energies

Special Issue Reprint

AI Facilitated Cyber–Physical Energy Systems

Planning, Operation, and Markets

Edited by
Stefanie Kuenzel and Xiao-Yu Zhang

mdpi.com/journal/energies

MDPI

AI Facilitated Cyber–Physical Energy Systems—Planning, Operation, and Markets

AI Facilitated Cyber–Physical Energy Systems—Planning, Operation, and Markets

Guest Editors

Stefanie Kuenzel
Xiao-Yu Zhang

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors
Stefanie Kuenzel
Department of Electronic Engineering
Royal Holloway, University of London
Egham
UK

Xiao-Yu Zhang
School of Artificial Intelligence
Anhui University
Hefei
China

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Energies* (ISSN 1996-1073), freely accessible at: https://www.mdpi.com/journal/energies/special_issues/95BI0XA5VK.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-8879-5 (Hbk)
ISBN 978-3-7258-8880-1 (PDF)
https://doi.org/10.3390/books978-3-7258-8880-1

Contents

About the Editors

Stefanie Kuenzel

Stefanie Kuenzel received M.Eng. and Ph.D. degrees from the Imperial College London, London, U.K., in 2010 and 2014, respectively. She is currently the Head of the Power Systems Group and a Senior Lecturer with the Department of Electronic Engineering, Royal Holloway, University of London, Egham, U.K. Her current research interests include renewable generation and transmission, including HVdc and smart meters. Dr. Kuenzel served as an Editor for the *IEEE Transactions on Sustainable Energy* and the *IEEE Power Engineering Letters*.

Xiao-Yu Zhang

Xiao-Yu Zhang received a B.Eng. degree from North China Electric Power University, Beijing, China, in 2016; an M.S. degree (Hons.) in electrical power systems from the University of Birmingham, U.K., in 2017; and a Ph.D. in electrical engineering from the University of London, U.K., in 2022. He is currently an Associate Professor with the School of Artificial Intelligence, Anhui University, Hefei, China. His research interests include smart grids, deep reinforcement learning technology, and data analytics.

Preface

The integration of cyber and physical infrastructures has advanced the digital and low-carbon evolution of modern energy systems, enabling the development of resilient grids with high renewable penetration. Artificial intelligence (AI) has become a fundamental enabler to address critical challenges including stochastic system operation, potential safety risks, cybersecurity vulnerabilities, and computational limitations in conventional energy planning and simulation. AI techniques support full-cycle optimization in system planning, operational regulation, real-time monitoring, and market trading, facilitating autonomous intelligent decision-making for modern power grids. However, unresolved issues in power system planning, resilient fault control, and uncertainty modeling still hinder the large-scale deployment of intelligent energy technologies. This Reprint aims to fill these research gaps by compiling novel original studies and systematic reviews on state-of-the-art AI innovations for cyber–physical energy systems.

This Reprint comprises seven original research articles and two review papers, forming a complete technical framework covering risk assessment, wide-area monitoring, distribution network optimization, scenario generation, energy forecasting, privacy protection, high-performance simulation, and industrial implementation. All selected studies target key technical bottlenecks in smart energy systems, including hidden risk detection, limitations of traditional monitoring architectures, planning uncertainties induced by renewable integration, data privacy constraints in distributed energy systems, and intensive computational demands in power system simulation.

It is with profound sorrow that Guest Editor Dr. Stefanie Kuenzel passed away in July 2026. Dr. Kuenzel devoted tremendous effort and professional expertise throughout the compilation of this Reprint. We dedicate this Reprint in loving memory of her invaluable dedication, profound academic insights, and outstanding contributions to this collection and the broader field of intelligent cyber–physical energy systems.

The guest editors would like to acknowledge all authors for their valuable contributions. Gratitude is extended to anonymous reviewers for their constructive and timely comments, which have substantially improved the quality of this Reprint. Sincere thanks are also given to the editorial team of *Energies* for their professional support throughout the publication process. It is hoped that this Reprint can promote academic exchanges and technical innovation in AI-driven cyber–physical energy systems, and accelerate the development of efficient, secure, and sustainable modern energy systems.

Stefanie Kuenzel and Xiao-Yu Zhang
Guest Editors

energies

Editorial

Special Issue "AI Facilitated Cyber–Physical Energy Systems—Planning, Operation, and Markets"

Stefanie Kuenzel [1,*] and Xiao-Yu Zhang [2,*]

1 Department of Electronic Engineering, Royal Holloway, University of London, London TW20 0EX, UK
2 School of Artificial Intelligence, Anhui University, Hefei 230601, China
* Correspondence: stefanie.kuenzel@rhul.ac.uk (S.K.); zhangxiaoyu@ahu.edu.cn (X.-Y.Z.)

AI technologies are rapidly becoming a foundational enabler for modern cyber–physical energy systems, supporting more efficient planning, safer and more flexible operation, stronger cybersecurity, and more dynamic energy markets. Together, these advances lay the groundwork for next-generation smart grids characterized by high renewable penetration, real-time situational awareness, and closed-loop intelligent decision-making. This Special Issue collects seven original research articles and two comprehensive review articles that together cover the full pipeline of AI-driven innovation in cyber–physical energy systems, from risk mining and wide-area monitoring to distribution network planning, scenario generation, energy forecasting, and quantum-enhanced simulation.

The first paper in this Special Issue is a novel data mining approach for accident factor analysis in intelligent energy systems by Li et al. (Contribution 1). Focused on unstructured power accident text records, the authors propose a Sparse Coefficient Optimized Weighted FP-Growth (SCO-WFP) algorithm that combines Word2Vec semantic embedding, K-means clustering, cosine similarity weighting, and sparse factor enhancement. The method improves the detection of low-frequency but high-impact risk factors and ranks their importance relative to accident severity. The results offer strong support for safety early warning and risk prevention in smart grid operations.

The review paper by Ogbogu et al. (Contribution 2) provides a comprehensive overview of wide-area measurement systems (WAMS) and phasor measurement unit (PMU) technologies for smart grid fault mitigation and cybersecurity. The work compares WAMS with traditional SCADA systems, outlines system architectures, optimal PMU placement, communication frameworks, and AI-enhanced fault detection. It also addresses challenges in developing economies, aging infrastructure, and cyber–physical system (CPS) resilience, offering a clear roadmap for future deployment and research.

Next, a technical contribution on distribution network planning is presented by Ma et al. (Contribution 3). To handle the uncertainty and complexity introduced by distributed photovoltaics and wind generation, the authors develop a deep reinforcement learning (DRL) framework based on the Proximal Policy Optimization (PPO) algorithm. By modeling network planning as a Markov decision process with multi-objective rewards including investment cost, voltage deviation, and renewable accommodation, the method automatically generates optimal network topologies. When compared with heuristic approaches, the proposed solution achieves lower costs, higher renewable integration, and greater computational efficiency.

Another important contribution addressing uncertainty in energy systems is presented by Ma et al. (Contribution 4). The authors propose a multi-time-scale scenario generation method using a Temporal Generative Adversarial Network (TGAN) enhanced with multi-head attention and temporal convolutional networks. The model captures spatiotemporal correlations in source–load behaviors and generates realistic daily, weekly, and monthly scenarios. By combining Markov chain state transitions and conditional GAN generation, the approach improves scenario fidelity and supports robust long-term distribution network planning and operation analysis.

A privacy-preserving solution for distributed photovoltaic power prediction is introduced by Luo et al. (Contribution 5). Motivated by the non-IID data characteristics and privacy constraints in distributed PV systems, the authors propose a Personalized Federated Multi-Task Learning (PFL) framework combined with an improved CBAM-iTCN predictor. The approach keeps raw data local while enabling collaborative model training, and the attention-augmented temporal convolution network extracts key meteorological and power features more effectively. Experiments show substantial reductions in MAE, RMSE, and MAPE compared with centralized and conventional federated learning models.

Soltaninia and Zhan (Contribution 6) explore an emerging frontier in power system simulation—quantum computing. This work introduces quantum neural networks (QNNs) to solve the high-dimensional differential-algebraic equations involved in power system transient simulation. By designing sinusoidal-friendly and polynomial-friendly QNN architectures, the authors achieve more efficient approximation of dynamic system responses with fewer parameters than classical physics-informed neural networks. The results demonstrate the promising potential of quantum computing to accelerate computationally demanding tasks in future cyber–physical energy systems.

Turning to wind energy forecasting, Xu et al. (Contribution 7) present a novel wind power forecasting framework based on multi-graph neural networks considering external disturbances. Termed GCN-EIF, the model decouples inherent wind power patterns from external interference factors (EIFs) using a dual graph structure (geographic proximity and power correlation) and attention-enhanced LSTM. By removing disturbances during training and reintroducing predicted external factors at forecasting stage, the method reduces RMSE by 18.99% and MAE by 5.08% over state-of-the-art methods, with strong real-time performance for industrial applications.

The eighth paper provides a systematic review by Saldaña-González et al. (Contribution 8) on active distribution network (ADN) planning and generative AI applications. The work classifies core planning elements including time horizons, objectives, decision variables, uncertainty modeling, and optimal power flow formulations. It further reviews GANs, diffusion models, VAEs, flow-based models, and large language models for scenario generation, uncertainty quantification, and decision support, offering a unified view for future AI-driven distribution planning.

Closing this Special Issue, Ji et al. (Contribution 9) propose an integrated energy management strategy for fuel cell hybrid-powered unmanned aerial vehicles (UAVs) based on optimal path planning. Using CFD-simulated urban wind fields and LSTM power prediction, a hybrid A*-ACO algorithm generates energy-efficient flight trajectories. A Deep Q-Learning (DQN) controller then optimizes power allocation between fuel cell and battery to reduce hydrogen consumption and suppress power fluctuations.

Together, the papers included in this Special Issue advance the state of the art in AI for cyber–physical energy systems across planning, operation, monitoring, forecasting, security, and computation. The collection combines rigorous methodological innovation, extensive experimental validation, and clear practical relevance, making it a valuable reference for researchers and engineers working toward smarter, more resilient, and sustainable power grids.

As guest editors, we would like to express our sincere gratitude to all authors for their high-quality contributions, the reviewers for their constructive and timely feedback, and the editorial team of *Energies* for their professional support throughout the publication process.

Funding: This work was supported in part by the National Natural Science Foundation of China under grant numbers 62303005, W2421091, and U24A20243.

Conflicts of Interest: The authors declare no conflicts of interest.

List of Contributions

1. Li, R.; Zhang, J.; Deng, F. Accident Factors Importance Ranking for Intelligent Energy Systems Based on a Novel Data Mining Strategy. *Energies* **2025**, *18*, 716. https://doi.org/10.3390/en18030716.
2. Ogbogu, C.E.; Thornburg, J.; Okozi, S.O. Smart Grid Fault Mitigation and Cybersecurity with Wide-Area Measurement Systems: A Review. *Energies* **2025**, *18*, 994. https://doi.org/10.3390/en18040994.
3. Ma, L.; Si, C.; Wang, K.; Luo, J.; Jiang, S.; Song, Y. Deep Reinforcement Learning-Based Distribution Network Planning Method Considering Renewable Energy. *Energies* **2025**, *18*, 1254. https://doi.org/10.3390/en18051254.
4. Ma, L.; Jiang, S.; Song, Y.; Si, C.; Li, X. Multi-Time Scale Scenario Generation for Source–Load Modeling Through Temporal Generative Adversarial Networks. *Energies* **2025**, *18*, 1462. https://doi.org/10.3390/en18061462.
5. Luo, W.; Shen, Y.; Li, Z.; Deng, F. Distributed Photovoltaic Short-Term Power Prediction Based on Personalized Federated Multi-Task Learning. *Energies* **2025**, *18*, 1796. https://doi.org/10.3390/en18071796.
6. Soltaninia, M.; Zhan, J. Quantum Neural Networks for Solving Power System Transient Simulation Problem. *Energies* **2025**, *18*, 2525. https://doi.org/10.3390/en18102525.
7. Xu, X.; Luo, Z.; Feng, M. Wind Power Forecasting Based on Multi-Graph Neural Networks Considering External Disturbances. *Energies* **2025**, *18*, 2969. https://doi.org/10.3390/en18112969.
8. Saldaña-González, A.E.; Aragüés-Peñalba, M.; Gadelha, V.; Sumper, A. Review of Active Distribution Network Planning: Elements in Optimization Models and Generative AI Applications. *Energies* **2026**, *19*, 116. https://doi.org/10.3390/en19010116.
9. Ji, Y.; Ling, X.; Wu, X.; Hu, J. Energy Management for a Fuel Cell Hybrid-Powered Unmanned Aerial Vehicle Based on Optimal Path Planning. *Energies* **2026**, *19*, 1854. https://doi.org/10.3390/en19081854.

energies

Review

Review of Active Distribution Network Planning: Elements in Optimization Models and Generative AI Applications

Antonio E. Saldaña-González *, Mònica Aragüés-Peñalba, Vinicius Gadelha and Andreas Sumper

Centre d'Innovació Tecnològica en Convertidors Estàtics i Accionaments (CITCEA-UPC), Departament d'Enginyeria Elèctrica, Universitat Politècnica de Catalunya ETS d'Enginyeria Industrial de Barcelona, Avinguda Diagonal, 647, Pl. 2, 08028 Barcelona, Spain; monica.aragues@upc.edu (M.A.-P.); vinicius.gadelha@upc.edu (V.G.); andreas.sumper@upc.edu (A.S.)
* Correspondence: antonio.emmanuel.saldana@upc.edu

Abstract

Active distribution networks (ADNs) are rapidly evolving with the integration of distributed energy resources, flexible loads, and energy storage systems. Traditional planning methods, based on passive upgrades and worst-case scenarios, are no longer adequate for high DER penetration and dynamic system behavior. This review highlights the key evolution needs that will drive the evolution towards a more dynamic and optimized active distribution planning. Furthermore, this work reviews the core elements in ADN planning, covering time horizons, objectives, decision variables, uncertainty approaches, and optimal power flow formulations. This work also reviews recent generative AI models applied to active distribution networks, presenting a structured classification and definitions of each generative AI category.

Keywords: active distribution network planning; generative AI applications; optimization elements; traditional decision variables; flexible sources; emerging assets

1. Introduction

Active distribution networks represent an evolution of traditional distribution networks. The key factors are driven by the increasing integration of distributed energy resources (DERs), including energy storage systems (ESS), electric vehicles (EVs), and demand response mechanisms [1]. This evolution is driven by the global need for decarbonization. The efforts to reduce climate change through DERs offer benefits such as reduced power losses, minimized fuel costs, and lower greenhouse gas emissions.

Traditional distribution systems, characterized by unidirectional power flows and centralized control, are increasingly strained by the variability of renewable sources and fluctuating loads, necessitating flexible, bidirectional operations in ADNs [2]. The rising EV adoption and the need for reliable infrastructure in the face of demand growth and intermittency from sources such as solar PV and wind underscore the relevance of ADN planning in enabling sustainable, cost-effective grid modernization.

Existing literature on distribution system planning has primarily focused on traditional approaches, emphasizing passive infrastructure upgrades, single-objective optimizations, and deterministic models under worst-case scenarios [3]. Key frameworks include linear programming for basic expansion planning and mixed-integer models for asset sizing and placement, often addressing isolated elements such as substation upgrades or feeder reinforcements. Recent studies have begun incorporating DER integration and active management schemes, such as optimal power flow (OPF) formulations and stochastic optimization, to handle uncertainties in generation and demand. Dominant paradigms involve multi-stage planning horizons and bi-level optimizations for coordinating transmission and distribution networks. While several studies have explored DER integration or OPF techniques, they often address isolated aspects without a unified planning framework. Figure 1 illustrates the transition from traditional distribution network planning—characterized by passive actions, a "fit and forget" scheme, and worst-case scenarios to active planning, which emphasizes DER integration, flexibility utilization, and dynamic optimization. Shared aims include objectives, time horizons, network constraints, and investment constraints, while evolution needs encompass AC-OPF formulations, new asset modeling, and uncertainty management. Despite advancements, the literature falls short in providing holistic frameworks that integrate time horizons, decision variables, and constraints across low- and medium-voltage (LV/MV) contexts, particularly under high DER penetration and uncertainties such as load forecast errors, EV charging variability, and cyber threats [4]. Traditional planning's reliance on passive strategies, rigid design rules, and worst-case assumptions limits adaptability to real-time flexibility and dynamic optimization [5]. Few reviews provide a comprehensive view of the evolving needs and new trends in the application of AI to power system planning. In this work, beyond focusing on the emerging elements in the formulation of the ADN planning problem, the main novelty lies in the potential application of generative AI models in the decision-making processes of operation and planning.

Table 1 highlights the key characteristics between passive and active distribution networks. Passive networks operate with unidirectional power flow, limited monitoring, and rely almost exclusively on infrastructure upgrades to solve constraints. In contrast, ADN uses bidirectional flows, near real-time control, and monitoring, and the use of flexibility from demand response, EVs, storage, and DERs.

Table 1. Key characteristics of passive and active distribution networks.

Characteristic	Passive Distribution Network	Active Distribution Network
Power flow direction	Unidirectional	Bidirectional
Control and monitoring	Limited, mostly manual	Near real-time control
Flexibility sources	Almost none	Demand response, EVs, storage, DERs
Voltage and congestion control	Network reinforcement	DERs, storage, reactive power control
Investment driver	Infrastructure upgrades	Flexibility + selective reinforcement
System role	Passive	Active system participant

This literature review aims to define the common key elements of passive and active distribution system planning while identifying unresolved challenges and trends in the transition to ADNs. Furthermore, this work provides an overview of generative AI (GenAI) applications for enhancing the operation and planning of distribution networks. The main contributions of this literature review are:

1. A comprehensive review of key concepts in passive and active distribution system planning, covering time horizons, objectives, decision variables, and technical constraints. This is supplemented by a comparison of different OPF formulations, uncertainty techniques, and flexible planning tools.
2. A structured review and categorization of generative AI models for power systems, with emphasis on scenario generation, uncertainty modeling, optimization, and decision support, and a focused analysis of their application to planning in ADNs.

https://doi.org/10.3390/en19010116

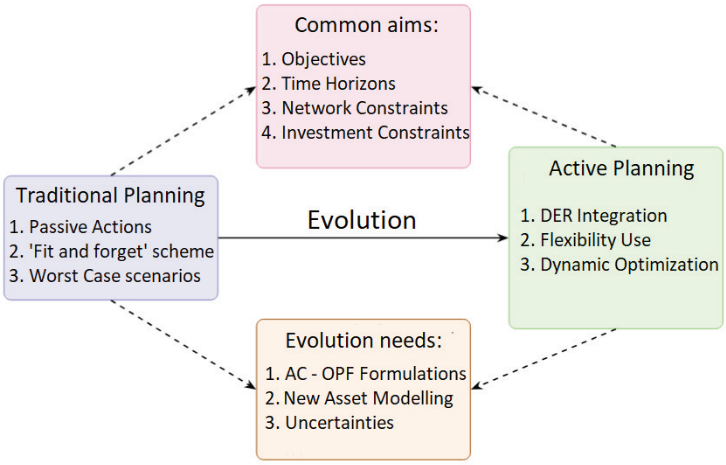


Figure 1. Evolution needs for active distribution network system planning.

2. Elements of the Distribution Network Planning Problem

Formulating a distribution network planning problem means translating complex real-world engineering and economic challenges into a structured, solvable mathematical model. The architectural integrity of this model depends entirely on the precise definition of its fundamental elements. These elements, which include the planning horizon, objectives, decision variables, constraints, and uncertainty, collectively define the problem's real applicability. Figure 2 shows a summary of the ADN planning elements that this review target. The specific choices made for each component are not trivial; they dictate the model's mathematical structure and the robustness of the resulting plan. The following subsections provide an examination of each of these elements.

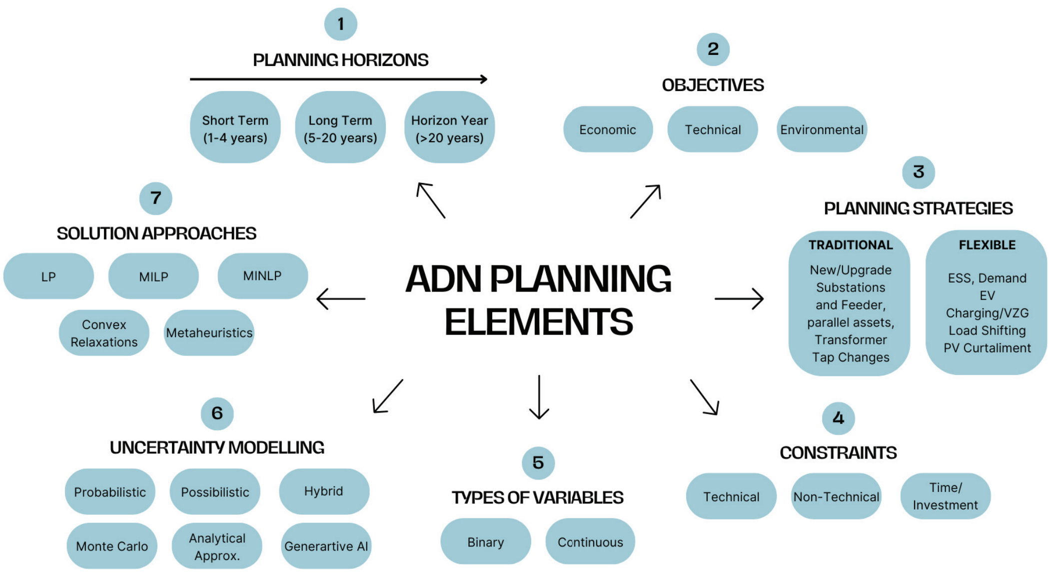


Figure 2. Summary of active distribution network elements and solutions approaches.

The following review employed a systematic search for this review on distribution network planning optimization elements and generative AI applications. The databases IEEE Xplore, Scopus, Web of Science, ScienceDirect, and Google Scholar were searched mainly from 2020–2025 using structured boolean keywords. Following PRISMA guidelines, about 500 records were initially identified and refined to 104 relevant references after screening, prioritizing simulation-based studies while excluding non-peer-reviewed sources.

2.1. Planning Horizons

Planning horizons in power systems can be classified by duration, reflecting different technical and economic considerations. Table 2 shows the time horizons, aims, and planning actions.

Table 2. Planning horizons.

Planning	Horizon	Aim	Actions
Short-Term	1 to 4 years	Expansion planning for immediate or near-term needs, operational improvements.	Conductor sizes, number of feeders, transformer sizes, and locations.
Long-Term	5 to 20 years	Developing infrastructure to meet future demand, aligning with medium-term goals.	System design standards, primary and secondary voltage classes, feeder configurations.
Horizon-Year	20+ years	Strategic, cost-effective infrastructure design to meet long-term consumer needs.	Comprehensive system design, integrating primary and secondary systems.

Short-term planning addresses immediate grid expansion to accommodate new demands. The system expansions include: upgrading feeders, transformers, and conductors to support new connections, such as electric vehicle charging stations. The integration of EV charging infrastructure often leads to multiple connection requests that must be addressed within months, necessitating rapid deployment of capacity. Some recent studies have adopted hybrid planning strategies that integrate traditional and flexible approaches to determine the optimal investment decisions for the next five years [6,7]. This combined perspective enables a more adaptive and cost-effective planning process capable of addressing increasing uncertainty in demand and generation patterns. The Nordic report in [8] highlights the role of short-term flexibility markets in providing immediate solutions for short term planning expansions. The use of distributed flexibility solutions, such as local batteries or demand response systems, can mitigate congestion and manage peak loads.

Long-term planning aims to reinforce the electric power infrastructure based on demand growth over the next 5 to 20 years. This horizon focuses on minimizing total investment costs, including the installation of new assets. Some of the reinforcements include the replacement of underground cables, the replacement of substations, and the deployment of energy storage to manage renewable energy integration and demand growth. The latter has yet to be regulated by DSOs in some countries. In [8] a Finnish project used battery energy storage in long-term investment applications to support reliability and energy markets, thus achieving sustainable grid modernization.

Horizon Year Planning focuses on strategic system design to meet energy needs for time horizons of more than 20 years. For such a long horizon, it is necessary to integrate and coordinate all primary and secondary systems to ensure seamless scalability and flexibility. Uncertainty is very high, and it is necessary to consider solutions that are prepared for widespread electrification, demographic and climate changes, and the integration of technologies aligned with future projections.

2.2. Planning Objectives

The planning objectives for distribution systems can be classified into economic, technical, and environmental objectives. Table 3 shows common planning objectives. The minimization of total investment costs is the most common objective in active distribution planning modeling [9]. The maximization of hosting capacity with a fixed budget and the maximization of net profit by effectively managing resources are also increasing in importance due to the need to integrate electric vehicles into the network.

Table 3. Planning objectives of distribution network models.

Category	Objective and Description
Economic [7,10–12]	1. Minimization of CAPEX and OPEX: Optimizing resources to reduce both initial investment and ongoing operational costs. 2. Maximization of assets capacity with fixed TOTEX: Increasing the ability of the asset capacity in the network to integrate DERs, EVs or load growth within a fixed TOTEX. 3. Maximization of net profit value: Enhancing revenue and profitability by effectively managing flexible resources.
Technical [13–16]	1. Maximization of system reliability: Ensuring consistent power delivery by minimizing system outages and enhancing redundancy. 2. Improvement of voltage deviation: Stabilizing voltage levels across the network to meet power quality standards and reduce flicker or voltage violations. 3. Minimization of network losses: Reducing energy losses in the network to improve overall efficiency.
Environmental [17–20]	1. Minimization of carbon emissions: Reducing the carbon footprint of the network by integrating clean energy sources and improving energy efficiency. 2. Maximization of renewable DG penetration: Increasing the share of renewables like solar and wind in the energy mix. 3. Maximization of EV charging stations: Supporting the adoption of electric vehicles by strategically placing charging infrastructure, which can also enhance renewable integration.

The objective functions can be single-objective, multi-objective, or bi-level, depending on the problem. Most multi-objective planning models consider weighted coefficient methods and Pareto-based methods. The weighted coefficient method aims to transform a multi-objective problem into a single-objective model by using different weighting approaches. Examples include user-defined fixed weights, the analytic hierarchy process, stochastic weights, fuzzy mathematical methods, and bargaining functions. On the other hand, the Pareto-based method generates a Pareto-optimal frontier by means of a non-dominated ranking algorithm to evaluate possible solutions. A key advantage of this method is that all objectives can be taken into account with equal attention and a reasonable set of optimal solutions can be provided to decision-makers. For this reason, this method is often preferred. However, it requires more computational time compared with the weighted coefficient method.

2.3. Decision Variables

In this subsection, a review of traditional and flexible or active planning decision variables is conducted. The most used decision variables for planning are shown in Figure 3. Furthermore, Table 4 summarizes the main decision variables considered in distribution network planning models. These variables define the elements subject to optimization when expanding, reinforcing, or modernizing the network. They include the locations and sizes of substations and feeders, which determine network topology, load distribution, and expansion capacity; the placement of reserve feeders and interconnection switches, which enhance reliability and flexibility during contingencies; and the integration of renewable distributed generation, energy storage systems, and electric vehicle charging stations, which support flexibility, stability, and decarbonization objectives. Additionally, the locations and sizes of voltage control devices play a critical role in maintaining voltage stability and regulatory compliance, particularly under high renewable penetration scenarios.

https://doi.org/10.3390/en19010116

Table 4. Decision variables for distribution network planning.

Decision Variable	Description
Locations and sizes of new substations	Optimizes voltage regulation and load distribution. Supports network expansion and enhances reliability [21–25].
Upgrade of existing substations for reinforcement	Upgrades ensure the capacity to handle increased demand [23–25].
Locations and sizes of new feeders	Expands network capacity and reduces the load on existing feeders. Placement optimized using GIS and power flow analyses [23–26].
Upgrade of existing feeders for reinforcement	Prevents overloads as demand grows. Employs sensitivity analysis and cost-benefit analysis [24].
Locations of reserve feeders and interconnection switches	Provides flexibility for re-routing power during contingencies. Enhances system reliability and resilience [24].
Locations, sizes, and types of renewable distributed generations	Manages demand fluctuations, providing operational flexibility. Placement optimized for stability and efficiency [23,24].
Locations of new EV charging stations	Strategically placed to prevent local network strain. Supports efficient load balancing with increasing EV adoption [27].
Locations, sizes, and types of ESS	Stabilizes load by storing excess energy. Assists in peak shaving and integration of renewable sources.
Locations and sizes of voltage control devices	Maintains stable voltage levels in high DER penetration areas. Essential for voltage stability and regulatory compliance.

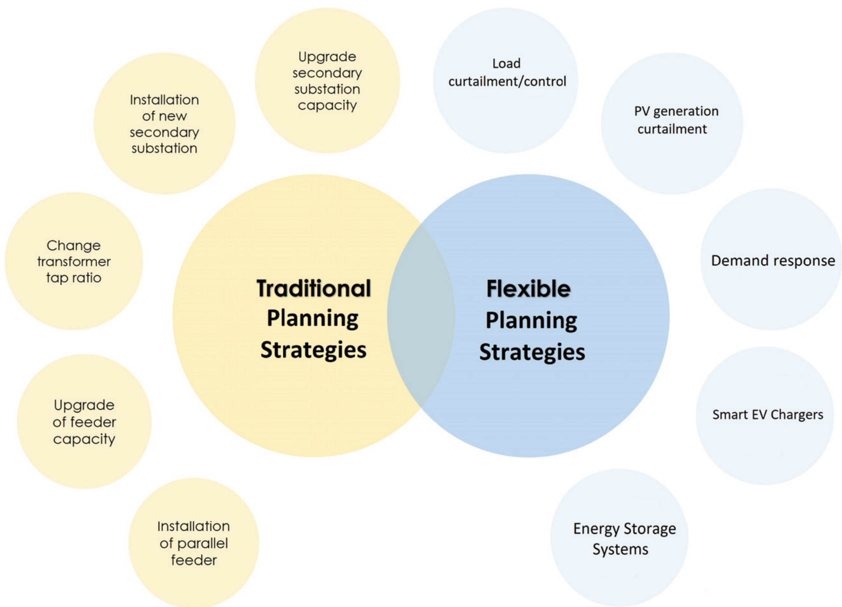


Figure 3. Traditional and flexible planning strategies in distribution networks planning.

2.3.1. Traditional Planning Strategies

Traditional distribution system planning strategies mainly involve upgrading existing infrastructure and adding new infrastructure. However, these approaches focus on adding infrastructure or enhancing existing components rather than adopting dynamic, real-time strategies. Figure 4 shows the upgrade of existing substations that address a critical overload at a secondary substation. This solution involves increasing the capacity of the existing substation, which also involves improving cooling systems. This strategy has the advantage of reinforcing the existing infrastructure without expanding the network, which can be more cost-effective and efficient for areas with space or environmental constraints. Another traditional distribution planning action to deal with capacity overloads in the existing infrastructure is to add new substations to redistribute the load. As shown in

https://doi.org/10.3390/en19010116

Figure 5, the system experiences a critical overload at a secondary substation connected to the external grid. When this occurs, a new secondary substation is added near the overloaded area, and some of this load is transferred from the existing substation to the new substation. This strategy is simple and effective in relieving immediate overload issues, is primarily reactive and requires continual infrastructure expansion by adding more substations as the demand increases to prevent future overloads. However, this approach may lead to a fragmented infrastructure and rising maintenance costs in the long term.

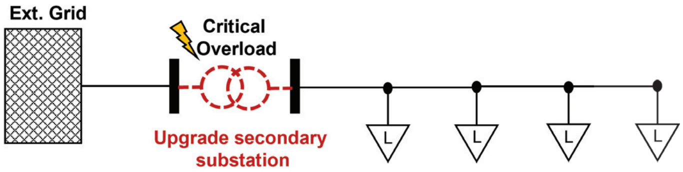


Figure 4. Upgrade of secondary substation.

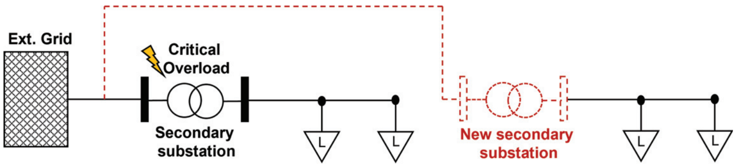


Figure 5. New secondary substation.

Figure 6 shows the feeder upgrade solution in red, which aims to increase the cross-sectional area for all overloaded branches. The increased feeder capacity allows for larger connected loads, which reduces the risk of outages, losses or voltage drops. Another planning action that addresses feeder overload conditions is adding a new feeder to support the existing infrastructure. As shown in Figure 7, when the main sections of the feeders experience overloads due to high demand from downstream loads, a new feeder is installed to divert a portion of this load. This feeder runs parallel to the existing network to alleviate the stress on the overloaded section of the original feeder, enhancing reliability and preventing potential failures. Lastly, the transformer tap adjustment is primarily used for voltage control, typically over seasonal or short-term planning horizons. In practice, during peak load periods, the tap is set to a higher position to boost downstream voltage and prevent undervoltage violations. During light-load periods or with reverse power flow from PVs, the tap is lowered to avoid overvoltage at the feeder ends.

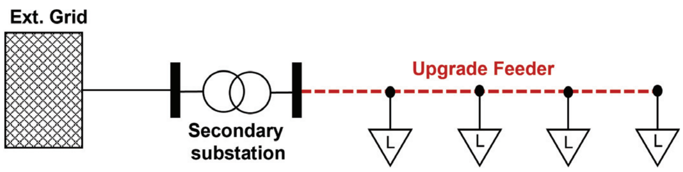


Figure 6. Upgrade of feeder.

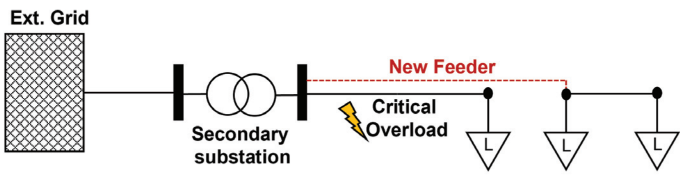


Figure 7. New feeder.

2.3.2. Flexible Planning Strategies

The flexibility in active distribution networks aims to increase the adaptability and responsiveness of the grid to manage the variable generation and demand. Table 5 compares the utilization of flexible sources and their impacts in active distribution networks.

Table 5. Overview of flexible sources in active distribution networks.

Flexible Source	Impact	Considerations
PV Generation	Voltage rise, reverse power flow under low load	Hosting capacity analysis, inverter-based voltage regulation, local control schemes
ESS	Peak shaving, arbitrage, reliability enhancement	Optimal siting/sizing, cost-benefit trade-offs, multi-period dispatch
EVCS and V2G	Rapid changes in load demand, potential feeder overload	EVCS siting, flexible charging strategies (time-of-use, V2G)
Demand Response	Demand shifting, improved load factor	Tariff and load control schemes

The solar PV integration is typically used for local voltage regulation, but it can also be triggered as a problem when generation exceeds local demand. However, modern inverters are able to control the reactive power to help mitigate these impacts. Wu et al. [28] address the challenges of long-term uncertainties and fluctuating load demand by considering flexible and dynamic strategies such as AC/DC feeders, voltage source converters for renewables, and static var generators. Electric vehicles present significant flexibility potential due to their ability to store energy in their batteries and the stationary nature of being parked more than 90% of the time. Vehicle-to-grid (V2G) technology enables the EV to inject stored energy from the battery into the grid. Proper charger siting and scheduling can unlock flexibility while minimizing distribution constraints. Ref. [29] demonstrates the economic value of EV flexibility in reducing peak demand levels and absorbing wind generation variability, and reveals that this value is enhanced with increasing electrification of the transport sector and increasing wind generation capacity. This work considered flexible EVs as a decision variable, and the V2G injections were optimally scheduled on the operational timescale.

Demand response initiatives reduce or shift load to off-peak periods. This flexible approach can be more cost-effective than major infrastructure upgrades in certain scenarios. Through the use of dynamic tariffs, incentives, or load control programs, demand response helps to avoid peak demand overloads and postpone high infrastructure investments. On the other hand, battery storage systems enable peak load shaving and efficient energy management. By storing excess energy during off-peak periods and releasing it during peak periods, this flexible source smooths load profiles and reduces stress on the grid. This is especially important in distribution networks with high levels of renewable energy. Ref. [30] proposed a flexible and stochastic coordinated planning model that uses different flexible sources, such as energy storage systems and demand response. This work also considers different uncertainties related to the loads, wind farms, and energy prices based on Monte Carlo simulations.

Figure 8 shows a new battery installation to address the critical overload at a secondary substation. The battery is optimally sized and placed to reduce the peak demands associated with the simultaneous charging of fast EV charging stations. The battery charging/discharging schedule permits avoiding overload in peak hours and charging the battery during low-demand hours. This planning strategy permits preventing the need for infrastructure upgrades. Another flexible planning alternative is shown in Figure 9.

https://doi.org/10.3390/en19010116

This graphical example shows how demand-side flexibility using dynamic tariffs can be coordinated with distributed energy resources.

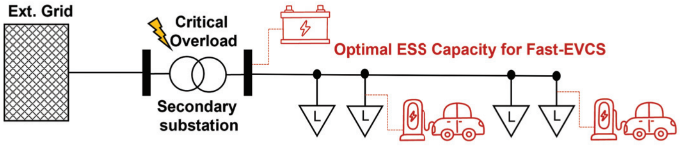


Figure 8. New Battery Installation for EVCS Peak Shaving.

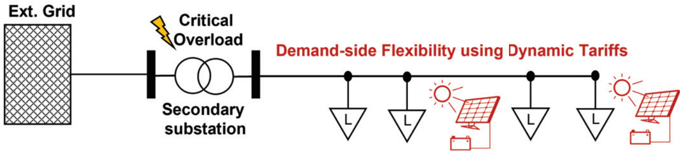


Figure 9. Demand-side flexibility.

2.4. Planning Constraints

The constraints are conditions that limit the solution space to obtain the optimal solution of the objective function. The number of constraints in distribution network planning models depends on the decision variables and scope of the problem. Table 6 shows four different groups of planning constraints and a representative constraints for each group. The groups are the following: technical constraints, non-technical constraints, time and investment constraints [3,31].

Table 6. Distribution network constraints.

Category	Constraints
Technical	Power balance equations, voltage magnitude, voltage angle, feeder/substation thermal limits, Battery state of charge limits, PV generation power, N-1 criterion.
Non-Technical	Location of assets, capacity of assets, quantity of assets per location.
Time and Investment	Capital expenditure, operational expenditure, static and dynamic states.

The most common constraints are related to technical constraints, including the bus voltage limits for nodes and buses, the power balance equations given by the Kirchoffs laws, the branch current flow constraints, thermal limits of the lines/substations, and the radiality of the distribution network [32–35]. Secondly, the non-technical constraints are limitations or requirements related to the practical aspects of distribution network planning. These include the logistical, environmental, allocation, quantity and capacity of transformers, feeders, substations, and circuit breakers, which must be optimized for efficient expansion [36,37]. Thirdly, the investment constraints involve budget limitations, purchased power restrictions, and fuel costs, which must be managed to minimize overall expenses. Also, the time constraints ensure operational stability by setting minimum periods for which a unit must remain on or off before changes in its operational state can occur, addressing scheduling and uncertainty considerations [32,38].

Lastly, the miscellaneous constraints consider the use of a single type of cable or transformer, distributed generation capacity and penetration level, and ramp rate limits for active and reactive power. They also address spatial area constraints, ensuring enough space for future equipment like transformers, and the reuse or removal of existing infrastructure [36,39].

https://doi.org/10.3390/en19010116

2.5. Type of Planning Variables

In optimal power flow (OPF) planning problems, understanding the types of variables involved is essential for problem formulation, computational efficiency, and the selection of solution techniques. Variables are typically classified into two categories: binary and continuous. This classification is based on their mathematical nature and role in representing different aspects of investment planning.

Table 7 shows the standard variables used in OPF planning problems. Binary variables represent decisions that involve only two states, such as whether to upgrade an existing asset, install a new asset, or activate a flexible asset. These variables are crucial for modeling investment choices that require a yes-or-no decision framework. Continuous variables represent quantities that can vary within a range of values, such as power losses, distributed generation injected power, branch power flow, bus voltage magnitudes, among others. These variables are critical for modeling the physical and operational behavior of the system, as well as for achieving efficiency and reliability in the network. By categorizing these variables, the figure highlights the diverse nature of the data and decisions involved in OPF planning. This structured approach facilitates the integration of technical, operational, and economic factors into the problem formulation.

Table 7. Types of planning variables.

Variable Type	Description
Binary	Installation of new secondary substations, upgrade of secondary substation capacity, installation of new ESS, installation of PV systems, installation of controlled EVCS, installation of parallel feeders, upgrade of feeder capacity, among others.
Continuous	Battery state of charge, substation power injection, voltage magnitude, voltage angle, branch power flow, generation power injection, load shedding, among others.

2.6. Uncertainty Modelling Techniques

This section reviews techniques for representing uncertainty in the planning of active distribution networks. Adequate planning of active distribution networks requires consideration of modeling uncertainty in their input parameters. Renewable generation, load, and EV charging, exhibit significant variability in magnitude, timing, and location [40,41].

The classification of uncertainty modelling techniques is shown in Table 8. At a high level, these can be divided into numerical and analytical approaches. Numerical methods often use Monte Carlo (MC) simulation, in which uncertain parameters are randomly sampled to generate a range of possible scenarios. Some MC simulation variants include sequential, pseudo-sequential, and non-sequential approaches. While these methods accurately capture complex stochastic behavior, they are computationally intensive. In contrast, analytical methods approximate probability distributions for faster evaluation. Examples include linearized expansions (e.g., Taylor series), point estimation techniques, cumulant-based approaches (balancing accuracy and efficiency), and unscented transformations used in active distribution network studies. Probabilistic methods assume that uncertainties, such as load growth and EV charging behavior follow known or estimated distributions, making them appropriate when historical data is available. Possibilistic (fuzzy) approaches use possibility distributions, ideal for sparse or incomplete data. Hybrid frameworks combine probabilistic and possibilistic elements to handle both well-characterized and poorly known uncertainties. Robust and stochastic optimization methods address uncertainty by optimizing network investments and operations under different scenarios, including

https://doi.org/10.3390/en19010116

worst-case conditions. Information Gap Decision Theory explores system performance under increasing uncertainty, making it valuable for long-term planning [41–44].

Table 8. Uncertainty modelling techniques.

Category	Methods
Probabilistic	1. Stochastic optimization 2. Robust optimization
Possibilistic	1. Possibilistic methods
Hybrid probabilistic-possibilistic	1. Combined probabilistic and possibilistic approaches
Information Gap Decision Theory	1. Decision-making under deep uncertainty
Monte Carlo simulations	1. Sequential: Simulation method with iterative updates 2. Pseudo-sequential: Hybrid approach between sequential and non-sequential 3. Non-sequential: Independent Monte Carlo runs without sequential updates
Analytical	1. Fuzzy-Monte Carlo 2. Fuzzy-scenario-based methods
Approximation of PDF	1. Convolution 2. Cumulants 3. Taylor Series Expansion 4. First-Order Second-Moment 5. Point Estimate Method 6. Unscented Transformation

In real use cases, DSOs must consider load forecast errors and EV charging uncertainties (such as arrival times, battery charge levels, and charging power). The choice of uncertainty modeling technique depends on data availability, computational resources, and the trade-off between accuracy and tractability [40,41].

3. Literature Review of OPF Planning Models

This section reviews various OPF formulations used in distribution network planning, along with studies that have employed these formulations. Deterministic and heuristic methods for solving OPF are significantly influenced by the underlying mathematical formulation, which may include non-linear programming, second-order cone programming, linear programming, or mixed-integer programming. Selecting the appropriate formulation depends on the application and the type of expansion planning problem. Additionally, the advantages and disadvantages of each OPF model are summarized in the final subsection.

3.1. Linear Programming

Linear Programming (LP) methods, such as DC-based OPF approaches, are widely used due to their simplicity, computational efficiency, and guaranteed convergence. These methods linearize the power flow equations, providing rapid solutions. However, the trade-off lies in reduced modeling accuracy, as linear approximations often fail to capture the non-linear and dynamic behavior of real-world power systems, particularly in distribution networks with high levels of distributed energy resources or electric vehicle integration.

A review of planning models related to substations and distribution feeders was studied in [45]. This work categorizes these models into two main areas: planning under standard operating conditions and planning for emergency scenarios. This classification allows for a structured analysis of how LP methods have been applied to various planning

https://doi.org/10.3390/en19010116

tasks. Moreover, the authors highlight potential research opportunities, particularly in enhancing the adaptability and robustness of these models for modern grid challenges. In [46], authors expand the application of LP methods in power system planning by categorizing distribution expansion models into four approaches based on time horizons and physical structure. For each category, they analyze the cost functions, constraints, and mathematical programming techniques employed. Their work provides valuable insights into the strengths of LP methods, such as computational efficiency and simplicity, as well as their limitations, including the inability to fully capture system non-linearities and complex interactions.

3.2. Mixed-Integer Linear Programming

Mixed-Integer Linear Programming (MILP) is a widely utilized optimization technique for problems that require modeling both continuous and discrete variables. In the context of distribution network planning, MILP is particularly effective for representing decisions such as the selection of asset types, capacities, and operational configurations. By incorporating binary or integer variables, MILP captures the discrete nature of many practical problems, including investment decisions, resource allocation, and system upgrades.

A bi-level optimization model using MILP to address distribution network expansion planning was developed by [47]. This approach integrates investment decisions, distributed energy resources, and energy storage systems into the optimization framework. This enables the model to propose solutions that are both cost-efficient and operationally feasible, using active network management strategies to accommodate real-time constraints. Similarly, ref. [48] expanded the application of MILP by introducing a multi-objective optimization framework that incorporates reliability and economic performance indices. This framework allows planners to evaluate trade-offs between minimizing costs and enhancing system reliability, offering a balanced and informed decision-making approach for distribution network planning. Authors in [49], introduced high penetration of plug-in electric vehicles (PEVs) in distribution networks. The study proposed a multi-stage bilevel MILP optimization model tailored to integrate PEVs into planning processes effectively. This hierarchical approach ensures that both long-term planning and short-term operational considerations are addressed, making it particularly suited to modern power systems with high levels of PEV adoption. The proposed model highlights the potential of MILP in enabling efficient integration of new technologies while maintaining economic and operational sustainability.

3.3. Mixed-Integer Nonlinear Programming

The earliest OPF formulations were inherently non-linear and accurately modeled power system characteristics, although they often relied on approximations, such as treating discrete variables as continuous. Modern non-linear formulations, however, frequently include discrete decision variables, resulting in Mixed-Integer Nonlinear Programming (MINLP). These advanced formulations address the inherent non-linearity while incorporating additional system constraints.

Uncertainties from large-scale deployment of electric vehicles and photovoltaic systems pose challenges to distribution network expansion planning. Work in [50], proposed an optimal planning method integrating PV-grid-EV transactions to mitigate these uncertainties. The study employs a peer-to-peer transactive market cleared via a decentralized algorithm, ensuring privacy and autonomy. Using multiple linearization techniques, the method ensures model convergence and enables cost-effective EV charging station planning while avoiding unnecessary PV curtailment. Results demonstrate improved integration of large-scale EVs and PV, enhancing network efficiency and security. In [51],

https://doi.org/10.3390/en19010116

authors presented an advanced MINLP model for handling the high penetration of electric vehicles in distribution networks, effectively addressing scalability challenges through scenario-based optimization. The study achieved a 42% reduction in computation time, demonstrating significant improvements in efficiency over traditional approaches.

3.4. Convex Relaxation Methods

Convex relaxations have gained prominence for their ability to achieve global optimality with faster convergence. These approaches reduce the solution space to find feasible, globally optimal solutions more reliably and can also provide bounds for the quality of local optima in the original problem. Among these, Second-Order Cone (SOC) relaxations are notable for their computational efficiency and widespread solver support. They are particularly suitable for radial network systems under certain constraints [52–54].

In [55], authors proposed a SOC relaxation to address the Optimal Power Flow problem for operational planning in active distribution networks. This approach integrates flexibility models, including On-Load Tap Changers and network reconfiguration, directly into the relaxed operational planning process. The model also considers the short-term economic impacts of utilizing these flexibilities, which are progressively applied in various test scenarios to achieve the goal of reducing operational costs. Authors in [56], the authors introduce a SOC optimization model aimed at minimizing the annualized social cost of the overall EV charging system using a two-step equivalence framework. This approach can solve the optimal planning problem of EVCSs, which comprise multiple types of charging facilities with different charging demand profiles. SOC relaxations are increasingly used for multi-stage ADN planning, ensuring tightness in solutions for DER-rich networks [14].

3.5. Metaheuristic Methods

Inspired by natural processes, meta-heuristic methods offer flexible solutions to complex problems. While they do not guarantee global optimality or error bounds, they can effectively handle discrete variables and large-scale systems. Popular meta-heuristic techniques include Genetic Algorithms, Particle Swarm Optimization, and Simulated Annealing, among others. These methods are particularly valuable for scenarios requiring robust solutions for operational planning. Ref. [57] propose a fuzzy-based meta-heuristic approach for multi-objective optimization in network planning, effectively addressing conflicting priorities like cost minimization and reliability enhancement. In [58] an optimal expansion of MV power networks to minimize the total investment costs was proposed. This work used a hybrid Tabu search/particle swarm optimization algorithm to optimize three electric distribution networks as case studies.

Table 9 presents the advantages and disadvantages of the distribution network planning formulations analyzed in this chapter. MINLP and MILP optimization models are the most used formulations for optimizing the total investment costs and meeting the operational constraints.

3.6. Dynamic OPF Planning Models and Tools with Flexibility

This section presents some examples of dynamic OPF planning models and tools that integrate flexible planning actions. Ref. [59] presents a comprehensive tool for optimal long-term grid planning considering power flow, demand flexibility model, and grid constraints. The parametrization of the generic flexibility model was based on real load demand data and selected individual demand flexibility resources. Furthermore, this work illustrates effects that are not captured by the more coarse-grained generic flexibility model but that are relevant for long-term grid planning purposes. Ref. [60] proposed FlexiPlan.jl, an open-source tool for holistic planning of transmission and distribution grids considering demand flexibility and storage use. This tool uses stochastic optimization to find robust decisions under different climate conditions and operating hours with respect power generation and demand.

Table 9. Comparison of mathematical formulations in OPF.

Formulation	Advantages	Disadvantages
Linear Programming	• Computationally efficient and scalable • Guarantees global optimality	• Low accuracy; oversimplifies non-linear power flows • Solutions may be physically infeasible
Mixed-Integer Linear Programming	• Handles discrete decisions (e.g., investment, switching) • Can be solved to global optimality (via branch-and-bound)	• High computational demand (NP-hard) • Relies on linearization, sacrificing physical accuracy
Mixed-Integer Non-Linear Programming	• High-fidelity: models AC power flow & discrete variables • Most realistic representation of the network	• Computationally intensive; complex to solve • Non-convex: risk of converging to local optima
Convex Relaxations	• Guarantees global optimality (if relaxation is exact) • Computationally tractable and reliable	• May have a "relaxation gap" (inaccurate solution) • Handling discrete variables requires MILP framework
Metaheuristic Methods	• Handles non-linear and non-convex models • Model-agnostic and good for large-scale problems	• No guarantee of global optimality • Solution quality is difficult to benchmark or prove

The FlexPlan methodology consists of three input modules. First, the grid data from the transmission and distribution networks is imported to model the AC and DC grids from PowerModels.jl. Second, a list of grid expansion options is provided using AC and DC grid expansions, demand flexibility, and storage investments to determine the optimal investments. Finally, the planning scenarios, number of renewable generation and demand time series are collected. Ref. [61] presented an open-source software tool called eGo, which is able to optimize grid and storage expansion at different voltage levels. Operating and investment costs are minimized by applying a multi-period linear optimal power flow considering the grid infrastructure. This tool performs simulation and optimization at the MV and LV levels to determine grid expansion costs, taking into account the allocation of curtailment requirements, storage integration, and traditional grid expansion measures. Ref. [62] proposes an automated distribution planning framework using the open-source library Pandapower. This library is a simulation model capable of running time series power flows to assess congestion and voltage deviation in the MV and LV network. A heuristic optimization that searches for the best combination of individual actions to solve the defined constraints using a set of reconfiguration, reinforcement and expansion actions is used to solve the potential problems in the network. This framework proposed the following planning actions: replacing existing lines and transformers, adding parallel lines to existing line paths, changing the switch configuration, finding new line paths, using advanced control functions on transformers and PV systems, replacing conventional transformers with On-Load-Tap-Changing transformers. Furthermore, the authors emphasise the benefits of using pandapower as a modular and flexible option, allowing constraints to be added such as, radiality, supply, n-1 and other topological constraints, load flow constraints for bus voltage, line load, transformer load, multiple worst case scenarios, load flow constraints for n-1 operation with optimal supply among others.

https://doi.org/10.3390/en19010116

4. ADN Trends: Generative AI Models, Applications and Future Asset Expansions

Active distribution networks must manage stochastic and dynamic uncertainties arising from variable renewable generation, fluctuating loads, and bidirectional power flows. Generative AI models have emerged as a powerful alternative to address these short-, mid-, and long-term challenges by generating realistic expansion scenarios, quantifying uncertainties, and supporting robust optimization. These models enable distribution system planners to optimize network expansion while integrating new technologies.

Table 10 provides an overview of the main categories of Generative AI models and their applications in energy systems planning and operation. The distinctive attributes and strengths of each category determine their suitability for particular applications. Generative Adversarial Networks (GAN)-based models are widely used for scenario generation due to their ability to learn complex, high-dimensional data distributions through adversarial training, which involves a generator and a discriminator. This setup enables GANs to produce highly realistic synthetic time series for renewable energy generation, electricity demand, or electric vehicle behavior. Their realistic nature and data-driven approach make them ideal for tasks such as data augmentation, cyber-resilience testing, and stochastic planning, where the objective is to expand existing datasets under different conditions [63–68]. The Sankey diagram in Figure 10 shows that different generative AI model families are connected to various application areas in the energy domain. The flow widths reflect their relative prevalence in the literature, showing that transformer-based models dominate across most applications—particularly planning and optimization, decision support, and energy management, while other generative models contribute more selectively to tasks such as scenario generation, data augmentation, resilience, and uncertainty modeling. For instance, the conditional style-based GAN proposed in [63] leverages meteorological variables to generate intraday renewable energy scenarios with spatiotemporal correlations, aiding stochastic planning by simulating volatile outputs without relying on extensive historical data. Notably, these models can produce diverse scenarios that capture rare but critical events, such as sudden weather changes. Authors in [64], demonstrated that the the cross-modal cGAN outperforms benchmark models on both PV and wind power datasets, even under missing data conditions. In a case study on stochastic day-ahead economic dispatch, scenarios generated by the cGAN yield an expected cost within 1.43% of the true scenarios, validating its practical applicability.

Diffusion-based models have recently emerged as a robust alternative to GAN models. They work by learning to reverse a gradual noising process: starting from pure random noise, the model iteratively predicts and removes noise in small steps until a coherent, high-fidelity image (or data sample) emerges. This iterative denoising makes them particularly effective at representing uncertainty and extreme events. Their probabilistic formulation naturally lends itself to applications such as uncertainty quantification, stress-testing, and resilience analysis in energy systems, where understanding variability and extreme scenarios is essential for reliability and planning under uncertainty [69–73]. As an example, ref. [69] used a denoising diffusion probabilistic approach to generate realistic EV charging scenarios at both battery-level and station-level, capturing temporal dynamics, station-specific patterns, and intractable uncertainties. Quantitative results show that the proposed DiffCharge model achieves the lowest marginal score, highest discriminative score, and best tail score, enabling accurate uncertainty modeling for distribution network planning. VAE-based models, including β-VAE, provide probabilistic frameworks for forecasting and optimization under uncertainty. Variational autoencoder (VAE)-based models combine deep learning and probabilistic inference in order to learn how to compress and reconstruct data in a lower-dimensional latent space. They consist of

https://doi.org/10.3390/en19010116

two main parts: an encoder, which maps input data into a distribution in the latent space (typically Gaussian), and a decoder, which reconstructs data samples from points drawn from this latent space. By learning this bidirectional mapping, VAEs can generate realistic and continuous variations of data while maintaining interpretability. Their probabilistic nature makes them well-suited for tasks involving scenario synthesis and forecasting under uncertainty—such as load or price prediction, energy efficiency planning, and voltage stability assessment [74–77]. Research in [74] applies a VAE-GAN hybrid for synthetic data generation in smart homes, leveraging the latent space to create new data points that preserve privacy while enabling efficient load and price forecasting for optimal power flows in ADNs. Furthermore, VAEs facilitate the compression of high-dimensional data, reducing the need for large datasets and making them suitable for edge computing in distributed ADN environments. Flow-based models, such as RealNVP, offer explicit density estimation for renewable and load scenarios. As explored in [78], these models generate interpretable renewable patterns with controllable diversity, supporting probabilistic OPF and expansion planning by modeling nonlinear correlations and geospatial dependencies in DER placement. Furthermore, this work showed that flow-based models improves coverage probabilities by up to 4 percentage points over VAEs and 3 points over GANs, while reducing interval widths by up to 7% compared to VAEs and 4% compared to GANs, indicating more reliable forecasts with sharper uncertainty intervals.

Table 10. Overview of generative AI models and applications.

Category	Type of Generative AI Model	Application	Ref.
GAN-based models	cGAN, WGAN-GP, CycleGAN, StyleGAN, ExGAN, BiGAN	Scenario generation (renewable, load, EV), data augmentation, cyber-resilience, fault diagnosis. Synthetic time-series for PV, wind, load, and EVs to support stochastic and long-term planning	[63–68]
Diffusion-based models	DiffCharge, DiffLoad, ExDiffusion, Conditional Diffusion	EV/load scenario generation, uncertainty quantification, extreme event modeling, resilience analysis. Stress-testing, extreme event simulation, reliability planning.	[69–73]
VAE-based models	β-VAE, Conditional VAE, BiVAE	Probabilistic forecasting, scenario synthesis, energy efficiency planning, voltage stability assessment. Probabilistic load/price forecasting and OPF optimization under uncertainty.	[74–77]
Flow-based models	RealNVP, Glow, Normalizing Flow	Renewable and load scenario generation, uncertainty modeling, probabilistic OPF, expansion planning.	[78–81]
Transformer-based models	GPT, LLaMA, eGridGPT, CPGA-BOT, PowerPulse	Planning assistants, data mining, simulation automation, energy management, human-in-the-loop decision support. Decision support via Large Language Models (LLMs), planning chatbots, automated data mining, and simulation scripting.	[82–101]

Unlike GANs, VAEs provide more stable training and a structured latent representation that can be directly exploited for optimization and decision support. Flow-based models use a sequence of invertible transformations to map simple probability distributions into complex data distributions and vice versa. Each transformation is mathematically reversible, allowing both exact likelihood computation and precise control over the generative process. This bidirectional and transparent structure enables direct sampling, density estimation, and interpretation of how data are formed. Such properties make flow-based models highly suitable for renewable and load scenario generation, probabilistic optimal power flow, and expansion planning, where accurate probabilistic modeling and explicit likelihood estimation are required [78–81].

Finally, Transformer-based models—such as GPT, LLaMA, and specialized versions like eGridGPT and PowerPulse—extend generative AI beyond data synthesis to include reasoning, automation, and decision support. They are built around self-attention mechanisms, which allow the model to weigh the importance of each element in a sequence relative to all others. In practice, this means the model learns which parts of the input are most relevant when generating or interpreting each output token, enabling it to capture long-range dependencies, contextual relationships, and nonlinear interactions across complex datasets. This capability makes transformers highly effective for planning assistance, data mining, simulation management, and human-in-the-loop decision support in energy systems. Their flexibility positions them as foundational tools for next-generation applications that merge language understanding with generative capabilities, such as intelligent planning assistants and simulation chatbots [82–101]. The eGridGPT framework [85] represents one of the first applications of LLMs in power grid control rooms, integrating large language models with digital twins for procedure analysis, scenario simulation, and holistic decision recommendations. Similarly, PowerPulse [100], fine-tuned on power sector data, automates responses to energy-related queries and supports adaptation in dynamic grids. A significant development is their integration into multi-agent systems for VPP coordination, where LLMs can process ambiguous data and enhance security in smart grids.

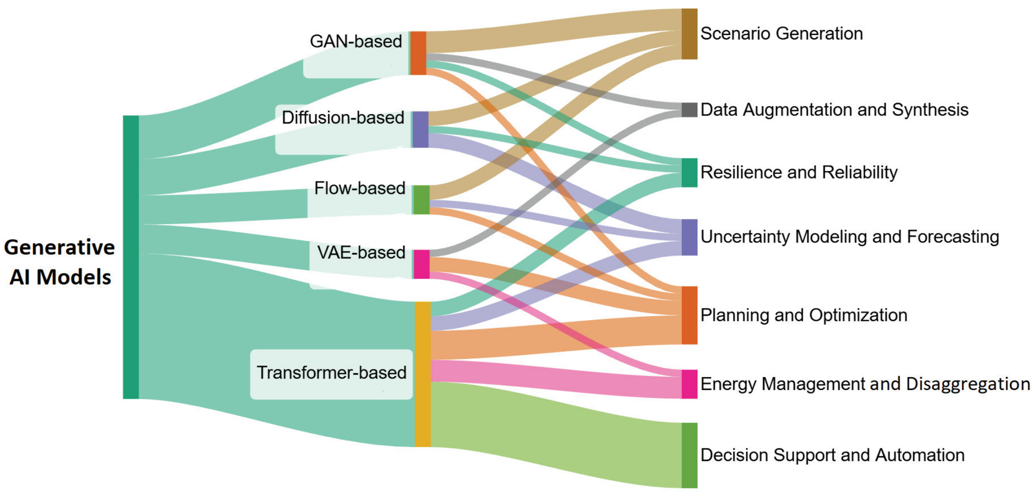


Figure 10. Sankey diagram of applications of generative AI models in power systems.

Each GenAI model category offers distinct strengths: GANs specialize in producing highly realistic data, diffusion models in modeling uncertainty and extreme events through iterative denoising, VAEs in generating interpretable probabilistic representations of complex systems, flow-based models in providing exact and transparent likelihood estimation, and transformers in capturing context, reasoning across data, and enabling automation—together forming a powerful, complementary framework for next-generation intelligent energy system planning and decision support. The ADN planning problem is mainly affected by long-term uncertainty from renewables, EVs, and flexible loads, making traditional planning no longer valid for realistic investment estimations. Generative AI can generate realistic scenarios by capturing complex spatio-temporal patterns from emerging assets, thereby improving robustness and decision-making for system operators. Recent works have integrated generative model into planning of active distribution networks by considering it into the methodology steps. In particular, ref. [102] proposes a cooperative planning method for renewable energy generations and multi-timescale flexible sources in ADN. It uses a Wasserstein generative adversarial network with gradient penalty based scenario generation and a second-order cone relaxation approach for optimization. The case study results in a 69 node network showed a 7.9% reduction over single-timescale planning. The main improvements were found in the reduction of the curtailment/shedding penalties by over 75% and voltage deviation by 2%. Ref. [103] proposes a GAN-assisted two-stage stochastic PV planning framework with coordinated multi-timescale Volt–Var control, validated on IEEE 37- and 123-node systems. Case studies show that, unlike deterministic planning, the stochastic model eliminates voltage violations under critical high-PV/low-load scenarios and reduces total power losses across most time slots, particularly around peak PV hours.

5. Conclusions

This work reviews the key elements of active distribution system planning, classifying critical components from the literature, including planning horizons, objectives, decision variables, constraints, uncertainty modeling, and OPF formulations. The review finds that traditional planning is characterized by passive "fit and forget" reinforcement and worst-case assumptions, whereas emerging active strategies prioritize DER integration, system flexibility, and dynamic optimization. This highlights a critical need for evolution, demanding advancements in AC-OPF formulations, new asset modeling frameworks, and robust uncertainty management. The literature also reveals that addressing complex ADN problems requires enhancing traditional planning with active decision variables, dynamic scenarios, probabilistic methods, and advanced optimization, such as MINLP and convex relaxation.

Furthermore, this review discusses applications of generative AI in active distribution planning, including the creation of realistic, synthetic load profiles to simulate future demand and optimize DER integration. This approach enables privacy-preserving data sharing, helping to bridge the gap between research and industry by facilitating innovative, adapted optimization models for modern power systems. Finally, the review reveals that active distribution planning faces critical challenges from the accelerating integration of data centers, AI factories, renewable energy communities, and fast-charging hubs. While functionally different, these assets create a common, critical threat: the introduction of high-magnitude, fast-ramping events. This new volatility lies outside the design parameters of traditional grid operations, leading to severe, new risks of congestion and instability.

Future research should prioritize the development of advanced integrated ADN planning models that address long-term uncertainty arising from renewables, EV charging, and flexible loads. These models should leverage generative AI for realistic spatio-temporal scenario generation to support robust and stochastic planning. Additionally, multi-timescale and cooperative planning frameworks are needed to jointly optimize investment and operational flexibility for all the planning horizons.

Author Contributions: Conceptualization, A.E.S.-G.; Literature search, A.E.S.-G.; Formal analysis and visualization, A.E.S.-G.; Validation, A.S., V.G. and M.A.-P.; Writing—original draft preparation, A.E.S.-G.; Writing—review and editing, A.S., V.G. and M.A.-P.; Supervision, A.S. and M.A.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was part of the project PLATON funded by the European Union (NextGenerationEU) and the Ministry for Digital Transformation and Public Administration, under the UNICO I + D Cloud call (Grant No. TSI-063100-2022-010), corresponding to the project "Federated Machine Learning for Electrical Distribution Networks" at the Polytechnic University of Catalonia.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AC	Alternating Current
ADN	Active Distribution Network
AI	Artificial Intelligence
CAPEX	Capital Expenditure
DC	Direct Current
DER	Distributed Energy Resources
DG	Distributed Generation
DR	Demand Response
DSO	Distribution System Operator
ESS	Energy Storage System
EV	Electric Vehicle
EVCS	Electric Vehicle Charging Station
GAN	Generative Adversarial Network
GenAI	Generative Artificial Intelligence
GIS	Geographic Information System
IGDT	Information Gap Decision Theory
LLM	Large Language Model
LP	Linear Programming
LV	Low Voltage
MC	Monte Carlo
MILP	Mixed-Integer Linear Programming
MINLP	Mixed-Integer Nonlinear Programming
MV	Medium Voltage
NLP	Nonlinear Programming
OPEX	Operational Expenditure
OPF	Optimal Power Flow
PDF	Probability Density Function
PEV	Plug-in Electric Vehicle
PV	Photovoltaic
SOC	Second-Order Cone
TOTEX	Total Expenditure
V2G	Vehicle-to-Grid
VAE	Variational Autoencoder
VPP	Virtual Power Plant

References

1. Jiang, M.; Xu, Y.; Zhang, D.; Liu, Y.; Du, Q.; Gao, X.; Qi, S.; Zou, H. Research on an Active Distribution Network Planning Strategy Considering Diversified Flexible Resource Allocation. *Processes* **2025**, *13*, 2254. [CrossRef]
2. Wang, K.; Huang, Y.; Liu, Y.; Huang, T.; Zang, S. A Review of Optimization Scheduling for Active Distribution Networks with High-Penetration Distributed Generation Access. *Energies* **2025**, *18*, 4119. [CrossRef]
3. Rastgou, A. Distribution network expansion planning: An updated review of current methods and new challenges. *Renew. Sustain. Energy Rev.* **2024**, *189*, 114062. [CrossRef]

4. Ghosh, P.; De, M. Resilience-oriented planning for active distribution systems: A probabilistic approach considering regional weather profiles. *Int. J. Electr. Power Energy Syst.* **2024**, *158*, 109976. [CrossRef]
5. Pan, C.; Chen, W.; Zhou, C.; Zhou, W. Distribution network planning considering active response of EVs and DTR of cables and transformers. *High Volt.* **2025**, *10*, 279–293. [CrossRef]
6. Anadón Martínez, V.; Sumper, A.; Saldaña-Gonzalez, A.E.; Gadelha, V. Planning fast-charging stations along highways using probability distribution functions and traffic data. *Sustain. Energy Technol. Assess.* **2025**, *82*, 104547. [CrossRef]
7. Saldaña-González, A.E.; Aragüés-Peñalba, M.; Sumper, A. Distribution network planning method: Integration of a recurrent neural network model for the prediction of scenarios. *Electr. Power Syst. Res.* **2024**, *229*, 110125. [CrossRef]
8. *Distributed Flexibility: Lessons Learned in the Nordics*; Nordic Energy Research: Oslo, Norway, 2022.
9. Li, R.; Wang, W.; Chen, Z.; Jiang, J.; Zhang, W. A review of optimal planning active distribution system: Models, methods, and future researches. *Energies* **2017**, *10*, 1715. [CrossRef]
10. Borousan, F.; Hamidan, M.A. Distributed power generation planning for distribution network using chimp optimization algorithm in order to reliability improvement. *Electr. Power Syst. Res.* **2023**, *217*, 109109. [CrossRef]
11. de Lima, T.D.; Soares, J.; Lezama, F.; Franco, J.F.; Vale, Z. A risk-based planning approach for sustainable distribution systems considering EV charging stations and carbon taxes. *IEEE Trans. Sustain. Energy* **2023**, *14*, 2294–2307. [CrossRef]
12. de Lima, T.D.; Tabares, A.; Arias, N.B.; Franco, J.F. Investment & generation costs vs. CO_2 emissions in the distribution system expansion planning: A multi-objective stochastic programming approach. *Int. J. Electr. Power Energy Syst.* **2021**, *131*, 106925. [CrossRef]
13. Hou, H.; Tang, J.; Zhang, Z.; Wang, Z.; Wei, R.; Wang, L.; He, H.; Wu, X. Resilience enhancement of distribution network under typhoon disaster based on two-stage stochastic programming. *Appl. Energy* **2023**, *338*, 120892. [CrossRef]
14. Zhou, S.; Han, Y.; Chen, S.; Yang, P.; Mahmoud, K.; Darwish, M.M.; Matti, L.; Zalhaf, A.S. A multiple uncertainty-based bi-level expansion planning paradigm for distribution networks complying with energy storage system functionalities. *Energy* **2023**, *275*, 127511. [CrossRef]
15. Ali, M.H.; Kamel, S.; Hassan, M.H.; Tostado-Véliz, M.; Zawbaa, H.M. An improved wild horse optimization algorithm for reliability based optimal DG planning of radial distribution networks. *Energy Rep.* **2022**, *8*, 582–604. [CrossRef]
16. Pal, A.; Bhattacharya, A.; Chakraborty, A.K. Planning of EV charging station with distribution network expansion considering traffic congestion and uncertainties. *IEEE Trans. Ind. Appl.* **2023**, *59*, 3810–3825. [CrossRef]
17. Ahmadi, S.E.; Kazemi-Razi, S.M.; Marzband, M.; Ikpehai, A.; Abusorrah, A. Multi-objective stochastic techno-economic-environmental optimization of distribution networks with G2V and V2G systems. *Electr. Power Syst. Res.* **2023**, *218*, 109195. [CrossRef]
18. Zare, P.; Dejamkhooy, A.; Davoudkhani, I.F. Efficient expansion planning of modern multi-energy distribution networks with electric vehicle charging stations: A stochastic MILP model. *Sustain. Energy, Grids Networks* **2024**, *38*, 101225. [CrossRef]
19. Subbaramaiah, K.; Sujatha, P. Optimal DG unit placement in distribution networks by multi-objective whale optimization algorithm & its techno-economic analysis. *Electr. Power Syst. Res.* **2023**, *214*, 108869.
20. Lin, J.; Qiu, J.; Tao, Y.; Sun, X. Planning of Electric Vehicle Charging Stations with PV and Energy Storage Using a Fuzzy Inference System. *IEEE Trans. Transp. Electrif.* **2023**, *10*, 5894–5909. [CrossRef]
21. Ganguly, S.; Sahoo, N.C.; Das, D. Mono-and multi-objective planning of electrical distribution networks using particle swarm optimization. *Appl. Soft Comput.* **2011**, *11*, 2391–2405. [CrossRef]
22. Ganguly, S.; Sahoo, N.; Das, D. Multi-objective planning of electrical distribution systems using dynamic programming. *Int. J. Electr. Power Energy Syst.* **2013**, *46*, 65–78. [CrossRef]
23. Xing, H.; Cheng, H.; Zhang, L.; Zhang, S.; Zhang, Y. Second-order cone model for active distribution network expansion planning. In Proceedings of the 2015 IEEE Power & Energy Society General Meeting, Denver, CO, USA, 26–30 July 2015; pp. 1–5.
24. Bagheri, A.; Monsef, H.; Lesani, H. Integrated distribution network expansion planning incorporating distributed generation considering uncertainties, reliability, and operational conditions. *Int. J. Electr. Power Energy Syst.* **2015**, *73*, 56–70. [CrossRef]
25. Xing, H.; Cheng, H.; Zhang, Y.; Zeng, P. Active distribution network expansion planning integrating dispersed energy storage systems. *IET Gener. Transm. Distrib.* **2016**, *10*, 638–644. [CrossRef]
26. Gitizadeh, M.; Vahed, A.A.; Aghaei, J. Multistage distribution system expansion planning considering distributed generation using hybrid evolutionary algorithms. *Appl. Energy* **2013**, *101*, 655–666. [CrossRef]
27. Lin, X.; Sun, J.; Ai, S.; Xiong, X.; Wan, Y.; Yang, D. Distribution network planning integrating charging stations of electric vehicle with V2G. *Int. J. Electr. Power Energy Syst.* **2014**, *63*, 507–512. [CrossRef]
28. Wu, Z.; Sun, Q.; Gu, W.; Chen, Y.; Xu, H.; Zhang, J. AC/DC hybrid distribution system expansion planning under long-term uncertainty considering flexible investment. *IEEE Access* **2020**, *8*, 94956–94967. [CrossRef]
29. Ramirez, P.G.; Papadaskalopoulos, D.; Strbac, G. Co-Optimization of Generation Expansion Planning and Electric Vehicles Flexibility. *IEEE Trans. Smart Grid* **2016**, *7*, 1609–1619. [CrossRef]

https://doi.org/10.3390/en19010116

30. Hamidpour, H.; Aghaei, J.; Pirouzi, S.; Dehghan, S.; Niknam, T. Flexible, reliable, and renewable power system resource expansion planning considering energy storage systems and demand response programs. *IET Renew. Power Gener.* **2019**, *13*, 1862–1872. [CrossRef]
31. Verma, M.K.; Mukherjee, V.; Yadav, V.K.; Ghosh, S. Constraints for effective distribution network expansion planning: An ample review. *Int. J. Syst. Assur. Eng. Manag.* **2020**, *11*, 531–546. [CrossRef]
32. Xie, S.; Hu, Z.; Yang, L.; Wang, J. Expansion planning of active distribution system considering multiple active network managements and the optimal load-shedding direction. *Int. J. Electr. Power Energy Syst.* **2020**, *115*, 105451. [CrossRef]
33. Nahman, J.M.; Perić, D.M. Radial distribution network planning under uncertainty by applying different reliability cost models. *Int. J. Electr. Power Energy Syst.* **2020**, *117*, 105655. [CrossRef]
34. Bhadoria, V.S.; Pal, N.S.; Shrivastava, V. Artificial immune system based approach for size and location optimization of distributed generation in distribution system. *Int. J. Syst. Assur. Eng. Manag.* **2019**, *10*, 339–349. [CrossRef]
35. Verma, M.K.; Yadav, V.K.; Mukherjee, V.; Ghosh, S. A multi-criteria approach for distribution network expansion through pooled MCDEA and Shannon entropy. *Int. J. Emerg. Electr. Power Syst.* **2019**, *20*, 20190043. [CrossRef]
36. Ugranlı, F. Analysis of renewable generation's integration using multi-objective fashion for multistage distribution network expansion planning. *Int. J. Electr. Power Energy Syst.* **2019**, *106*, 301–310. [CrossRef]
37. de Quevedo, P.M.; Munoz-Delgado, G.; Contreras, J. Impact of electric vehicles on the expansion planning of distribution systems considering renewable energy, storage, and charging stations. *IEEE Trans. Smart Grid* **2017**, *10*, 794–804. [CrossRef]
38. Mehrjerdi, H. Simultaneous load leveling and voltage profile improvement in distribution networks by optimal battery storage planning. *Energy* **2019**, *181*, 916–926. [CrossRef]
39. Pombo, A.V.; Murta-Pina, J.; Pires, V.F. Multiobjective formulation of the integration of storage systems within distribution networks for improving reliability. *Electr. Power Syst. Res.* **2017**, *148*, 87–96. [CrossRef]
40. Ehsan, A.; Yang, Q. State-of-the-art techniques for modelling of uncertainties in active distribution network planning: A review. *Appl. Energy* **2019**, *239*, 1509–1523. [CrossRef]
41. Aien, M.; Hajebrahimi, A.; Fotuhi-Firuzabad, M. A comprehensive review on uncertainty modeling techniques in power system studies. *Renew. Sustain. Energy Rev.* **2016**, *57*, 1077–1089. [CrossRef]
42. Aien, M.; Rashidinejad, M.; Fotuhi-Firuzabad, M. On possibilistic and probabilistic uncertainty assessment of power flow problem: A review and a new approach. *Renew. Sustain. Energy Rev.* **2014**, *37*, 883–895. [CrossRef]
43. Singh, V.; Moger, T.; Jena, D. Uncertainty handling techniques in power systems: A critical review. *Electr. Power Syst. Res.* **2022**, *203*, 107633. [CrossRef]
44. Ebeed, M.; Aleem, S.H.A. Overview of uncertainties in modern power systems: Uncertainty models and methods. In *Uncertainties in Modern Power Systems*; Academic Press: Cambridge, MA, USA, 2021; pp. 1–34.
45. Khator, S.K.; Leung, L.C. Power distribution planning: A review of models and issues. *IEEE Trans. Power Syst.* **1997**, *12*, 1151–1159. [CrossRef]
46. Temraz, H.K.; Quintana, V.H. Distribution system expansion planning models: An overview. *Electr. Power Syst. Res.* **1993**, *26*, 61–70. [CrossRef]
47. Kabirifar, M.; Fotuhi-Firuzabad, M.; Moeini-Aghtaie, M.; Pourghaderi, N.; Dehghanian, P. A bi-level framework for expansion planning in active power distribution networks. *IEEE Trans. Power Syst.* **2021**, *37*, 2639–2654. [CrossRef]
48. Qin, Y.; Zheng, J.; Wu, Q. Many-objective interactive optimization and decision making for distribution network expansion planning. *Control Eng. Pract.* **2021**, *116*, 104917. [CrossRef]
49. Abdi-Siab, M.; Lesani, H. Distribution expansion planning in the presence of plug-in electric vehicle: A bilevel optimization approach. *Int. J. Electr. Power Energy Syst.* **2020**, *121*, 106076. [CrossRef]
50. Yao, H.; Xiang, Y.; Gu, C.; Liu, J. Optimal planning of distribution systems and charging stations considering PV-Grid-EV transactions. *IEEE Trans. Smart Grid* **2024**, *16*, 691–703. [CrossRef]
51. Wang, X.; Shi, X.; Xu, Y.; Shen, X. Distribution network expansion planning approach for large scale electric vehicles accommodation. *IEEE Trans. Sustain. Energy* **2022**, *14*, 906–919. [CrossRef]
52. Saldaña-González, A.E.; Sumper, A.; Anadón-Martínez, V.; Aragüés-Peñalba, M. A machine learning approach for EVCS integration in distribution network based on optimal investment actions. *Electr. Power Syst. Res.* **2026**, *252*, 112376. [CrossRef]
53. Gadelha, V.; Bullich-Massagué, E.; Sumper, A. Optimal Power Flow in electrical grids based on power routers. *Electr. Power Syst. Res.* **2024**, *234*, 110581. [CrossRef]
54. Gadelha, V.; Sumper, A.; Bullich-Massague, E. Impact of converter losses in optimal power flow for hybrid AC-DC networks based on power routers. *CSEE J. Power Energy Syst.* **2025**, 1–12. [CrossRef]
55. Swaminathan, B.; Debusschere, V.; Caire, R. Short-term active distribution network operation with convex formulations of power flow. In Proceedings of the 2017 IEEE Manchester PowerTech, Manchester, UK, 18–22 June 2017; pp. 1–6.
56. Luo, L.; Gu, W.; Zhou, S.; Huang, H.; Gao, S.; Han, J.; Wu, Z.; Dou, X. Optimal planning of electric vehicle charging stations comprising multi-types of charging facilities. *Appl. Energy* **2018**, *226*, 1087–1099. [CrossRef]

https://doi.org/10.3390/en19010116

57. Gholizadeh-Roshanagh, R.; Zare, K.; Marzband, M. An a-posteriori multi-objective optimization method for MILP-based distribution expansion planning. *IEEE Access* **2020**, *8*, 60279–60292. [CrossRef]
58. Ahmadian, A.; Elkamel, A.; Mazouz, A. An improved hybrid particle swarm optimization and tabu search algorithm for expansion planning of large dimension electric distribution network. *Energies* **2019**, *12*, 3052. [CrossRef]
59. Bødal, E.F.; Lakshmanan, V.; Sperstad, I.B.; Degefa, M.Z.; Hanot, M.; Ergun, H.; Rossi, M. Demand flexibility modelling for long term optimal distribution grid planning. *IET Gener. Transm. Distrib.* **2022**, *16*, 5002–5014. [CrossRef]
60. Rossini, M.; Ergun, H.; Rossi, M. FlexPlan. jl-An open-source Julia tool for holistic transmission and distribution grid planning. In Proceedings of the 2023 Open Source Modelling and Simulation of Energy Systems (OSMSES), Aachen, Germany, 27–29 March 2023; pp. 1–8.
61. Müller, U.P.; Schachler, B.; Scharf, M.; Bunke, W.D.; Günther, S.; Bartels, J.; Pleßmann, G. Integrated techno-economic power system planning of transmission and distribution grids. *Energies* **2019**, *12*, 2091. [CrossRef]
62. Scheidler, A.; Thurner, L.; Braun, M. Heuristic optimisation for automated distribution system planning in network integration studies. *IET Renew. Power Gener.* **2018**, *12*, 530–538. [CrossRef]
63. Yuan, R.; Wang, B.; Sun, Y.; Song, X.; Watada, J. Conditional style-based generative adversarial networks for renewable scenario generation. *IEEE Trans. Power Syst.* **2022**, *38*, 1281–1296. [CrossRef]
64. Kang, M.; Zhu, R.; Chen, D.; Li, C.; Gu, W.; Qian, X.; Yu, W. A cross-modal generative adversarial network for scenarios generation of renewable energy. *IEEE Trans. Power Syst.* **2023**, *39*, 2630–2640. [CrossRef]
65. Wang, Z.; Hong, T. Generating realistic building electrical load profiles through the Generative Adversarial Network (GAN). *Energy Build.* **2020**, *224*, 110299. [CrossRef]
66. Sun, Y.; Wang, B.; Yuan, R.; Watada, J. Rolling unit commitment based on dual-discriminator conditional generative adversarial networks. *Electr. Power Syst. Res.* **2022**, *205*, 107770. [CrossRef]
67. Bhatia, S.; Jain, A.; Hooi, B. Exgan: Adversarial generation of extreme samples. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtually, 19–21 May 2021; Volume 35, pp. 6750–6758.
68. Kaselimi, M.; Doulamis, N.; Voulodimos, A.; Doulamis, A.; Protopapadakis, E. EnerGAN++: A generative adversarial gated recurrent network for robust energy disaggregation. *IEEE Open J. Signal Process.* **2020**, *2*, 1–16. [CrossRef]
69. Li, S.; Xiong, H.; Chen, Y. Diffcharge: Generating ev charging scenarios via a denoising diffusion model. *IEEE Trans. Smart Grid* **2024**, *15*, 3936–3949. [CrossRef]
70. Wang, Z.; Zhang, H. Customized load profiles synthesis for electricity customers based on conditional diffusion models. *IEEE Trans. Smart Grid* **2024**, *15*, 4259–4270. [CrossRef]
71. Dong, X.; Mao, Z.; Sun, Y.; Xu, X. Short-term wind power scenario generation based on conditional latent diffusion models. *IEEE Trans. Sustain. Energy* **2023**, *15*, 1074–1085. [CrossRef]
72. Wang, Z.; Wen, Q.; Zhang, C.; Sun, L.; Wang, Y. DiffLoad: Uncertainty quantification in electrical load forecasting with the diffusion model. *IEEE Trans. Power Syst.* **2024**, *40*, 1777–1789. [CrossRef]
73. Boulaguiem, Y.; Zscheischler, J.; Vignotto, E.; van der Wiel, K.; Engelke, S. Modeling and simulating spatial extremes by combining extreme value theory with generative adversarial networks. *Environ. Data Sci.* **2022**, *1*, e5. [CrossRef]
74. Razghandi, M.; Zhou, H.; Erol-Kantarci, M.; Turgut, D. Variational autoencoder generative adversarial network for synthetic data generation in smart home. In Proceedings of the ICC 2022-IEEE International Conference on Communications, Seoul, Republic of Korea, 16–20 May 2022; pp. 4781–4786.
75. Qi, Y.; Hu, W.; Dong, Y.; Fan, Y.; Dong, L.; Xiao, M. Optimal configuration of concentrating solar power in multienergy power systems with an improved variational autoencoder. *Appl. Energy* **2020**, *274*, 115124. [CrossRef]
76. Sharif, S.A.; Hammad, A.; Eshraghi, P. Generation of whole building renovation scenarios using variational autoencoders. *Energy Build.* **2021**, *230*, 110520. [CrossRef]
77. Langevin, A.; Carbonneau, M.A.; Cheriet, M.; Gagnon, G. Energy disaggregation using variational autoencoders. *Energy Build.* **2022**, *254*, 111623. [CrossRef]
78. Hu, S.; Zhu, R.; Li, G.; Song, L. Scenario forecasting for wind power using flow-based generative networks. *Energy Rep.* **2021**, *7*, 369–377. [CrossRef]
79. Ge, L.; Liao, W.; Wang, S.; Bak-Jensen, B.; Pillai, J.R. Modeling daily load profiles of distribution network for scenario generation using flow-based generative network. *IEEE Access* **2020**, *8*, 77587–77597. [CrossRef]
80. Cramer, E.; Paeleke, L.; Mitsos, A.; Dahmen, M. Normalizing flow-based day-ahead wind power scenario generation for profitable and reliable delivery commitments by wind farm operators. *Comput. Chem. Eng.* **2022**, *166*, 107923. [CrossRef]
81. Dumas, J.; Wehenkel, A.; Lanaspeze, D.; Cornélusse, B.; Sutera, A. A deep generative model for probabilistic energy forecasting in power systems: Normalizing flows. *Appl. Energy* **2022**, *305*, 117871. [CrossRef]
82. Yan, Z.; Zhenyuan, D.; Xu, Y.; Zhou, Z. Probabilistic PV power forecasting by a multi-modal method using GPT-agent to interpret weather conditions. In Proceedings of the 2024 IEEE 19th Conference on Industrial Electronics and Applications (ICIEA), Kristiansand, Norway, 5–8 August 2024; pp. 1–6.

83. Zhang, C.; Lu, J.; Zhao, Y. Generative pre-trained transformers (GPT)-based automated data mining for building energy management: Advantages, limitations and the future. *Energy Built Environ.* **2024**, *5*, 143–169. [CrossRef]
84. Emami, P.; Sahu, A.; Graf, P. Buildingsbench: A large-scale dataset of 900k buildings and benchmark for short-term load forecasting. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 19823–19857.
85. Choi, S.L.; Jain, R.; Emami, P.; Wadsack, K.; Ding, F.; Sun, H.; Gruchalla, K.; Hong, J.; Zhang, H.; Zhu, X.; et al. *eGridGPT: Trustworthy AI in the Control Room*; Technical report; National Renewable Energy Laboratory (NREL): Golden, CO, USA, 2024.
86. Zhang, C.; Zhang, J.; Zhao, Y.; Lu, J. Automated data mining framework for building energy conservation aided by generative pre-trained transformers (GPT). *Energy Build.* **2024**, *305*, 113877. [CrossRef]
87. Zhang, J.; Zhang, C.; Lu, J.; Zhao, Y. Domain-specific large language models for fault diagnosis of heating, ventilation, and air conditioning systems by labeled-data-supervised fine-tuning. *Appl. Energy* **2025**, *377*, 124378. [CrossRef]
88. Accurso, J.; Mendy, R.; Torres, A.; Tang, Y. A ChatGPT-like Solution for Power Transformer Condition Monitoring. In Proceedings of the 2023 International Conference on Machine Learning and Applications (ICMLA), Jacksonville, FL, USA, 15–17 December 2023; pp. 1716–1722.
89. Zaboli, A.; Choi, S.L.; Song, T.J.; Hong, J. Chatgpt and other large language models for cybersecurity of smart grid applications. In Proceedings of the 2024 IEEE Power & Energy Society General Meeting (PESGM), Seattle, WA, USA, 21–25 July 2024; pp. 1–5.
90. Jia, M.; Cui, Z.; Hug, G. Enabling large language models to perform power system simulations with previously unseen tools: A case of daline. *arXiv* **2024**, arXiv:2406.17215. [CrossRef]
91. Emami, P.; Li, Z.; Sinha, S.; Nguyen, T. SysCaps: Language interfaces for simulation surrogates of complex systems. *arXiv* **2024**, arXiv:2405.19653. [CrossRef]
92. Bonadia, R.S.; Trindade, F.C.; Freitas, W.; Venkatesh, B. On the potential of chatgpt to generate distribution systems for load flow studies using opendss. *IEEE Trans. Power Syst.* **2023**, *38*, 5965–5968. [CrossRef]
93. Huang, C.; Li, S.; Liu, R.; Wang, H.; Chen, Y. Large foundation models for power systems. In Proceedings of the 2024 IEEE Power & Energy Society General Meeting (PESGM), Seattle, WA, USA, 21–25 July 2024; pp. 1–5.
94. Yang, C.; Wang, X.; Lu, Y.; Liu, H.; Le, Q.V.; Zhou, D.; Chen, X. Large language models as optimizers. In Proceedings of the Twelfth International Conference on Learning Representations, Pensacola, FL, USA, 5–7 December 2023.
95. Wang, D.; Yin, C.; Liu, L.; Zhang, Z.; She, Y.; Yu, Y. Research on Intelligent Matching Technology of Power Grid Dispatching Automation Emergency Plan Based on Large Model and Small Sample LoRA Fine-Tuning Algorithm. In Proceedings of the 2024 9th Asia Conference on Power and Electrical Engineering (ACPEE), Shanghai, China, 11–13 April 2024; pp. 1390–1396.
96. Yan, Z.; Xu, Y. Real-time optimal power flow with linguistic stipulations: Integrating GPT-agent and deep reinforcement learning. *IEEE Trans. Power Syst.* **2023**, *39*, 4747–4750. [CrossRef]
97. Siino, M.; Tinnirello, I. Gpt prompt engineering for scheduling appliances usage for energy cost optimization. In Proceedings of the 2024 IEEE International Symposium on Measurements & Networking (M&N), Rome, Italy, 2–5 July 2024; pp. 1–6.
98. Sun, G.; Shen, R.; Jin, L.; Wang, Y.; Xu, S.; Chen, J.; Jiang, W. Instruction tuning text-to-sql with large language models in the power grid domain. In *Proceedings of the 2023 4th International Conference on Control, Robotics and Intelligent System*; Association for Computing Machinery: New York, NY, USA, 2023; pp. 59–63.
99. Wang, Q.; Zhou, Z. CPGA-BOT: A Customized Power Grid Assistant chatBOT Fine-Tuning in Large Language Model. In Proceedings of the Proceedings of the 2024 5th International Conference on Information Science, Parallel and Distributed Systems (ISPDS), Guangzhou, China, 31 May–2 June 2024; pp. 597–601.
100. Yin, C.; Du, K.; Nong, Q.; Zhang, H.; Yang, L.; Yan, B.; Zhang, X. PowerPulse: Power energy chat model with LLaMA model fine-tuned on Chinese and power sector domain knowledge. *Expert Syst.* **2024**, *41*, e13513. [CrossRef]
101. Cheng, Y.; Zhao, H.; Zhou, X.; Zhao, J.; Cao, Y.; Yang, C.; Cai, X. A large language model for advanced power dispatch. *Sci. Rep.* **2025**, *15*, 8925. [CrossRef]
102. Wang, C.; Liu, C.; Chen, J.; Zhang, G. Cooperative planning of renewable energy generation and multi-timescale flexible resources in active distribution networks. *Appl. Energy* **2024**, *356*, 122429. [CrossRef]
103. Xu, X.; Wang, M.; Xu, Z.; He, Y. Generative adversarial network assisted stochastic photovoltaic system planning considering coordinated multi-timescale volt-var optimization in distribution grids. *Int. J. Electr. Power Energy Syst.* **2023**, *153*, 109307. [CrossRef]

https://doi.org/10.3390/en19010116

energies

Review

Smart Grid Fault Mitigation and Cybersecurity with Wide-Area Measurement Systems: A Review

Chisom E. Ogbogu [1], Jesse Thornburg [1,2,*] and Samuel O. Okozi [3]

[1] College of Engineering, Carnegie Mellon University Africa, Kigali BP 6150, Rwanda; ceogbogu@andrew.cmu.edu
[2] Grid Fruit, LLC., Austin, TX 78758, USA
[3] Department of Electrical Engineering, School of Electrical Systems and Engineering Technology, Federal University of Technology Owerri, Owerri PMB 1526, Nigeria; samuel.okozi@futo.edu.ng
* Correspondence: jthornbu@andrew.cmu.edu

Abstract: Smart grid reliability and efficiency are critical for uninterrupted service, especially amidst growing demand and network complexity. Wide-Area Measurement Systems (WAMS) are valuable tools for mitigating faults and reducing fault-clearing time while simultaneously prioritizing cybersecurity. This review looks at smart grid WAMS implementation and its potential for cyber-physical power system (CPPS) development and compares it to traditional Supervisory Control and Data Acquisition (SCADA) infrastructure. While traditionally used in smart grids, SCADA has become insufficient in handling modern grid dynamics. WAMS differ through utilizing phasor measurement units (PMUs) to provide real-time monitoring and enhance situational awareness. This review explores PMU deployment models and their integration into existing grid infrastructure for CPPS and smart grid development. The review discusses PMU configurations that enable precise measurements across the grid for quicker, more accurate decisions. This study highlights models of PMU and WAMS deployment for conventional grids to convert them into smart grids in terms of the Smart Grid Architecture Model (SGAM). Examples from developing nations illustrate cybersecurity benefits in cyber-physical frameworks and improvements in grid stability and efficiency. Further incorporating machine learning, multi-level optimization, and predictive analytics can enhance WAMS capabilities by enabling advanced fault prediction, automated response, and multilayer cybersecurity.

Keywords: smart grid; power system monitoring; wide-area measurement systems; area measurement; cybersecurity; power system security; wide-area monitoring systems; cyber-physical systems; wide-area management systems; phasor measurement units

1. Introduction

Developing a resilient, reliable, and secure power system is central to modern economic development. Increasingly, such a grid also serves sociocultural and security goals. Yet, maintaining such a power system has become more challenging in recent years due to rising energy demand, the densification of population centers, climate change effects, and aging power infrastructure [1]. Grid stability and reliability are also threatened by the introduction of renewable power generation, which is often intermittent, nonsynchronous, or volatile [2]. While costly maintenance and upgrades are common responses to these challenges, more affordable improvement measures are often out of reach, especially for developing countries [3]. Over the years, power system restructuring has been the primary way of expanding power system operations. Complex measurement systems are essential for controlling the network in day-to-day generation, transmission, and distribution of energy [3,4].

Consequently, the modern-day focus of power system development is optimizing the present architecture responsible for accurate and adequate energy delivery [4,5]. On the software side, demand-side management and load shifting are being employed to these ends [6,7]. On the hardware side are many approaches involving relays, which have many shortcomings, including an inability to simultaneously report multiple parameters in real time. These shortcomings, in turn, were initially addressed using a supervisory control and data acquisition (SCADA) platform. SCADA could report multiple energy parameters to control towers across the network in real time [5]. SCADA measurements are typically updated in a ten-second window for high-level operations, which is unfortunately not granular enough to prevent certain faults [8,9].

SCADA's limitations have made us unable to reduce fault occurrence on power networks, and in most cases, SCADA increases the fault-clearing time [4]. Since neither relays nor generic SCADA address the rising issue of faults plaguing the power network, grid managers increasingly adopt synchrophasors. These increase the number of parameters like voltage, current, resistance, and impendence measured per unit time on the power system network [10]. The general introduction of synchrophasors, phasor measurement units (PMUs), and critical communication systems into the power system infrastructure results in it being termed a Wide-Area Measurement System (WAMS) [11,12].

WAMS infrastructure connected to the power system network relies on communication utilities to detect faults on the transmission lines using synchrophasors. WAMS then reports these data to the control centers for proper decisions to predict faults, correct faults, and improve consumer energy supply (as shown in Figure 1) [8,13,14]. This study reviews research on WAMS to apply them to developing countries' power systems, specifically to reduce fault-clearing time, improve cybersecurity, and make conventional power systems smart in these regards. To this end, this review analyzes the individual components of WAMS communication architecture. It details their potential for reducing fault-clearing time to increase the efficiency of energy delivery, all while maintaining cybersecurity.

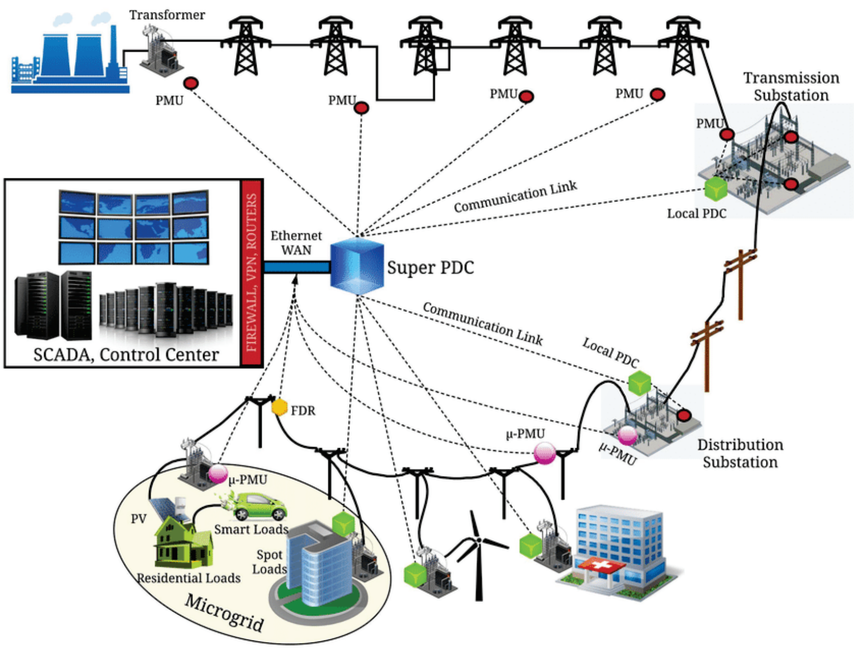


Figure 1. Schematic of a typical WAMS implementation [15].

Modern power systems are characterized by their intricate architectures, comprising numerous components and complex interconnections to ensure stable and reliable operation. However, the occurrence of faults—whether due to equipment failure, environmental factors, or other disruptions—can severely compromise energy flow and supply across the network. Even microgrids, which represent a transformative approach to power system design through the integration of Distributed Energy Resources (DERs), are not exempt from these vulnerabilities [16]. As depicted in Figure 1, microgrids are increasingly susceptible to faults propagated by conventional power systems, underscoring a critical challenge in achieving seamless interoperability between these two frameworks [16,17]. This issue has impeded the advancement of more resilient power system architectures and off-grid solutions, as the integration of microgrids with traditional grid infrastructure continues to present significant technical and operational hurdles. Faults are at some level inevitable, since key components are exposed to environmental and manmade stresses that affect the daily operation of a power network. WAMS are systems that provide grid managers with operational awareness while conducting key power system functions [10,14]. WAMS are, therefore, becoming well known as tools for improving fault clearing, thus enhancing reliability in power system architecture. The reliability setup described by Aminifar et al. consists of monitoring and control components, and their study measured its efficiency to accomplish both responsibilities [11]. In a normal power protection system, units are designed to accomplish these tasks whether through SCADA or conventional relay units [18,19]. However, these approaches are prone to malfunctions when overwhelmed by the faults they are meant to detect. Thus, in cases of high complexity, they fail to provide an accurate real-time assessment of the power system network situation [3,20].

1.1. Search Strategies and Selection

In this review, papers were selected that focus on WAMS development and its potential application to power system networks. Keywords like "WAMS", "Wide-Area Protection Systems", "PMU deployment", and "Systems for control of Wide-Area Power Systems" were used to search databases like ScienceDirect and Web of Science to select recent, well-cited research studies that cover the major variants of WAMS technologies and innovations. Papers with more focus on cybersecurity and power system optimization were prioritized.

1.2. Definitions of WAMs

Multiple definitions of WAMS have been published, and many studies have been conducted on their vast potential and practical results. In a SCADA system, monitoring and control responsibilities are not provided in a probabilistic power reliability assessment. In WAMS, by contrast, reliability assessments have shown that successful operation was more certain when WAMS elements were deployed and utilized optimally in a power system network [12]. Al-Odienat et al. deployed PMUs in the Jordanian National Grid and developed a WAMS architecture. They defined a WAMS as a network of PMUs and components that improves the situational awareness of a power system network for real-time monitoring and control [10]. PMUs, which make up the bulk of a WAMS, serve as connecting links when synchronized with a satellite-based Global Positioning System (GPS) [5]. Traditionally, WAMS that differ in certain components are deployed to increase the operational awareness of the system [21]. The integrity of the energy system is prioritized in these designs, and some research has even extended the WAMS infrastructure to under-frequency load-shedding schemes [22].

Researchers widely agree that any power system network, including those with WAMS technology, can be considered a smart grid even if they retain certain old, outdated components that are not hallmarks of modern smart grid networks [4,5]. WAMS was defined by

Premkumar et al. as an enabling technology that includes PMUs as measuring devices coupled with supporting communication infrastructure like phasor data controllers (PDCs) [1]. However, their study posited this definition before going on to suggest that a combination of a PMU and enabling information infrastructure is termed Synchrophasor Measurement Technology (SMT) [1]. In contrast, some studies have found SMT a component of WAMS infrastructure instead. Numerous research studies on European energy systems have highlighted the integration of smart systems into aging infrastructure, including SCADA and energy management systems (EMS) alongside PMUs to form WAMS [3–6,8,14,21,23–26]. In their study, Alhelou et al. included PDCs, which are designed to transmit data collected from the PMUs on the lines and the different interconnected units in the energy infrastructure [22]. All of these become imperative to articulate components that properly define a WAMS infrastructure and assess the position of numerous researchers on the advantages of these technologies in the WAMS infrastructure [12,18–20,27].

SMT consists of several arrays of PMUs on the grid and the PDCs whose role is to aggregate the data collected by the PMUs interconnected across the energy infrastructure [3,10]. The similarity in the components led Aminifar et al. to conclude that SMT is synonymous with WAMS [11]. Others disagree, as SMT includes data collection PMUs like WAMS, but an absence of communication infrastructure makes SMT fundamentally different [18]. Therefore, the integration of SMTs, including PMUs and PDCs, and a proper communication infrastructure is regarded as a WAMS. The definition of WAMS becomes imperative for a proper analysis of the development of technology through the integration of multiple advances in the modern world of technology attempting to increase the volume of data collected, synchronized, and communicated to Energy Control Centers (ECCs) (as shown in Table 1) [3–6,14,21,24,25,28–30].

Table 1. Table showing some areas of WAMS currently explored and the extensive academic interest in WAMS.

S/N	WAMS Technology Explored	Areas of Power System	Technology Considered	Reference
1.	**SMT**	Monitoring	PMUs+SCADA	[1,3,6]
2.	**Wireless Communication Technologies**	Monitoring, smart grid	PMUs+SCADA	[10]
3.	**Wide-Area Power Control and Monitoring (WAPCAM)**	Monitoring, fault detection, power quality, Smart Grid Architecture Model, and wide-area power system observability	SMTs+SCADA+RTUs+ PDCs+GPS	[31]
4.	**Wide-Area Measurement and Control Systems (WAMCS)**	Data quality, fault location, and multi-energy complementary systems	PMUs+SCADA+PDCs+GPS	[32,33]
5.	**Wide-Area Monitoring, Protection, and Control (WAMPAC)**	Monitoring, fault detection, and control	EMS+PMUs+SMTs	[3,29]

The main responsibility of WAMS infrastructure is to boost the real-time awareness of the power system network [18,24,25]. The observability and monitoring potential of WAMS cannot properly achieve the system's overall goal when proper communication infrastructure is absent [25]. This review shows the progression from SCADA infrastructure to modern applications of WAMS in increasing the operability of the energy management system following the recent deployment of artificial intelligence (AI) in real-time awareness, control, and efficiency enhancement [34]. Most of the literature studies the role of PMUs and

WAMS infrastructure as isolated technologies with limited objective goals, which are assessed to test the ability of the infrastructure to optimize the power system network [21,24,25,28].

While significant research has been conducted on WAMS technologies and smart grid innovations, limited studies have focused on the deployment of WAMS technologies in traditional power system networks, which remain predominant in developing economies and particularly in sub-Saharan Africa. This gap has resulted in a lack of comprehensive data necessary for conducting sensitivity analyses of wide-area power systems (WAPS). Such analyses are crucial for enabling sub-Saharan African countries to develop robust fault detection schemes and establish pathways for integrating renewable energy into conventional grids.

This review aims to highlight the progression of research in WAMS and identify key areas where these technologies could be effectively applied in traditional and aging power system networks within low- and middle-income economies. Additionally, it explores the potential role of AI in optimizing WAMS operations and examines the development of cyber-physical power system (CPPS) networks. According to the Smart Grid Architecture Model (SGAM), the CPPS represents a foundational framework for advancing smart grid development within conventional power grids, offering a structured approach to enhancing grid resilience, efficiency, and adaptability in emerging economies.

This review discusses WAMS operation and dynamic fault detection, benefits that AI can provide with WAMS, cyber-physical power systems (CPPS), and smart grid development, and concludes with recommendations for future research.

2. Background

The modern power network consists of various complex machines designed to play significant roles in generating, transmitting, and distributing electricity to numerous households connected to the grid. Dynamic state estimation (DSE) assessments show the complex systems that need to be operational to guarantee the optimal conversion, transformation, and delivery of electrical energy [18,35,36]. Figure 2 presents an infinite-bus power grid model consisting of a synchronous generator powered by a mechanical source in a three-phase circuit arrangement. The transformers and power measurement units serve to deliver optimal power to the asynchronous machine and other loads. During the operation of the power system network, the static state estimation cannot appropriately capture the dynamics of the operational environment [35,36]. This presents a huge problem because the power system circuitry consistently fluctuates to adjust to the evolving realities of the energy systems. Despite this, SSE methods are still used to design an EMS, but the approach proves insufficient for solving evolving problems and providing sufficient information needed to monitor power system architecture [36,37]. In asynchronous generators and machines, for example, EMS systems with the SSE configuration have been difficult to monitor and control since observability is often lagging and unresponsive in situations of optimal operation [30,37,38]. Dynamic state estimation (DSE) is capable of accurately capturing and assessing the dynamic system states and plays a crucial role in further monitoring and controlling by providing the maximum observability of the power system network [8,26,37,39]. For protection to be appropriately achieved, the dynamic states must be accurately captured to provide a unique state of awareness of the power system network [39]. The SCADA system was adopted to fulfill this responsibility in the differing state of the power system network. SCADA systems can assess two independent data points from the power system network and report the data to the control infrastructure [8,40]. One problem with SCADA systems, as stated earlier, is their limited number of parameters that can be assessed and observed at a time together within a period of 8–10 s before another cycle of data can be collected [8,37,41]. In a complex modern power system network, these limitations could cause significant damage as individual components are consistently exposed to fault occurrences on the line and other complex energy components whose damage could result in a partial or total system collapse of the grid. Additionally, present research into the role of varying DSE parameters and how they could be implemented in an EMS has increased the demand for a more robust communication and monitoring infrastructure to be developed [38].

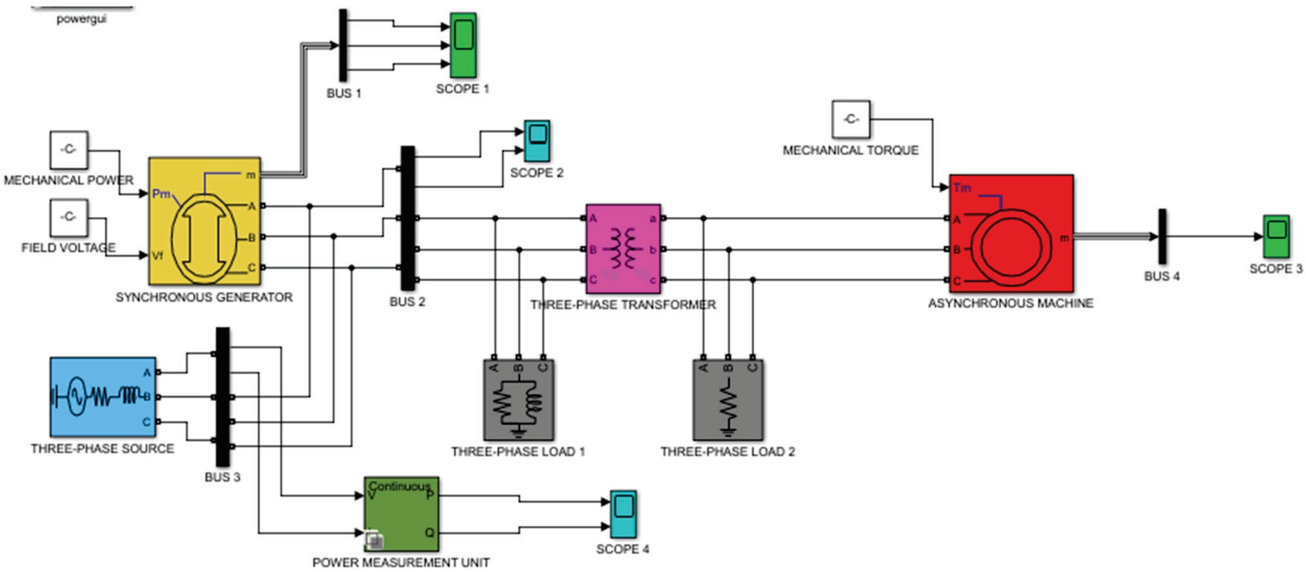


Figure 2. A typical infinite-bus power grid model.

2.1. WAMS Components

For full observability and monitoring capabilities, the WAMS architecture typically consists of the following components [3,42]:

1. PMUs;
2. Phasor data collectors;
3. Synchronizers like remote terminal units (RTUs) for PMU-based applications;
4. Communication infrastructure.

This WAMS infrastructure was developed as an alternative to SCADA systems. In addition to the basic components of the WAMS infrastructure, more modern model variants like Wide-Area Monitoring and Control (WAMC) or Wide-Area Protection, Control, and Monitoring (WAPCAM) systems have incorporated several components that help enhance some of its operations [4]. The observability operation of the WAMS infrastructure provides sufficient data to control the EMS properly and, in some cases, conduct an under-frequency load shedding (UFLS) of the entire power network in addition to fault detection [22,39,43]. The PMUs collate the data from power system components, including being integrated with conventional energy infrastructures like current transformers and voltage transformers, to develop a more robust power system architecture [10,42,43]. This responsibility of the PMUs becomes necessary because most countries already have functional power systems in need of system maintenance and upgrade without a complete overhaul of the system for mostly national security reasons [10,26,44]. The US, where the technology was innovated and developed, has not fully replaced its conventional protection arsenals for WAMS due to cybersecurity concerns though modern cyber-physical systems are being integrated at a rapid pace [14]. In practice, however, the WAMS architecture is designed to integrate with existing infrastructures and make them as smart as possible in an age of increasing rhetoric of smart technology and green energy innovation [10,22,44]. If any of these modern innovation concepts are feasible, energy losses must be minimal for energy conservation to be properly executed efficiently. PMUs and PDCs work in synergy with nearby components to achieve this. In our opinion, the deployment of WAMS is entirely dependent on the quality of the delivery system as well as the communication infrastructure. Some researchers have synonymized PMUs for WAMS. Many researchers disagree with this assessment, but the distinction stems from the comparison of WAMS with conventional SCADA infrastructure [10,22,27,44]. However, PMUs alone are also unable to properly analyze these data and relay them to the ECC for efficient actions and decisions to be taken [21,28,43]. WAMS, on the other hand, could consist of other energy infrastructures beyond PMUs, PDCs, and RTUs to fit and meet specific needs in a power system network [37,43]. Going further, the main potential of WAMS is their application. A model cannot be accepted if its architecture is not fully analyzed. The delicate nature of WAMS means that there will be significant training on their handling and on the deployment of the technology [10].

Sub-Saharan African countries, where off-grid systems coupled with renewable energy sources are increasingly developed, require the synchronization of devices to ensure the optimized operation of the grid. The synchronization of the PMUs and other protection and control devices is achieved through the same-time sampling of voltage and current waveforms using timing signals from a GPS (as shown in Figure 3) [1,2]. Synchronized phasor measurement elevates the standards of power system monitoring, control, and protection to a new level [21,24,45]. PMUs are the backbone of WAMS-based systems and have been proven instrumental in developing-world microgrid systems, especially with the integration of renewable energy sources [5,13,21]. This means that real-time system monitoring may be required to help reduce the impact of intermittency from renewable generation. Several protocols that consider the necessity of real-time monitoring in smart grid development have been codified in the SGAM, which has served as a framework for assessing smart grid technology development across five interoperability layers, domains, and zones, as shown in Figure 4 (more details in Section 2.2).

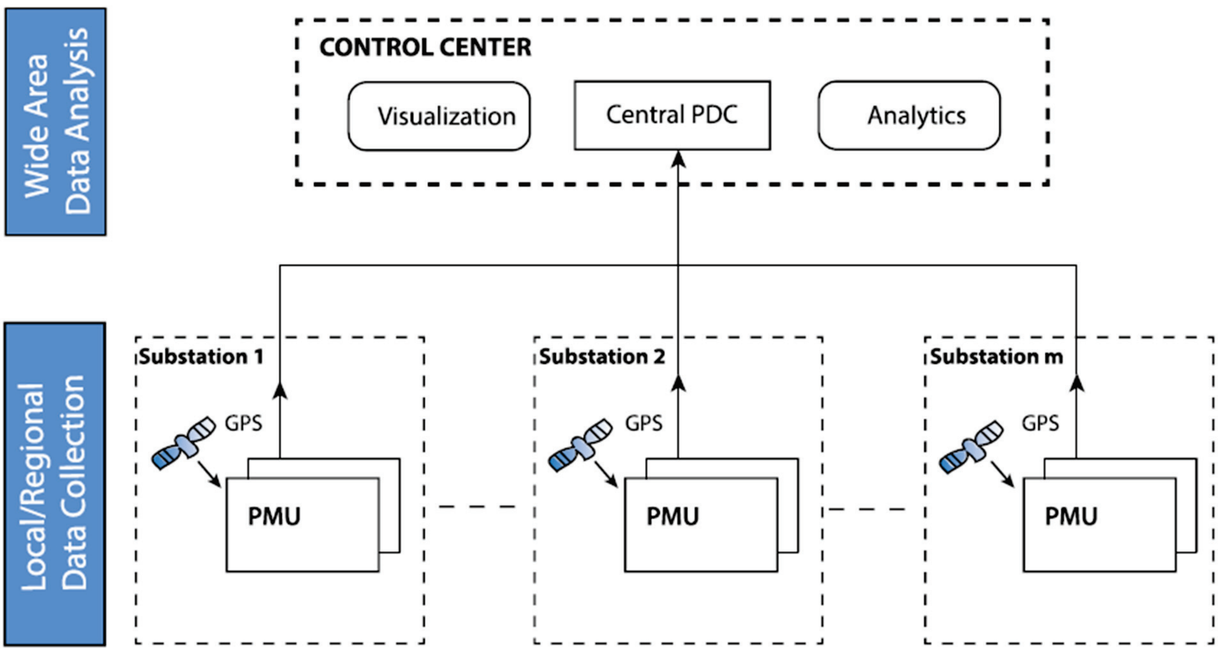


Figure 3. Centralized wide-area measurement system (WAMS) architecture [46].

In addition, another application is real-time system monitoring (RTSM) for stability assessment and state measurement. Phasor measurements at nodes help the system operators gain a dynamic view of the power system and initiate the necessary measures at the proper time [21]. This is performed following the latest IEEE standard (C37.118-2005), developed to standardize the data transmission format and sampling rates of PMUs [21,47,48]. The stability assessment algorithms, designed to take advantage of the phasor measurement information, can significantly support this [47]. In the past, post-event analysis was an application of synchrophasors (PMUs) without wide-area communication, where data were archived locally [21,24]. However, it was not a useful online (dynamic) control tool. Recently, the real-time control of WAMS has become a robust control and analysis tool that provides a new view of power systems [47,49]. This is achieved by improving the communication network capabilities while maintaining PMUs as central components [13,21,24,25,37,47]. Using PMUs for real-time control increases accuracy since the data are measured online [21,42]. Also, it will enhance the power system stability and delivery automation capabilities after the challenges of new data communication requirements across the system are resolved. The depth of observability is another advantage for PMUs. It means measuring the bus voltage phasor directly or calculating it using the nearest connected bus's PMU voltage and line current [47,48]. This is the cost-effective path since it reduces the number of data acquisition instruments and tools needed across the network, as the measuring line currents can extend the voltage measurements to buses with no PMUs [18,37,38,41,43].

2.2. WAMS and Smart Grid Architecture Model

The SGAM provides a comprehensive framework for analyzing and designing smart grid systems by addressing technical, informational, and operational challenges across five interoperability layers: component, communication, information, function, and business [17]. WAMS offers a transformative capability for modern power systems, playing a critical role in enhancing fault detection, reducing fault-clearing times, and facilitating the integration of renewable energy and DERs [50,51]. By integrating a WAMS into the SGAM framework, grid operators can achieve improved system observability, faster response times, and greater adaptability to the challenges posed by renewable energy intermittency. Panda and Das (2021) [17] assessed the SGAM framework across different areas, including peer-to-peer trading, distributed control, laboratory infrastructure, and power quality. They identified these areas as fundamentals of futuristic energy technologies. Other studies, including those by Aminifar et al., Ashok et al., and Calderon et al., showed that while innovations are being carried out, like in the SGAM, smart grid systems are consistently being designed to expand the frontiers of current power systems and connect underserved communities in remote areas, especially in developing countries [11,31]. As result of this, according to Ashok et al. and Jamil et al., current problems could still handicap these new innovations and compromise national security and may even result in blackouts across high-voltage lines [16,31]. This concern stems from the increasing cybersecurity threats posed because of AI development and increasing geopolitical tensions. Ashok et al. and Panda and Das, while commending the progress of the SGAM, also suggested that WAMS or WAPCAM technologies will be required to combat the growing threat of interoperability that may be faced with the ever-increasing wide-area power systems (WAPS) [17,31]. This review assesses key areas where WAMS technology may contribute to the SGAM across its five domains and zones:

1. **Component Layer: Enhanced Observability Through PMUs**
   PMUs provide synchronized, high-resolution measurements of key grid parameters such as voltage, current, and frequency. These devices enable the real-time monitoring of grid conditions, including renewable energy sources like wind turbines and solar panels. With the deployment of PMUs at critical locations, such as substations, transmission lines, and renewable generation sites, WAMS ensures comprehensive observability [51]. This capability is essential for monitoring the dynamic behavior of DERs and identifying anomalies that could indicate potential faults.

2. **Communication Layer: Enabling Real-Time Data Exchange**
   The communication layer of the SGAM supports the transmission of the vast amounts of data generated by WAMS. To achieve low latency and high reliability, WAMS leverages advanced communication technologies such as optical fiber networks, 5G, or power line communication (PLC) [50]. These technologies enable real-time data exchange between PMUs and centralized or distributed control centers, ensuring the timely identification of faults and rapid coordination of corrective measures [17]. This layer also facilitates the communication of control signals to distributed resources, ensuring that renewable energy sources can dynamically adjust to changing grid conditions.
3. **Information Layer: Transforming Data into Actionable Intelligence**
   The integration of a WAMS into the SGAM information layer highlights its role in data processing, storage, and analysis. WAMS data are analyzed using advanced algorithms, including AI and machine learning, to detect early warning signs of system instability, optimize the performance of DERs, and predict faults before they escalate. This layer ensures that actionable intelligence is generated to enhance grid reliability, while also incorporating data encryption and secure storage protocols to address cybersecurity challenges. By transforming raw data into meaningful insights, WAMS empowers grid operators to make informed decisions in real time.
4. **Functional Layer: Strengthening Grid Resilience**
   WAMS aligns closely with the SGAM functional layer by supporting advanced grid control and operational strategies. Its functionalities include voltage regulation, frequency control, and islanding detection, all of which are critical for the integration of renewable energy. WAMS enables dynamic load shedding and fast fault-clearing mechanisms, minimizing system disruptions and reducing fault-clearing times. Furthermore, its real-time fault detection capabilities enhance the grid's self-healing properties, ensuring that power systems remain operational even under adverse conditions. By supporting these functions, WAMS contributes to the development of a more resilient and adaptive grid.
5. **Business Layer: Enabling Cost-Efficiency and Compliance**
   At the business layer, WAMS provides significant economic and regulatory benefits. By enabling accurate and reliable grid operations, WAMS supports demand response strategies and enhances energy trading in markets that include renewable energy sources. It also ensures compliance with regulatory standards for grid reliability and renewable energy integration. These capabilities enable utilities to optimize their operations while achieving long-term sustainability goals. Moreover, WAMS facilitates the planning and operation of power systems in ways that align with market dynamics and future energy needs.

The integration of a WAMS into the SGAM framework represents a new frontier in modern power system operations. With its real-time monitoring and fault detection capabilities, WAMS addresses critical challenges in fault management, grid stability, and the integration of renewable energy sources. Through its alignment with the five SGAM layers, WAMS not only enhances grid reliability and resilience but also positions power systems for a sustainable and economically viable future. As the energy landscape continues to evolve, WAMS will remain a cornerstone technology in enabling the transition to smarter, more adaptive grids.

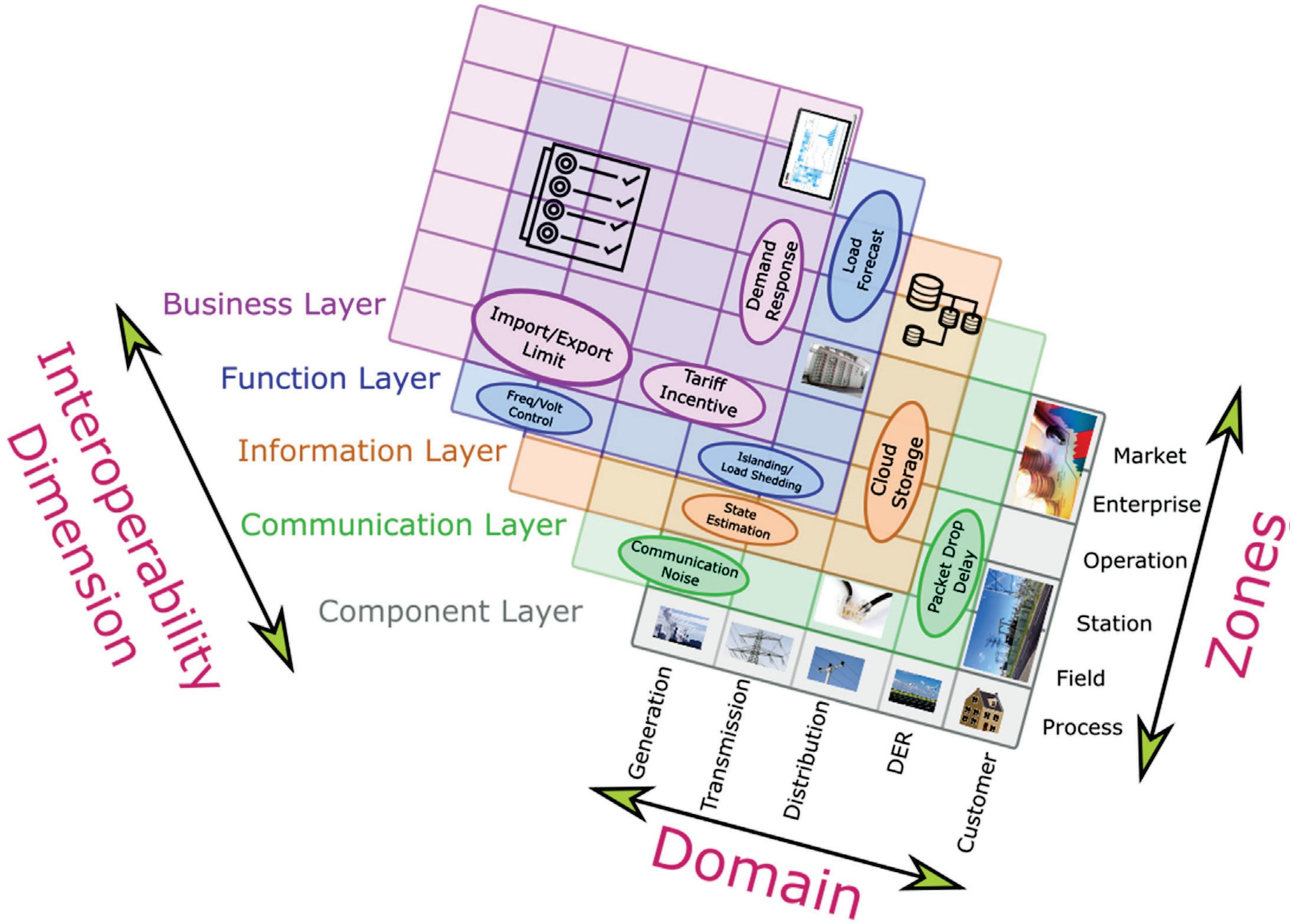


Figure 4. SGAM layers, zones, and domains [17].

3. WAMS Operation and Dynamic Fault Detection

WAMS are integral to modern power system management, particularly in enhancing grid observability, supporting network expansion, and aligning with the SGAM [31]. The ability of WAMS to perform real-time dynamic fault detection sets it apart from conventional protection technologies, including SCADA-based state estimation. While traditional systems rely on periodic measurements and interpolation techniques to estimate system states, WAMS leverages phasor measurement units (PMUs) to provide high-resolution, time-synchronized data, enabling more precise and timely fault identification (Figure 5). This capability significantly enhances grid resilience by facilitating faster fault detection, reduced fault-clearing times, and improved system stability [1,3,7].

Extensive research has explored the integration of WAMS into advanced smart grid architectures, with a particular emphasis on enhancing the incorporation of renewable energy sources [52–54]. WAMS enhances the adaptability of power networks to accommodate DERs such as wind and solar power by providing real-time insights into voltage stability, frequency variations, and power flow dynamics [38,43]. This integration is particularly crucial for the development of renewable, smart microgrids, which utilize WAMS for intelligent energy management, adaptive load balancing, and improved operational resilience [54].

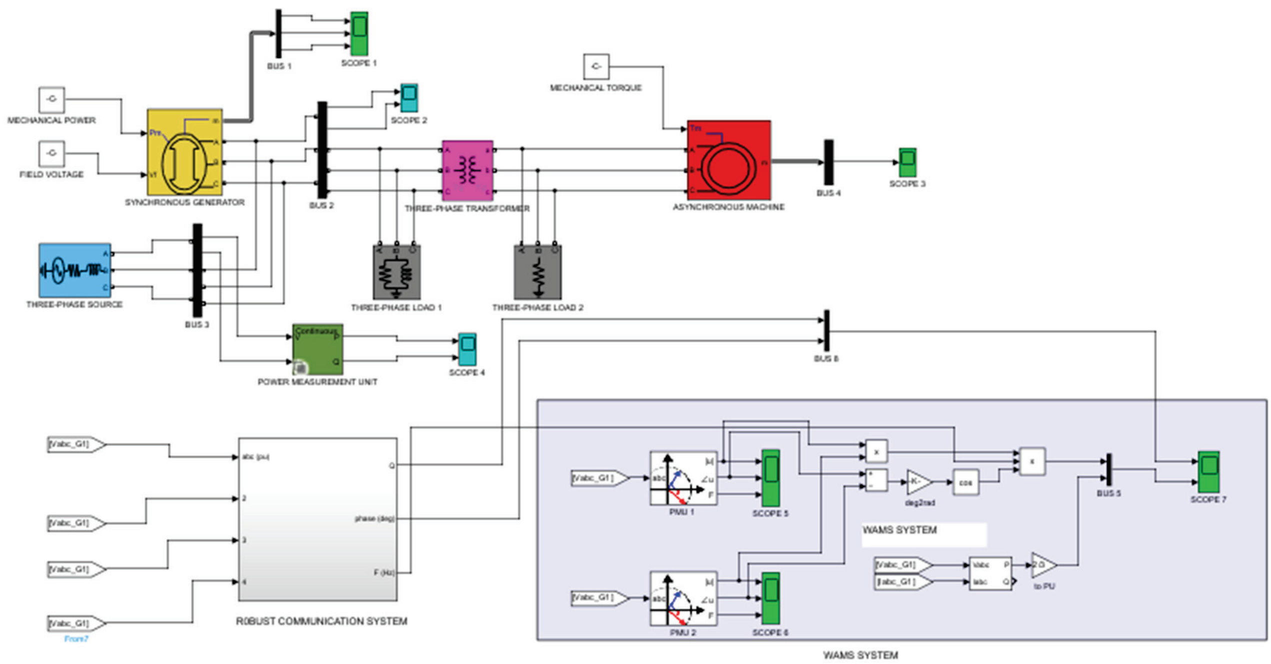


Figure 5. Operationalization of PMU-based WAMS in infinite-bus grid model with communication system.

Despite significant advancements in smart grid research, limited attention has been paid to the deployment of WAMS in aging power system networks, which are prevalent in developing economies. The lack of comprehensive research on the type of data that traditional power grids can provide for informed expansion remains a critical gap in the literature, potentially hindering the progress of renewable energy development. One key challenge in these regions is the frequent collapse of grid infrastructure. This issue, particularly prevalent in Nigeria and South Africa, has led to recurrent power outages, negatively impacting economic growth and deterring private investment in both the energy and industrial sectors.

Ashok et al. conducted a comprehensive analysis of blackouts in developed economies, outlining the operational states of a wide-area power system network—Normal, Alert, Emergency, In-Extremis, and Restorative—to examine the mechanisms through which blackouts occur and propagate within WAMS-integrated grids. Their study identified the preconditions, initiating events, cascading failures, and restoration processes that characterize blackout events [31]. While these insights have primarily been applied to modern smart grids, they are equally relevant to traditional power networks. Understanding how WAMS can be leveraged to mitigate systemic vulnerabilities in aging power infrastructures presents a crucial research avenue that warrants further investigation.

The effects of weather conditions, along with the varying oscillations within power systems, can collectively be classified as dynamic faults that impact grid stability. Through WAMS, these faults can be accurately characterized, predicted, and mitigated, enabling more efficient fault detection and facilitating the restoration of wide-area power systems (WAPS) in the event of disruptions. WAMS can enhance grid observability and enable proactive fault management by leveraging high-resolution, time-synchronized data from the power system network.

Additionally, WAMS can collect and analyze critical data on the impact of lightning strikes, governor responses, frequency instability (e.g., rate of change in frequency), oscillation issues, and interoperability challenges. These insights are particularly valuable for increasing the observability and operational efficiency of traditional power system networks in developing countries, where aging infrastructure and inconsistent grid performance remain persistent challenges. Through the integration of WAMS into these networks, utilities can develop data-driven strategies to enhance system resilience, optimize grid operations, and facilitate a smoother transition toward modernized, renewable energy-integrated power grids.

As illustrated in earlier sections, the issue of communication is central to the operational effectiveness of WAMS. The SCADA network, which is the most popular system, utilizes LAN systems to provide situational awareness [8]. On the other hand, WAMS aims to circumvent the insufficiencies of SCADA systems and provide more achievable situational awareness that relates real-time information from the power system network to the control room and from the control room relay instructions to automated infrastructure like isolators, circuit breakers, and CTs to observe and control the entire power network [14,21]. The PMUs scattered and deployed across different locations on the power system network must be interconnected to fulfill their responsibilities [14]. WAMS communication infrastructure, therefore, plays a crucial role in transforming a conventional grid into a smart grid network. Gore and Valsan (2018) [45] analyzed the challenges and efficiencies in deploying wireless communication infrastructure in WAMS. Premkumar et al. analyzed a WAMS architecture communication system and found that the precision of the communication system was determined by the distance of the PMUs from areas and, over time, the quality of the PMU placement on these energy infrastructures [1]. Generally, for the purpose of communication, the telecommunication infrastructure has been designed to be wireless to truly achieve wide-area operation. GPS transmission is the most significant data transmission infrastructure. The C37.118 standard format is divided into four packets, which can a transmit data frame, configuration frame, and header frame [1]. The data frame consists of the PMU-measured real-time phasor data in binary. The configuration frame is the subset of the available data in the data frame, which indicates the real-time data in the PMU [1,42]. These data can be transmitted only on a communication jacket in a way that synchronizes the PMU data. The GPS provides the communication infrastructure that synchronized the phasor measurements using PMUs placed in buses to complete the observability and controllability of the grid. Hassak et al. designed a communication infrastructure that deploys several wireless technologies to test their efficiency in monitoring and controlling WAMS operations. These infrastructures include wireless LAN, Wi-Fi, Wi-Max, GPRS, and GSM infrastructure [55]. Different communication infrastructures are selected to address the specific requirements of a given energy grid.

3.1. Optimal PMU Placement and Fault Detection

WAMS generally uses PMUs that are synchronized using a GPS clock to provide time-synchronized voltage and current phasors, including frequency measurements with sub-microsecond precision [31,41]. Alhelou et al. performed further analysis on WAMS communication infrastructure in different routing protocols when background traffic and link failure are present [22]. Aminifar's model design does not focus on the optimal communication infrastructure for PMU placement in a power grid. His team identified this limitation, where the optimal WAMS design was dependent on the optimal placement of the PMUs on the power grid [11]. The identified role played by the Wi-Fi and other wireless systems while guaranteeing the effective operation of the WAMS architecture was still necessitated to include the complexities of the IEEE 14 bus systems; however, the rapid transformation of conventional power system networks with such technologies raises concerns for the eventual cyber-physical power system (CPPS) to be developed. Ivanov and Dimova proposed an analysis of real-time data delivery in WAMS [53]. In their proposed model, the real-time potential of the PMUs in the WAMS architecture was analyzed, and they concluded that for the cost-effective deployment of PMUs to provide the maximum operation of the PMUs, the communication backbone network is required to determine the shortest path using a routing algorithm and channel information through an optical fiber in a ground wire without any delay analysis [53]. Premkumar et al. presented a model in which PMUs themselves may not effectively disperse the advantages of the WAMS architecture. In their argument, the placement of the PMUs and the communications infrastructure is required to ensure the optimal functioning of WAMS. This becomes necessary if the energy grid is designed to continually become flexible for integrating newer models of protection and communication infrastructure. The communication infrastructure in this context is synchronized with the placement of the PMUs for the optimal delivery of PMU data. These indications expose the study of the PMU placement problem [1]. The authors argued that PMUs are costly devices and, therefore, are to be deployed in a way that maximizes their benefit even with limited numbers. They proposed a deployment that would enable the further integration of technologies with no need for a radical alteration to the PMUs deployed or the energy grid itself. They relate the cost of PMU deployment to its effectiveness in the power grid with the following equation [1]:

$$min \sum_{i=1}^{n} c_i x_i \tag{1}$$

$$\text{s.t. } f(X) \geq b_i$$

where

x_i is a binary decision variable vector with entries defined as

$$x_i = \begin{cases} 1 \ if \text{ PMU is installed at bus } i \\ 0 \text{ otherwise} \end{cases}$$

c_i is the cost of the PMU installed at bus i;

b is the minimum observability matrix.

Fault detection is expressed as follows:

$$V_f = V_i - Z_{ij}^d I_i$$

where

$$V_f = Voltage\ at\ the\ fault\ point$$

$$V_i = Voltage\ at\ bus\ i,$$

$$Z_{ij}^d = Impedance\ of\ the\ line\ segment\ up\ to\ the\ fault\ point,$$

$$I_i = Current\ flowing\ from\ bus\ i.$$

When a fault occurs across the power system, distance d is nominally calculated using the relationship

$$d = \frac{|V_i - V_j|}{|I_i| \cdot Z_{ij}}$$

where

$$Z_{ij} = Total\ impedance\ of\ the\ line\ i - j,$$

$$V_j = Voltage\ at\ bus\ j$$

Note that d is mostly calculated using impedance measurements across the line. The values of V_j and V_i are obtained from PMU data, leveraging their ability to provide time-synchronized phasor measurements with high precision. These data are critical for identifying and localizing faults in real time. However, when faults are dynamic and evolve rapidly, such as in the case of transient faults or evolving system instabilities, the sudden changes in V_j and V_i may not be efficiently captured and processed, particularly if the sampling rate or data synchronization is inadequate. This has been a limitation in traditional SCADA systems, which operate at much lower sampling rates (typically seconds) and lack real-time synchronization.

Optimal PMU placement is therefore not just about ensuring observability but also about capturing the dynamic changes in V_j and V_i with high fidelity. This involves strategically locating PMUs at critical nodes where voltage deviations due to faults are most significant or where fault propagation is likely to impact grid stability. It is in this design that metaheuristic-based optimization algorithms are often explored to achieve overall grid stability. Furthermore, the placement must account for areas of the network prone to high fault incidence or regions where transient behavior can significantly affect operational decisions.

To enhance the effectiveness of PMU placement, it is essential to consider the dynamic characteristics of the system. While PMUs provide near-instantaneous measurements, they assume a relatively stable frequency or phase angle value across the grid. During faults, the grid frequency and phase angles may deviate, but these deviations are generally slower compared to the rapid voltage changes. Thus, the PMU placement strategy should aim to maximize the ability to monitor V_j and V_i dynamics under these conditions [56].

Additionally, advanced data processing techniques, such as adaptive filtering and real-time dynamic state estimation, can complement optimal PMU placement. These methods help refine the measurement of V_j and V_i during fast-changing fault scenarios, ensuring that the system remains accurately observable even under extreme conditions. Therefore, the design of optimal PMU placement should not only focus on achieving full-network observability but also on capturing and processing the dynamic changes in voltage magnitudes at critical points in the grid. Doing so ensures that the system can efficiently respond to and mitigate the impact of rapidly evolving faults, maintaining stability and operational integrity.

3.2. Limitation of PMU Placement in Developing Countries

In recent studies on WAMS, the PMU and PDC roles have been explored extensively vis à vis the cost-effective means of their deployment [5,13,14,21,25,48]. The main reason WAMS has not been deployed in energy infrastructures in less developed countries is the initial cost [14]. In these economies, power systems are significantly plagued with issues captured in wide-area monitoring systems that SCADA systems cannot detect [26]. Undoubtedly, authors [14,24,25] that know about the infrastructure mechanisms in WAMS are beginning to identify cost-effective means by which the communication infrastructure can accommodate the deployment of PMUs and PDCs on the grid to maximize the potential of WAMS. Premkumar et al. provided one example that illustrates the tendency for PMUs to be deployed and produce lower efficiency compared to others with the same communication jacket (CJ). The node–node interconnection achieved with PMUs is unveiled as the characteristic vehicle for transmitting the data collected on the grid to the control centers [1,21,25]. Though wireless, the potency remains subject to the sophistication and technicality of the PMU deployment [1].

Mohammadi and Zamani et al. supported this position by analyzing different models for deploying PMUs on the grid using the same communication infrastructure [14,42]. In practice, Mohammadi's team identified the N-1 contingency when a single-line or single-PMU outage occurs. They operated the WAMS architecture as one designed to maximize the communication infrastructure [42]. They showed that different algorithms are to be used to determine the efficiency of both PMU-PDC complexes in a WAMS [42]. During the WAMS application processes, Mohammadi's team suggested and analyzed several algorithms to determine the node–node iteration and provide a viable N-1 contingency for the effective operation of WAMS. Regarding the non-dominated sorting genetic algorithm, binary particle swarm optimization, and binary imperialistic competition algorithm, they found these inefficient in determining the safest and most efficient single path for PMU placement [42]. After analyzing different studies and their various means of placing PMUs, Mohammadi et al. identified the need to consider the communication infrastructure required by those designs and make the entire WAMS design and deployment both efficient and cost-effective [42]. In accomplishing these tasks, Premkumar et al. used the Dijkstra single-source shortest-path algorithm to determine the effective method of PMU-PDC deployment at a considerable communication infrastructure cost for WAMS to be advantageous in practice [1].

4. WAMS and Artificial Intelligence

As discussed above, WAMS addresses the fundamental issue of lacking situational awareness in a conventional SCADA infrastructure. Especially in a widely independent–interdependent energy infrastructure, the WAMS architecture can make the grid situationally aware and the conventional grid smarter without the need to overhaul a significant part of the design. In terms of complexity, the PMU, PDC, synchrophasor, and communications infrastructure technologies have seen different obstacles being presented that mitigate the deployment of the WAMS architecture on a grid [24]. As pointed out during the analysis of different communication infrastructures, the various algorithms needed for their deployment are as complex as WAMS [1,42]. In every engineering endeavor, the global optimum operation of every design is paramount, especially in terms of efficiency and cost. In the case of WAMS, Dijkstra's algorithm is used, and an N-1 contingency may arise from the binary imperialistic competition algorithm. This must be duly calculated and analyzed [1,42]. The research in [42] also showed that communication infrastructure design can conform to the efficient PMU design and thus become more cost-effective. In practice, these deployments have been known to change based on load differences and load potentials developing across the different load centers on the grid [8,24]. AI, therefore, becomes necessary to conduct these complex calculations and algorithms to ensure a proper and effective means of deploying WAMS along the grid with the least number of PMUs to achieve the maximum situational awareness of the grid. The Dijkstra algorithm could be modified at any time with an AI infrastructure by considering other factors like fault lines and vulnerable sections in optimizing PMU node placement [42]. With AI technology, the effective means of maximizing the communication infrastructure becomes feasible and cost-effective. Once this is considered, multiple studies have shown that WAMS adoption would increase drastically across economies that lack the financial incentives to deploy a smart grid [1,5,11,44]. In the race to deploy renewable energy systems in transitioning to a green economy, especially in largely sunny sub-Saharan Africa, the deployment of solar generation is paramount [57]. The intermittency of renewable energy sources remains a significant challenge that must be addressed to prevent major blackouts and ensure a reliable power supply. This can be achieved through strategies such as minimizing energy losses and improving consumption efficiency [52]. WAMS, when integrated with AI, offer a transformative solution by enabling the real-time determination of the safest and most efficient single-PMU paths for energy transmission across the grid [58,59]. This integration leverages AI's advanced computational capabilities to enhance the speed, accuracy, and reliability of grid monitoring and fault management.

AI significantly enhances the functionality of WAMS by refining metaheuristic optimization algorithms, including immune genetic algorithms, search trees, and simulated annealing, while simultaneously ensuring the robustness and integrity of the WAMS architecture [60,61]. These optimization techniques are critical in addressing fault detection, stability assessment, and system reliability, particularly in complex and evolving power grids.

Alhamrouni et al. identified several metaheuristic-based algorithms that form the backbone of AI deployment in power system networks. Their evaluation of four primary categories of optimization techniques—artificial neural networks (ANNs), fuzzy logic, evolutionary methods, and expert systems—highlighted the unique advantages of evolutionary-based algorithms, particularly those leveraging swarm intelligence, for power system optimization and fault detection [62,63]. Evolutionary algorithms offer self-adaptive mechanisms, global search capabilities, and robust convergence properties, making them well suited for dynamic power system environments where real-time decision making is essential. Among these, particle swarm optimization (PSO) has been widely recognized for its effectiveness in optimizing grid stability, designing protection schemes, and diagnosing faults [46,62,64]. PSO operates by mimicking the social behavior of particles (or agents) in a multidimensional space, allowing for the efficient exploration of search spaces and the identification of optimal solutions in power system applications. This technique has been particularly effective in enhancing fault detection precision, reducing response times, and improving grid resiliency. Polster et al. further emphasized the application of AI-driven optimization in WAMS-based systems by exploring hybrid optimization methods for fault detection [65]. Their research introduced a hybrid approach combining PSO techniques with Newton's method, comparing it against traditional linearization techniques that have historically dominated the field. By leveraging PSO's global optimization capabilities alongside Newton's rapid convergence properties, their approach significantly improved fault localization accuracy, reduced computational complexity, and enhanced the overall adaptability of WAMS in dynamic grid environments [65]. These findings illustrate the growing role of AI-driven optimization in modern power system monitoring, demonstrating AI's potential to revolutionize fault management, stability enhancement, and operational efficiency in smart grid architectures.

Despite their consistent use, the PSO algorithm still suffers from several limitations. PSO has limitations in handling fast-changing voltage and current parameters, which are common in dynamic grid conditions. On the other hand, ant colony optimization (ACO) and Bee Colony Optimization (BCO) overcome these limitations by offering adaptive protection schemes capable of responding to real-time changes in the grid. This adaptability makes ACO particularly suitable for ensuring grid stability under rapidly fluctuating conditions [62].

These advanced optimization techniques also provide a foundation for training AI models to detect and classify new fault patterns. AI systems, equipped with these models, can rapidly deploy adaptive protection schemes and reconfigure the grid in response to evolving conditions, thereby ensuring overall grid stability. The continuous learning capability of AI allows it to evolve alongside the grid, improving resilience and performance over time.

Furthermore, the integration of AI into WAMS has broader implications for the future of power systems. AI can serve as the cornerstone for a flexible AC transmission system (FACTS), enabling precise control of power flows and improving grid flexibility. It is also central to smart grid development, facilitating efficient energy management, demand response, and the integration of distributed energy resources. In modern CPPS, AI's ability to process vast amounts of data in real time ensures enhanced situational awareness, fault detection, and mitigation. AI's role in WAMS technology extends beyond simple optimization. It provides a comprehensive framework for enhancing grid stability, reliability, and adaptability in the face of increasing complexity and variability in modern energy systems. This synergy between AI and WAMS paves the way for a more resilient, efficient, and sustainable power grid capable of meeting the demands of renewable energy integration and evolving load dynamics. Optimization problems in WAMS can be addressed for the simultaneous modification and application of WAMS in degenerate power system networks [61]. Advanced communication infrastructure using fiber optics technology can also be flexibly employed in the WAMS architecture to achieve more cyber protection nominally guaranteed with GPS [25]. In simple terms, due to the high costs of PMUs and PDCs, it is not financially imperative to deploy PMUs on all system buses. With AI, the best possible path could be determined [61,66].

5. Cyber-Physical Power Systems and Smart Grid Development

The potential of the WAMS network to transform conventional power system networks is well established, but it comes with risks and concerns that potentially compound issues that smart grids were designed to solve [67]. The rise of modern communications systems means that existing power networks will inevitably be integrated to develop a more coherent smart grid architecture responsive to dynamic changes in the grid (as shown in Figure 6a) [67,68]. Despite the inevitability of the design, conventional power systems (especially those of sub-Saharan Africa) are naturally prone to conventional faults on a scale that threatens power system stability and operation. With a poor cybersecurity platform, highly volatile communications systems may be exploited to cause more damage to these networks [67].

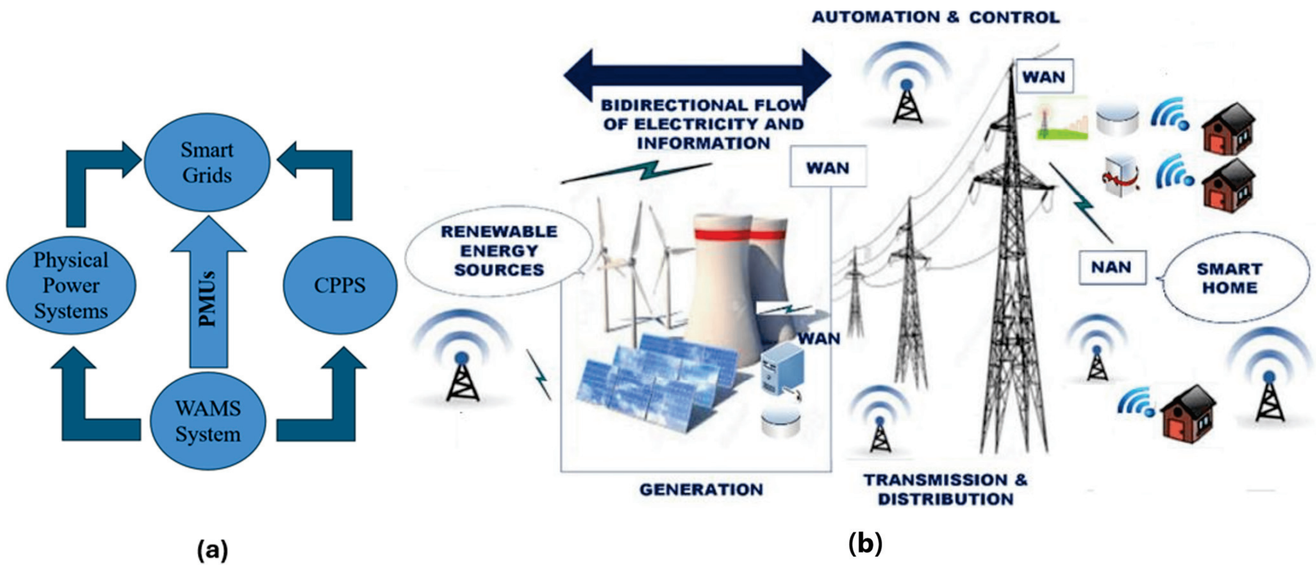


Figure 6. (**a**) Block diagram of how WAMS integrates into physical power system through strategic PMU deployment to become CPPS. (**b**) Schematic of potential application in multi-energy system with wide-area network (WAN) and neighborhood-area network (NAN).

Xu et al. pointed out different forms of cybersecurity attacks on CPPS integrated with Industry 4.0 tools, industrial ethernet, internet of things (IoT), and TCP/IP with the potential of testing the resilience of the CPPS [69,70]. With an analysis of the impact of cyberattacks on the Ukrainian power systems and the impact of natural hazards on CPPS, Xu et al. concluded that cybersecurity attacks on CPPS are categorized based on intent, which will then determine the form of protection strategies to be deployed on the network [69]. A CPPS, which is based on an ICS with high security requirements, can still fall prey to cybersecurity threats targeting integrity, availability, and confidentiality [67,69,71].

This paper's analysis of the future role of AI in WAMS and CPPS makes this threat more concerning. In mitigating these looming issues, advancements in WAMS design (and especially its communication infrastructure) should be built on cybersecurity models [67,71]. An absence of this may cause chaos, especially in cases where natural hazards like storms can derail most wireless communication systems at specific frequencies. Xu et al. emphasized the impact of natural hazards on physical systems while mostly focusing on wireless compromising power system networks. They encourage a holistic view of natural hazards, citing the impact of 2008's storms in China's southern province where 678 power lines (500 kV) and 1432 towers (220 kV) collapsed, compromising the existing CPPS and making it tenuous for integrating potential smart grid infrastructures [69]. Disruptions to the physical power system network often impact the cyber infrastructure, which is also complex to maintain and replace since WAPS are mostly integrated and coupled together (as shown in Figure 6b) [67,72]. Consequently, cyber threats arise for the PMUs and WAMS infrastructure, e.g., from hackers introducing malware into low-risk systems incorporated into the larger networks responsible for regulating voltage and current levels in the system. In this situation, a greater impact can mean that DERs, which are constantly being integrated into the conventional grid, can become a conduit for more comprehensive cyber-attacks in CPPS [72]. Ashok et al. also highlighted that when considering WAPS, attention must be paid to their framework [31]. Cybersecurity threats are not only in software areas [16]. Their study pointed out that physical damage to the power system could be the basis of further cyber threats that can evolve and compromise the entire system. A WAMS, as shown in Figure 6b, could help facilitate this transition from traditional power system networks to comprehensive CPPS in sub-Saharan Africa. With this understanding, cybersecurity measures should address both the software and hardware of a power system network.

Yohanandhan et al. assessed different models of visualizing CPPS interaction models. In each instance of the graphical, mechanical, probabilistic, and simulation models, they identify the Petri net model as the only one of the graphical models capable of visualizing sCPPS models to define both the interaction phenomena between the continuous nature of conventional power systems and emphasize the dynamic nature of cyber systems integrated with CPPS [71]. This emphasis is needed if accurate simulations provide the assessment of possible cyber threats to potential WAMS-based models. Similarly, research studies by Karamdel et al. followed a similar position, suggesting that potential CPPS may also include cloud-based energy management systems. They argue that financially motivated attacks will therefore seek to compromise the energy systems, since cloud-based systems have the potential to reduce financial losses in smart grid systems [67]. Aside from these attacks, coordinated cyber-physical attacks are potential ways that CPPS could be compromised and infiltrated [51,73]. If WAMS are to be deployed in conventional power system networks to transform them into credible CPPS, optimization models for deploying PMUs on the smart grid networks are to be prioritized since most dynamic load-altering attacks exploit PMU placement on WAMS [51,67,71]. Mitigation strategies can include several layers of optimization models, including multilayer models that ensure that the different categories of cyber-physical attacks on PMUs, RTUs, and SCADA systems can be identified and accurately addressed to ensure overall CPPS integrity [53,68,69]. State estimation models, the foundation of most energy management systems, can be compromised using frequency transmission techniques like false data injection attacks targeting simple RTU-PMU placement on the lines, which can ensure isolation at critical physical and cyber conditions that can mitigate the collapse of erstwhile conventional systems and current smart grid architecture [67,71]. The asynchronous, dynamic, and concurrent nature of CPPS in the operation means that PMU and WAWS operation will both be dynamic and designed to be responsive, especially in cloud-based systems; therefore, in such cases, state estimation may be performed spontaneously for other dynamic assessments to be conducted [67,68,71]. In this scenario, a multi-level optimization could make the system respond to dynamic changes in transient systems, which are mostly avenues for compromising smart grid systems. In simple terms, a multi-level CPPS attack will require a multi-level optimized mitigation technique to address potential damage to both system integrity and financial models of CPPS [67]. These complex optimization techniques can be operated well using AI systems integrated into CPPS and smart grid architecture. In such AI-integrated smart grid systems, machine learning (ML) models can be designed to test for the Karush–Kuhn–Tucker (KKT) conditions, which are essential in processing the nonlinear programming data reported from the dynamic system operations [67,68].

6. Conclusions

SCADA systems have become an integral part of conventional power systems. Emerging economies including Nigeria and other African countries have adopted SCADA systems to integrate their multiple protection components, including current transformers, power transformers, and relays [74]. The SCADA systems in these countries promise more sensitivity and awareness. Still, the incessant reports of faults on the grid have made them mostly incapable of participating in recent conversations around clean energy and renewable energy technologies. WAMS using PMUs, RTUs, and, in some cases, some SCADA infrastructures have been able to fill the gap of power system development and expansion comprehensively. The rise of AI has sparked expert conversations about the evolution of WAMS to accommodate cybersecurity challenges that may compromise clean energy development and renewable energy integration into the grid. With the increase in energy consumption due to population and the high impact of climate change (heat wave), power system expansion and potential CPPS are required to meet this demand [72]. This demand can greatly compromise existing energy infrastructure, especially conventional power system infrastructure in Nigeria and sub-Saharan African countries. Most of these energy grid systems are marred by inflexibility making them unresponsive to dynamic changes occurring in the grid. While smart grid systems have been installed in most of these cases to integrate DERs into the grid, CPPS designs remain the feasible solution for the overall grid [68]. Environmental factors compromise some aged systems, making them prone to faults and grid collapse during energy crises. Energy experts and researchers have identified WAMS as the significant way of making these conventional systems into smart grids and potentially transforming them into cyber-physical power system networks operating via optical-fiber-powered systems. This revolution in energy configuration will become the needed factor in placing these developing-world economies to contribute and participate in modern renewable energy conversations. However, WAMS are known to be expensive, and their flexibility depends on the design model applied in distributing the PMUs across the power system infrastructure. AI can play a significant role in ensuring that the WAM infrastructure is distributed along the line in a way that is both affordable and operationally effective. However, the CPPS and smart grid systems may also become prone to cyber-physical attacks, which could heavily compromise power system operation. Optimization techniques are the safest approach to deploy PMUs on the grid cost-effectively and to improve the integrity of the overall systems. These issues of optimization are quite concerning since the data collected are numerous and dynamic. The dynamic nature of the CPPS and smart grids will mean that highly accurate stochastic analysis will be required to identify cyber threats and still carry out control functions. AI could be efficient in such roles: rapidly retrieving and processing data, flagging cyber threats, and stacking up CPPS and smart grid behaviors simultaneously. AI is already informing power system expansion, the growth of the smart grid, and algorithms to maximize energy efficiency [34,42]. The observability and controllability of the power system network will also be greatly increased since AI models can improve the collection and analysis of the PMU data along the line and provide recommendations to control system operators for decisive steps to ensure power system integrity. AI could rapidly respond to security breaches and shut down sensitive infrastructure that may compromise long-term operations.

In conclusion, integrating AI with WAMS offers transformative improvement in fault mitigation. While traditional SCADA systems have long served as the backbone for monitoring and controlling electrical grids, their limitations in real-time data processing and prediction are increasingly evident. By contrast, AI-enhanced WAMS leverage sophisticated ML algorithms and analytics to offer unprecedented situational awareness and fault prediction. These systems can process vast amounts of data from PMUs at high speeds, providing operators with real-time insights and predictive analytics that enable proactive fault detection and swift corrective actions.

Moreover, the adaptability of AI ensures that these systems can continuously learn from historical data and improve their predictive models, making them highly effective in anticipating and mitigating potential faults before they escalate into significant issues. This continuous learning starkly contrasts with the static nature of SCADA systems, which rely heavily on predefined rules and manual interventions. AI-driven WAMS can identify patterns and anomalies that may not be immediately apparent to human operators, thus providing grid reliability and resilience.

Finally, integrating AI with WAMS facilitates advanced applications in CPPS, such as automated fault location, isolation, and system restoration, which are critical for minimizing downtime and maintaining smart grid stability. The ability to rapidly respond to and recover from disturbances reduces the economic impact of power outages and improves the power system's overall efficiency. In essence, the synergy between AI and WAMS mitigates faults more effectively than SCADA systems can. The combination ushers in a new era of intelligent, self-responsive, CPPS that far surpass the capabilities of traditional SCADA systems and WAMS individually. The future of power system management lies in this integration, promising more resilient, reliable, and efficient energy infrastructure.

7. Recommendations

The increasing complexity of modern and wide-area power systems, particularly in developing economies, highlights the need for advanced monitoring and control mechanisms to ensure grid reliability and support renewable energy integration. WAMS and its extended frameworks, including Wide-Area Monitoring and Control (WAMC), Wide-Area Protection and Control (WAMPAC), and Wide-Area Power, Control, and Monitoring (WAPCAM) systems, have emerged as key enablers of this transition. However, despite significant advancements in smart grid technology, limited research has explored the deployment of WAMS and their variations in traditional power system networks, especially in sub-Saharan Africa, where aging infrastructure and unstable grids persist. To address this gap, a structured approach is necessary, integrating real-time monitoring, AI-driven optimization, cybersecurity frameworks, and scalable financial models. This multi-prong approach is necessary for a robust sensitivity analysis to engage with regional stakeholders on the importance of WAMS in sustainable energy development.

A critical recommendation is the development of comprehensive models for deploying WAMS and WAMC, WAMPAC, and WAPCAM systems in WAPS across developing countries such as Nigeria and South Africa. These models should aim to enhance grid observability, improve fault detection, and enable large-scale renewable energy integration. Given the increasing penetration of DERs, such as solar photovoltaics (PVs) and wind power, future grid systems must leverage real-time synchrophasor data analytics to mitigate power fluctuations and ensure system stability. The integration of hybrid WAMS-SCADA frameworks can facilitate the transition from traditional power networks to adaptive smart grids, mitigating challenges associated with renewable energy intermittency and voltage fluctuations.

Another crucial area of focus is cybersecurity in CPPS and WAMS deployment. As power grids become more digitalized, the increasing vulnerability of PMUs, communication networks, and control centers to false data injection (FDI), denial-of-service (DoS) attacks, and malware infiltration remains a significant challenge. To enhance the security of WAMS-based systems, future grid architectures should integrate blockchain-based security mechanisms, AI-driven anomaly detection, and self-healing network protocols to safeguard against cyber threats. Furthermore, the interdependence between physical infrastructure and cyber networks must be addressed, as disruptions in physical grid components can trigger cascading failures in digital control systems. Additionally, a more robust cybersecurity system should be developed to strengthen the communication infrastructure required for WAMS deployment. This includes the adoption of security network protocols such as Transport Layer Security (TLS) to ensure encrypted data transmission and mitigate unauthorized access risks. As phishing attacks become more prevalent in PMU-based systems, increased efforts should be made to develop resilient authentication and intrusion detection frameworks to further secure power grid networks.

Ensuring the cost-effectiveness of WAMS deployment in developing economies is another key recommendation of this review. Some research suggests that the cost of deploying a PMU is more expensive than the PMUs themselves. This area must be swiftly improved to ensure a broader adoption of the technology. While industrialized nations have successfully integrated WAMS into their national grids, many developing countries face financial and infrastructural constraints. Strategies involving the development of low-cost PMUs, scalable fiber optic and 5G-enabled communication networks, and innovative financing models should be explored. Additionally, public–private partnerships and government-backed incentives can play a pivotal role in ensuring sustainable investments in energy infrastructure.

AI will be essential for optimizing WAMS functionality, particularly in fault detection, predictive maintenance, and grid resilience. The application of deep learning, reinforcement learning, and quantum computing will enable more accurate fault localization, reduced system downtime, and improved operational efficiency. By leveraging AI-enhanced WAMS, developing economies can optimize power flow dynamics, reduce transmission losses, and enhance the overall adaptability of their power grids.

Finally, policy standardization and regulatory frameworks must be established to ensure the interoperability and long-term sustainability of WAMS deployment. The absence of global WAMS implementation standards has resulted in fragmented adoption and compatibility issues across different grid infrastructures. Future efforts should prioritize the development of interoperability guidelines, cybersecurity compliance frameworks, and international research collaborations to ensure harmonized WAMS deployment across power networks.

Future Work

Although significant research has explored WAMS applications in modernized power systems, several gaps remain regarding its adaptation for traditional power grids, particularly in developing economies. Future research should focus on refining WAMS models that can be tailored to aging power infrastructures, improving fault detection mechanisms, and developing data-driven methodologies for optimizing grid expansion planning. This is especially crucial in Nigeria and South Africa, where frequent grid collapses and unreliable energy supply have hindered economic growth and industrial development.

A key area for future research is the expansion of WAMS to support renewable energy microgrid development. Many developing economies rely on off-grid and microgrid solutions to address energy access challenges [75,76]. Future studies should investigate how WAMS-enabled microgrid expansion can improve electrification in remote areas, ensuring grid stability while supporting decentralized renewable energy systems. Additionally, the WAMS-based optimization of distributed energy storage could play a crucial role in mitigating power supply fluctuations and enhancing overall system reliability.

Another pressing research avenue is the integration of AI-driven cybersecurity frameworks into WAMS-based power networks. With increasing cyber threats targeting critical power infrastructures, future work should explore adaptive cyber-physical security models that leverage machine learning and multi-agent systems to detect, mitigate, and prevent attacks in real time. Additionally, the development of self-learning cybersecurity architectures that can dynamically adjust to evolving threats will be instrumental in ensuring grid security in CPPS environments. Furthermore, enhanced protections against phishing attacks in PMU-based systems should be explored, as these attacks continue to pose a growing risk to data integrity and network security.

SGAM will also play a key role in the future development of WAMS. Future studies should establish methodologies for integrating WAMS within the SGAM's five-layer framework, enhancing grid interoperability, observability, and efficiency. Identifying scaling challenges and proposing solutions for implementing SGAM in traditional power grids, particularly in sub-Saharan Africa, will be crucial for modernizing outdated grid infrastructure.

Finally, economic and policy research should be expanded to support large-scale WAMS deployment. Future studies should analyze financing models for widespread WAMS adoption, including opportunities for international investments in smart grid infrastructure. Additionally, investigating the long-term economic benefits of WAMS deployment, particularly in relation to grid stability, renewable energy penetration, and industrial growth, will provide crucial insights for policymakers and energy stakeholders.

Author Contributions: Conceptualization, C.E.O. and S.O.O.; methodology, C.E.O.; software, C.E.O. and J.T.; validation, C.E.O., J.T. and S.O.O.; formal analysis, C.E.O. and S.O.O.; investigation, C.E.O. and J.T.; resources, C.E.O., J.T. and S.O.O.; data curation, C.E.O.; writing—original draft preparation, C.E.O.; writing—review and editing, C.E.O., J.T. and S.O.O.; visualization, C.E.O. and J.T.; supervision, J.T. and S.O.O.; project administration, J.T. and S.O.O.; funding acquisition, J.T. and C.E.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Carnegie Mellon University Africa and CMU Libraries.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: Author Jesse Thornburg is an owner of Grid Fruit, LLC. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACO	Ant colony optimization;
AI	Artificial intelligence;
ANN	Artificial neural network;
CPPS	Cyber-physical power system;
DER	Distributed energy resource;

EMS	Energy management system;
FACTS	Flexible AC transmission system;
GPS	Global Positioning System;
IoT	Internet of things;
IP	Internet protocol;
KKT	Karush–Kuhn–Tucker;
NAN	Neighborhood area network;
PDC	Phasor data controller;
PMU	Phasor measurement unit;
PSO	Particle swarm optimization;
RTU	Remote terminal unit;
SCADA	Supervisory control and data acquisition;
SGAM	Smart Grid Architecture Model;
SMT	Synchrophasor measurement technology;
WAMS	Wide-area measurement system;
WAMC	Wide-area monitoring and control;
WAMPAC	Wide-area monitoring, protection, and control;
WAN	Wide-Area network;
WAPCAM	Wide-area power control and monitoring;
WAPS	Wide-area power system.

References

1. Premkumar, S.; Susithra, M.; Saminadan, V. Performance Analysis of Smart Grid Communication Network Architecture for WAMS. In *Innovations in Electronics and Communication Engineering, Proceedings of the 6th ICIECE 2017, Hyderabad, India, 17–18 August 2018*; Lecture Notes in Networks and Systems; Springer: Singapore, 2018; pp. 347–354.
2. Wilson, V.; Li, F.; Thornburg, J.; Mohammadi, J.; Martinez, J. Energy Savings Through Refrigeration Load Control with Assessment of Commercial Potential: Lowering peak power and electricity bills through optimal demand scheduling. *IEEE Electrif. Mag.* **2024**, *12*, 66–76. [CrossRef]
3. Siyoi, V.; Nthontho, M.; Chowdhury, S.; Chowdhury, S.P. Wide Area Monitoring for Power System Protection—A Review. In Proceedings of the IEEE International Conference on Power System Technology, Auckland, New Zealand, 30 October–2 November 2012.
4. Dileep, G. A survey on smart grid technologies and applications. *Renew. Energy* **2020**, *146*, 2589–2625. [CrossRef]
5. Hojabri, M.; Dersch, U.; Papaemmanouil, A.; Bosshart, P. A Comprehensive Survey on Phasor Measurement Unit Applications in Distribution Systems. *Energies* **2019**, *12*, 4552. [CrossRef]
6. Goodman, C.; Thornburg, J.; Mohammadi, J. Optimization of Refrigeration Defrost Schedules for Demand Shifting in Commercial Buildings. In Proceedings of the 2021 IEEE Green Technologies Conference (GreenTech), Denver, CO, USA, 7–9 April 2021; pp. 201–208. [CrossRef]
7. Goodman, C.; Thornburg, J.; Ramaswami, S.K.; Mohammadi, J. Load Forecasting of Food Retail Buildings with Deep Learning. In Proceedings of the 2021 IEEE PES Innovative Smart Grid Technologies Conference-Latin America (ISGT Latin America), Lima, Peru, 15–17 September 2021; pp. 1–5. [CrossRef]
8. Castellani, F.; Astolfi, D.; Natili, F. SCADA data analysis methods for diagnosis of electrical faults to wind turbine generators. *Appl. Sci.* **2021**, *11*, 3307. [CrossRef]
9. Peeters, C.; Verstraeten, T.; Nowé, A.; Helsen, J. Wind Turbine Planetary Gear Fault Identification Using Statistical Condition Indicators and Machine Learning. In Proceedings of the 2019 International Conference on Prognostics and System Health Management (PHM), Paris, France, 2–5 May 2019.
10. Al-Odienat, A.; Alawasa, K.; Abu-Jaradeh, B.; Al-Walidi, K.; Dawood, M. Implementation of Wide-Area Monitoring System (WAMS) in Jordanian National Grid. In Proceedings of the 2021 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Amman, Jordan, 16–18 November 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 108–113.
11. Aminifar, F.; Fotuhi-Firuzabad, M.; Shahidehpour, M.; Safdarian, A. Impact of WAMS Malfunction on Power System Reliability Assessment. *IEEE Trans. Smart Grid* **2012**, *3*, 1302–1309. [CrossRef]
12. Zhao, J.; Tan, J.; Wu, L.; Zhan, L.; Yao, W.; Liu, Y. Impact of the Measurement Errors on Synchrophasor-Based WAMS Applications. *IEEE Access* **2019**, *7*, 143960–143972. [CrossRef]

13. Song, Z.; Hackl, C.M.; Anand, A.; Thommessen, A.; Petzschmann, J.; Kamel, O.; Braunbehrens, R.; Kaifel, A.; Roos, C.; Hauptmann, S. Digital Twins for the Future Power System: An Overview and a Future Perspective. *Sustainability* **2023**, *15*, 5259. [CrossRef]
14. Zamani, R.; Panahi, H.; Abyaz, A.; Alhelou, H.H. Introduction to WAMS and Its Applications for Future Power System. In *Power Systems*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 45–69.
15. Sundararajan, A.; Khan, T.; Moghadasi, A.; Sarwat, A.I. Survey on synchrophasor data quality and cybersecurity challenges, and evaluation of their interdependencies. *J. Mod. Power Syst. Clean Energy* **2018**, *7*, 449–467. [CrossRef]
16. Jamil, N.; Qassim, Q.S.; Bohani, F.A.; Mansor, M.; Ramachandaramurthy, V.K. Cybersecurity of Microgrid: State-of-the-Art review and possible directions of future research. *Appl. Sci.* **2021**, *11*, 9812. [CrossRef]
17. Panda, D.K.; Das, S. Smart grid architecture model for control, optimization and data analytics of future power networks with more renewable energy. *J. Clean. Prod.* **2021**, *301*, 126877. [CrossRef]
18. Manoj, K.S. *Power System Automation*; Notion Press: Chennai, India, 2021.
19. Nuthalapati, S. *Power System Grid Operation Using Synchrophasor Technology*; Springer: Berlin/Heidelberg, Germany, 2018.
20. Sayed, K.; Abo-Khalil, A.G.; Eltamaly, A.M. Wind Power Plants Control Systems Based on SCADA System. In *Green Energy and Technology*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 109–151.
21. Paramo, G.; Bretas, A.; Meyn, S. Research Trends and Applications of PMUs. *Energies* **2022**, *15*, 5329. [CrossRef]
22. Alhelou, H.H.; Golshan, M.E.H.; Njenda, T.C.; Hatziargyriou, N.D. An Overview of UFLS in Conventional, Modern, and Future Smart Power Systems: Challenges and Opportunities. *Electr. Power Syst. Res.* **2020**, *179*, 106054. [CrossRef]
23. Suhaimy, N.; Radzi, N.A.M.; Ahmad, W.S.H.M.W.; Azmi, K.H.M.; Hannan, M.A. Current and Future Communication Solutions for Smart Grids: A Review. *IEEE Access* **2022**, *10*, 43639–43668. [CrossRef]
24. Appasani, B.; Mohanta, D.K. A review on synchrophasor communication system: Communication technologies, standards and applications. *Prot. Control. Mod. Power Syst.* **2018**, *3*, 37. [CrossRef]
25. Alhelou, H.H.; Abdelaziz, A.Y.; Siano, P. *Wide Area Power Systems Stability, Protection, and Security*; Springer Nature: Berlin/Heidelberg, Germany, 2020.
26. Sospiro, P.; Harper, J.; Easton, M. Smart Grid in China, EU, and the US: State of Implementation. *Energies* **2021**, *14*, 5637. [CrossRef]
27. Leibovich, P.; Issouribehere, F.; Barbero, J. IoT Platforms for WAMS Systems: A complete Synchrophasor Measurement System in the Cloud. In Proceedings of the 2022 IEEE Power and Energy Society General Meeting (PESGM), Denver, CL, USA, 17–21 July 2022.
28. Chintakindi, R.; Mitra, A. WAMS challenges and limitations in load modeling, voltage stability improvement, and controlled island protection-A review. *Energy Rep.* **2022**, *8*, 699–709. [CrossRef]
29. Vahidi, S.; Ghasemi, H.; Shamsollahi, P.; Mazloom, A.; Amini, B. Security of Wide-Area Monitoring, Protection, and Control (WAMPAC) Systems of the Smart Grid: A Survey on Challenges and Opportunities. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 1294–1335. [CrossRef]
30. Khan, M.W.; Siddiqui, I.; Ahmad, Z. Optimal Control and Communication Strategies in Multi-Energy Generation Grid. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 2599–2653. [CrossRef]
31. Ashok, V.; Yadav, A.; Abdelaziz, A.Y. A Comprehensive Review on Wide-Area Protection, Control and Monitoring Systems. In *Wide Area Power Systems Stability, Protection, and Security*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 1–43.
32. Talpur, S.; Lie, T.T. PMU based WAMC application in multi-modular HVDC based large scale solar system. In Proceedings of the 2017 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia), Auckland, New Zealand, 4–7 December 2017; pp. 1–3. [CrossRef]
33. Rao, A.N.; Priya, P.V.; Kowsalya, M.; Gnanadass, R. Wide area monitoring for energy system: A review. *Int. J. Ambient. Energy* **2017**, *40*, 537–553. [CrossRef]
34. Goodman, C.; Thornburg, J.; Ramaswami, S.K.; Mohammadi, J. Building power grid 2.0: Deep learning and federated computations for energy decarbonization and edge resilience. In *Deep Learning Applications, Volume 3*; Advances in Intelligent Systems and Computing; Wani, M.A., Raj, B., Luo, F., Dou, D., Eds.; Springer Nature: Berlin/Heidelberg, Germany, 2022; Volume 1395, pp. 215–230. [CrossRef]
35. Zhou, Z.; Zhang, X.; Zhang, J.; Taylor, M. Comprehensive review on dynamic state estimation techniques with cybersecurity applications. *IET Smart Grid* **2024**, *7*, 370–385. [CrossRef]
36. Nadeem, M.; Qureshi, S.; Taylor, P. Dynamic State Estimation of Nonlinear Differential Algebraic Equation Models of Power Networks. *IEEE Trans. Power Syst.* **2023**, *38*, 2539–2552. [CrossRef]
37. Darmis, O.; Korres, G. A Survey on Hybrid SCADA/WAMS State Estimation Methodologies in Electric Power Transmission Systems. *Energies* **2023**, *16*, 618. [CrossRef]
38. Saleem, M.U.; Ashraf, I.; Iqbal, S. Design, Deployment and Performance Evaluation of an IoT Based Smart Energy Management System for Demand Side Management in Smart Grid. *IEEE Access* **2022**, *10*, 15261–15278. [CrossRef]
39. Jamal, S.; Hussain, A.; Rehman, Z. A Review of Energy Management and Power Management Systems for Microgrid and Nanogrid Applications. *Sustainability* **2021**, *13*, 10331. [CrossRef]

40. Pandit, R.; Astolfi, D.; Hong, J.; Infield, D.; Santos, M. SCADA data for wind turbine data-driven condition/performance monitoring: A review on state-of-art, challenges and future trends. *Wind Eng.* **2022**, *47*, 422–441. [CrossRef]
41. Heidemann, J.; Kumar, M.; Basu, S. Research Challenges and Applications for Underwater Sensor Networking. *IEEE Wirel. Commun. Netw. Conf. (WCNC)* **2006**, *1*, 228–235.
42. Mohammadi, M.B.; Hooshmand, R.-A.; Fesharaki, F.H. A New Approach for Optimal Placement of PMUs and Their Required Communication Infrastructure in Order to Minimize the Cost of the WAMS. *IEEE Trans. Smart Grid* **2016**, *7*, 84–93. [CrossRef]
43. Biswal, C.; Sharma, A.; Singh, N. Real-Time Grid Monitoring and Protection: A Comprehensive Survey on the Advantages of Phasor Measurement Units. *Energies* **2023**, *16*, 4054. [CrossRef]
44. Vincent, E.N.; Yusuf, S.D. Integrating Renewable Energy and Smart Grid Technology into the Nigerian Electricity Grid System. *Smart Grid Renew. Energy* **2014**, *5*, 220–238. [CrossRef]
45. Gore, R.N.; Valsan, S.P. Wireless communication technologies for smart grid (WAMS) deployment. In Proceedings of the 2022 IEEE International Conference on Industrial Technology (ICIT), Lyon, France, 20–22 February 2018; Volume ii, pp. 1326–1331. [CrossRef]
46. Cruz, M.A.R.S.; Rocha, H.R.O.; Paiva, M.H.M.; Segatto, M.E.V.; Camby, E.; Caporossi, G. An algorithm for cost optimization of PMU and communication infrastructure in WAMS. *Int. J. Electr. Power Energy Syst.* **2018**, *106*, 96–104. [CrossRef]
47. Saoud, A.; Recioui, A. PMU Reporting Rate Optimization for Data Transfer Enhancement in Wide Area Monitoring Systems. In Proceedings of the 2020 International Conference on Electrical Engineering (ICEE), Istanbul, Turkey, 5–27 September 2020.
48. Goldstein, A.; Maragal, D. Synchrophasor Measurement System Industry Standards and Guidance: A Comprehensive Overview. *IEEE Trans. Power Deliv.* **2024**, *39*, 1687–1696. [CrossRef]
49. Bhonsle, J.; Junghare, A. Optimal placing of PMUs in a constrained grid: An approach. *Turk. J. Electr. Eng. Comput. Sci. Elektr.* **2016**, *24*, 4508–4516. [CrossRef]
50. Almalaq, A.; Albadran, S.; Mohamed, M. Deep Machine Learning Model-Based Cyber-Attacks detection in smart power systems. *Mathematics* **2022**, *10*, 2574. [CrossRef]
51. Avatefipour, O.; Al-Sumaiti, A.S.; El-Sherbeeny, A.M.; Awwad, E.M.; Elmeligy, M.A.; Mohamed, M.A.; Malik, H. An intelligent secured framework for cyberattack detection in electric vehicles' CAN bus using machine learning. *IEEE Access* **2019**, *7*, 127580–127592. [CrossRef]
52. Morales, J.M.; Conejo, A.J.; Madsen, H.; Pinson, P.; Zugno, M. *Integrating Renewables in Electricity Markets*; Springer Science and Business Media: Berlin/Heidelberg, Germany, 2013.
53. Ivanov, I.; Dimova, R. Testbed for One-Way IP Packet Delay Measurement in Smart Grid WAMS Applications Based on IEEE C37.118. *Comput. Sci. Commun.* **2014**, *1*, 78–85.
54. Rouhani, S.H.; Su, C.-L.; Mobayen, S.; Razmjooy, N.; Elsisi, M. Cyber resilience in renewable Microgrids: A Review of Standards, Challenges, and Solutions. *Energy* **2024**, *309*, 133081. [CrossRef]
55. Hassak, S.D.; Leger, A.S.; Oh, H.; Opila, D.F. Implementing a PMU Based Linear State Estimator on a Single Board Computer. In Proceedings of the 2022 North American Power Symposium (NAPS), Salt Lake City, UT, USA, 9–11 October 2022.
56. Khalid, H.M.; Flitti, F.; Mahmoud, M.S.; Hamdan, M.M.; Muyeen, S.M.; Dong, Z.Y. Wide area monitoring system operations in modern power grids: A median regression function-based state estimation approach towards cyber attacks. *Sustain. Energy Grids Netw.* **2023**, *34*, 101009. [CrossRef]
57. Santos, F.C.D.; Thornburg, J.; Ustun, T.S. Automated Planning of Rooftop PV Systems with Aerial Image Processing. In Proceedings of the 2018 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Kota Kinabalu, Malaysia, 7–10 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 736–740. [CrossRef]
58. Zhao, S.; Blaabjerg, F.; Wang, H. An Overview of Artificial Intelligence Applications for Power Electronics. *IEEE Trans. Power Electron.* **2021**, *36*, 4633–4658. [CrossRef]
59. Shi, Z.; Yao, W.; Li, Z.; Zeng, L.; Zhao, Y.; Zhang, R.; Tang, Y.; Wen, J. Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions. *Appl. Energy* **2020**, *278*, 115733. [CrossRef]
60. Mabrouki, J.; Darghouth, R.; Hamdi, N. *Advanced Technology for Smart Environment and Energy*; Springer Nature: Berlin/Heidelberg, Germany, 2023.
61. Abdelaziz, A.Y.; Aleem, S.H.E.A.; Yadav, A. *Artificial Intelligence Applications in Electrical Transmission and Distribution Systems Protection*; Springer Nature: Berlin/Heidelberg, Germany, 2021.
62. Alhamrouni, I.; Kahar, N.H.A.; Salem, M.; Swadi, M.; Zahroui, Y.; Kadhim, D.J.; Mohamed, F.A.; Nazari, M.A. A Comprehensive review on the role of Artificial intelligence in power system stability, control, and Protection: Insights and future Directions. *Appl. Sci.* **2024**, *14*, 6214. [CrossRef]
63. Almalaq, A.; Albadran, S.; Alghadhban, A.; Jin, T.; Mohamed, M.A. An Effective Hybrid-Energy Framework for Grid Vulnerability Alleviation under Cyber-Stealthy Intrusions. *Mathematics* **2022**, *10*, 2510. [CrossRef]
64. Hadi Abdulwahid, A.; Wang, S. A Novel Method of Protection to Prevent Reverse Power Flow Based on Neuro-Fuzzy Networks for Smart Grid. *Sustainability* **2018**, *10*, 1059. [CrossRef]

65. Polster, S.; Renner, H.; Schürhuber, R.; Friedl, K. On WAMS-based emergency outage detection using optimization. *Int. J. Electr. Power Energy Syst.* **2021**, *131*, 107006. [CrossRef]
66. El-Ela, A.A.A.; Mouwafi, M.T.; Elbaset, A.A. *Modern Optimization Techniques for Smart Grids*; Springer Nature: Berlin/Heidelberg, Germany, 2022.
67. Karamdel, S.; Liang, X.; Faried, S.O.; Mitolo, M. Optimization models in cyber-physical power systems: A review. *IEEE Access* **2022**, *10*, 130469–130486. [CrossRef]
68. Kumar, S.; Soni, M.K.; Jain, D.K. Cyber security threats in synchrophasor system in WAMS. *Int. J. Comput. Appl.* **2015**, *115*, 17–22. [CrossRef]
69. Xu, L.; Guo, Q.; Sheng, Y.; Muyeen, S.; Sun, H. On the resilience of modern power systems: A comprehensive review from the cyber-physical perspective. *Renew. Sustain. Energy Rev.* **2021**, *152*, 111642. [CrossRef]
70. Chen, J.; Mohamed, M.A.; Dampage, U.; Rezaei, M.; Salmen, S.H.; Al Obaid, S.; Annuk, A. A Multi-Layer Security Scheme for Mitigating Smart Grid Vulnerability against Faults and Cyber-Attacks. *Appl. Sci.* **2021**, *11*, 9972. [CrossRef]
71. Yohanandhan, R.V.; Elavarasan, R.M.; Manoharan, P.; Mihet-Popa, L. Cyber-Physical Power System (CPPS): A Review on Modeling, Simulation, and Analysis with Cyber Security Applications. *IEEE Access* **2020**, *8*, 151019–151064. [CrossRef]
72. Paul, S.; Ding, F.; Utkarsh, K.; Liu, W.; O'Malley, M.J.; Barnett, J. On Vulnerability and Resilience of Cyber-Physical Power Systems: A Review. *IEEE Syst. J.* **2021**, *16*, 2367–2378. [CrossRef]
73. Calderón, D.; Folgado, F.J.; González, I.; Calderón, A.J. Implementation and experimental application of industrial IoT architecture using automation and IoT Hardware/Software. *Sensors* **2024**, *24*, 8074. [CrossRef]
74. Ahaotu, M.C.; Ogbogu, C.E.; Thornburg, J.; Akwukwaegbu, I.O. Simulation of PEM Electrolyzer Power Management with Renewable Generation in Owerri, Nigeria. *Energies* **2025**, *18*, 208. [CrossRef]
75. Thornburg, J.; Krogh, B.H. A tool for assessing demand-side management and operating strategies for isolated microgrids. *Energy Sustain. Dev.* **2021**, *64*, 15–24. [CrossRef]
76. Thornburg, J.; Ustun, T.S.; Krogh, B. Smart microgrid operation simulator for management and electrification planning. In Proceedings of the 2016 IEEE PES PowerAfrica, Livingstone, Zambia, 28 June–3 July 2016; pp. 1–5. [CrossRef]

energies

MDPI

Article

Deep Reinforcement Learning-Based Distribution Network Planning Method Considering Renewable Energy

Liang Ma [1,*], Chenyi Si [2], Ke Wang [2], Jinshan Luo [1], Shigong Jiang [1] and Yi Song [1]

[1] State Grid Economic and Technological Research Institute Co., Ltd., Beijing 102209, China; luojinshan@chinasperi.sgcc.com (J.L.); jiangshigong@chinasperi.sgcc.com.cn (S.J.); songyi@chinasperi.sgcc.com.cn (Y.S.)
[2] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; sichenyi2021@163.com (C.S.); walker_wang@tju.edu.cn (K.W.)
* Correspondence: liang_ma276@163.com

Abstract: Distribution networks are an indispensable component of modern economic societies. Against the background of building new power systems, the rapid growth of distributed renewable energy sources, such as photovoltaic and wind power, has introduced many challenges for distribution network planning (DNP), including different source-load compositions, complex network topologies, and varied application scenarios. Traditional heuristic algorithms are limited in scalability and struggle to address the increasingly complex optimization problems of DNP. The emergence of new artificial intelligence provides a new way to solve this problem. Based on the above discussion, this paper proposes a DNP method based on deep reinforcement learning (DRL). By defining state space and action space, a Markov decision process model tailored for DNP is formulated. Then, a multi-objective optimization function and a corresponding reward function including construction costs, voltage deviation, renewable energy penetration, and electricity purchase costs are designed to guide the generation of network topology schemes. Based on the proximal policy optimization algorithm, an actor-critic-based autonomous generation and adaptive adjustment model for DNP is constructed. Finally, the representative test case is selected to verify the effectiveness of the proposed method, which indicates that the proposed method can improve the efficiency of DNP and promote the digital transformation of DNP.

Keywords: DNP method; DRL; Markov decision process model; proximal policy optimization

1. Introduction

As an important component of modern power systems, distribution networks play a crucial role in promoting economic development, improving livelihoods, and facilitating energy transitions [1–3]. With the development of distributed renewable energy sources (such as photovoltaics (PVs) [4] and wind turbine generators (WTGs) [5]) and the emergence of new load types (such as electric vehicle), distribution systems are characterized by diverse source-load configurations and a complex grid structure. In addition, the variability of solar irradiance and wind speed in different regions brings uncertainties, which present new challenges for the task of distribution network planning (DNP). However, traditional planning methods primarily rely on human experience, which are influenced by planners' expertise. These methods have the drawbacks of low efficiency and suboptimal quality. Therefore, how to choose appropriate methods to achieve efficient DNP schemes generation is a worthy issue.

As an effective tool, heuristic algorithms [6–8] (such as the simulated annealing algorithm (SAA), genetic algorithm (GA), ant colony optimization (ACO), and particle swarm optimization (PSO)) are widely used in the field of distribution networks. By integrating the steepest descent method with the SAA, ref. [9] proposed a DNP approach in which the capital recovery, energy loss, and undelivered energy costs were considered simultaneously. By combining a GA with the k-nearest neighbors structure, a hybrid GA-based radial distribution network reconfiguration method was displayed in [10] which reduced the search space while enhancing planning performance. Ref. [11] employed fuzzy logic systems to process the objective function and further utilized ACO to determine the optimal solution, which solved the optimal configuration of switching devices in the distribution network. A two-stage heuristic structure was constructed in [12] to choose the switches, and the reconfiguration of radial distribution networks was solved. By employing heuristic algorithms to determine the service areas of substations, ref. [13] addressed the substation planning problem for large-scale distribution networks. Additionally, the unreliability of feeders and substations was studied simultaneously. Ref. [14] proposed a new hybrid optimization approach combining PSO and Tabu Search to deal with the expansion planning of large-scale electric distribution networks. Ref. [15] solved the multi-objective reconfiguration of distribution networks with distributed power sources, in which an improved binary PSO algorithm was designed to enhance global search ability and accelerate convergence. By combining binary PPO for network investment decisions and inner optimization for flexibility procurement, cost-effective and adaptable distribution network expansion planning was achieved in [16]. A GA and PSO were combined in [17] to determine the best locations and sizes of distributed generation units, which reduced power losses and enhance voltage stability while respecting safe constraints. Based on PSO, an enhanced fault-recovery reconfiguration method for DC distribution networks was proposed in [18] which ensured that power can be quickly restored after a fault occurs. However, heuristic algorithms have limited solving capabilities, and the solutions generated by heuristic algorithms exhibit low generalization, which makes it difficult to address the increasingly complex optimization problems of DNP. Nowadays, artificial intelligence technologies have been applied in the entire spectrum of power systems, including generation, transmission, substations, distribution, and consumption. Moreover, many complex tasks, such as load forecasting [19], voltage stability control [20], and emergency control [21], have been implemented by utilizing artificial intelligence technologies.

In recent years, the artificial intelligence technologies represented by deep reinforcement learning (DRL) have experienced rapid development [22]. By combining deep learning and reinforcement learning, DRL can effectively deal with complex decision problems. The main idea of DRL is that agents continuously interact with their environment to obtain optimal policies. Recently, DRL has been extensively applied in the field of power systems, such as voltage control [23], frequency control [24], and load shedding control [25]. In [23], uncertainties in operating conditions and system physical parameters were considered, and a DRL-based voltage control approach was proposed for maintaining voltage within a safe range. Based on multi-agent DRL, a data-driven frequency control method for multi-area power systems was proposed in [24] which effectively reduced the frequency fluctuations caused by load variations and renewable energy sources. By designing a novel voltage-based reward function, the deep deterministic policy gradient (DDPG) algorithm was utilized to achieve load shedding control in [25], improving the ability of anti-disturbance. Refs. [26–28] utilized soft actor-critic to solve issues in electric vehicle charging, dynamic network reconfiguration, and volt-VAR control, respectively. By utilizing the DDPG to distribute torque between two motors to enhance energy efficiency, a hierarchical energy management strategy structure was constructed in [29]. In [30], a multi-agent DRL approach was developed which combined discrete DQN and continuous DDPG on different time scales to optimize long-term voltage deviations. In [31], risk sensitivity and the uncertainties of renewable generation were integrated in TRPO, and a systematic way for multiple microgrids to cooperate was proposed. Proximal policy optimization (PPO) [32] was proposed by OpenAI in 2017, which is a reinforcement learning method based on the actor-critic (AC) architecture. By integrating trust region policy optimization with the advantage AC, the PPO algorithm prevents performance degradation caused by excessively large policy updates. In [33], a two-level planning model based on PPO was developed, and a grid structure optimization algorithm based on electrical coupling degree was proposed which reduced the search space and alleviated the difficulty of acquiring splitting sections. In [34], a multi-agent PPO was introduced to construct the multi-layer energy management system to deal with challenges in mechanism modeling and supply-demand uncertainty. PPO was introduced to learn the optimal policy for flexible battery energy storage systems in [35], in which the working time and the balancing performance were optimized simultaneously. Ref. [36] presented a PPO-based decentralized framework for the voltage control of microgrids. Ref. [37] took cyber contingencies into consideration, designing a graph-based DRL framework in which the graph information was embedded into the PPO algorithm to alleviate the impact of cyber contingencies. Stackelberg–Nash PPO was constructed in [38] to deal with the influence of neglecting consumers' interests and temporal relationships between participant actions. Based on the above discussion, DRL has been applied in multiple fields such as voltage control, reactive power dispatch, load shaping, demand response, and energy storage optimization. However, there are still some issues that need to be resolved; for example, few DNP studies consider the topology structure of distribution networks, while some DRL frameworks only focus on one main objective and neglect other factors. Thus, we propose a DRL-based DNP method considering renewable energy in this paper. The contribution of this paper is as follows:

(1) By designing the action space and state space, we construct a Markov decision process (MDP) for DNP. Through this transformation, the problem of DNP can thus be integrated into the DRL framework.
(2) Considering line construction costs, voltage deviation, renewable energy subsidies, and electricity purchasing costs, a multi-objective optimization function is designed. Additionally, a multi-reward function is developed to guide the agent to learn the optimal policy. After sufficient training, the agent can generate optimized planning schemes.
(3) Based on the AC architecture, a PPO-based DNP algorithm (PPODNPA) is designed. The actor network (AN) generates planning schemes upon receiving reward signals and environmental states, while the critic network (CN) further evaluates these schemes to optimize the AN, which achieves the autonomous generation and adaptive tuning of the planning scheme. Simulation results validate the superiority of the proposed algorithm.

The remaining part of the paper is as follows: In Section 2, the MDP for DNP and planning model is defined. In Section 3, the PPODNPA is designed, which achieves the autonomous generation and adaptive tuning of the DNP scheme. In Section 4, the simulation results are displayed and heuristic algorithms are introduced to demonstrate the superiority of PPODNPA. A flow chart of this paper is displayed in Figure 1.

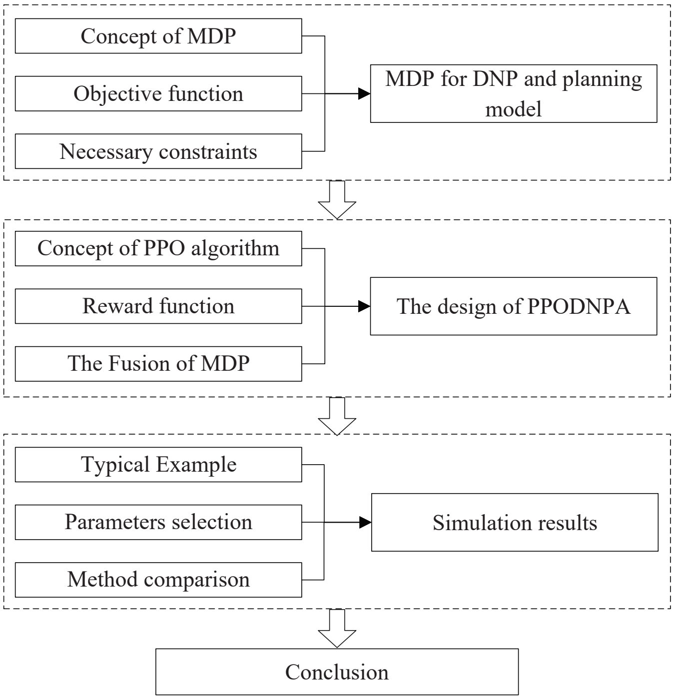


Figure 1. Flow chart.

2. MDP for DNP and Planning Model

2.1. MDP for DNP

In general, MDP can be represented by a quintuple (S, A, P, R, γ) [39], where S denotes the set of states, A denotes the set of actions, P represents the state transition probabilities, and R is the reward. The diagram of the MDP is illustrated in Figure 2, where S_i corresponds to the ith state, P_{ij}^k denotes the probability of transitioning from state s_i to state s_j after taking action a_k, and R_i^k indicates the immediate reward generated by taking action a_k in state s_i.

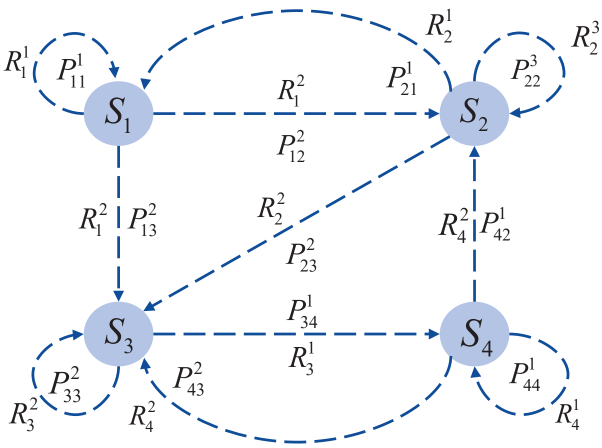


Figure 2. Diagram of MDP.

Considering the characteristics of DNP tasks, we assign physical meanings to various parameters of the MDP model. First, we define the state space as

$$S = \left\{ \begin{array}{l} P^{WTG}, Q^{WTG}, P^{PV}, Q^{PV}, \\ L_m, U_i, I_{ij}, D^{WTG}, D^{PV}, S_{ij} \end{array} \right\} \quad (1)$$

where P^{WTG} and Q^{WTG} represent the active power and reactive power of WTG, P^{PV} and Q^{PV} represent the active power and reactive power of PV, L_m represents the construction status of the lines, U_i represents the node voltage, I_{ij} represents the line current, D^{WTG} and D^{PV} represent the installation status of PV and WTG, and S_{ij} represents the result of load flow calculation. Then, we define the action space:

$$A = \{a_l, a_{WTG}, a_{PV}\} \tag{2}$$

where a_l represents the construction of lines, a_{WTG} represents the construction of WTG, and a_{PV} represents the construction of PV.

Remark 1. *It should be noted that load fluctuations, uncertainties in solar irradiance, and regional variations in wind power are factors that need to be considered. In this study, we use normal distribution to simulate uncertainties in PV output and load, and use Weibull distribution to simulate uncertainties in WTG output.*

After defining the state space and action space, we establish the transition of the distribution network framework: By applying an action to the initial network, both its topology and the installation status of renewable energy sources will be changed. Due to the inherent uncertainties associated with renewable energy, state variables such as power, line current, and node voltage will be converted into other states with certain probabilities. This definition will be utilized in the subsequent algorithm design. In the next section, the planning model will be further proposed, which includes the objective function and some necessary constraints.

2.2. Planning Model

The objective function is the economic cost of DNP considering renewable energy, which includes investment cost, electricity purchase cost, and renewable energy subsidies. It is defined as follows:

$$\min : C_{\text{total}} = N_{\text{cost}} \frac{k(1+k)^y}{(1+k)^y - 1} + \beta_1 E_{\text{buy}} - \beta_2 E_{\text{sub}} \tag{3}$$

where N_{cost} represents the initial investment cost of the grid structure, k is the discount rate, y is the service life, the overhead line is taken as 30 years, the cable is taken as 40 years, and the factor $\frac{k(1+k)^y}{(1+k)^y-1}$ can convert the initial investment into the equivalent year cost. $\beta_1 = 0.394\ \text{CNY/kW} \cdot \text{h}$ represents the purchase price for purchasing electricity from the external grid, E_{buy} is the amount of electricity purchased from the external grid to satisfy local demand. $\beta_2 = 0.2\ \text{CNY/kW} \cdot \text{h}$ represents the subsidy unit price of renewable energy, and E_{sub} represents the amount of renewable energy generation.

To ensure safe, reliable, and efficient operation of the DNP, the following constraints are incorporated:

(1) Power balance constraints [40]

$$\begin{aligned}
&U^{\text{ac}}_{i',t} = U^{\text{ac}}_{i,t} - \left(r^{\text{ac}}_{ii'} H^{\text{ac}}_{ii',t} + x^{\text{ac}}_{ii'} G^{\text{ac}}_{ii',t}\right) / U^{\text{ac}}_0, \forall i, i', t : \beta^{\text{ac}}_{ii',t} \\
&P^{\text{de}}_{i,t} - L^{\text{ac}}_{i,t} + \sum_{j_1 \in \Omega(i_1)} \sum_{l \in \mathcal{L}} \left(P^{\text{ac}}_{i_1 j_1, l, t} - P^{\text{acdc}}_{i_1 j_1, l, t}\right) = \sum_{i' \in \pi(i)} H^{\text{ac}}_{ii',t} \\
&\quad - \sum_{i'' \in \delta(i)} H^{\text{ac}}_{i'' i,t} - \left(P^{\text{w}}_{i,t} - \rho^{\text{w}}_{i,t}\right), \forall i, t : \beta^{\text{ac2}}_{i,t} \\
&Q^{\text{de}}_{i,t} - Q^{\text{ac}}_{i,t} = \sum_{i' \in \pi(i)} G^{\text{ac}}_{ii',t} - \sum_{i'' \in \delta(i)} G^{\text{ac}}_{i'' i,t}, \forall i, t : \beta^{\text{ac3}}_{i,t}
\end{aligned} \tag{4}$$

where $U^{ac}_{i,t}$ is the voltage magnitude, $r^{ac}_{ii'}$ and $x^{ac}_{ii'}$ are the resistance and reactance, respectively. $H^{ac}_{ii',t}$ and $G^{ac}_{ii',t}$ are the active and reactive power flow, respectively. $P^{de}_{i,t}$ and $L^{ac}_{i,t}$ are the output of the distributed generation and active load, respectively. $P^{w}_{i,t}$ and $\rho^{w}_{i,t}$ are the predicted wind power and curtailed wind power, respectively. $Q^{de}_{i,t}$ and $Q^{ac}_{i,t}$ are the reactive power of diesel engines and the reactive load, respectively. Equation (4) ensures that the power flowing into each node balances the load, generation, and line losses, which reflects Kirchhoff's current and voltage laws.

(2) The line current constraint and voltage constraint are calculated as follows:

$$\begin{gathered} I_{j\,\min} \leq I_j \leq I_{j\,\max} \\ U_{i\,\min} \leq U_i \leq U_{i\,\max} \end{gathered} \tag{5}$$

where $I_{j\,\max}$ and $I_{j\,\min}$ are the maximum and minimum values of the current, respectively, which reflect the thermal limits of the overhead line and cable. $U_{i\,\max}$ and $U_{i\,\min}$ are the maximum and minimum values of the voltage, respectively, which ensure power quality. Considering Equation (5) can prevent the overheating of lines and maintain the node voltage within safe ranges.

(3) The active power constraint is calculated as follows:

$$|P_i| \leq P_{i\,\max} \tag{6}$$

where P_i is the active power of the line and $P_{i\,\max}$ is the maximum active power. This constraint ensures that no line is influenced by real-power flows beyond its rated capability.

(4) Line flow constraints

$$S_{ij} \leq S_{ij}^{\max} \tag{7}$$

where S_{ij} is the line flow, and $S_{ij}^{\max}$ is the maximum value of the line flow. Considering Equation (7) can limit the power flow to safe levels considering both real power and reactive power.

(5) Radial grid constraint

$$n = m + 1 \tag{8}$$

where n is the number of nodes in the distribution network, and m is the sum of the original and future lines. This constraint ensures a radial topology of grid structure, which is a common structural requirement in DNP.

The investment cost, electricity purchase cost from external sources, and the subsidy income generated by renewable energy generation are comprehensively considered in the planning model. The objective is to minimize the total cost under the power balance constraints, line current and voltage constraints, active power constraints, line flow constraints, and radial grid constraints. By obtaining an equilibrium solution between economic viability and renewable energy utilization, this planning model provides a systematic framework for DNP, and the resulting solution can give a more sustainable, cost-effective, and reliable DNP scheme.

3. The Design of the PPODNPA

3.1. PPO Algorithm

PPO algorithm is an actor-critic method which includes CN and AN. As illustrated in Figure 3, AN is responsible for generating action probability distributions, and CN provides evaluated reward values. Consequently, the AN does not need to compute cumulative rewards for each episode. Instead, it learns from the reward evaluations supplied by the CN, which enables single-step updates and enhances training efficiency. The CN fits more accurate value functions through temporal difference (TD) errors to help the AN in selecting actions.

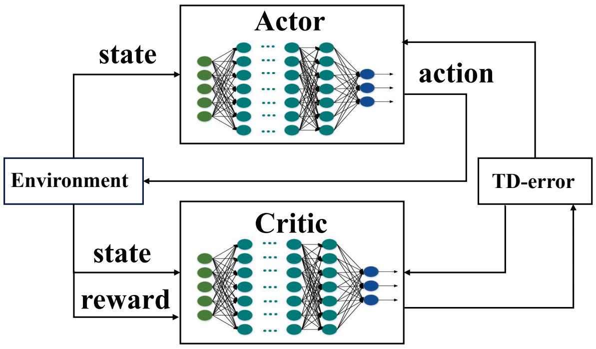


Figure 3. AC framework.

The basic principle of the PPO algorithm is to use policy gradient for updating, and the update objective and rule of policy gradient are displayed as follows:

$$L^{PG}(\theta) = \hat{E}_t\left[\log \pi_\theta(a_t \mid s_t)\hat{A}_t\right] \tag{9}$$

$$\hat{g} = \hat{E}_t\left[\nabla_\theta \log \pi_\theta(a_t \mid s_t)\hat{A}_t\right] \tag{10}$$

where π_θ is a stochastic policy and $\hat{A}_t$ is the estimated value of advantage function. However, the policy gradient algorithm relies on the update step size. Therefore, the trust domain method (TRPO) measures the distribution difference between new and old policies through KL divergence and constrains it. Then, the objective optimization function is modified to

$$\text{maximize}_\theta \quad \hat{E}_t\left[\frac{\pi_\theta(a_t \mid s_t)}{\pi_{\theta old}(a_t \mid s_t)}\hat{A}_t\right] \tag{11}$$

$$\hat{E}[KL[\pi_{\theta\ \text{old}}\,(a_t \mid s_t), \pi_\theta(\cdot \mid s_t)]] \leq \delta \tag{12}$$

where θ_{old} is the parameter of old policy. Directly solving the TRPO constraint problem will bring significant computational complexity. In order to improve computational efficiency, PPO has further simplified the TRPO objective function. By defining the probability ratio of the distribution of new and old strategies $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta\ \text{old}}\,(a_t|s_t)}$, the object function can be written as

$$L(\theta) = \hat{E}_t\left[\frac{\pi_\theta(a_t \mid s_t)}{\pi_{\theta\ \text{old}}\,(a_t \mid s_t)}\hat{A}_t\right] = \hat{E}_t\left[r_t(\theta)\hat{A}_t\right]. \tag{13}$$

In addition, PPO prunes the probability distribution ratio to limit its policy updates:

$$L^{CLIP}(\theta) = \hat{E}_t\left[\min\left(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t\right)\right]. \tag{14}$$

For a complete turn trajectory $T = \{s_0, a_0, r_0; s_1, a_1, r_1; s_2, \ldots, a_{n-1}, r_{n-1}, s_n\}$ and strategy network $\pi_\theta(a_t \mid s_t)$, use state action function $Q^\pi(s_t, a_t)$ as the evaluation value of state action pairs, and use state value function $V^\pi(s_t)$ as the evaluation baseline; then, the following advantage function is defined

$$A^\pi(s_t, a_t) = r + \gamma V^\pi(s_{t+1}) - V^\pi(s_t). \tag{15}$$

The loss function of AN is

$$L(\theta) = -E\left[\min\left(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t\right)\right] \tag{16}$$

And the loss function of CN is

$$L(\omega) = -E\Big([r_t + \gamma V(s_{t+1}) - V(s_t)]^2\Big). \tag{17}$$

In order to prevent overfitting caused by the strategy network being too fast, which leads to premature convergence of the agent and inability to effectively explore environmental information, regularization entropy is added as a penalty term in the final loss function. Therefore, the loss function of the AC PPO algorithm framework can be expressed as follows:

$$\text{Loss} = aL(\theta) + bL(\omega) - c\,\text{ENT}(\pi_\theta) \tag{18}$$

where $a = 1, b = 0.5, c = 0.01$ are the loss proportion coefficients.

Overall, PPO has the following advantages: By modifying the objective function to penalize large deviations between new and old policies, the trust region of the policy update can be ensured, which can achieve more stable training. Unlike TRPO, the second-order derivatives and large matrix inversions are not needed, which is relatively easy to implement. The hyperparameters of PPO are relatively robust, which makes PPO perform outstandingly in various complex tasks.

3.2. PPODNPA

By integrating the MDP of DNP into the AC-based PPO algorithmic framework, we propose the PPODNPA algorithm, which is illustrated in Figure 4. In this approach, the DNP environment is first initialized with a set of states. Based on these initial states, the AN generates a series of actions, such as building lines and installing renewable energy sources in the power grid. Influenced by these actions, the environment states transition to a next set of states according to a probabilistic distribution and an immediate reward is produced. The CN receives the reward and states, evaluates the quality of the actions, and then provides the action advantage function back to AN, thereby further training the AN. After sufficient training, the actions generated by the agent (which dictate the configuration of the grid) are optimized to meet the specified planning objectives. The reward function in the PPODNPA is designed as

$$R_{\text{total}} = -\alpha_1 C_{\text{line}} - \alpha_2 C_{\text{vol}} - \alpha_3 C_{\text{buy}} + \alpha_4 C_{\text{sub}} + R_5 \tag{19}$$

where the parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ serve as the respective weighting factors. These factors can be adaptively adjusted to accommodate various practical requirements. Notably, the parameter R_5 plays a critical role: if no renewable energy resources are installed in the resulting grid configuration, R_5 will impose a penalty. This mechanism ensures a certain level of renewable energy penetration in the final plan.

In the PPODNPA, the distribution network planning environment (DNPE) represents the distribution network prior to the construction of lines and the integration of renewable energy sources. When agent actions are applied to the DNPE, it performs power flow calculations to calculate the node voltages, line currents, and power flows. In addition, the DNPE also calculates line construction costs, distributed renewable energy subsidies, and the costs of purchasing electricity from the up-level grid. These calculated parameters are then transmitted back to the CN, which can guide the AN updates. The procedure of the PPODNPA is displayed in Algorithm 1.

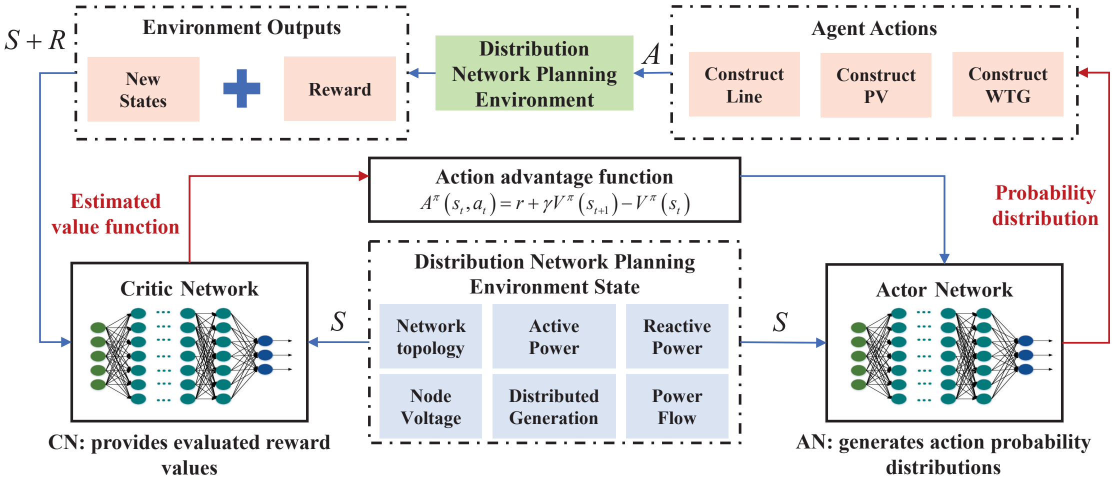


Figure 4. Diagram of PPODNPA.

Algorithm 1 PPODNPA Algorithm Procedure

Algorithm Output: Actor policy network π_θ which can output optimized actions of line construction and renewable energy installation, Critic value network Q_ω which can approximate the state value function $V^\pi(s_t)$.
Randomly initialize the policy network parameters and the value network parameters.
for episode = 1, 2, . . . **do**
  (a) Data Collection:
  Use the old policy $\pi_{\theta_{\text{old}}}$ to interact with the DNP environment.
  In each time step t, sample action $a_t \sim \pi_{\theta_{\text{old}}}(a \mid s_t)$, update the grid/renewable configuration, run power flow calculation, and obtain next state s_{t+1} and immediate reward r_t.
  Continue until reaching the episode end or a terminal condition.
  (b) Advantage Computation:
  For each transition (s_t, a_t, r_t, s_{t+1}), compute the TD advantage function A_t using

$$A_t = r_t + \gamma V^\pi(s_{t+1}) - V^\pi(s_t),$$

  (c) PPO Loss Calculation:
  According to Equation (18), compute the loss function.
  (d) Update Actor:
  Perform gradient descent on the actor parameters θ.
  (e) Update Critic:
  Perform gradient descent on the critic parameters ω.
  After the updates, set $\theta_{\text{old}} \leftarrow \theta$ {Update old policy}
end for

In this paper, the proposed PPO-based distribution network planning algorithm (PPODNPA) includes the following important processes: environment simulation, action selection, power flow computation and reward evaluation, and PPO strategy updating. First, the environment simulation builds and updates the system state space (which includes node voltages, line currents, and network topology) in response to different agent actions (which include constructing new lines and installing renewable energy). Second, the AN gives a probability distribution of possible actions. By sampling the probability distribution, the local optima phenomenon can be avoided. Third, the power flow computations and reward evaluations can measure system constraints and evaluate economic factors, and generate reward signals for AN and CN. Finally, the PPO strategy updates via advantage functions and the clipping mechanism, and the AN and CN are iteratively refined through backpropagation. The main computational burdens emerge from repeated power flow analyses, and parallel or distributed computation can help maintain computational overhead. Repeated episodes allow the PPODNPA to converge to the near-optimal solutions which balance investment cost, purchase price for purchasing electricity from the external grid, and renewable utilization subsidies within acceptable training times. Overall, environment simulation, neural network training, iterative power flow, and multi-round updates collectively illustrate the execution flow and resource requirements scale of the PPODNPA and provide guidance for actual applications.

4. Simulation Results

In this section, the simulation results will be presented. A typical DNP case is selected to verify the effectiveness of the PPODNPA, and the relevant parameters are listed in Table 1. The first column of Table 1 displays the node number, the second column shows each node's coordinates (in kilometers), the third and fourth columns show the active and reactive power of the nodes, respectively (in kW and KVAR). The hidden layer size of the AN and CN was selected as 512. Larger hidden layers can help the neural network learn the complex relationships between state and action. However, larger hidden layers may also lead to higher computational costs and increase the risk of overfitting with limited training data. The learning rate of AN is selected as 0.001 and the learning rate of CN is selected as 0.001. A higher learning rate may cause the policy to oscillate, and a lower rate can slow convergence. Therefore, we select the learning rates for both AN and CN as 0.001 to achieve a balance between convergence speed and stability. The discount factor is chosen as 0.9. The discount factor can balance the short-term and long-term rewards; selecting the discount factor as 0.9 can ensure that cumulative returns are sufficiently reflected, even in relatively shorter training sequences. The clipping parameter is set to 0.1. The clipping parameter is a common hyperparameter in the PPO algorithm that restricts the policy's deviation from the old policy. Selecting the clipping parameter as 0.9 can maintain enough policy improvement while avoiding overly large policy updates. After collecting each trajectory, 20 optimization steps are performed. The number of total episodes is 20,000, and each episode includes 10 steps, which ensures that the information gathered from each sample is effectively leveraged. The DNP lines are selected as LGJ-185 and the construction cost is 1 million Chinese Yuan (CNY)/kilometer. The PV installation capacities are 600 kW, 900 kW, 400 kW, and 700 kW, while the WTG are 500 kW, 600 kW, and 400 kW. Moreover, we select PSO [41], GA [42], ACO [43], and SAA [44] for comparison, which demonstrates the superiority of the proposed algorithm.

The simulation results are presented in Figures 5–12. Figure 5 illustrates the training process curve of the PPODNPA, which includes the instantaneous rewards for each episode and the average rewards (calculated every 200 episodes). From Figure 5, we can conclude that after sufficient training, the agent's rewards approach convergence. Figure 6 depicts the grid topology structure generated under the PPODNPA, in which the red nodes represent the node connected to the main grid, the blue nodes are load nodes, the yellow nodes are PV nodes, and the green nodes are WTG nodes. Additionally, we select PSO, GA, ACO, and SAA for comparison, and the line construction costs (LLC), renewable energy subsidies (RES), and electricity purchasing costs (EPC) from the upper-level grid as comparative indicators. The grid topology structures generated under PSO, GA, ACO, and SAA are shown in Figures 7–10. In addition, we also use DQN as the comparative algorithm; Figure 11 illustrates the training process curve of DQN and Figure 12 depicts the grid topology structure generated under DQN. The relevant indicators are shown in Table 2. As shown in Table 2, the PPODNPA effectively reduces the costs of construction and electricity purchasing while achieving higher subsidies for renewable energy sources. It should be noted that, as the number of loads and power sources on a single branch increases, higher currents can lead to more pronounced voltage drops along that branch. This problem will be one of our future research directions.

Table 1. The parameters of the DNP case.

Node Number	Node Coordinates/km	Active Load/kW	Reactive Load/kvar
1	(1.976, 1.090)	0	0
2	(1.056, 1.026)	439.92	273.78
3	(0.480, 1.304)	421.20	262.08
4	(1.928, 1.798)	318.24	196.56
5	(0.196, 1.076)	430.56	266.76
6	(3.640, 0.474)	374.40	231.66
7	(0.524, 0.914)	402.48	250.38
8	(2.876, 1.808)	383.76	238.68
9	(0.184, 1.602)	570.96	353.34
10	(1.008, 1.586)	589.68	365.04
11	(0.664, 1.822)	421.20	262.08
12	(3.360, 0.904)	477.36	231.66
13	(0.548, 0.430)	580.32	360.36
14	(0.916, 0.182)	374.40	231.66
15	(3.424, 1.192)	458.64	285.48
16	(2.856, 0.182)	336.96	208.26
17	(2.488, 0.272)	402.48	250.38
18	(3.272, 1.738)	439.92	273.78
19	(2.876, 1.560)	402.48	250.38
20	(3.112, 1.394)	683.28	423.54
21	(2.348, 0.112)	449.28	278.46
22	(2.128, 0.334)	486.72	301.86
23	(3.300, 0.474)	318.24	196.56
24	(3.440, 1.490)	393.12	243.36
25	(2.304, 1.556)	276.12	170.82
26	(1.172, 0.354)	159.12	98.28
27	(2.388, 0.506)	313.56	194.22
28	(2.944, 1.196)	480.18	402.32
29	(3.616, 0.718)	159.12	98.28

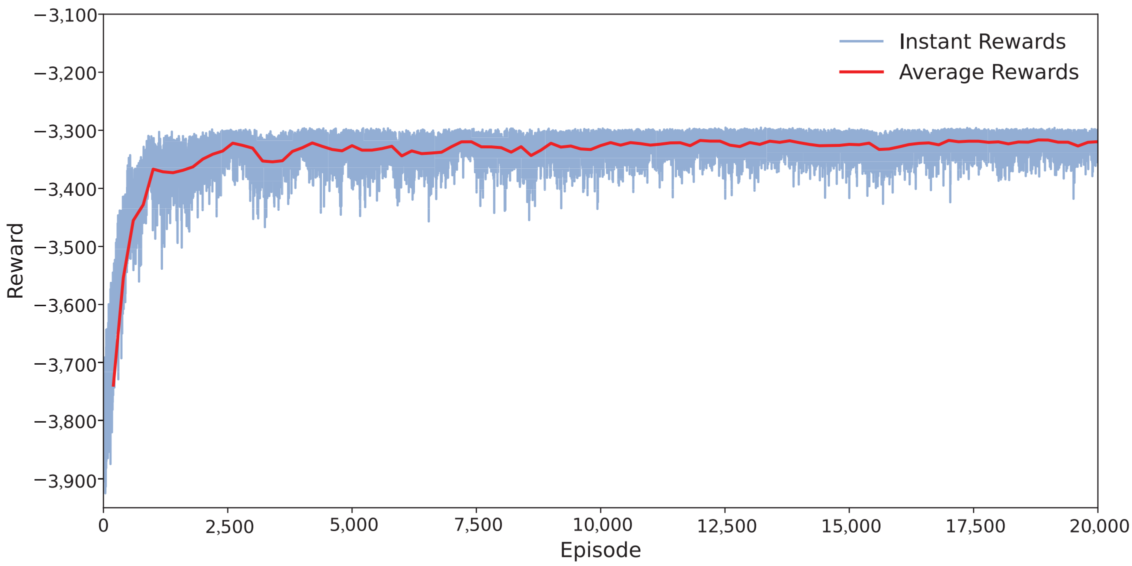


Figure 5. Instant rewards and average rewards of PPODNPA.

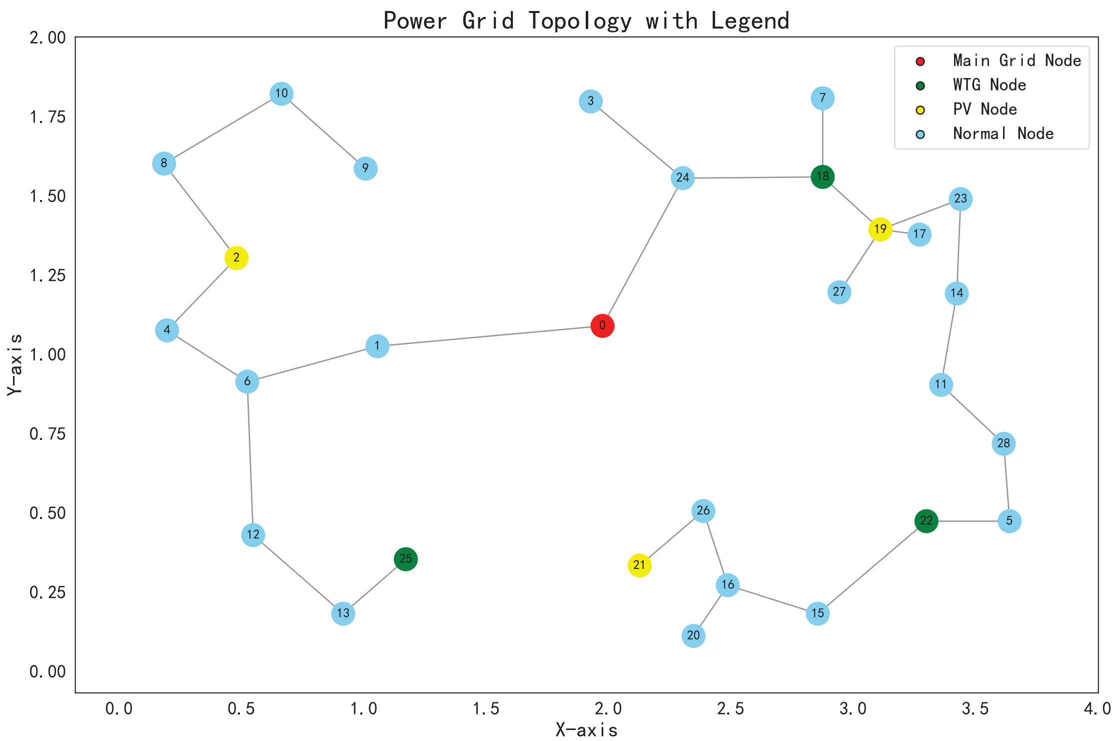


Figure 6. Grid topology diagram under PPODNPA.

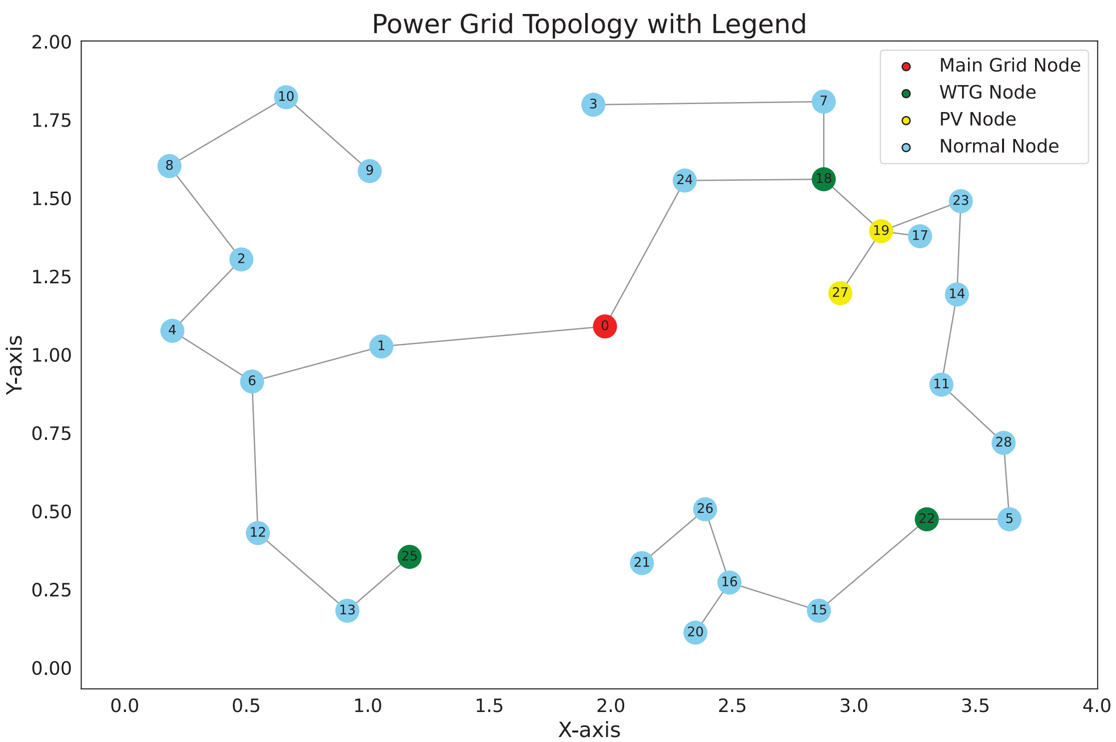


Figure 7. Grid topology diagram under PSO.

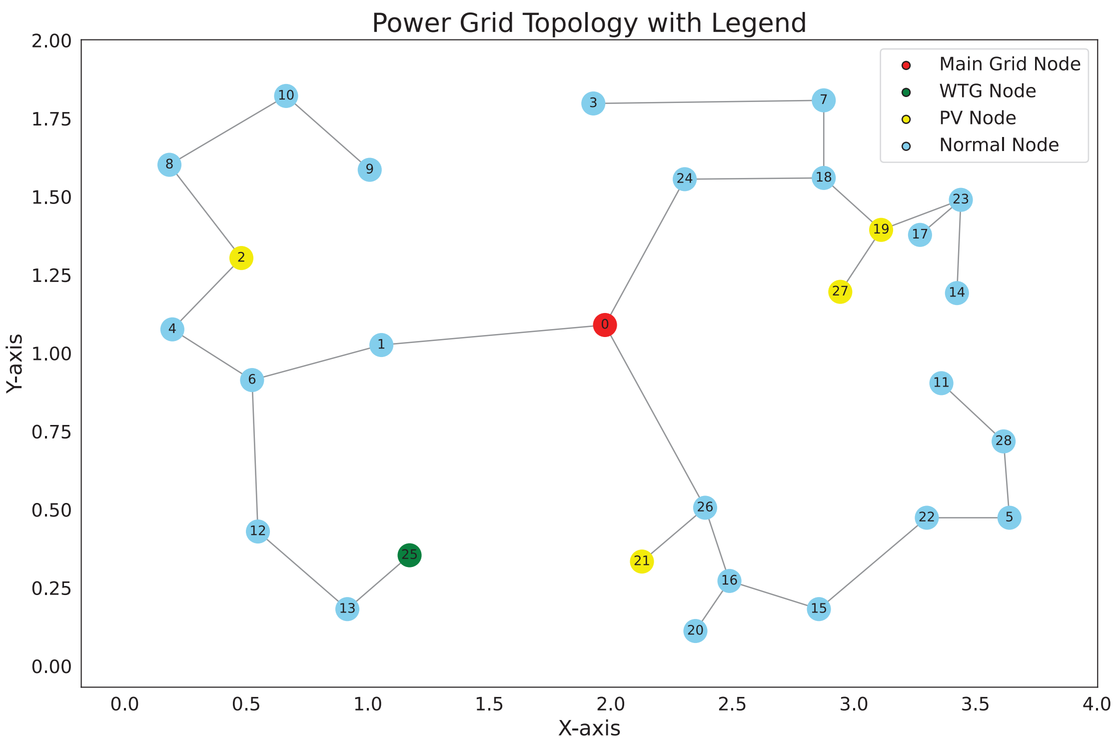


Figure 8. Grid topology diagram under GA.

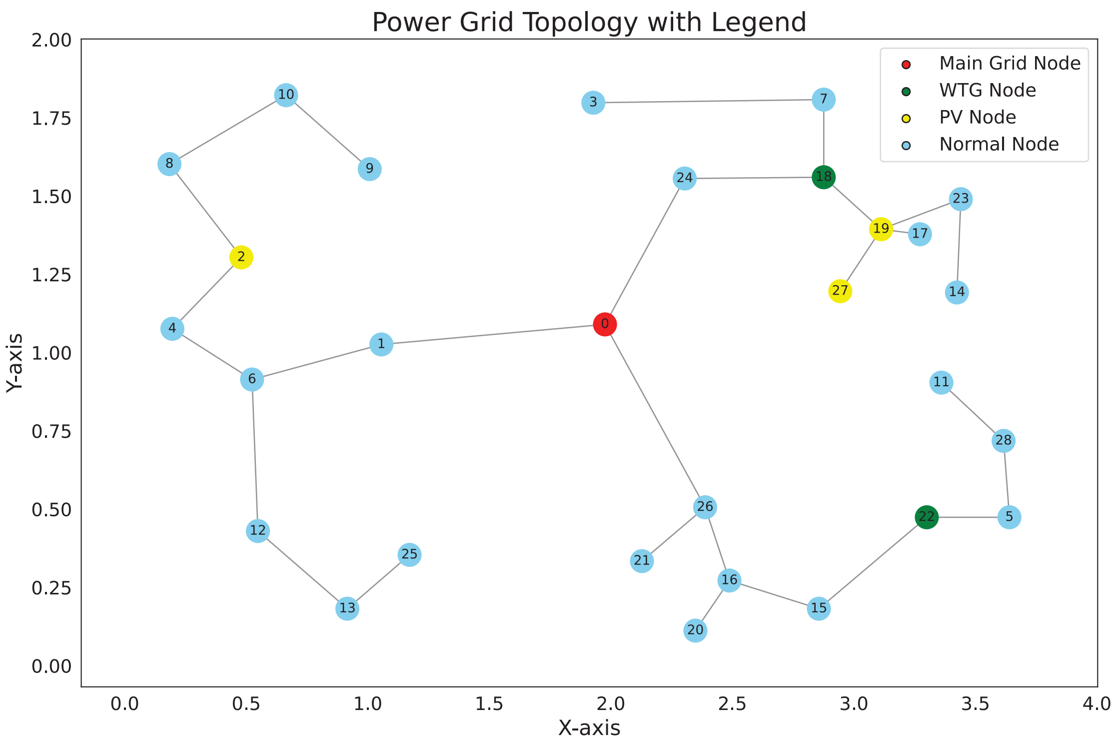


Figure 9. Grid topology diagram under ACO.

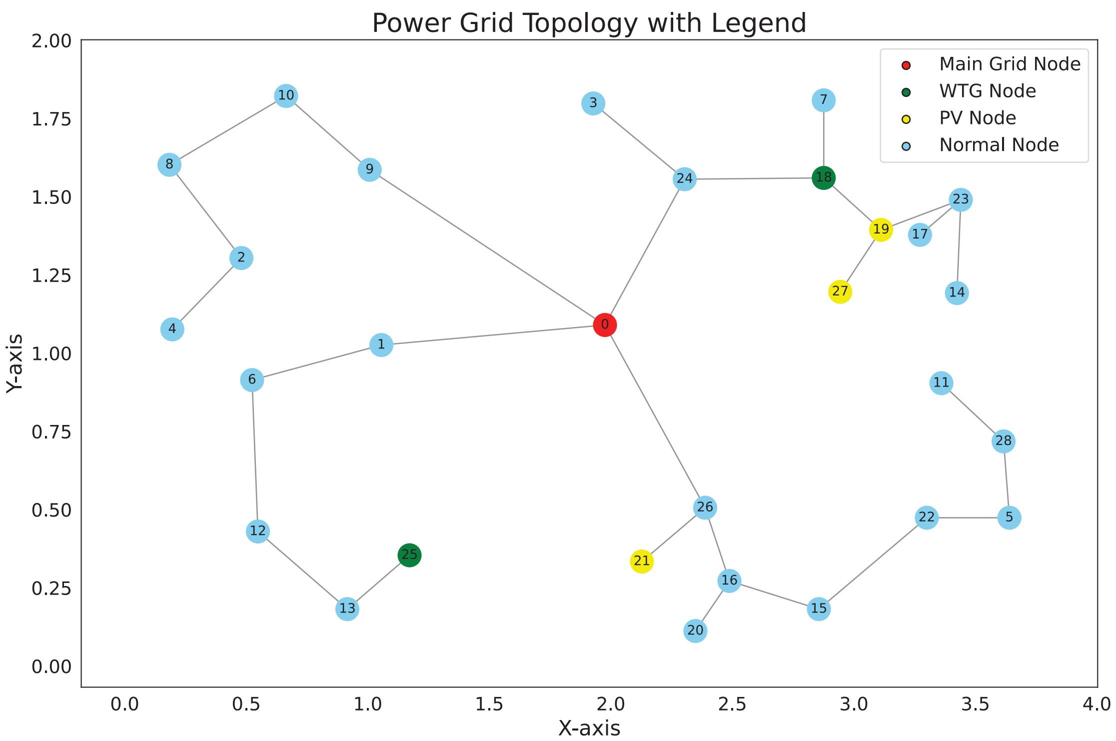


Figure 10. Grid topology diagram under SAA.

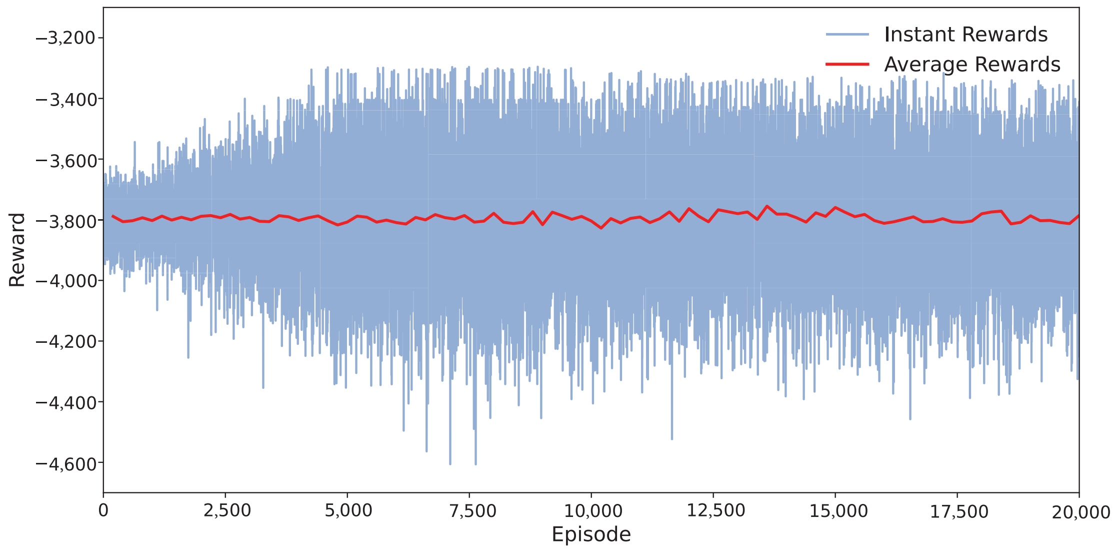


Figure 11. Instant rewards and average rewards of DQN.

Table 2. Comparative indicators.

Method	LCC/million CNY	VD	EPC/CNY	RES/CNY
PPODNPA	10.87	2.0744	78,857	16,320
PSO	11.37	2.3352	83,298	14,880
GA	11.83	0.5897	85,184	14,400
ACO	11.79	0.5105	77,620	15,840
SAA	12.05	0.3003	81,699	13,920
DQN	11.59	2.6364	96,241	9120

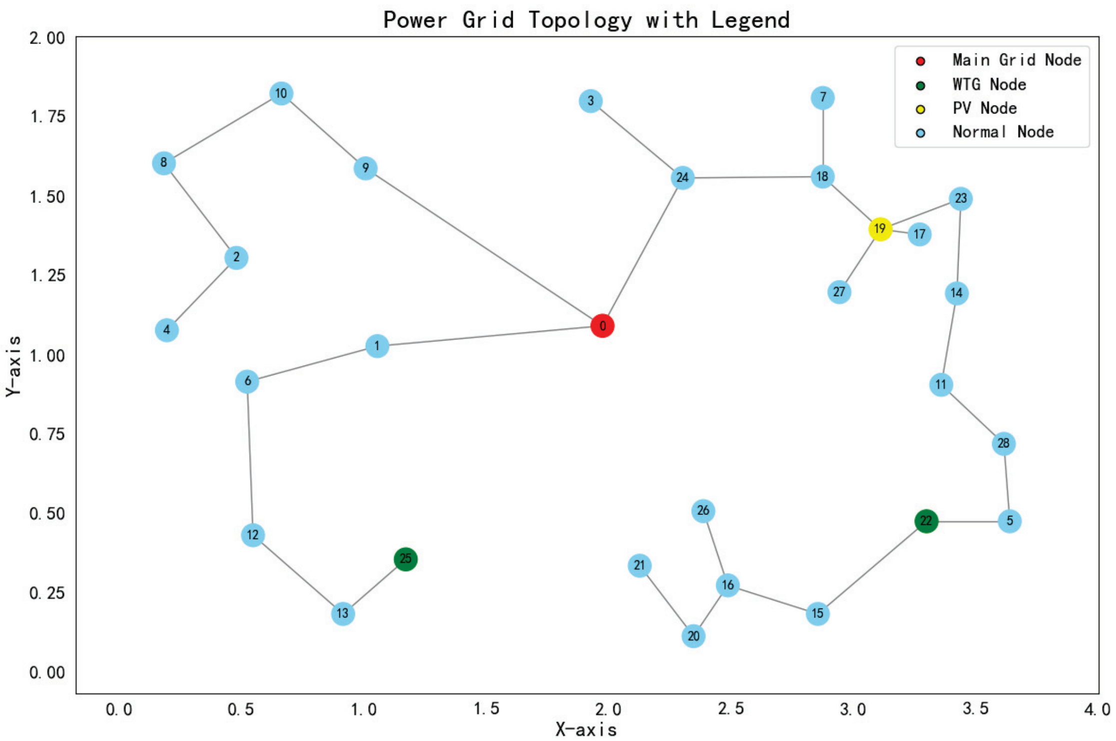


Figure 12. Grid topology diagram under DQN.

5. Conclusions

In this paper, a PPODNPA has been proposed and the MDP for the task of DNP has been constructed. The objective function and reward function have been formulated, and the PPO algorithm has been employed to train the agent. Furthermore, PSO, GA, ASO, and SAA algorithms have been utilized to validate the effectiveness of the proposed method. In future work, we will investigate the application of DRL in high-voltage DNP, and more factors such as voltage deviation and station location corridors will be considered. Moreover, we will consider further improving DRL algorithms such as SAC and DDPG to further optimize planning performance.

Author Contributions: Conceptualization, L.M.; methodology, L.M., S.J. and Y.S.; software, C.S. and Y.S.; validation, S.J.; formal analysis, K.W. and J.L.; investigation, C.S. and J.L.; resources, J.L.; data curation, C.S.; writing—original draft preparation, C.S. and S.J.; writing—review and editing, K.W.; visualization, K.W.; supervision, Y.S.; project administration, L.M.; funding acquisition, L.M., S.J., Y.S. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Science and Technology Project of the State Grid Corporation of China (5400-202456175A-1-1-ZN).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Liang Ma, Jinshan Luo, Shigong Jiang and Yi Song were employed by the company State Grid Economic and Technological Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Wang, P.; Yi, J.; Zangiabadi, M.; Lyons, P.; Taylor, P. Evaluation of voltage control approaches for future smart distribution networks. *Energies* **2017**, *10*, 1138. [CrossRef]
2. Liao, H. Review on distribution network optimization under uncertainty. *Energies* **2019**, *12*, 3369. [CrossRef]

3. Shen, X.; Shahidehpour M.; Zhu, S.; Han, Y.; Zheng, J. Multi-stage planning of active distribution networks considering the co-optimization of operation strategies. *IEEE Trans. Smart Grid* **2018**, *9*, 1425–1433. [CrossRef]
4. Zhang, L.; Liu, Y.; Liang, D.; Kou, P.; Wang, Y.; Gao, Y. Local and remote cooperative control of hybrid distribution transformers integrating photovoltaics in active distribution networks. *IEEE Trans. Sustain. Energy* **2022**, *13*, 2012–2026. [CrossRef]
5. Ding, T.; Liu, S.; Yuan, W.; Bie, Z.; Zeng, B. A two-stage robust reactive power optimization considering uncertain wind power integration in active distribution networks. *IEEE Trans. Sustain. Energy* **2015**, *7*, 301–311. [CrossRef]
6. Pucuhuayla, F.; Correa, C.; Huatuco, D.; Rodriguez, Y. Optimal reconfiguration of electrical distribution networks using the improved simulated annealing algorithm with hybrid cooling (ISA-HC). *Energies* **2024**, *17*, 4477. [CrossRef]
7. Cadenovic, R.; Jakus, D.; Sarajcev, P.; Vasilj, J. Optimal distribution network reconfiguration through integration of cycle-break and genetic algorithms. *Energies* **2018**, *11*, 1278. [CrossRef]
8. Valderrama, D.; Alonso, J.; de Mora, C.; Robba, M. Scenario generation based on ant colony optimization for modelling stochastic variables in power systems. *Energies* **2024**, *17*, 5293. [CrossRef]
9. Nahman, J.; Peric, D. Optimal planning of radial distribution networks by simulated annealing technique. *IEEE Trans. Power Syst.* **2008**, *23*, 790–795. [CrossRef]
10. Jo, S.; Oh, J.; Lee, J.; Oh, S.; Moon, H.; Zhang, C.; Gadh, R.; Yoon, Y. Hybrid genetic algorithm with k-nearest neighbors for radial distribution network reconfiguration. *IEEE Trans. Smart Grid* **2024**, *15*, 2614–2624. [CrossRef]
11. Falaghi, H.; Haghifam, M.; Singh, C. Ant colony optimization-based method for placement of sectionalizing switches in distribution networks using a fuzzy multiobjective approach. *IEEE Trans. Power Del.* **2009**, *24*, 268–276. [CrossRef]
12. Harsh, P.; Das, D. A simple and fast heuristic approach for the reconfiguration of radial distribution networks. *IEEE Trans. Power Syst.* **2023**, *38*, 2939–2942. [CrossRef]
13. Mazhari, S.; Monsef, H.; Romero, R. A hybrid heuristic and evolutionary algorithm for distribution substation planning. *IEEE Syst. J.* **2015**, *9*, 1396–1408. [CrossRef]
14. Ahmadian, A.; Elkamel, A.; Mazouz, A. An improved hybrid particle swarm optimization and tabu search algorithm for expansion planning of large dimension electric distribution network. *Energies* **2019**, *12*, 3052. [CrossRef]
15. Lu, D.; Li, W.; Zhang, L.; Fu, Q.; Jiao, Q.; Wang, K. Multi-objective optimization and reconstruction of distribution networks with distributed power sources based on an improved BPSO algorithm. *Energies* **2024**, *17*, 4877. [CrossRef]
16. Martinez, M.; Mateo, C.; Gomez, T.; Alonso, B.; Frias, P. A hybrid particle swarm optimization approach for explicit flexibility procurement in distribution network planning. *Int. J. Electr. Power Energy Syst.* **2024**, *161*, 110215. [CrossRef]
17. Moradi, M.; Abedini, M. A combination of genetic algorithm and particle swarm optimization for optimal DG location and sizing in distribution systems. *Int. J. Electr. Power Energy Syst.* **2012**, *34*, 66–74. [CrossRef]
18. Yang, M.; Li, J.; Li, J.; Yuan, X.; Xu, J. Reconfiguration strategy for DC distribution network fault recovery based on hybrid particle swarm optimization. *Energies* **2021**, *14*, 7145. [CrossRef]
19. Huang, C.; Bu, S.; Chen, W.; Wang, H.; Zhang, Y. Deep reinforcement learning-assisted federated learning for robust short-term load forecasting in electricity wholesale markets. *IEEE Trans. Netw. Sci. Eng.* **2024**, *11*, 5073–5086. [CrossRef]
20. Hossain, R.; Huang, Q.; Huang, R. Graph convolutional network-based topology embedded deep reinforcement learning for voltage stability control. *IEEE Trans. Power Syst.* **2021**, *36*, 4848–4851. [CrossRef]
21. Chen, Y.; Zhu, J.; Liu, Y.; Zhang, L.; Zhou, J. Distributed hierarchical deep reinforcement learning for large-scale grid emergency control. *IEEE Trans. Power Syst.* **2023**, *39*, 4446–4458. [CrossRef]
22. Al-Ani, O.; Das, S.; Reinforcement learning: Theory and applications in HEMS. *Energies* **2022**, *15*, 6392. [CrossRef]
23. Toubeau, J.; Zad, B.; Hupez, M.; De Greve, Z.; Vallee, F. Deep reinforcement learning-based voltage control to deal with model uncertainties in distribution networks. *Energies* **2020**, *13*, 3928. [CrossRef]
24. Yan, Z.; Xu, Y. A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system. *IEEE Trans. Power Syst.* **2020**, *35*, 4599–4608. [CrossRef]
25. Li, J.; Chen, S.; Wang, X.; Pu, T. Load shedding control strategy in power grid emergency state based on deep reinforcement learning. *CSEE J. Power Energy* **2021**, *8*, 1175–1182.
26. Jin, J.; Xu, Y. Optimal policy characterization enhanced actor-critic approach for electric vehicle charging scheduling in a power distribution network. *IEEE Trans. Smart Grid* **2021**, *12*, 1416–1428. [CrossRef]
27. Wang, R.; Bi, X.; Bu, S. Real-time coordination of dynamic network reconfiguration and volt-VAR control in active distribution network: A graph-aware deep reinforcement learning approach. *IEEE Trans. Smart Grid* **2024**, *15*, 3288–3302. [CrossRef]
28. Wang, W.; Yu, N.; Gao, Y.; Shi, J. Safe off-policy deep reinforcement learning algorithm for volt-VAR control in power distribution systems. *IEEE Trans. Smart Grid* **2020**, *11*, 3008–3018. [CrossRef]
29. Ruan, J.; Wu, C.; Liang, Z.; Liu, K.; Li, B.; Li, W.; Li, T. The application of machine learning-based energy management strategy in a multi-mode plug-in hybrid electric vehicle, part II: Deep deterministic policy gradient algorithm design for electric mode. *Energy* **2023**, *269*, 126972. [CrossRef]

30. Zhang, J.; Li, Y.; Wu, Z.; Rong, C.; Wang, T.; Zhang, Z.; Zhou, S. Deep-reinforcement-learning-based two-timescale voltage control for distribution systems. *Energies* **2021**, *14*, 3540. [CrossRef]
31. Zhu, Z.; Gao, X.; Bu, S.; Chan, K.; Zhou, B.; Xia, S. Cooperative dispatch of renewable-penetrated microgrids alliances using risk-sensitive reinforcement learning. *IEEE Trans. Sustain. Energy* **2024**, *15*, 2194–2208. [CrossRef]
32. Qi, T.; Ye, C.; Zhao, Y.; Li, L.; Ding, Y. Deep reinforcement learning based charging scheduling for household electric vehicles in active distribution network. *J. Mod. Power Syst. Clean Energy* **2023**, *11*, 1890–1901. [CrossRef]
33. Sun, X.; Han, S.; Wang, Y.; Shi, Y.; Liao, J.; Zheng, Z.; Wang, X.; Shi, P. Proximal policy optimization-based power grid structure optimization for reliable splitting. *Energies* **2024**, *17*, 834. [CrossRef]
34. Fang, X.; Hong, P.; He, S.; Zhang, Y.; Tan, D. Multi-layer energy management and strategy learning for microgrids: A proximal policy optimization approach. *Energies* **2024**, *17*, 3990. [CrossRef]
35. Meng, J.; Yang, F.; Peng, J.; Gao, F. A proximal policy optimization based control framework for flexible battery energy storage system. *IEEE Trans. Energy Convers.* **2024**, *39*, 1183–1191. [CrossRef]
36. Wang, Y.; Qiu, D.; Wang, Y.; Sun, M.; Strbac, G. Graph learning-based voltage regulation in distribution networks with multi-microgrids. *IEEE Trans. Power Syst.* **2024**, *39*, 1881–1895. [CrossRef]
37. Zhao, Y.; Liu, J.; Liu, X.; Yuan, K.; Ding, T. Enhancing the tolerance of voltage regulation to cyber contingencies via graph-based deep reinforcement learning. *IEEE Trans. Power Syst.* **2024**, *39*, 4661–4673. [CrossRef]
38. Nie, Y.; Liu, J.; Liu, X.; Zhao, Y.; Ren, K.; Chen, C. Asynchronous multi-agent reinforcement learning-based framework for bi-level noncooperative game-theoretic demand response. *IEEE Trans. Smart Grid* **2024**, *15*, 5622–5637. [CrossRef]
39. Tahir, I.; Nasir, A.; Algethami, A. Optimal control policergy management of a commercial bank. *Energies* **2022**, *15*, 2112. [CrossRef]
40. Liang, Z.; Chung, C.; Zhang, W.; Wang, Q.; Lin, W.; Wang, C. Enabling high-efficiency economic dispatch of hybrid AC/DC networked microgrids: Steady-state convex bi-directional converter models. *IEEE Trans. Smart Grid* **2025**, *16*, 45–61. [CrossRef]
41. Vlachogiannis, J.; Lee, K. A comparative study on particle swarm optimization for optimal steady-state performance of power systems. *IEEE Trans. Power Syst.* **2006**, *21*, 1718–1728. [CrossRef]
42. Mohamed, A.; Berzoy, A.; Mohammed, O. Design and hardware implementation of FL-MPPT control of PV systems based on GA and small-signal analysis. *IEEE Trans. Sustain. Energy* **2017**, *8*, 279–290. [CrossRef]
43. Priyadarshi, N.; Ramachandaramurthy, V.; Padmanaban, S.; Azam, F. An ant colony optimized MPPT for standalone hybrid PV-Wind power system with single cuk converter. *Energies* **2019**, *12*, 167. [CrossRef]
44. Ye, H. Surrogate affine approximation based co-optimization of transactive flexibility, uncertainty, and energy. *IEEE Trans. Power Syst.* **2018**, *33*, 4084–4096. [CrossRef]

energies

MDPI

Article

Multi-Time Scale Scenario Generation for Source–Load Modeling Through Temporal Generative Adversarial Networks

Liang Ma [1,*], Shigong Jiang [1], Yi Song [1], Chenyi Si [2] and Xiaohan Li [2]

[1] State Grid Economic and Technological Research Institute Co., Ltd., Beijing 102209, China; shigong_jiang0106@163.com (S.J.); yi_song0112@163.com (Y.S.)
[2] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; sichenyi2021@163.com (C.S.); lixiaohan20011129@163.com (X.L.)
* Correspondence: ma_liang276@163.com

Abstract: With the large-scale integration of distributed power sources, distribution network planning is undergoing significant transformations. To further enhance the efficiency and practicality of distribution network planning, it is essential to model the uncertainties in source–load dynamic scenarios. However, traditional scenario generation methods struggle with high-dimensional variables and complex spatiotemporal characteristics, posing severe challenges for distribution network planning. To address these issues, this paper proposes a multi-time scale source–load scenario generation method based on temporal convolutional networks and multi-head attention mechanisms within a temporal generative adversarial network framework. This algorithm not only enhances the richness and robustness of source–load scenarios in distribution networks but also serves as a valuable reference for medium-long-term analysis and planning. Finally, the results present a set of daily, weekly, and monthly multi-time scale source–load scenarios, and multiple evaluation indicators are utilized to evaluate the quality of the generated scenarios; the accuracy of the generated scenarios is increased by about 2%.

Keywords: source–load scenario generation; temporal generative adversarial network; multi-time scale; scenario evaluation

1. Introduction

In the current landscape, distributed energy resources dominated by wind and solar power along with new energy storage systems and charging facilities are advancing at an unprecedented pace. By 2025, distribution networks will need to support the integration of approximately 500 gigawatts of distributed energy resources and around 12 million charging stations. In this context, traditional single-source–load scenarios are increasingly inadequate for meeting the evolving planning requirements of distribution networks as they are unable to effectively manage the uncertainties introduced by operational simulations and planning schedules [1,2]. Therefore, it is crucial to solve the randomness and fluctuation of source–load in modern distribution networks, accurately simulate the change trend of source–load power, and ensure the safe operation of the power system [3].

Scenario generation is a powerful method for characterizing source–load uncertainties and can be used to construct random source–load scenarios that reflect spatiotemporal characteristics [4,5]. This enables the evaluation of distribution network performance under different scenarios, enhancing the adaptability and robustness of distribution network. From a time perspective, random scenarios can be classified into ultra-short-term [6], short-term [7], and medium-long-term scenarios [8,9]. As a means of assessing source–load

 https://doi.org/10.3390/en18061462

uncertainties across multi-time scales, source–load scenarios provide a critical foundation for generation planning, unit maintenance planning, and distribution network planning. However, the generation of source–load scenarios across multi-time scales is challenging, mainly due to the increasing data dimensions, constraints imposed by meteorological factors, and the frequent interactions source–load [10]. For example, in Ref. [11], complex factors such as high variable dimensions and random spatial and temporal correlations affect the integrated multi-energy complementary planning and long-term scheduling of water, wind, and solar energy. In Ref. [12], accurate modeling of multiple wind farms is affected by multiple spatiotemporal characteristics. In Ref. [13], renewable energy load and power generation level predictions are also affected by meteorological factors as well as spatial factors, which leads to a serious decline in prediction and modeling accuracy. These factors lead to low source–load interaction modeling accuracy and poor quality of generated source–load scenarios, which cannot provide planners with detailed and informative reference data.

Scenario analysis mainly consists of scenario generation and scenario reduction [14,15]. For example, in the work of [16], a low-dimensional scene generation method based on typical days of photovoltaic and wind power was used for short-time scale modeling. In Ref. [17], based on the Monte Carlo method, the joint scenario generation of photovoltaic and wind power was realized. In Refs. [18,19], scenario generation methods were used to assess the complementarity of future wind and solar photovoltaic hybrid energy resources. However, existing multi-time scale scenario generation methods primarily rely on statistical approaches such as probabilistic models, mean models, and quantile regression to generate sets of typical future source–load temporal scenarios. Although these methods effectively and intuitively describe the uncertainty of multi-time scale source–load power, they require a large amount of prior knowledge, and the quality and richness of the generated scenes cannot be guaranteed. If prior knowledge is lacking, it is more challenging to generate complex and multi-time scale source–load scenarios. On the other hand, scene reduction is a common means to remove noisy information from a scene set, minimizing the computational effort of multi-time scale analysis, while retaining the critical information from the generated scenario set. Common methods include backward reduction [20], scenario tree methods, clustering [21,22], and optimization approaches [23,24]. Therefore, selecting an appropriate scenario reduction algorithm to construct a flexible scenario reduction model is of great significance for analyzing source–load power information with high value.

Due to their powerful data representation capabilities, deep learning algorithms not only show excellent performance in processing high-dimensional nonlinear historical data, but also reduce the importance of prior knowledge. They provide a powerful means to effectively capture source–load uncertainty and provide diversity, richness, and accuracy for multi-time scale scenario generation of distribution networks. Currently, data-driven approaches are the mainstream method for scene generation, and generative adversarial networks (GANs) are particularly popular as one type of data-driven method. GANs can generate new samples that are very close to the real data distribution through the adversarial training of two neural networks [25,26]. However, traditional GANs have shortcomings such as unstable training and mode collapse. Therefore, it is necessary to use Wasserstein distance and a clipping technique to increase the stability of GANs. For example, in Ref. [27], a scenario generation method based on WGAN-GP was used for wind and solar power scenario generation. Likewise, in Ref. [28], WGAN-GP was used to achieve the generation of single and multiple wind power scenarios. In order to further improve the quality and accuracy of scene generation, experts and scholars added condition information to a GAN [29,30]. For instance, in Ref. [31], the author clustered the source and load information and used it as a condition to guide a generation of energy storage planning scenarios. The study [32] applied conditional generative adversarial networks (CGANs) to scenario generation for wind and solar power output using monthly information as labels. The literature [33] used wind and solar power forecasting data as conditions and employed CGAN to generate predictive scenario sets, providing accuracy for future wind and solar power scenario simulations. References [34,35] addressed the multi-time scale issue of source–load scenarios by proposing a monthly source–load scenario generation method based on a progressively growing CGAN. The study [36] clustered historical meteorological data and used a denoising variational autoencoder to generate multi-source–load scenarios, but it did not consider multi-time scale modeling. Although these methods have made significant progress in generating source–load scenarios and effectively fitting source–load data characteristics, most of them do not construct source–load scenario models at multi-time scales, thus leading to limitations in quantifying source–load uncertainty and the multi-time correlation characteristics in the source–load dynamic scene cannot be fully captured.

In response to these challenges, this paper comprehensively considers source–load uncertainty across multi-time scales, and proposes a source–load scenario generation method based on the temporal generative adversarial network (TGAN) by introducing multi-head attention mechanisms (MHAM)s and temporal convolutional networks (TCNs). The main contributions of this paper are as follows:

(1) Clustering techniques are used to extract representative daily source–load power states, and Markov chains are established to model source–load state transitions;
(2) A spatiotemporal feature extraction unit for TGAN is constructed by incorporating MHAM and TCN, using clustering information as labels to guide the generation of source–load power time series;
(3) A comprehensive evaluation index system is established based on the generated source–load scenario sets to assess the high fidelity of the scenarios.

The remaining sections of this article are organized as follows. In Section 2, we constructed a source–load time correlation model. Section 3 is mainly focused on TGAN source–load scenario generation, including MHAM, TCN, and CGAN applications. Sections 4 and 5 mainly include model training and scenario quality evaluation. This article concludes with Section 6. It should be noted that this paper uses the source–load data set of power distribution systems in a county and district in Shandong, and code support is provided in Appendix A.

2. Source–Load Time Correlation Modeling

Compared to the day-ahead source–load power, the variables in long-term source–load power scenario simulations increase several times. To reduce the modeling dimensionality and improve the accuracy of simulating the probabilistic characteristics and spatiotemporal correlations of source–load power, this section clusters historical source–load data and establishes a source–load power scenario set across multiple temporal and spatial scales.

Time Correlation Modeling Based on Markov Chains

To reveal the transition patterns of daily source–load power states and improve the simulation accuracy over longer time scales, the source–load historical data are first divided by month, and the K-means clustering technique is used to select typical daily source–load power states. Then, a Markov chain model is established based on these source–load power states, where Markov chain represents the transition process from one source–load state to another source–load state. The probability of this transition is only related to the current source–load state and is not affected by the previous event sequence.

Let $X = \{X_1, X_2, \ldots, X_n\}$ represent the source–load daily power sample set for n_p months, where X_i denotes the power matrix on day i, with dimensions $n_p \times 24$, and n is the sample size. In order to capture the different temporal characteristics of source–load interactions and thus generate a more accurate time-correlation model, the following steps for model construction are outlined:

(1) In light of the historical data of different source–load every month, $X \times \Theta$ state intervals are supported by the sample set X and the clustering results.
(2) According to the source–load daily power states and their frequencies obtained through clustering, the Markov chain state transition probability matrix $\mathcal{P}_t$ is calculated as described below:

$$\mathcal{P}_t = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,\Theta} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,\Theta} \\ \vdots & \vdots & \ddots & \vdots \\ P_{\Theta,1} & P_{\Theta,2} & \cdots & P_{\Theta,\Theta} \end{bmatrix} \tag{1}$$

where the element $P_{\theta,l}$ represents the probability of the source–load daily power state transitioning from state θ to state l $(\theta, l = 1, 2, \ldots, \Theta)$, and it is denoted as an element

$$P_{\theta,l} = \frac{f_{\theta,l}}{\sum_{\theta \in X} f_{\theta,l}} \tag{2}$$

where $f_{\theta,l}$ represents the frequency of the source–load power transitioning from state θ to state l.

(3) In light of the statistically derived source–load state transition frequency matrix, the cumulative state transition probability matrix $\mathcal{P}_c$ can be further obtained

$$\mathcal{P}_c = \begin{bmatrix} Q_{1,1} & Q_{1,2} & \cdots & Q_{1,\Theta} \\ Q_{2,1} & Q_{2,2} & \cdots & Q_{2,\Theta} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{\Theta,1} & Q_{\Theta,2} & \cdots & Q_{\Theta,\Theta} \end{bmatrix} \tag{3}$$

where $Q_{\theta,l} = \sum_{m=1}^{l} Q_{\theta,m}$. The cumulative state transition probability matrix summarizes the probability of transitioning from the current source–load state to another source–load state over time in a Markov process. Each entry in the matrix represents the total probability of transition from the initial source–load state to the final source–load state. It provides a comprehensive view of the likelihood that source–load information will evolve from one state to another over a given period of time.

Based on the cumulative state transition probability matrix obtained in step (3), the Markov chain Monte Carlo (MCMC) method is used to randomly generate an initial scenario set that contains multiple long-term source–load power state transition processes. The state transition probability matrix and the cumulative state transition probability matrix are shown in Figures 1 and 2 , respectively.

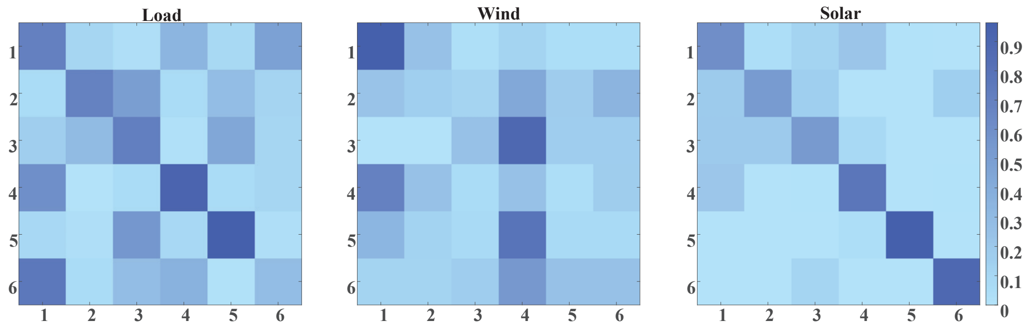


Figure 1. The state transition probability matrix $\mathcal{P}_t$ for the six representative source–load typical days extracted by clustering for each month.

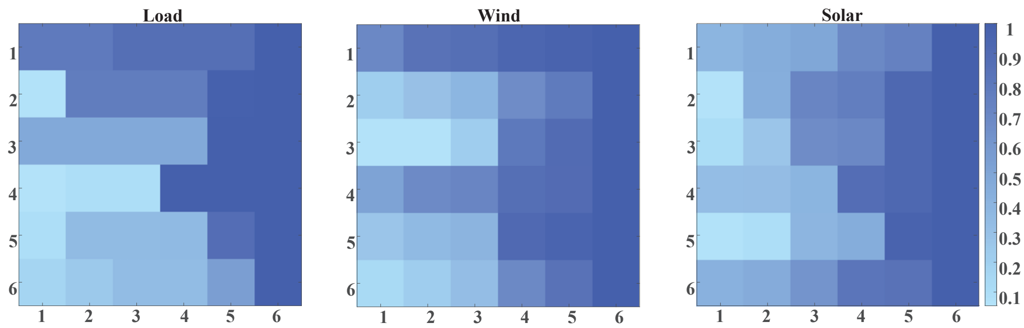


Figure 2. The cumulative state transition probability matrix $\mathcal{P}_c$ for the six representative source–load typical days extracted by clustering for each month.

3. Source–Load Scenario Generation Based on TGAN

To address the limitations of traditional GAN in extracting temporal features of source–load power, this paper proposes an improved CGAN model, which introduces an MHAM and TCN into the generator and discriminator to form the TGAN framework. This model generates intra-day power curves under the supervision of source–load daily state labels. MHAMs enhance the spatial feature extraction capability of TGAN for source–load data, while the TCN fully captures the temporal characteristics and correlations of source–load data, thereby improving the accuracy of source–load scenario generation.

3.1. Source–Load Spatial Feature Extraction Unit Based on MHAM

The MHAM can be used to evaluate the importance of different features of the historical source–load power matrix in scenario generation. Compared to single-head attention mechanisms, the primary advantage of the MHAM lies in its ability to process the relationships between source–load data, capturing diverse features of the data, thereby offering a more robust capacity for extracting spatial features of source–load power. In this paper, the weight calculation formula for the *i*-th head of the MHAM is given as follows (see Figure 3):

$$\alpha_i = \text{softmax}\left(\frac{x_a W_i^Q (x_b W_i^K)^T}{\sqrt{d_k}}\right) \tag{4}$$

where x_a and x_b can be seen as input representations from different sources. These input matrices are projected into different spaces using the linear projection matrices W_i^Q and W_i^K. The softmax function is then applied to the scaled results to produce a set of weights α_i.

The weights α_i are multiplied with the input matrix to obtain the output matrix of a single attention mechanism, which calculates the similarity between different source–load features

$$A_i(x_a) = \alpha_i\left(x_b W_i^V\right) \quad (5)$$

Through multiple output matrices, the different in source–load information can be identified, thereby proposing the spatial correlation features of source–load data.

$$A(x_a) = \text{Concat}\left(A_1(x_a), \ldots, A_h(x_a)\right) W^o \quad (6)$$

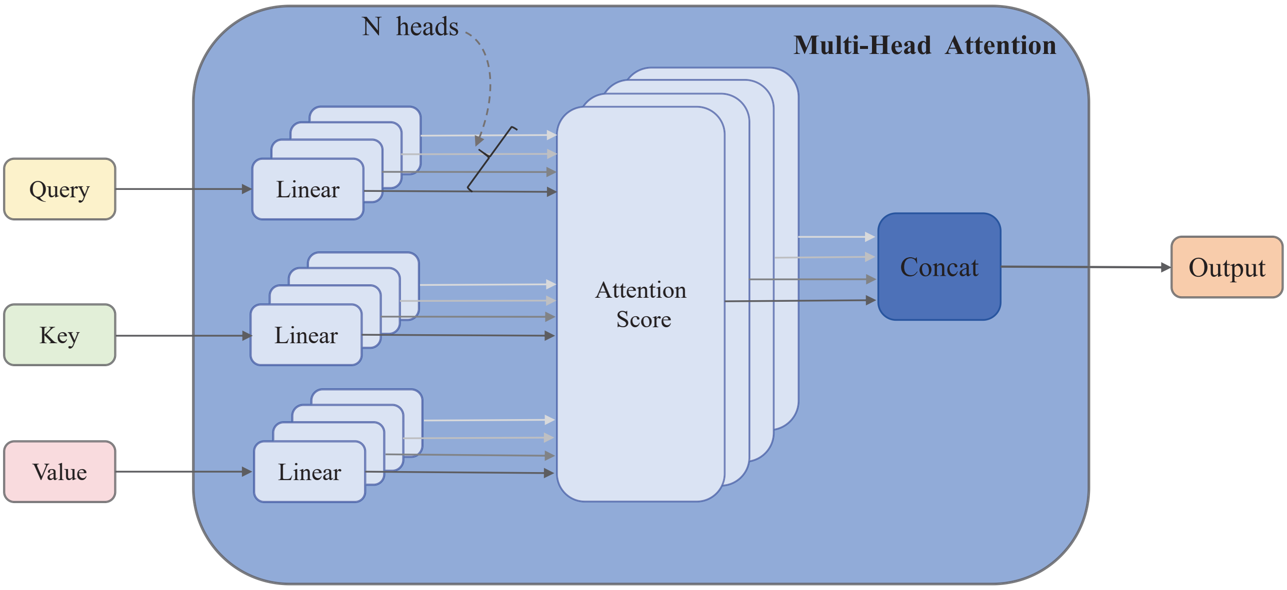


Figure 3. Workflow diagram of MHAM. The input source and load data X undergo multiple linear transformations (using W^Q, W^K, and W^V matrices for each head) to generate the query Q, key K, and value V matrices. Each attention head independently computes the attention output A_i. The results of each head are concatenated together and then projected by the matrix W_O into the final output A.

3.2. Source–Load Temporal Feature Extraction Unit Based on TCN

A TCN is a neural network that processes sequential data through convolution operations. Its basic working principle is to capture dependencies between different time steps in a time series through causal convolution and dilated convolution. Compared to recurrent neural networks that simulate source–load temporal power, TCN is more adept at handling high-dimensional source–load feature information and demonstrates significant advantages in training speed. In this paper, TCN is used as the main structure to build the temporal feature extraction unit called TGAN.

According to the source–load temporal data, the value y_r at time t is determined solely by the historical information at time $0 - t$, that is

$$y_t = g_{\text{causal}}(x_0, x_1, \ldots, x_t), t = 0, 1, \ldots, T \quad (7)$$

where x_t represents the input power at time t, g_{causal} refers to causal convolution, and T is the total number of time slices.

When processing long-term source–load historical data sequences, the network depth or convolution size may increase sharply, leading to issues such as gradient vanishing and inefficiency during training. The dilated convolution operation enables the capture of a wider range of information, allowing the TCN to extract essential information without needing to delve deeply into the entire source–load history (see Figure 4). The basic principle of dilated convolution is

$$D_c(x) = (x *_\delta f)(x) = \sum_{i_f=0}^{k-1} f(i_f) \cdot x_{x-\delta \cdot i_f} \tag{8}$$

where $D_c(x)$ represents the result of the filter performing a dilated convolution operation on the element x in the historical power vector x, $x*_\delta$ refers to the dilated convolution, $f(i_f)$ represents the i_f filter, δ is the dilation rate, and k is the filter size.

In the convolution operation above, a residual connection is further incorporated to form a residual module. The output of the residual module, which combines historical information and convolutional information, enhances the algorithm's ability to express power characteristics in source–load scenario generation, thereby obtaining the temporal feature extraction unit.

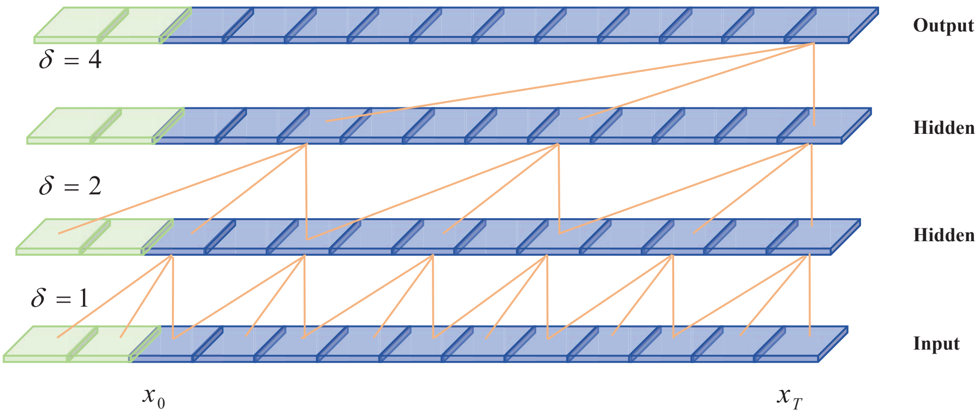


Figure 4. Diagram of causal convolution dilation. The diagram illustrates a causal dilated convolution operation with two hidden layers, where the initial data is set to 0, the direction of information flow is indicated by the connecting lines, and the output at time t represents the scenario power value.

3.3. Spatiotemporal Feature Extraction Unit for Source–Load Power Based on TGAN

CGANs extend the traditional GAN framework by integrating both supervised and unsupervised learning, resulting in enhanced generalization for specific data types. Unlike conventional GANs, CGANs incorporate conditional inputs into both the generator and discriminator, allowing for more controlled generation. This conditional input guides the generator to learn the mapping between data sample probability distributions and specified conditions, leveraging label information to improve data alignment with given parameters. In this paper, based on the CGAN framework, an MHAM and TCN are introduced to construct a TGAN, thereby enhancing the spatiotemporal feature extraction capabilities for source–load data. The generator takes Gaussian white noise Z and clustered daily source–load state labels C as inputs, and the output data are multi-spatiotemporal scale source–load scenarios S, where d_z is the noise dimension. The discriminator inputs are the source–load historical data $\mathrm{P}^{\mathrm{obs}}$, integrated with state labels, and the generated scenario S. By analyzing the features of the input source–load data, the discriminator outputs the classification results on the authenticity of the input data.

Considering that the generator and discriminator are relatively independent structures, their loss functions can be expressed as, respectively,

$$L_G = -\mathbb{E}_{Z,C}[D(S(Z|C))], \tag{9}$$

$$L_D = -\mathbb{E}_{\mathrm{P}^{\mathrm{obs}},C}[\log D(\mathrm{P}^{\mathrm{obs}}|C) + \mathbb{E}_{Z,C}[\log(1 - D(S(Z|C)|C))] \tag{10}$$

where $\mathbb{E}(\cdot)$ represents the expected value of the corresponding random variable, $D(\cdot)$ is the discriminator. The generator network takes random noise and label conditions as inputs and outputs synthesized source–load data. The discriminator network receives historical source–load data sets, generated source–load data, and label conditions as inputs, and outputs assessments distinguishing between real and fake data. The generator aims to create realistic data, while the discriminator seeks to improve the quality of the generated data by differentiating between genuine and synthetic data. This mutual competition ultimately ensures the high quality of the source–load data. Therefore, the training process of TGAN can be described as a min-max game, and its objective function is

$$\min_G \max_D V(G,D) = \mathbb{E}_{\mathrm{P^{obs}},C}[\log D(\mathrm{P^{obs}}|C)] + \mathbb{E}_{Z,C}[\log(1 - D(S(Z|C)|C))] \tag{11}$$

In the original GAN framework, the discriminator's loss function is dominated by Jensen–Shannon (JS) divergence, which can introduce training challenges, such as mode collapse. When the generated sample distribution is inconsistent with the real sample distribution, JS divergence remains static and cannot provide meaningful feedback. For generating source–load scenarios across multiple spatiotemporal scales, it is essential for the generator to capture the distribution patterns of source–load scenarios accurately. However, using JS divergence as the discriminator's loss function hinders effective measurement of distribution differences, often causing gradient vanishing during back propagation and complicating network training, thereby reducing the accuracy of the generated scenarios. By adopting the Wasserstein distance in place of the JS divergence as the discriminator's loss function, issues such as training instability and mode collapse can be effectively mitigated, resulting in a more precise fit to the target probability distributions. Consequently, the Wasserstein distance can be defined as follows:

$$W_C(P_{\mathrm{P^{obs}}|C}, P_{S(Z|C)|C}) = \inf_{\gamma \in \Pi(P_{\mathrm{P^{obs}}|C}, P_{S(Z|C)|C})} \mathbb{E}_{(\mathrm{P^{obs}},S(Z|C))\sim\gamma}[\|\mathrm{P^{obs}} - S(Z|C)\|_1] \tag{12}$$

where the discriminator adheres to 1-Lipschitz continuity, meaning that the maximum absolute value of its gradient is bounded by 1. By introducing the gradient penalty function D within its defined domain, we ensure the discriminator approximately meets the 1-Lipschitz continuity requirement, thus the Wasserstein distance can be accurately represented by the discriminator's loss function. Consequently, the objective function for the CGAN framework is reformulated as follows:

$$\min_G \max_D V(G,D) = \mathbb{E}_{\mathrm{P^{obs}},C}[D(\mathrm{P^{obs}}|C)] - \mathbb{E}_{Z,C}[D(S(Z|C)|C)] + \lambda \mathbb{E}_{\hat{P}\sim P_{\hat{P}}}[(\|\nabla_{\hat{P}} D(\hat{P}|C)\|_2 - 1)^2] \tag{13}$$

Unlike traditional CGAN, the advantage of using Wasserstein distance as a discriminator is that it can not only distinguish generated samples from real samples, but also fit Wasserstein distances between sample sets to accurately describe the differences between their distributions. Therefore, it provides an accurate training direction for the generator to fit the data distribution. This enables it to map conditional probability distributions rather than simply generate data similar to a real sample.

3.4. Stochastic Generation Method for Source–Load Scenarios at Multi-Time Scales

By combining the aforementioned source–load day-to-day transition process and intraday temporal power simulation method, a stochastic generation method for source–load power scenarios at multi-time scales is established. This method can randomly generate a specified number of multi-time scale N_s source–load scenarios with spatiotemporal correlation (see Figure 5). The specific generation steps are as follows:

(1) The monthly historical source–load data set X is divided into Θ categories by the K-means clustering technique.
(2) Based on the typical daily power states obtained from clustering, matrices $\mathcal{P}_t$ and $\mathcal{P}_c$ are calculated. Then, the initial source–load scenario set T_{in} is randomly generated using the MCMC algorithm.
(3) Taking the daily power states during the monthly power generation state transition in set T_{in} as labels and driven by Gaussian white noise Z, the source–load daily power curves for the corresponding states are generated based on TGAN.
(4) Based on the order determined by the source–load state transition process at each time scale, the daily power curves of the source and load are concatenated to form the long-term generation scenario set S.

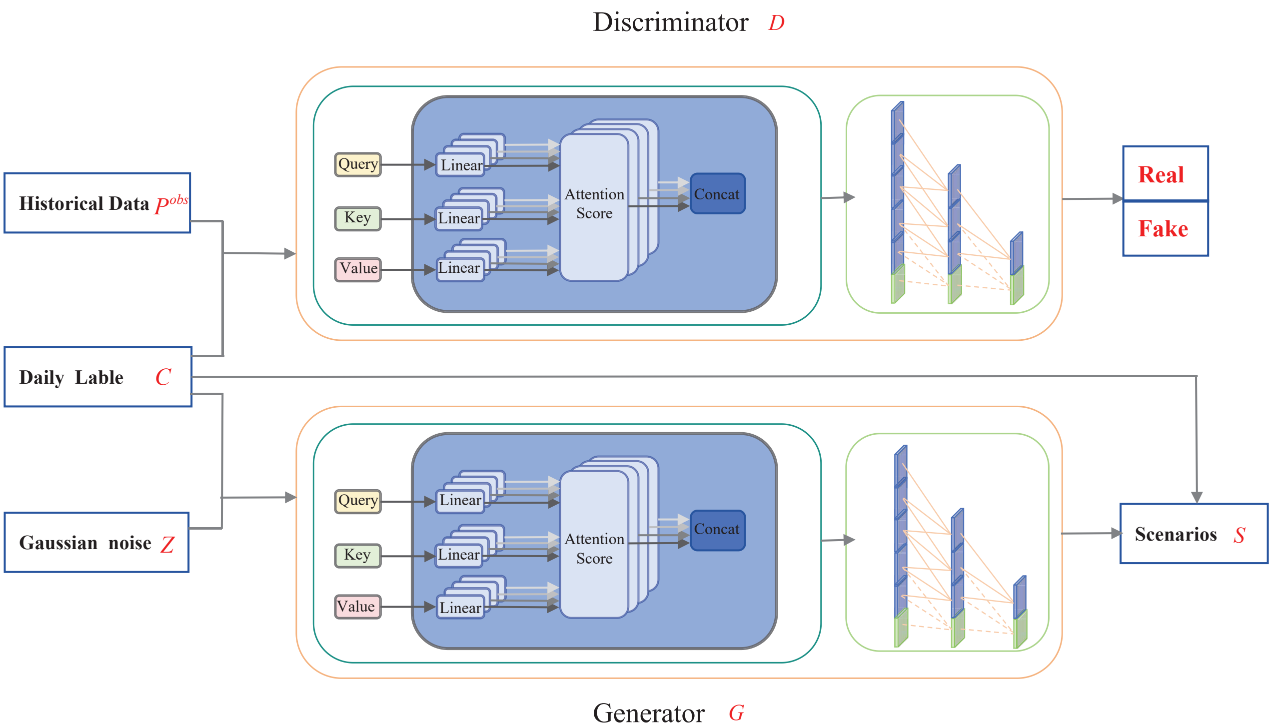


Figure 5. Framework of source–load scenario generation. The diagram depicts using historical data P^{obs}, daily label C, and Gaussian noise Z as network inputs, with spatiotemporal feature extraction performed by the MHAM and TCN, ultimately producing scenario S as the output.

4. Model Training

The validation of the proposed algorithm is supported by the source–load historical data set from the distribution network in Shandong, China, where the data sampling frequency is every 15 min. The number of representative daily power generation states is determined as $K = 6$ using the commonly used K-means method for evaluating clustering effectiveness. The experiments were conducted using Python 3.8.0 and the deep learning framework PyTorch 2.0.0. The computer hardware specifications include an Intel i9-10920X CPU (Intel Corporation, Santa Clara, CA, USA) with a frequency of 3.50 GHz and 24 cores, an NVIDIA GeForce RTX 3080 Ti GPU (NVIDIA Corporation, Santa Clara, CA, USA) , and 128 GB of RAM. During training, the learning rate is 0.0005, and the batch size is 32 (see Figure 6). The generator ultimately produces 1800 typical daily source–load scenarios per month, which are concatenated to form a monthly-scale source–load scenario set of 60 months (see Figures 7–9).

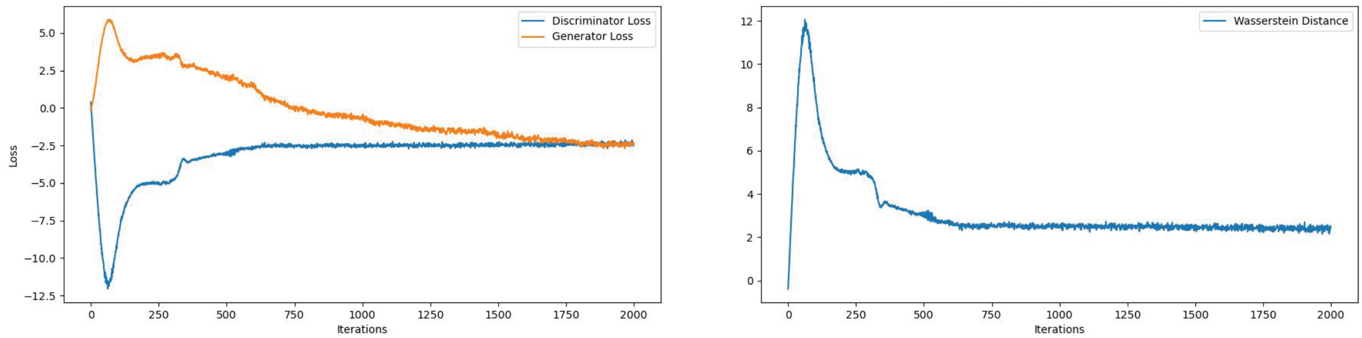


Figure 6. Convergence trend of the discriminator and Wasserstein distance.

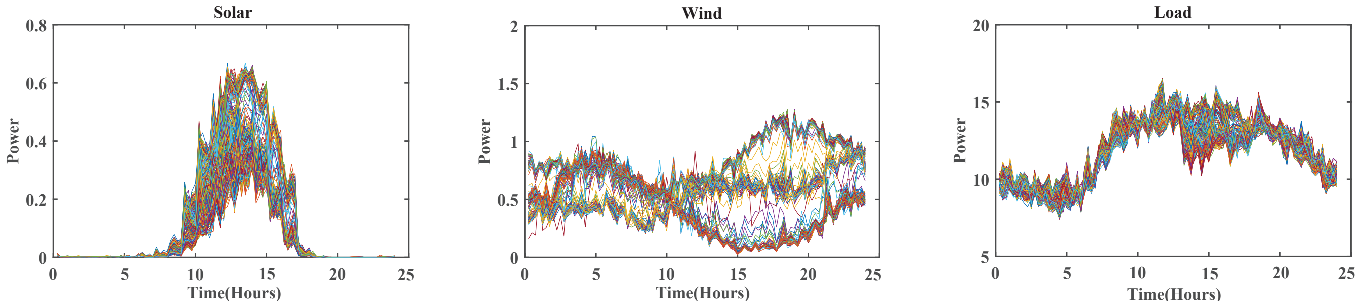


Figure 7. The diagram depicts the solar, wind power, and load scenario sets generated by the TGAN, where 24 h represents the daily time scale.

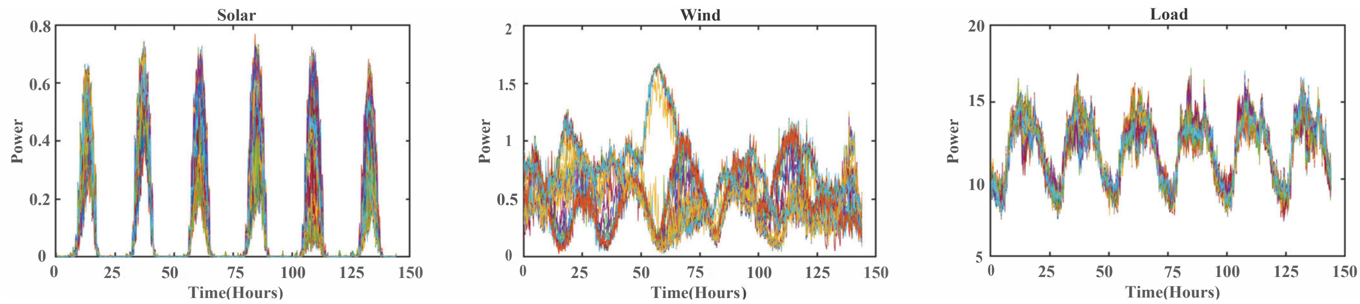


Figure 8. The diagram depicts the solar, wind power, and load scenario sets generated by the TGAN, where 144 h represents the weekly time scale.

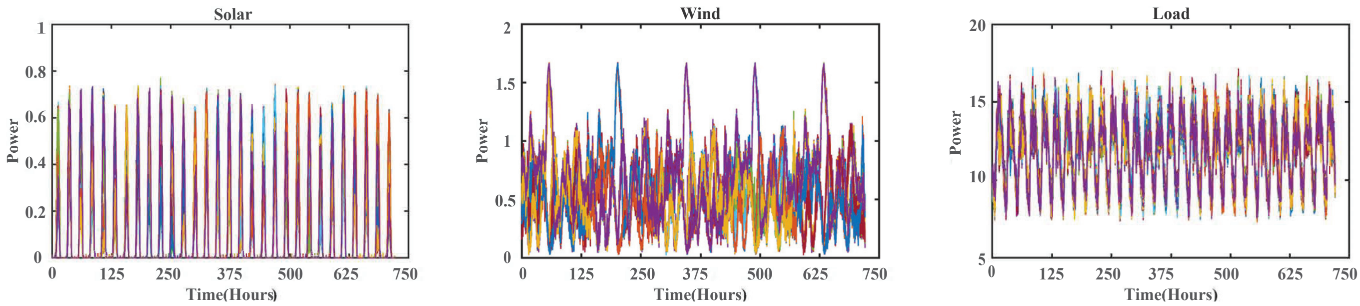


Figure 9. The diagram depicts the solar, wind power, and load scenario sets generated by the TGAN, where 720 h represents the monthly time scale.

As shown in Figure 6, during the process of generating source–load scenarios using the TGAN, we can observe that the initial loss of the discriminator dropped rapidly and flattened after about 500 iterations, which shows that the discriminator can quickly learn to differentiate real data with generated data. At the same time, the generator losses also decreased significantly in the early stages, but their downward trend was relatively slow, and it did not gradually stabilize until about 1500 iterations, which reflects that the generator shows gradual improvement in the ability to generate realistic data. In addition, the Wasserstein distance has a brief upward phase at the beginning of training, then rapidly declines and reaches a stable state after about 500 iterations, which is a phenomenon that shows that the model effectively narrows the gap between the real data distribution and the generated data distribution. To sum up, these trends show that the model has good convergence, and the losses of both discriminators and generators reach a stable state after an appropriate number of iterations. At the same time, the change in Wasserstein distance also confirms this, and after the initial learning stage, the distance between distributions remains at a stable level.

To emphasize the advantages of using the MHAM, we have added comparative experiments and used root mean square error (RMSE) and mean absolute error (MAE) as evaluation metrics. As shown in Table 1, compared to "AM + TCN", "MHAM + TCN" demonstrates a more significant advantage in improving the quality of generated data. Specifically, for solar data and wind data, the quality improvement is approximately 2%. For load data, the improvement is approximately 3%. Therefore, we can firmly conclude that method MHAM has a notable advantage in enhancing data quality and spatial feature extraction. Based on the results presented in Table 1, we analyzed the errors of methods "MHAM + TCN" and "MHAM + RNN". The comparison revealed that method "MHAM + TCN" has certain advantages, although these advantages are not immediately apparent. To further illustrate the benefits of "MHAM + TCN", we introduced method "AM + RNN"; it was found that the quality of generated source–load data using "MHAM + TCN" showed significant improvements over historical data. Specifically, the enhancements in data quality are evident when comparing the performance metrics between different methods, demonstrating the notable advantages of method "MHAM + TCN" in improving data quality and spatial feature extraction.

Table 1. The comparative results of ablation experiments.

	Solar		**Wind**		**Load**	
	RMSE	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**
MHAM + TCN	0.056	0.054	0.099	0.069	0.132	0.079
AM + TCN	0.078	0.071	0.110	0.087	0.161	0.106
MHAM + RNN	0.060	0.061	0.103	0.075	0.141	0.085
AM + RNN	0.088	0.082	0.124	0.096	0.171	0.115

Figures 7–9 illustrate the source–load scenarios generated on daily, weekly, and monthly time scales using the TGAN. Each set of scenarios not only clearly reflects the fluctuations and trends in source–load power under different cycles but also uncovers subtle patterns within complex spatiotemporal features. Therefore, it can be concluded that TGAN has unique advantages in handling complex and multi-time source–load characteristics, providing more precise data support for the planning and scheduling of power systems.

5. Evaluation of Generated Scenario Quality

The source–load scenarios generated by the TGAN should closely match the characteristics of real scenarios. To ensure high fidelity of the generated scenarios, this paper conducts a comprehensive evaluation from three aspects: time correlation, clustering effectiveness, and probability distribution characteristics.

5.1. Temporal Correlation Indicators

In this paper, the autocorrelation coefficient (ACF) is used as the time correlation index to simulate the generated source–load scenario set, and the time correlation is compared with the historical data set to evaluate the performance of the generated scenario. The

ACF reflects the correlation degree of source–load data over time, and the autocorrelation coefficient decreases gradually with the increase of hysteresis. Therefore, the performance of the source–load scenario set in capturing temporal characteristics can be measured by analyzing the magnitude and variation patterns of the ACF. The formula is as follows:

$$\rho_k = \frac{\Sigma_{t=1}^{m-j}(\mathcal{Q}_t - \bar{\mathcal{Q}})(\mathcal{Q}_{t+j} - \bar{\mathcal{Q}})}{\Sigma_{t=1}^{m}(\mathcal{Q}_t - \bar{\mathcal{Q}})^2} \tag{14}$$

where $\mathcal{Q}_t$ represents the source–load power at time t, j represents the time interval, $\bar{\mathcal{Q}}$ represents the average source–load power, and ρ_k represents the source–load autocorrelation coefficient with a time lag of j, which ranges from $[-1, 1]$. It should be noted that 1, -1, and 0, respectively, mean that there is a positive correlation, a negative correlation, and no correlation.

In Figure 10, the source–load historical scenario ACF is in stark contrast to the generated data ACF, providing a strong scenario support for us to evaluate the quality and fidelity of the generated scenario. The ACF curve of the generated solar data closely matched the actual data, exhibiting a similar trend. The correlation decreased significantly over time, reaching a negative value at around 12 h before returning to zero. This arrangement suggests that the TGAN successfully captures the typical cyclical pattern of solar scenario generation, most likely reflecting daily variations in sunlight. We can also find that the ACF of the generated wind scenario follows a similar pattern to the real scenario, although the correlation values are slightly biased over time, the overall correlation trend does not change significantly, which may be caused by the highly variable and random nature of wind scenario generation. Therefore, we can also conclude that the TGAN effectively learns the general time-dependent patterns of the wind scenario. For the load scenario, the ACF of the generated scenarios also aligns well with the real scenario, showing a similar dip in correlation around midday. This pattern reflects typical daily load cycles, with demand peaking at certain times and decreasing during off-peak hours. The close alignment of the ACF curves between the generated and actual load scenario suggests that the TGAN accurately models daily load fluctuation patterns. Overall, the data generated by TGAN shows high fidelity to the real scenario in terms of temporal correlation, as evidenced by the tight alignment of the ACF curves. The minimal deviation between the real and generated scenarios indicates that the generated scenario effectively captures the fundamental time characteristics of the historical source–load scenario. The high fidelity of this temporal correlation shows that the TGAN is well suited for generating real source–load scenarios.

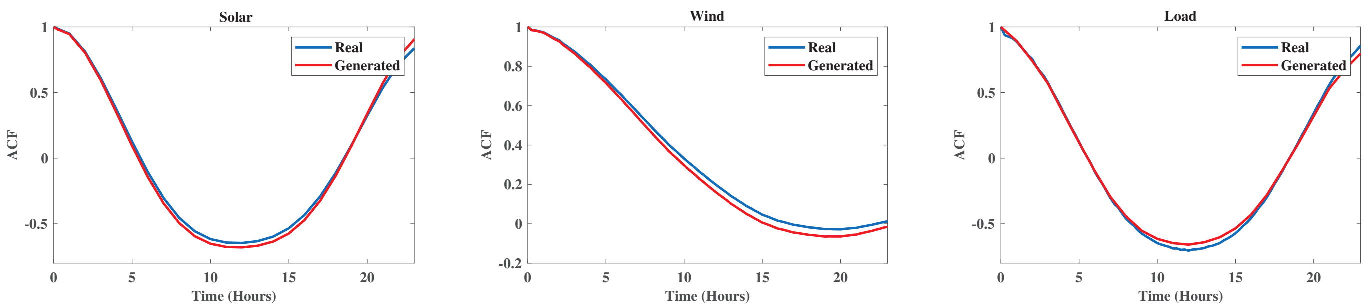


Figure 10. Time correlation trend chart of source–load scenario.

In addition, the ACF value of the source–load scenarios dropped rapidly in the early stage, which indicates that there is a strong short-term dependency. The high consistency between the actual scenarios and the ACF value of the generated scenarios indicates that the model can effectively capture the time pattern of the source–load scenarios. However, the overall decline rate of wind power data is slower than that of solar scenarios and load scenarios, and the matching degree is not so perfect, which shows that the TGAN needs further improvements in the generation of long-term wind power scenarios.

5.2. Clustering Effectiveness Indicators

The source–load scenarios generated by the TGAN typically consist of large amounts of multi-spatiotemporal data. On one hand, clustering algorithms can categorize and summarize these scenarios, reducing redundant data and lowering the computational complexity of subsequent processing and analysis. On the other hand, by using clustering algorithms to extract representative typical scenarios, the system's analysis and decision-making processes are enhanced.

Figures 11–13 show the power scenarios of *K*-means clustering of solar, wind, and load data after dimensionality reduction. Each group represents typical power patterns over different time periods, providing a representative view of daily, weekly, and monthly power fluctuations for each category. It is clear that the solar power generation scenario has a clear periodicity, with clear peaks corresponding to daily solar irradiance characteristics, which is consistent with typical solar power generation patterns. In contrast, wind power scenarios are more irregular and exhibit greater randomness due to the natural variability in wind speed. However, the TGAN managed to capture wind power at different intensities. Although the fluctuation patterns are dispersed, this is consistent with the randomness of changes in wind speed. On the other hand, load scenarios show cyclical variations, forming peaks and troughs during periods of high and low demand, which is consistent with typical daily load cycles. Additionally, to provide a more intuitive description of the realism of the generated source–load scenarios, Figure 14 illustrates the source–load scenarios divided into multiple clusters through clustering. Each cluster represents a group of source–load data patterns with similar characteristics, such as variations in solar power generation or electricity load. Different colors are used to distinguish these clusters, clearly showcasing various typical patterns and trends, such as peak load periods or power output under specific weather conditions. This enhances the understanding and simulation accuracy of real-world application scenarios. In summary, TGAN generates a high-fidelity representative power scenario that preserves the core pattern of the original scenarios, making it suitable for further analysis or as an input to models that require typical everyday patterns, thereby increasing efficiency without significant information loss.

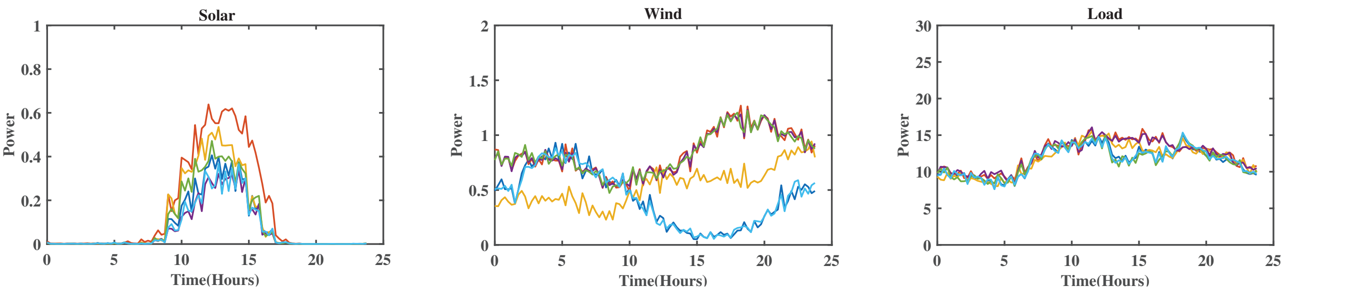


Figure 11. The source–load scenario diagram on a daily time scale after clustering reduction.

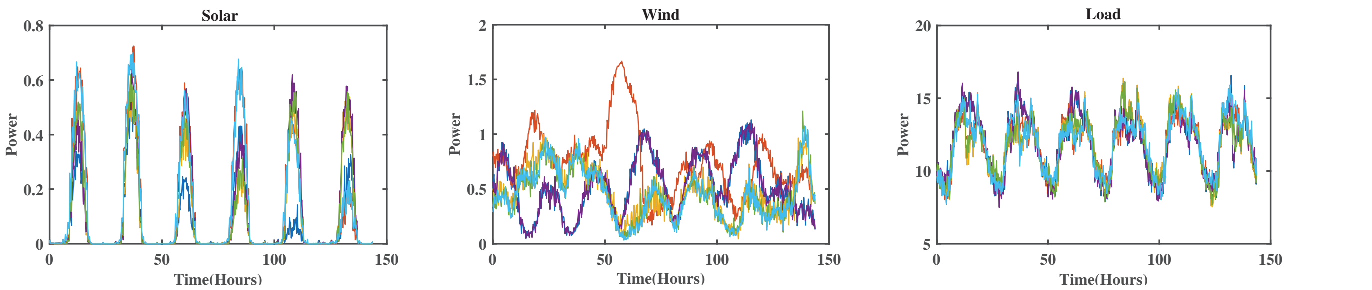


Figure 12. The source–load scenario diagram on a weekly time scale after clustering reduction.

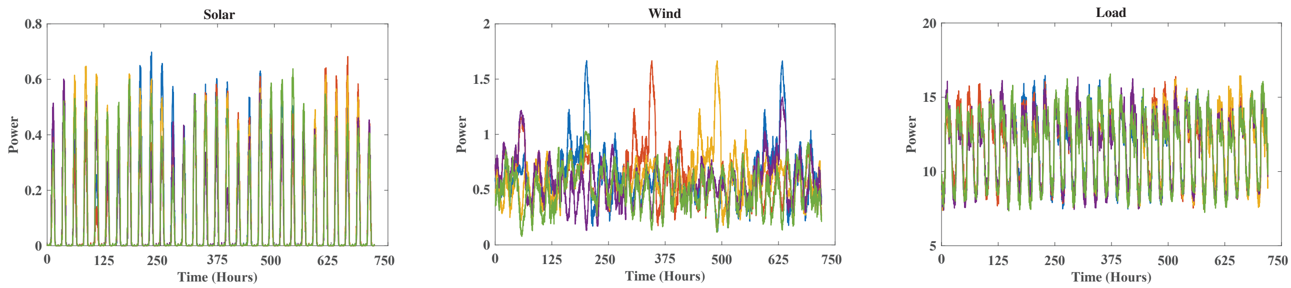


Figure 13. The source–load scenario diagram on a monthly time scale after clustering reduction.

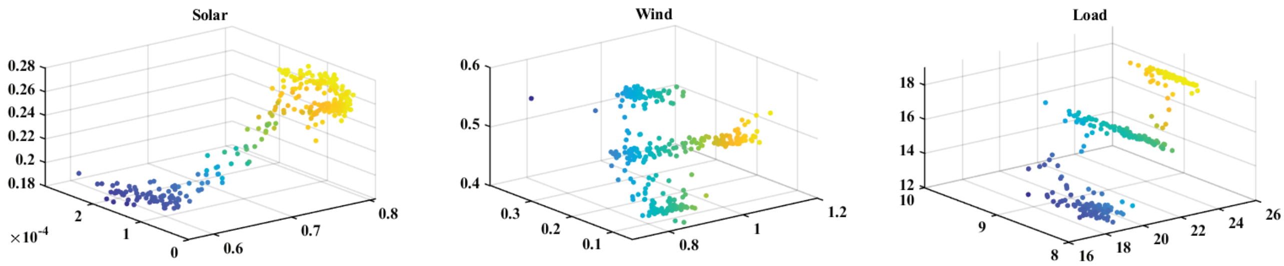


Figure 14. The distribution of source–load data after clustering.

5.3. Probability Density Distribution Characteristics

In order to further evaluate the quality of the generated source–load scenario, the cumulative distribution function (CDF) and probability density function (PDF) are introduced in this paper to visually represent the probability distribution of the source–load scenarios, which is convenient to compare the real scenarios.

As shown in Figure 15, the CDF curve of the generated scenarios are generally highly consistent with the actual data. Especially in the range of solar power 0–0.6 kW, wind power 0.4–0.8 kW, and load demand 8–16 kW, the actual scenarios almost completely coincide. This shows that the generated scenarios effectively capture the cumulative probability distribution characteristics of the real scenarios. Figure 16 provides some essential insights. The PDF curve of the generated photovoltaic scenarios aligns closely with the real scenarios in the low-power range (0–0.1 kW), showing a distinct peak characteristic. However, outside this low-power range, the generated scenarios exhibit increased fluctuation, indicating a slight deviation from the real distribution in less frequent power ranges, which may be due to the inaccurate model in handling low power output cases. For wind scenarios, the generated PDF curve is basically consistent with the actual scenarios in the main peak range (0–0.4 kW), and there is a slight difference above 0.4 kW, which indicates that the generated scenarios effectively capture the main distribution characteristics of wind scenarios, and the performance is poor in the case of high power. But it performs poorly at high-power conditions, suggesting that the model may not accurately capture the distribution of wind scenario output, especially near peaks. In addition, the load scenarios present a bimodal pattern, and the generated scenarios closely match the actual scenarios for the 8 kW and 12 kW peaks, accurately reflecting typical daily load variations.

It is important to note that a limitation of this study is that it does not consider source–load scenario generation under extreme conditions or the influence of meteorological factors. In future work, we will focus on the impact of extreme conditions and meteorological factors on source–load scenario models. Our aim is to improve the adaptability, robustness, and richness of these models, thereby providing more accurate technical support for planning personnel.

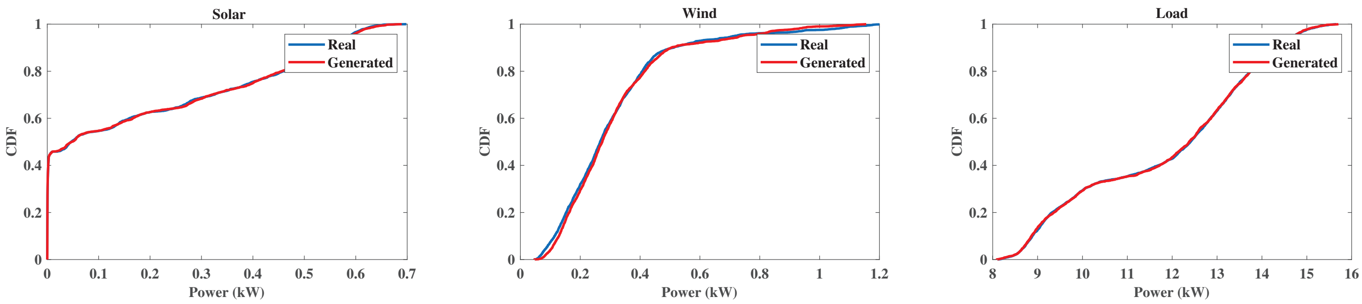


Figure 15. CDF of source–load scenarios.

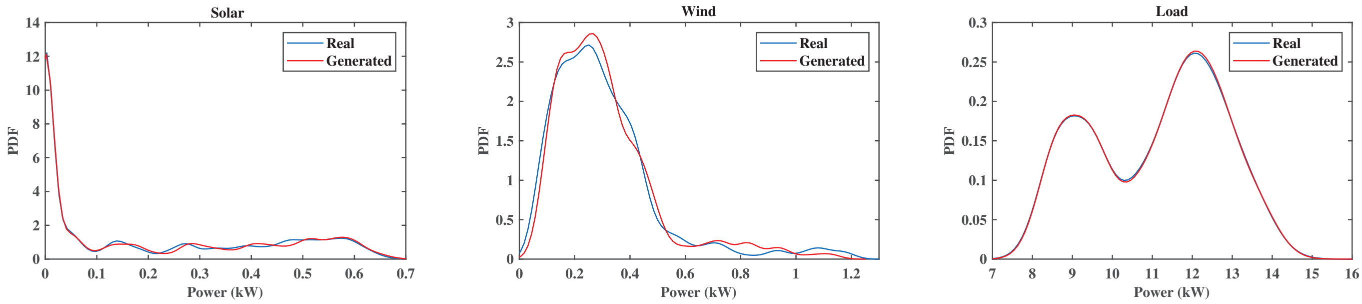


Figure 16. PDF of source–load scenarios.

6. Conclusions

This paper proposes a multi-time source–load power scenario generation method based on a TGAN. To meet the long-term analysis needs of distribution networks, a comprehensive evaluation system for source–load scenarios across various temporal and spatial scales is established. Using historical source–load power data from a county in Shandong, China, the proposed method is validated for effectiveness and accuracy. The main conclusions are as follows:

(1) Dividing typical daily source–load power states and simulating the state transition patterns with a Markov chain helps to uncover the probabilistic characteristics of source–load power across multiple temporal and spatial scales.
(2) The introduction of an MHAM and TCN enhances the TGAN's capability to capture the spatiotemporal features of source–load power, enabling the generated multi-scale source–load scenarios to better reflect the probabilistic and correlation characteristics of source–load power.
(3) The evaluation system assesses the quality of the generated source–load scenarios, providing detailed and reliable analytical data to support long-term power and energy balancing as well as operational planning in distribution networks.

Additionally, the TGAN can effectively simulate the actual energy generation and consumption mode, which provides a powerful tool for source–load prediction, further optimizing distribution network management, and can better forecast future electricity consumption trends and formulate more scientific and reasonable investment plans and infrastructure upgrade strategies. In future work, we will focus on improving model algorithms based on the limitations of existing models to improve accuracy in these critical areas. At the same time, other factors affecting the change of source–load will be considered, which will further improve the accuracy of source–load scene. Finally, we will select the appropriate algorithm to automatically adjust and optimize the model parameters to make the TGAN model adapt to different geographical and climatic conditions.

Author Contributions: Conceptualization, methodology, and visualization, L.M., S.J. and Y.S.; validation, L.M. and X.L.; formal analysis, investigation, resources, and data curation, L.M., S.J. and

C.S.; writing—original draft preparation, L.M., S.J. and Y.S.; writing—review and editing, L.M., S.J., Y.S. and C.S.; supervision, L.M. and Y.S; project administration, L.M.; funding acquisition, L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Science and Technology Project of State Grid Corporation of China (5400-202356678A-3-3-JC).

Data Availability Statement: The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare that this study received funding from State Grid Corporation of China. Authors Liang Ma, Shigong Jiang and Yi Song were employed by the State Grid Economic and Technological Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Abbreviations

The following abbreviations are used in this manuscript:

GAN	Generative adversarial network
CGAN	Conditional generative adversarial network
TGAN	Temporal generative adversarial network
MHAM	Multi-head attention mechanism
TCN	Temporal convolutional network
AM	Attention mechanism
RNN	Recurrent neural network
ACF	Autocorrelation coefficient
CDF	Cumulative distribution function
PDF	Probability density function
kW	Kilowatt
kWh	Kilowatt-hour
MCMC	Markov chain Monte Carlo
RMSE	Root mean square error
MAE	Mean absolute error

Appendix A

The code in this article is improved on the open source code [32]. Due to the confidentiality agreement, this article only provides open source code links "https://github.com/chennnnnyize/Renewables_Scenario_Gen_GAN" (accessed on 17 June 2024).

References

1. Li, H.; Ren, Z.; Fan, M. A review of scenario analysis methods in planning and operation of modern power systems: Methodologies, applications, and challenges. *Electr. Power Syst. Res.* **2022**, *205*, 107722. [CrossRef]
2. Shao, C.; Wang, Y.; Feng, C. Mid/long-term power system production simulation considering multi-energy generation. *Proc. CSEE* **2020**, *40*, 4072–4080.
3. Yang, M.; Shi, C.; Liu, H. Day-ahead wind power forecasting based on the clustering of equivalent power curves. *Energy* **2021**, *218*, 119515. [CrossRef]
4. Zhao, W.; Shao, Z.; Yang, S.; Lu, X. A novel conditional diffusion model for joint source-load scenario generation considering both diversity and controllability. *Appl. Energy* **2025**, *377*, 124555. [CrossRef]
5. Gu, J.; Liu, S.; Hu, Y.; Meng, L. Optimal allocation of intermittent distributed generation based on deep convolutions generative adversarial network in scenario generation. *Power Syst. Technol.* **2021**, *45*, 1742–1751.
6. Chen, G.; Wu, Y.; Yang, L.; Xu, K.; Lin, G.; Zhang, Y.; Zhang, Y. Ultra-short-term load dynamic forecasting method considering abnormal data reconstruction based on model incremental training. *Energies* **2022**, *15*, 7353. [CrossRef]

7. Sheng, Y.; Wang, H.; Yan, J.; Liu, Y.; Han, S. Short-term wind power prediction method based on deep clustering-improved temporal convolutional network. *Energy Rep*. **2023**, *9*, 2118–2129. [CrossRef]
8. Yao, W.; Huo, Z.; Zou, J.; Wu, C.; Wang, J.; Wang, X.; Lu, S.; Xie, Y.; Zhuo, Y.; Liang, J.; et al. Medium-and long-term power system planning method based on source-load uncertainty modeling. *Energies* **2024**, *17*, 5088. [CrossRef]
9. Liu, D.; Wang, W.; Zhang, H. Mid-long term available quantity of electricity forecasting with error calibration by hidden markov model in photovoltaic and application of dispatching plan. *High Volt. Eng.* **2023**, *49*, 840–848.
10. Mayer, M.; Biro, B.; Szvcs, B. Probabilistic modeling of future electricity systems with high renewable energy penetration using machine learning. *Appl. Energy* **2023**, *336*, 120801. [CrossRef]
11. M, C.; Zhang, J.; Cheng, C.; Xu, Y.; Yang, Y. High-dimensional uncertainty scenario generation method for basin hydro-wind-solar multi-energy complementary system considering spatio-temporal correlation. *Power Syst. Technol.* **2024**, *48*, 3614–3623.
12. Li, P.; Liu, C.; Huang, Y. Modeling correlated power time series of multiple wind farms based on hidden markov model. *Proc. CSEE* **2019**, *39*, 5683–5691.
13. Sidor, K.; Miller, P.; Małkowski, R.; Izdebski, M. Optimization of division and reconfiguration locations of the medium-voltage power grid based on forecasting the level of load and generation from renewable energy sources. *Energies* **2024**, *17*, 4933. [CrossRef]
14. Li, D.; Ren, Z.; Yan, W. Month-ahead wind power curve probabilistic prediction based on factor analysis and quantile regression neural network. *Proc. CSEE* **2017**, *37*, 5238–5247.
15. Camal, S.; Teng, F.; Michiorri, A.; Kariniotakis, G.; Badesa, L. Scenario generation of aggregated wind, photovoltaics and small hydro production for power systems applications. *Appl. Energy* **2019**, *242*, 1396–1406. [CrossRef]
16. Densing, M.; Wan, Y. Low-dimensional scenario generation method of solar and wind availability for representative days in energy modeling. *Appl. Energy* **2022**, *306*, 118075. [CrossRef]
17. Monforti, F.; Huld, T.; Bódis, K.; Vitali, L.; D'Isidoro, M.; Lacal-Arántegui, R. Assessing complementarity of wind and solar resources for energy production in Italy. A Monte Carlo approach. *Renew. Energy* **2014**, *63*, 576–586. [CrossRef]
18. Costoya, X.; deCastro, M.; Carvalho, D.; Gómez-Gesteira, M. Assessing the complementarity of future hybrid wind and solar photovoltaic energy resources for North America. *Renew. Sustain. Energy Rev*. **2023**, *173*, 113101. [CrossRef]
19. Hu, J.; Li, H. A transfer learning-based scenario generation method for stochastic optimal scheduling of microgrid with newly-built wind farm. *Renew. Energy* **2022**, *185*, 1139–1151. [CrossRef]
20. Li, J.; Lan, F.; Wei, H. A scenario optimal reduction method for wind power time series. *IEEE Trans. Power Syst*. **2016**, *31*, 1657–1658. [CrossRef]
21. Hu, J.; LI, H. A new clustering approach for scenario reduction in multi-stochastic variable programming. *IEEE Trans. Power Syst*. **2019**, *34*, 3813–3825. [CrossRef]
22. Smilde, A.; Kiers, H.; Bijlsma, S. Matrix correlations for high-dimensional data: The modified RV-coefficient. *Bioinformatics* **2009**, *25*, 401–405. [CrossRef] [PubMed]
23. Ma, X.-Y.; Sun, Y.-Z.; Fang, H.-L. Scenario generation of wind power based on statistical uncertainty and variability. *IEEE Trans. Power Syst*. **2013**, *4*, 894–904. [CrossRef]
24. Li, J.; Zhou, J.; Chen, B. Review of wind power scenario generation methods for optimal operation of renewable energy systems. *Appl. Energy* **2020**, *280*, 115992. [CrossRef]
25. Liao, W.; Yang, D.; Ge, L. Renewable energy scenario generation based on pixel convolutional generative networks. *High Volt. Eng.* **2022**, *48*, 1320–1331.
26. Zhu, R.; Liao, W.; Wang, Y. Stochastic scenarios generation for wind power and photovoltaic system based on generative moment matching network. *High Volt. Eng.* **2022**, *48*, 374–384.
27. Ma, X.; Liu, Y.; Yan, J.; Wang, H. A WGAN-GP-Based scenarios generation method for wind and solar power complementary study. *Energies* **2023**, *16*, 3114. [CrossRef]
28. Zhu, J.; Ai, Q.; Chen, Y.; Wang, J. Single-location and multi-locations scenarios generation for wind power based on WGAN-GP. *J. Phys. Conf. Ser*. **2023**, *2452*, 012022. [CrossRef]
29. Zhang, Y.; Ai, Q.; Xiao, F.; Hao, R.; Lu, T. Typical wind power scenario generation for multiple wind farms using conditional improved Wasserstein generative adversarial network. *Int. J. Electr. Power Energy Syst*. **2020**, *114*, 105388. [CrossRef]
30. Tang, J.; Liu, J.; Wu, J.; Jin, G.; Kang, H.; Zhang, Z.; Huang, N. RAC-GAN-Based scenario generation for newly built wind farm. *Energies* **2023**, *16*, 2477. [CrossRef]
31. Wang, Q.; Zhang, X.; Xu, D. Source-load scenario generation based on weakly supervised adversarial learning and its data-driven application in energy storage capacity sizing. *IEEE Trans. Sustain. Energy* **2023**, *14*, 1918–1932. [CrossRef]
32. Chen, Y.; Wang, Y.; Kirschen, D. Model-free renewable scenario generation using generative adversarial networks. *IEEE Trans. Power Syst*. **2017**, *33*, 3265–3275. [CrossRef]
33. Dong, X.; Sun, Y.; Pu, T. Day-ahead scenario generation of renewable energy based on conditional GAN. *Proc. CSEE* **2020**, *40*, 5527–5535.

34. Li, H.; Ren, Z.; Hu, B. A sequential generative adversarial network based monthly scenario analysis method for wind and photovoltaic power. *Proc. CSEE* **2022**, *42*, 537–547.
35. Zhu, C.; Liu, D.; Huang, J.; Han, X.; Gao, Y.; Sun Y. Monthly source and load scenario generation based on progressive growing of generative adversarial networks. *High Volt. Eng.* **2024**, *50*, 3955–3964.
36. Huang, N.; Wang, W.; Cai, G.; Yang, D.; Huang, D.; Song, X. The joint scenario generation of multi source-load by modular denoising variational autoencoder considering the complex coupling characteristics of meteorology. *Proc. CSEE* **2019**, *39*, 2924–2934.

energies

MDPI

Article

Accident Factors Importance Ranking for Intelligent Energy Systems Based on a Novel Data Mining Strategy

Rongbin Li [1], Jian Zhang [1] and Fangming Deng [2,*]

[1] Huizhou Power Supply Bureau, Guangdong Power Grid Corporation, Huizhou 516000, China; hfcyj_lib@126.com (R.L.); zff990616@163.com (J.Z.)
[2] School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China
* Correspondence: 2464@ecjtu.edu.cn

Abstract: As global energy networks expand and smart grid technology evolves rapidly, the volume of historical power accident data has increased dramatically, containing valuable risk information that is essential for building efficient public safety early warning systems. This paper introduces an innovative text analysis method, the Sparse Coefficient Optimized Weighted FP-Growth Algorithm (SCO-WFP), which is designed to optimize the processing of power accident-related textual data and more effectively uncover hidden patterns behind accidents. The method enhances the evaluation of sparse risk factors by preprocessing, clustering analysis, and calculating piecewise weights of power accident data. The SCO-WFP algorithm is then applied to extract frequent itemsets, revealing deep associations between accident severity and risk factors. Experimental results show that, compared to traditional methods, the SCO-WFP algorithm significantly improves both accuracy and execution speed. The findings demonstrate the method's effectiveness in mining frequent itemsets from text semantics, facilitating a deeper understanding of the relationship between risk factors and accident severity.

Keywords: smart grid; power security; text analysis; data mining strategy; FP-Growth algorithm

1. Introduction

Th electric power industry plays a key role in the life and production of modern society [1], but the inherent conditions in the implementation of electric power production are complex, the accidents are frequent, and the situation of safety production is grim [2,3]. As the scale of the power grid continues to expand, the data on electric power personal accidents have increased explosively. Researchers have found a certain regularity in the occurrence of power accidents in historical data, but the historical data of power accidents are still not fully explored [4]. In order to explore the relationship between different risk factors and the levels of power accidents, it is of great practical significance to mine the risk factors of power accidents that have occurred based on historical records of accidents.

Against the background of the wide range of challenges in the power sector, many researchers have devoted themselves to solving different problems using data mining techniques. Gholami et al. [5] addressed the problem of large-scale outages occurring in a distribution network with distributed energy sources and proposed a method to effectively find the root cause of large-scale outages by aggregating, classifying, and mining the available data. Wu et al. [6] analysed historical data in terms of grid operation stability and mined the association rules between different time periods, different voltages, and different risk levels through specific grid operation cases and the rules conformed to the grid operation rules. Wang et al. [7]

generated typical operating scenarios in a photovoltaic (PV) power system by correlating PV output scenarios with load scenarios and considered the weather factor to make the generated typical scenarios interpretable using the weather factor. Zhou et al. [8] used a link prediction method to explore the association rules of coal mine accident hidden danger texts in order to effectively mine valuable coal mine accident hidden danger information from massive text data and make predictions. Shen et al. [9] combined a comprehensive model of text mining and machine learning based Bayesian networks and used sensitivity analysis to determine the key risk factors of safety accidents in metro construction. However, in the current field of power accidents, the research on systematically mining the specific risk factors of power accidents is still insufficient, which to a certain extent restricts the accurate decision-making of power safety risks and the formulation of effective prevention and control strategies. In order to effectively prevent and timely deal with power accidents, knowing how to mine effective risk factors of accidents from a massive amount of recorded data is the current challenge. Thus, this paper comprehensively analyses the risk factors of electric power accidents based on a data mining algorithm.

The word vector generation model Word2Vec has been applied to the electric power domain several times, e.g., Liu et al. [10] combined the Word2Vec model and convolutional neural network (CNN) for electric power equipment fault diagnosis and mined the contextual semantic features of the words through the Word2Vec model to represent the descriptive text of equipment faults. However, generating word vectors using Word2Vec alone is not sufficient to fully reveal the deep structure and patterns in power incident texts. To solve this problem, this paper further introduces the K-means algorithm for cluster analysis. Among all clustering methods, the K-means algorithm is a commonly used unsupervised machine learning clustering method that is mainly used for grouping or classifying data sets [11]. In the field of electric power, K-means clustering is used to extract key factors from accident reports and classify the reports into different categories based on their contents. Wang et al. [12] applied K-means clustering to analyse electric power marketing information, which solved the limitations of the traditional BERT model and achieved high-precision and fast clustering identification of information. Therefore, in this paper, a K-means clustering method is used to achieve accurate clustering of text before association rule mining and provide more structured input data for subsequent association rule mining.

Association rule mining, as one of the important research methods in the field of data mining [13,14], is used to discover association relationships between items in a dataset. For power accident text data [15], association rule mining methods analyse the hidden patterns and connections in unstructured natural language information. The more typical of this method are a priori test class algorithms that generate candidate sets and Frequent Pattern-Growth (FP-Growth) algorithms that does not generate candidate sets. Agrawal et al. [16] first proposed an a priori algorithm for the shopping basket problem, which is an iterative search method that scans the original data several times, constructs and filters the candidate sets, and exhausts the items in the dataset. Liu et al. [17] used a binary logic "and" operation methos to improve the a priori to analyse the association rules of risk elements in the field of civil aviation. However, the a priori algorithm needs to scan the dataset several times before generating frequent items, which brings a large I/O load and low efficiency, and at the same time generates a large number of candidate sets, which faces high computational complexity and memory consumption. Based on the above problems, Han et al. [18] first proposed an FP-Growth algorithm. FP-Growth algorithms only need to scan the database twice and construct Frequent Pattern Tree (FP-Tree) to compress the data. The frequent itemsets are obtained by recursively mining the conditional FP-Tree to reduce the complexity of storage and computation. Therefore, in this paper, the FP-Growth algorithm is used to reveal the potential patterns and key influencing factors of accidents.

Due to the different importance of different items in the dataset, many studies have used weighting to mine frequent itemsets. Xiao et al. [19] introduced a time decay factor to assign weights to the transactions and designed a weighting function to mine the relevant frequent items of the most recent time transactions. Yu et al. [20] applied entropy weighting method to calculate the risk index of different food types and used FP-Growth algorithm with constraints for association rule mining of food risk factors. However, existing weighted association rule mining methods often neglect the unique structure and importance of text data, and the use of a traditional FP-Growth algorithm in the electric power field makes it difficult to fully mine potential risk factors and accident patterns when dealing with complex text data [21]. In addition, most of the association rule mining methods do not take into account the weighted differences between risk factors, and they also lack effective treatment of sparse data. Therefore, the existing methods have deficiencies in computational efficiency, data processing depth, and result accuracy in the specific risk factor analysis of power accidents. Based on this, this paper chooses to improve on the traditional FP-Growth algorithm by introducing a weighted algorithm to mine power accident records for frequent items. Risk factors with very low frequency of occurrence, but also with strong correlation with accidents, are added to the calculation of the importance of accident records. Then the algorithm calculates the weighted support degree of risk factor feature items and sorts them in descending order of weighted support degree. In this way, an FP-Tree is constructed to mine the frequent itemsets to highlight the importance of the concern data and fully explore the potential power risk factors.

Inspired by the aforementioned research, this paper addresses the limitations of traditional association rule mining methods in fully utilizing text data. This paper proposes a novel text-weighted FP-Growth association rule mining scheme for power accidents to solve the problem of an insufficient basis for the analysis of power accidents, so as to provide a theoretical basis and decision-making support for risk management of the electric power industry and to promote safe and stable operation of the electric power system and sustainable development. The main contributions are as follows:

1. Enhanced text utilization: In this paper, we propose a text-weighted FP-Growth based association rule mining scheme for power accidents that improves the utilisation of text data by incorporating semantic information into the association rule mining process.
2. Word vector and classification: The cosine similarity is used to calculate the sub-word weights and quantify the semantic correlation between each word and accident level, thus optimising the classification of accident types and risk levels.
3. Mining sparse factors: Sparse factors are introduced to address the challenge that sparse risk factors in power accidents are not easily mined. The experimental results show that this approach improves the identification of uncommon but important risk factors by accurately assessing the importance of accident records.

2. Basic Theory

2.1. Word2vec Model

Word2Vec is a word vector generation model that learns semantic knowledge from a large text corpus in an unsupervised manner that is mainly applied in the field of Natural Language Processing (NLP), such as text classification, sentiment analysis, etc. [22–24]. The method converts words in text into numerical vectors that are mainly divided into two architectures: the Continuous Bag of Words (CBOW) model and the Skip-gram model. The CBOW model predicts the target word based on its contextual words, while the Skip-gram model operates in the reverse direction, predicting the context with the target word. The CBOW and Skip-gram model architectures are shown in Figure 1.

Word2vec training is usually accelerated by negative sampling techniques and the softmax objective function is optimised to minimise the prediction error. After training the word vectors, similar words will be closer in the high dimensional space as well. In this paper, a Word2vec model is used to generate word vectors for power accident dataset to characterise the semantic information by means of word vectors in order to facilitate the computer to understand their semantic and contextual relationships.

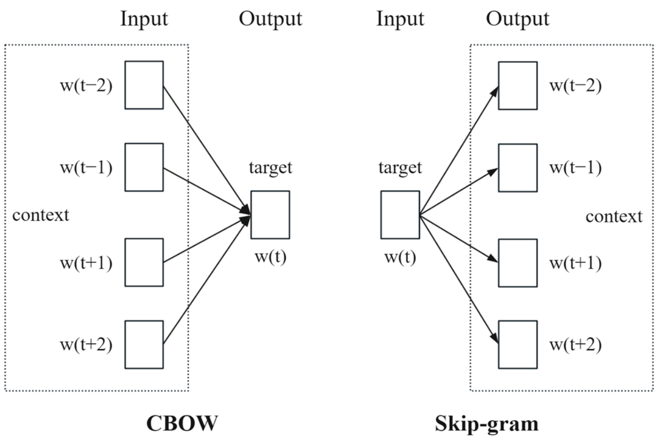


Figure 1. CBOW and Skip-gram model architectures.

2.2. K-Means Clustering

K-means clustering is performed by determining the similarity between samples, where similarity is usually measured using the distance between sample attributes [12,25]. The algorithm is based on an iterative process that divides observations (sample points) into a predetermined number of clusters (k centres), each consisting of data points closest to its centre of mass, which is also known as the mean or centre position of the cluster.

In order to perform association rule mining on power accident data, quantitative textual data need to be transformed into qualitative data, i.e., accident level data are risk graded and the sample data are transformed into the form of accident levels. The number of clusters k is determined according to the columns of accident levels in the dataset, which are Minor, Moderate, Major, and Severe. So, in this paper, k is set to 4, so that all the segmentation words belong to the same class as the representative words of one accident level. In this paper, the K-means algorithm is used to divide the dataset into four clusters representing four accident classes, which facilitates the subsequent mining of association rules between accident classes and risk factors. The main steps of the algorithm implementation are as follows:

(1) Initialisation: Four initial centres of mass (centroids) are randomly selected as representatives of each cluster.
(2) Assignment: The Euclidean distance d from each sample point to all centres of mass is computed and the sample point is assigned to the cluster with its closest distance. The Euclidean distance between the sample points $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ and $x_j = \left(x_{j1}, x_{j2}, \ldots, x_{jn}\right)$ is calculated as follows:

$$d\left(x_i, x_j\right) = \sqrt{\sum_{i=1,j=1}^{n} \left(x_i - x_j\right)^2} \tag{1}$$

(3) Update: For each cluster, recalculate its centre of mass c (the mean value of all points in the cluster) as follows:

$$c_j = \frac{\sum_{x_i \in C_j} x_i}{\left|C_j\right|} \tag{2}$$

where C denotes category, $C_j = \{C_1, C_2, C_3, C_4\}$.

(4) Iteration: If there is a change in the centre of mass of the cluster, repeat steps 2 and 3 until the centre of mass no longer changes or the preset maximum number of iterations is reached.

2.3. Cosine Similarity

In fields such as NLP, information retrieval, and data analysis, cosine similarity is a statistical measure of the angle between two non-zero vectors [26]. It is based on the vector space model and calculates the cosine of the angle between the directions of these two vectors, taking values from -1 to 1.

When two vectors are in the same direction, the angle is 0 degrees and the cosine similarity is 1. If they are perpendicular (orthogonal) to each other, the angle is 90 degrees and the cosine similarity is 0. If they are in the opposite direction, the angle is 180 degrees and the cosine similarity is -1. In this way, we can quantify the degree of similarity of the contents of any two texts in a dataset, and the smaller the angle is, the more similar the contents of these two texts are. In this paper, the cosine similarity approach is utilised to measure the similarity between accident descriptions, thereby revealing potentially similar events or risk factors. By calculating the cosine value between each word vector and the accident level and then multiplying the corresponding coefficients according to the accident level, we obtain the segmented words weights. The formula of this algorithm is as follows:

$$Similarity(A, B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} (A_i \times B_i)}{\sqrt{\sum_{i=1}^{n} {A_i}^2} \times \sqrt{\sum_{i=1}^{n} {B_i}^2}} \tag{3}$$

where A and B denote two different vectors; A·B denotes the dot product of vectors A and B; and ‖A‖ and ‖B‖ denote the modes of vectors A and B, respectively.

2.4. FP-Growth

Let $I = \{i_1, i_2, \ldots, i_{\mathrm{d}}\}$ denote the itemset consisting of all items I and $D = \{T_1, T_2, \ldots, T_{\mathrm{n}}\}$ denote the set of all transactions T. Each transaction T_{i} contains a set of items that are a subset of I, and the supports (s) are the ratio of the support count (*scount*) to the length of the transaction dataset $|D|$, as following:

$$s = \frac{scount}{|D|} \tag{4}$$

Confidence (c) represents the value that reflects the degree of confidence in the association rule, and the formula is as follows:

$$c(U \Rightarrow V) = \frac{s(U \cup V)}{s(U)} \tag{5}$$

where $s(U \cup V)$ is the support of the concatenated set of itemset U and itemset V, which represents the probability of U and V occurring at the same time and $s(U)$ denotes the support of itemset U.

The algorithm first introduces the FP-Tree, where each node in the tree corresponds to an item in the frequent itemset. Given the original dataset D and a specified minimum support threshold, the FP-Growth algorithm only needs to traverse the original data twice to generate FP-Tree data that compresses the original data. The implementation of the algorithm is mainly divided into two stages: constructing the FP-Tree and recursively mining of the FP-Tree, as illustrated in Figure 1.

(1) Construct the FP-Tree: Identify all frequent items from each transaction dataset and merge them to form the FP-Tree. In the FP-Tree, each internal node represents a frequent item, and the relationships between nodes reflect the order of item occurrences. The paths of the tree represent transactions with the same prefix. The specific steps are as follows:
   - Generate frequent itemset. The original transaction dataset D is scanned for the first time and all items in D are counted. The frequent 1-item set and the frequent item list F-list are filtered according to the minimum support and sorted in descending order.
   - Construct the tree structure. The original transaction dataset D is scanned for the second time, the dataset is arranged in the order of the frequent items in the F-list, and the FP-Tree is constructed for compressing and storing the dataset after ordering by frequent items.

(2) Recursively mining FP-Tree: Frequent patterns and corresponding support are generated by scanning the FP-Tree, followed by mining association rules that satisfy the minimum confidence threshold. The specific method is as follows:

The FP-Growth algorithm is based on the constructed FP-Tree to mine frequent items in a bottom-up principle. The total mining task is divided into several independent subtasks, which construct and recursively mine sub-FP-Trees to mine local frequent items. The local frequent items are connected with the suffix frequent items to generate longer frequent itemsets until no new frequent itemsets are generated. Figure 2 shows the flow of the FP-Growth algorithm for mining frequent itemsets.

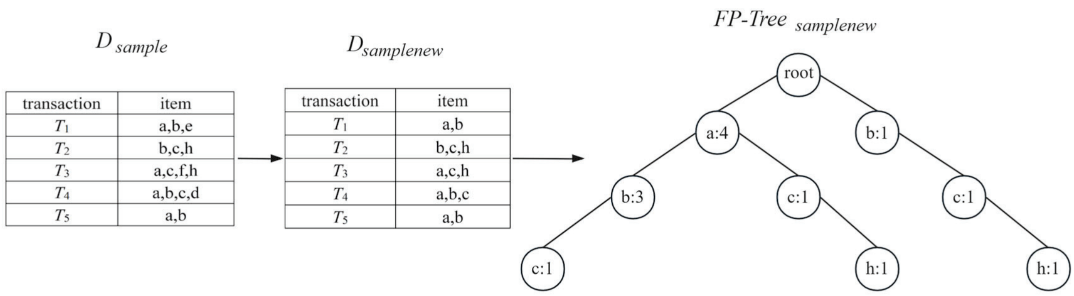


Figure 2. Process of mining FP-Growth frequent itemsets.

3. FP-Growth Algorithm Based on a Text-Weighted Method

To address the insufficient research on specific risk factors of electric power accidents, this paper explores the potential correlations between these risk factors and accident levels using an association rule mining method. For accident risk factors described in text. This paper designs the FP-Growth algorithm based on text weighting to analyse the impact and degree of association of different risk factors on accident level. The process of mining risk factors is shown in Figure 3.

In this paper, we firstly constructed a power accident dataset and preprocessed the text using a word-splitting method. Then the word vectors are generated by the CBOW model of Word2Vec, where the silhouette coefficient of the combination of the important parameters, vector_size (the dimension of the word vectors) and window (the size of the context window), are calculated to obtain the optimal combination of the parameters. Since the generated high-dimensional word vectors reduce the K-means clustering performance, the t-distributed stochastic neighbour embedding (t-SNE) algorithm is used to reduce the dimensionality before clustering, and then the cosine similarity is used to calculate the sub-word weights. Since the risk factors within a specific sparse range are not easy to mine, this paper introduces sparse coefficients, which together with the segmentation weights determine the importance of accident-related risk factors. Finally, frequent terms and association rules are mined based on the text-weighted FP-Growth algorithm defined in this paper. In addition, this paper optimizes the algorithm's running process to achieve higher efficiency.

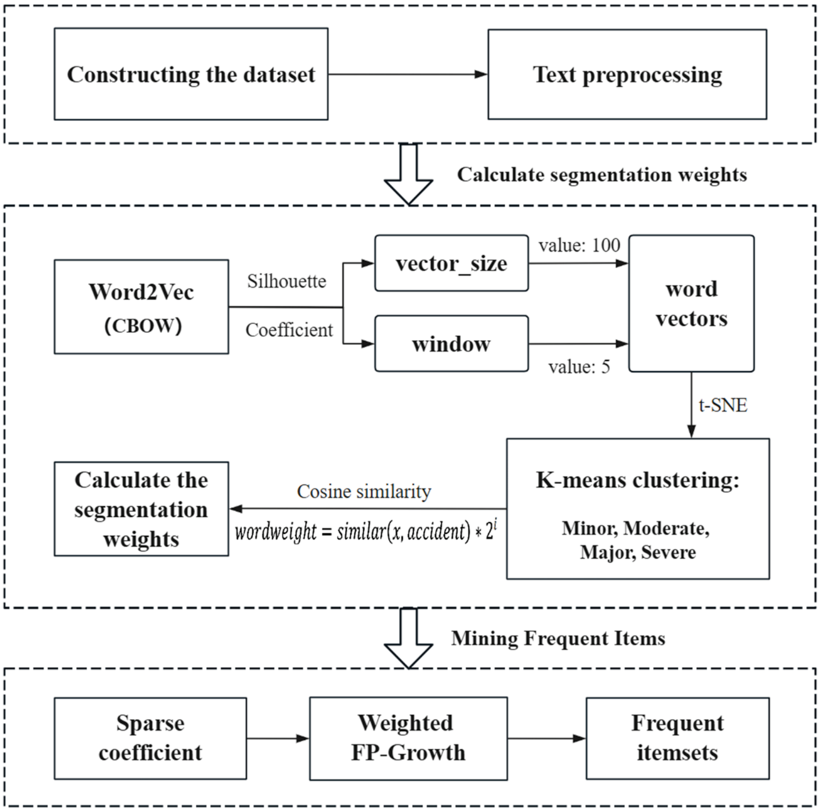


Figure 3. Process of mining power accident risk factors.

3.1. Design and Calculation of Word Segmentation Weights

3.1.1. Text Preprocessing

In order to convert the original text that does not conform to the input rules into the content that can be recognised by the weighted FP-Growth algorithm, it is necessary to carry out text preprocessing. Aimed at the original text corpus and feature item text description, this paper selects the jieba lexical library for the lexical processing of the text [27] and chooses the precise mode to ensure the accuracy of the lexical results to restore the semantic structure of the text to the maximum extent.

In this paper, we focus on the noise in text data to further improve the quality and efficiency of word splitting and take measures to deal with it accordingly. Firstly, an exclusive deactivation word list was constructed using the Chinese deactivation word list of Harvard University, and all deactivated words were eliminated in the process of lexical segmentation, which effectively avoided the interference of words with no practical significance with the thematic content of the text. Secondly, we cleaned up the punctuation marks, irrelevant symbols (e.g., emoticons), and formatting problems in the text to ensure that the text data are more standardised and to provide a clear basis for subsequent analysis. Through these methods, this paper effectively reduces the interference of noise in the analysis of text data, thus improving the accuracy and reliability of the analysis results and helping to focus on the important content of the text.

3.1.2. Design and Calculation of Weights

After the word segmentation process, this paper converts sentences in the original corpus into contextual sentences composed of segmented words and uses the segmented sentence lists to train the Word2Vec model. Word2Vec employs a shallow neural network to embed words into a low-dimensional vector space in order to capture the semantic information of the words. In small datasets, the vocabulary is usually limited and the contextual information of high-frequency words is relatively sufficient. Among the two architectures of the Word2Vec model, the CBOW model trains quickly on small datasets and makes better use of this contextual information. Since the dataset constructed in this paper is small, CBOW is chosen to train the Word2Vec model. The two important parameters of Word2Vec are vector_size and window, and the appropriate combination of these parameters is selected based on silhouette coefficient to generate word vectors.

In this paper, we use the t-SNE algorithm for dimensionality reduction of generated high dimensional word vectors [28]. This algorithm is a nonlinear technique for dimensionality reduction that keeps similar data similar in low dimensional space. Then, the dimensionality reduced word vectors are clustered using K-means clustering. Based on the above process, the cosine similarity is used to calculate the segmentation weights with the following formula:

$$wordweight = similar(x,\ accident) \times 2^i \tag{6}$$

where x denotes the current segmentation; $accident$ denotes the accident level; $similar(x,\ accident)$ denotes the cosine similarity between the current segmentation x and $accident$, and the current segmentation x belongs to the same clustering category as the accident level; and i is assigned different values based on $accident$ (when the $accident$ is Minor, Moderate, Major and Severe, the corresponding i is 0, 1, 2, and 3, respectively).

3.2. Text-Weighted FP-Growth Algorithm

Based on the design and calculation of segmentation weights, the improved FP-Growth algorithm presented in this paper primarily mines frequent terms by combining sparse coefficients to calculate the importance of accident records. At the same time, this paper introduces parallel computing and improves the relevant data structure during the execution of the algorithm to optimise the operation efficiency of the algorithm.

3.2.1. Sparse Coefficients

In this paper, 10 risk factors were mined as feature terms, and it was found that some feature terms with low participle weights in the accident dataset also responded to potential accident level correlations. These sparse feature terms appear with low frequency but may also have strong correlations with accidents. Therefore, this paper designs sparse coefficients to be added to the calculation of importance [29]. The calculation formula is as follows:

$$p_{avg} = \frac{1}{c}\sum p_i,\ p_i = \frac{n}{N},\ \alpha \le p_i \le \beta \tag{7}$$

$$coef = \frac{1}{1 - p_{avg}} \tag{8}$$

where p_i denotes the sparsity rate of the i-th feature term; n represents the number of missing accidents; N represents the total number of accidents; p_{avg} denotes the average sparsity of feature terms within the specified range α and β; and $coef$ denotes the sparse coefficient.

3.2.2. Definition of the Weighted FP-Growth Algorithm

We define the feature term set associated with accidents as $I = \{I_1,\ I_2,\ I_3, \ldots,\ I_n\}$, and the accident records set $R = \{R_1, R_2, R_3, \ldots, R_m\}$, where R_i consists of I_i. In the accident record R_i, each feature item contains multiple segmentation weights $wordweight$. The importance of the accident records (Record Importance, RI) is calculated based on the word weights and sparse coefficient, with the calculation formula is as follows:

$$RI = \frac{\sum wordweight}{n} \times coef \tag{9}$$

where n is the number of segmentation weights.

The weighted FP-Growth algorithm filters the frequent itemsets based on the weighted support thresholds ε, and the formula for calculating the weighted support of feature items is as follows:

$$Sup(I_i) = \frac{\sum_{j=0}^{n} \left(I_i \in R_j\right) \times RI_j}{\sum_{j=0}^{n} RI_j} \tag{10}$$

where $Sup(I_i)$ denotes the weight support of a feature term and the conditional expression $I_i \in R_j$ denotes whether the feature term I_i exists in the incident record R_j. If it exists, the conditional expression returns 1, otherwise it returns 0.

This algorithm calculates the weight support of the risk factor feature terms by scanning the power accident dataset and sorts them in descending order of weight support. Subsequently, the FP-Tree is constructed by each accident record feature term in the sorted order. Finally, the conditional FP-Tree is constructed recursively to mine the frequent termsets. The specific algorithm is described as follows (Algorithm 1):

Algorithm 1 Weighted FP-Growth Algorithm

Inputs: incident record set ***R****, minimum weight support thresholds* $\boldsymbol{\epsilon}$
Output: frequent item set ***FIS***
1. ***FP-Growth (R. ϵ)***
2. *Construct header* ***tableTable****: scan* ***R****, compute* $\boldsymbol{Sup}(I_i)$ *and construct header table in descending order*
3. *Construct* ***FP-tree****: scan* ***R*** *and initialise the root node of the tree as an empty node,* i.e., $\boldsymbol{Root} = \boldsymbol{null}$
4. ***While***(R_i):
5. *The* R_i *the feature items in* I_j *from* ***Root*** *into the* ***FP*** *tree in sorted order*
6. ***If(node ==*** $\boldsymbol{I_j}$***)****: Merge node weights together;*
7. ***else:*** *create new node, add node address to* ***Table***
8. *Iterate over* ***the Table****:*
9. *Constructed* I_i *The conditional pattern base of the* $\boldsymbol{R_c}$*: traverses the* $\boldsymbol{Table}(\boldsymbol{I_i})$ *indexes the corresponding tree nodes to construct the conditional schema base*
10. *Recursion:* $\boldsymbol{FIS_c}$ ***= FP-Growth (***$\boldsymbol{R_c}$***, ϵ)***
11. *Combined:* ***FIS = FIS***$\bigcup$ $\boldsymbol{FIS_c}$
12. *Return* ***FIS***

3.2.3. Optimisation of Operational Efficiency

Firstly, this paper enhances the algorithm through the implementation of hash tables. The hash table constructs the header table in Step 2, enabling rapid indexing and addition of node addresses during the FP-tree construction in Step 3. Secondly, the construction of the FP-tree necessitates ordering the feature items in the dataset by their weighted support in descending order. In building the conditional pattern base, a pointer link is established for each feature item node to the corresponding hash table based on its support value. This facilitates efficient ranking by directly comparing support values, thereby minimizing traversal overhead. Lastly, the recursive mining of frequent items is executed in parallel using Open Multi-Processing (OpenMP), which is a multi-threaded concurrent programming API that supports cross-platform shared-memory models. This approach significantly enhances operational efficiency by allowing parallel traversal of the header table for the recursive mining of frequent items.

4. Experimental Results and Discussions

4.1. Experimental Dataset

The power accident dataset presented in this paper comprises a corpus of power accidents within the Southern Power Grid, transcribed from the National Compendium of Power Accidents and Power Safety Events (2022) and documentation from a power supply bureau. The dataset organizes and classifies power accidents and safety events into categories such as personal injury and fatality incidents, equipment failures, and power safety events. Each accident is elaborated upon from several aspects, including a brief description, detailed accounts, root causes, identified issues, preventive measures, and corrective actions. In total, the dataset contains 373 records, encompassing both original text entries and extracted dimensions of accident causation with corresponding descriptions. An example of a training sample is provided, as shown in Table 1.

Table 1. Specific examples of the accident dataset.

Description of the Incident	Extractable Feature Text
(1) Brief description of the accident: While carrying out a comprehensive improvement of the 10 kV xx water line at the Linxx Power Bureau, which is under the xx Power Supply Bureau of the xx Power Grid, an operation and maintenance worker of the distribution network mistakenly climbed up to pole No. 06 of the 10 kV xx, which was electrically charged near the work site, and caused an electrocution accident, resulting in serious injuries. (2) How the accident happened: The XX Power Supply Bureau was carrying out a comprehensive improvement of the distribution network of the 10 kV xx line. The staff came from 4 power supply offices and several external construction units. The work was divided into two groups, involving a total of three work tickets. 07:00, 10 kV xx line 01 knife gate after the section of the line to repair. Without fully explaining safety matters, members of the first team (Qin X and others) signed off and went to the job site. Qin X and three others planned to replace porcelain bottles at poles 04–08 of 10 kV xx line. at 0956 h, after the rain abated, Qin X and others reached the vicinity of pole 06 of 10 kV xx, which was mistakenly thought to be the work site. Without verifying the identification of the pole, Qin X started to climb the pole. He was electrocuted and injured in the process of boarding the pole. at 10:25 a.m., Su x went up the pole to confirm that there was no voltage and then rescued Qin x and sent him to the hospital for treatment. (3) Causes of the accident: 1. Direct cause: The field worker had a weak sense of safety and mistakenly boarded the electrified No. 06 without checking the name of the line and the number of the pole tower or conducting a power test. 2. Indirect Causes: (a) The safety explanation was a mere formality. The general explanation given by the leader of the on-site coordination group, My, and the person in charge of the work ticket to all workers was not specific, and the leader of the subgroup did not give a safety explanation to the workers in the subgroup. (b) There was a vacuum in the on-site safety supervision service. The person in charge of the work did not explicitly set up a special-purpose supervisor in the work permit. (c) Wrong placement of materials and work equipment. The safety belt bag, electric tester pen, foot buckle, lanyard, etc. were placed in the wrong position, which became the cause of misleadingly climbing the tower by the working staff.	• Command error: the safety explanation is in form; • Operational error: mistakenly boarding pole No. 06 of the 10 kV xx line, which was running with electricity; • Supervision error: vacuum in site safety supervision duties; • Improper use of work equipment: wrong placement of materials and work equipment; • Accident level: general.

This paper specifically examines the correlations between the risk factor characteristics across ten dimensions and the four grades of accident occurrence: command error, operational error, supervisory error, insufficient knowledge and skills, inappropriate measures, improper use of equipment, inadequate regulations and procedures, ineffective dual prevention mechanisms for risks, insufficient risk management and control, and inadequate safety training. The four accident grades are categorized as Minor, Moderate, Major, and Severe.

4.2. Experiments on Word2Vec-Based Word Vector Generation

In this paper, we design experiments on accidental text using jieba participle and a deactivated word list to remove participle noise and generate word vectors based on Word2Vec model for the participle of the text after participle. Since vector_size and window are important parameters affecting the quality of word vectors, this experiment searches for the optimal combination by adjusting these two parameters.

Firstly, the parameter space is defined: the value range of vector_size is [60, 120], and the step size is 10; the value range of window is [3, 9], and the step size is 2. Then, all the combinations of parameters are traversed by using the grid search method, and the word vectors are obtained by training the Word2Vec model. The word vectors were dimensionality reduced and then K-means clustering was performed to classify the word vectors into four classes, representing the different accident levels. Each group of experiments is repeated five times; if any one of the clustering attempts fails, the statistical result is failure. The clustering results are shown in Table 2, where T indicates clustering success and F indicates clustering failure.

Table 2. Clustering results for different parameter combinations.

Window \ Vector_Size	60	70	80	90	100	110	120
3	F	F	F	F	F	F	F
5	F	F	F	T	T	F	F
7	F	F	F	F	F	F	F
9	F	F	F	F	F	F	F

From Table 2, it can be seen that when vector_size is 90 or 100 and window is 5, it can successfully cluster the word vectors into four classes. The two parameter combinations (90, 5) and (100, 5) were further evaluated using the silhouette coefficient S, calculated as follows:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{11}$$

where $a(i)$ is the average value of the degree of dissimilarity from vector i to other points within the same cluster and $b(i)$ is the minimum value of the average degree of dissimilarity from vector i to other clusters.

The results show that the profile coefficient of the combination (100, 5) is 9.76% higher than that of the parameter combination (90, 5), so the vector_size of 100 and window of 5 are selected as the final parameter settings in this experiment. Based on this parameter combination, the visualisation results of word vector clustering are shown in Figure 4, indicating that the word vectors in the power accident dataset have been successfully classified into four classes.

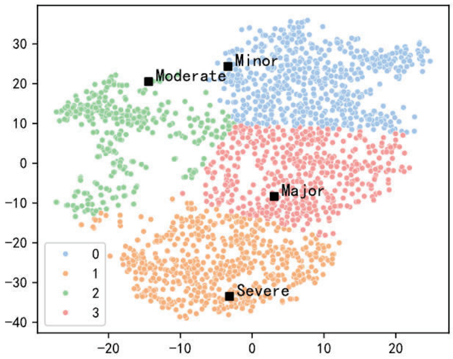


Figure 4. Segmentation clustering visualisation.

4.3. Experiments on the FP-Growth Algorithm Based on Text-Weighting

This algorithm is divided into two parts: text weights and sparse coefficients. First, the experiment explores the effect of text weighting on the algorithm results. In this experiment, in order to mine as many frequent itemsets as possible, the minimum support threshold is set to 0.01. The support degree of association rules shows the frequent association relationship between risk factors and accident levels, and the higher the support degree, the more frequent the rules appear. The change of support degree of each risk factor feature term before and after text-weighted is shown in Figure 5.

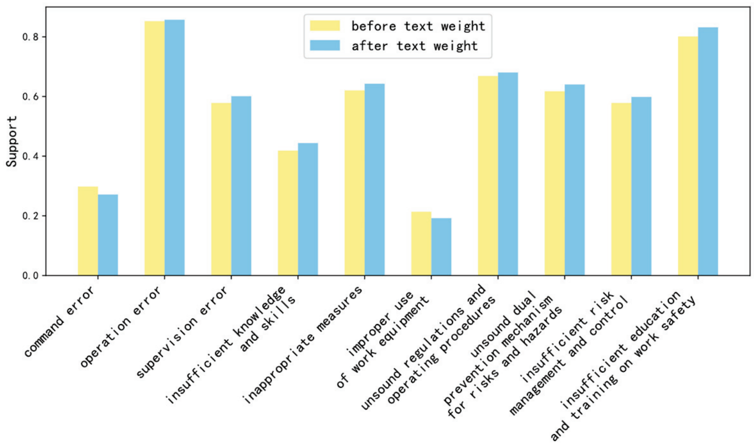


Figure 5. Change in support of feature items before and after text weight.

As illustrated in Figure 5, with the exception of the two feature items "command error" and "improper use of work equipment", the support levels of the remaining eight feature items have improved following text weighting. Notably, the support levels for "operation error" and "insufficient safety education and training" both exceed 0.8 before and after weighting. Among these, "insufficient safety education and training" experienced the most significant increase, rising by 3.75 percent. The support levels of the other six characteristics are primarily concentrated within the range of 0.4 to 0.7, showing a slight increase post-weighting.

Because the frequency of "command error" and "improper use of work equipment" is low in the dataset, the order of support is small, which leads to a further decrease in the weighted support, of which the weighted support for "improper use of work equipment" is only 0.19. If the support threshold is set to 0.2, the item cannot be mined into the relevant frequent items. For this reason, this paper introduces sparse coefficients and experimentally explores the effect of combining sparse coefficients on the algorithm results on the basis of text weights. The support changes of feature items before and after combining sparse coefficients are shown in Figure 6.

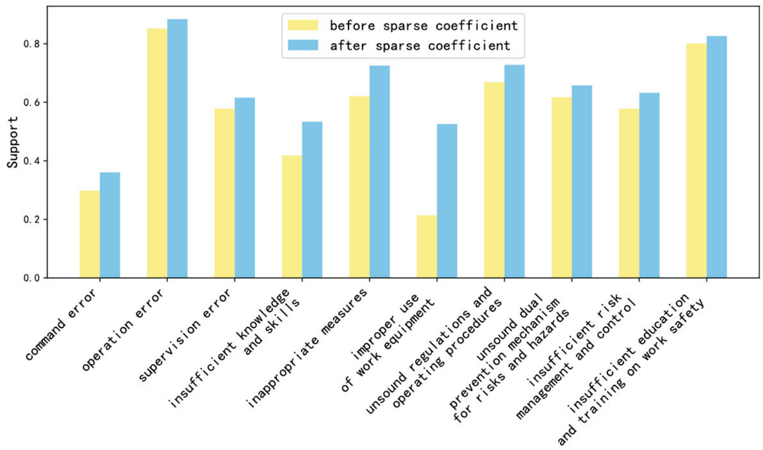


Figure 6. Change in support of feature terms before and after combining sparse coefficients.

As illustrated in Figure 6, the support for the sparse feature term "improper use of work tools" significantly improves after the incorporation of the sparse coefficients, increasing from 0.21 to 0.52. The support for the sparse feature term "command error" also shows a 20% improvement. This indicates that increasing the support threshold enhances the likelihood of mining relevant frequent items from sparse feature terms. Additionally, with the exception of the item "inadequate safety education and training," which exhibits no change in support before and after the application of the sparse coefficients, the support levels of all other feature items, when combined with sparse coefficients and text weighting, surpass those achieved through text weighting alone.

In order to verify the effectiveness of the proposed method, under the same dataset, this paper compares the original FP-Growth algorithm, the improved FP-Growth algorithm, and the Pearson correlation coefficient method [30]. The Pearson's correlation coefficient method is a classical method traditionally used to measure the linear relationship of variables and is widely used in correlation analysis. The results of the degree of correlation between the 10 risk factors and the accident level are shown in Figure 7.

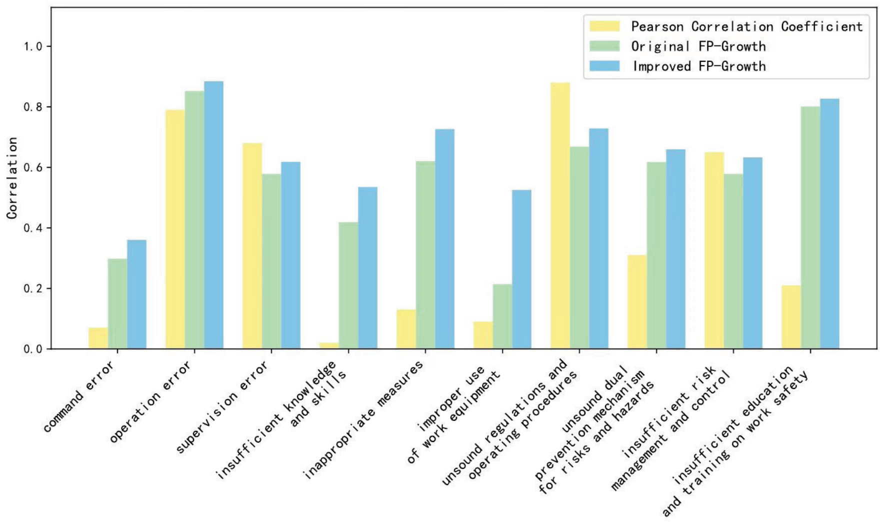


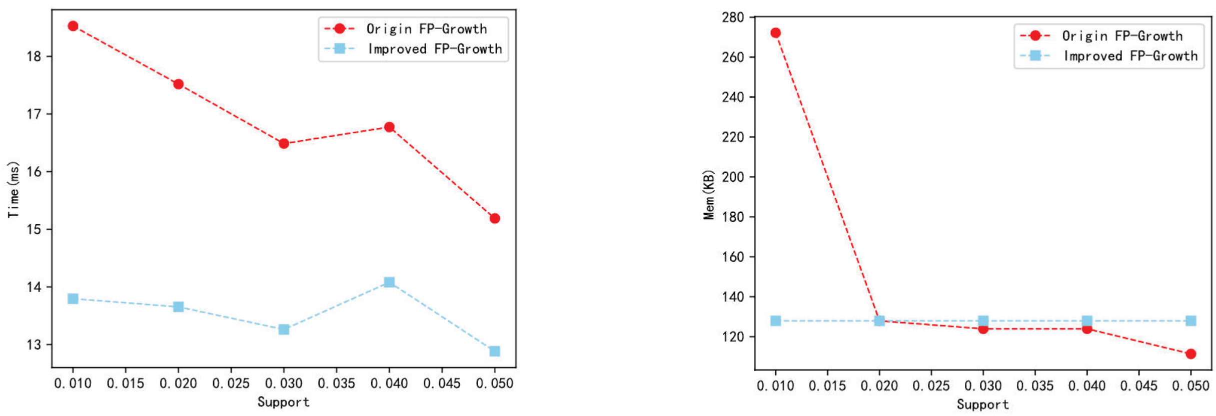
Figure 7. Comparison of the degree of association of the three methods.

Figure 7 shows that the Pearson correlation coefficient method associations obtained in the three risk factors "supervision error", "inadequate regulations and operating procedures", and "inadequate risk control" are similar to the other two algorithms. However, the association degree of the four risk factors "command error", "insufficient knowledge and skills", "improper measures", and "improper use of work equipment" is seriously insufficient and lower than 0.2. This is because the method assumes a linear relationship between the variables and considers only numerical data, ignoring the complex non-linear relationship that may exist between risk factors and accident levels and the influence of semantic information.

The FP-Growth algorithm based on text weighting proposed in this paper is not only able to deal with complex text data but also gives different weights to different risk factors through text weighting, which effectively avoids the insufficiency of Pearson's correlation coefficient method in dealing with sparse data and nonlinear relationships. Therefore, using this algorithm can more comprehensively mine the potential risk factors, make the association rule mining more in line with the reality, and provide more accurate risk decision support for the power accident early warning.

4.4. Algorithm Running Efficiency Analysis

This experiment is executed on an AMD Ryzen 7 5800H CPU purchased from China, with 32 GB of RAM and Windows 11 operating system. The core implementation of the algorithm is based on C++17. The running time of the algorithm under different support thresholds is shown in (a) of Figure 8, and the memory usage is shown in (b) of Figure 8.



(**a**) (**b**)

Figure 8. (**a**) Comparison of running time and (**b**) comparison of memory usage under different support thresholds.

From Figure 8, it can be seen that the improved FP-Growth algorithm proposed in this paper has less running time than the original FP-Growth algorithm at different support thresholds and possesses higher running efficiency. Meanwhile, this method occupies less memory when the support threshold is lower, i.e., mining more frequent items. The improved FP-Growth algorithm saves time and space costs and verifies the efficiency of the method.

4.5. Analysis of the Correlation Between Accident Risk Levels

After the FP-Growth algorithm based on text weights and sparse coefficients obtains the frequent itemset and their support, the resultant statistics on the frequent 2-itemset are performed in order to derive the degree of association of each accident risk feature with different accident classes. The degree of association of different risk features with different accident classes is shown in Figure 9. As can be seen from Figure 9, the degree of association between the moderate accident class and each risk feature is relatively high, which is due to the high percentage of accidents in the moderate class and the low percentage of accidents in the remaining three accident classes, which is in line with the objective law. Among them, operation error has the highest degree of association with the moderate accident class, indicating that both have the highest weight support after weighting in the whole dataset, which demonstrates the importance and high frequency of operation error behaviour in the characteristics of accidents. As for the severe accident class, eight dimensions have a similar degree of association, indicating that the occurrence of severe accidents is often inextricably linked to all of these factors.

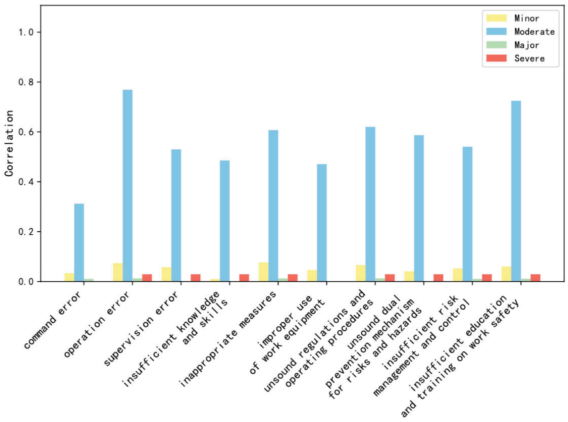


Figure 9. Degree of correlation between different feature terms and accident levels.

By adding up the degree of association between each risk feature and different accident levels in Figure 9 and arranging them in descending order, a comprehensive ranking of risk features can be obtained, which reveals the comprehensive degree of association between different risk features and accident levels, as shown in Figure 10. As shown in Figure 10, the feature item "operation error" has the greatest impact on power accidents, with an association degree of 0.88, followed by the feature item "insufficient safety education and training", with an association degree of 0.83. The lowest impact of "command error" is less than 0.4 (0.36), and the impact of the other seven risk factors on power accidents is within the range of 0.4–0.8. The experimental results show that the FP-Growth algorithm based on text weight and sparse coefficients proposed in this paper effectively exploits the risk factors of power accidents and analyses the degree of association of each risk factor on accidents.

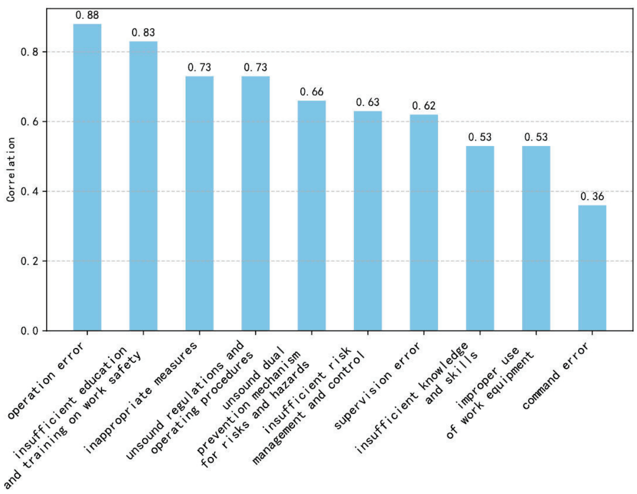


Figure 10. Comprehensive ranking of risk features.

4.6. Association Rule Mining Based on the Improved FP-Growth Algorithm

The confidence level of the association rule reveals the causal relationship between the risk factors and the accident level; the higher the confidence level, the greater the possibility that the former term will trigger the latter term. To achieve the early warning of power accidents, the causal relationship between the former and latter items can be analysed in depth to find the risk factors that lead to the strength of the risk level. Some of the association rules mined by this algorithm are shown in Table 3, which are sorted according to the level of confidence.

Table 3. Confidence level of the association rule between feature items and accident level.

Antecedent	Consequent	Confidence
Insufficient knowledge and skills	moderate accident	0.90
Unsound dual prevention mechanism for risks and hazards	moderate accident	0.89
Improper use of work equipment	moderate accident	0.89

In the historical accident data, the correlation degree between these risk factors is very high; by analysing the correlation rules in Table 3, it can be seen that "insufficient knowledge and skills" leads to a higher probability of moderate accidents, so the relevant electric power industry should focus on the knowledge and skills of the staff training. Similarly, "unsound dual prevention mechanism for risks and hazards" and "improper use of work equipment" also lead to a higher probability of moderate accidents, and the electric power companies should pay attention to the dual prevention mechanism of risks and hazards and the use of work equipment by employees. According to the mined association rules, the algorithm can provide effective risk decision support for the early warning of electric power accidents.

5. Conclusions

The amount of data on power accidents has risen dramatically, posing unprecedented challenges for risk assessment and safety improvement of power accidents. Accidents in smart energy systems usually involve multidimensional factors and it is difficult to capture these complex risk factors directly in unstructured textual data. Aiming at the textual description of accident risk features, this paper firstly performs word segmentation on the independently constructed datasets about electric power accidents, followed by the generation of word vectors through the Word2Vec model. The K-means clustering algorithm is employed to classify the accident records into four categories corresponding to different accident levels, following dimensionality reduction of the word vectors. Cosine similarity is then utilized to calculate text weights, leading to the definition of the FP-Growth algorithm based on text-weighting. Building on this, an improved FP-Growth algorithm explores the associations between various accident risk features and risk classes, with an emphasis on sparse feature terms. Additionally, the mining speed of frequent terms is enhanced by optimizing certain structural and procedural aspects of the FP-Growth algorithm, alongside the implementation of parallel processing strategies. Finally, the effectiveness of this algorithm is demonstrated through experimental results, providing risk decision support for the early warning of electric power personal accidents.

This method is applicable within the domain of electric power accidents, facilitating the analysis of correlations between different feature dimensions and accident levels. Safety managers are able to gain a deeper understanding of the key risk factors that influence the outcome of an incident, thus supporting faster, more accurate decision-making and aiding in the development of public safety early warning systems.

Author Contributions: Conceptualization, R.L. and J.Z.; methodology, R.L.; software, F.D.; validation, J.Z.; formal analysis, R.L.; investigation, F.D.; resources, J.Z.; data curation, R.L.; writing—original draft preparation, F.D.; writing—review and editing, R.L.; visualization, J.Z.; supervision, R.L.; project

administration, F.D.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 52377103; the General Program of National Natural Science Foundation of China, grant number 52277148; and Southern Power Grid Corporation Technology Project Funding, grant number 031300KK52222091.

Data Availability Statement: The original contributions presented in the study are included in the article and further inquiries can be directed to the corresponding author.

Conflicts of Interest: R.L. and J.Z. were employed by Guangdong Power Grid Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare that this study received funding from Southern Power Grid Corporation. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

References

1. Zhang, C.Y. Network-based Infrastructure as Media: Energy, Transportation and Information(ETI) Networks Integration. *Glob. J. Media Stud.* **2023**, *10*, 56–70. [CrossRef]
2. Zhang, K. Research on Key Technologies and Applications of Intelligent Operation and Inspection of Power Grid. Ph.D. Thesis, University of Science and Technology of China, Hefei, China, 2023.
3. Tang, G.F.; Zhou, J.; Pang, H.; Lin, J.J.; Fan, Z.; Wu, Y.N.; He, Z.Y.; Ma, S.C.; Xue, F.; Zhou, B.R. Strategic Framework for New Electric Power System Development under the Energy Security Pattern. *Strateg. Study CAE* **2023**, *25*, 79–88.
4. Yan, Y.Q.; Zhang, S.; Liang, Z.X.; Sheng, W. Statistics and Analysis of Electric Power Enterprises Personal Accidents in China During 2016–2021. *Saf. Secur.* **2023**, *44*, 46–51.
5. Gholami, A.; Srivastava, A.K. ORCA: Outage Root Cause Analysis in DER-Rich Power Distribution System Using Data Fusion, Hierarchical Clustering and FP-Growth Rule Mining. *IEEE Trans. Smart Grid* **2024**, *15*, 667–676. [CrossRef]
6. Wu, Y.; Du, X.S.; Fu, Y.H. Stable Operation of Power Grid Based on Association Rule Mining Method of Energy Big Data. *Opt. Optoelectron. Technol.* **2022**, *20*, 139–144.
7. Wang, X.H.; Liu, X.X.; Zhong, F.C.; Li, Z.L.; Xuan, K.G.; Zhao, Z.L. A Scenario Generation Method for Typical Operations of Power Systems with PV Integration Considering Weather Factors. *Sustainability* **2023**, *15*, 15007. [CrossRef]
8. Zhou, X.; Ma, L.N.; Ma, X.G.; Hao, Q.; Jia, H.Y.; Liu, H.; Bai, W.X.; Zhang, H.Z.; Li, S.J.; Yang, Q.F. Research on text analysis of hidden dangers of coal mine accidents based on link prediction. *J. Saf. Sci. Technol.* **2024**, *20*, 26–34.
9. Shen, J.H.; Liu, S.P.; Zhang, J. Using Text Mining and Bayesian Network to Identify Key Risk Factors for Safety Accidents in Metro Construction. *J. Constr. Eng. Manag.* **2024**, *150*, 04024052. [CrossRef]
10. Liu, J.F.; Ma, H.Z.; Xie, X.L.; Cheng, J. Short Text Classification for Faults Information of Secondary Equipment Based on Convolutional Neural Networks. *Energies* **2022**, *15*, 2400. [CrossRef]
11. Ruan, G.C.; Xie, F.; Tu, S.W. Application Research Based on Word2vec Diversity in Library Recommender System. *Libr. J.* **2020**, *39*, 124–132.
12. Wang, H.W.; Yin, P.; Duan, Z.T.; Li, Y. Research on power marketing data mining and clustering techniques based on Bert and k-meas. In Proceedings of the International Conference on Power Electronics and Artificial Intelligence (PEAI), Xiamen, China, 19 January 2024.
13. Kasihmuddin, M.S.M.; Jamaludin, S.Z.M.; Mansor, M.A.; Wahab, H.A.; Ghadzi, S.M.S. Supervised Learning Perspective in Logic Mining. *Mathematics* **2022**, *10*, 915. [CrossRef]
14. Gan, W.S.; Lin, J.C.W.; Yu, P.S. A Survey of Utility-Oriented Pattern Mining. *IEEE Trans. Knowl. Data Eng.* **2021**, *33*, 1306–1327. [CrossRef]
15. Diaz-Garcia, J.A.; Ruiz, M.D.; Martin-Bautista, M.J. A survey on the use of association rules mining techniques in textual social media. *Artif. Intell. Rev.* **2023**, *56*, 1175–1200. [CrossRef] [PubMed]
16. Agrawal, R.; Tomasz, I.; Arun, S. Mining association rules between sets of items in large databases. *Proc. 1993 ACM SIGMOD Int. Conf. Manag. Data* **1993**, *22*, 207–216. [CrossRef]
17. Liu, W.W.; Wang, H.W.; Hou, Z.G. Analysis on risk association rules of civil aviation aircraft maintenance based on BL-Apriori. *J. Saf. Sci. Technol.* **2024**, *20*, 27–33.
18. Han, J.; Pei, J. Mining frequent patterns without candidate generation. *ACM* **2000**, *29*, 1–12.

19. Xiao, Y.; Meng, L.H.; Zhang, Y.R.; Gu, Y.T. An Improved FP-Growth Algorithm with Time Decay Factor and Element Attention Weight. In Proceedings of the 2024 IEEE 4th International Conference on Power, Electronics and Computer Applications (ICPECA), Shenyang, China, 5 August 2024.
20. Yu, J.B.; Ma, X.Y.; Zhao, Z.Y. Research on association analysis of food risk factors based on the improved FP-growth algorithm. *Food Sci.* 2024; *in press*.
21. Lei, X.; Cheng, G.; Zhang, Y.J.; Guo, L.; Zhang, F.C. Association analysis of alarm information based on power network situation awareness platform. *Comput. Eng. Sci.* **2023**, *45*, 1197–1208.
22. Wu, D.P.; Hua, G. Text Case Classification of Safety Production Accidents Based on Word2Vec Word Embedding and Clustering Model. *Comput. Syst. Appl.* **2021**, *30*, 141–145.
23. Zhong, G.F.; Pang, X.W.; Sui, D. Text Classification Method Based on Word2Vec and AlexNet-2 with Improve Attention Mechanism. *Comput. Sci.* **2024**, *49*, 288–293.
24. Yan, F.X.; Wang, J.H. Sentiment recognition model of Weibo comments based on SVM and Word2vec. *Mod. Comput.* **2024**, *30*, 60–64.
25. Lv, J.; Qiu, X.L. A noisy label deep learning algorithm based on K-means clustering and feature space augmentation. *CAAI Trans. Intell. Syst.* **2024**, *19*, 267–277.
26. Öztürk, M.M. A cosine similarity-based labeling technique for vulnerability type detection using source codes. *Comput. Secur.* **2024**, *146*, 104059. [CrossRef]
27. Gao, P.; Li, F.; Peng, Y.H.; Zhang, C.H.; Peng, H.J. Accurate Classification Method of Power Customers Based on Jieba Chinese Word Segmentation. *Hunan Electr. Power* **2023**, *43*, 151–154.
28. Yagahara, A.; Uesugi, M.; Yokoi, H. t-SNE Visualization of Vector Pairs of Similar and Dissimilar Definition Sentences Created by Word2vec and Doc2vec in Japanese Medical Device Adverse Event Terminology. *Stud. Health Technol. Inform.* **2022**, *290*, 1058–1059. [PubMed]
29. Li, Y.H.; Hu, L.; Gao, W.F. Multi-Label Feature Selection Based on Sparse Coefficient Matrix Reconstruction. *Chin. J. Comput.* **2022**, *45*, 1827–1841.
30. Li, N.; Yang, F.; Wu, H.B.; Zhang, Y.X.; Yin, L. Study on the correlation between dual-control policies on energy consumption and energy use efficiency in energy-consuming industries. *High Volt. Eng.* **2023**, *49*, 215–220.

energies

MDPI

Article

Distributed Photovoltaic Short-Term Power Prediction Based on Personalized Federated Multi-Task Learning

Wenxiang Luo, Yang Shen, Zewen Li and Fangming Deng *

School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330032, China; 15297890421@139.com (W.L.); baoquan203@163.com (Y.S.); 3386@ecjtu.edu.cn (Z.L.)
* Correspondence: 2464@ecjtu.edu.cn

Abstract: In a distributed photovoltaic system, photovoltaic data are affected by heterogeneity, which leads to the problems of low adaptability and poor accuracy of photovoltaic power prediction models. This paper proposes a distributed photovoltaic power prediction scheme based on Personalized Federated Multi-Task Learning (PFL). The federal learning framework is used to enhance the privacy of photovoltaic data and improve the model's performance in a distributed environment. A multi-task module is added to PFL to solve the problem that an FL single global model cannot improve the prediction accuracy of all photovoltaic power stations. A cbam-itcn prediction algorithm was designed. By improving the parallel pooling structure of a time series convolution network (TCN), an improved time series convolution network (iTCN) prediction model was established, and the channel attention mechanism CBAMANet was added to highlight the key meteorological characteristics' information and improve the feature extraction ability of time series data in photovoltaic power prediction. The experimental analysis shows that CBAM-iTCN is 45.06% and 42.16% lower than a traditional LSTM, Mae, and RMSE. Compared with FL, the MAPE of the PFL proposed in this paper is reduced by 9.79%, and for photovoltaic power plants with large data feature deviation, the MAPE experiences an 18.07% reduction.

Keywords: personalized federated learning; multi-task; deep learning; photovoltaic power prediction

1. Introduction

With the advent of the "dual carbon" targets, renewable energy has become a primary direction for future international energy development [1]. In this context, the installed capacity of distributed photovoltaic (DPV) systems has surged, making accurate photovoltaic power forecasting significantly important for optimizing grid scheduling and facilitating the high-proportion grid integration of renewable energy sources. However, DPV systems are widespread, and DPV power data are affected by geographical location differences, photovoltaic array sizes, and local meteorological conditions, leading to non-uniform distribution and complex correlations of the data, exhibiting significant non-IID characteristics. This significantly challenges the improvement of adaptability in photovoltaic prediction models within distributed environments.

With the advancement of artificial intelligence research, deep neural networks have been widely applied in photovoltaic power generation prediction [2]. The photovoltaic power station uploads the data to the server, and the server uses the data to train the prediction model to predict the photovoltaic power performance in the future [3]. At present, the deep learning centralized training model has been fully developed. However, the most widely used DPV systems are characterized by a dispersed installation of photovoltaic stations and the influence of geographical and meteorological conditions, causing temporal and spatial variations in photovoltaic data and resulting in heterogeneous data [4,5]. Therefore, it is impossible to establish a model with strong generalization ability and high prediction accuracy in the DPV system.

Federated learning is a novel distributed model-building framework. Different from centralized training, FL adopts a decentralized structure to retain data on local devices and jointly develop a global model through collaboration [6,7]. In the context of DPV systems, FL enables secure information sharing and knowledge distillation among DPV stations and can also enhance model performance in a distributed environment. However, during the training process of FL, influenced by the non-IID nature of photovoltaic (PV) data, traditional FL exhibits poor convergence and tends to bias model features toward PV stations with abundant data, resulting in varied performance of the global model among different PV stations [8]. Consequently, the conventional FL strategy, which involves training a solitary global forecasting model, proves insufficient to fulfill the distinct predictive demands of each individual PV station. Alternatively, the traditional FL methodology of developing a single, all-encompassing prediction model fails to address the specific prediction needs of every PV station adequately.

In this paper, a personalized federated learning approach is introduced for the prediction of DPV power. The innovations presented herein are as follows:

1. This paper introduces a framework for DPV power prediction, which is grounded in PFL. In this framework, local prediction models are trained at individual PV stations, and the model parameters are updated using federated learning algorithms on a cloud server. This enables secure information sharing and knowledge distillation among multiple PV stations.
2. In the process of Personalized Federated Multi-Task Learning (PFL), a multi-task training method is proposed. Clients have the capability to maintain the confidentiality of the Weight Normalization (WN) layer and develop distinct personalized models tailored to their unique data characteristics. This addresses the limitation of conventional federated learning (FL), wherein a unified global model fails to enhance prediction accuracy across all PV stations.
3. An improved CBAM-iTCN PV power prediction model is proposed. The parallel pooling structure of the TCN is modified, and the attention mechanism CBAM net is added. By assigning different weights to different features extracted by the hidden layers of the TCN, key meteorological feature information is emphasized, enhancing the feature extraction capability for time series data in PV power prediction.

The remainder of this paper is structured as follows. In Section 2, a comprehensive summary and review of the relevant literature in the field addressed by this article is provided. Section 3 presents the proposed methodology outlined in this article. Section 4 introduces the implementation process of the PFL algorithm. Section 5 introduces the CBAM iTCN prediction model. Section 6 validates the effectiveness of the proposed method. Section 7 summarizes the advantages and future development directions of this article's plan.

2. Related Works

2.1. Photovoltaic Power Prediction Based on Deep Learning

Traditional prediction methods include statistical methods, physical models, and traditional machine learning methods. Statistical methods include time series analysis based on historical power data (such as ARIMA and SARIMA) and regression model prediction using statistical methods such as linear regression and multiple regression. Its advantage is low computational complexity, which is suitable for short-term forecasting, but it has difficulty dealing with nonlinear relations and cannot adapt to complex weather. Physical models include simulation prediction and indirect prediction methods based on physical formulas, which have the advantages of clear physical meaning and can be customized with equipment parameters to expand the prediction range but rely too much on high-precision data, resulting in large errors in practical applications. Traditional machine learning methods include support vector machines, random forest, gradient lifting, etc. Its advantages are that it can adapt to some nonlinear problems through function selection and has a strong anti-noise ability, but it has difficulty dealing with large-scale data and consumes a lot of model storage and calculation resources.

Deep learning (DL) can establish the prediction model between meteorological factors and photovoltaic power. Because of its strong nonlinear mapping ability, DL has been widely used in this field. While convolutional neural networks (CNNs) excel at processing the spatial correlation features present in weather image data for photovoltaic power forecasting [9,10], they are not proficient at capturing the long-range dependencies inherent in time series data. Therefore, the CNN is generally combined with Informer, Gating Unit (Gru), and other networks in photovoltaic power prediction [11–13]. The Cyclic Neural Network (RNN) has the ability to capture time dependence. The long-term and short-term memory network (LSTM), BiLSTM, and other networks [14–16] have good performance in photovoltaic power prediction; however, in the context of long-term series prediction, the challenge lies in avoiding gradient vanishing and effectively learning the long-range dependencies. Furthermore, researchers have also suggested that the integrated prediction methods, combined with a variety of neural networks, such as CNN-LSTM, GBDT-BiLSTM, and other integrated models [17–19], can realize photovoltaic power prediction; nevertheless, the computational expense is significant, leading to prolonged model training durations. The above paper aims at centralized photovoltaic power prediction, but the DPV system is more widely used.

The power prediction of the DPV system is affected by geographical diversity, data dispersion, as well as non-IID data, which impacts the performance of the prediction model, which faces challenges. Cheng et al. [20,21] proposed photovoltaic power prediction based on satellite image data sources. Through cloud image processing, short-term photovoltaic power generation prediction can be achieved, but processing satellite cloud images requires a lot of computing resources. Asiri et al. [22] introduced a prediction method for distributed photovoltaic power generation at the regional scale by dividing the region into different clusters and selecting a representative site in each cluster to realize photovoltaic power prediction. Some scholars use the idea of clustering to achieve photovoltaic power prediction, such as SOM [23] and K-means [24]. However, clustering only classifies the data samples without considering the learning ability of the model.

The above photovoltaic power prediction method usually adopts centralized model training and enhances the prediction accuracy solely by refining the algorithm's architecture but does not take into account the differences in data. However, in a distributed environment, due to the scattered installation of photovoltaic power stations, the PV data are affected by factors such as the size, location, and meteorological conditions of PV arrays, and the distribution of photovoltaic power data in time and space is heterogeneous. Therefore, it is impossible to establish a model with strong generalization ability and high prediction accuracy in the DPV system.

2.2. Federated Learning

Since federal learning was proposed, its data privacy protection and other advantages have attracted the attention of many scholars. McMahan et al. [25,26] proposed a federal averaging (FedAvg) method. The model parameters or gradients are uploaded to the cloud server, and the cloud server distributes the weights of all clients equally. At the same time, researchers [27] optimized and improved federated learning from the perspective of client selection, communication, and security. However, the above methods have a poor convergence effect on highly heterogeneous data.

To improve the convergence of FL, scholars have tried to personalize the local devices and models. Prevalent methods for personalization encompass Federated Transfer Learning (FTL), Federated Meta-Learning (FML), and Adaptive Federated Learning (AFL), among others. Zhang et al. [28] used FTL to train devices with limited computational capabilities, thereby safeguarding data privacy and significantly enhancing the efficiency and adaptability of model training. Deng et al. [29] introduced an AFL framework, adding an adaptive algorithm to the original FL algorithm to improve the training efficiency and reduce the communication cost. Chen et al. [30] introduced a dynamic approach to FML for the purpose of dynamically forecasting small data samples. These personalization strategies are applied to customize the global model. Initially, a global model is constructed, followed by client-specific customization. However, such methods tend to prolong the model training duration and pose challenges in enhancing the accuracy of the global model when dealing with non-independent and identically distributed (non-IID) data. Therefore, the existing federal learning strategies cannot be directly applied to DPV power prediction [31].

In the current research on DPV power prediction, due to the dispersion of the DPV system, the PV array size and the meteorological environment of PV power stations in different locations are different, resulting in non-IID PV data, and it is difficult to train a prediction model with strong adaptability. Therefore, this paper proposes the PFL of multi-task strategies to realize the regional collaborative training prediction model and improve the adaptability of the prediction model.

3. Overall Framework

To validate the resilience and generalization potential of the OGGWO algorithm in practical implementations, this investigation establishes two distinct experimental frameworks. The first case study examines multi-objective optimization in integrated energy management systems incorporating carbon–economic constraints, while the subsequent evaluation focuses on renewable energy penetration maximization under grid-stability prerequisites.

Consider a DPV system within a specified region. The comprehensive framework is illustrated in Figure 1. This system comprises N decentralized PV stations and a cloud server; these power stations are distributed in different geographical locations, such as roofs, mountains and open areas; every power station is furnished with a photovoltaic power prediction model and has local historical power generation data and meteorological data. In the PFL prediction framework proposed in this paper, firstly, each photovoltaic power station uses its own data to train the prediction model and upload the trained parameter gradient to the cloud server; the cloud server uses the PFL algorithm to aggregate and generate a global model to develop the collaborative training framework for power prediction models across each DPV power station, while ensuring the privacy protection of photovoltaic data; the multi task strategy is added to the PFL to make each photovoltaic power station generate a personalized model separately, solve the problem of the FL single global model being inadequate for accurately predicting the performance of all power stations, and improve the prediction accuracy and adaptability of the photovoltaic power model. The PFL prediction framework presented in this paper, illustrated in Figure 2, is primarily structured into the subsequent steps:

1. For each photovoltaic power station, the local historical meteorological data and photovoltaic power data are preprocessed, including filling in missing data, abnormal data, and normalization.
2. In response to task requests from the cloud server, each photovoltaic power station receives the global model parameters and utilizes its local historical photovoltaic and meteorological data to train the prediction model.
3. Upon the completion of training, each photovoltaic power station will transmit its model parameters to the cloud server and retain its personalized update for the current round. The ECS then collects the received model parameters and updates the global model accordingly. If the accuracy threshold is satisfied, the FL training process is concluded, resulting in the final global model. The last round of private patches is pushed back to a personalized model to improve the prediction accuracy and model adaptability of photovoltaic power plants.
4. If the accuracy of the aggregated global prediction model does not meet the standard, the global model will be re-distributed to each photovoltaic power station, and steps 2 and 3 will be repeated until it meets the standard.

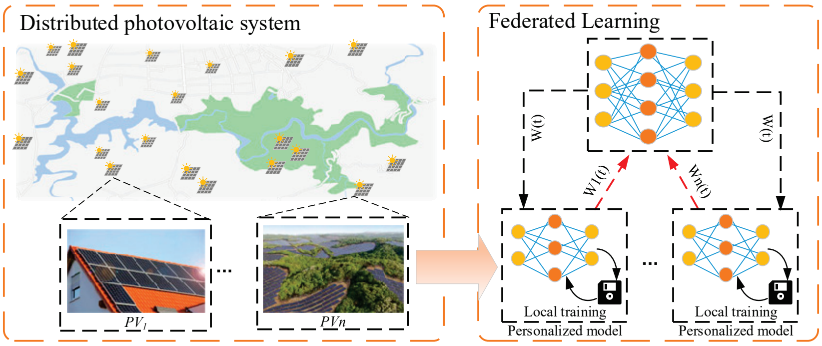


Figure 1. Overall Framework.

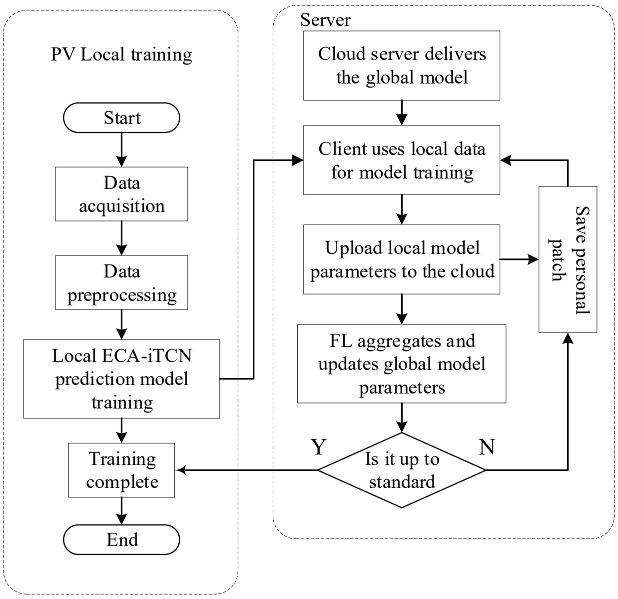


Figure 2. Flowchart of prediction system.

In the PFL prediction framework proposed in this paper, the edge server integrates the global model with the individual patch and trains the personalized model to optimize its applicability for the predictive endeavor of photovoltaic power plants. In the PFL process, the multi-task collaborative training mode provides a structured solution for DPV power prediction.

4. Personalized Federated Learning

4.1. Federated Learning

The introduction of federal learning disrupts the existing paradigm of data silos. The specific framework is shown in Figure 3. During the training process, all photovoltaic power station data are not interconnected, effectively ensuring the privacy of power data.

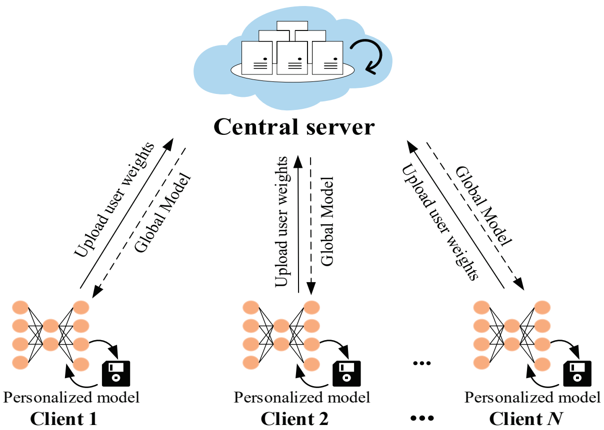


Figure 3. Personalized Federated Learning.

Federal learning represents, in essence, a distinct variant of distributed learning whereby each PV facility contributes model parameters to collaboratively enhance a global model. Owing to the non-IID characteristics of photovoltaic data, the localized prediction models developed by individual PV power stations exhibit significant divergence. As a result, the global model will exhibit a tendency toward the attributes of the photovoltaic power station during aggregation with a large amount of data in terms of model characteristics, resulting in poor user accuracy (UA) of some photovoltaic power stations. In this paper, the PFL with a multi-task strategy is used to avoid the impact of non-IID photovoltaic data on the performance of the model by training the personalized model locally.

4.2. Personalized Federated Learning

The Federal Average (FedAvg) algorithm stands as the foundational algorithm in FL. During the execution of FedAvg, each client performs one or multiple local parameter updates, as illustrated by Formula (1). Herein, η denotes the learning rate and (assume the original symbol for the local gradient was omitted; use $\nabla \mathrm{F_i}$ to represent it) signifies the local gradient.

$$\omega_i^t = \omega_i^{t-1} - \eta \nabla F_i(\omega) \tag{1}$$

In contrast to the gradient averaging algorithm, the FedAvg approach allows for multiple gradient descent iterations based on the computational capacity and communication latency of individual clients. Subsequently, the local model ω_i^t is updated and uploaded to the central server, where the next iteration of the global model is computed using Formula (2).

$$\omega^t = \sum_{i=1}^{N} \frac{n_i}{n} \omega_i^t \tag{2}$$

N represents the total data of all participants, the number of FL clients by n, and the data set of the ith client by n_i. The central server's objective function is formulated as Formula (3).

$$F_i = \min_{i \in R^k} \left\{ \sum_{i=1}^{N} \frac{n_i}{n} F_i(\omega) \right\} \tag{3}$$

But in the PFL of this paper, during each round of model updating, the client will update the personalized model using local data in addition to training the global model. Consequently, the loss function in PFL comprises both the local and personalized models and is represented by Formula (4).

$$f_i(\theta_i) + \frac{\lambda}{2}\|\theta_i - \omega\|^2 \tag{4}$$

θ_i represents the personalized model of client training, whereas ω represents the global model from the preceding training round. If it is in the first round of training, ω represents the initial model issued by the ECS. λ indicates the impact intensity of the global model on the personalized model λ. The size of it is directly proportional to the client's computing power and the extent of data heterogeneity. The PFL objective function can be rewritten into Formula (5) according to Formula (3).

$$F_i(\omega) = \min_{\theta_i \in R^d}\left\{f_i(\theta_i) + \frac{\lambda}{2}\|\theta_i - \omega\|^2\right\} \min_{\omega \in R^d}\left\{F(\omega) = \frac{1}{N}\sum_{i=1}^{N} F_i(\omega)\right\} \tag{5}$$

In this study, we employ the WN layer of the convolutional neural network as the patching layer to tailor the fl to individual needs. There are two reasons for choosing the WN layer as private: (1) the parameter storage cost of the WN layer is low; (2) the local model parameters obtained from FL training exhibit characteristics akin to those of the data in the WN layer. Consequently, the client has the option to store these parameters within the WN layer, enabling their utilization in the subsequent round of federated training.

Initially, it is necessary to apply whitening to the input data of the WN layer, as illustrated by Formulas (6) and (7).

$$\mu^t = \frac{1}{H}\sum_{i=1}^{H} x_i^t \tag{6}$$

$$\sigma^t = \sqrt{\frac{1}{H}\sum_{i=1}^{H} (x_i^t - \mu^t)^2} \tag{7}$$

μ^t and σ^t represent the mean and variance, respectively, of the neuron x_i. Here, H denotes the number of nodes in a hidden layer, and I indicates the total count of MLP layers. The weighted moving average values for μ^t and σ^t, as recorded by the WN layer during training, can be directly utilized in the inference process to expedite the model's reasoning time. By applying transformations to these two parameters, we can obtain data that follow a standard distribution with a mean of 0 and a variance of 1. Consequently, the input formulation for the preprocessed WN layer is presented in Formula (8).

$$\hat{x}_i = \frac{x^t - \mu^t}{\sqrt{(\sigma^t)^2 + \tau}} \tag{8}$$

τ is the minimum value introduced to prevent invalid calculation when the variance is 0.

In the convolutional neural network employed in this study, the relationship between the input and the output of the WN layer is not a straightforward sequential one. For any given node at time t, its input comprises a combination of the hidden layer state at time

$t-1$, denoted as h^{t-1}, and the input data at time $t-1$, represented as c_t. Therefore, the input representation for the WN layer is formulated as shown in Formula (9).

$$\alpha^t = W_{hh}h^{t-1} + W_{ch}C^t W_{hh} \in R^{4d_h \times d_n} W_{ch} \in R^{4d_h \times d} \tag{9}$$

W_{hh} denotes the recursive hidden layer weight, while W_{ch} signifies the bottom-up input to the hidden layer weight. The state of the hidden layer h_t and the cell input c_t of the WN layer at a specific time step t can be described by Formulas (10) and (11).

$$c^t = \sigma(f^t) \odot c^{t-1} + \sigma(i^t) \odot \tanh(g^t) \tag{10}$$

$$h_t = \sigma(ot) \odot \tanh(LN(c_t)) \tag{11}$$

During the execution of a CNN, numerous multiplications are carried out. When data values are less than 1, they progressively converge toward 0 after undergoing multiple multiplications, leading to the vanishing gradient problem. To address this, the incorporation of an activation function becomes imperative to introduce nonlinearity into the network layers, thereby enhancing the learning capability of the CNN. Given that the activation function is denoted as f, the ultimate output of the WN layer is expressed by Formula (12).

$$WN(\overline{x}_i) = f\left(\frac{g}{\sqrt{(\sigma^t)^2 + \tau}}\right) \odot (a^t - \mu^t) + b) \tag{12}$$

g and b, in the process of neural network training, serve as the reconstruction parameters. The normalization applied to the WN layer has an impact on the feature distribution of the subsequent layers. By introducing these reconstruction parameters, it is possible to restore the original feature distribution of the network, thereby ensuring that the model's expressive capacity is not compromised as a result of normalization.

5. Improved CBAM-iTCN

The proposed CBAM-iTCN prediction algorithm, aiming at the power model prediction of a distributed photovoltaic power station in the region, first analyzes the meteorological factors strongly related to the photovoltaic power and introduces the CBAM net attention mechanism to highlight the key meteorological characteristic information. The improved CBAM-iTCN is used to realize the photovoltaic power prediction. The prediction process is shown in Figure 4.

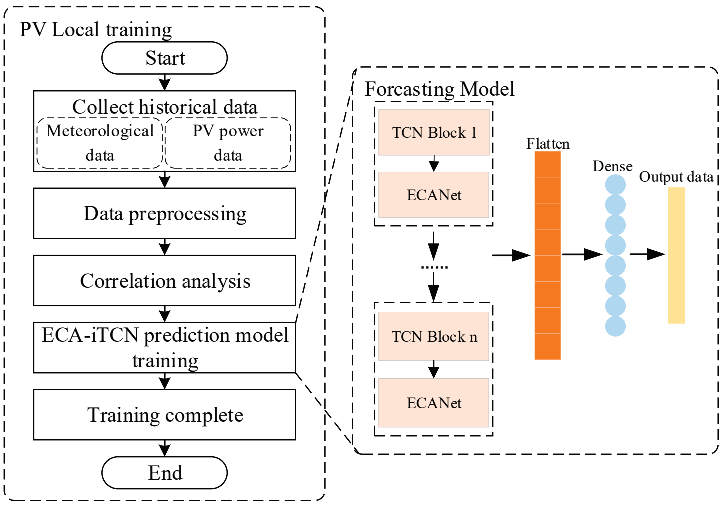


Figure 4. CBAM-iTCN forecast flowchart.

5.1. Temporal Convolutional Networks

In this paper, the TCN model based on CNNs is used to predict the photovoltaic output, and the parallel pooling structure (iTCN) is improved. By introducing the parallel pooling structure into the residual module of the TCN, the activation function ReLU is replaced by GELU. Figure 5 shows the comparison between the original TCN and the improved TCN. Compared with LSTM network, iTCN is not only more stable but also can capture longer sequence dependencies than LSTM, further improving the accuracy of the model, and it has a simple structure, which can process data and improve the efficiency of model training [32].

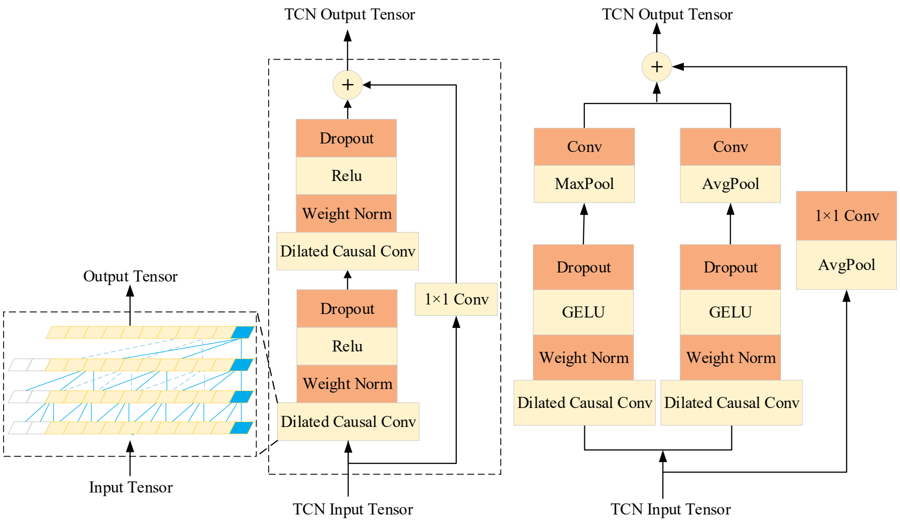


Figure 5. iTCN model structure.

The calculation of unfolding a convolutional network is Formula (13). The size of the convolutional network is k. d is the unfolding coefficient, and (s-di) is the convolutional kernel.

$$F(x_s) = (x * f_d)(s) = \sum_{i=0}^{k-1} f(i) x_{(s-d_i)}, i \in (0, 1, \ldots, k-1) \tag{13}$$

The TCN introduces extended convolution, as shown in Figure 6. Unlike conventional convolution, dilated networks allow for input sampling intervals, which are controlled by *d* in the graph. Generally, the higher the level, the larger the value of *d*. Therefore, TCN can capture longer dependencies in temporal data by using fewer network layers.

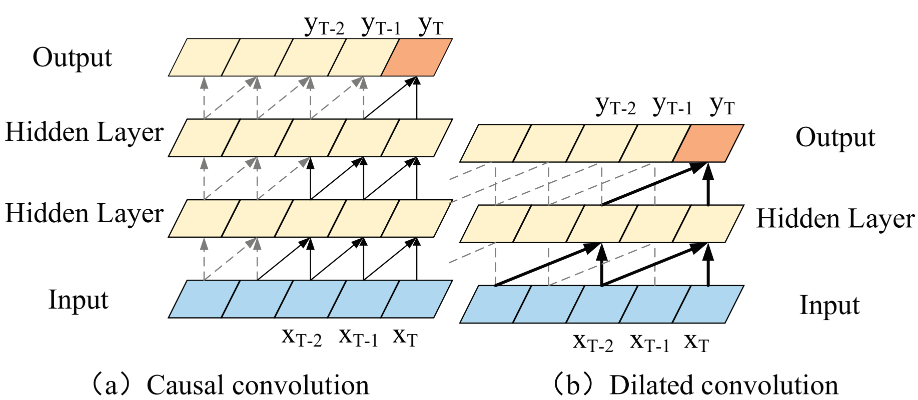


Figure 6. TCN extended convolution.

The increase in network model parameters will lead to the disappearance or explosion of gradient. In order to ensure the stability of the TCN, the residual module is introduced. The residual formula is (1), the input x of the model is weighted and fused into the power $F(x)$ of the model, and finally, the power o of the TCN is obtained.

$$o = RELU[x + F(x)] \tag{14}$$

Using iTCN to establish the photovoltaic prediction model, iTCN can capture the photovoltaic power data, which depend on the light system with a longer time span, but TCN gives each input feature the same weight. However, in practice, factors such as humidity have little impact on the photovoltaic power. Therefore, by introducing the attention mechanism, the sensitivity of the TCN to important features is enhanced, and the prediction performance is improved.

5.2. Attention Mechanism

The Convolutional Block Attention Module (CBAM) is a channel attention mechanism. By using one-dimensional convolution to dynamically allocate weights to different channels, it can effectively capture local cross-channel interactions and allocate different weights, which can enhance the sensitivity of the TCN to important features. Its structure is shown in Figure 7, the CBAMnet structure diagram.

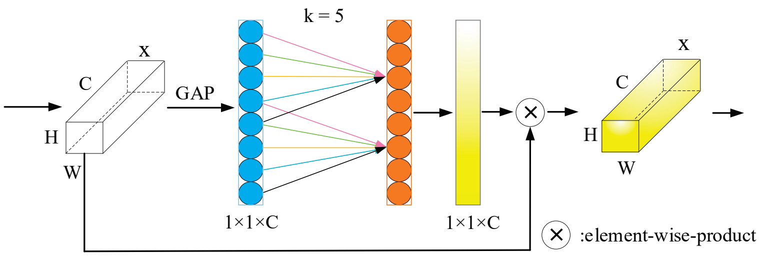


Figure 7. CBAMnet structure diagram.

The global average pool layer gap calculation formula of CBAM is (15), where U is the weighted sequence, T is the time step of the sequence, and the convolution kernel size k is adjusted adaptively; Formula (16), where γ = 2 and b = 1.

$$GAP(U) = \frac{1}{T}\sum_{i=1}^{T} U_{i,j} \tag{15}$$

$$k = \left| \frac{\log_2(c)}{\gamma} + \frac{b}{\gamma} \right| \tag{16}$$

Adding the CBAM attention mechanism to the TCN by focusing on the strongly correlated data features, the model can more effectively extract and use the key information in time series data. This not only optimizes the expression of features and improves the accuracy of prediction but also does not significantly increase the computational cost due to the efficiency of the CBAM.

6. Experimental Results and Analysis

To assess the efficacy of the PFL and DPV power prediction techniques presented in this article, experimental validation is conducted. Specifically, this paper examines and validates the PFL algorithm alongside the CBAM-iTCN prediction model. The experimental setup involves four computers within the same local area network (LAN), which are utilized to emulate DPV power stations (clients). Each of these power stations is equipped with an identical TCN prediction model for consistency.

See Table 1 for the experimental configuration.

Table 1. Experimental platform.

Category	Edition
Operating system	Windows10
CPU (Central server)	Intel Core i9-9900K Processor (Beijing, China)
CPU (Local server)	Intel Core i5-10200H Processor (Beijing, China)
GPU (Central server)	NVIDIA GeForce RTX 3080 Ti Graphics Card (Beijing, China)
GPU (Local server)	NVIDIA GeForce GTX 1660 Graphics Card (Beijing, China)
RAM	32 Gb

6.1. Dateset

The data used in this paper are from photovoltaic power generation companies in North China. The data we selected were collected from 1 January 2018 to 31 December 2019 at an interval of 5 min. The data set includes the meteorological characteristics and photovoltaic power data of photovoltaic power stations. To achieve reliable prediction, the data set is preprocessed first, and the original data are repaired differentially by layering the missing values. Secondly, to reduce the noise influence caused by sensor errors, sudden weather change, and shadow occlusion, the data are processed by multi-modal noise. Finally, the collection time of the processed data set is unified to every hour. And let each photovoltaic power station have different training data sets to better simulate the data deviation in the actual situation.

6.2. Evaluating Indicator

Mean Absolute Error (MAE), Mean Absolute Error Ratio (MAPE), and Root Mean Square Error (RMSE) are selected to evaluate the performance of the prediction model. The Mae formula is (17), the MAPE formula is (18), and the RMSE formula is (19), where $\widetilde{y}_i$ refers to the actual value, and y_i refers to the predicted value in the ith hour, respectively.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widetilde{y}_i| \tag{17}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widetilde{y}_i}{y_i}\right| \tag{18}$$

$$RASE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widetilde{y}_i)^2} \tag{19}$$

6.3. Model Training and Experimental Results

Table 2 presents the correlation between photovoltaic power and various meteorological characteristics. According to the Grey Relational Analysis (GRA), the correlation coefficient for wind speed is less than 0.1, indicating a weak relationship between surface wind speed and photovoltaic power. The primary factors influencing active power are solar radiation and temperature.

Table 2. Correlation analysis.

Meteorological Factors	GRA Correlation
Active power	1.000
Solar radiation	0.918
Temperature	0.56
Wind speed	0.056
Relative humidity	−0.408
Rainfall	−0.056

By comparing the LSTM, TCN, CBAM-TCN, and CBAM-iTCN algorithms on the test data set, the effectiveness of the CBAM-iTCN prediction algorithm is verified. The experiment was conducted on typical sunny, cloudy, and rainy days.

Meteorological factors will vary greatly with seasons, but in the same season, meteorological factors on different days will also vary greatly. At the same time, in order to prevent the influence of seasonal deviation and verify the adaptability of the CBAM-iTCN model, the experiment considers two seasons, spring and summer, which are set in the same season and have different weather conditions, namely, sunny (18 April and 2 July 2018), cloudy (25 April and 8 July 2018), and rainy (19 April and 13 July 2018). The experimental results are shown in Table 3.

Table 3. Models training results under different weather conditions.

Models			Spring			Summer		
			Sunny	Cloudy	Rain	Sunny	Cloudy	Rain
Centralized Learning	TCN	MAE	0.7764	1.0652	1.1115	0.6352	0.8956	1.0325
		RMAE	0.8756	1.5483	1.6168	0.7356	1.2497	1.4629
	LSTM	MAE	0.3895	0.9546	1.2057	0.3036	0.7936	0.9832
		RMAE	0.4928	1.1719	1.4274	0.3982	1.0425	1.2952
	CBAM-TCN	MAE	0.3255	0.8454	1.1456	0.2891	0.7052	0.9826
		RMAE	0.4565	0.9964	1.2564	0.4092	0.8925	1.1748
	Transformer	MAE	0.2951	0.7562	0.9098	0.2581	0.6982	0.8791
		RMAE	0.3858	0.9076	1.2048	0.3287	0.8672	1.1349
	CBAM-iTCN	MAE	0.2641	0.6423	0.7158	0.2273	0.6013	0.6891
		RMAE	0.3562	0.8547	1.1259	0.3159	0.7903	0.9864
Federated Learning	FL-Transformer	MAE	0.2731	0.6891	0.8314	0.2243	0.6142	0.8032
		RMAE	0.3418	0.8158	1.0314	0.3014	0.7631	0.9868
	FL-CBAM-iTCN	MAE	0.2519	0.5981	0.6482	0.2107	0.5572	0.6194
		RMAE	0.3215	0.7139	0.9381	0.2981	0.6739	0.8971
	PFL-CBAM-iTCN	MAE	0.2117	0.5341	0.5982	0.1982	0.4832	0.5531
		RMAE	0.2971	0.6538	0.7891	0.2541	0.5936	0.7013

The experimental results show that the average Mae of the cbam-itcn model is reduced by 45.06%, 36.37%, and 29.97%, respectively, under different weather conditions; the average value of the RMSE decreased by 42.16%, 21.9%, and 13.75%, respectively. In order to more intuitively illustrate the performance advantages of the cbam-itcn model under different weather conditions, the predicted and true values of each model under different weather conditions are shown in Figure 8. In addition, the experimental results show that the MAE and RMSE values of the PFL-CBAM-iTCN model and the FL-CBAM-iTCN model are lower than those of the FL-Transformer model in both spring and summer.

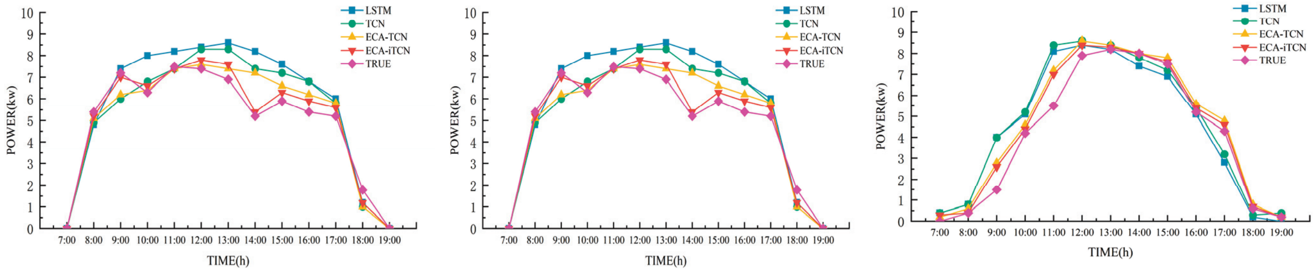


Figure 8. Prediction results under different weather conditions.

To reflect the feasibility of the privatization of the WN layer in multi-task federated learning, during the FL aggregation process, we examine the activation variations in CNN neurons within the local client test set. When the WN layer is designated as the personalized private component in federated learning during model training, it exerts a greater inhibitory effect on neuron activation compared to a scenario where the WN layer is not utilized as a patch. This adjustment brings the neuron distribution closer to its pre-aggregation state. As illustrated in Figure 9, the incorporation of the WN layer between each CNN layer demonstrates a distinct approach compared to conventional FL, regularly constraining the mutation of mean and variance in the process of propagation and ultimately making the network power closer to the pre-aggregated power, as shown in Figure 9.

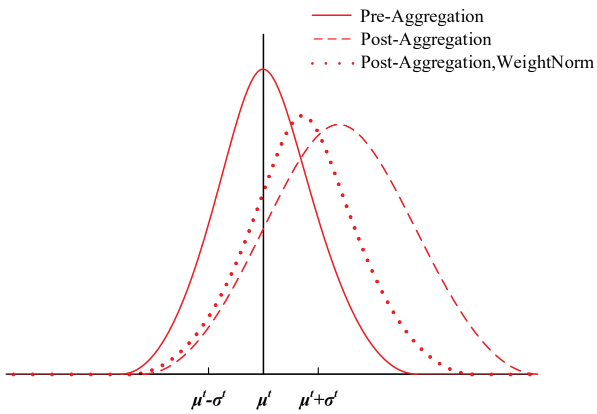


Figure 9. Distribution of neuronal activation before and after polymerization.

To verify the superiority of the PFL algorithm optimization strategy used in this paper, assuming that the average UA of the target is 0.85, the training rounds required by different optimization strategies to reach UA are shown in Table 4. The PFL and FedAvg Adam distributed optimization strategies adopted in this paper can be seen in the table.

Table 4. The communication times required in different optimization strategies.

Optimization Strategy	FedAvg		FedAdam		FedAvg-Adam	
	Epoch = 1	Epoch = 5	Epoch = 1	Epoch = 5	Epoch = 1	Epoch = 5
FL	256	171	124	73	22	18
PFL	53	48	31	24	12	9

The aforementioned analysis reveals that by privatizing the WN layer in FL and developing a distinct personalized model for each client, a significant enhancement in UA can be achieved. The efficacy of personalized federated learning (PFL) frameworks becomes particularly pronounced when addressing expanding client populations and heterogeneous data distributions across decentralized nodes. Compared with conventional federated learning (FL) paradigms, PFL implementations incorporating Adam-optimized federated averaging demonstrate enhanced convergence properties through adaptive gradient-based coordination mechanisms. This hybrid approach synergizes individualized model personalization with collaborative parameter aggregation, enabling accelerated convergence to predefined accuracy thresholds while minimizing inter-node communication overhead.

Next, this paper examines the performance of PFL using a real-world photovoltaic data set. All designated clients engage in the fl process, with the communication rounds set to a total of 75, and we set the training data set with large data feature deviation between clients to better simulate the non-IID data under the actual DPV power.

To reflect the advantages of the PFL algorithm, we compare the training results of pfedme, FL, and PFL. In the FedAvg approach, there is no provision for a client-specific private layer, and the training results are shown in Figure 10.

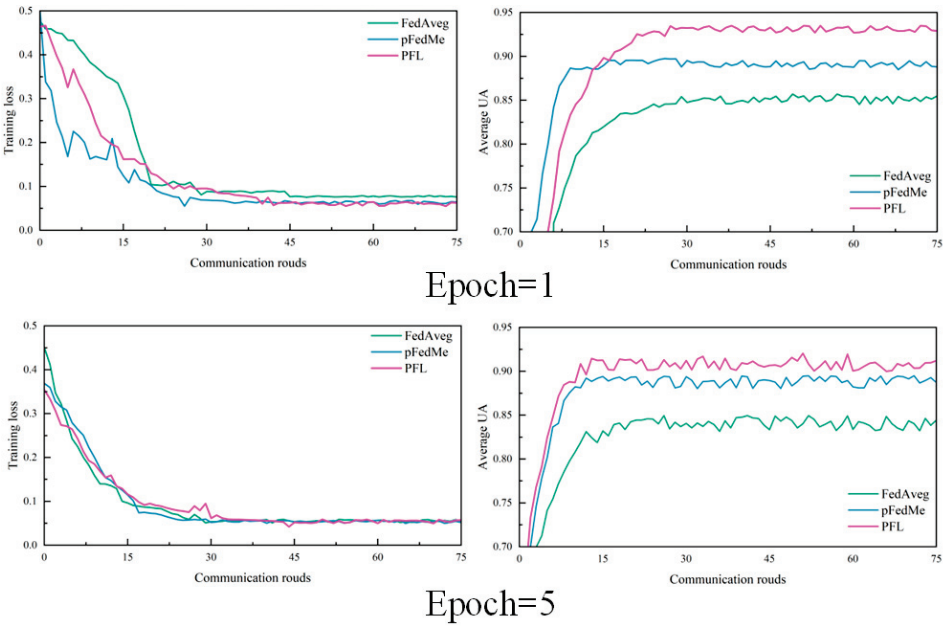


Figure 10. Comparison of training results between PFL and other algorithms.

Figure 10 shows that when the Epoch = 1, the method proposed in this paper can achieve a higher UA than other methods, but due to the fact that the training loss during the personalization phase encompasses both the local model and the personalized model, a greater number of training rounds are required for the model to achieve accuracy convergence. Therefore, the performance of FedAvg in terms of training loss is not as good as FedAvg.

When Epoch = 5, the PFL algorithm is still better than other algorithms, and increasing the number of epochs leads to a reduction in the training loss of the personalized model, enabling it to reach convergence more rapidly while preserving a high level of UA.

In order to show the advantages of PFL over the FL method in detail, a comparative analysis of the training outcomes for the personalized model and the global model is presented in Table 5.

Despite achieving model convergence via multi-round collaborative training, substantial parameter divergence persists across participating nodes in federated learning frameworks, primarily attributed to statistical heterogeneity inherent in client-specific data distributions. Nevertheless, the private WN layer in PFL effectively preserves the unique characteristics of each client, allowing for the tailoring of a proprietary personalized model based on the specific attributes of local data. This, in turn, helps to maintain the prediction error within an acceptable range for all parties involved. When compared to the global model, the personalized model exhibits a 9.79% reduction in MAPE and a 7.06 kW decrease in RMSE. For clients with significant deviations in data features, the MAPE reduction can reach up to 18.07%.

Table 5. Performance comparison between the global model and the personalized model.

	MAPE%		RASE/KW	
	Global Model	**Personalized Model**	**Global Model**	**Personalized Model**
Client 1	7.82	5.79	8.81	2.55
Client 2	19.13	6.08	9.75	2.67
Client 3	11.87	5.84	9.58	2.62
Client 4	24.47	6.41	10.61	2.68
Average	15.82	6.03	9.69	2.63

7. Conclusions

This paper proposes a distributed photovoltaic power prediction method based on personalized federated learning. The PFL collaborative training prediction model is adopted to solve the problems of poor generalization ability and the low accuracy of prediction models caused by a high non-IID of photovoltaic data in a distributed environment. The PFL with a multi-task strategy improves the adaptability of the model by training a separate personalized model. At the same time, by improving the parallel pool structure of the TCN and introducing the attention mechanism CBAMnet, the photovoltaic power prediction model is established to improve the efficiency and accuracy of capturing multi-scale time series features. Lastly, the practicality and efficacy of the proposed methodology are substantiated through experimental analysis conducted on real DPV power data. Compared with the traditional prediction method and single TCN prediction method, the proposed CBAM-iTCN is 45.06% and 42.16% lower than LSTM, MAE, and RMSE. In the real DPV prediction scenarios, the PFL-iTCN algorithm achieves an overall reduction of 9.79% in the MAPE of the predicted values, and for photovoltaic power plants with large data feature deviation, the MAPE is reduced by 18.07%. Experiments show that the proposed scheme can improve the model accuracy in a DPV environment.

However, this paper does not consider the model's performance when there is significant noise in highly heterogeneous data or meteorological data and ignores the prediction time and cost while improving the prediction accuracy. Therefore, how to improve the forecasting performance of newly built photovoltaic power plants with few samples is one of the key research directions in the future. In the future, this research can be further expanded in the following aspects:

(1) Improve the personalized federated learning algorithm to reduce computational cost;
(2) Enhance the generalization ability of federated learning to highly heterogeneous environments or noise data;
(3) Introduce continuous learning strategies and update model parameters regularly.

Author Contributions: Conceptualization, F.D.; data curation, W.L.; formal analysis, Y.S.; investigation, Z.L.; methodology, Y.S.; project administration, F.D.; resources, W.L. and Y.S.; software, W.L.; supervision, F.D.; validation, W.L. and Z.L.; writing—original draft, Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: Natural Science Foundation of China (52167008, 52377103), Natural Science Foundation of Jiangxi Province (20232BAB204064).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

PFL	Personalized Federated Multi-Task Learning
iTCN	improved time series convolution network
TCN	time series convolution network
DPV	Distributed Photovoltaic
WN	Weight Normalization
FL	Federated Learning
CNN	Convolutional Neural Network
LSTM	long-term and short-term memory network
FedAvg	federal averaging
FML	Federated Meta-Learning
AFL	Adaptive Federated Learning
MAE	Mean Absolute Error
MAPE	Mean Absolute Error Ratio
GRA	Grey Relational Analysis

References

1. United Nations. Department of Economic and Social Affairs. *The Sustainable Development Goals: Report 2022*; UN: New York, NY, USA, 2022.
2. Dai, Q.; Huo, X.; Su, D.; Cui, Z. Photovoltaic power prediction based on sky images and tokens-to-token vision transformer. *Int. J. Renew. Energy Dev.* **2023**, *12*, 1104–1112. [CrossRef]
3. Kim, G.G.; Choi, J.H.; Park, S.Y.; Bhang, B.G.; Nam, W.J.; Cha, H.L.; Park, N.; Ahn, H.K. Prediction model for PV performance with correlation analysis of environmental variables. *IEEE J. Photovolt.* **2019**, *9*, 832–841. [CrossRef]
4. Nelega, R.; Greu, D.I.; Jecan, E.; Rednic, V.; Zamfirescu, C.; Puschita, E.; Turcu, R.V.F. Prediction of Power Generation of a Photovoltaic Power Plant Based on Neural Networks. *IEEE Access* **2023**, *11*, 20713–20724. [CrossRef]
5. Ma, D.; Xie, R.; Pan, G.; Zuo, Z.; Chu, L.; Ouyang, J. Photovoltaic Power Output Prediction Based on TabNet for Regional Distributed Photovoltaic Stations Group. *Energies* **2023**, *16*, 5649. [CrossRef]
6. Zhou, N.R.; Zhou, Y.; Gong, L.H.; Jiang, M.L. Accurate prediction of photovoltaic power power based on long short-term memory network. *IET Optoelectron.* **2020**, *14*, 399–405.
7. Chen, J.; Li, J.; Huang, R.; Yue, K.; Chen, Z.; Li, W. Federated Transfer Learning for Bearing Fault Diagnosis with Discrepancy-Based Weighted Federated Averaging. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–11. [CrossRef]
8. Zhang, H.; Zeng, K.; Lin, S. FedUR: Federated Learning Optimization Through Adaptive Centralized Learning Optimizers. *IEEE Trans. Signal Process.* **2023**, *71*, 2622–2637. [CrossRef]
9. Abaoud, M.; Almuqrin, M.A.; Khan, M.F. Advancing Federated Learning Through Novel Mechanism for Privacy Preservation in Healthcare Applications. *IEEE Access* **2023**, *11*, 83562–83579. [CrossRef]
10. Saraswat, R.; Jhanwar, D.; Gupta, M. Sky Image Classification Based Solar Power Prediction Using CNN. *Trait. Signal* **2023**, *40*, 1731–1738. [CrossRef]
11. Jakoplić, A.; Franković, D.; Havelka, J.; Bulat, H. Short-Term Photovoltaic Power Plant Output Forecasting Using Sky Images and Deep Learning. *Energies* **2023**, *16*, 5428. [CrossRef]
12. Wu, Z.; Pan, F.; Li, D.; He, H.; Zhang, T.; Yang, S. Prediction of Photovoltaic Power by the Informer Model Based on Convolutional Neural Network. *Sustainability* **2022**, *14*, 13022. [CrossRef]
13. Yuqi, F.; Hui, L.I.; Lijuan, L.I. Voltage trajectory prediction of photovoltaic power station based on CNN-GRU. *Electr. Power* **2022**, *55*, 163–171.
14. He, Y.; Gao, Q.; Jin, Y.; Liu, F. Short-term photovoltaic power forecasting method based on convolutional neural network. *Energy Rep.* **2022**, *8*, 54–62. [CrossRef]
15. Ahn, H.K.; Park, N. Deep RNN-based photovoltaic power short-term forecast using power IoT sensors. *Energies* **2021**, *14*, 436. [CrossRef]
16. Succetti, F.; Rosato, A.; Araneo, R.; Panella, M. Deep Neural Networks for Multivariate Prediction of Photovoltaic Power Time Series. *IEEE Access* **2020**, *8*, 211490–211505. [CrossRef]

17. Lin, W.; Zhang, B.; Li, H.; Lu, R. Multi-step prediction of photovoltaic power based on two-stage decomposition and BILSTM. *Neurocomputing* **2022**, *504*, 56–67. [CrossRef]
18. Wang, S.; Ma, J. A novel GBDT-BiLSTM hybrid model on improving day-ahead photovoltaic prediction. *Sci. Rep.* **2023**, *13*, 15113. [CrossRef]
19. Liu, L.; Sun, Q.; Wennersten, R.; Chen, Z. Day-Ahead Forecast of Photovoltaic Power Based on a Novel Stacking Ensemble Method. *IEEE Access* **2023**, *11*, 113593–113604. [CrossRef]
20. Xiao, Z.; Huang, X.; Liu, J.; Li, C.; Tai, Y. A novel method based on time series ensemble model for hourly photovoltaic power prediction. *Energy* **2023**, *276*, 127542. [CrossRef]
21. Cheng, L.; Zang, H.; Wei, Z.; Ding, T.; Sun, G. Solar Power Prediction Based on Satellite Measurements—A Graphical Learning Method for Tracking Cloud Motion. *IEEE Trans. Power Syst.* **2021**, *37*, 2335–2345. [CrossRef]
22. Varma, R.K.; Akbari, M. Simultaneous Fast Frequency Control and Power Oscillation Damping by Utilizing PV Solar System as PV-STATCOM. *IEEE Trans. Sustain. Energy* **2019**, *11*, 415–425. [CrossRef]
23. Asiri, E.C.; Chung, C.Y.; Liang, X. Day-Ahead Prediction of Distributed Regional-Scale Photovoltaic Power. *IEEE Access* **2023**, *11*, 27303–27316. [CrossRef]
24. Nitisanon, S.; Hoonchareon, N. Solar power forecast with weather classification using self-organized map. In Proceedings of the 2017 IEEE Power & Energy Society General Meeting, Chicago, IL, USA, 16–20 July 2017; pp. 1–5.
25. Pan, C.; Tan, J. Day-Ahead Hourly Forecasting of Solar Generation Based on Cluster Analysis and Ensemble Model. *IEEE Access* **2019**, *7*, 112921–112930. [CrossRef]
26. Witt, L.; Heyer, M.; Toyoda, K.; Samek, W.; Li, D. Decentral and Incentivized Federated Learning Frameworks: A Systematic Literature Review. *IEEE Internet Things J.* **2022**, *10*, 3642–3663. [CrossRef]
27. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Aguera y Arcas, B. Communication-efficient learning of deep networks from decentralized data. Artificial intelligence and statistics. *Proc. Mach. Learn. Res.* **2017**, *54*, 1273–1282.
28. Nishio, T.; Yonetani, R. Client selection for federated learning with heterogeneous resources in mobile edge. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–7.
29. Deng, F.; Zeng, Z.; Mao, W.; Wei, B.; Li, Z. A Novel Transmission Line Defect Detection Method Based on Adaptive Federated Learning. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–12. [CrossRef]
30. Chen, B.; Chen, T.; Zeng, X.; Zhang, W.; Lu, Q.; Hou, Z.; Zhou, J.; Helal, S. DFML: Dynamic federated meta-learning for rare disease prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2023**, *21*, 880–889.
31. Zhang, P.; Sun, H.; Situ, J.; Jiang, C.; Xie, D. Federated Transfer Learning for IIoT Devices with Low Computing Power Based on Blockchain and Edge Computing. *IEEE Access* **2021**, *9*, 98630–98638. [CrossRef]
32. Chen, Q.; Liu, Y.B.; Ge, M.F.; Liu, J.; Wang, L. A Novel Bayesian-Optimization-Based Adversarial TCN for RUL Prediction of Bearings. *IEEE Sens. J.* **2022**, *22*, 20968–20977.

energies

Article

Quantum Neural Networks for Solving Power System Transient Simulation Problem

Mohammadreza Soltaninia † and Junpeng Zhan *,†

Department of Electrical Engineering, Alfred University, Alfred, NY 14802, USA; ms69@alfred.edu
* Correspondence: zhanj@alfred.edu
† These authors contributed equally to this work.

Abstract: Quantum computing, leveraging principles of quantum mechanics, represents a transformative approach in computational methodologies, offering significant enhancements over traditional classical systems. This study tackles the complex and computationally demanding task of simulating power system transients through solving differential-algebraic equations (DAEs). We introduce two novel Quantum Neural Networks (QNNs): the Sinusoidal-Friendly QNN and the Polynomial-Friendly QNN, proposing them as effective alternatives to conventional simulation techniques. Our application of these QNNs successfully simulates two small power systems, demonstrating their potential to achieve good accuracy. We further explore various configurations, including time intervals, training points, and the selection of classical optimizers, to optimize the solving of DAEs using QNNs. This research not only marks a pioneering effort in applying quantum computing to power system simulations but also expands the potential of quantum technologies in addressing intricate engineering challenges.

Keywords: differential-algebraic equations; power system transient simulation; quantum neural network

1. Introduction

Quantum computing represents a paradigm shift in computational capabilities, leveraging principles of quantum mechanics to process information in fundamentally different ways from classical computing. Traditional computers use binary bits as the basic unit of data, which can either be a zero or a one. In contrast, quantum computers use quantum bits, or qubits, which can exist in multiple states simultaneously, thanks to the phenomena of superposition and entanglement [1]. This allows quantum computers to process a vast number of possibilities concurrently, offering potential exponential speed-ups for certain computational tasks, for example, in factoring large integer numbers [2] or in searching unstructured database [3,4].

Recent advancements in quantum hardware [5–7] and the development of sophisticated algorithms [8–11] have made quantum computing a highly promising area of research. These advancements make quantum computing a promising field for a variety of applications, particularly in areas that require handling complex, high-dimensional datasets or simulations where classical computers struggle.

One of the pivotal developments in this field is the Quantum Neural Network (QNN) [12,13]. QNNs integrate the principles of quantum computing with the architectural concepts of classical neural networks. This hybrid approach has been explored for various applications, including drug discovery, financial modeling, and, notably, complex system simulations such as those needed in power systems.

Quantum computing has been explored for various power system applications, including power flow analysis [14], unit commitment [15], system reliability [16], and stability assessment [17]. There is a burgeoning body of literature, as indicated by several review papers [18–20], that discusses the integration of quantum computing into power system operations, underscoring the participation of entities such as the Department of Energy (DOE) and various utility companies in these research efforts.

Simulation of power systems is of paramount importance for ensuring their stability, efficiency, and reliability [21–24], particularly as the integration of renewable energy sources like wind and solar continues to expand. These renewables are incorporated into power systems through power electronic devices, such as inverters, which necessitate the simultaneous handling of both small and large time steps in simulations. This is due to the rapid responses of the power electronic devices contrasted with the slower-moving dynamics of synchronous generators. Accurate modeling of these diverse dynamics is essential for effective grid management and planning, highlighting the critical need for robust simulation tools capable of managing complex interactions across varying time scales.

The simulation of power systems typically involves solving differential-algebraic equations (DAEs), which describe both the dynamic behavior and the algebraic constraints within the electrical network. Traditional approaches for these simulations include numerical methods such as Euler's and Runge–Kutta methods [25], which discretize the equations to approximate solutions, and semi-analytical methods [23,26], which combine analytical and numerical techniques for improved accuracy. More recently, Physics-Informed Neural Networks (PINNs) [27] have been applied to solve differential equations (DEs) by training neural networks to adhere to the underlying physics of the systems, thus providing a data-driven approach to simulation.

However, each of these methods has its limitations. Numerical and semi-analytical methods can become computationally intensive and may not scale efficiently with the increasing complexity of modern power systems. PINNs, while innovative, often require extensive data and computational resources, as they rely on classical neural networks that utilize a large number of parameters to capture complex functions.

In this context, QNNs present a promising alternative. QNNs leverage the principles of quantum mechanics to perform computations, utilizing a smaller number of parameters than classical neural networks to represent nonlinear functions [12]. This reduction in parameters can potentially lead to more efficient simulations, as QNNs could process complex computations faster and with higher precision due to quantum superposition and entanglement.

Beyond the reduced parameterization and expressive power of QNNs, an emerging body of research suggests that QNNs possess strong generalization abilities and resilience against overfitting–especially in nonlinear settings [28]. This is highly relevant in power system simulations, where dynamic behaviors vary across operating scenarios and overfitting to specific conditions can lead to unreliable models. These properties, combined with the potential for quantum-enhanced speedups, make QNNs a promising alternative to conventional solvers for DAEs in power systems.

Quantum computing has been applied to solve DEs and Partial Differential Equations (PDEs) [29]. However, its potential to solve DAEs remains largely unexplored. Therefore, this paper takes a pioneering step by investigating the use of QNNs to address the challenges in power system simulation. This exploration is the first of its kind and aims to demonstrate how quantum-enhanced computational models can solve the simulation of complex power systems.

The rest of the paper is organized as follows. Section 2 provides a description of the DAEs for the two small power systems simulated. Section 3 offers a general overview of the QNN. Section 4 details the two types of QNNs used to solve the power system simulation problems. Section 5 presents the simulation results. Finally, conclusions are given in Section 6.

2. Problem Description: DAE

DAEs are integral to modeling dynamic systems where constraints intertwine the derivatives of some variables with the algebraic relationships among others. DAEs are prevalent across a diverse array of fields, including electrical circuit design, mechanical system simulation, chemical process modeling, and economic dynamics. This wide applicability underlines the importance of understanding their structure, challenges, and solution methods.

DAEs are characterized by their differentiation index, a critical measure that indicates the complexity of numerical solutions. The differentiation index describes the number of times an equation must be differentiated to convert it into an ordinary differential equation (ODE), affecting the choice of numerical methods for solving the DAE. Common forms of DAEs include semi-explicit, where some equations involve derivatives explicitly, and fully implicit, where derivatives may not be explicitly present.

Boundary and initial conditions play pivotal roles in solving DAEs, providing the necessary constraints to ensure unique and stable solutions. Proper formulation of these conditions is crucial, as inconsistent or inadequate initial conditions can lead to non-converging solutions.

Here, we provide a generalized expression of DAEs.

We define the set S_t as follows:

$$S_t = \left\{ t, y_1(t), y_1'(t), \ldots, y_1^{(m_1)}(t), \ldots, y_n(t), y_n'(t), \ldots, y_n^{(m_n)}(t) \right\}$$

where

- t represents the independent variable, typically time.
- $y_i(t)$ denotes the ith dependent variable as a function of t.
- $y_i'(t), y_i''(t), \ldots, y_i^{(m_i)}(t)$ represent the first, second, and up to the m_ith order derivatives of $y_i(t)$, where m_i is the highest order of derivative for the ith variable.

Subsequently, the DAE system can be succinctly expressed as:

$$\begin{cases} F_1(S_t) = 0, \\ F_2(S_t) = 0, \\ \quad \vdots \\ F_{n_e}(S_t) = 0, \end{cases} \tag{1}$$

where

- $F_j(S_t)$ symbolizes the functions delineating the interrelations within the DAE system. Each F_j corresponds to the jth equation in the system.
- n_e denotes the number of equations in the DAE system, indicating the size of the system.

Boundary conditions are essential for constraining the solution space of DAE systems. They are articulated as a series of equations that the solution must satisfy at predetermined points within the domain, not limited to its boundaries. Formally, these conditions are given by:

$$\begin{cases} B_1(S_{t_*}) = 0, \\ B_2(S_{t_*}) = 0, \\ \vdots \\ B_{n_b}(S_{t_*}) = 0, \end{cases} \tag{2}$$

where

- $B_k(S_{t_*})$ denotes the kth boundary condition function, which applies to the system state at a critical point t_* within the domain.
- n_b represents the total number of boundary conditions enforced on the system.

The DAEs for two power systems are given in Appendices A and B.

3. A Unified Quantum Neural Network Framework for DAEs

In solving DAEs, our approach employs QNNs, aiming to approximate the solutions $y_i(t)$. These QNNs encompass several components: data encoding, a parameterized ansatz, measurement, and a classical optimizer. Each of these components are elaborated on in later sections of this paper. The methodology primarily involves iteratively updating the circuit parameters. This process is continued until the error in the computed solution, based on the circuit's output, is reduced to a satisfactory level within the context of the DAE equations. An in-depth illustration of the QNN's architecture, particularly its iterative parameter update mechanism, is presented in Figure 1.

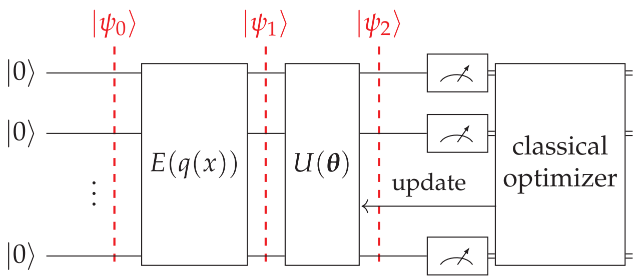


Figure 1. Structure of the QNN used to solve the DAE, including a classical optimizer that updates $\boldsymbol{\theta}$.

From a high-level perspective, this approach comprises two distinct segments: the quantum segment and the classical segment.

3.1. Quantum Segment

To initiate the process, we start with a quantum state represented as:

$$|\psi_0\rangle = |0\rangle^n \tag{3}$$

where n is the number of qubits in the circuit. We then embed our classical input x into the circuit through a quantum operator $E(q(x))$, as expressed in Equation (4).

$$|\psi_1(x, q)\rangle = E(q(x))|\psi_0\rangle \tag{4}$$

The encoding function $q(x)$ serves as a preprocessing step for the classical input before it is embedded into the quantum circuit. Next, a parametric multi-layer quantum ansatz operator $U(\boldsymbol{\theta})$ is applied to make the circuit trainable.

$$|\psi_2(x, q, \boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\psi_1(x, q)\rangle \tag{5}$$

The $\boldsymbol{\theta}$ is a matrix of parameters and can be written as

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_{00} & \boldsymbol{\theta}_{01} & \dots & \boldsymbol{\theta}_{0L} \\ \boldsymbol{\theta}_{10} & \boldsymbol{\theta}_{11} & \dots & \boldsymbol{\theta}_{1L} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\theta}_{n0} & \boldsymbol{\theta}_{n1} & \dots & \boldsymbol{\theta}_{nL} \end{bmatrix}, \tag{6}$$

where $\boldsymbol{\theta}_{ij}$ represents the parameter in the ith qubit and in the jth layer.

3.2. Classical Segment

The solution $f(x, q, \boldsymbol{\theta})$ is considered a function of the expectation value of an observable $\hat{O}$ from the final state of the circuit ($|\psi_2\rangle$). Mathematically, this is given by:

$$M(x, q, \boldsymbol{\theta}) = \langle \psi_2(x, q, \boldsymbol{\theta}) | \hat{O} | \psi_2(x, q, \boldsymbol{\theta}) \rangle \tag{7}$$

$$f(x, q, \boldsymbol{\theta}) = g(M(x, q, \boldsymbol{\theta})) \tag{8}$$

where g represents the post-measurement function applied to M, approximating the value of the function. Subsequently, the quantum models are trained by updating and optimizing the parameter $\boldsymbol{\theta}$ with a classical optimizer guided by a loss function. This loss function comprises two parts, providing a measure of the solution's accuracy.

The first part assesses the discrepancy between the estimated and true boundary values. This is quantified using (9), where n_b denotes the total boundary points.

$$\text{loss}_b = \sum_{i=1}^{n_b} (B_i(S_{t_*}) - f^{\text{model}}(S_{t_*}))^2 \tag{9}$$

Here, $B_i(S_{t_*})$ is derived from the boundary conditions specified by problem (1), while f^{model} is obtained from the quantum circuit, as described in (8).

The second part assesses the degree to which the approximated solutions satisfy the system of DAEs (1) for n_p training points. It is expressed as:

$$\text{loss}_p = \sum_{i=1}^{n_p} \sum_{j=1}^{n_e} F_j(S_{t_i}) \tag{10}$$

Here, n_e denotes the number of DAEs in the system.

Finally, the total loss is computed as a weighted sum of the two components, reflecting the relative importance of boundary accuracy and training point adherence. This is expressed as:

$$\text{loss}_{total} = \lambda_1 \times \text{loss}_b + \lambda_2 \times \text{loss}_p \tag{11}$$

In this equation, λ_1 and λ_2 are tuning coefficients. These coefficients are adjustable parameters that allow for the fine-tuning of the model, enabling the prioritization of either boundary accuracy (λ_1) or adherence at the training points (λ_2). The choice of these coefficients depends on the specific requirements and goals of the model, as well as the characteristics of the problem being addressed.

3.3. Classical Optimizer

We explored several classical optimizers, including Stochastic Gradient Descent (SGD), Adam optimizer, Broyden–Fletcher–Goldfarb–Shanno (BFGS), Limited-Memory BFGS (L-BFGS), and Simultaneous Perturbation Stochastic Approximation (SPSA).

Among these optimizers, BFGS stands out for its fast convergence and robustness in handling non-convex landscapes, making it advantageous in scenarios with saddle points and flat regions [30]. SGD is computationally efficient but may converge slowly [31]. The Adam optimizer offers adaptive learning rates but can exhibit erratic behavior in high-dimensional landscapes [32]. SPSA, on the other hand, is a gradient-free optimizer suitable for optimizing circuits subject to noise [33].

Given the nature of our problem, we opted for the BFGS optimizer to enhance the performance of our QNN circuit. Its capability to navigate non-convex landscapes aligns seamlessly with the intricacies and challenges inherent in our problem domain. Nevertheless, it is imperative to underscore that the choice of an optimizer should be contingent upon the specific demands and constraints of the given problem.

For instance, if the hardware introduces noise, the SPSA optimizer might prove more effective. It is crucial to evaluate the unique attributes of the problem and select an optimizer accordingly to ensure optimal performance. In this study, we utilized ideal simulators devoid of noise, and we opted for BFGS based on the elucidated reasons.

4. QNN Architectures for Power System Simulation

The implementation of the proposed approach necessitated a pivotal consideration: the selection of a quantum model. This model should not only exhibit expressive capabilities but also closely align with the function it is intended to approximate. In the realm of quantum computing, certain global approximator models, such as Haar Unitary, have been proposed. However, their extensive complexity and the vast parameter space required for training them make them less effective. These models often exhibit weak performance, are highly susceptible to encountering saddle-point issues, and may become trapped in local minima during the optimization and training process [34,35].

In this paper, we leveraged the physics information inherent in power system equations to design quantum models. These models feature generative functions that closely mimic the shapes of power system simulation functions. Consequently, the optimization process becomes significantly more efficient. Power systems involve two distinct types of functions: those characterized by fluctuations and sinusoidal patterns, requiring a model adapted to these features, and those amenable to polynomial fitting, such as power-series and Chebyshev functions.

For sinusoidal-friendly models, we employed our previously proposed Sinusoidal-Friendly QNN (SFQ). Additionally, for polynomial-friendly functions, we introduced a novel and promising quantum model named Polynomial-Friendly QNN (PFQ). PFQ harnesses the power of quantum superposition, offering the potential for quantum advantage.

To clearly link the proposed QNNs with the power system transient simulation problem, we emphasize how each QNN approximates solutions to specific DAE components. The DAEs governing power systems, as shown in Equations (A1) to (A3) for the SMIB system and Equations (A4) to (A9) for the 3-machine system, model the time evolution of rotor angle $\delta(t)$ and speed deviation $\omega(t)$.

The QNNs were trained to learn these time-dependent state variables by approximating the functions $\delta(t)$ and $\omega(t)$ using quantum circuits that minimized DAE residuals and boundary mismatches, as defined in Equations (9)–(11). The Sinusoidal-Friendly QNN (SFQ) was particularly suited to learn the sinusoidal behaviors, which reflect oscillatory generator dynamics. In contrast, the Polynomial-Friendly QNN (PFQ) better matched the slowly varying or polynomial-like behaviors.

Thus, each QNN architecture was designed with physical interpretability in mind, aligning its structure with the mathematical form of power-system DAEs, ensuring the learned quantum model respected the physics of the underlying system.

In the upcoming section, we delve into the properties of quantum segments for these specific QNNs utilized in our study.

4.1. Structure 1—Sinusoidal Friendly QNN (SFQ)

Quantum Segment

The first QNN we utilized was the one we previously explored in [12] for sinusoidal-friendly functions. We proved that this QNN could efficiently approximate sinusoidal and fluctuating functions. Here, we explain the structure of the circuit we used and leave the details of the expressivity of the model to the paper we mentioned [12].

For SFQ, we employed two different kinds of embedding: R_y embedding (Figure 2), and $\sin^{-1}$ embedding (Figure 3).

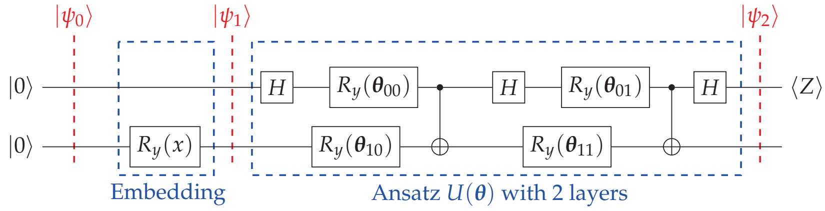


Figure 2. SFQ circuit with R_y embedding and two layers

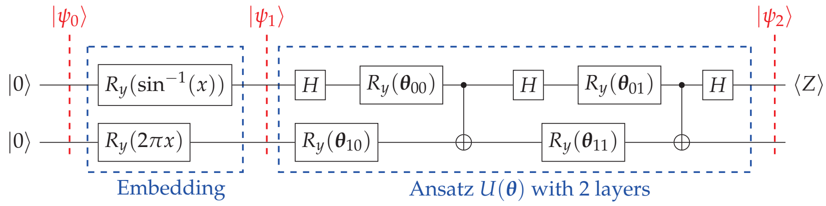


Figure 3. SFQ circuit with arcsin embedding and two layers.

4.1.1. $\sin^{-1}$ Embedding

For this type of embedding, we considered two qubits with

$$q(x) = \begin{bmatrix} \sin^{-1}(x) \\ 2\pi x \end{bmatrix} \tag{12}$$

and

$$E(q(x)) = R_y(\sin^{-1}(x)) \otimes R_y(2\pi x) \tag{13}$$

4.1.2. R_y Embedding

For the R_y embedding in the SFQ structure, we also considered two qubits, but this time

$$q(x) = \begin{bmatrix} I \\ x \end{bmatrix} \tag{14}$$

and,

$$E(q(x)) = I \otimes R_y(x) \tag{15}$$

To implement the quantum ansatz $U(\boldsymbol{\theta})$, we used a series of layers to train the circuit, fitting it to approximate the solution of our DAE. The ansatz consisted of L layers, and the flexibility and variance of our circuit depended on L [12]. The more layers, the more time and data points we needed to train it, so the demand on the number of layers depended on the complexity of our function to approximate [12].

The observable $\hat{O}$ for this structure is $HZ \otimes I$, where H is a Hadamard gate, Z is the Pauli-Z matrix, and I is the identity matrix. The post-processing function g in this structure is [12]:

$$g(u) = \tau_0 + \tau_1 \cdot u + \tau_2 \cdot u^2. \tag{16}$$

Assuming that the expectation value of our observable is $\langle Z_0 \rangle$, then we have

$$f(x, q, \boldsymbol{\theta}) = \tau_0 + \tau_1 \cdot \langle Z_0 \rangle + \tau_2 \cdot \langle Z_0 \rangle^2. \tag{17}$$

4.2. Structure 2—Polynomial Friendly QNN (PFQ)

We propose an innovative model harnessing quantum principles in QNNs, which exhibit exceptional suitability for approximating functions represented by power series and polynomials. Assume the ultimate solution of a DAE is expressed as:

$$h(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^m. \tag{18}$$

Subsequently, we define $|\psi_0\rangle$ as

$$|\psi_0\rangle = |0\rangle^{\otimes n},$$

where n is the number of qubits. Afterward, we define the quantum embedding operator $\hat{E}$ as:

$$|\psi_1(x)\rangle = \hat{E}|0\rangle = \frac{\begin{bmatrix} 1 & x & \ldots & x^m \end{bmatrix}^T}{\left\| \begin{bmatrix} 1 & x & \ldots & x^m \end{bmatrix}^T \right\|}, \tag{19}$$

where $m = 2^n$.

The construction of the operator $\hat{E}$ is described in [36], where researchers developed an effective method for quantum state preparation using measurement-induced steering, enabling the initialization of qubits in arbitrary states on quantum computers. Ultimately, a quantum ansatz operator $U(\boldsymbol{\theta})$ is utilized to produce a subset of the power series, as expressed in Equation (20).

$$|\psi_2(x, \boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|\psi_1(x)\rangle \tag{20}$$

4.3. Simplified PFQ Form with $m = 1$

To illustrate this process for a single qubit, let us consider an R_y gate as in Equation (21).

$$|\psi_2(x, \theta)\rangle = R_y(\theta)|\psi_1(x)\rangle \tag{21}$$

The corresponding quantum circuit with one rotation gate is depicted in Figure 4.

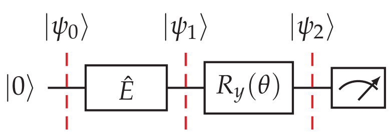


Figure 4. PFQ circuit with one rotation gate.

Assuming we set m to 1 in Equation (19), we obtain:

$$|\psi_1(x)\rangle = \frac{1}{\sqrt{1 + x^2}} \begin{bmatrix} 1 \\ x \end{bmatrix}. \tag{22}$$

The $R_y(\theta)$ is defined as:

$$R_y(\theta) = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}. \tag{23}$$

By substituting Equations (22) and (23) into Equation (21), we obtain:

$$|\psi_2(x,\theta)\rangle = \frac{1}{\sqrt{1+x^2}} \begin{bmatrix} \cos\frac{\theta}{2} - x\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} + x\cos\frac{\theta}{2} \end{bmatrix}. \tag{24}$$

Now, let us calculate the probability $P(|1\rangle)$:

$$P(|1\rangle) = \frac{1}{1+x^2}(\sin^2\frac{\theta}{2} + x^2\cos^2\frac{\theta}{2} + 2x\sin\frac{\theta}{2}\cos\frac{\theta}{2}).$$

After some trigonometric simplifications on $P(|1\rangle)$, we obtain:

$$P(|1\rangle) = \frac{1}{1+x^2}(\sin^2\frac{\theta}{2} + \sin\theta\, x + \cos^2\frac{\theta}{2}\, x^2). \tag{25}$$

Multiplying $P(|1\rangle)$ by $(1+x^2)$ yields:

$$(1+x^2)P(|1\rangle) = \sin^2\frac{\theta}{2} + \sin\theta\, x + \cos^2\frac{\theta}{2}\, x^2. \tag{26}$$

Equation (26) can be abstracted as:

$$h_1(x) = b_0 + b_1 x + b_2 x^2 \tag{27}$$

where

$$b_0 = \sin^2\frac{\theta}{2},\ b_1 = \sin\theta,\ b_2 = \cos^2\frac{\theta}{2}.$$

This expression can be viewed as a segment of a quadratic power series. However, it encounters some limitations, which we address subsequently.

4.4. Enhanced PFQ Version with $m = 1$

Initially, the limitation of a simple PFQ concerns the restricted range of b_0, specifically, $0 \le b_0 \le 1$. Secondly, the range of b_1 is also constrained within $-1 \le b_1 \le 1$.

To mitigate these constraints, we propose the incorporation of two classical adjustable parameters, enhancing the adaptability of the approximating function. Consequently, we define the output as:

$$h_2(x) = \tau_1 \cdot h_1(x) + \tau_0$$

Here, τ_0 stands as a classical trainable parameter, whereas τ_1 may either be a physics-informed constant, tailored to the problem's boundary conditions, or a trainable parameter. In cases where the boundary conditions are undefined, opting for a trainable τ_1 is advisable.

With these adjustments, we obtain:

$$b_0' = \tau_0 + \tau_1 \cdot b_0,$$

and

$$b_1' = \tau_1 \cdot b_1.$$

As a result, the range of b'_0 extends to the entire set of real numbers $\mathbb{R}$, and b'_1 falls within the interval $-\tau_1 \leq b'_1 \leq \tau_1$.

The third limitation is that the range of b_2 is limited, (i.e., $0 \leq b_2 \leq 1$). With the solution of the second limitation, we mitigate the problem of not having negative coefficients for x^2 since we have added $\tau_1 b_2$, which can create a negative coefficient for x^2. However, we still need to improve it and make it more expressive to have both negative and positive coefficients for b_2. To do this, we can add some rotation gates to add the degree of freedom for the coefficients. As an example, consider we add an R_z gate after the R_y gate, as shown in Figure 5. Then, we have:

$$|\psi_3(x,\theta)\rangle = R_z(\theta)R_y(\theta)|\psi_1(x)\rangle = R_z(\theta)|\psi_2(x,\theta)\rangle \tag{28}$$

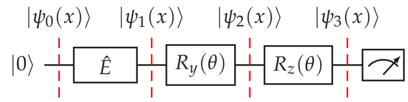


Figure 5. PFQ circuit with two rotation gates.

We can write R_z as:

$$R_z(\theta) = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix} \tag{29}$$

or

$$R_z(\theta) = \begin{bmatrix} \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} \end{bmatrix}. \tag{30}$$

By substituting $|\psi_2(x,\theta)\rangle$ from Equation (24) to (28), we obtain:

$$|\psi_3(x,\theta)\rangle = \frac{1}{\sqrt{1+x^2}} R_z(\theta) \begin{bmatrix} \cos\frac{\theta}{2} - x\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} + x\cos\frac{\theta}{2} \end{bmatrix} \tag{31}$$

Consequently, substituting $R_z(\theta)$ from Equation (30) into (31) and multiplying both sides by $\sqrt{1+x^2}$ yield:

$$\begin{aligned} \sqrt{1+x^2}|\psi_3(x)\rangle &= \begin{bmatrix} \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} \end{bmatrix} \begin{bmatrix} \cos\frac{\theta}{2} - x\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} + x\cos\frac{\theta}{2} \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\frac{\theta}{2} - x\cos\frac{\theta}{2}\sin\frac{\theta}{2} - i\sin\frac{\theta}{2}\cos\frac{\theta}{2} + ix\sin^2\frac{\theta}{2} \\ \sin\frac{\theta}{2}\cos\frac{\theta}{2} + x\cos^2\frac{\theta}{2} + i\sin^2\frac{\theta}{2} + ix\sin\frac{\theta}{2}\cos\frac{\theta}{2} \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\frac{\theta}{2} - \frac{x}{2}\sin(\theta) - \frac{i}{2}\sin(\theta) + ix\sin^2(\frac{\theta}{2}) \\ \frac{1}{2}\sin(\theta) + x\cos^2\frac{\theta}{2} + i\sin^2\frac{\theta}{2} + \frac{1}{2}ix\sin(\theta) \end{bmatrix}, \end{aligned}$$

which shows that we have $x\cos^2\left(\frac{\theta}{2}\right)$ and $ix\sin(\theta)$ in $\sqrt{1+x^2}|\psi_3(x,\theta)\rangle$. This implies that when calculating the probability of $|1\rangle$, terms like $i\cos^2\left(\frac{\theta}{2}\right)\sin(\theta)x^2$ can appear, resulting in both negative and positive coefficients of x^2.

For simplicity, we assumed the same θ for both R_y and R_z. However, it is also possible to consider different θ values (e.g., see Figure 6) to enhance the model's trainability.

Additionally, the introduction of more gates generates additional terms, enabling the model to express a broader range of power series with the same degree.

$$|0\rangle \; — \boxed{\hat{E}} — \boxed{R_y(\boldsymbol{\theta}_{00})} — \boxed{R_z(\boldsymbol{\theta}_{01})} — \boxed{R_y(\boldsymbol{\theta}_{02})} — \; \langle Z\rangle$$

Figure 6. Single-qubit PFQ circuit with three rotation gates.

To evaluate the function to be approximated through this method, we use the same approach as in Equation (8), where O is the expectation value of Z for the first qubit. The expectation value can be calculated as follows:

$$\langle Z\rangle = P(|0\rangle) - P(|1\rangle) \tag{32}$$

Here, we have both $P(|0\rangle)$ and $P(|1\rangle)$, incorporating different coefficients for power series of degree 2.

4.5. Scalability and Advantage of PFQ

PFQ's scalability is facilitated by increasing the number of qubits, as outlined in Equation (19). Specifically, augmenting the qubit count, n, exponentially increases the polynomial's degree, m, where $m = 2^n$. Furthermore, the introduction of entanglement between qubits, alongside additional rotation gates, enhances the degree of freedom for the coefficients of the generated independent terms in our polynomials. In fact, this approach enables the model to effectively tackle more complex systems by employing extra qubits, thereby extending the degrees of the power series.

Another critical aspect to note is that the output from our PFQ circuit undergoes post-processing, expressed as

$$\text{output}(x, \tau_3, \tau_4, \boldsymbol{\theta}) = \tau_3 \times \langle Z\rangle + \tau_4,$$

where τ_3 and τ_4 are classical parameters, and $\boldsymbol{\theta}$ represents quantum parameters, with $\langle Z\rangle$ denoting the expectation value of the observable Z for the first qubit. Consequently, the algorithm involves two classical parameters and a flexible number of quantum parameters, enhancing its quantum scalability.

In real-world power systems, DAEs frequently involve strong nonlinearities (e.g., trigonometric relationships in rotor dynamics) and discontinuities (e.g., abrupt changes during faults or switching events). Our proposed QNN-based framework addresses these challenges through two mechanisms:

Nonlinearity adaptability: Both the SFQ and PFQ architectures are designed to approximate nonlinear behaviors. SFQ targets sinusoidal oscillations, while PFQ, with its polynomial basis, captures smooth nonlinear trends. The quantum encoding and trainable parameterized circuits enable the representation of complex function spaces with fewer parameters than classical NNs, enhancing approximation in nonlinear regions.

Discontinuity management via domain decomposition: For discontinuous phenomena (e.g., faults), we divide the simulation into sequential time segments—pre-fault, fault-on, and post-fault—with QNNs trained independently in each. This approach allows the model to reset its approximation locally, mitigating the impact of discontinuities.

Importantly, recent work (e.g., [28]) has shown that QNNs exhibit strong generalization ability and robustness to overfitting, enabling them to maintain performance across training and testing phases even in complex nonlinear settings.

5. Results

This section presents the results of using various QNNs to solve DAEs. The QNNs were implemented on Pennylane [37,38] on the NCSA Delta high-performance computer (HPC). The BFGS optimizer was employed to optimize the parameters in the QNNs. Our simulations primarily addressed DAEs in power systems, focusing on one-machine and three-machine test systems. The details of their respective DAEs are given in Appendices A and B. In this study, no additional control mechanisms (e.g., excitation systems or PSS) were included. This design choice allowed us to focus on the fundamental response of the power system models and the QNNs' capacity to approximate their dynamics. Results detail the performance of different QNN structures with different embedding methods (Section 5.1), the setting of time interval and the number of training points (Sections 5.2 and 5.3), and the solution of the entire period of the DAE (Section 5.4).

5.1. QNN Structure and Embedding

To analyze the efficacy of various QNN architectures, we initially focused on a single-machine DE over a 0.5-second interval (the setting of the time interval is discussed in the next subsection), utilizing 20 training points. The experimental conditions were kept constant, varying only in the QNN structure applied. As depicted in Figure 7, the SFQ structure excelled in capturing the sinusoidal behavior inherent in the one-machine solution, achieving notably lower error compared to the PFQ structure, which exhibited less precision in replicating the solution dynamics.

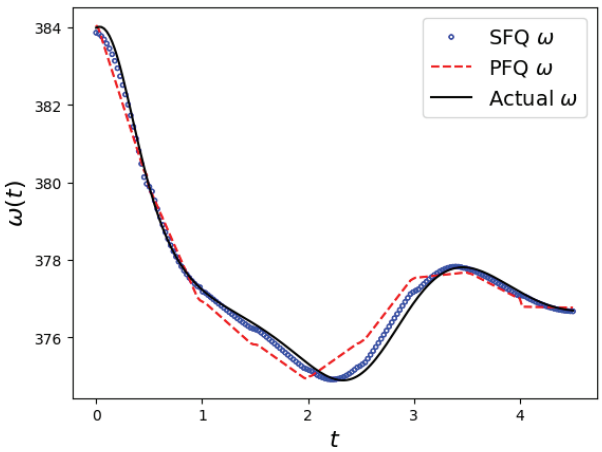


Figure 7. Comparison of actual ω and the ω generated by PFQ and SFQ on a 1-machine test system.

Transitioning to more complex systems, we investigated the performance of these QNN structures on DAE governing the three-machine system. This evaluation was conducted over a shorter interval of 0.2 s with 10 training points. Our results, as summarized in Table 1, demonstrate that all models effectively solved the DAE with minimal error. Note that ω and δ in Tables 1 and 2 represent the rotor speed and rotor angle, respectively. Notably, the PFQ model, illustrated in Figure 6, was particularly adept at managing the complexities of the three-machine system, displaying significant convergence improvements and an eightfold increase in training speed compared to the SFQ model.

To further investigate the robustness of these models, we extended the experimental time interval to 2 s and increased the number of training points to 50. The expanded results, presented in Table 2, clearly indicate that the PFQ structure outperformed the others. It should be noted that due to the restricted domain of the arcsin function, which requires inputs within the range $(-1, 1)$, implementing arcsin embedding within the specified architecture did not successfully solve the equations, resulting in an error, as shown in the third column of Table 2. While rescaling the input through normalization could address this issue, our objective was to maintain the same setup for comparison purposes.

Based on these findings, we suggest selecting the PFQ structure for solving the DE and DAE in power systems, and we used it in the rest of the paper, due to its accuracy and shorter training time.

Table 1. Mean square error for the DAE of the 3-machine WSCC system between the actual solution and solutions obtained by various QNN architectures for $0 \leq t \leq 0.2$ (10 points).

	SFQ-Ry (Figure 2)	SFQ-Arcsin (Figure 3)	PFQ (Figure 6)
$\omega_1(t)$	1.48×10^{-12}	1.02×10^{-12}	9.49×10^{-13}
$\omega_2(t)$	1.57×10^{-11}	5.79×10^{-12}	9.40×10^{-11}
$\omega_3(t)$	9.30×10^{-11}	1.45×10^{-11}	5.47×10^{-11}
$\delta_1(t)$	1.36×10^{-7}	1.44×10^{-7}	1.42×10^{-7}
$\delta_2(t)$	4.02×10^{-7}	3.72×10^{-7}	3.12×10^{-7}
$\delta_3(t)$	9.96×10^{-7}	1.16×10^{-6}	1.26×10^{-6}
Average	2.56×10^{-7}	2.79×10^{-7}	2.86×10^{-7}

Table 2. Mean square error for the DAE of the 3-machine WSCC system between the actual solution and solutions obtained by various QNN architectures for $0 \leq t \leq 2$ (50 points).

	SFQ-Ry (Figure 2)	SFQ-Arcsin (Figure 3)	PFQ (Figure 6)
$\omega_1(t)$	1.40×10^{-9}	failed	8.62×10^{-10}
$\omega_2(t)$	8.15×10^{-9}	failed	2.11×10^{-9}
$\omega_3(t)$	2.43×10^{-7}	failed	7.31×10^{-9}
$\delta_1(t)$	7.76×10^{-6}	failed	5.20×10^{-7}
$\delta_2(t)$	3.92×10^{-4}	failed	2.37×10^{-6}
$\delta_3(t)$	1.27×10^{-2}	failed	6.67×10^{-5}
Average	2.18×10^{-3}	-	1.16×10^{-5}

5.2. Time Interval for DAE Solving

In this paper, the DAE was solved by dividing the time domain into specific intervals, using the final state of one interval as the initial state for the next. This approach simplified the challenge of function fitting within each interval.

The selection of an appropriate time interval is critical. A shorter time interval can enhance solution accuracy by limiting the exploration space. However, while a smaller span increases the frequency of DAE resolutions required, it does not necessarily lead to longer overall runtime, as training over shorter spans can be faster, especially for complex models. Determining the best time span involves a trade-off between accuracy and computational efficiency, typically resolved through trial and error.

Figure 8 demonstrates the effect of different time spans on the solution accuracy for $\omega(t)$ in the one-machine system using 20 training points. The results indicate that a time span of $t_s = 0.2$ s delivered the most precise outcomes. However, a span of $t_s = 0.5$ s also provided satisfactory results, and to minimize the number of computational runs, we selected $t_s = 0.5$ s. Following similar experimental guidance, a span of $t_s = 2$ s was chosen for the three-machine DAEs, balancing precision and computational demand.

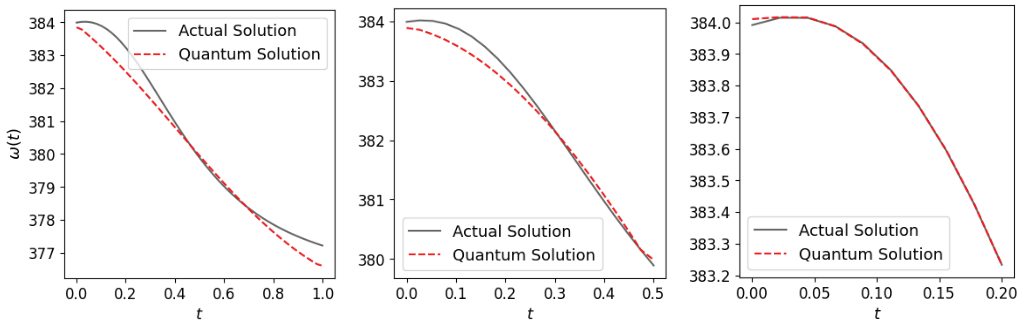


Figure 8. Comparison of quantum and actual solutions for an SMIB system using different time spans: left panel t_s = 1, middle panel t_s = 0.5, right panel t_s = 0.2.

5.3. Number of Training Points

The number of training points for a quantum circuit is determined by balancing accuracy against computational constraints. Increasing the number of points typically improves precision but also requires more computational time. Hence, optimizing the number of training points involves striking a balance between computational efficiency and desired accuracy. Figure 9 illustrates this relationship for the SMIB system, showing how an increased number of training points reduces the error between the quantum approximation and the actual solution for $\omega(t)$ over the interval $t \in [0, 0.5]$.

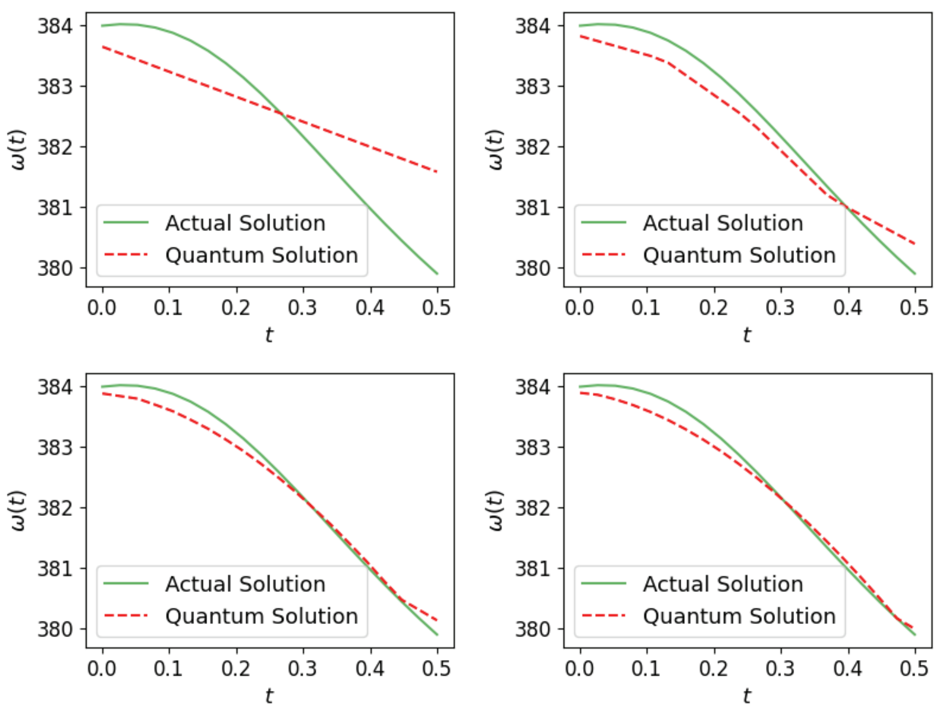


Figure 9. Comparison of quantum and actual solutions for an SMIB system using various training points: top left: 3 points, top right: 5 points, bottom left: 10 points, bottom right: 20 points.

5.4. Complete Simulation

By aggregating results over various time intervals, we successfully solved the DAEs for both one-machine and three-machine systems. In the one-machine scenario, we set the time span to 8 s with a time interval of $t_s = 0.5$ s, using 20 training points per interval. We utilized the SFQ structure with R_y embedding, and the complete solution is presented in Figure 10.

For the three-machine system, we set the time span to 20 s, maintaining a time interval of $t_s = 2$ s, and similarly using 20 training points per interval. The complete solution achieved with the PFQ structure is illustrated in Figure 11.

The results affirmed that the quantum solutions closely aligned with the actual solutions for both δ and ω variables, demonstrating the ability of the SFQ and PFQ to accurately simulate the dynamics of power systems in the one-machine and three-machine cases, respectively.

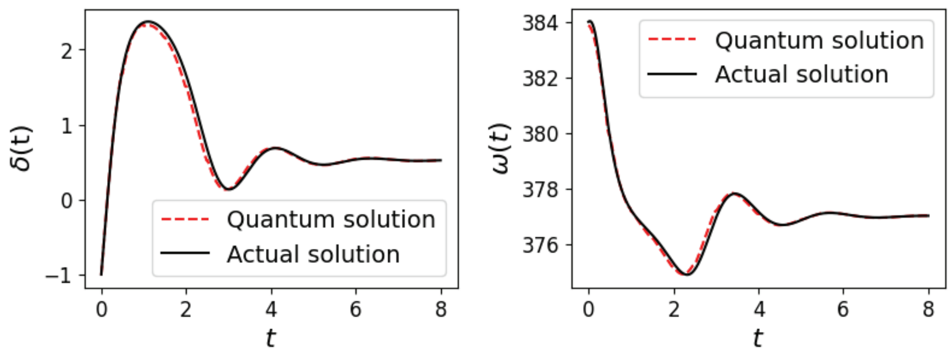


Figure 10. Quantum solution vs. actual solution for the SMIB system: left panel: rotor angle $\delta(t)$, right panel: rotor speed $\omega(t)$.

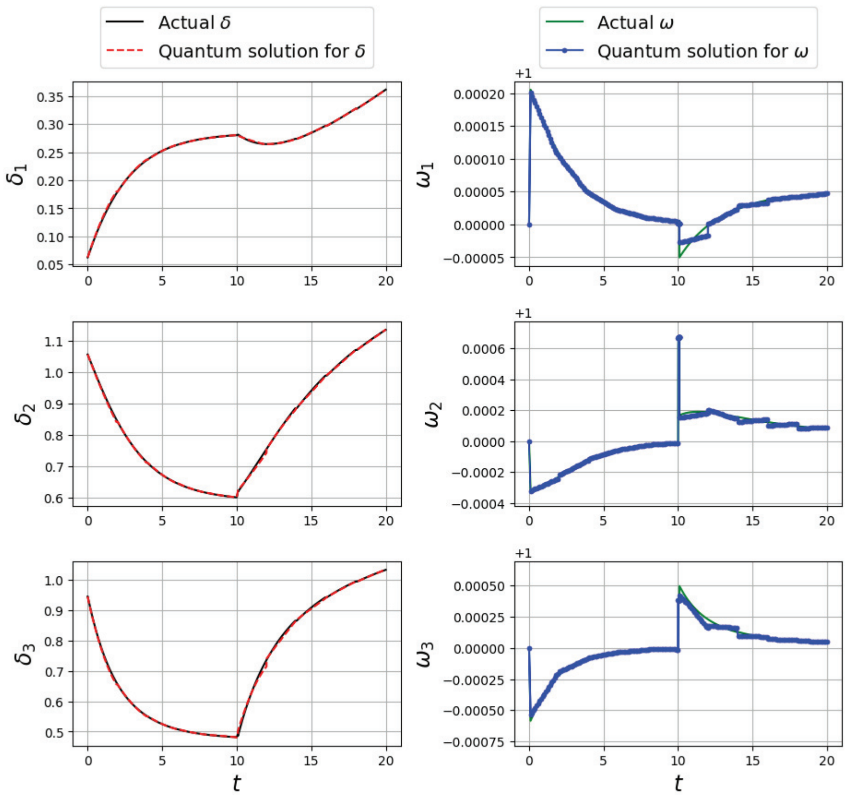


Figure 11. Quantum solution vs. actual solution for the WSCC 3-machine system: the three panels on the left are the rotor angles δ of the three machines, respectively; the three panels on the right are the rotor speed ω of the three machines, respectively.

6. Conclusions

This study demonstrated the potential of Quantum Neural Networks (QNNs) in enhancing the simulation of power system transients by effectively solving differential-algebraic equations (DAEs). This study presented a novel application of Quantum Neural Networks (QNNs) for solving differential-algebraic equations (DAEs) in power system transient simulations. We developed two specialized QNN architectures—the Sinusoidal-Friendly QNN (SFQ) and the Polynomial-Friendly QNN (PFQ)—that incorporated domain knowledge about the underlying physics of power systems to enhance training efficiency and accuracy. Our results on benchmark systems (SMIB and WSCC 3-machine) demonstrated that QNNs could accurately approximate DAE solutions, with mean square errors as low as 10^{-5} in multi-phase transient scenarios. Several benefits of the proposed QNN approach were observed: (1) efficient learning with fewer parameters than classical neural networks; and (2) tailored architectures (SFQ and PFQ) that matched the functional characteristics of power system responses.

While the current implementation was simulated on noiseless quantum backends, the proposed models are hardware-compatible and set the stage for real-time quantum-assisted simulation tools in power grid analysis. Future research will aim to incorporate advanced power system controllers, scale the proposed methodology to larger and more complex power networks, and validate the approach on near-term quantum devices with noise-resilient training strategies to rigorously assess scalability and computational benefits. This work advances the frontier of applying quantum machine learning to critical engineering problems and contributes a new paradigm for the data-efficient, structure-aware simulation of power systems using quantum resources.

Author Contributions: Conceptualization, M.S. and J.Z.; methodology, M.S. and J.Z.; software, M.S.; validation, M.S. and J.Z.; formal analysis, M.S. and J.Z.; investigation, M.S. and J.Z.; resources, M.S. and J.Z.; data curation, M.S. and J.Z.; writing—original draft preparation, M.S.; writing—review and editing, J.Z.; visualization, M.S. and J.Z.; supervision, J.Z.; project administration, J.Z.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the NSF ERI program, under award number 2138702. This work used the Delta system at the National Center for Supercomputing Applications through allocation CIS220136 and CIS240211 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296.

Data Availability Statement: The data will be made available upon request.

Acknowledgments: We acknowledge the use of IBM Quantum services for this work. The views expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Quantum team. We thank I. Tutul for providing the Julia code used to solve the simulation problem of the two systems using classical methods.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Single-Machine Infinite Bus (SMIB) System

This part details the ODE for the SMIB system, as described below [39]:

$$\frac{d\delta}{dt} = \omega - \omega_s \tag{A1}$$

$$\frac{d(\omega - \omega_s)}{dt} = K_1 - K_2 \sin(\delta) - K_3(\omega - \omega_s) \tag{A2}$$

where $K_1 = \frac{\omega_s}{2H} T_m^0$, $K_2 = \frac{\omega_s}{2H}\left(\frac{E_c V}{X}\right)$, $K_3 = \frac{\omega_s}{2H} D$, E_c is equal to the magnitude of the internal voltage of the machine, X is the sum of reactance, and T_m^0 is the constant mechanical torque. The initial δ_0 was -1, and the initial speed difference $(\omega_0 - \omega_s)$ was 7. $K_1 = 5$, $K_2 = 10$, $K_3 = 1.7$. These data were taken from [39].

We combined Equations (A1) and (A2) into a single DE by eliminating ω:

$$\frac{d^2\delta}{dt^2} = K_1 - K_2 \sin(\delta) - K_3 \frac{d\delta}{dt} \tag{A3}$$

Appendix B. WSCC Three-Machine System

Here, we provide the DAEs of the WSCC three-machine nine-bus power system. The DAE of the ith generator of the power system is given below [40]:

$$\frac{d\delta_i}{dt} = \omega_s \Delta\omega_i \tag{A4}$$

$$\frac{d\Delta\omega_i}{dt} = \frac{1}{2H_i}(P_{mi} - P_{ei} - D_i \Delta\omega_i) \tag{A5}$$

$$P_{ei} = e_{xi} i_{xi} + e_{yi} i_{yi} \tag{A6}$$

$$\begin{bmatrix} e'_{xi} \\ e'_{yi} \end{bmatrix} = \begin{bmatrix} \sin\delta_i & \cos\delta_i \\ -\cos\delta_i & \sin\delta_i \end{bmatrix} \begin{bmatrix} 0 \\ e'_{qi} \end{bmatrix} \tag{A7}$$

$$I_t = \begin{bmatrix} i_{x1} \\ i_{y1} \\ \vdots \\ i_{xn} \\ i_{yn} \end{bmatrix} = Y \begin{bmatrix} e'_{x1} \\ e'_{y1} \\ \vdots \\ e'_{xn} \\ e'_{yn} \end{bmatrix} \tag{A8}$$

$$\begin{bmatrix} e_{xi} \\ e_{yi} \end{bmatrix} = \begin{bmatrix} e'_{qi}\cos\delta_i \\ e'_{qi}\sin\delta_i \end{bmatrix} - \begin{bmatrix} R_{ai} & -X'_{di} \\ X'_{di} & R_{ai} \end{bmatrix} \begin{bmatrix} i_{xi} \\ i_{yi} \end{bmatrix} \tag{A9}$$

where n is the total number of generators, δ_i and $\Delta\omega_i$ are the rotor angle and rotor speed deviation from the nominal value of generator i, respectively, H_i and D_i are the inertia and damping constants of generator i, respectively, P_{mi} is the mechanical power of generator i, P_{ei} is the electric power of generator i e'_{xi} and e'_{yi} are the internal bus voltages of generator i in the non-rotating coordinate, respectively, e'_{qi} is the field voltage of generator i, i_{xi} and i_{yi} are the terminal currents of generator i, Y is the admittance matrix, R_{ai} and X'_{di} are the source impedance of generator i, respectively, e_{xi} and e_{yi} are the terminal voltages for the x and y axes of generator i, respectively.

Initial values for δ_i, ω_i, ω_s, H_i, P_{mi}, D_i, e'_{qi}, δ_i, R_{ai}, X'_{di} were obtained from the machine dataset given in [40]. Utilizing these data, the admittance matrix Y was calculated. Additionally, P_{ei} was derived from the network equations based on these initial values.

For the disturbance simulation, we simulated a three-phase fault near bus 7 at the end of lines 5–7, which was cleared within five cycles (0.083 s) by opening lines 5–7. The simulation included three distinct operational phases: pre-fault, fault-on, and post-fault conditions. The pre-fault phase spanned from 0 to 10 s, the fault-on phase from 10.0 to 10.083 s, and the post-fault phase from 10.083 to 20 s. Each phase initialized with different starting points: pre-fault initial values were sourced from machine data to calculate P_{ei}; fault-on initial values were derived from the terminal values of the pre-fault simulation at 10 s, with adjustments made to accommodate the new fault-on Y matrix; and post-fault initial values were taken from the end of the fault-on phase at 10.083 s, incorporating modifications for a subsequent Y matrix change. The respective Y matrices for each condition are documented in [40].

Table A1. Parameter data of the three machines in the WECC system

Parameter	Generator 1	Generator 2	Generator 3
H	23.64	6.40	3.01
D	23.64	6.40	3.01
R_a	0	0	0
X'_d	0.0608	0.1198	0.1813
P_m	0.7164	1.6300	0.8500
e'_q	1.0566	1.0502	1.0170
$\delta(0)$	0.0626	1.0567	0.9449
$\Delta e(0)$	0	0	0
$\text{Real}(I_t)$	0.6889	1.5799	0.8179
$\text{Imag}(I_t)$	-0.2601	0.1924	0.1730
I_d	0.2872	0.3523	0.0178
I_q	0.6780	1.5521	0.8358

References

1. Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information: 10th Anniversary Edition*; Cambridge University Press: Cambridge, UK, 2011. [CrossRef]
2. Shor, P.W. Algorithms for quantum computation: Discrete logarithms and factoring. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA, 20–22 November 1994; pp. 124–134. [CrossRef]
3. Grover, L.K. A fast quantum mechanical algorithm for database search. In Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing (STOC '96), Philadelphia, PA, USA, 22–24 May 1996; ACM: New York, NY, USA, 1996; pp. 212–219. [CrossRef]
4. Grover L.K. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **1997**, *79*, 325. [CrossRef]
5. Madsen, L.S.; Laudenbach, F.; Askarani, M.F.; Rortais, F.; Vincent, T.; Bulmer, J.F.F.; Miatto, F.M.; Neuhaus, L.; Helt, L.G.; Collins, M.J.; et al. Quantum computational advantage with a programmable photonic processor. *Nature* **2022**, *606*, 75–81. [CrossRef]
6. Hyyppä, E.; Kundu, S.; Chan, C.F.; Gunyhó, A.; Hotari, J.; Janzso, D.; Juliusson, K.; Kiuru, O.; Kotilahti, J.; Landra, A.; et al. Unimon qubit. *Nat. Commun.* **2022**, *13*, 1–14. [CrossRef]
7. Arrazola, J.M.; Bergholm, V.; Brádler, K.; Bromley, T.R.; Collins, M.J.; Dhand, I.; Fumagalli, A.; Gerrits, T.; Goussev, A.; Helt, L.G.; et al. Quantum circuits with many photons on a programmable nanophotonic chip. *Nature* **2021**, *591*, 54–61. [CrossRef]
8. Lloyd, S.; De Palma, G.; Gokler, C.; Kiani, B.; Liu, Z.-W.; Marvian, M.; Tennie, F.; Palmer, T. Quantum algorithm for nonlinear differential equations. *arXiv* **2020**, arXiv:2011.06571. [CrossRef]
9. Jordan, S.P. Quantum Algorithm Zoo. Available online: https://quantumalgorithmzoo.org (accessed on 22 April 2011).
10. Bharti, K.; Cervera-Lierta, A.; Kyaw, T.H.; Haug, T.; Alperin-Lea, S.; Anand, A.; Degroote, M.; Heimonen, H.; Kottmann, J.S.; Menke, T.; et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **2022**, *94*, 015004. [CrossRef]
11. Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. *Nature* **2017**, *549*, 195–202. [CrossRef]
12. Liao, Y.; Zhan, J. Expressibility-Enhancing Strategies for Quantum Neural Networks. *arXiv* **2022**, arXiv:2211.12670v1. [CrossRef]

13. Skolik, A.; McClean, J.R.; Mohseni, M.; van der Smagt, P.; Leib, M. Layerwise learning for quantum neural networks. *Quantum Mach. Intell.* **2021**, *3*, 5. [CrossRef]
14. Feng, F.; Zhou, Y.; Zhang, P. Quantum Power Flow. *IEEE Trans. Power Syst.* **2021**, *36*, 3810–3812. [CrossRef]
15. Feng, F.; Zhang, P.; Bragin, M.A.; Zhou, Y. Novel Resolution of Unit Commitment Problems Through Quantum Surrogate Lagrangian Relaxation. *IEEE Trans. Power Syst.* **2023**, *38*, 2460–2471. [CrossRef]
16. Nikmehr, N.; Zhang, P. Quantum-Inspired Power System Reliability Assessment. *IEEE Trans. Power Syst.* **2023**, *38*, 3476–3490. [CrossRef]
17. Zhou, Y.; Zhang, P. Noise-Resilient Quantum Machine Learning for Stability Assessment of Power Systems. *IEEE Trans. Power Syst.* **2023**, *38*, 475–487. [CrossRef]
18. Gao, F.; Wu, G. Application of Quantum Computing in Power Systems. *Energies* **2023**, *16*, 2240. [CrossRef]
19. Golestan, S.; Habibi, M.R.; Mousavi, S.Y.M.; Guerrero, J.M.; Vasquez, J.C. Quantum computation in power systems: An overview of recent advances. *Energy Rep.* **2023**, *9*, 584–596. [CrossRef]
20. Zhou, Y.; Tang, Z.; Nikmehr, N.; Babahajiani, P.; Feng, F.; Wei, T.-C.; Zheng, H.; Zhang, P. Quantum computing in power systems. *iEnergy* **2022**, *1*, 170–187. [CrossRef]
21. Chow, J.H. *Power System Coherency and Model Reduction*; Springer: New York, NY, USA, 2013; Volume 94.
22. Barret, J.-P.; Bornard, P.; Meyer, B. *Modelling and Simulation Techniques for Power System Engineering*; Electricite de France: Paris, France, 2024.
23. Sun, K. *Power System Simulation Using Semi-Analytical Methods*; Wiley: Hoboken, NJ, USA, 2023. [CrossRef]
24. Fan, L.; Miao, Z. *Modeling and Stability Analysis of Inverter-Based Resources*; CRC Press: Boca Raton, FL, USA, 2024. [CrossRef]
25. Wanner, G.; Hairer, E. *Solving Ordinary Differential Equations II*; Springer: Berlin, Germany, 1996; Volume 375.
26. Liu, Y.; Sun, K. Solving Power System Differential Algebraic Equations Using Differential Transformation. *IEEE Trans. Power Syst.* **2020**, *35*, 2289–2299. [CrossRef]
27. Moya, C.; Lin, G. DAE-PINN: A physics-informed neural network model for simulating differential algebraic equations with application to power networks. *Neural Comput. Appl.* **2023**, *35*, 3789–3804. [CrossRef]
28. Jiang, J.; Zhao, Y.; Li, R.; Li, C.; Guo, Z.; Fan, B.; Li, X.; Li, R.; Zhang, X. Strong generalization in quantum neural networks. *Quantum Inf. Process.* **2023**, *22*, 428. [CrossRef]
29. Oz, F.; San, O.; Kara, K. An efficient quantum partial differential equation solver with chebyshev points. *Sci. Rep.* **2023**, *13*, 7767. [CrossRef]
30. Zhou, W.J.; Li, D.H. A globally convergent BFGS method for nonlinear monotone equations without any merit functions. *Math. Comput.* **2008**, *77*, 2401–2419. [CrossRef]
31. Bottou, L. Large-scale machine learning with stochastic gradient descent. In Proceedings of the COMPSTAT'2010: 19th International Conference on Computational Statistics, Paris, Italy, 22–27 August 2010; pp. 177–186. [CrossRef]
32. Reddi, S.J.; Kale, S.; Kumar, S. On the convergence of Adam and beyond. In Proceedings of the 6th International Conference on Learning Representations (ICLR), New Orleans, LA, USA, 6–9 May 2019. [CrossRef]
33. Wiedmann, M.; Hölle, M.; Periyasamy, M.; Meyer, N.; Ufrecht, C.; Scherer, D.D.; Plinge, A.; Mutschler, C. An Empirical Comparison of Optimizers for Quantum Machine Learning with SPSA-based Gradients. *arXiv* **2023**, arXiv:2305.00224v1. [CrossRef]
34. Holmes, Z.; Sharma, K.; Cerezo, M.; Coles, P.J. Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus. *PRX Quantum* **2022**, *3*, 010313. [CrossRef]
35. McClean, J.R.; Boixo, S.; Smelyanskiy, V.N.; Babbush, R.; Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.* **2018**, *9*, 4812. [CrossRef] [PubMed]
36. Volya, D.; Mishra, P. State Preparation on Quantum Computers via Quantum Steering. *IEEE Trans. Quantum Eng.* **2024**, *5*, 1–14. [CrossRef]
37. Bergholm, V.; Izaac, J.; Schuld, M.; Gogolin, C.; Ahmed, S.; Ajith, V.; Alam, M.S.; Alonso-Linaje, G.; AkashNarayanan, B.; Asadi, A.; et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv* **2022**, arXiv:1811.04968. [CrossRef]
38. Soltaninia, M.; Zhan, J. Comparison of Quantum Simulators for Variational Quantum Search: A Benchmark Study. *arXiv* **2023**, arXiv:2309.05924. [CrossRef]
39. Sauer, P.W.; Pai, M.A.; Chow, J.H. *Power System Dynamics and Stability: With Synchrophasor Measurement and Power System Toolbox*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2017.
40. Wang, B.; Liu, Y.; Sun, K. Power System Differential-Algebraic Equations. *arXiv* **2021**, arXiv:1512.05185v3. [CrossRef]

energies

Article

Wind Power Forecasting Based on Multi-Graph Neural Networks Considering External Disturbances

Xiaoyin Xu [1], Zhumei Luo [1,*] and Menglong Feng [2]

[1] Faculty of Metallurgical and Energy Engineering, Kunming University of Science and Technology, Kunming 650093, China; 18807769398@163.com
[2] School of Energy Power and Mechanical Engineering, North China Electric Power University (NCEPU), Beijing 102206, China; dh18810387092@163.com
* Correspondence: luozhumei@163.com

Abstract: Wind power forecasting is challenging because of complex, nonlinear relationships between inherent patterns and external disturbances. Though much progress has been achieved in deep learning approaches, existing methods cannot effectively decompose and model intertwined spatio-temporal dependencies. Current methods typically treat wind power as a unified signal without explicitly separating inherent patterns from external influences, so they have limited prediction accuracy. This paper introduces a novel framework GCN-EIF that decouples external interference factors (EIFs) from inherent wind power patterns to achieve excellent prediction accuracy. Our innovation lies in the physically informed architecture that explicitly models the mathematical relationship: $P(t) = P_{\text{inherent}}(t) + \text{EIF}(t)$. The framework adopts a three-component architecture consisting of (1) a multi-graph convolutional network using both geographical proximity and power correlation graphs to capture heterogeneous spatial dependencies between wind farms, (2) an attention-enhanced LSTM network that weights temporal features differentially based on their predictive significance, and (3) a specialized Conv2D mechanism to identify and isolate external disturbance patterns. A key methodological contribution is our signal decomposition strategy during the prediction phase, where an EIF is eliminated from historical data to better learn fundamental patterns, and then a predicted EIF is reintroduced for the target period, significantly reducing error propagation. Extensive experiments across diverse wind farm clusters and different weather conditions indicate that GCN-EIF achieves an 18.99% lower RMSE and 5.08% lower MAE than state-of-the-art methods. Meanwhile, real-time performance analysis confirms the model's operational viability as it maintains excellent prediction accuracy (RMSE < 15) even at high data arrival rates (100 samples/second) while ensuring processing latency below critical thresholds (10 ms) under typical system loads.

Keywords: wind power forecasting; multi-graph convolution networks; external interference factor; LSTM; attention

1. Introduction

Wind power generation has become a critical component in sustainable energy portfolios worldwide, owing to its renewable nature, zero-emission operation, and decreasing implementation costs. However, the inherent variability and stochastic characteristics of wind power generation, which are heavily influenced by meteorological conditions and geographical factors, pose significant challenges to power system integration, transmission network management, and overall grid stability [1]. As a result, accurate wind power forecasting has become crucial for efficient power system scheduling and operation, the optimal utilization of wind generation resources, the maintenance of reliable electricity supply to meet demand variations, and the strategic planning of energy storage systems and grid infrastructure [2].

Despite its obvious volatility, wind power generation exhibits underlying spatio-temporal patterns that make it potentially predictable through sophisticated modeling approaches. In past decades, extensive studies have been conducted on wind power prediction, such as autoregressive integrated moving average (ARIMA), support vector regression (SVR), and artificial neural networks [3–5].

Wind power usually demonstrates inherent patterns and thus high potential predictability, laying the foundation for our method to separate inherent patterns from external disturbances. Recently, owing to their strong ability to capture spatial or temporal dependencies, deep neural networks exhibit better performance than shallow models in wind power prediction [6]. Meanwhile, several factors influence wind power, including temporal aspects (such as increased travel during holiday periods), weather conditions (like an unexpected rainstorm during evening rush hours), and social activities (for instance, a pop concert scheduled to take place downtown) [7]. The impact of these factors on wind power is referred to as disturbance.

Nevertheless, state-of-the-art (SOTA) deep learning methods have two main limitations. On the one hand, wind power is time series data distributed across different farms with a non-Euclidean structure. Therefore, the spatial characteristics acquired by Convolutional Neural Networks (CNNs) are not the most effective for capturing structural representations. Conversely, a significant amount of research utilizes wind power data for training deep neural networks, inevitably introducing noise into the model and restricting prediction accuracy. Importantly, historical wind power data encompass not only standard wind power trends but also considerable disturbances. As a result, reducing the influence of various disturbances from historical wind power data before inputting them into prediction models is essential.

As illustrated in Figure 1, this paper proposes a model called GCN-EIF to capture features of different nodes and uncover the disturbances for wind power prediction. The contributions are summarized below:

- A two-graph convolutional network is proposed for extracting spatial features from different farms. Specifically, two different adjacency matrices are constructed to describe similarity among nodes, indicating that more similar nodes have more similar characteristics.
- To capture the temporal features of a farm, a Long Short-Term Memory (LSTM) network is designed, which uses an attention mechanism to distinguish the importance of different temporal features.
- A combinational mechanism is designed to further incorporate the EIF into wind power prediction. First, the EIF is eliminated from the historical wind power to represent inherent wind power patterns. Next, the wind power without the EIF is fed into LSTM-Attention. Finally, to better account for future disturbances, the final prediction is derived by integrating the EIF of the prediction time interval with the output of LSTM-Attention. A comprehensive series of experiments is carried out to confirm the efficacy of the EIF modeling approach and the combinatorial mechanism.

Note: The Related Work section expands upon concepts introduced in the Introduction with more technical depth and comparative analysis of methods, while the Introduction section focuses on providing the necessary context and motivation for our research problem.

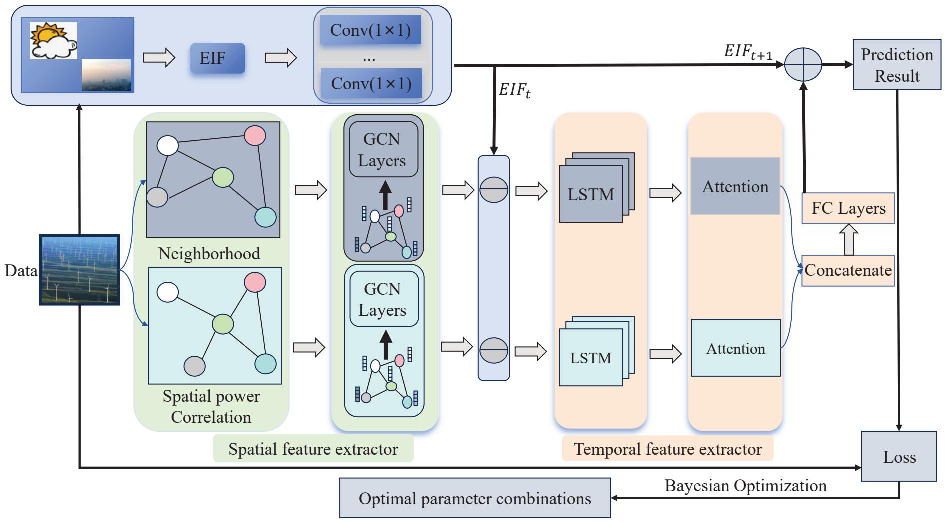


Figure 1. Proposed prediction model employing multi-graphs incorporating EIF.

2. Related Work

Wind power forecasting has been a significant focus of research due to its importance in integrating renewable energy sources into power systems. Different methodologies have been investigated over the years, spanning from classical statistical approaches to sophisticated machine learning methods, including those that capture both temporal and spatial features of wind power data [8,9].

The ARIMA model is one of the most commonly used traditional methods for time series forecasting, including wind power prediction. It focuses on modeling linear relationships among past data points to predict future values. Though ARIMA performs well with stationary time series, it faces challenges when dealing with the nonlinear and volatile nature of wind power data. Recent studies have shown that while ARIMA models can be useful for short-term predictions, their performance degrades substantially for long-term forecasts [3,10,11].

To overcome the limitations of traditional statistical models, machine learning techniques such as SVR and random forest regression have been widely applied to wind power forecasting [4,12]. SVR is particularly effective in capturing nonlinear relationships within the data, providing improved forecasting accuracy over linear models. However, SVR's performance can be limited by the computational complexity involved, especially with large datasets, and it may fail to fully capture the temporal dependencies critical for wind power forecasting [13].

Neural networks, especially those designed to capture temporal dependencies, have gained popularity in wind power forecasting [5]. The Recurrent Neural Network (RNN) is highly suitable for time series forecasting due to its ability to process sequential data [14]. However, standard RNNs suffer from issues such as vanishing gradients, which limit their effectiveness in capturing long-term dependencies. LSTM networks and Gated Recurrent Units (GRUs) were presented to overcome the aforementioned limitations [15,16]. These models have been extensively used in wind power forecasting, demonstrating superior performance in capturing complex temporal patterns. The Temporal Convolutional Network (TCN) has also become a promising alternative to RNNs for capturing temporal dependencies. It utilizes convolutional layers to model patterns over various time scales, achieving a balance between complexity and accuracy. It has shown potential in improving wind power forecasting by effectively capturing both short-term and long-term dependencies [17].

Graph Neural Networks (GNNs) have gained much attention for modeling spatial dependencies in wind power data. By representing wind farms or measurement points as nodes in a graph and their interactions as edges, GNNs can capture complex spatial relationships that traditional models often ignore. Attributed to this capability, GNNs are particularly effective for spatially aware wind power forecasting and can significantly improve prediction accuracy [18,19].

Recent advancements in wind power forecasting have led to the development of hybrid models that combine the strengths of both temporal and spatial models. These spatio-temporal models, such as Convolutional LSTM (ConvLSTM) [6] and Spatio-Temporal Graph Convolutional Networks (ST-GCNs), integrate convolutional or recurrent layers with GNNs to simultaneously capture spatial and temporal dependencies. These models have demonstrated excellent performance in wind power forecasting by effectively modeling the intricate interactions between time and space [20,21].

Another prominent trend in wind power forecasting is to incorporate external factors such as meteorological conditions, geographical information, and environmental parameters. These external factors have a large impact on wind patterns, making their inclusion in predictive models crucial for improving accuracy. Recent models that incorporate meteorological data into spatio-temporal modeling have shown great improvements in forecasting performance [7,22].

3. Methodology

3.1. Spatial Features and Multi-Graph Convolution Networks

$G = (V, \boldsymbol{A})$ is an undirected graph for representing the correlations among various nodes, where V is the node set and $\boldsymbol{A} \in \mathbb{R}^{|V|\times|V|}$ is the adjacent matrix indicating connections among these nodes.

(1) Multi-GraphConstruction: To encode varying types of correlation between nodes, two graphs are constructed using distinct similarity measures. $G_N = (V, \boldsymbol{A}^N)$ is a neighborhood graph for representing the spatial distance between different nodes, while the spatial power correlation $G_S = (V, \boldsymbol{A}^S)$ depicts the power correlations among the nodes.

- Neighborhood graph $G_N = (V, \boldsymbol{A}^N)$: The neighborhood graph performs information aggregation from neighboring nodes and utilizes their physical proximity to assess their correlation. Generally, a shorter distance between two nodes indicates a stronger correlation between them. Therefore, the elements in the adjacent matrix $A^N_{h,w}$ are defined using the Euclidean distance between nodes:

$$A^N_{h,w} = -\left(\frac{||v_h - v_w||^2}{2\sigma^2}\right) \quad \forall h, w \tag{1}$$

  where v_h is the position of node h, $||v_h - v_w||$ represents the Euclidean distance between nodes h and w, and the scaling parameter σ controls the sensitivity of the similarity measure to distance variations. The optimal value of σ is determined through cross-validation on the training dataset.

- Spatial power correlation graph $G_S = (V, \boldsymbol{A}^S)$: The Pearson coefficient is employed to examine the spatial correlation between the power across various nodes. Therefore, it is utilized to establish the adjacent matrix as

$$A^S_{h,w} = \frac{cov(\boldsymbol{X}_h, \boldsymbol{X}_w)}{\sigma_{\boldsymbol{X}_h}\sigma_{\boldsymbol{X}_w}} \quad \forall h, w \tag{2}$$

  where $cov(\cdot)$ is the covariance operator, σ denotes the standard deviation, and $\boldsymbol{X}_h$ and $\boldsymbol{X}_w$ represent the power of nodes h and w, respectively.

Let $\boldsymbol{A}$ represent the adjacency matrix for a previously constructed graph, and power $\boldsymbol{X}$ is the other input parameter. As illustrated in Figure 2, a graph convolutional network is employed to capture the spatial characteristics. The procedure is described in detail below.

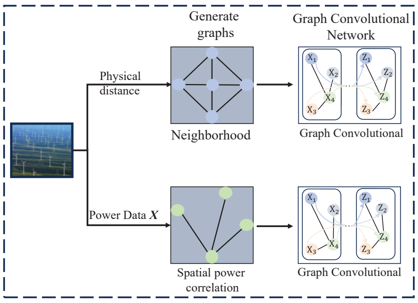


Figure 2. Graph convolutional networks.

(2) Graph Convolutional Networks: Given the adjacent matrix $\boldsymbol{A} \in \mathbb{R}^{V \times V}$ of a constructed graph, we can calculate a degree matrix $\boldsymbol{D}$ with elements $\boldsymbol{D}_{ii} = \sum_j \boldsymbol{A}_{ij}$. $\boldsymbol{A}$ is symmetric, so $\boldsymbol{D}$ is a diagonal matrix. Then, the normalized Laplacian matrix is derived as $\boldsymbol{L} = \boldsymbol{I}_V - \boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2}$, where $\boldsymbol{I}_V$ is the identity matrix. Through eigenvalue decomposition $\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$, the eigenvector matrix $\boldsymbol{U}$ and a diagonal matrix $\boldsymbol{\Lambda}$ can be obtained.

As $\boldsymbol{L}$ is also symmetric and has V linearly independent eigenvectors, $\boldsymbol{U}$ is orthogonal, providing a basis set of Fourier transforms. Consequently, with $\boldsymbol{U}$, Fourier transform can be conducted on the input feature $\boldsymbol{X}$: $\hat{\boldsymbol{X}} = \boldsymbol{U}^T\boldsymbol{X}$, with $(\cdot)^T$ indicating the transpose.

By Equation (4), a filter $g_\theta(\boldsymbol{\Lambda})$ can be established by the Chebyshev polynomial $T_k(\cdot)$: $g_\theta(\boldsymbol{\Lambda}) = \sum_{k=0}^{K-1} \theta_k T_k(\hat{\boldsymbol{\Lambda}})$, where $\hat{\boldsymbol{\Lambda}} = \frac{2}{\lambda_{\max}}\boldsymbol{\Lambda} - \boldsymbol{I}_V$, and $\lambda_{\max} = \max\{\boldsymbol{\Lambda}_{ii}\}$ is the largest eigenvalue. Subsequently, a graph convolution operation, $*_\mathcal{G}$, is conducted between the established filter $g_\theta(\boldsymbol{\Lambda})$ and input feature $\mathbf{X}$. We have

$$g_\theta(\boldsymbol{\Lambda}) *_\mathcal{G} \boldsymbol{X} = \boldsymbol{U} g_\theta(\boldsymbol{\Lambda}) \boldsymbol{U}^T \boldsymbol{X}. \tag{3}$$

(3) Spatial feature extractor output: In taking the adjacent matrix and each node's power as inputs, the spatial feature extractor obtains two outputs, as shown in the middle of Figure 1, that is,

$$G_n = GCN(A_{h,w}^N, \boldsymbol{X}) \tag{4}$$

$$G_s = GCN(A_{h,w}^S, \boldsymbol{X}) \tag{5}$$

3.2. Modeling External Interference Factors

External interference factors affecting power include pressure, humidity, temperature, wind speed and direction, etc. Among these external factors, our analysis indicates that wind speed and direction have the greatest impact on prediction accuracy, with correlation coefficients of 0.78 and 0.65, respectively. Temperature effects showed moderate impact (0.42), while pressure variations exhibited the least influence (0.28). Measurement uncertainties in these parameters are managed through a robust normalization process that standardizes input ranges while preserving relative relationships. Here, 2D convolution (Conv2D) is used to map the feature of the EIF, as demonstrated at the top of Figure 1. Using the historical EIF and the EIF for predicted time, two outputs EIF_t and EIF_{t+1} can be obtained, respectively.

Since the power encompasses both typical power and various disturbances, mitigating the impact of these disturbances is crucial for modeling the power patterns. Thus, during the prediction phase, the EIF is removed from the historical power to reveal the complex dependencies between inherent power patterns and various disturbances for power prediction:

$$\bar{G}_n = GCN(A^N_{h,w}, \boldsymbol{X}) - EIF_t \tag{6}$$

$$\bar{G}_s = GCN(A^S_{h,w}, \boldsymbol{X}) - EIF_t \tag{7}$$

3.3. LSTM Networks and Attention Mechanisms

The temporal features of the power data are captured with LSTM networks, and the importance of different times is extracted using the attention mechanisms.

(1) LSTM Networks: LSTM networks have a specialized recurrent architecture designed to overcome the fundamental limitations of conventional RNNs. By implementing gating mechanisms, LSTMs effectively mitigate gradient instability issues, allowing for the robust modeling of extended temporal dependencies. Its basic structure is called a cell. As depicted in Figure 3, each cell involves a forget gate, an input gate, and an output gate. The inputs to these three gates include the previous output h_{t-1}, the current input x_t, and the previous cell state c_{t-1}.

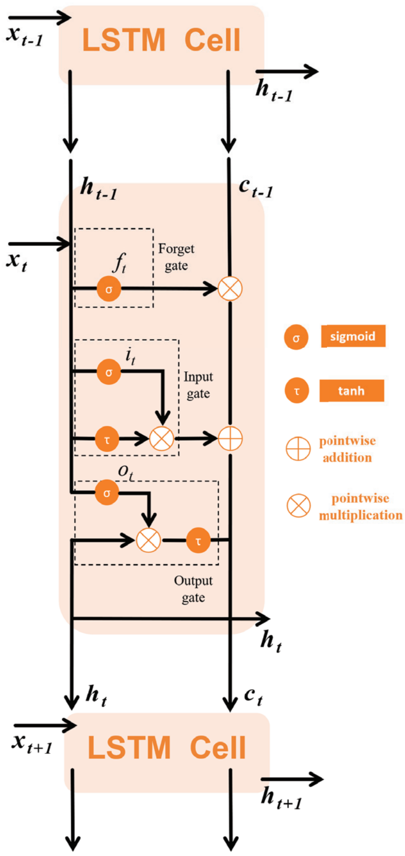


Figure 3. LSTM cell.

The function of the forget gate f_t is to filter information, judging whether to discard or keep the information from the previous time step. The input gate i_t determines whether to add the new information to the current cell state. The output gate o_t selects the effective information from the current cell and is directly related to the output h_t. The calculation formulas for each gate and the update formula for the cell state are as follows:

$$i_t = \sigma(\boldsymbol{\Omega} xix_t + \boldsymbol{\Omega} hih_{t-1} + \boldsymbol{\Omega} ci \diamond ct - 1 + \beta_i), \tag{8}$$

$$c_t = f_t \diamond c_{t-1} + i_t \diamond \tanh(\boldsymbol{\Omega} xcx_t + \boldsymbol{\Omega} hch_{t-1} + \beta_c), \tag{9}$$

$$f_t = \sigma(\mathbf{\Omega} xf x_t + \mathbf{\Omega} hf h_{t-1} + \mathbf{\Omega} cf \diamond ct - 1 + \beta_f), \tag{10}$$

$$o_t = \sigma(\mathbf{\Omega} xo x_t + \mathbf{\Omega} ho h_{t-1} + \mathbf{\Omega} co \diamond ct - 1 + \beta_o), \tag{11}$$

$$h_t = o_t \diamond \tanh(c_t), \tag{12}$$

where Ω_{xi}, Ω_{hi}, Ω_{ci}, Ω_{xf}, Ω_{hf}, Ω_{cf}, Ω_{xo}, Ω_{ho}, Ω_{co}, Ω_{xc}, and Ω_{hc} represent the weight matrices, $\odot$ denotes the element-wise product operation, and β_i, β_f, β_o, and β_c indicate the bias parameters of the network.

(2) Attention Mechanisms: The attention mechanism, as an optimization model algorithm, is introduced to investigate the inherent features of data and enhance the efficiency of information processing, by leveraging its ability to distinguish the importance of different data. Our prediction model recognizes that the information provided by the data at different times may hold varying levels of importance for prediction performance. Nevertheless, the standard LSTM fails to identify the important parts for a power data sequence. To address this issue, we integrate the attention mechanism into the LSTM system, enabling the model to establish direct dependencies between states across different times.

At time slot t, for dataset $\mathcal{X}$, the scores $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ represent the importance of the traffic flow sequence, which is calculated as

$$s_t = U_s \tanh(W_{xs}\mathcal{X} + W_{hs} h^s_{t-1} + b_s), \tag{13}$$

with U_s, W_{xs}, and W_{hs} being the learnable parameters, b_s being the bias parameter, and h^s_{t-1} being the hidden output generated by the LSTM network. The weights at varying moments are called attention weights and are represented as

$$\alpha_k = \frac{\exp(s_k)}{\sum_{n=1}^{n} \exp(s_k)}, \tag{14}$$

The scores are normalized so that a larger weight indicates a more important parameter. At each time slot t, the output of the LSTM hidden state h^s_t, which relies on the attention mechanisms, can be obtained as a weighted summation

$$H_t = \sum_{k=1}^{n} \alpha_k h^s_k, \tag{15}$$

where the attention weight α is affected by both the input $\mathcal{X}$ and the hidden variable h^s_{t-1}, so it relies on both the present and previous moments. The attention weight α can be understood as the activation of the flow selection gate, with the set of gates controlling the amount of information collected from each flow to input to the LSTM network. A greater attention weight signifies a larger influence on the prediction outcomes.

(3) Temporal feature extractor output: After $\bar{G}_n$ and $\bar{G}_s$ have gone through the LSTM network and the attention mechanism, the output L_n and L_s can be obtained, as shown in Figure 1:

$$L_n = Attention(LSTM(G_n)), \tag{16}$$

$$L_s = Attention(LSTM(G_s)). \tag{17}$$

To fuse the features of L_n and L_s, the Concatenate and FC layers are fused to obtain the inherent output $\bar{Y}_t$ as

$$\bar{Y}_t = FC(Concatenate(L_n, L_s)). \tag{18}$$

Finally, the inherent output $\bar{Y}_t$ is combined with the EIF of time interval $t+1$ to account for the future disturbances

$$\hat{Y}_{t+1} = \bar{Y}_t + EIF_{t+1}, \tag{19}$$

where Y represents the final prediction, which includes both typical traffic patterns and various disturbances at time interval $t+1$.

3.4. Parameter Learning

The training of the framework can be conducted in an end-to-end manner through reducing the mean squared error (MSE) between the prediction results and the ground truths as much as possible:

$$\theta = \arg\min_{\theta}\left(Y_{t+1} - \hat{Y}_{t+1}\right)^2, \tag{20}$$

where θ denotes all parameters that can be learned in our model, which can be optimized using the Adam optimizer through backpropagation.

4. Experiment

4.1. Dataset Description

The dataset contains power information and weather information. The dataset used in this study is secretly available from our lab. Other researchers can access relPower.mat and NWP files for validation purposes.

The file relPower.mat contains information about 40 offshore wind farms, including wind station IDs and actual power for all stations from January 2022 to August 2022. As listed in Table 1, the first column is the timestamp, and the second to last column is the actual power data for each station. The file Nwp contains the weather data, with each station corresponding to one NWP, and the file is named after the ID of the station, where each weather file contains eight columns. As shown in Table 2, the first two columns show the latitude and longitude information of the wind farms, the third column is the timestamps, and the remaining five columns are the pressure, humidity, temperature, wind speed, and wind direction, respectively. The time of the data is UTC0 time, which needs to be transformed into Beijing time by adding 8. The time resolution of the data is 1 h.

Table 1. Power data.

Timestamp	1	2	...	40
2022/1/1 0:00	6	5		1.568
2022/1/1 1:00	2	10	...	2.491
...		...		
2023/8/31 23:00	48.23	69.9	...	29.62

Table 2. Weather data.

Feature	Sample
Longitude	32.6°
Latitude	121.1°
Timestamp	2022/1/1 0:00
Pressure	103,138.92 Pa
Wind Speed	8.16 m/s
Wind Direction	29.34
Humidity	39.96%
Temperature	279.69 K

4.2. Data Process

First, the UTCO time was converted to BST, and the power data and weather data were normalized to [0,1] via min-max normalization. When performing evaluation, the predicted traffic flows were re-scaled to normal ones.

The latitude and longitude of each wind farm were utilized to calculate the distance between them, and the power data were employed to calculate the Pearson correlation coefficient. Figure 4 shows the Pearson correlation heat map for 10 of these wind farms. It can be observed that their correlations are all above 0.9, indicating that they are highly correlated in terms of power generation.

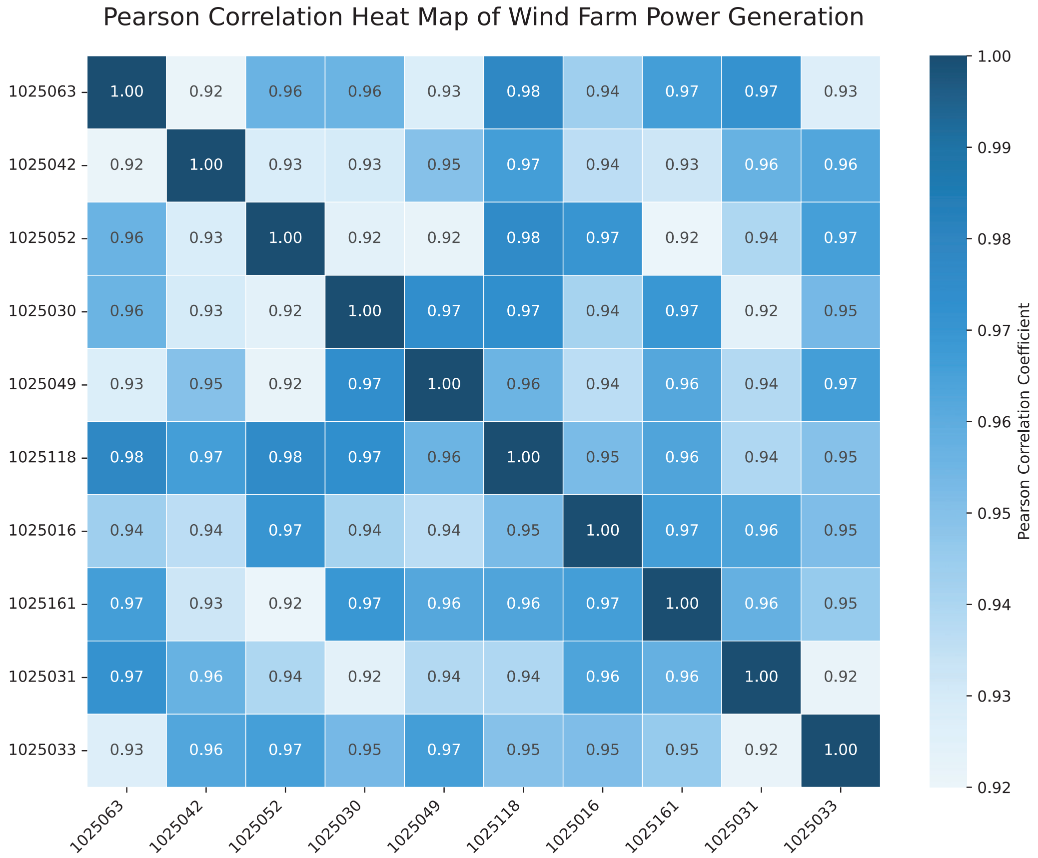


Figure 4. Pearson correlation heat map showing strong spatial relationships between wind farms. The correlation coefficient values range from 0.92 to 0.98, indicating high spatial dependencies. Latitude and longitude data are included in our correlation analysis as they provide essential context for understanding spatial relationships between wind farms.

The 10 farms depicted in the heatmap were selected from the total 40 farms based on their geographical distribution to offer a representative sample of different correlation patterns. This visualization helps illustrate the importance of modeling spatial dependencies in wind power forecasting.

Notably, latitude and longitude data are included in our correlation analysis as they provide an essential context for understanding spatial dependencies between wind farms. The correlation coefficient values range from 0.92 to 0.98, showing strong spatial relationships.

4.3. Hyperparameter Optimization

Bayesian optimization is employed to obtain the optimal combination of hyperparameters. Based on the experimental results, this study chose one GCN layer, two LSTM network layers, and four Conv2D layers; the learning rate was set to 0.05, the batch size was set to 16, and the dropout rate was set to 0.01. The impact of these hyperparameters will be discussed in the next section.

4.4. Evaluation Indicators

The performance of our model was evaluated using the following indicators:

$$RMSE = \sqrt{\frac{1}{t}\sum_{t}(Y_{t+1} - \hat{Y}_{t+1})^2}, \quad (21)$$

$$MAE = \frac{1}{t}\sum_{t}|Y_{t+1} - \hat{Y}_{t+1}|, \quad (22)$$

$$R^2 = 1 - \frac{\sum_t \left(Y_{t+1} - \hat{Y}_{t+1}\right)^2}{\sum_t \left(\bar{Y}_{t+1} - \tilde{Y}_{t+1}\right)} \quad (23)$$

where Y_{t+1} and $\hat{Y}_{t+1}$, respectively, represent the available ground truths and the prediction values; t denotes the number of all available ground truths.

5. Results and Analysis

The experiments on the dataset include four aspects:

- Comparison with baselines. We begin by comparing the general predictive performance between GCN-EIF and baselines; next, we analyze our model's computational complexity and real-time performance; third, we examine the model performance on various days, including workdays, weekends, and holidays.
- Computational efficiency analysis. We analyze the theoretical complexity of GCN-EIF and perform comprehensive empirical evaluations of runtime performance, processing latency, and memory scaling in different operational scenarios to verify the suitability of the model for real-time applications.
- Comparison with variants of GCN-EIF. First, we validate the efficacy of the GCN-EIF modeling approach; subsequently, we investigate the influence of the elimination and combination approaches; lastly, we assess the impact of multiple context factors.
- Impact of hyperparameters. We also explore the influence of critical hyperparameters in GCN-EIF, including the number of LSTM layers, historical time intervals, and selection of similar nodes for graph construction.

5.1. Comparison with Baselines

(a) General predictive performance comparison: We assess the general predictive performance of GCN-EIF against eight baseline algorithms. As illustrated in Table 3, GCN-EIF obtains the lowest RMSE (12.71), MAE (7.842), and highest R^2 (0.935) among all tested algorithms, showing improvements of 18.99% (RMSE), 5.08% (MAE), and 4.24% (R^2) compared to the best-performing baseline algorithm. Notably, ARIMA (8), SVR (9), RNN (62), LSTM (63), and GRU (64) exhibit subpar performance because they focus only on temporal dependencies. CNN-LSTM (16) performs convolution operations to capture richer characteristics of time; CNN-LSTM-Attention (65) further uses the attention mechanism to obtain the importance of varying time periods. Therefore, they perform better. Nevertheless, these models ignore the disturbances present in historical power data. For instance, although STGCN utilizes multiple context factors and generally excels among baselines, it directly learns power and context features from the raw power and context data and subsequently integrates them to make predictions. In comparison, as listed in Table 4, our GCN-EIF effectively models the EIF by eliminating the EIF from the historical power to improve the learning of inherent power patterns, and then combines the EIF at the prediction time interval to account for future disturbances. Consequently, our GCN-EIF significantly surpasses STGCN.

Table 3. Performance comparison of different prediction algorithms.

Algorithm	RMSE	MAE	R^2
ARIMA	25.55	15.68	0.821
SVR	22.25	14.87	0.833
RNN	19.56	13.55	0.854
LSTM	18.45	12.29	0.862
GRU	18.34	12.31	0.865
CNN-LSTM	17.69	10.39	0.876
CNN-LSTM-Attention	17.25	9.58	0.882
STGCN	15.69	8.26	0.897
GCN-EIF	12.71	7.84	0.935

To confirm the convergence of the proposed prediction model, we present the loss function in each training and testing epoch in Figure 5. It is evident that the value of the loss function decreases rapidly and converges to a small value, showing relatively stable loss performance after 20 epochs. Additionally, both training and testing losses become stable after 40 epochs, suggesting that GCN-EIF converges effectively and demonstrates efficiency in the training process.

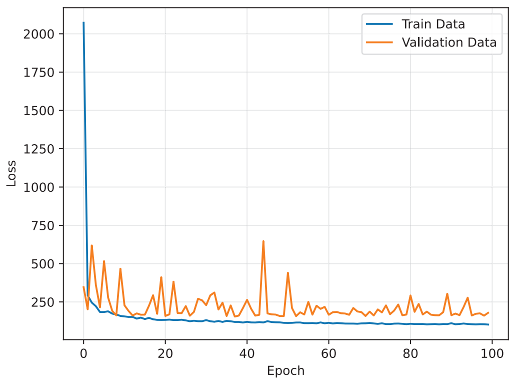


Figure 5. The convergence illustration.

Table 4. RMSE on workdays/weekends/all days.

Algorithms	Workdays	Weekends	All Days
GRU	19.39	18.67	18.34
CNN-LSTM	17.38	18.26	17.69
CNN-LSTM-Attention	16.55	17.32	17.25
STGCN	14.63	15.21	15.69
GCN-EIF	12.96	13.34	12.71

(b) Computational complexity and real-time performance analysis: In addition to prediction accuracy, computational efficiency is crucial for practical wind power forecasting. This section performs a comprehensive analysis of GCN-EIF's computational performance compared with the baselines.

The computational complexity of GCN-EIF can be analyzed by decomposing the model into its principal components:

$$\mathcal{C}_{total} = \mathcal{C}_{GCN} + \mathcal{C}_{LSTM-Att} + \mathcal{C}_{Conv2D} + \mathcal{C}_{fusion} \tag{24}$$

For the multi-graph convolutional networks, the complexity is

$$\mathcal{C}_{GCN} = O(|V|^2 \cdot F \cdot F' + |E| \cdot F') \tag{25}$$

where $|V|$ is the number of nodes (wind farms), $|E|$ is the number of edges in the graph, and F and F' denote the input feature dimension and output feature dimension, respectively.

The LSTM-Attention component exhibits complexity:

$$\mathcal{C}_{LSTM-Att} = O(T \cdot (4H^2 + 4HF + H \cdot T)) \tag{26}$$

where T is the sequence length, H denotes the hidden state dimension, and the last term accounts for the attention mechanism's computation.

For the Conv2D operations used in EIF modeling,

$$\mathcal{C}_{Conv2D} = O(K^2 \cdot C_{in} \cdot C_{out} \cdot H_{out} \cdot W_{out}) \tag{27}$$

where K represents the kernel size, C_{in} and C_{out} stand for the input and output channels, and H_{out} and W_{out} are the output spatial dimensions. The Conv2D operation for EIF mapping takes a tensor of shape [batch_size, time_steps, features] as input, where features represent the various meteorological variables. Here, 2D convolution with a kernel size of 3 × 3 is used to extract spatial–temporal patterns across the weather variables. Then, temporal convolution is performed across time slices of meteorological variables to capture their dynamic evolution.

The overall computational complexity of GCN-EIF is dominated by the graph convolution operations, leading to $O(|V|^2 \cdot F \cdot F' + T \cdot H^2 + K^2 \cdot C_{in} \cdot C_{out} \cdot H_{out} \cdot W_{out})$. In contrast, the complexity of the baselines is as follows: ARIMA: $O(T^2)$; SVR: $O(T^3)$; RNN/LSTM/GRU: $O(T \cdot H^2)$; CNN-LSTM: $O(K^2 \cdot C_{in} \cdot C_{out} \cdot H_{out} \cdot W_{out} + T \cdot H^2)$; and STGCN: $O(|V|^2 \cdot F \cdot F' + T \cdot K^2 \cdot C_{in} \cdot C_{out})$.

The training and inference time of GCN-EIF was evaluated and compared to that of the baselines on a system equipped with an NVIDIA Tesla V100 GPU and Intel Xeon E5-2680 v4 CPU. As listed in Table 5, though GCN-EIF requires more computational resources than simpler methods, its inference time is still within the practical limits for real-time applications.

Table 5. Empirical runtime performance comparison.

Algorithm	Training Time (min/ep)	Inference Time (ms/samp)	GPU Memory (MB)	Model Size (MB)
ARIMA	0.5	2.3	N/A	0.8
SVR	3.2	1.8	N/A	4.5
LSTM	2.5	3.5	128	12.4
CNN-LSTM	3.7	5.2	246	18.7
STGCN	4.1	7.6	312	23.5
GCN-EIF	5.8	8.3	358	26.2

The inference time reported is based on a batch size of 32 samples. For real-time applications, this means that the processing capability is about 120 predictions per second, which is sufficient to meet operational wind power forecasting needs.

To evaluate the real-world applicability of GCN-EIF, experiments were conducted to compare our model with progressively simpler variants: GCN-LSTM-Attention, GCN-LSTM, and the basic GCN model. Figure 6 presents the comparative performances across three critical dimensions: (a) prediction accuracy vs. data arrival rate, (b) processing latency vs. system load, and (c) performance–efficiency trade-offs.

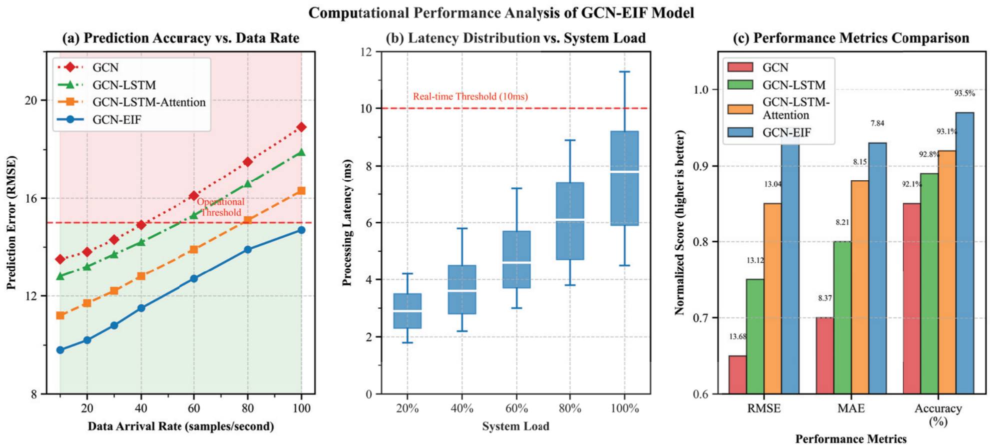


Figure 6. Computational performance analysis of GCN-EIF compared to model variants: (**a**) prediction accuracy (RMSE) vs. data arrival rate, demonstrating GCN-EIF's excellent performance across all data rates; (**b**) processing latency vs. system load, showing that all models maintain real-time performance under typical operational conditions; (**c**) performance–efficiency trade-off visualization where a smaller bubble size indicates a lower MAE (better), suggesting that GCN-EIF achieves the optimal balance between accuracy and computational efficiency.

Figure 6a illustrates that GCN-EIF maintains excellent prediction accuracy across all data arrival rates. Even at high data rates (100 samples/second), the RMSE of GCN-EIF remains below 15, while GCN-LSTM-Attention exceeds this operational threshold at about 80 samples/second. The simpler models (GCN-LSTM and GCN) exceed this threshold at even lower data rates (60 and 40 samples/second), making them unsuitable for high-throughput applications. This indicates that our EIF modeling approach significantly enhances prediction stability under demanding operational conditions.

The processing latency analysis presented in Figure 6b reveals that though GCN-EIF exhibits a slightly higher computational overhead compared to the simpler variants, it still maintains latency well below the critical 10 ms threshold required for real-time grid operations. At 80% system load, GCN-EIF's latency is about 6.8 ms, compared to 6.3 ms for GCN-LSTM-Attention, 5.8 ms for GCN-LSTM, and 5.0 ms for GCN. This marginal increase in processing time (less than 1.8 ms compared to the basic GCN) is negligible in practical applications, yet it yields significant accuracy improvements, validating the computational efficiency of our architecture.

Figure 6c comprehensively shows the performance–efficiency trade-off achieved by each model. This visualization plots relative processing speed against prediction accuracy (R^2), with a small bubble size indicating a lower MAE (better). GCN-EIF achieves the optimal balance, with the highest R^2 value (0.935), lowest RMSE (12.71), and lowest MAE (7.84, indicated by the largest bubble). Notably, although GCN-EIF processes slightly fewer operations per second than the basic GCN model (124 vs. 145 operations/second), it maintains competitive throughput with substantially improved prediction accuracy. This suggests that the additional computational complexity of our model is efficiently implemented and translated directly to meaningful performance improvements.

These results conclusively indicate that GCN-EIF achieves an ideal balance between computational efficiency and prediction accuracy. The model's ability to maintain real-time performance with excellent accuracy makes it highly suitable for operational wind power forecasting systems where both precision and responsiveness are critical requirements. The analysis confirms that our EIF modeling method has minimal computational overhead compared to the significant accuracy gains it provides, indicating that it is easy to implement in practical forecasting applications.

(c) Performance on varying days: For performance evaluation on varying days, the RMSE values of GCN-EIF and four counterparts (GRU, CNN-LSTM, CNN-LSTM-Attention, STGCN, and ST-ResNet) are listed in Table 4. To save space, only the RMSE is presented. Similar conclusions are drawn regarding MAE and R^2. We omitted the results of the other seven baselines because they performed very poorly. The results indicate that GCN-EIF consistently surpasses the other methods across various days, i.e., workdays, weekends, and holidays, highlighting our method's robustness. In Figure 7, the advantage of our proposed method is visually displayed for different days.

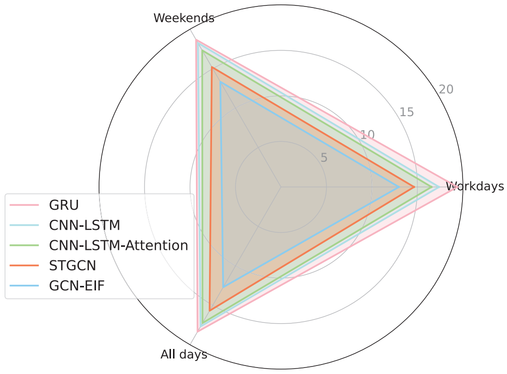


Figure 7. Performance on different days.

5.2. Comparison with Variants of GCN-EIF

(a) Performance comparison on different graphs: Figure 8 illustrates the performance of the GCN when various types of graphs are combined. It is observed that incorporating more graphs leads to greater performance. Table 6 demonstrates a more specific performance comparison, where the RMSE of the two graphs (G_n + G_s) is 4.79% and 13.47%, respectively, higher than that of the neighborhood graph (G_n) and spatial similarity graph (G_s). Meanwhile, there is a minimum value of MAE for the two graphs, and a maximum value of R^2.

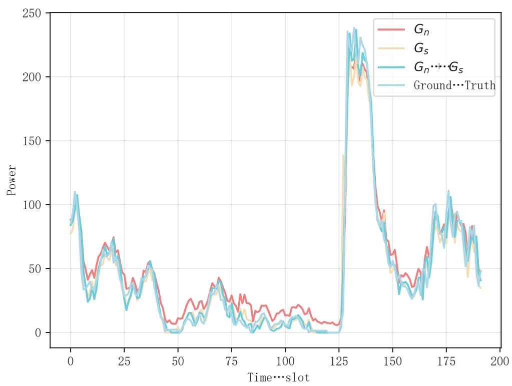


Figure 8. Performance Comparison of GCN with Combined Graph Types.

Table 6. Performance indicators of different graphs.

Algorithm	RMSE	MAE	R^2
G_n	14.69	8.62	0.915
G_s	13.35	8.21	0.921
$G_n + G_s$	12.71	7.84	0.935

(b) Performance comparison on different combinations: From Table 7, it can be seen that the GCN algorithm that only captures the spatial features performs the worst. GCN-

LSTM captures both spatial and temporal features with higher performance. GCN-LSTM-Attention can extract the importance of different points in time and performs better than both of the above algorithms. However, none of them consider the effect of external inference factors on the prediction, so GCN-EIF performs the best. For instance, the RMSE of GCN-EIF is 12.71, which is 0.97, 0.56, and 0.33 higher than that of GCN, GCN-LSTM, and GCN-LSTM-Attention, respectively.

Table 7. Performance indicators of different combinations.

Algorithm	RMSE	MAE	R^2
GCN	13.68	8.37	0.921
GCN-LSTM	13.12	8.21	0.928
GCN-LSTM-Attention	13.04	8.15	0.931
GCN-EIF	12.71	7.84	0.935

(c) Effect of elimination and combination: During the prediction phase, the EIF is eliminated from the historical power, and the STD is combined at the prediction time interval. To confirm the advantages of these techniques, experiments were conducted with or without elimination or combination steps. As listed in Table 8, NoElimination + NoCombination indicates without considering the EIF and directly inputting the historical traffic flow to the GCNs for prediction, and it leads to the lowest performance. Elimination + NoCombination and NoElimination + Combination imply that both elimination and combination lead to a performance higher than that of NoElimination + NoCombination. This confirms the value of eliminating the EIF in historical time intervals and combining it at the prediction time interval. Ultimately, Elimination + Combination achieves the best performance.

Table 8. Effect of elimination and combination.

Method	RMSE	MAE	R^2
NoElimination + NoCombination	14.25	8.66	0.915
Elimination + NoCombination	13.25	8.26	0.925
NoElimination + Combination	13.21	8.19	0.929
Elimination + Combination	12.71	7.84	0.935

5.3. Ablation Study on Attention Mechanisms

To specifically evaluate the contribution of our adaptive attention component, a detailed ablation study was conducted to compare different attention mechanisms and analyze their impact across various temporal horizons, as illustrated in Figure 9.

(a) Comparison of Attention Variants: Figure 9a shows a comprehensive comparison of our GCN-LSTM model under five different attention configurations. The baseline "No Attention" approach simply aggregates LSTM outputs without using any weighting mechanism, leading to an RMSE of 13.68, MAE of 8.37, and R^2 of 0.921. The "Simple Weighting" approach, which applies fixed weights based on temporal proximity, provides a modest performance improvement of 1.7% in RMSE (13.45).

The conventional "Vanilla Attention" mechanism, which computes attention scores using a single-layer perceptron, further reduces RMSE by 1.7% to 13.22. The "Self Attention" approach, which models relationships between different time steps, obtains an RMSE of 13.12, showing an improvement of 0.8% over vanilla attention. Finally, our proposed "Adaptive Attention" mechanism, which dynamically adjusts attention weights based on both temporal position and feature characteristics, contributes to the best performance with an RMSE of 13.04, MAE of 8.15, and R^2 of 0.931.

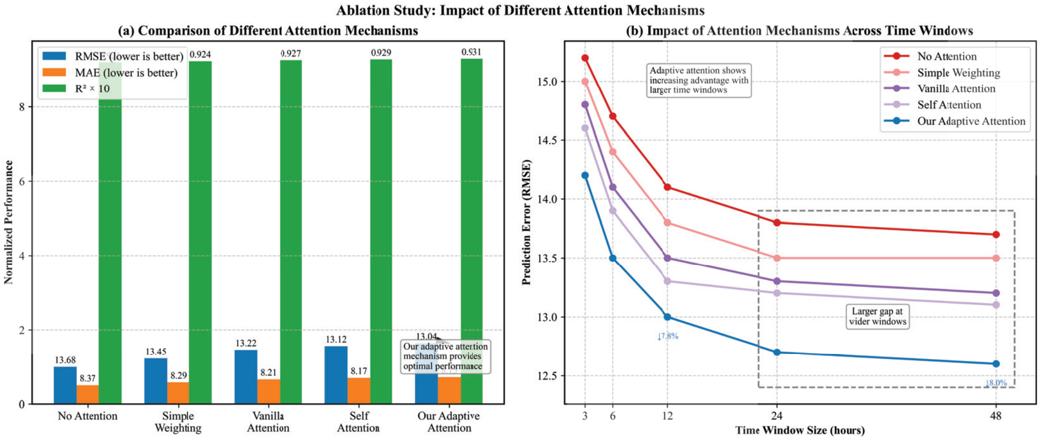


Figure 9. Ablation study on different attention mechanisms: (**a**) performance comparison across five attention variants, demonstrating incremental improvements with more sophisticated attention mechanisms; (**b**) impact of different attention mechanisms across various time window sizes, demonstrating the greater advantage of adaptive attention with larger temporal contexts.

The incremental performance gains across these mechanisms highlight the importance of advanced attention modeling for wind power forecasting. Notably, though all attention variants surpass the baseline, our adaptive approach offers the most substantial benefit, particularly regarding RMSE reduction. This underscores the necessity of dynamically adjusting feature importance based on both temporal context and feature content for capturing the complex dependencies in wind power data.

(b) Impact Across Different Time Windows: The effectiveness of attention mechanisms usually relies on the temporal context length. Figure 9b demonstrates how different attention approaches perform across varying time window sizes, from 3 h to 48 h of historical data.

All models benefit from increased temporal context, as confirmed by the downward slope of all curves. Nevertheless, the performance gap between our adaptive attention mechanism and other approaches becomes larger as the time window increases. With a 3 h window, our approach surpasses the no-attention baseline by 6.6% in RMSE reduction, but this advantage increases to 8.0% with a 48 h window. This widening performance gap indicates that our adaptive attention mechanism becomes more prominent as the temporal complexity of the input data increases.

The enlarged region that highlights the 24 h and 48 h windows indicates that though simpler attention mechanisms show diminishing returns with extremely large windows (as shown by the flattening curves), our adaptive approach can still extract valuable information from extended temporal contexts. This characteristic is particularly crucial for wind power forecasting, where long-range dependencies and recurrent patterns may extend over days instead of hours.

These findings clearly indicate that our adaptive attention mechanism is not only a minor improvement over standard approaches, but a significant enhancement that greatly improves model performance, especially when processing complex temporal sequences. The mechanism's capability to adaptively weight features based on both position and content helps overcome the specific challenges of wind power forecasting, where relevant patterns may emerge at irregular and unpredictable intervals within the historical data.

5.4. Impact of Hyperparameters

(a) Impact of LSTM layers and historical time interval: Figure 10a indicates that the RMSE first decreases and then increases as the number of LSTM layers increases. It can be seen that when six LSTM layers are adopted, our method performs the best. Generally, when the number of LSTM layers further increases, performance drops. A possible explanation for this is that deeper or more complex networks make training more challenging due to the larger number of parameters that need to be learned.

Figure 10b shows the effect of varying the number of historical time intervals. Overall, as the number of historical intervals increases, the RMSE tends to decrease while the training time increases. This outcome is quite expected; incorporating more temporal information generally enhances predictive performance, although it requires a longer training duration.

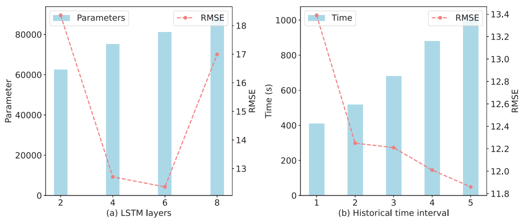


Figure 10. The impact of LSTM layers and historical time interval.

The performance degradation with deeper LSTM networks can be due to overfitting, as shown by our validation loss trend analysis. Figure 10 presents the training and validation loss curves for different LSTM depths, clearly indicating that models with more than six layers demonstrate increasing validation loss despite continued decreases in training loss. To address potential overfitting concerns, we implemented early stopping, dropout regularization (rate = 0.01), and cross-validation across different wind farm clusters. For deployment considerations, we conducted comprehensive computational efficiency analysis, as detailed in Section 5.1b.

(b) Impact of similar nodes:

Figure 11 shows how performance is affected by the number of neighboring nodes and similar nodes that use electricity. When the number of nodes is set to 1, the RMSE is 14.36 and 13.21, which is the same as the case of making predictions without using information from other nodes. Evidently, the network performs the best when there are four neighboring nodes and six similar nodes that use electricity. This is because similar nodes that share power and are close to each other perform similarly in terms of features. To improve its prediction performance, the network will extract these characteristics. When continuing to increase the amount of data, performance does not increase, suggesting that selecting a moderate amount of similar data is beneficial to model training and improves prediction performance generalization ; if too many relevant regions are selected, the network tends to overfit, leading to a sharp decline in prediction performance.

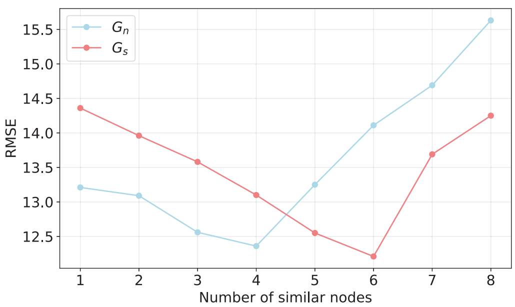


Figure 11. The impact of similar nodes.

6. Conclusions

This paper proposes GCN-EIF for power prediction by modeling spatio-temporal dependencies among inherent power patterns and disturbances. GCN-EIF consists of three phases. First, spatial features are extracted by creating a suitable neighborhood matrix to generate the topological map structure. Then, the temporal characteristics of the data are extracted using the LSTM network and attention mechanism. Finally, the EIF is extracted using convolutional networks. In the prediction, the EIF is eliminated from the historical power to avoid disturbances, thereby capturing inherent power patterns, and the future EIF is combined to account for the disturbances at the prediction time interval. Experiments indicate that GCN-EIF surpasses SOTA methods under diverse conditions, e.g., prediction on various days, varying time intervals, and similar nodes.

Though our GCN-EIF framework exhibits significant improvements over existing methods, there are still several limitations for the present work. First, the model complexity increases computational requirements, posing challenges to deployment in resource-constrained environments. Second, our evaluation is limited to offshore wind farms, and further evaluation is needed to confirm its generalizability to onshore installations with different meteorological dynamics. Third, the current implementation does not consider rare extreme weather events that may cause large deviations from predicted patterns. Future work will overcome these limitations through model compression techniques, use expanded datasets including diverse wind farm types, and adopt specialized modules for extreme event detection and handling.

Author Contributions: Methodology, X.X. and Z.L.; Software, M.F.; Investigation, M.F.; Writing—original draft, X.X. and Z.L.; Writing—review & editing, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: The authors thank the National Natural Science Foundation of China (Grant No. 52469017) and the Applied Basic Research Key Project of Yunnan Province (Grant No. 202401AS070058) for their financial support.

Data Availability Statement: The datasets presented in this article are not readily available because they contain proprietary information from operational wind farms and are subject to confidentiality agreements with the data providers. The raw wind power and weather data include sensitive commercial information that cannot be publicly shared. Additionally, some aspects of the data are part of ongoing research projects at Kunming University of Science and Technology.

Acknowledgments: The authors wish to express their many thanks to the reviewers for their useful and constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Bazionis, I.K.; Georgilakis, P.S. Review of deterministic and probabilistic wind power forecasting: Models, methods, and future research. *Electricity* **2021**, *2*, 13–47. [CrossRef]
2. Wang, Y.; Zou, R.; Liu, F.; Zhang, L.; Liu, Q. A review of wind speed and wind power forecasting with deep neural networks. *Appl. Energy* **2021**, *304*, 117766. [CrossRef]
3. Biswas, A.K.; Ahmed, S.I.; Bankefa, T.; Ranganathan, P.; Salehfar, H. Performance analysis of short and mid-term wind power prediction using ARIMA and hybrid models. In Proceedings of the 2021 IEEE Power and Energy Conference at Illinois (PECI), Urbana, IL, USA, 1–2 April 2021; pp. 1–7.
4. Deng, Y.-C.; Tang, X.-H.; Zhou, Z.-Y.; Yang, Y.; Niu, F. Application of machine learning algorithms in wind power: A review. *Energy Sources Part A Recover. Util. Environ. Eff.* **2021**, *47*, 4451–4471. [CrossRef]
5. Chen, J.; Zhu, Q.; Li, H.; Zhu, L.; Shi, D.; Li, Y.; Duan, X.; Liu, Y. Learning Heterogeneous Features Jointly: A Deep End-to-End Framework for Multi-Step Short-Term Wind Power Prediction. *IEEE Trans. Sustain. Energy* **2020**, *11*, 1761–1772. [CrossRef]
6. Sun, Z.; Zhao, M. Short-Term Wind Power Forecasting Based on VMD Decomposition, ConvLSTM Networks and Error Analysis. *IEEE Access* **2020**, *8*, 134422–134434. [CrossRef]

7. Saint-Drenan, Y.-M.; Besseau, R.; Jansen, M.; Staffell, I.; Troccoli, A.; Dubus, L.; Schmidt, J.; Gruber, K.; Simões, S.G.; Heier, S. A parametric model for wind turbine power curves incorporating environmental conditions. *Renew. Energy* **2020**, *157*, 754–768. [CrossRef]
8. Ozpineci, B. AI/ML and Cybersecurity in Power Electronics. *IEEE Power Electron. Mag.* **2022**, *9*, 38–40. [CrossRef]
9. Zhao, S.; Wang, H. Enabling Data-Driven Condition Monitoring of Power Electronic Systems With Artificial Intelligence: Concepts, Tools, and Developments. *IEEE Power Electron. Mag.* **2021**, *8*, 18–27. [CrossRef]
10. Gupta, A.; Sharma, K.C.; Vijayvargia, A.; Bhakar, R. Very short term wind power prediction using hybrid univariate ARIMA-GARCH model. In Proceedings of the 2019 8th International Conference on Power Systems (ICPS), Jaipur, India, 20–22 December 2019; pp. 1–6.
11. Xu, T.; Du, Y.; Li, Y.; Zhu, M.; He, Z. Interval prediction method for wind power based on VMD-ELM/ARIMA-ADKDE. *IEEE Access* **2022**, *10*, 72590–72602. [CrossRef]
12. Alkesaiberi, A.; Harrou, F.; Sun, Y. Efficient wind power prediction using machine learning methods: A comparative study. *Energies* **2022**, *15*, 2327. [CrossRef]
13. Zhang, Y.; Sun, H.; Guo, Y. Wind power prediction based on PSO-SVR and grey combination model. *IEEE Access* **2019**, *7*, 136254–136267. [CrossRef]
14. Shi, Z.; Liang, H.; Dinavahi, V. Direct Interval Forecast of Uncertain Wind Power Based on Recurrent Neural Networks. *IEEE Trans. Sustain. Energy* **2018**, *9*, 1177–1187. [CrossRef]
15. Niu, Z.; Yu, Z.; Tang, W.; Wu, Q.; Reformat, M. Wind power forecasting using attention-based gated recurrent unit network. *Energy* **2020**, *196*, 117081. [CrossRef]
16. Pan, C.; Wen, S.; Zhu, M.; Ye, H.; Ma, J.; Jiang, S. Hedge Backpropagation Based Online LSTM Architecture for Ultra-Short-Term Wind Power Forecasting. *IEEE Trans. Power Syst.* **2024**, *39*, 4179–4192. [CrossRef]
17. Liu, M.; Qin, H.; Cao, R.; Deng, S. Short-Term Load Forecasting Based on Improved TCN and DenseNet. *IEEE Access* **2022**, *10*, 115945–115957. [CrossRef]
18. Song, Y.; Tang, D.; Yu, J.; Yu, Z.; Li, X. Short-Term Forecasting Based on Graph Convolution Networks and Multiresolution Convolution Neural Networks for Wind Power. *IEEE Trans. Ind. Inform.* **2023**, *19*, 1691–1702. [CrossRef]
19. Liao, W.; Wang, S.; Bak-Jensen, B.; Pillai, J.R.; Yang, Z.; Liu, K. Ultra-short-term Interval Prediction of Wind Power Based on Graph Neural Network and Improved Bootstrap Technique. *J. Mod. Power Syst. Clean Energy* **2023**, *11*, 1100–1114. [CrossRef]
20. Li, Z.; Ye, L.; Song, X.; Luo, Y.; Pei, M.; Wang, K.; Yu, Y.; Tang, Y. Heterogeneous Spatiotemporal Graph Convolution Network for Multi-Modal Wind-PV Power Collaborative Prediction. *IEEE Trans. Power Syst.* **2024**, *39*, 5591–5608. [CrossRef]
21. Tang, J.; Liu, Z.; Hu, J. Spatial-Temporal Wind Power Probabilistic Forecasting Based on Time-Aware Graph Convolutional Network. *IEEE Trans. Sustain. Energy* **2024**, *15*, 1946–1956. [CrossRef]
22. Pryor, S.C.; Barthelmie, R.J.; Bukovsky, M.S.; Leung, L.R.; Sakaguchi, K. Climate change impacts on wind power generation. *Nat. Rev. Earth Environ.* **2020**, *1*, 627–643. [CrossRef]

energies

MDPI

Article

Energy Management for a Fuel Cell Hybrid-Powered Unmanned Aerial Vehicle Based on Optimal Path Planning

Yunpeng Ji, Xingpeng Ling, Xiaojuan Wu * and Jiangping Hu

School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 610056, China; 202522060721@std.uestc.edu.cn (Y.J.); 18200376238@163.com (X.L.); hjp_lzu@163.com (J.H.)

* Correspondence: xjwu@uestc.edu.cn

Abstract

Unmanned Aerial Vehicles (UAVs) present a promising solution for urban logistics, where an effective energy management strategy guided by optimal path planning is crucial for reducing operational costs and extending system lifespan. This study begins by analyzing the wind field distribution in a specific urban area of Chengdu using Computational Fluid Dynamics, and establishes a data-driven power prediction model to evaluate UAV energy consumption. A hybrid wind-field-aware A* with Ant Colony Optimization algorithm is subsequently proposed to compute the optimal flight path that balances energy consumption and distance, generating corresponding power demand profiles for the ensuing energy management strategy. Finally, a Deep Q-Learning (DQN)-based energy management strategy is implemented to regulate power distribution between the fuel cell and the battery, aiming to minimize hydrogen consumption and stabilize the power output of the primary source. Experimental results demonstrate that the proposed path planning method can effectively reduce energy consumption across different scenarios while causing only a marginal increase in travel distance. In addition, the DQN-based strategy significantly suppresses fuel cell power fluctuations at the cost of only a slight increase in hydrogen consumption, thereby demonstrating the effectiveness of the path-planning-informed energy management strategy.

Keywords: fuel cell powered hybrid UAV; wind field distribution; path planning; energy management strategy

1. Introduction

Unmanned Aerial Vehicles (UAVs) are a promising solution for urban "last-mile" logistics, owing to their high maneuverability and independence from ground traffic [1]. However, their power systems must meet conflicting mission requirements: high power density for dynamic maneuvers such as takeoff and landing, and high energy density for sustained cruising. While lithium-ion batteries offer high power density, their energy density is limited; conversely, fuel cells provide high energy density but suffer from slow power response [2]. As a result, a hybrid power system that integrates both energy sources represents the most viable architecture to satisfy the full mission profile [3].

The performance of a UAV hybrid power system hinges critically on its Energy Management Strategy (EMS), which governs the intelligent power allocation between energy sources to minimize operational costs and extend system lifespan. Existing EMS approaches for UAVs can be broadly categorized into rule-based [4–8], optimization-based [9–13], and learning-based methods [14–16]. Due to their dependence on predetermined rules for power distribution, rule-based EMS strategies often fail to adapt effectively to complex flight conditions. Optimization-based methods aim to find globally optimal solutions to allocate the power, but they generally suffer from high computational burdens. Learning-based strategies, especially deep reinforcement learning (RL), have been introduced into energy management for UAV hybrid power systems due to their ability to achieve a superior trade-off between control optimality and computational efficiency, effectively handling system nonlinearity, environmental uncertainty, and high-dimensional state spaces. A health-aware strategy based on a Safe Deep Reinforcement Learning framework was proposed to minimize component degradation costs, including fuel cell voltage drop and battery capacity fade [14]. A Clipped Proximal Policy Optimization agent was used to optimize power distribution while ensuring engine operational safety and reducing speed fluctuations [15]. An RL-based framework was developed to minimize overall energy consumption by dynamically optimizing the vehicle's power control system based on mission objectives and environmental conditions [16].

Current EMS research is largely confined to standardized cyclic flight patterns and consequently cannot be directly extended to mission-specific trajectories. Extensive research has been conducted on path planning to ensure the operational efficiency and safety of UAVs in complex logistics scenarios [17,18]. However, there has been limited research on integrating path planning with EMS to enhance overall system performance. A discrete graph was constructed by sampling the boundaries of quiet zones, and a sampling-based algorithm was utilized to simultaneously optimize the flight path and energy management by scheduling the mode switching between gasoline and electric power [19]. An adaptive Sequential Convex Programming method was proposed to solve the coupled model of trajectory optimization and energy management, which jointly optimizes the flight path to maximize solar energy capture while regulating the power allocation among the photovoltaic system, fuel cell, and battery [20]. A coordinated optimization method was developed to integrate flight trajectory and energy management, utilizing Fuzzy Neural Network Sequential Convex Programming to maximize solar energy utilization and a Convex Quadratic Programming Model Predictive Control strategy for real-time power allocation [21].

However, these studies overlook the impact of the wind field on path planning, which can lead to suboptimal power curves. Moreover, current integrated strategies rarely consider the impact of trajectory-induced power fluctuations on the degradation of the hybrid power system, leading to reduced component lifespan and operational reliability. Given the aforementioned issues, this research makes three key contributions to the field of urban drone logistics:

Novel Holistic Optimization Framework: A holistic framework that integrates wind-inclusive path planning with health conscious, multi-objective energy management is developed.

Power-Cognizant Trajectory Planning: By incorporating a wind-influence factor, hybrid A*-Ant Colony Optimization (HA*-ACO) algorithm is proposed to generate the optimized power demand profiles that serve as critical inputs for subsequent energy management.

Component-Aware Energy Strategy: The resulting energy management system implicitly incorporates component degradation factors, thereby enhancing both economic viability and operational reliability.

The remainder of this paper is organized as follows. Section 2 analyzes the problem and outlines the operational workflow of the logistics mission. Section 3 details the wind-affected flight models and introduces the proposed hybrid path planning method. Section 4 establishes the hybrid power system model and describes the DQN-based energy

https://doi.org/10.3390/en19081854

management strategy. Section 5 presents the experimental validation and results, and the paper is concluded in Section 6.

2. Problem Analysis

The workflow for the fuel cell hybrid UAV performing urban logistics delivery missions, as depicted in Figure 1, can be broadly divided into two main stages:

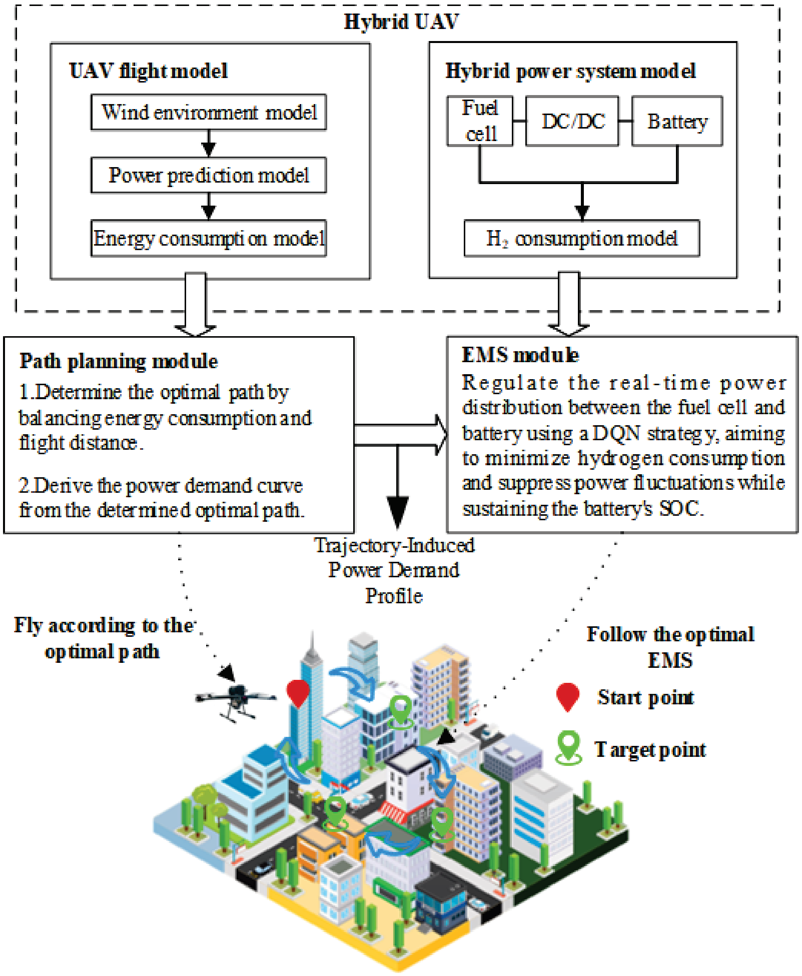


Figure 1. Problem illustration.

Path Planning Stage: A comprehensive UAV flight model is established, incorporating wind environment simulation, power prediction, and energy consumption models to derive the power requirements of the aircraft. Given start and destination points, the optimal flight path is computed by balancing energy consumption and travel distance. This stage also determines the visiting sequence and routes for all delivery points.

EMS Stage: A hybrid power system model is formulated, characterizing the fuel cell and battery to establish an H_2 consumption model. The energy management strategy first configures system capacities based on the UAV power curve. Subsequently, it utilizes the power demand curve to distribute power between the fuel cell and battery, with the objectives of minimizing hydrogen consumption, reducing fuel cell power fluctuations, and constraining battery state of charge (SOC) variations.

Although the workflow is divided into two sequential stages, it is important to emphasize that trajectory planning and energy management are not independent modules. The optimal flight trajectory directly determines the time-varying power demand profile

https://doi.org/10.3390/en19081854

through velocity evolution, altitude changes, and wind disturbances, and this trajectory-induced power demand profile serves as the primary input to the EMS stage.

3. System Modeling and Path Planning

3.1. Wind-Affected UAV Flight Model

Wind speed and wind direction are key factors influencing UAV power consumption. This study employs computational fluid dynamics (CFD) simulation to characterize the wind field distribution in the local urban environment, which includes the continuity equation, momentum conservation equations, and energy conservation equation [22]. The continuity equation describes the transport of fluid mass, while the momentum equations represent the fundamental physical processes of fluid motion, incorporating the effects of pressure gradients, viscous stresses, gravity, and external body forces. The energy equation characterizes the energy transfer processes within the fluid system. The CFD inputs include building geometry, inlet wind velocity, and atmospheric boundary layer conditions, and the wind speed and the wind direction are simulated through the following steps:

(1) Based on the three-dimensional urban voxel grid model, a CFD computational domain of 1100 m $\times$ 700 m $\times$ 360 m is constructed. Following standard CFD modeling requirements, the domain boundaries are positioned at least 5H (300 m) away from the building cluster, where H represents the maximum building height (60 m). This configuration ensures adequate flow development space and prevents boundary effects from influencing the flow field calculations within the target area.

(2) The computational domain is spatially discretized using Finite Volume Method with hexahedral structured mesh. A two-level refinement strategy is implemented: minimum mesh size of 5 m $\times$ 5 m within the building group area, and 10 m $\times$ 10 m mesh size in surrounding areas.

(3) Select the standard k-ε turbulence model and the boundary conditions are set as follows: velocity inlet with 5 m/s uniform wind on the windward side, pressure outlet on leeward side, no-slip walls for ground and building surfaces, and symmetry boundaries for top and lateral domain sides.

(4) Governing equations are solved using Ansys Fluent 2023 R1 with SIMPLE algorithm for pressure-velocity coupling and second-order upwind scheme for spatial discretization. The convergence criteria requires a residual of less than 10^{-4}.

3.2. Power Prediction

To accurately evaluate UAV energy consumption within complex urban wind fields, a data-driven power prediction model is established using a Long Short-Term Memory (LSTM) network. Unlike traditional Recurrent Neural Networks, the LSTM architecture effectively mitigates the vanishing gradient problem, enabling the capture of long-term temporal dependencies inherent in nonlinear flight power profiles. The schematic of the proposed framework is illustrated in Figure 2.

UAV power consumption results from the coupling of kinematic states and external environmental disturbances. To characterize this, the input feature vector X_t at time step t is constructed by integrating flight dynamic parameters with CFD-simulated wind field data:

$$X_t = \left[v_t, a_t, \phi_t, \theta_t, \psi_t, m, v_{wind,t}, \theta_{wind,t}\right]^T \tag{1}$$

where v_t and a_t denote the magnitude of ground velocity and acceleration, respectively; ϕ_t, θ_t, ψ_t represent the roll, pitch, and yaw Euler angles; and m is the total aircraft mass. Critically, the environmental inputs $v_{wind,t}$ and $\theta_{wind,t}$ (wind speed and direction) are

https://doi.org/10.3390/en19081854

derived directly from the CFD simulation, allowing the model to account for aerodynamic power variations caused by local wind field non-uniformity.

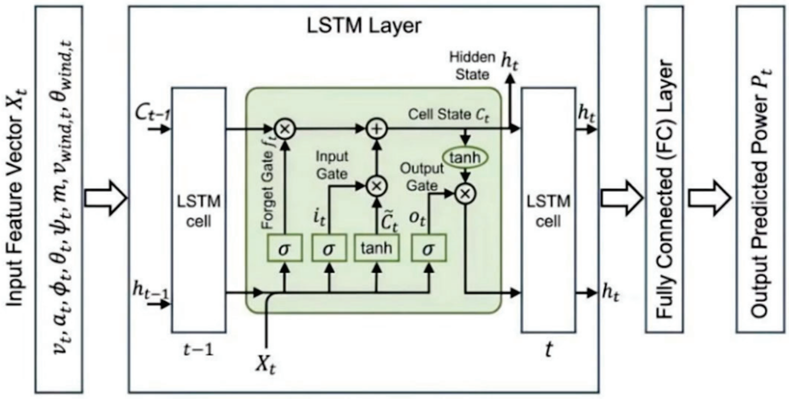


Figure 2. Schematic diagram of LSTM network architecture.

The prediction framework comprises three functional modules: the input layer, the LSTM feature extraction layer, and the fully connected output layer. The sequence of input vectors X_t is fed into the LSTM layer. Leveraging its internal gating mechanisms (input, forget, and output gates), the LSTM unit regulates the information flow, selectively retaining historical flight states in the cell state C_t. This process generates a hidden state h_t that encodes the complex, time-varying correlations between flight maneuvers, wind disturbances, and power demand.

The LSTM model establishes a nonlinear mapping from the input feature vector X_t to the predicted instantaneous power P_t, which can be expressed as:

$$P_t = f_{LSTM}(X_t) \tag{2}$$

All weight matrices and bias vectors involved in the LSTM and fully connected layers are trainable parameters, which are automatically optimized during the network training process. The model is trained using the Adam optimization algorithm with a learning rate of 0.001 and a batch size of 128, optimizing the network parameters to minimize the Mean Squared Error between the predicted and actual measured power profiles.

The LSTM model is developed and evaluated using a documented power consumption dataset for a DJI Matrice 100 UAV [23]. The first 100 flight records from this dataset are partitioned into training and validation sets with a ratio of 8:2, while two distinct flight records, separate from these, are reserved as an independent test set to evaluate the model's ability to generalize. The input to the LSTM model at each time step comprises the UAV's previous flight state variables—including velocity (0.0–18.5 m/s), acceleration (-3.8 to 4.2 m/s^2), yaw angle ($-82.7°$ to $65.9°$), pitch angle ($-21.1°$ to $26.5°$), and roll angle ($-10.5°$ to $14.4°$)—along with environmental conditions such as wind speed (0.0–9.6 m/s) and wind direction (0.0–359.9°).

The prediction performance on the independent test set is presented in Figure 3. This figure illustrates the comparison between the actual recorded power consumption and the model's predicted power over a series of time steps. Power values near 0 W correspond to periods when the UAV is on the ground either pre-takeoff or post-landing. During flight operations, the power consumption generally fluctuates between approximately 400 W and 800 W. The lower end of this range typically represents hovering or stable, low-speed flight, while the higher values indicate more demanding maneuvers such as rapid acceleration, climbing, or compensating for strong winds. The figure visually confirms that the model accurately tracks these actual power variations, maintaining high fidelity even

https://doi.org/10.3390/en19081854

during periods of significant and rapid fluctuation. This strong predictive performance is quantitatively validated by an RMSE of 42.99 W and an R^2 of 0.95 on the test set, confirming the model's high accuracy and generalization capability. The LSTM power prediction model is trained and evaluated using flight data covering wind speeds from 0.0 to 9.6 m/s and wind directions from 0.0° to 359.9°. Higher wind speeds increase aerodynamic drag and require more power to maintain speed or altitude, while strong headwinds lead to higher power consumption during flight, whereas tailwinds can reduce power needs.

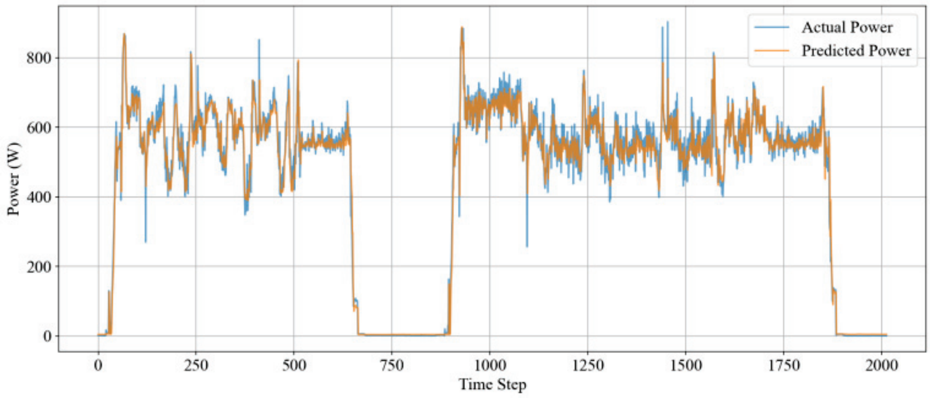


Figure 3. Power prediction on test set.

3.3. UAV Energy Consumption Model

In the context of logistics delivery missions within urban environments, the total energy consumption $E(n_u, n_v)$ between any two consecutive path points n_u and n_v is calculated by integrating power consumption over different flight phases, which is calculated as follows:

$$E(u,v) = E_{acc} + E_{const} = \frac{(P^u_{UAV} + P^v_{UAV})}{2} \cdot \Delta t_{acc} + P^{uv}_{UAV} \cdot (\Delta t - \Delta t_{acc}) \tag{3}$$

where P^u_{UAV} and P^v_{UAV} represent the UAV power at the start and end of the acceleration/deceleration phase, respectively; P^{uv}_{UAV} denotes the power maintained during the constant velocity phase; Δt_{acc} is the duration of the acceleration/deceleration phase; and Δt represents the total time elapsed between the two path nodes.

3.4. UAV Path Planning

As shown in Figure 1, given a set of delivery target nodes, the UAV is required to perform a logistics delivery mission that starts from a start node, visits all target nodes, and finally returns to the origin. In order to minimize the energy consumption and the flight distance during the delivery mission, the path planning task is decomposed into two hierarchical optimization problems: (1) local path optimization between pairs of target nodes; (2) global route optimization that determines the visiting sequence of target nodes.

For local path planning, the start node and the target node are given by the delivery task. Let n_0 denote the start node, and let n_i and n_j refer to the i-th and j-th target nodes, respectively, where $i, j \in \{0, 1, \ldots, K\}$, $i \neq j$, and K is the total number of target nodes in the mission. For any two mission nodes n_i and n_j, the objective is to determine a feasible path p from n_i to n_j that minimizes the accumulated edge cost. Therefore, the local path cost is defined as:

$$C_{ij} = \min_{p \in P(n_i, n_j)} \sum_{(u,v) \in p} c(u,v) \tag{4}$$

https://doi.org/10.3390/en19081854

where $P(n_i, n_j)$ represents the set of all feasible paths from node n_i to node n_j, and p is a feasible path in this set; u and v denote two adjacent nodes along path p; $c(u, v)$ is the cost of traversing the edge between two adjacent nodes u and v, which is calculated as follows:

$$c(u, v) = \omega e D(u, v) + (1 - \omega)E(u, v) \tag{5}$$

where $D(u, v)$ is the Euclidean distance between the two voxel nodes; $E(u, v)$ is the estimated energy consumption between the nodes, calculated using Equation (3); ω ∈ [0, 1] is the weighting coefficient that balances distance and energy objectives, e is a normalization factor used to scale distance and energy to comparable magnitudes.

In this study, this local path optimization problem is solved using a wind-aware A* algorithm. In a conventional A* algorithm, the Euclidean distance is typically used as the heuristic function. In this paper, a wind field factor w_f is introduced to modify the heuristic function, enabling the search process to account for wind-induced energy variations. The evaluation function of the local path search is defined as follows:

$$f(x) = \sum_{(u,v) \in p} c(u, v) + w_f(v_r, \theta)\sqrt{(x_g - x_{n_j})^2 + (y_g - y_{n_j})^2 + (z_g - z_{n_j})^2} \tag{6}$$

where p denotes the path from the mission node n_i to the current node g; the coefficient $w_f(v_r, \theta)$ is a wind-field adjustment factor that modifies the heuristic distance according to the relative wind speed v_r and the angle θ between the UAV's flight direction and the wind direction; (x_g, y_g, z_g) denotes the coordinates of the current voxel node g, and (x_{n_j}, y_{n_j}, z_{n_j}) denotes the coordinates of the target node n_j.

After obtaining the locally optimal path using the A* algorithm, the remaining task is to determine the optimal visiting sequence of these nodes so as to minimize the total mission cost of the delivery tour:

$$\min_{\pi} F_{tour} = \sum_{i=0}^{K}\sum_{j \neq i}^{K} C_{ij}x_{ij} \tag{7}$$

where π is the visiting sequence of target nodes which is defined as $\pi = \{n_0, n_1, n_2, \ldots, n_K, n_0\}$; n_0 is the starting nodes; $n_1, n_2, \ldots, n_K$ are the delivery target nodes; x_{ij} is a binary decision variable defined as:

$$x_{ij} = \begin{cases} 1, & if\ the\ path\ segment\ from\ n_i\ to n_j\ is\ selected \\ 0, & otherwise \end{cases} \tag{8}$$

The global route optimization problem then becomes a traveling salesman problem with energy-aware edge costs. To solve this problem, the ACO algorithm is employed to search for the optimal visiting sequence of the delivery nodes. The probability that ant k moves from node i to node j is defined as [24]:

$$p_{ij}^{(k)} = \frac{[\tau_{ij}(t)]^{\alpha}[\eta_{ij}]^{\beta}}{\sum\limits_{l \in K}[\tau_{il}(t)]^{\alpha}[\eta_{il}]^{\beta}} \tag{9}$$

where $\tau_{ij}(t)$ denotes the pheromone concentration on edge $i \to j$; α and β control the relative influence of pheromone and heuristic information; K represents the set of selectable

https://doi.org/10.3390/en19081854

nodes; η_{ij} is the heuristic value which is defined as the inverse of the path cost between two nodes:

$$\eta_{ij} = \frac{1}{\sum\limits_{(u,v)\in p} c(u,v)} \tag{10}$$

Initially, pheromone values are distributed uniformly across all edges, allowing the ants to explore different candidate routes. As the iterations proceed, pheromone concentrations on high-quality paths are reinforced, enabling the algorithm to gradually converge toward an optimal visiting sequence. After each iteration, the pheromone values are updated according to the following rule:

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \sum_{k=1}^{n} \Delta\tau_{ij}^{k} \tag{11}$$

where ρ is the pheromone evaporation coefficient; $\Delta\tau_{ij}^{(k)} = \frac{Q}{F_{tour}^{(k)}}$ represents the pheromone increment deposited by ant k, and $F_{tour}^{(k)}$ denotes the tour cost obtained by ant k.

During the path planning process, the generated trajectories must satisfy the UAV's kinematic and operational constraints to ensure flight feasibility and safety. In particular, the UAV velocity and acceleration are restricted within their allowable ranges. These constraints describe the physical flight limits of the UAV and guarantee that the planned paths can be executed by the vehicle during real flight operations. The corresponding constraints can be expressed as:

$$\begin{gathered} v_{\min} \le v_t \le v_{\max} \\ a_{\min} \le a_t \le a_{\min} \end{gathered} \tag{12}$$

where v_{min} and v_{max} represent the UAV's minimum and maximum velocity, respectively. Similarly, the UAV's horizontal acceleration and vertical acceleration at n_u must not exceed their respective maximum values. The instantaneous velocity v_t and acceleration a_t are calculated as:

$$v_t = \sqrt{\dot{x}_t^2 + \dot{y}_t^2 + \dot{z}_t^2} \tag{13}$$

$$a_t = \frac{v_t - v_{t-1}}{\Delta t} \tag{14}$$

where $\dot{x}_t$, $\dot{y}_t$ and $\dot{z}_t$ are the velocity components of UAV at time t which can be expressed as:

$$\dot{x}_t = \frac{x_t - x_{t-1}}{\Delta t} \tag{15}$$

$$\dot{y}_t = \frac{y_t - y_{t-1}}{\Delta t} \tag{16}$$

$$\dot{z}_t = \frac{z_t - z_{t-1}}{\Delta t} \tag{17}$$

where x_t, y_t and z_t denote the spatial coordinates of the UAV at time t, which are taken from the optimal sequence of spatial nodes generated by the path planning stage.

Based on this discrete trajectory, the attitude angles θ_t, ψ_t and ϕ_t representing the pitch, yaw, and roll angles, respectively, can also be extracted from the trajectory geometry as follows [25]:

$$\theta_t = \arctan\left(\frac{\dot{z}_t}{\sqrt{x_t^2 + y_t^2}}\right) \tag{18}$$

$$\psi_t = \mathrm{arctan2}(\dot{y}_t, \dot{x}_t) \tag{19}$$

$$\phi_t = \arctan\left(\frac{v_t - \dot{\psi}_t}{g}\right) \tag{20}$$

where g denotes the gravitational acceleration.

Furthermore the localized wind variables $v_{wind,t}$ and $\theta_{wind,t}$ are obtained by mapping each path point (x, y, z) to the CFD-based wind field model established in Section 2. These variables, including the UAV velocity v_t, acceleration a_t, attitude angles θ_t, ψ_t and ϕ_t, local wind speed $v_{wind,t}$, and wind direction $\theta_{wind,t}$, are then fed into the pre-trained LSTM power prediction model to derive the corresponding UAV power demand profile $P_{req}(t)$.

4. DQN-Based EMS

4.1. Hybrid Power System Model

The topology of the hybrid power system is illustrated in Figure 4. Proton Exchange Membrane Fuel Cell (PEMFC) exhibits high efficiency and specific power characteristics, requires a short start-up time, and operates at low temperatures. Therefore, this study selects PEMFC as the main power source for the UAV power system. The lithium battery pack acts as the auxiliary/peak power source and energy buffering unit. The output voltage of the fuel cell stack is generally lower than the bus voltage, so a unidirectional DC/DC converter is used to boost the voltage before connecting to the bus. Meanwhile, the power battery not only supplies power to the system but also needs to absorb energy from the fuel cell for recharging when its SOC is low. Therefore, a bidirectional DC/DC converter is employed to enable two-way energy flow between the battery and the bus.

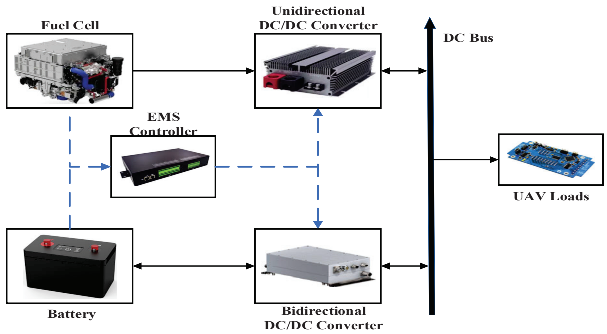


Figure 4. Hybrid power system topology. Black solid arrows indicate the main power flow paths in the hybrid power system, while blue dashed arrows indicate the control or signal interactions associated with the EMS controller.

4.1.1. Fuel Cell

Complex nonlinear multi-input multi-output PEMFC models are difficult to be used for real-time EMS, Therefore, using cftool fitting in MATLAB R2022a, the power-efficiency relationship is established as follows:

$$\eta_{fc}(P_{fc}(t)) = -21.779474P_{fc}^3(t) + 37.499668P_{fc}^2(t) - 33.23206P_{fc}(t) + 61.943587 \tag{21}$$

where P_{fc} is output power of the fuel cell. Accordingly, the instantaneous hydrogen consumption rate is expressed as [26]:

$$\dot{m}_{H_2,\,fc}(t) = \frac{P_{fc}(t)}{\eta_{fc}(P_{fc}(t))LHV_{H_2}} \tag{22}$$

https://doi.org/10.3390/en19081854

where LHV_{H_2} is the lower heating value of hydrogen, representing the chemical energy released per unit mass of hydrogen during combustion (typically 120 MJ/kg). The total hydrogen consumption over the mission horizon is calculated as:

$$m_{H_2,\,fc} = \sum_{t=0}^{T-1} \dot{m}_{H_2,\,fc}(t)\Delta t \tag{23}$$

where Δt is the sampling interval and T is the total number of time steps.

4.1.2. Lithium Battery

The SOC of the lithium battery reflects the ratio of its remaining capacity to its capacity and is modeled as in Equation (22):

$$SOC(t) = SOC(t-1) - A_n \frac{P_{bat}(t)\Delta t}{E_{bat}},\ A_n = \begin{cases} \eta_{bat} & \text{charging} \\ 1/\eta_{bat} & \text{discharging} \end{cases} \tag{24}$$

where $SOC(t)$ represents the battery state of charge at time step t; P_{bat} is the lithium battery power; η_{bat} is the lithium battery charge/discharge efficiency (98%); E_{bat} is the lithium battery capacity.

4.1.3. DC/DC Converter

The DC/DC converters are mainly responsible for energy transfer and voltage regulation. However, this process inevitably involves some energy loss, which is expressed as follows:

$$\eta_{DC/DC} = \frac{P_{\text{in}} - P_{\text{loss}}}{P_{\text{in}}} \tag{25}$$

where $\eta_{DC/DC}$ is the converter efficiency; P_{in} is the input power, and P_{loss} is the total power loss of the converter.

After considering converter losses, the bus-side powers of the fuel cell and the battery can be respectively expressed as:

$$P_{fc,\,bus}(t) = \eta_{unDC} P_{fc}(t) \tag{26}$$

$$P_{bat,\,bus}(t) = \begin{cases} \eta_{BidDC} P_{bat}(t), & P_{bat}(t) \geq 0(discharging) \\ \frac{P_{bat}(t)}{\eta_{BidDC}}, & P_{bat}(t) < 0(charging) \end{cases} \tag{27}$$

where η_{unDC} and η_{BidDC} are the efficiencies of the unidirectional and bidirectional DC/DC converters, respectively.

4.2. EMS Optimization Objectives and Constraints

The objective of the EMS is to optimally allocate the power demand between the fuel cell system and the battery, so as to achieve high hydrogen economy while maintaining battery charge sustainability and reducing fuel cell load fluctuations. Therefore, the EMS problem is formulated as a multi-objective optimization problem as follows:

$$\min_{u(t)} J = \sum_{t=0}^{T-1} \left[\alpha \dot{m}_{H_2,\,fc}(u(t))\Delta t + \beta (SOC(u(t)) - SOC_{ref})^2 + \gamma |u(t)| \right] \tag{28}$$

where $u(t)$ is the control input; α, β and γ are weighting coefficients, and SOC_{ref} is the reference SOC. The first term represents the hydrogen consumption over the control interval. Reducing this term helps improve the fuel economy of the hybrid power system. The second term penalizes the deviation of the battery SOC from its reference value, so as

https://doi.org/10.3390/en19081854

to maintain the battery within a charge-sustaining operating range rather than allowing continuous depletion or excessive charging during the mission. The third term represents the variation in fuel cell output power, which can suppress abrupt load changes and reduce the dynamic operating stress of the fuel cell system, and can be calculated as follows:

$$u(t) = \Delta P_{fc}(t) \tag{29}$$

Thus, the fuel cell power is updated by:

$$P_{fc}(t) = P_{fc}(t-1) + u(t) \tag{30}$$

Based on Equations (22) and (30), the hydrogen consumption rate can be rewritten as:

$$\dot{m}_{H_2,\ fc}(u(t)) = \frac{P_{fc}(t-1) + u(t)}{\eta_{fc}(P_{fc}(t))LHV_{H_2}} \tag{31}$$

The required power $P_{req}(t)$, generated during the path-planning stage, is supplied by the fuel cell and the lithium battery. Therefore, the EMS should satisfy the following power balance:

$$P_{req}(t) = P_{fc,\ bus}(t) + P_{bat,\ bus}(t) \tag{32}$$

Based on Equations (26), (30) and (32), the battery bus-side power can be calculated as follows:

$$P_{bat,bus}(t) \ = P_{req}(t) - \eta_{unDC}\Big(P_{fc}(t-1) + u(t)\Big) \tag{33}$$

By substituting Equation (33) into Equation (27), the battery power can be further expressed explicitly as a function of the control input $u(t)$:

$$P_{bat}(u(t)) = \begin{cases} \frac{P_{req}(t) - \eta_{unDC}\left(P_{fc}(t-1)+u(t)\right)}{\eta_{BidDC}}, & P_{bat,\ bus}(t) \geq 0 \\ \eta_{BidDC}\Big(P_{req}(t) - \eta_{unDC}\Big(P_{fc}(t-1) + u(t)\Big)\Big), & P_{bat,\ bus}(t) < 0 \end{cases} \tag{34}$$

Then, substituting $P_{bat}(t)$ into the battery SOC dynamics in Equation (24), the SOC update can be written explicitly as:

$$SOC(t) = \begin{cases} SOC(t-1) - \frac{\Delta t}{\eta_{bat}\eta_{BidDC}E_{bat}}\Big(P_{req}(t) - \eta_{unDC}\Big(P_{fc}(t-1) + u(t)\Big)\Big), P_{bat,\ bus}(t) \geq 0 \\ SOC(t-1) - \frac{\eta_{bat}\eta_{BidDC}\Delta t}{E_{bat}}\Big(P_{req}(t) - \eta_{unDC}\Big(P_{fc}(t-1) + u(t)\Big)\Big), P_{bat,\ bus}(t) < 0 \end{cases} \tag{35}$$

From the above derivations, each term in the optimization objective is related to $u(t)$. In the optimization process of Equation (28), the following constraints must be satisfied:

(1) Fuel cell power increment constraint

$$0 \leq P_{fc}(u(t)) \leq P_{fc}^{\max} \tag{36}$$

where $P_{fc}^{\max}$ is the maximum output power of the fuel cell system. This constraint ensures that the fuel cell always operates within its allowable power range.

(2) Battery power constraint

$$P_{bat}^{\min} \leq P_{bat}(u(t)) \leq P_{bat}^{\max} \tag{37}$$

where $P_{bat}^{\min} < 0$ and $P_{bat}^{\max} > 0$ denote the charging and discharging power limits of the battery, respectively. This constraint prevents the battery from exceeding its allowable charging and discharging power limits.

https://doi.org/10.3390/en19081854

(3) Battery SOC constraint

$$SOC^{\min} \leq SOC(u(t)) \leq SOC^{\max} \tag{38}$$

where $SOC^{\min}$ and $SOC^{\max}$ are its lower and upper bounds, respectively. This constraint keeps the battery SOC within a safe operating range to avoid over-charge and over-discharge.

4.3. Design of DQN-Based EMS

To address the energy management optimization problem within a continuous state space, this study adopts DQN algorithm. DQN integrates the strengths of Q-Learning with deep neural networks by approximating the action-value function $Q(s, a)$ via a neural network, thereby enabling the effective handling of high-dimensional state inputs and the learning of optimal control strategies.

Key variables reflecting the UAV flight status and the health of the hybrid power system are selected to construct the state vector. At time step t, the state s_t is defined as:

$$s_t = \left\{ SOC(t), P_{fc}(t), P_{req}(t), v(t) \right\} \tag{39}$$

To ensure smooth control and protect the fuel cell from rapid power fluctuations, the action a_t is defined as the change in the fuel cell's output power ΔP_{fc}. The continuous range of this change is uniformly discretized into 9 distinct actions spanning the interval [−0.2, 0.2] kW, as presented in Equation (40):

$$a_t = \{-0.2, -0.15, -0.1, -0.05, 0, 0.05, 0.1, 0.15, 0.2\}(\mathrm{kW}) \tag{40}$$

The agent selects an action a_t at each time step, and the fuel cell power for the subsequent moment is updated as:

$$P_{fc}(t) = P_{fc}(t-1) + a_t \tag{41}$$

The reward function is critical as it encodes the multi-objective optimization problem into the RL framework. It is formulated as the negative of a weighted cost function that reflects the system's lifecycle operating cost, which is given as follows:

$$r_t = -(\left[\alpha\ \dot{\mathrm{m}}_{H_2,\ fc}(t)\Delta t + \beta(SOC(\mathrm{t}) - SOC_{ref})^2 + \gamma\left|\Delta P_{fc}(t)\right|\right]) \tag{42}$$

The objective of the DQN algorithm is to learn the optimal action–value function, which represents the expected cumulative discounted reward obtained by taking action a in state s and subsequently following the optimal policy:

$$Q^*(s, a) = \mathbb{E}[r_t + \gamma \max_{a\prime} Q(s_{t+1}, a\prime)\Big| s_t = s, a_t = a] \tag{43}$$

In this study, the DQN-based EMS follows the standard off-policy learning procedure illustrated in Figure 5. At each time step, the agent first observes the current system state s_t. Based on an ε-greedy policy, an action a_t, representing the increment of fuel cell power $\Delta P_{fc}(t)$, is selected from the discrete action space. The selected action is then applied to the hybrid power system model, resulting in the next state s_{t+1} and an immediate reward $\mathrm{r_t}$. The transition tuple (s_t, a_t, r_t, s_{t+1}) is stored in the replay buffer.

https://doi.org/10.3390/en19081854

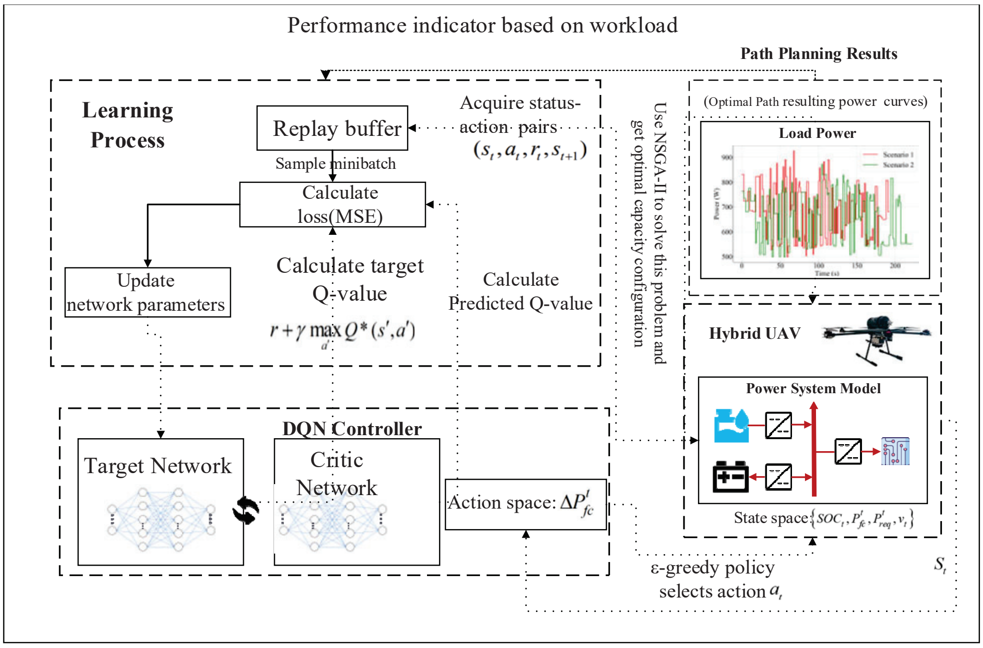


Figure 5. Logical flow of DQN-based EMS.

During training, minibatches of transitions are randomly sampled from this buffer to reduce temporal correlation among data samples. For each sampled transition, the target Q-value y_t is computed using the target network:

$$y_t = r_t + \gamma \underset{a\prime}{max} \hat{Q}(s\prime, a\prime; \theta\prime) \tag{44}$$

where $\gamma \in (0, 1)$ is the discount factor that determines the importance of future rewards; $\hat{Q}$ is the Q-value of the target network; $s\prime$ represents the next state; $a\prime$ denotes a candidate action in the action space; $\theta\prime$ denotes the parameters of the target network. The mean squared error between the predicted Q-value and the target Q-value is defined as follows:

$$L(\theta) = \mathbb{E}_{(s,a,r,s\prime)\sim D}[(y_t - Q(s, a; \theta))^2] \tag{45}$$

where $L(\theta)$ is the loss function; θ denotes the parameters of the main Q-network and it is updated by minimizing $L(\theta)$; $\mathbb{E}$ represents the expectation; $(s, a, r, s\prime) \sim D$ means that the transition tuple is randomly sampled from this buffer; $Q(s, a; \theta)$ is the predicted Q-value produced by the main Q-network.

5. Experiments and Analysis of Results

5.1. Path Planning Results

The proposed method is evaluated in a residential area of Chengdu with dimensions of 800 × 400 × 80 m, which is discretized into a three-dimensional voxel grid. A detailed 3D voxel-based geometric model is constructed using Ansys SpaceClaim, as illustrated in Figure 6. CFD simulations are conducted under representative north wind conditions to characterize the urban wind environment. The simulation results indicate distinct wind field distributions across four altitude intervals: 0–26 m, 27–36 m, 37–60 m, and 61–80 m. Figure 7 illustrates the horizontal wind field distributions corresponding to these height ranges and under the north-wind condition, the steady-state simulation required approximately 4.8 h to converge to a residual threshold of 10^{-4}. The results reveal that lower altitudes show high non-uniformity with channeling effects and wake zones, while higher altitudes approach uniform free-stream conditions as building influence diminishes.

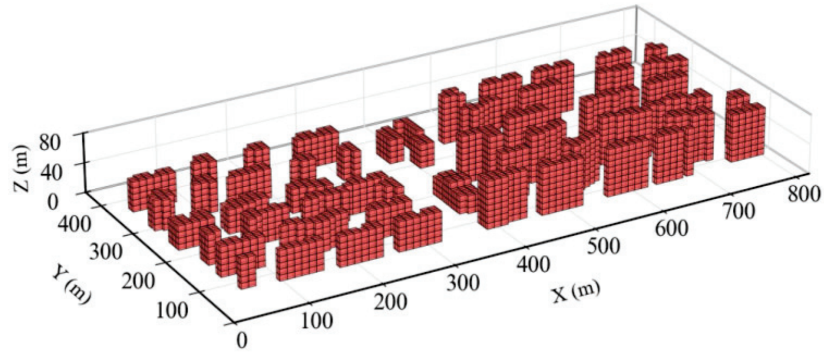


Figure 6. Local urban voxel obstacle model.

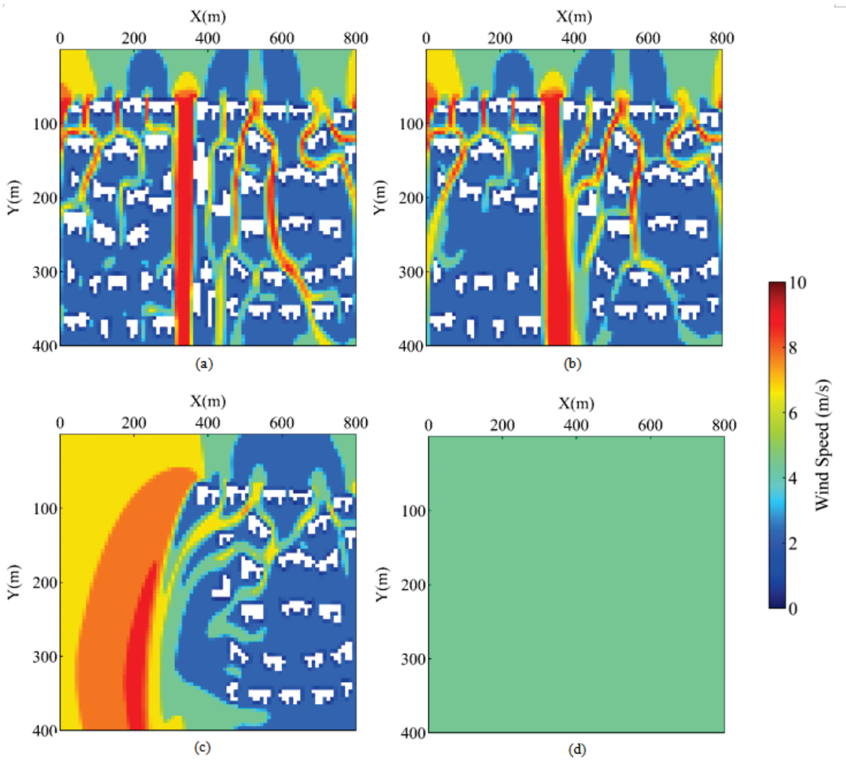


Figure 7. Discrete wind field distribution at various heights under north wind condition: (**a**) 0–26 m; (**b**) 27–36 m; (**c**) 37–60 m; (**d**) 61–80 m.

Two test scenarios are designed to evaluate HA*-ACO's wind-aware path planning performance. The dense delivery scenario features clustered task points causing rapid power fluctuations, with start point (200, 95, 16) and task points at (500, 180, 32), (550, 150, 32), (650, 205, 24), (700, 160, 32), and (750, 190, 24), exhibiting maximum differences of 250 m (x-axis), 45 m (y-axis), and 8 m (z-axis). The multi-layer delivery scenario has significant altitudinal variations leading to higher peak power demands, with start point (160, 60, 16) and task points at (600, 150, 16), (650, 200, 40), (700, 180, 56), and (750, 150, 24), showing a maximum altitude difference of 40 m.

Performance comparison is made against standard ACO without wind consideration. HA*-ACO parameters include: population size 30–100, maximum 200 iterations, 10 initial samples, pheromone importance factor 1–5, heuristic factor 1–5, decay rate 0.1–0.9, weight 0.5, and normalization factor 0.01. Standard ACO uses identical parameter ranges. The computational complexity of the HA*-ACO framework mainly depends on the scale of the voxel search space in the wind-aware A* stage, as well as the ant population size, the number of iterations, and the number of delivery nodes in the ACO stage. Under the

https://doi.org/10.3390/en19081854

parameter settings adopted in this study, the overall computational scale is on the order of 7.2×10^5 basic operations.

Figures 8 and 9 illustrate that the HA*-ACO algorithm generates trajectories with significantly reduced altitude oscillations and fewer abrupt directional changes compared with the wind-agnostic ACO approach. In Scenario 1, the proposed method maintains relatively stable cruising altitudes and avoids unnecessary climb–descent cycles, thereby reducing vertical acceleration demands. In Scenario 2, instead of multiple fragmented altitude adjustments, the HA*-ACO strategy performs a single structured climb followed by a controlled descent.

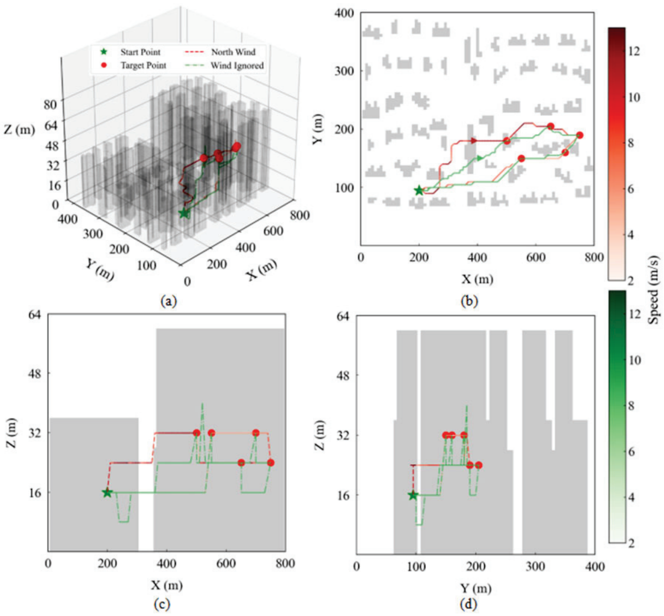


Figure 8. Comparison of UAV path planning results in a dense delivery scenario: (**a**) 3D view; (**b**) Z-axis projection; (**c**) X-axis projection; (**d**) Y-axis projection.

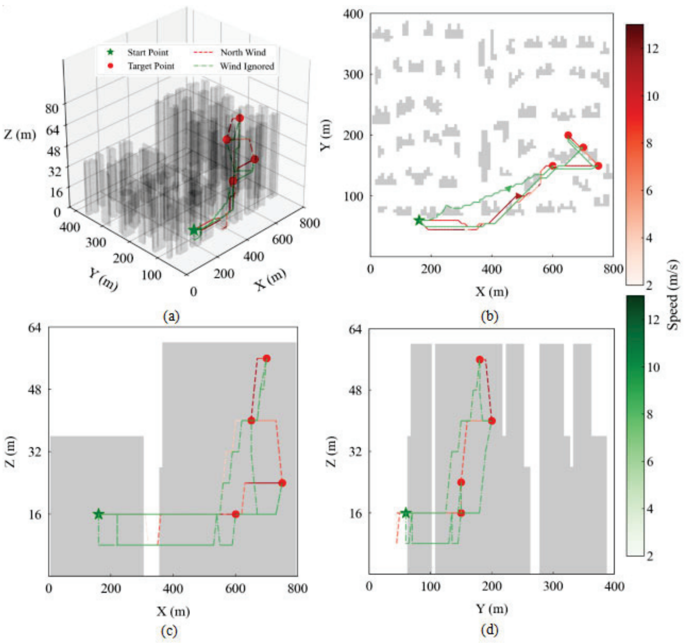


Figure 9. Comparison of UAV path planning results in a multi-layer delivery scenario: (**a**) 3D view; (**b**) Z-axis projection; (**c**) X-axis projection; (**d**) Y-axis projection.

https://doi.org/10.3390/en19081854

These trajectory characteristics directly influence the power demand profile. Frequent altitude changes and sharp maneuvering typically result in rapid thrust variations and high transient power spikes. By suppressing unnecessary vertical motion and smoothing directional transitions, wind-aware planning produces a more continuous and gradually varying power demand curve. This smoother load profile provides a robust foundation for the subsequent DQN-based energy management strategy.

After obtaining the optimal paths for the UAV under both delivery scenarios, the optimal power curves of the UAV can be derived, which are presented in Figure 10. Scenario 1 has frequent power fluctuations, and Scenario 2 has a greater peak power demand. The erratic trajectory of the standard ACO algorithm results in highly fluctuating power curves, with significant peaks in excess of 1300 W. However, the HA*-ACO path produces a more stable power profile.

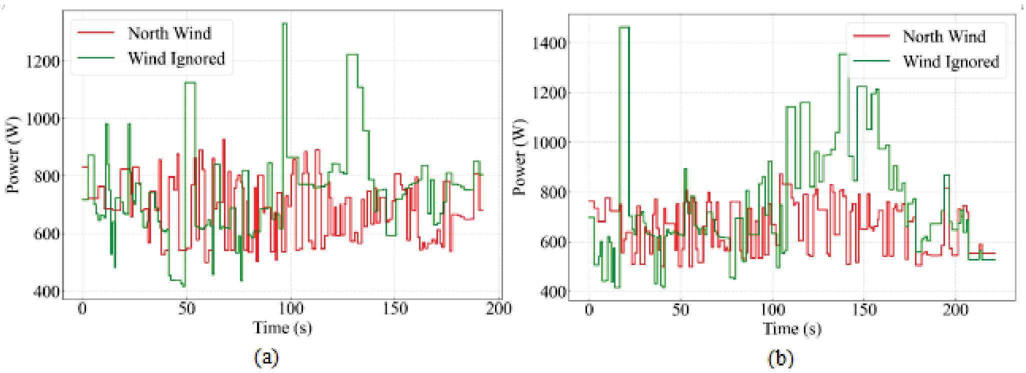


Figure 10. Comparison of power curves in different scenarios: (**a**) dense delivery scenario; (**b**) multi-layer delivery scenario.

Table 1 quantitatively summarizes the comparative performance between wind-aware HA*-ACO planning and the standard wind-agnostic ACO approach. In Scenario 1 (dense delivery), the proposed HA*-ACO algorithm achieves an 8.8% reduction in total energy consumption while incurring only a marginal 0.9% increase in flight distance. This indicates that minor path length adjustments, when guided by wind-field awareness, can significantly reduce aerodynamic resistance and unnecessary vertical maneuvers, thereby improving overall energy efficiency. In Scenario 2 (multi-layer delivery), the advantage becomes even more pronounced. The proposed method reduces energy consumption by 15%, whereas the flight distance increases by only 2.8%. This demonstrates that, particularly in missions involving substantial altitude variations, incorporating wind effects into trajectory optimization yields considerable energy savings with negligible compromise in mission feasibility. Overall, the results confirm that wind-aware trajectory planning provides a favorable trade-off between energy efficiency and travel distance, making it practically superior to conventional planning strategies that ignore environmental disturbances.

Table 1. Energy consumption and distance results in two delivery scenarios.

Scenario ID	Energy (KJ) (Wind Ignored)	Energy (KJ) (North Wind)	Energy Reduction	Distance (m) (Wind Ignored)	Distance (m) (North Wind)	Distance Increase
1	143.74	131.11	8.8%	1222.19	1233.55	0.9%
2	170.03	144.52	15%	1412.11	1451.72	2.8%

https://doi.org/10.3390/en19081854

5.2. EMS Results

The proposed DQN-based energy management strategy is benchmarked against a hydrogen-consumption-oriented DP approach. In the EMS simulations, the constraint bounds were set as: P_{fc}^{max} = 1 kW, P_{bat}^{min} = −600 W, P_{bat}^{max} = 600 W, SOC^{min} = 0.2 and SOC^{max} = 0.8. The DQN agent was trained for 500 epochs with a replay buffer size of 10,000 and a learning rate of 0.001. The training was conducted offline using an NVIDIA RTX 3060 GPU with 32 GB RAM, the total training time was approximately 1.6 h and the sampling interval Δt was 1s. A smaller time step provides finer control resolution but increases computational cost, whereas a larger time step reduces computational burden at the expense of control precision. Through this iterative interaction-learning process, the agent gradually converged to an effective energy management policy that balances hydrogen consumption, fuel cell power smoothness, and battery SOC regulation. As illustrated in Figure 11, the DQN-based strategy yields a substantially smoother fuel cell power profile than the DP approach. This improvement is particularly evident in the dense delivery scenario during the interval of 75–125 s, where rapid and potentially harmful power fluctuations are effectively mitigated.

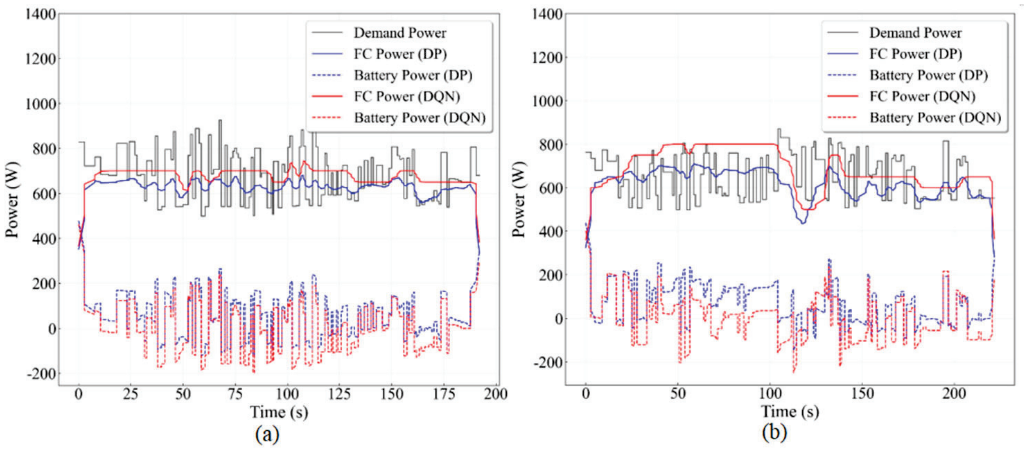


Figure 11. Power allocation results for optimal workload based on DP and DQN based strategies for different scenarios: (**a**) dense delivery scenario; (**b**) multi-layer delivery scenario.

In addition, the lithium battery absorbs short-term power variations induced by transient load demands. In the multi-layer delivery scenario, the battery power under the DQN strategy remains within ±200 W during the period of 100–150 s, whereas the DP-based strategy exhibits pronounced and irregular power oscillations. These results demonstrate the superior capability of the DQN-based strategy in smoothing power distribution while accommodating transient power demands.

The superior performance of the DQN strategy in maintaining the battery's SOC is demonstrated in Figure 12. In Scenario 1, the DQN strategy successfully arrests the SOC decline after t = 100 s and initiates a gradual recovery, stabilizing it near the initial 0.600 level. The advantage is even more pronounced in Scenario 2, where the DQN strategy maintains the SOC within a tight band around its initial value. In stark contrast, the DP-based approach, solely focused on minimizing fuel use, leads to a continuous and significant depletion of the SOC in both scenarios. In addition, the DQN-based strategy maintains the battery SOC within a narrow band around its initial level as shown in Figure 12, whereas the DP-based strategy results in continuous SOC depletion. Maintaining SOC within a moderate operating range helps avoid deep discharge and extreme SOC conditions, thereby contributing to extend the battery lifespan.

https://doi.org/10.3390/en19081854

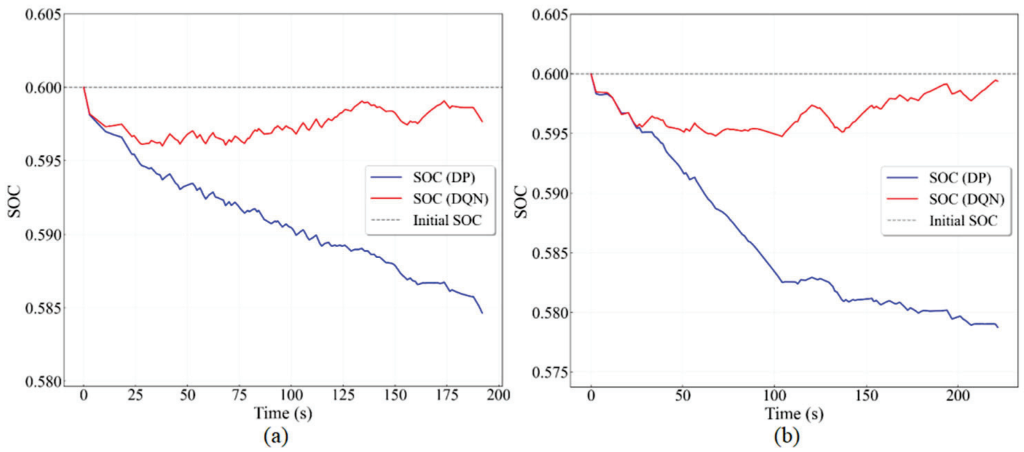


Figure 12. Comparison of Lithium battery SOC curves for DP and DQN strategies in different scenarios: (**a**) dense delivery scenario; (**b**) multi-layer delivery scenario.

To further demonstrate the synergy between mission planning and energy management, the performance of the DQN-based strategy is also evaluated on non-optimal flight paths, which are generated using wind-agnostic path planning results. The corresponding quantitative comparison is summarized in Table 2, which systematically quantifies the trade-offs among the three strategies in terms of hydrogen consumption and fuel cell power fluctuation. While the DP-based strategy marginally excels in minimizing instantaneous hydrogen consumption, the DQN-based strategy achieves near-optimal fuel economy while delivering substantially smoother power allocation. As summarized in Table 2, the DQN strategy reduces the average fuel cell power change rate from 7.54 W/s to 4.69 W/s in Scenario 1, corresponding to a 37.8% reduction relative to the DP benchmark. Similarly, in Scenario 2, the average power change rate achieved by the DQN strategy is reduced to 5.86 W/s, corresponding to a 36.8% decrease compared with the DP strategy's 9.27 W/s.

Table 2. EMS performance comparison on the optimal path.

Scenario	Strategy	Demand Power	Hydrogen Consumption (g)	Avg. Power Change (W/s)	Hydrogen Change (%)	Avg. Power Change (%)
Dense delivery	DP	Optimal	1.823	7.54	0%	0%
	DQN	Optimal	1.872	4.69	+2.7%	−37.8%
	DQN	Non-optimal	2.210	5.31	+21.2%	−29.6%
Multi-layer delivery	DP	Optimal	1.990	9.27	0%	0%
	DQN	Optimal	2.093	5.86	+5.2%	−36.8%
	DQN	Non-optimal	2.540	6.77	+27.6%	−27.0%

In terms of hydrogen consumption, the DQN-based strategy incurs only a modest increase of 2.7% in Scenario 1 and 5.2% in Scenario 2 relative to the DP-based benchmark. In contrast, operating the DQN-based energy management strategy on non-optimal flight paths leads to a pronounced performance degradation. Specifically, in the dense delivery scenario, hydrogen consumption increases by 18.0% (from 1.872 g to 2.210 g), corresponding to an energy increase from approximately 224,640 J to 265,200 J, accompanied by a 13.2% increase in the average power change rate (from 4.69 W/s to 5.31 W/s). A similar deterioration is observed in the multi-layer delivery scenario. These results explicitly demonstrate that optimal path planning not only improves fuel economy but also generates a less aggressive power demand profile, which is essential for enhancing the effectiveness and durability of the hybrid power system.

https://doi.org/10.3390/en19081854

In conclusion, when compared to a benchmark DP strategy, the DQN-based approach achieves a significant reduction in fuel cell power fluctuations while ensuring robust state-of-charge maintenance for the lithium battery. This is accomplished at the cost of a marginal increase in hydrogen consumption. From a system lifecycle perspective, the DQN strategy presents a more economical choice, as the significant cost of fuel cell replacement far outweighs the minor expense of increased daily hydrogen use. Furthermore, the results prove that an energy management strategy informed by optimal path planning is essential for simultaneously achieving low operational costs and extended system lifespans. From a system lifecycle perspective, the DQN-based strategy represents a more sustainable operational choice. Although it incurs a marginal increase in hydrogen consumption, the substantial reduction in fuel cell power fluctuations significantly lowers long-term degradation rates. Considering that fuel cell stack replacement costs far exceed daily hydrogen expenses, prioritizing power smoothness ultimately leads to improved economic viability and enhanced system durability.

6. Conclusions

This paper targets the urban drone logistics delivery scenario and addresses the critical impact of wind fields on flight performance. It presents an integrated optimization approach that first establishes an energy-efficient flight path considering both energy consumption and distance. Then, based on the resulting power profile, it devises an optimal energy management strategy. This strategy facilitates the entire drone delivery process by optimizing for minimal hydrogen consumption and power fluctuation. The main research contributions and findings are as follows:

Develop the wind-aware UAV power prediction model. To accurately characterize the influence of complex urban wind fields on UAV energy consumption, this study first establishes a high-fidelity three-dimensional wind environment model based on CFD. Leveraging the CFD-simulated wind speed and direction data, a data-driven UAV power prediction framework based on a LSTM network is developed. By jointly incorporating flight kinematic states and localized wind field features, the proposed model is capable of capturing the nonlinear and time-varying coupling between environmental disturbances and power demand. Experimental validation on real flight data demonstrates that the proposed model achieves a root mean square error of 42.99 W and a coefficient of determination of 0.95, confirming its high prediction accuracy and strong generalization capability. This wind-aware power prediction model provides a reliable foundation for both trajectory optimization and subsequent energy management strategy design.

For solving the multi-objective path planning problem under urban wind disturbances, a novel hybrid path planning algorithm, termed HA*-ACO, is proposed. The method integrates a wind-field-aware enhanced A* algorithm for local trajectory optimization with the ACO algorithm for global waypoint sequencing. By explicitly embedding wind-induced energy consumption into the heuristic evaluation, the proposed algorithm generates energy-efficient flight paths that mitigate unnecessary altitude changes and abrupt maneuvers. Simulation results in representative urban delivery scenarios show that the HA*-ACO algorithm reduces total energy consumption by up to 15% compared with wind-agnostic planning methods, while incurring only a marginal increase of 0.9–2.8% in flight distance. These results demonstrate that incorporating wind awareness into trajectory planning is essential for improving energy efficiency without compromising mission feasibility.

Formulate an energy management strategy based on optimal path analysis. Building upon the optimized power demand profiles obtained from wind-aware path planning, this study further develops a DQN-based energy management strategy for a fuel cell–battery hybrid UAV power system. The energy management problem is formulated as a Markov

https://doi.org/10.3390/en19081854

Decision Process, with the objective of minimizing lifecycle operating costs by jointly considering hydrogen consumption, fuel cell power fluctuation, and battery state-of-charge regulation. Compared with a benchmark DP strategy, the proposed DQN-based EMS a achieves a 36.8–37.8% reduction in the average fuel cell power change rate at the cost of only a 2.7–5.2% increase in hydrogen consumption. Moreover, operating the EMS on an optimally planned trajectory enables further performance gains: hydrogen consumption is reduced by up to 17.6%, and power fluctuations are suppressed by up to 13.4% compared with non-optimal paths. These results clearly indicate that trajectory optimization and energy management are strongly coupled, and that optimal path planning plays a critical role in alleviating power stress and extending fuel cell system lifespan.

Overall, the proposed integrated framework demonstrates that trajectory optimization and energy management are strongly coupled through the power demand profile. Wind-aware path planning reshapes the UAV's velocity and altitude evolution, thereby directly influencing the temporal characteristics of power demand. In turn, the EMS responds to this trajectory-induced load pattern by allocating power between the fuel cell and battery to balance hydrogen consumption and component stress. By jointly optimizing these two stages rather than treating them independently, the proposed framework achieves synergistic improvements in energy efficiency, power smoothness, and component durability. This coupling-aware design provides a practical and effective solution for urban hydrogen-powered UAV logistics.

Future work will extend this framework to multi-UAV cooperative delivery systems, where additional challenges arise in coordinated trajectory planning, real-time conflict resolution, and system-level energy optimization under dynamic mission and environmental uncertainties.

Author Contributions: Conceptualization, Y.J. and J.H.; methodology, Y.J.; software, Y.J.; validation, X.L.; formal analysis, Y.J.; investigation, X.L.; resources, J.H.; data curation, X.L.; writing—original draft preparation, Y.J.; writing—review and editing, X.W. and J.H.; visualization, X.L.; supervision, X.W. and J.H.; project administration, X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the International Science and Technology Innovation Cooperation Project of the Science and Technology Department of Sichuan Province (Grant No. 2025YFHZ0134).

Data Availability Statement: The data presented in this study are available on request from the corresponding author, the data are not publicly available due to project confidentiality restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jazairy, A.; Persson, E.; Brho, M.; von Haartman, R.; Hilletofth, P. Drones in last-mile delivery: A systematic literature review from a logistics management perspective. *Int. J. Logist. Manag.* **2025**, *36*, 1–62. [CrossRef]
2. Abdulrahman, A.G.A.Q.; Qasem, N.A.A.; Abdelrahman, W.G.; Alkhulaifi, Y.M.; Almatrafi, E.; Alqahtani, N. A review of powering unmanned aerial vehicles by clean and renewable energy technologies. *Sustain. Energy Technol. Assess.* **2025**, *73*, 104150. [CrossRef]
3. Kwasiborska, A.; Stelmach, A.; Jabłońska, I. Quantitative and Comparative Analysis of Energy Consumption in Urban Logistics Using Unmanned Aerial Vehicles and Selected Means of Transport. *Energies* **2023**, *16*, 6467. [CrossRef]
4. Quintana, J.A.; Bordons, C.; Esteban, S.; Delgado, J. Hybrid Powerplant Design and Energy Management for UAVs: Enhancing Autonomy and Reducing Operational Costs. *Energies* **2025**, *18*, 3101. [CrossRef]
5. Engin, K.O.; Cicek, D.B.; Ozkan, A. A rule-based energy management technique considering altitude energy for a mini UAV with a hybrid power system consisting of battery and solar cell. *Energy Convers. Manag.* **2023**, *287*, 117106.
6. Lei, Z.; Gu, C.; Li, F.; Zhou, Y.; Yuan, W. An optimal fuzzy logic-based energy management strategy for a fuel cell/battery hybrid power unmanned aerial vehicle. *Aerosp. Sci. Technol.* **2022**, *129*, 107821.

https://doi.org/10.3390/en19081854

7. Lee, J.; Yu, B.; Choi, K.; Kim, T. Power management strategy of hybrid fuel cell drones for flight performance improvement based on various algorithms. *Energies* **2022**, *15*, 1932.
8. Gao, Y.; Qiao, Z.; Pei, X.; Wu, G.; Bai, Y. Design of energy-management strategy for solar-powered UAV. *Sustainability* **2023**, *15*, 14972. [CrossRef]
9. Liu, H.; Yao, Y.; Wang, J.; Qin, Y.; Li, T. A control architecture to coordinate energy management with trajectory tracking control for fuel cell/battery hybrid unmanned aerial vehicles. *Int. J. Hydrogen Energy* **2022**, *47*, 15236–15253. [CrossRef]
10. Cheng, Z.; Liu, H.; Yu, P.; Zhu, L.; Sun, T.; Yao, Y. Energy management for fuel cell/battery hybrid unmanned aerial vehicle. *Int. J. Electrochem. Sci.* **2021**, *16*, 210919. [CrossRef]
11. Tian, W.; Liu, L.; Zhang, X.; Yang, D. Double-layer fuzzy adaptive NMPC coordinated control method of energy management and trajectory tracking for hybrid electric fixed wing UAVs. *Int. J. Hydrogen Energy* **2022**, *47*, 39239–39254. [CrossRef]
12. Narimanov, J.; Abdujabarov, N. Optimization strategies for energy management systems of solar-powered unmanned aerial vehicles. *Acad. Open* **2025**, *10*, 10638. [CrossRef]
13. Tian, W.; Liu, L.; Zhang, X.; Shao, J.; Ge, J. Adaptive hierarchical energy management strategy for fuel cell/battery hybrid electric UAVs. *Aerosp. Sci. Technol.* **2024**, *146*, 108938. [CrossRef]
14. Gao, Q.; Lei, T.; Yao, W.; Zhang, X. A health-aware energy management strategy for fuel cell hybrid electric UAVs based on safe reinforcement learning. *Energy* **2023**, *283*, 129092. [CrossRef]
15. Chen, Y.; Wang, W.; Yang, C.; Liang, B.; Liu, W. An efficient energy management strategy of a hybrid electric unmanned aerial vehicle considering turboshaft engine speed regulation: A deep reinforcement learning approach. *Appl. Energy* **2025**, *390*, 125837. [CrossRef]
16. Wang, Z.; Xing, Y. Energy consumption optimisation for unmanned aerial vehicle based on reinforcement learning framework. *Int. J. Powertrains* **2024**, *13*, 75–94. [CrossRef]
17. Wu, Y.; Low, K.H. An adaptive path replanning method for coordinated operations of drone in dynamic urban environments. *IEEE Syst. J.* **2021**, *15*, 4600–4611. [CrossRef]
18. Shao, J.; Cheng, J.; Xia, B.; Yang, K.; Wei, H. A novel service system for long-distance drone delivery using the 'Ant Colony+A*' algorithm. *IEEE Syst. J.* **2021**, *15*, 3348–3359. [CrossRef]
19. Manyam, S.G.; Casbeer, D.W.; Darbha, S.; Weintraub, I.E.; Kalyanam, K. Path planning and energy management of hybrid air vehicles for urban air mobility. *IEEE Robot. Autom. Lett.* **2022**, *7*, 10176–10183. [CrossRef]
20. Tian, W.; Liu, L.; Zhang, X.; Shao, J. Flight trajectory and energy management coupled optimization for hybrid electric UAVs with adaptive sequential convex programming method. *Appl. Energy* **2024**, *364*, 123166. [CrossRef]
21. Tian, W.; Liu, L.; Zhang, X.; Shao, J.; Ge, J. A coordinated optimization method of energy management and trajectory optimization for hybrid electric UAVs with PV/Fuel Cell/Battery. *Int. J. Hydrogen Energy* **2024**, *50*, 1110–1121. [CrossRef]
22. Santiago, J.L.; Martilli, A. A dynamic urban canopy parameterization for mesoscale models based on computational fluid dynamics Reynolds-averaged Navier-Stokes microscale simulations. *Bound.-Layer Meteorol.* **2010**, *137*, 417–439. [CrossRef]
23. Rodrigues, T.A.; Patrikar, J.; Choudhry, A.; Feldgoise, J.; Arcot, V.; Gahlaut, A.; Lau, S.; Moon, B.; Wagner, B.; Matthews, H.S.; et al. In-flight positional and energy use data set of a DJI Matrice 100 quadcopter for small package delivery. *Sci. Data* **2021**, *8*, 8. [CrossRef]
24. Dorigo, M.; Maniezzo, V.; Colorni, A. Ant system: Optimization by a colony of cooperating agents. *IEEE Trans. Syst. Man Cybern. B* **1996**, *26*, 29–41. [CrossRef]
25. Beard, R.W.; McLain, T.W. *Small Unmanned Aircraft: Theory and Practice*; Princeton University Press: Princeton, NJ, USA, 2012.
26. Quan, S.; Wang, Y.-X.; Xiao, X.; He, H.; Sun, F. Real-time energy management for fuel cell electric vehicle using speed prediction-based model predictive control considering performance degradation. *Appl. Energy* **2021**, *304*, 117845. [CrossRef]

https://doi.org/10.3390/en19081854

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Energies Editorial Office
E-mail: energies@mdpi.com
www.mdpi.com/journal/energies

MDPI

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-8880-1