Journal of Electronic Research and Application, 2026, Volume 10, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

The Application of Large AI Models in Smart Grids: Prospects, Challenges, and Future Outlooks

Wen Chen

School of Automation, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, China

Abstract: With the rapid progress of AI technology, AI big models with transformer architecture as the core has made great progress in natural language processing, computer vision and other fields. Smart grid is a modern power system integrated with advanced information, communication and control technology. The complexity and variability of the system and the massive reference data provide application scenarios for the application of AI large model. This paper systematically expounds the key technologies matching with AI large model and its adaptability to the core links of smart grid, and focuses on the role of AI large model in smart grid construction, such as new energy grid connection, equipment management, grid topology optimization and dispatching decision, such as specific application modes and cases in load forecasting, real-time dispatching and multi-energy complementarity. At the same time, this paper deeply analyzes the key challenges in data, technology, engineering and security faced by the application of AI large model in various fields of power, and puts forward the corresponding optimal solutions. Finally, combined with typical cases, the future development direction of the integration of digital twins, generative AI and other technologies is conceived, which provides a theoretical reference and practical path for promoting the autonomous and efficient development of smart grid.

Keywords: AI large models; Smart grid; Transformer; Renewable energy integration; Power dispatch; Digital twin

Online publication: April 3, 2026

1. Introduction

Amid the global wave of energy transition, renewable energy sources such as wind and solar power are being integrated into the power grid on a large scale and in high proportions. Their intermittent and fluctuating characteristics pose significant challenges to the stable operation of the grid. At the same time, as power systems continue to expand in scale and grow increasingly complex, the interaction among generation, grid, load, and storage is intensifying. Traditional analysis methods based on physical models and classical optimization are no longer sufficient to meet the demands for real-time, precise, and intelligent operation. In this context, AI large models, with their powerful general cognition, contextual understanding, and few-shot/zero-shot learning capabilities, provide new solutions to these challenges. Applying these models to smart grids can deeply leverage the value of massive power data, enable precise perception of complex system states, forecast trends, and support intelligent decision-making. More importantly, it serves as a key driving force for advancing the power grid from automation to intelligence and autonomy, which is of great significance for ensuring national energy security.

Internationally, Google DeepMind has partnered with the UK National Grid to use deep learning models to enhance wind power forecasting accuracy and reduce grid balancing costs, becoming an early example of this technology. In recent years, power companies have focused on developing large models for the power industry (such as Power GPT), exploring their applications in scenarios such as generating equipment maintenance reports and understanding and generating dispatch instructions. Enterprises such as the State Grid and China Southern Power Grid are actively deploying these technologies, widely practiced in areas like load forecasting and equipment fault diagnosis. Some provincial power grids have piloted large AI models to assist in dispatch decision-making, with initial results evident. However, overall, current research and applications remain at an early stage, mostly consisting of individual breakthroughs, and a systematic application ecosystem covering the entire grid has yet to be established. Challenges remain in model interpretability and deep integration with existing systems.

This article focuses on summarizing the core theories of AI large models and the technical requirements of power grids. It elaborates on the application scenarios and model construction methods of large models from the two major dimensions of construction and scheduling. Through case analysis and future prospects, it outlines a development blueprint, aiming to systematically establish a theoretical framework and practical approach for the application of AI large models in smart grids.

2. Core theoretical foundation

2.1. Key technologies of large AI models

Transformer's self-attention mechanism can simultaneously process sequence data and effectively capture long-distance dependencies [1]. This property is very suitable for the high spatio-temporal correlation of power data (such as the interaction of different node voltages and the periodicity of load in time), which makes it suitable for complex tasks such as load sequence prediction and power grid transient stability analysis.

The core of transformer is as follows:

(1) Powerful presentation learning ability, which can automatically extract high-level features from massive and multi-source power data (Supervisory Control and Data Acquisition (SCADA), Phasor Measurement Unit (PMU), meteorology, text reports, etc.);

(2) Excellent context learning and generalization ability, which can adapt to power grid operation status or fault types in real time under a small number of samples;

(3) Multimodal fusion capability, which can process multi-modal data such as numerical value, text and image (such as infrared thermograph) cooperatively, and support the comprehensive diagnosis of equipment [2,3].

In order to cope with the challenges brought by the limited computing power of the equipment at the edge of the power grid, it is very important to establish a grid-oriented edge computing power optimization and distributed data storage and processing model [4]. The large-scale investment and use of such technologies can compress the huge cloud model into a lightweight model suitable for deployment at edge nodes such as substations and distribution terminals, which is helpful to realize localized real-time reasoning.

2.2. Core aspects and technical requirements of the smart grid

The smart grid construction end includes key nodes such as the planning and access of new energy power stations, the selection and full life cycle management of power grid equipment, power grid topology and infrastructure planning. The core link of smart grid construction requires AI model to have the ability of long-term prediction, multi-objective optimization and simulation planning. The AI model should cover the real-time scheduling control at the second/minute level, the day ahead/ hour scheduling plan, and the emergency recovery under emergency faults. At the same time, AI model is required to have high precision, high reliability and real-time response ability. The first special requirement for AI model is high reliability and robustness, where any wrong decision may lead to cascading failures. The second is strict interpretability, which needs to enable the dispatcher to understand the decision basis of the model. Finally, data security and privacy protection. Power data involves national security and user privacy. User privacy and the confidentiality of core data are also considered the capabilities of AI model.

3. Application of large AI models in smart grid construction

3.1. Optimized application of new energy grid connection

Based on Transformer's prediction model, AI large-scale model can simultaneously integrate multi-source information such as high-precision numerical weather forecast, historical power data and satellite cloud images, dynamically weight the importance of different features by using attention mechanism, significantly improve the accuracy and certainty of wind power prediction in the next 0–4 hours. At the same time, the large model can learn a large amount of historical operation data and simulation data, build a dynamic model of the interaction between the new energy station and the power grid, generate the optimal reactive power compensation and energy storage charging and discharging strategies in real time, and contribute a more reasonable scheme in terms of suppressing voltage fluctuations and improving the stability of the grid.

3.2. Full life cycle management of power grid equipment

Large-scale vision-language models can be developed to interpret multimodal data related to power equipment monitoring, including transformer oil chromatographic data, partial discharge signals, infrared inspection images, and patrol text records. By integrating these heterogeneous data sources, the model can perform comprehensive assessments of equipment health conditions and enable early detection of potential abnormalities [5]. Furthermore, leveraging equipment operational time-series data, the sequence modeling capability of large models can be utilized to provide early fault warnings and predict the remaining useful life of equipment. Such models can estimate performance degradation trajectories, thereby facilitating the transition from traditional periodic maintenance to predictive maintenance strategies. This approach also enables more efficient scheduling of maintenance activities and improved optimization of maintenance resources and spare parts management [6].

3.3. Power grid topology and infrastructure optimization

Combined with regional economic development planning, population migration, climate change and other macro factors, the large-scale model is used for long-term load forecasting in the next decade, providing solutions for infrastructure planning such as power grid expansion and new lines. In the distribution network with high proportion of distributed generation, the large model can calculate the optimal network reconfiguration scheme in real time, guide power flow distribution, reduce network loss and improve power supply reliability.

4. Application of large AI models in smart grid dispatch decision-making

4.1. Load forecasting and scheduling plan development

The large-scale model can deeply integrate all kinds of information such as temperature, humidity, holidays, and even social media events to achieve high-precision short-term load forecasting and lay the data foundation for power generation planning. Based on accurate load and new energy forecasting, the large-scale model can quickly solve complex optimization problems with multiple objectives, such as unit commitment, line safety constraints, and carbon emission constraints, and generate economic, safe, and implementable scheduling plans.

4.2. Real-time dispatching and emergency decision

On the second/minute time scale, the large model can quickly analyze the frequency deviation and power shortage of the power grid, coordinate AGC units, energy storage system, controllable loads and other resources, and achieve fast and accurate frequency and peak shaving. When the power grid fails, the large model can instantly process massive protection action signals and SCADA data, quickly locate the fault point, and generate the optimal load transfer and power supply recovery path based on the current power grid topology, which can greatly shorten the outage time.

4.3. Multi-energy complementary scheduling optimization

In today's era, the integrated energy system has been put into use on a large scale. The large model can uniformly dispatch various energy forms such as power, heat, refrigeration and energy storage, tap the complementary characteristics of different energy systems, and improve the overall energy efficiency. Recently, the proportion of new energy installed capacity has increased and the power balance in some provinces has become tighter, and the system supply guarantee risk has an obvious upward trend. China Southern Airlines 110kV substation uses AI-driven forecasting + optimal dispatching scheme. During peak hours, the regulation scheme is deployed in advance through AI forecasting, which effectively reduces peak load and reduces the overall operation cost by about 8.5%. During the extreme high temperature period, the fluctuation amplitude of grid load decreases by about 14%, and the system operation is more stable. The response delay of the dispatching system to sudden load changes is shortened from an average of 15 minutes to about minutes, and the response speed is increased by 21% [7].

5. Key challenges and solutions

At the technical level, the problem of data fragmentation and privacy protection can be solved by building a distributed training framework based on Federated learning, so that the model can learn from the data of power companies without centralizing the original data and protect data privacy [8]. The black box and interpretability of the model are insufficient. To solve this problem, interpretable AI technologies such as attention visualization and counterfactual interpretation can be introduced, and a hybrid model of symbolic reasoning and neural network for power knowledge can be developed to enhance the transparency of the decision-making process [8].

In terms of the compatibility with the existing dispatching system, microservice architecture and standardized API interface can be used to encapsulate the large model as an independent intelligent service module, which is loosely coupled with the existing EMS, DMS and other systems. The computing power of edge devices is limited and deployed. By continuously promoting the model lightweight technology and exploring the cloud edge end collaborative reasoning architecture, the complex computing is carried out in the cloud, and the edge side only performs lightweight reasoning.

In terms of security and compliance, the model resists attacks and improves its robustness. By introducing a large number of antagonistic samples in the training, the robustness training can be carried out for many times, and the real-time monitoring and manual intervention mechanism for the decisions made by the model can be established. At the same time, power data security and compliance use should establish a strict data classification management system and access control strategy, and use a variety of cutting-edge technologies such as differential privacy and homomorphic encryption to achieve the effect of protecting sensitive information in the process of data use.

6. Case analysis

6.1. AI large model dispatching system of a domestic municipal power grid

The project builds a unified load forecasting and dispatching decision-making model for the whole city based on transformer. The model integrates more than 100 features such as meteorology, economy and calendar, and improves the accuracy of short-term load forecasting to more than 97.5% [7]. In the process of dispatching planning, the system can generate multiple feasible schemes within one minute to provide dispatchers with choices, and simulate the economy and security of the scheme, which can effectively reduce the pressure of dispatchers and improve the economic benefits of power grid operation.

6.2. Deepmind and national grid

Deepmind has improved the prediction accuracy of wind farms in the UK by about 20% through its own deep learning model [9]. The project not only verifies that AI has great potential in improving the consumption of new energy, but also provides the prediction results and suggestions provided by its man-machine collaborative landing mode. The final decision-making power is still in the hands of the dispatcher, which also provides an important reference for the safe implementation of subsequent projects.

6.3. American NREL Egridgpt auxiliary system

As a top research institution in the energy field, the National Renewable Energy Laboratory (NREL) of the United States released the technical report “Egridgpt: Trustworthy AI in the control room” in 2024, proposing that Egridgpt, the world’s first trusted decision support system that applies generative AI to the grid control room, constructs a new human-computer collaborative intelligent dispatching mode by deeply integrating the large language model with the physical laws of the power system [9].

7. Future prospects

7.1. Direction of technological development

By integrating large models with digital twin power grids, AI large models will become the brain of the digital twin of the power grid in the future, realizing advanced control and independent decision-making of the physical power grid through real-time simulation, deduction, and optimization in virtual space, which is conducive to the construction and development of smart grids. The State Grid Corporation of China coordinates artificial intelligence and top-level design, and forms a 6541 overall planning layout with the Guangming Power Model as the core, that is, it fully covers more than 600 scenarios in 6 major business areas such as planning and construction and power grid operation, and strives to improve the five technical capabilities of perception intelligence, cognitive intelligence, decision-making intelligence, embodied intelligence, and scientific intelligence, and continues to consolidate the four core elements of samples, computing power, platform, and model, and builds a two-level collaborative operation mechanism with the Guangming Power Model as the core. The implementation of AI technology should be systematically promoted in the power field [10]. By integrating generative AI into grid design, it can be used to automatically generate optimal grid line layout and substation design schemes based on geographic information, load demand, and other conditions, greatly improving the efficiency of grid planning and design [11].

7.2. Expansion of application scenarios and construction of industrial ecology

When put into practical application, it is necessary to achieve integrated collaborative scheduling of source, grid, load, and storage, use large models to break through the boundaries of traditional scheduling, realize the aggregation and collaborative scheduling of resources in the whole link of source-grid-load-storage, and build an energy balance system in a wide area of time and space [12]. By realizing cross-border cross-regional power grid joint scheduling, the large model can coordinate the dispatching strategies of different regions in a larger range of interconnected power grids, optimize cross-regional power trading and backup sharing, and ensure the overall safety and economic operation of the large power grid.

When building an industrial ecology, it is necessary to help formulate standards and build open source communities, and the supply side needs to accelerate the formulation of data, models, and interface standards for large power models, and encourage the construction of open source communities, gather industry wisdom, avoid duplicate construction, and lower the application threshold. To build a compound talent training system, we should vigorously cultivate compound talents who are proficient in both power system expertise and AI technology, and provide manual support for the continuous innovation and development of smart grids.

8. Conclusion

This paper systematically demonstrates the feasibility of AI large models in the construction and development of smart grids, and also fully demonstrates their huge application potential. The study finds that from the precise grid-connected regulation of new energy sources such as wind power and photovoltaics, to the health control of the whole life cycle of power grid core equipment, to the refined prediction of user side load and the real-time optimization and scheduling of power grid operation status, AI large models are gradually becoming the core force to promote the intelligent upgrading of power grids and break through the bottlenecks of traditional power grid operation with their excellent feature extraction capabilities, complex scenario modeling level and adaptive decision-making advantages. Although AI large models have development potential, many core technical difficulties must be overcome to fully implement them in the power field. First of all, the power system has extremely high requirements in terms of safety, and whether the reliability of large models can meet the standards under various extreme working conditions still needs to be fully verified. Moreover, the black box nature of the large model makes its decision-making process not open and transparent, which makes it difficult to meet the requirements of the power industry in supervision and operation and maintenance. Finally, most of the existing power grid systems are heterogeneous architectures, and there are many technical problems to be overcome when large models are compatible with these traditional systems. Future research can focus on the following directions:

(1) To create a more robust and explainable power-specific large model, integrate the physical laws of the power system into the model, and optimize the structure and training strategy of the model;

(2) To build an efficient and secure cloud-edge-end collaborative deployment architecture, so that the computing power of the large model can more accurately match the real-time response needs of the power grid;
(3) To explore the in-depth combination of large models with new technologies such as digital twins and blockchain, with the help of digital twins to realize the full-scenario simulation and deduction of power grid operation, and use blockchain to ensure the trusted sharing of data and transaction security.

As long as research breakthroughs can be made in these aspects, a safe, efficient, green, and intelligent next-generation power system can finally be built, providing strong support for the energy revolution and the realization of the dual carbon goal.

Disclosure statement

The author declares no conflict of interest.

References

[1] Xu L, Zeng Y, Yuan J, et al., 2025, Time Series Interpolation Computer Engineering based on Self-Attention Mechanism. 51(11): 90–99.

[2] Li B, Zhang F, Shen Y, 2020, Research on Edge Computing Power Optimization and Distributed Data Storage and Processing Model for Power Grid. Guangdong Electric Power, 33(9): 92–99.

[3] Dong Y, Hou H, Gong M, et al., 2025, Multimodal Industrial Defect Detection and Intelligent Decision-Making Driven by Visual Language Large Model. Computer Engineering and Applications, 1–15.

[4] Wei W, 2025, AI-based Smart Grid Load Prediction and Scheduling Optimization. Proceedings of the 2nd Academic Symposium on Intelligent Engineering and Economic Construction, 2025(1): 956–959.

[5] An Y, Cui M, Han Y, et al., 2025, Evaluation Method of Distributed Photovoltaic Active Adjustable Capability based on Adaptive Federated Learning for Data Privacy Protection. Transactions of the CSE, 45(21): 8281–8295.

[6] Wang X, Dou J, Liu Z, et al., 2024, A Review and Prospect of the Application of Explainable Artificial Intelligence in Power Systems. Power System Automation, 48(4): 169–191.

[7] Zhang H, Li H, 2025, Short-Term Power Load Prediction based on MSCSO-Transformer-BiLSTM. Journal of Jiamusi University (Natural Science Edition), 43(11): 15–20.

[8] Ding Z, Ding Z, Zhu L, 2025, Application of Large Language Model in Load Prediction and Optimal Scheduling in Smart Grid Environment. Electrotechnical Technology, 2025(S1): 284–286.

[9] Sparkes M, 2023, DeepMind AI Predicts the Weather. New Scientist, 260(3465): 9.

[10] State Grid, 2025, State Grid Released the First 100 Billion-Level Multi-Modal Power Industry Model in China. Industrial Control Computer, 38(1): 137.

[11] Wu Q, Zhao W, 2025, A Review of the Research on Digital Twin and Generative Artificial Intelligence to Empower the Operation and Control of New Power Systems. Power Automation Equipment, 45(11): 59–71.

[12] Wei Y, Li G, 2025, Integration of Source, Grid, Load and Storage: The Innovation Drive and Development Path of New Power Systems. Energy Research and Management, 17(3): 23–29.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.