Advanced Distribution Measurement Technologies and Data Applications for Smart Grids: A Review

Antonio E. Saldaña-González 1,* , Andreas Sumper 1 , Mònica Aragüés-Penalba 1 and Miha Smolnikar 2

1 Centre d’Innovació Tecnològica en Convertidors Estàtics i Accionaments (CITCEA-UPC), Departament d’Enginyeria Elèctrica, Universitat Politècnica de Catalunya ETS d’Enginyeria Industrial de Barcelona, Avinguda Diagonal, 647, Pl. 2, 08028 Barcelona, Spain; andreas.sumper@upc.edu (A.S.); monica.aragues@upc.edu (M.A.-P.)
2 COMSENSUS, Brezje pri Dobu 8, 1233 Dob, Slovenia; miha.smolnikar@comensus.eu
* Correspondence: antonio.emmanuel.saldana@upc.edu; Tel.: +34-631-506-910

Received: 19 May 2020; Accepted: 15 July 2020; Published: 20 July 2020

Abstract: The integration of advanced measuring technologies in distribution systems allows distribution system operators to have better observability of dynamic and transient events. In this work, the applications of distribution grid measurement technologies are explored in detail. The main contributions of this review are: (a) a comparison of eight advanced measurement devices for distribution networks, based on their technical characteristics, including reporting periods, measuring data, precision, and sample rate; (b) a review of the most recent applications of micro-Phasor Measurement Units, Smart Meters, and Power Quality Monitoring devices used in distribution systems, considering different novel methods applied for data analysis; and (c) an input-output table that relates measured quantities from micro-Phasor Measurement Units and Smart Meters needed for each specific application found in this extensive review. This paper aims to serve as an important guide for researches and engineers studying smart grids.

Keywords: advanced measuring devices; micro-phasor measurement units; power quality monitors; smart grid technologies; smart meters

1. Introduction

The integration of advanced measuring technology, such as Smart Meters, micro-Phasor Measurement Units (µPMUs), and Power Quality Monitors (PQM) in distribution systems, allows distribution system operators to have better observability of the electrical distribution system. High-precision measurements, rapid communication, and remote storage of the extracted data are some of the characteristics of these measuring devices.

The urgent reason to enhance the observability with the deployment of new technology devices is mainly driven by the increasing integration of distributed energy resources (solar energy, wind energy, bio-energy) and flexible loads (electric vehicles and air conditioning systems) in distribution grids. These devices have a significant effect on the operation, stability, and quality of energy distribution networks. Customers are able to exchange active power with the electric grid in two-way directions, increasing the complexity and uncertainty of the distribution system operation [1–3].

Recent works have reviewed the general applications of measurement technologies in distribution systems. In Reference [4,5], the authors described the technology architecture used in smart grids, including the metering and communication systems for transmission and distribution systems. In Reference [6,7], the authors presented an overview of measurement technology, including smart meters, smart sensors, smart power meters, Phasor Measurement Units (PMUs), Phasor
Data Concentrators (PDCs), and Supervisory Control And Data Acquisition (SCADA) systems, for monitoring, protection, and control in smart grid networks. However, none of these overviews mentioned the required data characteristics for each application or the types of measuring devices used for distribution systems.

Relevant overviews of \( \mu \)PMUs applications for distribution grids were described in Reference \[8–11\], which include monitoring, diagnostic, and control applications. In fact, these papers have not reviewed recent research work related to \( \mu \)PMUs data applications. In recent reviewed papers \[12,13\], authors have focused on \( \mu \)PMUs applications, including state awareness, event detection, adaptive protection, and network reconfiguration. However, these reviews do not include a list of \( \mu \)PMUs applications based on the input data, methods, and visualization of each application.

Furthermore, several review papers have studied the applications of smart meter data for distribution networks \[14–17\]. In Reference \[14\], the authors reviewed the smart meter data techniques and methodologies developed for different applications. They also discussed the big data issues, the transition of energy systems, data privacy, and security. In Reference \[15\], the authors reviewed the methods and techniques for using smart meter data, such as forecasting, clustering, classification, and optimization. However, these works do not mention the data inputs necessary to implement each of these methods.

Some research papers have studied power quality applications. Reference \[18\] analyzed the harmonic impact of the integration of renewable sources into the distribution network. Some other recent applications are for optimal location of PQM in distribution systems, due to the limitation of measuring devices \[19\]. The development of optimal placement techniques and energy data are discussed in Reference \[20–24\]. To the best knowledge of the authors, this is the first time that a work integrates an overview of the applications of PQM for distribution systems. In this review the main contributions are:

- A comparison of eight advanced measurement devices for distribution networks, based on their technical characteristics, including reporting periods, observability, measure quantities, precision, and sample rate.
- A review of the most recent applications of micro-Phasor Measurement Units, Smart Meters, and Power Quality Monitoring devices used in distribution systems, considering different novel methods applied for data analysis.
- An input-output table that relates the measured quantities from \( \mu \)PMUs and smart meters needed for each specific application found in this review. To our knowledge, this is the first time that a review integrates the input data according to each application.

The organization of this paper is as follows. In Section 2, a comparison among different types of advanced measurement devices used in distribution systems is addressed. In Section 3, a general architecture of communication systems is presented based on different types of networks. Section 4 reviews and classifies recent applications of advanced measurement technologies based on an extended literature review. Finally, Section 5 emphasizes the main conclusions from the work.

2. Advanced Measurement Devices for Active Distribution Networks

2.1. Monitoring Active Distribution Networks

The integration of renewable energy sources, energy storage devices and charging stations for electric vehicles are some of the emerging technologies in active distribution networks. Active distribution networks are mainly characterized by bi-directional power flows, meaning that consumers can generate energy locally and inject it into the distribution network (also known as prosumers). Active components can provide flexibility to the network, which is useful for improving network stability (voltage, transient, and dynamic), improving power quality and optimally planning network expansion. This is an economically viable alternative for energy operators because it means investment savings to reinforce the network.
However, high penetration of distributed generation in active distribution networks can cause considerable impacts on the operation of transmission networks. In order to simulate the uncertainty of distributed generation it is essential to use precise models based on data-driven solutions to determine the future behaviour of the power grid, as well as improving the impacts on transmission networks. For this reason, it is important to know the different measurement devices used in active distribution networks today, which allow algorithm developers or researchers to know the types of measurement devices that can provide the required resolution for each application considering the costs of the devices. The development of these tools will allow to improve the current monitoring systems (e.g., SCADA) that operate at distribution and transmission level.

Monitoring devices help active distribution network operators to monitor two-way power flows, energy consumption, and distributed generation in near real time. This allows Distribution System Operators (DSO) to monitor, operate and plan active distribution networks remotely with higher energy efficiency and establish optimal investment planning (such as maintenance operations and new equipment) for short and long-term scenarios. Monitoring devices also help to identify the areas of the network that are not performing at acceptable levels of quality. This can provide greater protection to consumers against over-voltages or congestions in lines, reducing incidents on the power grid. Modern meters used in active distribution networks actively communicate with a central system, which can provide information about the location and magnitude of network incidents. Generally, smart metering devices are integrated into an informatic platform so they can be managed centrally and remotely by DSOs [25,26].

Future active distribution networks will need to reinforce or integrate advanced metering systems in order to increase the observability, security and reliability. The integration of $\mu$PMUs, PQMs, and smart meters are a viable solution to enhance wide-area visibility at the medium and low voltage level as seen in Figure 1. In Reference [27], the authors investigated the roles that human operators will take in highly automated systems. They conclude that the future role of operators will be as fault managers. They will be mainly focused on identifying network failures and providing immediate technical support to maintain network operation limits.

![Figure 1. High observability in an active distribution network.](image)

2.2. Types of Measurement Devices in Distribution Systems

In this section, the technical characteristics and main uses of measurement technologies for distribution systems are described. In Table 1, a technical comparison of eight advanced measurement devices used to observe the behavior of the electrical distribution network are presented. The objective
of this comparison table is to identify which of them are useful for observing events in a stable, dynamic, and transitory state, considering their technical measurement parameters. The investment cost of monitoring devices is another important aspect considered in the comparison table; this is especially important when such instruments have to be employed for wide area distribution monitoring networks and thus need a high number of measurement instruments to be installed. This costs do not entail the cost of their installation at distribution level. In general, the measurement devices that present better performance results according in their accuracy, sample rate, reporting period, and amount of measured quantities are μPMUs and PQM devices. Therefore, this article aims to review the main applications of their measured data, including their methods and techniques. Recent developed methods are based on Machine Learning (ML) and deep learning techniques capable of analyzing multiple dynamic events in the distribution network.
Table 1. Comparison table of advanced measurement devices in distribution networks.

<table>
<thead>
<tr>
<th>Power Sensors</th>
<th>Smart Meter</th>
<th>µPMU</th>
<th>PMUs</th>
<th>Power Quality Monitors</th>
<th>Substation Meters</th>
<th>Digital Fault Recorders</th>
<th>Wireless Power Line Sensors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reporting period:</td>
<td>1 sample each (2–4) s</td>
<td>1 sample each 1–60 min</td>
<td>Up to 120 samples each second</td>
<td>10–60 samples each second</td>
<td>1 sample each second</td>
<td>1 sample each min</td>
<td>Record the event Up to 120 s</td>
</tr>
<tr>
<td>System Observability:</td>
<td>Steady State</td>
<td>Load Profiles</td>
<td>Dynamic and Transient State</td>
<td>Steady and Dynamic State</td>
<td>Transient events</td>
<td>Dynamic State</td>
<td>Transient and Disturbance Events</td>
</tr>
<tr>
<td>Network Monitoring:</td>
<td>Medium Voltage</td>
<td>Low Voltage</td>
<td>Medium and Low Voltage</td>
<td>High and Medium voltage</td>
<td>Medium and Low Voltage</td>
<td>Medium Voltage</td>
<td>Medium Voltage</td>
</tr>
<tr>
<td>Measure Quantities</td>
<td>Voltage (RMS)</td>
<td>Active Power, PF Reactive Power, Frequency, Current, RMS Voltage and Power Quality data</td>
<td>3 Ph-Voltage phasors</td>
<td>3 Ph-Voltage phasors</td>
<td>Frequency, Voltage, Current, THD, Harmonics, Flicker, Unbalance</td>
<td>Active and reactive power flow, complex current and complex voltage</td>
<td>Voltage, Currents, frequency, P, Q, S, PF, Harmonics, Symmetrical Components</td>
</tr>
<tr>
<td>Sensor Accuracy</td>
<td>±0.5%</td>
<td>Active Power: ±1% Reactive Power: ±2%</td>
<td>Amplitude: ±0.05% Angle: ±0.01%</td>
<td>Amplitude: ±1% Angle: ±1%</td>
<td>±0.2% and ±0.1%</td>
<td>±0.5%</td>
<td>±0.1%</td>
</tr>
<tr>
<td>Sample Rate</td>
<td>-</td>
<td>-</td>
<td>Up to 30,720 s/s</td>
<td>Up to 2880 s/s</td>
<td>Up to 100,000 s/s</td>
<td>-</td>
<td>Up to 25,600 s/s</td>
</tr>
<tr>
<td>Cost Level *</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Reference:</td>
<td>[28–30]</td>
<td>[31–34]</td>
<td>[35]</td>
<td>[36,37]</td>
<td>[30,38,39]</td>
<td>[30,40]</td>
<td>[30,41]</td>
</tr>
</tbody>
</table>

* Low Cost (Range $200–$3500 USD), Medium Cost (Range $4000–$15,000 USD), High Cost (Range $22,000–$80,000 USD) [43–46].
- **Power Sensor**: The power sensor provides the supervision, back-up management, control, and regulation of power grids usually for SCADA systems. This sensor is able to measure the Root Mean Square (RMS) value of the voltage, active, and reactive power measurements, each 2–4 s. Then, this information is sent to a master station that allows the operator to monitor and control the grid. In recent years, hybrid SCADA systems based on distribution phasor measurement units have been proposed to enhance the observability on distribution networks [47,48]. In Reference [49], the authors proposed a SCADA system based on internet of things, which had integrated a fog router and a cloud architecture that takes care of outage management. This fog router was able to locate faults in a distribution system using voltages and currents from smart meters and line sensors.

- **Smart Meter**: Smart meters are electronic devices that measure active and reactive power, current, voltage RMS, and, in some cases, power quality data, with a measurement accuracy of $\pm 1\%$ or $\pm 0.5\%$, with an adjustable reporting rate from 1 to 60 min. They are usually installed in the low voltage side of the distribution system. The main functionalities of smart meters are for automatic billing and energy management. The two-way communication is the most important difference that distinguishes advanced smart meters from conventional smart meters [50]. In smart distribution grids, smart meters provide consumers and energy providers knowledge about the price information generally every 15 to 30 min. This enables the consumer to know when is the cheapest time to use energy and also to manage their own consumption. Electric service companies control smart meters remotely and use this information to forecast their daily production. In Section 4, recent smart meters data application are analysed, such as Non-Technical Loss (NTL) detection, customer characterization, and load clustering with recent novel techniques.

- **Micro-Phasor Measurement Unit**: The $\mu$PMU is currently the most advanced measurement device used in medium and low voltage level for distribution grids. It is able to measure different quantities, such as 3-phase voltage phasors, 3-phase current phasors, active power, reactive power, and frequency, with a reporting rate up to 120 samples each second, in a 60 Hz system. This last feature allows the operator to observe dynamic and transient events on the distribution system with a sampling rate of over 30 kHz. Furthermore, the high sampling rate allows to have an angle accuracy of $\pm 0.010^\circ$ and an amplitude accuracy of $\pm 0.05\%$. The $\mu$PMU has precise time stamps that are synchronized with the phase angle between multiple locations. In Reference [51], an experimental microgrid was developed using seven $\mu$PMUs located at the Lawrence Berkeley National Laboratory. This was the first pilot network at distribution-level using $\mu$PMUs.

- **Phasor Measurement Unit**: The PMU is a measurement device that was originally developed to observe the dynamic operation of transmission power grids. However, the use of PMUs in active distribution grids, can also be a very helpful tool to estimate the state variables of a distribution system. This device has the ability to measure 3-phase voltage phasors, 3-phase current phasors, active power, reactive power, frequency, and power factor, with a reporting period of 10–60 samples each second. Although the operator is able to observe steady and dynamic state conditions, it lacks in observing transient events. Most of the commercial PMUs have a maximum sample rate of 2880 Hz, angle accuracy of $\pm 1^\circ$, and amplitude accuracy of $\pm 1\%$. The PMUs also have precise time stamps that compare and synchronize the phase angle between different locations. PMUs are used in Wide-Area Measurement Systems (WAMS) to improve the monitoring, protection, and control of power networks, some of these applications are discussed in Reference [52,53].

- **Power Quality Monitor**: The PQM is an advanced measurement device that collect different quantities, such as frequency, RMS voltage, currents, total harmonic distortion, individual harmonics, and flickers. The PQM have a maximum sample rate of 4 MHz, which is the best performance of all the presented devices, having a sensor accuracy of $\pm 0.1\%$ and $\pm 0.2\%$. This key feature allows to track transient events and harmonic sources in a distribution grid with high precision. The PQMs devices are being widely deployed in distribution systems due to
the growing penetration of renewable sources, energy storage systems, and electronic devices installed in distribution networks. Some recent applications of PQM devices used for power quality distribution monitoring are presented in Section 4.

- Substation Meters: Substation meters are wireless sensors located on transmission and distribution substations. Their main function is to measure active power, reactive power, complex current and complex voltage typically at a frequency rate of 1 Hz (1 measurement each second). A recent work has developed an online learning algorithm for energy management and energy balance for a three-phase distribution feeder. This algorithm uses sensor fusion to incorporate output equations from real measurements, such as active and reactive power flow, complex bus voltage measurements from distribution substation measurements, and residential smart meter measurements [40].

- Digital Fault Recorder (DFR): The DFR is an automatic recorder capable of storing transient fault events from different protection relays distributed at the substation level (Medium Voltage). Its main functionality is to automatically record and store events in a local database so that the operator has all the information to perform fault and post-mortem disturbance analysis with higher resolution and accuracy than power line sensors. The Digital fault recorder have a precise measurement recorder, sensor accuracy of ±0.1%, and a sampling rate of 512 samples per cycle (25.6 kHz), ideal for monitoring transients from flexible alternating current transmission systems and switching operational events [41].

- Wireless Power Line Sensor (PLS): The overhead line sensor is a low investment cost device used primarily to monitor the status of distribution lines, including line faults, line loads, power quality, conductor temperature and can also optimize the distribution network topology. Power line sensors report line status up to every 5 min at a sampling rate of 12.8 kHz and sensor accuracy of ±0.5%. The sensor can be charged with a minimum current flow of 1 ampere, saving the use of batteries. The main advantages are low maintenance, remote configuration and easy installation on live networks. The PLS can also send warning messages to distribution system operator when a fault event occurs [42].

3. Communication of Advanced Measurement Technology in Distribution Systems

High-speed communication on smart grids allows instantaneous monitoring of the network. This allows to anticipate possible incidents, manage those that occur more rapidly, and improve the quality of service to customers. Recently, communication infrastructure has been designed using broadband power line, Wi-Fi, and fiber optics for distribution network management services. Authors in Reference [54] compared the speed performances of each type of communication. The experimental results demonstrated that fiber optics is the fastest communication network, capable of providing a two-way latency of 3 ms. An important factor to consider in communication systems is infrastructure investment costs. Sometimes slower solutions, such as power line communications, could be more suitable for an specific application or when a high number of measurement devices have to be installed in distribution systems. Power Line Communication (PLC) is an example of a versatile and cost-effective means of communication for smart grid monitoring because it allows the power line to be used for both power and communication, thus eliminating the need for special cables to carry control and data signals. In Reference [55], different overhead communication PLC lines were compared for in several applications for electrical distribution networks. This section briefly describes the communication infrastructure of μPMUs, smart meters, and PQM devices.

3.1. μPMU Storage and Communication System

The μPMU network requires central data storage, precise synchronization, efficient analytical tools, and high-speed communications systems to collect and transport the data. Since μPMU data is sampled at a rate of 100/120 Hz, when these measurements are stored, it collects a considerable
amount of data over short periods of time. Therefore, an efficient server infrastructure must be able to support high data volumes and high-speed searches with advanced automated support [56].

The communication of µPMU requires high data rate, which can be obtained with optic fibers or high data rate wireless networks. These networks also require high infrastructure costs, installation and service fees. Sometimes a solution is chosen to reinforce the existing communication infrastructure in order to reduce investment costs. Some other communication systems have been designed. In Reference [57], authors proposed a data concentrator for synchrophasors based on a new logic performance scheme to minimize introduced latency, without corrupting or damaging the data measurements. This work also compared in terms of reliability, determinism and latency with traditional telecommunication infrastructures against fiber optic links (15 PMUs) and 4G LTE wireless network (10 PMUs). The experimental results show that the proposed logic is characterized by the lowest latency.

Works in Reference [8,58,59] have developed a database storage system to concentrate micro-PMU data with nanosecond-precision timestamps, from different distribution systems. This database storage system has an open-source software that can be run from hardware or from the cloud (where data can be stored). It stores large sets of data, with capacity to read over 19.8 million snapshots per second, that can be used to analyze big data problems using different processing techniques for real-time applications. Figure 2 shows the architecture design for a database network.

Figure 2. Micro-Phasor Measurement Units (µPMU) data system architecture (adapted from Reference [59]).

Furthermore, the Berkeley database offers different services for customers, i.e., it provides a plotting visualization service for clients, based on a web application. They also facilitate the remote access of historical and real-time data for the development, analysis, and implementation of novel processing techniques programmed in MATLAB, Python, and C++. Recently, predictive platforms have been developed for the analysis of real-time data using machine learning algorithms. In Reference [60], a platform was developed to process large data sets from µPMUs. This platform provides the opportunity for engineers and researchers to program lines of code and analyze the data stored from the web with recent machine learning techniques.

3.2. Smart Meter Communication

The key feature of smart meters is the two-way communication system between smart meters and consumers, and also from smart meters to utility providers. This interaction has several advantages in both sides because users are able to manage their own consumption by knowing their near-real time power consumption and variability of prices. On the other hand, energy providers are able to forecast future scenarios of consumption or generation, allowing utility operators to optimize their operations under different planning horizons.

The communication of data can be transmitted via wired or wireless connections. The main advantages of wired networks are the speed of communication and high security, but these require a high investment in their infrastructure. On the other hand, wireless communication based on the use the Internet of Things which has several advantages, i.e., the high-speed of data transmission, easy remote access, storage in the cloud of data, and low cost of infrastructure. Nevertheless, it can
present security problems and complexity in the integration of this system over the existing ones. Recent technologies are being deployed using the concept of internet for energy for distribution network communication. In Reference [61,62], the authors classified smart meter networks into Home Area Networks (HAN), Neighborhood Area Networks (NAN), and Wide Area Networks (WAN). The HAN is a local communication network between different devices around a household (e.g., appliances, home illumination, Electric Vehicles) and can also be controlled for efficient energy consumption. The NAN is the communication network between nearby households which are interconnected to a local data concentrator by a Wi-Fi network, Cellular, or power line communication. The WAN is the largest type of network, capable of covering long distances for monitoring those networks across municipal or regional boundaries. WAN are formed by sets of NANs, which are connected to a global concentrator or directly to a Network Operation Center using optical fiber, satellite, WiMAX, cellular, or digital subscriber lines. Then, the smart meter data is concentrated and stored in a big database cluster with all the power consumption and geographic information of customers. This data is then transmitted to a Network Operation center using fiber optic, cable, or Wi-Fi. Finally, different services are provided with smart meter data and GIS information, including energy billing, planning, demand side management, and outage side management. Figure 3 provides an overview of the advanced metering infrastructure in distribution networks. Some works have proposed innovative smart metering architecture for smart metering in distribution power networks based on the use of Power Line Communication (PLC) couplers. Authors in Reference [63] proposed a communication approach considering PLC signal concentrators, avoiding wireless solutions with the intrinsic installation and service provider costs. The low cost of this solution guarantees easy scalability, favoring its vast employment in modern smart cities.

Figure 3. Smart Meter Architecture Diagram (adapted from Reference [64]).

3.3. Power Quality Monitor Communication

The PQM is composed mainly by: (1) Power Quality (PQ) monitoring devices, (2) Network devices (for data transmission), (3) Server, (4) Monitoring centers, (5) PQ website, (6) PQ database, and (7) Users. In Figure 4, the general architecture of a PQ system is shown, and the process is briefly described below.
The PQM integrates two blocks, which are, a signal card block and a Field Programmable Gate Array (FPGA). The signal card block is used to collect the measured signals of currents and voltages. The FPGA processes these signals with a digital processing technique to determine the power quality indices, such as Total Harmonic Distortion (THD), distortion index, and power crest factor, among others. Firstly, the data and indices are transmitted instantly from a remote network device to the monitoring center by a User Datagram Protocol (UDP)/IP communication protocol, which means that the data packages are transferred by either an IP local Ethernet or using only an internet connection. Secondly, the data passes through a security network system and then is selected with an automatic software and stored in a database on the server. Finally, from a website, it is possible to access to the data collected; this enables administrators and users to connect where the Internet connection is possible. Power quality data files include real-time measurements and historical records.

4. Distribution Measurement Technologies: Application of Data

In this section, a review of recent applications and techniques used for data processing of μPMUs, smart meters, and PQM devices is shown, based on a literature review of the last four years. In addition, few relevant articles were considered due to their high impact of the research over the last few years. Figure 5 summarizes the overall groups of applications found in this work, considering only three advanced measurement devices (Micro-PMUs, Smart meters, and PQMs). It is important to mention that each measurement device work with different resolutions; therefore, the applications are also oriented to monitoring events with the same time response.
Recent applications of $\mu$PMU data are shown in this subsection. Table 2 shows a general summary of the applications, methods and input data found in more than 25 articles. A total of eight application groups were obtained, and two large groups can be highlighted, situational awareness and state estimation. These two groups require a high sampling rate of the measured data in order to visualize transitory events in the distribution network. Table 3 lists the input data ($\mu$PMUs measurements) required for each specific application. Most of the described methods are based on synchronized current and voltage phasors (magnitude and phase) to identify, analyze, and monitor possible failures in the distribution network. A brief description of each of these application groups and methods found for data analysis is provided below.

- **Situational Awareness**: Most of applications are driven to aware distribution operator of transient events due to the high sampling frequency and communication abilities of $\mu$PMUs. In Reference [66], the authors proposed a method based on the compensation theorem to detect abnormal events in distribution systems. This method generates an equivalent circuit using the current and voltage phasors captured by $\mu$PMUs. Similarly, a cumulative sum (CUSUM) algorithm was proposed in Reference [67] to detect anomalies having limited micro-PMUs. Validation results showed the effectiveness of this algorithm to voltage, current, and active power changes in the distribution system. In Reference [68], the authors proposed the multi-class Support Vector Machine (SVM) method to detect and classify abnormal events based on large volumes of data. A total of 1.2 billion real measurements of two micro-PMU installed in a distribution feeder were analyzed to evaluate their actual performance and were validated with two different methods, which are K-Nearest Neighbor (k-NN) and Decision Tree (DT). The results showed that the proposed technique can accurately identify a total of 10,700 events, outperforming the other two evaluated techniques. In Reference [69], a generalized Graph Laplacian Matrix (GLM) to visualize different voltage and current events in a real test feeder was proposed. Moreover, a kernel principle component analysis and a partially Support Vector Machine (pSVM) was used in Reference [70] for voltage sags detection based on data index and reconstruction error.
The effectiveness of these methods were tested on a real distribution network with \(\mu\)PMUs. In Reference [71], the authors proposed a Granger causality technique to analyze the frequency event propagation from the transmission network to the main feeder of a distribution network using \(\mu\)PMUs measurements. The authors also proposed a sparse coding method to determine the spectral frequency of abnormal events. The proposed approach was tested with real-time data from a public network located in Riverside, California. In Reference [72], the authors proposed a state estimator to identify faults in distribution lines using micro-PMUs. This estimator determines the error, using a weighted residual metric. Validation tests have shown that the proposed estimator correctly detects and locates distribution line failures in presence of bi-directional flows.

In Reference [73], an experimental analysis of lightning strikes was proposed using \(\mu\)PMU data collected during a day of rainstorms. The main interest of the study was to analyze the transient response of a 7.5 MW PV farm and its associated substation. Results showed the high resolution of micro-PMU to capture transients of current and voltage phasors during lightning-induced events.

Reference [74] proposed a parametric sparsity method to detect and locate events from distribution grids. An optimization algorithm based on particle swarm was proposed in Reference [75] to coordinate overcurrent relays installed in microgrids and distribution networks. In addition, a technique to identify uncertainties in real-time was also proposed. Authors in Reference [76] proposed a method to synthesize steady state models for multiple-sections of active distribution networks (unbalanced) using real-time PMU data. Additionally, a Kalman filtering technique was proposed to extract the quasi-steady state components, noise filtration, and outliers from PMUs. The results from two simulated events demonstrate that the proposed technique can produce an accurate model for any feeder configuration located between PMUs installed in active distribution networks. Authors in Reference [77] evaluated the transmission characteristics of a Rogowski electronic current transformer and an electronic voltage transformer (EVT) in a simulated and real testing platform. Experiment results showed that the EVT and the traditional power transformer have similar performance in the transient process of disconnecting switch breaking. Additionally, the power transformer was not affected by temperature changes, while that in the electronic voltage transformer the temperature had a great influence impact.

- **Topology Verification:** Distribution networks models are often imprecise or outdated. Topology identification is essential for monitoring and control distribution systems. The \(\mu\)PMUs devices are able to extract measurements from network nodes in real-time in order to track topology changes. In Reference [78], the authors proposed a technique to estimate impedances through a reduced Kron matrix also called “subKron” form. Additionally, a recursive clustering algorithm was implemented to reconstruct the topology of radial networks from line impedances. The results of the simulation showed that this technique is robust to measurements with additive noise that is generally captured by micro-PMUs; however, it has limitations when applied to large distribution networks. In Reference [74], the authors proposed an adaptive lasso technique to identify changes in topology caused by permanent failures in distribution systems. This technique is able to locate faults geographically in real time using PMUs that capture voltage and current phasors with high accuracy. The results of this work demonstrated the efficiency of this technique in different case studies. In Reference [79], the authors proposed a method to detect topology changes in distribution networks based on the Time-Series Signature Verification (TSV) method. This method considers the relationship that occur when there are changes in network topology. Validation results showed that the proposed method works satisfactorily with the partial knowledge of the state of the network. Authors in Reference [80] proposed a data driven approach based on the projection of a norm tren vector in to a topology library. This method was able to detect over 32 possible topology scenarios in a distribution grid.
Table 2. Recent application groups of $\mu$PMU Data.

<table>
<thead>
<tr>
<th>Application Group</th>
<th>Input Data</th>
<th>Methods</th>
<th>Output Visualization</th>
<th>Year (Reference)</th>
<th>Is It Real $\mu$PMU Data?</th>
<th>Simulation Data?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Situational Awareness</strong></td>
<td>3-Ph voltage and currents (magnitudes and angles), frequency, active power, reactive power.</td>
<td></td>
<td>Voltage magnitude change, current magnitude change, active power change.</td>
<td></td>
<td>1. Yes</td>
<td>1. No</td>
</tr>
<tr>
<td><strong>Classification of Events</strong></td>
<td>3 Ph-Voltage and currents (magnitudes &amp; angles)</td>
<td>1. Multi-class SVM</td>
<td>Event location and disruptive classes</td>
<td>1. 2019 [90]</td>
<td>1. Yes</td>
<td>1. No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. NN based autoencoders (Softmax)</td>
<td></td>
<td>2. 2019 [81]</td>
<td>2. No</td>
<td>2. Yes</td>
</tr>
<tr>
<td><strong>State Estimation</strong></td>
<td>Voltages, currents, line admittances, loads.</td>
<td></td>
<td>Estimation of voltage magnitude, loads, currents and errors from actual states.</td>
<td></td>
<td>1. 2020 [83]</td>
<td>1. No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. WLS, WLS with NR</td>
<td></td>
<td>2. 2020 [84]</td>
<td>2. No</td>
<td>2. Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Integer linear programming</td>
<td></td>
<td>2. 2019 [99]</td>
<td>2. No</td>
<td>2. Yes</td>
</tr>
<tr>
<td><strong>Model Calibration</strong></td>
<td>Frequency, voltage and current phasors.</td>
<td>1. Non-Linear Estimation</td>
<td>Calibrated parameters</td>
<td>1. 2020 [90]</td>
<td>1. No</td>
<td>1. Yes</td>
</tr>
<tr>
<td><strong>Operation Events</strong></td>
<td>Voltage and currents (magnitudes and angles)</td>
<td>1. Fuzzy C-means</td>
<td>Switching event, operational parameters (real and reactive power flow), voltage and current (feeders and loads)</td>
<td>1. 2017 [91]</td>
<td>1. Yes</td>
<td>1. No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Data driven analysis based on RLC Model</td>
<td></td>
<td>2. 2017 [92]</td>
<td>2. Yes</td>
<td>2. No</td>
</tr>
</tbody>
</table>
Table 3. Specific applications of Micro-PMU data.

<table>
<thead>
<tr>
<th>Micro-PMU Data</th>
<th>Specific Applications in Distribution Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Ph-I (Phasors):</td>
<td>Analysis of Transient Load Behaviors [91]</td>
</tr>
<tr>
<td>3 Ph-Currents (Phasor) &amp; Y (Admitance matrix):</td>
<td>Optimal number and localization of μPMUs in buses [84,89]</td>
</tr>
<tr>
<td>3-Ph-V (Phasors):</td>
<td>Estimation of Voltage Magnitude error [84] Identification of switching actions and new topology scenarios [74,79,80]</td>
</tr>
<tr>
<td>3-Ph-V (Phasors) &amp; Loads:</td>
<td>Load and Voltage Estimation Error [85]</td>
</tr>
<tr>
<td>3 Ph-V &amp; 3 Ph-I (Phasors):</td>
<td>Single and 3 phase current &amp; voltage event detection [69] Response of a PV farm (Current and Voltage) of 3 lightning events [73]. Detection of capacitor bank switching [92]</td>
</tr>
<tr>
<td>3 Ph-V &amp; 3 Ph-I (Phasors):</td>
<td>Fault position with accuracy, sensitivity to noise level [72] Distinguishing between two disruptive events and the normal load changing event [82]. Identify event location and disruptive classes [74,81]</td>
</tr>
<tr>
<td>3 Ph-V &amp; 3 Ph-I (Phasors):</td>
<td>Event Location and Identification [66] Event detection of voltage sag based on Data Index and Reconstruction Error [70]. Fault currents to coordinate relays [75]</td>
</tr>
<tr>
<td>3 Ph-V &amp; 3 Ph-I (phases), frequency:</td>
<td>Estimation of sub-transient generator model variables [90]. Cyclic frequency trend and anomaly signals detection [71]</td>
</tr>
<tr>
<td>3 Ph-V &amp; 3 Ph-I (Phasors), loads, Y (admitance matrix):</td>
<td>Tracking State Estimation [86,87]</td>
</tr>
</tbody>
</table>
• **Classification of Events**: The classification of disturbing events is responsible for quantifying abnormal events that occur in the system. Recent approaches of event classification have been explored. In Reference [68], the authors proposed the multi-class Support Vector Machine (SVM) method to classify abnormal events based on large volumes of data. A total of 1.2 billion real measurements of two micro-PMU installed in a distribution feeder were analyzed to evaluate their actual performance, and were validated with two different methods (K-Nearest Neighbor (k-NN) and Decision Tree (DT)). The results showed that the proposed technique can accurately identify a total of 10,700 events, outperforming the other two evaluated techniques. A neural network approach was proposed in Reference [81], using autoencoders along with soft-max classifiers to distinguish two disruptive events. The performance of the algorithm was tested to identify if a capacitor bank switching has a normal load change or if it has a malfunctioned switching. In Reference [82], authors proposed two different algorithms to classify disruptive events in distribution networks. The first algorithm was based on a hybrid combination of Principle Component Analysis (PCA) together with a multi-class SVM, and the second algorithm was with an auto-encoder along with soft-max classifier. Validation results showed the superiority of the second algorithm over the first algorithm in term of accuracy. The data for training and testing was simulated in the IEEE 13-bus distribution system.

• **State Estimation (SE)**: Distribution system state estimation (DSSE) is the minimum set of variables that can be used to describe the dynamic behavior of the system, advanced measurement devices are useful to quantify these variables. In Reference [83], the authors proposed a decentralized state estimator to improve the operating privacy in active distribution networks and microgrids. In this work, the iterative procedure based on quadratic programming was demonstrated, which used micro-PMUs as main inputs. The studies demonstrated a high accuracy of this proposed approach for different scenarios. In Reference [84], the authors proposed a regularized estimator to accurately identify the operating state of the system in a short time. This estimator operated with different measuring devices with different resolutions using data mainly from SCADA-type systems and micro-PMUs. This fusion of data allowed to provide greater robustness of the estimator to noise and less error in the estimation of states. The authors of Reference [85] proposed a weighted least square-based for distribution system state estimation, in which voltages and loads are chosen as state variables to compensate insufficient real-time measurements in medium voltage distribution systems. In Reference [86], the authors proposed a method based on the compensation circuit theory to generate an equivalent circuit. This method was able to estimate and follow the system states when sudden load changes occurred. This method used real measurement from µPMUs in a distribution system. In Reference [87], the authors proposed a simple method to determine the state variables based on power line data and bus voltage phasors from micro-PMUs installed in a distribution network. The authors showed that the proposed method can be robust to noise measurements, high levels of distributed generation, and a reduced number of measurements. Authors in Reference [88] proposed an open testbed to evaluate and compare PMU estimation algorithms accuracy under experimental conditions, considering the noise propagation in order to quantify the uncertainty contributions and their impact on the estimates of the variables.

• **Optimal Placement**: The optimal placement of µPMUs aims to maximize the observability of the distribution network while minimizing investment costs. In Reference [84], the authors propose a D-Weighted Total Variation State Estimation (WTVSE) algorithm to estimate system states with a reduced time scale (every 15 min), considering the observations of a SCADA system and micro-PMUs. In addition, a semi-defined scheduling model was proposed to optimally locate micro-PMUs and thus improve state estimation. The results of the simulation of a 95 bus distribution network showed that this proposal presents a great accuracy in the estimation of states under a diversity of scenarios, in addition to its low computational complexity. In Reference [89], the authors proposed a linear programming model to optimally locate phasor measurement
units in distribution networks. The aim of this model was to ensure observability during possible changes in topology by operational actions. The results obtained from a medium voltage distribution network in southern China showed that the proposed method is efficient and robust to topology changes. In Reference [67], a greedy search algorithm was proposed for optimal µPMUs placement. This algorithm uses an optimal location criterion to achieve maximum observability and therefore increase the monitoring range considering different event scenarios.

- **Model Calibration:** Dynamic models can be calibrated based on real-time advanced measurements. This is a new field of application that promises to improve the current models. In Reference [90], the authors proposed a methodology to enhance the synchronous generator model, based on PMU measurements. First, the estimation of the variables (frequency, voltage, and current phasors) of the dynamic state were obtained. Then, the authors calibrated the inertia constant and the reactances of the model. Finally, the performance results were obtained under different perturbation scenarios. The authors conclude that the calibration of parameters in real time requires high accuracy of advanced measurement devices.

- **Operation Events:** The high resolution of the µPMU data allows to observe the dynamics (transient) of operational events that generally occur in distribution networks, such as the reconnection of microgrids, the connection of loads, and/or the connection of capacitor banks. In Reference [91], the authors proposed a method to analyze the transient behaviors caused by the addition of flexible loads/generation in distribution feeders. This approach modeled the load profiles based on the collection of data from various µPMU located at the low voltage level. The authors demonstrated that it is possible to compromise network reliability if several flexible regulation resources are located on the same feeder. In Reference [92], the authors analyzed the switching events of a three-phase capacitor bank to determine the operational parameters and the flow of reactive energy from a capacitor. The authors conducted an experimental study based on real measurements from µPMUs that were installed in an electrical distribution network. The results showed that the magnitude of the transient current of the feeder depends on the initial condition and the phase angle at the time of capacitor switching.

### 4.2. Smart Meter Data Applications

The following subsection describe the applications of smart meter data. The objective is to present the methods developed in recent years, based mainly on machine learning techniques for the processing, prediction and monitoring of the distribution network. Table 4 shows a summary of the applications, methods and general input/output data of 37 relevant articles published in recent years. The classification of smart meters applications were divided into eight groups; however, the most prominent groups are the forecasting group and the topology identification group. These applications are mainly used by operators to monitor and control the electrical distribution network. On the other hand, Table 5 shows a relationship between the required input data, such as consumption, voltage, and current profiles (captured by Smart Meters), with multiple applications found in the literature. Most of the applications are based on household consumption profiles for fraud identification, prediction of energy consumption (short and long term), and topologies identification.
Table 4. Recent application groups of Smart Meter Data.

<table>
<thead>
<tr>
<th>Application Groups</th>
<th>Input Data</th>
<th>Methods</th>
<th>Output Visualization</th>
<th>Year (Reference)</th>
<th>Is It Real SM Data?</th>
<th>Simulation Data?</th>
</tr>
</thead>
</table>
Anomaly Detection: Anomaly detection techniques are useful to detect abnormal conditions in smart meters data, such as suspicious consumption, missing data, or unplanned events. In Reference [93], authors proposed an algorithm based on isolation forest to detect the presence of anomalies in smart grid using non-labeled data. Authors also used a Principal Component Analysis (PCA) to compress the volume of data in a distribution system. In Reference [94], authors implemented a support vector regression with a random forest method using sliding windows, to identify anomalous consumptions of real-world data. In Reference [95], the authors developed an anomaly detector system, capable of analyzing smart meter measurements from a data concentrator in near real-time. The design of this system can detect abnormal conditions at medium and low voltage levels using a quasi-linear classifier. Reference [96] proposed a technique with a Lambda architecture to detect anomalies. This technique was based on real-time consumptions that were analyzed through supervised learning. Additionally, a threshold index for the detection of suspicious consumption was considered. The evaluation of this algorithm in real scenarios demonstrated its effectiveness and precision in detecting anomalies.

Compression of Data: Data compression techniques help to reduce the volume of data collected from advanced measurement devices; they also help to improve the transmission speeds from multiple measurement points. In Reference [97], the authors proposed a deep learning technique with a convolutional dispersion auto-encoder for data compression. This method keeps more information than Singular Value Decomposition (SVD) and PCA methods, at the same coding speed, preserving details of the original power, and the calculation times are lower. In Reference [98], the authors proposed a neural network based on an automatic encoder to compress household consumption data in a distribution network. This proposed encoder must be installed on the user’s side to compress the smart meter readings. Compared to some existing linear compression models, such as PCA, DWT, and SVD, the SAE compressor has lower % errors according to a study carried out with real data from China and Ireland. Similarly, in Reference [99], a methodology using the SVD technique for data compression was presented. This methodology was used in a test system with data from different substations of a UK company. This technique achieves a significant reduction in the volume of data to be transmitted, with minimal error in its reconstruction. Reference [100] proposed a SVD sparse coding technique to compress smart meter data. This dispersion technique extracts the information using linear combinations from load clusters. The results obtained comparing 4 techniques showed that the proposed technique obtains the least loss of information.

Customer Characterization (Socio-demographic): Predictive analysis can also be applied to determine the characteristics of network consumers, for example, predicting the number of unemployed people, number of occupants in a building and/or predicting daily household activities. In Reference [101], authors compared six machine learning models to determine the number of unemployed people in a household. The overall results showed that the most accurate models were the multi-layered perception and distance-weighted discrimination approach. Similarly, in Reference [104], a neural model was proposed to determine the employment situation of consumers. This type of information can help governments to reduce unemployment levels and also help to improve their economy. In Reference [102], the authors implemented a genetic algorithm to identify the number of occupants in a residential building using smart meter data. Validation results showed that this algorithm can optimally predict the number of occupants in households. In Reference [103], an automatic learning model was proposed to identify characteristics of residential occupants, e.g., people living in the household, average age, and daily activities in the household, from the daily electricity consumption of the users. The validation of this model was implemented in a real distribution network in Ireland, in which technical characteristics allow obtaining this information.

Forecasting: The optimal generation planning requires that operators have tools to predict demand growth in a short and medium terms. In recent years, techniques based on machine
learning have been developed, considering multiple variables in the prediction models, including electricity consumption, weather conditions, electricity tariff costs, and population growth that contributes to generate accurate prediction models. In addition, with the growing technological development, it is possible to concentrate all this data in real time, which allows the parameters of the models to be systematically updated. In Reference [106], the authors designed a short-term prediction model based on a Q-learning scheme that used meteorological data and smart meters as input variables. This scheme was composed of ten deterministic prediction models and four probabilistic heuristic models which were selected based on their accuracy. The results presented demonstrate a higher accuracy of Q-learning predictions than traditional approaches. In Reference [105], an artificial neural network with a multiple regression technique was proposed to predict load consumption using temperature and solar irradiation variables in the model to obtain more accurate predictions. The validation of this proposal was demonstrated in a real data set of smart meters that included photovoltaic generation. In Reference [107], the authors proposed a recurrent neural network technique to predict consumption in short-term scenarios. This proposed approach was tested on a public data set on real residential consumption and compared with other techniques for validation. In Reference [108], the authors proposed a Nest-Bcktr algorithm for short-term load forecasting. A comparison between other six machine learning algorithms were made, in terms of RMSE indices and absolute errors. The validation was programmed in Python using a 2-year data set from smart meters. The results showed that this proposed algorithm predicts consumption with lower errors. In Reference [109], a load ensemble method to forecast aggregated loads was proposed. This method produces multiple training and prediction models with different sub-profiles. In addition, a weighted optimization is used to combine and determine the best prediction. In Reference [111], an additive regression model was proposed to forecast the distribution of electricity consumption added to the network. This model generates different probability scenarios that help operators to plan and operate the network in the future. In Reference [112], the authors used a method based on kernel density estimation to forecast future growth of electricity consumption. This method considered predictions of electricity costs for different tariffs, which means potentially important savings for users.

- **Load Classification:** Refers to the grouping of electrical consumption including residential, commercial and industrial loads. In recent years, various grouping techniques have been proposed with data from smart meters that have provided useful information to distribution system operators. In Reference [113], electricity consumption and energy tariff variability were analysed in four different seasons. In this analysis, several statistical tools were applied to analyze different energy consumption’s using real data from smart meters. The study showed that users consume more energy when they do not know the variability of energy costs. The authors recommended that consumers learn about tariff dynamics in order to minimize energy consumption costs. In Reference [114], the authors proposed a prediction technique based on a trained auto-encoder that analyzed smart meter data and also grouped them using a self organizing map. In Reference [115], a finite mixture model based on a variant Gaussian distribution to identify non-typical behavior in the distribution system was proposed. This model classified different customer profiles according to their load levels and variability. In Reference [116], the authors proposed a k-means clustering algorithm for phase identification of interconnected customers in the network. This algorithm uses as inputs the voltage signals from smart meters and SCADA measurement system. The test results obtained from a distribution network in California showed that the algorithm has an overall accuracy above 90%.
Table 5. Specific Application of Smart Meter Data.

<table>
<thead>
<tr>
<th>Smart Meter Input Data</th>
<th>Specific Application in Distribution Network</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household Load Profiles (kWh):</strong></td>
<td>Identification of anomalous consumption patterns based on classifiers [94]</td>
</tr>
<tr>
<td></td>
<td>Storage and transmission of large volumes of power consumption data [97,98,100]</td>
</tr>
<tr>
<td></td>
<td>Consumption profiling and prediction of energy consumption [110,114]</td>
</tr>
<tr>
<td></td>
<td>Short-Term Load Forecast in Residential Buildings [107–109]</td>
</tr>
<tr>
<td></td>
<td>Online Anomaly consumption detection [96,118]</td>
</tr>
</tbody>
</table>

| Household Load Profiles (kWh): | Identification of time periods during the day and seasonality [112,115] |
|                                 | Phase identification based on clustering [116]                                |
|                                 | Non-technical losses identification [119]                                    |
|                                 | Load profiling based on the energy tariffs across four climatic conditions [113] |
|                                 | Forecasting uncertainty in electricity data [111]                           |

| Load Profiles and geographical data: | Detection of Non-Technical Losses [110]                                    |
|                                      | Identification of network topology and load phase connectivity [124]         |
|                                      | Topology Identification via Graphical Modeling [126]                         |

| Voltage Magnitudes and Household Load Profiles (kWh): | Network Topology identification [127]                                      |

| Voltage and Current Magnitudes, load profiles and admittance matrix of the network: | Inferring operation modes of distribution networks [122]                    |
|                                                                                      | Identifying the connections, as well as the voltage correlations between different buses [122] |
|                                                                                      | Distribution system Low-Voltage Circuit Topology Estimation [125]            |
|                                                                                      | Approximation the missing cable information in LV networks (cable's cross section area (XSA) data). [123] |

| Load profiles and socio-demographic attributes of households: | Unemployment prediction of single household occupants [101]                 |
|                                                               | Household Load Forecasting based on clustering [105]                       |
|                                                               | Predict the home occupancy status of households [102]                      |
|                                                               | Predict multiple household characteristics (e.g., age of person, household income, cooking style, etc.) [103,104] |
• **Non-technical loss detection:** Detection of non-technical losses are basically electricity theft consumers, faulty meters or billing errors. In Reference [110], a hybrid deep neuronal network to detect non-technical losses in smart meters was proposed. This algorithm was tested with real smart meter data from the largest electric utility in Spain. Validation results showed the accuracy of this approach to identify anomalies in distribution systems. In Reference [117], the authors proposed a deep convolutional neural network to detect non technical losses in distribution grids. This approach detected manipulations of consumer energy readings that falsely overloaded the power company. The results obtained in this work indicate that the fusion of multiple data, including smart meters, SCADA systems, and meteorological reports, contributes to the accurate detection of energy theft consumers. Reference [118] proposed a methodology based on the hybrid combination of decision tree and support vector machine classifiers to detect fraudulent consumption. In Reference [119], the authors proposed a classifier based on the optimal-path forest algorithm to detect anomalies and non-technical losses in distribution networks. This machine learning technique requires training from regular consumer profiles in order to generate a sample group base, and, when a new consumer connects with irregular profiles, he is automatically identified. Validation results showed that this technique is robust and accurate for classifying different types of consumers.

• **Sensor Fusion:** Sensor fusion is the integration of data from smart meters with other measurement devices and is intended to improve the observability and accuracy of monitoring distribution systems. In Reference [120], the authors proposed a state estimator to identify harmonic sources in an unbalanced distribution system. The state estimator was based on neural networks and Bayesian learning, and the input signals were captured by smart meters and micro-PMUs. Validation results showed the high accuracy of the estimator even in presence of distributed generation. In Reference [40], the authors proposed an algorithm to disaggregate loads from a distribution feeder into N components. The main objective was to separate network losses and reactive power injections from capacitors. This algorithm was based on a learning approach and used multiple measurement sensors to determine the technical feasibility of separation. Validation results indicate that data fusion of reactive power measurements in the algorithm can improve the accuracy in the prediction of the network behavior up to 32%. In Reference [121], the authors proposed a mixed integer linear programming algorithm to determine fault locations and prediction of outage regions. This algorithm requires of smart meter data and remote fault indicators measurements in near real-time in order to support distribution system operation in a precise time-step.

• **Topology Identification:** Information on the topology of the distribution network helps the operator to make optimal decisions when unexpected events occur. Authors in Reference [122] proposed a physical probabilistic network model to identify the connections using voltage correlations between different buses. This method was compared with a lasso regression method. In Reference [123], a tree-based search methodology was proposed to approximate the missing cable information in low voltage distribution networks. In Reference [124], a method to identify the connectivity between load phases in distribution networks was proposed. Additionally, the presence of technical losses and some errors that may arise during measurements (missing data, synchronization) were considered. This method implemented the principal component analysis to infer the topology of the use of smart meter measurements. This method proved to be robust in the presence of distributed generation. In Reference [125], the authors presented an algorithm for topology estimation based on voltage measurements from smart meters. Validation results showed that 9 out of 10 of the estimates were correct in secondary circuits of a Georgia Tech distribution system, even in noisy environments. The authors mentioned that it is extremely important to have ultra-precise measurement devices for correct estimation of voltage drop based topology, especially when analyzing short lines feeding small loads. In Reference [126], a graphical model to identify distribution topologies based on a probabilistic relationship between different
voltage measurements was proposed. Additionally, the authors proposed an expansion-tree based algorithm aimed at minimizing the Kullback-Leibler divergence in a distribution system. In Reference [127], an algorithm to correct connectivity errors of smart meters and meters on distribution feeders was developed. This algorithm identified the neighboring meters through a voltage profile correlation analysis.

4.3. Power Quality Monitoring Applications

In this last section, recent applications of PQM devices are shown, considering different methods. Table 6 shows a general summary of the applications, methods, and PQM input/output data obtained from 18 articles published in recent years. The PQM applications are divided into six groups, of which the optimal placement group can be highlighted by the number of publications in recent years. One of the groups included in this table was the power quality monitoring systems, which is basically the application of PQM in the distribution network of some countries that have carried out projects to improve the quality of transported energy.

- **Optimal Placement**: This large application group describes some recent approaches to determine the optimal positioning of PQM, with the aim of minimizing network investment costs. In Reference [128], the authors implemented the TLBO algorithm to optimally locate PQMs by considering degradation in large distribution networks. The objective of this approach was to minimize the number of PQMs in order to minimize the costs of assets in the monitoring system. In Reference [129], the authors proposed the MEAT optimization algorithm to find the best locations to install advanced PQM in distribution network. This proposed approach had multiple objectives, such as minimizing monitoring investment costs, minimizing voltage drops, and maximizing system observability. The authors recommend this approach for those electricity companies that need to evaluate the investments they will make to optimally improve network observability. In Reference [130], authors proposed the seeker optimization algorithm to find the optimal locations of PQM devices in a 14-bus test system. The test system results showed that with few locations of the PQMs the values of the harmonic state were accurately estimated. In Reference [131], an optimization algorithm based on Bayesian network models was proposed. The objective was to minimize the investment costs of PQ monitoring devices and to maximize the observability of the distribution network. Evaluation results showed that this algorithm significantly reduced the uncertainty of PQ values on unsupervised feed links. In Reference [132], the authors proposed a probabilistic method to observe the uncertainty associated with high/low impedance faults in distribution systems. The objective was to determine the optimal location of PQM devices to maximize observability in the system. In addition, two indices were proposed in this work to quantify the robustness of distribution networks with different voltage drops. The authors in Reference [133] mathematically analyzed the impact of the accuracy of state estimation (with power meters) by varying the spatial distribution and number of devices installed in the network. The objective was to minimize the number of devices to be installed and to identify the optimal location in the distribution networks, ensuring a desired accuracy in the estimation of voltage and current. The results show that the proposed mathematical framework is a useful tool for the design of optimal device placement strategies in current monitoring systems.

- **Fault Location**: Due to the high sampling rate and precision of this device, some authors have proposed algorithms to track faults in distribution systems. In Reference [134], the authors proposed a PQ disturbance predictor based on a Multi-Hidden Markov Model (MHMM). This predictor analyzes large volumes of data, including local weather variables to improve forecast accuracy, and also incorporates a Hadoop system that reduces calculation times for very complex systems. The forecast of this model can be adapted to different resolutions (minutes, hours, days, or up to 3 weeks). In Reference [135], the authors proposed a multivariate data analysis technique to locate line failures in unbalanced distribution systems.
<table>
<thead>
<tr>
<th>Application Group</th>
<th>Input Data</th>
<th>Methods</th>
<th>Output Visualization</th>
<th>Year (Reference)</th>
<th>Is It Real PQM Data?</th>
<th>Simulation Data?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data error detection</strong></td>
<td>Voltage and current phasors, THD, TDD, and short term flicker.</td>
<td>1. Correlation Analysis</td>
<td>Detection and correction error in PQ monitoring data.</td>
<td>1. 2017 [143]</td>
<td>1. Yes</td>
<td>1. No</td>
</tr>
<tr>
<td><strong>Load Modeling</strong></td>
<td>RMS voltage and current, Active and Reactive Power, during disturbances on the upstream networks.</td>
<td>1. Load parameter derivation</td>
<td>Derive, test, and verify the dynamic load model parameters.</td>
<td>1. 2013 [144]</td>
<td>1. Yes</td>
<td>1. Yes</td>
</tr>
</tbody>
</table>
This technique was based on PQM devices installed in distribution substations, line parameters and the topology configuration. The authors concluded that the proposed technique benefits operators to accelerate the tasks of system restoration (permanent faults). In Reference [136], the authors designed a power quality software to identify and locate faults in distribution feeders. This software performs a short circuit analysis based on historic and real measurements of PQM. The validation results showed that this software has better accuracy for locating faults in distribution feeders than some commercial software.

- **Harmonic Analysis**: These applications refer to detecting harmonics or abnormal behaviors in distribution systems using advanced PQ devices. The authors in Reference [137] presented multiple techniques to locate harmonic sources in distribution grids using PQ data. The objective was to design strategies to mitigate potential problems. Additionally, in this work a harmonic compliance index was presented, which allows to give a quick indication about violations of the permissible harmonic limit in a particular site. In addition, a graphical method based on harmonic reports showed a wide detail of harmonic performance in many sites in a compact form. The authors in Reference [138] presented several digital processing techniques to detect missing or abnormal data. The validation tests were implemented in 8 German networks with residential, commercial and mixed customer loads. The authors concluded that: "The identification of useful information cannot be manual anymore and requires a comprehensive set of intelligent and automated analysis tools". Authors in Reference [139], compared the robustness, flexibility, and limitations of a composite bus index and an aggregate bus index. These two indices were proposed and validated in a test system to evaluate the PQ of buses installed in distribution networks. The authors concluded that these indices were closely related and it is important to provide an adequate weighting in order to have a greater flexibility between them.

- **Power Quality Monitoring System**: Power Quality Monitoring Systems (PQMS) have been implemented in several countries to improve the power quality in distribution systems. In Reference [65], the authors designed a power quality monitoring software based on real-time data. This software was capable to analyze complex power quality problems using a FPGA-type hardware that worked as an independent integrated system. The authors concluded that these modern monitoring systems substantially improve the life of the assets that make up the smart grids. Authors in Reference [140] presented a project report of a power quality monitoring system which has been in operation in Australia since 2002. The objective of this report was to provide a general overview of the main problems found during the development of the PQMS. The authors concluded that monitoring systems with advanced measurement devices capable of providing PQ indices will rapidly increase in future power grids. Authors in Reference [141] designed a PQ monitoring system for a new generation of substations located in Shanghai, China. This system was based on international communication standards that allow remote monitoring of harmonics at a frequency of 12.8 kHz. This system is capable to analyze complex power quality problems, as well as the location of harmonic sources in the distribution networks. An improved hardware/software architecture with a real-time monitoring and control system for the integration of micro grids into MV distribution networks was presented in Reference [142]. The proposed system is capable of estimating the power flows of the medium voltage branch by means of load power measurements and a suitable load flow algorithm. The proposed system was considered more efficient than SCADA implementation.

- **Data error detection**: Data error detection is the process where the system automatically detects and corrects errors in PQ monitoring data. In Reference [143], an automatic detection and correction of errors system was developed based on data captured by PQ meters installed on a UK smart grid. The objective of the automatic system was to reduce the number of errors caused by various factors, such as poor installation of the devices, poor synchronization between multiple PQMs or by non-captured data (missing data), and maximizing the useful data for future
network operations. The authors conclude that the number of errors can be reduced considerably by adopting a correct installation procedure for PQ monitoring devices.

- **Load Modeling:** To represent the loads of an electrical network, mathematical models are used to simulate the dynamic or static behavior considering the active and reactive power of the load with respect to the variation of the voltage and frequency. In Reference [144], the authors derived the parameters of a dynamic load model of an 11 kV distribution network using a power quality monitoring system. The measurements used in the load model includes voltages, currents, active power, and reactive power at a sampling frequency of 1.6 kHz. To validate these results, the distribution network was simulated in a software using the parameters of the load model obtained and the load response to a disturbance was compared against a real disturbance in the distribution system captured by the PQMs. The general conclusion of this work was that 30–40% of the commercial loads considered in the distribution network are composed of induction motor loads, and if you want to make an accurate load model at the distribution level it is imperative to consider them.

5. Conclusions

This work made a comparison of eight advanced measurement devices for distribution networks based on their technical characteristics, including the sample frequency, reporting periods, measuring data, costs, precision, and time response. The comparative results showed that micro-phasor measurement unit and power quality monitor devices have the best performance overall to track dynamic and transitory events in distribution systems, due to their high-precision measurements, communication systems and remote storage of the extracted data.

This work also reviewed the most recent applications of µPMU, smart meters, and PQM data, considering novel methods and techniques. In addition, an input-output table that relates measured quantities from µPMU and smart meters needed for each specific application was developed in this review. From the extended literature reviewed in this work, the following conclusions are drawn:

1. The dominant applications of interest for µPMU data is currently leaning towards analyzing situational awareness events and estimating the state variables of the system in near real-time. With the extremely high resolution (sampling rate up to 30,720 s/s), amplitude accuracy of 0.05%, and angle accuracy of 0.01%, it is possible to visualize transitory events in the distribution network. The sensor accuracy can have a strong influence on the uncertainty of the quantities to be measured and thus can highly impacting in the algorithms performance.

2. The dominant application of interest of smart meter data is currently driven to forecast future load consumption in a short term horizon based on artificial intelligence, machine learning, and deep learning techniques. Topology identification is also of current interest due to the limited knowledge about the topology of low voltage networks. Some novel methods are related to correlation techniques and graph theory methods.

3. The most recent applications of PQM devices are related to find the optimal placement of the PQM based on multiple objectives, focusing on minimizing the cost of monitoring, minimizing topological ambiguity and maximizing the load monitoring.

The integration of µPMU, PQM, and smart meters is an alternative to improve visibility, precision, and security in active distribution systems. However, the large amount of data generated with the use of these devices is a challenge that demands high computational complexity and the development of efficient algorithms with the ability to process information in real-time. Data connectivity with different resolutions, parameters, and locations is a challenge that requires further investigation.
Author Contributions: A.E.S.-G. and A.S. conceived the scope of the paper; M.S. and M.A.-P. provided technical support and bibliographic recommendations of the latest research topics. A.E.S.-G. wrote the paper with support from A.S., M.A.-P. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been funded under the European Project RESOLVD of the Horizon 2020 Research and Innovation Program (topic LCE-01-2016-2017), grant agreement number: 773715.

Acknowledgments: M.A.-P. is lecturer of the Serra Húnter program.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- ANN: Artificial Neural Network
- ARMA: Auto-Regressive Moving Average
- CA: California
- CART: Classification and Regression Trees
- CCAD-SW: Collective Contextual Anomaly Detection using a Sliding Window
- CNN: Convolutional Neuronal Network
- CUSUM: Cumulative Sum
- DFT: Discrete Fourier Transform
- DFR: Digital Fault Recorder
- DNN: Deep Neural Network
- DSSE: Distribution System State Estimation
- DSTE: Distribution System Topology Estimation
- DT: Decision Tree
- DWD: Distance Weighted Discrimination
- DWT: Discrete Wavelet Transform
- FF-ANN: Feed-Forward Artificial Neural Network
- FPGA: Field Programmable Gate Array
- GBM: Gradient Boosting Machines
- GLM: Graph Laplacian Matrix
- GMM: Gaussian mixture model
- HAN: Home Area Network
- Hz: Hertz
- k-NN: k-Nearest Neighbor
- KF: Kalman Filter
- kV: Kilovolts
- LAMDA: Learning Algorithm for Multivariable Data Analysis
- LSTM: Long Short Term Memory
- LWLS-SE: Linear Weighted Least Squares State Estimator
- MHMM: Multi-Hidden Markov Model
- ML: Machine Learning
- MLR: Multiple Linear Regression
- MW: Megawatt
- NAN: Neighborhood Area Network
- NARX: Non-linear AutoRegressive with eXogenous
- NN: Neuronal Network
- NR: Normalized Residuals
- NTLs: Non-Technical Losses
- PAA: Piece-Wise Aggregate Approximation
- PCA: Principle Component Analysis
- PDCs: Phasor Data Concentrators
- PF: Path Forest
PLS Power Line Sensor
PMRA Probabilistic Monitor Research Area
PMU Phasor Measurement Unit
PQM Power Quality Monitor
PQMS Power Quality Monitoring System
pSVM Partially Support Vector Machine
PV Photovoltaic
R-NESE Regularized version of the Normal Equations based State Estimation
RF Random Forest
RMS Root Mean Square
SAE Stacked Autoencoder
SCADA Supervisory Control And Data Acquisition
SCSAE Stacked Convolutional Sparse auto-encoder
SM Smart Meter
SOM Self Organizing Map
SVD Singular Value Decomposition
SVM Support Vector Machine
SVR Support Vector Regression
THD Total Harmonic Distortion
TLBO Teaching Learning Based Optimization
TSV Time-Series Signature Verification
UDP User Datagram Protocol
WAMS Wide-Area Measurement Systems
WAN Wide Area Network
WLS Weighted Least Squares
WSMW Window Size with a Moving Window
WT Wavelet Transform
WTVSE Weighted Total Variation State Estimation
µPMU Micro-Phasor Measurement Unit

References


64. Ikpehai, A.; Adebiyi, B.; Rabie, K.M. Broadband PLC for clustered advanced metering infrastructure (AMI) architecture. *Energies* 2016, 9, 569. [CrossRef]


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).