Citation Information: 10.1016/j.rser.2023.114151

A Review of Distribution Network Applications Based on Smart Meter Data Analytics

Christos L. Athanasiadis[a,b], Theofilos A. Papadopoulos[a,*], Georgios C. Kryonidis[c] and Dimitrios I. Doukas[b]

[a]Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi,, 67100, Greece
[b]NET2GRID BV, Thessaloniki, Greece
[c]School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece

ARTICLE INFO

Keywords:
Asset management
data analytics
distribution networks
load forecasting
non-technical losses
power system operation and analysis
power system planning
smart meters
topology identification

ABSTRACT

The large-scale roll-out of smart meters allows the collection of a vast amount of fine-grained electricity consumption data. Once analyzed, such data can enable cutting-edge data-driven services to enhance power systems efficiency and sustainability. In this paper, a comprehensive literature overview of the state-of-the-art distribution network-oriented applications employing smart meter data is conducted and potential areas for future research are identified. The most recent innovations are outlined and discussed with an emphasis on six key areas, namely load forecasting, non-technical losses, asset management, power system planning, topology identification, and power system operational analysis. It is anticipated that energy retailers, service providers and distribution system operators would find the taxonomy and related applications, as assessed and presented in this study, helpful in identifying emerging technology trends regarding smart meter data analytics.
Word count: 6018

Highlights

- Comprehensive overview of distribution network applications employing smart meter data
- Literature review has been categorized into six key areas
- Discussion on future research areas

Nomenclature

AI	Artificial Intelligence
AMI	Advanced Metering Infrastructure
AMR	Automated Meter Reading
ANN	Artificial Neural Network
BES	Battery Energy Storage
CNN	Convolutional Neural Network
DL	Deep Learning
DN	Distribution Network
DNN	Deep Neural Network

*Corresponding author, thpapad@ee.duth.gr
cathanas@ee.duth.gr (C.L. Athanasiadis); thpapad@ee.duth.gr (T.A. Papadopoulos); kryonidi@ece.auth.gr (G.C. Kryonidis); dimitrios@net2grid.com (D.I. Doukas)
ORCID(s):

DR	Demand Response
DRESs	Distributed Renewable Energy Sources
DSM	Demand Side Management
DSO	Distribution System Operator
DT	Decision Tree
EV	Electric Vehicle
GIS	Geographic Information System
HEMS	Home Energy Management System
HIF	High Impedance Fault
HVAC	Heating, Ventilation, and Air-Conditioning
IWF	Inter-turn Winding Fault
LV	Low-Voltage
MIP	Mixed-Integer Programming
ML	Machine Learning
MV	Medium-Voltage
NILM	Non-Intrusive Load Monitoring
NTLs	Non-Technical Losses
OLTC	On-Load Tap Change
PCA	Principal Component Analysis
PV	Photovoltaic
RTUs	Remote Terminal Units
SM	Smart Meter
SVM	Support Vector Machine
TB	Technical Brochure
TLs	Technical Losses
WLS	Weighted Least Squares

1. Introduction

Digitalization is nowadays emerging as a necessity for energy utilities and companies globally. One of the most important milestones of this transformation is the widespread adoption of smart meters (SMs). At its first step, SM technology was limited to automated readings by the energy provider in order to simplify billing and lower the labor expenses associated with on-site visits, similar to the preceding metering system, i.e., the automated meter reading (AMR) technology. However, the vast amount of fine-grained SM data provides numerous benefits for all energy stakeholders, e.g., distribution system operators (DSOs), retailers, consumers, and aggregators, by paving the way for new energy services and data-driven business models.

A detailed overview of the current status and functions of SMs is presented in the technical brochure (TB) "Utilization of data from smart meter system" [1]. This TB presents the results of a survey conducted by CIGRE focusing on two key points. The first objective was to draw the general outline of the SM systems and their specifications in various geographical regions. This includes SM measuring capabilities such as type of data (e.g., active/reactive power, energy, voltage, current, etc.), measurement period, and transmission period as summarized in Table 1. The second objective was to examine possible applications for the usage of SM data analytics to facilitate and enhance the operation of distribution networks (DN); these included the reduction of labor costs, improvement of the accuracy of energy measurements, reduction of non-technical losses (NTLs), and optimization of network operation and maintenance.

In the literature, there is a number of research activities regarding the global SM deployment, seeking to exploit and create added value from the collected data through SM data analytics. Detailed analyses and comprehensive comparisons of these solutions are presented in relevant review articles [2, 3, 4]. Specifically, in [2], a consumer-centric perspective of SM analytics is discussed focusing on energy consumption awareness, ambient assisted living, consumption anomalies, load profiling, and demand-side flexibility. Similarly, in [3], end-user-oriented applications, such as energy efficiency and home energy management systems (HEMS) are presented. Finally, in [4], an extensive review of various SM applications is presented ranging from data-related concerns (cleaning and imputation of missing data, data privacy, and compression) to customer characterization for personalized services. It is worth mentioning that in [3] and [4], the potential exploitation of SM data for DN-oriented applications is also discussed. Nevertheless, therein analysis considers a limited taxonomy of DN-oriented applications focusing mainly on NTLs detection, demand forecasting, and demand side management (DSM).

From the above, it can be seen that these review articles are mostly related to SM applications for end-users and/or consumers, without thoroughly assessing the newly emerging research studies on DN-oriented applications. To this end, this paper aims to conduct a comprehensive overview regarding the applications enabled by using SM data explicitly from the DN perspective. Besides commonly reviewed topics, e.g., forecasting and NTLs detection, additional applications, such as impedance estimation, phase grouping, remote switching, and hosting capacity are examined, among others.

In particular, over 110 research papers have been examined; the number of papers published in various mediums is summarized in Fig. 1. In addition, in Fig. 2, the number of examined papers published per year is depicted.

Based on the outcome of the literature review, the various applications are grouped into six key categories, namely load forecasting, NTLs detection, asset management, power system planning, topology identification, and power system operation and analysis. These categories along with their corresponding sub-categories are presented in Fig. 3 and thoroughly discussed in the following sections. A red asterisk beside an application indicates that a complex data-driven model is required. On the contrary, no asterisk implies raw SM data calculations. A brief summary of the examined literature per sub-category is analysed in Table 2.

The paper is structured as follows. In Section 2, applications regarding various types of forecasting, i.e., demand, peak demand, and flexibility, are presented. Section 3 focuses on NTLs detection and, specifically, electricity theft, whereas in Section 4, applications for asset management, such as outage management, remote switching, and fault detection, are described. Section 5 discusses hosting capacity and operating envelopes as part of long-term and short-term power system planning, and in Section 6, applications for topology identification by means of SM data are presented. Section 7 describes the utilization of SM data in power system operation and analysis and, finally, Section 8, discusses future research areas and concludes the paper.

Table 1
SM measurement capabilities per region [1].

Region	Measurement	Measurement period	Transmission period
Japan	Wh	30 min	30 min
China, South Korea	W, Wh, VAR, VARh, V, I	15 min	15 min
Northern-Western Europe	W, Wh, VAR, VARh, V, I	1 s - 1 h	6 h, 1 d
Southern-Eastern Europe	W, Wh, VAR, VARh, V, I	1 s - 1 d	15 min, 1 d, 1 mo
Brazil, South Africa	-	15 min, 30 min	1 h - 1 d

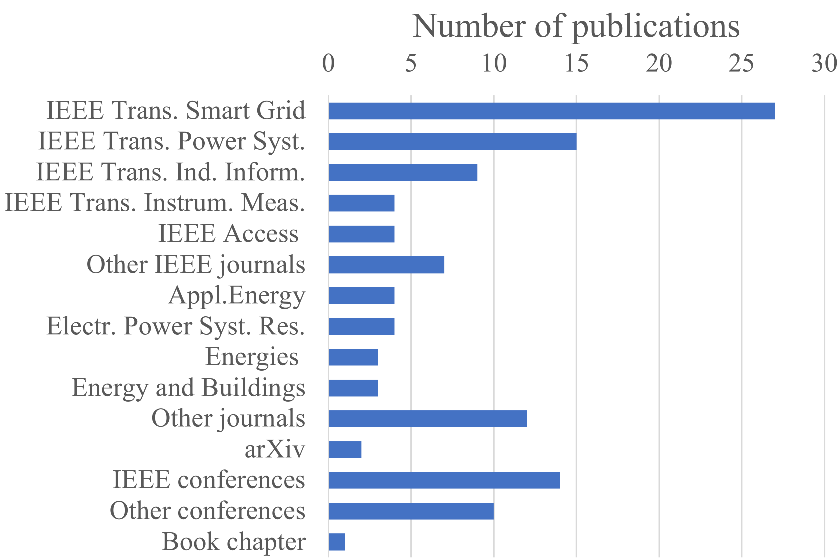

Figure 1: Number of publications per journal/conference.

2. Load Forecasting

Load forecasting has gained significant attention in recent years as the energy sector has been highly affected by various factors. In many countries, the COVID-19 pandemic led to reduction in energy demand, particularly in services and industry. Furthermore, the reduced gas deliveries of 2022 have developed a gas supply crisis mainly in European gas markets, causing them to diversify their source of energy imports.

Moreover, the unprecedented data availability, e.g., from SMs, remote terminal units (RTUs), and many other measurement systems, and the great progress of artificial intelligence (AI) and machine learning (ML) paved the way for advanced data analytics and techniques for load forecasting. In this section, three different types of forecasting are discussed, namely demand forecasting, peak demand forecasting, and demand flexibility forecasting. The typical data requirements for this category of applications are summarized in Table 3.

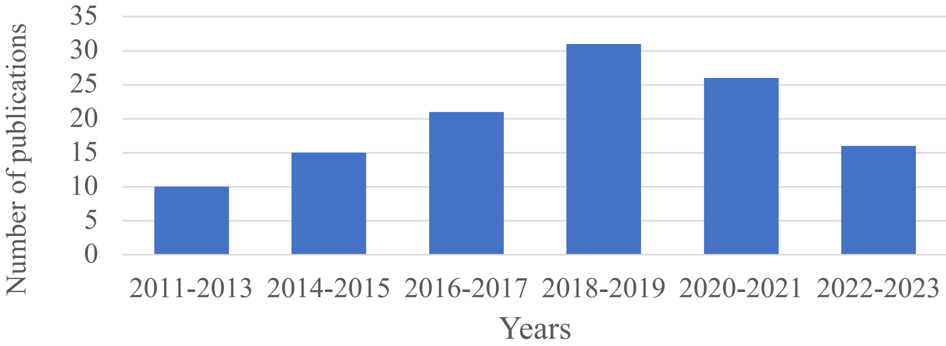

Figure 2: Number of publications per year.

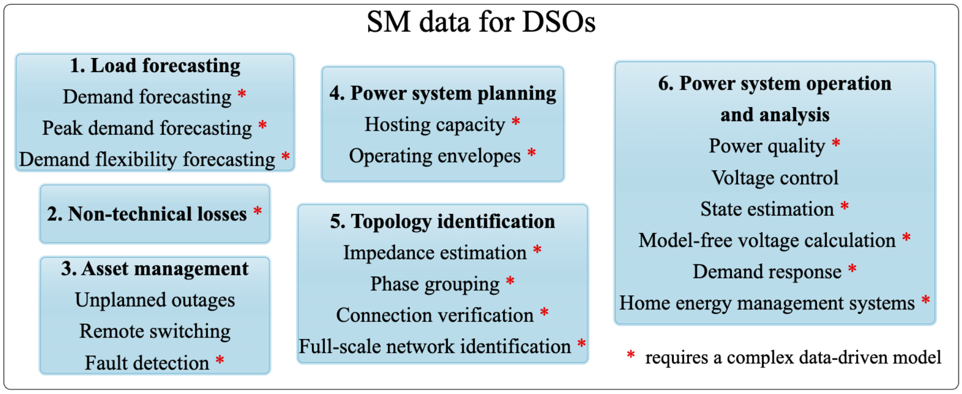

Figure 3: Taxonomy of SM data analytics for DSOs.

2.1. Demand Forecasting

Demand forecasting is a useful tool in modern power systems to achieve supply-demand equilibrium, support operations and planning processes and improve retailers decision-making about pricing, procurement, and hedging [4].

SMs may assist DSOs and retailers to better understand and forecast the load of an individual house or building [4] and can also improve forecasting accuracy at an aggregated level. However, forecasting at a more granular level, e.g., household, building, or neighborhood, is more challenging as the load timeseries are more volatile.

2.1.1. Household Level

SM data enable forecasting of individual houses or buildings allowing the extraction of useful information to improve the aggregated forecast. Sparse coding [5], ML and deep learning (DL) [6, 7, 8, 9] approaches have been investigated considering behavioral patterns [10] and socio-economic factors [11], e.g., number and age of occupants, employment status, etc. The concept of "similarity" is widely used in many works, identifying similar days in historical data of a household and using them to boost the accuracy based on the observation that individuals express their patterns of energy consumption behavior at different times on different days [12, 13]. Similarly, spatio-temporal approaches exploiting trends and interactions between data from a target house and the surrounding houses have also been proposed [14]. Finally, bottom-up approaches for household-level forecasting have also been tested aiming to predict the consumption of each appliance and aggregate the results [15, 16].

Table 2
Summary of the Examined Literature.

Category	Sub-category		References
Load forecasting	Demand forecasting	Building level	[5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]
		Aggregated level	[17, 18, 19, 20]
	Peak demand forecasting	Building level	[21, 22, 23, 24]
		Aggregated level	[25]
	Demand flexibility forecasting		[26]
NTLs	Data-oriented		[27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37]
	Network-oriented		[38, 39, 40, 41, 42, 43]
	Hybrid		[44, 45, 46, 47]
Asset management	Unplanned outage management		[48, 49, 50, 51, 52]
	Remote switching		[53]
	Fault detection		[54, 55]
Planning	Hosting capacity		[56, 57, 58, 59]
	Operating envelopes		[60, 61, 62]
Topology identification	Impedance estimation		[63, 64, 65, 66]
	Phase grouping	Mixed-integer programming	[67, 68, 69]
		Voltage-based	[70, 71, 72, 73, 74, 75, 76, 77, 78]
		ML power-based	[79, 80, 81, 82]
	Connection verification		[83, 84, 85, 86, 87]
	Full-scale network identification		[88, 89, 90, 91, 92, 93, 94]
Power system operation and analysis	Power quality		[95, 96, 97, 98]
	Voltage control		[99, 100, 101]
	State estimation		[102, 103, 104, 105, 106, 107, 108]
	Model-free voltage calculations		[109, 110]
	Demand response (DR) and HEMS	Residential	[111, 112, 113, 114, 115, 116, 117, 118]
		Microgrid	[119, 120, 121, 122]

2.1.2. Aggregated Level

Aggregated forecasting on the substation level has been extensively researched in the literature [17]. Fine-grained SM and demographic data can be utilized to improve the forecasts of aggregated models by extracting valuable information from clusters of households with similar characteristics. These characteristics may be demographics, appliance information, household information [18], and energy consumption [19, 20]. Once the clusters are formed, cluster-specific forecasting models are employed using SM consumption, weather, and calendar data. Eventually, the forecasts are aggregated to form the substation-level forecast.

2.2. Peak Demand Forecasting

Another noteworthy type of forecasting is peak demand, i.e., predicting the level and time of the highest demand. Accurate peak load estimation can play a pivotal role in planning studies and is the key driver for determining the capacity of electric power delivery equipment, such as substations and feeders [25]. Moreover, energy providers can properly schedule their production, ensure energy balance by optimally utilizing the costly and non-renewable peak

Table 3
Data Requirements for Forecasting.

	Demand Forecasting	Peak Demand Forecasting	Flexibility Forecasting
Input	Historical SM data	Historical SM data	Aggregated forecast of active and reactive power
Output	Demand forecasts	Peak demand forecasts	Forecasts of aggregated flexible active power
Sampling period	30 min	30 min	1 min or higher
Other (non-SM) input	Weather, end-user metadata	Weather, end-user metadata	Sub-measurements of appliances

load power plants, avoid grid congestion during peak hours, and ensure the economic benefits and stability of the power grid [123].

Building-level peak forecasting based on SM data has been examined using ML [21, 22] and fuzzy logic [23, 24] methods. Moreover, various approaches have been proposed in the literature for estimating the aggregated coincident peak demand [123]. However, the use of SM data has not been thoroughly examined. A data-driven probabilistic peak demand estimation framework using SM and sociodemographic data is developed in [25] performing customer clustering.

2.3. Demand Flexibility Forecasting

With the evolution of the smart grid, DR has been envisioned as one of the potentially cost-effective options for operating the power system. SMs unlock new opportunities for DR. For example, high-granular SM data can provide advanced end-user behavior profiling with regard to the usage of individual appliances [26].

In this context, the DR potential, i.e., the flexibility of an aggregated group of residential users is quantified and analyzed in [26]. It is assumed that several households own SMs with sub-metering capabilities, i.e., measurements of specific appliances are available. All monitored appliances are grouped into controllable and uncontrollable loads; controllable loads pertain to space/water heating and laundry activities. An artificial neural network (ANN) is trained to calculate the participation of controllable loads in the total aggregated consumption of the monitored households. Eventually, the trained model is used for the day-ahead forecast of controllable loads projected to the whole customer base.

It should be noted that the SM sub-metering capability can also be replaced by advanced analytics, such as non-intrusive load monitoring (NILM) [124, 125] which allows to break down the electricity consumption on an appliance level by analyzing SM measurements.

3. Non-technical Losses

In power systems, losses can be generally categorized into technical losses (TLs) and NTLs. TLs are the expected losses attributed to cables, overhead lines, transformers, and other substation equipment that is used to transfer electricity during the operation of a power system. NTLs correspond to the unaccounted energy that is neither measured nor allocated to TLs and may arise from electricity theft, measurement errors, metering faults, etc [126]. Furthermore, the term pertains to any action that causes incorrect billing, such as unlicensed distributed renewable energy sources (DRESs), and violations of licensed capacity. The most common source of NTLs is electricity theft, i.e, malicious measurement manipulation by consumers or other parties.

The impact of NTLs on the optimal operation, management, and cost of DNs is significant. It has been reported that NTLs may reach even 50% of the estimated consumption, and the overall cost of electricity theft worldwide reaches more than $25 billion every year [47]. In this context, NTLs detection is of utmost importance and can be classified into data-oriented, network-oriented, and hybrid [127]. The typical data requirements are presented in Table 4, although in practice requirements may vary depending on the approach.

Table 4
Data Requirements for NTLs detection.

Input	Historical SM data, voltage measurements
Output	Customers with NTLs
Sampling period	15 min or higher, even daily
Other (non-SM) input	Network topology

3.1. Data-oriented Methods

Data-oriented methods utilize SM and end-user-related data, e.g., personal, spatial, or financial information. They are divided into supervised and unsupervised based on the existence of labels (known positive/fraud and negative/non-fraud classes) or not, respectively, and are usually applied to the end-user level, characterizing each one independently as fraudulent or not. This is achieved by taking into account various features such as maximum/minimum power, contracted power, etc.

The most common supervised approaches rely on ML, e.g., support vector machines (SVMs) [27], ANNs [28], decision trees (DTs) [29], and fuzzy logic [30]. Advanced DL models have also been proposed, such as convolutional neural networks (CNNs) [31, 32], and recurrent neural networks [33, 34]. Although most research works consider NTLs solely in the consumption domain, i.e., reduce the observed energy consumption, there are cases of cyberattacks in distributed generation domain as well. In these cases, cyberattacks aim to manipulate SM and increase the reported generated energy. In [35], regressor trees are developed to detect such attacks using solar irradiance, temperature, and SM readings.

Additionally, unsupervised methods have been proposed that do not require labeled samples presenting inferior performance compared to supervised. Such approaches rely on game theory [36] and the expertise of inspectors [37].

3.2. Network-oriented Methods

Besides SM data, network-oriented methods utilize additional DN information, e.g., topology and additional measurements from RTUs and observer meters, i.e., meters on the secondary side of the distribution transformer. These methods are based on power flow analysis, state estimation, and sensor placement.

In power flow methods, TLs are initially calculated and subsequently NTLs are estimated by subtracting TLs from the total losses [38, 39]. However, estimating TLs requires knowledge of the network topology and cable impedances, which might not be readily available. For example, in [38, 40], additional algorithmic solutions are needed to estimate the cable impedances before calculating TLs. State estimation approaches [41, 42] have been applied to calculate the loading of distribution transformers from three-phase voltage, current, active, and reactive power measurements. In case of a mismatch between measured and estimated values, NTLs may be considered. The use of dedicated sensors for detecting fraud is also proposed in [43].

3.3. Hybrid Methods

Hybrid methods adopt a combination of data- and network-oriented approaches. For example, in [44, 45], TLs are calculated through power flow, and in case of high mismatch between produced and consumed energy, SVMs and DTs are used to detect fraudulent customers. In [46], SM data, wavelet-based feature extraction, and fuzzy c-means along with extra observer meter data are utilized to detect electricity theft. In [47], three modules are combined for NTLs detection. The first one uses endogenous features extracted from SM data to train an SVM for fraud detection. The second module calculates the voltage self-sensitivities from SM measurements and compares them to their theoretical values extracted from the network topology and power flow analysis. Finally, the third module solves an optimization problem aiming to minimize losses for detecting the location, extent, and time of NTLs per consumer.

Table 5
Data Requirements for Asset Management.

	Unplanned Outage Management	Remote Switching	Fault Detection
Input	Event-triggered energized states of SMs	Event-triggered energized states of SMs	Harmonics (HIFs), active/reactive power (IWFs)
Output	Outage/restoration events	Switching actions	Detected faults
Sampling period	Event-triggered	Event-triggered	kHz (HIFs), 15 min (IWFs)
Other (non-SM) input	Network topology	Network topology	Network topology

4. Asset Management

Asset management is one of the most important chapters in the operation of power systems. Poor asset management can lead to increased costs and unreliability. In this section, three groups of asset management applications using SM data analysis are discussed and typical data requirements are presented in Table 5.

4.1. Unplanned Outage Management

An unplanned power outage is defined as an electricity supply failure caused by short circuits, station failure, or distribution line damage [4]. Outage management is the most significant SM data application behind billing as SMs enable automatic outage notifications by last-gasp messages without the need of end-user calls. Furthermore, SMs allow outage confirmation by checking if other SMs in the area operate normally. Such application is crucial since it has been reported that about 70% of trouble calls are single service outages [128]. Similarly, SMs allow restoration verification to ensure the proper operation of the system after an outage. In these cases, on-demand SM states and location of each SM are required.

Several works in the literature focus on identifying the outage location. The basic idea in [48] is the usage of multiple SMs in a neighborhood. For a single service outage, neighbor SMs should operate normally, in contrast to a mass power outage. In [49], given the DN tree structure, an outage detection method is developed by combining the use of real-time power flow measurements on the edges with load forecasts at the nodes. The authors in [50] use the outage reports from SMs as input for a multiple-hypotheses method to quickly determine the most credible outage scenario based on an integer programming optimization model. In [51], a hierarchical framework is proposed for multi-level anomaly detection, e.g., momentary faults, transient and temporary faults, short outages, as well as long-standing faults and outages that last for many days. The framework can efficiently integrate the large-scale data collected from SMs at the customers' premises and transform them into actionable real-time insights with regard to the anomaly of interest and its severity.

The large-scale penetration of DRESs in DNs is considered in [52]. Outage detection methods relying on end-users' reports and SM last-gasp signals present poor performance since DRESs provide power even during an outage. To this end, a data-driven outage monitoring approach is proposed based on the hypothesis that voltage measurements exhibit significant statistical changes after outages.

4.2. Remote Switching

Besides remote reading, SM can offer to DSOs the possibility of remote switching. By sending a remote signal to the breaker of the SM, DSOs can connect/disconnect end-users from the grid [53]. This feature can be cost-efficient reducing labor costs. For example, the DSO can remotely disconnect end-users that have delayed their payments or do not have any contract and reconnect them as problems have been resolved. Moreover, it is feasible to disconnect selected end-users during peak load crises to avoid overloading of the network lines. Remote switching can also be applied at the end user's request, e.g., when moving to a new apartment.

Table 6
Data Requirements for Power System Planning.

	Hosting Capacity	Operating Envelopes
Input	Historical data of active/reactive power for all end-users	Active/reactive power for all end-users
Output	Hosting capacity of PVs and EVs	Operating envelopes for end-users engaged with the aggregator
Sampling period	15 min or higher	5 min or higher
Other (non-SM) input	Network topology	Network topology, voltage at the head of the feeder

4.3. Fault Detection

SM data can be used to detect and locate specific types of faults, e.g., high-impedance faults (HIFs). In [54], a HIF detection method is developed based on the even harmonics contained in the high-frequency voltage data of SMs. The authors tested the proposed algorithm assuming both partial and total SM penetration in the DN.

Inter-turn winding faults (IWFs) in single-phase distribution transformers are investigated in [55]. Instead of measuring the transformer secondary voltage by sensors, the proposed method uses SM measurements. Results obtained from simulations as well as experimental data show that such measurements can be utilized to achieve very high detection accuracy while maintaining low costs.

5. Power System Planning

Power system planning pertains to the development and design of the system and its elements aiming to satisfy future needs. Proper planning in developing countries has become more difficult due to the ever-increasing penetration of DRESs and electric vehicles (EVs) posing unprecedented technical challenges and jeopardizing the reliable operation of power systems. SM can play a pivotal role in aiding DSOs to prepare for future challenges and plan their network development. In this section, works focused on PV/EV hosting capacity and operating envelopes [62] are presented. Typical data requirements for this category are presented in Table 6.

5.1. Hosting Capacity

As solar PV penetration continues to grow, technical challenges, such as overvoltage and congestion are expected to occur. To this end, approaches to estimate PV penetration limits for the long-term planning of the power system have been investigated [56]. The extent to which low-voltage (LV) DNs can host solar PV is the hosting capacity.

In [56], a SM-driven method is introduced for the fast estimation of the hosting capacity without requiring complex and detailed network studies. Using SM data, a regression model is trained to estimate the PV capacity that can be hosted without causing voltages outside an upper limit. In [57, 58], probabilistic tools are used to perform power flow analyses for possible future PV integration scenarios. As long as operational constraints are not exceeded, more PV units can be added to the DN.

Besides DRESs, EV hosting capacity has also been investigated since the expected increase in peak demand can compromise the network integrity and pose significant technical challenges, such as asset congestion or voltage drop issues [59]. The EV hosting capacity is assessed in [59] by exploring multiple EV scenarios and considering their time-varying behavior during the peak demand day.

5.2. Operating Envelopes

The high penetration of residential DRESs in DNs has enabled households to provide bottom-up services through aggregators. The use of operating envelopes, i.e., individualized, time-varying import/export limits, has been proposed to facilitate such services better while ensuring network integrity and allowing more efficient short-term planning of the DN [60]. The envelopes are calculated by the DSOs using key network information, such as network model, head of

Table 7
Data Requirements for Topology Identification.

	Impedance Estimation	Phase Grouping	Connection Verification / Full-scale Network Identification
Input	Active/reactive power and voltage measurements for all end-users	Active power and voltage measurements for all end-users	Active/reactive power and voltage measurements for all end-users
Output	Line impedance	Phase connectivity of end-users	Network topology
Sampling period	15 min or higher	15 min or higher	15 min or higher
Other (non-SM) input	Network topology, measurements at the head of feeder	Measurements at the head of feeder	Partial topology, measurements from extra devices

feeder voltage, and net demand of end-users engaged with the aggregator. Once calculated, these operating envelopes are broadcasted to DRES aggregators, which use them as constraints when managing their portfolio.

In [61], the authors explore the use of operating envelopes calculated by the distribution company using a three-phase optimal power flow-based algorithm. The voltage regulation capability of the substation on-load tap changer (OLTC) is also used to enhance the operating envelopes. The work of [62] proposes a framework for operating envelopes in the presence of prosumers. The prosumer's intended operation is periodically submitted and power flow analysis is performed by the DSO to check for possible operational violations.

6. Topology Identification

Knowledge of the DN topology and parameters, e.g., line impedance and phase connection, is necessary for the thorough analysis and application of control schemes. However, such information is sometimes not known or inaccurate in DNs. Various methods utilizing SM data have been proposed to tackle similar issues. Typical data requirements are presented in Table 7.

6.1. Impedance Estimation

To enhance the observability of DNs, impedance estimation using SM measurements of active/reactive power at each node has gained significant attention assuming a known DN topology [63]. Several methods have been developed including particle swarm optimization [64], non-linear and non-convex optimization [65], and multi-linear regression [66].

6.2. Phase Grouping

Phase grouping is the process of determining the phase connection of end-users. Several data-driven approaches based on SM data have been introduced [129, 130] and can be classified into mixed-integer programming (MIP), voltage-based and ML power-based approaches.

6.2.1. MIP Approaches

In MIP approaches, optimization problems are formed using SM measurements and the distribution transformer supply. On the basis of the law of conservation of power, the connection phase of each end-user is determined, as the load measured at the feeder level must be equal to the aggregated consumption of all SMs connected to that feeder plus the unmetered load, e.g., street lights, and TLs. The optimization aims to minimize the difference between the total feeder demand and the transformer supply as presented in [67, 68]. In [69], both active and reactive power as well as the connection of PV units are taken into consideration. Such approaches are inferior when dealing with missing data and assume that all end-users have SMs installed and that there are no NTLs.

6.2.2. Voltage-based Approaches

Voltage measurements are also used for phase grouping. In [70, 71], the Pearson correlation between an end-user's voltage time series and a reference voltage time series (voltage of the transformer) is calculated; the end-user is assigned to the reference phase with the highest correlation. The advantage of such a method is that it is less error-prone due to missing data. A similar approach is followed in [72], but instead of using measurements at the substation downstream, the voltage of a three-phase end-user is used as reference. This avoids using additional measuring devices at the transformer level, but the connection of a three-phase end-user must be known.

Clustering approaches have also been adopted aiming to form groups of end-users connected to the same phase by using voltage measurements from SMs and applying k-means [73, 74], and spectral clustering [75, 76]. Spectral clustering combined with a sliding window ensemble approach is also proposed in [77] to handle missing data and deal with seasonality. Finally, in [78], principal component analysis (PCA) along with k-means is used focusing on DNs with high PV penetration.

6.2.3. ML Power-based Approaches

Finally, some approaches use power measurements from SMs in combination with ML techniques. Such methods can handle missing data, thus can mitigate the disadvantages of MIP methods. Moreover, power measurements are typically available compared to voltage since they are used for billing purposes. In [79, 80], PCA along with graph theory and power conservation have been utilized. In [81], the phase connectivity is identified by means of a modified k-means clustering algorithm and the correlation between the consumption time series and the aggregated consumption of each phase at the substation. In [82], spectral and saliency analysis is performed to extract features from the SM power measurements. Subsequently, correlation analysis between end-user features and phase features at the substation is used to determine phase connectivity.

6.3. Connection Verification

This topic concerns connection verification and detection of reconfigurations based on known topology information obtained through the geographic information system (GIS).

A MIP-based topology identification model is proposed in [83] to determine the topology configuration with weighted least squares (WLS) using active power, reactive power, and voltage measurements at each node. In [84], a generalized state estimation approach for the identification of topology changes is proposed. In [85], an algorithm for correcting connectivity errors in the GIS representation of the DN is developed that leverages SM measurements. This algorithm is based on voltage correlation to identify neighboring meters and predict end-users' upstream and downstream locations. The task of topology verification is posed as maximum-likelihood and maximum a-posteriori probability detection problem in [86]. Similarly, in [87], the basic topology information is obtained through GIS, and the states of unmonitored switches are identified based on a two-stage topology identification framework. A split expectation-maximization method is proposed for the topology identification problem on the historical batch data and classifiers, such as DTs, SVMs, and ANNs, are trained to predict the real-time topology efficiently.

6.4. Full-scale Network Identification

Regarding the estimation of the overall distribution grid topology linear-coupled power flow models [88, 89] and graph theory [90] have been widely used based on the assumptions of a radial DN structure and SM availability only at terminal nodes (end-users). In [91], a latent tree model is proposed to provide probabilistic representation for all possible topologies based on SM data, and the Bayesian information criterion is used to find the optimal topology model. The Markov random field method is implemented in [92] to perform a nodal correlation analysis using data from RTUs and SMs at end-users premises and an iterative screening method is developed to generate the DN topology.

In [93], only historical voltage measurements of SMs are used and a probabilistic graphical model is utilized to capture the statistical dependencies amongst bus voltages. The bus connectivity and grid topology estimation problems, in radial and mesh structures are formulated as a linear regression with a least absolute shrinkage regularization on grouped variables. Finally, in [94], the network topology is identified by reconstructing a weighted Laplacian matrix of DNs.

Table 8
Data Requirements for Power System Operation and Analysis.

	Power Quality	Voltage Control	State Estimation	Model-free Voltage Calculations	DR and HEMS
Input	Voltage and current measurements	Active/reactive power or voltage measurements	Active/reactive power and voltage measurements	Active/reactive power, voltage measurements only for training	Active/reactive power or voltage measurements
Output	Type of disturbance	Control action	State variables of the system	Voltage for all nodes	Control signals, recommendations
Sampling period	kHz	15 min or higher	1 min or higher	30 min	1 s or higher
Other (non-SM) input	-	Topology	Topology	Topology (if voltage data not available for training)	Appliance direct control, temperature

7. Power System Operation and Analysis

Operational analysis applications pertain to methods of examining and improving the performance of power systems, reducing costs, and facilitating better data-driven decision-making for proper day-to-day management. In this section, the concepts listed in Table 2 are discussed. Typical data requirements are presented in Table 8.

7.1. Power Quality

Power quality refers to the degree to which the voltage characteristics of the power supply system, e.g., voltage magnitude, frequency, harmonics, etc., conform to established specifications. Poor power quality means that there are non-stationary disturbances that can cause significant malfunctioning of the electrical equipment, financial losses, and low quality of the electricity that is delivered to consumers. SM data can be valuable in the detection of these disturbances.

In [95], the authors use feature extraction and ML models to classify a number of disturbances, e.g., voltage sags, swells, etc. The overall system is developed to run on the edge inside SMs. Similarly, in [96], a real-time power quality monitoring system for SM level is proposed to detect and classify any type of disturbance. Discrete wavelet transform is used for feature extraction and a SVM for segregation between regular and abnormal data. The classification of disturbances is based on a multi-class SVM.

An edge-based architecture running on a low-cost Raspberry Pi 3 is proposed in [97] offering enormous potential for real scenarios. The classification is performed through a DL model consisting of convolutional, long short-term memory, and dense layers. Besides detecting power quality issues, a power quality-based tariff scheme build on the basis of analyzing the techno-economic consequences of consumers' reactive power and harmonics profiles is presented in [98].

7.2. Voltage Control

SMs allow two-way communication enabling sending and receiving commands in real-time and consequently voltage monitoring and control. SM data can be used to detect in real-time if voltage regulation schemes should be applied to mitigate voltage violations. Such control strategies could be OLTC, Volt-VAr control or capabilities of modern inverters.

In [99], the potential use of SM data as part of an OLTC voltage control strategy is theoretically discussed aiming to solve voltage problems caused by DRES. The SMs can provide the necessary end-users voltage measurements to a control center within short time period. Using these measurements, the voltage set point for OLTC voltage control can be determined by means of optimal power flow.

In [100], a control scheme is developed to encounter voltage stability issues solved at the end-user side by reactive power support using both utility-scale and residential DRESs. To maintain the system frequency and voltage magnitude close to rated values, a multi-objective optimization model is proposed in [101]. The adopted control strategy is developed under the Nash game framework and the controllable loads of the end-users are utilized to optimize the nodal-power variations.

7.3. State Estimation

State estimation is a digital scheme that processes available imperfect information of the power system state and produces the best possible estimate of the true system state. State estimation of DNs is nowadays essential to enable the smart management of medium voltage (MV) and LV grids and is considered the foundation of a variety of key applications, e.g., voltage control, system reconfiguration, and DSM.

SM data have been widely used for state estimation of LV DN. In [102], a combination of WLS and the Levenberg-Marquardt algorithm with an integrated power flow formulation is used. The methodology is applicable for real-time state estimation and uses information provided by SMs. WLS is also used in [103], where a cloud-based SM architecture allowing scalability and interoperability among different off-the-shelf meters is proposed. Moreover, a suitable design of the estimation algorithm using the uncertainty propagation theory is proposed to improve accuracy. To avoid inaccurate modeling due to measurement uncertainties, which can lead the state estimation algorithms to deviate from the true operating states, in [104], an interval state estimation approach is proposed.

SMs have also been utilized for state estimation of MV DN [105, 106, 107, 108] to accurately model pseudo-measurements that are required due to limited number of measurement devices.

7.4. Model-free Voltage Calculations

Model-free voltage calculations pertains to the process of calculating voltages at network nodes without the need of electrical models. This can be achieved by capturing the nonlinear relationship between SM data (demand and voltages) and the corresponding LV feeder. Such an approach based on a deep neural network (DNN) and single-phase SM data is proposed in [109, 110]. The authors aim to replace the traditional power flow analysis (where the topology of the DN is known) with a DNN, since traditional power flow is allegedly expensive, time-consuming, and not 100% accurate. The inputs of the model are the active/reactive power measurements of all end-users and the outputs are the voltages. For training purposes, active/reactive power along voltage data are required. These can be obtained through power flow analyses or through SMs. Once the model has been trained, what-if scenarios can be evaluated by simulating cases of interest, e.g., PV, battery energy storage (BES) units, or EV penetration.

7.5. Demand Response / Home Energy Management Systems

DR refers to a change in the power consumption of a user to match the demand with supply [131]. Typically, a signal is broadcasted by a utility to the user containing a price change or a command for load shedding [132]. Based on this, the end-user can adjust the power demand by postponing selected activities that require large amounts of electric power. A HEMS combines hardware and software components to efficiently manage home energy under DR strategies [133].

7.5.1. Residential DR

Residential DR and HEMS have been widely investigated in the literature. Various approaches rely on direct appliance control by programming the set-points of heating, ventilation, and air-conditioning (HVAC) units and water heaters to provide thermal and hot water comfort [111, 112] while minimizing the energy cost. Similarly, in [113], a HEMS for the co-optimization of cost and comfort is presented based on time of use/dynamic tariffs and penalties/payments.

Appliance scheduling under day-ahead pricing has also been investigated [114] to plan the operation of specific appliances for the next day aiming to minimize the cost and maximize end-user thermal satisfaction. Similar criteria are considered in [115], where a methodology for smoothing power fluctuations resulting from DRES integration, using the DR of a large number of residential appliances is presented. Besides cost and thermal optimization, a deep reinforcement learning approach in [116] integrates time-shiftable appliances, energy utilization of EV and BES systems, and transformer degradation.

To extract end-user habits and provide optimal day-ahead scheduling, NILM has been integrated into HEMSs. In [117, 118], NILM is used to extract appliance-level end-user habits from SM data. The results are integrated into a HEMS to create an efficient and user-centered system by scheduling the appliance usage without user intervention.

7.5.2. Microgrids

DR for microgrids has also been studied using energy management systems. A distributed data-driven coordinated design is proposed in [119]. To achieve efficient energy management of a residential grid, controllable distributed resources, such as EVs and thermostatically controlled loads, are adjusted by balancing the end-use electricity cost, charging preference, and thermal comfort. The idea of [117] was extended in [120], where an advanced system-level energy management system designed for a residential microgrid has been proposed. The approach of [117] was used for each household as the first level of optimization to ensure that both consumers' bills were reduced and their comfort levels were not affected. In the second level, the optimum operation of the microgrid was ensured by considering the generation and consumption units of the microgrid.

In [121], a cooperative and decentralized reinforcement learning method is proposed for demand flexibility assuming a cluster of buildings aiming to minimize thermal and humidity discomfort through heaters and HVAC. Building agents can submit flexibility offers to satisfy as much as possible the grid operator's flexibility request, by taking also into account the building objectives and constraints, with a minimum communication burden. Finally, in [122], a NILM model is used to detect real-time EV activity for all end-users. The DSO can select end-users close to critical nodes and persuade them to shift their charging session during off-peak hours providing additional flexibility and alleviating undervoltage and line congestion issues in the DN.

8. Conclusions and Future Research

In this paper, a comprehensive overview of DN applications based on SM data is conducted. Aside from automated energy consumption metering, there are various purposes for integrating SMs in a DN. Specifically, from the literature review, six key categories of DN-oriented applications have been identified, namely load forecasting, NTLs detection, asset management, power system planning, topology identification, and power system operation and analysis; the latest developments for each case have been reported and discussed.

It should be noted that according to [1], for the time being, most DSOs are focused on applications, e.g., reduction of labor costs, and improvement of SM accuracy, that do not require complex models and analytics, but rather raw measurements. In this context, it is clear that currently, DSOs do not fully exploit their valuable existing data and profit only from a small part of the actual SM potential. However, as the number of installed meters increases globally, it is expected that DSOs will utilize SM data for new fields outside their core business.

Based on the conducted literature review, the most popular data-driven applications are demand forecasting, NTLs detection, and DR/HEMS. Moreover, other promising areas for future research are summarized below:

- *Modeling*: Currently, DN modeling applications rely on several assumptions and simplifications, e.g., radial network structure, approximation of line parameters, etc. More accurate and reliable data-driven models can be achieved via fine-grained SM data from thousands of endpoints within the DN.

- *Planning*: The proliferation of energy-intensive and flexible appliances, e.g., EVs and heat pumps, creates significant potentialities for DR and short-term planning. Since not all EV owners use public charging stations, new solutions for coordinated home charging of EVs can be developed. Furthermore, future research is needed on how SM data can provide smart insights and recommendations about the optimal operation of heat pumps including efficient pre-heating or pre-cooling.

- *Operation*: Active assets, such as EVs, PV, and BES units, can provide additional services to DN operation and help alleviate voltage violations or congestion issues. With the uptake of these assets in the residential sector, the use of already-installed SMs should be further investigated to enable HEMS, DR, and voltage control strategies and elucidate their potential impact on the power system.

- *Monitoring*: Fault detection is one of the most challenging tasks in DN monitoring and SMs can be the key to unlocking new solutions, e.g., for HIF or residential neutral fault detection. Additionally, the next-generation advanced metering infrastructure (AMI 2.0) enables real-time DN monitoring based on distributed edge computing, leading to new business cases, access to high-frequency data, and cost minimization.

9. Acknowledgement

We acknowledge the support of this work by the project "EVOLUTION: rEView Of smart meter data anaLytics for distribUTIOn Networks".

References

[1] Cigre, Utilization of data from smart meter system (2019).
URL https://e-cigre.org/publication/782-utilization-of-data-from-smart-meter-system

[2] B. Völker, A. Reinhardt, A. Faustine, L. Pereira, Watt's up at Home? Smart Meter Data Analytics from a Consumer-Centric Perspective, Energies 14 (3) (2021). doi:10.3390/en14030719.

[3] N. Efkarpidis, et al., Smart Metering Applications, Springer International Publishing, Cham, 2022, pp. 13–124.

[4] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges, IEEE Trans. Smart Grid 10 (3) (2019) 3125–3148. doi:10.1109/TSG.2018.2818167.

[5] C.-N. Yu, P. Mirowski, T. K. Ho, A Sparse Coding Approach to Household Electricity Demand Forecasting in Smart Grids, IEEE Trans. Smart Grid 8 (2) (2017) 738–748. doi:10.1109/TSG.2015.2513900.

[6] R. E. Edwards, J. New, L. E. Parker, Predicting future hourly residential electrical consumption: A machine learning case study, Energy Build. 49 (2012) 591–603. doi:https://doi.org/10.1016/j.enbuild.2012.03.010.

[7] E. Mocanu, P. H. Nguyen, M. Gibescu, W. L. Kling, Deep learning for estimating building energy consumption, Sustain. Energy Grids Netw. 6 (2016) 91–99. doi:https://doi.org/10.1016/j.segan.2016.02.005.

[8] M. N. Fekri, H. Patel, K. Grolinger, V. Sharma, Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network, Appl. Energy 282 (2021) 116177. doi:https://doi.org/10.1016/j.apenergy.2020.116177.

[9] M. N. Fekri, et al., Distributed load forecasting using smart meter data: Federated learning with Recurrent Neural Networks, Int. J. Electr. Power Energy Syst. 137 (2022) 107669. doi:https://doi.org/10.1016/j.ijepes.2021.107669.

[10] Y.-H. Hsiao, Household Electricity Demand Forecast Based on Context Information and User Daily Schedule Analysis From Meter Data, IEEE Trans. Industr. Inform. 11 (1) (2015) 33–43. doi:10.1109/TII.2014.2363584.

[11] Y. Han, X. Sha, E. Grover-Silva, P. Michiardi, On the impact of socio-economic factors on power load forecasting, in: 2014 IEEE Int. Conf. on Big Data (Big Data), 2014, pp. 742–747. doi:10.1109/BigData.2014.7004299.

[12] T. Teeraratkul, D. O'Neill, S. Lall, Shape-Based Approach to Household Electric Load Curve Clustering and Prediction, IEEE Trans. Smart Grid 9 (5) (2018) 5196–5206. doi:10.1109/TSG.2017.2683461.

[13] M. Chaouch, Clustering-Based Improvement of Nonparametric Functional Time Series Forecasting: Application to Intra-Day Household-Level Load Curves, IEEE Trans. Smart Grid 5 (1) (2014) 411–419. doi:10.1109/TSG.2013.2277171.

[14] A. Tascikaraoglu, B. M. Sanandaji, Short-term residential electric load forecasting: A compressive spatio-temporal approach, Energy Build. 111 (2016) 380–392. doi:https://doi.org/10.1016/j.enbuild.2015.11.068.

[15] S. Wang, et al., A bottom-up short-term residential load forecasting approach based on appliance characteristic analysis and multi-task learning, Electr. Power Syst. Res. 196 (2021) 107233. doi:https://doi.org/10.1016/j.epsr.2021.107233.

[16] Z. Zheng, H. Chen, X. Luo, A Kalman filter-based bottom-up approach for household short-term load forecast, Appl. Energy 250 (2019) 882–894. doi:https://doi.org/10.1016/j.apenergy.2019.05.102.

[17] A. A. Mamun, et al., A Comprehensive Review of the Load Forecasting Techniques Using Single and Hybrid Predictive Models, IEEE Access 8 (2020) 134911–134939. doi:10.1109/ACCESS.2020.3010702.

[18] B. Goehry, Y. Goude, P. Massart, J.-M. Poggi, Aggregation of Multi-Scale Experts for Bottom-Up Load Forecasting, IEEE Trans. Smart Grid 11 (3) (2020) 1895–1904. doi:10.1109/TSG.2019.2945088.

[19] Z. Khan, D. Jayaweera, Smart Meter Data Based Load Forecasting and Demand Side Management in Distribution Networks With Embedded PV Systems, IEEE Access 8 (2020) 2631–2644. doi:10.1109/ACCESS.2019.2962150.

[20] Y. Wang, Q. Chen, M. Sun, C. Kang, Q. Xia, An Ensemble Forecasting Method for the Aggregated Load With Subprofiles, IEEE Trans. Smart Grid 9 (4) (2018) 3906–3908. doi:10.1109/TSG.2018.2807985.

[21] C. Fan, F. Xiao, S. Wang, Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques, Appl. Energy 127 (2014) 1–10. doi:https://doi.org/10.1016/j.apenergy.2014.04.016.

[22] E. U. Haq, X. Lyu, Y. Jia, M. Hua, F. Ahmad, Forecasting household electric appliances consumption and peak demand based on hybrid machine learning approach, Energy Rep. 6 (2020) 1099–1105, 2020 The 7th Int. Conf. on Power and Energy Syst. Eng. doi:https://doi.org/10.1016/j.egyr.2020.11.071.

[23] J. Yu, H. Lee, Y. Jeong, S. Kim, Linguistic fuzzy modeling approach for daily peak load forecasting, in: 2013 Int. Conf. on Fuzzy Theory and Its Applications (iFUZZY), 2013, pp. 116–121. doi:10.1109/iFuzzy.2013.6825420.

[24] J. R. G. Sarduy, K. G. Di Santo, M. A. Saidel, Linear and non-linear methods for prediction of peak load at University of São Paulo, Measurement 78 (2016) 187–201. doi:https://doi.org/10.1016/j.measurement.2015.09.053.

[25] M. Sun, Y. Wang, G. Strbac, C. Kang, Probabilistic Peak Load Estimation in Smart Cities Using Smart Meter Data, IEEE Trans. Ind. Electron. 66 (2) (2019) 1608–1618. doi:10.1109/TIE.2018.2803732.

[26] J. Ponoćko, J. V. Milanović, Forecasting Demand Flexibility of Aggregated Residential Load Using Smart Meter Data, IEEE Trans. Power Syst. 33 (5) (2018) 5446–5455. doi:10.1109/TPWRS.2018.2799903.

[27] S. Y. Han, J. No, J.-H. Shin, Y. Joo, Conditional abnormality detection based on AMI data mining, IET Gener. Transm. Distrib. 10 (2016) 3010–3016(6).

[28] I. U. Khan, N. Javaid, C. J. Taylor, X. Ma, Robust Data Driven Analysis for Electricity Theft Attack-Resilient Power Grid, IEEE Trans. Power Syst. 38 (1) (2023) 537–548. doi:10.1109/TPWRS.2022.3162391.

[29] A. Jindal, et al., Decision Tree and SVM-Based Data Analytics for Theft Detection in Smart Grid, IEEE Trans. Industr. Inform. 12 (3) (2016) 1005–1016. doi:10.1109/TII.2016.2543145.

[30] K. Blazakis, T. Kapetanakis, G. Stavrakakis, Effective Electricity Theft Detection in Power Distribution Grids Using an Adaptive Neuro Fuzzy Inference System, Energies 13 (12) (2020). doi:10.3390/en13123110.

[31] M. N. Hasan, et al., Electricity Theft Detection in Smart Grid Systems: A CNN-LSTM Based Approach, Energies 12 (17) (2019). doi:10.3390/en12173310.

[32] Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, Y. Zhou, Wide and Deep Convolutional Neural Networks for Electricity-Theft Detection to Secure Smart Grids, IEEE Trans. Industr. Inform. 14 (4) (2018) 1606–1615. doi:10.1109/TII.2017.2785963.

[33] B. Kabir, et al., Detection of Non-Technical Losses Using MLP-GRU Based Neural Network to Secure Smart Grids, in: Complex, Intelligent and Software Intensive Systems, Springer International Publishing, 2021, pp. 383–394.

[34] F. Shehzad, et al., A Robust Hybrid Deep Learning Model for Detection of Non-Technical Losses to Secure Smart Grids, IEEE Access 9 (2021) 128663–128678. doi:10.1109/ACCESS.2021.3113592.

[35] M. Shaaban, et al., Data-Driven Detection of Electricity Theft Cyberattacks in PV Generation, IEEE Syst. J. 16 (2) (2022) 3349–3359. doi:10.1109/JSYST.2021.3103272.

[36] T.-S. Zhan, et al., Non-technical loss and power blackout detection under advanced metering infrastructure using a cooperative game based inference mechanism, IET Gener. Transm. Distrib. 10 (2016) 873–882(9).

[37] J. I. Guerrero, et al., Improving Knowledge-Based Systems with statistical techniques, text mining, and neural networks for non-technical loss detection, Knowl. Based Syst. 71 (2014) 376–388. doi:https://doi.org/10.1016/j.knosys.2014.08.014.

[38] M. Tariq, H. V. Poor, Electricity Theft Detection and Localization in Grid-Tied Microgrids, IEEE Trans. Smart Grid 9 (3) (2018) 1920–1929. doi:10.1109/TSG.2016.2602660.

[39] E. A. Aranha Neto, J. Coelho, Probabilistic methodology for Technical and Non-Technical Losses estimation in distribution system, Electr. Power Syst. Res. 97 (2013) 93–99. doi:https://doi.org/10.1016/j.epsr.2012.12.008.

[40] A. L. Shah, W. Mesbah, A. T. Al-Awami, An Algorithm for Accurate Detection and Correction of Technical and Nontechnical Losses Using Smart Metering, IEEE Trans. Instrum. Meas. 69 (11) (2020) 8809–8820. doi:10.1109/TIM.2020.2999175.

[41] W. Luan, et al., Energy theft detection via integrated distribution state estimation based on AMI and SCADA measurements, in: 2015 5th Int. Conf. on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), 2015, pp. 751–756. doi:10.1109/DRPT.2015.7432350.

[42] S. A. Salinas, P. Li, Privacy-Preserving Energy Theft Detection in Microgrids: A State Estimation Approach, IEEE Trans. Power Syst. 31 (2) (2016) 883–894. doi:10.1109/TPWRS.2015.2406311.

[43] Y. Zhou, et al., A Dynamic Programming Algorithm for Leveraging Probabilistic Detection of Energy Theft in Smart Home, IEEE Trans. Emerg. Topics Comput. 3 (4) (2015) 502–513. doi:10.1109/TETC.2015.2484841.

[44] P. Jokar, N. Arianpoo, V. C. M. Leung, Electricity Theft Detection in AMI Using Customers' Consumption Patterns, IEEE Trans. Smart Grid 7 (1) (2016) 216–226. doi:10.1109/TSG.2015.2425222.

[45] A. Jindal, et al., Decision Tree and SVM-Based Data Analytics for Theft Detection in Smart Grid, IEEE Trans. on Industr. Inform. 12 (3) (2016) 1005–1016. doi:10.1109/TII.2016.2543145.

[46] R. Qi, J. Zheng, Z. Luo, Q. Li, A Novel Unsupervised Data-Driven Method for Electricity Theft Detection in AMI Using Observer Meters, IEEE Trans. Instrum. Meas. 71 (2022) 1–10. doi:10.1109/TIM.2022.3189748.

[47] G. M. Messinis, A. E. Rigas, N. D. Hatziargyriou, A Hybrid Method for Non-Technical Loss Detection in Smart Distribution Grids, IEEE Trans. Smart Grid 10 (6) (2019) 6080–6091. doi:10.1109/TSG.2019.2896381.

[48] K. Kuroda, R. Yokoyama, D. Kobayashi, T. Ichimura, An Approach to Outage Location Prediction Utilizing Smart Metering Data, in: 2014 8th Asia Modelling Symposium, 2014, pp. 61–66. doi:10.1109/AMS.2014.23.

[49] R. A. Sevlian, et al., Outage Detection Using Load and Line Flow Measurements in Power Distribution Systems, IEEE Trans. Power Syst. 33 (2) (2018) 2053–2069. doi:10.1109/TPWRS.2017.2727979.

[50] Y. Jiang, et al., Outage Management of Distribution Systems Incorporating Information From Smart Meters, IEEE Trans. Power Syst. 31 (5) (2016) 4144–4154. doi:10.1109/TPWRS.2015.2503341.

[51] R. Moghaddass, J. Wang, A Hierarchical Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter Data, IEEE Trans. Smart Grid 9 (6) (2018) 5820–5830. doi:10.1109/TSG.2017.2697440.

[52] Y. Liao, et al., Quick Line Outage Identification in Urban Distribution Grids via Smart Meters, CSEE J. Power Energy Syst. 8 (4) (2022) 1074–1086. doi:10.17775/CSEEJPES.2020.04640.

[53] Y. Arafat, et al., Remote switching of multiple smart meters and steps to check the effect on the grid's power quality, in: 2014 IEEE PES T&D Conf. and Exposition, 2014, pp. 1–5. doi:10.1109/TDC.2014.6863273.

[54] S. Chakraborty, S. Das, Application of Smart Meters in High Impedance Fault Detection on Distribution Systems, IEEE Trans. Smart Grid 10 (3) (2019) 3465–3473. doi:10.1109/TSG.2018.2828414.

[55] K. Ashok, et al., Online Detection of Inter-Turn Winding Faults in Single-Phase Distribution Transformers Using Smart Meter Data, IEEE Trans. Smart Grid 12 (6) (2021) 5073–5083. doi:10.1109/TSG.2021.3102101.

[56] A. T. Procopiou, et al., Smart meter-driven estimation of PV hosting capacity, in: CIRED 2020 Berlin Workshop (CIRED 2020), Vol. 2020, 2020, pp. 128–131. doi:10.1049/oap-cired.2021.0287.

[57] V. Klonari, et al., Estimating the Photovoltaic Hosting Capacity of a Low Voltage Feeder Using Smart Meters' Measurements, in: Smart Metering Technology and Services, IntechOpen, 2016, Ch. 7. doi:10.5772/62418.

[58] A. K. Jain, et al., Quasi-Static Time-Series PV Hosting Capacity Methodology and Metrics, in: 2019 IEEE PES Innov. Smart Grid Technol. Conf. (ISGT), 2019, pp. 1–5. doi:10.1109/ISGT.2019.8791569.

[59] J. Zhu, W. J. Nacmanson, L. F. Ochoa, B. Hellyer, Assessing the EV Hosting Capacity of Australian Urban and Rural MV-LV Networks, Electr. Power Syst. Res. 212 (2022) 108399. doi:https://doi.org/10.1016/j.epsr.2022.108399.

[60] M. Z. Liu, et al., Grid and Market Services From the Edge: Using Operating Envelopes to Unlock Network-Aware Bottom-Up Flexibility, IEEE Power and Energy Mag. 19 (4) (2021) 52–62. doi:10.1109/MPE.2021.3072819.

[61] M. Z. Liu, L. F. Ochoa, P. K. Wong, J. Theunissen, Using OPF Based Operating Envelopes to Facilitate Residential DER Services, IEEE Trans. Smart Grid (2022) 1–1doi:10.1109/TSG.2022.3188927.

[62] K. Petrou, et al., Ensuring Distribution Network Integrity Using Dynamic Operating Limits for Prosumers, IEEE Trans. Smart Grid 12 (5) (2021) 3877–3888. doi:10.1109/TSG.2021.3081371.

[63] J. Zhang, P. Wang, N. Zhang, Distribution Network Admittance Matrix Estimation With Linear Regression, IEEE Trans. Power Syst. 36 (5) (2021) 4896–4899. doi:10.1109/TPWRS.2021.3090250.

[64] S. Han, et al., An Automated Impedance Estimation Method in Low-Voltage Distribution Network for Coordinated Voltage Regulation, IEEE Trans. Smart Grid 7 (2) (2016) 1012–1020. doi:10.1109/TSG.2015.2489199.

[65] S. Iakovlev, R. J. Evans, I. Mareels, Low-Voltage Distribution Network Impedances Identification Based on Smart Meter Data (2018). doi:10.48550/ARXIV.1809.06657.
URL https://arxiv.org/abs/1809.06657

[66] Y. Wang, M. Z. Liu, L. F. Ochoa, Assessing the effects of DER on voltages using a smart meter-driven three-phase LV feeder model, Electr. Power Syst. Res. 189 (2020) 106705. doi:https://doi.org/10.1016/j.epsr.2020.106705.

[67] V. Arya, T. S. Jayram, S. Pal, S. Kalyanaraman, Inferring Connectivity Model from Meter Measurements in Distribution Networks, in: Proc. of the 4th Int. Conf. on Future Energy Syst., e-Energy '13, Association for Computing Machinery, 2013, p. 173–182. doi:10.1145/2487166.2487186.

[68] X. Tang, J. V. Milanovic, Phase Identification of LV Distribution Network with Smart Meter Data, in: 2018 IEEE Power & Energy Society General Meeting (PESGM), 2018, pp. 1–5. doi:10.1109/PESGM.2018.8586483.

[69] A. Heidari-Akhijahani, A. Safdarian, F. Aminifar, Phase Identification of Single-Phase Customers and PV Panels via Smart Meter Data, IEEE Trans. Smart Grid 12 (5) (2021) 4543–4552. doi:10.1109/TSG.2021.3074663.

[70] H. Pezeshki, P. Wolfs, Correlation based method for phase identification in a three phase LV distribution network, in: 2012 22nd Australasian Universities Power Eng. Conf. (AUPEC), 2012, pp. 1–7.

[71] B. K. Seal, M. F. McGranaghan, Automatic identification of service phase for electric utility customers, in: 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1–3. doi:10.1109/PES.2011.6039623.

[72] H. Pezeshki, P. J. Wolfs, Consumer phase identification in a three phase unbalanced LV distribution network, in: 2012 3rd IEEE PES Innov. Smart Grid Technol. Europe (ISGT Europe), 2012, pp. 1–7. doi:10.1109/ISGTEurope.2012.6465632.

[73] V. Arya, R. Mitra, Voltage-based clustering to identify connectivity relationships in distribution networks, in: 2013 IEEE Int. Conf. on Smart Grid Commun. (SmartGridComm), 2013, pp. 7–12. doi:10.1109/SmartGridComm.2013.6687925.

[74] R. Mitra, et al., Voltage Correlations in Smart Meter Data, in: KDD '15: Proceedings of the 21th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA, 2015, p. 1999–2008. doi:10.1145/2783258.2788594.

[75] Y. Ma, et al., Phase Identification of Smart Meters by Spectral Clustering, in: 2018 2nd IEEE Conf. on Energy Internet and Energy System Integration (EI2), 2018, pp. 1–5. doi:10.1109/EI2.2018.8582318.

[76] S. Liu, X. Cui, Z. Lin, Z. Lian, Z. Lin, F. Wen, Y. Ding, Q. Wang, L. Yang, R. Jin, H. Qiu, Practical Method for Mitigating Three-Phase Unbalance Based on Data-Driven User Phase Identification, IEEE Trans. Power Syst. 35 (2) (2020) 1653–1656. doi:10.1109/TPWRS.2020.2965770.

[77] L. Blakely, M. J. Reno, W.-c. Feng, Spectral Clustering for Customer Phase Identification Using AMI Voltage Timeseries, in: 2019 IEEE Power and Energy Conference at Illinois (PECI), 2019, pp. 1–7. doi:10.1109/PECI.2019.8698780.

[78] A. Simonovska, L. F. Ochoa, Phase Grouping in PV-Rich LV Feeders: Smart Meter Data and Unconstrained k-Means, in: 2021 IEEE Madrid PowerTech, 2021, pp. 1–6. doi:10.1109/PowerTech46648.2021.9495097.

[79] S. P. Jayadev, A. Rajeswaran, N. P. Bhatt, R. Pasumarthy, A novel approach for phase identification in smart grids using Graph Theory and Principal Component Analysis, in: 2016 American Control Conference (ACC), 2016, pp. 5026–5031. doi:10.1109/ACC.2016.7526150.

[80] S. J. Pappu, N. Bhatt, R. Pasumarthy, A. Rajeswaran, Identifying Topology of Low Voltage Distribution Networks Based on Smart Meter Data, IEEE Trans. Smart Grid 9 (5) (2018) 5113–5122. doi:10.1109/TSG.2017.2680542.

[81] Z. S. Hosseini, A. Khodaei, A. Paaso, Machine Learning-Enabled Distribution Network Phase Identification, IEEE Trans. Power Syst. 36 (2) (2021) 842–850. doi:10.1109/TPWRS.2020.3011133.

[82] M. Xu, R. Li, F. Li, Phase Identification With Incomplete Data, IEEE Trans. Smart Grid 9 (4) (2018) 2777–2785. doi:10.1109/TSG.2016.2619264.

[83] Z. Tian, W. Wu, B. Zhang, A Mixed Integer Quadratic Programming Model for Topology Identification in Distribution Network, IEEE Trans. Power Syst. 31 (1) (2016) 823–824. doi:10.1109/TPWRS.2015.2394454.

[84] G. N. Korres, N. M. Manousakis, A state estimation algorithm for monitoring topology changes in distribution systems, in: 2012 IEEE Power and Energy Society General Meeting, 2012, pp. 1–8. doi:10.1109/PESGM.2012.6345126.

[85] W. Luan, et al., Smart Meter Data Analytics for Distribution Network Connectivity Verification, IEEE Trans. Smart Grid 6 (4) (2015) 1964–1971. doi:10.1109/TSG.2015.2421304.

[86] G. Cavraro, V. Kekatos, S. Veeramachaneni, Voltage Analytics for Power Distribution Network Topology Verification, IEEE Trans. Smart Grid 10 (1) (2019) 1058–1067. doi:10.1109/TSG.2017.2758600.

[87] L. Ma, L. Wang, Z. Liu, Topology Identification of Distribution Networks Using a Split-EM Based Data-Driven Approach, IEEE Trans. Power Syst. 37 (3) (2022) 2019–2031. doi:10.1109/TPWRS.2021.3119649.

[88] D. Deka, S. Backhaus, M. Chertkov, Learning Topology of Distribution Grids using only Terminal Node Measurements (2016). doi:10.48550/ARXIV.1608.05031.
URL https://arxiv.org/abs/1608.05031

[89] S. Park, D. Deka, M. Chcrtkov, Exact Topology and Parameter Estimation in Distribution Grids with Minimal Observability, in: 2018 Power Systems Computation Conference (PSCC), 2018, pp. 1–6. doi:10.23919/PSCC.2018.8442881.

[90] G. Cavraro, R. Arghandeh, K. Poolla, A. von Meier, Data-driven approach for distribution network topology detection, in: 2015 IEEE Power & Energy Society General Meeting, 2015, pp. 1–5. doi:10.1109/PESGM.2015.7286490.

[91] H. Zhang, J. Zhao, X. Wang, Y. Xuan, Low-Voltage Distribution Grid Topology Identification With Latent Tree Model, IEEE Trans. Smart Grid 13 (3) (2022) 2158–2169. doi:10.1109/TSG.2022.3146205.

[92] J. Zhao, L. Li, Z. Xu, X. Wang, H. Wang, X. Shao, Full-Scale Distribution System Topology Identification Using Markov Random Field, IEEE Trans. Smart Grid 11 (6) (2020) 4714–4726. doi:10.1109/TSG.2020.2995164.

[93] Y. Liao, Y. Weng, G. Liu, R. Rajagopal, Urban MV and LV Distribution Grid Topology Estimation via Group Lasso, IEEE Trans. Power Syst. 34 (1) (2019) 12–27. doi:10.1109/TPWRS.2018.2868877.

[94] Y. Guo, et al., Distribution Grid Modeling Using Smart Meter Data, IEEE Trans. Power Syst. 37 (3) (2022) 1995–2004. doi:10.1109/TPWRS.2021.3118004.

[95] F. A. S. Borges, R. A. S. Fernandes, I. N. Silva, C. B. S. Silva, Feature Extraction and Power Quality Disturbances Classification Using Smart Meters Signals, IEEE Trans. Industr. Inform. 12 (2) (2016) 824–833. doi:10.1109/TII.2015.2486379.

[96] I. Parvez, et al., Online power quality disturbance detection by support vector machine in smart meter, J. Mod. Power Syst. Clean Energy (2018) 1–12.

[97] W. L. Rodrigues Junior, F. A. Borges, A. F. da S. Veloso, R. de A.L. Rabêlo, J. J. Rodrigues, Low voltage smart meter for monitoring of power quality disturbances applied in smart grid, Measurement 147 (2019) 106890. doi:https://doi.org/10.1016/j.measurement.2019.106890.

[98] M. R. Balwani, K. Thirumala, V. Mohan, S. Bu, M. S. Thomas, Development of a Smart Meter for Power Quality-Based Tariff Implementation in a Smart Grid, Energies 14 (19) (2021). doi:10.3390/en14196171.

[99] C. Gao, M. A. Redfern, A Review of Voltage Control in Smart Grid and Smart Metering Technologies on Distribution Networks, in: 2011 46th Int. Universities Power Eng. Conf. (UPEC), 2011, pp. 1–5.

[100] N. Duan, C. Huang, C.-C. Sun, L. Min, Smart Meters Enabling Voltage Monitoring and Control: The Last-Mile Voltage Stability Issue, IEEE Trans. Industr. Inform. 18 (1) (2022) 677–687. doi:10.1109/TII.2021.3062628.

[101] Y. Wang, D. Qi, J. Zhang, Multi-objective optimization for voltage and frequency control of smart grids based on controllable loads, Global Energy Interconnection 4 (2) (2021) 136–144. doi:https://doi.org/10.1016/j.gloei.2021.05.006.

[102] E. B. Alzate, et al., Distribution System State Estimation to Support Coordinated Voltage-Control Strategies by Using Smart Meters, IEEE Trans. Power Syst. 34 (6) (2019) 5198–5207. doi:10.1109/TPWRS.2019.2902184.

[103] M. Pau, et al., Low voltage system state estimation based on smart metering infrastructure, in: 2016 IEEE Int. Workshop on Appl. Meas. for Power Syst. (AMPS), 2016, pp. 1–6. doi:10.1109/AMPS.2016.7602804.

[104] M. Huang, et al., Interval State Estimation for Low-Voltage Distribution Systems Based on Smart Meter Data, IEEE Trans. Instrum. Meas. 68 (9) (2019) 3090–3099. doi:10.1109/TIM.2018.2877549.

[105] A. Angioni, et al., Impact of Pseudo-Measurements From New Power Profiles on State Estimation in Low-Voltage Grids, IEEE Trans. Instrum. Meas. 65 (1) (2016) 70–77. doi:10.1109/TIM.2015.2454673.

[106] Q. Chen, D. Kaleshi, Z. Fan, S. Armour, Impact of Smart Metering Data Aggregation on Distribution System State Estimation, IEEE Trans. Industr. Inform. 12 (4) (2016) 1426–1437. doi:10.1109/TII.2016.2573272.

[107] B. P. Hayes, J. K. Gruber, M. Prodanovic, A Closed-Loop State Estimation Tool for MV Network Monitoring and Operation, IEEE Trans. Smart Grid 6 (4) (2015) 2116–2125. doi:10.1109/TSG.2014.2378035.

[108] Y. Liu, et al., State Estimation of Three-Phase Four-Conductor Distribution Systems With Real-Time Data From Selective Smart Meters, IEEE Trans. Power Syst. 34 (4) (2019) 2632–2643. doi:10.1109/TPWRS.2019.2892726.

[109] V. Bassi, L. Ochoa, T. Alpcan, Model-Free Voltage Calculations for PV-Rich LV Networks: Smart Meter Data and Deep Neural Networks, in: 2021 IEEE Madrid PowerTech, 2021, pp. 1–6. doi:10.1109/PowerTech46648.2021.9494847.

[110] V. Bassi, L. F. Ochoa, T. Alpcan, C. Leckie, Electrical Model-Free Voltage Calculations Using Neural Networks and Smart Meter Data, IEEE Trans. Smart Grid (2022) 1–1doi:10.1109/TSG.2022.3227602.

[111] H. Mortaji, et al., Load Shedding and Smart-Direct Load Control Using Internet of Things in Smart Grid Demand Response Management, IEEE Trans. Ind. Appl. 53 (6) (2017) 5155–5163. doi:10.1109/TIA.2017.2740832.

[112] X. Kou, et al., A Scalable and Distributed Algorithm for Managing Residential Demand Response Programs Using Alternating Direction Method of Multipliers (ADMM), IEEE Trans. Smart Grid 11 (6) (2020) 4871–4882. doi:10.1109/TSG.2020.2995923.

[113] M. Shafie-Khah, P. Siano, A Stochastic Home Energy Management System Considering Satisfaction Cost and Response Fatigue, IEEE Trans. Industr. Inform. 14 (2) (2018) 629–638. doi:10.1109/TII.2017.2728803.

[114] Y. Liu, et al., Pricing-Based Demand Response for a Smart Home With Various Types of Household Appliances Considering Customer Satisfaction, IEEE Access 7 (2019) 86463–86472. doi:10.1109/ACCESS.2019.2924110.

[115] F. Elghitani, E. El-Saadany, Smoothing Net Load Demand Variations Using Residential Demand Management, IEEE Trans. Industr. Inform. 15 (1) (2019) 390–398. doi:10.1109/TII.2018.2852482.

[116] A. A. Amer, K. Shaban, A. M. Massoud, DRL-HEMS: Deep Reinforcement Learning Agent for Demand Response in Home Energy Management Systems Considering Customers and Operators Perspectives, IEEE Trans. Smart Grid 14 (1) (2023) 239–250. doi:10.1109/TSG.2022.3198401.

[117] J. M. Çimen, Halil others, A Microgrid Energy Management System Based on Non-Intrusive Load Monitoring via Multitask Learning, IEEE Trans. Smart Grid 12 (2) (2021) 977–987. doi:10.1109/TSG.2020.3027491.

[118] Y.-H. Lin, M.-S. Tsai, An Advanced Home Energy Management System Facilitated by Nonintrusive Load Monitoring With Automated Multiobjective Power Scheduling, IEEE Trans. Smart Grid 6 (4) (2015) 1839–1851. doi:10.1109/TSG.2015.2388492.

[119] P. Wang, et al., Data-driven energy management in residential areas leveraging demand response, Energy Build. 269 (2022) 112235. doi:https://doi.org/10.1016/j.enbuild.2022.112235.

[120] H. Çimen, et al., An online energy management system for AC/DC residential microgrids supported by non-intrusive load monitoring, Appl. Energy 307 (2022) 118136. doi:https://doi.org/10.1016/j.apenergy.2021.118136.

[121] L. A. Hurtado, et al., Enabling Cooperative Behavior for Building Demand Response Based on Extended Joint Action Learning, IEEE Trans. Industr. Inform. 14 (1) (2018) 127–136. doi:10.1109/TII.2017.2753408.

[122] C. L. Athanasiadis, et al., A Benchmarking Testbed for Low-Voltage Active Distribution Network Studies, IEEE Open Access Journal of Power and Energy 10 (2023) 104–115. doi:10.1109/OAJPE.2022.3232186.

[123] S. Dai, et al., Electrical peak demand forecasting-A review, arXiv preprint arXiv:2108.01393 (2021).

[124] C. L. Athanasiadis, T. A. Papadopoulos, D. I. Doukas, Real-time non-intrusive load monitoring: A light-weight and scalable approach, Energy Build. 253 (2021) 111523. doi:https://doi.org/10.1016/j.enbuild.2021.111523.

[125] C. Athanasiadis, et al., A Scalable Real-Time Non-Intrusive Load Monitoring System for the Estimation of Household Appliance Power Consumption, Energies 14 (3) (2021). doi:10.3390/en14030767.

[126] T. Ahmad, Non-technical loss analysis and prevention using smart meters, Renewable and Sustainable Energy Reviews 72 (2017) 573–589. doi:https://doi.org/10.1016/j.rser.2017.01.100.

[127] G. Messinis, N. Hatziargyriou, Review of non-technical loss detection methods, Electr. Power Syst. Res. 158 (2018) 250–266. doi:https://doi.org/10.1016/j.epsr.2018.01.005.

[128] H. Tram, Technical and operation considerations in using Smart Metering for outage management, in: 2008 IEEE/PES Transm. and Distrib. Conf. and Exposition, 2008, pp. 1–3. doi:10.1109/TDC.2008.4517273.

[129] F. Therrien, L. Blakely, M. J. Reno, Assessment of Measurement-Based Phase Identification Methods, IEEE Open Access Journal of Power and Energy 8 (2021) 128–137. doi:10.1109/OAJPE.2021.3067632.

[130] A. Hoogsteyn, M. Vanin, A. Koirala, D. Van Hertem, Low voltage customer phase identification methods based on smart meter data, Electr. Power Syst. Res. 212 (2022) 108524. doi:https://doi.org/10.1016/j.epsr.2022.108524.

[131] P. Siano, Demand response and smart grids—a survey, Renewable and Sustainable Energy Reviews 30 (2014) 461–478. doi:https://doi.org/10.1016/j.rser.2013.10.022.

[132] P. Palensky, D. Dietrich, Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads, IEEE Trans. Industr. Inform. 7 (3) (2011) 381–388. doi:10.1109/TII.2011.2158841.

[133] B. Zhou, W. Li, K. W. Chan, Y. Cao, Y. Kuang, X. Liu, X. Wang, Smart home energy management systems: Concept, configurations, and scheduling strategies, Renewable and Sustainable Energy Reviews 61 (2016) 30–40. doi:https://doi.org/10.1016/j.rser.2016.03.047.