Cell Reports
Physical Science

CellPress
OPEN ACCESS

Perspective

Unleashing the benefits of smart grids by overcoming the challenges associated with low-resolution data

Rui Yuan,[1,*] S. Ali Pourmousavi,[1,4] Wen L. Soong,[1] Andrew J. Black,[2] Jon A.R. Liisberg,[3] and Julian Lemos-Vinasco[3]

SUMMARY

Smart meters have been widely deployed worldwide, but there is an often-overlooked problem that remains unresolved: the data collected from these meters is of relatively low time resolution, hindering the realization of smart grid benefits. This perspective unfolds the roadblocks to achieving high-resolution data from a smart metering infrastructure. We highlight the loss of critical information, essential for many smart grid applications, due to low-resolution readings of residential consumer data. We then outline the main reasons behind the lack of high-resolution data, the tetralemma on balancing data collection, transmission, warehousing, and privacy concerns. Finally, we hypothesize a framework for data collection, maintenance, communication, and storage of smart utility meters data to obtain high-resolution records by tackling the challenges using a dictionary-based compression method and separately maintaining the compressed products at the user ends and data center.

INTRODUCTION

During the last two decades, the deployment of smart meters has gained momentum around the world, with the market size of smart meters estimated at 23.1 billion USD and expected to reach 36.3 billion USD in 2028. The Australian Energy Market Commission has recommended 100% uptake of smart meters by 2030,[1] and the EU expects 225 million smart electricity meters installed by 2024, corresponding to a penetration rate of 77%.[2] This is timely with the advancements in big data, machine learning, and demand-side management techniques that enable service providers to tap into the new energy data stream to create value-added services.[3] Out of this paradigm shift, the concept of energy digitalization was born[4] to develop tools for data-driven decision-making through data transformation into actionable insights. As a result, smart meters, as the cornerstone of energy digitalization, were expected to offer numerous benefits to end users, such as energy monitoring, alerts for overconsumption, lower moving costs due to remote power on/off control, flexible pricing, such as time of use (ToU) and, more importantly, new and expanded services that could only be envisaged in a digital energy system ushering in a smart grid era.

Although some of these benefits may have been realized in some jurisdictions, the widespread adoption of smart meters did not necessarily lead to full-scale energy digitalization, particularly in terms of delivering financial benefits to consumers. According to the Mission:data Coalition,[5] less than 3% of today's smart meters in the United States fulfill the 2009 promises of customer savings in the American Recovery and Reinvestment

[1]The University of Adelaide, School of Electrical and Mechanical Engineering, Adelaide, SA, Australia

[2]The University of Adelaide, School of Computer and Mathematical Sciences, Adelaide, SA, Australia

[3]Watts A/S, Research & Department of Data Science, Køge, Zealand, Denmark

[4]X (formerly Twitter): @AliPourmousavi

*Correspondence: r.yuan@adelaide.edu.au

https://doi.org/10.1016/j.xcrp.2024.101830

	Penetration	Resolution of reading	Major smart meter manufacturer	Meter owner	Accessible to third parties	Communication protocol	Reported year	References
SWEDEN	100%	1 hour	Kamstrup	DSO	Accessible		2020	[12, 13]
ESTONIA	99%	1 hour	Landis+Gyr	DSO	Accessible		2020	[15, 29, 30]
CANADA	82%	1 hour	Kamstrup	IESO, Utilities	On request		2022	[14, 16, 17]
DENMARK	80%	1 hour	Kamstrup	DSO	On request	In EU: M-bus, Zigbee	2020	[15, 18, 27]
CHINA	80%	15 minutes	Sanxing Medical Electric	Energy providers	Not accessible		2022	[19, 29]
U.S.A	78%	15 minutes-1 hour	Duke Energy	Suppliers, Users	Based on utilities	In Asia Pacific: PLC, Zigbee, Cellular, RF	2023	[20, 21, 29]
JAPAN	64%	15-30 minutes	TEPCO	Suppliers, Users	Based on utilities		2022	[23, 24, 29]
UK	52%	30 minutes	Sensus	Suppliers	On request		2022	[8, 15, 25, 29]
AUSTRALIA	24%	30 minutes	Intellihub	Suppliers	On request (charged)		2023	[1, 26, 28]

Figure 1. A summary of the current status of the smart meters rollout in selected countries
The figure outlines usage, time resolution, and meter details across nine selected countries.[1,8,12–30]

Act.[6] In Australia, the Australian Energy Council highlighted that the current metering arrangements fail to effectively realize the key potential benefits to customers.[7] One report pointed out that, in the UK, 7% of meters are smart meters in the traditional mode in which they lost simple and smart functionality, such as automatic readings and bidirectional communications.[8] The European Commission reported that 7 out of the 16 countries undergoing widespread deployment of smart meters lacked the ability to even relay consumption data back to consumers.[9] In Canada, wireless smart meters are read every 30 or 60 days,[10] which is not much different from traditional manually read electricity meters. We summarized the current state of the smart meter rollout in selected countries in Figure 1, including the ownership of the smart meters, the accessibility of the data from the smart meters, and the reading resolution, suggesting that current smart meter implementations are discriminatory, lack interoperability, and produce low-resolution data. This is despite that the internal sampling rate of smart meters is typically between 1 Hz and 1 MHz.[11]

In this perspective, we focus on issues and challenges related to data granularity on utility meters for residential consumers, which is the main barrier to seeing the benefits of the smart grid. As shown in Figure 1, smart utility meters predominantly provide readings within timescales ranging from 15 min to an hour. Taking into account the proliferation of distributed energy resources (DERs) and electric vehicles (EVs) in modern power systems, we show that the current resolution of smart meter readings does not support a broad spectrum of smart grid applications and services due to information loss. To do so, first, the challenges of the existing smart metering infrastructure with low-resolution data will be explored in the subsequent section. Then, we identify the key reasons and roadblocks behind achieving high-resolution readings. Subsequently, we propose a potential solution to enable modern smart meters to be operated in a truly smart way. This framework facilitates the acquisition of high resolution, possibly real-time energy consumption data through a dictionary-based compression method. The data are compressed into daily patterns and their representations with only a fraction of information transmitted to the data center. It offers several advantages over the existing metering infrastructure, including low storage and communication requirements, minimal privacy concerns, and robustness, which are addressed in the paper, respectively.

ISSUES WITH LOW-RESOLUTION SMART METER DATA

The main objective of a smart grid is to improve the reliability, efficiency, and sustainability of the electricity supply and demand by using high-resolution data through advanced monitoring and control systems with bidirectional communication. This will facilitate new smart grid technologies and applications for managing electricity usage and grid performance by enabling a quick response to changes in demand and supply,[31] as well as offering new benefits to consumers. Hence, accessing high-granular electricity consumption data is the key enabler of smart grid benefits, including direct benefits to consumers, benefits to distribution networks, and to the grid as a whole. For example, as a direct benefit to consumers, non-intrusive load monitoring (NILM) techniques help residential consumers optimize their energy use by providing detailed insights into their energy consumption at the appliance level, which could potentially lower their electricity bill.[32] NILM tools can also be developed to detect abnormal usage patterns or energy theft, allowing timely intervention and maintenance of faulty appliances.[33] Conducting NILM studies, however, requires granular smart meter data with resolutions from 1 Hz to 1 MHz[32,33] As benefits to the distribution network, the dynamic operating envelope provides dynamic upper and lower bounds on import or export power in a given time interval for individual DER assets or a connection point. This contributes to the increased adoption and compliance of various DER assets installed in distribution networks. It also helps to dynamically control residential DER instead of a fixed export limit that can lead to an unnecessary amount of curtailed rooftop photovoltaic (PV) generation.[34] To facilitate a dynamic operating envelope, however, high-resolution data through (near) real-time bidirectional communication are needed to update the export limits for DERs.[34] In addition, as benefits to the grid as a whole, the aggregated service allows residential users with rooftop PV systems, EVs, and batteries to participate in the energy market as small generation aggregators in the context of the Australian electricity market.[35]

Currently, the applications of smart meters mentioned above require special meters in the absence of high-resolution real-time data exchange in smart meters. A comprehensive list of smart grid applications and their requirements is provided in Table 1. Comparing the requirements for many smart grid services with the existing sampling intervals reported in Figure 1, we can see a substantial gap between the application requirements and the available data from the existing smart metering infrastructure. This gap will prevent us from achieving the goals of the smart grid due to relying on the data currently collected by the existing smart metering infrastructure. This problem becomes increasingly severe with behind-the-meter DERs and energy storage, as they need to be separately measured to provide usable data that fulfill some of the application requirements, which presents a 2-fold challenge. First, the amount of data that needs to be processed and transmitted could be doubled or tripled, placing additional strain on communication networks and the utility infrastructure. This is important because it is not the smart meters that are the largest cost of smart grid implementation, but rather the required communication infrastructure.[36] Second, the intermittency of DERs introduces more fluctuations in the net demand, which cannot be accurately captured with low-resolution readings. This will cause more generation and demand mismatch, which has to be fixed by more expensive services; hence, higher electricity prices for consumers. To better understand the limitations of low-resolution data, two examples are presented using actual data from residential consumers. The first example comes from an unidentified residential consumer in Adelaide with a 10.6-kW solar PV system, an air conditioning system drawing up to 8.4 kW, and a private meter installed

Table 1. Smart grid applications, required sampling interval, and benefits

Application	Required sampling interval	Stakeholder	Benefits	Beneficiary	Reference
Load signature detection	925 Hz to 1.2 kHz	end users and third parties	understand users' load, load forecasting, improved power quality, and reliability	grid as a whole	Huchtkoetter and Reinhardt[38]
Non-intrusive load modeling	1 Hz to 1 MHz	end users	provide diagnostic information useful for identifying energy-consuming or malfunctioning appliance	consumers	Wang et al.[3,32]; Himeur et al.[33]
Preventive maintenance and malfunction detection	hourly to 1 MHz	utilities	predictive maintenance can also help optimize maintenance techniques and improve program planning	distribution network	Wang et al.[3]; Haq et al.[39]
Cybersecurity	hourly to 1 MHz	end users and third parties	protect the smart grid from cyberattacks and other security threats	consumers	Campbell[31]; Wang et al.[40]
Citizen energy communities as microgrids	1 s	utilities	islanding detection and seamless operation, frequency, and voltage regulation	distribution network	Wang et al.[3]; Campbell[31]
Grid-interactive buildings	1 s	utilities	enable buildings to provide ancillary services to the grid	distribution network	Li et al.[41]; Kapoor et al.[42]
Electric vehicle integration	1 s to 30 min	end users and third parties	enable the integration of electric vehicles into the grid, including charging stations and management systems	grid as a whole	Muratori[43]
Identification of outages	1 to 30 min	distribution system operators	verification of power restoration	distribution network	Australian Energy Market Commission[1]; McKenna et al.[44]; Andrysiak et al.[45]
Energy management systems	5 min	end users and third parties	optimize the use of energy resources across a range of applications, including buildings, industries, and transportation	consumers	Táczi et al.[36]; Collinson[46]
Dynamic operating envelopes	5 min	utilities, end users, DOS, and third parties	increase network visibility	distribution network	Australian Energy Market Commission[1]; AEMO[35]; DEIP Dynamic Operating Envelopes Working Group[47]
Energy theft detection	30 min	utilities	prevent meter tampering, direct theft, and tariff misuse, help improve power quality and overall profitability	distribution network	Australian Energy Market Commission[1]; Wang et al.[3]
Energy marketplaces	subject to the services (e.g., 6 s for regulation service)	end users and third parties	allow customers to participate in a dynamic energy marketplace, where they can buy and sell energy and other related products and services	grid as a whole	Kapoor et al.[42]
Energy trading platforms	subject to the market (e.g., 15 min for energy market)	end users and third parties	more efficient and transparent pricing and allocation of energy resources	consumers	Wang et al.[3]

(*Continued on next page*)

Table 1. ***Continued***

Application	Required sampling interval	Stakeholder	Benefits	Beneficiary	Reference
Demand response	subject to the market (e.g., 6 s for regulation service)	end users and utilities	help demand-supply balance save electricity expense for users	consumers	Huchtkoetter and Reinhardt[38]
Virtual power plants	subject to the market (e.g., 6 s for regulation service)	end users and third parties	provide reliable and flexible power to the grid	distribution network	Campbell[31]; Muratori[43]
Aggregated service	subject to the market (e.g., 6 s for regulation service)	utilities, end users, DOS, and third parties	aggregate residential users distributed energy resources and participate in the market as an aggregator	grid as a whole	Wang et al.[3]; Australian Energy Market Commission[34]
Smart home technologies	subject to data availability (e.g., 5 min for batteries)	end users and third parties	optimize energy usage in the home	consumers	Green Button Data[48]; i-Hub, Smart building data clearing house[49]
Distribution automation	subject to data availability (e.g., 1 s for power quality)	utilities, end users, DOS, and third parties	improve reliability, reduce outage duration and frequency, improve power quality, and reduce operational costs	distribution network	IEEE PES Distribution Subcommittee[50]
Load scheduling	subject to data availability (e.g., 1 min for EVs)	end users and third parties	improve power quality and reliability by automated load control and loading forecasting	consumers	Wang et al.[3]; Muratori[43]

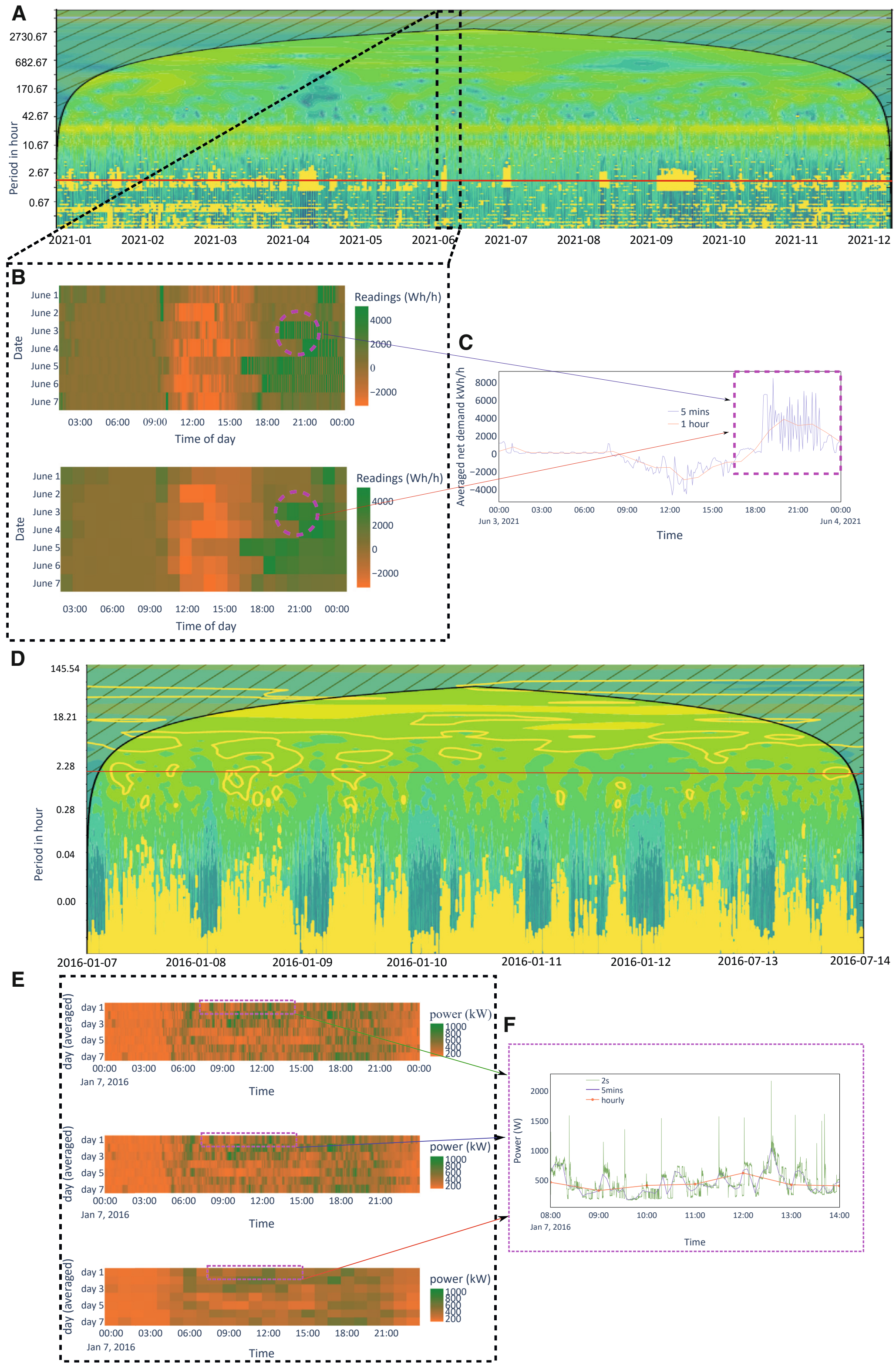
A
Period in hour
2730.67
682.67
170.67
42.67
10.67
2.67
0.67
2021-01 2021-02 2021-03 2021-04 2021-05 2021-06 2021-07 2021-08 2021-09 2021-10 2021-11 2021-12
B
Date
June 1 June 2 June 3 June 4 June 5 June 6 June 7
Readings (Wh/h)
4000 2000 0 −2000
03:00 06:00 09:00 12:00 15:00 18:00 21:00 00:00
Time of day
C
Averaged net demand kWh/h
8000 6000 4000 2000 0 −2000 −4000
5 mins
1 hour
00:00 Jun 3, 2021 03:00 06:00 09:00 12:00 15:00 18:00 21:00 00:00 Jun 4, 2021
Time
D
Period in hour
145.54
18.21
2.28
0.28
0.04
0.00
2016-01-07 2016-01-08 2016-01-09 2016-01-10 2016-01-11 2016-01-12 2016-07-13 2016-07-14
E
day (averaged)
day 1 day 3 day 5 day 7
power (kW)
1000 800 600 400 200
00:00 Jan 7, 2016 03:00 06:00 09:00 12:00 15:00 18:00 21:00 00:00
Time
F
Power (W)
2000 1500 1000 500
2s
5mins
hourly
08:00 Jan 7, 2016 09:00 10:00 11:00 12:00 13:00 14:00
Time

Figure 2. Data comparison of a single house in Australia and average consumption of 12 houses in Germany
The data with different temporal sampling rates are compared with wavelet analysis (A and D), heatmaps (B and E), and detailed comparison of time series patterns (C and F). The red line in the wavelet graph (A and D) indicates a 2-h period, where all wavelet information below the red line is invisible for hourly data based on Nyquist sampling theorem, i.e., the data must be sampled at an interval at most half of the smallest sampling period present in the signal. The top figure in (B) is the heatmap of 1-week's electricity consumption stacked by days with a 5-min resolution, while the bottom figure shows hourly resolution. The area in the purple box is the period containing the air-conditioning load signature with 5-min data (blue) and hourly data (red). Similarly, the three figures in (E) are heatmaps by days with a 2-s sampling interval, 5-min interval, and hourly interval. Time series comparison is shown in (F), with 2-s data (green), 5-min data (blue), and hourly data (red), respectively.

to measure solar generation and the user's energy consumption at a 5-min sampling rate. The second case study is from the NOVAREF project in Germany[37] containing 1 week of data at 2-s sampling intervals. Twelve houses are monitored in this project, and the dataset contains the average power consumption of these houses. We compared the data with different sampling intervals of 2-s, 5-min, and 1-h resolution (plotted in Figure 2). Three types of figures are produced for each example, covering various time ranges and offering distinct information about the samples. The wavelet analysis provides a multi-resolution view of the data, showing frequency components, their scales, and how the frequency components change over time. The heatmap provides the intensity of electricity consumption and its dependency and correlation over time for a week. The time series plot shows the detailed load signature over several hours.

From domain knowledge, we know that an air conditioner operates in cycles, turning on and off to maintain the desired temperature in a room, called the load signature and shown in the purple box in Figure 2C, where we see that the load signature can be clearly identified in 5-min sampled data but not in the hourly data. This can be distinguished by the rapid fluctuations in energy consumption at night. Similarly, by comparing the demands from three different resolutions shown in the purple box in Figure 2F, we observe that the daily peak demands are reduced in the low-resolution data. Therefore, any smart grid application that relies on accurate peak demand values, for example, new residential tariffs with demand charge or dynamic export limit, will fail due to the lack of appropriate data. Looking at Table 1, we see that most of the smart grid applications with direct or indirect benefits to consumers are severely affected by the potential loss of information due to low-resolution metering.

REASONS FOR LOW-RESOLUTION SMART METER DATA

Although high-resolution readings are essential and possible using the capability of the existing smart meters, end users and smart grid service providers do not have access to them because of three major reasons presented in the following subsections.

Data transmission

The first obstacle that prevents widespread high-resolution data reading is the limited band-width available for communication between smart meters and metering service providers (utility companies or distribution system operators). For example, the main communication protocols for European utilities smart meters are Zigbee and M-bus, both have limited range and band-width, that is, around 100 m and up to 250 kbps at 868 MHz,[51,29] which makes it challenging to transmit larger amounts of data generated by high-resolution readings, leading to delays or interruptions in the data transmission process. The communication protocols of smart meters in the Asia Pacific region rely on computer networking technologies, such as power line communication (PLC) or cellular services, to transmit the collected data to the metering service provider.[11] The large number of low bit rate modems and the delays for long-distance transmitted data pose challenges for PLC in supporting smart grid communication,[52] while cellular services face the risks of data loss during network congestion.[11] This can affect the accuracy and reliability of the data readings. The problem becomes increasingly severe for higher-resolution data, as a higher volume of numbers produces more delays or disruptions in the data transmission process.

Data warehousing

Data warehousing is another impediment to measuring and collecting high-resolution electricity consumption data. A common arrangement for smart meter data warehousing, such as in Australia,[53] is a centralized approach in which the metering data provider or distribution network service providers (DNSPs), exercising a monopoly over an area, receive and maintain data exclusively. Then, utilities get access to the data and some utilities make them available to end users through their portals and applications. The problem is that the current level of resolution is not imperative for billing purposes although it is not enough to enable smart grid applications (see Table 1). DNSPs exhibit minimal interest in capturing data at higher resolutions for all residential users, as they will face more challenges and difficulties. One such issue is the high cost of data warehousing, which increases significantly with the storage requirement. For example, an article in the Amazon News Blog shows that it costs US$19,000 to US$25,000 per terabyte per year to build and run a data warehouse.[54] This means that, for a DNSP maintaining a one-terabyte warehouse, increasing the data resolution from 1 h to 2 s will require 1,800 times the storage, hence a US$36 million additional cost on the data warehousing alone. Certain utilities in the United States divested from their electricity generation plants to focus solely on distribution, which could earn a guaranteed return for utility shareholders.[55] It would be difficult to convince them to spend even more to enable higher-resolution data on smart meters without the prospect of a good return on investment. Thus, they show no incentive to collect high granular energy consumption data.

In addition, since DNSPs or utility companies need to deal with a constantly changing group of residential users and other stakeholders, both in terms of their energy-related equipment and the number of customers, the high cost of data storage, maintenance, backup, migration, and building communication links to provide data access to customers or third parties may prevent them from increasing the granularity of current meter reading.[56]

Data privacy

Research studies have identified serious privacy concerns about high-resolution electricity consumption data from end users.[44,57] In essence, any processing method applicable to electricity service providers to analyze the habits and lifestyles of residential electricity users can be applied equally by attackers.[58] For instance, using simple NILM algorithms, attackers can estimate the number/types of electrical appliances, estimate the composition of a home, gather behavioral knowledge to commit criminal activities, infer users' activities at home, and perform location tracking based on EV usage patterns.[59] Other studies have also reported the detection of human behavior, occupancy, holidays, air conditioners, and swimming pools using data from residential electricity consumption.[59] Even law enforcement agencies use electricity consumption data in the courts. For example, the *Columbus Dispatch* reported that, in central Ohio, law enforcement agencies issue at least 60 subpoenas every month for energy usage records of individuals suspected of operating indoor marijuana-growing operations.[60]

In Australia, data on electricity usage from electricity customers was shared with third parties, including debt collectors and mail houses, who can get insight into when a house is occupied.[61] Furthermore, despite the general belief that smart meter data are anonymous, there is evidence indicating that malicious third parties or hackers can link personal data records without identifiers to individuals through external knowledge.[62] For example, a study[62] shows that 68% of the records in the dataset of 180 households were re-identified using simple statistical metrics. Beyond the threat of re-identification, mass human behavioral analysis, malicious data engineering, and real-time surveillance are possible by hackers accessing fine granular electricity data. For these reasons, the smart meter rollouts also face resistance from unexpected intermediary actors, e.g., municipalities in France, who are the official owners of the electricity meters.[63] They perceived the risk of using electricity data divided between privacy, terrorism, and commercial use by third parties.[63] Furthermore, concerns are raised among several consumer advocacy groups about the extensive household data that smart meters collect and transmit to DNSPs or utilities, e.g., how real-time electricity consumption data can be used to identify the number of residents in a home.[9] In response to the above concerns, the EU's energy efficiency legislation has led to a slowdown in the deployment of smart meters. The Netherlands first envisaged smart meters in 2006, but the Smart Metering Act was rejected by the Dutch First Chamber in 2009 due to privacy concerns raised by the smart meter legislation.[57] In France, lawyers filed a class action lawsuit against utility companies because of remote monitoring of users' gas and electricity usage.[46]

Other concerns

Under the circumstances explained above, smart grid service providers, e.g., virtual power plant operators, choose to install private meters for their users to collect high-resolution data. However, this leads to more issues regarding data accessibility, cost, and privacy. For instance, retail suppliers do not have access to the private metered data installed by the home battery system, but they have to manage price risks associated with the load-serving entity's obligations with only low-resolution data from utility meters. In some cases, customers do not have access to their high-resolution data collected from their houses, as their private meters are controlled by a third party who wants to preserve its exclusive rights to the valuable data. Furthermore, the growth in data collected by various entities poses difficulties in the integration of data systems.[64] Due to the lack of policies, regulations, and common standards on the collection, transmission, and usage of private meters compared with smart utility meters. Interoperability issues are unavoidable, making it difficult to collate the data from different appliances behind the meter for integrated solutions, such as home energy management systems. Also, privacy concerns will grow this way because, instead of only one entity responsible for data collection and warehousing, there could be several entities with partial or full access to consumers' data. This means that there are more vulnerabilities and access points to attack and a higher chance for one entity to fail to protect the data.

In addition to the concerns discussed in this section, some new arguments have emerged questioning the necessity of replacing dumb meters with smart meters. For instance, cost-benefit analysis in some countries, e.g., Germany and France, shows that large-scale deployment of smart meters may not be ecologically or economically beneficia.[51] This is mainly due to the fact that many current smart meters are only used for automatic billing purposes. As a result, the uni-directional communication and low-resolution data prevent the development of new services that could justify the upfront capital cost of the smart metering infrastructure.

Categories of potential solutions	Low privacy concerns	Low communication requirement	Low data warehousing requirement	Lossless	Additional requirements
Hardware-based solutions	✗	✗	✗	✓	Extensive communication requirement
Mathematical and AI modelling	✓	✗	✓	✗	High edge computing capacity
Binary-to-binary data compressions	✗	✓	✓	✓	Already implemented as data encoding approaches
Smart Meter Framework 2.0	✓	✓	✓	✓	Accumulating energy consumption per tier is required for billing purposes

Figure 3. Potential solutions for enabling the high-resolution data analysis
The requirements of several potential solutions, including hardware-based, mathematical and AI modeling, binary-to-binary compression, and Smart Meter Framework 2.0, are shown.

In recent years, more efforts have been made in response to data availability issues. Generally, these projects aim to provide third-party companies and end users with easy and secure access to energy usage information, e.g., Green Button Data in the United States,[48] i-Hub in Australia,[49] DataHub[65] and Center Denmark[66] in Denmark, and Estfeed[30] and Synergies[67] for European users. However, such initiatives face new challenges, such as additional storage requirements, data synchronization, and data transmission, which become increasingly significant with high-resolution readings. These challenges must be effectively addressed to ensure that these projects will succeed in the long term.

POTENTIAL SOLUTIONS FOR LOW-RESOLUTION SMART METER DATA

In summary, three main challenges must be addressed to measure, transmit, and maintain high-resolution electricity data, which are consumers' privacy, communication requirements, and data warehousing. In addition, several minor issues must be considered, as listed in Figure 3. The best solution will be the one addressing all major and minor issues simultaneously. The existing methodologies in the literature that could potentially be viable solutions are categorized and compared in Figure 3.

The first category is to use robust and reliable hardware solutions, e.g., internet of things hardware and improved telecommunication protocols, to support high-volume data transmission required for high-resolution smart meter readings.[11,68] However, these solutions do not address concerns regarding data warehousing and the cost of the required communication infrastructure, which is reported to be the largest cost, rather than the smart meters themselves.[36]

The second category involves mathematical or machine learning (ML) techniques to model electricity usage data such that we communicate and store the model instead of the actual data. Specifically, the former uses statistical models, such as the non-homogeneous Markov chain and statistical adjusted engineering model.[69] The latter uses ML techniques, such as federated learning[70] or autoencoder,[71] to model smart meter data and store the models produced in the data center. Mathematical models have greater interpretability than ML-based solutions, but they require detailed knowledge of the system in dealing with complex and non-linear data. Among the candidate solutions based on ML, federated learning appears to be the most promising one, as it facilitates data processing in a distributed manner on the user side, preserving privacy by having each user contribute to creating the model in the training stage.[72] However, it requires high computational capacity at the end users' side and frequent communications among all end users and the data center in the training phase for aggregating the updates into the model, which is currently an obstacle for smart meters. Furthermore, a large amount of memory is needed to store the users' time series data for modeling at the edge of the grid, which could be beyond the available resources by edge computing devices.

The third category of potential solutions is data compression techniques, particularly time series compression. These algorithms (encoders) take time series data and return its compressed representation of a smaller size. From the compressed representation, the original time series can be reconstructed using a decoder. If the reconstructed data are identical to the original data, the algorithm is lossless; otherwise, it is lossy. Modern smart metering data exchange adheres to the International Electrotechnical Commission 62056 standard, where smart meters send ASCII data via a serial port.[73] As a result, current studies of compression techniques for smart grid data focus on binary-to-binary lossless compression. Although these approaches could compress the data at high rates, thus reducing communication and storage burden, they are not privacy preserving. Other approaches are proposed using audio signal binary-to-binary compression or text-to-text compression techniques,[32] such as LZMA,[74] Lempel-Ziv,[75] and BZIP2.[75] For example, the Lempel-Ziv approach compresses a stream of bits of data by using fewer bits to present the repeated segments. Similarly, some research studies use sequential algorithms, including Huffman coding, delta encoding, run-length encoding, and Fibonacci binary encoding, to compress the data.[76] To conclude, time series compression techniques are effective in reducing the communication and storage burden of smart grid data, but are not privacy-preserving. These data compression techniques mainly focus on lossless encoding having already been applied in the data encryption stage to preserve accurate data for billing purposes, while the amount of data transmitted depends on the resolution of the readings. However, when high-resolution data need to be transmitted, the volume of data increases linearly. Therefore, alternative solutions are required to further reduce the size of transmitted data while addressing the privacy concerns of conventional compression techniques.

Proposed solution: Smart Metering Framework 2.0

Figure 4A illustrates how smart meter data are processed and transmitted in the existing frameworks, which we refer to as Smart Metering Framework 1.0. Lossless data compression is facilitated with data encryption when the numerical values of the readings are converted into binary values. The data are then transmitted via a utility-maintained concentrator and the headend system to the data center. At the data center, trusted third parties (TTPs) can request access to data.

In comparison, we propose a potential solution, hereafter called Smart Metering Framework 2.0, that could fulfill all the criteria mentioned in the previous section. The framework is summarized in Figures 4B and 4C. In essence, it consists of two phases, namely data storage and data reconstruction. In data storage, smart meters can store a single accumulated reading of imported and exported energy (or as many tiers in a ToU tariff) for billing purposes while retaining a high-resolution time series for use in smart grid applications. We propose to apply a lossy but shape-preserving approach that can compress and separate the information so that users only keep partial information for privacy preservation. The other part of the information is transmitted to the data center, so neither side keeps the entire time series of the smart meter data. One idea that inspires us most to achieve this

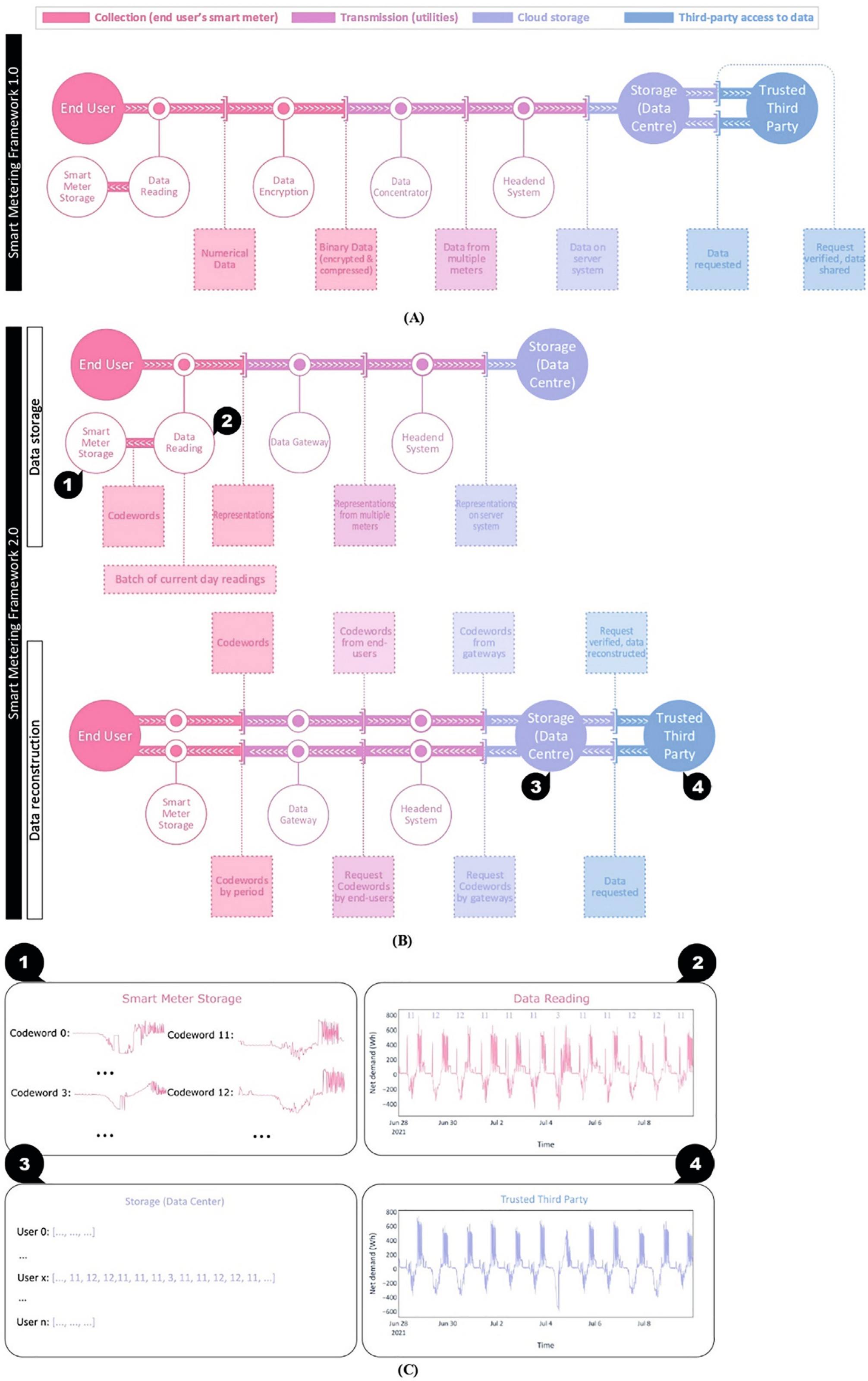
Collection (end user's smart meter)
Transmission (utilities)
Cloud storage
Third-party access to data
Smart Metering Framework 1.0
End User
Smart Meter Storage
Data Reading
Data Encryption
Data Concentrator
Headend System
Storage (Data Centre)
Trusted Third Party
Numerical Data
Binary Data (encrypted & compressed)
Data from multiple meters
Data on server system
Data requested
Request verified, data shared
(A)
Smart Metering Framework 2.0
Data storage
End User
Smart Meter Storage
Data Reading
Data Gateway
Headend System
Storage (Data Centre)
Codewords
Representations
Representations from multiple meters
Representations on server system
Batch of current day readings
Data reconstruction
Codewords
Codewords from end-users
Codewords from gateways
Request verified, data reconstructed
End User
Smart Meter Storage
Data Gateway
Headend System
Storage (Data Centre)
Trusted Third Party
Codewords by period
Request Codewords by end-users
Request Codewords by gateways
Data requested
(B)
Smart Meter Storage
Codeword 0:
Codeword 11:
Codeword 3:
Codeword 12:
Data Reading
Net demand (Wh)
Time
Storage (Data Center)
User 0: [..., ..., ...]
User x: [..., 11, 12, 12,11, 11, 11, 3, 11, 11, 12, 12, 11, ...]
User n: [..., ..., ...]
Trusted Third Party
Net demand (Wh)
Time
(C)

Figure 4. Flow chart of the existing metering data infrastructure, including Smart Metering Framework 1.0 and the proposed framework, Smart Metering Framework 2.0, with a single user example

(A) In the current advanced metering infrastructure, the smart meter data are collected at the users' premises as numerical values. Then it is encrypted and compressed into binary data before being sent to the data concentrator located in physically secured locations, e.g., substations or transformer stations.[77] The data concentrator receives multiple smart meters data via the neighborhood area network and uses an Ethernet Wide Area Network connected to the headend system. The headend system has an off-the-shelf server that is connected to the meter data nanagement (MDM) system database.[77,78].

(B) The proposed Smart Metering Framework 2.0 involves two phases: data storage and data reconstruction. For data storage, the smart meter data are collected in a temporal batch and compressed into codewords and representations on a daily basis. The codewords are stored in the smart meter storage, while the representations are communicated to the data center via a local gateway and a headend system. When the TTP requests data for data reconstruction, it must send its authentication, requested users, period, and resolution. The data center will forward the information via the transmission stage, get the codewords from corresponding end users, and reconstruct the data with the required period and resolution.

(C) 1–4: the codewords stored in the smart meter of a user, data collection, representations for all users in the data center, and the reconstructed data for TTP.

goal is the codebook method. The codebook method was first conceptualized in 2004[79] and formulated in 2008 for time series similarity analysis.[80] It contains a set of representative patterns, called codewords, and a set of indexes to the codewords for time series reconstruction.[80] The optimal codewords are usually obtained from the entire time series data, i.e., offline compression, with the generalized Lloyd algorithm (GLA) based on the nearest neighbors condition and with iterative computations.[80] However, for smart grid applications, the entire electricity time series data for compression does not exist because the smart meters require at least daily compressed data to transmit to the data centers. Therefore, the GLA is not applicable. In this regard, we believe that a dictionary-based time series compression method[76] could be a compelling solution by combining the compression algorithm and the codebook method. More specifically, instead of finding codewords from iterations and optimizations, the codebook could be built with batch compression without requiring the entire time series data. Essentially, in Smart Metering Framework 2.0, the codebook for each residential user contains two sets of data, namely codewords and "representations." The codewords are unique, daily high-resolution patterns. The representations are the indices that point to the codewords that represent the demand of consumers each day. Each index in representations shows one day's pattern by pointing to the codewords in a daily sequence. To be more specific, the smart meters measure imported and exported energy at a high sampling rate, with the measurements for the current day being stored in a temporary batch, as shown in the Figure 4B. Then it uses the codewords stored at the user's end to compute daily representations and update the existing codewords when new unique patterns are identified.

With the above compression method, each meter only sends the daily representation to the data center on a daily basis in the data storage phase. This is shown in the top flowchart of Figure 4B. Since the representation cannot provide much information without codewords and does not require much communication bandwidth, we propose a simple gateway instead of data concentrators to upload data from the neighborhood area network (NAN) to the wide area network (WAN). For data reconstruction, as shown in the bottom flowchart of Figure 4B, when a TTP requests access to high-resolution data from users' smart meters for a given time period, they can query the codewords according to the representations of the requested period. Obviously, a consumer's consent is required before the TTP's access is granted. To be more specific, an authenticated TTP initiates a request to the data center, specifying the desired users using national meter identifier numbers and the period of time. Subsequently, the data center undertakes the processing of this request by collecting representations for the requested time period from the database. Through this process, a distinct set of representations can be computed, indicating the indices of codewords essential for subsequent data reconstruction. Upon determining the required codewords, the data center transmits a request to the headend system via WAN, thereby confirming the user IDs and associated codewords. The headend system then relays this confirmation to the respective gateways in the NAN, which subsequently forwards the request to the intended end users' smart meters. Then, the end users' smart meter will send the queried codewords to the data center based on the indices received, similar to the example shown in Figure 4C. When real-time data are requested by the service provider, i.e., the data being processed at the collection stage, the end users' meter will send the data from the temporary current day batch. In this way, the service provider receives the requested real-time data in a lossless manner.

The proposed framework offers a number of advantages. Smart meters only store codewords without time or day notion, while the data center retains compressed representations of the users. As a result, reconstructing the data from one side alone becomes impossible. This approach provides a robust solution to address privacy concerns, as attackers cannot obtain information that does not exist. For example, physical attacks to the end user's meter storage can only attain the codewords that present partial daily patterns without the time signature that made it. Since the disclosed information does not indicate which pattern corresponds to which day, the attackers cannot get much information, e.g., to know which days the user is usually not at home. Similarly, the adversarial attack during data communication cannot access complete information. This is because the proposed framework exclusively transmits representations in the data storage phase and codewords corresponding to the requested time period in the data reconstruction phase. Moreover, obtaining even this limited information (codewords) is more challenging for hackers when compared with Smart Metering Framework 1.0, where the data are transmitted at fixed intervals. The proposed solution communicates codewords upon request, ensuring that attackers remain unaware of the timing of TTP queries. In addition, the Smart Metering Framework 2.0 has low storage and minimum communication requirements. For example, assuming a consumer with ten codewords, the daily demand profile of the 100th day could be best represented by codeword 1. Therefore, the representation for the 100th day is 1, and all m = 100 days of data can be stored at the user's end in the form of n = 10 codewords. This will result in $\frac{m-n}{m}$ = 97% memory saving. At the date center end, all representations are stored for each user for these 100 days. The space-saving ratio can increase over time for users with relatively stable electricity usage patterns. In addition, the accumulated energy consumption could be handled separately for billing purposes by smart meters knowing the current tariff structure of users. Therefore, this framework could enable trustworthy bidirectional communications. It also enables a more efficient billing of energy consumption for both consumers and utilities compared with Smart Metering Framework 1.0 because of higher resolution. The increased accuracy in tracking users' consumption patterns makes it possible for utilities to charge their users more accurately based on their peak demands and generations, which cannot be adequately captured in low-resolution data, as shown in Figure 2. Furthermore, Smart Metering Framework 2.0 is more robust in terms of suffering from invalid data or faulty readings, since they would be replaced by data from the codeword.

KEY AREAS FOR FUTURE RESEARCH

Different aspects of the proposed method require extensive research, which is summarized in the following four sections.

Similarity metrics for time series compression

First, a key challenge in codebook methods is quantifying pattern similarities based on their shapes. Most research studies in this area use Euclidean distance and dynamic time warping for quantification of similarity.[81] However, these research studies developed shape-based similarity identification techniques that only work for clustering/classification of consumers, where their methodologies are not suitable for reconstructing time series.[82] Furthermore, similarity measurements for smart meter data compression and different smart grid applications have rarely been discussed or applied. As a result, similarity metrics that target specific applications and are more efficient in compression and reconstruction tasks can significantly advance the proposed framework.

Smart meter functionalities

Since the proposed framework comprises a lossy compression method, a separate function to accumulate energy consumption per tariff tier is essential for accurate billing purposes. There are several factors that need to be considered, including the reliability of the accumulated result and transmission frequency under different tariff structures, e.g., ToU or real-time prices. Currently, smart meters cannot deliver this functionality; hence, further research and development are needed to address this issue in existing and future smart meters. Furthermore, the smart meter should be able to compute codewords and representations based on the shape-based similarity metric above. This requires smart meters capable of edge computing tasks.

Key performance indicators

Key performance indicators (KPIs) are needed to quantitatively measure and compare the performance of different potential solutions on different challenges we discuss in this perspective. Some potential KPIs could be general statistics, memory saving rate, algorithm complexity, recovered time series complexity using Shannon information theory, load duration, and frequency domain analysis, to just name a few. However, more research is needed to identify or develop the right and systematic KPIs for effectiveness measurements.

Regulations and policies change

As the energy landscape evolves with the deployment of the proposed Smart Metering Framework 2.0, it is necessary to review regulations and policies to effectively govern the collection, transmission, storage, and usage of energy data. Current regulations and policies may not fully account for the unique characteristics and challenges associated with the proposed framework. For example, monopoly utilities owning a large database might refuse to be merged into the new framework. The possible reason is that they may prefer to fully own data in a low resolution rather than accessing high-resolution data with other utilities under the restrictions. Therefore, it is essential to conduct research and engage in discussions with policymakers, industry stakeholders, and other relevant parties to ensure that regulations and policies are updated to reflect the changing landscape of energy data management. This may involve considerations such as data privacy and security, data sharing and interoperability, data quality and accuracy, and fairness and equity in data usage.

CONCLUSIONS AND OUTLOOK

The current implementations of smart meters lack the required data granularity and are not fulfilling their potential benefits. In this perspective, we argue that achieving high-resolution energy data by addressing associated challenges is crucial in realizing the true benefits of smart grids for all stakeholders. We discuss the issues with low-resolution data and the issues preventing smart meters from operating in a truly smart way. We then list potential solutions and introduce a new framework to enable high-resolution data. The proposed framework consists of a codebook for each residential user, which contains two sets of data: codewords and representations. Codewords are unique high-resolution patterns of the user over a day, while representations are indexes of codewords pointing toward historical data. In the data compression stage, the smart meter measures the imported/exported energy at high resolution, computes daily representations using the codewords and sends only the representation of the day to the data center. The codewords are updated when new unique patterns are found. In the data reconstruction phase, service providers can query codewords based on the representations of the user and a certain temporal range. Smart meters store codewords, while the data center stores compress representations of users, ensuring that user data remain secure.

ACKNOWLEDGMENTS

This project is funded jointly by the University of Adelaide industry-PhD grant scheme and Watts A/S, Denmark.

AUTHOR CONTRIBUTIONS

Conceptualization, R.Y. and S.A.P.; methodology, R.Y. and S.A.P.; software, R.Y.; formal analysis, R.Y.; investigation, R.Y.; writing – original draft, R.Y. and S.A.P.; writing – review & editing, R.Y., S.A.P., W.L.S., A.J.B., J.A.R.L., and J.L.-V.; resources, W.L.S.; supervision, S.A.P., W.L.S., A.J.B., J.A.R.L., and J.L.-V.; project administration, S.A.P.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT and Grammarly in order to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

REFERENCES

1. Australian Energy Market Commission (2021). Review of the Regulatory Framework for Metering SerVice Review. https://www.aemc.gov.au/market-reviews-advice/review-regulatory-framework-metering-services.

2. European Commission (2022). Smart Grids and Meters. https://energy.ec.europa.eu/topics/markets-and-consumers/smart-grids-and-meters_en.

3. Wang, Y., Chen, Q., and Kang, C. (2017). Smart Meter Data Analytics*42*. https://doi.org/10.1145/3004295.

4. Singh, S., and Yassine, A. (2018). Big data mining of energy time series for behavioral analytics and energy consumption forecasting. Energies *11*, 452. https://doi.org/10.3390/en11020452.

5. Mission:data Coalition (2022). DEACTIVATED: How Electric Utilities Turned off the Data-Sharing Features of 14 Million Smart Meters. https://www.missiondata.io/reports.

6. U.S. Government (2009). Public Law 111 - 5 - American Recovery and Reinvestment Act of 2009. https://www.govtrack.us/congress/bills/111/hr1.

7. Keen, B. (2023). Rushing to the Finish Line: Can We Clear the Hurdles on the Way to 100by 2030?. https://www.energycouncil.com.au/analysis/rushing-to-the-finish-line-can-we-clear-the-hurdles-on-the-way-to-100-smart-meters-by-2030/.

8. Kerai, M. (2022). Smart meter statistics in Great Britain: quarterly report to end June 2022. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1099629/Q2_2022_Smart_Meters_Statistics_Report.pdf.

9. Wynn, G. (2014). Smart Meters Are Poor Value, Find 10 Eu Countries. Climate Home News. https://www.climatechangenews.com/2014/06/19/eu-efficiency-set-back-as-countries-delay-smart-meter-rollout/.

10. City of Nelson, British Columbia (2022). Meters & Meter Reads. https://www.nelson.ca/933/Meters-Meter-Reads.

11. Zhu, C., and Reinhardt, A. (2019). Reliable streaming and synchronization of smart meter data over intermittent data connec- tions. In 2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, SmartGridComm 2019. https://doi.org/10.1109/SmartGridComm.2019.8909705.

12. (2018). Smart meters in sweden – lessons learned and new regulations. In 3rd AIEE Symposium on Current and Future Challenges to Energy Security3rd AIEE Symposium on

Current and Future Challenges to Energy Security, p. 12. December.

13. Swenman, M., Energy, S., and Inspectorate, M. (2021). Metering as an enabler for consumer-focused flexibility in sweden. Tech. Rep. November. Swedish Energy Markets Inspectorate. https://www.iea-isgan.org/wp-content/uploads/2022/06/WG9-Metering-as-Enabler-for-Flexibility.pdf.

14. Østergaard, J., Weckesser, J.T.G., Bondy, D.E.M., Bondy, D.E.M., Nørgård, P.B., Pinson, P., Martini, L., and Madsen, H. (2018). Smart grids innovation challenge. Tech. rep.

15. Garcia, D.. How Do Smart Meters Communicate. https://www.emnify.com/blog/how-smart-meters-communicate.

16. The Independent Electricity System Operator (IESO), Smart metering entity (2022). [Currently hosted at URL https://www.ieso.ca/en/Sector-Participants/Smart-Metering-Entity/Data-Sharing]

17. Office of the Auditor General of Ontario (2014). Annual Report of the Office of the Auditor General of ontario:3.11 Smart Metering Initiative (Tech. rep.). http://www.auditor.on.ca/en/reports_en/en14/311en14.pdf.

18. Smart Meter, TriPica. Deployment in Eu: Who Is Leading and Who Is Trailing. 2020. https://www.tripica.com/blog/smart-meter-deployment-the-impact-on-eu-households.

19. ReportLinker. Global and china Smart Meters Industry Report. https://www.globenewswire.com/news-release/2022/06/17/2464866/0/en/Global-and-China-Smart-Meters-Industry-Report-2022-2027.html.

20. Trabish, H.K. (2022). 97% of smart meters fail to provide promised customer benefits. can $3b in new funding change that? https://www.utilitydive.com/news/97-of-smart-meters-fail-to-provide-promised-customer-benefits-can-3b-in/632662/.

21. Statista Research Department. (2023). Leading Electric Utilities Based on Number of Smart Meters Installed in the United States as. of 2019. https://www.statista.com/statistics/246455/leading-us-electric-utilities-based-on-number-of-smart-meters/.

22. International Energy Agency. Japan 2021 energy policy reviewTech. rep. https://www.iea.org/reports/japan-2021.

23. Research and markets. Japan smart meter market overview, 2027Tech. rep. https://www.researchandmarkets.com/reports/5648616/japan-smart-meter-market-overview-2027.

24. Japanese T&D World.. Utility Reaches 20 Million Meter Milestone. https://www.tdworld.com/smart-utility/metering/article/20972343/japanese-utility-reaches-20-million-meter-milestone.

25. Fortune Business Insights.. Awareness to Save Electricity to Bolster the Market for Top 10 Smart Electric Meter Companies. https://www.fortunebusinessinsights.com/blog/top-smart-electric-meter-companies-10691.

26. Yuan, R., Pourmousavi, S.A., Soong, W.L., Nguyen, G., and Liisberg, J.A. (2023). Irmac: Interpretable refined motifs in binary classification for smart grid applications. Eng. Appl. Artif. Intell. *117*, 105588. https://doi.org/10.1016/j.engappai.2022.105588.

27. Yuan, R., Pourmousavi, S.A., Soong, W.L., Black, A.J., Liisberg, J.A.R., and Lemos-vinasco, J. (2023). A synthetic dataset of Danish residential electricity prosumers. Sci. Data *10*, 371. https://doi.org/10.1038/s41597-023-02271-3.

28. Thompson, C.. Five Minute Settlement Explained. https://help.amber.com.au/hc/en-us/articles/4411273786637-Five-minute-settlement-explained.

29. Smart Energy International. AMI Penetration and Communication Protocols. https://www.smart-energy.com/magazine-article/ami-penetration-and-communication-protocols/.

30. Elering; Elering; Estonian Data Hub (2017). Guide for Using and Joining the Estonian Data Hub. https://www.elering.ee/sites/default/files/attachments/DataHub_guide%20for%20using_102017_EN%20ED.pdf.

31. Campbell, R.J.. The Smart Grid: Status and Outlook. https://crsreports.congress.gov/product/pdf/R/R45156.

32. Wang, Y., Chen, Q., Hong, T., and Kang, C. (2019). Review of smart meter data analytics: Applications, methodologies, and challenges. IEEE Trans. Smart Grid *10*, 3125–3148. arXiv:1802.04117. https://doi.org/10.1109/TSG.2018.2818167.

33. Himeur, Y., Alsalemi, A., Bensaali, F., Amira, A., and Al-Kababji, A. (2022). Recent trends of smart nonintrusive load monitoring in buildings: A review, open challenges, and future directions. Int. J. Intell. Syst. *37*, 7124–7179. https://doi.org/10.1002/int.22876.

34. Australian Energy Market Commission. Integrating Energy Storage Systems into the NEM (December 2020). https://www.aemc.gov.au/rule-changes/integrating-energy-storage-systems-nem.

35. AEMO. Small Generation Aggregators in the NEM. https://aemo.com.au/-/media/files/initiatives/integrating-energy-storage-systems-project/factsheet-nem-small-generation-aggregators-march-2023-release.pdf.

36. Táczi, I., Sinkovics, B., Vokony, I., and Hartmann, B. (2021). The challenges of low voltage distribution system state estimation—an application oriented review. Energies *14*, 5363. https://doi.org/10.3390/en14175363.

37. Anvari, M., Proedrou, E., Schäfer, B., Beck, C., Kantz, H., and Timme, M. (2022). Data-driven load profiles and the dynamics of residential electricity consumption. Nat. Commun. *13*, 1–12. https://doi.org/10.1038/s41467-022-31942-9.

38. Huchtkoetter, J., and Reinhardt, A. (2019). A study on the impact of data sampling rates on load signature event detection. Energy Inform. *2*, 24. https://doi.org/10.1186/s42162-019-0096-9.

39. Haq, A., and Jacobsen, H.A. (2018). Prospects of appliance-level load monitoring in off-the-shelf energy monitors: A technical review. Energies *11*, 189. https://doi.org/10.3390/en11010189.

40. Wang, Y., Su, Z., Zhang, N., Chen, J., Sun, X., Ye, Z., and Zhou, Z. (2021). Spds: A secure and auditable private data sharing scheme for smart grid based on blockchain. IEEE Trans. Ind. Inf. *17*, 7688–7699. https://doi.org/10.1109/TII.2020.3040171.

41. Li, H., Wang, Z., and Hong, T. (2021). A synthetic building operation dataset. Sci. Data *8*. 213-13. https://doi.org/10.1038/s41597-021-00989-6.

42. Kapoor, S., Sturmberg, B., and Shaw, M. (2021). A review of publicly available energy data sets, Tech. Rep. The Aus- tralian National University, 00120.

43. Muratori, M. (2018). Impact of uncoordinated plug-in electric vehicle charging on residential power demand. Nat. Energy *3*, 193–201. https://doi.org/10.1038/s41560-017-0074-z.

44. McKenna, E., Richardson, I., and Thomson, M. (2012). Smart meter data: Balancing consumer privacy concerns with legitimate applications. Energy Pol. *41*, 807–814. https://doi.org/10.1016/j.enpol.2011.11.049.

45. Andrysiak, T., Saganowski, Ł., and Kiedrowski, P. (2017). Anomaly detection in smart metering infrastructure with the use of time series analysis. J. Sens. *2017*, 1–15. https://doi.org/10.1155/2017/8782131.

46. Collinson, P. (2017). Is Your Smart Meter Spying on You? The Guardian. https://www.theguardian.com/money/2017/jun/24/smart-meters-spying-collecting-private-data-french-british.

47. DEIP Dynamic Operating Envelopes Working Group. Dynamic operating envelopes working group outcomes re- port. Tech. Rep. March, Distributed Energy Integration Program 2022. https://arena.gov.au/knowledge-bank/deip-dynamic-operating-envelopes-workstream-final-outcomes-webinar/.

48. Green Button Data. Green button data: secure, anonymous, digital. https://www.greenbuttondata.org/.

49. i-Hub (2021). Smart building data clearing house. https://www.ihub.org.au/ihub-initiatives/smart-building-data-clearing-house/.

50. IEEE PES Distribution Subcommittee. Smart grid for distribution systems: The benefits and challenges of distribution automation (da). https://grouper.ieee.org/groups/td/dist/da/doc/IEEE%20Distribution%20Automation%20Working%20Group%20White%20Paper%20v3.pdf.

51. European Commission; Directorate-General for Energy, Alaton, C., and Tounquet, F. (2020). Benchmarking smart metering deployment in the EU-28 – Final report. Publications Office. https://data.europa.eu/doi/10.2833/492070.

52. Segatto, M.E.V., de Oliveira Rocha, H.R., Silva, J.A.L., Paiva, M.H.M., and do Rosário Santos Cruz, M.A. (2018). Telecommunication technologies for smart grids: Total cost optimization. In Advances in Re- newable Energies and Power Technologies, I. Yahyaoui, ed. (Elsevier), pp. 451–478. https://doi.org/10.1016/B978-0-12-813185-5.00007-3.

53. Department of Energy. Environment and Climate Action of Australia, Smart Meters Help

You Understand and Manage Your Energy Use and Costs. 2022. https://www.energy.vic.gov.au/for-households/save-energy-and-money/smart-meters.

54. Barr, J. (2012). Amazon Redshift – the New Aws Data Warehouse. AWS News Blog. https://aws.amazon.com/blogs/aws/amazon-redshift-the-new-aws-data-warehouse/.

55. McCoy, M. (2022). Monopolistic Utility Companies Suppress the Use of Customer Data — Episode 155 of Local Energy Rules. Institute for Local Self-Reliance.. https://ilsr.org/utilities-customer-data-portability-ler155/

56. Hariharan, N.K. (2021). Trends in data warehousing. International Journal of Innovations in Engi- neering Research and Technology *6*, 7–14. https://doi.org/10.1016/b978-012088469-8.50109-1.

57. Pallas, F. (2012). Beyond gut level – some critical remarks on the german privacy approach to smart metering. European Data Protection: Coming of Age *5*, 313–345.

58. Perspective, C.-c., Völker, B., Reinhardt, A., Faustine, A., and Pereira, L. (2021). Watt's up at home? smart meter data analytics from a consumer-centric perspective. Energies, 1–21.

59. Armoogum, S., and Bassoo, V. (2019). Privacy of Energy Consumption Data of a Household in a Smart Grid. In Smart Power Distribution Systems: Control, Communication, and Optimization, Q. Yang, T. Yang, and W. Li, eds. (Elsevier), pp. 163–177.

60. Smith, M. (2011). Are Smart Meters Real-Time Surveillance Spies? CSO Online. https://www.csoonline.com/article/546876/microsoft-subnet-are-smart-meters-real-time-surveillance-spies.html.

61. Vince, C., Craig, B., and Henrietta, C. (2012). Smart Meter Data Shared Far and Wide. The Sydney Morning Herald. https://www.smh.com.au/technology/smart-meter-data-shared-far-and-wide-20120922-26dvp.html.

62. Buchmann, E., Böhm, K., Burghardt, T., and Kessler, S. (2013). Re-identification of smart meter data. Personal Ubiquitous Comput. *17*, 653–662.

63. Chamaret, C., Steyer, V., and Mayer, J.C. (2020). "hands off my meter!" when municipalities resist smart meters: Linking arguments and degrees of resistance. Energy Pol. *144*, 111556. https://doi.org/10.1016/j.enpol.2020.111556.

64. Thiris, C.C., Rakotojaona, L., Kundevski, M., Lacroix, O., Lelong, P., Son Tan, D., and Farrow, H.. Data Opportunities for Smarter Networks Energy Networks Australia. https://www.energynetworks.com.au/resources/reports/2020-reports-and-publications/data-opportunities-for-smarter-networks/.

65. Energinet. What is datahub? https://en.energinet.dk/energy-data/datahub/.

66. Cener Denmark Intelligent Energy. Learn about Center Denmark's unique data platform. https://www.centerdenmark.com/en/the-data-platform/.

67. Synergies, Synergies Solution (2023). https://energydataspaces.eu/.

68. Perez, N.R., Domingo, J.M., Lopez, G.L., Avila, J.P.C., Bosco, F., Croce, V., Kukk, K., Uslar, M., Madina, C., and Santos-Mugica, M. (2022). ICT architectures for TSO-DSO coordination and data exchange: a European perspective. IEEE Trans. Smart Grid. https://doi.org/10.1109/TSG.2022.3206092.

69. Grandjean, A., Adnot, J., and Binet, G. (2012). A review and an analysis of the residential electric load curve models. Renew. Sustain. Energy Rev. *16*, 6539–6565. https://doi.org/10.1016/j.rser.2012.08.013.

70. Wang, Y., Jia, M., Gao, N., Von Krannichfeldt, L., Sun, M., and Hug, G. (2022). Federated clustering for electricity consump- tion pattern extraction. IEEE Trans. Smart Grid *13*, 2425–2439. https://doi.org/10.1109/TSG.2022.3146489.

71. Bregere, M., and Bessa, R.J. (2020). Simulating tariff impact in electrical energy consumption profiles with conditional varia- tional autoencoders. IEEE Access *8*, 131949–131966. https://doi.org/10.1109/ACCESS.2020.3009060.

72. Mitra, A., Ngoko, Y., and Trystram, D. (2021). Impact of federated learning on smart buildings. In Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS, *2021*, pp. 93–99. https://doi.org/10.1109/ICAIS50930. 2021.9395938.

73. Hanif, M. Introduction to IEC Standards. International Electrotechnical Commission. https://doi.org/10.13140/RG.2.2.35290.62405.

74. Bin Tariq, Z., Arshad, N., and Nabeel, M. (2015). Enhanced lzma and bzip2 for improved energy data compression, SMART- GREENS 2015 - 4th International Conference on Smart Cities and Green ICT Systems. Proceedings, 256–263. https://doi.org/10.5220/0005454202560263.

75. Abuadbba, A., Khalil, I., and Yu, X. (2018). Gaussian approximation-based lossless compression of smart meter readings. IEEE Trans. Smart Grid *9*, 5047–5056. https://doi.org/10.1109/TSG.2017.2679111.

76. Chiarot, G., and Silvestri, C. (2021). Time series compression: a survey. Preprint at arXiv. https://doi.org/10.48550/arXiv.2101.08784.

77. Langer, L., and Kammerstetter, M. (2015). The Evolution of the Smart Grid Threat Landscape and Cross-Domain Risk Assessment, pp. 49–77. https://doi.org/10.1016/B978-0-12-802122-4.00003-1.

78. Faisal, M.A., Aung, Z., Williams, J.R., and Sanchez, A. (2015). Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study. IEEE Syst. J. *9*, 31–44. https://doi.org/10.1109/JSYST.2013.2294120.

79. Megalooikonomou, V., Li, G., and Wang, Q. (2004). A dimensionality reduction technique for efficient similarity analysis of time series databases. In Proceedings of the thirteenth ACM international conference on Information and knowledge management, pp. 160–161.

80. Wang, Q., and Megalooikonomou, V. (2008). A dimensionality reduction technique for efficient time series similarity analysis. Inf. Syst. *33*, 115–132. https://doi.org/10.1016/j.is.2007.07.002.

81. Zhang, H., Dong, Y., Li, J., and Xu, D. (2022). Dynamic time warping under product quantization, with applications to time- series data similarity search. IEEE Internet Things J. *9*, 11814–11826. https://doi.org/10.1109/JIOT.2021.3132017.

82. Jayanth Krishnan, K., and Mitra, K. (2022). A modified kohonen map algorithm for clustering time series data. Expert Syst. Appl. *201*, 117249. https://doi.org/10.1016/j.eswa.2022.117249.