The Use of Big Data Technology in Distributed Photovoltaic Power Generation Operation and Management--Take China as an Example

Tingxu Pu [1,*], Yitong Liu [2], Xina Leng [1], Chuanyong Pu [3]

[1] School of Electrical and Electronic Information, Xihua University, Chengdu, China

[2] School of Economics, Xihua University, Chengdu, China

[3] Power Supply Company of Leshan, State Grid Corporation of China, Leshan, China

* Corresponding Author Email: putting0282@foxmail.com

Abstract. Given the lack of distributed PV power generation operation and management capability, this paper profoundly analyzes the current situation of the application of big data technology in the process and management of PV power generation in China and proves the necessity and importance of promoting new equipment, new technology, and new mode based on traditional power system operation and management, using the extensive data information formed in the process and management of power plants, meteorology, environment, and power grid, and launching ample data analysis research. In addition, the typical application scenarios of big data technology in the operation and management of distributed PV power generation are explored, and a new mode of operation and management of distributed PV power generation at the power plant, grid, and government levels is explored and constructed. The study shows that with the dramatic increase in the number of distributed PV power generation, the use of big data technology in scenarios such as the power generation side and grid side has excellent potential to provide adequate support for PV power generation operation, grid scheduling, market-based transaction settlement, and government decision-making.

Keywords: Big data, photovoltaic power generation, operation management, marketing.

1. Introduction

Photovoltaic power generation technology is one of the most sustainable renewable energy technologies; after years of development, it has now been widely used in some domestic areas rich in solar energy resources, the rapid growth of installed PV power generation capacity resulting in the increasing amount of data within PV power generation, the daily data collection volume shows a blowout growth. In China, for example, according to the National Energy Administration statistics [1], the annual installed PV power generation capacity of 54.88GW in 2021, of which 29.28GW of distributed PV capacity, accounting for about 53.4% of all newly installed PV power generation, for the first time more than centralized power plants. In the first three quarters of 2022, PV power generation added 52.06GW of grid-connected capacity, including 17.27GW of centralized PV power plants and 35.33GW of distributed PV. The cumulative installed capacity of distributed PV power plants in China has also increased year by year, from 29.66GW in 2017 to 107.51GW in 2021, when the cumulative installed capacity of domestic distributed PV exceeded 100GW. In the first three quarters of 2022, the country's cumulative installed PV capacity of 358.07GW, including 215.64GW of centralized PV power plants and 142.43GW of distributed PV. the first three quarters of 2022, the country's new grid-connected capacity of installed PV power generation by the province in Hebei, Zhejiang, Shandong and other regions saw the most growth.

As the power industry continues to undergo innovative reforms, technologies such as the Internet of Things, blockchain, and artificial intelligence continue to integrate with the power system [2-5]. There are still many problems in the production O&M management for PV generation and grid-connected processes.

(1) For campus-type PV power plants, problems such as data monitoring are not perfect, operation and maintenance management efficiency needs to be improved, and comprehensive assessment methods for power plants need to be improved.

(2) For small distributed PV, its operation characteristics will significantly impact the distribution network system's operation, monitoring, control, protection, and power quality.

(3) Power grid enterprises also have problems such as insufficient monitoring of PV power generation data and insufficient operation monitoring capability.

Therefore, both power plant operators and grid enterprises should effectively use the information technology means of extensive data analysis, fully exploit the massive monitoring data formed by PV power generation and grid operation, build effective algorithm models, accurately monitor and control distributed PV power generation equipment and grid operation status through extensive data technology analysis and application, ensure operational capability, and guarantee the safe, reliable and stable operation of distributed PV power generation grid connection.

However, most enterprises do not pay enough attention to the operation and management of distributed photovoltaic power generation, and the research on digital technology of energy systems based on big data technology and the construction of intelligent management systems are lagging, which are not enough to support the requirements of high-speed development of photovoltaic industry in the future. Given the above problems, this paper analyzes big data technology's development status and application scenarios in photovoltaic power generation and summarizes and outlooks it accordingly.

2. Distributed photovoltaic power generation and big data

2.1. Basic information

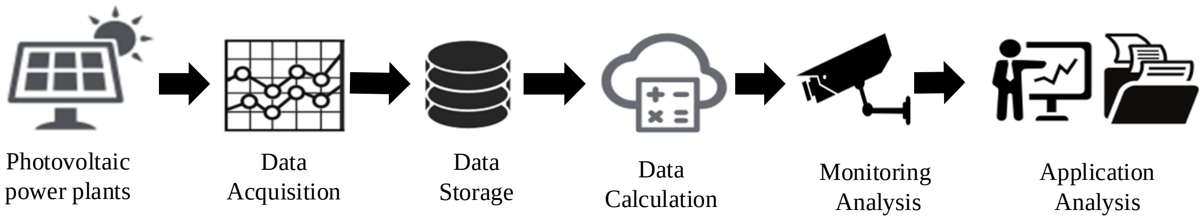

Figure 1. Extensive data platform application process

The PV industry involves a wide range of data, including (1) upstream and downstream industry chains such as the crystalline silicon industry, PV modules, power equipment, information and communication, energy storage equipment, etc.; (2) the amount of data generated by PV power generation is significant, in addition to the power plant's equipment parameters and operational data generated by power generation, there are also power metering and collection data generated by grid enterprises on and off the grid, as well as grid operational data.

The above situation shows that the distributed PV power generation industry is suitable for analysis and research by applying big data technology. With the iterative update of collection terminals and collection systems, as well as the improvement of equipment level and control technology, the massive data generated in various ways and dimensions, such as power plant operation data, power plant equipment monitoring data, power grid operation data, and the value and benefits that may be created from them, lay the foundation for the application of big data in distributed power.

According to the literature, we total the conventional big data platform into data collection, storage, computation, monitoring and control, and analysis and application (as shown in Figure 1).

For large and medium-sized photovoltaic power plants, generally installed in situ with the station safety operation of the monitoring system, used to monitor the station string line, converter box, inverter, transformer, circuit breaker, and other equipment operating status, on the one hand, these monitoring systems, although part of the information interaction with the grid dispatch. Still, only to meet the daily operation of the power station management, in-depth data mining, analysis and application did not achieve better results. On the other hand, the monitoring system of the power station needs to collect a variety of data, including analog, status, pulse, and environmental information, etc. and the amount of data is extensive, and no scientific method to form effective results. For household and general commercial and industrial types of small distributed photovoltaic power supply, because the amount of investment in the capital, workforce, and time is significant, the power grid enterprises or local governments also do not have a more mature, centralized and unified monitoring platform to monitor the PV power generation in the region, not to mention the analysis and research of PV big data.

Based on this, there is an urgent need to build a set of practical tools for monitoring, operation and maintenance, control, and analysis of regional PV power generation through power plant big data or even multi-system fusion big data technology to provide adequate support for PV power generation operation, grid dispatching, market-based trade settlement and government decision-making.

2.2. Foreign PV big data use

Comprehensive literature [6] to [9], we found that: since the 1990s, the United States, Japan, Germany, and other countries have formulated PV power generation plans and invested heavily in promoting the PV power generation market and PV industry [10]. At the same time, they proceeded to establish demonstration projects and bases, researched microgrid structure, control methods, relay protection, market-based transaction settlement, and other related issues, and achieved specific results. All these studies are based on analyzing and processing a large amount of data and model construction.

2.3. Domestic PV extensive data application status

2021 is the strongest year for supporting the development of distributed power in China, with relevant government departments releasing a series of policy documents to support the development of PV. The PV power generation industry is an important part of China's new energy development strategy. At the same time, the data generated by PV power generation has grown exponentially with the installed PV capacity. Grid companies, government departments, power plant owners, financial institutions and other industry players need accurate, complete and timely data as business support, yet the current PV power generation has problems such as poor quality of monitoring data, low level of digital operation and maintenance, low accuracy of quantitative assessment, and prominent information barriers between players, which affect the technical progress of the PV power generation industry and the healthy development of the PV industry.

Currently China's domestic research institutions, power grid companies and other units have carried out and practiced some representative projects one after another. The State Grid Ningxia Electric Power Research Institute and China Electric Power Research Institute Co., Ltd. focus on key technologies such as PV power plant monitoring and prediction, PV power plant operation and maintenance decision making, and PV power plant operation online evaluation, with the integration of massive data from multiple sources as the core, and analysis, integration and publication as the main line. It carries out the design and construction of service systems for all relevant social entities in the PV industry. State Grid Electric also uses the new generation of information technology represented by big data to empower the development of the PV poverty alleviation industry, fully exploiting the massive monitoring data of PV poverty alleviation power stations over the years to inject new momentum into the poverty alleviation management work in less developed rural areas, and has now fully supported 4,431 management users at all levels of poverty alleviation management departments, governments at all levels and power companies. The application of photovoltaic big data is of great significance to China in ensuring energy security, coping with climate change and achieving energy saving and emission reduction.

3. Typical application scenarios of big data technology in distributed photovoltaic power generation operation and management

3.1. Application scenarios of big data technology on the power generation side

3.1.1 Power generation forecasting

Using big data deep learning algorithms, we construct power generation prediction models for power plants to achieve annual, monthly, daily, hourly and minute power generation forecasts for power plants. The general approach is to refer to historical meteorological information, historical power generation information, and other data by collecting primary data such as string mode (PV array), quantity and parameters of PV modules, as well as inverter equipment parameters and equipment operation status, and data provided by weather forecasting system (such as sunlight, temperature, humidity, wind, pollution, dust accumulation, sunset and sunrise time and location, etc.), with year, month, day, hour and minute (e.g., every 15 minutes) as a period, and construct a PV power generation forecasting model using time series (e.g., ARMA), machine learning and deep learning algorithms (e.g., BP neural network and its optimization algorithm), etc. to forecast power generation. At the same time, it is also possible to set the power generation warning threshold value based on this power generation measurement value, compare the real-time power generation calculated from the metering data collected by the collection terminal with it, conduct power generation warning, comprehensively study and judge the causes of abnormalities, carry out operation and maintenance processing in real-time, and improve power generation efficiency [11-13]. In addition, real-time data collected by real-time light intensity (e.g., radiometer, etc.) collectors, temperature collectors, and equipment operation status collectors can be used to compare the power generation of each square array of the power station horizontally to achieve real-time state assessment of the power generation capacity of each square array and the whole station.

3.1.2 Operation and maintenance monitoring

Operation and maintenance monitoring mainly monitor the status and fault of PV power generation equipment to realize intelligent operation and maintenance, which applies to large and medium-sized centralized PV power plants. Photovoltaic power generation equipment has its life cycle; with the aging of electronic components and electrical equipment such as photovoltaic modules, convergence boxes, cables, inverters, transformers, and switchgear, the reliability will gradually decrease. The data collected by various terminals of the PV power stations are complicated. Some data have no substantial effect on power station operation and maintenance monitoring, so extracting the critical data useful for PV operation and maintenance work from these massive data in a targeted manner for analysis is necessary. These vital data include string current, voltage, active power, reactive power, time, analog quantities, irradiance data, ambient temperature data, wind speed, wind direction, dust, and different environmental information quantities. The general practice is to select these big data, build a data processing and calculation model, form the theoretical operation effect that should be achieved, measure the theoretical calculation value, and then compare and analyze with the real-time collection data to determine whether the power generation capacity reaches the target value through the difference between the theoretical calculation and the actual value, to discover the equipment defects of abnormal operation and carry out the elimination treatment in time. At the same time, according to the accumulated historical big data formed by operation, maintenance and fault, relying on the auxiliary analysis and algorithm, the research is equipped with artificial intelligence algorithm and image recognition technology to assist drones' intelligent inspection, intelligent cleaning robot, Beidou satellite positioning and other in-depth integration technology to improve the level of intelligent operation and maintenance of power stations.

3.1.3 Fault prediction

The power generation and daily power generation of the power station are analyzed year-on-year and year-over-year according to the time series. The algorithm is analyzed by combining the probability and mathematical statistics principles. Then, the warning is carried out at different levels.

3.1.4 Comprehensive assessment of power plant operation

The production and operation indexes of PV power plants generally include operation and maintenance evaluation indexes, solar resource indexes, power indexes, energy consumption indexes, equipment operation level indexes, overall system efficiency, power loss rate, power generation performance evaluation, and operation efficiency indexes. These indicators require operators to comprehensively assess power plant generation quality and efficiency from different dimensions, professions, and businesses. It is possible to explore extensive data analysis through multi-disciplinary and multi-business integration to improve power plant operation and management level; for example, for power index, it is possible to carry out analysis of string power generation variance, analysis of square array power generation variance, analysis of inverter power generation variance, analysis of power generation loss, analysis of power loss from dispatching power restriction and abandonment, etc.

3.2. Application of big data technology on the grid side

In addition to large and medium-sized centralized photovoltaic power generation, more non-centralized distributed photovoltaic power generation users are connected to the distribution grid through low-voltage methods, which are large in number and more difficult to monitor. These power points are more dispersed, with small individual capacities, and their power output is greatly affected by environmental factors, low energy density, poor stability, and poor regulation ability. These characteristics will bring a more complex impact on distribution network operation, such as power quality, tide distribution, system protection, power islanding, operation control, distribution network design, planning, and function. In addition, after many distributed PV power sources are connected to the distribution grid, users also participate in the main body of market-based transactions and settlements, changing the traditional fee calculation mode and settlement methods considerably. So far, the power grid enterprises do not have a more mature data system to centralize and unify the management of PV power generation customers in their jurisdiction. As the number of distributed photovoltaic power generation continues to grow, it will bring many challenges to the safe and stable operation of the entire power system and power marketing management. Then, it is essential to combine the relevant requirements of power market-oriented trading reform and tariff formation mechanism reform, actively explore the new mode and mechanism of new energy grid-connected operation management by using big data technology, develop distribution grid upgrade and technology upgrade, closely track the development situation of distributed PV, and optimize the management mode of grid-connected service, scheduling management, transaction settlement, operation and maintenance control, etc.

To build a green, low-carbon, safe, and efficient energy system, scientific and orderly speed up the development of light, wind, and other resources can explore the local government's leadership, the government, power grids, power stations, suppliers, and other parties to jointly carry out the construction of photovoltaic evidence experimental base, to create an open public service platform for distributed power big data, advanced equipment, product performance, technical solutions, etc. to carry out exploratory testing, empirical comparisons, optimization To promote the application of new PV technologies, new materials, and new models. Explore new energy grid monitoring, take wind power and photovoltaic power plants as the main monitoring objects, and conduct in-depth research on a new energy grid structure, control methods, communication methods, relay protection, grid efficiency, intelligent operation and maintenance, economic evaluation, fresh energy consumption, settlement mode, etc. to provide scientific reference for new energy planning, grid operation and maintenance, photovoltaic industry development and even government decision-making.

3.2.1 Grid-connected service, marketing, and management

The application of big data technology in power marketing management is of great significance to electric power enterprises. Develop PV grid-connected service management system, carry out distributed PV power grid-connected business process processing, grid-connected equipment management, grid-connected operator management, meter reading management, settlement management, subsidy management, etc., and form data interaction with local government data sources to realize centralized management and resource sharing of ample data information of PV grid-connected users. At the same time, relying on the big data thus formed to carry out data analysis, the new energy value-added service business about daily operation analysis, weather impact analysis, power generation efficiency prediction, etc., is launched to explore a new mode of innovative power marketing management development of power supply enterprises.

3.2.2 Intelligent control and scheduling

With the improvement of equipment level and information technology, integrating distributed PV power generation into grid dispatch management will be an inevitable trend for future development. There is an urgent need to establish proper planning, access, efficient operation, and maintenance management standards, and measurement and control management means for distributed PV power farms to realize complete volume monitoring and grid dispatch management of PV power generation data in the region and build a new mode of distribution network operation that can adapt to a large number of distributed PV power sources. Grid companies can develop a centralized monitoring system for distributed PV power generation to realize the monitoring of multiple PV power generation units in their jurisdictions. The system can be composed of the following parts: First, the PV power collection system. Based on the collection terminal and power collection information system, real-time data is collected at 15-minute intervals to monitor the operation status of PV equipment, grid operation conditions, and equipment fault conditions in real time. Second is the PV power monitoring system. The monitoring system aims to ensure the safe operation of power stations and power grids, mainly monitoring the quality of grid-connected power and grid currents and remote tripping and closing control of faulty grid-connected points and abnormal running of grid-connected switch equipment to ensure the safe and stable operation of power grids. Third, the PV power generation data analysis and management model. Using historical mass data, we can create research topics such as grid-connected access point selection, distributed power distribution network planning, grid access capacity and consumption capacity analysis, energy-saving analysis, and the impact of small and micro power sources on the grid through big data technology, laying a solid foundation for future grid optimization operation, new electric energy settlement mode under power market reform, PV virtual power plant cloud computing, and even the development of industry standards. At the same time, it can also integrate multi-source system data such as electricity consumption information collection system, power marketing business system, financial system, and historical meteorological information to analyze regional PV power generation big data and study and analyze to solve the impact and development trend of PV power supply on distribution grid caused by a large-scale grid connection.

4. Problems and Opportunities

The application of big data technology in distributed PV power generation management and operation still reveals some problems, as follows.

(1) Poor data quality and missing data. Most PV operators do not install collectors on some equipment to save costs, so they cannot achieve real-time and full-volume data collection, which is not enough to support extensive data analysis. With the promotion of PV intelligent operation and comprehensive data technology applications, the demand for full-volume data access is becoming increasingly urgent.

(2) The timeliness and accuracy of the collected data are not enough. Most power stations do not pay attention to data utilization, and the accuracy and stability of the collection devices used are not enough, so the actual operating conditions of the equipment differ significantly from the collected data, which cannot accurately reflect the actual functional status of the equipment, and is not conducive to data analysis and processing.

(3) Big data integration and mining capacity are not strong. In recent years, photovoltaic big data technology application has made some development also accumulated a certain amount of historical data, but most of the application results remain at the level of maintaining the primary production and operation of the photovoltaic power plant; the application of big data technology scenarios are few, based on the application of in-depth analysis of big data capacity to be improved, for how to make a large amount of data to create more excellent value still needs to be further enhanced.

(4) Multi-business integration or cross-professional integration is not enough. There is no cross-business analysis and advanced application of PV big data, and new technology applications based on extensive data analysis, such as intelligent cleaning robots, precise fault identification, and precise fault positioning derived from the BeiDou positioning system, were not carried out.

(5) PV technology suppliers cannot provide high-quality products and technologies. Most suppliers have not studied the difficulties in PV power generation, grid operation, and management. The various types of PV operation solutions (technologies) developed are not adapted to the intelligent management needs of power stations and the development requirements of the new power market. The level of support for enterprise production and operation management needs to be strengthened.

5. Conclusion

Some sizeable centralized PV power plants have done extensive data application work and achieved some results. The construction of these monitoring systems based on big data technology has provided some support for the operation and management of PV power generation. However, the data application is not deep enough, and the monitoring system is not fully utilized for high-quality data mining and exploring new application scenarios. Innovative development of comprehensive data technology applications is still needed on the power generation side and grid side to achieve efficient and accurate power generation forecasting, operation and maintenance monitoring, fault prediction, comprehensive evaluation, and other functions, and at the same time, combined with the relevant requirements of power market-oriented trading reform and tariff formation mechanism reform, the use of It is necessary to actively explore new models and mechanisms of new energy grid-connected operation management by using big data technology.

In the future, the relevant authorities should pay attention to the interaction between big data technology and distributed photovoltaic power generation and increase support in industrial development planning and standard setting, market cultivation, and other aspects. At the same time, energy and grid service companies should make full use of advanced information technology and intelligent management technology to carry out extensive data analysis of PV power generation creatively and even renewable energy operation and management and deeply integrate various information systems to help "source-grid-load-storage" coordinate development and integrate and complement each other, to build a new low-carbon clean power generation system with high security, reliability, and efficiency. It can create a new low-carbon, clean, intelligent energy system with increased safety, reliability, and high efficiency.

Acknowledgments

The authors gratefully acknowledge the financial support from Chinese Student Innovation and Entrepreneurship Training Program Funding funds (Grant No. 202210623029).

References

[1] National Energy Board. http://www.nea.gov.cn/.

[2] LIN H Y. Will We Adopt AI Like We Adopted Electricity? [J]. Computer, 2021, 54(3): 48-51.

[3] LUO F J, DONG Z Y, LIANG G Q, et al. A Distributed Electricity Trading System in Active Distribution Networks Based on Multi-Agent Coalition and Blockchain [J]. Ieee Transactions on Power Systems, 2019, 34(5): 4097-108.

[4] NIET I, VAN EST R, VERAART F. Governing AI in Electricity Systems: Reflections on the EU Artificial Intelligence Bill [J]. Frontiers in Artificial Intelligence, 2021, 4.

[5] NOUR M, CHAVES-AVILA J P, SANCHEZ-MIRALLES A. Review of Blockchain Potential Applications in the Electricity Sector and Challenges for Large Scale Adoption [J]. Ieee Access, 2022, 10: 47384-418.

[6] Cao Wei, Dong Haoyang, Li Yun, et al. Foreign experience and practical insights of distributed PV high ratio access [J]. Electrical Drives, 2022, 52(04): 3-11. DOI:10.19457/j.1001-2095.dqcd22454.

[7] Xin Songxu. Foreign experience in distributed photovoltaic development and inspiration for China [J]. Energy and Environment, 2013(06): 42-43+46.

[8] Ma Di. Foreign photovoltaic power generation into the development of intelligent grid [J]. Power Technology and Environmental Protection, 2012, 28(05): 53-54.

[9] Zhu Zhenyu. Foreign solar photovoltaic industry development policy reference and inspiration [J]. China Market, 2010(13): 76-78.

[10] Chen JG, Liu CHANG, Zhu TX. Analysis of production and operation data of grid-connected photovoltaic power plants [J]. Solar Energy, 2020(01): 47-54.

[11] WANG J Y, QIAN Z, WANG J Y, et al. Hour-Ahead Photovoltaic Power Forecasting Using an Analog Plus Neural Network Ensemble Method [J]. Energies, 2020, 13(12).

[12] WANG Y F, ZHU L, XUE H. Ultra-Short-Term Photovoltaic Power Prediction Model Based on the Localized Emotion Reconstruction Emotional Neural Network [J]. Energies, 2020, 13(11).

[13] ZHOU N R, ZHOU Y, GONG L H, et al. Accurate prediction of photovoltaic power output based on long short-term memory network [J]. It Optoelectronics, 2020, 14(6): 399-405.