Ahsan *et al.*
Protection and Control of Modern Power Systems (2023) 8:43
https://doi.org/10.1186/s41601-023-00319-5

Protection and Control of Modern Power Systems

REVIEW **Open Access**

Data-driven next-generation smart grid towards sustainable energy evolution: techniques and technology review

Faiaz Ahsan[1], Nazia Hasan Dana[1], Subrata K. Sarker[1*], Li Li[2], S. M. Muyeen[3], Md. Firoj Ali[1], Zinat Tasneem[1], Md. Mehedi Hasan[1], Sarafat Hussain Abhi[1], Md. Robiul Islam[1], Md. Hafiz Ahamed[1], Md. Manirul Islam[1], Sajal K. Das[1], Md. Faisal R. Badal[1] and Prangon Das[1]

Abstract

Meteorological changes urge engineering communities to look for sustainable and clean energy technologies to keep the environment safe by reducing CO_2 emissions. The structure of these technologies relies on the deep integration of advanced data-driven techniques which can ensure efficient energy generation, transmission, and distribution. After conducting thorough research for more than a decade, the concept of the smart grid (SG) has emerged, and its practice around the world paves the ways for efficient use of reliable energy technology. However, many developing features evoke keen interest and their improvements can be regarded as the next-generation smart grid (NGSG). Also, to deal with the non-linearity and uncertainty, the emergence of data-driven NGSG technology can become a great initiative to reduce the diverse impact of non-linearity. This paper exhibits the conceptual framework of NGSG by enabling some intelligent technical features to ensure its reliable operation, including intelligent control, agent-based energy conversion, edge computing for energy management, internet of things (IoT) enabled inverter, agent-oriented demand side management, etc. Also, a study on the development of data-driven NGSG is discussed to facilitate the use of emerging data-driven techniques (DDTs) for the sustainable operation of the SG. The prospects of DDTs in the NGSG and their adaptation challenges in real-time are also explored in this paper from various points of view including engineering, technology, et al. Finally, the trends of DDTs towards securing sustainable and clean energy evolution from the NGSG technology in order to keep the environment safe is also studied, while some major future issues are highlighted. This paper can offer extended support for engineers and researchers in the context of data-driven technology and the SG.

Keywords Data-driven technology, Smart grid, Sustainable energy evolution, Next-generation smart grid, Intelligent management, And Machine learning technique

1 Introduction

Data-driven technologies have become a widely used set of techniques in the field of scientific research and engineering where data are being used for understanding, maintaining, and turning typical systems into smart sustainable systems. The use of data-driven techniques (DDTs) is gaining popularity in various engineering sectors because of their appearance in decision making, transparency, reliability, and sustainability. For example, data-driven machine learning (ML) techniques are used

*Correspondence:
Subrata K. Sarker
subrata@mte.ruet.ac.bd
[1] Department of Mechatronics Engineering, Rajshahi University of Engineering and Technology, Rajshahi, Bangladesh
[2] School of Electrical and Data Engineering, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, Australia
[3] Electrical Engineering Department, Qatar University, Doha, Qatar

for analysis, prediction, control and diagnosis in medical research [1], precise agriculture [2], quantum finance [3], risk management in the supply chain [4], etc. These techniques may be supervised, semi-supervised or unsupervised depending on the availability and condition of collected data, and can obtain a higher rate of success than typical methods used in the various fields of science and business. Because of the increasing trends of data-driven methodologies, researchers have started contemplating the presence of DDTs in conventional power systems. This has enabled the construction of a next-generation smart grid (NGSG) from the typical smart grid (SG). It also accelerates the traditional SG to unlock the full potential of future SGs with zero carbon emission and lifelong sustainability.

The conventional SG is an improved version of the traditional power grid and microgrid, where advanced technologies are used to enable communication, simulation, sensing, decision-making, etc. A comparative study of the microgrid and different versions of SG in terms of technical features associated with them is of great importance. An SG allows the components of the grid, e.g., smart meters, renewable energy sources (RESs), advanced communication systems, closed-loop feedback systems, distributed generation, storage, etc., to communicate with each other. The grid ensures the production of sufficient high-quality power while integrating other benefits such as self-healing capabilities, fault assessment, consumer friendliness, cyber and physical security [5]. Because of the extended features as compared to the microgrid, some countries have successfully enabled the SG with good annual growth rate, as shown in Fig. 1 [6]. This is feasible because the existing SGs all around the globe are still operated based on conventional power systems to produce power from the kilowatt to gigawatt scale.

The conventional SG cannot fully meet the requirements as it continuously changes with emerging advanced technologies. The need for clean energy has increased globally over the past decade as a result of changing environmental conditions and expanding populations and technology that may impose non-linear dynamics on the SG. The non-linearity in the smart power grid transmission and distribution systems may add new congestion, outages, fluctuation in voltage and frequency, that lead to blackouts as a result of the increasing demand for electricity [7]. Non-renewable energy sources though being an easier, quicker, and cheaper path to generate power, they are a direct obstruction to the green environment because of high emissions [8]. Renewable energy sources are on the rise to reduce dependency on fossil fuel-based

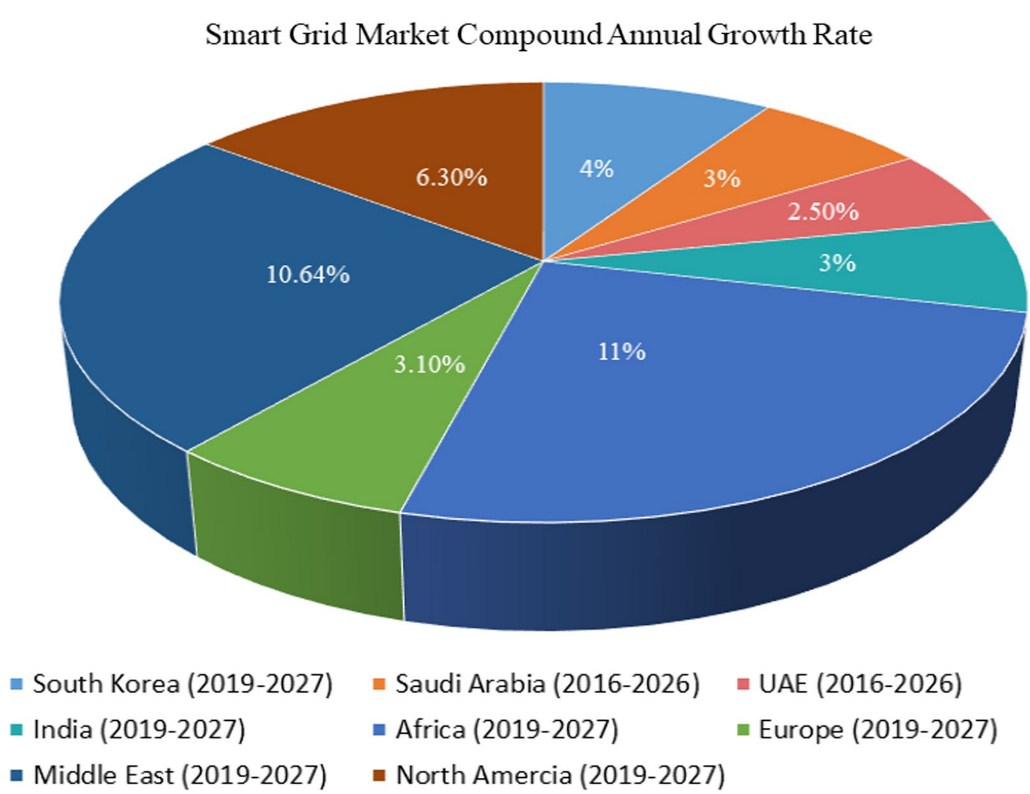


Fig. 1 Countries contributing to SG world market [6]

Table 1 Differences between conventional SG and NGSG

Difference based on	Conventional SG	NGSG
Setup assembly	Digital electronics, microprocessors	Edge computing enabled IoT devices and components
Power generation	Distributed power generation plants	Multiple distributed renewable energy sources and each assigned with agent-oriented software
Communication system	Involving microprocessor-based digital technology which allows data communication among the devices of the system and makes the remote control possible	Energy devices can communicate with the power grid and exchange information by giving access to the dynamic communication system
Security system	SGs provide automated protection	The security system will be supported with next-generation blockchain technology
Control system	Fast conventional control measurements are provided	Each parameter such as frequency, voltage, restoration etc., is controlled by different DDTs so that optimal control can be ensured. Hence, internet-based inverter control technology may be used there
Monitoring system	SG involves sensor-based self-monitoring of power distribution	Monitoring in the NGSG is divided to every stage, components and systems within the NGSG with the goal of preventing any type of power disruption by analyzing monitored data and predicting outcomes
Failure and restoration	SG has minimal self-healing properties. At the time of SG failure, power can be rerouted through alternate paths	The advanced self-healing capabilities of NGSG include the use of data from every stage of the power grid for detecting failures and troubleshooting to quickly fix them
Environmental effects	SG involves renewable energy integration which reduces the impacts on the environment, such as emission of CO_2 and global warming	DDTs in the NGSG enable a way to reach sustainable energy evolution with lower carbon emission, less waste and reduced global warming by collecting and analyzing grid data

power generation [9]. However, the uncertainty and complexity of SGs are increasing with the addition of more distributed generation (DG), increased market size, and renewable sources [10].

Again, the existing SGs are not yet sustainable in long-term power generation and distribution, because of the lack of absolute compatibility between grid components [11], programmable sensors deployment [12], fast real-time monitoring, analysis and decision making with minimized latency [13], and integrating maximum intermittent generation [14]. To make a sustainable SG operation, researchers are interested in formulating the next-generation smart grid (NGSG). An NGSG will have the ability to address the above shortcomings through the integration of advanced DDTs, blockchain technology, and other edge computing techniques based on collecting and analyzing conventional SG data. As the datasets are getting massive because of increasing complexity in the SG systems, a better storage system with secured high-speed data transfer system may also need to be integrated in an NGSG, where the data storage should be encrypted with blockchain technology and managed with advanced data management algorithms.

Further, the preservation of data privacy and data security also needs advancement in the conventional SG domain, where the security of massive amounts of datasets in the NGSG domain will be handled with next-generation blockchain technology and data techniques [15, 16]. Additionally, an NGSG may also have several extended features including interoperability, less transmission loss, decreased latency, large sources handling capability, grid mobility, ease of renovation, and advanced resilience, all features that are quite dependent on the adaptation of data-driven technology.

Thus, it can be concluded that an NGSG is the improved version of the existing SG which enables some extended features to work on minimizing the shortcomings of the conventional SG. It can be an automated grid driven by data where the control operation, energy management, condition monitoring, forecasting, fraud characterization, energy transaction and its security may be done in an improved manner on the basis of collecting and analyzing data, and implementing advanced data-driven techniques. A comparative study between the conventional SG and an NGSG is reported in Table 1 in terms of operation and technologies. From Table 1, it can be seen that the use of highly computationally efficient DDTs, edge computing devices, next-generation blockchain technology, advanced interoperability, and agent-oriented techniques in the NGSG framework makes explicit differences between the conventional SG and the NGSG. The purpose of considering these technologies is to ensure sustainable energy evolution in the NGSG. Thus, it can be stated that the framework of NGSG focuses on sustainable energy technologies.

An NGSG may be largely dependent on the use of DDTs to achieve sustainable energy evolution worldwide. Sustainable evolution refers to the integration of DDTs in data analysis from datasets of multiple decentralized RESs and energy storage systems (ESSs), enabling internet of things (IoT) devices, load forecasting, energy trading, security systems, grid faults, and losses. The ongoing research in the SG domain states that DDTs have been successfully implemented in characterizing grid faults and energy trading. However, it may impose new challenges in terms of security constraints as the energy demand increases, as well as gradually increased cyber security threats around the world. The solution to these challenges requires a revision in the SG structure based on enabling data-driven modeling and planning. The primary benefit of the data-driven NGSG is the availability of faster and more reliable operation and more accurate data that authorize the use of advanced DDTs towards enabling efficient and sustained electricity flow from generation to distribution. Additionally, increased management and monitoring capabilities across the entire power system, as well as more affordable, adaptable, and effective operation, are presented by revolutionary developments in data-driven analysis models and algorithms, mostly inspired by advanced data science.

From the critical surveys addressed in Table 2, it can be seen that there exists much scope for and many applications of DDTs in the SG domain. The purpose of DDTs is to enable advanced features towards securing the sustainable operation for energy evolution from the NGSG, as the absence of these features may hinder the scalability, availability, security, and other issues in the SG. Many of them show the additional challenges that may arise while implementing DDTs in an NGSG. At present, there are many loop-holes in SG systems and it is necessary to study these drawbacks to remove them by improving the present SG technology. The main contributions of this study are:

- *Studying conventional SG features* A study of the technical features of a conventional SG is done to explore improvement potential. Also, some current SG projects around the world with their capacities are studied.
- *Developing a technical framework for a data-driven NGSG* First, a technical framework for an NGSG is developed by integrating new advanced technical features into the SG domain. A study is then pre-

Table 2 Comparison between the current study and related existing literature

References	Study the technical architecture and feature of SG	Investigate the developing aspect of SG through integrating advanced technical feature	Discuss the conceptual framework of data-driven NGSG	Study the prospects of DDTs in NGSG with their adaption challenges	Discuss the trends of DDTs in NGSG towards sustainable energy evolution
[17]	Yes	Yes	No	No	No
[18]	Yes	Yes	No	No	No
[19]	Yes	Yes	No	No	No
[20]	Yes	Yes	No	No	No
[7]	No	Yes	No	No	No
[21]	Yes	Yes	No	No	No
[22]	Yes	Yes	No	No	No
Current study	Yes	Yes	Yes	Yes	Yes

sented on the development of a data-driven NGSG along with the necessary analytics required to be performed before the implementation of DDTs.

- *Investigating the scope of data-driven techniques in the NGSG* This study also explores the possible prospects of DDTs in an NGSG and discusses the adaptation challenges of data-driven NGSGs in reality.
- *Exploring the role of DDTs in sustainable energy evolution* A brief discussion about the trends of DDTs towards obtaining sustainable energy evolution from an NGSG is also incorporated in this study to highlight the significance of data-driven SG modeling.

2 Smart grid at present: technical architecture

An SG enables bidirectional flow of electricity between the utility and its end users, with its smart framework structured by combining information, power technologies, and telecommunication with the prevailing electricity system. This energy technology also supports automation mechanization for efficient power distribution, storage elements, fault detection, electric vehicles, grid data supervision, combination of hybrid RESs, and flexibility of grid networks [23]. The various components shown in Fig. 2 can be used to build the SG energy technology. They include renewable sources, a smart

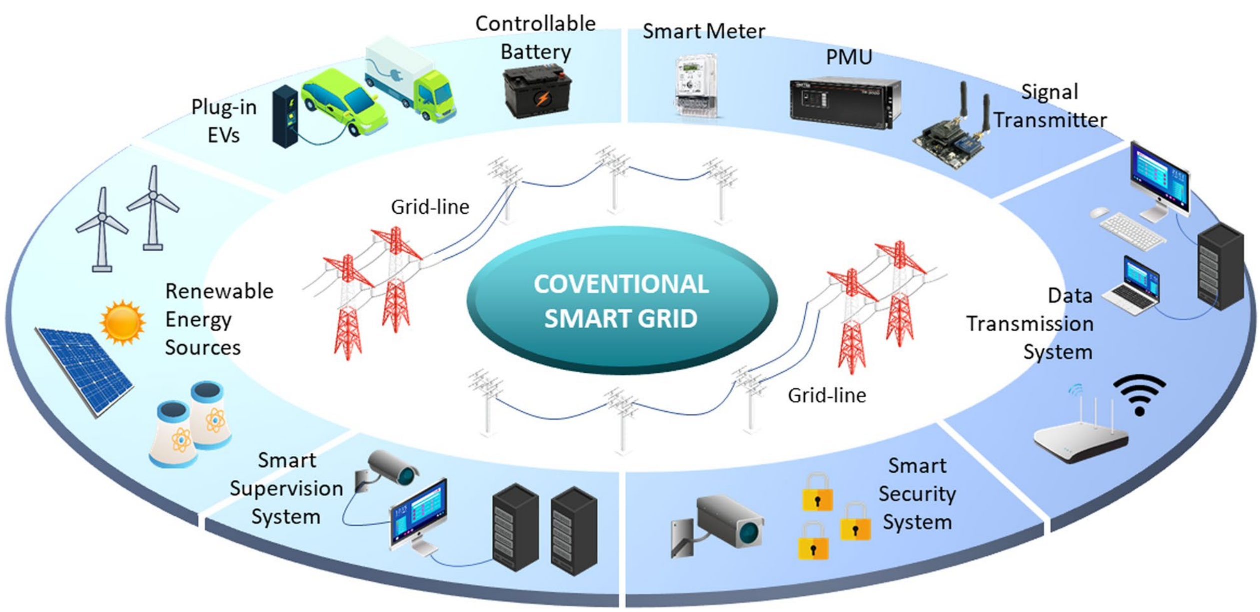


Fig. 2 Conventional smart grid architecture

supervision system, a smart information system, an advanced storage system, a smart security system, sensors, and grid-lines.

2.1 Smart distributed generation sources

An SG uses a "smart distributed generation" unit which refers to the process of producing electricity efficiently in small-scale implementations close to the place of consumer usage. The primary technologies for SG application are RESs in addition to ESSs. It offers excellent prospects for controlling frequency and voltage deviations, responding to emergency situations when the load exceeds the generation, and decarbonizing targeted areas. Plug-in hybrid electric vehicles (PHEVs) have the potential to reduce emissions while also lowering transportation costs [24]. The potential of PHEVs to integrate onboard energy storage devices with the power grid can increase grid efficiency and dependability. The power grid can also increase its acceptance of intermittent renewable energy generation with the sole use of energy storage devices like battery ESSs. To achieve this, effective coordination among ESSs, the grid, and renewable generation units is needed [25].

A crucial prototype for power generation is the DG units that have improved reliability and power quality, and can lower system capacity margin. Executing DGs in practice may be difficult for several reasons including: (1) large fluctuation in terms of availability of RESs; (2) very different generation and demand patterns; and (3) higher execution costs of DGs than the conventional power plants [26]. The development of DG units has also introduced the idea of a virtual power plant (VPP) that collects capacities of diverse DERs to increase electricity generation. In a VPP, a controller controls a large group of DGs, and thus, VPPs provide more efficiency and flexibility, and can handle fluctuations better than conventional power plants. However, VPPs require complex optimization, secure communication and intelligent control [27].

An SG consists of many DG units, and therefore electricity generation flexibility increases while the flow control becomes complex There are two domestic electricity distribution systems, i.e., (1) AC (Alternating Current) power distribution; and (2) DC (Direct Current) power dispatch [28]. The DC power distribution is more practicable because it makes domestic power distribution well organized and easier to control. Several technologies including microgrid and vehicles to grid (V2G), have emerged to distribute DC power. The microgrid can generate electricity of low voltage, even if it is islanded from the main grid. In the islanded mode, the users do not get electricity from any external sources. Microgrid disentangles execution of SG functions, e.g., better dependability, significant renewable energy penetration, self-healing, and effective load control systems [29]. V2G usually enables getting power from stored electricity like vehicles running in battery packs. It enables a novel method of storing and delivering electrical energy and enhances power quality by providing electrical energy stored in PHEV batteries to the grid during peak hours.

2.2 Smart metering, measurement and monitoring

Any information technology that is concerned with distributed automation, such as data exchange compatibility and combination with current and future devices or systems should be addressed in SG technology. As a result, in the framework of an SG, a smart information subsystem is employed to enable information production, simulation, analysis, integration, and optimization.

Smart metering technology is the most vital means of obtaining information from consumers. The advanced metering infrastructure (AMI) uses automatic meter reading (AMR) technology to logically fit with an SG. The AMR system works on automatic data gathering, diagnostics, as well as collecting data from smart metering devices and sending data to the main database for accounting, troubleshooting, and analysis. In contrast to a typical AMR, AMI allows bidirectional communication with the meters [30]. The advantage of advanced smart metering is that the consumers can predict their approximate bills and manage the power usage to lower bills. It is also beneficial for utilities because smart metering enables real-time pricing [31].

Again, measuring and monitoring the system's current status at various places are essential for the smooth operation of the SG system. The topology of phasor measurement units (PMUs) and sensors is important for advanced monitoring. The status of an electrical grid is measured by PMUs to be used to analyze system health. A high number of PMUs as well as the capability to compare the measurements taken from the grid can enable use of the collected data to track the state of the power system and rapidly respond to system circumstances. The existing frequency monitoring network system architecture is designed to handle large amounts of data flows, processing, storage, and usage [32]. Accordingly, the sensor networks provide practicable and low cost sensing as well as communication media for distant monitoring and identification of the system.

2.3 Smart management of information

An SG can manage big datasets efficiently by extracting the most efficient information and rejecting the false data. Data management is a process of examining, evaluating, integrating, and optimizing data obtained from a large network of data-gathering devices. The goal of data modeling is to make information interchangeable among multiple devices that are standard for diverse working environments and conditions. Data modeling is required for device forward and backward compatibilities, which means that the device is compatible with its previous and future versions. The goal of information integration is to combine data from several sources with distinct theoretical, contextual, and graphical representations. Information optimization is a technique for increasing the effectiveness of information. Singular value decomposition analysis is used to investigate the coupling architecture of an energy grid in order to uncover chances for lowering network traffic by determining which data must be exchanged between portions of the infrastructure to implement a control action [33].

2.4 Smart data transmission

Modern technology has enabled the availability of various commutation modules suited for SG systems. It is complicated to choose a suitable model as SGs tend to have different preferences for data transmission. However, the data transmission system of an SG must be of the utmost high quality to support quality of service. The data transmitted should be accurate, secure, complete, and private. Wireless data transmission uses radio waves to transmit signals and data. Wireless data transmission holds several advantages compared to wired data transmission, including remote access, low maintenance and installation cost, high-speed data transfer, etc. The wireless data transmission category is subdivided into four subcategories described in the following sub-sections [34].

1. *Wireless mesh system* A wireless mesh system follows the method of mesh topology. A mesh ensures that all the data transmission modules are interconnected. The modules form nodes and gateways. In a particular area, a wireless mesh system will provide a very cost-effective communication system that needs little to no mobility. This data transmission system is highly reliable for communication. The mesh network provides a large coverage area and a high data transfer rate [34].
2. *Cellular data transmission technology* A cellular network is a network for communication that is distributed over a large area. This wireless network system uses GSM, 3G, and 4G technology for transferring data at a very low cost. Time delay can be efficiently reduced using cellular data transmission technology. The latest 5G technology can be used for even faster data transfer [35].
3. *Satellite data transmission* Satellite data transmission is highly suitable for covering a very large area. A satellite is connected to a ground station through radio signals, and the satellite communicates in a straight line to the ground station. A satellite stays as a repeater above the earth as it orbits the earth in a geosynchronous position. Satellite technology can be used for rural and remote infrastructures. The downside of using satellites is that the performance of data transmission may deteriorate depending on weather conditions [36].
4. *Direct data transmission* Direct data transmission or point-to-point data transmission refers to sending data from one specific point to another directly. It is usually done using microwave signals. It is a cheap and conventional method for data transmission that has been used for the last two decades. A similar technology, free-space optical (FSO) data transmission, is done by propagating light through free space. This wireless data transmission method is suitable for remote places where complications arise from other data transmission methods. FSO data transmission is highly feasible in urban destinations whereas microwave communications face blockades in particular places [34].

2.5 Smart supervision/regulation technology

Supervision of an SG is essential for high-performance output and efficient management of all the subsystems. The flow of energy and information, being bidirectional, needs to be handled by ensuring the completion of various supervision objectives. An SG is easier to manage than typical power grids. This is mainly because SG enables bidirectional flows of electricity and information. The active participation of electricity customers is also the main feature of an SG. Supervision of an SG can be done based on electricity demand, rather than supply. The objectives of SG supervision and management may include ensuring maximum efficiency, enhancing power production, easy monitoring, and analysis, control of emissions, waste management, gaining maximum profit, etc. Reference [37] proposes an optimized control technique from analyzing the profiles of a large group of customers to shave off energy consumption, while [38] presents a pricing method to incentivize customers.

To properly manage the supervision objectives of an SG, different methods have been adopted ranging from game theory to machine learning. Optimization of an SG using both convex and dynamic programming is proposed in [39], while another technique for optimization, swarm intelligence, shows promising performance in the field of energy distribution resources optimization, which has no dimensional limitation. Data gathered from an SG using sensors and PMUs may also be used to predict the behavior of the SG system through properly developed machine learning algorithms.

2.6 Smart security system

Avoiding cyber security breaches is imperative to ensure the security of an SG. A smart security system protects the information of an SG and increases the integrity of privacy. The security of an SG can be maintained in three categories: solidity, failure detection and protection. The solidity of a system promises that the system can perform consistent behavior in various situations and changed working conditions. The integration of locally generated power can ensure fewer future failures, including both electrical and mechanical failures [40]. Again, an SG has integrated failure detection methods that can detect failures when they occur. This also helps to diagnose and recover from failures in an SG. The failures can be branched out to various faults which occur in an SG, and their protection using digital methods. All the above-mentioned features pave the way to feasibly adapt an SG in the real-time environment. Some countries have already implemented SG projects, as shown in Table 3.

3 Add-on technology towards the next-generation smart grid

NGSGs have the possibility of enabling enhanced features in the SG landscape as compared to conventional SG technologies. The security and privacy issues of the current SG systems may be better covered by an NGSG in the context of integrating more advanced features. The advance of an NGSG entirely depends on the use of data-driven techniques in its different parts. A conceptual framework of an NGSG is illustrated in Fig. 3. From Fig. 3, it can be seen that the framework of an NGSG may consist of integrating edge computing devices, IoT enabled inverters, blockchain-based energy trading, and computationally efficient DDTs in monitoring, controlling, and forecasting. It can also be noted that a data center may appear in an NGSG to collect data from the interconnected technologies and share the data among them to ensure its interoperability. By applying DDTs, the collected data from the different sources can be analyzed intelligently to help make decisions towards sustainable energy evolution. The detailed explanation of the intelligent technologies used in an NGSG framework can be found in the following sub-sections.

3.1 Intelligent agent-based modeling of energy sources

To digitize the energy generation process in an NGSG, there has been a significant rise in the use of

Table 3 Current smart grid projects in different countries

Project name	Location	Activation date	Capacity	Total cost
Narara Ecovillage Smart Grid [41]	New South Wales, Australia	September 2016	0.471 MW	USD 4.73 M
Berrimal Wind [42]	Western Victoria, Australia	2034	72 MW	USD 135.72 M
Mortlake South Wind Farm [43]	Victoria, Australia	2022 (Anticipated)	157.5 MW	N/A
Aldoga Solar Farm [44]	Central Queensland, Australia	2024 (Anticipated)	600 MW	USD 0.5B
Lilyvale Solar PV Plant [45]	Central Queensland, Australia	2019	126.2 MW	USD 283.57 M
Harapaki Wind Farm [46]	New Zealand	2024 (Anticipated)	176 MW	USD 395 M
AEP Ohio [47]	Ohio, USA	2009	2954.03 MW	USD 133.77 M
Detriot Edison's [48]	Michigan, USA	N/A	11,084 MW	USD 10.88 M
Pacific Northwest Smart Grid Project [49]	Washington, USA	2009	47 MW	USD 179 M
The Roscoe Wind Farm [50]	Texas, USA	2009	781.5 MW	USD 1B
Glencore RAGLAN Mine Renewable Electricity Smart-Grid [51]	Quebec, Canada	2015	20.1 GW	USD 7.8 M
Hebei Shahe Power Plant [52]	Hebei, China	2013	1200 MW	USD 750 M
Xiangjiaba shanghai power company [53]	Shanghai, China	2020	6400 MW	N/A
APDCL [54]	Assam, India	2010	90,000 MW	USD 3.67 M
CESC [54]	Mysore, India	N/A	151,890 MW	USD 4 M
HPSEB [54]	Himachal Pradesh	N/A	533,000 MW	USD 3 M
UGVCL [54]	Gujarat	2014	1,700,000 MW	USD 10 M
TSECL [54]	Tripura	2005	128,730 MW	USD 9.8 M
Jeju Island Smart Grid Project [55]	Jeju Island, South Korea	2009	N/A	USD 208 M
Setana Osato Wind Power Plant [56]	Hokkaido, Japan	2020	51.2 MW	USD 115.84 M

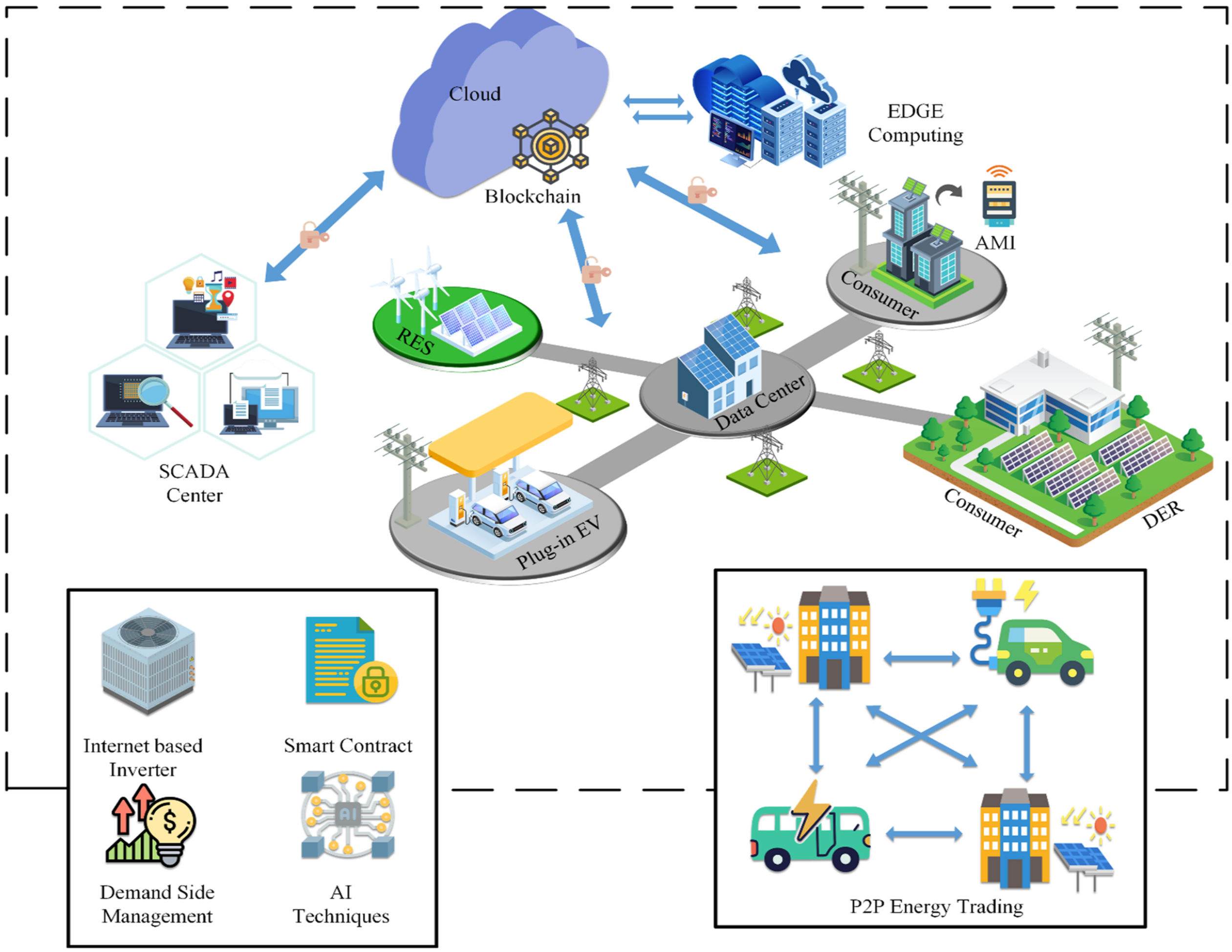


Fig. 3 A conceptual framework for next-generation smart grid energy system

agent-oriented software. An NGSG can be modularized by assigning the data-driven autonomous software that may virtually control the individual components of an NGSG and convert the centralized SG technology to a scalable and adaptable decentralized technology. For a multi-agent system, the complete NGSG is not needed to be recognized at any single point of a node, while the individual components can work towards predefined goals to achieve optimized performance, where the agent-backed components can interact with the system as well as each other [57]. However, the characteristics of an agent depend on the goal, which can be assigned to be cooperative or competitive with the other agent's characteristics. The aggregated characteristics of the agents may be able to determine the generation characteristics of the whole NGSG system. These characteristics of agents can be tweaked and redefined for better optimization of an NGSG. Then the failure of components in an SG does not result in total system failure because the individual agent works automatically with the initial knowledge it possesses.

3.2 Intelligent agent-oriented energy conversion unit integration

Agent-based energy models can be used to optimize energy conversion to minimize loss and maximize output. A conventional SG is designed to convert energy on different levels. However, in each step of conversion, energy loss may occur. Analyzed data gathered from such energy conversion systems can be used to construct agent software for efficient energy conversion. At the time of converting energy from one state to another, sometimes losses occur in the form of energy other than the required output energy. Such losses can be reverted by reusing the excess energy through efficient agent-based conversion. The synchronization between different conversion devices may have to be done using a multi-agent system for both monitoring and integrating the devices into the main power system [58].

3.3 Edge computing for energy data management

To eradicate the issues from IoT in conventional SG, edge computing (EC) is a technology of great significance for an NGSG. The IoT approach for SG collects massive datasets that are difficult to process because the cloud servers are situated in a distant geographic area. The networking system is stressed when raw data collected from IoT devices are transmitted to the cloud because of the increases in latency and reaction time. The data collected from an SG may contain private data, and as the data are sent to a third-party cloud server it may pose the risk of privacy breach. The EC solution shows huge potential to remove these problems which are presented by typical IoT systems. It takes the data close to the collection point where they are to be processed [59]. Another perk of EC is that it can reduce the network load to a great extent by shrinking the volume of transmitted data. This creates a low-latency high-response network system essential for the forthcoming SG systems.

EC creates a hierarchical architecture, as shown in Fig. 4. The architecture consists of multiple processing layers where all the IoT devices in the SG are located. In EC, some data processing tasks are shifted from the clouds to the multiple layers. The processing is done at lower-level layer, unless it needs more computation than that when it is offloaded to the higher layers. For example, some embedded IoT devices can perform small prepossessing like noise filtration. However, the computation capabilities of these devices are limited, and sometimes they cannot fully process the data. Thus, the data are sent for processing to a higher layer with more computational power and gateways. The gateways in EC provide local computation work with the IoT devices parallel to their conventional work. Different data-driven models may have to be used to process data in the EC structure, such as a prediction algorithm based on reinforcement learning for energy price estimation and home scheduling [60], and the heuristic evolutionary model for advanced demand side management by load shifting, a model which aims to reduce peak load and cost in the SG domain [61].

3.4 Interoperability between multiple energy hubs

The connectivity between different energy components in an NGSG will play a vital role in sustainable energy evolution. Market interoperability also needs to be explored to achieve overall connected operations over the entire system. A system's interoperability refers to its capacity to collaborate with other systems in order to share resources [62]. The multiple levels of interoperability in an NGSG can be divided into the following segments:

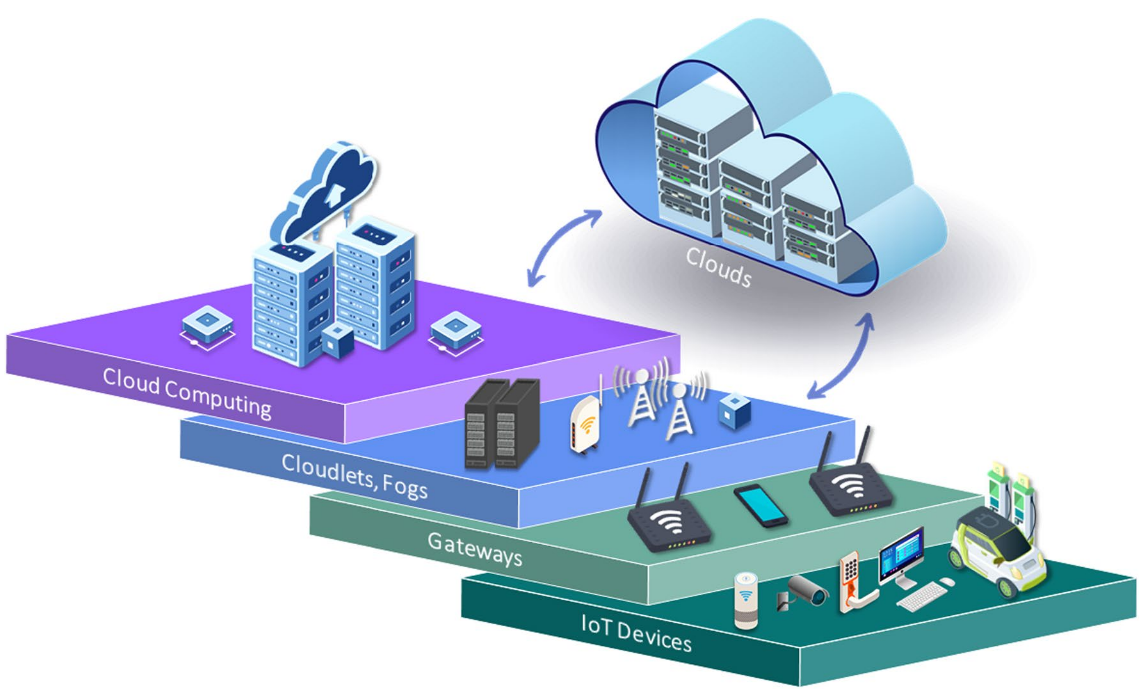


Fig. 4 Hierarchical architecture of edge computing consisting of multiple processing layers

1. *User interoperability* to ensure that there are options for customers to choose among various commercial and technological options.
2. *Commercial interoperability* to ensure that value can flow to where it is needed. Driven by market forces, it is important to confirm that incentives are matched across the energy system.
3. *Interoperability of data* to ease portability and data sharing between the components of energy sources, consumers and suppliers.
4. *Equipment interoperability* to ensure that the equipment is replaceable or exchangeable when there are changing demands, to allow energy consumers to make intelligent and informed choices.
5. *Vector interoperability* to make sure that timely coordination takes place and that energy provisions across various components of the energy system are compatible with each other.

3.5 Internet-based inverter control technology

Intelligent data-driven inverter technology plays a significant role in the root-level controlling of an NGSG by ensuring the mutual connection between generators and loads. These smart inverters have the capacity to connect with IoT devices with more embedded intelligent data-driven software. This emerging technology ensures the devices perform more intelligently in relation to quick response, effective fault diagnosis, automated maintenance, etc. [63]. The inverters in an NGSG will work autonomously without intervention and take a sophisticated step towards the control of power conversion. Smart inverters will be aware of their adjacent environment and guarantee quick adaptation to sudden changes in the context of an SG. They will also have the ability to learn from the accumulated data to enhance future adaptability and control management.

3.6 Self-healing grid enabled by agent-based control

The most crucial traits of an SG include self-healing capacity in the presence of unexpected conditions. When defects are found, the power system networks may have the ability to automatically restore the information. Although it is inevitable to have defects and disruptions in power systems, the potential dangers mainly depend on the fault magnitude, nature, duration, and location. The integration of sensors, self-operating sophisticated controllers, and cutting-edge software tools make up the agent-based self-healing grid. It will use real-time data to locate and isolate issues, restructure the system and reduce the number of impacted consumers. To attain the control of self-healing under faulty conditions, an agent-oriented control technique based on optimization is required for the SG domain which will mitigate the effect of over-voltage by enabling the automatic restoration of the sound condition of the power network [64]. In terms of the multi-agent control systems, fuzzy logic is used to make decisions.

3.7 Agent-based holonic approach on the demand side

To balance the demand and supply sides of an NGSG, multi-agent-based holarchies consisting of various abstraction layers of the distribution grid may have to be proposed as a holonic approach [65]. The holon concept may be applied as a holonic multi-agent approach to manage the information technology-based infrastructure of NGSG. This leads the path to efficient data transfer and robust communication security.

4 Data-driven next-generation smart grid

4.1 Critical steps for data-driven NGSG development

The framework of a data-driven NGSG may depend on the forming of the critical steps as shown in Fig. 5, which demonstrates how a data-driven NGSG solves critical issues and develops the final model for a data-driven NGSG. The bottom of the pyramid is the first step and the top is the last step of the process. Every step in developing the NGSG framework shown in Fig. 5 is discussed in detail in the following sub-sections.

4.1.1 Identifying problems

First and foremost, the SG power system needs to be thoroughly studied to understand the issues to be solved for system sustainability. Understanding the problem plays an important role in data management modeling. The SG power system may produce a large number of datasets that can be analyzed using different data science tools. Most of the data may prove irrelevant when coming to the goal of data science modeling, and thus, datasets related to the problem to be analyzed are of the most significance [66]. Intensive studies of the power system incentivize data collection, as it simplifies understanding of the type of data that is needed for further analyzing the data algorithms.

4.1.2 Data requirement and data collection

Data science methods need a huge amount of data to properly analyze certain system characteristics. The more data are available from a system, the easier it is to generate the final output. Data from an SG can be generated by enabling smart meters, sensors, and PMUs. Automation in data collection is an important aspect in the sustainable and robust modeling of the data science method [66]. The required data can be found in the first step, “Identifying problems”. Additional datasets, such as the power system configuration, voltage and current levels,

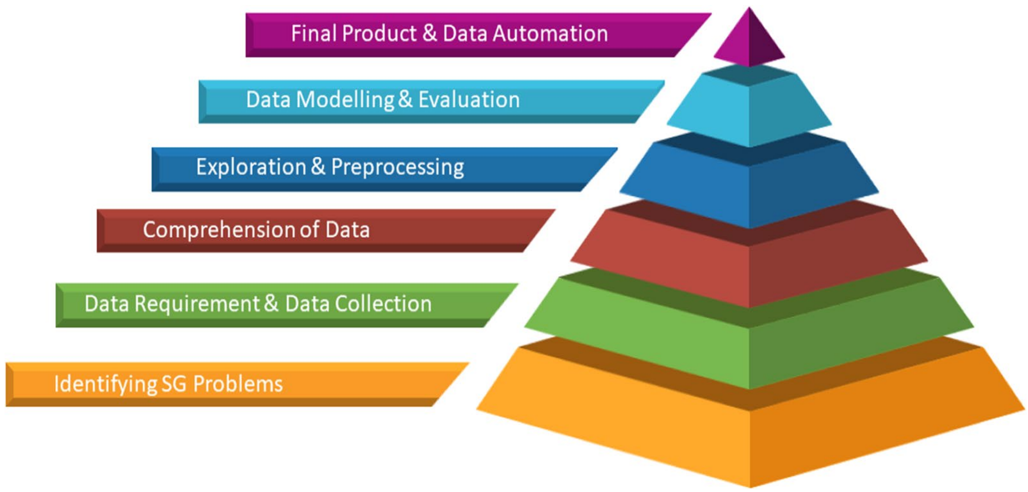


Fig. 5 Critical steps to develop a data-driven next-generation smart grid

transformer and generator information, security system, load flow, etc. may need to be added to improve the data science modeling.

4.1.3 Comprehension of data

The data should be studied after collection, and categorized based on the different characteristics of the system. The accuracy of data measured or collected should be high because data science methods require accurate data for smooth analysis. If the measured datasets are far from their actual values, the final output from the data algorithms may not be satisfactory [66]. The comprehension of data will significantly enhance the process of acquiring data as well as understanding which data are needed most for the system model. Nonetheless, several characteristics of data such as data type, data quantity, data accessibility, data features, the combination of multiple datasets, previous datasets, etc., should be given attention for better data-driven NGSG modeling.

4.1.4 Exploration and pre-processing of data

Exploration of data involves the analysis of a dataset to summarize its key aspects. Data are explored at first to understand the essence of data towards assessing the quality and characteristics of the data. Various statistical representations can be used to process these datasets with different points of interest. This helps to understand initial trends and attributes of the data. The quality of data may be further enhanced by using various pre-processing methods consisting of noise reduction, finding missing data, smart labeling of data, data filtration, and data formatting. One of the main goals of data preprocessing is to solidify the quality of data by correcting, reformatting, and combining datasets [66]. Some of the processes for enriching the available data include data cleansing, data transformation, finding missing values, unbalanced data handling, bias issues handling, distribution of data, detecting anomalies in data, etc.

4.1.5 Data modeling and evaluation

Different forms of data-driven and machine-learning models should be chosen for data analysis with the best fitting of the data according to the type of analytics. The typical process for separating data into training data and test data is either done by dividing the available datasets into a ratio of 8:2 or using the k-fold method for data splitting. To maximize model performance, it is necessary to split data and observe [67]. To test model performance, several model validation and assessment benchmarks can be used. These can help data scientists choose or build the learning method or model. These benchmarks include true positive, true negative, false positive, false negative, error rate, accuracy, precision, recall, receiver operating characteristic analysis, f-score, applicability analysis, etc. In addition, researchers may use sophisticated analytics, which may include feature selection and extraction, feature engineering, tuning algorithms, ensemble methods, modification of existing models, etc. to improve the final data-driven model for smart decision-making to handle specific system problems.

4.1.6 Final product and data automation

The final product is the outcome of the system after processing and analyzing all the data. It can be a recommendation, a comprehension, or a forecast. The obtained data product is used to make the best decision on various problems. In practical application, several data products have made considerable contributions to make the system intelligent and self-activated [67]. In the case of energy trade, information gained through data analysis, such as churn prediction and customer segmentation, can be used to make smart decisions towards sustainable energy trade. Finally, the whole process of collection, comprehension, processing, and modeling data should be run through an automated algorithm system, thus eliminating the need for manual handling and ultimately reducing data processing time and increasing efficiency.

4.2 Data-driven techniques used in NGSGs

Properly designed data-driven techniques can have the ability to make the updated version of an SG and solve the existing problems related to insufficient, incorrect, and unreliable data. These techniques consist of different types of algorithms, which are broadly divided into three categories: supervised, semi-supervised, and unsupervised [67–81]. A summary of various data-driven algorithms used in SG for executing and improving different functions is reported in "Appendix 1.1". However, the theory behind the development of DDTs can also be split into numerous categories, as discussed in the following sub-sections.

4.2.1 Bayes concept-based learning technique

As a practical data-driven technique, the theory of the Bayes concept establishes the connection between the model and the dataset. Deep learning-driven processes adhere to the Bayesian framework, and its methods exist to measure uncertainty. The Bayesian approaches can be applied to forecast net load in NGSG systems, while a deep long-short-term memory (LSTM) and Bayesian theory can be combined to anticipate the aggregated load in SG systems. A recurrent neural network (RNN) with memory cells which can store important information for a long time can perform effectively for the loads based on long-term reliance, significant volatility, and unpredictability. Conversely, completely Bayesian inference can be used to pick models for both evidence-based and predictive frameworks. The models for both frameworks can be chosen using fully Bayesian inference. Several studies have shown that the predictive approach, which displays data overfitting, does not perform as well as the evidence framework in this area [82].

4.2.2 Probabilistic learning technique

The probabilistic learning concept for smart energy systems includes binary and Bernoulli, univariate Gaussian, and multinomial and categorical distributions. The binomial distribution expresses the probability of a certain value among one or more independent values for a given set of parameters. The probability distribution of the intelligent power system has been significantly influenced by the binary and the Bernoulli distribution model. For plug-in electric vehicles, several methods have been developed to ascertain the probability distribution for their charging patterns at various periods of usage [83]. To reduce uncertainty and volatility in power systems, most studies have recently embraced grey Bernoulli approaches, and as a result, prediction now takes less functional data and research, especially when predicting long-term development.

4.2.3 Common univariate distribution technique

Probability studies typically address common distributions individually when it comes to the data-driven process, e.g., Student-t-, Gamma, Cauchy, and Beta distributions, Laplace irradiance, etc. The Cauchy distribution is heavily used in the analysis of power system harmonics, estimation of wind power uncertainty, prediction-based models, and real-time dispatch of wind-based power plants. The Gamma and Weibull distributions are two methods that are widely used to determine wind speed in dispersed generation [83].

4.2.4 Optimized learning technique

Power systems frequently provide diverse optimization strategies for various issues such as non-linearity, sensitive to uncertainty, and large-scale. The constrained [84], bound and blackbox free optimizations [85] are some of the techniques used in the SG domain. There are also first-order and second-order approaches. The first-order optimization approach is widely used in the classification of numerical optimization strategies that use the first-derivative methodology, while the second-order approach, often called the Newton technique, applies the second derivative in a scalar problem. These modifications have a significant impact on the power system's optimal power flow problem [83]. Optimal power flow is an optimization tool for running power systems and controlling energy. The linear programming, Karush–Kuhn Tucker conditions, quadratic programming, and estimation of wind power uncertainty can be applied to SG systems in many ways, including but not limited to power generation planning, power system expansion, advanced energy systems, power flow analysis modeling and heuristic methodologies, threats, unpredictability measures, and demand response.

4.3 Key analytics to adopt data-driven techniques in NGSG

Numerous processes of analytics shown in Fig. 6 can significantly aid the data-driven techniques used for an SG [86]. Figure 6 shows the process of prescriptive analytics, predictive analytics, decision intelligence, and data mining.

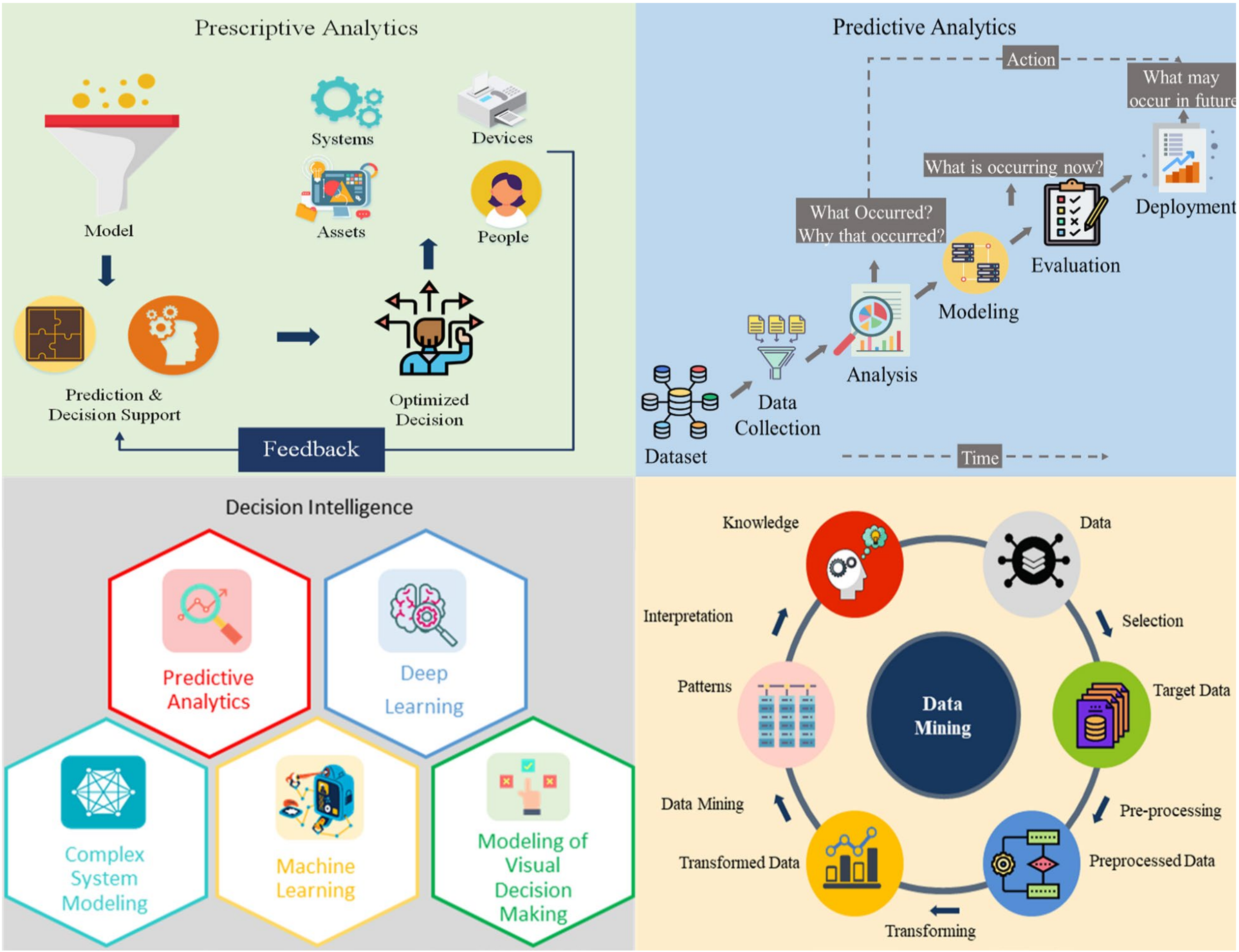


Fig. 6 Variable processes of advanced analytics using data-driven techniques

4.3.1 Predictive analytics

Predictive analytics, a form of advanced analytics, uses statistical modeling, past data, data mining methods, machine learning, etc., to forecast future events. In Fig. 6, data are gathered from numerous datasets and analyzed to comprehend the reasons for and results of every occurrence. A pattern is created, and then all the data are evaluated statistically. Finally, predictive analytics predicts the outcome.

4.3.2 Prescriptive analytics

The practice of using data to decide the best action is known as prescriptive analytics. This form of analysis generates recommendations for the next moves by taking into account all the essential aspects. As shown in Fig. 6, prescriptive analytics methods analyze the model and extract knowledge from the data. Then, possible future outcomes are generated. Observing the possible outcomes, prescriptive analytics gives optimized decision to the systems, or devices, or people, on the action they should take.

4.3.3 Data mining

The process of going through massive datasets to uncover patterns and links in order to forecast outcomes by data analysis is known as data mining. The process of data mining from collection and selection of data to acquiring knowledge via processing target data and interpreting patterns is shown in Fig. 6.

4.3.4 Cohort and cluster analytics

Cluster analytics refers to the grouping of similar data into a number of finite clusters. This is a type of behavioral analytics that divides data into clusters before analysis. The clusters carry similar characteristics or experiences over a period of time. Multiple clusters are made when the substance of a group of data varies from another group of data. Being an unsupervised analysis method, cluster analysis does not assure the number of clusters beforehand, while the number of clusters is only revealed after the clustering algorithm is completed. It tends to find a core similarity within data thus separating a group of data based on differences among the groups.

4.3.5 Decision intelligence analytics
Decision intelligence is a trending method that uses data-driven techniques to make a decision based on cause and effect. It uses various models and algorithms of data science assisted by social and managerial sciences. This method is important for designing, modeling, and tuning the decision-making process of power systems. Figure 6 shows that decision intelligence blends multiple decision-making methodologies with AI, ML, automation, and relevant information.

4.3.6 Operationalizing and scaling
Operationalizing refers to the materialization of an abstract idea or concept into a measurable form. This method is valuable for collecting data on abstract or unobservable systems, e.g., future power systems, in a systematic way. This can quantify different parameters of an NGSG as such power systems are not yet available in a practical environment and can only be observed as an idea or a simulation. Conversely, the scaling of an NGSG refers to a comparatively small size prototype and is essential for running various operations on a small scale to identify the characteristics of that test before conducting it on a large scale in the original power system.

5 Data-driven techniques in NGSG: prospects and adaptation challenges

5.1 Prospect of data-driven techniques in NGSG

5.1.1 DDTs in intelligent energy materials processing
Energy materials production is on the verge of a breakthrough as per the advancement in data-driven technologies for materials research. Significant growth in the field of materials science can be found in [87]. The recent improvements in data-driven techniques for materials engineering show that ML innovates intelligent energy materials' production and design process. Additionally, it can be used for measuring the electronic properties of power systems. In using data-driven techniques like ML, the first step is to gather the objectives and set goals to achieve them. This is the most important step as the goals must be specific and achievable from the available datasets or information. The data-driven ML models are useful for enabling a low-cost and reliable approach toward predictions where computational or experimental approaches increase expense.

5.1.2 DDTs in intelligent energy systems component
A smart energy system is made of multiple components for the generation, storage, distribution, and consumption of energy. These aspects of energy systems can all be subjected to data-driven techniques such as ML or artificial intelligence (AI) for the performance improvement of an NGSG. With the enabling of this technology, data gathering from connected devices has provided a better understanding of system characteristics and improvement in various details. The data-driven ML methods have the ability to allow better simulations to construct prediction and forecasting models. The energy storage system of an NGSG should be improved for efficient charging and discharging of the storage devices [88].

5.1.3 DDTs towards intelligent demand-side management
Data-driven ML technologies play a significant role in demand-side management by allowing energy consumers to try out different market mechanisms in practical scenarios. A sophisticated integration of demand side devices, such as solar PV, battery storage, and smart meters, is done through linking with the internet, being associated with ML techniques, and following advance in data collection and data sharing. The concept of "smart homes" is very popular now and the number of smart homes has seen a spike in recent years [89]. A Swedish pilot project was done on reducing peak energy usage significantly by implementing data-driven ML techniques in the field of demand response management [90], in which a multi-agent approach offers demand responses in the NGSG by allowing coordination among its components. Energy devices can communicate with the power grid and exchange information by giving access to the dynamic communication system [90]. The demand response programs are categorized into two groups, i.e., price-based and incentive-based. Real-time pricing, rate per usage, critical peak pricing, etc., are included in price-based demand response, whereas emergency response, direct load control, ancillary market services, market capacity arrangement, and buyback programs are included in incentive-based demand responses. These categories can be subjected to data-driven management techniques for better demand-side management of an NGSG.

5.1.4 DDTs towards smart manufacturing in NGSG
The fourth industrial revolution has enabled the production and collection of data from connected machines in industry. Data-driven ML techniques can be used to analyze the collected data as an approach to smart manufacturing [91]. Some of the various ML models used for smart manufacturing include: (1) support vector machine

(SVM); (2) k-nearest neighbors; (3) Bayesian networks; (4) artificial neural networks; (5) decision tree; (6) multiple logistic regression; (7) k-means; (8) random forest; (9) gradient boosted; and (10) additive models. The newer business models also require smart manufacturing. This is enabled by the technical advance of Industry 4.0. In a data-driven smart manufacturing system, the benefits of real-time data analysis, advanced decision-making, better plant efficiency, and increased production may be crucial for NGSG modeling.

5.1.5 DDT in intelligent energy resource planning

Energy forecasting and management is a significant field of interest for energy resource allocation and demand-side handling [92]. Decision-makers can be assisted by different data-driven decision-making techniques constructed by data experts. These contribute a lot to designing energy plans, choosing optimal decisions, and finding alternatives. Robust energy systems enabled by intelligent planning allow the use of data-driven algorithms to identify market conditions and aid the building of advanced energy devices. The real-time applications of data-driven methods in the field of energy are commonly seen in various energy systems. A key aspect of data-guided techniques is the use of AI to improve NGSG performance [91]. The incorporation of the IoT in intelligent energy planning and management is also one of the most significant aspects of data-driven techniques used in the energy industry. The IoT can enable access to remote access and control of an NGSG with a smart tracking system. Here, smart meters inform consumers about the volume of energy consumption, while local infrastructures like microgrids can be connected to cloud servers to exchange information to enable significantly better load forecasting.

5.1.6 DDTs in integrating the large-scale heterogeneous energy sources

Policy makers have already been focusing on the up-scaling of renewable energy. This will affect the energy market. Thus, power grid operators and engineers are putting emphasis on data-driven techniques and models to achieve a seamless transition from fossil fuel to renewable energy. Harnessing energy from renewables on a large scale requires enabling multiple green sources of energy at the same time, which signifies the importance of heterogeneous energy sources. The synchronization of such sources can be guaranteed using data-driven algorithms, including collecting and analyzing data from the sources with specific ML models. For example, solar and wind power plants already generate a huge amount of data which allows data-guided techniques to forecast different levels of energy with the help of sensor integration [93]. These energy consumption datasets can be analyzed to predict peak and low demand times, and design the production rate to minimize losses. However, the up-scaling of green energy sources also opens a door in an NGSG for cyber attackers. Thus, the security of an NGSG should be ensured by updating the data-driven ML models regularly to increase integrity.

5.2 Challenges to implementing DDTs in the NGSG

The development of a data-driven smart grid system toward achieving sustainable energy transition has some challenges from various points of view. In the following sub-sections, a thorough discussion on the challenges during the adaptation of DDTs in the NGSG is conducted.

5.2.1 Engineering point of view

1. *Overfitting mechanism* When a model tries to forecast a trend in excessively noisy data, overfitting may occur. This is the result of a model that is too complicated and may have a large number of parameters because it does not accurately reflect the reality in the data. A typical data-driven ML network may contain millions of variables. The training data model typically consists of a large number of records. However, even when a network recognizes the training set and gives answers that are hundred percent precise and correct, it may entirely fail when faced with new data. This mechanism is known as overfitting, and is one of the limitations of data-driven techniques [94].
2. *Installation of intelligent energy processing unit* Intelligent processing methods need complex thermo-chemical operations and multi-component frameworks. These generate a lot of data quickly. The best scenario is for operators to receive rapid data on the properties of energy material manufacture and process parameters in real-time, allowing them to identify novel processes and phenomena more quickly and react effectively and efficiently. Existing techniques, however, provide "postmortem" data yearly after the manufacturing process has ended. To improve and assess the production process, data-driven ML techniques can be applied [95]. However, the existing data-driven techniques may demand a revision in their structure to maintain the energy materials and electric infrastructure at the energy distribution level.
3. *Feasible energy storage material* The enormous amount of background data and the increasing complexity of energy storage systems provide significant hurdles for the current methodologies and algorithms. For greater precision, stability, and efficiency,

emerging cutting-edge technologies can address the shortcomings of traditional approaches. First, the development of energy storage encompasses invention and breakthrough, long-term storage, a high amount of protection for electro-chemical backups, and cheap cost. This low-cost technology is also necessary for high efficiency and physical storage. Secondly, research is focused on modeling energy storage and streamlining the procedure in different energy systems, supporting the use of energy storage technologies, and developing innovative structures and thorough evaluations for modernizing and advertising energy storage [96].

5.2.2 Technology point of view

1. *Tech advancement* Argonne scientists are trying to develop optimization approaches that combine ML and AI to simulate the intricacy of various electrical system challenges much more quickly than the current methodologies. The primary focus is to accelerate load flow analysis and daily computation of the electricity system [97].
2. *Improved energy efficiency* Future difficulties in sustainable 5G and 6G power management hold significant potential for data-driven methods. For the cost-effective design and optimization of network operations, data-driven ML approaches, like federated learning, deep learning, and optimization may be considered. By gaining flexible network structure and altering traffic conditions, it is possible to construct 5G or 6G air interfaces. Using a variety of 5G and 6G technologies, including SG, intelligent transmission and distribution of network lines, smart buildings, and industrial automation, data-driven ML will be more widespread and crucial than simply conserving energy. On the other hand, these approaches typically require coordination and computing, which can pose significant challenges for the design and implementation of power-efficient data-driven techniques and for upcoming 5G and 6G networks [83].

5.2.3 Decision-making point of view

1. *Decision-making* When making decisions on energy distribution, data-driven techniques can improve intelligent system performance. Any situation involving decision-making in a dynamic environment can benefit from reinforcement learning. Agricultural production optimization, robotics, automatic control and adjusting (i.e., heating, air conditioning, and ventilation), and supply chain optimization are examples in which DDTs can assist. In the future, renewable energy sources should be employed fast and in ways suitable for their unpredictable nature. Here, light and wide energy usages are offered in the interim by the placement of smart meters. The efficient use and analysis of the data can present new load forecasting options where proper decision-making can be difficult [98].
2. *Decision on demand response unit* Demand response representatives should first operate in a barely visible environment, which means they cannot properly understand the working process of the demand response unit. The structure and administration for demand response are designed to monitor and use real-time data on energy consumption to offer energy pricing for thousands of customers via the utility power grid. Customers can adjust their energy consumption in response to grid conditions and the rates. By assisting end-users to think about how they need power grid improvements, ongoing growth can increase reliability, cost-effectiveness, and sustainability. The prospective integration of renewable energy directly into the power grid will be encouraged by the corresponding knowledge and such resilience [99].

5.2.4 Others

1. *Economic challenges* The energy storage industry is now facing difficulties in several countries, including weak legislative support, high price, doubt in value, unsound business practices, etc. In the coming years, it will be crucial because of two factors: first, the suggestion of substitutes to the energy storage plan including power generators and electrical firms; and second, the development of a suitable business competitive structure and arrangement of sufficient funding schemes for fresh data-driven advanced technologies [100]. According to Woori's forecast, the cost of energy storage will increase globally by 26 percent in a year. Although there are various market variables for energy storage, the primary obstacles continue to be high costs, poor subsidy programs, a median cost configuration, and lack of a business prototype.
2. *Trained consumers* Many companies have to deal with the challenges of training their consumers on how to use cutting-edge technologies. The same task may be required of data engineers. Investors, developers, and managers overestimate the existing capabilities of data-driven techniques, while anticipating

that the algorithms will comprehend difficult issues with ease and make reliable predictions.

3. *Lack of expert manpower* Even though the market for data-driven methodologies is attractive to many people and the energy sector, to further develop this research field it requires more expert manpower. In the energy sector, power utilities face challenges when innovating new technologies because of a lack of skilled employees.

6 Trends of DDT towards sustainable energy evolution in NGSG

Utility firms may allocate resources more effectively, reduce costs, and find better ways to serve customers with the help of the proper analytical platform. Additionally, the appropriate data analytic platform enables them to maximize the value of the generated data. This can help the sustainable energy evolution through the improvement of the following aspects in an NGSG.

6.1 Securing reliable control operation

The most crucial aspect of SG energy systems is their ability to operate securely and reliably. The SG has already benefited from the involvement of data driven techniques in terms of stability, security, and dependability [90]. It is well recognized for providing timely and efficient stability analysis which claims the implementation of automatic control. The use of data-driven techniques, like machine learning, reinforcement learning, and deep learning in stability and control analysis has been the subject of extensive research in recent decades, as shown in Table 4. It is realized that the implementation of the data-driven techniques in an NGSG may offer a reliable solution to address the control issues in terms of frequency, voltage, preventive and restorative measures, and enable a sustainable energy evolution through the reduction of CO_2 emission in the environment.

6.2 Definitive energy management

Energy management is associated with the control, planning, and monitoring of energy-related processes to conserve energy resources, reduce energy costs, and safeguard the environment by minimizing CO_2 emissions. Energy management through advanced data-driven methodologies has already started in SGs as shown in Table 5. The advantage of using the advanced methods is the ability to perform work in less time, while offering a realistic solution to manage energy over a small amount of data. This is done through enabling DDTs in SG planning and management, including grid synchronization, active and reactive power management, ancillary services, and techno-economic modeling. From Table 5, it may be predicted that the emergence of DDTs in an NGSG also paves the way to contributing to sustainable energy evolution [112].

6.3 Precise asset condition monitoring

Old assets are a prime cause for uncertainty in load and demand management, affecting optimal operation and the overall health of the NGSG. Thus, constant monitoring of all the assets of an NGSG is needed to reduce the risk of equipment failure [120]. Obsolete technologies are also to be replaced with advanced technologies. Various data-guided methods, as shown in Table 6, can be prime examples of asset monitoring systems where data taken from an NGSG are analyzed to understand the asset conditions.

6.4 Accurate fault prediction and characterization

Traditional fault detection algorithms, like impedance based and wave-based techniques, cannot adjust with the penetration of distributed renewable power generation [130]. On the other hand, AI-based data-driven approaches can bypass challenging modeling and fault mechanism analysis. A fault classification approach based on a data-driven CNN fed with features retrieved by the Hilbert-Huang Transform (HHT) in power distribution systems is proposed in [131]. This approach performs admirably in fault classification thanks to the CNN's strong feature learning capabilities. Another data-driven Graph Convolutional Networks (GCN)-based method for addressing fault location is suggested in [132]. It keeps the spatial information of buses in the GCN structure, which allows improved fault detection accuracy. To achieve fault detection and location, the voltage and frequency signals are used, respectively. Additionally, a fault contour map that groups the buses into several tiers based on the severity of the impacts is provided. A short summary on the recent progress of data-driven techniques for precise SG fault characterization, detection, and location identification is shown in Table 7. It is seen that the data-driven approaches can satisfactorily perform fault diagnosis, though their performances may suffer because of a lack of sufficient data. By developing the data-driven NGSG infrastructure, data can be gathered from various sources and then combined and used to increase the precision of defect diagnosis. This can help improve sustainable energy evolution.

6.5 Accurate forecasting and uncertainty estimation

The increasing integration of RESs, such as tidal, solar, wind, etc., demands more effort to schedule and operate an SG. Load forecasting (LF) is a crucial component for planning and running modern power systems since it helps to preserve stability, and keep the environment

Table 4 Secure control operation techniques

Type of control operation	Data-driven techniques	Applied system	Robustness	Advantages	Problems and challenges
Frequency control [101]	IRL (Internal Reinforcement Learning)	Single Area System	Yes	Tackles unknown dynamical environments, and acute load deviations	Weak results may occur due to overloading
Frequency control [102]	Genetic algorithm	New England 10 generators (applied for 1st test), 39-bus system (2nd test)	Yes	(1) Simple design, (2) Flexibility of performance, (3) Reaches steady state faster, handles discontinuities, and noisy functions	(1) Mutation operator may cause legal issues, (2) Repair function has to be applied individually
Frequency control [103]	Neuro Hybrid Fuzzy Logic	Four area (Inter connected Hydro Thermal) power system	N/A	(1) Fastly controls the non-linearity, (2) Lessens FPD (frequency peak deviation), time imprecision, and tie-line power	Needs regular update for proper functioning
Frequency control [104]	RADP (Robust Approximate Dynamic programming)	Multi machine power System (New England 10-machine 39-bus system)	N/A	Recovers system's frequency	System dynamics is unknown at the time of training datasets
Preventive [105]	MLP (Multi-layer perceptron)	(IEEE 6, IEEE 118), Real 33-bus Bodaibo subsystem	No	Prevents voltage instability and longer overload, and has smart preventive security	Independent variables are affected by dependent variables to an unknown extent
Restorative [106]	Q learning	IEEE 57-bus, IEEE 118-bus and IEEE 300-bus systems	Yes	(1) Avoids risk of new failures (cascading), (2) Restores with better power grid topology	(1) Works on one individual power grid at a time, (2) Takes high computational time when recovery sequence increases
General control [107]	Q-learning, WoLF-PHC WoLF (win or learn fast) PHC (policy hill climbing)	Secondary users (SUs)	No	Improves anti-jamming performance	Works better with only discrete and finite datasets
General control [108]	Wolf-Pack	Multi-agent systems	Yes	(1) Efficient coordinated control, (2) Virtual tribes control optimizes electricity generation as wind, solar, and electric vehicles increase	Time delay occurs while obtaining optimal strategy
General control [109]	Lazy Learning	Parallel cyber physical–social energy systems	No	(1) Highest control performance, (2) Reduced frequency deviation, (3) Predicts next state	Lacks optimization algorithm
Voltage control [110]	Teaching–learning, Sugeno fuzzy logic	AVR system	Yes	Delivers adequate robust performance and satisfactory dynamic responses over parametric fluctuations of system	Cannot overcome OPD (optimal power dispatch) problems
Voltage control [111]	DQN/DDPG	200-bus test system	No	Automatically and efficiently control voltage according to situation	Lacks multiple control operation

Table 5 Energy management techniques

Data guided management models	Methods and techniques	Purpose	Constraint	Advantage	Disadvantage
Residential home energy management models [113]	Binary backtracking search algorithm, Mixed integer linear programming, Internal genetic algorithm, Outer particle swarm optimization, Integer linear programming	Energy consumption optimization and Electricity bill reduction	Consumer comfort is compromised	(1) Increased energy efficiency, (2) correctness and calculation speed, (3) Reduced electricity consumption	Solution is often not optimized
Household appliance scheduling management [114]	Shuffled frog leaping algorithm, Teaching and learning based optimization algorithm	Gross bill reduction and peak power reduction	Consumer comfort is compromised	Problems related to combinatorial optimization are easily solved	Convergence operation is sometimes slow and premature
Residential power scheduling management [115]	Integer linear programming	Establishment of sophisticated trade-off between bill and pay	Peak-to-average ratio is ignored	Capable of high-quality decision making	Unable to handle non-linear data and uncertainty
Network based energy management models [116]	ZigBee, Wi-Fi, Z-Wave	Carbon footprint reduction	Consumer comfort is compromised	(1) Improves interoperability and energy efficiency, (2) Lessens carbon emissions, (3) Cost friendly	Transmission rate is lower comparatively
Heuristic home energy management models [117]	Heuristic algorithms	Electricity cost optimization and execution time reduction	N/A	(1) Affordable waiting time, (2) Reduces peaks in demand and electricity cost	Takes longer time to make optimum decision
Demand response management approaches [118]	Time of use pricing scheme, Real-time pricing scheme and Demand Response Algorithms	Potential social benefits and bill reduction	Consumer comfort is compromised	Ensures optimal consumption and prices	Requires high computational power
Fuzzy control based management system [119]	Time of use pricing scheme, Real-time pricing scheme, Inclined block rate pricing scheme	Peak-to-average ratio optimization	High computational power needed	Energy consumption reduction	The system needs frequent update

Table 6 Asset condition monitoring techniques

Component	Technique	Method	Condition monitoring	Advantages	Problems and challenges
Transformer	IoT [120]	A sophisticated monitoring system that will send a notification to a specific device for further action	Successful	Two-way communication within connected devices is enabled	Creates a complex array of connected devices which may become difficult to control
	SVM, K-Nearest Neighbor (KNN), Naive Bayes (NB), Random Forest (RF), Artificial Neural Network (ANN), Adaptive Boosting (AdaBoost), and Decision Tree [121]	To handle the issue of reliability and uncertainty of Health Index of power transformer using Artificial Intelligence based algorithms	Successful	Weak classifiers with less accuracy are strengthened for improved results	The classifiers should be high quality classifiers
	IoT [122]	A health monitoring system that uses temperature sensor to monitor transformers and send data for remote analysis	Successful	Two-way communication within connected devices is enabled	Creates a complex array of connected devices which may become difficult to control
	Adaptive Neuro FuzzyInference System (ANFIS) [123]	A comparison between fuzzy model and adaptive neuro fuzzy model to the Health Index (HI) of transformers based on various parameters	Successful	Able to capture nonlinear structure with high adaptability and learning capability	Faces difficulty in handling large input datasets
Inverter	SVM [124]	A condition monitoring system gained by training an SVM model with characteristics of DC-link capacitors in a three-phase inverter	Successful	Efficient handling of nonlinear data samples	Requires extended memory and longer training time
	Convolutional Neural Network (CNN) [125]	A condition monitoring system gained by using an CNN model to analyze characteristics of DC-link capacitors in a three-phase inverter	Successful	Automated detection of unique features in training data samples	Huge quantity of data samples is needed for completing analysis
PV	Machine learning and deep learning models [126]	Different condition checking systems based on machine learning models	Successful	Integration of different data processing tools bears strong output	Requires large amount of nonlinear abstractions for meaningful representation
	Deep learning, Reinforcement learning, Transfer learning, Ensemble learning [126]	Use of deep learning models to tackle different issues of condition monitoring	Successful	Presents an accurate, stable and robust algorithm	Interpretation ability is reduced in this method

Table 6 (continued)

Component	Technique	Method	Condition monitoring	Advantages	Problems and challenges
Wind turbine	ANN, Bayesian network, Support vector regression, RF, KNN [127]	Data on vibration taken from wind turbine is combined with data acquired from supervisory control and data acquisition systems (SCADA) which is analyzed using machine learning methods building a condition monitoring system	Successful	Able to develop individual prediction using historical data samples	Needs high computational power
	Bidirectional gated recurrent unit (BiGRU), CNN [128]	CNN and BiGRU methods are used on data acquired from supervisory control and data acquisition systems (SCADA) for condition monitoring	Successful	Quick response while using very little memory	Slight inaccuracies due to quick processing of data
	Deep convolutional generative adversarial networks (DCGAN) [129]	A health condition monitoring (HCM) system for wind turbine using DCGAN	Successful	Generates high quality artificial data which further enhances the training sequence	Requires large quantity of data and is also difficult to train

Table 7 Fault prediction and characterization techniques

References	Technical approach	Fault detection	Fault characterization
[133]	Dissimilarity learning method based on Clustering	✓	✓
[134]	IoT based model	✓	×
[135]	Power line communication-based data transmission algorithm	✓	✓
[136]	Machine learning algorithm	✓	✓
[137]	Neural networks, SVM, Decision tree	✓	×
[138]	Data extraction from smart meters and sensors	✓	×
[139]	Feature extraction method from big data	✓	×
[140]	Machine learning algorithm	✓	✓
[141]	Big data approach towards data processing from smart meters	✓	×
[142]	ANN	✓	✓
[143]	Convolutional sparse autoencoder	✓	✓
[144]	Supervised data-driven topic model consisting of heterogeneous network of information	✓	×
[145]	Field programmable gate arrays (FPGAs) based higher order statistical method	✓	✓
[146]	Data-mining based model	✓	✓
[147]	Deep neural network	✓	✓
[148]	Data-driven Multivariate Exponentially Weighted Moving Average (MEWMA)	✓	✓
[149]	Sparse self-encoding neural network	✓	✓
[150]	ANN, Multiplier-based method (MBM)	✓	✓
[151]	Decision tree	✓	✓
[152]	Neural network	✓	×
[153]	Multivariate Statistical Analysis (MVA)	✓	×
[154]	Combined data analysis	✓	✓
[155]	Multi-agent model	✓	×
[156]	Mobile Edge Computing (MEC), IoT-based Solutions	✓	×
[157]	Neural network	✓	×
[158]	Neural network	✓	×
[159]	SVM	✓	×
[160]	Holonic multi-agent approach	✓	×

safe by reducing CO_2. Faultless load forecasting is useful for decreasing production costs, as it enables reducing utility risks by predicting future consumption of products that the utility will transport or deliver. However, it is highly challenging as the load is stochastic in nature [161]. Conventional forecasting models frequently do not disclose the degree of uncertainty in their forecasting, which can result in expensive and dangerous choices, and compromise attempts to develop dependable SG systems [162]. Before digging into the data-driven deep learning approaches of load forecasting, it is essential to categorize load forecasting techniques. The objective of short-time load forecasting (STLF) is to measure the load over a few weeks starting at one hour [163]. STLF is essential for the generation, transmission, and distribution of SG power. The data-driven techniques in Table 8 are used for improving STLF. The methods of Table 9 are used for analyzing the data for very-short-time load forecasting (VSTLF). For longer periods, such as medium-time load forecasting (MTLF) and long-time load forecasting (LTLF), the techniques shown in Tables 10 and 11 are used, respectively. It can be shown that the data-driven method can provide accurate forecasting for the NGSG model, and can also conveniently improve the possibility of achieving sustainable energy evolution.

6.6 Precise fraud characterization

Electricity utilities must deal with non-technical losses incurred by fraud and theft committed by their customers or third parties. Certain approaches have been developed to developed to detect potential scammers among consumers and third-party interference as listed in Table 12. Many data analysis-based approaches are taken toward detecting and diminishing fraud. Table 12 shows that fraud characterization may become more accurate and convenient by enabling the data-driven SG model. This can create a reliable security layer in diminishing

Table 8 STLF techniques

LF type	Year	Technical approach	Contribution	Challenges
STLF	2022	Hybrid method consisting of Prophet model, ARIMA (Autoregressive Integrated Moving Average) model, and LSTM model, and BPNN (Back Propagation Neural Network) [164]	Overcomes different technical gaps of load forecasting with low computational time and fast convergence	1. Complex calculation 2. High Calculation time 3. Expensive 4. Insufficient data 5. Clustering of data 6. Management of structured and unstructured data
		Extreme Gradient Boosting (XGBoost) [165]	Forecasts loads specifically for warehouses and logistics consumption	
		Enhanced decision tree classifier (EDTC) [166]	Accurate prediction of the stability of the smart grid	
		Mix-up and transfer learning [167]	A reliable model for load forecasting designed for new houses	
		Integrated CEEMDAN-IGOA-LSTM (Complete Ensemble Empirical Mode Decomposition With Adaptive Noise, Improved Grasshopper Optimization Algorithm, and Long Short-Term Memory Network) [168]	Aggregates different data techniques to effectively forecast load	
	2021	CNN, RNN [169]	Works quickly and smoothly in noisy systems	
		Asynchronous Deep Deterministic Policy Gradient (ADDPG) Adaptive Early Forecasting (AEF) Reward Incentive Mechanism (RIM) model [170]	Solves the problem of excessive temporal connection and high convergence instability	
		Integrated CNN and LSTM Network [171]	Highly precise and accurate STLF which can analyze long sequence time-series data of electric load	
	2020	LSTM, Reinforcement learning, DQN, BPNN [172]	Similar day identification and selection based on reinforcement learning on BPNN	
		RBN, MI-ANN, Genetic Wind Driven Optimization (GWDO) [173]	Load forecasting for linear and non-linear power systems	
		Singular Spectrum Analysis(SSA), Fuzzy ARTMAP, Neuro-fuzzy, BP [174]	Reducing the cost of computational energy and data requirements	
		Ensemble Empirical Mode Decomposition(EEMD), Multivariable Linear Regression(MLR) [175]	Analyzing large datasets for electric load	
		Kalman Filtering, Clustering techniques, Weightless Neural Network (WNN) [176]	Use of different clustering techniques to cluster load forecasting data	
		ELM, Genetic Algorithm, Support Vector Machine, XBoost, decision Tree [177]	Tuning hyper parameter and extracting features for load forecasting	
	2019	BP, LSTM, CNN [178]	Using LTSM and CNN for coupling electric load	
		XG-Boost, Decision Tree, Support Vector Regression (SVR) BP, CNN [179]	Predicting load forecasting within the price of electricity	
		WaveNet, CNN, BP, LSTM [180]	Improving performance of different error detection of load forecasting	
		LSTM, Ensemble learning, Quantile forecasting, Quantile method, ENN, Parallel computing [181]	Diminishing the need for feature extraction in load forecasting	

Table 8 (continued)

LF type	Year	Technical approach	Contribution	Challenges
		Unsupervised Learning, BP, Auto Encoders, Denoising Autoencoders [182]	Error reduction for unsupervised load forecasting	
		Dropout technique, Fuzzy logic, CNN [183]	Feature extraction improvement with high accuracy and the over-fitting issue resolved	
		WaveNet, CNN, BP, LSTM [180]	Improving performance of different error detection of load forecasting	
	2018	SVR, Auto Encoders, Denoising Autoencoders [161]	Achieving high features of load forecasting from lower-level datasets	

Table 9 VSTLF techniques

LF type	Year	Technical approach	Contribution	Challenges
VSTLF	2022	ANN [184]	Load forecasting with optimal asset management	1. Weak performance on unstructured and sparse data 2. Improper short time intervals 3. Insufficient data 4. High calculation time 5. Random and big data 6. Over-fitting problem 7. Management of structured and unstructured data
		Extreme Gradient Boosting (XGBoost) [165]	Forecasts loads specifically for warehouses and logistics consumption	
	2021	Markov-chain mixture distribution (MCM) model [185]	Develops a standard model for household power consumption	
	2020	FFNN, Neuro-fuzzy, Fuzzy Multi-Objective Decision Making (F-MODM) [186]	Develops load forecasting 1 h ahead based on weather data	
		RNN, GRU, BP [187]	Predicting load demand of residential infrastructure for a short period	
		DML, Apache Spark, Apache Hadoop, Linear Regression, Generalized Linear Regression, Decision Tree, Random Forest, Gradient-boosted trees, Distributed computing [188]	Reduces training time and testing time of load forecasting	
		CNN, Mutual Information (MI), MI-ANN, Relief F, Kernel Principal Component Analysis (KPCA), BP [189]	Over-fitting issue reduction and computational time reduction using CNN, KPCA, MI etc.	
		LTSM, Bayesian deep learning, Bayesian Theory [82]	Probabilistic-residential load forecasting for PV systems	
	2019	BPNN, Bayesian Regularization, Levenberg–Marquardt algorithm [190]	Load forecasting for individual district buildings	
		DBN, BP, Phase Space, Reconstruction PSR, Levenberg–Marquardt algorithm [191]	Predicting load forecasting of bus-load forecasting and distributed energy penetration	
		KNN-ANN, FFNN, Euclidean theory [192]	Load forecasting for hydro-thermal unit generation combining ANN and KNN	
	2018	Neuro-fuzzy, ANFIS, Genetic algorithm, Particle Swarm Optimization [193]	Decreasing execution or training time as well as reducing feature selection complexity	

and characterizing the fraud. This itself may accelerate the sustainable energy evolution process.

6.7 Safe energy trading (blockchain)

The highest priorities of every system are security, privacy, and trust. In the same vein, the upcoming SG should have a good level of security, including: 1) ensuring that an unauthorized third party cannot acquire any information; (2) ensuring established cryptographic techniques; (3) preventing information changes from unauthorized entities; (4) denying access without permission; and (5) ensuring authorized access to those with rights and privileges. Reference [229] presents a revolutionary consensus technique that makes Bitcoin the most popular application of blockchain to date, resolving the issue of creating trust in a distributed system. Additional approaches are also being used, including cryptographically secured data structures, digital signatures, time stamps, and incentive schemes. The majority of current solutions are based on centralized models. To make decentralized energy trading, blockchain technology has emerged and successfully trades energy among consumers, prosumers, and suppliers. Although these technologies are mature and functioning properly, the existing blockchain-enabled SG system has a number of problems, including consumer priority, security, and time consumption. Table 13 indicates the blockchain-based techniques and algorithms for safe energy trading. It is concluded that implementing DDTs in an NGSG can drive the world to sustainable energy evolution.

7 Future research directions

All the research conducted on DDTs and their results for various aspects of SG highlight the significance of methods to achieve sustainability as a whole for an NGSG. Reliable control operations powered by data-driven technologies may cover all the control problems of a future SG. The management models used in an SG can be improved by increasing computational capability to analyze large datasets simultaneously. This improvement can also ensure even lower carbon emissions and energy consumption, ultimately aiding the goal of sustainability. The next-generation blockchain enabled trading eradicates the chance of energy theft by keeping decentralized records of all the simultaneous energy transactions happening in a certain time frame. Further, the advancement

Table 10 MTLF techniques

LF type	Year	Technical approach	Contribution	Challenges
MTLF	2022	LSTM and NARX neural network [194]	Hourly energy demand prediction of a municipality	1. Over-fitting issue 2. Systems precision iii. Huge calculation time
		SARIMA (seasonal auto-regressive integrated moving average) and ES (Exponential Smoothing) [195]	Predicts yearly consumption of electricity for the agriculture sector	
		ISSA-SVM (improved sparrow search algorithm-Support Vector Machine) [196]	Error index of load forecasting is kept optimal which results in better prediction accuracy	
	2021	LSTM network [197]	Load forecasting with minimal error for industrial power consumption	
		Support Vector Regression (SVR) [198]	Mean absolute percentage error (MAPE) and root mean square error (RMSE) are kept to a minimum	
	2020	BPNN, Singular Spectrum Analysis (SSA), Weightless Neural Network (WNN), Cuckoo Search algorithm [199]	Surveying load forecasting for wavelet disintegration to learn about the reduction of stochastic part	
		Grasshopper Optimization Algorithm, BP, Regressive Model [200]	Daily and hourly continuous load forecasting	
		Load Range Discretization (LRD), CNN, BP [201]	Probability distribution generation for load forecasting	
		Mutual Information-ANN, Jaya algorithm [202]	Removes feature selection redundancy	
		LSTM, Cascade NN, Edited Nearest Neighbor (ENN), Ensemble Learning. Levenberg–Marquardt algorithm [203]	Decreasing mean absolute percentage error by integrating cascade neural network in load forecasting	
		CNN, BP, Image encoding, Gramian Angular field, Recurrence Plot, Markov Transition field [204]	Single residential user load forecasting using CNN on time series datasets	
		DML, Apache Hadoop, Apache Spark, Linear Regression, Generalized Linear Regression, Decision Tree, Gradient-boosted trees, Random Forest, Distributed computing [188]	Development of a Distributed Machine Learning approach for reducing training time and test time with higher accuracy	
	2019	KNN-ANN, BPNN, Spark [205]	Handling multivariate data and multiple time series while predicting load forecasting outputs	
		LSTM, BPNN, Adaptive Moment Estimation [206]	Load forecasting prediction by analyzing electricity price	
		Parallel deep learning [207]	Ensuring control of hybrid energy storing system in a distributed system using parallel deep learning	
		LSTM, GRU [208]	Predicting load forecasting by training GRU and LTSM with various time scale sequences	
	2018	FFNN, Particle Swarm Optimization, MLP [209]	Mid-term load forecasting in terms of green environment and peak load	
		BPNN [210]	Identification of max power load at photovoltaic power generation and power capacity	

in data-enabled asset monitoring can confirm a robust energy grid by eliminating the chance of component failure, improving NGSG integrity, and prolonging its lifetime. However, the development of a techno-economic model for a data-driven NGSG system in terms of operational cost, time consumption, manufacturing cost, and computational efficiency imposes additional challenges which open the following research platforms for further improvement.

1. *Increased robustness in techniques* Future SGs based on multiple renewable energy sources will need to depend on data techniques that satisfy multidisciplinary constraints as the system complexity is increas-

Table 11 LTLF techniques

LF type	Year	Technical approach	Contribution	Challenges
LTLF	2021	Improved ANN model with an Adaptive Backpropagation Algorithm (ABPA) [211]	Fixes deviations between trained datasets and newly collected forecast datasets	1. Randomness 2. Uncertainty of output
		Hybrid Support Vector Regression (HSVR) [212]	Long-term load forecasting for real industrial power consumption in China	
		Feature-fusion-kernel-based Gaussian process model [213]	Converts one dimensional time-series data into multidimensional features to minimize the gap between original datasets and forecasting	
	2020	Takagi–Sugeno model, RFNN, Fuzzy Rules, Nonlinear System, BP [214]	Retaining temperature data from weather stations with LTLF process and holiday feature management	
		FFNN. BPNN [215]	Mean square error reduction for smart grid consisting of low voltage	
		LSTM, ANN [216]	Enhancing system marginal price using ANN	
		DML, Apache Spark, Apache Hadoop, Linear Regression, Generalized Linear Regression, Decision Tree, Random Forest, Gradient-boosted trees, Distributed computing [188]	Single residential user load forecasting using CNN on time series datasets	
		CNN, Mutual Information (MI), MI-ANN, Relief F, Kernel Principal Component Analysis (KPCA), BP [189]	Over-fitting issue reduction and computational time reduction using CNN, KPCA, MI etc.	
	2019	Parallel deep learning, DC-DC converter [207]	Ensuring control of hybrid energy storing system in a distributed system using parallel deep learning	
	2018	Neuro-fuzzy, ANFIS, BPNN, Levenberg–Marquardt algorithm [217]	Effectively predicting long term load forecasting using ANN	
		BPNN [210]	Identification of max power load at photovoltaic power generation and power capacity	

ing gradually as per changing requirements. Failure to satisfy any of the requirements of an NGSG may result in disruption of power generation and transmission, increased operational cost, damaged components, and long blackouts.

2. *Enhanced data preprocessing and handling efficiency* Various circumstances such as climate change, tax, regulation, and economic growth, etc., can affect the supply and demand requirements for energy in the future. This will differentiate the data acquired from an SG which may vary from the previously acquired data. This can cause the data techniques trained on historical datasets to be unable to generate accurate results. The variations of collected data can have new information unknown to the algorithm that may be analyzed after advanced preprocessing with higher efficiency. This will increase the demand for better and quicker preprocessing techniques.
3. *Consideration of local environment* Most established data-driven techniques are trained on the data available at a global scale or in a specific area. However, the data generated at a local level may vary greatly from the global datasets. The data techniques should be flexible without losing robustness on the adaptation to new systems in different environments.
4. *Optimization of management system* The existing management algorithms may need further improvement to ensure the maximum sustainability of an NGSG. The improvement includes the management of every component, supply chain, security, demand response and all other aspects of an NGSG by introducing synchronization and interoperability between them. The data acquired from one sector of an NGSG may be used to improve other sectors. This is a key aspect of future management techniques.
5. *International policy optimization* Research on data-driven technology for SG systems is confined to small solutions for large problems. This is where energy policies can offer flexibility in the research field of DDTs and force an NGSG to pursue and implement data techniques to decrease carbon emission, cost, waste, etc., as well as increasing generation, efficiency, resilience, and overall sustainability.

8 Conclusion

With the advance of technologies, the need for a sustainable and green environment is increasing. As well as increasing the amount of intermittent renewable generation, a data-driven technology may boost the capacity of clean energy sources, like solar, wind, and photovoltaic systems. An NGSG promotes energy-efficient power systems and improves the effectiveness of power consumption and energy sustainability. In this paper, the

Table 12 Fraud characterization techniques

Attack type	Objective	Data driven techniques	Advantages	Disadvantages
Network Integrity and Confidentiality Violation [218]	Intrusion detection	Fuzzy Logic (FL), DAL (Domain-Adversarial Learning) game theory, RL, Data loss prevention, Distributed Network Protocol 3, Public Key Infrastructure (PKI), Transport Layer Security, Secure Sockets Layer	Rarity of sample data and shift of data distribution are handled for properly detecting attacks	Performance falls when transfer is not needed
Malicious attack on voltage stability [219]	Diagnosis	ANN	Ability to process data parallel with high tolerance towards faults	Lower process time often fails to give optimum results
CCDA (Covert cyber deception assault) [220]	Diagnosis	SVM, Isolation Forest	Handles nonlinear data efficiently using SVM technique	Needs large memory with lengthy training time
Theft of Electricity [221]	Detection	CNN, Random Forest	Unique features in training data samples are found in an automated process	A large number of data samples are required to complete analysis
Cyber Attack [222]	Identification	KNN, SVM, SDAE, RL, ANN	It is possible to seamlessly add datasets with existing datasets	Unable to handle larger datasets and sensitive to noisy datasets
Survey on traffic, SSS-IP (Social Engineering canning), Scanning of modbus network [223]	Privacy Conformity, Verification, Encryption	Distributed Network Protocol 3 (DNP3), Public Key Infrastructure (PKI), Transport Layer Security, Secure Sockets Layer, Security information and event management	Multiple modes of operation can be used in a flexible environment	Unable to work with high bandwidth signals
Trojan horse, Virus [224]	Intrusion Detection	Security information and event management, Data loss prevention, AV(Anti-Virus)	Cost is minimal	Noise can drastically reduce system efficiency
DoS (Denial of Service) of AMI [224]	Intrusion Detection, Calculation of Collapsed Transmission, Calculation of Time, Checking strength of Signal	Security information and event management	Ease of deployment	Poor noise-handling capability
Channel Jamming of PMU, HMI Popping in EMS substation and SCADA [224]	Intrusion Detection, Privacy Conformity	AJ (Anti-Jamming), Security information and event management, Data loss prevention, AV(Anti-Virus),	Cost is minimal	Noise can drastically reduce system efficiency
Attack of Masquerade on PLC [225]	Intrusion Detection, Verification, Encryption	Security information and event management, Data loss prevention, Distributed Network Protocol 3, Public Key Infrastructure (PKI), Transport Layer Security, Secure Sockets Layer	Effortless setup	Ineffective noise management skills
Backdoor attack on SCADA [225, 226]	Intrusion Detection	Security information and event management, AV(Anti-Virus)	Ease of deployment	Poor noise-handling capability
Distributed denial of service (DDoS) attack [227]	Identification and defense	Random forest (RF) and Naive Bayes (NB)	Divides the security measures in three different levels for better threat aversion	Self-awareness of smart grids is hampered for huge networks of grid infrastructures

Table 12 (continued)

Attack type	Objective	Data driven techniques	Advantages	Disadvantages
False data injection (FDI) attacks [227]	Anomaly Detection	Unsupervised, Semi-supervised and Supervised ML Approach	Divides the security measures in three different levels for better threat aversion	Self-awareness of smart grids is hampered for huge networks of grid infrastructures
Hidden Cyber Attack [228]	Detection	Dynamic Bayesian network (DBN)	Efficiently uses hidden layer for robust classification	Unable to deal with cyber threats that are evolving dynamically

Table 13 Blockchain-based safe energy trading techniques

Application	Used technical approaches	Description	Type	Advantages	Disadvantages
Transactive energy [230]	Smart Contract	Enabling DERs to trade energy quickly and securely via blockchain-based AMI	Public Blockchain	No intermediaries or brokers needed for agreement confirmation which eliminates third party interference	Difficult to change the error and chance of loopholes
Decentralized Demand Response Regulation Program [231]	Smart Contract, Proof of Stake (PoS), Ethereum platform	A distributed system with blockchain technology to store smart meter data and use it to balance supply and demand	Public Blockchain	High scalability with fast transactions	Privacy security is poor
Transparent and Traceable Energy [232]	Smart Contract Group signature, Edge computing, Covert Channel Authorization Technique, Pseudo names, Voting-based consensus	Using an authorized blockchain in a SG to assure privacy as well as energy security (traceable and transparent energy usage	Permissioned Blockchain	Better response time and less latency between all devices	Large storage capacity is needed which drives up cost
Information on Energy demand and supply [233]	Smart Contract, PoS consensus, Lagrange Relaxation algorithm	A privacy protecting efficient scheduling strategy is used for energy service providers	N/A	Removes different grid constraints	Requires multiple iterations for a good result
Peer-to-peer (P2P) energy trading [234]	E-wallet (Credit-based), Stackelberg Game Theory	Decentralized power trading between producers and consumers and consumers, support for renewable energy production	Consortium Blockchain	Provides strong security for data	Not suitable for large scale energy trading
DCT (Decentralized energy trading) and cost [235]	Proof of Work (PoW), Multi-signature, Anonymous Messaging Streams, Elliptic Curve Digital Signature Algorithm (ECDSA)	Implementation of secure energy exchanges enabled by blockchain and confidentiality strategies for energy price negotiation	N/A	Overcomes transaction limitations and ensures optimal pricing	Extreme scenarios are not tested
SG Energy Trading [236]	Smart Contract, PoS	A blockchain consortium with effective, adaptable, and protected energy trading	Consortium Blockchain	Ensures transaction security without relying on third party	Large scale data replication faces problems
Energy trading in Vehicle to Grid (V2G) configuration [237]	PBFT consensus mechanism, (Elliptic curve cryptography)	An EV incentive program with multilevel authentication scheme enabled by blockchain for energy transactions on V2G networks	Consortium Blockchain	No risk of centralization of data because of increased frequency of transaction	Security implementation for lengthy datasets faces difficulty
Power trading between Electric Vehicles [238]	Smart contracts, Energy Coins	Privacy protection and trading of surplus energy between for EVs	Consortium Blockchain	Faster processing time with reduced cost	The data are immutable throughout the whole process
P2P power trading [239]	Smart contract, Redundant Byzantine Fault Tolerance (RBFT)	An industrial control architecture based on blockchain to guarantee effective data, an ICS BlockOpS system consistency	Permissioned Blockchain	A robust system is used for ensuring increased security	Efficiency falls in the case of closed loop system

Table 13 (continued)

Application	Used technical approaches	Description	Type	Advantages	Disadvantages
Commercial Cyber-Physica System [240]	Access Control Lists (ACL), PoW	Blockchain is used widely for access control list in cost per sale	Private Blockchain	High security with simple mechanism	Requires high energy usage
EV charging providers [241]	Contract theory, Reputation based DBFT consensus, Smart Contract	Blockchain is widely used in EV charging	Permissioned Blockchain	Decreased tolerance towards malicious users strengthens overall security	Multi-operator market is not tested
Safety and Privacy protecting technique [242]	Bloom Filter, Data Aggregation, authentication techniques	User data privacy and protection	Private Blockchain	Proper utilization of space removes false results	Cannot remove or delete input data
Power Production and Distribution [243]	Smart contract, dApps, control of distortion using power electronics devices	Ensure defense from cyber-attacks and abnormality control assessment	Consortium Blockchain	Better adaptability with extra privacy protection for users' data	Difficulty arises in maintenance and scaling

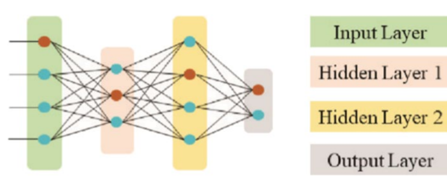
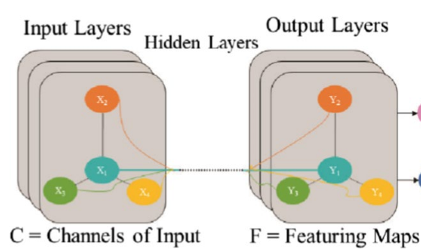
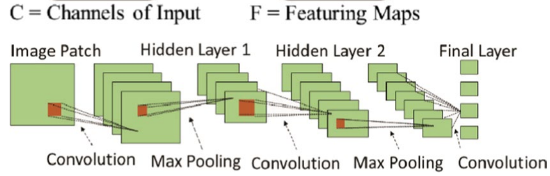
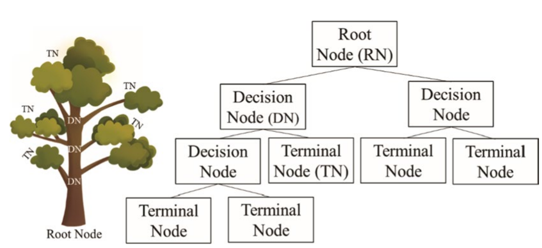
conceptual data-driven NGSG framework for sustainable energy evolution is discussed. The main findings of this paper can be summarized as:

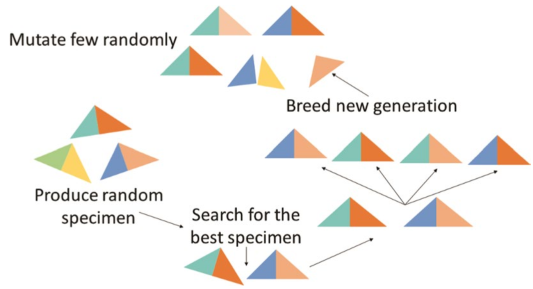
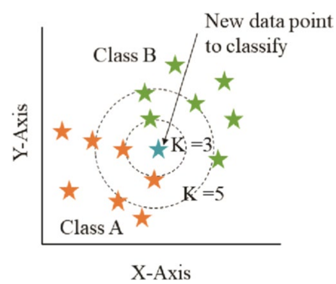
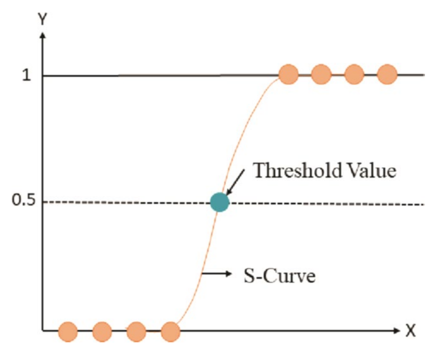
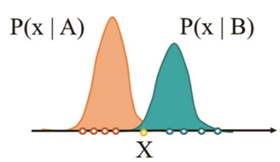
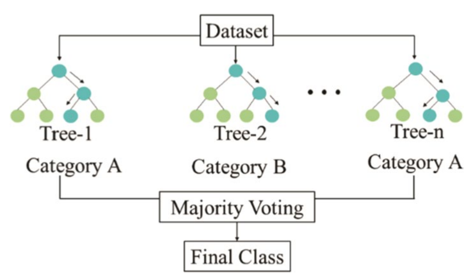
- A comparative study on the conventional SG and NGSG is explicitly done here in terms of their operation and technology. Also, the critical steps to build the data-driven NGSG are also demonstrated and briefly discussed.
- All the intelligent features of a data-driven NGSG are reported and discussed to identify the scope of DDTs.
- Several challenges in initiating the implementation of DDTs are explored and addressed for the growth towards sustainable evolution in an NGSG.
- Advanced DDTs in the conventional SG for the management, condition monitoring, fault prediction, advanced forecasting, and precise fraud characterization are summarized. These lead to the purpose of using DDTs in an NGSG.

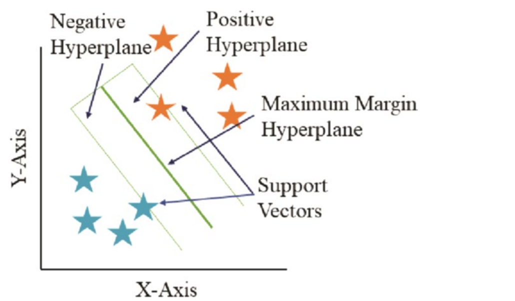
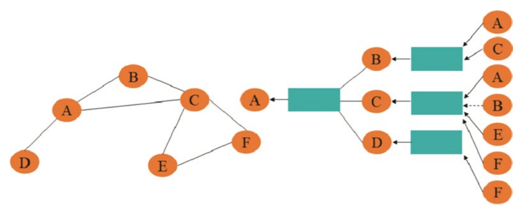
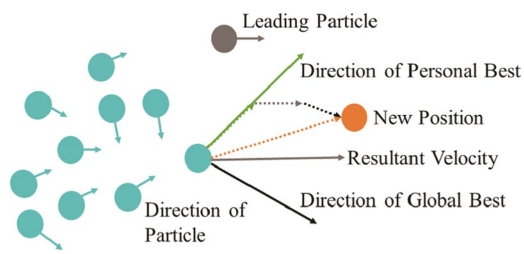
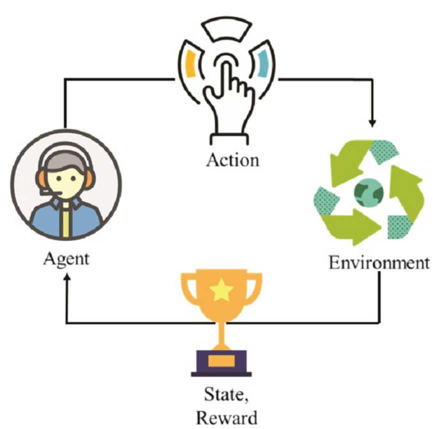
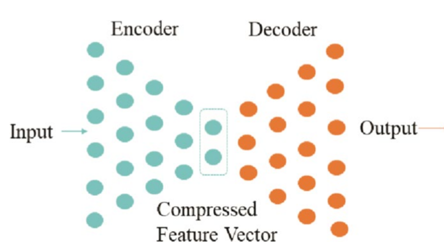
In conclusion, it can be seen that a variety of challenging problems in NGSGs, problems which resist even the most determined efforts of conventional mechanism-based solutions, are successfully resolved by data-driven techniques. These techniques improve NGSG security, increase effectiveness, and reduce dependency on labor and knowledge-intensive human tasks.

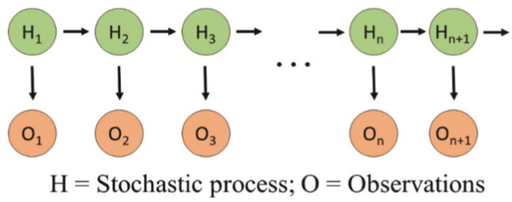
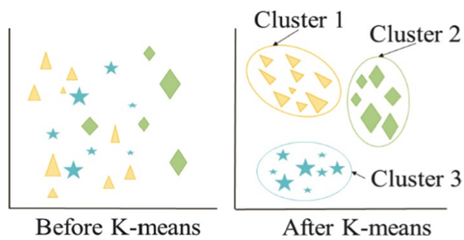
Appendix

Data-driven techniques in SGs

Algorithms	Description	Figure	Applications
Supervised learning			
ANN [67]	ANN is a network consisting of multiple nodes that take input and perform simple functions and send the data to an adjacent node		(1) Load forecasting, (2) Power grid stability assessment, (3) Fault detection, (4) Smart grid security
GCN [68]	GCN, when learning representations from data, takes into consideration the knowledge about the data's structure and generates strong representations; nonetheless, the robustness of the GCN depends on the caliber of the feature matrix and the original graph		Electric parameter identification
CNN [69]	CNN is beneficial when the reduction of parameters is necessary in ANN. An important aspect of CNN is that the problems may not have spatially dependent features		Electricity theft detection
Decision Tree [70]	Decision tree is a tree structured algorithm that is useful for classification and regression. A decision tree consists of three parts: internal nodes, branches and leaves. The dataset attributes are represented in nodes		Small-signal stability analysis

Algorithms	**Description**	**Figure**	**Applications**
Genetic Algorithm [71]	Genetic Algorithm is an efficient searching and meta-heuristic method that replicates the behavior of naturally occurring genetic materials by its selection, mutation and crossover operation		Optimal demand response
KNN [72]	KNN algorithm is a nonparametric classification algorithm based on the proximity of data. The classification method includes Euclidean distance for calculating the nearest neighbors		Power consumption prediction
Logistic Regression [73]	Logistic regression is a method used for linear classification and binary classification problems. Depending on a collection of independent variables, logistic regression calculates the probability of a particular event		Smart grid stability prediction
Naïve Bayes [74]	Naive Bayes algorithm is based on the popular Bayes theorem and is one of the prominent probabilistic robust classification techniques used in machine learning and data analytics		Demand side management
Random Forest [73]	Random Forest algorithm consists of multiple decision trees that are the subsets of the collected data		Smart grid stability prediction

Algorithms	Description	Figure	Applications
SVM [75]	SVM is used for classification as well as regression problems. SVM is popular in the sectors of data mining, machine learning and pattern recognition because of its remarkable generalization ability		Stealthy false data injection detection
Semi-supervised learning			
Graph Neural Network (GNN) [76]	GNN has been proposed as a new deep learning model to learn non-Euclidean material		False data injection attack detection
Q-Learning [77]	Updates are made via bootstrapping in the off policy algorithm known as Q-learning		Vulnerability analysis
Particle Swarm Optimization [78]	The Particle Swarm Optimization technique is easy to implement and use, adaptable, and has a small number of controlling parameters (cognitive ratio, inertia weight, and social ratio)		Energy consumption monitoring
Unsupervised learning			
Deep Autoencoder [79]	Deep Autoencoder consists of two deep belief networks that are symmetrical		Anomaly detection of electricity theft cyberattacks

Algorithms	Description	Figure	Applications
Hidden Markov Model (HMM) [80]	The capacity of HMMs to connect chains of observations with an inherent Markov process—whose unseen states serve as the focus of inference—explains their widespread use. Because HMMs can handle discontinuous time series, such as hourly data, they are particularly well suited for describing and forecasting failures		Islanding prediction
K-means clustering [81]	K-means clustering is the most basic, widely used, and computationally efficient clustering technique This method has been heavily applied in a variety of fields, including the categorization of documents, ride data analysis, in-depth call record analysis, customer classification, criminal network analysis, and others		Privacy preserving

Abbreviations

SG	Smart grid
NGSG	Next-generation smart grid
DDT	Data-driven technique
IoT	Internet of things
ML	Machine learning
RES	Renewable energy sources
PHEV	Plug-in hybrid electric vehicle
ESS	Energy storage system
AMI	Advanced metering infrastructure
PMU	Phasor measurement unit
STLF	Short time load forecasting
VSTLF	Very-short-time load forecasting
MTLF	Medium-time load forecasting
LTLF	Long-time load forecasting

Acknowledgements
Not applicable.

Author contributions
All authors have the equal contribution to prepare, revise and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The data of this study are available on request from the corresponding author.

Declarations

Competing interests
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 article.

Received: 24 October 2022 Accepted: 25 August 2023
Published online: 06 September 2023

References

1. Garg, A., & Mago, V. (2021). Role of machine learning in medical research: A survey. *Computer Science Review, 40*, 100370.
2. Akhter, R., & Sofi, S. A. (2021). Precision agriculture using IoT data analytics and machine learning. *Journal of King Saud University-Computer and Information Sciences, 34*, 5602–5618.
3. Pistoia, M., Ahmad, S. F., Ajagekar, A., Buts, A., Chakrabarti, S., Herman, D., Hu, S., Jena, A., Minssen, P., Niroula, P., & Rattew, A. (2021). Quantum machine learning for finance ICCAD special session paper. In *2021 IEEE/ACM international conference on computer aided design (ICCAD)* (pp. 1–9). IEEE.
4. Schroeder, M., & Lodemann, S. (2021). A systematic investigation of the integration of machine learning into supply chain risk management. *Logistics, 5*(3), 62.
5. Hassan, R., & Radman, G. (2010). Survey on smart grid. In *Proceedings of the IEEE SoutheastCon 2010 (SoutheastCon)* (pp. 210–213). IEEE.
6. Retrieved September 29, 2022, from https://www.mordorintelligence.com/
7. Mollah, M. B., Zhao, J., Niyato, D., Lam, K.-Y., Zhang, X., Ghias, A. M., Koh, L. H., & Yang, L. (2020). Blockchain for future smart grid: A comprehensive survey. *IEEE Internet of Things Journal, 8*(1), 18–43.
8. Ramos, L., Colnago, M., & Casaca, W. (2022). Data-driven analysis and machine learning for energy prediction in distributed photovoltaic generation plants: A case study in Queensland, Australia. *Energy Reports, 8*, 745–751.
9. Tan, K. M., Babu, T. S., Ramachandaramurthy, V. K., Kasinathan, P., Solanki, S. G., & Raveendran, S. K. (2021). Empowering smart grid: A comprehensive review of energy storage technology and application with renewable energy integration. *Journal of Energy Storage, 39*, 102591.
10. Xu, C., Liao, Z., Li, C., Zhou, X., & Xie, R. (2022). Review on interpretable machine learning in smart grid. *Energies, 15*(12), 4427.
11. Dharmadhikari, S. C., Gampala, V., Rao, C. M., Khasim, S., Jain, S., & Bhaskaran, R. (2021). A smart grid incorporated with ml and IoT for a secure management system. *Microprocessors and Microsystems, 83*, 103954.
12. Slama, S. B. (2022). Prosumer in smart grids based on intelligent edge computing: A review on artificial intelligence scheduling techniques. *Ain Shams Engineering Journal, 13*(1), 101504.
13. Qureshi, N. M. F., Siddiqui, I. F., Unar, M. A., Uqaili, M. A., Nam, C. S., Shin, D. R., Kim, J., Bashir, A. K., & Abbas, A. (2019). An aggregate mapreduce

data block placement strategy for wireless IoT edge nodes in smart grid. *Wireless Personal Communications, 106*(4), 2225–2236.

14. Sabadini, F., & Madlener, R. (2023). A smart grid with renewable energy sources, e-vehicles, and storage systems: Operational and economic aspects. In *Smart grids for renewable energy systems, electric vehicles and energy storage systems* (pp. 141–165). CRC Press.
15. Uddin, S. S., Joysoyal, R., Sarker, S. K., Muyeen, S., Ali, M. F., Hasan, M. M., Abhi, S. H., Islam, M. R., Ahamed, M. H., Islam, M. M., & Das, S. K. (2022). Next-generation blockchain enabled smart grid: Conceptual framework, key technologies and industry practices review. *Energy and AI, 12*, 100228.
16. Behara, R. K., & Saha, A. K. (2022). Artificial intelligence methodologies in smart grid-integrated doubly fed induction generator design optimization and reliability assessment: A review. *Energies, 15*(19), 7164.
17. Ponnusamy, V. K., Kasinathan, P., Madurai Elavarasan, R., Ramanathan, V., Anandan, R. K., Subramaniam, U., Ghosh, A., & Hossain, E. (2021). A comprehensive review on sustainable aspects of big data analytics for the smart grid. *Sustainability, 13*(23), 13322.
18. Bhattacharya, S., Chengoden, R., Srivastava, G., Alazab, M., Javed, A. R., Victor, N., Maddikunta, P. K. R., & Gadekallu, T. R. (2022). Incentive mechanisms for smart grid: State of the art, challenges, open issues, future directions. *Big Data and Cognitive Computing, 6*(2), 47.
19. Panda, D. K., & Das, S. (2021). Smart grid architecture model for control, optimization and data analytics of future power networks with more renewable energy. *Journal of Cleaner Production, 301*, 126877.
20. Massaoudi, M., Abu-Rub, H., Refaat, S. S., Chihi, I., & Oueslati, F. S. (2021). Deep learning in smart grid technology: A review of recent advancements and future prospects. *IEEE Access, 9*, 54558–54578.
21. Malik, S. A., Gondal, T. M., Ahmad, S., Adil, M., & Qureshi, R. (2019). Towards optimization approaches in smart grid a review. In *2019 2nd international conference on computing, mathematics and engineering technologies (iCoMET)* (pp. 1–5). IEEE.
22. Shi, Z., Yao, W., Li, Z., Zeng, L., Zhao, Y., Zhang, R., Tang, Y., & Wen, J. (2020). Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions. *Applied Energy, 278*, 115733.
23. Ghosn, S. B., Ranganathan, P., Salem, S., Tang, J., Loegering, D., & Nygard, K. E. (2010). Agent-oriented designs for a self healing smart grid. In *2010 first IEEE international conference on smart grid communications* (pp. 461–466). IEEE.
24. Worighi, I., Maach, A., Hafid, A., Hegazy, O., & Van Mierlo, J. (2019). Integrating renewable energy in smart grid system: Architecture, virtualization and analysis. *Sustainable Energy, Grids and Networks, 18*, 100226.
25. Medora, N. K. (2017). Electric and plug-in hybrid electric vehicles and smart grids. In B. W. D'Andrade (Ed.), *The power grid* (pp. 197–231). Elsevier.
26. Lombardi, P., Powalko, M., & Rudion, K. (2009). Optimal operation of a virtual power plant. In *2009 IEEE power & energy society general meeting* (pp. 1–6). IEEE.
27. Molderink, A., Bakker, V., Bosman, M. G., Hurink, J. L., & Smit, G. J. (2010). Management and control of domestic smart grid technology. *IEEE Transactions on Smart Grid, 1*(2), 109–119.
28. Takuno, T., Koyama, M., & Hikihara, T. (2010). In-home power distribution systems by circuit switching and power packet dispatching. In *2010 first IEEE international conference on smart grid communications* (pp. 427–430). IEEE.
29. Lasseter, R. H. (2011). Smart distribution: Coupled microgrids. *Proceedings of the IEEE, 99*(6), 1074–1082.
30. U. S. F. E. R. Commission. (2008). Federal energy regulatory commission assessment of demand response & advanced metering. Research report 2017, Technical Report.
31. D. G. Photovoltaics and E. Storage. (2011). IEEE guide for smart grid interoperability of energy technology and information technology operation with the electric power system (eps), end-use applications, and loads.
32. Gungor, V. C., Lu, B., & Hancke, G. P. (2010). Opportunities and challenges of wireless sensor networks in smart grid. *IEEE Transactions on Industrial Electronics, 57*(10), 3557–3564.
33. Rusitschka, S., Eger, K., & Gerdes, C. (2010). Smart grid data cloud: A model for utilizing cloud computing in the smart grid domain. In *2010 first IEEE international conference on smart grid communications* (pp. 483–488). IEEE.
34. Gharavi, H., & Hu, B. (2011). Multigate communication network for smart grid. *Proceedings of the IEEE, 99*(6), 1028–1045.
35. Akyol, B. A., Kirkham, H., Clements, S. L., & Hadley, M. D. (2010). A survey of wireless communications for the electric power system. Pacific Northwest National Lab. (PNNL), Richland, WA (United States), Technical Report.
36. Deep, U. D., Petersen, B. R., & Meng, J. (2009). A smart microcontroller-based iridium satellite-communication architecture for a remote renewable energy source. *IEEE Transactions on Power Delivery, 24*(4), 1869–1875.
37. Efthymiou, C., & Kalogridis, G. (2010). Smart grid privacy via anonymization of smart metering data. In *2010 first IEEE international conference on smart grid communications* (pp. 238–243). IEEE.
38. Liu, X., & Xu, W. (2010). Minimum emission dispatch constrained by stochastic wind power availability and cost. *IEEE Transactions on Power Systems, 25*(3), 1705–1713.
39. Anderson, R. N., Boulanger, A., Powell, W. B., & Scott, W. (2011). Adaptive stochastic control for the smart grid. *Proceedings of the IEEE, 99*(6), 1098–1115.
40. Chen, X., Dinh, H., & Wang, B. (2010). Cascading failures in smart grid-benefits of distributed generation. In *2010 first IEEE international conference on smart grid communications* (pp. 73–78). IEEE.
41. Narara ecovillage smart grid. Retrieved September 28, 2022, from https://arena.gov.au/projects/narara-ecovillage-smart-grid/
42. Berrimal wind farm, Australia. Retrieved September 28, 2022, from https://www.power-technology.com/marketdata/berrimal-wind-farm-australia/
43. Mortlake south wind farm. Retrieved September 28, 2022, from https://www.acciona.com.au/projects/mortlake-south-wind-farm/?adin=02021864894
44. Aldoga solar farm. Retrieved September 28, 2022, from https://www.acciona.com.au/projects/aldoga-solar-farm/?adin=02021864894
45. Lilyvale solar pv park, Australia. Retrieved September 28, 2022, from https://www.power-technology.com/marketdata/lilyvale-solar-pv-park-australia/
46. Harapaki wind project. Retrieved September 28, 2022, from https://www.meridianenergy.co.nz/power-stations/wind/harapaki
47. Aep ohio. Retrieved September 28, 2022, from https://aepretirees.com/2014/01/15/aep-ohio-completes-gridsmart-demonstration-project/
48. Project: Detroit Edison. Retrieved September 28, 2022, from https://www.smartgrid.gov/project/detroit-edison-advanced-implementation-energy-storage-technologies
49. Pacific northwest smart grid demonstration project. Retrieved September 28, 2022, from https://www.bpa.gov/energy-and-services/efficiency/smart-grid/pacific-northwest-demo-project
50. The roscoe wind farm project. Retrieved September 28, 2022, from https://www.power-technology.com/projects/roscoe-wind-farm/
51. Retrieved September 28, 2022, from http://mission-innovation.net/our-work/mission-innovation-breakthroughs/glencores-raglan-mine-a-wind-turbine-success/
52. Hebei shahe power plant. Retrieved September 28, 2022, from https://www.power-technology.com/marketdata/hebei-shahe-power-plant-china/
53. Retrieved September 28, 2022, from https://www.hitachienergy.com/case-studies/xiangjiaba---shanghai
54. Retrieved September 28, 2022, from https://indiasmartgrid.org/nsgm.php
55. Retrieved September 28, 2022, from https://en.wikipedia.org/wiki/Jeju-Smart-Grid-Demonstration-Project-in-Korea
56. Retrieved September 28, 2022, from https://www.power-technology.com/marketdata/setana-osato-wind-power-plant-japan/
57. Li, L., Wang, J., Zhong, X., Lin, J., Wu, N., Zhang, Z., Meng, C., Wang, X., Shah, N., Brandon, N., & Xie, S. (2022). Combined multi-objective optimization and agent-based modeling for a 100% renewable island energy system considering power-to-gas technology and extreme weather conditions. *Applied Energy, 308*, 118376.
58. El Zerk, A., Ouassaid, M., & Zidani, Y. (2022). Decentralised strategy for energy management of collaborative microgrids using multi-agent system. *IET Smart Grid, 5*, 440–462.

59. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE internet of things journal, 3*(5), 637–646.
60. O'Neill, D., Levorato, M., Goldsmith, A., & Mitra, U. (2010). Residential demand response using reinforcement learning. In *2010 first IEEE international conference on smart grid communications* (pp. 409–414). IEEE.
61. Logenthiran, T., Srinivasan, D., & Shun, T. Z. (2012). Demand side management in smart grid using heuristic optimization. *IEEE Transactions on Smart Grid, 3*(3), 1244–1252.
62. Daneshvar, M., Mohammadi-Ivatloo, B., & Zare, K. (2022). A novel transactive energy trading model for modernizing energy hubs in the coupled heat and electricity network. *Journal of Cleaner Production, 344*, 131024.
63. Ahmad, T., & Zhang, D. (2021). Using the internet of things in smart energy systems and networks. *Sustainable Cities and Society, 68*, 102783.
64. Guan, L., Chen, H., & Lin, L. (2021). A multi-agent-based self-healing framework considering fault tolerance and automatic restoration for distribution networks. *IEEE Access, 9*, 21522–21531.
65. Subkhankulova, D. (2019). Exploring future opportunities and challenges of demand side management with agent based modelling. Ph.D. dissertation, UCL (University College London).
66. Kusiak, A. (2009). Innovation: A data-driven approach. *International Journal of Production Economics, 122*(1), 440–448.
67. Omitaomu, O. A., & Niu, H. (2021). Artificial intelligence techniques in smart grid: A survey. *Smart Cities, 4*(2), 548–568.
68. Wang, Z., Xia, M., Lu, M., Pan, L., & Liu, J. (2021). Parameter identification in power transmission systems based on graph convolution network. *IEEE Transactions on Power Delivery, 37*, 3155–3163.
69. Lepolesa, L. J., Achari, S., & Cheng, L. (2022). Electricity theft detection in smart grids based on deep neural network. *IEEE Access, 10*, 39638–39655.
70. da Cunha, G. L., Fernandes, R. A., & Fernandes, T. C. C. (2022). Small-signal stability analysis in smart grids: An approach based on distributed decision trees. *Electric Power Systems Research, 203*, 107651.
71. Jeyaranjani, J., & Devaraj, D. (2022). Improved genetic algorithm for optimal demand response in smart grid. *Sustainable Computing: Informatics and Systems, 35*, 100710.
72. Tiwari, S., Jain, A., Ahmed, N. M. O. S., Alkwai, L. M., Dafhalla, A. K. Y., & Hamad, S. A. S. (2022). Machine learning-based model for prediction of power consumption in smart grid-smart way towards smart city. *Expert Systems, 39*(5), e12832.
73. Bashir, A. K., Khan, S., Prabadevi, B., Deepa, N., Alnumay, W. S., Gadekallu, T. R., & Maddikunta, P. K. R. (2021). Comparative analysis of machine learning algorithms for prediction of smart grid stability. *International Transactions on Electrical Energy Systems, 31*(9), e12706.
74. Babar, M., Tariq, M. U., & Jan, M. A. (2020). Secure and resilient demand side management engine using machine learning for IoT-enabled smart grid. *Sustainable Cities and Society, 62*, 102370.
75. Esmalifalak, M., Liu, L., Nguyen, N., Zheng, R., & Han, Z. (2014). Detecting stealthy false data injection using machine learning in smart grid. *IEEE Systems Journal, 11*(3), 1644–1652.
76. Takiddin, A., Atat, R., Ismail, M., Boyaci, O., Davis, K. R., & Serpedin, E. (2023). Generalized graph neural network-based detection of false data injection attacks in smart grids. *IEEE Transactions on Emerging Topics in Computational Intelligence, 7*, 618–630.
77. Yan, J., He, H., Zhong, X., & Tang, Y. (2016). Q-learning-based vulnerability analysis of smart grid against sequential topology attacks. *IEEE Transactions on Information Forensics and Security, 12*(1), 200–210.
78. Xue, X., & Tsai, P.-W. (2022). Integrating energy smart grid's ontologies through multi-objective particle swarm optimization algorithm with competitive mechanism. *Sustainable Energy Technologies and Assessments, 53*, 102442.
79. Takiddin, A., Ismail, M., Zafar, U., & Serpedin, E. (2022). Deep autoencoder-based anomaly detection of electricity theft cyberattacks in smart grids. *IEEE Systems Journal, 16*(3), 4106–4117.
80. Kumar, D., & Bhowmik, P. S. (2019). Hidden Markov model based islanding prediction in smart grids. *IEEE Systems Journal, 13*(4), 4181–4189.
81. Wang, Y., Ma, J., Gao, N., Wen, Q., Sun, L., & Guo, H. (2023). Federated fuzzy k-means for privacy-preserving behavior analysis in smart grids. *Applied Energy, 331*, 120396.
82. Sun, M., Zhang, T., Wang, Y., Strbac, G., & Kang, C. (2019). Using Bayesian deep learning to capture uncertainty for residential net load forecasting. *IEEE Transactions on Power Systems, 35*(1), 188–201.
83. Ahmad, T., Madonski, R., Zhang, D., Huang, C., & Mujeeb, A. (2022). Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm. *Renewable and Sustainable Energy Reviews, 160*, 112128.
84. Li, X., Tian, Y.-C., Ledwich, G., Mishra, Y., Han, X., & Zhou, C. (2018). Constrained optimization of multicast routing for wide area control of smart grid. *IEEE Transactions on Smart Grid, 10*(4), 3801–3808.
85. Wkedzik, A., Siewierski, T., & Szypowski, M. (2019). The use of black-box optimization method for determination of the bus connection capacity in electric power grid. *Energies, 13*(1), 41.
86. Deshpande, P. S., Sharma, S. C., & Peddoju, S. K. (2019). Predictive and prescriptive analytics in big-data era. In *Security and data storage aspect in cloud computing* (pp. 71–81. Springer.
87. Liu, Y., Esan, O. C., Pan, Z., & An, L. (2021). Machine learning for advanced energy materials. *Energy and AI, 3*, 100049.
88. Mahmud, K., Khan, B., Ravishankar, J., Ahmadi, A., & Siano, P. (2020). An internet of energy framework with distributed energy resources, prosumers and small-scale virtual power plants: An overview. *Renewable and Sustainable Energy Reviews, 127*, 109840.
89. López, K. L., Gagné, C., & Gardner, M.-A. (2018). Demandside management using deep learning for smart charging of electric vehicles. *IEEE Transactions on Smart Grid, 10*(3), 2683–2691.
90. Vázquez-Canteli, J. R., & Nagy, Z. (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied energy, 235*, 1072–1089.
91. Kotsiopoulos, T., Sarigiannidis, P., Ioannidis, D., & Tzovaras, D. (2021). Machine learning and deep learning in smart manufacturing: The smart grid paradigm. *Computer Science Review, 40*, 100341.
92. Ahmad, T., Zhang, H., & Yan, B. (2020). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Sustainable Cities and Society, 55*, 102052.
93. Magazzino, C., Mele, M., & Schneider, N. (2021). A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO_2 emissions. *Renewable Energy, 167*, 99–115.
94. Jeyaraj, P. R., & Nadar, E. R. S. (2021). Computer-assisted demand-side energy management in residential smart grid employing novel pooling deep learning algorithm. *International Journal of Energy Research, 45*(5), 7961–7973.
95. O'Driscoll, E., Kelly, K., & O'Donnell, G. E. (2015). Intelligent energy based status identification as a platform for improvement of machine tool efficiency and effectiveness. *Journal of Cleaner Production, 105*, 184–195.
96. Mishra, R. K., Verma, K., Mishra, V., & Chaudhary, B. (2022). A review on carbon-based phase change materials for thermal energy storage. *Journal of Energy Storage, 50*, 104166.
97. Remmen, P., Lauster, M., Mans, M., Fuchs, M., Osterhage, T., & Müller, D. (2018). Teaser: An open tool for urban energy modelling of building stocks. *Journal of Building Performance Simulation, 11*(1), 84–98.
98. Seyedzadeh, S., Rahimian, F. P., Oliver, S., Rodriguez, S., & Glesk, I. (2020). Machine learning modelling for predicting non-domestic buildings energy performance: A model to support deep energy retrofit decision-making. *Applied Energy, 279*, 115908.
99. Javaid, N., Naseem, M., Rasheed, M. B., Mahmood, D., Khan, S. A., Alrajeh, N., & Iqbal, Z. (2017). A new heuristically optimized home energy management controller for smart grid. *Sustainable Cities and Society, 34*, 211–227.
100. Lu, Y., & Zheng, X. (2020). 6G: A survey on technologies, scenarios, challenges, and the related issues. *Journal of Industrial Information Integration, 19*, 100158.
101. Abouheaf, M., Gueaieb, W., & Sharaf, A. (2019). Load frequency regulation for multi-area power system using integral reinforcement learning. *IET Generation, Transmission & Distribution, 13*(19), 4311–4323.
102. Daneshfar, F., & Bevrani, H. (2012). Multiobjective design of load frequency control using genetic algorithms. *International Journal of Electrical Power & Energy Systems, 42*(1), 257–263.

103. Prakash, S., & Sinha, S. (2014). Simulation based neuro-fuzzy hybrid intelligent pi control approach in four-area load frequency control of interconnected power system. *Applied Soft Computing, 23*, 152–164.
104. Shi, Z., Wang, Z., Luo, Y., & Ye, D. (2018). Supplementary frequency control for multi-machine power system based on adaptive dynamic programming. In *International symposium on neural networks* (pp. 677–685). Springer.
105. Tomin, N. V., Kurbatsky, V. G., & Reutsky, I. S. (2019). Hybrid intelligent technique for voltage/VAR control in power systems. *IET Generation, Transmission & Distribution, 13*(20), 4724–4732.
106. Wu, J., Fang, B., Fang, J., Chen, X., & Chi, K. T. (2019). Sequential topology recovery of complex power systems based on reinforcement learning. *Physica A: Statistical Mechanics and its Applications, 535*, 122487.
107. Xiao, L., Li, Y., Liu, J., & Zhao, Y. (2015). Power control with reinforcement learning in cooperative cognitive radio networks against jamming. *The Journal of Supercomputing, 71*(9), 3237–3257.
108. Xi, L., Yu, T., Yang, B., Zhang, X., & Qiu, X. (2016). A wolf pack hunting strategy based virtual tribes control for automatic generation control of smart grid. *Applied Energy, 178*, 198–211.
109. Yin, L., Li, S., & Liu, H. (2020). Lazy reinforcement learning for real-time generation control of parallel cyber–physical–social energy systems. *Engineering Applications of Artificial Intelligence, 88*, 103380.
110. Chatterjee, S., & Mukherjee, V. (2016). PID controller for automatic voltage regulator using teaching–learning based optimization technique. *International Journal of Electrical Power & Energy Systems, 77*, 418–429.
111. Duan, J., Shi, D., Diao, R., Li, H., Wang, Z., Zhang, B., Bian, D., & Yi, Z. (2019). Deep-reinforcement-learning-based autonomous voltage control for power grid operations. *IEEE Transactions on Power Systems, 35*(1), 814–817.
112. Abubakar, I., Khalid, S., Mustafa, M., Shareef, H., & Mustapha, M. (2017). Application of load monitoring in appliances' energy management: A review. *Renewable and Sustainable Energy Reviews, 67*, 235–245.
113. Jiang, X., & Xiao, C. (2019). Household energy demand management strategy based on operating power by genetic algorithm. *IEEE Access, 7*, 96414–96423.
114. Paterakis, N. G., Erdinc, O., Bakirtzis, A. G., & Catalão, J. P. (2015). Optimal household appliances scheduling under day-ahead pricing and load-shaping demand response strategies. *IEEE Transactions on Industrial Informatics, 11*(6), 1509–1519.
115. Ma, K., Yao, T., Yang, J., & Guan, X. (2016). Residential power scheduling for demand response in smart grid. *International Journal of Electrical Power & Energy Systems, 78*, 320–325.
116. Erol-Kantarci, M., & Mouftah, H. T. (2011). Wireless sensor networks for cost-efficient residential energy management in the smart grid. *IEEE Transactions on Smart Grid, 2*(2), 314–325.
117. Hafeez, G., Islam, N., Ali, A., Ahmad, S., Usman, M., & Saleem Alimgeer, K. (2019). A modular framework for optimal load scheduling under price-based demand response scheme in smart grid. *Processes, 7*(8), 499.
118. Hu, M., Xiao, J.-W., Cui, S.-C., & Wang, Y.-W. (2018). Distributed real-time demand response for energy management scheduling in smart grid. *International Journal of Electrical Power & Energy Systems, 99*, 233–245.
119. Khalid, R., Javaid, N., Rahim, M. H., Aslam, S., & Sher, A. (2019). Fuzzy energy management controller and scheduler for smart homes. *Sustainable Computing: Informatics and Systems, 21*, 103–118.
120. Hazarika, K., Katiyar, G., & Islam, N. (2021). IoT based transformer health monitoring system: A survey. In *2021 international conference on advance computing and innovative technologies in engineering (ICACITE)* (pp. 1065–1067). IEEE.
121. Rediansyah, D., & Prasojo, R. A. (2021). Study on artificial intelligence approaches for power transformer health index assessment. In *2021 international conference on electrical engineering and informatics (ICEEI)* (pp. 1–4). IEEE.
122. Perumal, B., Nagarai, P., Venkatesh, R., Muneeswaran, V., GopiShankar, Y., SaiKumar, A., Koushik, A., & Anil, B. (2022). Real time transformer health monitoring system using IoT in r. In *2022 international conference on computer communication and informatics (ICCCI)* (pp. 1–5). IEEE.
123. Samal, P. K., Sharma, V., & Kumar, R. (2022). Condition assessment of transformer health by using intelligent technique. In *2022 IEEE international conference on distributed computing and electrical circuits and electronics (ICDCECE)* (pp. 1–6). IEEE.
124. McGrew, T., Sysoeva, V., Cheng, C.-H., Miller, C., Scofield, J., & Scott, M. J. (2022). Condition monitoring of DC-link capacitors using time–frequency analysis and machine learning classification of conducted EMI. *IEEE Transactions on Power Electronics, 37*(10), 12606–12618.
125. McGrew, T., Sysoeva, V., Cheng, C.-H., & Scott, M. (2021). Condition monitoring of DC-link capacitors using hidden Markov model supported-convolutional neural network. In *2021 IEEE applied power electronics conference and exposition (APEC)* (pp. 2323–2330). IEEE.
126. Berghout, T., Benbouzid, M., Bentrcia, T., Ma, X., Djurović, S., & Mouss, L.-H. (2021). Machine learning-based condition monitoring for PV systems: State of the art and future prospects. *Energies, 14*(19), 6316.
127. Black, I. M., Richmond, M., & Kolios, A. (2021). Condition monitoring systems: A systematic literature review on machine-learning methods improving offshore-wind turbine operational management. *International Journal of Sustainable Energy, 40*(10), 923–946.
128. Xiang, L., Yang, X., Hu, A., Su, H., & Wang, P. (2022). Condition monitoring and anomaly detection of wind turbine based on cascaded and bidirectional deep learning networks. *Applied Energy, 305*, 117925.
129. Chen, P., Li, Y., Wang, K., Zuo, M. J., Heyns, P. S., & Baggeröhr, S. (2021). A threshold self-setting condition monitoring scheme for wind turbine generator bearings based on deep convolutional generative adversarial networks. *Measurement, 167*, 108234.
130. Yang, D., Pang, Y., Zhou, B., & Li, K. (2019). Fault diagnosis for energy internet using correlation processing-based convolutional neural networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49*(8), 1739–1748.
131. Guo, M.-F., Yang, N.-C., & Chen, W.-F. (2019). Deeplearning-based fault classification using Hilbert–Huang transform and convolutional neural network in power distribution systems. *IEEE Sensors Journal, 19*(16), 6905–6913.
132. Chen, K., Hu, J., Zhang, Y., Yu, Z., & He, J. (2019). Fault location in power distribution systems via deep graph convolutional networks. *IEEE Journal on Selected Areas in Communications, 38*(1), 119–131.
133. De Santis, E., Rizzi, A., & Sadeghian, A. (2018). A cluster-based dissimilarity learning approach for localized fault classification in smart grids. *Swarm and Evolutionary Computation, 39*, 267–278.
134. Saleem, Y., Crespi, N., Rehmani, M. H., & Copeland, R. (2019). Internet of things-aided smart grid: Technologies, architectures, applications, prototypes, and future research directions. *IEEE Access, 7*, 62962–63003.
135. Madeti, S. R., & Singh, S. (2017). A comprehensive study on different types of faults and detection techniques for solar photovoltaic system. *Solar Energy, 158*, 161–185.
136. Tokel, H. A., Al Halaseh, R., Alirezaei, G., & Mathar, R. (2018). A new approach for machine learning-based fault detection and classification in power systems. In *2018 IEEE power & energy society innovative smart grid technologies conference (ISGT)* (pp. 1–5). IEEE.
137. Hare, J., Shi, X., Gupta, S., & Bazzi, A. (2016). Fault diagnostics in smart micro-grids: A survey. *Renewable and Sustainable Energy Reviews, 60*, 1114–1124.
138. Dileep, G. (2020). A survey on smart grid technologies and applications. *Renewable Energy, 146*, 2589–2625.
139. Phan, S. K., & Chen, C. (2017). Big data and monitoring the grid. In *The power grid* (pp. 253–285). Elsevier.
140. Ibrahim, M. S., Dong, W., & Yang, Q. (2020). Machine learning driven smart electric power systems: Current trends and new perspectives. *Applied Energy, 272*, 115237.
141. Munshi, A. A., & Yasser, A.-R.M. (2017). Big data framework for analytics in smart grids. *Electric Power Systems Research, 151*, 369–380.
142. Bangalore, P., & Tjernberg, L. B. (2015). An artificial neural network approach for early fault detection of gearbox bearings. *IEEE Transactions on Smart Grid, 6*(2), 980–987.
143. Chen, K., Hu, J., & He, J. (2016). Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder. *IEEE Transactions on Smart Grid, 9*(3), 1748–1758.
144. Sun, H., Wang, Z., Wang, J., Huang, Z., Carrington, N., & Liao, J. (2016). Data-driven power outage detection by social sensors. *IEEE Transactions on Smart Grid, 7*(5), 2516–2524.
145. Martinez-Figueroa, G. D. J., Morinigo-Sotelo, D., Zorita-Lamadrid, A. L., Morales-Velazquez, L., & Romero-Troncoso, R. D. J. (2017). FPGA-based

smart sensor for detection and classification of power quality disturbances using higher order statistics. *IEEE Access, 5*, 14259–14274.
146. Mishra, D. P., Samantaray, S. R., & Joos, G. (2015). A combined wavelet and data-mining based intelligent protection scheme for microgrid. *IEEE Transactions on Smart Grid, 7*(5), 2295–2304.
147. James, J., Hou, Y., Lam, A. Y., & Li, V. O. (2017). Intelligent fault detection scheme for microgrids with waveletbased deep neural networks. *IEEE Transactions on Smart Grid, 10*(2), 1694–1703.
148. Harrou, F., Sun, Y., Taghezouit, B., Saidi, A., & Hamlati, M. E. (2018). Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches. *Renewable Energy, 116*, 22–37.
149. Xi, P., Feilai, P., Yongchao, L., Zhiping, L., & Long, L. (2017). Fault detection algorithm for power distribution network based on sparse self-encoding neural network. In *2017 international conference on smart grid and electrical automation (ICSGEA)* (pp. 9–12). IEEE.
150. Kumar, D., & Bhowmik, P. S. (2018). Artificial neural network and phasor data-based islanding detection in smart grid. *IET Generation, Transmission & Distribution, 12*(21), 5843–5850.
151. Achlerkar, P. D., Samantaray, S. R., & Manikandan, M. S. (2016). Variational mode decomposition and decision tree based detection and classification of power quality disturbances in grid-connected distributed generation system. *IEEE Transactions on Smart Grid, 9*(4), 3122–3132.
152. Pertl, M., Douglass, P. J., Heussen, K., & Kok, K. (2018). Validation of a robust neural real-time voltage estimator for active distribution grids on field data. *Electric Power Systems Research, 154*, 182–192.
153. Jiang, Y., Yin, S., & Kaynak, O. (2018). Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond. *IEEE Access, 6*, 47374–47384.
154. Wu, H., Liu, J., Liu, Y., Qiu, G., & Taylor, G. A. (2017). Power system transmission line fault diagnosis based on combined data analytics. In *2017 IEEE power & energy society general meeting* (pp. 1–5). IEEE.
155. Rahman, M. S., Isherwood, N., & Oo, A. (2018). Multi-agent based coordinated protection systems for distribution feeder fault diagnosis and reconfiguration. *International Journal of Electrical Power & Energy Systems, 97*, 106–119.
156. Al Ridhawi, I., Otoum, S., Aloqaily, M., Jararweh, Y., & Baker, T. (2020). Providing secure and reliable communication for next generation networks in smart cities. *Sustainable Cities and Society, 56*, 102080.
157. Ahmadipour, M., Hizam, H., Othman, M. L., Radzi, M. A. M., & Murthy, A. S. (2018). Islanding detection technique using slantlet transform and ridgelet probabilistic neural network in grid-connected photovoltaic system. *Applied Energy, 231*, 645–659.
158. Ahmadipour, M., Hizam, H., Lutfi Othman, M., & Amran Mohd Radzi, M. (2018). An anti-islanding protection technique using a wavelet packet transform and a probabilistic neural network. *Energies, 11*(10), 2701.
159. Ahmadipour, M., Hizam, H., Othman, M. L., Mohd Radzi, M. A., & Chireh, N. (2019). A fast fault identification in a grid-connected photovoltaic system using wavelet multi-resolution singular spectrum entropy and support vector machine. *Energies, 12*(13), 2508.
160. Howell, S., Rezgui, Y., Hippolyte, J.-L., Jayan, B., & Li, H. (2017). Towards the next generation of smart grids: Semantic and holonic multi-agent management of distributed energy resources. *Renewable and Sustainable Energy Reviews, 77*, 193–214.
161. Tong, C., Li, J., Lang, C., Kong, F., Niu, J., & Rodrigues, J. J. (2018). An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders. *Journal of Parallel and Distributed Computing, 117*, 267–273.
162. Ashwini Kumari, P., & Geethanjali, P. (2020). Artificial neural network-based smart energy meter monitoring and control using global system for mobile communication module. In *Soft computing for problem solving* (pp. 1–8). Springer.
163. Jacob, M., Neves, C., & Vukadinović Greetham, D. (2020). *Forecasting and assessing risk of individual electricity peaks*. Springer.
164. Bashir, T., Haoyong, C., Tahir, M. F., & Liqiang, Z. (2022). Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN. *Energy Reports, 8*, 1678–1686.
165. Ribeiro, A. M. N., do Carmo, P. R. X., Endo, P. T., Rosati, P., & Lynn, T. (2022). Short-and very short-term firm-level load forecasting for warehouses: A comparison of machine learning and deep learning models. *Energies, 15*(3), 750.
166. Alquthami, T., Zulfiqar, M., Kamran, M., Milyani, A. H., & Rasheed, M. B. (2022). A performance comparison of machine learning algorithms for load forecasting in smart grid. *IEEE Access, 10*, 48419–48433.
167. Lu, Y., Wang, G., & Huang, S. (2022). A short-term load forecasting model based on mixup and transfer learning. *Electric Power Systems Research, 207*, 107837.
168. Hu, H., Xia, X., Luo, Y., Zhang, C., Nazir, M. S., & Peng, T. (2022). Development and application of an evolutionary deep learning framework of LSTM based on improved grasshopper optimization algorithm for short-term load forecasting. *Journal of Building Engineering, 57*, 104975.
169. Eskandari, H., Imani, M., & Moghaddam, M. P. (2021). Convolutional and recurrent neural network based model for short-term load forecasting. *Electric Power Systems Research, 195*, 107173.
170. Zhang, W., Chen, Q., Yan, J., Zhang, S., & Xu, J. (2021). A novel asynchronous deep reinforcement learning model with adaptive early forecasting method and reward incentive mechanism for short-term load forecasting. *Energy, 236*, 121492.
171. Rafi, S. H., Deeba, S. R., & Hossain, E. (2021). A short-term load forecasting method using integrated CNN and LSTM network. *IEEE Access, 9*, 32436–32448.
172. Park, R.-J., Song, K.-B., & Kwon, B.-S. (2020). Short-term load forecasting algorithm using a similar day selection method based on reinforcement learning. *Energies, 13*(10), 2640.
173. Hafeez, G., Alimgeer, K. S., & Khan, I. (2020). Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid. *Applied Energy, 269*, 114915.
174. Müller, M., Gaio, G., Carreno, E., Lotufo, A., & Teixeira, L. (2020). Electrical load forecasting in disaggregated levels using fuzzy ARTMAP artificial neural network and noise removal by singular spectrum analysis. *SN Applied Sciences, 2*(7), 1–10.
175. Li, J., Deng, D., Zhao, J., Cai, D., Hu, W., Zhang, M., & Huang, Q. (2020). A novel hybrid short-term load forecasting method of smart grid using MLR and LSTM neural network. *IEEE Transactions on Industrial Informatics, 17*(4), 2443–2452.
176. Aly, H. H. (2020). A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid. *Electric Power Systems Research, 182*, 106191.
177. Ahmad, W., Ayub, N., Ali, T., Irfan, M., Awais, M., Shiraz, M., & Glowacz, A. (2020). Towards short term electricity load forecasting using improved support vector machine and extreme learning machine. *Energies, 13*(11), 2907.
178. Zhu, R., Guo, W., & Gong, X. (2019). Short-term load forecasting for CCHP systems considering the correlation between heating, gas and electrical loads based on deep learning. *Energies, 12*(17), 3308.
179. Zahid, M., Ahmed, F., Javaid, N., Abbasi, R. A., Zainab Kazmi, H. S., Javaid, A., Bilal, M., Akbar, M., & Ilahi, M. (2019). Electricity price and load forecasting using enhanced convolutional neural network and enhanced support vector regression in smart grids. *Electronics, 8*(2), 122.
180. Pramono, S. H., Rohmatillah, M., Maulana, E., Hasanah, R. N., & Hario, F. (2019). Deep learning-based short-term load forecasting for supporting demand response program in hybrid energy system. *Energies, 12*(17), 3359.
181. Yang, Y., Hong, W., & Li, S. (2019). Deep ensemble learning based probabilistic load forecasting in smart grids. *Energy, 189*, 116324.
182. Liu, P., Zheng, P., & Chen, Z. (2019). Deep learning with stacked denoising auto-encoder for short-term electric load forecasting. *Energies, 12*(12), 2445.
183. Sadaei, H. J., e-Silva, P. C. L., Guimaraes, F. G., & Lee, M. H. (2019). Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. *Energy, 175*, 365–377.
184. Velasco, L. C. P., Arnejo, K. A. S., & Macarat, J. S. S. (2022). Performance analysis of artificial neural network models for hour-ahead electric load forecasting. *Procedia Computer Science, 197*, 16–24.
185. Munkhammar, J., van der Meer, D., & Widén, J. (2021). Very short term load forecasting of residential electricity consumption using the Markov-chain mixture distribution (mcm) model. *Applied Energy, 282*, 116180.
186. Yundra, E., Surabaya, U. N., Kartini, U., Wardani, L., Ardianto, D., Surabaya, U., Surabaya, U., & Surabaya, U. (2020). Hybrid model combined fuzzy multiobjective decision making with feed forward neural network

(F-MODMFFNN) for very short-term load forecasting based on weather data. *International Journal of Intelligent Engineering & Systems, 13*(4), 182–195.
187. Wen, L., Zhou, K., & Yang, S. (2020). Load demand forecasting of residential buildings using a deep learning model. *Electric Power Systems Research, 179*, 106073.
188. Syed, D., Refaat, S. S., & Abu-Rub, H. (2020). Performance evaluation of distributed machine learning for load forecasting in smart grids. In *2020* cybernetics *& informatics (K&I)* (pp. 1–6). IEEE.
189. Adil, M., Javaid, N., Daood, N., Asim, M., Ullah, I., & Bilal, M. (2020). Big data based electricity price forecasting using enhanced convolutional neural network in the smart grid. In *Workshops of the international conference on advanced information networking and applications* (pp. 1189–1201). Springer.
190. Dagdougui, H., Bagheri, F., Le, H., & Dessaint, L. (2019). Neural network model for short-term and very-short-term load forecasting in district buildings. *Energy and Buildings, 203*, 109408.
191. Shi, T., Mei, F., Lu, J., Lu, J., Pan, Y., Zhou, C., Wu, J., & Zheng, J. (2019). Phase space reconstruction algorithm and deep learning-based very short-term bus load forecasting. *Energies, 12*(22), 4349.
192. Kartini, U. T., Ardianto, D., & Wardani, L. (2019). Very short term load forecasting based on meteorological with modelling k-NN-feed forward neural network. *Journal of Electrical Systems, 15*(1), 1–16.
193. Semero, Y. K., Zhang, J., Zheng, D., & Wei, D. (2018). An accurate very short-term electric load forecasting model with binary genetic algorithm based feature selection for microgrid applications. *Electric Power Components and Systems, 46*(14–15), 1570–1579.
194. Matrenin, P., Safaraliev, M., Dmitriev, S., Kokin, S., Ghulomzoda, A., & Mitrofanov, S. (2022). Medium-term load forecasting in isolated power systems based on ensemble machine learning models. *Energy Reports, 8*, 612–618.
195. Sharma, M., Mittal, N., Mishra, A., & Gupta, A. (2022). Analytical machine learning for medium-term load forecasting towards agricultural sector. In *Proceedings of second doctoral symposium on computational intelligence* (pp. 581–592). Springer.
196. Li, J., Lei, Y., & Yang, S. (2022). Mid-long term load forecasting model based on support vector machine optimized by improved sparrow search algorithm. *Energy Reports, 8*, 491–497.
197. Jiang, Y., Huang, Q., Zhang, K., Lin, Z., Zhang, T., Hu, X., Liu, S., Jiang, C., Yang, L., & Lin, Z. (2021). Medium-long term load forecasting method considering industry correlation for power management. *Energy Reports, 7*, 1231–1238.
198. Rai, S., & De, M. (2021). Analysis of classical and machine learning based short-term and mid-term load forecasting for smart grid. *International Journal of Sustainable Energy, 40*(9), 821–839.
199. Yuan, Z., Wang, W., Wang, H., & Mizzi, S. (2020). Combination of cuckoo search and wavelet neural network for midterm building energy forecast. *Energy, 202*, 117728.
200. Talaat, M., Farahat, M., Mansour, N., & Hatata, A. (2020). Load forecasting based on grasshopper optimization and a multilayer feed-forward neural network using regressive approach. *Energy, 196*, 117087.
201. Huang, Q., Li, J., & Zhu, M. (2020). An improved convolutional neural network with load range discretization for probabilistic load forecasting. *Energy, 203*, 117902.
202. Samuel, O., Alzahrani, F. A., Hussen Khan, R. J. U., Farooq, H., Shafiq, M., Afzal, M. K., & Javaid, N. (2020). Towards modified entropy mutual information feature selection to forecast medium-term load using a deep learning model in smart homes. *Entropy, 22*(1), 68.
203. Wang, L., Mao, S., Wilamowski, B. M., & Nelms, R. (2020). Ensemble learning for load forecasting. *IEEE Transactions on Green Communications and Networking, 4*(2), 616–628.
204. Estebsari, A., & Rajabi, R. (2020). Single residential load forecasting using deep learning and image encoding techniques. *Electronics, 9*(1), 68.
205. Talavera-Llames, R., Pérez-Chacón, R., Troncoso, A., & Martínez-Álvarez, F. (2019). MV-kWNN: A novel multivariate and multi-output weighted nearest neighbours algorithm for big data time series forecasting. *Neurocomputing, 353*, 56–73.
206. Mujeeb, S., Javaid, N., Ilahi, M., Wadud, Z., Ishmanov, F., & Afzal, M. K. (2019). Deep long short-term memory: A new price and load forecasting scheme for big data in smart cities. *Sustainability, 11*(4), 987.
207. Yang, S., Wu, J., Qin, H., Xie, Q., Xu, Z., & Hua, Y. (2021). Distributed buildings energy storage charging load forecasting method considering parallel deep learning model. *Concurrency and Computation: Practice and Experience, 33*(12), e5580.
208. Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. (2019). Single and multi-sequence deep learning models for short and medium term electric load forecasting. *Energies, 12*(1), 149.
209. Heydari, A., Keynia, F., Garcia, D. A., De Santoli, L. (2018). Mid-term load power forecasting considering environment emission using a hybrid intelligent approach. In *2018 5th international symposium on environment friendly energies and applications (EFEA)* (pp. 1–5). IEEE.
210. Song, Y., Chen, H., Yuan, K., Sun, C., Xue, Z., Jin, X., Liu, W., & Han, J. (2018). Medium and long term load forecasting considering the uncertainty of distributed installed capacity of photovoltaic generation. In *2018 13th IEEE conference on industrial electronics and applications (ICIEA)* (pp. 1691–1696). IEEE.
211. Mohammed, N. A., & Al-Bazi, A. (2022). An adaptive backpropagation algorithm for long-term electricity load forecasting. *Neural Computing and Applications, 34*(1), 477–491.
212. Wang, Z., Zhou, X., Tian, J., & Huang, T. (2021). Hierarchical parameter optimization based support vector regression for power load forecasting. *Sustainable Cities and Society, 71*, 102937.
213. Guan, Y., Li, D., Xue, S., & Xi, Y. (2021). Feature-fusion-kernel-based Gaussian process model for probabilistic long-term load forecasting. *Neurocomputing, 426*, 174–184.
214. Wen, Z., Xie, L., Fan, Q., & Feng, H. (2020). Long term electric load forecasting based on TS-type recurrent fuzzy neural network model. *Electric Power Systems Research, 179*, 106106.
215. Masoumi, A., Jabari, F., Ghassem Zadeh, S., & Mohammadi-Ivatloo, B. (2020). Long-term load forecasting approach using dynamic feed-forward back-propagation artificial neural network. In *Optimization of power system problems* (pp. 233–257). Springer.
216. Yudantaka, K., Kim, J.-S., & Song, H. (2019). Dual deep learning networks based load forecasting with partial real-time information and its application to system marginal price prediction. *Energies, 13*(1), 148.
217. Ammar, N., Sulaiman, M., & Nor, A. F. M. (2018). Long-term load forecasting of power systems using artificial neural network and ANFIS. *ARPN Journal of Engineering and Applied Sciences, 13*(3), 828–834.
218. Zhang, Y., & Yan, J. (2020). Semi-supervised domain adversarial training for intrusion detection against false data injection in the smart grid. In *2020 international joint conference on neural networks (IJCNN)* (pp. 1–7). IEEE.
219. Kosek, A. M. (2016). Contextual anomaly detection for cyber-physical security in smart grids based on an artificial neural network model. In *2016 joint workshop on cyber-physical security and resilience in smart grids (CPSR-SG)* (pp. 1–6). IEEE.
220. Ahmed, S., Lee, Y., Hyun, S.-H., & Koo, I. (2019). Unsupervised machine learning-based detection of covert data integrity assault in smart grid networks utilizing isolation forest. *IEEE Transactions on Information Forensics and Security, 14*(10), 2765–2777.
221. Li, S., Han, Y., Yao, X., Yingchen, S., Wang, J., & Zhao, Q. (2019). Electricity theft detection in power grids with deep learning and random forests. *Journal of Electrical and Computer Engineering, 2019*, 1–12.
222. Haghnegahdar, L., & Wang, Y. (2020). A whale optimization algorithm-trained artificial neural network for smart grid cyber intrusion detection. *Neural computing and applications, 32*(13), 9427–9441.
223. Rawat, D. B., & Bajracharya, C. (2015). Cyber security for smart grid systems: Status, challenges and perspectives. *SoutheastCon, 2015*, 1–6.
224. Knapp, E. D., & Samani, R. (2013). *Applied cyber security and the smart grid: Implementing security controls into the modern power infrastructure*. Newnes.
225. Faisal, M. A., Aung, Z., Williams, J. R., & Sanchez, A. (2014). Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study. *IEEE Systems journal, 9*(1), 31–44.
226. Hashemi, S., & Zarei, M. (2021). Internet of things backdoors: Resource management issues, security challenges, and detection methods. *Transactions on Emerging Telecommunications Technologies, 32*(2), e4142.
227. Tufail, S., Parvez, I., Batool, S., & Sarwat, A. (2021). A survey on cybersecurity challenges, detection, and mitigation techniques for the smart grid. *Energies, 14*(18), 5894.

228. Karimipour, H., Dehghantanha, A., Parizi, R. M., Choo, K. K. R., & Leung, H. (2019). A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids. *IEEE Access, 7*, 80778–80788.
229. Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*, 21260.
230. Mylrea, M., & Gourisetti, S. N. G. (2017). Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 resilience week (RWS)* (pp. 18–23). IEEE.
231. Pop, C., Cioara, T., Antal, M., Anghel, I., Salomie, I., & Bertoncini, M. (2018). Blockchain based decentralized management of demand response programs in smart energy grids. *Sensors, 18*(1), 162.
232. Gai, K., Wu, Y., Zhu, L., Xu, L., & Zhang, Y. (2019). Permissioned blockchain and edge computing empowered privacy-preserving smart grid networks. *IEEE Internet of Things Journal, 6*(5), 7992–8004.
233. Tan, S., Wang, X., & Jiang, C. (2019). Privacy-preserving energy scheduling for ESCOs based on energy blockchain network. *Energies, 12*(8), 1530.
234. Abdella, J., & Shuaib, K. (2018). Peer to peer distributed energy trading in smart grids: A survey. *Energies, 11*(6), 1560.
235. Li, Z., Kang, J., Yu, R., Ye, D., Deng, Q., & Zhang, Y. (2017). Consortium blockchain for secure energy trading in industrial internet of things. *IEEE Transactions on Industrial Informatics, 14*(8), 3690–3700.
236. Aitzhan, N. Z., & Svetinovic, D. (2016). Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. *IEEE Transactions on Dependable and Secure Computing, 15*(5), 840–852.
237. Zheng, D., Deng, K., Zhang, Y., Zhao, J., Zheng, X., & Ma, X. (2018). Smart grid power trading based on consortium blockchain in internet of things. In *International conference on algorithms and architectures for parallel processing* (pp. 453–459). Springer.
238. Kang, J., Yu, R., Huang, X., Maharjan, S., Zhang, Y., & Hossain, E. (2017). Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains. *IEEE Transactions on Industrial Informatics, 13*(6), 3154–3164.
239. Garg, S., Kaur, K., Kaddoum, G., Gagnon, F., & Rodrigues, J. J. (2019). An efficient blockchain-based hierarchical authentication mechanism for energy trading in v2g environment. In *2019 IEEE international conference on communications workshops (ICC workshops)* (pp. 1–6). IEEE.
240. Wan, J., Li, J., Imran, M., & Li, D. (2019). A blockchainbased solution for enhancing security and privacy in smart factory. *IEEE Transactions on Industrial Informatics, 15*(6), 3652–3660.
241. Su, Z., Wang, Y., Xu, Q., Fei, M., Tian, Y.-C., & Zhang, N. (2018). A secure charging scheme for electric vehicles with smart communities in energy blockchain. *IEEE Internet of Things Journal, 6*(3), 4601–4613.
242. Wang, Y., Luo, F., Dong, Z., Tong, Z., & Qiao, Y. (2019). Distributed meter data aggregation framework based on blockchain and homomorphic encryption. *IET Cyber-Physical Systems: Theory & Applications, 4*(1), 30–37.
243. Singh, K., & Choube, S. (2018). Using blockchain against cyber-attacks on smart grids. In *2018 IEEE international students' conference on electrical, electronics and computer science (SCEECS)* (pp. 1–4). IEEE.