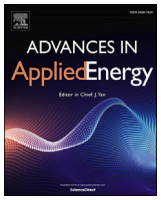
Advances in Applied Energy 14 (2024) 100170



Contents lists available at ScienceDirect

Advances in Applied Energy

journal homepage: www.elsevier.com/locate/adapen





Data-driven energy management of virtual power plants: A review

Guangchun Ruan [a,*], Dawei Qiu [b,*], S. Sivaranjani [c], Ahmed S.A. Awad [d], Goran Strbac [b]

[a] Laboratory for Information & Decision Systems, Massachusetts Institute of Technology, 32 Vassar Street, Boston, MA 02139, United States of America
[b] Department of Electrical & Electronic Engineering, Imperial College London, London SW7 2AZ, UK
[c] School of Industrial Engineering, Purdue University, West Lafayette, IN 47907, United States of America
[d] Department of Electrical & Computer Engineering, Sultan Qaboos University, Muscat 123, Oman

ARTICLE INFO

Keywords:
Distributed energy resource
Renewable energy
Microgrid
Battery
Big data
Machine learning
Reinforcement learning
Game theory
Peer-to-peer
Blockchain
Uncertainty
Risk management
Communication
Privacy
Electricity market
Real-world projects

ABSTRACT

A virtual power plant (VPP) refers to an active aggregator of heterogeneous distributed energy resources (DERs), which creates a promising pathway to expand renewable energy and demand-side electrification for deep decarbonization. The VPP market is projected to have a significant growth potential, with the global investment surging from $6.47 billion in 2022 to $16.90 billion by 2030. Up to now, VPPs still face technical challenges in dealing with the inherent uncertainty of DERs, and data emerge as a promising and essential resource to handle this issue. This paper makes the first effort to review the development of VPP technologies from a data-centric perspective, and then analyze the major role of data within every decision phase of VPPs. We examine the VPP energy management through a data lifecycle lens, and extensively survey the progress in data creation, data communication, data-driven decision support, data sharing and privacy, as well as technical solutions stemming from reinforcement learning, peer-to-peer sharing, blockchain, and market participation. In addition, we offer a unique overview of open data and recent real-world projects around the world to showcase the latest VPP practices. We finally discuss the major challenges and future opportunities in detail, with a focus on topics such as public benchmarks, internet of things, 5G, explainable artificial intelligence, and federated learning. We highlight the need for technical advances in data management and support systems for the growing scale of future VPP systems, and suggest VPPs delivering more ancillary grid services in the future.

1. Introduction

1.1. Background

The rapid surge of distributed energy resources (DERs) has transformed the current energy landscape. There is no doubt that regional aggregation and local balancing will become more and more universal in future energy systems. DER units are highly heterogenous with complex dynamics and inherent uncertainties, and they pose significant operational challenges in large-scale coordination due to inherent intermittency, partial controllability, and inconsistent preferences among stakeholders [1]. Unfortunately, the conventional infrastructures and passive network controllers struggle to utilize these resources at an aggregate level, so the energy community is now actively searching for the next-generation technologies for DER management.

Virtual power plants (VPPs) offer a promising solution to manage large-scale DERs, especially distributed renewable energy and flexible end-users. Coordinating these DERs at scale is essential to promote the transition for a cleaner and more affordable way of energy use.

A VPP is technically defined as a connected and controllable aggregation of DER units. For illustration, a typical VPP system is composed of several kinds of generators (e.g., solar panels, wind turbines, combined heat and power plants, geothermal generators, diesel generators, small hydro-plants), active users (e.g., buildings, responsive users, smart home appliances), storage facilities (e.g., batteries, pumped storage, electric vehicles), and communication channels (e.g., WiFi, 5G base stations). The components of VPPs in some case may also include prosumers (both produce and consume energy) as well as microgrids. Two well-known categories are technical and commercial VPPs [2], depending on whether distribution network constraints are considered. There are other VPPs for specific applications, such as building VPPs, industrial park VPPs, and community VPPs.

The concept of microgrids might be similar to VPPs, but the key differences can be summarized in terms of general sizes, operational

* Corresponding authors.
E-mail addresses: gruan@mit.edu (G. Ruan), d.qiu15@imperial.ac.uk (D. Qiu).

https://doi.org/10.1016/j.adapen.2024.100170
Received 8 December 2023; Received in revised form 12 January 2024; Accepted 8 February 2024
Available online 5 March 2024
2666-7924/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

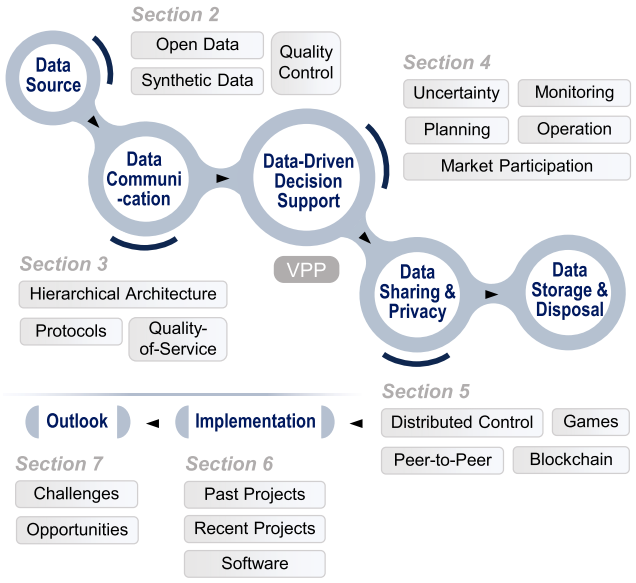


Fig. 1. A typical lifecycle of VPP data with five phases and several key technical details being highlighted. The initial four phases will be fully discussed in the subsequent sections. This figure also showcases the structure and organization for different sections.

patterns, reliability performance, and reliance on storage [3]. In fact, the concept of VPP has evolved and expanded consistently ever since its inception. For instance, distributed multi-energy devices (e.g., heat pumps, power-to-gas facilities, electrolysis equipments) are now within the scope of VPP research [4].

Apart from academic efforts, VPPs are also attracting significant attention from the energy industry and experiencing rapid growth in practical projects or applications. Many world-renowned companies have made their ambitious investment plans in VPP projects because of the potential economic and environmental benefits. As for investment, VPPs are projected to expand their global market share from $6.47 billion in 2022 to around $16.90 billion by 2030, achieving a dramatic 2.5x growth within only eight years [5].

Indeed, there are great opportunities for VPPs in the era of deep decarbonization of the energy sector. In this paper, we are motivated to provide an extensive survey of the progress in the data-driven VPP management, and discuss several exciting directions of future research. Our work is unique to offer new understanding of VPP technologies and implementation.

1.2. Data lifecycle in VPPs

Data and deep learning (DL) technologies play a special role in supporting VPP energy management. Sensor networks collecting real-time dataflows can provide crucial information to address the existing challenges in VPPs, including endogenous uncertainties, irrational user preferences, and complicated system patterns. In this context, the use of data has potential to revolutionize VPP energy management in terms of decision quality and operational efficiency.

The data science community has developed the concept and language of data lifecycle management, i.e., to process and analyze data throughout the entire procedure from creation to destruction [6].

Fig. 1 contextualizes the data flow in VPP systems using the language of data lifecycle management. In this figure, the cycles track the entire flow involving data source/creation, data communication, data-driven decision support, data sharing & privacy, and data storage & disposal. This paper is focused on the initial four phases, as shown in Fig. 1, and present the key technologies and challenges for each phase. These phases also involve a series of support systems involving databases, communication networks, market analysis and information processing systems.

Data lifecycle management opens up a novel perspective for VPP operation, but has not yet been well documented in the literature. This observation motivates us to reflect the recent developments of VPP energy management from the lens of data lifecycle management, which is a key distinguishing feature of this review. Big data thinking and data-driven decision making are particular informative to explore the next technical directions for VPPs.

1.3. Contributions

While there are some existing review articles such as [7–10] covering either specific conceptual (category or hierarchical structure details) or mathematical (optimization formulations) aspects of VPP management, it is still unclear how to understand the role of data at every single phase. In this paper, we take a data-centric perspective and put special focus on how data are generated, transferred, processed, utilized, and protected in VPPs. Our work promotes a systematic and deep understanding of data-driven energy management of VPPs.

To sum up, this review paper makes the following contributions to the current body of knowledge:

1. We make the first effort to summarize the development of VPP energy management through a lens of data lifecycle, including data creation, data communication, data-driven decision support, and data sharing and privacy. This novel perspective unlocks a series of data-driven insights through a comprehensive literature review of recent progress, especially over the past five years.
2. Different support systems for VPP energy management are broadly investigated, including databases, communication networks, information processing systems, and market analysis systems. We highlight the emerging use of advanced data science and machine learning in these systems.
3. Recent real-world VPP projects (since 2017) and open datasets are fully documented in this paper. We also highlight the major challenges, opportunities, and several promising solutions at the end.

The remainder of this paper is organized to demonstrate the data lifecycle of VPPs, and Fig. 1 illustrates the overall logic flow for different sections. Section 2 collects the popular data sources and summarizes the progress in data creation, Section 3 analyzes the major solutions in data communication, Section 4 conducts a systematic overview of data-driven decision support technologies, and Section 5 moves to the data sharing and privacy issues in VPPs with a specific focus on decentralized decision making. Real-world projects are then summarized in Section 6. We further identify the key challenges and future opportunities in Section 7, and make the concluding remarks in Section 8.

2. Data sources in VPPs

This section will explore the data landscape of VPP energy management, and collect potential resources from open or synthetic data. Quality control will then be discussed as a remedy for data noises and sample biases.

2.1. Data landscape

Data from VPPs present a broad diversity in types, dataset sizes, and quality [11]. VPP systems are often comprised of heterogeneous components, but the data quality control and sensor maintenance are seldom stringent.

In total, there are at least nine categories of data from VPPs, i.e., network topology, renewable energy, controllable generator, baseline load, flexible load, energy storage, multi-energy facility, energy price,

Table 1
Data category, examples, and accuracy levels in VPPs.

Category	Typical Examples	Accuracy
Network topology	IEEE 13/33/39-bus system, CIGRE low-voltage network, user-specified systems (often small-scale).	Perfectly precise or slightly noisy (need identification)
Renewable energy	Wind power from PJM, wind power from NREL M2 Tower, wind power from preset Weibull distributions, solar power from pvoutput.org.	Very noisy, with potentially large errors
Controllable generator	Coal-fired/gas-fired generators from IEEE testbeds, hydropower generators from NYISO, user-specified setup for specification (often).	Highly precise
Baseline load	Load profiles from PJM, household energy use from the REFIT program, energy consumption from preset normal distributions.	Precise (with small fluctuations)
Flexible load	Residential energy use from Pecan Street program, business and household energy use from the Irish CER smart metering project, demand response data from the low carbon London project, user-specified setup for flexibility and responsiveness.	Precise (full control) or noisy (incentive driven)
Energy storage	Tesla lithium-ion battery pack, Nissan electric vehicle, pump storage, user-specified setup (often).	Highly precise
Multi-energy facility	Heat pumps, power-to-gas facility, natural gas pipeline, hydrogen/natural gas/thermal storage, all with preset or realistic setups.	Highly precise (often) or slightly noisy
Energy price	Day-ahead/real-time energy prices from PJM, reserve/regulation prices from CAISO, preset or real-world time-of-use tariffs, preset or real-world fixed pricing schemes.	Perfectly precise (fixed and preset prices) or very noisy (others)
Weather observation	Temperature/humidity/wind speed observations from meteorological stations, other user-specified setup.	Slightly noisy

Note: “IEEE” and “CIGRE” are two well-known professional organizations; “NREL” is a U.S. national lab; “NYISO”, “PJM”, “CAISO” are three U.S. electricity markets; several companies and projects are involved as well.

and weather observation. Table 1 presents a detailed overview of different categories and typical examples. Most uncertain data in the VPP literature appear to be the renewable energy outputs, flexible loads (or demand response), and energy prices.

These data can be further divided into the categories of structured and unstructured data [12]. Here, the most popular structured data are time-series data, which are consistently collected at a preset frequency and then stored in a format of data frames or data cubes. On the contrary, unstructured data have no preset formats, and they are generally difficult to handle or analyze. Typical examples of unstructured data include graphical topologies, census data, economic coefficients, text and image data streams.

Data management in VPP refers to the integration and processing of diverse data sources. In practice, each DER unit may be equipped with its own sensors, so the sampling parameters and API interfaces of these local devices might be very different. This makes it challenging for VPP operators to create effective management schemes, which motivates the use of a compatible data interface for the entire workflow.

2.2. Open data

Open data may come from separate sources or from a single collection. For the separate kind, there are some review papers mentioning about relevant datasets. For example, the authors of [13] collected 24 data sources, including the famous low-carbon London project, the Pecan Street dataset, and the EPRI distributed PV monitoring project. Reference [14] instead focused on several energy-economic models such as TIMES, whose data sources were analyzed and tracked accordingly.

As for the full collections, a useful resource turns out to be microgrid datasets, but these datasets require further customization to match the setup of VPPs. In [15], a real dataset was collected from a microgrid in the University of California, San Diego. The system involved buildings, solar panels, electric vehicles (EVs), a combined heat and power plant, electricity and thermal storage. Another multi-year microgrid data were collected in [16] and the test system was a research building in Japan with solar arrays and a battery. The authors of [17] proposed a general method to extend the SimBench dataset to establish a microgrid benchmark system involving solar panels, distributed generators, storage, and electric load. In [18], three microgrid benchmarks were demonstrated: a campus-level microgrid, a utility-level microgrid, and a CIGRE microgrid.

Apart from the above datasets, one may also find several supplemental materials containing data for VPP systems, such as storage-based VPPs [19], flexible VPPs [20], market-based VPPs [21], and building VPPs [22]. Most of these datasets are not collected from real-world projects and are only intended for simulation purposes. Data matching and alignment have been adopted to make these data ready to use.

Despite several open data initiatives, only a limited number of datasets are now publicly shared in the VPP literature. It is important for the energy community to take care of this issue, since high quality open-source data is key to driving innovation in algorithms and decision-making for VPP systems.

2.3. Synthetic data

Synthetic data do not come from real world and are mostly created by algorithms. These data serve as an alternative to real-world data and offer an opportunity to expand the dataset size without any costly sensor networks. In many practical use cases, the size of real datasets may be moderate to small because of limited data release, potential data biases or incompleteness, and a lack of labeled data [23]. Given that small datasets cannot guarantee reliable and robust results from data-driven decision making, synthetic data bridge the gap by increasing the dataset size and data availability [24].

Synthetic data are statistically similar to real data, but they are by no means the same. Although synthetic data are created based on realistic distributions and domain knowledge, estimation errors will result in an issue of distribution shift [25]. From another aspect, synthetic data can be viewed as a kind of machine-generated data with noises.

In VPP research, the most popular approach for synthetic data generation is the Monte-Carlo simulation and other extensions. Reference [26] used the Monte-Carlo method to generate price data by a log-normal distribution and load data by a normal distribution, then finally, ran backward scenario reduction. Similarly, a sampling technique called Markov Chain Monte-Carlo was applied in [27] to develop time-series data for renewable energy outputs. The authors of [28] proposed an approach to obtain two time series data with different resolutions. Here, weekly and hourly data were well captured, with the chronological order being preserved. Then in [29], the correlation in those generated data was modeled using a multivariate Copula function, and these data were found to reproduce multi-source renewable productions very well.

Statistical regression models are another option to generate synthetic data. Reference [30] created a group of probabilistic models of

household appliances and then generated time-extended stochastic load profiles accordingly. Synthetic scenarios of wind speed were created in [31] by using Fourier series to capture seasonal trends and an autoregressive moving average (ARMA) model to capture the correlations in residue series.

2.4. Data quality control

Data quality control refers to a procedure that inspects and improves the data quality to meet certain requirements. The priority is to avoid defects in data sources, and then reduce those adverse effects caused by noises or sensor failure. In general, there are two common types of data quality issues: data anomalies and missing data [32].

Data anomalies or outliers are those suspicious and inconsistent data in a database. The top reasons for these anomalies include excessive data noises, malfunctioned sensors, unauthorized processing, human misoperations, or even cyber attacks. Reference [33] presented the result of detecting outliers and missing values in short-term solar forecasting in VPPs. Cluster analysis was applied in [34] to find and exclude outlier data in a large-scale power quality dataset. In [35], the authors proposed a data-driven Bayesian optimization algorithm to handle the uncertainty factors in parameter prediction.

Missing data is another typical problem of data loss, caused by sampling/communication interruptions or outlier removal. A common remedy for this issue is the interpolation approaches such as linear interpolation, piecewise polynomial interpolation, and cubic spline interpolation. Reference [36] developed a prediction model for interrupted data, and the idea was to select a neighboring device to benchmark the vulnerability assessment and forecasting results.

Data quality can generally be improved at different phases. However, controlling the quality of data sources is often more effective and should be given priority, while post-processing solutions may be necessary as a backup option.

3. Data communication in VPPs

This section will review the latest progress in data transmission and communication. We will outline the classical architectures and protocol options for VPP communication systems. The quality of service metrics will then be used to evaluate the communication quality.

3.1. Layers and classical architectures

Data communication in VPPs serve as the backbone to build up coordination and connections between DER units. In this context, the abstract layers are the most important components of a communication system determining the pattern of data transmission. A VPP communication system is technically layered in the same way as computer networks and can be characterized by delving into each layer sequentially. Such a system often consists of one or more physical and functional layers. Here, physical layers are associated with devices and communication channels, while functional layers could be more abstract regarding high-level decisions and applications. It should be noted that these basic setups are very similar across the literature, even if the layers are named differently in specific applications.

For instance, the authors of [37] developed five layers that include, from the bottomon upwards, VPP components, a communication layer, a data layer, a management layer, and a framework layer. Similar in [38], another set of four layers were developed, involving a business layer, a services layer, an operation layer, and a resource layer.

One key issue is how to configure the communication layer for reliable performance. A unique feature of VPPs is the integration of DER units that are hierarchically distributed. To coordinate these units, the communication layer should support multi-layer and high-throughput communication of data.

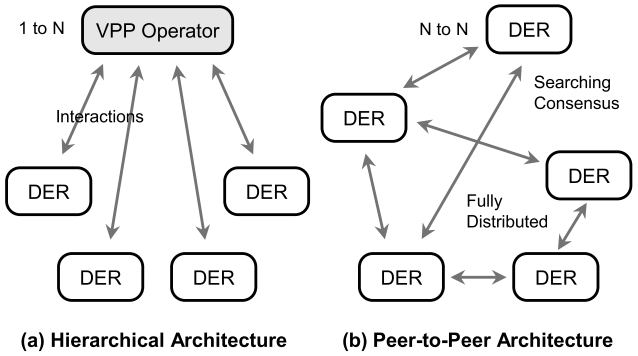


Fig. 2. Two popular architectures for data communication within VPP systems. The hierarchical architecture is 1-to-N with the VPP operator serving as a point of coordination. The peer-to-peer architecture abandons such a centralized operator and instead implements a fully distributed scheme.

Fig. 2 displays two classical architectures for data communication, namely, a hierarchical architecture and a peer-to-peer (P2P) architecture. The main difference lies in whether a system coordinator is present (hierarchical) or absent (P2P). Hierarchy is a flexible structure to support data communication needs. A two-tier architecture developed in [39] took into account both individual control and coordination between VPPs, while in [40], an extended architecture was proposed with prosumers in the middle connecting grid operators and end-users. But a P2P architecture is distinct, as shown in Fig. 2. Such an architecture was implemented in [41] to run smart contracts on blockchain.

The internal structure of a communication layer is of interest as well, and the main focus is on the hardware components and modules required to support communication traffics during busy hours. Reference [42] covered a broad range of topics in communication channel design for VPPs. In particular, the authors showcased the upstream and downstream communication modules, and then demonstrated the core modules, functionalities, and database accordingly.

3.2. Communication protocols

Building upon a given architecture, the next step is to select and configure the communication protocols, i.e., a set of technical rules describing how to format, connect, transmit, and exchange data. These protocols are critical to determine how DER units contact each other and how to mitigate data conflicts as well as transmission failures.

In general, VPP operators choose to implement the full-fledged protocols. Protocol selection begins with an analysis of communication requirements, followed by picking a protocol that adequately aligns with these needs. In [43], the authors made efforts to estimate the communication requirements of medium data-intensive devices. They then proposed several data-communication-enabled solutions and enhancement strategies.

Communication protocols can be formally classified into wired/wireline and wireless protocols [44]. According to [45], the main technologies for wired protocols are optical fiber, twisted pairs, and power transmission lines. Alternatively, the wireless communication mainly relies on WiFi, Zigbee, Bluetooth, and cellular 3G/4G/5G.

Table 2 merges the collections from [44,46] and extends to cover more commonly-used protocols in practice. These protocols guide the design of local communication, automation systems, control/management systems, cybersecurity, physical interfaces, and integration, respectively. They well identify the best practices to ensure the quality and reliability of VPP systems.

IEC 61850 is one of the most popular and advanced wired protocols nowadays. This protocol provides far more functionalities than the previous generation IEC 60870-5-104 [47], such as the support for real-time GOOSE and SMV messaging, configuration language, and information modeling methods [48]. Using IEC 61850, the authors of [49]

Table 2
A list of communication protocols implemented in VPP systems.

Standard	Year	Description	Application
IEC 802	1980	A family (802.1–802.24) of international standards for local area networks, personal area network, and metropolitan area networks. Allow communication packets of variable sizes. Focus on the data link and physical layers.	WiFi, Ethernet
IEC 61850	1995	An international standard that defines communication for intelligent electronic devices at substation automation systems. Include specific demonstrations in communication requirements, structure, service mapping, and network engineering.	Substations, microgrids
IEC 61968	2013	A family (61968-1–61968-14) of international standards that define information exchanges between electrical distribution networks. Intended to enable middleware services to collect data from different applications and interfaces.	System operators, utility companies, retailers
IEC 61970	2022	An international standard for application program interfaces for energy management systems. Specify the integration, information model, interface design to facilitate the management.	System operators, utility companies, microgrids
IEC 62264	2020	An international standard for enterprise control system integration. Specify the object models, attributes, and activity models of manufacturing operations.	System operators, utility companies
IEC 62351	2023	An international standard for communication security, including the digital signatures, access authentication, eavesdropping prevention, playback and spoofing prevention, and intrusion detection.	ICT systems
IEC 62443	2018	A family of international standards for cybersecurity in automation and control systems. Target at system operators, service providers, and system manufactures.	System operators, utility companies, retailers
IEEE 1547	2003	An international standard to provide a group of criteria and requirements for the integration of DERs. Provide requirements regarding the performance, operation, tests, safety, and maintenance for DER integration.	System operators, utility companies, microgrids
IEEE 1646	2005	An international standard that defines communication requirements of information exchange within and external to the protection, control, and data acquisition systems of substations.	Substations
IEEE 2030	2011	A standard to develop guidance for smart grid interoperability of energy technology and information technology operation in power systems and end-use applications.	System operators, utility companies, retailers, end users
OpenADR	2009	A standard for open automated demand response led by North American research companies and labs. Specify the data model, interfaces, and features of a demand response automation service.	Demand response

were able to enhance the interaction between a VPP operator and DER units. In [50], further efforts were made to enable a federation in multi-domain communication networks within VPPs.

Wireless communication is also widely implemented in practice, and the most popular protocol is IEEE 802. In the context of local area networks, the IEEE 802.11 has been widely known to support WiFi and WiFi variants. In [51], a communication and a computing architecture were proposed to virtualize several renewable power plants in a VPP. The technologies in use included IEEE 802.11 and long-term evolution.

OpenADR is another common protocol for VPP systems. OpenADR originally acted as a public domain standard to meet automation goals for demand response, but has been widely applied to coordinate DER units afterwards. For example, DERs were controlled in [52] to provide ancillary services to the system, while the data exchange was realized in public Internet infrastructure using OpenADR 2.0b.

Other typical protocols involve IEC 61968, IEEE 1547, and IEEE 1646, and a detailed comparison can be found in [45]. Project managers could follow IEC 61968 for data exchange between devices and networks in distribution networks, follow IEEE 1547.x to interconnect DERs with power grids or microgrids, and then apply IEEE 1646 for substation automation. All of these protocols turned out to succeed in VPP communication and networking.

Lastly, there are some efforts at protocol-level innovation, and the main idea is to design a customized protocol aligning with VPP characteristics. For example, a novel geographic routing protocol was developed in [53] to support dynamic control in VPPs for seamless enrollment and dis-enrollment of prosumers.

3.3. Communication quality

Beyond full connectivity, a data communication system must ensure the quality of data traffic to avoid channel-level distortions or unpredictable errors. Any unreliable or unauthorized channels might become a weak point under cyber attacks; therefore, it is essential to monitor the quality of data communication. Reference [54] pointed out that any data exchange between a VPP operator and EVs needs to be reliable, responsive, and secure.

Existing research on communication quality puts emphasis on the following two issues: how to develop efficient systems to maximize the communication traffics as much as possible; and how to design customized approaches to match the limited communication conditions for energy management. We will next overview the progress on these two issues one after another.

Quality of service (QoS) is an important metric used to measure, analyze, and evaluate the quality of data communication and exchange. QoS represents a group of network control strategies to manage the traffic and realize the preferred performance with limited network bandwidth. In particular, QoS strategies detect several aspects of communication performance, including latency/delay, accuracy, packet loss rates, retransmissions, bandwidth, amount of traffic and message patterns [55].

The core idea of QoS is to implement priority and congestion management so that the scarce bandwidths could be first allocated to critical tasks [56]. A dynamic spectrum allocation strategy was presented in [57] to enable the multi-user multi-channel cognitive radio for EVs in a VPP. Results showed a drop in the packet loss probability when using dynamic spectrum. In [58], the authors carried out a virtual private network to guarantee high communication quality between distributed load resources and VPP operators. Similarly, smart meter gateways were used in [59] to provide a secure communication channel for VPPs.

Another solution to handle an imperfect condition for communication is to redesign the existing approaches to match the restrictions. There are several efforts and results following this or similar ideas. For example, in order to address the issue of limited bandwidth, the authors of [60] designed a communication-efficient dispatch strategy using the distributed primal-dual sub-gradient algorithm. An edge-side economic dispatch model for VPPs was established in [61] to avoid heavy communication burden in the centralized cloud computing. Reference [62] discussed a new feeder flow control scheme that occupied less communication bandwidth. Communication delay was another issue in practice. Reference [63] broadly discussed the impacts of communication delays, channel noises, and time-varying topology on communication networks. A dynamic update policy served as a solution to control the potential fluctuations. In [64], the communication delays and DER capacity limitations were two influential factors for the power tracking control of VPP generators. The authors of [65] found that disturbances in communication systems might influence the distributed decisions and resulted in the limited extent of effective ancillary services. Likewise,

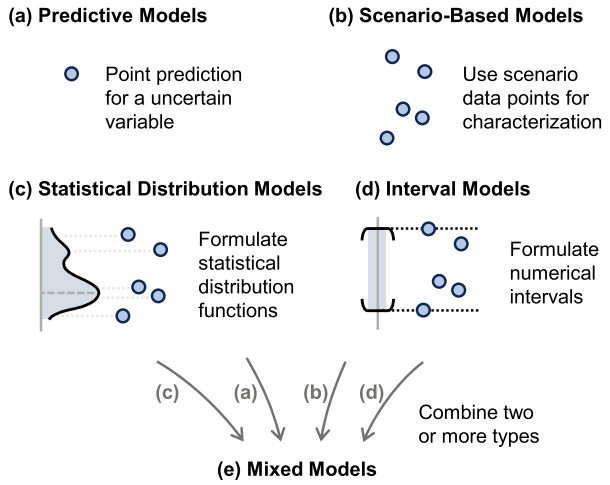


Fig. 3. Five types of uncertainty analysis models. Different types of models are visualized to show their key characteristics, and the mixed models are generally a combination of two or more types.

in [66], unreliable communication conditions such as fragile channels were evaluated. Using a four-state discrete time Markov chain to simulate the packet loss probability, the authors finally came up with a multi-time-scale coordination model to reduce the impacts of data loss.

4. Data-driven decision support for VPPs

This section will demonstrate the advances in data-driven decision making for VPPs, involving a full discussion of uncertainty analysis, monitoring, planning, operation, control, benefit-cost analysis, and market participation. This is perhaps the most active VPP area with a significant amount of research efforts in recent years. Here, the role and benefits of data in VPP management are our major focus.

4.1. Uncertainty analysis

With a large number of DER units, VPPs are highly susceptible to uncertainty and may encounter many unexpected operating patterns. Proficiency in uncertainty analysis and mitigation measures is thus important for VPP energy management.

Uncertainty modeling and analysis have been extensively discussed in the literature, and a specialized review article is [67]. Uncertainty sources in VPPs generally include renewable power outputs (typical wind and solar energy), energy market prices, and energy demand. We next list five mainstream models for uncertainty characterization: predictive models, scenario-based models, statistical distribution models, interval models, and hybrid models.

Fig. 3 showcases these five models, and the distinction lies in how they quantify the potential changes and outcomes. This taxonomy provides a useful perspective to discuss the current modeling practices.

Predictive models typically make point predictions for one or some statistical properties of a random variable (e.g., mean/maximum/minimum). Although these models are mostly deterministic with high computational efficiency, the potential variations are overlooked; therefore, final decisions based on these models may be sub-optimal and risky when dealing with large uncertainties. This category of models is now widely used in industrial projects, and may be informative in analyzing extreme or selected use cases as well. For instance, a bi-directional long short-term memory network was applied in [68] to estimate the uncertain renewable energy outputs and EV behaviors. Similarly, an autoencoder neural network was developed in [69] to handle uncertainty factors in generation and load profiles. Reference [70] turned to another point estimate method that captured the uncertainty conditions using three points for each input random variable.

Scenario-based models collect a scenario set to capture the possible fluctuations of random variables. These models often rely on a large amount of data that may not be available in practice, and the scenario data generation is a widely-used but case-specific solution for rescue. Reference [71] formulated an autoencoder model in a multi-energy VPP to generate a variety of scenarios, which would then be reduced by the fast forward reduction algorithm. In [72], Monte-Carlo simulation was running several times to construct a scenario dataset, and the number of scenarios was cut down by a k-means clustering technique.

Statistical distribution models apply typical distributional priors to quantify uncertainty. Popular options include: the Gaussian distribution (or normal distribution), the exponential distribution, and the Weibull distribution (mostly for wind). It is clear that these models are ready to use if a prior density function is well defined. But they may be very inaccurate and biased when the distributional assumption does not hold true. For uncertainty descriptions, the authors of [73] assumed the probability distribution functions for wind speed, solar irradiation, and day-ahead prices to be Weibull, Gamma, and Gaussian. It was similar in [74] that the wind and load data were set to follow a Weibull and a normal distribution respectively.

Interval models construct mathematical intervals and bounds to express the value range of random variables. Perhaps these models have the simplest formulations, and they are practically popular in analyzing renewable energy outputs (especially wind power). A primary concern is the ignorance of distribution details, so the decisions coming from these models could be rather conservative. A budget-based box uncertainty set was established in [75] using the Dirichlet process mixture model (DPMM) and variational inference algorithm. This set was a high-dimensional extension of the classical intervals, and the correlations among multiple uncertainty factors were fully handled. Likewise, a compact uncertainty set was developed in [76] based on a data-driven Bayesian nonparametric approach, which leveraged the spatio-temporal properties from historical observations.

There are also several mixed models that simply combine two or more types of models together. An example of this kind was [77]: in this work, the authors constructed three types of uncertainty sets, i.e., intervals, scenarios, and a mix of both. Different parameters used for risk management were then analyzed in an out-of-sample case.

4.2. Monitoring

VPP monitoring refers to the process of detecting, tracking, and evaluating the operating conditions of an entire system and each equipment over time. At present, system identification and anomaly detection are two common tasks for VPP monitoring.

System identification is mainly focused on how to use real-time data to calibrate and update the system models in a dynamic fashion. Fig. 4 demonstrates a typical identification process to determine the model parameters from DER data. The parameters in Fig. 4 should be dynamically updated because the available VPP capacity turns out to be dependent on the time-varying DER outputs. Reference [78] developed a model to evaluate the VPP capacity based on a distributionally robust optimization approach. Reference [79] focused on how to predict the available flexibility of VPPs, and a neural network was found useful in estimating the longevity of household flexibility. In [80], a low-order multiple-input-multiple-output model was identified for secondary voltage control, with the consideration of reactive power, solar irradiance, power demand, and voltage levels. A similar identification problem was formulated in [81] as well, and the focus was using reinforcement learning to identify the uncertain grid components accurately. Reference [82] discussed non-intrusive load modeling via multitask learning. The end-user satisfaction and daily behaviors were both taken into account.

The aim of anomaly detection lies in how to identify faults or sub-optimal operating conditions from measurements. Power quality is an important issue of anomaly detection, and typical examples include

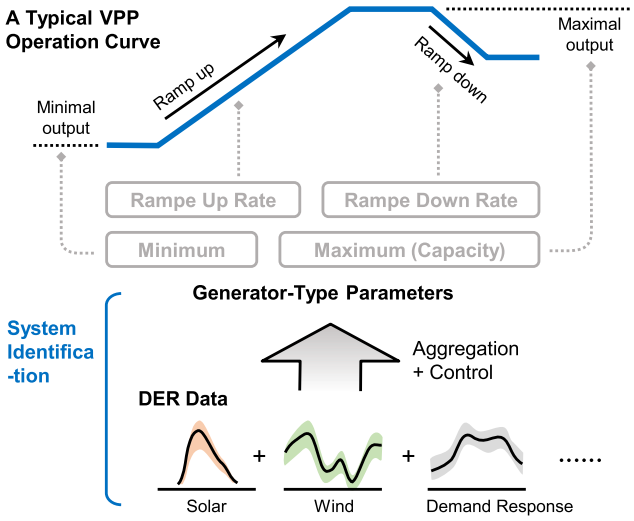


Fig. 4. An illustrative example for data-driven system identification for VPP systems. The generator-type parameters are useful to reflect the holistic flexibility characteristic.

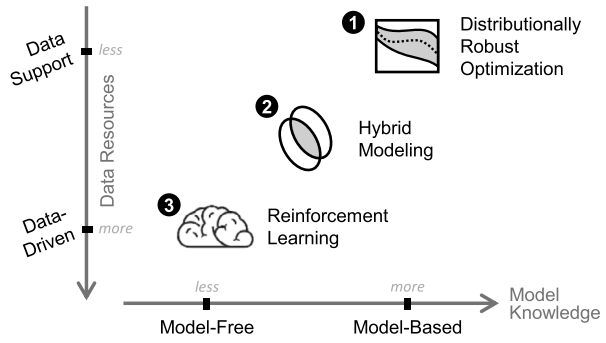


Fig. 5. Three categories of approaches for data-driven planning, operation, and control. We compare them in terms of the reliance on model knowledge (model-based to model-free) and data resources (data support to data-driven).

voltage transients, harmonics, and supply interruptions. Power quality in VPPs might be deteriorated by renewable energy resources. This issue was studied in [83] through 26-week multipoint synchronic measurements, and the authors conducted clustering analysis to detect possible anomalies. The next task was to monitor the short-term working conditions of VPPs that were less common in practice [84]. Other power quality factors of voltage magnitudes and phase angles were discussed in [85] to prevent undesirable power quality issues.

Fault detection is an important step for fault management and anomaly detection. There could be a number of fault events, such as open-circuit or short-circuit faults. To detect and classify faults, the authors of [86] applied wavelet transform to extract effective features and formulated a decision tree model to analyze these features in an intelligent protection scheme. A similar detection issue was studied in [87], and the key techniques involved were the Hilbert transform and AdaBoost classification. Reference [88] presented a multi-output deep learning algorithms to mitigate the risks of photovoltaic system faults, with the required number of sensors being cut down by 50%. Reference [89] combined the signal processing and data mining approaches for fault protection, and the data for both looped and radial configurations were applied in the analysis pipeline. Rapid localization of faults was another critical issue. Reference [90] proposed a data-driven approach in a fault locator using the Gaussian process regression.

4.3. Planning, operation, and control

Planning, operation, and control of VPPs involve a wide range of topics and an abundant line of research efforts. Here we focus on the use of data analytics and data-driven decision-making in different management tasks.

We begin with a quick summary of the top three approaches in this area, i.e., distributionally robust optimization (DRO), hybrid modeling, and reinforcement learning (RL). Fig. 5 collects and compares these techniques in terms of model and data aspects, with DRO typically locating on the model-based end and RL locating on the model-free side.

DRO is a special kind of robust optimization that considers the worst case over a family of uncertainty distributions. The formulation based on Wasserstein ambiguity sets is most commonly used in the literature [91]. Hybrid modeling is another decision making approach that integrates model-based and data-driven components within one workflow [92]. Finally, RL is typically utilized as a model-free learning-based paradigm, whose primary focus is taking optimal actions to maximize cumulative rewards [93]. RL can sequentially capture system uncertainties and learn from unknown system dynamics.

Data-driven planning is focused on how to make realistic planning decisions by capturing the historical trends. Historical data make it possible to return more reasonable and reliable results in planning. Reference [94] made data-driven resource adequacy for VPPs, and the overall model was a multistage stochastic program with uncertain renewable inputs. In [95], a hybrid model was established by combining a physical-economical model and a data-driven model to determine the optimal system size, the nonlinear battery degradation, and variant EV charging profiles. Reference [96] was different from these works, and the authors utilized an RL model to optimize VPP portfolios by learning the market conditions and charging demand features.

Data-driven operation broadly discusses accelerating the decision making process, improving constraint modeling, and adapting to diverse and complicated applications. In [97], the optimal energy management model of VPPs was solved by a differential evolution algorithm, which explored evolutionary processes and data for the global optimum. Alternatively, hybrid modeling was implemented in [98] and [99]. In particular, an estimated distribution was used in [98] to formulate a group of data-driven nonparametric chance constraints that were later embedded in the energy management model. Similarly in [99], a neural network was applied to approximate frequency constraints to ensure adequate frequency response and inertia support. A data-driven aggregation approach was proposed in [100] to compute the thermal dynamics of building clusters, while a virtual storage model was integrating these features into the decision making pipeline. RL is another versatile approach to handle hard operational decisions and automatic energy management in real time. Reference [101] proposed a deep RL model to empower reliable online economic dispatch of VPPs. The carbon emission issue was solved by an improved deep RL approach scheduling the carbon capture and power-to-gas facilities [102]. An empirical example in China's carbon market was then given in [103]. There were some other extensions of RL methodologies, such as [104] and [105]. Reference [104] extended to consider a novel soft actor-critic RL based on sharpness-aware minimization, and the benefit of this model was its high robustness and adaptability to different applications. In [105], a deep neural network and an RL model were paired to make decisions in retail pricing and energy management together.

Data-driven control intends to design innovative control strategies and calibrate controller parameters in real time to mitigate operational uncertainties. Typical examples in VPP control include: a decision tree model [106] to determine the active power reduction of VPPs for mitigating over-frequency, and a k-vector nearest neighbor model [107] to update the frequency control coefficients for a doubly-fed induction generator wind turbine for stability analysis. Two control approaches based on neural networks were proposed in [108] for improving backtracking search and particle swarm optimization. These neural network approaches were found to be superior in cost reduction and decision

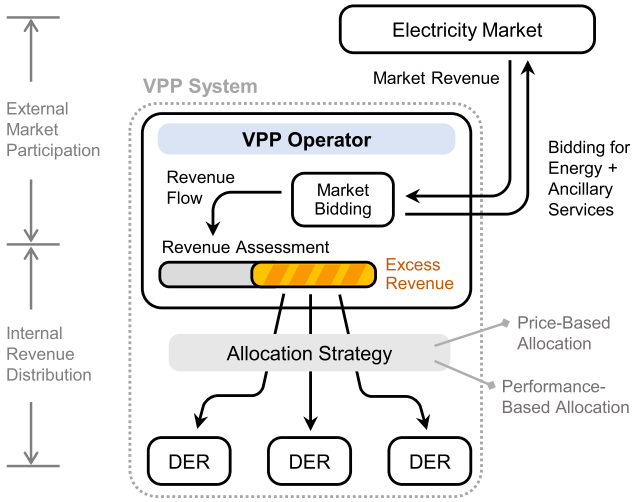


Fig. 6. External and internal decision making for a VPP system. The revenue from market bids is distributed inside VPP and assigned to different DER units based on a preset allocation strategy.

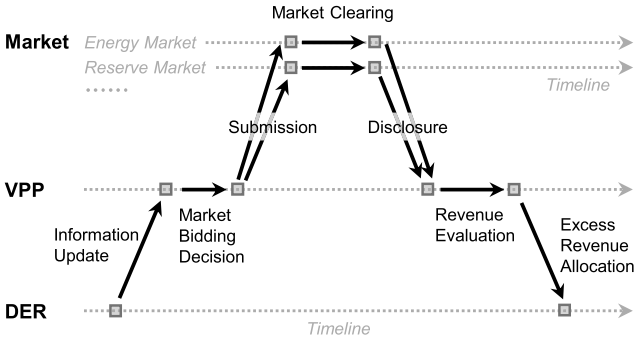


Fig. 7. Timeline for the sequential decisions made by DERs, VPP operators, and market operators.

efficiency. Data-driven controllers were also applied in [109] to schedule EVs, while the system damping and robustness were automatically detected for voltage and frequency regulation.

4.4. Benefit-cost analysis and market

Data resources from electricity markets are utilized to better measure the energy cost, evaluate the potential benefits, and make financial settlements with all individual DER units. Excess revenue is a key concept in the benefit-cost analysis of VPPs [110]. It refers to the extra revenue gain from coordinating DERs in a VPP system compared with a non-coordinated case. A positive excess return is a clear evidence for the value of VPP management, and VPPs could share part of the revenue to incentivize DER units. To get a positive excess return, VPP operators need to take a series of internal and external actions.

Fig. 6 showcases the excess revenue as an intermediate element within a revenue flow running through a VPP system. This flow plays a vital role in connecting a VPP operator with both internal DER units and external markets. Fig. 7 further illustrates the timing and coordination between different stakeholders and different decision steps.

Using a market analysis system, VPP operators are required to carry out internal revenue allocation using either a price-based or a performance-based strategy. In price-based allocation, different price levels for DER units will result in differentiated income surplus among all. A data mining approach was used in [111] to cluster and classify DER units into different categories by their bidding behaviors, and the incentive prices for demand response were configured based on DER categories. Reference [112] analyzed the competitive DER pricing problem in a Markov decision process, and a deep RL was formulated to decide competitive and robust pricing decisions. In performance-based allocation, the entire excess revenue will be distributed to DER units according to their contributions. Much like dividends in financial markets, the allocation policy here becomes the most important detail. This policy was modeled in [113] using cooperative game theory, and the Shapley values were used to decide the allocation portion. Reference [114] used another policy from Nash-Harsanyi bargaining to allocate the profit within a VPP-wind cooperation. This was an asymmetric bargaining game and the contributions of different users were well distinguished by a bargaining power score. Likewise in [115], a customized pricing mechanism was developed for profit allocation to prosumers within a VPP. The authors used a Stackelberg game to capture the heterogeneous dynamics and solved the game with an evolutionary algorithm. Note that a fair distribution should follow the principle of incentive compatibility to ensure positive motivations for each DER unit.

VPP research has long been concerned about external market participation, i.e., the optimal design of market purchase and sale plans based on DER features. The review paper [10] provided a complete summary of different market types, market participation, and optimization models for VPP operators. When zooming in the use of machine learning, the authors from [116] discussed the optimal VPP bidding for a joint market for energy and frequency regulation, and established a bidirectional long short-term memory network to estimate the status of renewable energy generators and EV charging demand. A similar neural network model was implemented in [117] to predict the inputs for Monte-Carlo optimization for the day-ahead market. Another data-driven bidding model was proposed in [118] with a hybrid deep learning forecaster and a data-driven uncertainty set capturing the uncertain renewable power generation. Reference [119] established a data-driven bidding package model for the day-ahead market and a virtual auction model for the real-time market.

The last topic is about the evaluation of benefits and costs, and the research question of interest is how to reliably measure the excess revenue. Such evaluation often involves techno-economic analysis. For example, the techno-economic impacts of massive integration of DERs into VPP systems were analyzed in [120], and different operation strategies were taken into consideration as well. Reference [121] conducted a benefit-cost analysis on a VPP in western Australia, and the results validated the promising financial opportunity with an energy cost reduction of 24% and an internal rate of return of 11%. In [122], another case was tested in Germany, where the revenue of VPPs was projected to increase 11–30% in 2030 if a market-oriented operation mode was utilized.

5. Data sharing and privacy for VPPs

This section will extend the previous decision making models in a decentralized manner to protect data privacy. This is another active area, and the key technologies involve distributed and game-theoretic control, peer-to-peer energy sharing, and blockchain.

5.1. Distributed control and games

VPP data are owned by multiple stakeholders. The core concern is how to merge these data through decentralized collaboration. Such merging is practical to extend connectivity and avoid the curse of isolated data resources. Technically, VPPs take the form of multi-agent systems [123] with different DER units acting as the agents.

Multi-agent systems and distributed control are closely related to game theory, and all these three concepts often appear simultaneously in the literature. The key idea is to protect data privacy by separating data from computing, so that user data can only be accessed locally

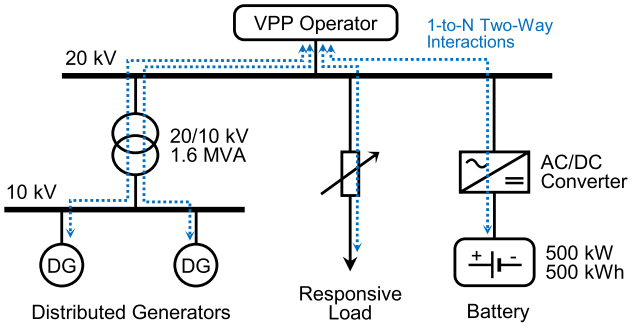


Fig. 8. An illustrative Poland VPP system using the hierarchical interactions and distributed control. In this VPP, different DER communicates with the operator independently.

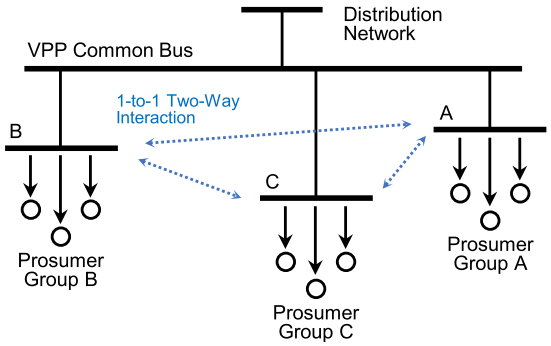


Fig. 9. An illustrative VPP system used for the VirtElect P2P trading platform. In this VPP, three prosumer groups are able to communicate with each other.

with no data uploads to the centralized operator. Below, we will dive into more details about typical options.

Perhaps the most classical example is the hierarchical architecture shown in Fig. 2(a), and the Stackelberg game is a popular model that fits such an architecture. In this VPP, there is a level for DER units and another for a VPP operator. DER units will conduct local computing and deliver the results upward, while the VPP operator coordinates different submissions by broadcasting the coordination signals (e.g., prices) downward. A typical 20kV-level VPP in Poland was reported to use this hierarchical architecture [124] for distributed control, which was further visualized in Fig. 8.

A key concept within this architecture is the Nash equilibrium, which is defined as an optimal outcome where no participant can gain by changing their own actions. In general, these equilibrium states can be obtained by running a centralized optimization model. Reference [125] constructed a Stackelberg game model to maximize the VPP profit. This bilevel model was later transformed into a classical mathematical program using KKT conditions and then solved by commercial solvers. A similar Stackelberg game for prosumers was modeled in [126] using a different approach, namely, the column-and-constraint generation algorithm.

Another way to finding the Nash equilibrium is to implement distributed optimization algorithms such as the alternating direction method of multipliers (ADMM). Different kinds of improvements have been implemented in the literature. Reference [127] applied ADMM to find optimal operating conditions for a group of DERs, while an over-relaxation strategy was adopted to get further speedup. ADMM together with a robust feedback optimization were implemented in [128] to conduct distributed volt/var control for distribution networks with multiple VPPs. Then, in [129], the authors employed the augmented Lagrange function to formulate an accelerated ADMM algorithm with dual acceleration. Reference [130] used a neural network to update Lagrange multipliers in a Stackelberg game, and this strategy could reduce the iteration steps by 63% on average.

Some complicated Nash equilibrium states in VPPs exhibit non-convexity, which cause classical optimization models to failure. The authors of [131] applied a genetic algorithm to find the equilibrium among multiple VPPs, where the game theoretic dispatch model was a large-scale nonlinear program.

RL is also useful to solve for complicated Nash equilibria in VPPs. Reference [132] utilized a deep RL model to solve a Stackelberg game within an EV-integrated VPP. Specifically, a soft actor-critic algorithm was implemented to learn how to sell electricity to EVs, and a twin delay deep deterministic policy gradient algorithm was developed for charging and discharging scheduling. A similar method was used in [133] to approximate agent decisions via an actor-critic network. A decomposition strategy to parallelize the agent training process was also established. Because of abundant control actions and massive controllable appliances, the authors of [134] developed a novel deep RL algorithm by the credit assignment mechanism to accommodate the DERs and residential load effectively. Reference [135] also applied RL to operate DER units, while the algorithm performances were then investigated under different conditions of uncertain renewable generation and load consumption.

5.2. Peer-to-peer sharing

In addition to the hierarchical architecture, VPPs and other multi-agent systems can also operate in a P2P paradigm. A prominent advantage of such a flat architecture is the absence of coordinating centers, which might be prone to single-point failures. This architecture follows the current trend of energy democratization and has become more popular in recent years. A typical P2P architecture in the literature is given in Fig. 2(b), and a specific case is the P2P trading platform VirtElect [136]. Fig. 9 presents the physical network (radial) and communication (P2P) layout of a test case used in [136].

In general, the P2P sharing allows any two agents sharing energy and exchanging information as needed. Reference [137] validated that P2P greatly incentivized DER units to actively aggregate and expand the pool of flexibility resources. Similarly, the concept of energy-as-a-service was proposed in [138] to secure the energy balance and resilience of VPPs using incomplete information.

P2P energy management greatly relies on individual-level decisions, and a major issue is how to guarantee the optimality and efficiency for the whole system. A P2P energy transaction model was formulated in [139] to minimize the community operating costs, and this model was directly solved by a Matlab simulator. A similar P2P model was developed in [140] to decide the pricing and deployment for prosumers in a VPP system. Within this model, the user preference was characterized by a Hotelling model, and a sensitivity analysis demonstrated those relevant influential factors. Reference [141] formulated a stochastic optimization model to analyze the P2P energy trading among VPPs. Such an optimization model could be well solved by most commercial or open-source solvers.

Note that all the above works merely solve the equilibrium solutions, but it is unclear how to reach the equilibrium through iterative control or optimization. In this context, the consensus algorithm is a simple-but-powerful option for linear convergence to the equilibrium. In each round, DER units update their estimations of the coupling variables based on their neighbors' guess until the entire system converges.

Most DER units only communicate with several connected units. In a sparse network, it is impossible to collect full system-wide information, and data privacy is thus protected. Reference [142] studied the distributed energy sharing problem, and the social welfare was maximized by a consensus-based fully distributed algorithm. In [143], a visual pool was designed for prosumer trading and information exchange. Here, the consensus algorithm was utilized to explore the bilaterally-agreed trading price and quantity over multiple time slots. The impacts of time-varying delays on the consensus algorithm were analyzed in [144]. Results revealed that the convergence time was determined by the max-

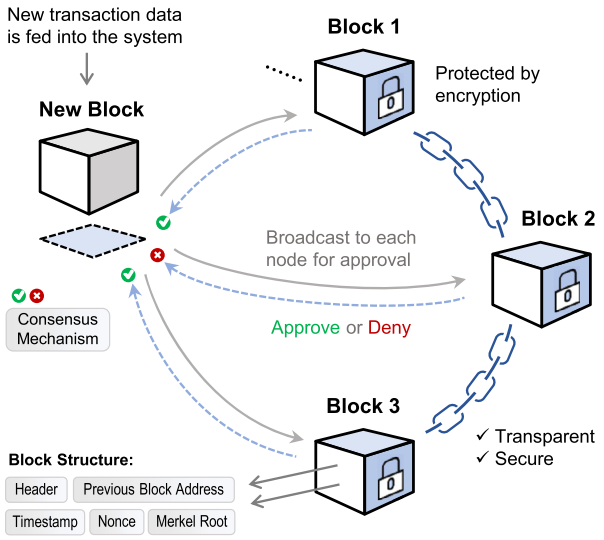


Fig. 10. A typical blockchain workflow and a standard block structure. Any edits such as adding data records require all the blocks reaching a consensus of approval. Encryption techniques of diverse kinds are also implemented to protect data privacy.

imum delays and the number of DER units. Reference [145] developed a distributed economic dispatch model against cyber attacks and solved this model by an isolation rule (to cut out the problematic components) and a consensus algorithm (to reach a consensus).

P2P energy management can be solved by RL algorithms as well, and a typical example is [146]. Here, a constrained deep RL model was formulated to determine optimal energy trades for multiple VPPs.

5.3. Blockchain

Either distributed control or P2P sharing works with raw data flows during agent interactions. These frequent data transmissions might be attacked by data theft or tampering. Blockchain is a new security technology based on P2P sharing, and places two additional add-on features (encryption and mutual verification) into the management workflow. Blockchain successfully upgrades the level of data security and privacy protection. Fig. 10 illustrates a typical pipeline and a standard block structure for blockchain.

Several key concepts are well-known, including the distributed ledger technology, proof of work/energy mechanism, Ethereum, and smart contract technology (see surveys [147] and [148]). Blockchain systems can be categorized into public chains, private chains, and consortium/federated chains. The authors of [149] present a detailed comparison of all the types. Among these three, federated chains generally became the most widely-used option, with both private and public blockchain features.

Existing works in the literature have also tried the integration of blockchain technologies and VPP energy management. For example, a blockchain-based energy management platform was developed in [150] to manage various kinds of DERs in a VPP system. Reference [151] used a public blockchain system and the self-enforcing smart contracts to provide energy services for prosumers. The prosumer-level constraints of energy profiles and service requirements were encoded in the smart contracts accordingly to implement the chain of offer-operate-measure-remunerate actions.

The blockchain community typically cares about the mechanism design and practical applications. Attention has currently been paid to the reputation mechanism, similar to the notion of human reputation. Reference [152] applied a dual blockchain architecture with a reputation-based market segmentation mechanism. The users were treated and divided differently according to their reputation values. Reference [153] introduced the consensus reputation to increase the operating efficiency of blockchain systems, and the reputation was closely relevant to the transaction results and trading behaviors. Additionally, the proof-of-energy mechanism is an extension of the classical proof-of-work, i.e., a cryptographic proof requiring a certain amount of computation. Reference [154] established an energy trading and block mining protocol by integrating differential privacy. Extensive theoretical analysis and validations were provided to analyze market competence, market stability, and expected energy trades. In [155], a collaboration mechanism for blockchain was proposed to accelerate the consensus search, so this novel mechanism was particularly useful for large-scale DER users.

Regarding practical applications, the companies and operators often prefer deploying mature blockchain platforms. For example, the authors of [156] complied smart contracts using Solidity v0.6.6 and Remix v0.9.4. Both tools were well-known for their full functionalities to streamline the system development process of testing, debugging, and deployment. Reference [157] chose to use the Avalanche blockchain platform for P2P trading between VPPs. Avalanche was an open-source platform for deploying business-wide blockchain installations in a highly scalable setup, and had seemingly addressed the Ethereum trilemma with a unique design of consensus mechanism. As for the asymmetric encryption, ElGamal was implemented in [158] with blockchains for auction protections in a VPP system. Reference [159] developed a web-based software to make fog-enabled transactions between DERs. Other efforts were studying the corresponding hardware systems for implementation. Examples of this kind are rare in the literature, and a typical instance is [160]. Here, the authors aggregated DER units using a blockchain-assisted reserve framework, and then made the demonstrations with seven Rock Pi devices.

6. Real-world VPP implementation

This section will collect practical VPP implementations and demonstrations around the world. Representative projects are selected to demonstrate the latest advances in recent years.

6.1. Past projects

Over the past 20 years, quite a few VPP projects of diverse scales, structures, and technologies have been initiated globally [161]. Europe has actively promoted VPP projects since the mid-2000s, and many of the most pioneering projects were originated there. The VPP development in the U.S. was historically slow, but has caught up due to recent interest and increased commercial investment. Asian countries such as China have also launched VPP projects over the past few years, with the preliminary demonstrations dating back to around 2018.

Existing review papers, such as [161,162], have made summary reports of VPP projects, but most of them appeared about 10 years ago. In these reports, Fenix [163] and Edison [164] have been widely recognized as the top two projects with international renown.

The Fenix (Flexible electricity networks to integrate the expected energy evolution) project [163] was funded by the European Commission to boost DERs by maximizing their potential contributions to power systems. From October 2005 to September 2009, this project had implemented five work packages: system solutions for DER integration and demand response through VPPs, electrical and information system architecture adapted to large-scale VPPs, commercial framework for power system operation and control, large-scale VPP concept demonstration, stakeholders advisory group, dissemination and training. This project involved 21 partners in total and had an entire budget of €14.76 million. Two large field deployments were conducted in the U.K. (northern case, 6.8 MW DERs) and in Spain (southern case, 150 MW DERs) to assess two future scenarios with the realistic DER penetration [161].

The Edison project [164] was partially funded by the Danish transmission system operator Energinet.dk through a research grant called

Table 3
Comparison of the selected VPP projects around the world.

Project	Country	Duration	Key Components	VPP Functions
Fenix	Spain, U.K.	2005–2009	Solar, wind, micro CHP	Reactive power control, congestion management, balancing and ancillary services
Edison	Denmark	2009–2012	Solar, wind, EV, micro CHP	Vehicle-to-grid, energy balance
Twenties	Denmark	2010–2013	Solar, wind, hydropower, micro CHP, industrial load, EV, heat pump, building	Voltage control, reserves, and fast frequency demand response
Web2Energy	Germany, Poland	2010–2015	Solar, wind, CHP, biogas, hydropower	Smart energy supply
Smart Power Hamburg	Germany	2011–2015	CHP, heat storage	Energy balance, thermal storage
City-zen	Netherlands	2014–2019	Solar, home battery, residential load	Local energy balancing
Insulae	Denmark	2019–2023	Solar, wind, biogas, EV, thermal storage	Flexibility support
Statkraft VPP	Germany, U.K.	2019–2023	Solar, wind, hydropower, hydrogen, storage, EV, industrial load	Balancing services
Vpp4Islands	France, Germany, U.K, etc.	2020-2024	Solar, wind, DG, battery	Renewable energy integration, reactive power, frequency and voltage control
POSITYF	France, Spain, etc.	2020–2023	Renewable energy	Ancillary services
Sunrun VPP in NE	U.S.	2019–2022	Solar, battery	Capacity reserve services
OhmConnect VPP	U.S.	2020–2023	Solar, wind, residential load	Smart thermostat control, peak load reduction
Shelter Valley VPP	U.S.	2022–2023	Solar, DG, battery, thermostat	Reduce outages, improve reliability, demand shift and reduction
SPEEDIER	Canada	2018–2022	solar, battery, EV, controllable water tanks	Energy balance, reduce peak load
AGL VPP in SA	Australia	2017–2022	Solar, battery	Demand-side management, voltage control, charging fleet management
Tesla VPP in SA	Australia	2020–2023	Solar, battery (Powerwall)	Voltage support, fast frequency response, inertia/emergence response
Northern Hebei VPP	China	2019	DG, storage, EV, heat pump, industrial/commercial users	Demand response, energy trading, ancillary services
Shenzhen VPP	China	2020	Solar, storage, EV	Energy balance, voltage control
Miyakojima VPP	Japan	2021–2024	Solar, battery (Powerwall)	Solar energy integration, self-sufficient energy supply

Note: “NE” is New England region, “SA” is Southern Australia; “CHP” is combined heat and power generator, “EV” is electric vehicle, “DG” is distributed generator, “Powerwall” is a residential battery product from Tesla.

Forskel. The total budget was around 49 million Danish Krone, and the project partners broadly included energy companies, technology suppliers, and research institutes. This project was mainly focused on how to integrate large-scale EVs through VPPs to support grid operation and decarbonization. The Danish island of Bornholm was selected for VPP implementation because it contained an isolated power system with the high wind and EV penetration. This case succeeded in validating the coordination and balancing strategies between wind turbines and EV smart charging.

There are also some other typical projects including Twenties [165], Web2Energy [166], Smart Power Hamburg [167], and City-zen [168]. Table 3 presents a detailed summary and comparison of all the projects involved in this section.

6.2. Recent and ongoing projects (since 2017)

We next select several recent and ongoing projects since 2017 and summarize their most latest progress. The details of these projects are presented in Table 3 as well.

6.2.1. Projects in Europe

Europe still plays a leading role in utilizating VPPs in various applications.

The Insulae project [169] was one of the Horizon 2020 projects funded by the European Union during 2019–2023. With a €12.07 million budget, this project was committed to developing innovative solutions for European island decarbonization, and the key outcome was an investment planning tool for island management. The research studies from this project included VPP control strategies, flexibility potential assessment, models and simulation, feasibility tests, and several on-site implementations. A use case was undertaken in the Danish Bornholm island to investigate how a renewable-rich VPP system coordinated with other energy sectors (gas, heat, transportation) to deliver flexibility services to island energy systems.

The Vpp4Islands (VPP for interoperable and smart islands) project [170] was another Horizon 2020 project spanning from 2020 to 2024, and the entire budget was €7.22 million. This was also a island-related project focusing on exploiting intermittent renewable energy such as solar and wind through VPP systems. The ultimate goal was to revolutionize conventional VPPs by using virtual energy storage, digital twin, and distributed ledger technologies to enable smart energy communities on islands. Two demonstration sites were selected: one was Gokceada island (Turkey) to validate the mixed storage solution, and the other site was Formentera island (Spain) to test the feasibility of smart energy communities. Besides, there were three replication cases in Bozcaada island (Turkey), Grado island (Italy), and Bornholm island (Denmark) as well.

The POSITYF (Powering system flexibility in the future through renewable energy sources) project [171] was also funded by the Horizon 2020 programme for a duration of 2020–2023. 11 partners were involved and the total budget was €4.73 million. This project aimed to develop a novel concept of dynamic VPPs [172], which provided ancillary services and flexibility support through a portfolio of DERs. The work plans for this project intended to design the architecture of dynamic VPPs and investigated the control and optimization models for system-level scheduling.

The Statkraft VPP project [173] was launched by the Statkraft company, which was an active service provider in European markets. Statkraft aimed to operate the largest VPP in Europe with more than 10 GW installed capacity from 1k power generators. Since November

2019, this VPP has entered the U.K. balancing market with 20 MW highly flexible gas generators in Southwest England. Such fast-acting flexibility capacity was realized through real-time communication within the VPP system.

6.2.2. Projects in North America

The U.S. and Canada have experienced a boom in the number of VPP projects because of an increasing interest in commercial investments in North America.

The Sunrun VPP project [174] in New England, the U.S. has been initiated since 2019 by the Sunrun company, a clean energy technology provider of household solar, battery storage, and intelligent energy services. Aggregating local solar energy, the project team won the first cost-competitive bid in the New England electricity market by substituting polluting fossil fuel power plants. In the summer months of 2019, the VPP system returned 1.8 GWh energy back to the grid and successfully reduced peak energy demand. This project provided convincing evidences that local solar and batteries could support wholesale markets and reduce the grid costs for public good.

The OhmConnect VPP project [175] was implemented in California, the U.S. by the OhmConnect company. This project was partially funded with $80 million by the Sidewalk infrastructure partners. Starting in 2020, OhmConnect rewarded residential users with cost-saving plans by controlling smart plugs and thermostats. Located in California, this project aimed to gather residential networks to form the largest residential VPP, namely Resi-Station, in North America with a total capacity approaching 550 MW. In August 2020, this VPP managed to reduce 1 GWh energy demand by coordinating 150k residential customers.

The Shelter Valley VPP pilot project [176] was an initiative proposed by the San Diego Gas & Electric company and Electric Power Research Institute (EPRI) in order to build a resilient power grid by boosting high-level reliability and emergency preparedness. This project was launched in 2022 in the unincorporated community of Shelter Valley in East San Diego County. Over an 18-month period, this VPP project investigated how DERs (smart thermostats, load controllers, and battery storage facilities) could act as a resource for energy balance on the grid. In this project, selected participants could receive incentives if they signed a demand response contract and allowed a free installation of smart thermostats or smart plugs in their house.

The SPEEDIER (smart, proactive, enabled, energy distribution, intelligently, efficiently, responsive) project [177] was funded with over C$3.7 million by Natural Resources Canada. This four-year project spanning from 2018–2022 was mainly focused on the DER management of load and variable renewable generators. This project succeeded in automating the management of solar, energy storage, EVs and many other resources to reduce heavy loading and reduce dependencies on long-distance energy generation.

6.2.3. Projects in Australia

Australia is another hotspot for successful development of VPP projects, especially in the populous Southern Australian region.

The AGL VPP project [178] in Southern Australia was led by the AGL Energy limited company and partially funded by the Australian Renewable Energy Agency (ARENA) from 2017 to 2022. This project was the next phase of a previous three-year project and managed a total budget of AU$19.52 million (AU$5.3 million from ARENA). The project developed a cloud-based control system to connect a number of solar+battery systems to operate as a 5 MW solar power plant across 1k residential and business premises in Adelaide, South Australia. Results showed that the residential energy storage installed behind the meter could offer grid services through intelligent control. This project also shared knowledge around user perceptions of VPPs and marketing programs for commercial use.

The Tesla VPP project [179] in Southern Australia was led by the Tesla Motors Australia company and partially funded by ARENA during 2020–2023. The total budget of this project was AU$64.17 million, including AU$8.2 million from ARENA. In this project, a network of 50k solar and Tesla Powerwall household battery systems would construct a large regional VPP system. Tesla typically used WiFi and advanced control to manage the Powerwall batteries and perform energy arbitrage in the Australian national energy market. This project aimed to demonstrate the commercial viability and technical feasibility of a VPP solution in Southern Australia. The VPP demonstration also raised interest in how to provide grid stability services by VPPs.

In addition, the Australia energy market operator has made the pioneering demonstrations during 2019–2021 to inform changes to regulatory frameworks and operational outcomes by applying eight VPP portfolios (31 MW, 7150 consumers) in the national markets [180]. The four knowledge-sharing reports provided a number of technical details that showed a successful participation of VPPs in offering contingency frequency support, responding to price signals, and delivering local network services (sometimes more than one service simultaneously).

6.2.4. Projects in Asia

China, Japan, and many other Asian countries initially trailed in VPP development, but substantial progress has taken place in recent years. Below, three typical projects are selected and described in detail.

The Northern Hebei VPP project [181] was funded by the state grid Jibei electric power company around 2019 to validate the DER values in decarbonization and flexibility support. Located in Northern Hebei, China, this VPP system involved distributed generators, flexible loads (EVs, heat pumps, and industrial & commercial users), and energy storage. The overall system capacity was 358 MW, and 204 MW (57.0%) were flexible resources.

The Shenzhen VPP project [182] was launched in the Shenzhen power grid (city-level system), China around 2020 to utilize flexible DERs to support the grid operation. Regional VPPs were formulated based on the controllable units to handle the voltage issues caused by overloading. This VPP system typically involved a 5 MW charging station, a 10.4 MW bay eco-park, and a 0.7 MW industrial park.

The Miyakojima VPP project [183] has been initiated by the Miyakojima Mirai Energy company and implemented in Miyako island, Okinawa Prefecture, Japan since 2021. This project has constructed one of the largest commercial VPPs in Japan by connecting and extending the installation of Tesla Powerwall batteries to 600 units or more. The aggregation of mass charging and discharging brought benefits to stabilize the island grid effectively.

6.3. Energy management software

The rapid growth of VPP projects has triggered an emerging need for software services. Reference [184] summarized the typical requirements and functions of VPP software. Currently, there are a few VPP software products available on the market, including the ABB Optimax software [185], the Next Kraftwerke NEMOCS software [186], the AutoGrid Flex software [187], and the GE AEMS software [188]. An extended list of relevant software tools were collected in [189] (page xi, Table 1) as well.

Most of the above software tools originate from the DER management modules, and implement the prevalent design of platform-as-a-service (PaaS) or software-as-a-service (SaaS). In addition, they are upgraded with customized functions for utility-level and enterprise-level VPP projects. It is clear that this type of software tools accelerates the deployment of VPP projects.

7. Challenges and opportunities

This section will discuss the major challenges and future opportunities in the area of data-driven energy management of VPPs. Fig. 11 gives a high-level overview that covers all the data phases and the emerging technologies for VPP energy management.

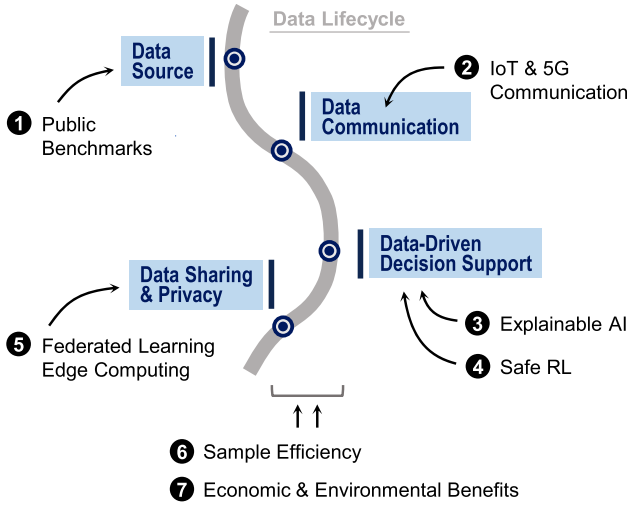


Fig. 11. Major challenges and opportunities through the entire data lifecycle of VPP energy management. A series of new technologies and solutions are labeled accordingly.

7.1. Public benchmarks

Evaluating energy management schemes for VPPs requires a collection of consistent criteria, test data, and simulation systems. However, most of the existing case studies are running diverse forms of simulation and may not follow a rigorous process to validate their results. This issue makes it challenging to reproduce and compare the existing schemes for VPP management.

Public benchmarks and open data are powerful resources to solve the above challenge of reproducibility and validation. However, the existing public benchmarks in VPP research are less popular. For example, Section 2 has summarized a list of publicly available data for VPPs, but those datasets are not common in the research community. Although it is impossible to establish an all-inclusive benchmark that is universally applicable for any VPP modeling, we can instead start with fundamental tasks and make progress step by step.

In the future, the VPP research community should promote reproducibility policies and culture through the task forces or committees for standardization. The benchmark design should fully consider the different needs of VPP modeling and simulation, such as active & reactive power control, a broad range of temporal resolutions, and different support systems (multi-energy conversion or communication infrastructures). The benchmarks and data sources for VPPs should be compatible with the interfaces of popular simulators and programming languages. In specific, electrical engineers often prefer compatible data formats for OpenDSS, Homer, and Matpower. Regarding the sensitive data, using data masking or anonymization [190] is a common option, and there should be a proper balance between data usability and security. Another popular guidance for data formatting is the FAIR (Findable, accessible, interoperable, and reusable) principle. First published in 2016 [191], FAIR principle described how data should be organized for easy access and use. Furthermore, it would be helpful to integrate iterative development and active maintenance to troubleshoot and streamline the use of VPP data.

7.2. Internet of things and 5G communication

High-quality wide-area communication is a major support technology for VPP systems. The communication module is responsible for ensuring the real-time or quasi-real-time data interaction between a VPP operator and massive controllable DER units. It is tough to support the fast-growing communication demand from new sensors and smart terminals. The potential communication limits and bottlenecks may degrade or even disable the services supplied by VPPs. Therefore, the efficient data communication with low latency and high reliability becomes the key to unlocking the flexibility potential of VPPs.

Internet of things (IoT) and 5G communication are two novel technologies for safe and ubiquitous connectivity. Reference [192] discussed the use of 5G networking slicing to control DERs for providing frequency regulation services. There are also several recent works, such as [193,194], attempting to utilize the cloud/fog architecture in the literature, but it is generally rare to come across such efforts. In fact, the heterogeneity of communication components in VPPs appears to be a major difficulty, and it is important to simulate the energy-communication coordination in VPPs.

There are several future directions for VPP communication. It is important to study the dynamic networking and coordination approaches that match the plug-and-play characteristics of DERs. Precise identification and efficient resource allocation will be achieved by the differentiated service priority for short- and long-distance communication. Another topic is the capacity of networked-and-heterogeneous channels. More efforts should be put on the advanced gateway devices, differentiated QoS strategies, and modular software platforms for the open services gateway initiative (OSGi).

7.3. Explainable artificial intelligence

Big data and AI technologies could provide a range of new opportunities by utilizing the VPP data. Deep learning and other AI-based approaches can help identify hidden operation patterns for VPPs; however, the black box nature of learning models makes their performance and outcomes hard-to-interpret and fragile to adversarial examples. In VPP systems, such undesirable features of AI may lead to unknowingly irrational decisions and biases in different levels of coordination.

In practice, these AI or deep learning models can only be trusted by VPP operators if their outputs are fully understandable and reliable. Interpretability is thus the key and an explainable solution is crucial to maintain transparency and fairness in AI-based decision making for VPPs.

Explainable artificial intelligence (XAI) is a recent paradigm to address the above challenge. XAI is a group of tools or frameworks to interpret the outputs and express model reasoning of deep learning models. Technically, the performance can be analyzed in terms of transparency and explanations. Transparency refers to the level of human understanding of the internal processes in a given learning model. Besides, several forms of explanations have been discussed in the literature, such as text/visual explanations, local explanations, explanations by use cases, and feature relevance.

Note that the existing applications of XAI in VPP research are still at a preliminary stage, and it is currently unclear how to apply XAI in VPP energy management. Reference [195] made a tentative effort to construct an ensemble model that provided interpretable uncertainty forecasts for VPP load and renewable energy. In [196], an explainable framework was employed to demonstrate how a neural network detected the cyber anomalies from the voltage and frequency control signals. Beyond these attempts, a number of other applications are expected to become possible by using XAI.

The future energy management of VPPs may include the XAI technology in the entire workflow. A particular interest is in the advanced methodologies and practical software platforms for VPP use cases. For example, it is possible to extend the local interpretable model-agnostic explanation (LIME) and Shapley additive explanation (SHAP) technologies by integrating counterfactual analysis and physical constraints for VPPs. This will contribute to identifying the key factors in an explainable manner. Building upon useful tools such as the what-if tool (WIT), a validation flowchart is able to test the robustness and fairness of VPP system models.

7.4. Safe reinforcement learning

Reinforcement learning is an attractive and popular option for VPP energy management. But a major problem of RL is the physical infeasibility of the model outputs because RL is essentially an unconstrained learning approach that ignores system-level constraints of VPPs. This leads to potential security risks and keeps RL out of real-world VPP projects. In theory, this characteristic is an endogenous defect that cannot be easily solved with better data volume or quality. It further deteriorates when training RL models with noisy data measurements from VPP systems.

Instead, the safe RL aims develop algorithms that explicitly satisfy the physical constraints. The exploration of safe RL in VPPs has been initiated in the literature. Reference [197] developed an adversarial safe RL approach based on a two-player game framework, in which a target agent maximized the cumulative rewards and an adversary agent intended to destroy the rewards. Reference [198] then implemented safe RL in strategic bidding tasks by a projection-based safety shield and a shield penalty in the reward function. It is evident that safe RL still needs further refinement for practical VPP applications.

In the future, research attention should be concentrated on the feasibility restoration for hard constraints in VPP systems. Multiple types of restoration strategies will be needed for VPP energy management, such as projection-based approaches and physics-informed learning. It is also imperative to take into account whether feasibility restoration steps might have adverse impacts on the model generalization performance and sample complexity. In addition, combining RL with robust optimization is a promising option to take uncertainties into consideration and adapt to use cases with different levels of safety requirements. Here, the key issue is to design customized reward functions and stochastic learning policies.

7.5. Federated learning and edge computing

Data security and privacy have not yet become a major issue for VPP operators, so research and implementation progress is slow in this area. However, they are undoubtedly important because VPPs need to interact with a number of stakeholders that show very diverse features and preferences. Failure to protect data privacy in this case will result in the loss of trust and support from these stakeholders. Unfortunately, many existing VPP solutions have technological limitations in database maintenance and user privacy protection.

Section 5 mentioned the solutions of distributed control, P2P sharing, and blockchain for VPPs. In fact, there are another two emerging options for distributed decision making, i.e., federated learning and edge computing.

Federated learning (FL), also known as collaborative learning, is a decentralized learning approach that does not require exchanging local datasets. Heterogeneous datasets could be fed into a FL model to expand the cloud-side model pool and contribute to improving the whole system performance. FL provides a novel perspective for VPP privacy protection, but there are very few discussions in the literature at present. Reference [199] was one of these studies that trained energy management models across multiple VPPs by federated reinforcement learning. A new design was the robust gradient filter to relieve the impacts of operational disturbances. Clearly, the use of FL in VPP systems is still very preliminary, so more research and practical inputs are urgently needed.

Edge computing refers to a distributed paradigm that allows local data processing close to the sources. Technically, cloud and edge computing share a few similarities with federated learning, and they are highly promising for VPP energy management. Except for some preliminary efforts such as [200], this line of research is still very limited.

Regarding future directions, one typical issue is how to apply FL and edge computing in VPP energy management to align with the heterogeneous local datasets and learning capabilities of DER units. The convergence of FL and cloud-edge coordination are still unsatisfactory in practice because of the potential oscillations. An interesting topic is the two-layer FL architecture. In this architecture, the group learners and global learner can interact with each other. Adaptive grouping might make this architecture flexible and robust against unexpected disturbances. Furthermore, both FL and cloud/edge computing could be enhanced by encryption technologies such as blockchain or differential privacy to seek a higher level of privacy protection.

7.6. Sample efficiency of deep learning

Real-world VPP data are never perfect, and one of the common issues is the limited amount of data available for practical use. This seems obvious for new DER participants whose historical data might be absent, but it is also the case for some others with changing behaviors or preferences in energy use (existing data are useless). This lack of data will have negative impacts on any deep learning approach for VPP management. Training deep learning models with small datasets is prone to overfitting risks or a reduced level of generalization performance. Such models cannot guarantee reliable performance and may mislead VPP operators to biased decisions, especially for out-of-sample inferring.

The concept of sample efficiency is introduced to describe how efficiently a learning model leverages the available data. Sample efficiency is formally defined as the minimal amount of data required to achieve an expected level of performance. It is evident that sample-efficient deep learning models are preferred and should be prioritized for VPP use cases.

Currently, the sample efficiency issue is still rarely studied in VPP research. For example, a knowledge-augmented training framework was established in [201] to improve sample efficiency by integrating domain knowledge in the training procedure. Reference [202] used experience replay and a semi-deterministic phase to increase the sample efficiency in reinforcement learning. Beyond the priority experience replay, the authors of [203] used transfer learning to further boost the learning performance.

In the future, the VPP research community should make efforts to develop novel categories of hybrid models that reduce the data requirement by integrating domain knowledge. Another alternative is data augmentation, which can be accomplished by creating synthetic data or implementing experience replay strategies. Measuring and predicting the data volume threshold to achieve a given performance level is also a critical issue. This threshold is useful to determine when to select a simplified analytical model instead of deep learning for a given VPP system.

7.7. Economic and environmental impacts

VPPs are expected to bring significant economic and environmental benefits through the deep integration of distributed resources, but the detailed accounting of these benefits might be very complicated. Typical reasons include a broad range of influential factors and the problems of high uncertainty and risks.

For example, it is never an easy work to get a robust measurement of the economic efficacy of VPP projects. Using a Poland VPP system, the authors of [204] demonstrated that a VPP was able to achieve a positive net present value (or economically viable) but this value was mostly affected by the energy prices. As for the environmental benefits, a common issue is how much carbon emission is reduced due to a large penetration of distributed renewable energy. In [205], a VPP in California was found to achieve a cost drop of 29.5%, a carbon emission reduction of 79%, and a NOx reduction of 83%. Similar benefits of cost saving and emission reduction were also detected for a sustainable VPP in [206]. There are another line of efforts focusing on carbon trading and electricity-carbon joint dispatch. Reference [207] translated the environmental cost into a monetary penalty using the carbon trading price.

The previous sections have provided rich evidences that data can improve energy management efficiency and thereby enhance the economic and environmental benefits of VPPs. However, it is still very preliminary to identify the additional benefits brought by data-driven decision making. It is also unclear how to perform sensitivity analysis to find out the group of major factors.

The future work needs to dive into the full-cycle economic and carbon accounting for VPPs. It is thus possible to build up connections between decision-making efficiency and economic & environmental benefits. There is a need to expand the accounting data sources by further digitizing VPP systems with a larger sensor network. Higher frequency offers potential for high data rates and more operational details for benefit analysis as well. Note that the data management and metering subsystems also consume energy and cause an increase in the cost and emissions, which should also be considered in the accounting.

8. Conclusion

Virtual power plant is a promising solution for large-scale DER management, attracting extensive attention from both academia and industry. Many evidences have shown that VPPs hold real potential to decarbonize energy production with distributed renewable energy.

This review paper has taken a data-centric perspective to summarize the progress of VPP energy management. In particular, we pay attention to the development in the major phases of data source/creation, data communication, data-driven decision-making, data sharing, and privacy protection. The use of modern data science and machine learning in different support systems are broadly inspected. This paper further makes a collection of VPP projects across the globe, and validates the unique role of data in promoting the grid integration of large-scale DERs. Finally, the key challenges, future opportunities, and potential solutions have been fully discussed.

According to our systematic review, support systems such as data quality control and communication platforms are less developed and require full attention. In terms of practical implementation, the majority of existing projects are working on local energy balancing, while some recent pioneers have explored other ancillary services. In the future, persistent efforts will be made to enable a broad range of grid services and establish a pool of flexibility resources by implementing advanced control.

VPP systems are expected to realize their full potential as a novel networking and decarbonization technology with support from modern data science and deep learning. This paper is intended as a holistic outlook and navigation of this future trend.

CRediT authorship contribution statement

Guangchun Ruan: Writing – review & editing, Writing – original draft, Visualization, Conceptualization. **Dawei Qiu:** Writing – review & editing, Writing – original draft. **S. Sivaranjani:** Writing – review & editing, Writing – original draft. **Ahmed S.A. Awad:** Writing – review & editing. **Goran Strbac:** Writing – review & editing, Supervision.

Declaration of competing interest

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

Data availability

No data was used for the research described in the article.

References

[1] Adu-Kankam Kankam O, Camarinha-Matos Luis M. Towards collaborative virtual power plants: trends and convergence. Sustain Energy Grids Netw 2018;16:217–30.
[2] Saboori Hedayat, Mohammadi M, Taghe R. Virtual power plant (VPP), definition, concept, components and types. In: 2011 Asia-Pacific power and energy engineering conference. IEEE; 2011. p. 1–4.
[3] Yavuz Levent, Onen Ahmet, Muyeen SM, Kamwa Innocent. Transformation of microgrid to virtual power plant–a comprehensive review. IET Gener Transm Distrib 2019;13(11):1994–2005.
[4] Zhang Jiahui, Xu Zhiyu, Xu Weisheng, Zhu Feiyu, Lyu Xiaoyu, Fu Min. Bi-objective dispatch of multi-energy virtual power plant: deep-learning-based prediction and particle swarm optimization. Appl Sci 2019;9(2):292.
[5] Skyquest team. Global virtual power plant market insights. Available: https://www.skyquestt.com/report/virtual-power-plant-market, 2023 [Online].
[6] IBM team. What is data lifecycle management? Available: https://www.ibm.com/topics/data-lifecycle-management, 2023 [Online].
[7] Nosratabadi Seyyed Mostafa, Hooshmand Rahmat-Allah, Gholipour Eskandar. A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems. Renew Sustain Energy Rev 2017;67:341–63.
[8] Rouzbahani Hossein Mohammadi, Karimipour Hadis, Lei Lei. A review on virtual power plant for energy management. Sustain Energy Technol Assess 2021;47:101370.
[9] Sierla Seppo, Pourakbari-Kasmaei Mahdi, Vyatkin Valeriy. A taxonomy of machine learning applications for virtual power plants and home/building energy management systems. Autom Constr 2022;136:104174.
[10] Naval Natalia, Yusta Jose M. Virtual power plant models and electricity markets-a review. Renew Sustain Energy Rev 2021;149:111393.
[11] Kim Mijeom, Choi Jungin, Yoon Jaeweon. Development of the big data management system on national virtual power plant. In: 2015 10th international conference on P2P, parallel, grid, cloud and Internet computing (3PGCIC). IEEE; 2015. p. 100–7.
[12] Syed Dabeeruddin, Zainab Ameema, Ghrayeb Ali, Refaat Shady S, Abu-Rub Haitham, Bouhali Othmane. Smart grid big data analytics: survey of technologies, techniques, and applications. IEEE Access 2020;9:59564–85.
[13] Kapoor Shubhankar, Sturmberg Bjorn, Shaw Marnie. A review of publicly available energy data sets. The Australian National University; 2020.
[14] Arribas Luis, Lechon Yolanda, Perula Alberto, Dominguez Javier, Ferres Manuel, Navarro Jorge, et al. Review of data and data sources for the assessment of the potential of utility-scale hybrid wind–solar PV power plants deployment, under a microgrid scope. Energies 2021;14(21):7434.
[15] Silwal Sushil, Mullican Colton, Chen Yi-An, Ghosh Avik, Dilliott John, Kleissl Jan. Open-source multi-year power generation, consumption, and storage data in a microgrid. J Renew Sustain Energy 2021;13(2).
[16] Vink Karina, Ankyu Eriko, Koyama Michihisa. Multiyear microgrid data from a research building in Tsukuba, Japan. Sci Data 2019;6(1):1–9.
[17] Sarajlic Džanan, Uhlenbrock Simon, Strunck Christoph, Rehtanz Christian. Modelling of microgrid benchmark networks based on the simbench data set. In: 2021 IEEE PES innovative smart grid technologies Europe (ISGT Europe). IEEE; 2021. p. 01–5.
[18] Farhangi Hassan, Joos Geza. Microgrid benchmarks; 2019.
[19] Qi Ning. Portfolio optimization of generic energy storage-based virtual power plant under decision-dependent uncertainties. Available: https://data.mendeley.com/datasets/gz6trjhdcr/1, 2022 [Online].
[20] Wang Siyuan, Wu Wenchuan, Chen Qizhan, Yu Junjie, Wang Peng. Supplemental file for stochastic flexibility evaluation for virtual power plants by aggregating distributed energy resources. Available: https://figshare.com/articles/dataset/Supplemental_File_for_Stochastic_Flexibility_Evaluation_for_Virtual_Power_Plants_by_Aggregating_Distributed_Energy_Resources/23717610, 2022 [Online].
[21] Baringo Ana, Baringo Luis, Arroyo Jose M. An adjustable robust optimization approach for the expansion planning of a virtual power plant. Available: https://ieee-dataport.org/documents/adjustable-robust-optimization-approach-expansion-planning-virtual-power-plant-data-case, 2022 [Online].
[22] Royapoor Mohammad, Pazhoohesh Mehdi, Davison Peter J, Patsios Charalampos, Walker Sara. Data for: building as a virtual power plant, magnitude and persistence of deferrable loads and human comfort implications. Available: https://data.mendeley.com/datasets/g523kyy4v7/1, 2020 [Online].
[23] Jordon James, Szpruch Lukasz, Houssiau Florimond, Bottarelli Mirko, Cherubin Giovanni, Maple Carsten, et al. Synthetic data–what, why and how? ArXiv preprint arXiv:2205.03257, 2022.
[24] Raghunathan Trivellore E. Synthetic data. Annu Rev Stat Appl 2021;8:129–40.
[25] Seib Viktor, Lange Benjamin, Wirtz Stefan. Mixing real and synthetic data to enhance neural network training–a review of current approaches. ArXiv preprint arXiv:2007.08781, 2020.
[26] Karimyan Peyman, Abedi Mehrdad, Hosseinian Seyyed Hossein, Khatami Roohallah. Stochastic approach to represent distributed energy resources in the form of a virtual power plant in energy and reserve markets. IET Gener Transm Distrib 2016;10(8):1792–804.
[27] Han Dongho, Lee Jay H. Two-stage stochastic programming formulation for optimal design and operation of multi-microgrid system using data-based modeling of renewable energy sources. Appl Energy 2021;291:116830.
[28] Wang Xiaobo, Huang Wentao, Tai Nengling, Shahidehpour Mohammad, Li Canbing. Two-stage full-data processing for microgrid planning with high penetrations of renewable energy sources. IEEE Trans Sustain Energy 2021;12(4):2042–52.

[29] Camal Simon, Teng Fei, Michiorri Andrea, Kariniotakis Georges, Badesa L. Scenario generation of aggregated wind, photovoltaics and small hydro production for power systems applications. Appl Energy 2019;242:1396–406.
[30] Hosseini S, Kelouwani Sousso, Agbossou Kodjo, Cardenas Alben, Henao Nilson. A semi-synthetic dataset development tool for household energy consumption analysis. In: 2017 IEEE international conference on industrial technology (ICIT). IEEE; 2017. p. 564–9.
[31] Chen Jun, Rabiti Cristian. Synthetic wind speed scenarios generation for probabilistic analysis of hybrid energy systems. Energy 2017;120:507–17.
[32] Chen Wen, Zhou Kaile, Yang Shanlin, Wu Cheng. Data quality of electricity consumption data in a smart grid environment. Renew Sustain Energy Rev 2017;75:98–105.
[33] Poplawski Tomasz, Dudzik Sebastian, Szelag Piotr, Baran Janusz. A case study of a virtual power plant (vpp) as a data acquisition tool for pv energy forecasting. Energies 2021;14(19):6200.
[34] Jasinski Michal. Combined correlation and cluster analysis for long-term power quality data from virtual power plant. Electronics 2021;10(6):641.
[35] Dong Guangzhong, Chen Zonghai. Data-driven energy management in a home microgrid based on Bayesian optimal algorithm. IEEE Trans Ind Inform 2018;15(2):869–77.
[36] Yuan Dongmei, Lu Zhigang, Zhang Jiangfeng, Li Xueping. A hybrid prediction-based microgrid energy management strategy considering demand-side response and data interruption. Int J Electr Power Energy Syst 2019;113:139–53.
[37] Bhuiyan Erphan A, Hossain Md Zahid, Muyeen SM, Fahim Shahriar Rahman, Sarker Subrata K, Das Sajal K. Towards next generation virtual power plant: technology review and frameworks. Renew Sustain Energy Rev 2021;150:111358.
[38] Roozbehani Mohammad Mohammadi, Heydarian-Forushani Ehsan, Hasanzadeh Saeed, Ben Elghali Seifeddine. Virtual power plant operational strategies: models, markets, optimization, challenges, and opportunities. Sustainability 2022;14(19):12486.
[39] Asl Saeed Abrisham Foroushan, Bagherzadeh Leila, Pirouzi Sasan, Norouzi Mohammadali, Lehtonen Matti. A new two-layer model for energy management in the smart distribution network containing flexi-renewable virtual power plant. Electr Power Syst Res 2021;194:107085.
[40] Pasetti Marco, Rinaldi Stefano, Manerba Daniele. A virtual power plant architecture for the demand-side management of smart prosumers. Appl Sci 2018;8(3):432.
[41] Seven Serkan, Yao Gang, Soran Ahmet, Onen Ahmet, Muyeen SM. Peer-to-peer energy trading in virtual power plant based on blockchain smart contracts. IEEE Access 2020;8:175713–26.
[42] Zajc Matej, Kolenc Mitja, Suljanovic Nermin. Virtual power plant communication system architecture. In: Smart power distribution systems. Elsevier; 2019. p. 231–50.
[43] Sood Vijay K, Fischer Daniel, Eklund JM, Brown Tim. Developing a communication infrastructure for the smart grid. In: 2009 IEEE electrical power & energy conference (EPEC). IEEE; 2009. p. 1–7.
[44] Kabalci Yasin. A survey on smart metering and smart grid communication. Renew Sustain Energy Rev 2016;57:302–18.
[45] Serban Ioan, Cespedes Sandra, Marinescu Corneliu, Azurdia-Meza Cesar A, Gomez Juan S, Hueichapan Doris Saez. Communication requirements in microgrids: a practical survey. IEEE Access 2020;8:47694–712.
[46] Nwauka Oliver, Telukdarie Arnesh, Enslin Johan. Virtual power plant basic requirements for integration of distributed energy resources, driven by industry 4.0. IEOM 2018;2:511–23.
[47] Kolenc Mitja, Nemček Peter, Gutschi Christoph, Suljanovic Nermin, Zajc Matej. Performance evaluation of a virtual power plant communication system providing ancillary services. Electr Power Syst Res 2017;149:46–54.
[48] Schwarz Karlheinz. Difference between IEC 60870-5-104 and IEC 61850. Available: https://blog.nettedautomation.com/2019/02/difference-between-iec-60870-5-104-and.html, 2019 [Online].
[49] Etherden Nicholas, Vyatkin Valeriy, Bollen Math HJ. Virtual power plant for grid services using IEC 61850. IEEE Trans Ind Inform 2015;12(1):437–47.
[50] Nadeem Furquan, Aftab Mohd Asim, Hussain SM Suhail, Ali Ikbal, Tiwari Prashant Kumar, Goswami Arup Kumar, et al. Virtual power plant management in smart grids with XMPP based IEC 61850 communication. Energies 2019;12(12):2398.
[51] Rinaldi Stefano, Pasetti Marco, Ferrari Paolo, Massa Giovanni, Della Giustina Davide. Experimental characterization of communication infrastructure for virtual power plant monitoring. In: 2016 IEEE international workshop on applied measurements for power systems (AMPS). IEEE; 2016. p. 1–6.
[52] Kolenc Mitja, Ihle Norman, Gutschi Christoph, Nemček Peter, Breitkreuz Thomas, Goedderz Karlheinz, et al. Virtual power plant architecture using OpenADR 2.0 b for dynamic charging of automated guided vehicles. Int J Electr Power Energy Syst 2019;104:370–82.
[53] Rekik Mouna, Chtourou Zied, Mitton Nathalie. A scalable geographic routing protocol for virtual power plant communications. In: 2015 IEEE 15th international conference on environment and electrical engineering (EEEIC). IEEE; 2015. p. 1742–7.
[54] Jansen Bernhard, Binding Carl, Sundstrom Olle, Gantenbein Dieter. Architecture and communication of an electric vehicle virtual power plant. In: 2010 first IEEE international conference on smart grid communications. IEEE; 2010. p. 149–54.
[55] Tang Jia, Zhang Xi. Quality-of-service driven power and rate adaptation over wireless links. IEEE Trans Wirel Commun 2007;6(8):3058–68.
[56] Sun Wei, Yuan Xiaojing, Wang Jianping, Han Dong, Zhang Chongwei. Quality of service networking for smart grid distribution monitoring. In: 2010 first IEEE international conference on smart grid communications. IEEE; 2010. p. 373–8.
[57] Zhou Bin, Zhang Kuan, Chan Ka Wing, Li Canbing, Lu Xi, Bu Siqi, Gao Xiang. Optimal coordination of electric vehicles for virtual power plants with dynamic communication spectrum allocation. IEEE Trans Ind Inform 2020;17(1):450–62.
[58] Suo Siliang, Chen Liming, Cheng Renli, Kuang Xiaoyun, Zou Jin. Design of secure access to distributed load resources of virtual power plant based on virtual communication private network. In: 2021 IEEE sustainable power and energy conference (iSPEC). IEEE; 2021. p. 4142–9.
[59] Heimgaertner Florian, Menth Michael. Distributed controller communication in virtual power plants using smart meter gateways. In: 2018 IEEE international conference on engineering, technology and innovation (ICE/ITMC). IEEE; 2018. p. 1–6.
[60] Yang Hongming, Yi Dexin, Zhao Junhua, Dong Zhaoyang. Distributed optimal dispatch of virtual power plant via limited communication. IEEE Trans Power Syst 2013;28(3):3511–2.
[61] Fang Dawei, Guan Xin, Lin Lin, Peng Yu, Sun Di, Hassan Mohammad Mehedi. Edge intelligence based economic dispatch for virtual power plant in 5G internet of energy. Comput Commun 2020;151:42–50.
[62] Chang Jae-Won, Moon Hee Seung, Moon Seung-Il, Yoon Yong Tae, Glick Mark B, Kim Seung Wan. Improved feeder flow control method for a virtual power plant with various resources to reduce communication dependency. IEEE Access 2020;8:206820–34.
[63] Cao Chi, Xie Jun, Yue Dong, Huang Chongxin, Wang Jixiang, Xu Shuyang, et al. Distributed economic dispatch of virtual power plant under a non-ideal communication network. Energies 2017;10(2):235.
[64] Chen Yutian, Chen Junru, Ge Chenchen, Zhong Weilin, Liu Muyang. Scheduled power tracking control of the virtual power plant for its internal contingency considering the communication delay and the unit capacity limitation. Electr Power Syst Res 2023;221:109402.
[65] Oest Frauke, Radtke Malin, Blank-Babazadeh Marita, Holly Stefanie, Lehnhoff Sebastian. Evaluation of communication infrastructures for distributed optimization of virtual power plant schedules. Energies 2021;14(5):1226.
[66] Hu Bo, Wang Huajian, Niu Tao, Shao Changzheng, Li Chunyan. Multi-time-scale coordinated optimal dispatch of virtual power plant under unreliable communication. CSEE J Power Energy Syst 2022.
[67] Yu Songyuan, Fang Fang, Liu Yajuan, Liu Jizhen. Uncertainties of virtual power plant: problems and countermeasures. Appl Energy 2019;239:454–70.
[68] Gougheri Saleh Sadeghi, Jahangir Hamidreza, Golkar Mahsa A, Ahmadian Ali, Golkar Masoud Aliakbar. Optimal participation of a virtual power plant in electricity market considering renewable energy: a deep learning-based approach. Sustain Energy Grids Netw 2021;26:100448.
[69] Aguilar Juan, Bordons Carlos, Arce Alicia. Chance constraints and machine learning integration for uncertainty management in virtual power plants operating in simultaneous energy markets. Int J Electr Power Energy Syst 2021;133:107304.
[70] Shayegan-Rad Ali, Badri Ali, Zangeneh Ali. Day-ahead scheduling of virtual power plant in joint energy and regulation reserve markets under uncertainties. Energy 2017;121:114–25.
[71] Alabi Tobi Michael, Lu Lin, Yang Zaiyue. Data-driven optimal scheduling of multi-energy system virtual power plant (mevpp) incorporating carbon capture system (ccs), electric vehicle flexibility, and clean energy marketer (cem) strategy. Appl Energy 2022;314:118997.
[72] Rezaei Navid, Pezhmani Yasin, Mohammadiani Rojiar Pir. Optimal stochastic self-scheduling of a water-energy virtual power plant considering data clustering and multiple storage systems. J Energy Storage 2023;65:107366.
[73] Tajeddini Mohammad Amin, Rahimi-Kian Ashkan, Soroudi Alireza. Risk averse optimal operation of a virtual power plant using two stage stochastic programming. Energy 2014;73:958–67.
[74] Sheidaei F, Ahmarinejad A. Multi-stage stochastic framework for energy management of virtual power plants considering electric vehicles and demand response programs. Int J Electr Power Energy Syst 2020;120:106047.
[75] Fang Fang, Yu Songyuan, Xin Xiuli. Data-driven-based stochastic robust optimization for a virtual power plant with multiple uncertainties. IEEE Trans Power Syst 2021;37(1):456–66.
[76] Yin Mingjia, Li Kang, Yu James. A data-driven approach for microgrid distributed generation planning under uncertainties. Appl Energy 2022;309:118429.
[77] Rahimiyan Morteza, Baringo Luis. Real-time energy management of a smart virtual power plant. IET Gener Transm Distrib 2019;13(11):2015–23.
[78] Babaei Sadra, Zhao Chaoyue, Fan Lei. A data-driven model of virtual power plants in day-ahead unit commitment. IEEE Trans Power Syst 2019;34(6):5125–35.
[79] MacDougall Pamela, Kosek Anna Magdalena, Bindner Hendrik, Deconinck Geert. Applying machine learning techniques for forecasting flexibility of virtual power plants. In: 2016 IEEE electrical power and energy conference (EPEC). IEEE; 2016. p. 1–6.
[80] Escobar Eros D, Manrique Tatiana, Isaac Idi A. Campus microgrid data-driven model identification and secondary voltage control. Energies 2022;15(21):7846.
[81] She Buxin, Li Fangxing, Cui Hantao, Zhang Jingqiu, Bo Rui. Fusion of microgrid control with model-free reinforcement learning: review and vision. IEEE Trans Smart Grid 2022.

[82] Cimen Halil, Cetinkaya Nurettin, Vasquez Juan C, Guerrero Josep M. A microgrid energy management system based on non-intrusive load monitoring via multitask learning. IEEE Trans Smart Grid 2020;12(2):977–87.
[83] Jasinski Michal, Sikorski Tomasz, Kaczorowska Dominika, Rezmer Jacek, Suresh Vishnu, Leonowicz Zbigniew, et al. A case study on data mining application in a virtual power plant: cluster analysis of power quality measurements. Energies 2021;14(4):974.
[84] Jasinski Michal, Sikorski Tomasz, Kaczorowska Dominika, Rezmer Jacek, Suresh Vishnu, Leonowicz Zbigniew, et al. A case study on a hierarchical clustering application in a virtual power plant: detection of specific working conditions from power quality data. Energies 2021;14(4):907.
[85] Liang Zheming, Alsafasfeh Qais, Jin Tao, Pourbabak Hajir, Su Wencong. Risk-constrained optimal energy management for virtual power plants considering correlated demand response. IEEE Trans Smart Grid 2017;10(2):1577–87.
[86] Mishra Debi Prasad, Samantaray Subhransu Ranjan, Joos Geza. A combined wavelet and data-mining based intelligent protection scheme for microgrid. IEEE Trans Smart Grid 2015;7(5):2295–304.
[87] Baloch Shazia, Muhammad Mannan Saeed. An intelligent data mining-based fault detection and classification strategy for microgrid. IEEE Access 2021;9:22470–9.
[88] Mustafa Zain, Awad Ahmed SA, Azzouz Maher, Azab Ahmed. Fault identification for photovoltaic systems using a multi-output deep learning approach. Expert Syst Appl 2023;211:118551.
[89] Baloch Shazia, Samsani Sunil Srivatsav, Muhammad Mannan Saeed. Fault protection in microgrid using wavelet multiresolution analysis and data mining. IEEE Access 2021;9:86382–91.
[90] Srivastava Adhishree, Parida SK. Data driven approach for fault detection and Gaussian process regression based location prognosis in smart ac microgrid. Electr Power Syst Res 2022;208:107889.
[91] Liu Huichuan, Qiu Jing, Zhao Junhua. A data-driven scheduling model of virtual power plant using Wasserstein distributionally robust optimization. Int J Electr Power Energy Syst 2022;137:107801.
[92] Ruan Guangchun, Zhong Haiwang, Zhang Guanglun, He Yiliu, Wang Xuan, Pu Tianjiao. Review of learning-assisted power system optimization. CSEE J Power Energy Syst 2020;7(2):221–31.
[93] Chen Tao, Cui Qiushi, Gao Ciwei, Hu Qinran, Lai Kexing, Yang Jianlin, et al. Optimal demand response strategy of commercial building-based virtual power plant using reinforcement learning. IET Gener Transm Distrib 2021;15(16):2309–18.
[94] Liang Huishi, Ma Jin. Data-driven resource planning for virtual power plant integrating demand response customer selection and storage. IEEE Trans Ind Inform 2021;18(3):1833–44.
[95] Wei Yifan, Han Tianyi, Wang Shuoqi, Qin Yudi, Lu Languang, Han Xuebing, et al. An efficient data-driven optimal sizing framework for photovoltaics-battery-based electric vehicle charging microgrid. J Energy Storage 2022;55:105670.
[96] Al-Gabalawy Mostafa. Reinforcement learning for the optimization of electric vehicle virtual power plants. Int Trans Electr Energy Syst 2021;31(8):e12951.
[97] Borisoot Kanjanapon, Liemthong Rittichai, Srithapon Chitchai, Chatthaworn Rongrit. Optimal energy management for virtual power plant considering operation and degradation costs of energy storage system and generators. Energies 2023;16(6):2862.
[98] Ciftci Okan, Mehrtash Mahdi, Kargarian Amin. Data-driven nonparametric chance-constrained optimization for microgrid energy management. IEEE Trans Ind Inform 2019;16(4):2447–57.
[99] Zhang Yichen, Chen Chen, Liu Guodong, Hong Tianqi, Qiu Feng. Approximating trajectory constraints with machine learning–microgrid islanding with frequency constraints. IEEE Trans Power Syst 2020;36(2):1239–49.
[100] Cui Xueyuan, Liu Shu, Ruan Guangchun, Wang Yi. Data-driven aggregation of thermal dynamics within building virtual power plants. Appl Energy 2024;353:122126.
[101] Lin Lin, Guan Xin, Peng Yu, Wang Ning, Maharjan Sabita, Ohtsuki Tomoaki. Deep reinforcement learning for economic dispatch of virtual power plant in internet of energy. IEEE Int Things J 2020;7(7):6288–301.
[102] Tan Yongqiang, Shen Yanxia, Yu Xinyan, Lu Xin. Low-carbon economic dispatch of the combined heat and power-virtual power plants: a improved deep reinforcement learning-based approach. IET Renew Power Gener 2023;17(4):982–1007.
[103] Wu Gengqi, Hua Haojun, Niu Dongxiao. Low-carbon economic dispatch optimization of a virtual power plant based on deep reinforcement learning in China's carbon market environment. J Renew Sustain Energy 2022;14(5).
[104] Yi Zhongkai, Xu Yinliang, Wang Xue, Gu Wei, Sun Hongbin, Wu Qiuwei, et al. An improved two-stage deep reinforcement learning approach for regulation service disaggregation in a virtual power plant. IEEE Trans Smart Grid 2022;13(4):2844–58.
[105] Du Yan, Li Fangxing. Intelligent multi-microgrid energy management based on deep neural network and model-free reinforcement learning. IEEE Trans Smart Grid 2019;11(2):1066–76.
[106] Moutis Panayiotis, Hatziargyriou Nikos D. Decision trees-aided active power reduction of a virtual power plant for power system over-frequency mitigation. IEEE Trans Ind Inform 2014;11(1):251–61.
[107] Kazemi Mohammad Verij, Sadati Seyed Jalil, Gholamian Seyed Asghar. Adaptive frequency control of microgrid based on fractional order control and a data-driven control with stability analysis. IEEE Trans Smart Grid 2021;13(1):381–92.
[108] Abdolrasol Maher GM, Hannan Mahammad Abdul, Hussain SM Suhail, Ustun Taha Selim, Sarker Mahidur R, Ker Pin Jern. Energy management scheduling for microgrids in the virtual power plant system using artificial neural networks. Energies 2021;14(20):6507.
[109] Khemmook Panya, Prompinit Krisada, Surinkaew Tossaporn. Control of a microgrid using robust data-driven-based controllers of distributed electric vehicles. Electr Power Syst Res 2022;213:108681.
[110] Saeed Rahmani-Dabbagh, Sheikh-El-Eslami Mohammad Kazem. A profit sharing scheme for distributed energy resources integrated into a virtual power plant. Appl Energy 2016;184:313–28.
[111] Luo Zhe, Hong SeungHo, Ding YueMin. A data mining-driven incentive-based demand response scheme for a virtual power plant. Appl Energy 2019;239:549–59.
[112] Xu Zheng, Guo Ye, Sun Hongbin, Tang Wenjun, Huang Wenqi. Deep reinforcement learning for competitive DER pricing problem of virtual power plants. CSEE J Power Energy Syst 2023.
[113] Wang Yafei, Gao Weijun, Qian Fanyue, Li You. Evaluation of economic benefits of virtual power plant between demand and plant sides based on cooperative game theory. Energy Convers Manag 2021;238:114180.
[114] Chen Wen, Qiu Jing, Zhao Junhua, Chai Qingmian, Dong Zhao Yang. Bargaining game-based profit allocation of virtual power plant in frequency regulation market considering battery cycle life. IEEE Trans Smart Grid 2021;12(4):2913–28.
[115] Chen Wen, Qiu Jing, Zhao Junhua, Chai Qingmian, Dong Zhao Yang. Customized rebate pricing mechanism for virtual power plants using a hierarchical game and reinforcement learning approach. IEEE Trans Smart Grid 2022;14(1):424–39.
[116] Sadeghi Saleh, Jahangir Hamidreza, Vatandoust Behzad, Golkar Masoud Aliakbar, Ahmadian Ali, Elkamel Ali. Optimal bidding strategy of a virtual power plant in day-ahead energy and frequency regulation markets: a deep learning-based approach. Int J Electr Power Energy Syst 2021;127:106646.
[117] Sarathkumar Tirunagaru V. Virtual power plant generating units and energy arbitrage economics in the electricity market using machine learning approach. SSRN; 2022.
[118] Qiu Dawei, Dong Zihang, Ruan Guangchun, Zhong Haiwang, Strbac Goran, Kang Chongqing. Strategic retail pricing and demand bidding of retailers in electricity market: a data-driven chance-constrained programming. Adv Appl Energy 2022;7:100100.
[119] Xinhe Chen, Wei Pei, Wei Deng, Hao Xiao. Data-driven virtual power plant bidding package model and its application to virtual vcg auction-based real-time power market. IET Smart Grid 2020;3(5):614–25.
[120] Dietrich Kristin, Latorre Jesus M, Olmos Luis, Ramos Andres. Modelling and assessing the impacts of self supply and market-revenue driven virtual power plants. Electr Power Syst Res 2015;119:462–70.
[121] Behi Behnaz, Baniasadi Ali, Arefi Ali, Gorjy Arian, Jennings Philip, Pivrikas Almantas. Cost–benefit analysis of a virtual power plant including solar pv, flow battery, heat pump, and demand management: a western Australian case study. Energies 2020;13(10):2614.
[122] Losner Martin, Bottger Diana, Bruckner Thomas. Economic assessment of virtual power plants in the German energy market–a scenario-based and model-supported analysis. Energy Econ 2017;62:125–38.
[123] Hernandez Luis, Baladron Carlos, Aguiar Javier M, Carro Belen, Sanchez-Esguevillas Antonio, Lloret Jaime, Chinarro David, et al. A multi-agent system architecture for smart grid management and forecasting of energy demand in virtual power plants. IEEE Commun Mag 2013;51(1):106–13.
[124] Kaczorowska Dominika, Rezmer Jacek, Jasinski Michal, Sikorski Tomasz, Suresh Vishnu, Leonowicz Zbigniew, et al. A case study on battery energy storage system in a virtual power plant: defining charging and discharging characteristics. Energies 2020;13(24):6670.
[125] Wu Hongbin, Liu Xin, Ye Bin, Xu Bin. Optimal dispatch and bidding strategy of a virtual power plant based on a Stackelberg game. IET Gener Transm Distrib 2020;14(4):552–63.
[126] Yin Shuangrui, Ai Qian, Li Zhaoyu, Zhang Yufan, Lu Tianguang. Energy management for aggregate prosumers in a virtual power plant: a robust Stackelberg game approach. Int J Electr Power Energy Syst 2020;117:105605.
[127] Afrasiabi Mousa, Mohammadi Mohammad, Rastegar Mohammad, Kargarian Amin. Multi-agent microgrid energy management based on deep learning forecaster. Energy 2019;186:115873.
[128] Li Siyun, Wu Wenchuan, Lin Yi. Robust data-driven and fully distributed volt/var control for active distribution networks with multiple virtual power plants. IEEE Trans Smart Grid 2022;13(4):2627–38.
[129] Lin Wen-Ting, Chen Guo, Li Chaojie. Risk-averse energy trading among peer-to-peer based virtual power plants: a stochastic game approach. Int J Electr Power Energy Syst 2021;132:107145.
[130] Ruan Guangchun, Zhong Haiwang, Wang Jianxiao, Xia Qing, Kang Chongqing. Neural-network-based Lagrange multiplier selection for distributed demand response in smart grid. Appl Energy 2020;264:114636.
[131] Wang Yao, Ai Xin, Tan Zhongfu, Yan Lei, Liu Shuting. Interactive dispatch modes and bidding strategy of multiple virtual power plants based on demand response and game theory. IEEE Trans Smart Grid 2015;7(1):510–9.
[132] Wang Jianing, Guo Chunlin, Yu Changshu, Liang Yanchang. Virtual power plant containing electric vehicles scheduling strategies based on deep reinforcement learning. Electr Power Syst Res 2022;205:107714.

[133] Loschenbrand Markus. Modeling competition of virtual power plants via deep learning. Energy 2021;214:118870.
[134] Qin Zhaoming, Liu Di, Hua Haochen, Cao Junwei. Privacy preserving load control of residential microgrid via deep reinforcement learning. IEEE Trans Smart Grid 2021;12(5):4079–89.
[135] Samadi Esmat, Badri Ali, Ebrahimpour Reza. Decentralized multi-agent based energy management of microgrid using reinforcement learning. Int J Electr Power Energy Syst 2020;122:106211.
[136] Jogunola Olamide, Tsado Yakubu, Adebisi Bamidele, Virtelect Mohammad Hammoudeh. A peer-to-peer trading platform for local energy transactions. IEEE Int Things J 2021;9(8):6121–33.
[137] Morstyn Thomas, Farrell Niall, Darby Sarah J, McCulloch Malcolm D. Using peer-to-peer energy-trading platforms to incentivize prosumers to form federated power plants. Nat Energy 2018;3(2):94–101.
[138] Mishra Sambeet, Crasta Cletus John, Bordin Chiara, Mateo-Fornes Jordi. Smart contract formation enabling energy-as-a-service in a virtual power plant. Int J Energy Res 2022;46(3):3272–94.
[139] Park Dae-Hyun, Park Jong-Bae, Lee Kwang Y, Son Sung-Yong, Roh Jae Hyung. A bidding-based peer-to-peer energy transaction model considering the green energy preference in virtual energy community. IEEE Access 2021;9:87410–9.
[140] Zeng Yu, Wei Xuan, Yao Yuan, Xu Yinliang, Sun Hongbin, Chan Wai Kin Victor, et al. Determining the pricing and deployment strategy for virtual power plants of peer-to-peer prosumers: a game-theoretic approach. Appl Energy 2023;345:121349.
[141] Lyu Xiaoyu, Xu Zhiyu, Wang Ning, Fu Min, Xu Weisheng. A two-layer interactive mechanism for peer-to-peer energy trading among virtual power plants. Energies 2019;12(19):3628.
[142] Liu Nian, Guo Bin, Liu Zifa, Wang Yongli. Distributed energy sharing for pvt-hp prosumers in community energy internet: a consensus approach. Energies 2018;11(7):1891.
[143] Wang Zhenyu, Zhang Xiaoyu, Wang Hao. Consensus-based decentralized energy trading for distributed energy resources. Energy Convers Econ 2021;2(4):221–34.
[144] Lin Chengrong, Hu Bo, Shao Changzheng, Niu Tao, Cheng Qian, Li Chunyan, et al. An analysis of delay-constrained consensus-based optimal algorithms in virtual power plants. ISA Trans 2022;125:189–97.
[145] Li Peikai, Liu Yun, Xin Huanhai, Jiang Xichen. A robust distributed economic dispatch strategy of virtual power plant under cyber-attacks. IEEE Trans Ind Inform 2018;14(10):4343–52.
[146] Li Jiangnan, Amin M Asim, Shi Jun, Cheng Lanfen, Lu Feifan, Geng Bo, et al. Energy trading of multiple virtual power plants using deep reinforcement learning. In: 2021 international conference on power system technology (POWERCON). IEEE; 2021. p. 892–7.
[147] Siano Pierluigi, De Marco Giuseppe, Rolan Alejandro, Loia Vincenzo. A survey and evaluation of the potentials of distributed ledger technology for peer-to-peer transactive energy exchanges in local energy markets. IEEE Syst J 2019;13(3):3454–66.
[148] Guo Mingxing, Zhang Ke, Wang Su, Xia Jinlei, Wang Xiaohui, Lan Li, et al. Peer-to-peer energy trading and smart contracting platform of community-based virtual power plant. Front Energy Res 2023;10:1007694.
[149] Chu Tianfeng, An Xingchen, Zhang Wuyang, Lu Yan, Tian Jiaqi. Multiple virtual power plants transaction matching strategy based on alliance blockchain. Sustainability 2023;15(8):6939.
[150] Yang Qing, Wang Hao, Wang Taotao, Zhang Shengli, Wu Xiaoxiao, Wang Hui. Blockchain-based decentralized energy management platform for residential distributed energy resources in a virtual power plant. Appl Energy 2021;294:117026.
[151] Cioara Tudor, Antal Marcel, Mihailescu Vlad T, Antal Claudia D, Anghel Ionut M, Mitrea Dan. Blockchain-based decentralized virtual power plants of small prosumers. IEEE Access 2021;9:29490–504.
[152] Zhang Xiaohong, Song Zilong, Moshayedi Ata Jahangir. Security scheduling and transaction mechanism of virtual power plants based on dual blockchains. J Cloud Comput 2022;11(1):1–26.
[153] Yin Shuangrui, Ai Qian, Li Jiamei, Li Da, Guo Qinglei. Trading mode design for a virtual power plant based on main-side consortium blockchains. Appl Energy 2022;325:119932.
[154] Hassan Ul Muneeb, Rehmani Mubashir Husain, Chen Jinjun. Vpt: privacy preserving energy trading and block mining mechanism for blockchain based virtual power plants. ArXiv preprint arXiv:2102.01480, 2021.
[155] Wang Dewen, Wang Zhao, Lian Xin. Research on distributed energy consensus mechanism based on blockchain in virtual power plant. Sensors 2022;22(5):1783.
[156] Gough Matthew, Santos Sergio F, Almeida Artur, Lotfi Mohamed, Javadi Mohammad S, Fitiwi Desta Z, et al. Blockchain-based transactive energy framework for connected virtual power plants. IEEE Trans Ind Appl 2021;58(1):986–95.
[157] Seven Serkan, Yoldas Yeliz, Soran Ahmet, Alkan Gulay Yalcin, Jung Jaesung, Ustun Taha Selim, et al. Energy trading on a peer-to-peer basis between virtual power plants using decentralized finance instruments. Sustainability 2022;14(20):13286.
[158] Yadav Radhika, Kumar Balla Manoj, Baid Saurav, Padma Priya R. Energy trading in virtual power plant enabled communities using double auction technique and blockchain technology. Intell Soft Comput Syst Green Energy 2023:111–25.
[159] Song YJ, Lee Jae-Kyu. A blockchain and internet of things based architecture design for energy transaction. J Syst Manag Sci 2020;10:122–40.
[160] Li Hongyi, Hui Hongxun, Zhang Hongcai. Blockchain-assisted virtual power plant framework for providing operating reserve with various distributed energy resources. iEnergy 2023;2(2):1–10.
[161] Nafkha-Tayari Wafa, Ben Elghali Seifeddine, Heydarian-Forushani Ehsan, Benbouzid Mohamed. Virtual power plants optimization issue: a comprehensive review on methods, solutions, and prospects. Energies 2022;15(10):3607.
[162] Zhang Jiatong. The concept, project and current status of virtual power plant: a review. J Phys Conf Ser 2022;2152:012059.
[163] Kieny Christophe, Berseneff Boris, Hadjsaid Nouredine, Besanger Yvon, Maire Joseph. On the concept and the interest of virtual power plant: some results from the European project fenix. In: 2009 IEEE power & energy society general meeting. IEEE; 2009. p. 1–6.
[164] Binding Carl, Gantenbein Dieter, Jansen Bernhard, Sundstrom Olle, Andersen Peter Bach, Marra Francesco, et al. Electric vehicle fleet integration in the Danish edison project-a virtual power plant on the island of Bornholm. In: IEEE PES general meeting. IEEE; 2010. p. 1–8.
[165] Twenties project team. Twenties project final report. Available: https://www.ewea.org/fileadmin/files/library/publications/reports/Twenties.pdf, 2013 [Online].
[166] Web2Energy project team. Web2Energy - the intelligent energy network of the future. Available: https://www.entega.ag/web2energy, 2023 [Online].
[167] Vuthi Petrit Patrick, Lorenzen Peter, Schaefers Hans, Raths Stephan, Krengel Stefan, Sudeikat Jan, et al. Smart power Hamburg: a virtual power plant for Hamburg. In: International ETG congress 2015; Die Energiewende-Blueprints for the new energy age. VDE; 2015. p. 1–8.
[168] City-zen project team. City-zen: virtual power plant. Available: https://amsterdamsmartcity.com/updates/project/city-zen-virtual-power-plant, 2016 [Online].
[169] Gabderakhmanova Tatiana, Engelhardt Jan, Zepter Jan Martin, Sorensen Thomas Meier, Boesgaard Knud, Ipsen Hans Henrik, et al. Demonstrations of dc microgrid and virtual power plant technologies on the Danish island of Bornholm. In: 2020 55th international universities power engineering conference (UPEC). IEEE; 2020. p. 1–6.
[170] VPP4Islands team. About our VPP4Islands project. Available: https://vpp4islands.eu/, 2021 [Online].
[171] POSITYF team. POSITYF European project. Available: https://posytyf-h2020.eu/project-overview/posityf-european-project, 2020 [Online].
[172] Marinescu Bogdan, Gomis-Bellmunt Oriol, Dorfler Florian, Schulte Horst, Sigrist Lukas. Dynamic virtual power plant: a new concept for grid integration of renewable energy sources. IEEE Access 2022;10:104980–95.
[173] Tranninger Judith. Statkraft's virtual power plant participates in UK's balancing mechanism. Available: https://www.statkraft.com/newsroom/news-and-stories/2020/vpp-balancing-services/, 2020 [Online].
[174] Sunrun team. Sunrun activates nation's first residential virtual power plant in wholesale market. Available: https://investors.sunrun.com/news-events/press-releases/detail/273/sunrun-activates-nations-first-residential-virtual-power, 2022 [Online].
[175] OhmConnect team. Resi-station: California's largest virtual power plant. Available: https://www.ohmconnect.com/about-us/news/resi-station, 2023 [Online].
[176] Sempra Energy utility. Demand response emerging technologies program semi-annual report 2022. Available: https://www.dret-ca.com/wp-content/uploads/2022/04/SDGE-Semi-Annual-EMT-DR-Report-2021-Q4-2022-Q1.pdf, 2022 [Online].
[177] Natural Resources Canada. NRCan smart grid project overview. Available: https://natural-resources.canada.ca/sites/nrcan/files/environment/Smart%20Grig_E_2021_accessible.pdf, 2021 [Online].
[178] Australian Renewable Energy Agency. AGL virtual power plant. Available: https://arena.gov.au/projects/agl-virtual-power-plant/, 2022 [Online].
[179] Tesla team. What is South Australia's virtual power plant. Available: https://www.tesla.com/en_au/sa-virtual-power-plant, 2023 [Online].
[180] Australian Energy Market Operator. Virtual power plant demonstrations. Available: https://aemo.com.au/en/initiatives/major-programs/nem-distributed-energy-resources-der-program/der-demonstrations/virtual-power-plant-vpp-demonstrations, 2021 [Online].
[181] Wang Xuanyuan, Liu Dunnan, Liu Zhen, Liu Mingguang, Wang Jianni, Gao Yuan, et al. Operation mechanism and key technologies of virtual power plant under ubiquitous internet of things. Power Syst Technol 2019;43(9):3175–83 (in Chinese).
[182] Song Xinming, Deng Yongsheng, Jiao Fengshun, Shi Jing, Cheng Meng, Xiang Qianliang, et al. Virtual power plant implementation scheme in Shenzhen city. Environ Prog Sustain Energy 2021;40(5):e13598.
[183] Fox Eva. Tesla builds the Miyakojima VPP in Japan, becoming one of largest in the country. Available: https://www.tesmanian.com/blogs/tesmanian-blog/tesla-built-the-miyakojima-vpp-in-japan-which-has-become-one-of-the-largest-in-the-country, 2022 [Online].
[184] Mao Tian, Guo Xiaoxuan, Xie Pingping, Zhou Jin, Zhou Baorong, Han Shuai, et al. Virtual power plant platforms and their applications in practice: a brief review. In: 2020 IEEE sustainable power and energy conference (iSPEC). IEEE; 2020. p. 2071–6.
[185] ABB Company. ABB launches OPTIMAX solutions for power generation, industrial and commercial companies. Available: https://pes.eu.com/misc/abb-launches-

optimax-solutions-for-power-generation-industrial-and-commercial-companies/, 2023 [Online].
[186] Next Kraftwerke Company. NEMOCS: our platform for your virtual power plant. Available: https://img1.wsimg.com/blobby/go/09912b22-3c31-42d5-b85f-f9992d33f5ab/downloads/nemocs-brochure-next-kraftwerke.pdf, 2023 [Online].
[187] AutoGrid Company. AutoGrid Flex: intelligent energy management. Available: https://www.auto-grid.com/solution/derms/, 2023 [Online].
[188] GE Company. Advanced EMS - platform. Available: https://www.ge.com/digital/sites/default/files/download_assets/advanced-ems-platform-from-ge-digital-brochure_0.pdf, 2023 [Online].
[189] Giz GmbH. Opportunities for virtual power plant in India. Available: https://www.energyforum.in/fileadmin/user_upload/india/media_elements/publications/20191121_Virtual_Power_Plants/Virtual_Power_Plants_Report.pdf, 2019 [Online].
[190] Haradat Keiya, Ohnot Yuta, Nakamurat Yuichi, Nishit Hiroaki. Anonymization method based on sparse coding for power usage data. In: 2018 IEEE 16th international conference on industrial informatics (INDIN). IEEE; 2018. p. 571–6.
[191] Wilkinson Mark D, Dumontier Michel, Aalbersberg Jan, Appleton Gabrielle, Axton Myles, Baak Arie, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016;3(1):1–9.
[192] Feng Cheng, Chen Qixin, Wang Yi, Ma Jiaqi, Wang Xuanyuan. Frequency regulation service provision for virtual power plants through 5g ran slicing. IEEE Trans Smart Grid 2022;13(6):4943–56.
[193] Zhang Yifeng, Chen Zeqi, Ma Kaixuan, Chen Feifan. A decentralized IoT architecture of distributed energy resources in virtual power plant. IEEE Int Things J 2022;10(10):9193–205.
[194] Pal Poushali, Parvathy AK, Devabalaji KR, Antony S Joseph, Ocheme Simon, Sudhakar Thanikanti Babu, et al. IoT-based real time energy management of virtual power plant using PLC for transactive energy framework. IEEE Access 2021;9:97643–60.
[195] Wang Qianchao, Pan Lei, Wang Haitao, Wang Xinchao. Artificial intelligence based interpretable uncertainty forecasting framework in the configuration of a virtual power plant towards carbon. Available at SSRN 4164367, 2022.
[196] Khan Asad Ali, Beg Omar A, Jin Yu-Fang, Ahmed Sara. An explainable intelligent framework for anomaly mitigation in cyber-physical inverter-based systems. IEEE Access 2023.
[197] Yi Zhongkai, Xu Ying, Wu Chenyu. Model-free economic dispatch for virtual power plants: an adversarial safe reinforcement learning approach. IEEE Trans Power Syst 2023.
[198] Stanojev Ognjen, Mitridati Lesia, di Prata Riccardo de Nardis, Hug Gabriela. Safe reinforcement learning for strategic bidding of virtual power plants in day-ahead markets. ArXiv preprint arXiv:2307.05812, 2023.
[199] Feng Bin, Liu Zhuping, Huang Gang, Guo Chuangxin. Robust federated deep reinforcement learning for optimal control in multiple virtual power plants with electric vehicles. Appl Energy 2023;349:121615.
[200] Pop Claudia, Antal Marcel, Cioara Tudor, Anghel Ionut, Salomie Ioan, Bertoncini Massimo. A fog computing enabled virtual power plant model for delivery of frequency restoration reserve services. Sensors 2019;19(21):4688.
[201] Ruan Guangchun, Kirschen Daniel S, Zhong Haiwang, Xia Qing, Kang Chongqing. Estimating demand flexibility using Siamese LSTM neural networks. IEEE Trans Power Syst 2021;37(3):2360–70.
[202] Nakabi Taha Abdelhalim, Toivanen Pekka. Deep reinforcement learning for energy management in a microgrid with flexible demand. Sustain Energy Grids Netw 2021;25:100413.
[203] Arwa Erick O, Folly Komla A. Reinforcement learning techniques for optimal power control in grid-connected microgrids: a comprehensive review. IEEE Access 2020;8:208992–9007.
[204] Sikorski Tomasz, Jasinski Michal, Ropuszynska-Surma Edyta, Weglarz Magdalena, Kaczorowska Dominika, Kostyla Pawel, et al. A case study on distributed energy resources and energy-storage systems in a virtual power plant concept: economic aspects. Energies 2019;12(23):4447.
[205] Arslan Okan, Karasan Oya Ekin. Cost and emission impacts of virtual power plant formation in plug-in hybrid electric vehicle penetrated networks. Energy 2013;60:116–24.
[206] Darvishi Mostafa, Tahmasebi Mehrdad, Shokouhmand Ehsan, Pasupuleti Jagadeesh, Bokoro Pitshou, Raafat Jwan Satei. Optimal operation of sustainable virtual power plant considering the amount of emission in the presence of renewable energy sources and demand response. Sustainability 2023;15(14):11012.
[207] Liu Zuoyu, Zheng Weimin, Qi Feng, Wang Lei, Zou Bo, Wen Fushuan, et al. Optimal dispatch of a virtual power plant considering demand response and carbon trading. Energies 2018;11(6):1488.