KAMPALA INTERNATIONAL UNIVERSITY
KIU
Journal of Science, Engineering and Technology

KJSET Research Article

A review of SCADA-based predictive models for distribution network load forecasting

Kajjuba Hubaine G[1], Ibrahim Sani Madugu[2], Abdullahi Bala Kunya[3]

[1,2,3] Department of Electrical Engineering, Kampala International University, Uganda
g.kajjuba@studwc.kiu.ac.ug, maduguemir@kiu.ac.ug, and abkunya@kiu.ac.ug
Corresponding Author: Kajjuba Hubaine G; g.kajjuba@studwc.kiu.ac.ug

Paper history:
Received 22 April 2026
Accepted in revised form 04 June 2026

Keywords
SCADA Systems; Predictive Load Forecasting; Power Distribution Networks; Smart Grid Analytics; Machine Learning; Hybrid Forecasting Models; Distribution Load Profiling.

Abstract

The increasing complexity of modern power distribution networks, driven by renewable energy integration, distributed generation, and dynamic electricity demand patterns, has created a strong need for intelligent and adaptive load forecasting systems. This review examines the role of Supervisory Control and Data Acquisition (SCADA) systems in predictive load forecasting for power distribution networks, emphasizing their contribution to real-time monitoring, operational automation, and data-driven decision-making. The study evaluates conventional statistical approaches, machine learning techniques, deep learning architectures, and hybrid forecasting models used for predictive load profiling. The review highlights that machine learning-based methods generally offer superior forecasting accuracy and adaptability compared to traditional statistical models, particularly in handling nonlinear and high-dimensional power system data. Hybrid forecasting methods further improve robustness and predictive performance by integrating multiple analytical techniques and leveraging SCADA-derived operational information. However, despite significant progress, key challenges remain, including limited feeder-level forecasting, weak real-time implementation, fragmented SCADA integration, cybersecurity vulnerabilities, scalability constraints, and poor interpretability of advanced artificial intelligence models. The study identifies emerging opportunities in explainable AI, federated learning, edge computing, digital twins, transfer learning, reinforcement learning, and cyber-resilient forecasting as critical directions for future research and implementation. Overall, SCADA-based predictive load forecasting is positioned as a foundational component of next-generation smart grids, capable of improving energy efficiency, enhancing reliability, enabling adaptive control, and supporting sustainable power distribution in increasingly decentralized electricity networks.

Nomenclature and units

MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MSE	Mean Squared Error
MV	Medium Voltage
PMU	Phasor Measurement Unit
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
RTU	Remote Terminal Unit
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SCADA	Supervisory Control and Data Acquisition

1. Introduction

Modern power distribution systems have evolved from conventional manually operated networks into highly automated, intelligent, and interconnected infrastructures capable of supporting real-time monitoring, distributed generation, and advanced control mechanisms. This transformation has been driven by increasing electricity demand, urbanization, integration of renewable energy resources, and the need for improved reliability, resilience, and power quality within electricity networks (Afonso et al., 2021; Azizan, 2024; Samuel et al., 2025). Distribution systems are now central to smart grid development, requiring enhanced reliability analysis, adaptive control strategies, and effective load management techniques to ensure stable power delivery (Aljohani, 2014; Lopez-Prado et al., 2020; Wang et al., 2023). As these networks become increasingly complex, intelligent forecasting approaches are necessary to support proactive planning, operational efficiency, and improved system stability.

The growing adoption of Supervisory Control and Data Acquisition (SCADA) systems has significantly improved visibility, automation, and control in power distribution systems. SCADA architecture enables continuous monitoring of voltage, current, fault conditions, transformer loading, and breaker status, thereby enhancing system reliability and decision-making processes (Iacob et al., 2010; Abou El-Ela et al., 2019). In modern smart grids, SCADA systems also generate large volumes of real-time operational data that can be integrated with predictive analytics and time-series forecasting models for better demand estimation and load behavior analysis (Balakrishna & Jeyan, 2024; Seifi & A. R., 2017). At the same time, increasing penetration of renewable energy resources, electric vehicles, distributed generation, and decentralized grid operations has introduced greater uncertainty and variability into electricity demand patterns (Lu & Chen, 2024; Roy et al., 2024). These challenges highlight the growing need for adaptive and real-time forecasting techniques capable of handling dynamic grid behavior.

Despite substantial progress in load forecasting research, several critical limitations remain in existing predictive models. Many traditional forecasting approaches largely focus on transmission-level forecasting or aggregated system demand, while providing limited attention to distribution-level forecasting where load patterns are more variable and localized (Fan & Hyndman, 2011; Singh et al., 2021). Similarly, conventional load profiling methods often struggle to capture high-resolution operational variability in smart grid environments (Tazi et al., 2018; Kewo et al., 2023). Although statistical and machine learning models have improved forecasting accuracy, many still rely heavily on historical datasets while underutilizing high-frequency SCADA measurements that provide richer operational insights (Bui et al., 2020; Vivas et al., 2020; Shariatzadeh & A., 2019). Furthermore, most forecasting models lack real-time adaptability in handling rapidly changing demand conditions, especially under high renewable penetration and uncertain load dynamics (Farah et al., 2024; Lu & Chen, 2024). Another major challenge is the weak integration of cybersecurity considerations despite the increasing vulnerability of SCADA systems to cyber-attacks, false data injection, and adversarial manipulation that can compromise forecasting reliability and grid security (Chen et al., 2019; Upadhyay & Sampalli, 2020; Mohammadi et al., 2023). These gaps justify the need for a more comprehensive review of SCADA-driven predictive load forecasting techniques.

This review therefore aims to critically examine the role of SCADA systems in predictive load forecasting for power distribution networks. Specifically, it investigates SCADA architecture and its importance in automation, monitoring, and data acquisition for distribution networks (Abou El-Ela et al., 2019). It also reviews conventional statistical, machine learning, deep learning, and hybrid forecasting approaches applied to load prediction (Fan & Hyndman, 2011; Vivas et al., 2020). Additionally, the study analyzes the integration of SCADA-based real-time data into predictive forecasting models, evaluates implementation limitations such as scalability, cybersecurity, data quality, and interpretability, and identifies research gaps that require further exploration (Balakrishna & Jeyan, 2024; Upadhyay & Sampalli, 2020). By synthesizing recent developments, this review seeks to support the development of robust, adaptive, and intelligent forecasting frameworks for future smart distribution systems.

The remainder of this paper is organized as follows. Section 2 presents the review methodology, including search strategy, study selection criteria, and analysis procedures. Section 3 discusses the fundamentals of power distribution systems and their operational behavior. Section 4 explores load profiling techniques and their significance in demand analysis. Section 5 reviews SCADA architecture and its role in modern distribution automation. Section 6 examines traditional, statistical, machine learning, deep learning, and hybrid forecasting models. Section 7 focuses specifically on SCADA-based predictive load forecasting methods and implementation challenges. Section 8 presents comparative findings from previous studies. Section 9 highlights major research gaps and future research directions, while Section 10 proposes a conceptual SCADA-based predictive load profiling framework. Finally, Section 11 concludes the study.

2. Methodology of the Review

2.1 Review Design

This study adopts a systematic literature review methodology to critically evaluate SCADA-based predictive load forecasting models for power distribution networks. A systematic review approach was selected because it provides a structured, transparent, and reproducible process for identifying, screening, analyzing, and synthesizing relevant scholarly studies related to forecasting techniques, SCADA integration, and smart distribution system applications. Unlike traditional narrative reviews that may rely on selective interpretation, a systematic review reduces bias by applying predefined inclusion and exclusion criteria during study selection. The review primarily focuses on understanding how SCADA-generated real-time data has been integrated into predictive load forecasting frameworks and how different forecasting approaches have evolved to address operational uncertainty, demand variability, and grid complexity. In addition, the review combines elements of narrative synthesis to explain trends, methodological gaps, implementation barriers, and future opportunities in SCADA-based forecasting research. This hybrid systematic-narrative design therefore allows both rigorous study selection and detailed critical interpretation of existing literature.

2.2 Search Strategy

A comprehensive literature search strategy was employed to identify high-quality studies related to SCADA-driven predictive load forecasting in distribution systems. Multiple well-established scientific databases were selected to ensure broad coverage of engineering, smart grid, automation, and machine learning research. These databases included IEEE Xplore, which provides extensive literature in electrical engineering, power systems, and automation; Scopus, known for multidisciplinary peer-reviewed indexing; ScienceDirect, which contains high-impact journals in energy and computational sciences; SpringerLink, which includes technical and applied engineering studies; and Google Scholar, used to capture additional conference papers, theses, and emerging scholarly contributions. The search process was performed using combinations of targeted keywords and Boolean operators to improve retrieval accuracy and reduce irrelevant results. Core search terms included SCADA load forecasting, distribution network forecasting, machine learning power load forecasting, predictive load profiling, smart grid SCADA analytics, real-time load prediction, and hybrid forecasting in power systems. Keyword combinations such as SCADA AND machine learning AND load forecasting, or distribution system forecasting AND smart grid analytics, were used to expand the scope while maintaining relevance. This strategy ensured that both foundational and recent studies were captured for analysis.

2.3 Inclusion Criteria

To maintain consistency, relevance, and methodological quality, predefined inclusion criteria were applied during the literature screening process. Studies were included if they were published between 2010 and 2025, a period chosen to capture the evolution of smart grid technologies, modern SCADA architectures, and advanced forecasting models. Selected papers had to specifically focus on SCADA systems, power distribution automation, distribution-level load forecasting, predictive load profiling, or real-time forecasting applications within smart grids. In addition, studies were considered if they employed statistical, machine learning, artificial intelligence, deep learning, or hybrid forecasting approaches, since the review aims to compare forecasting paradigms across traditional and intelligent models. Only peer-reviewed journal articles, conference proceedings, and high-quality academic publications were included to ensure credibility and scientific rigor. Studies presenting measurable forecasting performance, implementation frameworks, or analytical insights into data-driven load forecasting were prioritized. These criteria helped ensure that the final literature pool directly contributed to the objectives of the review.

2.4 Exclusion Criteria

To improve reliability and eliminate irrelevant or weak studies, several exclusion criteria were also defined. Duplicate studies retrieved from multiple databases were removed during the screening stage to prevent repetition and analytical bias. Research focused exclusively on transmission-level optimization, bulk power dispatch, or unrelated energy market forecasting without clear relevance to distribution networks or SCADA-based forecasting was excluded. Since the review emphasizes practical predictive modeling, studies lacking sufficient technical detail regarding methodology, datasets, forecasting algorithms, or evaluation metrics were omitted because they provided limited analytical value. In addition, non-English publications were excluded to maintain consistency in interpretation and avoid translation-related inaccuracies. Papers that were purely conceptual without experimental, analytical, or comparative forecasting evidence were also excluded where they did not contribute substantially to understanding SCADA-based predictive modeling. These exclusion criteria improved the precision of study selection and ensured a stronger evidence base for analysis.

2.5 Data Extraction and Analysis

After screening and selecting relevant studies, a structured data extraction and analytical framework was applied to systematically compare forecasting approaches. For each selected paper, key technical and methodological variables were extracted to support comparative synthesis. These included the forecasting technique used, such as regression-based models, ARIMA, support vector machines, neural networks, deep learning architectures, or hybrid models. The review also examined input variables used in forecasting, including historical load demand, weather conditions, voltage, current, time-series features, transformer loading, and SCADA-derived operational parameters. Another major factor analyzed was the SCADA integration level, which assessed whether studies used SCADA data directly, indirectly, or minimally within forecasting models. Dataset characteristics such as real-world utility datasets, synthetic datasets, high-frequency sensor streams, or historical archives were also evaluated. To compare predictive performance, common accuracy metrics including Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and forecasting reliability indicators were analyzed. Finally, each study’s limitations were documented, including issues such as poor scalability, limited interpretability, weak real-time adaptability, insufficient cybersecurity integration, and lack of generalization across different grid environments. This structured extraction approach strengthened analytical consistency and supported meaningful cross-study comparisons.

2.6 Review Workflow

The overall review workflow followed a systematic multi-stage process to ensure methodological rigor, transparency, and reproducibility. The process began with problem identification, where the need to investigate SCADA-based predictive load forecasting in distribution networks was established based on increasing grid complexity, renewable integration, and forecasting limitations in conventional methods. This was followed by the literature search stage, during which relevant studies were collected from selected academic databases using predefined keywords and search strings. The next phase involved screening and eligibility assessment, where titles, abstracts, and full texts were reviewed using the inclusion and exclusion criteria as presented in Figure 1. After identifying suitable studies, data extraction and classification were performed to organize information related to forecasting methods, datasets, SCADA applications, accuracy measures, and implementation challenges. The extracted data was then subjected to comparative analysis and synthesis, where forecasting techniques were evaluated across statistical, machine learning, deep learning, and hybrid categories. Finally, the review concluded with identification of research gaps, critical limitations, and future research opportunities, providing insight into unresolved issues in SCADA-based predictive forecasting.

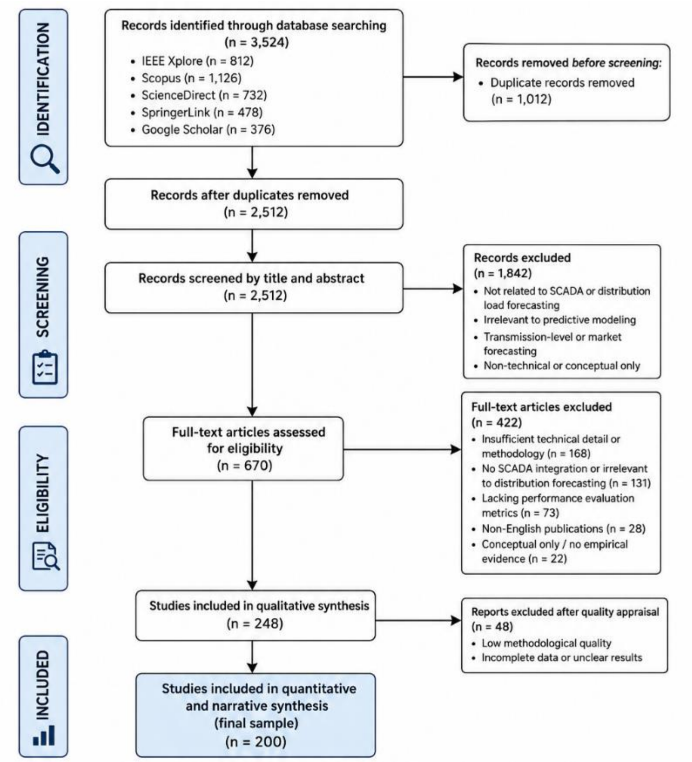

Figure 1. PRISMA Framework Applied to this Review.

3 Fundamentals of Power Distribution Systems

3.1 Overview of Power Distribution Systems

Power distribution systems are the portion of the electrical power network responsible for delivering electricity from transmission systems or substations to end users (residential, commercial, industrial). They encompass all equipment and infrastructure between medium-voltage delivery points and customer service points. Distribution systems must satisfy requirements for safety, voltage regulation, reliability, and quality of service under varying loads and environmental conditions. Recent literature emphasizes that distribution systems are undergoing transformation: increased automation, integration of distributed generation, and tighter power quality requirements (e.g. through smart grid technologies) (Review of Reliability Evaluation in Distribution Networks with Microgrids, 2020) (Lopez-Prado, Vélez, & Garcia-Llinas, 2020) power quality control technologies.

3.2 Components of AC Power Distribution Systems

The distribution of electrical power is the final and most important step in the journey of electricity from generating facilities to consumers. AC power distribution systems are designed to provide electricity to users in the residential, commercial, and industrial sectors in a safe, efficient, and reliable manner as sown in Figure 2. This section delves into the major components of AC power distribution systems, including distribution lines, distribution transformers, circuit breakers and switchgear, distribution substations, and voltage regulators. Table 1 present the major components of AC power distribution systems and their functions.

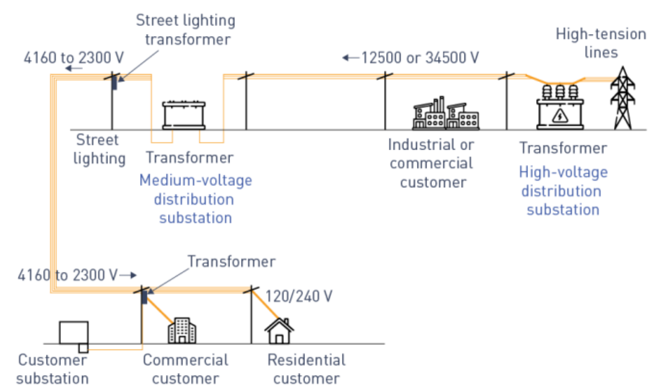

Figure 2. Power Distribution Architecture (Lopez-Prado et al., 2020)

AC distribution networks include several major components:

Distribution Lines (Feeders and Service Lines): These carry power from substations to distribution transformers and onward to end users. They are usually medium voltage (e.g., 11-33 kV) in primary feeders, and lower in secondary/supply lines. They may be overhead or underground, with trade-offs in cost, maintenance, reliability, and exposure to faults (Lopez-Prado et al., 2020)

Distribution Transformers: These step-down medium voltage to utilization voltages (e.g., 400/230 V in many regions) for final delivery. Transformer losses, temperature, loading, and correct sizing are important for system efficiency and reliability (Afonso et al., 2021).

Circuit Breakers, Switchgear & Protection Devices: These allow the network to isolate faults, protect against overloads or short circuits, and restore service. Modern systems also deploy automatic reclosers, smart relays, etc., to enhance response and reduce outage durations (Aljohani, 2014).

Distribution Substations: These are nodes where voltage is transformed downward from transmission levels, switching and protection devices are housed, and routing / control of the distribution network occurs. Substations are critical for voltage regulation and for the adapting of the network under load changes or faults (H. Wang et al., 2023).

Voltage Regulators / Tap Changers / Reactive Power Compensation Elements: To maintain acceptable voltage levels along lines (which suffer voltage drop due to line impedance and varying loads), devices like on-load tap changers, voltage regulators, capacitor banks, etc., are used. Also, in recent literature, power-electronics based solutions are being employed to dynamically regulate voltage and reactive power (Samuel, Toyin, Somefun, Awelewa, & Abba-Aliyu, 2025). These components collectively ensure that power delivered meets voltage, safety, and operational reliability requirements.

Table 1. Major Components of AC Power Distribution Systems and Their Functions (Samuel et al., 2025)

Component	Primary Function	Typical Operating Voltage	Monitoring Parameters (SCADA)	Remarks
Distribution Substation	Voltage transformation	132/33 kV, 66/11 kV	Voltage, current, breaker status	Main node for voltage conversion
Distribution Lines	Power delivery	33 kV, 11 kV, 0.4 kV	Line voltage, current, faults	Can be overhead or underground
Distribution Transformer	Step-down to end-use	11/0.4 kV, 33/0.4 kV	Load current, temperature	Reduces voltage for consumers
Switchgear & Breakers	Protection and isolation	All voltage levels	Trip signals, relay feedback	Safety and network control
Voltage Regulator	Voltage stabilization	11–33 kV	Tap position, voltage levels	Ensures voltage quality
End-User Interface	Consumption metering	230/400 V	kWh, load factor	Interface for billing and monitoring

3.3 Process of Power Distribution to End-Users

The flow of power from generation to end users in AC distribution typically follows this generalized sequence (noting that exact voltages, configurations, and component details vary by region):

1. High-Voltage Supply from Transmission / Bulk Power Sources: Power is generated at centralized plants, stepped up to high voltages for efficient long-distance transmission.
2. Incoming to Distribution Substations: Transmission lines deliver to distribution substations where voltage is reduced to distribution levels (often medium voltage).
3. Primary Distribution via Feeders: From substations, primary feeders carry medium-voltage power to distribution transformers closer to load centers.
4. Voltage Step-Down for End Use: Distribution transformers reduce voltage to utilization levels (e.g. 400/230 V or similar depending on region) which are safe and usable for final consumption.
5. Final Delivery to Customers: Via low-voltage lines (overhead or underground), through service drops/laterals, meters, and then into buildings / loads. Measurement (metering) often at this last stage for billing, load monitoring, etc.

Throughout this process, protection, monitoring, and control devices (switchgear, breakers, sensors) are involved to ensure safety, detect/clear faults, maintain voltage stability, and manage load variations as presented in Figure 3 Automation / smart-grid features are increasingly integrated to reduce response times and improve reliability indices (SAIDI, SAIFI, etc.) (Afonso et al., 2021; Azizan, 2024).

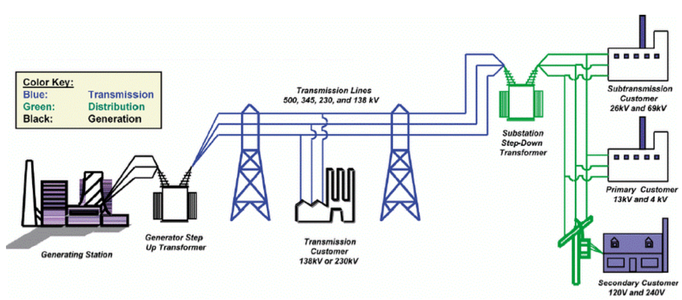

Figure 3. Schematic Flow of Power Distribution to End-Users (Afonso et al., 2021)

3.4 Power Quality and Reliability Considerations

Maintaining acceptable power quality and high reliability is central to distribution system performance. The literature identifies several technical measures and strategies, as well as metrics used to evaluate performance.

i. Power Quality Issues and Mitigation: Typical disturbances include voltage sags/swells, harmonics, flicker, voltage unbalance, and transient events. Mitigation strategies include harmonic filters (passive, active, hybrid), power factor correction, voltage regulation (via regulators and tap changers), and reactive power compensation (capacitor banks, STATCOMs etc.) (Azizan, 2024).
ii. Reliability Strategies: To ensure reliability (availability, low outage frequency / duration) several practices are documented: network configuration (e.g., radial vs ring vs meshed/topology), automation (fault detection, remote switching), distributed generation and storage, preventive / predictive maintenance. Integration of smart grid technologies and real-time monitoring helps reduce outage duration and improve reliability metrics (Aljohani, 2014; H. Wang et al., 2023).
iii. Reliability Metrics & Evaluation: Commonly used indices include SAIDI (System Average Interruption Duration Index), SAIFI (System Average Interruption Frequency Index), ENS/EENS (Energy Not Supplied / Expected Energy Not Supplied), CAIDI, availability, etc. Techniques such as Monte Carlo simulation, analytical probabilistic models, FMEA (Failure Modes & Effects Analysis) are used to evaluate and compare system designs or improvements (Aljohani, 2014; H. Wang et al., 2023).
iv. Role of Automation / Information Flow: Increasingly, research shows that not only the physical components but also the quality of communication, control, and measurement systems (e.g. Remote Terminal Units, Automatic Switches, SCADA/DA systems) significantly affect reliability. For example, errors or delays in communication or control operations can degrade outage detection or fault isolation, thereby worsening reliability indices (Luo, Ge, & Xu, 2023).

4. Load Profiling in Power Distribution Networks

Load profiling is a critical analytical tool in modern power distribution systems, enabling utilities to understand, predict, and manage electricity consumption patterns across different user segments. By analyzing how and when electricity is consumed, utilities can optimize grid operations, enhance demand-side management (DSM), and support the integration of renewable energy sources.

4.1 Definition and Importance

Load profiling involves the collection and analysis of electrical consumption data over time to create representative consumption patterns or "load profiles." These profiles are essential for various applications, including forecasting demand, designing tariffs, and implementing demand response programs. Accurate load profiles allow utilities to anticipate peak demand periods, reduce operational costs, and improve system reliability (Y. Wang et al., 2015).

4.2 Load Profiling Methodologies

This involves techniques used to analyze and predict electricity consumption patterns. They are mainly categorized into bottom-up approaches, which use detailed user-level data, and top-down approaches, which rely on aggregated system data (Kewo, Manembu, & Nielsen, 2023). Advanced methods like clustering further enhance accuracy, enabling efficient energy planning, demand-side management, and tailored tariff design. Figure 4 illustrates the two primary approaches used in load profiling: bottom-up and top-down methodologies, alongside supporting analytical techniques. The bottom-up approach estimates electricity demand using detailed appliance- or user-level data, enabling highly accurate and granular modeling but requiring extensive data collection (Y. Wang et al., 2015). In contrast, the top-down approach relies on aggregated data from feeders or substations, making it more scalable and cost-effective, though less detailed. The figure also highlights advanced analytical tools, particularly clustering techniques such as k-means and hierarchical clustering, which group consumers based on similar consumption patterns. These techniques enhance demand-side management (DSM) strategies and enable personalized tariff design, thereby improving energy efficiency, planning, and decision-making in modern power systems.

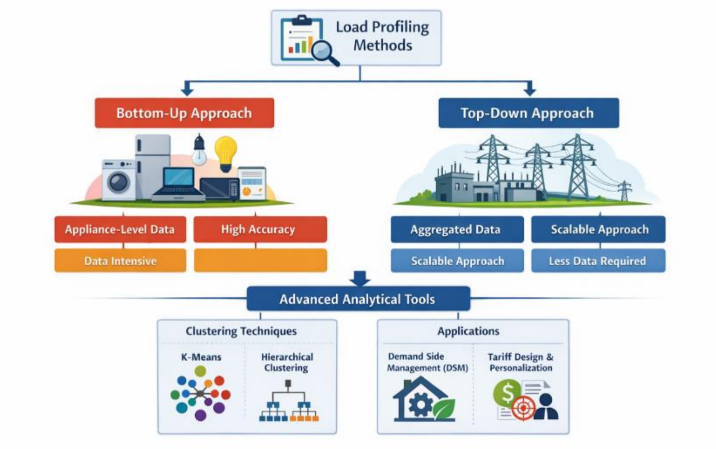

Figure 4. Methodologies for Load Profiling

4.3 Applications of Load Profiling

Load profiling serves multiple purposes in power distribution systems:

1. Demand Forecasting: By analyzing historical consumption data, utilities can predict future demand trends, facilitating better planning and resource allocation.
2. Tariff Design: Utilities use load profiles to develop time-of-use tariffs that encourage consumers to shift their usage to off-peak periods, thereby flattening the demand curve and reducing peak load (Kewo et al., 2023; Y. Wang et al., 2015).
3. Demand Response Programs: Load profiling is instrumental in identifying opportunities for demand response, where consumers adjust their usage in response to price signals or incentives, helping to balance supply and demand (Kewo et al., 2023).
4. Integration of Renewable Energy: Understanding load profiles assists in aligning renewable energy generation with consumption patterns, improving the efficiency of renewable integration into the grid (Tazi, Abbou, & Abdi, 2018).

4.4 Challenges

Despite its benefits, load profiling faces several challenges:

1. Data Privacy and Security: The collection of detailed consumption data raises concerns about user privacy and data security, necessitating robust protection measures.
2. Data Quality and Availability: Inaccurate or incomplete data can lead to misleading load profiles, affecting the effectiveness of DSM strategies.
3. Dynamic Consumption Patterns: Changes in consumer behavior, such as the adoption of electric vehicles or smart appliances, can alter load profiles, requiring continuous monitoring and model updates.

5. SCADA Systems in Power Distribution

SCADA systems are critical for the monitoring, control, and automation of industrial processes, including power distribution networks. They allow operators to collect real-time data from geographically dispersed sites, make informed decisions, and control field devices remotely. SCADA systems improve operational efficiency, enhance reliability, and enable predictive maintenance in modern power distribution infrastructure (Upadhyay, 2020; El-Sehiemy et al., 2019). The increasing integration of smart grids and renewable energy sources further underscores the importance of SCADA in managing dynamic loads and ensuring system stability.

5.1 SCADA Architecture Overview

A clear understanding of SCADA systems begins with recognizing their multi-layered architecture. SCADA architecture is not merely a network of connected devices; it is a structured ecosystem designed to ensure scalability, security, interoperability, and real-time control across all levels of operation as presented in Figure 5 Each layer contributes uniquely to the seamless flow of data and control commands, forming an integrated framework that underpins mission-critical decision-making in modern power systems.

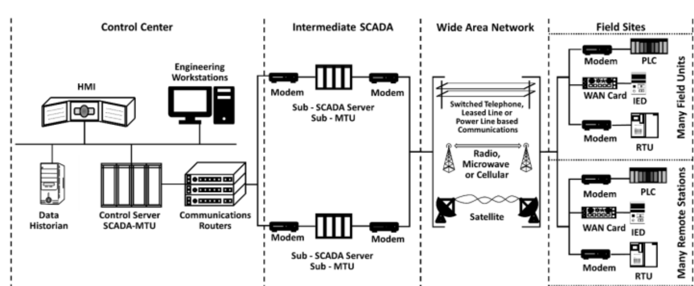

Figure 5. General Architecture of a SCADA System in Power Distribution (Upadhyay & Sampalli, 2020)

i. Field Devices (Level 0): These include sensors, actuators, and transducers that collect real-time process data such as voltage, current, and breaker status. They also perform direct control actions like opening circuit breakers or adjusting voltage regulators.

ii. Control Devices (Level 1): Remote Terminal Units (RTUs) and Programmable Logic Controllers (PLCs) serve as the bridge between the field and supervisory levels. They aggregate and preprocess field data, execute control logic, and transmit information to higher tiers. RTUs are typically deployed in remote substations, while PLCs are used in localized industrial environments.

iii. Supervisory Systems (Level 2): This level encompasses the SCADA master station, Human-Machine Interface (HMI), and SCADA servers. These components handle real-time visualization, event management, alarm notification, and operator control, enabling comprehensive supervision of the distribution network (Iacob & Andreescu, 2015).

iv. Enterprise Systems (Level 3): At this top layer, SCADA integrates with enterprise-level platforms such as Historian databases, Outage Management Systems (OMS), and Geographic Information Systems (GIS). These integrations facilitate long-term analytics, system optimization, and strategic decision-making, bridging the gap between operational technology (OT) and information technology (IT).

Table 2 summarizes key SCADA components and their functions. It highlights elements such as control equipment, industrial automation, local control, RTUs, SCADA software, user interfaces, and wireless communication. Together, these components enable real-time monitoring, data acquisition, control, and efficient operation of industrial and power systems.

Table 2. Key SCADA Components and Functions (El-Sehiemy et al., 2019)

Component	Function	Example
Control Equipment	Devices such as breakers, valves, switches, and relays	Executes commands received from RTUs or PLCs
Industrial Automation	Use of control systems industrial processes	Improves operational efficiency and reduces manual intervention
Local Control	On-site decision-making via PLCs or embedded logic	Enables rapid response without relying on central systems
Remote Terminal Units (RTUs)	Field devices that collect data and send control signals to equipment	Interface between field sensors/actuators and SCADA master station
SCADA Software	Centralized applications for monitoring, command execution, and logging	Manages data visualization, alarms, trending, and control logic.
User Interfaces	HMIs (Human-Machine Interfaces) and dashboards	Allow operators to interact with system data and issue commands
Wireless Communication	Radio, cellular, or satellite links between remote sites and control centers	Enables real-time remote control and data transmission

5.2 Benefits of SCADA in the Electric Power Industry

The implementation of SCADA technology has revolutionized the management of electric power systems by enabling real-time visibility, control, and intelligence across the grid. More than just a monitoring tool, SCADA empowers utilities to transition from reactive maintenance toward predictive and proactive grid management, ensuring greater resilience and cost efficiency.

i. Enhanced Reliability: Continuous, real-time monitoring allows early fault detection and automatic isolation of affected network sections, reducing the scope and duration of outages.

ii. Operational Efficiency: SCADA minimizes the need for on-site interventions by providing remote access and control, significantly reducing operational costs and response times, especially for geographically dispersed assets.

iii. Improved Safety: Operators can remotely monitor and control high-voltage equipment, thereby reducing physical exposure to hazardous conditions.

iv. Data-Driven Decision-Making: The system's historical and real-time datasets support load forecasting, asset management, and predictive maintenance, improving planning accuracy and infrastructure longevity.
v. Regulatory and Compliance Support: Automated data logging, event recording, and reporting functions assist utilities in meeting industry regulations, audit requirements, and performance standards.

In summary, SCADA systems provide a comprehensive digital backbone for modern electric utilities enhancing grid intelligence, enabling rapid response to system disturbances, and supporting the transition toward smart, self-healing power distribution networks.

5.3 Applications in Power Distribution

SCADA systems provide multiple benefits in power distribution, including:

1. Real-Time Monitoring: Operators can track system voltages, currents, and frequencies continuously, allowing quick identification of anomalies and load variations (Muruga Lal Jeyan & Balakrishna, 2024).
2. Fault Detection and Isolation: SCADA enables immediate detection and localization of faults, facilitating rapid isolation and minimizing service interruptions (El-Sehiemy et al., 2019).
3. Load Management: Provides data for load balancing, shedding, and optimization strategies, especially during peak periods or contingency events.
4. Data Logging and Analysis: Historical SCADA data supports performance assessment, predictive maintenance, and operational optimization (Upadhyay, 2020).
5. SCADA systems therefore play a critical role in ensuring reliability, efficiency, and operational visibility in modern distribution networks.

5.4 Integration with Modern Technologies

SCADA systems are increasingly integrated with advanced digital technologies to improve efficiency, reliability, and intelligence. Cloud Computing enables scalable data storage, remote system access, and reduced infrastructure costs by minimizing dependence on on-site hardware. The Internet of Things enhances monitoring by deploying numerous smart sensors and devices, providing real-time, high-resolution data for precise control and decision-making. Additionally, Machine Learning and data analytics techniques support predictive maintenance, accurate load forecasting, and early anomaly detection using both historical and real-time datasets (Balakrishna & Jeyan, 2024). As connectivity increases, robust cybersecurity measures become essential to protect systems from cyber threats, unauthorized access, and data breaches, ensuring the safe and reliable operation of modern industrial and power systems.

5.5 Challenges with Integration of Modern Technologies

While SCADA systems offer significant benefits, they face challenges such as:

1. Cybersecurity Risks: Increased interconnectivity exposes SCADA systems to potential cyber-attacks.
2. System Complexity: Integration with multiple technologies requires specialized expertise for operation and maintenance.
3. Data Overload: Vast amounts of real-time data require advanced data management and analytics to avoid overwhelming operators (Iacob & Andreescu, 2015).

6. Predictive Models for Load Forecasting

Predictive load profiling involves the use of historical consumption data, environmental variables, and sometimes socio-economic factors to forecast future electricity demand. Traditional methods, such as time series analysis and regression models, have been widely used but often struggle with the complexity and non-linearity of modern load patterns. Recent advancements have introduced machine learning techniques, such as artificial neural networks (ANNs), support vector machines (SVMs), and deep learning models, to enhance forecasting accuracy. Table 3 presents a comparison of traditional and Machine Learning-based load forecasting models. It highlights differences in accuracy, data requirements, adaptability, and computational complexity, showing how modern techniques improve prediction performance and flexibility, thereby supporting more efficient load profiling, energy planning, and demand-side management in power systems.

Table 3. Comparison of Traditional and Machine Learning-Based Load Forecasting Models (Balakrishna & Jeyan, 2024)

Aspect	**Traditional Load Forecasting Models**	**Machine Learning (AI)-Based Load Forecasting Models**
Main Techniques	(S)ARIMA, SWDP2A, System Identification, Time-based Data-driven Schemes, IEMD–ARIMA–WNN–FOA Hybrid, Probabilistic Load Forecasting (PLF)	Decision Trees, SVR, ANN, RNN, LSTM, MLSTM, Fuzzy Logic, Federated Learning (FL), Deep Learning (DL), Feature Engineering (FE), IPFE, CHAID
Key Features	Based on statistical and mathematical time-series models; often incorporate weather, time, and occupancy features	Utilize data-driven learning from large datasets; capable of nonlinear modeling and feature extraction automatically
Highlights	- SWDP2A algorithm provides accurate hourly forecasts - Occupancy-based STLF for non-inhabitant buildings - System identification achieves 90% accuracy in 60 s - Time-based weather model enhances demand prediction - IEMD–ARIMA–WNN–FOA hybrid offers high accuracy and stability - PLF applies probabilistic quantile scoring	- Decision tree (C4.5) achieves 92–93% accuracy - Merged SVR–ANN models best for heating/cooling prediction - Simplified LF models with minimal features - RNNs for cooling control and occupant behavior - LSTM and MLSTM outperform linear methods and traditional ML - FL and IPFE methods enhance privacy and efficiency - AI-based schemes detect electricity theft accurately
Pros	- High accuracy (e.g., hybrid models) - Effective for short-term and hourly forecasts - Low computation time for simpler models - Strong interpretability of parameters	- Superior accuracy for complex and nonlinear data - Can handle unbalanced, multivariate, and large datasets - Capable of automatic feature learning - Incorporates privacy-preserving techniques (e.g., FL) - Effective for energy theft detection and smart grid management
Cons	- Require heavy data preprocessing - Often limited to specific contexts (e.g., non-inhabitant buildings) - Struggle with extreme or rare events - Complex hybrid models are hard to implement - High data dependency for accuracy	- Model-specific limitations (e.g., C4.5 only, unbalanced data focus) - Generalization may be weak across datasets or domains - Some techniques (e.g., RNN, LSTM) ignore external influences - Complexity and high computational demands - Certain models target niche applications (e.g., net-metering privacy)
Data Requirements	Structured, preprocessed time-series data (often from smart meters or occupancy systems)	Large, diverse, and labeled datasets for effective training; may include behavioral, temporal, and environmental data
Forecasting Scope	Mostly short-term load forecasting (STLF); limited long-term applicability	Applicable to short-, medium-, and long-term forecasts; also suitable for dynamic system prediction and optimization
Interpretability	High models like ARIMA, PLF are transparent and explainable	Often lower — black-box nature of deep learning; explainability tools may be needed
Practicality	Easier to deploy with small datasets and simple systems	Better suited for smart grids, IoT-enabled, and big data environments
Privacy Handling	Generally, lacks built-in privacy mechanisms	Advanced privacy preservation through FL, encryption, and IPFE

The method is organized into several key stages. The first stage involves data collection and preparation, where relevant datasets are gathered, verified, and structured to ensure consistency and reliability for subsequent analysis. The second stage focuses on data exploration, which aims to uncover underlying trends, patterns, and relationships within the dataset through visualization and statistical analysis. In the third stage, data preprocessing is performed to make the dataset suitable for machine learning applications. This includes cleaning, normalization, feature selection, and data splitting to ensure that models can learn effectively. The fourth stage involves implementing and evaluating multiple machine learning algorithms such as Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and Gated Recurrent Units (GRU) to perform short-term electrical load forecasting. The performance of each model is assessed using various evaluation metrics, allowing for a comprehensive comparison of their predictive capabilities. Finally, the best-performing model is identified and further refined through optimization and hyperparameter tuning to enhance its accuracy and robustness. An overview of this step-by-step methodology is presented in Figure 6, which visually summarizes the entire process from data preparation to model optimization.

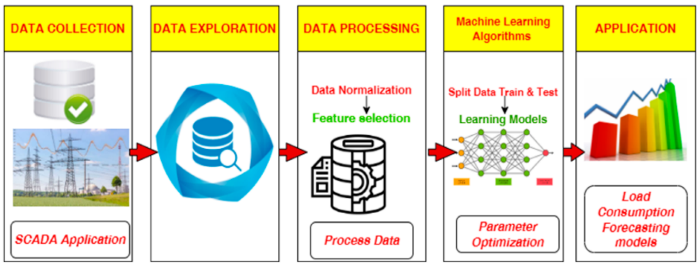

Figure 6. Basic workflow for electric load forecasting models.

6.1 Evolution of Load Forecasting Models

The evolution of load forecasting models reflects the increasing complexity of modern power systems and the transition from deterministic, rule-based approaches to data-driven intelligent forecasting frameworks. Early forecasting techniques were primarily based on statistical and time-series models, which assumed relatively stable and linear relationships in electricity demand patterns. These methods were widely used due to their simplicity and interpretability but were often limited in handling nonlinear behavior and sudden load variations. As power systems became more dynamic with the introduction of distributed generation, renewable energy sources, and changing consumption behaviors, machine learning (ML) approaches emerged as more flexible alternatives capable of learning complex patterns from historical data. Subsequently, deep learning (DL) models further advanced forecasting capability by automatically extracting hierarchical features from large-scale time-series datasets, significantly improving prediction accuracy in highly nonlinear environments (Vivas et al., 2020; Fan & Hyndman, 2011). More recently, hybrid models combining statistical, ML, and signal decomposition techniques have been developed to address limitations in individual approaches, offering improved robustness, adaptability, and forecasting stability in smart grid applications (Bui et al., 2020; Lu & Chen, 2024). This progression demonstrates a clear shift toward intelligent, data-intensive forecasting systems suitable for SCADA-integrated smart distribution networks.

6.2 Statistical Forecasting Models

Statistical forecasting models form the foundation of load prediction techniques and are widely used for short-term electricity demand forecasting due to their simplicity and interpretability. The Autoregressive Integrated Moving Average (ARIMA) model is one of the most commonly applied techniques, which captures temporal dependencies by combining autoregression, differencing, and moving average components to model time-series data. ARIMA performs well under stable and linear load conditions but struggles with nonlinear patterns and external influencing factors such as weather variability and renewable integration. Similarly, regression-based models establish mathematical relationships between load demand and explanatory variables such as temperature, time of day, and seasonal factors, offering a straightforward and interpretable forecasting framework. However, their linear assumptions limit performance in highly dynamic grid environments. Exponential smoothing methods, including Holt-Winters models, improve forecasting by assigning exponentially decreasing weights to older observations, making them effective for capturing trends and seasonality in load data. Despite their usefulness in traditional systems, statistical models are generally less effective in modern smart grids where load behavior is increasingly nonlinear and influenced by multiple interacting variables (Fan & Hyndman, 2011; Wang et al., 2015).

6.3 Machine Learning Models

Machine learning models represent a significant advancement over statistical methods by enabling systems to learn complex nonlinear relationships from historical data without strict assumptions about data distribution. Artificial Neural Networks (ANNs) are widely used in load forecasting due to their ability to approximate nonlinear functions and capture complex interactions between input variables such as temperature, historical load, and SCADA measurements. Support Vector Machines (SVMs) are effective for regression-based forecasting tasks, particularly in high-dimensional spaces, where they construct optimal hyperplanes to minimize prediction error while maintaining generalization capability. Random Forest models, which are ensemble-based methods, improve forecasting accuracy by aggregating multiple decision trees to reduce variance and prevent overfitting, making them robust against noisy SCADA data. Similarly, XGBoost (Extreme Gradient Boosting) has gained popularity due to its high computational efficiency and strong predictive performance, especially in large-scale datasets. These ML models significantly improve forecasting accuracy compared to traditional approaches; however, they still require feature engineering and may struggle with capturing long-term temporal dependencies unless combined with sequence-based architectures (Vivas et al., 2020; Bui et al., 2020; Farah et al., 2024).

6.4 Deep Learning Models

Deep learning models have transformed load forecasting by enabling automatic feature extraction and superior handling of nonlinear and time-dependent relationships in large-scale datasets. Recurrent Neural Networks (RNNs) are designed to process sequential data and are capable of capturing temporal dependencies in electricity demand; however, they often suffer from vanishing gradient issues when dealing with long sequences. To address this limitation, Long Short-Term Memory (LSTM) networks were introduced, incorporating memory cells that allow the model to retain long-term dependencies, making them highly effective for short-term load forecasting in SCADA-based systems. Gated Recurrent Units (GRUs) provide a simplified alternative to LSTMs with fewer parameters while maintaining comparable performance, making them more computationally efficient for real-time forecasting applications. More recently, Transformer-based forecasting models have emerged as a powerful alternative, utilizing self-attention mechanisms to capture long-range dependencies without relying on sequential processing. These models are particularly effective in handling multivariate SCADA datasets and large-scale smart grid environments due to their scalability and parallel processing capability. Despite their superior accuracy, deep learning models often require large datasets and high computational resources, which can limit deployment in resource-constrained utility environments (Lu & Chen, 2024; Vivas et al., 2020; Roy et al., 2024).

6.5 Hybrid Forecasting Models

Hybrid forecasting models integrate multiple analytical techniques to overcome the limitations of individual approaches and improve prediction accuracy, robustness, and adaptability in complex power systems. One common approach is the combination of Empirical Mode Decomposition (EMD) with Artificial Neural Networks (ANNs), where EMD is used to decompose non-stationary load signals into simpler components that are then modeled using ANN for improved forecasting performance. Similarly, ARIMA–LSTM hybrid models combine the strengths of statistical and deep learning approaches, where ARIMA captures linear components of the time series while LSTM models nonlinear and long-term dependencies, resulting in enhanced forecasting accuracy. Another important hybrid approach involves integrating SCADA data with weather and socio-economic variables, allowing forecasting models to account for multiple external influences such as temperature variations, consumer behavior, economic activity, and grid operational states. These hybrid models are particularly effective in smart grid environments where load patterns are influenced by diverse and dynamic factors. Overall, hybrid forecasting frameworks provide improved generalization, better stability under uncertain conditions, and higher predictive accuracy compared to standalone models, making them highly suitable for SCADA-based predictive load forecasting in modern distribution networks (Bui et al., 2020; Seifi & A. R., 2017; Farah et al., 2024; Shariatzadeh & A., 2019).

6.6 Predictive Load Profiling: Applications, Benefits, and Challenges

Table 4 provides a summary of the key applications, benefits, and challenges associated with predictive load profiling models. These models play a crucial role in modern power systems by enhancing grid reliability, supporting renewable energy integration, and improving economic efficiency. However, their practical implementation is often constrained by issues related to data quality, model interpretability, and scalability. The table below highlights these aspects for a clearer understanding of their significance and limitations.

Table 4. Summary of Applications, Benefits, and Challenges of Predictive Load Profiling Models (Upadhyay, 2020)

Category	Description
Applications and Benefits	

Enhanced Grid Management	Accurate load forecasts improve grid planning and operations, reducing outage risks and enhancing overall system reliability.
Integration of Renewable Energy	Forecasting demand patterns helps balance supply and demand, supporting the integration of variable renewable sources like solar and wind.
Economic Efficiency	Precise load forecasting enables optimized generation scheduling and resource allocation, lowering operational costs.
Demand Response Programs	Understanding consumption behavior allows utilities to design effective demand response strategies, leading to more efficient and flexible energy use.
Challenges	
Data Quality and Availability	Reliable and high-resolution data is crucial for accurate forecasting but may be limited due to data gaps, inconsistencies, or collection constraints.
Model Interpretability	Advanced models such as deep learning may produce accurate predictions but often lack transparency, making it hard for operators to trust or explain results.
Scalability	Predictive models must handle large-scale, complex datasets as power systems evolve, posing computational and integration challenges.

7 SCADA-Based Predictive Models for Load Forecasting

Supervisory Control and Data Acquisition (SCADA) systems provide high-resolution, real-time data from power distribution networks, which is increasingly leveraged for load forecasting. SCADA data includes voltage, current, frequency, transformer loadings, and breaker status, providing a detailed view of system behavior. The integration of SCADA data into predictive load profiling models improves forecast accuracy, enabling utilities to optimize operations, plan maintenance, and implement demand response strategies (Upadhyay, 2020; Muruga Lal Jeyan & Balakrishna, 2024).

7.1 Role of SCADA Data in Load Forecasting

i. Data analytics plays a vital role in Supervisory Control and Data Acquisition (SCADA) systems used in power generation and transmission. By leveraging analytical techniques, SCADA systems can transform raw, real-time data into actionable insights that enhance grid reliability, operational efficiency, and decision-making. The following points highlight the key reasons why data analytics is essential in power SCADA systems:

ii. Real-Time Monitoring and Control: SCADA systems continuously collect real-time data from numerous sensors and devices distributed across power generation and transmission networks. Through data analytics, this information is processed and visualized to provide operators with an accurate view of the grid's current status. Such insights enable informed decision-making for system control, stability, and efficient operation as illustrated in Figure 6.

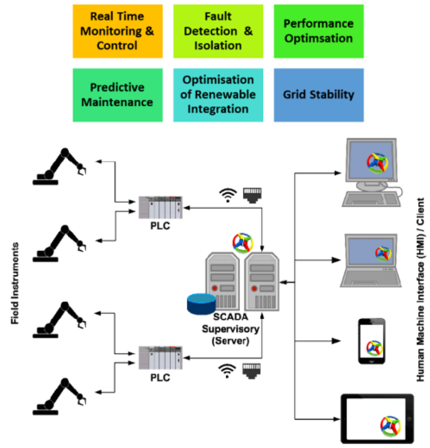

Figure 7. Need for Data Analytics in Power SCADA & a typical SCADA system (Balakrishna & Jeyan, 2024).

iii. Fault Detection and Diagnostics: Analyzing SCADA data helps identify anomalies, equipment malfunctions, and potential system failures in the generation and transmission infrastructure. Early detection through analytics allows operators to initiate timely corrective measures, minimizing equipment downtime and preventing large-scale service interruptions.

iv. Performance Optimization: By examining historical SCADA data, data analytics can uncover patterns and trends in system performance. These insights can be used to optimize resource utilization, improve energy efficiency, and enhance equipment reliability ultimately leading to more sustainable and cost-effective power system operations.

v. Predictive Maintenance: Data analytics can process key operational variables such as temperature, pressure, voltage, and current to forecast potential equipment failures. This enables the adoption of predictive maintenance strategies, where maintenance tasks are planned proactively. As a result, utilities can reduce unexpected outages, extend asset lifespans, and minimize maintenance costs.

vi. Optimization of Renewable Energy Integration: With the growing share of renewable energy sources in modern grids, data analytics helps optimize the integration of these intermittent resources. By combining SCADA data with weather forecasts and renewable generation profiles, utilities can better predict renewable output and manage supply-demand balance more effectively.

vii. Grid Stability and Resilience: Through continuous analysis of SCADA data, utilities can assess grid stability, detect vulnerabilities, and evaluate system resilience. Data-driven insights enable operators to design strategies that strengthen grid reliability, mitigate operational risks, and enhance the system’s ability to recover from disturbances.

7.2 SCADA Data Types Used in Forecasting

SCADA systems provide a rich set of real-time operational variables that form the foundation of predictive load forecasting in modern distribution networks. These data types capture both electrical behavior and system operational states, enabling more accurate and adaptive forecasting models.

1. Voltage is one of the most critical SCADA measurements used in forecasting as it reflects system stability and load variation across the network. Fluctuations in voltage levels indicate changes in demand, feeder loading, and reactive power conditions, making it a key predictor of system stress and consumption trends.
2. Current measurements provide direct information about power flow and loading conditions within feeders, transformers, and distribution lines. Since current is closely related to real-time energy consumption, it is widely used in forecasting models to estimate demand levels and identify peak load periods.
3. Breaker status data indicates the operational condition of circuit breakers (open/closed) in the network. This information is essential for understanding network topology changes due to switching actions, faults, or maintenance operations. It helps forecasting models adjust predictions when load redistribution occurs.
4. System frequency is a global indicator of power balance between generation and demand. Deviations from nominal frequency values signal system stress or load imbalance, making frequency a valuable indirect input for detecting and forecasting demand variations in real time.
5. Transformer temperature reflects loading conditions, thermal stress, and environmental effects on equipment. Rising temperature trends often correspond to increased load demand, making it useful for both forecasting consumption patterns and supporting predictive maintenance strategies.
6. Harmonic data represents distortions in voltage and current waveforms caused by non-linear loads such as power electronics and renewable energy systems. These measurements are important for capturing power quality issues and improving the accuracy of forecasting models in modern smart grids.
7. Event logs record time-stamped system activities such as faults, alarms, breaker operations, and protection actions. These logs provide contextual and behavioral insights that help distinguish between normal load variations and abnormal system disturbances, improving the robustness of forecasting models.

7.3 Methods of Using SCADA Data

SCADA data is employed in load forecasting through several approaches:

i. Statistical Models: Time series models (ARIMA, exponential smoothing) can use SCADA measurements for short-term forecasts.

ii. Machine Learning Models: Neural networks, support vector regression, and ensemble methods use SCADA data to capture non-linear load patterns.

iii. Hybrid Approaches: Combining SCADA data with weather, calendar, and socio-economic data improves forecasting accuracy by addressing multiple factors affecting load (Vivas et al., 2020).

Table 5. Examples of SCADA-Based Load Forecasting Approaches (Vivas et al., 2020)

Approach	Description	Benefits
Statistical	ARIMA, Exponential Smoothing	Simple, interpretable
Machine Learning	ANN, LSTM, Random Forest	High accuracy, captures non-linear trends
Hybrid	Combination of SCADA + Weather/Calendar Data	Improved predictive performance

7.4 Benefits and Challenges

Benefits of SCADA Data for Load Forecasting

Integrating data analytics into power SCADA systems offers significant advantages for modern power grid operations. It enhances the accuracy of load forecasting, particularly for short-term planning, enabling operators to make timely and informed decisions. This improved forecasting capability directly contributes to the reliability and efficiency of power distribution networks by allowing better resource allocation and demand management. Furthermore, by aligning generation schedules with predicted demand, analytics facilitates the integration of renewable energy sources, ensuring smoother operation despite their intermittent nature. In addition, SCADA-based analytics supports predictive maintenance and optimized grid management, helping utilities minimize equipment downtime, reduce operational costs, and maintain a more resilient power system.

Challenges of SCADA Data for Load Forecasting

Despite its benefits, applying data analytics in power SCADA systems presents several technical and operational challenges. One major issue is data quality, as SCADA data often contains missing values, noise, or measurement errors, which must be addressed through preprocessing to ensure analytical accuracy. The large volume of high-resolution data generated by SCADA networks also poses scalability challenges, demanding efficient data storage, processing, and management frameworks. Additionally, cybersecurity remains a critical concern, since any compromise in SCADA data integrity can have severe consequences for grid operations and reliability. Therefore, robust data validation, secure communication protocols, and continuous monitoring are essential to safeguard these systems (Upadhyay & Sampalli, 2020).

8. Review of Similar Studies

Previous studies have extensively explored the integration of SCADA data into load forecasting models. For instance, (Balakrishna & Jeyan, 2024) emphasized the necessity of data analytics in Power SCADA systems, highlighting its role in optimizing power grid operations. Their research focused on enhancing real-time monitoring and control capabilities through advanced analytical techniques.

Nickelson et al. (2014) adopted a comprehensive review methodology to analyze SCADA architectures and their role in power systems. The study concluded that SCADA significantly improves monitoring and operational control efficiency. However, it did not incorporate predictive analytics or forecasting capabilities, leaving a gap in integrating SCADA data into intelligent load forecasting frameworks.

Teixeira et al. (2010) used cyber-attack modeling and simulation techniques to evaluate vulnerabilities in SCADA systems. Their findings showed that SCADA infrastructures are highly susceptible to stealthy cyber-attacks, which can compromise system reliability. The study emphasized the need for improved security mechanisms but lacked practical, scalable real-time defense solutions suitable for modern smart grid environments.

Ziel & Liu (2016) applied LASSO-based statistical modeling for probabilistic load forecasting. Their results demonstrated improved forecasting accuracy compared to traditional regression approaches, especially in handling variability. However, the method required well-structured datasets and did not effectively utilize real-time SCADA data streams, limiting its applicability in dynamic and data-intensive smart grid systems.

Bianchi et al. (2017) employed recurrent neural networks (RNNs) to model electricity load time-series data. The study concluded that RNNs outperform conventional models in capturing temporal dependencies and nonlinear patterns. Despite improved prediction accuracy, the approach faced challenges related to high computational cost, complex training processes, and lack of interpretability, which hinder practical deployment in real-time SCADA environments.

Ji et al. (2018) utilized data-driven probabilistic forecasting models for appliance-level load prediction. The results showed high accuracy in short-term forecasting and better uncertainty representation. However, the study was limited to small-scale applications and lacked scalability to large power systems, particularly those integrating SCADA-based real-time data across distribution networks.

Chen et al. (2019) used adversarial attack analysis to investigate vulnerabilities in load forecasting models. The study revealed that manipulated SCADA data can significantly degrade forecasting accuracy. It concluded that cybersecurity is essential for reliable forecasting systems, but did not propose comprehensive or integrated frameworks to simultaneously address prediction accuracy and system security.

Kong et al. (2019) adopted deep learning techniques, particularly LSTM networks, for residential load forecasting. Their findings indicated superior performance in capturing complex consumption patterns compared to traditional models. However, the approach required large datasets and high computational resources, making it difficult to deploy in resource-constrained environments and limiting integration with real-time SCADA systems.

Vivas et al. (2020) conducted a systematic comparative review of statistical and machine learning forecasting models. The study concluded that machine learning approaches outperform traditional methods in accuracy and adaptability. However, it highlighted gaps in real-time implementation and insufficient integration with SCADA data streams for operational decision-making.

Bui et al. (2020) proposed a statistical data-filtering methodology to improve load forecasting accuracy by removing noise and outliers from SCADA datasets. The results showed enhanced model reliability and prediction precision. Nonetheless, the study did not address how filtered data could be integrated into advanced machine learning pipelines for real-time forecasting applications.

Singh (2021) applied machine learning algorithms to SCADA datasets for wind power forecasting. The study demonstrated improved prediction accuracy and better handling of renewable energy variability. However, it focused primarily on wind generation and did not consider broader system-level load forecasting or integration with other energy sources.

Effenberger & Ludwig (2022) used dataset collection and categorization methods to analyze open-source energy datasets. Their findings emphasized the importance of accessible datasets for model validation and benchmarking. However, the study identified a significant gap in the availability of standardized, high-quality SCADA datasets, limiting the development and comparison of advanced forecasting models.

Mohammadi et al. (2023) developed a deep learning-based intrusion detection system for smart grids using SCADA data. The results showed improved detection accuracy and enhanced system security. Despite these advances, challenges related to real-time deployment, scalability, and integration with forecasting systems remain unresolved.

Balakrishna & Jeyan (2024) employed data analytics techniques within SCADA systems to enhance real-time monitoring and grid optimization. The study concluded that analytics significantly improve operational efficiency and decision-making. However, it lacked a unified framework integrating forecasting, cybersecurity, and real-time SCADA data processing.

Farah et al. (2024) applied data-driven and hybrid forecasting models incorporating environmental and socio-economic variables. The results demonstrated improved forecasting accuracy and adaptability. Nevertheless, the study did not fully integrate SCADA data streams or address real-time system implementation challenges.

Lu & Chen (2024) used advanced machine learning models for short-term load forecasting in systems with high renewable penetration. The study concluded that combining SCADA data with AI improves handling of variability and uncertainty. However, gaps remain in developing scalable, secure, and fully integrated real-time forecasting frameworks.

Despite significant advances in leveraging SCADA data for load forecasting, several research gaps remain unaddressed. Many previous studies (e.g., Balakrishna & Jeyan, 2024; Vivas et al., 2020; Partanen et al., 1990) have primarily focused on either improving forecasting accuracy or enhancing real-time monitoring capabilities, but few have proposed a comprehensive framework that simultaneously integrates SCADA-based measurements with predictive modeling for distribution-level load profiling. While existing approaches have successfully demonstrated the utility of machine learning and statistical models for short-term load forecasting (Bui et al., 2020; Vivas, Allende-Cid, & Salas, 2020), most are designed for transmission-level or aggregated demand data, which limits their applicability to low-voltage distribution systems where consumption patterns are more volatile and heterogeneous. Moreover, although studies such as (Singh, 2021) and (Lu & Chen, 2024) explored SCADA's role in renewable integration and variability management, they often overlooked the detailed spatiotemporal profiling of consumer loads essential for operational planning and reliability assessment in modern distribution networks.

Additionally, prior works tend to treat SCADA data primarily as an input source rather than a core element of predictive modeling for dynamic load profiling. Few studies have explicitly utilized high-resolution SCADA measurements for identifying consumer behavior patterns or developing adaptive, data-driven forecasting schemes tailored to specific feeder characteristics. Furthermore, research addressing the challenges of data quality, latency, and cybersecurity in SCADA-integrated forecasting remains fragmented (Chen, Tan, & Zhang, 2019; Mohammadi, Aflaki, Kavousifard, & Gitizadeh, 2023). Hence, there is a critical need for an integrated modeling approach that not only enhances prediction accuracy but also improves decision support for distribution system operators through continuous learning and model refinement. The uniqueness of the present study lies in its focus on developing a predictive load profiling model grounded in SCADA-based measurements, enabling detailed, real-time characterization of load behavior at the distribution level. By combining data analytics, predictive modeling, and system feedback mechanisms, this study bridges the gap between traditional forecasting approaches and smart grid operational requirements. Table 6 Present the summary of the reviewed related literature.

Table 6. Summary of the Reviewed Related Literature

Authors	Methodology	Forecasting Horizon	Dataset Type	Strengths	Limitations	Key Findings
Mohammadi et al. (2019)	Deep learning-based anomaly detection for smart grids	Real-time monitoring	Smart grid/SCADA datasets	Effective cyber-threat detection and anomaly identification	Forecasting performance not evaluated	Deep learning improves SCADA cybersecurity and operational reliability.
Chen et al. (2019)	Adversarial attack analysis on forecasting models	Short-term	Simulated and SCADA-based datasets	Highlights cybersecurity risks in forecasting systems	Limited mitigation strategies	Load forecasting models are vulnerable to adversarial attacks that may compromise operational decisions.
Abou El-Ela, El-Sehiemy & El-Shebiny (2019)	Review of SCADA applications in distribution automation	N/A	Distribution automation case studies	Comprehensive review of SCADA functions	Limited forecasting focus	SCADA plays a critical role in automation, monitoring, and fault management.
Bui et al. (2020)	Statistical filtering and preprocessing techniques	Short-term	SCADA load datasets	Improves data quality and forecasting accuracy	Scalability not fully evaluated	Data cleansing significantly improves forecasting performance.
Upadhyay & Sampalli (2020)	SCADA architecture and application review	N/A	Industrial SCADA systems	Broad coverage of SCADA functionalities	Limited predictive analytics discussion	SCADA remains foundational for data acquisition and grid automation.
Vivas et al. (2020)	Systematic review of statistical and machine learning forecasting models	Short-term	Multiple benchmark datasets	Comprehensive comparison of forecasting methods	Limited real-time implementation analysis	Machine learning methods outperform traditional statistical approaches in most cases.
Rafi et al. (2021)	Integrated CNN–LSTM deep learning framework	Short-term	Historical load datasets	Captures nonlinear and temporal load	Computationally intensive	CNN-LSTM significantly improves forecasting accuracy over standalone models.

				patterns effectively		
Farsi et al. (2021)	Parallel Deep LSTM-CNN architecture	Short-term	Utility load datasets	High predictive accuracy and robust feature extraction	High training complexity	Hybrid deep learning models outperform conventional ML approaches.
Singh (2021)	LSTM-based forecasting using SCADA data	Short-term	SCADA load demand datasets	Effective for nonlinear load forecasting	High computational requirements	LSTM networks improve short-term forecasting accuracy using SCADA data.
Lin et al. (2021)	Attention-enhanced LSTM network	Short-term	Power consumption time-series datasets	Improves feature selection and temporal dependency learning	Reduced interpretability	Attention mechanisms enhance forecasting accuracy and stability.
Cai et al. (2021)	Bidirectional LSTM (Bi-LSTM) deep learning model	Short-term	Power system load datasets	Captures forward and backward temporal dependencies	Computational complexity	Bi-LSTM achieves higher forecasting precision than conventional LSTM models.
Stratigakos et al. (2021)	SSA-LSTM hybrid forecasting model	Short-term	High-frequency net-load datasets	Handles seasonality and complex load patterns effectively	Increased model complexity	Combining Singular Spectrum Analysis and LSTM improves forecast accuracy.
Xie et al. (2021)	Two-stage LSTM-MLP forecasting framework	Short-term	Utility load datasets	Reduces prediction error through staged learning	Requires extensive training data	Hybrid architectures outperform single-model approaches.
Sharma & Jain (2021)	LSTM with enhanced feature engineering	Short-term	Electrical load datasets	Improved feature representation enhances accuracy	Sensitive to feature selection quality	Feature-space optimization improves forecasting reliability.
Zhou et al. (2021)	Appliance-level LSTM forecasting	Short-term	Residential smart meter datasets	Fine-grained appliance forecasting	Limited scalability to utility-level systems	Appliance-level forecasting improves demand-side management.
Farah, Farou, Kouahla & Seridi (2024)	Data-driven forecasting using historical load, weather, and socio-economic factors	Short- and medium-term	Algerian utility dataset (2008–2020)	Multi-factor forecasting framework with strong predictive capability	Limited regional generalization	Combining historical demand and external variables improves forecasting performance.
Balakrishna & Jeyan (2024)	SCADA analytics and time-series forecasting framework	Short- to medium-term	SCADA operational datasets	Highlights importance of SCADA-driven analytics	Limited real-time deployment evaluation	SCADA analytics significantly improve forecasting effectiveness and operational awareness.
Lu & Chen (2024)	Hybrid ML framework for renewable-rich grids	Short-term	SCADA and renewable generation datasets	Handles renewable variability effectively	Limited explainability	Hybrid forecasting models improve prediction accuracy under high renewable penetration.
Quansah & Tenkorang (2024)	PSO-optimized Attention CNN-LSTM model	Short-term	Real electricity demand datasets	High accuracy and automatic hyperparameter optimization	Computationally demanding	Attention-based CNN-LSTM achieves superior forecasting performance with low MAPE.
Emerging Studies (2025)	Explainable AI (XAI), Graph Neural Networks (GNNs), Transformer-based forecasting	Short- and medium-term	High-frequency SCADA and IoT datasets	Improved scalability and interpretability	Limited industrial deployment evidence	Transformer and graph-based architectures show strong potential for next-generation smart-grid forecasting.

Emerging Studies (2026)	SCADA-integrated Digital Twin and Federated Learning frameworks	Real-time and adaptive forecasting	Distributed utility and edge-computing datasets	Enhanced privacy, scalability, and real-time adaptability	Early-stage implementation and validation	Digital twins and federated learning are emerging as promising directions for secure and adaptive forecasting in smart distribution networks.

9. Research Gaps and Future Directions

9.1 Research Gaps

Despite the growing application of SCADA-driven predictive load forecasting in modern power distribution systems, several significant research gaps remain unresolved. One of the most notable limitations is the lack of feeder-level forecasting models. Most existing forecasting studies focus on aggregated system-level or transmission-level demand prediction, which often overlooks localized load behavior, feeder congestion patterns, and distribution-specific variability. Since feeder-level forecasting is essential for demand-side management, outage mitigation, and distributed energy optimization, this remains a critical gap in practical distribution forecasting research (Fan & Hyndman, 2011; Singh et al., 2021; Wang et al., 2015).

Another major challenge is the weak real-time implementation of forecasting models. Although many machine learning and hybrid forecasting approaches demonstrate high predictive accuracy under offline simulations, they often fail to support real-time deployment in operational SCADA environments. Distribution systems generate high-frequency streaming data, requiring adaptive models capable of responding to rapid demand fluctuations, renewable intermittency, and fault-driven operational changes. However, many current forecasting frameworks remain computationally intensive and poorly suited for real-time grid automation (Balakrishna & Jeyan, 2024; Farah et al., 2024; Lu & Chen, 2024).

A further research gap lies in cybersecurity integration within forecasting systems. SCADA infrastructures are increasingly vulnerable to cyber-attacks such as false data injection, adversarial manipulation, denial-of-service attacks, and unauthorized access. Yet most forecasting models prioritize prediction accuracy while neglecting cyber-resilience, anomaly detection, and secure data validation. This creates risks where compromised SCADA signals may directly affect forecasting reliability, operational planning, and grid stability (Chen et al., 2019; Upadhyay & Sampalli, 2020; Mohammadi et al., 2023).

Another unresolved issue is the limited interpretability of advanced forecasting models, particularly deep learning and hybrid artificial intelligence architectures. While neural networks, transformers, and ensemble learning techniques can significantly improve prediction accuracy, many operate as "black-box" systems that provide limited transparency into how predictions are generated. In critical energy infrastructures, operators require explainable and trustworthy decision support systems for effective adoption, auditing, and fault diagnosis. Poor interpretability therefore remains a barrier to practical deployment in utility-scale SCADA environments (Vivas et al., 2020; Roy et al., 2024).

Scalability also remains a major concern in modern forecasting research. Many predictive models are developed and validated on small or isolated datasets, often under controlled experimental conditions. These models may struggle when extended to large-scale smart grids with multiple substations, distributed energy resources, microgrids, and heterogeneous sensor infrastructures. Poor scalability limits model robustness, computational efficiency, and applicability across diverse network topologies (Lopez-Prado et al., 2020; Luo et al., 2023; Wang et al., 2023).

Finally, there is fragmented SCADA integration in forecasting frameworks. Although SCADA systems provide rich operational datasets such as voltage, current, breaker status, transformer loading, and fault signals, many forecasting studies only partially integrate SCADA-derived information. Instead, models frequently depend primarily on historical demand or weather-based variables. This underutilization reduces forecasting depth and limits the benefits of real-time situational awareness and adaptive decision-making in smart distribution systems (Abou El-Ela et al., 2019; Shariatzadeh & A., 2019; Seifi & A. R., 2017).

9.2 Future Directions

To address these gaps, future SCADA-based predictive load forecasting research should focus on several emerging intelligent and resilient technologies.

One promising direction is the adoption of Explainable Artificial Intelligence (XAI). XAI techniques can improve transparency in forecasting models by helping operators understand which variables most influence predictions, thereby increasing trust, accountability, and interpretability in mission-critical SCADA systems. This is especially important for utility operators who require explainable decisions before deploying AI-driven forecasting tools in real-time power management (Vivas et al., 2020; Roy et al., 2024).

Another emerging direction is federated learning, which allows forecasting models to be trained across decentralized utility systems without directly sharing sensitive operational data. Since SCADA infrastructures often contain critical infrastructure information, federated learning can enhance privacy, improve collaboration between distributed substations, and reduce cybersecurity exposure while enabling robust distributed forecasting frameworks (Upadhyay & Sampalli, 2020; Mohammadi et al., 2023).

The development of Edge AI for SCADA systems also offers strong potential. Edge-based intelligence enables forecasting models to process and analyze data directly at substations, feeders, or smart sensors rather than relying entirely on centralized cloud platforms. This can reduce latency, improve real-time responsiveness, enhance local fault detection, and support adaptive forecasting in highly dynamic distribution environments (Balakrishna & Jeyan, 2024; Luo et al., 2023).

Another transformative area is the integration of digital twins in power distribution forecasting. Digital twin technology can create virtual real-time replicas of SCADA-enabled power networks, enabling utilities to simulate demand fluctuations, test forecasting strategies, analyze disturbances, and optimize system behavior before real-world deployment. This can significantly improve predictive maintenance, reliability analysis, and operational planning (Aljohani, 2014; Lopez-Prado et al., 2020).

Future studies should also investigate transfer learning to improve forecasting generalization across different utility regions, feeder architectures, and grid topologies. Transfer learning can reduce dependence on large localized datasets by adapting previously trained models to new forecasting environments, particularly useful where SCADA data availability is limited or inconsistent (Farah et al., 2024; Lu & Chen, 2024).

Similarly, reinforcement learning (RL) presents strong opportunities for adaptive forecasting and autonomous energy optimization. RL can allow forecasting systems to continuously learn from real-time grid responses, operational constraints, and demand changes, making models more adaptive in smart grid control, demand-side response, and dynamic energy balancing (Roy et al., 2024; Afonso et al., 2021).

Finally, future research should prioritize cyber-resilient forecasting frameworks that integrate predictive analytics with anomaly detection, adversarial robustness, encrypted communication, and secure SCADA validation protocols. Such models would not only forecast demand but also maintain reliability under malicious attacks, data corruption, and infrastructure compromise. Cyber-resilient forecasting will be essential for next-generation smart grids where predictive intelligence and cybersecurity must operate as integrated components rather than independent systems (Chen et al., 2019; Upadhyay & Sampalli, 2020; Mohammadi et al., 2023).

Overall, advancing SCADA-based predictive load forecasting requires a shift from static, isolated forecasting models toward explainable, scalable, decentralized, secure, and adaptive intelligent systems capable of supporting future resilient smart distribution networks. These directions provide a strong foundation for both academic research and real-world utility implementation.

11. Conclusion

The integration of Supervisory Control and Data Acquisition (SCADA) systems into modern power distribution networks has significantly transformed how utilities monitor, analyze, and manage electricity demand. SCADA systems play a critical role in enabling real-time data acquisition, system visibility, automation, and operational control, making them essential for predictive load forecasting in increasingly complex smart grid environments. By providing high-resolution measurements such as voltage, current, transformer loading, breaker status, and fault signals, SCADA enhances forecasting depth and supports proactive decision-making, reliability improvement, and efficient resource management across distribution systems.

This review demonstrates that machine learning-based forecasting approaches generally outperform traditional statistical methods in handling nonlinear relationships, dynamic load patterns, and large multidimensional datasets. Techniques such as artificial neural networks, support vector machines, deep learning models, and transformer-based architectures have shown improved forecasting accuracy and adaptability, especially in environments characterized by renewable energy variability, distributed generation, and changing consumption behaviors. Their ability to learn complex dependencies makes them particularly suitable for modern distribution-level forecasting challenges.

In addition, hybrid forecasting methods have emerged as one of the most promising approaches for SCADA-based predictive load profiling. By combining statistical techniques, machine learning models, optimization algorithms, and domain-specific operational insights, hybrid models can leverage the strengths of multiple forecasting paradigms while minimizing weaknesses such as overfitting, low interpretability, or poor generalization. These models often demonstrate better robustness, improved prediction stability, and enhanced performance under uncertain and highly variable grid conditions.

Despite these advancements, several unresolved challenges continue to limit widespread implementation of SCADA-based predictive forecasting systems. These include inadequate feeder-level forecasting, weak real-time deployment capabilities, fragmented SCADA data integration, cybersecurity vulnerabilities, poor scalability across large smart grid infrastructures, and limited interpretability of complex artificial intelligence models. Addressing these issues is essential to ensure reliability, trust, and operational feasibility in real-world utility environments.

Looking forward, the future of SCADA-based predictive load forecasting lies in the deployment of intelligent, adaptive, and resilient smart-grid solutions. Emerging technologies such as explainable artificial intelligence, federated learning, edge AI, digital twins, reinforcement learning, and cyber-resilient forecasting frameworks are expected to improve scalability, transparency, autonomy, and security. As smart grids continue to evolve toward decentralized and data-driven architectures, integrating advanced forecasting models with SCADA systems will be critical for optimizing energy efficiency, enhancing grid reliability, supporting renewable energy integration, and enabling sustainable power distribution.

References

Abou El-Ela, A., *et al.* (2019). Review of SCADA system for distribution power system automation. *ERJ. Engineering Research Journal*, *42*(2), 93-98.

Afonso, J. L., *et al.* (2021). A review on power electronics technologies for power quality improvement. *Energies*, *14*(24), 8585.

Aljohani, T. M. (2014). *Distribution system reliability analysis for smart grid applications.* University of Southern California,

Azizan, M. M. (2024). Introductory Chapter: Power Quality. In *Power Quality-New Insights*: IntechOpen.

Balakrishna, R., & Jeyan, J. (2024). SCADA-Need for Data Analytics & Time Series Analysis for Effective Load Forecasting. *J. Electr. Syst*, *22*, 552-560.

Bui, D. M., *et al.* (2020). A statistical data-filtering method proposed for short-term load forecasting models. *Journal of Electrical Engineering & Technology*, *15*(5), 1947-1967.

Chen, Y., *et al.* (2019). *Exploiting vulnerabilities of load forecasting through adversarial attacks.* Paper presented at the Proceedings of the tenth ACM international conference on future energy systems.

Fan, S., & Hyndman, R. J. (2011). Short-term load forecasting based on a semi-parametric additive model. *IEEE transactions on power systems*, *27*(1), 134-141.

Farah, R., *et al.* (2024). Data driven models on load forecasting: Case study Algeria. *Data in Brief*, *52*, 109854.

Iacob, M., *et al.* (2010). Supervisory control and data acquisition laboratory. *Telfor Journal*, *2*(1), 49-54.

Kewo, A., *et al.* (2023). A rigorous standalone literature review of residential electricity load profiles. *Energies*, *16*(10), 4072.

Lopez-Prado, J. L., *et al.* (2020). Reliability evaluation in distribution networks with microgrids: Review and classification of the literature. *Energies*, *13*(23), 6189.

Lu, W., & Chen, X. (2024). Short-term load forecasting for power systems with high-penetration renewables based on multivariate data slicing transformer neural network. *Frontiers in Energy Research*, *12*, 1355222.

Luo, F., *et al.* (2023). Power supply reliability analysis of distribution systems considering data transmission quality of distribution automation terminals. *Energies*, *16*(23), 7826.

Mohammadi, M., *et al.* (2023). Deep Learning-Based Cyber-Attack Detection Model for Smart Grids. *arXiv preprint arXiv:2312.08810*.

Mouftah, M. E.-K. a. H. T. (2015). Energy-Efficient Information and Communication Infrastructures in the Smart Grid: A Survey on Interactions and Open Issues," IEEE Communications Surveys & Tutorials, vol. 17, no. 1, pp. 179–197.

Roy, A., *et al.* (2024). *Intelligent Computation and Analytics on Sustainable Energy and Environment*.

Samuel, I. A., *et al.* (2025). Evaluation and improvement of power quality of distribution network: a case study of Covenant University, Ota. *Frontiers in Energy Efficiency, 2*, 1389622.

Seifi, A. R. M. a. A. R. (2017). A probabilistic framework for short-term load forecasting using weather variables and SCADA data," Electric Power Systems Research, vol. 143, pp. 675–683.

Shariatzadeh, H. N. a. A. (2019). Utilizing SCADA data for Load Modeling and Identification in Distribution Networks," IET Generation, Transmission & Distribution, vol. 13, no. 24, pp. 5580–5587.

Singh, M. K. S. a. A. R. (2021). Limitations of Traditional Load Profiling Techniques in Smart Grid Environment," International Journal of Electrical Power & Energy Systems, vol. 125, pp. 106410.

Tazi, K., *et al.* (2018). *Load analysis and consumption profiling: An overview*. Paper presented at the International Conference on Electronic Engineering and Renewable Energy.

Upadhyay, D., & Sampalli, S. (2020). SCADA (Supervisory Control and Data Acquisition) systems: Vulnerability assessment and security recommendations. *Computers & Security, 89*, 101666.

Vivas, E., *et al.* (2020). A systematic review of statistical and machine learning methods for electrical power forecasting with reported mape score. *Entropy, 22*(12), 1412.

Wang, H., *et al.* (2023). Reliability evaluation method for distribution network with distributed generations considering feeder fault recovery and network reconfiguration. *IET Renewable Power Generation, 17*(14), 3484-3495.

Wang, Y., *et al.* (2015). Load profiling and its application to demand response: A review. *Tsinghua Science and Technology, 20*(2), 117-129.

Aljohani, T. M. (2014). *Distribution system reliability analysis for smart grid applications.* University of Southern California,

Bui, D. M., Le, P. D., Cao, T. M., Nguyen, H., Pham, T. T., & Pham, D. A. (2020). A statistical data-filtering method proposed for short-term load forecasting models. *Journal of Electrical Engineering & Technology, 15*(5), 1947-1967.

Chen, Y., Tan, Y., & Zhang, B. (2019). *Exploiting vulnerabilities of load forecasting through adversarial attacks.* Paper presented at the Proceedings of the tenth ACM international conference on future energy systems.

Lu, W., & Chen, X. (2024). Short-term load forecasting for power systems with high-penetration renewables based on multivariate data slicing transformer neural network. *Frontiers in Energy Research, 12*, 1355222.

Luo, F., Ge, N., & Xu, J. (2023). Power supply reliability analysis of distribution systems considering data transmission quality of distribution automation terminals. *Energies, 16*(23), 7826.

Mohammadi, M., Aflaki, A., Kavousifard, A., & Gitizadeh, M. (2023). Deep Learning-Based Cyber-Attack Detection Model for Smart Grids. *arXiv preprint arXiv:2312.08810*.

Singh, M. K. S. a. A. R. (2021). Limitations of Traditional Load Profiling Techniques in Smart Grid Environment," International Journal of Electrical Power & Energy Systems, vol. 125, pp. 106410.

Tazi, K., Abbou, F. M., & Abdi, F. (2018). *Load analysis and consumption profiling: An overview.* Paper presented at the International Conference on Electronic Engineering and Renewable Energy.

Upadhyay, D., & Sampalli, S. (2020). SCADA (Supervisory Control and Data Acquisition) systems: Vulnerability assessment and security recommendations. *Computers & Security, 89*, 101666.

Vivas, E., Allende-Cid, H., & Salas, R. (2020). A systematic review of statistical and machine learning methods for electrical power forecasting with reported mape score. *Entropy, 22*(12), 1412.

Wang, H., Lu, H., Sun, K., Wu, X., He, Y., & Du, X. (2023). Reliability evaluation method for distribution network with distributed generations considering feeder fault recovery and network reconfiguration. *IET Renewable Power Generation, 17*(14), 3484-3495.

Wang, Y., Chen, Q., Kang, C., Zhang, M., Wang, K., & Zhao, Y. (2015). Load profiling and its application to demand response: A review. *Tsinghua Science and Technology, 20*(2), 117-129.

International Journal of Engineering Research and Technology, 12(4), 521–530.

Bui, V. H., Hussain, A., & Kim, H. M. (2020). A statistical data-filtering method proposed for short-term load forecasting. *Energies, 13*(23), 6314.

Cai, M., Zhang, M., Wang, Y., & Liu, J. (2021). A bidirectional LSTM deep learning method for short-term load forecasting. *Applied Sciences, 11*(17), 8129.

Chen, S., Zhang, Y., & Wang, X. (2019). Exploiting vulnerabilities of load forecasting through adversarial attacks. *IEEE Transactions on Smart Grid, 10*(6), 6587–6595.

Farsi, B., Amayri, M., Bouguila, N., & Eicker, U. (2021). On short-term load forecasting using machine learning techniques and a novel parallel deep LSTM-CNN approach. *IEEE Access, 9*, 31191–31212. https://doi.org/10.1109/ACCESS.2021.3060290

Lin, Y., Chen, Z., & Wang, H. (2021). Attention-enhanced LSTM network for short-term load forecasting. *Electric Power Systems Research, 196*, 107251.

Lu, J., & Chen, H. (2024). Short-term load forecasting for power systems with high penetration of renewable energy. *Energy Reports, 10*, 1256–1268.

Mohammadi, M., Mohammadi, M., Hosseinzadeh, M., & Asadi, S. (2023). Deep learning-based cyber-attack detection model for smart grids. *Sustainable Energy, Grids and Networks, 34*, 101018.

Quansah, P. K., & Tenkorang, E. K. A. (2024). Short-term load forecasting using a particle-swarm optimized multi-head attention-augmented CNN-LSTM network. *arXiv Preprint arXiv:2309.03694*. Retrieved from https://arxiv.org/abs/2309.03694

Rafi, S. H., Al-Masood, N., Deeba, S. R., & Hossain, E. (2021). A short-term load forecasting method using integrated CNN and LSTM network. *IEEE Access*, *9*, 32436–32448. https://doi.org/10.1109/ACCESS.2021.3060654

Razghandi, M., Zhou, H., Erol-Kantarci, M., & Turgut, D. (2021). Short-term load forecasting for smart home appliances with sequence-to-sequence learning. *arXiv Preprint arXiv:2106.15348*. Retrieved from https://arxiv.org/abs/2106.15348

Singh, R. (2021). SCADA system dataset exploration and machine learning for wind power forecasting. *Renewable Energy*, *178*, 1082–1094.

Singh, R. (2021). Short-term load forecasting using LSTM and SCADA data. *International Journal of Electrical Power & Energy Systems*, *130*, 106936.

Stratigakos, A., Papageorgiou, D., Dikaiakos, C., & Hatziargyriou, N. (2021). Short-term net load forecasting using singular spectrum analysis and LSTM networks. *Energies*, *14*(14), 4107.

Xie, J., Chen, Y., Hong, T., & Wang, S. (2021). A two-stage LSTM-MLP model for short-term load forecasting. *Energies*, *14*(18), 5873.

Zhou, Y., Nair, A. S., Ganger, D., Tripathi, A., Baone, C., & Zhu, H. (2021). Appliance level short-term load forecasting via recurrent neural network. *arXiv Preprint arXiv:2111.11998*. Retrieved from https://arxiv.org/abs/2111.11998