deschrijver, "Detection of unidentified appliances in non-intrusive load monitoring using siamese neural networks," Int. J. Electr. Power Energy Syst., vol. 104, no. December 2017, pp. 645–653, 2019,
7. He K, Stankovic L, Liao J, Stankovic V (2018) Non-Intrusive Load Disaggregation Using Graph Signal Processing. IEEE Trans. Smart Grid 9(3):1739–1747
8. Gillis JM, Alshareef SM, Morsi WG (2016) Nonintrusive load monitoring using wavelet design and machine learning. IEEE Trans. Smart Grid 7(1):320–328
9. Basu K, Debusschere V, Douzal-Chouakria A, Bacha S (2015) Time series distance-based methods for non-intrusive load monitoring in residential buildings. Energy Build. 96:109–117
10. Zoha A, Gluhak A, Imran MA, Rajasegarar S (2012) Non-intrusive Load Monitoring approaches for disaggregated energy sensing: A survey. Sensors (Switzerland) 12(12):16838–16866
11. Zeifman M, Roth K (2011) Nonintrusive appliance load monitoring: Review and outlook. IEEE Trans Consum Electron 57(1):76–84
12. A. Faustine, N. H. Mvungi, S. Kaijage, and K. Michael, "A Survey on Non-Intrusive Load Monitoring Methodies and Techniques for Energy Disaggregation Problem," 2017,
13. B. Najafi, S. Moaveninejad, and F. Rinaldi, "Data Analytics for Energy Disaggregation: Methods and Applications," Big Data Appl. Power Syst., no. January, pp. 377–408, 2018,
14. J. Gao, S. Giri, E. C. Kara, and M. Bergés, "PLAID: A public dataset of high-resolution electrical appliance measurements for load identification research," BuildSys 2014 - Proc. 1st ACM Conf. Embed. Syst. Energy-Efficient Build., pp. 198–199, 2014,
15. Ruano A, Hernandez A, Ureña J, Ruano M, Garcia J (2019) NILM techniques for intelligent home energy management and ambient assisted living: A review. Energies 12(11):1–29
16. H. Rashid, P. Singh, V. Stankovic, and L. Stankovic, "Can non-intrusive load monitoring be used for identifying an appliance's anomalous behaviour?," Appl. Energy, vol. 238, no. August 2018, pp. 796–805, 2019,
17. Sun G, Cong Y, Hou D, Fan H, Xu X, Yu H (2019) Joint household characteristic prediction via smart meter data. IEEE Trans. Smart Grid 10(2):1834–1844
18. Buzau MM, Tejedor-Aguilera J, Cruz-Romero P, Gomez-Exposito A (2019) Detection of non-technical losses using smart meter data and supervised learning. IEEE Trans. Smart Grid 10(3):2661–2670
19. N. Batra, M. Gulati, A. Singh, and M. Srivastava, "It's Different: Insights into home energy consumption in India," Proc. 5th ACM Work. Embed. Syst. Energy-Efficient Build., no. August, pp. 1–8, 2013,

20. Wang Y, Chen Q, Hong T, Kang C (2019) Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. IEEE Trans. Smart Grid 10(3):3125–3148
21. R. Gopinath, M. Kumar, C. Prakash Chandra Joshua, and K. Srinivas, "Energy management using non-intrusive load monitoring techniques – State-of-the-art and future research directions," Sustain. Cities Soc., vol. 62, no. June, p. 102411, 2020,
22. H. Liu, Non-intrusive load monitoring: Theory, technologies and applications. 2019.
23. "about-smart-meters @ eeslindia.org."
24. L. Nilm, P. Huber, A. Calatroni, A. Rumsch, and A. Paice, "Review on Deep Neural Networks Applied to," 2021.
25. Gao J, Kara EC, Giri S, Berges M (2016) "A feasibility study of automated plug-load identification from high-frequency measurements", 2015 IEEE Glob. Conf. Signal Inf. Process. Glob. 2015:220–224
26. K. Basu, A. Hably, V. Debusschere, S. Bacha, G. J. Driven, and A. Ovalle, "A comparative study of low sampling non intrusive load dis-aggregation," IECON Proc. (Industrial Electron. Conf., pp. 5137–5142, 2016,
27. N. Batra et al., "NILMTK: An open source toolkit for non-intrusive load monitoring," e-Energy 2014 - Proc. 5th ACM Int. Conf. Futur. Energy Syst., pp. 265–276, 2014,
28. "index @ dataport.pecanstreet.org."
29. A. U. Haq and H. A. Jacobsen, "Prospects of appliance-level load monitoring in off-The-shelf energy monitors: A technical review," Energies, vol. 11, no. 1, 2018,
30. I. E. Monitor, "The intelligent home monitor .",
31. "98d12044bdf7c73befa6613fc7173486e3722fec @ www.smappee.com."
32. C. Shin, S. Rho, H. Lee, and W. Rhee, "Data requirements for applying machine learning to energy disaggregation," Energies, vol. 12, no. 9, 2019,
33. A. Reinhardt et al., "On the accuracy of appliance identification based on distributed load metering data," 2012 Sustain. Internet ICT Sustain. Sustain. 2012, no. October 2014, 2012.
34. S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht, "Smart*: An Open Data Set and Tools for Enabling Research in Sustainable Homes," SustKDD, no. August, p. 6, 2012,
35. J.-P. Zimmermann et al., "Household Electricity Survey: A study of domestic electrical product usage," Intertek, p. 600, 2012,
36. C. Holcomb, "Pecan Street Inc.: A Test-bed for NILM," Int. Work. Non-Intrusive Load Monit., pp. 271–288, 2007.
37. S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajić, "AMPds: A public dataset for load disaggregation and eco-feedback research," 2013 IEEE Electr. Power Energy Conf. EPEC 2013, no. Section III, 2013,
38. C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, and S. Santini, "The ECO data set and the performance of non-intrusive load monitoring algorithms," BuildSys 2014 - Proc. 1st ACM Conf. Embed. Syst. Energy-Efficient Build., pp. 80–89, 2014,
39. Monacchi A, Egarter D, Elmenreich W, D'Alessandro S, Tonello AM (2015) "GREEND: An energy consumption dataset of households in Italy and Austria", 2014 IEEE Int. Conf. Smart Grid Commun. SmartGridComm 2014(1):511–516
40. B. Larson et al., "Residential building stock assessment: Metering study," Northwest Energy Effic. Alliance, 2014,
41. N. Batra, M. Gulati, A. Singh, and M. B. Srivastava, "It's Different," no. May 2016, pp. 1–8, 2013,
42. A. S. N. Uttama Nambi, A. Reyes Lua, and V. R. Prasad, "LocED," no. ii, pp. 45–54, 2015,
43. Johnson G, Beausoleil-Morrison I (2017) Electrical-end-use data from 23 houses sampled each minute for simulating micro-generation systems. Appl Therm Eng 114:1449–1456
44. Makonin S, Wang ZJ, Tumpach C (2018) RAE: The rainforest automation energy dataset for smart grid meter data analysis. Data 3(1):1–9
45. C. Shin, E. Lee, J. Han, J. Yim, W. Rhee, and H. Lee, "The enertalk dataset, 15 hz electricity consumption data from 22 houses in Korea," Sci. Data, vol. 6, no. 1 1, pp. 1–13, 2019,
46. Rashid H, Singh P, Singh A (2019) Data descriptor: I-BLEND, a campus-scale commercial and residential buildings electrical energy dataset. Sci. Data 6:1–12

47. Pullinger M et al (2021) The IDEAL household energy dataset, electricity, gas, contextual sensor data and survey data for 255 UK homes. Sci. Data 8(1):1–18
48. Pipattanasomporn M et al (2020) CU-BEMS, smart building electricity consumption and indoor environmental sensor datasets. Sci. Data 7(1):1–14
49. M. S. Clark, “Improving the feasibility of energy disaggregation in very high- and low-rate sampling scenarios,” no. September, 2015,
50. Machlev R, Belikov J, Beck Y, Levron Y (2019) MO-NILM: A multi-objective evolutionary algorithm for NILM classification. Energy Build. 199:134–144
51. Norford LK, Leeb SB (1996) Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. Energy Build. 24(1):51–64
52. T. T. H. Le and H. Kim, “Non-intrusive load monitoring based on novel transient signal in household appliances with low sampling rate,” Energies, vol. 11, no. 12, 2018,
53. Welikala S, Dinesh C, Ekanayake MPB, Godaliyadda RI, Ekanayake J (2019) Incorporating Appliance Usage Patterns for Non-Intrusive Load Monitoring and Load Forecasting. IEEE Trans. Smart Grid 10(1):448–461
54. Kong W, Dong ZY, Ma J, Hill DJ, Zhao J, Luo F (2018) An Extensible Approach for Non-Intrusive Load Disaggregation with Smart Meter Data. IEEE Trans. Smart Grid 9(4):3362–3372
55. Zeifman M, Akers C, Roth K (2015) “Nonintrusive monitoring of miscellaneous and electronic loads”, 2015 IEEE Int. Conf. Consum. Electron. ICCE 2015:305–308
56. Drenker S, Kader A (1999) Nonintrusive monitoring of electric loads. IEEE Comput. Appl. Power 12(4):47–51
57. Ahmadi H, Marti JR (2015) Load Decomposition at Smart Meters Level Using Eigenloads Approach. IEEE Trans Power Syst 30(6):3425–3436
58. J. Z. Kolter, S. Batra, and A. Y. Ng, “Energy disaggregation via discriminative sparse coding,” Adv. Neural Inf. Process. Syst. 23 24th Annu. Conf. Neural Inf. Process. Syst. 2010, NIPS 2010, pp. 1–9, 2010.
59. Henao N, Agbossou K, Kelouwani S, Dube Y, Fournier M (2017) Approach in Nonintrusive Type i Load Monitoring Using Subtractive Clustering. IEEE Trans. Smart Grid 8(2):812–821
60. H. Rafiq, X. Shi, H. Zhang, H. Li, and M. K. Ochani, “A deep recurrent neural network for non-intrusive load monitoring based on multi-feature input space and post-processing,” Energies, vol. 13, no. 9, 2020,
61. K. S. Barsim and B. Yang, “On the Feasibility of Generic Deep Disaggregation for Single-Load Extraction,” pp. 1–5, 2018,
62. W. Kong, Z. Y. Dong, B. Wang, J. Zhao, and J. Huang, “A Practical Solution for Non-Intrusive Type II Load Monitoring based on Deep Learning and Post-Processing,” IEEE Trans. Smart Grid, vol. PP, no. c, pp. 1–1, 2019,
63. J. Huchtkoetter and A. Reinhardt, “On the Impact of Temporal Data Resolution on the Accuracy of Non-Intrusive Load Monitoring,” BuildSys 2020 - Proc. 7th ACM Int. Conf. Syst. Energy-Efficient Build. Cities, Transp., pp. 270–273, 2020,
64. G. Running, S. M. House, L. T. Street, and L. Ecr, “DEEP NEURAL NETWORKS FOR APPLIANCE TRANSIENT CLASSIFICATION Peter Davies, Jon Dennis, Jack Hansom, William Martin, Aistis Stankevicius and Lionel Ward,” pp. 8320–8324, 2019.
65. D. Jakovetic, V. Stankovic, and L. Stankovic, “Post-processing for Event-based Non-intrusive Load Monitoring,” pp. 2–5.
66. Bonfigli R, Felicetti A, Principi E, Fagiani M, Squartini S, Piazza F (2018) Denoising autoencoders for Non-Intrusive Load Monitoring: Improvements and comparative evaluation. Energy Build. 158(November):1461–1474
67. A. M. A. Ahmed, Y. Zhang, and F. Eliassen, “Generative Adversarial Networks and Transfer Learning for Non-Intrusive Load Monitoring in Smart Grids,” in 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), 2020, pp. 1–7.
68. H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han, “Unsupervised disaggregation of low frequency power measurements,” Proc. 11th SIAM Int. Conf. Data Mining, SDM 2011, pp. 747–758, 2011,

69. J. Kim, T. Le, and H. Kim, "Nonintrusive Load Monitoring Based on Advanced Deep Learning and Novel Signature," vol. 2017, 2017.
70. R. Gopinath, "Performance Analysis of Similar Appliances Identification using NILM Technique under Different Data Sampling Rates," 2020.
71. Moradzadeh A, Zeinal-Kheiri S, Mohammadi-Ivatloo B, Abapour M, Anvari-Moghaddam A (2020) "Support vector machine-assisted improvement residential load disaggregation", 2020 28th Iran. Conf. Electr. Eng. ICEE 2020:1–6
72. Makonin S, Popowich F, Bajic IV, Gill B, Bartram L (2016) Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring. IEEE Trans. Smart Grid 7(6):2575–2585
73. Liao J, Elafoudi G, Stankovic L, Stankovic V (2015) "Non-intrusive appliance load monitoring using low-resolution smart meter data", 2014 IEEE Int. Conf. Smart Grid Commun. SmartGridComm 2014:535–540
74. M. Khazaei, L. Stankovic, and V. Stankovic, "Evaluation of low-complexity supervised and unsupervised NILM methods and pre-processing for detection of multistate white goods," NILM 2020 - Proc. 5th Int. Work. Non-Intrusive Load Monit., no. i, pp. 34–38, 2020,
75. F. Hidiyanto and A. Halim, "KNN Methods with Varied K, Distance and Training Data to Disaggregate NILM with Similar Load Characteristic," in Proceedings of the 3rd Asia Pacific Conference on Research in Industrial and Systems Engineering 2020, 2020, pp. 93–99.
76. J. Kelly and W. Knottenbelt, "Neural NILM: Deep neural networks applied to energy disaggregation," BuildSys 2015 - Proc. 2nd ACM Int. Conf. Embed. Syst. Energy-Efficient Built, pp. 55–64, 2015,
77. Zhang J, Chen X, Ng WWY, Lai CS, Lai LL (2019) New Appliance Detection for Nonintrusive Load Monitoring. IEEE Trans. Ind. Informatics 15(8):4819–4829
78. N. Batra et al., "Towards reproducible state-of-the-art energy disaggregation," pp. 193–202, 2019,
79. Devlin MA, Hayes BP (2019) Non-Intrusive Load Monitoring and Classification of Activities of Daily Living Using Residential Smart Meter Data. IEEE Trans Consum Electron 65(3):339–348
80. Linh NV, Arboleya P (2019) "Deep learning application to non-intrusive load monitoring", 2019 IEEE Milan PowerTech. PowerTech 2019:1–5
81. B. Zhao, M. Ye, L. Stankovic, and V. Stankovic, "Non-intrusive load disaggregation solutions for very low-rate smart meter data," Appl. Energy, vol. 268, no. October 2019, p. 114949, 2020,
82. "TensorFlow Hub." https://www.tensorflow.org/hub (accessed Jul. 15, 2021).
83. Kaselimi M, Doulamis N, Voulodimos A, Protopapadakis E, Doulamis A (2020) Context Aware Energy Disaggregation Using Adaptive Bidirectional LSTM Models. IEEE Trans. Smart Grid 11(4):3054–3067
84. A. Harell, S. Makonin, and I. V. Bajic, "Wavenilm: A Causal Neural Network for Power Disaggregation from the Complex Power Signal," ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc., vol. 2019-May, pp. 8335–8339, 2019,
85. Valenti M, Bonfigli R, Principi E (2018) Squartini, and Stefano, "Exploiting the Reactive Power in Deep Neural Models for Non-Intrusive Load Monitoring", in. International Joint Conference on Neural Networks (IJCNN) 2018:1–8
86. Kaselimi M, Protopapadakis E, Voulodimos A, Doulamis N, Doulamis A (2019) Multi-Channel Recurrent Convolutional Neural Networks for Energy Disaggregation. IEEE Access 7:81047–81056
87. "EU NILM 2019 - Dimitrios Doukas - YouTube." https://www.youtube.com/watch?v=v5XoLtQH9Uw&list=PLJrF-gxa0ImryGeNtil-s9zPJOaV4w-Vy&index=23 (accessed Jul. 15, 2021).