Doctoral Dissertation

Telecommunication Technologies and Computer Engineering
University of Las Palmas de Gran Canaria

Contributions to Electricity Demand Estimation in Remote Areas: A Neural Network Approach

---

Author

Stefano Sanfilippo

Supervisors

José Juan Hernández-Cabrera
José Évora-Gómez

November 2025

Acknowledgments

I am deeply grateful to José Juan Hernández-Cabrera and José Évora-Gómez for their unwavering support throughout my PhD journey, which began in 2020 amid the challenges of the COVID-19 pandemic. Their guidance and encouragement helped me stay focused and motivated during a period of uncertainty, isolation, and professional transitions, including my return to Italy. Beyond their academic mentorship, their empathy, friendship, and personal warmth, shown through shared experiences, meant a great deal to me. Their human presence and kindness provided comfort and a sense of belonging that went far beyond professional support, and I will never forget it.

I sincerely thank Bruno Péchiné, Enrique Kremers, David Eyler, Gregory Savidand, and Nurten Avci for believing in me and supporting my doctoral journey both academically and professionally. Their guidance made possible a scientific publication and my participation in several conferences.

I also thank José Juan Hernández-Gálvez, who supported me during his internship at EIFER and afterwards, contributing to and supporting the experimental phase of the research. He is an excellent professional with immense potential, and a person with a heart of gold.

I am deeply grateful to STAM, and in particular to Pietro De Vito, Enrico Maggiani, and Fabio Bagnoli, for offering me the opportunity to return to Italy and to pursue my professional growth within such a stimulating and inspiring environment.

Dedication

A very special dedication goes to my girlfriend, Giorgia, who always understands me and possesses extraordinary emotional intelligence. She has supported me through every moment, both the brightest and the darkest, always standing by my side. I owe her so much, as she has helped me grow immensely. I am incredibly happy that we have now bought our own apartment together, because she is the woman I envision by my side for life.

I also dedicate this achievement to my mother, who continues to inspire me despite her passing in 2011. I am also deeply grateful to my father, Bennardo, and his partner, Anna, for their love and support; to my brother, Domenico, his partner, Alice, and their child, Nicolò, of whom I am the godfather, for their closeness. I also thank my aunt Rosa for her care; she has always been a source of strength throughout my life. I am also grateful to my uncles and grandparents for their love. Though many of my grandparents are no longer with us, their example has taught me to live fully and pursue my passions.

I also want to express my affection for Giorgia's family: her mother Anna, an extraordinary woman; Nicolò, a remarkable architect; her grandmother Angela, whose infinite kindness reminds me of my own grandmothers; as well as her uncles and cousins, whose warmth and generosity make me feel part of the family. I feel a special bond with them.

To my friends from Karlsruhe, Genoa, Turin, Sommatino, Belfast, and from all over Italy and the world, thank you. Along the way, I have met many passionate people, and I am deeply grateful to everyone, whether from university, football, or calisthenics, who has been part of my journey during these intense years.

To all of you, thank you. This journey was never mine alone. It was ours.

Resumen

La previsión de la demanda eléctrica es un componente fundamental de la planificación energética, pero se vuelve especialmente desafiante en contextos con limitaciones de datos caracterizados por la escasez, la fragmentación, las dependencias no lineales entre variables socio-técnicas y ambientales, y la evolución no estacionaria de la demanda. Los enfoques tradicionales, ya sean estadísticos o puramente basados en datos, suelen presentar dificultades en estas condiciones, pues dependen de grandes conjuntos de datos estacionarios y no logran capturar las dinámicas complejas y cambiantes del comportamiento de consumo. Estos desafíos se hacen especialmente evidentes al emplear modelos de redes neuronales, que normalmente requieren conjuntos de datos extensos y de alta calidad para generalizar de forma eficaz y evitar el sobreajuste.

Abordar esta brecha requiere replantear la predicción de la demanda eléctrica como un proceso causal, flexible y evolutivo, en el que los mecanismos de aprendizaje capturen los verdaderos impulsores del consumo y no simples correlaciones estadísticas. La intuición central que guía esta tesis es que una predicción fiable en microrredes no puede ignorar ninguno de los tres desafíos fundamentales mencionados: la escasez y fragmentación de datos, las dependencias no lineales entre variables socio-técnicas y ambientales, y la evolución no estacionaria de la demanda.

Con este propósito, la tesis introduce un nuevo marco metodológico, denominado **ANGEL**, concebido para actuar como un sistema orientador del aprendizaje en condiciones de incertidumbre. ANGEL funciona como un agente adaptativo capaz de extraer significado en contextos con limitaciones de datos, proporcionando dirección donde los datos son escasos y revelando patrones que no son inmediatamente visibles. El nombre evoca la idea de un mediador entre lo conocido y lo desconocido, reflejando el papel del marco en tender un puente entre la evidencia limitada y la comprensión significativa. En este sentido, ANGEL encarna la noción de iluminación: la capacidad de revelar estructura y causalidad en contextos con pocos datos, transformando información fragmentada en conocimiento coherente y accionable.

El marco ANGEL integra tres pilares complementarios: **observación**, **representación** y **evolución**. La observación se centra en identificar y codificar los impulsores causales con significado físico y conductual, garantizando robustez cuando los datos son limitados o fragmentados. La representación captura dependencias no lineales mediante arquitecturas neuronales flexibles, concretamente la Red de Kolmogorov–Arnold, que descompone relaciones multivariadas complejas en componentes funcionales estables alineados con las dinámicas de la demanda. La evolución incorpora la adaptación temporal y el cambio socio-técnico a través del parámetro de Grado de Adopción (Degree of Adoption (DoA)) y técnicas de aprendizaje continuo, permitiendo que los modelos se ajusten de manera constante a medida que las comunidades se electrifican y los patrones de consumo evolucionan. Estos tres componentes son mutuamente indispensables: una predicción fiable en microrredes requiere abordarlos simultáneamente.

La validación empírica, basada en datos reales de El Espino (Bolivia), demuestra que la integración de estructura causal, representación flexible y aprendizaje adaptativo reduce el error de predicción y mejora la robustez. El marco también muestra un desempeño eficiente bajo recursos computacionales limitados, lo que sugiere su potencial como una solución desplegable y eficiente en el uso de datos para sistemas energéticos descentralizados.

Desde una perspectiva metodológica, la tesis conecta el razonamiento causal con el aprendizaje automático, contribuyendo al paradigma emergente de una inteligencia artificial confiable y eficiente en el uso de datos para sistemas energéticos sostenibles. Los resultados muestran que las mejoras en el rendimiento de la

predicción no provienen del aumento del volumen de datos o de la complejidad del modelo, sino de estructurar el aprendizaje en torno a la causalidad significativa, la representación flexible y la evolución adaptativa.

La investigación se adhiere a los principios de reproducibilidad y ciencia abierta. Todos los conjuntos de datos, scripts y flujos de trabajo experimentales fueron versionados, documentados y trazables, y los resultados se difundieron en tres publicaciones revisadas por pares que abarcan las sucesivas etapas de desarrollo del modelo: desde la estimación de carga y sensibilidad, hasta el modelado neuronal causal y la predicción adaptativa.

En conclusión, los resultados evidencian que es posible alcanzar predicciones de demanda fiables y de bajo error incluso frente a los desafíos combinados de escasez y fragmentación de datos, dependencias no lineales y evolución no estacionaria de la demanda. Al articular observación, representación y evolución dentro de la arquitectura metodológica coherente de ANGEL, la tesis demuestra que ninguna de estas dimensiones puede ser ignorada, y que, en conjunto, permiten el desarrollo de modelos interpretables, adaptativos y eficientes en el uso de datos, que impulsan una inteligencia artificial equitativa y sostenible para la transición energética global.

Abstract

Electricity demand forecasting is a fundamental component of energy planning but becomes particularly challenging in data-constrained contexts characterized by scarcity, fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution. Traditional approaches, whether statistical or purely data-driven, often struggle under such conditions, as they rely on large, stationary datasets and fail to capture the complex, evolving dynamics of consumption behavior. These challenges become particularly evident when employing neural network models, which typically require extensive, high-quality datasets to generalize effectively and avoid overfitting.

Addressing this gap requires rethinking electricity demand forecasting as a causal, flexible, and evolving process, where learning mechanisms capture the true drivers of consumption rather than mere statistical correlations. The central intuition guiding this thesis is that reliable forecasting in microgrids cannot disregard any of the three fundamental challenges identified above: data scarcity and fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution.

To this end, the thesis introduces a new methodological framework, **ANGEL**, conceived to act as a guiding system for learning under uncertainty. ANGEL functions as an adaptive agent capable of extracting meaning in data-constrained contexts, providing direction where data are limited, and uncovering patterns that are not immediately visible. The name evokes the idea of a mediator between what is known and what is unknown, reflecting the framework's role in bridging limited evidence and meaningful understanding. In this sense, ANGEL embodies illumination, the capacity to reveal structure and causality in data-scarce contexts, transforming fragmented information into coherent and actionable knowledge.

The ANGEL framework integrates three complementary pillars: **observation**, **representation**, and **evolution**. Observation focuses on identifying and encoding causal drivers that are physically and behaviorally meaningful, supporting robustness when data are limited or fragmented. Representation captures nonlinear dependencies through flexible neural architectures, specifically, the Kolmogorov–Arnold Network, which decomposes complex multivariate relations into stable functional components aligned with demand dynamics. Evolution incorporates temporal adaptation and socio-technical change through the inclusion of the Degree of Adoption (DoA) parameter and continual learning techniques, allowing models to adjust continuously as communities electrify and consumption patterns evolve. These three components are mutually indispensable: reliable forecasting in microgrids requires addressing all of them simultaneously.

Empirical validation based on real-world data from El Espino (Bolivia) demonstrates that embedding causal structure, flexible representation, and adaptive learning reduces forecasting error and improves robustness. The framework also performs efficiently under limited computational resources, suggesting its potential as a data-efficient and deployable solution for decentralized energy systems.

From a methodological standpoint, the thesis bridges causal reasoning and machine learning, contributing to the emerging paradigm of trustworthy and data-efficient artificial intelligence for sustainable energy systems. The findings show that improvements in forecasting performance stem not from increasing data volume or model complexity, but from structuring learning around meaningful causality, flexible representation, and adaptive evolution.

The research adheres to the principles of reproducibility and open science. All datasets, scripts, and experimental workflows were versioned, documented, and made traceable, with results disseminated through three peer-reviewed publications covering successive stages of model development, from sensitivity and load

estimation to causal neural modeling and adaptive forecasting.

In conclusion, the results provide evidence that reliable, low-error demand forecasting can be achieved even under the combined challenges of data scarcity, fragmentation, nonlinear dependencies, and nonstationary demand evolution. By articulating observation, representation, and evolution within the coherent methodological architecture of ANGEL, the thesis establishes that none of these dimensions can be neglected, and together they enable the development of interpretable, adaptive, and data-efficient models that advance equitable and sustainable artificial intelligence for the global energy transition.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Research is to see what everybody else has seen, and to think what nobody else has thought.

Albert Szent-Györgyi

Chapter Overview

This chapter introduces the overall context, motivation, impact, and research direction of the thesis. It begins by describing the characteristics and challenges of remote areas, emphasizing energy access as a foundation for sustainable development and identifying microgrids as a viable and resilient solution for electrification where grid extension is not feasible. The chapter highlights electricity demand estimation as a critical factor influencing microgrid design, operation, and long-term sustainability, illustrating its impact through an exploratory study conducted within the LEOPARD project.

Building on this empirical foundation, the chapter defines the central research problem, focusing on forecasting electricity demand in data-scarce, fragmented, nonlinear, and dynamically evolving contexts where conventional and data-intensive approaches such as neural networks often struggle to generalize. To address this challenge, the thesis introduces a data-efficient neural network framework grounded in three methodological principles: causality, flexibility, and adaptability, which together enable robust learning in data-constrained contexts.

The chapter then outlines the main research goals and contributions, spanning theoretical, empirical, conceptual, design, and methodological dimensions, and situates them within the author's academic and professional trajectory. It concludes with a description of the thesis structure, which progresses from the definition of the research problem and theoretical foundations to methodological development, experimental validation, and final synthesis.

1.1 Context

Remote areas are territories located far from major urban centers, typically characterized by low population density, limited accessibility, underdeveloped infrastructure, and reduced availability of essential services [108]. The degree of remoteness often translates into lower levels of socioeconomic development and restricted access to opportunities, making these areas vulnerable to depopulation and economic stagnation.

Recent work has formalized the spatial and socioeconomic dimensions of remoteness by developing indices that combine population size with weighted distances to surrounding settlements, thereby quantifying how disconnected a place is from larger agglomerations [142].

Addressing the challenges associated with planning and implementing infrastructure and development projects in remote areas requires tailored rather than standardized methodologies [63]. Understanding their specificities is key to designing interventions that address local realities. Lasting impact arises from self-sustaining value creation, where communities build capabilities, generate economic opportunities, and strengthen resilience in ways meaningful to their context [125].

From this perspective, empowerment should not be reduced to a top-down consultation exercise but conceived as enabling local actors to co-design solutions. Effective interventions therefore emphasize knowledge transfer, local ownership, and participatory governance, ensuring that benefits persist long after external support diminishes. This idea of empowering remote areas also applies to energy projects: for electrification to succeed in the long term, communities must do more than just use electricity, they should take an active role in planning, operating, and managing their energy systems [26].

Energy access is a fundamental enabler of sustainable development, underpinning the autonomy and empowerment dimensions outlined above [130]. In the absence of reliable electricity, opportunities for innovation, entrepreneurship, and service delivery are severely constrained. Conversely, electrification supports schools, clinics, businesses, and digital connectivity, while fostering the skills and institutions required for long-term resilience and prosperity in remote areas [101].

Microgrids (MGs) have become a widely recognized solution for electrifying remote areas, where the extension of the centralized grid is economically or technically infeasible [105]. They operate through localized generation units, ensuring stability and reliability [53]. They integrate generation sources such as Photovoltaic (PV), wind, hydro, or biomass, often combined with storage, and can operate either independently or in connection with the main grid [128].

MGs provide technical advantages, including improved reliability, resilience, and operational flexibility, as well as socioeconomic gains, such as reduced reliance on fossil fuels and greater local participation [14]. Recent research has advanced their design and operation through optimization approaches for resource sizing, placement, and dispatch, aiming to minimize investment and operational costs [89].

However, the effectiveness of such optimization critically depends on an understanding of electricity demand. Electricity demand, defined as the rate at which electrical energy is consumed, forms the foundation for appropriate system sizing, operation, and realistic financial modeling [60].

During the design phase, minimizing forecasting errors related to current electricity demand ensures appropriate system sizing, avoiding oversized or insufficient capacity and unnecessary costs. During operation, anticipating future variations in electricity demand enables planners to balance generation and consumption in real time, thereby maintaining system reliability and efficiency [65]. Large estimation errors compromise both technical performance and economic sustainability, threatening the long-term viability of electrification projects [40].

Oversized MGs require higher investment and result in underutilized assets and unnecessary costs, whereas undersized MGs compromise reliability, leading to insufficient capacity and supply failures [117]. For these reasons, demand estimation is a critical bottleneck: without it, both the technical and economic sustainability of MGs is undermined [157, 77].

1.2 Motivation and Impact

The LEOPARD project [73], part of the LEAP-RE program, a large-scale initiative that fosters research and innovation collaboration between Europe and Africa to advance renewable energy, aims to disseminate MG technology by designing, testing, and evaluating a containerized solution for renewable energy production in rural and peri-urban areas of West Africa. Work carried out within the LEOPARD project highlighted that achieving low-error electricity demand estimation is a crucial element in MG design and operation. Beyond serving as the motivation for this research, this work also generated a tangible impact within the LEOPARD project by improving the understanding of how demand estimation errors affect MG performance, configuration, and operational efficiency.

To further investigate this relationship, the author conducted an exploratory study within the LEOPARD project [117] to evaluate the effects of different demand scenarios on the configuration, operation, and cost of the real MG installed at the Songhaï agro-ecological center in Porto-Novo, Benin.

Based on the measured data, three representative demand scenarios were defined:

- Reference Load Profile (RLP): the reference load profile was built using the data collected in November and December (annual demand of 115 MWh).

- Overestimated Load Profile (OLP): derived from extrapolating only the November data (100 MWh, a 13% underestimation).
- Underestimated Load Profile (ULP): derived from extending the December data (130 MWh, a 13% overestimation).

Figure 1.1 displays the daily averages of the three profiles, while Table 1.1 lists their hourly values.

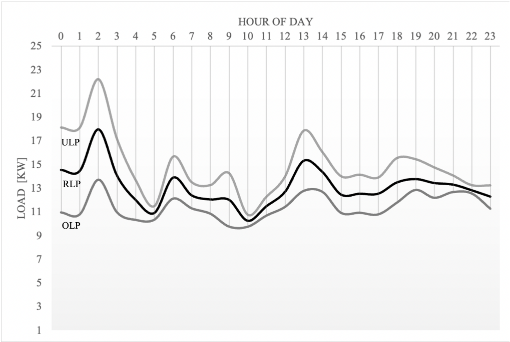


Figure 1.1: Daily average electricity demand for the RLP, OLP, and ULP scenarios at the Songhaï agro-ecological center in Porto-Novo.

The MG design was optimized using the Multi-Energy MOdeling for MGs (MemoGrid) platform, a simulation tool developed by European Institute for Energy Research (EIFER). The objective of the optimization was to minimize the Levelized Cost Of Electricity (LCOE) while satisfying specific constraints on diesel generation shares.

Four optimization scenarios for the MG configuration were executed, considering the RLP as the demand. Each scenario was defined by a maximum permissible share of diesel generation (Maximum Percentage of Electricity Production Accepted from Diesel engine (MPEPAD)) set at 75%, 50%, 25%, and 0%. These correspond to progressively stricter renewable integration targets and are referred to as Case 1, Case 2, Case 3, and Case 4, respectively.

The optimization outcomes show variations in PV and battery capacities across

Table 1.1: Load profiles [kW] for the Songhaï agro-ecological center in Porto-Novo under the OLP, RLP, and ULP scenarios.

Hour	OLP	RLP	ULP
0	10.9566	14.5384	18.1203
1	10.7940	14.4531	18.1122
2	13.7063	17.9549	22.2036
3	10.9543	14.0477	17.1411
4	10.3282	11.9979	13.6676
5	10.3505	10.9180	11.4854
6	12.1141	13.8881	15.6621
7	11.3003	12.3981	13.4958
8	10.8462	12.0504	13.2547
9	9.7667	12.0171	14.2675
10	9.7512	10.2543	10.7574
11	10.6867	11.4776	12.2685
12	11.4170	12.6952	13.9734
13	12.7637	15.3096	17.8555
14	12.6860	14.3704	16.0548
15	10.9213	12.4663	14.0113
16	10.9298	12.5336	14.1373
17	10.7929	12.5370	13.9011
18	11.7799	13.4879	15.5425
19	12.8455	13.7692	15.4435
20	12.1852	13.4337	14.7487
21	12.6666	13.3021	14.0723
22	12.5543	12.8270	13.2565
23	11.2726	12.2912	13.2319

the four cases. In Case 1, the optimization suggested a PV capacity of 43 kW with no battery storage. In Case 2, the same PV capacity of 43 kW was maintained, but a 55 kWh battery was added. In Case 3, the PV capacity increased to 68 kW, accompanied by a 140 kWh battery. Finally, in Case 4, the optimization resulted in the largest configuration, with 132 kW of PV capacity and 255 kWh of battery storage.

The optimal configurations of the MG for each of the four cases were applied to the OLP and ULP to assess deviations in annual energy losses, engine electricity

Table 1.2: Performance indicators for the MG configuration with 43 kW of PV capacity and no battery storage, evaluated under the OLP, RLP, and ULP demand scenarios at the Songhaï agro-ecological center in Porto-Novo.

	OLP	RLP	ULP	
Annual energy loss	27	23	20	MWh
Engine electricity share	61	63	64	%
LCOE	0.131	0.136	0.140	€/kWh

Table 1.3: Performance indicators for the MG configuration with 43 kW of PV capacity and a 55 kWh battery, evaluated under the OLP, RLP, and ULP demand scenarios at the Songhaï agro-ecological center in Porto-Novo.

	OLP	RLP	ULP	
Annual energy loss	10	7	4	MWh
Engine electricity share	45.6	49.8	54	%
LCOE	0.159	0.160	0.163	€/kWh

share, and LCOE with respect to the demand for which the MG configuration was optimized. Each application of an optimal design to the OLP and ULP is referred to as an experiment. Accordingly, Experiments 1 to 4 correspond to the optimal designs of Cases 1 to 4, respectively, when applied to both the OLP and ULP.

Across all configurations, the results revealed a strong sensitivity of MG performance to errors in demand estimation. Table 1.2 summarizes the results of Experiment 1, while Tables 1.3, 1.4, and 1.5 show the outcomes of Experiments 2, 3, and 4, respectively.

Table 1.4: Performance indicators for the MG configuration with 68 kW of PV capacity and a 140 kWh battery, evaluated under the OLP, RLP, and ULP demand scenarios at the Songhaï agro-ecological center in Porto-Novo.

	OLP	RLP	ULP	
Annual energy loss	18	15	11	MWh
Engine electricity share	17	24.7	31	%
LCOE	0.184	0.179	0.175	€/kWh

Table 1.5: Performance indicators for the MG configuration with 132 kW of PV capacity and a 255 kWh battery, evaluated under the OLP, RLP, and ULP demand scenarios at the Songhaï agro-ecological center in Porto-Novo.

	OLP	RLP	ULP	
Annual energy loss	95	81	66	MWh
Engine electricity share	0	0	0.7	%
LCOE	0.277	0.245	0.221	€/kWh

In Experiment 1, energy losses increased by 17% under the OLP and decreased by 13% under the ULP. The engine electricity share varies, decreasing by about 3% in the OLP and increasing by nearly 2% in the ULP. Similarly, the LCOE changes slightly, decreasing by about 4% in the OLP and increasing by nearly 3% in the ULP.

In Experiment 2, where a battery was introduced, the system became more sensitive to demand fluctuations: the OLP led to a 43% increase in losses, while the ULP achieved a comparable reduction. The engine electricity share decreases by more than 8% in the OLP, while it increases by a similar margin in the ULP. Notably, in the ULP, the objective of maintaining a maximum engine electricity share of 50% is no longer achieved, as it rises to 54%. The LCOE follows a trend similar to that observed in Experiment 1, with a reduction of less than 1% in the OLP and an increase of nearly 2% in the ULP. Overall, integrating a battery into the optimal design for this experiment reduces annual energy losses compared to the previous configuration, albeit at the expense of a higher LCOE.

In Experiment 3, the OLP increased losses by 20%, while the ULP reduced them by 27%. In the OLP, the engine electricity share decreases by more than 31%, while it increases by nearly 26% in the ULP. As in the previous case, it is noteworthy that in the ULP, the MPEPAD constraint is no longer satisfied, as the engine electricity share reaches 31%. The LCOE exhibits the opposite trend compared to the preceding experiment, increasing by almost 3% in the OLP and decreasing by more than 2% in the ULP.

In Experiment 4, the results indicate that annual energy losses increase by approximately 17% in the OLP and decrease by around 19% in the ULP. For the 0% MPEPAD case, the findings show that the system can meet all load profiles exclusively through renewable energy sources, although in the ULP this share is nearly, but not entirely, zero. Achieving this performance requires substantially larger PV and battery capacities to accommodate fluctuations in the load profiles. The LCOE exhibits a marked variation, increasing by about 13% in the OLP and decreasing by nearly 10% in the ULP. Although the LCOE in this experiment is relatively higher than in the previous one, it is important to account for the environmental advantages associated with relying on renewable energy sources.

The results of the four experiments indicate that overestimation (ULP) leads to oversizing and lower utilization of installed capacity, whereas underestimation (OLP) results in insufficient capacity, increased operational effort, and greater reliance on diesel generation. These findings clearly demonstrate that minimizing electricity demand estimation errors is essential for reliable and cost-effective MG planning.

1.3 Problem

Electricity demand forecasting is generally conducted through data-driven modeling, an evidence-based approach that relies on historical consumption data to identify temporal patterns, quantify relationships among variables, and predict future behavior. These models encompass a broad range of statistical and Machine Learning (ML) techniques, such as regression analysis, time-series models, or neural networks, that infer the relationship between past and future demand based on observed data.

In general, the effectiveness of such methods fundamentally depends on the availability of consistent, high-quality datasets that capture the diversity of consumption behaviors over time. When data are insufficient or inconsistent, model calibration becomes unreliable, parameter estimation is biased, and predictive error increases. In these circumstances, even sophisticated algorithms are unable to extract meaningful patterns, leading to large forecasting errors and high uncertainty in system design.

Ultimately, the challenge of electricity demand forecasting is a problem of data. This problem becomes particularly acute in remote areas, where the conditions required for reliable data collection and modeling are rarely met. As discussed below, the forecasting process in such contexts faces three problems:

1. **Data scarcity and fragmentation.** Historical demand records are typically unavailable, incomplete, or collected over short and irregular time spans [161]. *Data scarcity* stems from the limited number of monitored sites, short measurement periods, and the lack of continuous metering infrastructure to capture long-term electricity use. In many remote areas, monitoring systems operate only intermittently or at small scales, producing datasets with narrow spatial and temporal coverage. *Data fragmentation*, in turn, refers to the coexistence of heterogeneous, incomplete, and non-synchronized records originating from diverse devices, time scales, or aggregation levels. It typically arises from limited or unreliable monitoring infrastructure, intermittent data acquisition, and ad hoc measurement practices, leading to discontinuities and inconsistencies that prevent coherent temporal or spatial modeling [99]. Consequently, available datasets often suffer from low temporal resolution, missing values, and measurement errors, reflecting the technical and logistical constraints typical of remote environments [54].

2. **Nonlinear dependencies among socio-technical and environmental variables.** Electricity demand in remote areas is influenced by complex and interdependent relationships among socio-technical and environmental fac-

tors [114]. On the socio-technical side, variables such as income, household composition, appliance adoption, and grid reliability shape consumption behaviors [52]. On the environmental side, factors such as temperature and humidity further affect electricity use by altering comfort needs and appliance operation [139]. These relationships are inherently nonlinear: small variations in socioeconomic or environmental conditions can lead to disproportionate changes in demand patterns, while infrastructural or behavioral constraints may amplify or mitigate such effects. Capturing these nonlinearities requires modeling approaches capable of representing feedback loops, threshold effects, and coupled dynamics across multiple domains, ultimately achieving lower predictive errors than conventional linear regression-based or purely deterministic models [75].

3. **Nonstationary demand evolution.** In newly electrified areas, electricity demand changes rapidly as households and businesses progressively adopt new appliances, improve living standards, and expand productive activities. This dynamic evolution makes past data quickly lose predictive value, violating the stationarity assumptions underlying most conventional forecasting methods [138].

In remote areas, this combination of scarce and fragmented data, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution severely limits the applicability of traditional forecasting techniques. These classical modeling approaches typically require large, high-quality, and stationary datasets to identify meaningful patterns and avoid overfitting or spurious correlations [126, 31].

The challenge becomes even more pronounced when employing neural networks for demand forecasting. These models are inherently data-intensive: their performance and generalization capabilities rely on access to large and diverse training datasets that adequately capture the variability of real-world consumption. In data-constrained contexts, neural networks struggle to learn stable representations, becoming prone to overfitting and poor generalization. As a result, their prediction error increases sharply [151].

The limitations of neural network–based forecasting in data-constrained conditions define the central research problem of this thesis, motivating the development of data-efficient, adaptive neural network-based models capable of learning from limited and evolving information.

1.4 Research Goals

Building on these considerations, the present research advances the estimation of electricity demand for MGs in contexts characterized by (i) data scarcity and fragmentation, (ii) nonlinear dependencies among socio-technical and environmental variables and (iii) nonstationary demand evolution.

This thesis introduces a data-efficient neural network framework for forecasting electricity demand in MGs, as improving data efficiency is essential to ensure reliable forecasts. The approach aims to enhance the understanding and representation of demand, providing a foundation for reducing forecasting errors in the design, operation, and planning of MGs in remote areas. The proposed methodology integrates three complementary principles that underpin the learning process: causality, flexibility, and adaptability.

Causality is implemented through a feature engineering methodology, which operationalizes causal reasoning by constructing stable representations of electricity demand. This methodological approach identifies and models cause–effect relationships among variables, providing the theoretical foundation for stable and reliable learning. By extracting meaningful relationships among environmental, temporal, and behavioral factors influencing electricity use, feature engineering transforms empirical data into causally grounded inputs suitable for robust forecasting [122].

Flexibility is expressed through a modeling methodology, which captures nonlinear and context-dependent interactions within the data. This approach employs flexible architectures that adjust their response functions based on observed relationships rather than relying on fixed transformations. Unlike conventional neural networks with static activation functions, methods such as Kolmogorov–Arnold Network (KAN) allow the model to refine its internal representations as new information becomes available and to represent complex dependencies with lower error [76].

Adaptability is realized through an adjustment methodology, which maintains coherence between the model and the evolving conditions of electrification. By incorporating iterative updating and recalibration, this methodological component captures feedback loops, threshold effects, and behavioral shifts associated with progressive adoption [156].

By uniting the principles of causality, flexibility, and adaptability within a framework, this research contributes to the broader field of data-efficient Artificial Intelligence (AI) for sustainable energy systems. The resulting methodology represents both a theoretical advancement in electricity demand forecasting and a practical decision-support tool for planners and policymakers pursuing resilient, inclusive, and adaptive electrification.

In light of these reflections, this thesis pursues the following research goals:

1. To develop a neural network-based, data-efficient framework for forecasting electricity demand in MGs that remains reliable under conditions of data scarcity, fragmentation, nonlinearity, and nonstationarity.
2. To enhance the representation and understanding of demand dynamics by integrating causal reasoning, flexible modeling, and adaptability within a unified neural network–based methodological structure.

3. To operationalize three methodological approaches: feature engineering, modeling, and adjustment, corresponding respectively to the principles of causality, flexibility, and adaptability, thereby linking theoretical foundations with neural network–driven computational implementation.
4. To evaluate the proposed framework's robustness in data-constrained MGs, assessing its capacity to minimize predictive error as consumption patterns evolve.
5. To advance data-efficient AI for sustainable energy systems, providing a theoretical foundation for learning in data-constrained contexts and a practical decision-support tool for planners and policymakers pursuing resilient electrification.

1.5 Contributions

This doctoral research advances electricity demand forecasting in remote areas by introducing a neural network–based theoretical framework for data-efficient learning in MGs. The framework connects epistemological, theoretical, and computational dimensions of learning, providing an integrated rationale for robust forecasting under data scarcity, fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution. Its novelty lies in defining a coherent system of knowledge structured across four epistemic levels: Problem, Component, Design, and Methodology, which together link empirical challenges, conceptual organization, design foundations, and methodological realization.

At the theoretical level, the research formulates a unified framework that integrates causal reasoning, flexibility, and adaptability as the core principles of data-efficient forecasting. It establishes the theoretical basis for learning in data-constrained contexts, defining how valid and robust knowledge can be generated, represented, and updated. This framework constitutes the central theoretical innovation of the thesis, bridging epistemology, theory, and computation in a coherent structure for data-efficient neural network modeling.

At the empirical level, the research identifies and characterizes the key sources of forecasting uncertainty in MGs, namely data scarcity, fragmentation, nonlinear socio-technical dependencies, and nonstationary demand evolution. By positioning these as the epistemic context within which data-efficient forecasting must operate, the thesis clarifies why conventional models fail in remote areas and establishes the diagnostic foundation for new methodological development.

At the conceptual level, forecasting in remote areas is redefined as a causal, flexible, and adaptive process structured around three cognitive functions: **Observation**, addressing data scarcity and fragmentation through causal reasoning; **Representation**, capturing nonlinear dependencies by enabling flexible modeling of socio-technical and environmental relationships; and **Evolution**, accounting for nonstationary demand through adaptability and dynamic updating. This redefinition bridges empirical challenges and computational implementation, framing forecasting as an explanatory and learning process rather than a purely predictive one.

At the design level, the thesis formalizes three normative principles, namely causality, flexibility, and adaptability, that govern data-efficient learning. Causality grounds learning in invariant relationships among physical, environmental, and behavioral variables. Flexibility enables generalization through nonlinear, context-aware architectures. Adaptability maintains predictive coherence as system dynamics evolve. Together, these principles define the epistemological conditions for robust forecasting and provide a normative theory of data-efficient neural learning.

At the methodological level, the framework operationalizes these principles through three components: **Feature engineering**, implementing causal reasoning; **Modeling**, expressing flexibility through learned nonlinear functions; and **Adjustment**, realizing adaptability via iterative updating. This integration transforms theory into practice, yielding a reproducible and scalable forecasting pipeline that achieves low-error forecasts.

Beyond methodological innovation, the framework represents a broader scientific contribution to AI for sustainable energy systems. It offers a principled model of learning in remote areas that aligns forecasting performance with social relevance. By embedding causal reasoning, ensuring methodological openness, and adhering to open-science principles such as reproducibility and accessibility, the research advances a vision of AI as a tool for understanding and discovery rather than a black-box predictor. It contributes to the development of data-efficient and socially responsible models that support equitable and resilient energy transitions.

This thesis contributes:

- The development of a neural network–based theoretical framework that unifies epistemological, theoretical, and computational perspectives for data-efficient forecasting in MGs.
- The identification of empirical drivers of forecasting uncertainty in remote areas, providing diagnostic insight into data scarcity, fragmentation, nonlinearity, and nonstationarity.
- The conceptualization of forecasting as a causal, flexible, and adaptive learning process linking explanatory understanding with predictive modeling.
- The formalization of a normative foundation for data-efficient learning grounded in the principles of causality, flexibility, and adaptability.
- The implementation of an operational forecasting pipeline that integrates feature engineering, nonlinear modeling, and iterative adjustment to achieve

low-error, reproducible predictions.

- The advancement of transparent and socially responsible AI practices that promote equitable and sustainable energy transitions.

1.6 Personal Background

The author's academic and professional trajectory has progressively led to this doctoral research by identifying electricity demand estimation in remote areas as the critical bottleneck in MG design and operation. This journey unfolded as follows.

During his master's thesis in Energy and Nuclear Engineering, the author developed a simulation-based decision support tool for the Songhaï MG in Porto Novo, Benin, a remote rural community [116]. The tool integrated PV panels, batteries, biogas systems, gasifiers, and diesel generators to meet local electricity demand. Using a multi-method approach, the author incorporated local data, such as solar irradiation, ambient temperature, and biomass availability, to conduct a technical and economic feasibility analysis. This work initially revealed the critical importance of reliable demand estimation for effective MGs design and operation.

Subsequent professional experience at EIFER further built on this foundation through the simulation of fully renewable, multi-energy MGs across various regions, including Africa and Asia. These simulations incorporated diverse technologies such as fuel cells, hydrogen storage systems, electrolyzers, and batteries [104]. His experience reinforced the importance of achieving low error in demand estimation for MG design and operation, while also highlighting the challenges posed by scarce and fragmented data, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand patterns in remote areas. Moreover, this experience demonstrated that high-error input data in simulation and optimization tools can lead to erroneous design and operational outcomes

[117].

Currently, the author works in an innovation-oriented environment at STAM S.r.l. (STAM), where research activities have consolidated this expertise through the modeling and optimization of multi-energy systems, including MGs, Positive Energy Districts (PEDs), and Renewable Energy Communitys (RECs) [118, 119, 120]. Collectively, these experiences have converged on a central insight: minimizing forecasting error in electricity demand estimation is the foundational challenge for sustainable MGs in remote areas [2, 153].

1.7 Structure of the Thesis

This thesis is structured into six main chapters, each progressively contributing to a coherent narrative that moves from the definition of the research problem to methodological design, empirical validation, and final synthesis.

Chapter 1 presents the overall context, motivation, and impact of the research, emphasizing the role of electricity demand estimation in the design and operation of MGs for remote areas. It identifies the main challenges associated with data scarcity, fragmentation, nonlinearity, and nonstationarity, defines the research problem, and introduces the proposed methodological vision based on causality, flexibility, and adaptability. The chapter also outlines the research goals and key contributions and concludes with the structure of the thesis.

Chapter 2 provides a comprehensive review of the state of the art in electricity demand forecasting. Section 2.1 presents major classification frameworks, distinguishing approaches by forecasting horizon, system conceptualization, application domain, and data orientation. Section 2.2 reviews the principal forecasting methodologies, including time series, causal and econometric, statistical and probabilistic, end-use, Geographic Information System (GIS)-, Agent Based (AB)-, structural, scenario-based, and AI/ML-based methods, laying the groundwork for identifying research gaps. Section 2.3 reviews forecasting methods for MGs, analyzing statistical, machine learning, and deep learning approaches in terms of their ability to address data scarcity and fragmentation, nonlinear dependencies, and nonstationary demand. It concludes by outlining the methodological gaps that motivate the proposed forecasting framework.

In **Chapter 3**, Section 3.1 identifies the research gap motivating this work; Section 3.2 formulates the research questions that guide the thesis; Section 3.3 presents the proposed methodological approach, detailing the Conceptual Foundations, Architecture of ANGEL, Causal Feature Engineering, Flexible Modeling, and Adaptive Adjustment; Section 3.4 outlines the scientific and practical significance of the study; and Section 3.5 presents the expected contributions of the research.

Chapter 4 outlines the research methodology. Section 4.1 describes the overall research design; Section 4.2 states the research hypotheses; Section 4.3 details the research context and data sources; and Section 4.4 defines the experimental setup. Sections 4.5 and 4.6 introduce the evaluation metrics and validation strategy. Section 4.7 discusses the practical considerations related to implementation.

Chapter 5 reports the experimental validation. Section 5.1 explains the preparation of data and variables; Section 5.2 analyzes causal feature engineering; Section 5.3 examines flexible representation; Section 5.4 investigates latent temporal evolution; Section 5.5 evaluates integrative adaptive forecasting; and Section 5.6 synthesizes the experimental findings and critically examines their methodological implications and limitations.

Finally, **Chapter 6** concludes the thesis. Section 6.1 presents the empirical and theoretical synthesis, establishing how the research questions, principles, and findings converge into a unified framework. Section 6.2 details the methodological contributions that operationalize this framework into a reproducible and adaptive forecasting pipeline. Section 6.3 examines the practical and technological implications derived from the implementation of the proposed models, while Section 6.4 outlines directions for future research. Finally, Section 6.5 offers a concluding reflection on the overall significance and impact of the work.

Chapter 2

State of the art

Prediction is very difficult, especially if it's about the future.

Niels Bohr

Chapter Overview

This chapter provides a comprehensive review of the state of the art in electricity demand forecasting. It begins by outlining the main classification frameworks that organize forecasting approaches according to forecasting horizon, system conceptualization, application domain, and data orientation.

The second part reviews the principal methodological families, including time series, causal and econometric, statistical and probabilistic, end-use, GIS-, AB-, structural, scenario-based, and AI/ML-based methods.

The final part focuses on forecasting methods applied to MGs, critically examining how statistical, ML, and deep learning approaches address the main challenges of data scarcity and fragmentation, nonlinear dependencies, and nonstationary demand. The discussion highlights the growing prominence of neural-network-based methods and the ongoing shift from correlative to causal and adaptive models capable of operating in remote areas.

The chapter concludes by identifying the methodological and operational gaps that limit forecasting reliability in data-constrained contexts, providing the rationale for the data-efficient, causally grounded framework proposed in this thesis.

2.1 Classification Frameworks

Electricity demand forecasting has attracted extensive research attention, resulting in a wide spectrum of methodologies ranging from classical statistical and regression-based models to modern ML and hybrid approaches [50]. Despite continuous methodological advancements, no single model has proven universally optimal, as forecasting performance remains highly dependent on the characteristics of each power system and its underlying socioeconomic and environmental conditions. As energy systems evolve toward greater decentralization, electrification of end uses, and increased demand-side flexibility, the forecasting task has become more complex and critical [93].

From an ontological and taxonomic perspective, forecasting methods can be systematically organized according to distinct conceptual criteria. This classification aims to clarify the fundamental dimensions along which electricity demand forecasting approaches differ, providing a structured understanding of their theoretical scope, methodological assumptions, and practical applications.

Researchers have proposed multiple frameworks that categorize forecasting techniques based on key analytical dimensions, such as the forecasting time horizon, modeling paradigm, and application domain [68]. The following subsections examine each of these dimensions in detail, highlighting their conceptual basis and practical implications for electricity demand forecasting.

2.1.1 Time-Horizon-Based Forecasting Classification

Electricity demand forecasting methods are often categorized according to the time horizon of the forecast [12]. This classification distinguishes between three main categories: Short-Term Electricity Demand Forecasting (STF), which typically spans from a few hours to several days or weeks ahead; Medium-Term Electricity Demand Forecasting (MTF), which covers horizons from several months to a few years; and Long-Term Electricity Demand Forecasting (LTF), which extends from around five years to several decades into the future [93]. Each category, short-, medium-, and long-term forecasting, serves distinct objectives, relies on different data resolutions, and employs specific methodological approaches, reflecting the diverse requirements of operational, tactical, and strategic decision-making in power systems [132].

STF

In increasingly complex grids with high shares of renewables and distributed resources, STF is essential for reliable and cost-effective power system operation, as it directly affects grid stability and market efficiency. STF focuses on minimizing prediction errors caused by rapid fluctuations in load driven by daily cycles, weather variability, social activity, and short-term behavioral dynamics. Unlike medium- or long-term forecasts, which address structural and strategic planning, STF supports real-time operational decisions that require continuous adaptation to changing conditions [4].

Low error STF enables optimal generation scheduling, lowers reserve and balancing costs, and enhances market trading strategies. It also contributes to maintaining grid stability by anticipating demand variability and activating reserves or demand-response mechanisms [51]. With the growing integration of renewable sources such as wind and solar, STF has become increasingly crucial for forecasting net load (demand minus renewable generation) and supporting storage management and demand-side programs, ultimately improving peak shaving and overall system efficiency.

A diverse set of methods is employed for STF, including statistical methods, ML algorithms, hybrid approaches, and probabilistic forecasting techniques:

- Statistical methods: Methodologies such as Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and exponential smoothing capture temporal autocorrelation and seasonality in high-resolution datasets. Variants like ARIMA with exogenous variables (ARIMAX) integrate exogenous variables (e.g., weather, holidays) to reduce forecasting error [102].

- ML Techniques: Methods such as Artificial Neural Network (ANN), Support Vector Regression (SVR), and Gradient Boosting (GB) capture nonlinear dependencies in demand data. They accommodate multiple variables (e.g., temperature, humidity, weekday effects, events) to achieve low forecasting error [131].
- Hybrid methodologies: Combinations of time series and ML methods capture both linear and nonlinear patterns. For example, ARIMA residuals can be modeled with neural networks to address nonlinearities [18].
- Probabilistic Forecasting: Increasingly, STF integrates uncertainty quantification, producing forecast intervals rather than point forecasts. Techniques include quantile regression and Bayesian inference, which support risk-aware real-time decision-making [79, 27].

The advantages of STF are immediate and practical: it directly supports daily grid operations and market decisions. High-Frequency (HF) historical data facilitate low-error model training, allowing methods to capture rapid fluctuations and quickly adapt to system changes or anomalies. However, STF faces challenges: forecasts are highly sensitive to unexpected events (e.g., extreme weather, industrial shifts), and minute-scale fluctuations are noisy, requiring error-resilient methods and frequent recalibration [35].HF, particularly when using AI, ML or ensemble approaches, can also be computationally intensive [39]. Furthermore, integrating diverse data sources (e.g., renewables, electric vehicle charging, sensor data) adds complexity to model design and deployment [5].

MTF

MTF, covering horizons from several months to a few years, serves as a bridge between operational and strategic planning. It captures broader temporal patterns such as seasonal variations, economic cycles, technological shifts, and policy effects that influence demand evolution over intermediate periods [30]. Low-error MTF supports resource allocation, capacity planning, maintenance scheduling, and energy trading, enabling operators to optimize asset utilization and anticipate long-term trends [6].

A diverse set of methodologies is employed for MTF, encompassing time series, econometric, ML, and scenario-based approaches:

- Time Series Methods: Techniques such as SARIMA, Holt–Winters exponential smoothing, and multivariate models like ARIMAX capture seasonal and cyclical demand patterns, and remain fundamental for medium-term forecasting tasks [123].
- Econometric Methods: Regression-based models establish relationships between demand and macroeconomic indicators (e.g., Gross Domestic Product (GDP), industrial output, population, and energy prices), providing insights into structural drivers [44].
- ML Techniques: Algorithms such as Random Forest (RFS), GB, and ANN identify nonlinear interactions among multiple explanatory variables and reduce forecasting error in complex contexts [131].
- Scenario-Based Approaches: These explore uncertainties linked to economic growth, technological adoption, policy evolution, and climate variability, supporting resilient and adaptive planning [64].

The advantages of MTF are both operational and strategic: it integrates socioeconomic and policy factors, aligns near-term operations with long-term investment goals, and enhances flexibility through scenario analysis [95]. However, challenges persist: MTF depends on high-quality, multi-year datasets and is sensitive to structural shifts in technology, market behavior, and regulation. Moreover, it lacks the temporal granularity required for real-time balancing [30].

LTF

LTF, spanning horizons from approximately five years to several decades, serves primarily a strategic role. It supports long-term infrastructure development, policy formulation, and sustainable energy transition planning under conditions of deep uncertainty. By integrating economic, demographic, technological, and environmental drivers, LTF captures the structural factors that shape demand evolution over extended periods [42].

A diverse range of methodologies is employed for LTF, encompassing econometric, scenario-based, structural, and hybrid approaches:

- Econometric and Macroeconomic Methods: Regression and input–output frameworks link electricity demand with macroeconomic indicators such as GDP, industrial output, population growth, and sectoral activity, providing quantitative relationships between demand and socioeconomic development [43].
- Scenario-Based Approaches: Widely used to explore uncertainties in economic growth, technology adoption, policy evolution, and climate change, enabling resilient and adaptive long-term planning [64].
- Structural and Engineering Methods: Simulate the physical and technical characteristics of energy systems, including generation, transmission, and efficiency measures, to represent system-level behavior and capacity expansion needs [74].
- Hybrid Methods: Integrate econometric, scenario-based, structural, and,

> increasingly, AI- and ML-driven techniques to produce comprehensive, data-informed, and flexible long-term projections [61].

LTF advantages include strategic relevance, holistic integration of multiple drivers, and flexibility in exploring alternative futures. However, LTF faces challenges related to deep uncertainty, data requirements, model complexity, and limited temporal resolution [84]. The rationale behind LTF aligns closely with Top-Down Method (TDM) philosophies. Because long-term analyses emphasize aggregated trends, sectoral dynamics, and macro-level drivers, TDMs, estimating sectoral demand and distributing demand among subgroups, complement LTF by revealing broad trajectories and policy impacts.

In this context, forecasting methods can also be classified according to their underlying modeling paradigms, distinguishing between TDMs, Bottom-Up Methods (BUMs), and hybrid frameworks [86]. The following subsection examines these paradigms in detail.

2.1.2 System-Conceptualization-Based Classification

This classification offers distinct perspectives on system representation and data aggregation. TDMs models aggregate demand as a function of macroeconomic, demographic, or climatic variables, emphasizing broad system-level relationships. BUMs, conversely, simulate electricity demand by aggregating individual end-use or device-level behaviors, capturing detailed technological and behavioral dynamics. Hybrid approaches combine both paradigms, integrating macro- and micro-level insights to reduce error while improving adaptability. Each methodology serves different analytical objectives and data contexts, contributing to a comprehensive understanding of energy demand across scales and applications [8].

TDMs

TDMs begin with aggregate energy demand data at regional, national, or sectoral levels, which is then disaggregated into smaller subgroups using macro-level characteristics. Typical explanatory variables include GDP, population, household income, industrial activity, urbanization rates, and climatic factors [147]. Statistical or econometric methods, such as regression analysis, input-output frameworks, or econometric time series methods, are employed to establish relationships between total electricity demand and these macro variables. Once aggregate demand is estimated, it is distributed across sectors (e.g., residential, commercial, industrial), or even individual households or firms, using allocation rules, scaling factors, or proxy variables [34].

TDMs present several advantages. They are highly data-efficient, requiring only aggregate or sector-level data, which is generally easier to obtain than detailed micro-level information. Their simplicity makes them straightforward to implement, while results provide clear insights into overall demand trends and the relationship between electricity use and macroeconomic drivers. These methods are particularly suitable for medium- and long-term forecasting, infrastructure planning, and policy evaluation, where broad structural trends are more important than fine-scale detail. Moreover, their scalability allows application across large regions or entire countries [23].

Despite these strengths, TDMs face important limitations. Their granularity is limited, as they cannot capture detailed variations at household or appliance levels, potentially overlooking behavioral and technological effects. Allocation procedures often assume homogeneity within subgroups, failing to represent heterogeneous populations. They are also sensitive to structural shocks such as rapid technological adoption, energy efficiency improvements, or economic crises, which may introduce significant errors. Furthermore, they lack behavioral insights, as they do not explicitly account for user-level decision-making or Demand-Side Management (DSM) interventions [8].

BUMs

BUMs estimate electricity demand by aggregating demand at the most granular level possible, such as households, businesses, or specific appliances. Data sources include appliance inventories, household surveys, smart meters, and experimental measurements. Consumers are grouped into categories based on socioeconomic status, building type, or appliance ownership. Energy use for each unit or category is then estimated using operating schedules, usage profiles, or probabilistic methods, and results are aggregated to obtain sectoral, regional, or national forecasts [134].

The main advantage of BUMs is their high granularity. They capture appliance-level loads, behavioral variability, and detailed usage patterns, making them highly relevant for policy evaluation and targeted interventions. They are flexible, allowing explicit representation of new technologies, energy efficiency measures, or demand-response strategies. Scenario analysis is a key feature, as BUMs enable “what-if” experiments to evaluate the effects of appliance adoption, lifestyle changes, or renewable energy integration on overall demand [134].

However, BUMs are data-intensive, requiring detailed, high-quality information that is often costly or unavailable, particularly in remote regions. Modeling many devices or user categories increases computational effort and complexity, and scalability becomes challenging when the methodology is applied to large populations or entire countries. Moreover, while they are effective for structural and long-term analysis, BUMs may be less suited to operational short-term forecasting unless HF data is incorporated [7].

Hybrid methodologies

Hybrid methodologies combine the strengths of TDMs and BUMs to provide forecasts that are both granular and consistent with aggregate trends. Typically, a TDM estimates aggregate demand using sectoral or national data, providing macro-level constraints. A bottom-up component simultaneously captures detailed demand behavior at the household, appliance, or sector level. The two components are reconciled through calibration or scaling procedures, aligning micro-level estimates with aggregate totals. This integration ensures consistency across scales while preserving heterogeneity in demand patterns. Hybrid methods also support scenario analysis, enabling the evaluation of policy measures, efficiency programs, or technological adoption within realistic system-wide contexts [8].

Hybrid methodologies provide several advantages. They balance scalability with granularity, capturing both overall trends and local variation. They can adapt across different spatial and temporal scales, from households to national grids, and from STF to long-term horizons. By leveraging complementary insights, hybrid methods often achieve lower error than single-method approaches. They are particularly relevant for planning and policy-making, supporting targeted interventions while maintaining consistency with aggregate demand projections [46].

Nevertheless, hybrid methods are complex. They require expertise in both TDMs and BUMs, as well as substantial computational resources to reconcile heterogeneous datasets. Their data requirements are also high, as both detailed micro-level information and reliable aggregate statistics are necessary. Moreover, the integration of multiple paradigms can complicate communication with stakeholders [46].

Yet, the traditional division into TDMs, BUMs, and hybrid methodologies does not fully capture the diversity of modern forecasting practices.

2.1.3 Application-Oriented Classification

Electricity demand forecasting can also be classified according to its application domain, which defines the decision context, temporal scope, and performance criteria of the forecasting task. Each domain, operational, tactical, strategic, and sector specific, serves distinct functions within the broader energy management and planning landscape.

Operational Forecasting

Operational forecasting supports control, balancing, and demand-response mechanisms in power systems that operate on short-term or real-time horizons. It is essential for grid stability, market operations, and the integration of renewable and distributed energy resources. Low-error short-term forecasts enable operators to anticipate rapid demand fluctuations, schedule generation efficiently, and activate flexibility resources such as storage and DSM. Typical applications include smart grid operation, electric vehicle charging optimization, and local MG management, where forecasts often rely on HF data and adaptive ML models. These models must handle noise, volatility, and nonlinearities in real time, ensuring responsiveness and robustness under uncertainty [4].

Tactical Forecasting (Medium-Term)

Tactical forecasting spans horizons from several months to a few years and supports planning activities that connect operational efficiency with long-term strategies. Its primary goals include resource allocation, maintenance scheduling, energy trading, and mid-term market analysis. Forecasting models at this level combine time-series and econometric techniques with scenario-based and hybrid approaches to capture seasonality, economic cycles, and evolving policy frameworks. Data inputs often include meteorological variables, macroeconomic indicators, and demand profiles aggregated over time, providing a balance between temporal detail and computational feasibility. The outputs guide tactical decisions such as generation portfolio adjustments, procurement strategies, and hedging in energy markets [95].

Strategic Forecasting (Long-Term / Policy-Oriented)

Strategic forecasting underpins infrastructure development, policy formulation, and long-term sustainability planning. It focuses on structural and systemic drivers such as macroeconomic growth, demographic evolution, technological innovation, and regulatory change that shape electricity demand over decades. Models in this category employ econometric, structural, and scenario-based methodologies to explore alternative futures and assess long-term investment risks. Because strategic forecasts emphasize aggregated patterns over operational variability, they typically rely on low-frequency data and simplified representations of temporal dynamics. Applications include national energy system modeling, grid expansion planning, and evaluation of decarbonization pathways [42].

Sector-Specific Forecasting (Residential, Industrial, Commercial)

Sector-specific forecasting targets distinct demand patterns and behavioral dynamics across end-use sectors. Residential, commercial, and industrial domains exhibit heterogeneous demand drivers, ranging from occupant behavior and appliance usage to production cycles and process loads, that require differentiated modeling approaches. Residential load profiling, for example, benefits from high-frequency data and behavioral modeling [93, 47], while industrial forecasting often leverages production indicators and process-level variables. Sectoral differentiation reduces error by aligning model structure with the physical and behavioral characteristics of each user category. Such forecasts are increasingly valuable for targeted policy design, demand-side interventions, and distributed resource optimization [25].

2.1.4 Data-Oriented Classification

The application-oriented classification described above frames electricity demand forecasting in relation to its purpose, timescale, and decision-making context. However, forecasting performance also depends critically on the nature and quality of the data available for model development. To address this complementary dimension, this section introduces a data-oriented classification, developed in this research, which organizes forecasting methods according to the type, granularity, and completeness of the data they require. The classification identifies four main categories: microdata-based, macrodata-based, prototypical, and hybrid data methods.

Microdata-Based Methods

Microdata-based methods are grounded in the use of highly disaggregated datasets, often collected at the appliance or household level. These may include time-use surveys, occupant activity logs, or high-resolution smart meter and sensor data, which enable the reconstruction of individual demand behaviors with fine temporal granularity.

Typically, these methods are implemented through BUMs, either deterministic or stochastic. Monte Carlo simulations are often used to capture variability and uncertainty, AB modeling allows for detailed emulation of household decision-making, and activity-based methods link social routines with energy use [21].

The main strength of microdata-based methods lies in their ability to capture heterogeneity across users and devices. This makes them particularly valuable for DSM programs, targeted interventions, and MG design, where minimizing component-level error is essential. Their limitations, however, are significant: extensive data collection is costly and time-consuming, while the resulting methods require substantial computational resources [47]. For these reasons, their application is often confined to pilot projects or research studies rather than large-scale system planning.

Macrodata-Based Methods

Macrodata-based methods, often referred to as TDMs, rely on aggregated indicators such as regional or national electricity demand, GDP, urbanization rates, and demographic statistics [95]. They employ techniques such as econometric modeling, aggregate time series analysis, and, in some cases, non-intrusive load monitoring.

These methods are valued for their scalability and relatively modest data requirements, making them well suited for policy evaluation, regulatory design, and long-term infrastructure forecasting. For instance, they are commonly used to estimate the relationship between electricity demand and macroeconomic growth or to anticipate peak load requirements at the grid level [146].

However, their aggregated nature means they sacrifice granularity. They cannot capture the diversity of end-use behaviors or sudden shifts in demand patterns, which limits their reliability in contexts characterized by high heterogeneity or rapid technological and behavioral change. As a result, macrodata methodologies are best suited to broad planning and policy tasks, where overall demand trajectories are more relevant than detailed user-level profiles [9].

Prototypical Data Methods

Prototypical methods occupy a middle ground between the micro and macro perspectives. They are based on archetypal demand profiles or parameterized functions derived from representative datasets. These archetypes may be differentiated according to household size, building type, climate conditions, or socioeconomic characteristics [149].

By scaling these archetypes, planners can simulate demand in larger populations without requiring detailed microdata. They also provide a practical tool for scenario analysis, enabling exploration of what-if conditions such as retrofitting programs, changes in appliance ownership, or climate-driven shifts in demand [9].

The strength of this methodology lies in its pragmatic balance between feasibility and insight. However, its error level depends heavily on the representativeness of the archetypes and the validity of the underlying assumptions. Outdated or non-representative profiles can increase error, introducing bias and misalignment with real-world conditions. Consequently, prototypical methods are most frequently used in feasibility studies, energy planning, and preliminary assessments, where they provide a first-order approximation before more detailed data can be collected.

Hybrid Data Methods

Hybrid methods combine elements of microdata, macrodata, and prototypical methodologies. A common example is the calibration of archetypal demand profiles with sensor or smart meter data, or the integration of bottom-up simulations with macroeconomic trends to ensure system-level consistency [42, 34].

The appeal of hybrid methods lies in their ability to harness multiple data layers, thereby reducing error while improving robustness and adaptability. For instance, they can capture household-level behavior while maintaining alignment with national or regional demand trajectories. This makes them particularly relevant for modern applications such as smart grids, renewable integration, urban energy system planning, and digital twins of PEDs [37].

At the same time, hybrid methodologies are complex to design and operate. They require handling heterogeneous datasets, developing calibration procedures, and managing high computational loads. Errors in one component may propagate across the system, making careful methodological design essential. Despite these challenges, hybrid methods are increasingly recognized as the most promising avenue for bridging the gap between high-resolution modeling and large-scale planning [37].

2.2 Forecasting Methods

The classification frameworks presented above provide a structured foundation for understanding electricity demand forecasting across multiple dimensions, including temporal scope, system conceptualization, application domain, and data requirements. Together, they clarify why different forecasting approaches are used and under what conditions they are most effective. Building on this conceptual groundwork, this section shifts the focus from classification to methods, examining the main families of forecasting approaches and their underlying analytical principles.

These approaches differ in their conceptual foundations, data requirements, and analytical objectives, reflecting diverse ways of representing and predicting consumption behavior. The following sections present a structured overview of the main forecasting methodologies commonly applied in energy research, including time series, causal and econometric, statistical and probabilistic, end-use, GIS-based, AB, structural, scenario-based, and AI-based approaches, encompassing ML.

2.2.1 Time Series Methods

Time series methods constitute the most classical and widely applied family of electricity demand forecasting approaches. Their central assumption is that future demand can be forecast by extrapolating regularities observed in past data. In other words, demand is treated not as random but as a temporally structured process shaped by recurring social routines, climatic cycles, and long-term growth trends. When external conditions remain relatively stable, the past becomes a statistically reliable proxy for the future. [56]

A time series is defined as a sequence of observations y_t recorded at regular intervals of time $t = 1, 2, \ldots, T$. The forecasting task consists in estimating future values $\hat{y}_{t+h}$ for a given horizon h using past information. Time series analysis decomposes the observed data into three main components:

- Trend: The long-term evolution of the series, reflecting underlying drivers such as demographic growth, economic expansion, or technological diffusion.
- Seasonality: Recurrent patterns within fixed periods, such as daily, weekly, or annual cycles linked to routines or climatic variations.
- Irregular component: Random shocks and residual noise caused by exceptional events or unobserved factors.

The objective of time series forecasting is to identify and model these components to produce low-error predictions. Among the classical approaches, the ARIMA family remains one of the most established and effective methods [56].

The ARIMA model, denoted as ARIMA(p, d, q), combines autoregressive, differencing, and moving average processes, as expressed in Equation 2.1:

$$\phi_p(B)(1 - B)^d y_t = \theta_q(B)\epsilon_t, \tag{2.1}$$

where B is the backward shift operator, $\phi_p(B)$ and $\theta_q(B)$ are polynomials of order p and q, respectively, d is the degree of differencing applied to achieve stationarity, and ϵ_t is a white-noise error term. The autoregressive component, represented by $\phi_p(B)y_t$, captures dependence on previous demand values, while the moving average component, represented by $\theta_q(B)\epsilon_t$, accounts for the influence of past errors. Differencing $(1 - B)^d$ ensures stationarity by removing long-term trends.

To handle periodic fluctuations such as daily or seasonal cycles, the Seasonal ARIMA (SARIMA) extension includes an additional set of seasonal parameters $(P, D, Q)_s$, where s denotes the periodicity. When exogenous variables such as temperature, rainfall, or socioeconomic indicators are available, the ARIMA with exogenous regressors (ARIMAX) variant reduces forecasting error by incorporating these causal drivers [146].

Time series methods are applicable across all forecasting horizons. In short-term forecasting (STF), ranging from minutes to a week, they support grid operation, balancing, and tariff optimization. In the medium term (MTF), they guide maintenance planning and resource allocation. In the long term (LTF), they inform strategic decisions on capacity expansion, infrastructure investments, and policy design [126].

Their persistence in the energy domain stems from several strengths. Time series models allow planners to trace forecast outcomes back to historical patterns, a feature often missing in black-box ML approaches. They are computationally efficient, easy to implement with standard statistical tools, and perform remarkably well in systems with stable periodicity and consistent data records.

However, time series methods also face clear limitations, as their reliability depends heavily on the availability of long, high-quality, and stationary datasets. They struggle with abrupt structural changes such as rapid appliance adoption or demographic shifts, which invalidate past regularities. Even with extensions like ARIMAX, integrating diverse and complex datasets remains challenging compared to modern hybrid or ML-based methodologies.

Despite these limitations, time series methods continue to serve as essential baselines and benchmarks. In data-constrained contexts, especially in isolated or off-grid MGs, ARIMA and its seasonal variants are still the preferred choice due to their simplicity and modest computational requirements. They provide reliable short-term forecasts, ensure operational stability, and form a foundation for developing more advanced hybrid or data-augmented approaches [37].

2.2.2 Causal and Econometric Methods

Causal methods form a major family of forecasting methodologies in which electricity demand is explained as the result of identifiable external drivers rather than being inferred solely from its own past behavior. Electricity demand is influenced by measurable factors such as weather conditions, demographic trends, economic activity, technological change, and policy interventions. By explicitly linking demand to these drivers, causal methods generate not only forecasts but also explanatory insights into the mechanisms behind demand patterns [29].

Whereas time series techniques assume continuity of past patterns, causal methods remain valuable even when structural changes occur, since the relationships between demand and its determinants can evolve alongside new conditions. For this reason, they are particularly relevant for long-term forecasting, scenario analysis, and policy evaluation, where understanding the “why” behind demand is as important as anticipating its future trajectory [9].

The general form of a causal model can be expressed as shown in Equation 2.2:

$$D_t = \alpha + \sum_{i=1}^{n} \beta_i X_{i,t} + \epsilon_t, \tag{2.2}$$

where D_t is electricity demand at time t, $X_{i,t}$ are exogenous variables, β_i are coefficients capturing their influence, α is the intercept, and ϵ_t is a stochastic error term.

This formulation highlights the central principle of causal modeling: demand is explained as a function of identifiable external factors. For instance, temperature drives the use of heating and cooling systems, GDP growth affects industrial demand, and demographic changes shape household electricity use.

Several modeling strategies fall under the umbrella of causal methodologies. Econometric methods such as linear regression, generalized linear models, cointegration, and panel-data approaches link electricity demand to macroeconomic and demographic factors, often across multiple regions or sectors.

The strengths of causal methods are manifold. They move beyond pattern recognition, uncovering the underlying mechanisms that generate electricity demand. This explanatory power makes them particularly suitable for long-term forecasting, where structural dynamics such as economic growth, demographic change, or appliance adoption exert greater influence than short-term fluctuations. They are robust tools for scenario analysis, enabling sensitivity testing of alternative socioeconomic pathways or policy interventions, and they are directly applicable to policy evaluation, such as assessing efficiency standards, tariff reforms, or demand-response programs. Because they rely on explicit relationships between demand and its drivers, causal methods are especially relevant in decision-making contexts where justifications are as important as low forecasting error [96].

At the same time, causal methods face significant challenges. They require extensive and reliable exogenous data. Relationships between drivers are often nonlinear and context-specific, complicating model specification, calibration, and validation. Moreover, the exogenous variables on which these methods rely, such as GDP growth, demographic evolution, or appliance adoption, are themselves uncertain, and these uncertainties propagate into demand forecasts. Robust causal modeling therefore requires large datasets, statistical expertise, and interdisciplinary knowledge, resources that may be scarce in practice [96].

Despite these limitations, causal methods remain widely used in electricity demand forecasting. Utilities employ weather-sensitive regression models to anticipate peaks during extreme climatic conditions, while policymakers use econometric techniques to evaluate the impacts of efficiency regulations, tariff reforms, or electrification programs. Long-term energy system models, such as those developed by the International Energy Agency or national planning authorities, routinely rely on causal forecasts as inputs for transition pathways.

Where historical load data are sparse or fragmented, causal methodologies are particularly valuable because they can exploit contextual information from socioeconomic surveys, weather records, and appliance ownership statistics. In such settings, causal methods offer a realistic basis for forecasting, outperforming purely statistical extrapolations and delivering actionable insights for infrastructure design and electrification planning [21].

Causal methods occupy a unique position in electricity demand forecasting. Their explanatory nature bridges technical modeling and policy analysis, enabling forecasts that exhibit low error while remaining policy-relevant. However, their reliance on exogenous data and their sensitivity to specification errors limit their effectiveness in poorly monitored or highly dynamic environments. Future advances are likely to emerge from combining causal reasoning with ML. Such hybrid methodologies could provide low-error forecasts tailored to contexts ranging from national energy planning to MG electrification. Although causal methods reveal the mechanisms underlying demand, they often remain deterministic and strongly dependent on assumptions about external drivers. In practice, electricity demand also exhibits randomness and variability that cannot be fully explained by deterministic relationships [150]. This has motivated the use of statistical methods, which combine explanatory power with probabilistic reasoning and enable forecasts that explicitly account for uncertainty.

2.2.3 Statistical and Probabilistic Methods

Statistical methods represent one of the most established traditions in electricity demand forecasting, grounded in the formal application of statistical inference to model demand behavior. Their defining feature is the explicit representation of the relationship between electricity demand and explanatory variables through mathematical formulations that can be rigorously estimated and tested. Unlike purely historical methodologies, which rely mainly on autocorrelation and time-series structures, statistical methods provide a framework for quantifying how demand responds to multiple influencing factors while also generating measures of uncertainty around forecasts [146].

These methodologies are valued for their methodological rigor. They enable researchers and practitioners to test hypotheses, assess the relevance of predictors, and construct confidence intervals, capabilities that are particularly valuable for risk management, regulatory design, and long-term system planning. At the same time, their effectiveness depends heavily on the validity of assumptions such as linearity, stationarity, and homoscedasticity, as well as on the availability of sufficiently large and representative datasets.

A general linear regression model can be expressed as shown in Equation 2.3:

$$D_t = \beta_0 + \sum_{i=1}^{n} \beta_i x_{it} + \epsilon_t, \tag{2.3}$$

where β_0 is the intercept, D_t is demand at time t, x_{it} are explanatory variables, β_i are coefficients, and ϵ_t is the error term, typically assumed to follow a normal distribution with zero mean and constant variance.

Extensions of this framework include multiple regression, polynomial regression, and generalized linear methods, which relax the assumption of normality and allow nonlinear relationships. Panel-data methods extend the analysis across regions or sectors, capturing both temporal and cross-sectional variation. These methods are straightforward to estimate and apply, which explains their enduring role in demand analysis [146].

While point forecasts are useful, they fail to capture the inherent uncertainty of electricity demand. Probabilistic forecasting addresses this limitation by estimating the full conditional distribution of demand given explanatory variables, as shown in Equation 2.4:

$$P(y_t \mid x_{1t}, x_{2t}, \ldots, x_{nt}), \tag{2.4}$$

where $P(\cdot)$ denotes the probability distribution, y_t is the electricity demand at time t, $\mid$ indicates conditioning (that is, "given"), and $x_{1t}, x_{2t}, \ldots, x_{nt}$ represent the explanatory variables at time t, such as weather conditions, socioeconomic indicators, or temporal features. This formulation provides forecast intervals or probability distributions instead of single-point estimates, offering a more informative representation of prediction variability. Such an approach is particularly relevant in modern electricity systems, where volatility is amplified by high shares of renewable generation and increasingly dynamic consumption patterns.

Several probabilistic methods are commonly applied. Quantile regression estimates conditional quantiles, producing intervals that explicitly represent uncertainty [98]. Bayesian inference integrates parameter uncertainty directly into forecasts, yielding posterior distributions of demand [109]. Bootstrapping constructs empirical distributions by repeatedly resampling historical data, offering a non-parametric way to represent uncertainty without restrictive assumptions. Although methodologically diverse, these approaches share the goal of enabling risk-aware planning by providing more complete information about possible futures.

Statistical and probabilistic methods are applied across all forecasting horizons. In

STF, regression models enriched with weather variables such as temperature and humidity are widely used to capture hourly or daily fluctuations in demand, especially in cities where heating and cooling dominate variability. In the MTF, probabilistic forecasts provide forecast intervals that help operators plan procurement, maintenance, and reserve strategies while explicitly accounting for seasonal uncertainty. At the LTF, regression-based methodologies are used to uncover structural relationships between demand and macroeconomic drivers such as GDP, population growth, or industrial activity [68]. For example, temperature-sensitive regressions have been successfully used to capture summer cooling peaks in urban grids, while Bayesian probabilistic forecasts help system operators anticipate variability and prepare for extreme demand events. This versatility makes statistical methods indispensable for both operational decision-making and strategic planning.

The appeal of statistical methods lies in their combination of rigor and transparency. Coefficients provide direct insights into the influence of explanatory variables such as weather or income on demand, making results clear and understandable for researchers and policymakers.

Probabilistic extensions go further by explicitly quantifying uncertainty, supporting risk-aware strategies in systems where reliability is critical. Their solid theoretical foundation also provides robust tools for validation, hypothesis testing, and uncertainty quantification.

Nevertheless, important limitations remain. Many statistical methods rest on restrictive assumptions, linearity, stationarity, or homoscedasticity, that are frequently violated in practice. Electricity demand is often nonlinear, highly variable, and subject to structural change, making simple regression-based methodologies insufficient. Parameter estimation typically requires large and high-quality datasets, which are often unavailable. Statistical methods offer limited flexibility in capturing nonlinear interactions or sudden structural changes unless augmented with additional techniques. As a result, while statistical methodologies remain foundational, they are increasingly complemented by more flexible ML or hybrid approaches [10].

Statistical methods remain a cornerstone of electricity demand forecasting because of their transparency and capacity to quantify uncertainty. They are particularly valuable for long-term planning, where technology adoption, behavioral change, and efficiency improvements drive demand dynamics. However, their reliance on strong assumptions and large datasets constrains their applicability in data-constrained contexts.

Statistical methods incorporate explanatory variables and can quantify uncertainty through probabilistic forecasts, yet they typically assume that the underlying relationships remain stable over time. In practice, electricity demand is subject to irregular fluctuations driven by weather conditions, consumer behavior, and unforeseen events, which introduce additional layers of uncertainty beyond what traditional statistical frameworks capture.

Despite their analytical rigor, statistical and probabilistic methods describe electricity demand mainly at an aggregate level, offering limited insight into the mechanisms driving consumption within individual sectors or technologies. End-use approaches overcome this limitation by explicitly modeling the sources of demand, linking consumption to specific appliances, activities, or processes. This disaggregated perspective enables a deeper exploration of how changes in technology, behavior, or policy influence overall demand.

2.2.4 End-use Methods

Rather than treating electricity demand as a monolithic variable, these methods disaggregate it into components such as residential appliances, industrial machinery, commercial equipment, and public infrastructure. This disaggregation provides detailed insights into the drivers of electricity demand and allows for fine-grained analysis of how changes in technology, behavior, or policy affect demand patterns [145].

The central idea behind end-use methods is that total demand D_t at time t can be represented as the sum of demand from individual end-use devices or categories, as shown in Equation 2.5:

$$D_t = \sum_{i=1}^{N} N_i \cdot P_i \cdot U_{it}, \tag{2.5}$$

where:

- N_i is the number of devices of type i,
- P_i is the rated power or average load of device i,
- U_{it} is the utilization factor or load factor of device i at time t,
- N is the total number of end-use categories considered.

This formulation allows the analyst to link demand directly to the stock, efficiency, and usage patterns of specific devices. For instance, the adoption of efficient lighting technologies or smart appliances can be explicitly incorporated into forecasts [21].

End-use methods represent a BUM approach to electricity demand forecasting, decomposing total demand into contributions from specific end-use categories and sectors. The three most common categories are residential, commercial, and industrial users. Residential demand encompasses appliances such as refrigerators, washing machines, air conditioners, lighting, and entertainment devices, with forecasts influenced by household demographics, appliance ownership rates, and behavioral patterns. Commercial demand reflects demand in offices, retail spaces, and public buildings, where equipment such as Heating, Ventilation, and Air Conditioning (HVAC) systems, lighting, and computing devices dominate; usage patterns are largely shaped by building function, occupancy, and efficiency standards. Industrial demand, on the other hand, is primarily driven by production machinery, motors, and process heating, with forecasting linked to industrial activity levels, production schedules, and the pace of technology adoption [32]. By disaggregating demand into these categories, end-use methods provide sector-specific insights that are particularly valuable for targeted energy efficiency programs and DSM strategies.

The primary strength of end-use methodologies lies in their ability to support medium- and long-term forecasting, where structural drivers of demand are especially relevant. A central application is policy analysis, as end-use methods are frequently used to evaluate the effects of efficiency standards, appliance labeling programs, and subsidies for energy-efficient technologies. They are also instrumental in scenario development, allowing analysts to explore alternative demand trajectories under different assumptions regarding electrification, economic growth, or the uptake of emerging technologies such as electric vehicles and heat pumps. In addition, utilities and grid operators use end-use methods for infrastructure planning, as they provide detailed sector-level projections that inform grid expansion, reinforcement, and capacity allocation decisions. For instance, end-use methods have been employed to estimate the long-term impact of widespread electric vehicle adoption on household electricity demand patterns, while in the commercial sector they have been used to assess the peak-load reduction potential of efficiency programs targeting HVAC systems and lighting [145]. These applications highlight the role of end-use methodologies as a bridge between technical detail and policy relevance.

End-use methods offer several distinct advantages. They provide detailed insights by breaking down demand into its constituent drivers, thereby enabling targeted interventions at the device or sectoral level. Their alignment with policy objectives makes them particularly suitable for evaluating efficiency programs, technology adoption pathways, and DSM strategies. Moreover, they are flexible frameworks that can explicitly integrate the effects of new technologies, behavioral changes, or regulatory measures into forecasts. Despite these benefits, end-use methodologies also face important limitations. They require highly detailed device-level or sector-specific data. Their complexity can also be a drawback, as modeling multiple devices and sectors introduces numerous parameters that may increase model uncertainty if not supported by sufficient data. While well suited for capturing long-term structural dynamics, they are less effective for STF, where rapid demand fluctuations are better addressed by time-series or ML-based methods. As such, end-use methods are best regarded as instruments for strategic planning and policy evaluation rather than for real-time demand forecasting [21].

End-use methods provide a rich and policy-relevant framework for electricity demand forecasting. Their disaggregated structure makes them particularly valuable in long-term planning, where technology adoption, behavioral change, and efficiency improvements shape demand dynamics. However, the need for detailed device-level data and the complexity of model calibration remain major challenges, particularly in regions with limited data availability.

To overcome these limitations, hybrid methodologies that combine end-use disaggregation with statistical or AI-based approaches are attracting increasing attention. Such integrations can harness the clarity of end-use models together with the low forecasting error typical of data-driven techniques, offering a promising pathway toward more comprehensive and robust electricity demand forecasting [33].

2.2.5 Geographic Information System-based Methods

GIS-based methods have emerged as powerful tools for electricity demand forecasting by integrating spatial data with traditional forecasting techniques. These methodologies leverage geographic variables such as population density, land use, infrastructure distribution, and socioeconomic indicators to model spatial variations in energy demand. Unlike purely temporal or sectoral methodologies, GIS-based methods explicitly account for the geographic dimension of electricity demand, making them particularly valuable for regional and urban energy planning [140].

The fundamental principle of GIS-based methods is to relate electricity demand $D(s,t)$ at a given spatial location s and time t to geographic, demographic, and infrastructural factors, as formulated in Equation 2.6:

$$D(s,t) = f\left(P(s), L(s), I(s), E(s,t)\right) + \epsilon(s,t), \tag{2.6}$$

where:

- $P(s)$ represents population density at location s,
- $L(s)$ denotes land-use characteristics (e.g., residential, industrial, agricultural),
- $I(s)$ includes infrastructure variables such as road networks, transmission lines, or building footprints,
- $E(s,t)$ refers to temporal or environmental variables such as temperature, season, or income,
- $\epsilon(s,t)$ is a stochastic error term capturing unobserved variability.

GIS tools enable the integration and visualization of these heterogeneous datasets, producing spatially explicit forecasts that highlight geographic demand patterns.

A key feature of GIS-based methods is their ability to disaggregate total demand into finer spatial units. For instance, regional or national demand can be broken down into demand at the level of neighborhoods, municipalities, or grid cells.

This is achieved through spatial interpolation techniques, such as inverse distance weighting or kriging, which estimate demand at unobserved locations based on nearby observations [91].

By mapping demand forecasts onto geographic layers, planners can identify demand hot spots, evaluate grid adequacy, and prioritize investments in remote areas.

GIS have emerged as powerful tools for electricity demand forecasting, particularly in contexts where spatial heterogeneity plays a central role. By linking electricity demand to land-use categories, such as residential, commercial, or industrial zones, GIS-based methods have proven especially valuable in urban planning, where they support the projection of load growth in rapidly expanding cities. In remote areas, these methods are widely applied to estimate demand in off-grid or underserved communities by combining data on population density, building footprints, and socioeconomic indicators [90]. GIS methodologies are also instrumental in assessing the spatial alignment between demand and renewable energy resources, such as solar irradiation or wind availability, thereby informing strategies for integrating variable renewable generation into local energy systems. Additionally, GIS-based demand projections are crucial for infrastructure planning, supporting the siting of substations, transmission lines, and distributed generation units in alignment with anticipated demand growth. For instance, in Sub-Saharan Africa, GIS-based methods have been deployed to plan MGs in remote areas by integrating census information, satellite imagery, and geolocated household surveys, providing reliable forecasts of electricity needs in remote regions [140, 90]. These applications demonstrate how GIS techniques can bridge spatial data with energy system modeling, yielding insights that are both technically robust and highly relevant to policy and planning.

The strengths of GIS-based methods lie primarily in their spatial explicitness. Unlike aggregate demand methods, they provide detailed geographic insights into the distribution of demand, enabling location-specific planning that is particularly relevant for infrastructure development and electrification strategies in remote areas [90]. A second advantage is their capacity to integrate heterogeneous data sources within a unified analytical framework, combining demographic, infrastructural, environmental, and economic variables to generate rich, multidimensional forecasts. Their policy relevance further enhances their value, as they directly support the design of targeted interventions, including electrification programs in remote areas, smart city initiatives, and renewable energy deployment strategies. Despite these advantages, GIS-based methodologies face several important limitations. They are highly data-dependent, relying on the availability, resolution, and low error of spatial datasets, which may be outdated, incomplete, or inaccessible in certain regions. Their computational complexity is another challenge, since processing and analyzing large-scale spatial data requires advanced GIS tools and substantial computational resources. Scalability also remains a challenge: while highly effective at regional or local levels, extending GIS-based forecasts to national or continental scales requires harmonized datasets that are seldom available. These limitations suggest that while GIS-based methodologies are indispensable for spatially explicit demand forecasting, their successful implementation depends heavily on data availability, technical expertise, and computational capacity [91, 90].

GIS-based methods bring an essential spatial perspective to electricity demand forecasting, making them highly relevant for regional planning, infrastructure development, and electrification in remote areas. Their ability to integrate heterogeneous data sources and generate spatially explicit demand maps provides unique insights not captured by purely temporal or statistical methods. Nonetheless, their reliance on high-quality spatial datasets remains a critical bottleneck, particularly in developing regions with limited data availability. Future research should focus on integrating GIS-based methodologies with remote sensing, ML, and hybrid approaches to address data limitations and minimize forecasting error. Such integration would enable comprehensive, scalable, and actionable electricity demand forecasts that better align with the spatial realities of modern power systems [91].

2.2.6 Agent Based Methods

AB methods are computational simulation methodologies that represent electricity systems as the result of interactions among autonomous entities, or agents, such as households, businesses, utilities, and policymakers. Each agent is characterized by behavioral rules, decision-making processes, and interaction patterns that collectively generate emergent demand dynamics at the system level. AB methods are particularly well-suited for exploring heterogeneity in electricity demand, capturing complex behaviors, and analyzing the impacts of interventions such as demand-response programs or policy incentives.

The core idea of AB methods is to represent electricity demand D_t at time t as the aggregation of individual agents' demand decisions, as formulated in Equation 2.7:

$$D_t = \sum_{i=1}^{N} d_i(t), \tag{2.7}$$

where:

- $d_i(t)$ is the demand of agent i at time t,
- N is the total number of agents,
- each $d_i(t)$ is determined by internal states (e.g., preferences, income, device ownership) and external stimuli (e.g., prices, weather, peer influence).

Agents interact with one another and with their environment, which may include market signals, policy instruments, and infrastructure constraints. The resulting system-level demand is not explicitly programmed but emerges from these micro-level interactions [135].

A defining strength of AB methods is their capacity to represent heterogeneous agents and the diverse behaviors that shape electricity demand. In residential contexts, households may differ in size, appliance ownership, income levels, and willingness to adopt new technologies, producing highly variable demand patterns. Businesses, by contrast, exhibit distinct operational hours, load profiles, and sensitivities to electricity prices, while utilities and grid operators actively influence demand through pricing schemes, demand-response programs, and investment strategies. Behavioral rules governing these agents can range from simple deterministic “if–then” statements to stochastic decision-making processes or more sophisticated algorithms inspired by psychological or economic theory. This flexibility allows AB methods to capture both aggregate demand trends and the diversity and complexity of behaviors within populations, providing a richer representation of real-world demand dynamics than many traditional methodologies [159].

AB methods have been applied across a wide range of electricity demand forecasting and energy planning challenges, particularly where heterogeneity and interaction among agents are central. In demand-response analysis, for instance, AB methods simulate how households with varying socioeconomic characteristics respond to time-of-use tariffs or incentive-based programs, revealing uneven participation across different population segments. They are also widely used to explore technology adoption pathways, such as the diffusion of electric vehicles, rooftop solar panels, or smart appliances, where social influence and peer effects play an important role. At the policy level, AB methods provide a platform for evaluating the effectiveness of interventions such as subsidies, awareness campaigns, or regulatory reforms, helping to identify which groups are most likely to respond. AB methods have been applied in resilience studies, where they model how demand evolves under disruptions caused by blackouts, extreme weather, or economic shocks, highlighting emergent vulnerabilities and adaptive behaviors. For example, simulations of household load shifting under real-time pricing demonstrate how differences in income, appliance ownership, and behavioral preferences produce unequal responses to the same policy, with important implications for equity and system performance [135].

AB methods bring several distinctive advantages to electricity demand forecasting. Their ability to explicitly represent heterogeneity allows for detailed modeling of diverse demand patterns and behavioral responses across households, businesses, and utilities. They are also uniquely suited to capturing emergent system-level dynamics arising from micro-level interactions, which traditional TDM often overlook. Moreover, AB methods serve as valuable tools for policy experimentation, enabling the testing of "what-if" scenarios in silico before real-world implementation. However, these benefits are balanced by important limitations. AB methods are highly data-intensive, requiring detailed micro-level information on household profiles, appliance usage, and behavioral tendencies that may not be readily available. They also depend on extensive assumptions about agent behavior, which can introduce subjectivity, bias, or oversimplification if not carefully validated. Large-scale AB models with many interacting agents are computationally demanding and may face scalability challenges, particularly when integrated into broader energy system models. These strengths and limitations highlight both the promise and complexity of AB methods as tools for capturing behavioral dynamics in electricity demand forecasting [159].

AB methods represent a powerful methodology for electricity demand forecasting, offering unique insights into the role of heterogeneity, behavior, and interaction in shaping energy demand. Their ability to model emergent dynamics makes them invaluable for exploring complex systems and assessing policy impacts under realistic social and economic conditions. However, the challenges of data availability, model validation, and computational scalability limit their widespread use [135].

Future research should prioritize hybrid approaches that combine AB methods with statistical or AI-based models, achieving forecasts that are both behaviorally rich and low error. Such integration would harness the explanatory power of AB methods while ensuring practical applicability in real-world forecasting and planning contexts.

2.2.7 Structural Methods

Structural methods represent electricity demand using engineering-based formulations grounded in the physical and technical characteristics of energy systems. Unlike purely statistical or behavioral methodologies, these methods rely on fundamental engineering principles, such as thermodynamics, electrical load equations, and equipment specifications, to simulate energy demand patterns. Structural methods are especially relevant for technical studies where the focus is on the influence of physical infrastructure, building characteristics, or appliance performance on electricity demand.

The central principle of structural methods is to estimate electricity demand $D(t)$ at time t as a function of physical system parameters and operating conditions, as expressed in Equation (2.8):

$$D(t) = \sum_{k=1}^{K} P_k \cdot u_k(t), \tag{2.8}$$

where:

- P_k is the rated power demand of device or system k,
- $u_k(t)$ represents the utilization factor or operating schedule of k at time t,
- K is the total number of devices or subsystems considered.

In building energy modeling, for instance, $u_k(t)$ may be derived from occupancy schedules, climate conditions, or control strategies (e.g., thermostat setpoints). Thus, demand forecasts emerge directly from the technical operation of physical components rather than empirical correlations [9].

Structural methods are grounded in engineering principles and rely on detailed technical specifications to ensure that electricity demand forecasts are physically consistent with the capabilities of the systems being modeled. These methods incorporate building physics, including heat transfer equations, thermal inertia of walls, and the performance characteristics of HVAC systems. They also consider the electrical properties of systems, such as load flow calculations, device efficiency curves, and operational constraints, as well as appliance-level performance, using manufacturer data on rated capacity, standby power, and duty cycles. By integrating these elements, structural methods provide a physically coherent representation of demand dynamics. For example, the electricity demand of an HVAC system can be represented as shown in Equation (2.9):

$$P_{\mathrm{HVAC}}(t) = \frac{Q_{\mathrm{cooling}}(t)}{\mathrm{COP}}, \tag{2.9}$$

where $Q_{\mathrm{cooling}}(t)$ is the cooling load at time t and COP is the coefficient of performance. Such formulations ensure that forecasts exhibit low error and remain grounded in the operational realities of the equipment and infrastructure being modeled.

Structural methods are widely applied in contexts that demand detailed technical analysis. In building energy simulation, for example, they are used to estimate hourly or daily load profiles based on architectural design, material properties, and HVAC system configurations. Appliance-level analyses leverage structural methods to assess the impact of new efficiency standards or the introduction of novel devices on aggregate electricity demand. In the context of renewable energy integration, structural methods help evaluate how physical constraints of equipment affect the alignment of demand with variable generation sources, such as PV panels or wind. These methods are also invaluable for scenario analysis, enabling planners to test the effects of retrofits, electrification strategies, or equipment replacements on the evolution of demand. Widely used tools such as EnergyPlus and TRNSYS illustrate the practical application of these methods, forecasting building-level load curves under diverse climate and occupancy scenarios [92].

The primary strength of structural methods lies in their technical realism. Because they are grounded in engineering principles and actual system specifications, the resulting forecasts are physically consistent and can provide high-fidelity insights into energy demand patterns. They also offer scenario flexibility, allowing "what-if" analyses of equipment upgrades, retrofits, or other interventions. Moreover, their component-level granularity enables a detailed understanding of the contribution of individual devices or subsystems to overall demand. However, these advantages are balanced by important limitations. Structural methods often omit socioeconomic and behavioral factors, such as consumer preferences, adaptive behaviors, or usage patterns, which can significantly influence demand in real-world settings. They also require extensive technical data and operating schedules, which may be unavailable or costly to obtain. Large-scale structural methods can be computationally intensive, making calibration against observed demand data challenging. These constraints highlight the trade-off inherent in structural modeling: while they provide highly realistic, engineering-based insights, their scope may be limited when human behavior and socioeconomic factors play a dominant role in shaping electricity demand.

Structural methods provide a rigorous, engineering-based lens for electricity demand forecasting, making them indispensable for technical studies and infrastructure planning. Their ability to capture system-level impacts of physical changes (e.g., efficiency measures, equipment adoption) distinguishes them from statistical or behavioral methods. However, by focusing primarily on physical and technical characteristics, they risk overlooking human and socioeconomic drivers of demand [92].

A promising research direction lies in hybrid methodologies that integrate structural methods with behavioral, econometric, or AB methods. Such combinations would enable forecasts that are both technically sound and reflective of real-world decision-making, thereby supporting holistic energy planning.

2.2.8 Scenario-Based Methods

Scenario-based methods are designed to evaluate electricity demand under alternative possible futures, rather than producing a single deterministic forecast. These methods explicitly account for uncertainties in exogenous factors such as economic growth, population change, technology adoption, policy interventions, and climate variability. By generating multiple plausible demand trajectories, scenario-based methodologies support strategic planning, policy assessment, and long-term decision-making.

The fundamental principle is to represent the future as a set of scenarios $S_1, S_2, \ldots, S_n$, each characterized by a specific combination of drivers affecting electricity demand, as summarized in Equation (2.10):

$$D_i(t) = f(X_i(t)), \quad i = 1, \ldots, n \tag{2.10}$$

where:

- $D_i(t)$ is the forecasted electricity demand at time t for scenario S_i,
- $X_i(t)$ represents the set of exogenous factors (economic, demographic, technological, climatic) defining scenario S_i,

- $f(\cdot)$ is the forecasting function, which may be based on statistical, econometric, structural, or hybrid methods.

Each scenario represents a coherent and internally consistent narrative about the future, allowing stakeholders to explore the potential impact of different assumptions or uncertainties on electricity demand [155].

Scenario-based methodologies in electricity demand forecasting provide a structured framework for exploring how future demand might evolve under different conditions. The process begins with identifying key drivers that significantly influence demand, including macroeconomic variables such as GDP growth, population dynamics, electrification rates, and policy measures like energy efficiency programs. Once these drivers are established, plausible ranges or discrete levels for each variable are defined, reflecting low, medium, or high growth assumptions, technological adoption rates, or regulatory intensity. These assumptions are then combined into internally consistent scenarios, ensuring logical coherence across drivers. For each scenario, appropriate forecasting methods are selected, ranging from econometric methods for macro-level drivers to structural or engineering-based methods for capturing technical constraints. The resulting scenarios are analyzed to compare demand trajectories, identify risks and opportunities, and evaluate strategies robust across diverse futures.

Scenarios are commonly classified into three categories. Exploratory scenarios investigate a broad spectrum of plausible futures without assigning probabilities, providing insight into potential variations in demand. Normative scenarios focus on desired or target outcomes, often linked to policy objectives, such as decarbonization or energy access goals. Forecasting scenarios incorporate probabilistic information to estimate the likelihood of different futures, supporting more quantitative risk assessments. By systematically exploring these scenarios, planners and decision-makers can better understand uncertainties, trade-offs, and potential vulnerabilities in the electricity system [145].

Scenario-based methods are widely applied in energy planning and forecasting, particularly where uncertainty is high and multiple interacting factors must be considered. They are valuable for evaluating energy policies, such as subsidies, regulatory reforms, or decarbonization targets, by projecting their impact on future demand. In infrastructure planning, scenario analysis helps assess the need for grid expansion, generation capacity, and storage under varying demand trajectories. They are also used in climate and environmental studies to explore the effects of temperature changes, extreme weather events, and renewable integration on electricity demand patterns.

In emerging markets and rural electrification projects, scenario methodologies support planning for uncertain adoption rates, technology penetration, and seasonal demand variations, providing guidance in contexts where historical data may be limited or unreliable.

Scenario-based methodologies offer several key advantages. They are highly flexible, capable of incorporating a wide range of drivers, uncertainties, and planning horizons. By generating multiple plausible futures, they provide decision support that enables the identification of risks, opportunities, and robust strategies under uncertainty. They are also compatible with a variety of underlying modeling techniques, including statistical, structural, AB, and hybrid methods, allowing them to integrate insights from diverse analytical perspectives [59].

However, these methodologies also face inherent limitations. Their usefulness depends heavily on the validity and realism of the underlying assumptions, and generating or analyzing multiple scenarios can be computationally and resource-intensive. Furthermore, exploratory scenarios do not provide probabilistic guidance, which can limit their precision in risk assessment and make it challenging to prioritize interventions when probabilities of different futures are needed. Despite these constraints, scenario-based methods remain essential tools for strategic planning and policy evaluation in electricity demand forecasting.

Scenario-based methodologies are essential for long-term energy planning and policy evaluation, particularly under conditions of high uncertainty or rapid system transformation. By providing multiple demand trajectories, they allow planners to explore robustness and resilience in infrastructure and operational strategies. Nevertheless, careful scenario design is critical: poorly defined scenarios can mislead decision-makers. Hybrid methodologies that combine scenario-based methods with structural, econometric, or AI-driven forecasting enhance reliability by capturing both physical constraints and socioeconomic drivers, while offering multiple plausible futures.

2.2.9 Artificial Intelligence and Machine Learning Methods

AI/ML have emerged as transformative tools in electricity demand forecasting, offering powerful alternatives and complements to classical methodologies. Unlike traditional statistical or econometric methods, which rely on predefined functional forms or strong parametric assumptions, AI/ML methods learn directly from data, enabling them to capture highly nonlinear relationships, complex temporal dependencies, and interactions among heterogeneous variables. This data-driven flexibility positions AI/ML at the forefront of modern forecasting research, particularly as electricity systems become increasingly dynamic due to decentralized generation, demand-side flexibility, and the integration of renewable resources.

The growing availability of HF smart meter data, detailed weather records, and socioeconomic indicators has accelerated the adoption of AI/ML techniques. These methods have been successfully applied across all forecasting horizons, from STF, critical for system operation and balancing markets, to MTF and LTF, where planning, infrastructure investment, and policy evaluation dominate. By leveraging large datasets and advanced architectures, AI/ML methodologies frequently achieve superior forecasting performance compared to classical benchmarks, particularly in contexts characterized by nonlinearities, irregular demand patterns, or high variability [1].

Artificial Neural Network Models

ANNs are computational architectures inspired by biological neurons, designed to capture nonlinear mappings between inputs and outputs. They consist of interconnected layers of nodes, where each node applies a weighted sum of its inputs followed by a nonlinear activation function. A simple feedforward ANN with one hidden layer can be expressed as shown in Equation (2.11):

$$D_t = f\left(\sum_{j=1}^{M} w_j \cdot g\left(\sum_{i=1}^{N} v_{ij} x_{it} + b_j\right) + c\right), \tag{2.11}$$

where x_{it} are the input features (e.g., past demand, temperature, socioeconomic indicators), N is the number of input variables, and M is the number of neurons in the hidden layer. The parameters v_{ij} and w_j represent the connection weights between layers, while b_j and c are bias terms. The function $g(\cdot)$ denotes the activation function applied at the hidden layer (e.g., sigmoid, Rectified Linear Unit (ReLU)), and $f(\cdot)$ represents the output function. Finally, D_t is the forecasted electricity demand at time t.

Extensions such as Recurrent Neural Networks (RNNs) and Long Short-Term Memorys (LSTMs) are particularly well-suited for sequential data, as they capture time dependencies and long-range correlations in load patterns. These properties make neural networks highly effective in STF, where they can model rapid fluctuations caused by weather variability and consumer behavior [115].

Support Vector Machines and Ensemble Methods

Support Vector Machines (SVMs) are a class of supervised learning algorithms used for both classification and regression tasks. They operate by identifying an optimal hyperplane that separates data points of different classes with the maximum possible margin. When applied to regression problems, this principle leads to the development of SVR, a variant specifically designed for continuous output forecasting.

The objective of SVR is to determine a regression function that forecasts electricity demand within a predefined error margin ϵ, while maintaining model simplicity, as shown in Equation (2.12):

$$f(x) = \langle w, x \rangle + b, \tag{2.12}$$

where $\langle w, x \rangle$ denotes the inner (dot) product between the weight vector w and the input vector x, and b is the bias term that shifts the regression hyperplane. The constraint $|D_i - f(x_i)| \leq \epsilon$ defines the ϵ-insensitive loss function, where D_i is the observed electricity demand at data point i, $f(x_i)$ is the forecasted demand, and ϵ specifies the tolerance margin within which prediction errors are not penalized. Slack variables are introduced to allow deviations beyond this margin, enabling the model to handle cases where the data are not perfectly linearly separable. Slack variables are introduced to allow deviations beyond this tolerance, ensuring that the model can handle cases where the data are not perfectly linearly separable. By employing kernel functions, such as radial basis or polynomial kernels, SVR projects data into higher-dimensional spaces, enabling it to effectively model nonlinear relationships in electricity demand [124].

Ensemble methods, particularly RFSs, represent another powerful ML methodology. RFSs aggregate the forecasts of multiple decision trees, each trained on bootstrapped samples and random feature subsets. Averaging forecasts (for regression) or majority voting (for classification) enhances robustness and mitigates overfitting. RFSs are especially effective when dealing with heterogeneous forecasters, such as weather, demographics, socioeconomic indicators, and temporal cycles, making them well suited for electricity demand forecasting in both urban and remote contexts [110].

Hybrid AI/ML Methodologies

Hybrid AI/ML methodologies integrate multiple techniques to exploit their complementary strengths. For instance, neural networks can be combined with SVR, where the former captures nonlinear demand patterns and the latter enhances generalization. Ensemble frameworks may also merge ANNs, RFSs, and econometric methods, aligning macroeconomic insights with micro-level nonlinear interactions. In MTF, hybrid methodologies frequently combine RFSs with time-series methods to integrate structural drivers and historical patterns. At long-term horizons, AI methods are frequently paired with econometric methodologies to capture macrotrends such as GDP, demographic growth, and policy effects, while still modeling nonlinearities that traditional techniques may overlook.

These hybrid strategies are applied across forecasting horizons. AI and ML techniques more broadly bring significant advantages to electricity demand estimation. They offer low forecasting error, capture complex nonlinearities, and integrate heterogeneous datasets ranging from weather records to socioeconomic indicators and smart meter readings. However, challenges remain. Deep learning architectures often function as “black boxes,” limiting transparency in decision-making. Training requires substantial computational resources and high-quality datasets, while overfitting remains a persistent risk in data-scarce contexts [28].

2.3 Forecasting Methods in Microgrids

Electricity demand forecasting has become a rapidly evolving research area, particularly in data-constrained contexts. Achieving low forecasting error supports efficient energy management by optimizing generation scheduling, enhancing grid stability, reducing operational costs, and improving resilience. With the growing penetration of renewable energy sources such as PV and wind, forecasting methods must also address the variability and uncertainty inherent in distributed, bidirectional energy systems.

As highlighted by [85], the transition from traditional unidirectional grids to intelligent, interconnected MGs equipped with IoT devices, smart meters, and real-time analytics has significantly increased system complexity. Classical statistical approaches, while effective under stable conditions, often fail to capture nonlinear, nonstationary, or context-dependent behaviors. Consequently, recent developments have embraced AI-driven paradigms capable of leveraging multidimensional data, including meteorological, socioeconomic, and behavioral variables.

Literature distinguishes three major categories of forecasting approaches [80]:

- **Statistical models**, such as ARIMA, SARIMA, or exponential smoothing, assume stationarity and are effective for STF under stable conditions. However, they struggle with high volatility, nonlinear dependencies, and exogenous influences that characterize renewable-based MGs.
- **ML models**, including SVR, RFS, decision trees, and ensemble methods, learn nonlinear patterns directly from data without requiring explicit functional assumptions. These techniques outperform traditional models when rich datasets are available, particularly when exogenous variables such as temperature, humidity, and socioeconomic indicators are included.
- **Deep learning and hybrid models**, encompassing architectures such as

LSTM, Convolutional Neural Network (CNN), and hybrid frameworks, integrate AI with statistical or physics-informed methods. They capture complex temporal and spatial dependencies, enabling probabilistic forecasting and improving adaptability under variable conditions.

Table 2.1 summarizes how these three categories address key challenges in MG demand forecasting.

Table 2.1: Coverage of major forecasting approaches with respect to key challenges in MG demand forecasting.

Forecasting Approach	**Data Scarcity & Fragmentation**	**Nonlinear Dependencies**	**Nonstationary Evolution**
Statistical models	×	×	×
Machine Learning models	△	✓	×
Deep Learning and hybrid models	×	✓	△

Legend: ✓ = explicitly addressed; △ = partially addressed; × = not addressed.

Statistical models are effective only under stationary and well-monitored conditions. They generally fail to manage scarce or fragmented datasets or capture nonlinear dependencies among socio-technical and environmental factors [96].

ML methods, including Support Vector Regression, Random Forests, and Gradient Boosting, improve performance by learning nonlinear relationships directly from data, yet they remain sensitive to data quality and availability and generally lack mechanisms to adapt to nonstationary demand evolution, limiting their robustness in remote or newly electrified contexts [49].

Deep learning and hybrid models, including LSTM, CNN, and physics-informed or statistical–AI hybrid architectures, effectively capture complex nonlinear dependencies and temporal–spatial dynamics, achieving low forecasting errors under variable conditions. However, they remain heavily dependent on large, high-frequency datasets, making them unsuitable for data-scarce or fragmented environments, and only partially address nonstationary demand evolution due to their limited capacity for continual or online adaptation, while their computational burden further constrains real-time or edge deployment [143].

The analysis of the threefold challenge of data scarcity and fragmentation, nonlinear behavioral and environmental dependencies, and nonstationary demand evolution reveals a persistent methodological gap in current forecasting research. Existing approaches tend to address these dimensions separately, overlooking their interaction and the compounded uncertainty they generate in data-constrained contexts. Conversely, deep learning architectures achieve strong forecasting performance but depend on large, stationary datasets and often lack adaptability. As a result, no framework currently reconciles causality, flexibility and adaptivity [57].

By recognizing this limitation, a critical hypothesis is formulated: the efficiency and robustness of learning systems can be enhanced not merely through larger datasets or deeper architectures, but through the explicit integration of causal reasoning, adaptive mechanisms, and data-efficient design principles. This hypothesis reframes the problem of model improvement from an engineering challenge to an epistemological one, questioning how knowledge is represented, adapted, and transferred under data scarcity or evolving contexts.

In this sense, the move toward a data-efficient framework is not only a technical proposal but also an ontological hypothesis about the nature of learning systems in complex environments. It assumes that methodological economy, understood as achieving explanatory and predictive power with minimal data, is an emergent property of systems that align representation, causality, and adaptation. Testing this hypothesis requires the development and validation of a framework that operationalizes these principles across real-world data scenarios.

Chapter 3

Proposal

The purpose of models is not to fit the data but to sharpen the questions.

Samuel Karlin

Chapter Overview

This chapter introduces the conceptual foundations and research rationale of a data-efficient neural network–based framework for electricity-demand forecasting in MGs. It begins by outlining the core research gap: existing forecasting paradigms fail to operate effectively in remote areas constrained by scarce and fragmented data, nonlinear socio-technical dependencies, and non-stationary demand evolution. These interrelated challenges invalidate the assumptions of conventional models built on abundant, homogeneous, and stationary datasets, calling for new methods capable of learning reliably under uncertainty while remaining responsive to dynamic system conditions.

Building on this foundation, the chapter formulates the four Research Questions that address these methodological challenges and define the thesis trajectory. It then introduces ANGEL, the proposed framework that operationalizes the responses to these questions. ANGEL is structured across four epistemic levels: Problem, Component, Principle, and Methodology, each representing a distinct layer of knowledge. The Problem level captures the empirical conditions driving forecasting uncertainty. The Component level organizes these conditions into three cognitive functions: observation, representation, and evolution, which describe how information is perceived, modeled, and updated over time. The Principle level articulates the theoretical rationale through three guiding concepts: causality for structural grounding under data scarcity and fragmentation, flexibility for representing nonlinear and context-dependent interactions, and adaptability for maintaining coherence as demand evolves. Finally, the Methodology level operationalizes these principles through analytical strategies such as causal feature engineering, flexible modeling, and adaptive adjustment.

The chapter concludes by discussing the scientific and practical significance of the proposed approach and outlining the expected contributions of the study.

3.1 Research Gap

Building upon the limitations identified in the previous section, this study focuses on the unresolved methodological and contextual challenges that hinder the achievement of low-error demand forecasting in MGs. Despite notable progress in electricity demand forecasting, current methods remain poorly suited to the operational and data realities of MGs in remote areas. Most ML and deep learning approaches are designed for abundant, stationary, and homogeneous datasets, assuming that data are both representative and stable over time. These assumptions do not hold in MGs, where information is often scarce, fragmented, and subject to complex socio-technical and environmental interactions, evolving dynamically as electrification progresses [121]. Under such conditions, traditional forecasting models exhibit three critical limitations.

First, their data dependency makes them fragile in low-data regimes: they require extensive historical observations to identify reliable patterns, and when trained on small or irregular datasets, they tend to overfit noise rather than capture meaningful dynamics.

Second, their structural rigidity constrains their ability to represent nonlinear and feedback-driven relationships that dominate small-scale systems. Models built on linear or quasi-linear assumptions fail to capture threshold effects and coupled behaviors among technological, environmental, and behavioral factors, resulting in unstable or misleading predictions.

Third, their static nature prevents them from maintaining predictive coherence in nonstationary contexts. As demand evolves with changes in appliance ownership, usage habits, and community development, static models trained on past data rapidly lose validity [96].

Even deep learning models, while capable of learning complex patterns, generally operate as black boxes that approximate statistical regularities without embedding causal or adaptive reasoning. Consequently, their forecasts degrade when applied to unseen contexts or evolving systems, and their internal logic remains difficult to interpret or validate. [16]

These shortcomings point to a persistent methodological gap: existing forecasting paradigms lack the capacity to forecast in remote areas under the combined constraints of scarce and fragmented data, nonlinear dependencies, and nonstationary demand evolution [38].

Addressing this gap requires rethinking electricity demand forecasting as a causal, flexible, and evolving process, where learning mechanisms capture the true drivers of consumption rather than mere statistical correlations. The central intuition guiding this thesis is that reliable forecasting in remote areas cannot overlook any of the three fundamental challenges identified earlier.

To this end, the thesis proposes ANGEL, a data-efficient neural network–based framework for learning under uncertainty. ANGEL functions as an adaptive agent capable of deriving meaning in contexts affected by data scarcity and fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution. The framework aims to achieve robust forecasting performance in data-constrained contexts while maintaining efficiency. The name also evokes the idea of a mediator between what is known and what is unknown, reflecting its role in bridging limited evidence and meaningful understanding. In this sense, ANGEL embodies illumination, the capacity to reveal structure and causality in data-constrained contexts, transforming fragmentary information into coherent and actionable knowledge.

To operationalize this vision, the following Research Questions are formulated to address the key methodological challenges of causality, flexibility, and adaptability identified above.

3.2 Research Questions

Building on this vision, the thesis formulates four Research Questions (RQs), each addressing a methodological challenge in forecasting electricity demand under conditions of data scarcity and fragmentation, nonlinear socio-technical and environmental dependencies, and nonstationary demand evolution. These questions are summarized in Table 3.1.

Table 3.1: Conceptual structure of the Research Questions.

Dimension	**Invariant**	**Adaptability**
Causality	RQ1 Causal Feature Engineering	RQ3 Latent Representation
Flexibility	RQ2 Flexible Representation	RQ4 Continual Learning

Each quadrant represents a progressive layer of methodological development:

- **RQ1 (Invariant–Causality):** How can explanatory and contextual variables be transformed into stable causal representations that capture reliable relationships between electricity demand and its underlying drivers under conditions of data scarcity and fragmentation [94, 70]?
- **RQ2 (Invariant–Flexibility):** How can flexible neural architectures enhance the representation of complex, nonlinear interactions among socio-technical and environmental variables beyond the assumptions of linearity or additivity [144, 148]?
- **RQ3 (Adaptability–Causality):** How can latent or indirectly observable socio-technical processes, such as behavioral adaptation, appliance diffusion, and technological adoption, be represented to capture the dynamic evolution of electricity demand over time [97, 138]?

- **RQ4 (Adaptability–Flexibility):** How can adaptive learning mechanisms enable flexible representations to remain efficient and responsive to evolving demand patterns without requiring complete retraining [36, 15]?

These Research Questions delineate the methodological trajectory of the thesis: progressing from correlation to causation, from linear representation to flexible understanding, and from static estimation to adaptive learning. ANGEL is not a generic ML construct but a domain-specific methodology designed to address the distinctive conditions of remote MGs, characterized by data scarcity and fragmentation, nonlinear socio-technical and environmental dependencies, and nonstationary demand evolution.

3.3 Data-efficient framework

To address these four Research Questions, the thesis proposes ANGEL, a data-efficient framework for developing neural networks composed of three interdependent components: observation, which extracts meaningful information from scarce and fragmented data; representation, which captures the causal and nonlinear dependencies governing electricity demand; and adaptivity, which ensures that models remain coherent and reliable under nonstationary and dynamically changing conditions.

At the first component, observation, most forecasting approaches rely on incomplete or weakly representative data, capturing only superficial statistical regularities rather than the structural signals that describe demand behavior. In MGs, the scarcity, fragmentation, and heterogeneity of available records severely limit what can be empirically observed. As a result, models tend to overfit to transient correlations that fail to generalize when data distributions shift or new consumption patterns emerge [103]. Achieving robustness under such conditions requires learning from observations that reflect stable and transferable mechanisms, linking climatic, temporal, and behavioral drivers to electricity use. This calls for a principled approach to causal feature construction, ensuring that observed information retains meaning across different stages of electrification.

At the second component, representation, most forecasting models describe demand behavior through statistical associations rather than structural understanding. Their representations often capture surface-level correlations without accounting for the causal or nonlinear mechanisms that drive electricity consumption. In reality, demand emerges from the interaction of climatic variability, appliance use, and behavioral adaptation, processes that are inherently nonlinear and context dependent. Capturing such relationships requires architectures capable of representing functional complexity without losing robustness. Nonlinear modeling thus becomes not merely a means of decreasing the error, but a way of expressing how diverse drivers jointly shape demand, allowing forecasts to reflect the underlying dynamics of the system rather than its transient patterns [127].

At the third component, evolution, the challenge lies in capturing the temporal dynamics that govern how demand patterns and model behavior evolve over time. Demand patterns in MGs located in remote areas evolve as consumption habits, economic activities, and technological penetration change over time. Models must therefore include mechanisms to detect shifts in data regimes and adjust their learning dynamics accordingly. This entails the use of adaptive update strategies, such as online learning, transfer learning, or meta-learning, that enable timely adjustment without compromising stability. By modeling not only static dependencies but also the timing and rate at which they change, forecasting methods can maintain low error and resilience under nonstationary conditions [57].

In combination, observation, representation, and evolution define the scope of the forecasting problem in remote areas. Effective models must not only extract meaningful information from scarce and fragmented data, but also articulate the behavioral and environmental interactions that shape demand and remain responsive as these interactions evolve. The following section develops this perspective into a coherent methodological basis for forecasting under uncertainty and change.

3.3.1 Conceptual Foundations

Each of the three components of the approach corresponds to a distinct methodological principle that governs how forecasting knowledge is acquired, represented, and maintained over time.

Observation is grounded in causality, as it requires identifying reliable and transferable relationships within scarce and fragmented data. Representation relies on flexibility, enabling models to capture the complex and context-dependent interactions that shape electricity demand. Finally, evolution is sustained by adaptability, ensuring that learning processes remain coherent and responsive as demand patterns and socio-technical conditions change. Together, causality, flexibility, and adaptability provide structural, expressive, and temporal coherence respectively, allowing forecasting models to operate synergistically across the empirical, representational, and temporal dimensions of demand prediction.

Causality provides the structural foundation of learning. It guides the selection of features that reflect stable mechanisms rather than transient statistical associations. Traditional ML models are inherently correlational, they depend on empirical regularities in large datasets and implicitly assume temporal stationarity. In contrast, causal reasoning isolates invariant relationships that explain why demand changes, rather than merely describing how it behaves. Variables such as temperature, rainfall, or income exert causal influence across contexts, making them reliable drivers for learning even when data are scarce or fragmented. Embedding causal structure within forecasting models constrains their search space, steering them toward physically consistent and generalizable relationships. By focusing on structural dependencies instead of statistical regularities, causal modeling enhances robustness, interpretability, and transferability, enabling meaningful inference under uncertainty and supporting knowledge reuse across different stages of electrification [122].

Flexibility provides the expressive capacity to represent the complex and evolving dependencies that characterize electricity demand in remote MGs. Behavioral, climatic, and economic factors interact in nonlinear and context-dependent ways, generating thresholds, feedback loops, and saturation effects that linear methods cannot reproduce. This property is particularly relevant in electrification contexts, where demand evolves from basic to diversified uses as households acquire new technologies and modify their consumption habits. Flexible architectures, such as KANs, capture these functional interactions through modular mappings, enabling models to approximate real-world dynamics with both accuracy and interpretability [76].

Adaptability emerges as a transversal property of this interaction. From the causal perspective, it ensures that the stable mechanisms encoded in the model can evolve as the underlying socio-technical processes themselves change. From the flexibility perspective, it allows the model to adjust its internal mappings incrementally as new data become available, preserving expressiveness without loss of stability. Mechanisms such as Continual Learning (CL) and the inclusion of latent causal variables make this possible: the former enables the assimilation of new information without catastrophic forgetting, while the latter captures hidden processes, such as electrification maturity or appliance adoption, that drive long-term demand evolution [11].

Rather than functioning as separate components, these three principles form an integrated logic of learning: causal reasoning defines what to learn, flexible modeling defines how it is expressed, and adaptive mechanisms determine how it evolves and remains valid over time. They articulate the essence of data-efficient forecasting, a paradigm that is robust to limited and fragmented data, capable of representing complex dependencies, and resilient to continuous change.

3.3.2 Architecture

Figure 3.1 presents the architecture of ANGEL, structured across four epistemic levels, Problem, Component, Principle, and Methodology. Each level represents a distinct type of knowledge, ranging from empirical observation to theoretical justification and methodological implementation. They form a coherent rationale that links the conditions under which knowledge is produced with the mechanisms through which it becomes operational.

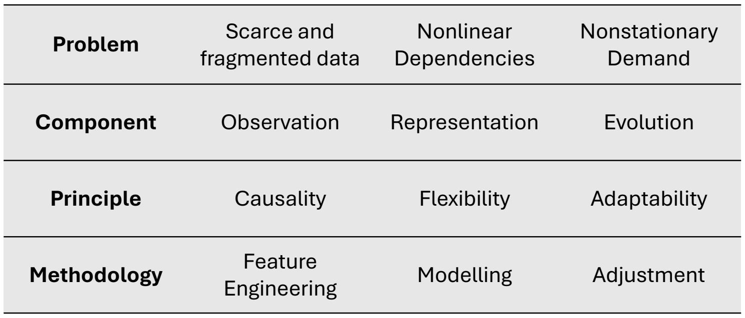
At the Problem level, knowledge is empirical. This layer captures the observable limitations that shape forecasting uncertainty in MGs: scarce and fragmented data, nonlinear dependencies among socio-technical variables, and nonstationary demand evolution. It defines the epistemic object of inquiry, the phenomena that forecasting seeks to understand and predict. Here, knowledge is descriptive and diagnostic: it characterizes how demand behaves and why conventional models fail to forecast it with low error.

The Component level articulates conceptual knowledge. It organizes the forecasting process into three cognitive functions, observation, representation, and evolution, that structure how information is perceived, interpreted, and updated. At this level, knowledge is schematic: it defines the conceptual architecture of forecasting, translating empirical problems into analytical perspectives. It bridges the gap between what is observed and what must be formulated, establishing the structure of the learning process itself.

The Principle level introduces theoretical knowledge. It provides the abstract rationale that governs each component: causality provides the structural foundation and stability of observation; flexibility allows representation to accommodate nonlinear and context-dependent interactions; and adaptability maintains coherence between models and evolving systems. These principles define the conditions under which forecasting knowledge can be considered valid, generalizable, and robust. Knowledge here is normative: it prescribes how learning should occur to remain scientifically consistent.

The Methodology level embodies procedural knowledge, the set of analytical processes that operationalize theoretical principles. Feature engineering implements causal reasoning through the construction of stable representations; modeling expresses flexibility by learning nonlinear functional forms; and adjustment realizes adaptability through iterative updating and recalibration [93]. At this level, knowledge becomes performative: it is encoded in reproducible methods that transform theory into practice.

These four levels define a continuous epistemic gradient, extending from empirical observation to methodological realization, through which ANGEL connects what is known with how it is learned and applied. This layered rationale ensures that each modeling decision rests on a coherent chain of reasoning, linking the empirical context of MGs with the theoretical and computational mechanisms that enable data-efficient forecasting.



Problem	Scarce and fragmented data	Nonlinear Dependencies	Nonstationary Demand
Component	Observation	Representation	Evolution
Principle	Causality	Flexibility	Adaptability
Methodology	Feature Engineering	Modelling	Adjustment

Figure 3.1: Epistemic architecture of ANGEL, structured across four levels of knowledge: empirical (Problem), conceptual (Component), theoretical (Principle), and procedural (Method).

Building on this structure, the following sections detail how ANGEL's theoretical principles are instantiated through corresponding methodological strategies. Specifically, they examine how causality, flexibility, and adaptability are operationalized within the proposed architecture through feature engineering, modeling, and adjustment. This discussion illustrates how ANGEL translates high-level design principles into actionable analytical procedures, clarifying the dynamic interplay between theory and method that underpins its internal logic.

3.3.3 Causal Feature Engineering

A key methodological contribution of this work is the integration of causal reasoning into feature design for electricity demand forecasting. This is not a generic application of causal inference but a domain-specific strategy tailored to conditions of data scarcity and fragmentation. Understanding the mechanisms that generate demand patterns is as critical as minimizing forecasting error. Thus, reliable forecasting under data scarcity and fragmentation requires moving beyond correlations toward representations of the causal-like mechanisms that shape demand [83].

Recent advances in data-efficient deep learning suggest that the integration of causal reasoning with neural architectures is becoming a defining trend, which deserves further exploration in this context [22]. Electricity consumption in MGs in remote areas emerges from a combination of environmental, temporal, and socio-behavioral drivers acting across different timescales and levels of observability. Environmental conditions such as temperature and humidity play a crucial role in shaping electricity demand. Temperature affects the use of refrigeration for food preservation and medical supplies, as well as fans and basic ventilation during warmer months. Humidity modulates the efficiency of appliances and affects comfort-driven use, reinforcing temperature effects. Temporal routines, captured by variables such as hour of day, month, and weekend, reflect daily cooking schedules, evening lighting, small-scale commercial or communal events, and local festivities or traditions that alter demand profiles in predictable ways. Many of these relationships are causal rather than merely correlative: they represent mechanisms through which environmental and social contexts drive electricity use. However, these mechanisms are often partially observable, due to limited metering coverage and the intermittent nature of data collection. Embedding causal reasoning in feature construction allows the model to focus on structurally stable relationships, for instance, between climatic cycles and refrigeration demand or between social routines and evening peaks, rather than on transient or spurious correlations [127].

Within ANGEL, causality is not imposed through explicit causal-graph architectures but encoded in the representation of inputs. Observable variables are transformed into features that express the direct or mediated effects of known drivers, such as meteorological conditions, temporal cycles, or community routines, according to domain knowledge about how demand responds to context. This process, termed causal abstraction, ensures that each input has a clear physical or socio-technical meaning [113].

In practical terms, causal feature engineering is implemented through a sequence of domain-informed transformations. First, raw variables are categorized according to their causal role (e.g., exogenous drivers such as temperature or humidity, and endogenous responses such as appliance usage). Then, derived features are generated to capture lagged effects, cumulative influences, or interaction terms that represent real-world mechanisms, for instance, temperature–humidity indices or temporal encodings for social routines. These transformations are guided by prior empirical evidence and validated through sensitivity analysis, ensuring that each engineered feature aligns with a plausible causal pathway rather than statistical convenience. A comparable approach is seen in recent prosumer forecasting research, where extensive feature generation and selection processes are employed to model complex energy behaviors using domain knowledge and temporal dynamics [152].

Grounding feature design in causal reasoning reduces overfitting to spurious correlations and improves model reliability. Because features correspond to identifiable processes, stakeholders can relate model behavior to known phenomena, such as climatic variations, improving trust. In this way, causal feature engineering bridges limited observations and the evolving mechanisms governing energy use in MGs, transforming forecasting into a task of explanation as well as foresight [152].

3.3.4 Flexible Modeling

Electricity demand in MGs exhibits inherently nonlinear dynamics that arise from multiple, overlapping mechanisms. First, physical nonlinearities emerge from the thermodynamic behavior of appliances and systems: the relationship between temperature and electricity use is non-proportional, as cooling and refrigeration loads grow exponentially beyond certain thermal thresholds. Similarly, the efficiency of electrical devices varies with voltage fluctuations and humidity, introducing nonlinear effects even under stable operating conditions [137].

Second, behavioral nonlinearities result from threshold-driven human decisions. Households often change usage patterns abruptly, for instance, switching on multiple appliances during meal preparation or social events, creating discontinuous jumps in demand. These responses are context-dependent, influenced by routine and access to energy services, leading to variable elasticities rather than smooth gradients [111].

Third, infrastructural and operational nonlinearities arise from the interaction of distributed generation, storage, and load management within the MG. Limited capacity or technical constraints can trigger saturation effects, feedback loops, and sudden curtailments, especially when several users draw power simultaneously [81].

Finally, evolutionary nonlinearities accompany the process of electrification itself:

as households acquire new appliances or engage in productive uses, the underlying relationships between environmental, social, and economic drivers of demand are redefined [9].

Capturing these diverse forms of nonlinearity requires models capable of representing complex, non-additive dependencies and of adapting to structural change over time. Conventional linear or additive models, based on constant coefficients, cannot express such dynamic and threshold-sensitive behavior. In contrast, neural architectures provide the flexibility to approximate these interactions, allowing the model to learn both local sensitivities and global transitions within the evolving energy system [28].

ANGEL therefore employs modeling, using flexible architectures capable of representing complex functional relationships. Nonlinearity here is not merely a mathematical property but a mechanism for representing feedback loops, thresholds, and behavioral discontinuities that emerge during electrification. The model evolves alongside the system it describes, maintaining representational fidelity and explanatory coherence. Nonlinearity thus becomes the mathematical expression of a broader epistemic stance: modeling electrification not as a fixed equilibrium but as a dynamic, feedback-driven process.

3.3.5 Adaptive Adjustment

The dynamic nature of electrification requires forecasting methods capable of capturing the continuous evolution of electricity demand. Static approaches, defined by fixed assumptions or parameters, are inherently inadequate in contexts where consumption behaviors, technologies, and infrastructures evolve over time. In remote MGs, socio-technical transformations occur rapidly as households acquire new appliances, adopt productive uses, and modify their energy practices. Long-term reliability therefore depends on adaptive mechanisms that allow forecasting processes to update their knowledge base and remain contextually aligned with evolving realities [57].

Adaptation in ANGEL unfolds as a continuous feedback process that links new evidence with prior understanding. At the data level, it involves the regular incorporation of updated information, such as new consumption records, climatic observations, or socio-economic indicators, to refine and expand the knowledge base. At the analytical level, adaptation ensures that the relationships inferred from historical data remain consistent with current and emerging conditions, adjusting weights or priorities among variables as contextual relevance evolves. At the structural level, adaptation captures deeper transformations within the electrification process itself, such as the diffusion of new appliances, changes in income-generating activities, or shifts in community-scale energy practices [156].

In practical terms, adaptive adjustment operates through mechanisms of iterative learning and transfer of knowledge. Each new observation contributes to refining previous insights rather than replacing them, allowing the forecasting process to evolve cumulatively over time. When significant contextual changes occur, such as seasonal transitions, infrastructural upgrades, or behavioral disruptions, the updating process integrates these discontinuities, maintaining continuity and coherence in how information is interpreted and used. This dynamic updating reduces temporal drift and ensures that forecasting remains relevant under changing socio-technical and environmental conditions [57].

A key element of this adaptive process is the consideration of latent causal factors that evolve slowly over time, such as electrification maturity, appliance diffusion, or community development. These latent dynamics serve as contextual anchors, ensuring that short-term adjustments remain consistent with long-term structural evolution. By linking observable changes to deeper causal processes, adaptation supports a coherent understanding of demand evolution rather than a purely reactive adjustment to fluctuations. Ultimately, adaptive adjustment transforms forecasting from a static estimation task into an evolving reasoning process. It sustains robustness and stability under nonstationary conditions by maintaining coherence between immediate variability and gradual socio-technical transformation. In doing so, it strengthens the capacity of forecasting to serve as a decision-support tool for energy planners and policymakers, aligning analytical reliability with the broader objectives of resilient and adaptive electrification in remote areas [156].

3.4 Scientific and Practical Significance

In the broader context of electricity demand forecasting, ANGEL operationalizes a coherent methodological triad. Causality provides structural stability by grounding learning in mechanisms that remain valid under data scarcity and fragmentation. Flexibility introduces expressive capacity to represent nonlinear socio-technical and environmental dependencies. Adaptability ensures temporal validity, allowing the forecasting process to remain coherent as demand patterns evolve. Together, these principles define a methodological paradigm capable of achieving low forecasting error in remote areas through causal grounding, while maintaining resilience under nonlinear dependencies and nonstationary demand evolution.

From a scientific perspective, closing this gap advances the frontier of data-efficient AI. By demonstrating how causal structure, flexible representation, and adaptive adjustment can be coherently integrated within a unified learning process, this research expands the methodological boundaries of ML toward contexts characterized by limited, fragmented, nonlinear, and evolving data. ANGEL contributes to the development of learning systems that maintain robustness, stability, and long-term adaptability in remote areas [69].

From a theoretical standpoint, ANGEL bridges two paradigms often treated as incompatible: causal inference and flexible representation. This synthesis introduces a principled way to embed domain knowledge into learning systems, ensuring that inferred relationships remain stable under distributional shifts and evolving conditions [127].

It also demonstrates that high predictive performance and adaptive flexibility can coexist when causality and nonlinearity are coherently integrated. Furthermore, by incorporating continuous updating and latent-variable abstraction, ANGEL aligns with the emerging paradigm of adaptive and self-evolving learning systems [11].

From a practical perspective, ANGEL directly supports the design and operation of MGs. Reliable and adaptive forecasting mitigates two persistent risks in MG planning: overestimation, which leads to oversized infrastructure and excessive capital expenditure, and underestimation, which results in frequent shortages and operational instability [117]. By embedding causal reasoning, flexible representation, and adaptive adjustment, ANGEL enables decision-making that is both evidence-based and context-sensitive, improving investment efficiency, operational reliability, and long-term sustainability.

Beyond the energy domain, these methodological advances are transferable to other sectors characterized by scarce data and dynamic evolution, such as agriculture, water management, and climate adaptation. In such systems, decision-makers face analogous challenges of uncertainty and change. A causal–flexible–adaptive approach provides a generalizable foundation for forecasting under these constraints, enhancing the robustness and interpretability of predictions [87].

At an institutional level, adaptive forecasting strengthens the governance and sustainability of MGs initiatives. Reliable demand estimation improves planning transparency, facilitates resource allocation, and supports the transition from externally managed projects to community-driven systems [57]. In this sense, ANGEL contributes to broader policy goals, enhancing energy equity, local empowerment, and climate resilience.

3.5 Expected Contributions

The expected contributions of this doctoral research correspond to four complementary epistemic levels, empirical, conceptual, theoretical, and methodological, and to its overarching commitment to responsible and transparent AI. Each level generates a distinct form of knowledge that, together, constitutes ANGEL, a coherent and testable framework for data-efficient forecasting in remote MGs. This multi-layered contribution connects empirical diagnosis, conceptual reasoning, theoretical formulation, and methodological realization within a unified epistemological architecture.

At the empirical level, the research produces a systematic characterization of the uncertainty factors that constrain forecasting in remote areas. By quantifying the effects of data scarcity, fragmentation, nonlinear socio-technical dependencies, and nonstationary demand evolution, it contributes explanatory knowledge about the structural causes of model unreliability and the conditions under which data-efficient approaches are required.

At the conceptual level, the thesis reconceptualizes forecasting as a *causal, flexible, and adaptive* reasoning process that integrates observation, representation, and evolution as core analytical dimensions. This contribution advances a framework that links empirical phenomena, such as behavioral adaptation or electrification maturity, with computational representation, thus transforming forecasting from a purely predictive task into a dynamic form of reasoning about evolving systems.

At the theoretical level, the research formalizes a normative model of data-efficient learning grounded in three epistemological principles, causality, flexibility, and adaptability. It defines the conditions under which forecasting knowledge can be generated, validated, and generalized, providing a theoretical basis for coherent learning in data-constrained contexts.

At the methodological level, the research operationalizes these principles through a reproducible computational pipeline that integrates feature engineering, flexible modeling, and adaptive adjustment. This level contributes procedural knowledge by demonstrating how theoretical principles can be instantiated in transparent, scalable, and resource-efficient forecasting methods applicable to real-world energy systems.

Beyond its technical scope, the thesis provides a broader scientific and ethical contribution to the field of AI for sustainable energy systems. It establishes a principled model for learning in data-constrained contexts, aligning computational efficiency with scientific transparency and social accountability. Adhering to open-science principles of reproducibility and accessibility, the work advances a vision of AI as an explanatory and ethical instrument that supports equitable and sustainable energy transitions.

The expected contributions include:

- Empirical understanding of the uncertainty mechanisms that limit forecasting reliability in remote MGs;
- Conceptual reframing of forecasting as a causal, flexible, and adaptive reasoning process;
- Theoretical grounding for data-efficient learning based on causality, flexibility, and adaptability;
- Methodological implementation of these principles through reproducible and scalable computational procedures; and
- Broader scientific and ethical advancement of responsible and transparent AI for sustainable energy systems.

Collectively, these contributions generate an integrated body of knowledge that connects empirical evidence, conceptual reasoning, theoretical rigor, and methodological precision in support of data-efficient and context-aware forecasting.

Chapter 4

Research Methodology

The best way to predict the future is to invent it.

Alan Kay

Chapter Overview

Building upon the theoretical and methodological principles established in Chapter 3, this chapter translates ANGEL into a concrete research design that defines the hypotheses, outlines the empirical and experimental validation strategy, and connects theoretical assumptions with practical implementation. It situates the ANGEL framework within its empirical and contextual foundations, describing the data sources, collection processes, and characteristics of the datasets used for analysis, while emphasizing the challenges of scarcity, fragmentation, and nonstationarity that typify remote MGs.

The discussion characterizes the data challenges associated with missing values, irregular sampling, and noise, and outlines the environmental, temporal, and socioeconomic features that capture the multifaceted drivers of electricity demand.

The chapter then presents the experimental setup, detailing the configuration of model architectures, training environments, and optimization strategies. Finally, it introduces the evaluation metrics and validation procedures adopted to ensure methodological rigor, robustness, and reproducibility.

This chapter provides the operational foundation for the empirical validation presented in Chapter 5, demonstrating how the principles of causality, flexibility, and adaptability are instantiated through systematic experimentation and transparent research practices.

4.1 Research Design

Designing a methodology for forecasting electricity demand in MGs located in remote areas requires a framework that is both empirically grounded and conceptually innovative. The contexts examined in this thesis are characterized by structural data scarcity and fragmentation [66], nonlinear dependencies among socio-technical and environmental variables [71], and nonstationary demand evolution [138], all of which jointly challenge the reliability and generalizability of conventional forecasting models. These conditions fundamentally constrain the availability, consistency, and representativeness of empirical evidence, rendering conventional research designs that rely on large datasets or single analytical paradigms inadequate.

To address these challenges, the research design integrates three complementary dimensions: empirical analysis, experimental evaluation, and validation. They form a coherent strategy that links abstract conceptual foundations with concrete methodological outcomes. This integrated approach ensures that each phase of the research, from data collection to analytical formulation, remains sensitive to the empirical limitations of remote electrification contexts while preserving theoretical rigor and methodological innovation.

The empirical foundation of this research is based on data measured and collected from a MG operating in a recently electrified area, capturing the complex and evolving dynamics of electricity demand. These data reveal nonlinear dependencies among socio-technical and environmental factors, such as appliance adoption, temperature, and humidity, that jointly shape consumption behavior. The interaction of these drivers produces coupled and context-dependent demand patterns, where small variations can trigger feedback loops and threshold effects typical of newly electrified systems.

The dataset captures the nonstationary evolution of electricity demand. Despite its inherent limitations, it provides an authentic depiction of demand dynamics in contexts where data scarcity, fragmentation, nonlinearity, and change are structural characteristics of the electrification process rather than anomalies. These measurements serve for validating forecasting approaches under real-world constraints. Complementary information sources, including meteorological records, further enrich the empirical basis. These heterogeneous inputs capture the interplay of environmental, technical, and behavioral drivers of electricity demand, highlighting the need for methods capable of integrating partial, asynchronous, and multi-domain information when direct measurement is limited or unavailable [133].

Building on this empirical foundation, the experimental design systematically evaluates alternative forecasting approaches under controlled conditions that mirror the limitations of real-world MG data. A series of experiments compare different neural architectures, including the triangle-shaped network, Deep Feedforward Neural Network (DFNN), Multilayer Perceptron (MLP), and KAN. This comparative setup enables a rigorous evaluation of how each approach reflects the underlying epistemic principles of ANGEL, causality, flexibility, and adaptability, within the constraints of real-world MG data. It assesses the capacity of each architecture to capture the structural relationships among socio-technical and environmental variables, to represent complex and context-dependent interactions, and to maintain coherence under evolving demand conditions. In doing so, it bridges empirical experimentation with the theoretical foundations of ANGEL, aligning practical evaluation with the broader rationale of data-efficient forecasting in data-constrained contexts.

Experiments are structured not only to benchmark forecasting error through metrics such as Mean Absolute Error (MAE) and Mean Squared Error (MSE), but also to assess robustness under noisy inputs, adaptability to sparse datasets, and clarity of the results. Importantly, experimentation follows an iterative process: early shortcomings expose structural weaknesses that inform successive refinements in feature construction, architectural selection, and validation protocols. In this way, the experimental analysis remains closely aligned with the conceptual evolution of the thesis. Comparable evaluation strategies have been successfully employed in prior studies that benchmarked deep learning architectures against classical and hybrid models for short-term and day-ahead forecasting in data-constrained environments [141].

The validation phase integrates the insights derived from empirical analysis and experimental evaluation to assess the coherence, efficiency, and robustness of the ANGEL framework as a whole. It examines whether the combined implementation of causal reasoning, structural flexibility, and temporal adaptability ensures responsiveness to evolving data distributions while maintaining stability over time. Validation thus confirms the methodological soundness and practical relevance of the proposed approach, establishing a direct link between theoretical principles, empirical evidence, and operational applicability.

These three dimensions of the research design are not independent but mutually reinforcing. Empirical analysis grounds the work in observed reality, experimental evaluation ensures methodological rigor, and validation consolidates the findings into a coherent and testable framework. This integrated design ensures that the proposed methods are assessed not only in terms of statistical performance but also in relation to the defining conditions of electricity demand in remote areas: data scarcity and fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution. By grounding the evaluation in these contextual realities, the research design bridges conceptual foundations and practical implementation, establishing the basis for the hypotheses presented in the next section and their empirical validation in the subsequent chapter.

4.2 Research Hypotheses

Building on the conceptual foundations of causality, flexibility, and adaptability presented in Section 3.3.1, this section formulates the research hypotheses (H) that operationalize these principles into testable propositions. Each hypothesis is structured around an observed empirical challenge (A), an explanatory hypothesis (H), and an abductive conjecture that motivates empirical validation. Together, these hypotheses form the methodological backbone of the thesis, guiding the design of feature engineering, modeling strategies, and adaptive learning mechanisms for forecasting electricity demand in MGs.

H1, Causal models enhance robustness under data scarcity and fragmentation.

Electricity demand forecasting in remote MGs often suffers from instability when data are scarce or fragmented. Models that rely solely on statistical correlations fail to capture the persistent relationships that govern demand behavior [106].

> Forecasting models grounded in causal reasoning, where input features represent stable physical and behavioral mechanisms such as temperature effects, humidity, or weekly routines, achieve greater robustness under uncertainty. By encoding structural dependencies rather than transient correlations, causal architectures preserve explanatory validity across incomplete or variable datasets.

This suggests that causality provides the necessary epistemic grounding for robust forecasting in data-constrained MG environments.

H2, Flexible representations improve the modeling of nonlinear dependencies.

Electricity consumption in evolving MGs reflects nonlinear, context-dependent interactions among environmental, temporal, and socio-technical drivers that cannot be captured by rigid functional forms [40].

> Forecasting models should adopt flexible representational architectures capable of adapting their functional form to the data structure. By enabling flexible mapping through locally adaptive transformations, such models can represent complex dependencies and reduce forecasting error without overfitting or excessive data requirements.

This hypothesis posits flexibility as the methodological counterpart to causality, allowing models to express heterogeneous and evolving demand behaviors within consistent structural boundaries.

H3, Modeling must account for the causal evolution of microgrids over time.

In newly electrified or developing communities, electricity demand evolves alongside social and technological adoption. Ignoring this temporal evolution leads to systematic bias and error accumulation [138].

> Forecasting frameworks should integrate temporal mechanisms that explicitly model the causal evolution of demand, linking changes in consumption patterns to stages of adoption and use. This temporal awareness enables models to maintain coherence as MGs mature, ensuring consistency between data collected during early electrification stages and the emerging consumption behaviors that follow.

This view extends causal reasoning to the time domain, recognizing that electricity demand follows structured developmental trajectories rather than random variation.

H4, Adaptive learning ensures sustained performance under nonsta-

tionarity.

In dynamic environments, where behavioral and environmental conditions shift continuously, static models lose validity in electricity demand forecasting and require frequent retraining [15].

> Forecasting models should incorporate adaptive and flexible learning mechanisms that update internal representations as new data arrive while preserving previously acquired causal knowledge. Such CL maintains temporal stability, minimizes error growth, and supports long-term adaptability in nonstationary MG conditions.

Adaptability unifies causality and flexibility into a continuous epistemic process, allowing the model to evolve coherently with the demand it represents.

These four hypotheses translate the thesis's theoretical pillars, causality, flexibility, and adaptability, into a structured empirical strategy. Each hypothesis addresses a specific dimension of the forecasting process, reflecting the interaction between the epistemic principles of causality and flexibility across two methodological regimes: invariance and adaptability. Accordingly, H1 relates to causality under invariance, H2 to flexibility under invariance, H3 to causality under adaptability, and H4 to flexibility under adaptability. Their combined implementation aims to minimize forecasting error while ensuring robustness, generalization, and operational relevance for MG planning and sustainable energy access.

In this sense, the hypotheses bridge conceptual reasoning and empirical validation, connecting methodological innovation in data-efficient AI with actionable insights for MG development and the broader objectives of Sustainable Development Goals (SDG) 7.

4.3 Research Context and Data Sources

Following the articulation of the research design and hypotheses, this section situates the ANGEL within its empirical setting, the forecasting of electricity demand in MGs in remote areas. In such contexts, **data scarcity and fragmentation** are structural rather than incidental, reflecting the technological and socioeconomic conditions that shape decentralized electrification. Unlike urban or industrial systems, where smart metering ensures HF and standardized data streams, remote MGs often depend on partial or manually collected measurements characterized by missing intervals, irregular sampling, and noise from outages or recording errors. Consequently, long, continuous time series, assumed by most conventional forecasting techniques, are rarely available, making it necessary to develop forecasting methods that can operate effectively under informational constraints.

This scarcity and fragmentation is compounded by infrastructural and institutional barriers. Financial limitations often prevent the deployment of permanent monitoring systems, while logistical challenges, such as geographic isolation and limited connectivity, hinder systematic data acquisition. Local utilities, when present, may lack the capacity to store or process records in standardized formats, and existing datasets typically originate from short-term pilot projects that fail to capture the long-term dynamics of electricity use. The result is a *fragmented and discontinuous mosaic of data*, marked by inconsistencies in temporal and spatial coverage, which complicates the construction of reliable and generalizable forecasting models [161].

Beyond data limitations, forecasting in remote MGs must also account for **nonlinear dependencies among socio-technical and environmental variables**. Electricity demand reflects the combined influence of behavioral, economic, and climatic factors whose interactions are inherently nonlinear. Income levels, household composition, appliance adoption, and grid reliability shape consumption patterns, while temperature and humidity affect energy use through comfort requirements and appliance operation. Small variations in these conditions produce shifts in demand, amplified or mitigated by infrastructural and social constraints. Capturing such relationships requires flexible modeling approaches capable of representing feedback loops, threshold effects, and coupled dynamics beyond the assumptions of linearity or additivity [75].

Furthermore, the challenge of **nonstationary demand evolution** further differentiates remote MGs from conventional electricity systems. As households and businesses progressively adopt new appliances, improve living standards, and diversify productive activities, electricity use evolves rapidly, causing past data to lose predictive relevance. Seasonal agricultural cycles, climatic variability, and community routines reinforce this nonstationarity, introducing abrupt peaks or troughs linked to market days, school schedules, or festive events [41].

To address these challenges, ANGEL draws upon multiple categories of data to approximate the underlying mechanisms that govern electricity demand. Environmental variables, such as temperature, humidity, capture physical influences on energy use. Temporal descriptors, hour of day, type of day, and season, encode cyclical and behavioral rhythms. Socioeconomic and demographic indicators, such as household size, income level, and productive activities, introduce contextual dimensions reflecting human practices. Together, these heterogeneous data layers support the construction of a causally informed representation of demand, enabling learning models to generalize from limited evidence.

The empirical foundation of this research is provided by the case of El Espino [13], a rural community in Bolivia that typifies the data and infrastructural constraints faced by MGs. The locality is characterized by modest energy infrastructure, intermittent electricity supply, and limited metering coverage. Available records originate from pilot projects and monitoring campaigns that, although incomplete, offer sufficient granularity to enable the experimental validation of data-efficient forecasting methods. Owing to its structural scarcity, data fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution, the El Espino dataset constitutes an exemplary testbed for evaluating ANGEL's capacity to perform under the realistic and complex conditions of decentralized electrification.

Comparable datasets were initially available for Bolivia, Namibia, Mexico, and Tanzania; however, many exhibited extensive gaps or irregular sampling intervals. After filtering incomplete entries, the consolidated dataset comprised 869 days of usable records, with El Espino providing the most complete subset, 578 consecutive days of electricity-demand observations. This made it the preferred case for model development and validation. Although modest compared to urban datasets, it represents one of the most complete and granular collections available for a MG. Supplementary metadata, such as temperature and humidity, and derived temporal features, such as weekend, enable the application of AI-based forecasting methods capable of capturing underlying causal relationships even under limited observation conditions.

Recognizing data scarcity, fragmentation, nonlinear dependencies, and nonstationary demand evolution as structural features rather than anomalies is central to this research. Building on this empirical foundation, ANGEL treats datasets not as exhaustive representations of reality but as partial traces of complex socio-technical, environmental, and economic dynamics. By integrating contextual variables that reveal the nonlinear and causal mechanisms underlying observed patterns, particularly the interdependent relationships among socio-technical and environmental variables, even sparse, fragmented, and nonstationary demand measurements can yield actionable insights. This perspective shifts the focus from data quantity to data quality and relevance, enabling meaningful forecasts under the very conditions that characterize electrification in remote areas. In doing so, ANGEL remains both scientifically rigorous and directly aligned with the operational realities of the communities it aims to support.

4.4 Experimental Setup

The experimental setup was designed to evaluate ANGEL in MGs in remote areas under conditions of structural data scarcity and fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution. Unlike conventional forecasting contexts, where abundant and HF datasets are available, the case of El Espino (Bolivia) represents a setting where records are incomplete, fragmented, and noisy. For this reason, the experiments were tailored to test how neural networks behave under such constraints and whether feature engineering, modeling, and adjustment mechanisms can deliver meaningful improvements.

All models were developed using the PyTorch framework, selected for its versatility in constructing custom neural network architectures and for its efficiency in gradient-based optimization. The use of PyTorch enabled an iterative workflow in which models could be rapidly prototyped, adjusted, and evaluated, allowing a constant alignment between conceptual goals and empirical implementation.

The experiments initially employed a feature set composed of variables that are both widely available in open datasets and known to influence electricity demand. These included temperature, measured in degrees Celsius, which directly affects the use of appliances for heating and cooling; humidity, measured in percentage, which shapes perceived thermal comfort and indirectly influences the use of cooling devices; the hour of the day, which reflects the strong temporal cycles of daily routines and distinguishes between morning and evening peaks; the month of the year, which captures seasonal patterns and longer climatic cycles; and a binary indicator of the type of day, distinguishing between weekdays and weekends to reflect differences in socio-economic activity. As the research progressed, this feature set was refined to improve its ability to represent the electricity demand.

The target variable for all experiments was hourly electricity demand per capita, expressed in kilowatts per person (kW/person).

During the training process, four optimizers were systematically compared: Adaptive Moment Estimation (ADAM), Adaptive Gradient (ADAGRAD), Adaptive Delta (ADADELTA), and Stochastic Gradient Descent (SGD) [136].

ADAM combines the advantages of momentum, which stabilizes parameter updates by considering past gradients, with adaptive learning rates that are individually adjusted for each parameter. This combination makes it especially effective when handling noisy datasets in which irregular peaks or troughs in demand might destabilize convergence [67].

ADAGRAD, by contrast, scales the learning rate based on the frequency of updates, assigning larger steps to rarely adjusted parameters. This property allows the model to capture infrequent but important variations, such as short-lived spikes in demand linked to agricultural irrigation or community events [17].

ADADELTA was included as a refinement of ADAGRAD: by restricting the accumulation of past gradients to a fixed window, it prevents the learning rate from shrinking excessively, thereby ensuring that the model remains adaptable when new demand patterns emerge, for instance during the adoption of refrigerators, pumps, or electric mobility [158].

SGD, the most classical of the optimizers, updates parameters using small random subsets of training data. Although it requires careful tuning and more iterations to converge, its simplicity often encourages generalization and reduces the risk of overfitting to short-lived irregularities [20].

All experiments were conducted on a personal computer equipped with an Intel Core i9-9980HK CPU running at 2.4 GHz and 32 GB of RAM. The training process did not require GPU acceleration or access to a high-performance computing cluster. This highlights the practicality of the proposed methodology, demonstrating that robust forecasting models can be trained with modest computational resources, which is especially relevant for deployment in environments where advanced hardware infrastructure is unavailable.

This setup established a rigorous and reproducible experimental design. By carefully selecting inputs, adopting a data split that ensures independence between training and evaluation, and systematically comparing different optimization strategies, the design ensured that the results would be both technically robust and practically relevant for electrification challenges in remote areas.

4.5 Evaluation Metrics

The evaluation of the proposed forecasting models relied on error metrics designed to quantify the discrepancy between forecasted and observed electricity demand. Because both average performance and the ability to avoid critical deviations are essential in MGs planning, two complementary measures were employed: the MAE and the MSE.

The MAE was used as the primary evaluation metric. It measures the average magnitude of forecasting errors without considering their direction and is formally defined in Equation (4.1):

$$\varepsilon_{\mathrm{MAE}} = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i| \ , \tag{4.1}$$

where $\varepsilon_{\mathrm{MAE}}$ denotes the error, n is the number of data points, Y_i are the observed values, and $\hat{Y}_i$ are the forecasted values. Because it is expressed in the same units as the target variable (kW/person), the MAE provides an intuitive and easily interpretable measure of average error. Unlike the MSE, it treats all deviations proportionally, making it less sensitive to outliers while still reflecting the overall forecasting error [78].

To complement this measure, the MSE was also employed. The MSE is formally defined in Equation (4.2):

$$\varepsilon_{\mathrm{MSE}} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{4.2}$$

By squaring each error before averaging, the MSE places stronger emphasis on large deviations. Together, the MAE and MSE provide a balanced evaluation: the former reflects the average magnitude of forecasting errors, while the latter emphasizes the impact of large deviations occurring during peaks or atypical demand conditions [55].

Validation in this study consisted of systematically comparing model forecasts against observed demand data from El Espino. Lower values of MAE and MSE indicate models with lower error and higher reliability, while higher values highlight weaknesses in capturing variability such as abrupt appliance adoption or seasonally driven changes in demand. The comparison of optimizers across these metrics further clarified how different training strategies influence forecasting error and stability.

4.6 Validation Strategy

The validation strategy unfolds progressively, mirroring the conceptual evolution of ANGEL. It begins with the establishment of benchmark models using classical forecasting approaches, such as a simple triangular network, which provide reference points for assessing subsequent methodological improvements through the error metrics defined above. Once these baselines are established, the causal feature engineering pipeline is introduced. In this stage, raw variables such as temperature, humidity, or day type are systematically enriched with derived constructs like seasonal cycles and socio-behavioral proxies. This allows evaluating whether models trained on causally meaningful features demonstrate higher robustness under data scarcity and fragmentation compared to those relying purely on statistical regularities. The first validation stage focuses on assessing the contribution of causal reasoning to model stability.

The next stage focuses on representational flexibility. Here, the validation examines how structurally adaptive models can capture nonlinear dependencies among socio-technical and environmental variables. This step evaluates whether models that can dynamically align their internal structure to the complexity of the input space achieve more stable and context-sensitive forecasts, thereby addressing the limitations of static or depth-based approaches. The goal is to test whether structural flexibility constitutes the primary determinant of performance in heterogeneous and nonlinear environments.

The following stage introduces temporal adaptability. The validation examines whether explicitly representing the evolution of electrification and the progressive change in consumption behavior allows forecasts to remain coherent across different stages of community development. This stage focuses on assessing ANGEL's ability to accommodate nonstationary demand evolution and maintain continuity over time, reflecting the dynamic nature of real-world energy systems.

Finally, the validation culminates in the integrated assessment of ANGEL, examining the combined effect of causal reasoning, structural flexibility, and temporal adaptability within a unified forecasting framework. This stage evaluates whether their joint implementation ensures responsiveness to evolving data distributions while maintaining predictive stability and coherence over time. It also investigates how the learning process can incorporate new information without compromising previously acquired knowledge, aligning with the overarching objective of achieving robust and data-efficient forecasting under real-world, nonstationary conditions.

Validation is conceived as an iterative and cumulative process. Each experimental cycle refines the configuration of features, architectures, and training strategies, progressively strengthening the coherence between conceptual hypotheses and empirical evaluation. Through this structured process, the El Espino case study serves as the empirical testbed where theoretical assumptions are systematically examined, ensuring that each methodological component contributes transparently to the overall design and integrity of ANGEL.

4.7 Practical Considerations

Beyond methodological rigor and empirical validation, the forecasting models for MGs in remote areas must take into account a number of practical considerations that determine their long-term impact and usability. These include reproducibility, which ensures that scientific progress can be verified and extended, as well as the broader implications that arise when such models are applied to community safety and energy resilience. In addition, ethical and policy-related aspects play an equally important role, as forecasting outcomes directly influence investment, planning, and the daily lives of people in remote areas [161].

Reproducibility is a cornerstone of robust scientific practice and an essential condition for building cumulative knowledge [48]. For this reason, all models developed in this thesis were implemented using open-source frameworks, specifically PyTorch, which guarantees accessibility. This choice not only reduces dependency on proprietary software but also ensures that other researchers and practitioners can replicate the architectures without restrictions. Similarly, the datasets used for training and evaluation, including the El Espino case study, were carefully curated and preprocessed. Metadata describing preprocessing steps, variable definitions, missing-value treatments, and the rationale behind the division into training, validation, and testing subsets were explicitly recorded to facilitate replication. The study enables independent verification of results and provides a foundation upon which others can build. This not only increases trust in the findings but also fosters collaboration, which is particularly important for advancing data-driven electrification initiatives in remote areas.

At the same time, forecasting electricity demand in MGs cannot be treated as a purely academic exercise. MGs are critical infrastructures that sustain basic services such as healthcare, education, water supply, and communication in remote communities. In these contexts, forecasting errors are not neutral: underestimation can lead to frequent outages, service interruptions, and the erosion of community trust in the electrification project, while overestimation risks unnecessary Capital Expenditure (CAPEX) and the inefficient use of financial resources [117]. The models developed in this thesis therefore carry direct implications for security and resilience. Low-error forecasting contributes to ensuring that hospitals maintain reliable power for medical equipment, that refrigeration for vaccines or food storage remains stable, and that schools, telecommunication systems, and water pumping facilities can operate without disruption. In short, forecasting quality becomes a determinant of both technical sustainability and human well-being.

Moreover, MGs are increasingly embedded within broader strategies for climate resilience and sustainable development. Their capacity to integrate renewable resources and reduce reliance on fossil fuels depends on forecasts that can anticipate variability, adoption dynamics, and disruptive events. If forecasts fail to capture these dynamics, communities may experience energy insecurity precisely at moments of greatest vulnerability, such as during extreme weather events, seasonal agricultural activities, or sudden increases in appliance adoption. By integrating causality, flexibility, and adaptability, ANGEL seeks to minimize such risks, providing low error forecasts. In this sense, demand forecasting becomes not only a technical exercise but also a governance tool for managing uncertainty in the transition to sustainable energy systems [19].

A further dimension concerns ethical and policy considerations. Forecasting models play a decisive role in shaping electrification strategies and guiding infrastructure investments. Ensuring their transparency and accountability is therefore essential, as decisions derived from opaque or poorly validated models can lead to inefficient resource allocation or misguided planning. Dissemination of models and datasets reinforces this ethical commitment by promoting inclusiveness, enabling local institutions, researchers, and Non-governmental Organizations (NGOs) to access, verify, and adapt forecasting tools to the specific contexts, thereby strengthening collective capacity for sustainable electrification [154].

Practical considerations also extend to deployment. Models that require only modest computational resources, as demonstrated in the experimental setup, are more feasible for application in real-world contexts where advanced hardware such as GPUs or clusters is unavailable. This scalability ensures that forecasting tools can be deployed not only by well-funded institutions but also by local utilities, cooperatives, and community organizations operating in resource-constrained environments. By designing models that are lightweight and efficient, the research contributes to creating technological solutions that are not only scientifically innovative but also socially equitable, economically viable, and operationally sustainable.

The practical considerations addressed in this thesis highlight the need to align technical innovation with principles of reliability, transparency, and ethical responsibility. Forecasting for MGs in remote areas carries direct social and infrastructural implications, as forecasting errors influence how energy systems are planned, deployed, and maintained. Emphasizing openness in research practice fosters trust and collaboration across scientific and operational communities, while awareness of safety and resilience implications situates forecasting within the broader mission of sustainable energy planning. Grounded in causality, flexibility, and adaptability, ANGEL aims to combine technical efficiency with social relevance, supporting the equitable and resilient expansion of clean energy access.

Chapter 5

Experimental Validation

The goal of ML should not be to eliminate humans from the loop, but to make humans more effective decision makers by revealing the causal structure of their environment.

Judea Pearl, computer scientist and Turing Award laureate

Chapter Overview

Building on the research design established in Chapter 4, this chapter presents the empirical validation of ANGEL. Using data from the El Espino MG, it assesses how successive methodological principles of causality, flexibility, latent temporal evolution, and adaptability enhance forecasting robustness under conditions of data scarcity, fragmentation, nonlinear socio-technical dependencies, and nonstationary demand evolution. The validation is organized through a series of experiments that progressively examine and integrate these dimensions, from causal feature engineering to adaptive learning, to demonstrate their individual and cumulative contributions to model performance and reliability.

The results show that each methodological dimension strengthens specific aspects of the forecasting process. Collectively, these findings establish the empirical foundation of ANGEL and demonstrate how methodological progression from isolated principles to their integration supports reliable forecasting in evolving MGs.

5.1 Data and Variable Preparation

As discussed in Section 4.3, this research builds on the El Espino dataset, publicly available on GitHub [13]. The dataset spans the period from January 1, 2016, to July 31, 2017, comprising 13,872 hourly measurements collected from a MG serving 128 households, a hospital, a school, and street-lighting systems. Each record contains wattage values with a 5-minute resolution, representing one of the few examples of high-frequency monitoring in MGs located in remote areas. Beyond its completeness, the dataset provides a unique testbed for evaluating forecasting methods under data-constrained conditions characterized by structural data scarcity and fragmentation, nonlinear socio-technical and environmental dependencies, and nonstationary demand evolution.

To gain a preliminary understanding of the demand structure, several graphical representations were generated. Figures 5.1 and 5.2 illustrate the distribution of demand across hours of the day and months of the year, respectively, with outliers marked as dots [119]. The hourly distribution highlights clear daily cycles, with a minimum around 8 a.m. and a pronounced evening peak at 8 p.m., reflecting household routines and communal activities. The monthly distribution, by contrast, reveals seasonal influences, with October showing the highest recorded demand. These visualizations confirm the importance of temporal metadata in capturing electricity demand dynamics.

Beyond static distributions, heatmaps were employed to capture the joint effects of temporal variables. Figure 5.3 presents average demand values during weekends across hours and months, highlighting the interaction between daily and seasonal cycles [119]. Demand peaks were observed in the evenings, with clear seasonal variations attributable to heating or cooling needs. Weekends also showed distinct patterns compared to weekdays, reflecting differences in social and economic

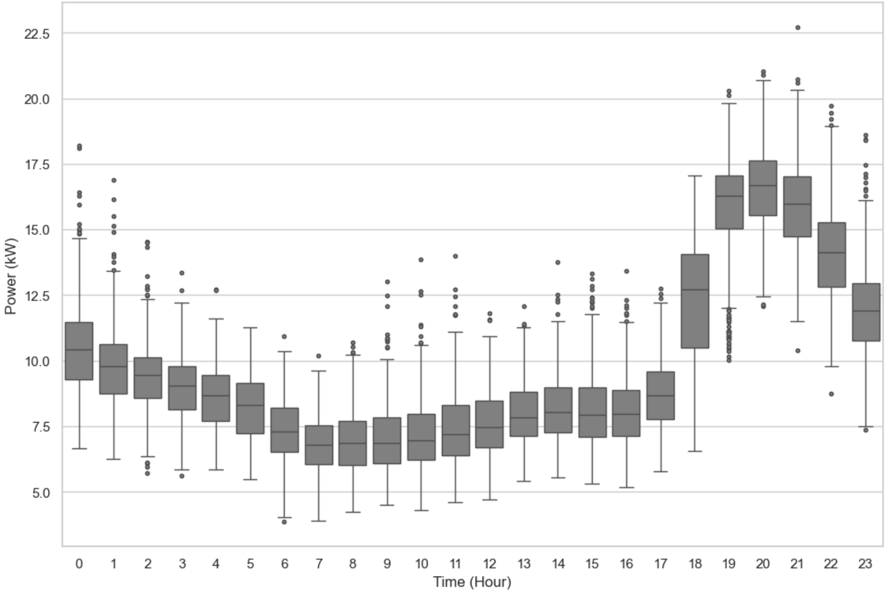


Figure 5.1: Hourly distribution of electricity demand in El Espino, represented as a boxplot highlighting variability and outliers.

activity.

Figure 5.4 complements this analysis by plotting the double standard deviation of demand across the same dimensions, which yields lower variability than single-standard-deviation plots [119]. This confirms the importance of including metadata such as hour of the day, day type, and month in the model training process: incorporating these features reduces unexplained variability and enhances the reliability of demand estimation.

These findings informed the feature selection adopted in the subsequent experiments described in Section 4.4, which includes temperature, humidity, hour of the day, month of the year, and day type (weekday or weekend) as key inputs for model development and validation.

Data preprocessing included the detection and removal of anomalous values in the

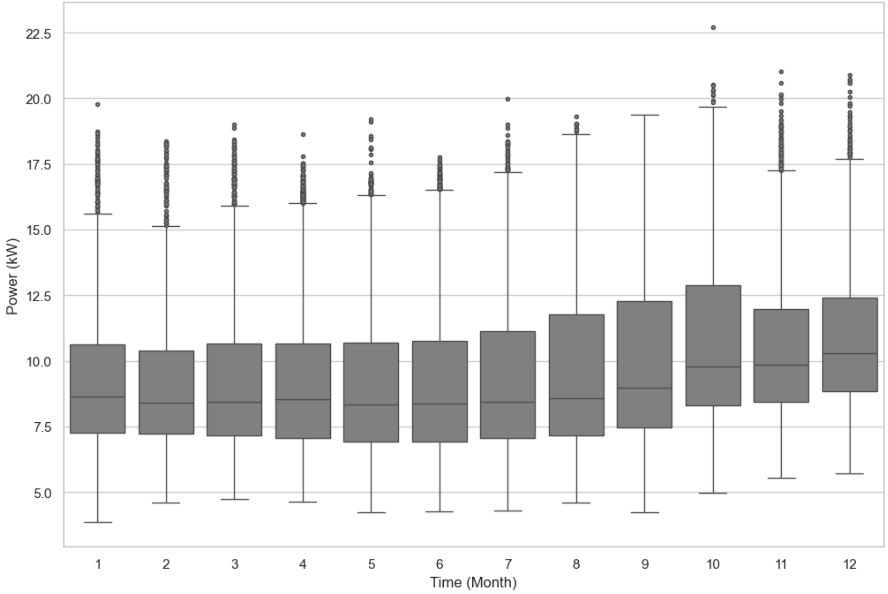
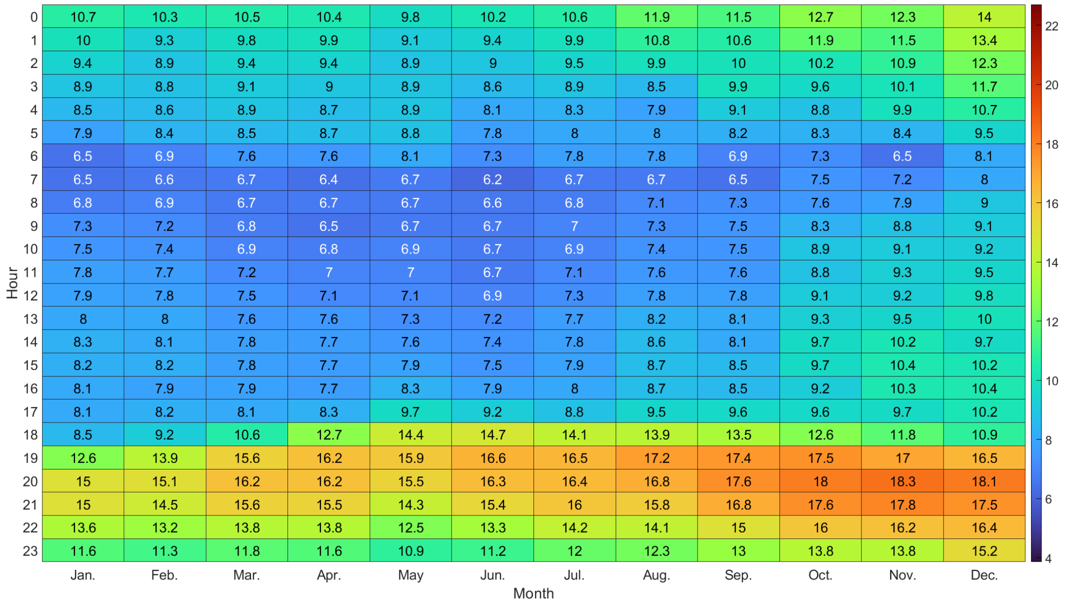


Figure 5.2: Boxplot of the monthly electricity demand distribution in El Espino, including outliers.

time series. Outliers may occur due to sensor malfunctions, temporary disconnections, or atypical events that do not reflect regular demand behavior. Their presence risks distorting the training of learning-based models, particularly in data-scarce contexts where each datapoint carries significant weight. Outliers were identified using the three-standard-deviation rule, as expressed in Equation 5.1:

$$x < \mu - 3\sigma \quad \text{or} \quad x > \mu + 3\sigma, \tag{5.1}$$

where μ represents the mean of the demand distribution and σ its standard deviation. This statistical criterion assumes that demand values approximately follow a normal distribution and classifies as anomalies any observations that deviate by more than three standard deviations from the mean [72]. By filtering values that exceed this threshold, the resulting dataset reduces the risk of unstable training dynamics such as exploding gradients, thereby ensuring a smoother and more



Hour	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
0	10.7	10.3	10.5	10.4	9.8	10.2	10.6	11.9	11.5	12.7	12.3	14
1	10	9.3	9.8	9.9	9.1	9.4	9.9	10.8	10.6	11.9	11.5	13.4
2	9.4	8.9	9.4	9.4	8.9	9	9.5	9.9	10	10.2	10.9	12.3
3	8.9	8.8	9.1	9	8.9	8.6	8.9	8.5	9.9	9.6	10.1	11.7
4	8.5	8.6	8.9	8.7	8.9	8.1	8.3	7.9	9.1	8.8	9.9	10.7
5	7.9	8.4	8.5	8.7	8.8	7.8	8	8	8.2	8.3	8.4	9.5
6	6.5	6.9	7.6	7.6	8.1	7.3	7.8	7.8	6.9	7.3	6.5	8.1
7	6.5	6.6	6.7	6.4	6.7	6.2	6.7	6.7	6.5	7.5	7.2	8
8	6.8	6.9	6.7	6.7	6.7	6.6	6.8	7.1	7.3	7.6	7.9	9
9	7.3	7.2	6.8	6.5	6.7	6.7	7	7.3	7.5	8.3	8.8	9.1
10	7.5	7.4	6.9	6.8	6.9	6.7	6.9	7.4	7.5	8.9	9.1	9.2
11	7.8	7.7	7.2	7	7	6.7	7.1	7.6	7.6	8.8	9.3	9.5
12	7.9	7.8	7.5	7.1	7.1	6.9	7.3	7.8	7.8	9.1	9.2	9.8
13	8	8	7.6	7.6	7.3	7.2	7.7	8.2	8.1	9.3	9.5	10
14	8.3	8.1	7.8	7.7	7.6	7.4	7.8	8.6	8.1	9.7	10.2	9.7
15	8.2	8.2	7.8	7.7	7.9	7.5	7.9	8.7	8.5	9.7	10.4	10.2
16	8.1	7.9	7.9	7.7	8.3	7.9	8	8.7	8.5	9.2	10.3	10.4
17	8.1	8.2	8.1	8.3	9.7	9.2	8.8	9.5	9.6	9.6	9.7	10.2
18	8.5	9.2	10.6	12.7	14.4	14.7	14.1	13.9	13.5	12.6	11.8	10.9
19	12.6	13.9	15.6	16.2	15.9	16.6	16.5	17.2	17.4	17.5	17	16.5
20	15	15.1	16.2	16.2	15.5	16.3	16.4	16.8	17.6	18	18.3	18.1
21	15	14.5	15.6	15.5	14.3	15.4	16	15.8	16.8	17.6	17.8	17.5
22	13.6	13.2	13.8	13.8	12.5	13.3	14.2	14.1	15	16	16.2	16.4
23	11.6	11.3	11.8	11.6	10.9	11.2	12	12.3	13	13.8	13.8	15.2

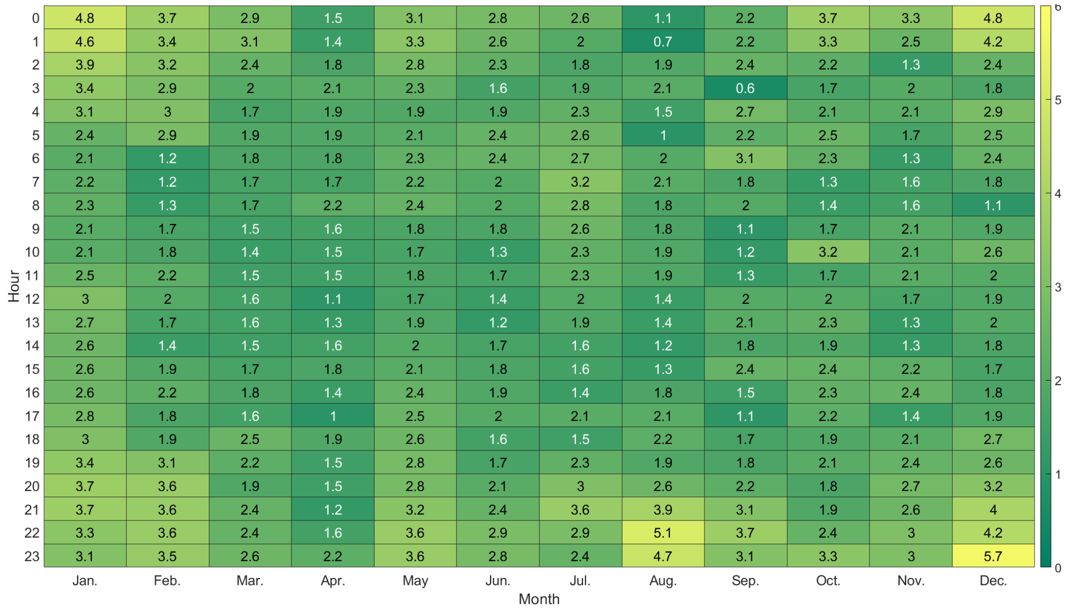
Figure 5.3: Heatmap of mean electricity demand in El Espino on weekends, showing variation across hours of the day and months of the year.

efficient learning process.

To prepare the data for neural network training, normalization was applied to all input variables. Electricity demand values were scaled using min–max normalization to the range $[0, 1]$, ensuring that all inputs contributed proportionally during training. This transformation, expressed in Equation 5.2, rescales each value based on the minimum and maximum observed in the training set:

$$x_{\text{scaled}} = \frac{x - \min(x_{\text{train}})}{\max(x_{\text{train}}) - \min(x_{\text{train}})}. \tag{5.2}$$

Although alternative scaling techniques such as z-score standardization are commonly used, min–max normalization was selected in this case because it preserves the original distribution shape and facilitates model convergence given the bounded activation functions used in the network [100]. Temporal variables such



Hour	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
0	4.8	3.7	2.9	1.5	3.1	2.8	2.6	1.1	2.2	3.7	3.3	4.8
1	4.6	3.4	3.1	1.4	3.3	2.6	2	0.7	2.2	3.3	2.5	4.2
2	3.9	3.2	2.4	1.8	2.8	2.3	1.8	1.9	2.4	2.2	1.3	2.4
3	3.4	2.9	2	2.1	2.3	1.6	1.9	2.1	0.6	1.7	2	1.8
4	3.1	3	1.7	1.9	1.9	1.9	2.3	1.5	2.7	2.1	2.1	2.9
5	2.4	2.9	1.9	1.9	2.1	2.4	2.6	1	2.2	2.5	1.7	2.5
6	2.1	1.2	1.8	1.8	2.3	2.4	2.7	2	3.1	2.3	1.3	2.4
7	2.2	1.2	1.7	1.7	2.2	2	3.2	2.1	1.8	1.3	1.6	1.8
8	2.3	1.3	1.7	2.2	2.4	2	2.8	1.8	2	1.4	1.6	1.1
9	2.1	1.7	1.5	1.6	1.8	1.8	2.6	1.8	1.1	1.7	2.1	1.9
10	2.1	1.8	1.4	1.5	1.7	1.3	2.3	1.9	1.2	3.2	2.1	2.6
11	2.5	2.2	1.5	1.5	1.8	1.7	2.3	1.9	1.3	1.7	2.1	2
12	3	2	1.6	1.1	1.7	1.4	2	1.4	2	2	1.7	1.9
13	2.7	1.7	1.6	1.3	1.9	1.2	1.9	1.4	2.1	2.3	1.3	2
14	2.6	1.4	1.5	1.6	2	1.7	1.6	1.2	1.8	1.9	1.3	1.8
15	2.6	1.9	1.7	1.8	2.1	1.8	1.6	1.3	2.4	2.4	2.2	1.7
16	2.6	2.2	1.8	1.4	2.4	1.9	1.4	1.8	1.5	2.3	2.4	1.8
17	2.8	1.8	1.6	1	2.5	2	2.1	2.1	1.1	2.2	1.4	1.9
18	3	1.9	2.5	1.9	2.6	1.6	1.5	2.2	1.7	1.9	2.1	2.7
19	3.4	3.1	2.2	1.5	2.8	1.7	2.3	1.9	1.8	2.1	2.4	2.6
20	3.7	3.6	1.9	1.5	2.8	2.1	3	2.6	2.2	1.8	2.7	3.2
21	3.7	3.6	2.4	1.2	3.2	2.4	3.6	3.9	3.1	1.9	2.6	4
22	3.3	3.6	2.4	1.6	3.6	2.9	2.9	5.1	3.7	2.4	3	4.2
23	3.1	3.5	2.6	2.2	3.6	2.8	2.4	4.7	3.1	3.3	3	5.7

Figure 5.4: Heatmap of the double standard deviation of weekend electricity demand in El Espino, highlighting variability across hours and months.

as hour of the day and month of the year were encoded using sine and cosine transformations. This cyclical encoding preserves continuity across boundaries (e.g., between 23:00 and 00:00 hours or between December and January), preventing artificial discontinuities that would otherwise hinder model learning [82, 120].

A total of 13,872 hourly demand measurements from El Espino formed the basis of the analysis and were partitioned into training (64%), validation (16%), and testing (20%) subsets, corresponding to 8,878, 2,220, and 2,774 hours, respectively. This partition supports balanced model evaluation and prevents overfitting, following standard deep learning practice [45].

All subsequent experiments (Sections 5.2–5.5) build directly on this prepared dataset and are based on the author's prior works [119, 120].

5.2 H1 Validation: Causal Feature Engineering

This section validates Hypothesis H1, which posits that observation grounded in causal reasoning enhances robustness under data scarcity and fragmentation. Operationally, this hypothesis tests whether causal feature engineering, the deliberate construction of input variables reflecting stable socio-technical and environmental relationships, improves forecasting reliability in data-constrained contexts.

Purely statistical models often overfit transient correlations that fail to generalize beyond the training domain, especially under data scarcity and fragmentation. In contrast, features derived from plausible causal mechanisms, such as the influence of temperature and time, introduce structural regularities that persist across contexts. Embedding these relations within the model's architecture is therefore expected to strengthen robustness and generalization, two qualities essential for energy planning under uncertainty.

To verify whether causally informed variables contribute statistically significant explanatory power, a one-way Analysis of Variance (ANOVA) was performed. This test quantifies how much of the variance in electricity demand can be attributed to specific input factors and their interactions relative to random noise, thereby identifying which drivers exert a stable influence on the target variable. The ANOVA model [24] is expressed in Equation 5.3:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij}, \tag{5.3}$$

where Y_{ij} is the observed electricity demand for the i-th and j-th factor levels, μ is the overall mean, α_i and β_j represent the main effects, γ_{ij} the interaction effect, and ϵ_{ij} the residual error term. The F-value measures the ratio of explained to unexplained variance, while the corresponding p-value assesses statistical significance at a 5% threshold.

Table 5.1 summarizes the results. Variables such as *Month* and *Hour* show extremely low p-values ($p < 2 \times 10^{-16}$), confirming pronounced seasonal and diurnal dependencies. Interaction terms, including *Month:Hour* and *Month:Weekend*, are also significant, highlighting that demand variability emerges from the combined influence of climatic conditions and human routines. These findings empirically demonstrate that causally grounded features capture meaningful structure in the data, providing a robust foundation for subsequent forecasting experiments.

Moreover, the ANOVA results reinforce the significance of key causal drivers, particularly *Month*, *Hour*, and *Temperature*, which consistently explain a large share of variance in demand. Higher-order interactions, such as *Month* × *Weekend* × *Temperature* × *Humidity*, were not significant, underscoring that flexible neural architectures are essential to capture deeper dependencies among socio-technical and environmental factors.

Building on this statistical evidence, three neural architectures were compared to evaluate the robustness contribution of causal reasoning: a baseline model without metadata enrichment and two causally informed configurations (MLP and DFNN). The baseline consisted of a simple triangular network trained solely on demand data, representing a purely statistical approach. Its performance, summarized in Table 5.2, indicates that the ADAM optimizer achieved the lowest error (MSE = 0.1339; MAE = 0.1345), outperforming ADAGRAD, ADADELTA, and SGD.

All enriched models developed in this research build upon a shared initial input–output structure. The input layer includes seven nodes corresponding to the explanatory variables defined in Section 5.1: temperature, humidity, month and hour (both encoded using sine–cosine transformations), and a weekend indicator. The output layer represents the normalized hourly electricity demand per person, which is later rescaled to its actual unit (kW/person).

Table 5.1: Statistical Significance of Variables and Interactions.

Variable	**F-value**	**p-value**
Month	379.251	$< 2 \times 10^{-16}$
Weekend	44.579	2.55×10^{-11}
Hour	4533.618	$< 2 \times 10^{-16}$
Temperature	875.502	$< 2 \times 10^{-16}$
Humidity	20.656	5.55×10^{-6}
Month:Weekend	8.320	9.70×10^{-15}
Month:Hour	19.389	$< 2 \times 10^{-16}$
Weekend:Hour	5.756	$< 2 \times 10^{-16}$
Month:Temperature	14.013	$< 2 \times 10^{-16}$
Weekend:Temperature	1.384	0.239
Hour:Temperature	19.521	$< 2 \times 10^{-16}$
Month:Humidity	39.199	$< 2 \times 10^{-16}$
Weekend:Humidity	13.819	0.000202
Hour:Humidity	6.461	$< 2 \times 10^{-16}$
Temperature:Humidity	0.316	0.573
Month:Weekend:Hour	0.987	0.548
Month:Weekend:Temperature	6.417	1.02×10^{-10}
Month:Hour:Temperature	1.510	4.01×10^{-7}
Weekend:Hour:Temperature	0.811	0.721
Month:Weekend:Humidity	4.846	1.64×10^{-7}
Month:Hour:Humidity	1.249	0.004701
Weekend:Hour:Humidity	0.263	0.999
Month:Temperature:Humidity	7.281	1.58×10^{-12}
Weekend:Temperature:Humidity	17.559	2.81×10^{-5}
Hour:Temperature:Humidity	1.359	0.116
Month:Weekend:Hour:Temperature	0.719	0.999
Month:Weekend:Hour:Humidity	0.604	1.000
Month:Weekend:Temperature:Humidity	4.961	9.70×10^{-8}
Month:Hour:Temperature:Humidity	1.040	0.320
Weekend:Hour:Temperature:Humidity	1.543	0.0467
Month:Weekend:Hour:Temperature:Humidity	0.811	0.987

Optimizer	MAE	MSE
ADAM	0.1345	0.1339
ADAGRAD	0.1818	0.1744
ADADELTA	0.1364	0.1962
SGD	0.1327	0.1403

Table 5.2: Performance of the baseline (triangular) architecture without causal enrichment.

The DFNN adopted a hierarchical architecture with five hidden layers (50, 250, 750, 300, and 150 neurons) using ReLU activations to capture complex nonlinearities, whereas the MLP comprised two hidden layers (50 and 150 neurons) with sigmoid and ReLU activations. Training was performed with a learning rate of 0.01 over 20 epochs, applying early stopping based on validation loss to prevent overfitting.

Figures 5.5 and 5.6 illustrate the implemented architectures. Under identical conditions, the MLP achieved an MAE of 0.1044 and the DFNN of 0.1144, both outperforming the baseline's best result (MAE = 0.1345). This reduction in error confirms that embedding causally informed features enhances predictive reliability under scarce and fragmented data.

These results validate H1 from complementary perspectives: the statistical analysis confirms that causally derived features possess explanatory power, while the predictive evaluation demonstrates that their inclusion yields lower forecasting error and improved robustness under data scarcity and fragmentation. Causal reasoning thus provides not only theoretical coherence but also measurable benefits in practice, establishing a stable foundation for subsequent investigations into representational flexibility (H2).

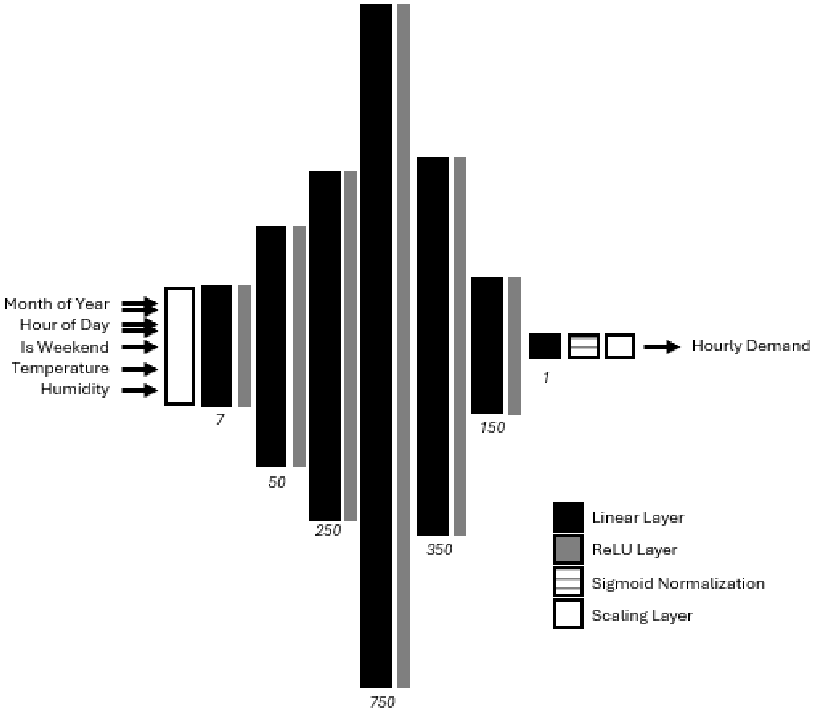


Figure 5.5: DFNN architecture with multiple hidden layers for electricity demand modeling.

5.3 H2 Validation: Flexible Representation

This section validates H2, which posits that representation through flexible neural network modeling improves the capture of nonlinear dependencies and reduces forecasting error. Building on H1, which demonstrated that embedding causal metadata enhances model robustness, the present validation examines whether the representational flexibility of neural architectures further improves the capture of nonlinear dependencies and reduces forecasting error. The central question is not only whether different architectures achieve lower error but also whether they embody distinct strategies of learning. In this regard, the MLP captures patterns through static activations, the DFNN leverages network depth to represent more

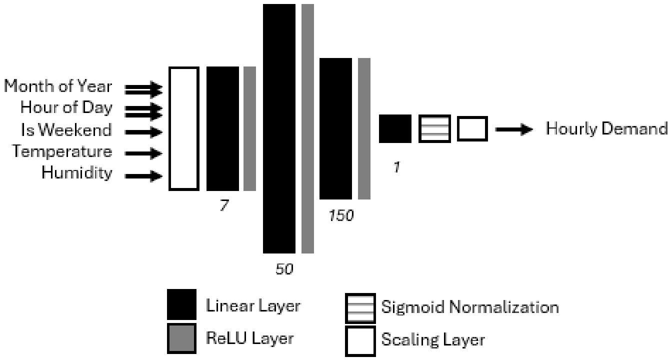


Figure 5.6: MLP architecture as a classical baseline with multiple hidden layers for electricity demand modeling.

complex mappings, and the KAN restructures the functional representation of data itself. The validation of H2 therefore examines how architectural choices affect both forecasting error and computational efficiency, particularly in low-data and high-variability environments such as MGs.

Model development involved several experimental iterations to identify the most suitable configuration. Different combinations of layer sizes, activation functions, and learning rates were tested to optimize convergence stability and generalization error, improving the model's ability to capture the nonlinear dependencies and contextual patterns within the El Espino dataset.

In Section 5.2, the MLP serves as the baseline architecture, characterized by static activation functions and limited structural flexibility. The DFNN, also introduced in that section, extends this design by increasing network depth to capture higher-order nonlinearities, though at the expense of higher computational cost. In contrast, the KAN architecture, illustrated in Figure 5.7, represents a fundamentally different modeling approach based on the Kolmogorov–Arnold Representation Theorem (Equation 5.4) and implemented through adaptive B-spline basis functions (Equations 5.5a–5.5b).

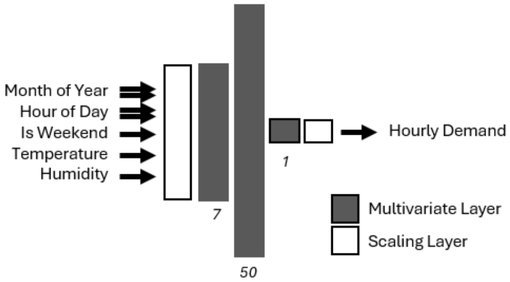


Figure 5.7: KAN architecture implementing Kolmogorov–Arnold decomposition with B-spline functions.

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{2n+1} g_i \left(\sum_{j=1}^{n} h_{ij}(x_j) \right), \tag{5.4}$$

$$B_{i,0}(x) = \begin{cases} 1 & \text{if } t_i \leq x < t_{i+1}, \\ 0 & \text{otherwise,} \end{cases} \tag{5.5a}$$

$$B_{i,k}(x) = \frac{x - t_i}{t_{i+k} - t_i} B_{i,k-1}(x) + \frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}} B_{i+1,k-1}(x), \tag{5.5b}$$

Unlike conventional neural networks that rely on fixed nonlinear activation functions such as sigmoid or ReLU, the KAN employs learnable, data-dependent activations that evolve during training, enabling the network to determine how nonlinearity should manifest based on the input structure [76]. This flexibility aligns with the methodological goal of this doctoral work, to move beyond correlation fitting toward architectures that embody structural adaptability. By combining functional decomposition with localized, adaptive activations, the KAN achieves a balance between expressive power and generalization efficiency, a combination that traditional ANNs often struggle to realize [160].

In comparative experiments, all architectures were trained using the ADAM optimizer, which provided the best trade-off between convergence and generalization. Under this configuration, the KAN achieved a MAE of 0.0537, markedly outperforming both the MLP (MAE = 0.1044) and the DFNN (MAE = 0.1144). This represents nearly a 50% reduction in error compared to the best prior model, demonstrating the effectiveness of adaptive functional representation in reducing forecasting error and capturing nonlinear, context-dependent dependencies.

These results empirically validate **H2**, confirming that neural architectures with flexible representational mechanisms achieve superior performance. The KAN's adaptive B-spline activations enable it to dynamically adjust to the structure of the input space, allowing the model to learn non-additive and context-dependent relationships that static or depth-based architectures fail to capture. This capacity directly accounts for the observed improvements in error stability and forecasting robustness in heterogeneous MG environments.

The validation of H2 therefore extends beyond numerical performance to reveal broader methodological implications: forecasting performance in low-data, nonlinear contexts depends not on network depth or data volume, but on structural adaptability, a principle that complements the causal robustness established in H1. Within the broader framework of this doctoral research, these findings confirm that architectural innovation, rather than data abundance alone, is the cornerstone of resilient forecasting, bridging causal reasoning and flexible representation for sustainable energy planning.

5.4 H3 Validation: Latent Temporal Evolution

This section validates H3, which posits that evolution informed by adaptive adjustment ensures coherence under nonstationary demand evolution. The validation focuses on the Degree of Adoption (DoA) parameter, which quantifies the maturity of electricity demand within a community and captures the progressive dynamics of electrification. Unlike exogenous metadata such as temperature or humidity that describe environmental influences, the DoA represents an endogenous and temporally evolving feature that reflects the internal trajectory of a community's transition from restricted to mature electricity use [3]. Its purpose is to model the behavioral and technological evolution of demand as households progressively acquire new appliances and integrate electricity more deeply into daily life. While H1 demonstrated that causal feature engineering enhances robustness through stable environmental and temporal drivers, the present validation focuses on the structural evolution of demand itself. In this sense, the DoA constitutes not merely an additional input but a conceptual advance in how electricity demand is represented and forecasted.

Importantly, the DoA was not predefined but emerged inductively from the causal forecasting analysis. It originated from behavioral patterns unexplained by environmental or temporal variables, revealing a latent dimension of adoption maturity that influences demand evolution. This discovery underscores the exploratory nature of the research, in which causal reasoning guided not only model design but also the identification of new explanatory mechanisms.

The theoretical foundation of the DoA derives from the literature on technology adoption and diffusion of innovations. Demand growth in newly electrified communities rarely follows a linear path; rather, it evolves along S-shaped diffusion curves [129]. In the early phase, adoption accelerates as electricity access expands and essential appliances are acquired. Growth then slows as saturation is approached, limited by income, infrastructure, and behavioral routines. To operationalize this behavior, the DoA was modeled using a logarithmic adoption equation (Equation 5.6), which maps elapsed time since electrification to an adoption index $a(t)$ bounded between 0 and 1:

$$a(t) = 10^{m \cdot \log(t) + n}, \tag{5.6}$$

where t is the time elapsed since electrification, m and n are empirically estimated parameters, and $a(t)$ denotes the DoA. For the El Espino dataset, calibration yielded m=0.1253 and n=-0.1143, effectively reproducing the observed trajectory from restricted, low-demand usage toward mature and diversified consumption. The resulting curve captures the initial acceleration of electricity use and its gradual deceleration as the system stabilizes, reflecting both theoretical precedent and empirical observation.

Experimental validation compared forecasting models trained with and without the inclusion of the DoA. The tested architectures (DFNN, MLP, and KAN), shown in Figures 5.8, 5.9, and 5.10, respectively, incorporate the DoA as an additional explanatory input variable compared with the previous experiments.

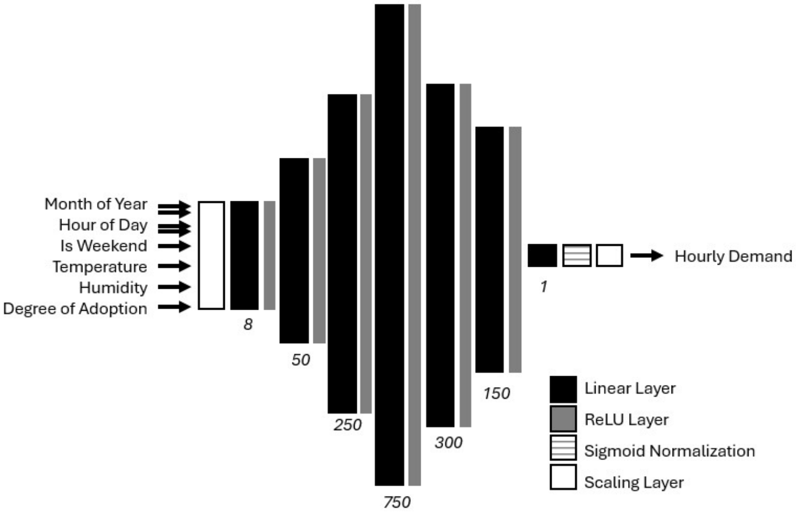


Figure 5.8: DFNN architecture with multiple hidden layers for electricity demand modeling, including the DoA variable as an additional input feature.

Figure 5.11 presents the results from 20 experiments conducted for each ANN architecture, DFNN, MLP, and KAN. Each experiment involved training the respective ANN for up to 20 epochs. The figure illustrates the relationship between training duration and mean error across the 20 experiments, with key data points

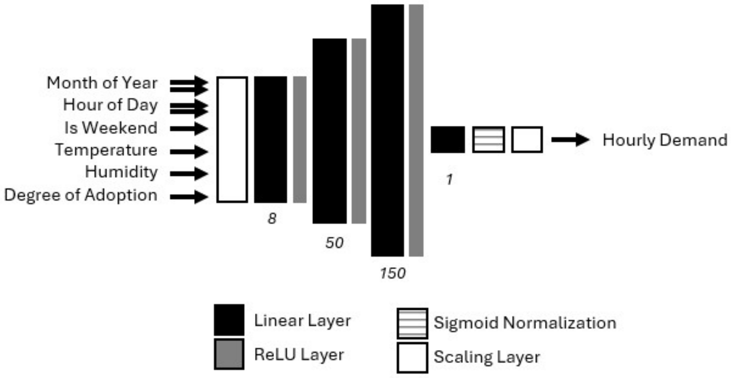


Figure 5.9: MLP architecture as a classical baseline with multiple hidden layers for electricity demand modeling, incorporating the DoA variable among the input factors.

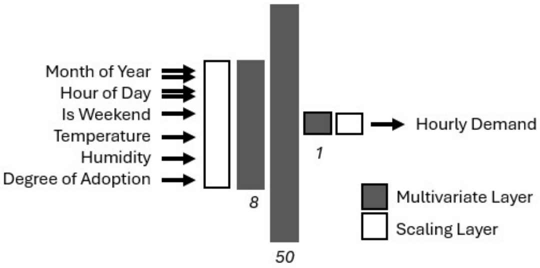


Figure 5.10: KAN architecture implementing Kolmogorov–Arnold decomposition with B-spline functions, extended to include the DoA variable as a dynamic explanatory input.

corresponding to specific epochs (1, 5, 10, 15, and 20). These points reveal consistent trends in the models' convergence behavior, highlighting differences in both efficiency and error.

The stopping criteria in neural network training can be defined in two ways: by specifying a target error and measuring the time required to reach it, or by fixing the training duration and evaluating the resulting error. In Figure 5.11, the first approach is applied, selecting the number of epochs based on the observation that the training and validation errors remained within acceptable bounds for demand

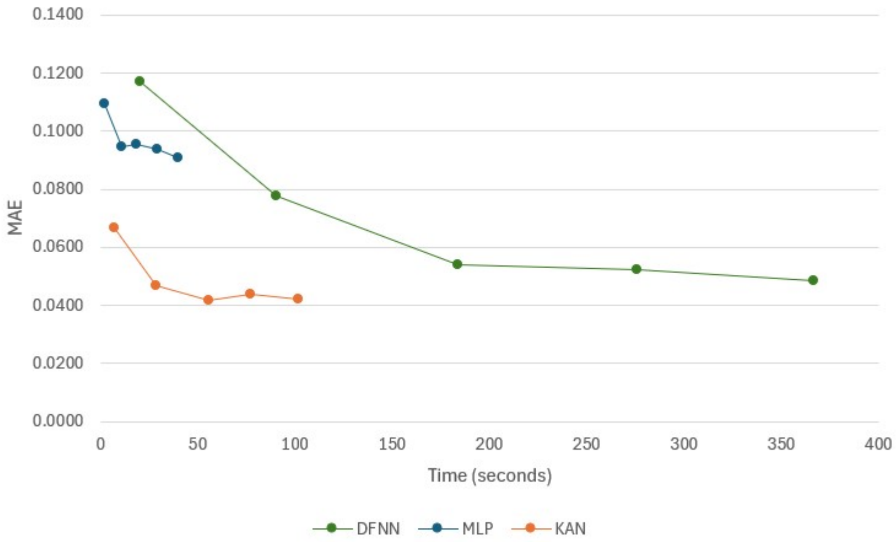


Figure 5.11: Comparison of ANN architectures: MAE and training time over 20 epochs.

estimation. This strategy minimizes error growth while avoiding unnecessary computational overhead. Conversely, Figure 5.12 illustrates the second criterion, where training time is limited to a maximum of 20 minutes to constrain overfitting and evaluate the model's error stability under restricted computational conditions.

Figure 5.11 demonstrates that the KAN achieves the lowest error, reaching 0.042 in under two minutes, significantly outperforming the other architectures in both speed and precision. The MLP converges faster but with a higher error (0.09), failing to match the low error level of the KAN, which reduces the error by nearly 54% compared to the MLP. The DFNN requires over six minutes to achieve an error of 0.049, approximately 17% higher than the KAN, demonstrating that increasing network depth does not necessarily improve forecasting performance. Figure 5.12 reinforces these observations, showing that even with extended training, the KAN remains the most efficient architecture, stabilizing around an error of 0.04, whereas the DFNN and MLP plateau at 0.044–0.049 and above 0.08, respectively.

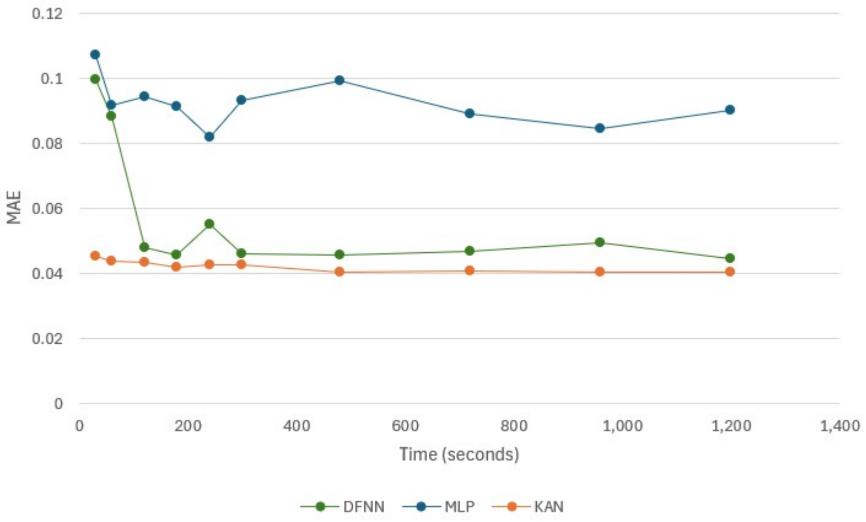


Figure 5.12: Comparison of ANN architectures: MAE and training time for up to 20 minutes.

The inclusion of the DoA significantly decreased the MAE, confirming the added value of the adoption-aware approach. With this feature, the DFNN achieved a MAE of 0.049, the MLP reached 0.090, and the KAN obtained 0.042, all improved compared to the configurations without DoA, where the best KAN result was 0.0537, followed by the MLP (0.1044) and the DFNN (0.1144). The integration of the DoA therefore enhanced the adaptability of all models and reduced forecasting error across architectures.

These results validate H3, confirming that modeling informed by adaptive and temporally evolving features enhances robustness under nonstationary demand evolution. The inclusion of the DoA demonstrated that representing latent temporal processes, such as the progressive adoption of electricity, enables models to maintain coherence as demand patterns evolve. Moreover, the superior convergence speed and reduced forecasting error achieved by the KAN indicate that flexibility in representational mechanisms, as postulated in H2, further strengthens this adaptability. These findings show that incorporating latent temporal evolution through causal and flexible modeling is essential to ensure predictive reliability in data-constrained contexts.

From a methodological perspective, the results highlight that managing nonlinearity is as crucial as representing it. The KAN adapts its nonlinear mappings to the data structure, allowing its complexity to evolve in parallel with that of the system being modeled. This adaptability, combined with adoption-aware features such as the DoA, produces forecasts that remain consistent across changing demand regimes.

5.5 H4 Validation: Adaptive Forecasting

This section validates H4, which posits that integrating causal, flexible, and adaptive principles, particularly through a CL mechanism, yields a data-efficient and robust forecasting methodology capable of maintaining predictive coherence under nonstationary demand conditions. Building upon the previous hypotheses, this validation examines whether the combined implementation of causal feature engineering (H1), flexible neural representation (H2), and adaptive adjustment (H3) produces a coherent and resilient learning framework capable of maintaining forecasting reliability under evolving demand conditions.

The experiments focus on the KAN architecture, which integrates these principles into a single architecture: causal features derived from domain-informed variables (including the DoA), flexible representation through B-spline functional decomposition, and adaptive learning mechanisms via DoA.

This configuration enables the model to capture nonlinear dependencies, maintain temporal coherence, and adapt continuously to evolving demand dynamics while minimizing computational costs. To operationalize this adaptive capability, new experiments were conducted using CL, a mechanism that incrementally updates the model as new data become available while preserving previously acquired knowledge [62]. This approach directly embodies the principle stated in H4, enabling efficient and adaptive forecasting under nonstationary conditions characteristic of evolving MGs.

Empirical results reveal a clear trade-off between computational efficiency and forecasting error, as shown in Figure 5.13. While CL substantially shortens training time, it introduces a modest increase in error compared to full training (Full Training (FT)). However, this increase is linear rather than exponential, indicating predictable and manageable degradation. Full retraining remains the most reliable option, producing the lowest MAE, but at a significantly higher computational cost.

In scenarios where responsiveness and limited resources are key, such as rural or remote MGs, CL provides a practical and efficient solution. It allows frequent model updates as new electricity demand data become available, maintaining coherence with evolving behavioral and technological conditions. The main challenge remains mitigating catastrophic forgetting [88], which can be addressed through hybrid strategies that combine CL for real-time adaptability with periodic FT for recalibration.

Two core insights emerge from the experiments. First, the KAN demonstrates that adaptive learning can be achieved without compromising stability. Its B-spline-based structure supports efficient integration of new data while preserving learned causal relationships, confirming the model's robustness under continuous updates. Second, although CL produces slightly higher error than FT, the linear increase is predictable and thus controllable. This allows operators to balance precision and responsiveness depending on operational priorities: in short-term forecasting, CL offers rapid, resource-efficient adaptability, whereas for strategic long-term analysis, FT remains preferable.

These results confirm H4 by demonstrating that the integration of causal, flexible,

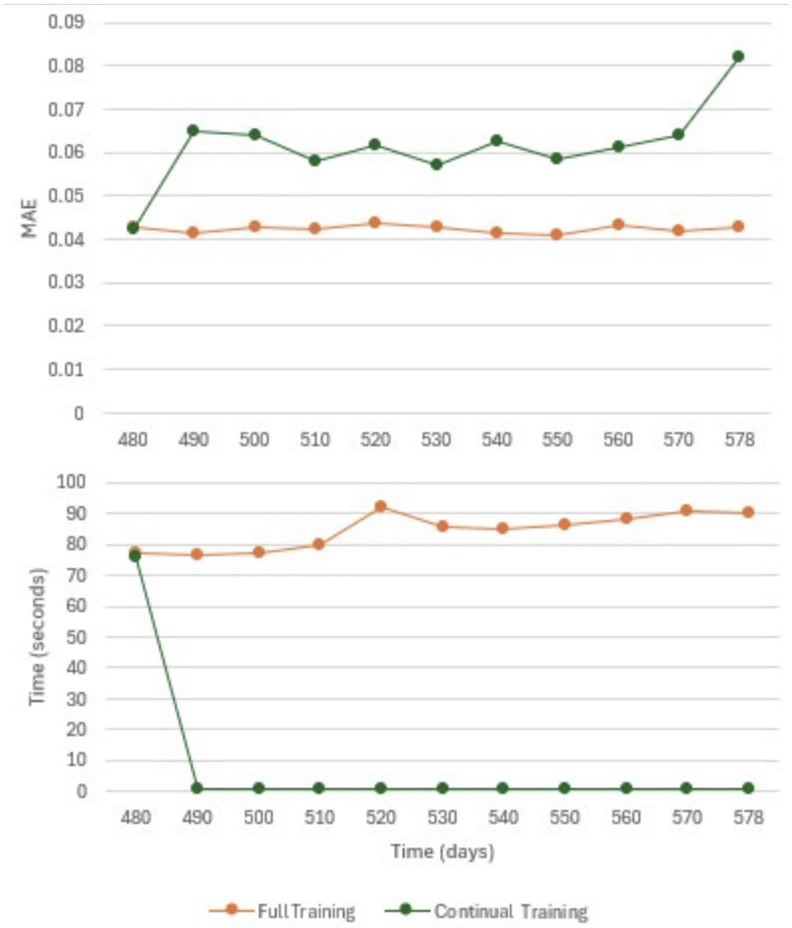


Figure 5.13: Comparison of full training vs. continual learning in KAN: mean absolute error and training time.

and adaptive principles yields a robust and computationally efficient forecasting methodology. Causal feature engineering directs learning toward physically and behaviorally meaningful dependencies; flexible representation through the KAN captures nonlinear and contextual variability; and adaptive adjustment via DoA and CL enables responsiveness to temporal evolution. Their joint implementation significantly improves training speed and scalability, even if accompanied by a minor increase in forecasting error, maintaining stability and continuity under nonstationary demand conditions.

From a methodological standpoint, this synthesis advances a holistic view of energy forecasting, where causality, flexibility, and adaptability operate as interdependent elements within a unified learning architecture. From a practical perspective, the proposed approach provides a deployable and resource-efficient solution for MGs, combining resilience and adaptability with computational parsimony. Nevertheless, further research is required to optimize the trade-off between speed and the error, ensuring that CL preserves performance consistency under more diverse operational scenarios.

5.6 Discussion and Limitations

This section discusses and critically interprets the empirical validation of the four hypotheses (H1–H4), emphasizing how the experiments contributed to developing a coherent methodological approach for forecasting electricity demand in MGs in data-constrained contexts. Beyond numerical validation, the research aimed to clarify how different modeling choices, namely causal, flexible, and adaptive, shape the reliability of data-driven forecasting in constrained environments. The experiments were designed as epistemic instruments rather than mere performance tests. Through iterative refinement, they enabled a transition from describing empirical regularities to uncovering mechanisms and integrating them within a unifying methodological structure. This process produced knowledge that is both constructive, by clarifying the relationships between variables, and operational, by translating those relationships into functional forecasting tools applicable to real systems.

Empirically, the results demonstrated that models informed by causal reasoning exhibit greater robustness under conditions of scarce and fragmented data. Incorporating physically and behaviorally meaningful variables reduced sensitivity to noise and improved stability. In turn, flexible neural representations, particularly the KAN, demonstrated the ability to capture nonlinear dependencies efficiently, achieving lower forecasting error and faster convergence than both shallow (MLP) and deeper (DFNN) architectures. Introducing temporal evolution through the DoA parameter and adaptability through CL routines enabled forecasts to remain coherent as demand evolved, ensuring responsiveness to both socio-technical and temporal dynamics.

These results indicate that methodological progress arises from the integration of complementary design principles rather than from algorithmic complexity alone. The empirical work established how each principle (causal, flexible, and adaptive) contributes to improving model reliability in data-constrained contexts. Despite their coherence, the results are bounded by several factors that define the empirical and methodological scope of the thesis.

Regarding internal validity, the available dataset, consisting of 578 days of demand data from El Espino, contains missing observations and measurement noise. Although preprocessing, causal feature selection, and regularization mitigated overfitting, residual risks remain that models captured dataset-specific patterns rather than fully invariant causal relationships.

Concerning external validity, the empirical analysis focused on a single MG, which limits generalization. Regional variations in climate, culture, policy, and economic structures could significantly alter the causal mechanisms driving demand. Extending validation to multiple locations and operational settings, such as agricultural communities in Sub-Saharan Africa or island MGs in Southeast Asia, will be essential to evaluate the framework's transferability and scalability. Differences in institutional models, tariff structures, or community governance can further influence technology adoption and consumption behavior, requiring local adaptation and re-validation of causal assumptions.

The DoA parameter, modeled as a smooth adoption curve, effectively represented gradual electrification but may overlook abrupt behavioral or infrastructural shifts caused by policy or economic shocks. Similarly, socio-technical transitions were expressed through proxy variables that simplify collective behaviors, introducing abstraction between model representation and real-world dynamics. Model optimization was heuristic, prioritizing feasibility and transparency over an exhaustive search; alternative hyperparameter configurations could further improve performance.

Finally, computational and operational constraints persist. Model calibration and evaluation were performed offline; hence, real-time performance in terms of responsiveness, latency, and stability in operational MGs remains untested. Although the KAN and CL models demonstrated efficiency and robustness, long-term deployment and scalability require empirical verification in live systems.

ANGEL achieves robust forecasting under challenging conditions, yet its conclusions are bounded by data quality, limited external validation, and modeling simplifications. Recognizing these limitations clarifies the scope of the contributions and provides a foundation for future work aimed at extending the ANGEL's scalability, adaptability, and real-world implementation.

Chapter 6

Conclusions and Future Work

We cannot solve our problems with the same thinking we used when we created them.

Albert Einstein

Chapter Overview

This chapter thus connects the empirical validation of the framework with its theoretical generalization and broader meaning. It begins by closing the empirical cycle established in Chapter 5, synthesizing the main findings and demonstrating how the integration of causal reasoning, flexible representation, and adaptive learning produced consistent reductions in forecasting error and enhanced methodological coherence. These results confirm that forecasting reliability in MGs arises from the principled combination of complementary mechanisms rather than from isolated algorithmic complexity.

The chapter then opens the theoretical cycle by interpreting these findings within a broader conceptual perspective. Electricity demand is reframed as a causal and evolving phenomenon shaped by nonlinear dependencies among socio-technical and environmental variables. This conceptualization establishes a paradigm in which forecasting is not merely a predictive task but a process of learning from dynamic causal structures under data-constrained conditions.

Finally, the chapter projects the epistemological cycle, reflecting on how this integrated approach transforms the understanding of both the phenomenon and the act of modeling it. By grounding learning in causal reasoning and adaptive mechanisms, the research contributes to a more transparent, reliable, and scientifically grounded form of AI for energy systems. The chapter concludes by acknowledging the study's limitations and outlining future directions, such as transfer learning, causal discovery, and CL, while emphasizing the broader significance of developing trustworthy, efficient, and context-aware AI to support the global transition toward sustainability.

6.1 Empirical and Theoretical Synthesis

The four research questions addressed in this thesis converge on a single conclusion: reliable electricity-demand forecasting in MGs requires the joint operation of causal reasoning, flexible representation, and adaptive learning. These three dimensions constitute the core of the ANGEL framework, a methodological architecture designed to guide neural network learning in data-constrained contexts affected by data scarcity, fragmentation, nonlinear socio-technical dependencies, and nonstationary demand evolution.

Each research question corresponds to one or more of these dimensions, linking a methodological principle to concrete empirical evidence and illustrating how the thesis connects theoretical reasoning with operational results:

- **RQ1 – (Invariant–Causality)**
  Underlying principle: Causal Feature Engineering.
  Empirical finding: Explanatory variables reflecting underlying physical and behavioral mechanisms increased model stability under scarce and fragmented data, confirming that causal reasoning enhances robustness and reduces forecasting error (H1).
- **RQ2 – (Invariant–Flexibility)**
  Underlying principle: Flexible Representation.
  Empirical finding: Adaptive architectures, particularly the KAN, efficiently captured nonlinear dependencies, achieving lower error and faster convergence than both shallow (MLP) and deep (DFNN) networks, validating H2.
- **RQ3 – (Adaptability–Causality)**
  Underlying principle: Latent Temporal Evolution.
  Empirical finding: Embedding causal structure within adaptive frameworks through the DoA parameter captured behavioral and technological evolu-

tion, maintaining temporal coherence under nonstationary demand, thereby validating H3.

- **RQ4 – (Adaptability–Flexibility)**
  Underlying principle: Integrative Adaptive Forecasting.
  Empirical finding: Incremental learning routines preserved prior knowledge while integrating new patterns, ensuring predictive stability and computational efficiency as demand evolved, validating H4.

These results validate the three pillars of ANGEL: causality, which secures robustness under scarce and fragmented data; flexibility, which captures nonlinear socio-technical relations; and adaptability, which maintains coherence as demand evolves. ANGEL thus establishes an epistemic sequence of observation, representation, and evolution, transforming data limitations into structured understanding rather than constraints.

From a theoretical standpoint, this synthesis advances three key ideas that extend beyond the empirical setting:

1. **Causality as a response to data scarcity and fragmentation.** Electricity consumption is not a random continuation of past usage but the emergent outcome of interacting environmental, temporal, and behavioural mechanisms. Embedding these mechanisms as causal processes provides stability when data are scarce or fragmented, enabling models to generalise beyond observed samples. Within ANGEL, causality anchors learning in meaningful relationships rather than correlations, ensuring robustness under incomplete information.

2. **Flexibility as a response to nonlinear dependencies.** Energy demand arises from complex interactions among socio-technical and environmental variables that are inherently nonlinear. ANGEL addresses this through flexible functional representation, particularly via the KAN, which decom-

poses multivariate relationships into adaptive, low-dimensional components. This structural flexibility enhances efficiency, allowing the model to capture nonlinear dynamics without excessive computational complexity or data requirements.

3. **Adaptability as a response to nonstationary demand evolution.** Consumption patterns evolve over time as technologies, behaviours, and external conditions change. ANGEL incorporates adaptability through CL and the modelling of adoption dynamics, allowing it to update incrementally as new information becomes available. This capacity to adjust while preserving previously learned knowledge ensures temporal coherence and resilience, transforming variability from a source of instability into a driver of sustained learning.

ANGEL stands as both an operational and theoretical model of learning in data-constrained contexts. It transforms limited evidence into coherent understanding, demonstrating that forecasting, when properly structured, becomes a process of knowledge formation rather than mere prediction. Its explanatory strength lies in integrating causal reasoning, flexible representation, and adaptive learning within a unified architecture that not only predicts but explains the dynamics of evolving energy demand. Through this integration, the framework bridges empirical evidence and theoretical insight, showing that reliability in forecasting arises from understanding the mechanisms that generate change rather than from algorithmic complexity or data abundance.

6.2 Methodological Contributions

This doctoral work translates the principles consolidated in the ANGEL framework into a concrete and reproducible forecasting pipeline for data-constrained contexts. The methodological contribution lies not in restating theoretical concepts, but in demonstrating how they can be operationalized through specific tools, procedures, and validation practices that ensure both rigor and deployability.

In line with the broader research frontier of data-efficient learning [58], which seeks to achieve high predictive performance under limited labeled data through minimal supervision and adaptive representations, this work contributes to extending data-efficient principles to the energy forecasting domain.

1. Functional decomposition for efficient nonlinearity. KANs were introduced and adapted to the energy forecasting domain to represent nonlinear dependencies with minimal computational depth. Their functional decomposition of multivariate relations into adaptive univariate components proved an effective way to capture complex socio-technical interactions while maintaining convergence speed and stability. This establishes KANs as a practical alternative to deep architectures when working with limited or fragmented data.

2. Behavioral quantification through the DoA. The DoA parameter operationalizes technology diffusion and behavioral evolution as measurable drivers of demand. Calibrated on empirical data through a logarithmic adoption model, it integrates socio-technical transitions directly into the forecasting process. This procedure bridges data-driven modeling with real-world adoption dynamics and can be generalized to other domains where usage behavior evolves over time.

3. Adaptive updating and continual learning. Adaptability was implemented through CL and incremental retraining routines that update model parameters as new data arrive, avoiding full retraining. This ensures temporal coherence and makes the forecasting system suitable for operational contexts such as MG management, where data streams are intermittent and conditions nonstationary.

4. Reproducible and transferable workflow. Every stage of the pipeline, from data preparation to validation, was automated, versioned, and documented in open-source environments using traceable workflows. Configuration files and logged metadata ensure that the models and results can be independently verified, reused, and adapted to other forecasting problems. Interoperable data formats and transparent documentation support accessibility, while the modular design facilitates adaptation across contexts and datasets. Beyond technical rigor, this commitment to openness reflects an ethical stance toward responsible innovation, enabling researchers and practitioners to examine, extend, and sustain the models as part of a broader open science ecosystem.

The methodological contribution of this research is a robust, open, and adaptive forecasting pipeline that embodies ANGEL's logic in practice. By combining functional decomposition, behavioral modeling, and CL within a transparent workflow, the thesis delivers a tangible toolset for developing data-efficient and operationally reliable forecasting systems across diverse energy contexts.

The methodological process and empirical findings underpinning this thesis have been openly disseminated through three peer-reviewed publications, each documenting a distinct phase of model development. The first, [117], explores how errors in demand estimation influence optimal MG design within the LEOPARD project. The second, [119], presents a causal neural network for electricity demand estimation in El Espino (Bolivia), highlighting the benefits of causal feature integration under incomplete data. The third, [120], introduces the DoA parameter and implements the KAN to capture nonlinear and evolving demand dynamics. Together, these works exemplify the continuity, validation, and transparency of the research pathway.

6.3 Practical and Technological Implications

The methodological framework developed in this thesis translates directly into operational, policy, and technological advances for the planning and management of future energy systems. Building on the ANGEL principles, it delivers forecasting models that combine causality, structural efficiency, and adaptive responsiveness, qualities particularly suited to MGs and other decentralized systems where minimizing forecasting error is vital for stability and sustainability.

6.3.1 Operational Implications for Microgrids

Low-error forecasting underpins the optimal sizing of distributed generation, storage capacity, and control strategies. The proposed framework enables planners and operators to move beyond correlation-based approaches, using causal and behavioral drivers to anticipate demand peaks, improve scheduling, and integrate renewable resources more effectively.

The adaptive efficiency of the implemented architectures, particularly the rapid convergence of KAN and the low cost of incremental retraining, makes real-time forecasting feasible in settings with limited processing capacity. This supports adaptive control strategies in which forecasts are continuously updated, enhancing resilience under uncertainty. Such capabilities are directly applicable to remote, islanded, or vulnerable regions, where reliable forecasting and minimal computational overhead are critical for sustainable electrification.

6.3.2 Alignment with Policy Frameworks

ANGEL is also consistent with European and global objectives for a sustainable and inclusive energy transition. By delivering adaptive and data-efficient forecasting tools, it directly supports the European Green Deal and the United Nations Sustainable Development Goals, specifically Goal 7 (Affordable and Clean

Energy), Goal 11 (Sustainable Cities and Communities), and Goal 13 (Climate Action). The framework offers a concrete blueprint for integrating explainable and adaptive forecasting modules into digital infrastructures, strengthening decision-support and governance capabilities for Europe's digital and sustainable energy transformation.

6.3.3 Scalability, Transferability, and Societal Impact

Beyond MGs, ANGEL can be scaled across higher-level systems such as RECs and PEDs. Its methodological core supports error minimization and scenario analysis in complex, data-constrained contexts.

This potential is exemplified by the project PEDAID3 [107], coordinated by the Universidad de Las Palmas de Gran Canaria (ULPGC), where the forecasting pipeline serves as a key analytical module for community-scale planning. Coupled with digital twins and agent-based simulations, this integration enables exploration of alternative policies and control strategies under real-world conditions.

Similarly, the methodological alignment with the multi-method simulation framework for RECs developed in Sanfilippo et al. [118] illustrates the framework's interoperability. There, demand forecasting operates as a key input for evaluating the technical and economic performance of REC. When embedded into hybrid simulation environments such as those in the PROBONO project [112], ANGEL enhances realism, robustness, and policy relevance.

Technologically, the research demonstrates that advanced forecasting does not require centralized or data-intensive infrastructures. Lightweight, causally grounded, and adaptively trained models can operate at the grid edge, continuously updating with new data. This supports distributed intelligence in digital energy platforms, enabling integration with simulation, decision-support, and participatory planning systems even in contexts with limited connectivity or hardware. The societal implications are equally significant. By reducing forecasting error and enhancing the resilience of demand forecasts, ANGEL helps address one of the main barriers to equitable electrification, the lack of reliable consumption data. Public authorities, development agencies, and utilities can use these tools to anticipate needs, plan infrastructure investments, and design inclusive energy policies. The explicit modeling of adoption dynamics ensures that forecasts reflect community growth trajectories and behavioral transitions, contributing to socially responsive electrification strategies.

In humanitarian or post-conflict contexts, where rapid deployment and minimal resources are vital, the models can be integrated into portable digital planning tools. By providing low-cost, and adaptive forecasting, they support timely decision-making and empower local actors to manage evolving energy needs effectively.

6.3.4 Interdisciplinary and Governance Implications

The implications of ANGEL extend beyond technology to governance and cross-disciplinary collaboration. Embedding causal reasoning in feature design bridges data science with energy engineering, while modeling adoption dynamics incorporates insights from behavioral and social sciences. This interdisciplinarity enhances the interpretability and accountability of AI-driven decision support.

By explaining not only what a model predicts but also why, the framework strengthens transparency and stakeholder trust, key prerequisites for deploying intelligent systems in socially sensitive domains. This causal and interpretable orientation aligns technical innovation with ethical and governance principles, promoting responsible and inclusive energy transitions.

6.4 Future Research

Building on the practical and technological advances discussed above, future research should extend the methodological and empirical scope of the ANGEL framework. The next stage involves consolidating its theoretical foundations, expanding its empirical validation, and enhancing its applicability across diverse energy and socio-technical contexts. These efforts aim to evolve ANGEL from a validated prototype into a generalizable, deployable architecture for adaptive and interpretable energy forecasting.

A primary direction involves expanding the diversity of data sources. Extending validation to heterogeneous geographical regions and socio-economic contexts will make it possible to test the generalizability of causal mechanisms and evaluate the adaptability of the model to different energy environments. Integrating multi-source datasets and additional temporal scales could further enhance the framework's robustness and scalability.

Another line of research concerns the refinement of behavioral and adoption dynamics. Introducing stochastic or agent-based representations of the DoA parameter would enable the modeling of discontinuous or policy-driven adoption behaviors. Incorporating richer climate and socio-economic variables could also improve the causal expressiveness of the model, supporting transferability across sectors and communities.

From a methodological perspective, optimization and uncertainty quantification deserve further exploration. Employing Bayesian, evolutionary, or multi-objective optimization methods could improve convergence, stability, and interpretability, while integrating uncertainty quantification would strengthen the transparency and reliability of forecasts.

Finally, the next step toward operational implementation is the deployment of

ANGEL in real-time environments. Testing CL under live conditions in MGs or decentralized energy systems will enable the evaluation of computational efficiency, latency, and resilience over extended periods. This transition from simulation to operation marks the path toward a new generation of ANGEL: a generalizable, data-efficient, and adaptive forecasting framework capable of supporting decision-making across heterogeneous and dynamic systems.

6.5 Final Reflection

This doctoral research has pursued a central aspiration: to demonstrate that reliable forecasting of electricity demand is achievable even under the combined challenges of data scarcity and fragmentation, nonlinear dependencies among socio-technical and environmental variables, and nonstationary demand evolution that characterize MGs in remote areas.

The work established that combining causal reasoning, flexible representation, and adaptive learning provides a coherent pathway toward this goal. The proposed ANGEL framework, structured around observation, representation, and evolution, translates these principles into practice through feature engineering, modeling, and adaptability.

Although the thesis is organized linearly, the research evolved iteratively. Each empirical insight refined data, models, and understanding, forming a feedback loop in which knowledge emerged from structuring meaning under uncertainty. This defines the essence of ANGEL: learning not from abundance, but from the ability to extract coherence where information is incomplete, acting as a mediator between what is known and what remains unknown.

This research shows that data efficiency in neural network learning does not stem from algorithmic complexity or data abundance, but from structuring learning around causal, flexible, and adaptive principles. Causality enables the model to learn why demand changes by identifying the physical, social, and environmental mechanisms that drive it. Flexibility allows it to represent nonlinear relationships, capturing how small variations in temperature or behavior can produce disproportionate effects on consumption. Adaptability ensures that the model evolves over time, incorporating new patterns as communities electrify or adopt technologies, maintaining coherence with a changing reality. Forecasting improves because the model not only anticipates demand but also explains it, transforming prediction into understanding.

Empirical validation confirms that even limited data can yield robust knowledge when guided by principled modeling. Just as meticulous astronomical observations enabled Kepler's discovery of planetary laws, carefully curated datasets here revealed the mechanisms governing energy demand.

Looking ahead, ANGEL points toward a vision of AI that learns meaningfully as well as efficiently, with models that adapt continuously, communicate transparently, and align predictive capability with interpretability and social relevance. Scientific modeling becomes, in this view, a constructive act: revealing the mechanisms that make complex phenomena intelligible.

Ultimately, the thesis affirms that the energy transition requires not only new technologies but also new epistemologies. By integrating observation, representation, and evolution into a unified causal architecture, this work establishes the foundation for forecasting models that support resilient and inclusive electrification.

ANGEL thus represents a form of coevolutionary learning: a process in which data, models, and socio-technical contexts evolve together, shaping a new generation of causal, flexible, adaptive, and sustainable AI.

Bibliography

[1] Aderibigbe, A. O., Ani, E. C., Ohenhen, P. E., Ohalete, N. C., and Daraojimba, D. O. (2023). Enhancing energy efficiency with ai: a review of machine learning models in electricity demand forecasting. *Engineering Science & Technology Journal*, 4(6):341–356.

[2] Adinkrah, J., Kemausuor, F., Tutu Tchao, E., Nunoo-Mensah, H., Agbemenu, A. S., Adu-Poku, A., and Kponyo, J. J. (2025). Artificial intelligence-based strategies for sustainable energy planning and electricity demand estimation: A systematic review. *Renewable and Sustainable Energy Reviews*, 210:115161.

[3] Agrawal, S., Harish, S., Mahajan, A., Thomas, D., and Urpelainen, J. (2020). Influence of improved supply on household electricity consumption - evidence from rural india. *Energy*, 211:118544.

[4] Akhtar, S., Shahzad, S., Zaheer, A., Ullah, H. S., Kiliç, H., Goňo, R., Jasinski, M., and Leonowicz, Z. (2023). Short-term load forecasting models: A review of challenges, progress, and the road ahead. *Energies*.

[5] Aldossary, M., Alharbi, H. A., and Ayub, N. (2024). Optimizing electric vehicle (ev) charging with integrated renewable energy sources: A cloud-based forecasting approach for eco-sustainability. *Mathematics*, 12(17).

[6] Amjady, N. and Keynia, F. (2008). Mid-term load forecasting of power

systems by a new prediction method. *Energy Conversion and Management*, 49(10):2678–2687.

[7] Anand, H., Nateghi, R., and Alemazkoor, N. (2023). Bottom-up forecasting: Applications and limitations in load forecasting using smart-meter data. *Data-Centric Engineering*, 4:e14.

[8] Angizeh, F., Ghofrani, A., and Jafari, M. A. (2022). Higher granular sectoral demand forecast under data scarcity: An integrated physics-based top–down and bottom–up approach. *IEEE Systems Journal*, 16(2):2923–2933.

[9] Atma, H., Ruzzenenti, F., and van den Broek, M. (2025). Exploring the evolution of long-term electricity demand and load curves in emerging economies: A case study of indonesia's energy transition. *Energy Strategy Reviews*, 60:101805.

[10] Awerkin, A., Verdejo, H., Becker, C., and Barbosa, K. (2017). Statistical parametric techniques for power residential demand forecasting. In *2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*, pages 1–7.

[11] Azeem, A., Ismail, I., Jameel, S. M., and Danyaro, K. U. (2024). Transfer-learning enabled adaptive framework for load forecasting under concept-drift challenges in smart-grids across different-generation-modalities. *Energy Reports*, 12:3519–3532.

[12] Bae, H.-J., Park, J.-S., hyeok Choi, J., and Kwon, H.-Y. (2025). Learning model combined with data clustering and dimensionality reduction for short-term electricity load forecasting. *Scientific Reports*, 15.

[13] Balderrama Subieta, S. L. (2022). *Optimal design and deployment of isolated energy systems: The Bolivian pathway to 100 % rural electrification.* PhD thesis, ULiège - Université de Liège, Liège, Belgium.

[14] Barelli, L., Bidini, G., Cherubini, P., Micangeli, A., Pelosi, D., and Tacconelli, C. (2019). How hybridization of energy storage technologies can provide additional flexibility and competitiveness to microgrids in the context of developing countries. *Energies*, 12(16).

[15] Bayram, F., Aupke, P., Ahmed, B. S., Kassler, A., Theocharis, A., and Forsman, J. (2023). Da-lstm: A dynamic drift-adaptive learning framework for interval load forecasting with lstm networks. *Engineering Applications of Artificial Intelligence*, 123:106480.

[16] Bedi, J. and Toshniwal, D. (2019). Deep learning framework to forecast electricity demand. *Applied energy*, 238:1312–1326.

[17] Belhaiza, S. and Al-Abdallah, S. (2024). A neural network forecasting approach for the smart grid demand response management problem. *Energies*, 17(10):2329.

[18] Bento, P., Pombo, J., Calado, M., and Mariano, S. (2021). Stacking ensemble methodology using deep learning and arima models for short-term load forecasting. *Energies*.

[19] Billanes, J., Jørgensen, B., and Ma, Z. (2025). A framework for resilient community microgrids: Review of operational strategies and performance metrics. *Energies*, 18:405.

[20] Bottou, L. (2012). Stochastic gradient descent tricks. In Montavon, G., Orr, G. B., and Müller, K.-R., editors, *Neural Networks: Tricks of the Trade – Second Edition*, volume 7700 of *Lecture Notes in Computer Science*, pages 421–436. Springer, Berlin, Heidelberg.

[21] Broehl, J. H. (1981). An end-use approach to demand forecasting. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(6):2714–2718.

[22] Carloni, G., Berti, A., and Colantonio, S. (2025). The role of causality in explainable artificial intelligence. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 15(2):e70015.

[23] Chin, H. (2016). An analytical evaluation of top-down versus bottom-up forecast in the electricity demand. In *2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, pages 1–2.

[24] Christensen, R. (1996). *One-Way ANOVA*, pages 79–93. Springer New York, New York, NY.

[25] Chung, M. and Park, H.-C. (2015). Comparison of building energy demand for hotels, hospitals, and offices in korea. *Energy*, 92:383–393. Sustainable Development of Energy, Water and Environment Systems.

[26] Coy, D., Malekpour, S., Saeri, A. K., and Dargaville, R. (2021). Rethinking community empowerment in the energy transformation: A critical review of the definitions, drivers and outcomes. *Energy Research & Social Science*, 72:101871.

[27] Dang, S., Peng, L., Zhao, J., Li, J., and Kong, Z. (2022). A quantile regression random forest-based short-term load probabilistic forecasting method. *Energies*.

[28] Daut, M. A. M., Hassan, M. Y., Abdullah, H., Rahman, H. A., Abdullah, M. P., and Hussin, F. (2017). Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. *Renewable and Sustainable Energy Reviews*, 70:1108–1118.

[29] de Assis Cabral, J., Legey, L. F. L., and de Freitas Cabral, M. V. (2017). Electricity consumption forecasting in brazil: A spatial econometrics approach. *Energy*, 126:124–131.

[30] De Felice, M., Alessandri, A., and Catalano, F. (2015). Seasonal climate forecasts for medium-term electricity demand forecasting. *Applied Energy*, 137:435–444.

[31] Demir, S., Mincev, K., Kok, K., and Paterakis, N. (2021). Data augmentation for time series regression: Applying transformations, autoencoders and adversarial networks to electricity price forecasting. *Applied Energy*, 304:117695.

[32] Dilaver, Z. and Hunt, L. C. (2011). Industrial electricity demand for turkey: A structural time series analysis. *Energy Economics*, 33(3):426–436.

[33] Dong, B., Li, Z., Rahman, S. M., and Vega, R. (2016). A hybrid model approach for forecasting future residential electricity consumption. *Energy and Buildings*, 117:341–351.

[34] Dortolina, C. and Nadira, R. (2005). Estimating future demand a top down/bottom up approach for forecasting annual growths. In *IEEE Power Engineering Society General Meeting, 2005*, pages 400–405 Vol. 1.

[35] Eren, Y. and Küçükdemiral, I. (2024). A comprehensive review on deep learning approaches for short-term load forecasting. *Renewable and Sustainable Energy Reviews*.

[36] Erlangga, G. and Cho, S.-B. (2025). Causally explainable artificial intelligence on deep learning model for energy demand prediction. *Engineering Applications of Artificial Intelligence*, 162:112620.

[37] Fatema, I., Lei, G., and Kong, X. (2023). Probabilistic forecasting of electricity demand incorporating mobility data. *Applied Sciences*, 13(11).

[38] Ferdaus, M. M., Dam, T., Sarkar, M. R., Uddin, M., and Anavatti, S. G. (2025). Foundation models for clean energy forecasting: A comprehensive review. *arXiv preprint arXiv:2507.23147*.

[39] Ferrouhi, E. M. and Bouabdallaoui, I. (2024). A comparative study of ensemble learning algorithms for high-frequency trading. *Scientific African*, 24:e02161.

[40] Few, S., Barton, J., Sandwell, P., Mori, R., Kulkarni, P., Thomson, M., Nelson, J., and Candelise, C. (2022). Electricity demand in populations gaining access: Impact of rurality and climatic conditions, and implications for microgrid design. *Energy for Sustainable Development*, 66:151–164.

[41] Fillol, L. T., Pinomaa, A., Stevanato, N., Mereu, R., and Honkapuro, S. (2025). What happens after electrification? exploring the evolution of appliance adoption in rural kenya. *Energy Research & Social Science*, 122:104015.

[42] Gebremeskel, D. H., Ahlgren, E. O., and Beyene, G. B. (2021). Long-term evolution of energy and electricity demand forecasting: The case of ethiopia. *Energy Strategy Reviews*, 36:100671.

[43] Ghalehkhondabi, I., Ardjmand, E., Weckman, G. R., and Young, W. A. (2017). An overview of energy demand forecasting methods published in 2005–2015. *Energy Systems*, 8(2):411–447.

[44] Ghelasi, P. and Ziel, F. (2025). From day-ahead to mid and long-term horizons with econometric electricity price forecasting models. *Renewable and Sustainable Energy Reviews*, 217:115684.

[45] Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.

[46] Grandjean, A., Adnot, J., and Binet, G. (2012). A review and an analysis of the residential electric load curve models. *Renewable and Sustainable Energy Reviews*, 16(9):6539–6565.

[47] Gruber, J. K. and Prodanovic, M. (2012). Residential energy load profile generation using a probabilistic approach. In *2012 Sixth UKSim/AMSS European Symposium on Computer Modeling and Simulation*, pages 317–322.

[48] Gürses-Tran, G., Oppermann, F., and Monti, A. (2023). Proloaf: Probabilistic load forecasting for power systems. *SoftwareX*, 23:101487.

[49] Haben, S., Arora, S., Giasemidis, G., Voss, M., and Vukadinović Greetham, D. (2021). Review of low voltage load forecasting: Methods, applications, and recommendations. *Applied Energy*, 304:117798.

[50] Hammad, M., Jereb, B., Rosi, B., and Dragan, D. (2020). Methods and models for electric load forecasting: A comprehensive review. *Logistics & Sustainable Transport*, 11:51 – 76.

[51] Haq, M. R. and Ni, Z. (2019). A new hybrid model for short-term electricity load forecasting. *IEEE Access*, 7:125413–125423.

[52] Hashemi, M. (2022). How would residential electricity consumers respond to reductions in power outages? *Energy for Sustainable Development*, 69:1–10.

[53] Hirsch, A., Parag, Y., and Guerrero, J. (2018). Microgrids: A review of technologies, key drivers, and outstanding issues. *Renewable and Sustainable Energy Reviews*, 90:402–411.

[54] Huang, Z. and Zhu, T. (2016). Real-time data and energy management in microgrids. In *2016 IEEE Real-Time Systems Symposium (RTSS)*, pages 79–88.

[55] Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688.

[56] Iftikhar, H., Mancha Gonzales, S., Zywiołek, J., and López-Gonzales, J. L. (2024). Electricity demand forecasting using a novel time series ensemble technique. *IEEE Access*, 12:88963–88975.

[57] Jaramillo, M., Pavón, W., and Jaramillo, L. (2024). Adaptive forecasting in energy consumption: A bibliometric analysis and review. *Data*, 9(1).

[58] Ju, W., Yi, S., Wang, Y., Long, Q., Luo, J., Xiao, Z., and Zhang, M. (2024). A survey of data-efficient graph learning. pages 8104–8113.

[59] Kanté, M., Li, Y., and Deng, S. (2021). Scenarios analysis on electric power planning based on multi-scale forecast: A case study of taoussa, mali from 2020 to 2035. *Energies*, 14(24).

[60] Karunarathne, M. G. I. U., Sudasinghe, P. B., Weeratunge, H. Y., and Herath, D. (2024). Improving microgrid energy demand forecasting using convolutional neural networks. *Engineer: Journal of the Institution of Engineers, Sri Lanka*, 57(4):37–45. Accessed: 2025-10-11.

[61] Kazemzadeh, M.-R., Amjadian, A., and Amraee, T. (2020). A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting. *Energy*, 204:117948.

[62] Ke, Z., Liu, B., Wang, H., and Shu, L. (2021). Continual learning with knowledge transfer for sentiment classification. *ECML-PKDD 2020*.

[63] Kestle, L., London, K. A., Bodger, P. S., and Storey, B. C. (2002). The development of a typology for remote construction projects. In Greenwood, D., editor, *18th Annual ARCOM Conference, 2–4 September 2002, University of Northumbria*, volume 1, pages 305–316. Association of Researchers in Construction Management (ARCOM).

[64] Khalifa, A., Caporin, M., and Di Fonzo, T. (2019). Scenario-based forecast for the electricity demand in qatar and the role of energy efficiency improvements. *Energy Policy*, 127:155–164.

[65] Khare, V. and Chaturvedi, P. (2023). Design, control, reliability, economic and energy management of microgrid: A review. *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, 5:100239.

[66] Khezri, R., Mahmoudi, A., Aki, H., and Muyeen, S. M. (2021). Optimal planning of remote area electricity supply systems: Comprehensive review, recent developments and future scopes. *Energies*, 14(18).

[67] Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015.

[68] Klyuev, R., Morgoev, I., Morgoeva, A., Gavrina, O., Martyushev, N., Efremenkov, E. A., and Qi, M. (2022). Methods of forecasting electric energy consumption: A literature review. *Energies*.

[69] Kondaiah, V. Y., Saravanan, B., Sanjeevikumar, P., and Khan, B. (2022). A review on short-term load forecasting models for micro-grid application. *The Journal of Engineering*, 2022(7):665–689.

[70] Kristjanpoller, W., Michell, K., Llanos, C., et al. (2025). Incorporating causal notions to forecasting time series: a case study. *Financial Innovation*, 11(15).

[71] Kumar, A., He, X., Deng, Y., Singh, A. R., Sah, B., Kumar, P., Bansal, R. C., Bettayeb, M., and Rayudu, R. (2022). A sustainable rural electrification based on a socio-techno-economic-environmental-political microgrid design framework. *Energy and Environmental Science*, 15(10):4213–4246. Publisher Copyright: © 2022 The Royal Society of Chemistry.

[72] Lehmann, R. (2013). 3sigma-rule for outlier detection from the viewpoint of geodetic adjustment. *Journal of Surveying Engineering*, 139:157–165.

[73] LEOPARD Consortium (2020). Leopard. EU Horizon 2020 LEAP-RE Project No 963530. Funded by the European Union's Horizon 2020 Research and Innovation programme under the LEAP-RE initiative.

[74] Levin, T., Bistline, J., Sioshansi, R., Cole, W., Kwon, J., Burger, S., Crabtree, G., Jenkins, J., O'Neil, R., Korpås, M., Wogrin, S., Hobbs, B., Rosner, R., Srinivasan, V., and Botterud, A. (2023). Energy storage solutions to decarbonize electricity through enhanced capacity expansion modelling. *Nature*

Energy, 8(11):1199–1208. Publisher Copyright: © 2023, Springer Nature Limited.

[75] Liu, J., Dietz, T., Carpenter, S., Alberti, M., Folke, C., Moran, E., Pell, A., Deadman, P., Kratz, T., Lubchenco, J., Ostrom, E., Ouyang, Z., Provencher, W., Redman, C., Schneider, S., and Taylor, W. (2007). Complexity of coupled human and natural systems. *Science (New York, N.Y.)*, 317:1513–6.

[76] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T. Y., and Tegmark, M. (2024). Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*.

[77] Lombardi, F., Balderrama, S., Quoilin, S., and Colombo, E. (2019). Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model. *Energy*, 177:433–444.

[78] Lotfipoor, A., Patidar, S., and Jenkins, D. P. (2024). Deep neural network with empirical mode decomposition and bayesian optimisation for residential load forecasting. *Expert Systems with Applications*, 237:121355.

[79] Löschenbrand, M., Gros, S., and Lakshmanan, V. (2021). Generating scenarios from probabilistic short-term load forecasts via non-linear bayesian regression. *2021 International Conference on Smart Energy Systems and Technologies (SEST)*, pages 1–6.

[80] Ma, J. and Ma, X. (2017). State-of-the-art forecasting algorithms for microgrids. In *2017 23rd International Conference on Automation and Computing (ICAC)*, pages 1–6.

[81] Machado, L. F. M., Santo, S. G. D., Junior, G. M., Itiki, R., and Manjrekar, M. D. (2021). Multi-source distributed energy resources management system based on pattern search optimal solution using nonlinearized power flow constraints. *IEEE Access*, 9:30374–30385.

[82] Mahajan, T., Singh, G., and Bruns, G. (2021). An experimental assessment of treatments for cyclical data. In *Proceedings of the 2021 Computer Science Conference for CSU Undergraduates*. Conference session, Virtual.

[83] Maisonnave, M., Delbianco, F., Tohmé, F., Milios, E., and Maguitman, A. (2024). Learning causality structures from electricity demand data. *Energy Systems*, pages 1–23.

[84] Makridakis, S. (1996). Forecasting: its role and value for planning and strategy. *International Journal of Forecasting*, 12(4):513–537.

[85] Martinez-Zapata, K. D., Ospina-Acero, D., Torres, J. J. G., Muñóz-Galeano, N., Gaviria-Gómez, N., Botero, J. F., and Gutiérrez, S. A. (2025). Data-driven load forecasting in microgrids: integrating external factors for efficient control and decision-making. *TechRxiv*. Preprint.

[86] Massidda, L. and Marrocu, M. (2025). Hybrid forecasting of demand flexibility: A top-down approach for thermostatically controlled loads. *Energy and AI*, 20:100487.

[87] Mawalagedara, R., Ray, A., Das, P., et al. (2025). Non-linear dynamical approaches for characterizing multi-sector climate impacts under irreducible uncertainty. *npj Climate and Atmospheric Science*, 8:329.

[88] McCloskey, M. and Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*, volume 24, pages 109–165. Elsevier.

[89] Meena, S., Liu, H., Tu, H., Lukic, S., and Tang, W. (2023). A benchmarking tool for state-of-the-art microgrid design approaches. In *2023 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 625–632.

[90] Mentis, D., Welsch, M., Fuso Nerini, F., Broad, O., Howells, M., Bazilian, M., and Rogner, H. (2015). A gis-based approach for electrification planning—a case study on nigeria. *Energy for Sustainable Development*, 29:142–150.

[91] Messoudi, L., Gouareh, A., Settou, B., Mokhtara, C., and Settou, N. (2024). Modeling and forecasting energy consumption in algerian residential buildings using a bottom-up gis approach. *Energy and Buildings*, 317:114370.

[92] Michalik, G., Khan, M., Bonwick, W., and Mielczarski, W. (1997). Structural modelling of energy demand in the residential sector: 1. development of structural models. *Energy*, 22(10):937–947.

[93] Mir, A. A., Alghassab, M., Ullah, K., Khan, Z. A., Lu, Y., and Imran, M. (2020). A review of electricity demand forecasting in low and middle income countries: The demand determinants and horizons. *Sustainability*, 12(15).

[94] Miraki, A., Parviainen, P., and Arghandeh, R. (2024). Electricity demand forecasting at distribution and household levels using explainable causal graph neural network. *Energy and AI*, 16:100368.

[95] Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., Lalas, D., Moschovits, M., Karagiannis, F., and Papakonstantinou, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. *Energy*, 31(2):208–227.

[96] Misiurek, K., Olkuski, T., and Zyśk, J. (2025). Review of methods and models for forecasting electricity consumption. *Energies*, 18(15).

[97] Mokilane, P., Debba, P., Yadavalli, V. S. S., and Sigauke, C. (2019). Bayesian structural time-series approach to a long-term electricity demand forecasting. *Applied Mathematics & Information Sciences*, 13(2):93–102.

[98] Mokilane, P., Galpin, J., Yadavalli, V., Debba, P., Koen, R., and Sibiya, S. (2018). Density forecasting for long-term electricity demand in south africa using quantile regression. *South African Journal of Economic and management Sciences*, 21.

[99] Morales, R., Marín, L. G., Roje, T., Caquilpan, V., Sáez, D., and Nuñez, A. (2024). Microgrid planning based on computational intelligence methods for rural communities: A case study in the josé painecura mapuche community, chile. *Expert Systems with Applications*, 235:121179.

[100] Muhammad Ali, P. (2022). Investigating the impact of min-max data normalization on the regression performance of k-nearest neighbor with different similarity measurements. *ARO-THE SCIENTIFIC JOURNAL OF KOYA UNIVERSITY*, 10:85–91.

[101] Mulugetta, Y., Ben Hagan, E., and Kammen, D. (2019). Energy access for sustainable development. *Environmental Research Letters*, 14(2):020201.

[102] N, S. G. and Sheshadri, G. S. (2020). Electrical load forecasting using time series analysis. In *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)*, pages 1–6.

[103] Nandakumar, A., Li, Y., Zheng, H., Zhao, J., Zhao, D., Zhang, Y., Hong, T., and Chen, B. (2024). Data-driven modeling of microgrid transient dynamics through modularized sparse identification. *IEEE Transactions on Sustainable Energy*, 15(1):109–122.

[104] Napoleoni, M. (2020). Building a decision aid tool in anylogic to simulate a completely renewable multi-energy μgrid. Master's thesis, Politecnico di Torino, Turin, Italy.

[105] Nayanathara, C. and Srilatha, R. (2018). Electrifying villages using mi-

crogrids. In *2018 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, pages 300–305.

[106] Otieno, F., Williams, N., and McSharry, P. (2018). Forecasting energy demand for microgrids over multiple horizons. In *2018 IEEE PES/IAS PowerAfrica*, pages 457–462.

[107] Partnership, D. U. T. D. (2024). Pedaid3: Positive energy district supported by ai and data-driven decisions. Project profile, DUT Partnership. Co-funded by the European Union under the Horizon Europe programme. Coordinated by the Universidad de Las Palmas de Gran Canaria (ULPGC).

[108] Perpiña Castillo, C., Ribeiro Barranco, R., Curtale, R., Kompil, M., Jacobs-Crisioni, C., Vallecillo Rodriguez, S., Aurambout, J.-P., Batista e Silva, F., Sulis, P., and Auteri, D. (2024). Are remote rural areas in europe remarkable? challenges and opportunities. *Journal of Rural Studies*, 105:103180.

[109] Pezzulli, S., Frederic, P., Majithia, S., Sabbagh, S., Black, E., Sutton, R., and Stephenson, D. (2006). The seasonal forecast of electricity demand: a hierarchical bayesian model with climatological weather generator. *Applied Stochastic Models in Business and Industry*, 22(2):113–125.

[110] Pinto, T., Praça, I., Vale, Z., and Silva, J. (2021). Ensemble learning for electricity consumption forecasting in office buildings. *Neurocomputing*, 423:747–755.

[111] Poblete-Cazenave, M. and Pachauri, S. (2020). A simulation-based estimation model of household electricity demand and appliance ownership. MPRA Paper 103403, University Library of Munich, Germany.

[112] PROBONO Consortium (2022). Probono: The integrator-centric approach for realising innovative energy efficient buildings in connected sustainable green

neighbourhoods. EU Horizon 2020 Project No 101037075. Funded by the European Union's Horizon 2020 Research and Innovation programme. Coordinated by Ingenieria Especializada Obra Civil e Industrial S.A. (Spain).

[113] Qiu, L. (2024). Deep learning approaches for building energy consumption prediction. *Frontiers in Environmental Research*, 2(3):11–17.

[114] Riva, F. and Colombo, E. (2020). System-dynamics modelling of the electricity-development nexus in rural electrification based on a tanzanian case study. *Energy for Sustainable Development*, 56:128–143.

[115] Román-Portabales, A., López-Nores, M., and Pazos-Arias, J. J. (2021). Systematic review of electricity demand forecast using ann-based machine learning algorithms. *Sensors*, 21(13):4544.

[116] Sanfilippo, S. (2025). Building a decision aid tool based on anylogic's models for the songhai's microgrid in porto novo. Master's thesis, Politecnico di Torino, Turin, Italy.

[117] Sanfilippo, S. and et al. (2023). Microgrid design optimization in benin within the leopard project: Evaluating the impact of inaccurate load profile estimation. In *2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*.

[118] Sanfilippo, S., Farina, L., Vito, P. D., Hernández-Cabrera, J. J., Hernández-Gálvez, J. J., and Évora-Gómez, J. (2025a). Multi-method simulation and optimisation for maximising benefits in renewable energy communities: A real-world case study from italy. *Baltic Journal of Modern Computing*, 13(1):200–220.

[119] Sanfilippo, S., Hernández-Cabrera, J. J., Kändler, C., Hernández-Gálvez, J. J., Évora Gómez, J., and Roncal-Andrés, O. (2025b). A neural network-

based causal model for electricity demand estimation in remote areas: A case study in el espino, bolivia. *Baltic Journal of Modern Computing*, 13(2):315–330.

[120] Sanfilippo, S., Hernández-Gálvez, J. J., Hernández-Cabrera, J. J., Évora Gómez, J., Roncal-Andrés, O., and Caballero-Ramirez, M. (2025c). Evolving electricity demand modelling in microgrids using a kolmogorov-arnold network. *INFORMATICA*.

[121] Sarker, M., Alam, M. J., Ramasamy, G., and Nasir Uddin, M. (2024). Energy demand forecasting of remote areas using linear regression and inverse matrix analysis. *International Journal of Electrical and Computer Engineering (IJECE)*, 14:129–139.

[122] Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., and Bengio, Y. (2021). Toward causal representation learning. *Proceedings of the IEEE*, 109(5):612–634.

[123] Semenoglou, A.-A., Spiliotis, E., Makridakis, S., and Assimakopoulos, V. (2021). Investigating the accuracy of cross-learning time series forecasting methods. *International Journal of Forecasting*, 37(3):1072–1084.

[124] Setiawan, A., Koprinska, I., and Agelidis, V. G. (2009). Very short-term electricity load demand forecasting using support vector regression. In *2009 International Joint Conference on Neural Networks*, pages 2888–2894. IEEE.

[125] Setokoe, T. J. and Ramukumba, T. (2022). Ability to do for self: A rural community's view of the importance of capacity building in sustainable tourism development. *GeoJournal of Tourism and Geosites*, 41(2):583–588.

[126] Shah, I., Iftikhar, H., and Ali, S. (2022). Modeling and forecasting electricity demand and prices: A comparison of alternative approaches. *Journal of Mathematics*, 2022:1–14.

[127] Sharma, K., Dwivedi, Y. K., and Metri, B. (2024). Incorporating causality in energy consumption forecasting using deep neural networks. *Annals of Operations Research*, 339(1):537–572.

[128] Shufian, A. and Mohammad, N. (2022). Modeling and analysis of cost-effective energy management for integrated microgrids. *Cleaner Engineering and Technology*, 8:100508.

[129] Shukla, A. K., Yadav, S. K., and Tiwari, V. (2015). Linear models for s-shaped growth curves. *Journal of Statistics Applications & Probability*, 4(1):113.

[130] Singh, M. and Balachandra, P. (2019). Microhybrid electricity system for energy access, livelihoods, and empowerment. *Proceedings of the IEEE*, 107(9):1995–2007.

[131] Solyali, D. (2020). A comparative analysis of machine learning approaches for short-/long-term electricity load forecasting in cyprus. *Sustainability*.

[132] Son, N. and Shin, Y. (2023). Short- and medium-term electricity consumption forecasting using prophet and gru. *Sustainability*.

[133] Sossou, S., Tossa, A. K., Vodounnou, S., Dumont, O., Sanchez-Solis, C., Balderrama, S., Lurie, N., Sehou, M., Semassou, C., and Quoilin, S. (2025). Calibration and extrapolation of a stochastic model for electricity demand forecasting in rural communities: The case of samionta, benin. In *Proceedings of the 38th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS 2025)*, pages 1–12, Paris, France. ECOS. Paper ID: 7822.

[134] Swan, L. G. and Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, 13(8):1819–1835.

[135] Tian, S. and Chang, S. (2020). An agent-based model of household energy consumption. *Journal of Cleaner Production*, 242:118378.

[136] Tian, Y., Zhang, Y., and Zhang, H. (2023). Recent advances in stochastic gradient descent in deep learning. *Mathematics*, 11(3).

[137] Toersche, H. A., Molderink, A., Hurink, J. L., and Smit, G. J. (2014). Cascaded column generation for scalable predictive demand side management. In *2014 IEEE International Energy Conference (ENERGYCON)*, pages 1228–1235.

[138] Tomas Fillol, L., Stevanato, N., Pinomaa, A., Mereu, R., and Honkapuro, S. (2025). Improving electricity demand growth estimation in rural mini-grids through data-driven appliance diffusion modeling. *Energy for Sustainable Development*, 89:101826.

[139] Tong, Y., Yang, W., and Zhan, D. (2019). Energy use prediction with information theory and machine learning technique. *IOP Conference Series: Earth and Environmental Science*, 291(1):012031.

[140] Torabi Moghadam, S., Toniolo, J., Mutani, G., and Lombardi, P. (2018). A gis-statistical approach for assessing built environment energy use at urban scale. *Sustainable Cities and Society*, 37:70–84.

[141] Tsoumplekas, G., Athanasiadis, C. L., Doukas, D. I., Chrysopoulos, A., and Mitkas, P. A. (2024). Few-shot load forecasting under data scarcity in smart grids: A meta-learning approach. *arXiv preprint arXiv:2406.05887*.

[142] Uhl, J. H., Leyk, S., Hunter, L. M., Talbot, C. B., Connor, D. S., Nieves, J. J., and Gutmann, M. (2022). A fine-grained, versatile index of remoteness to characterize place-level rurality. Annual meeting of the Population Association of America (PAA) 2021, May 5-8, 2021, virtual event. Extended abstract. 5 pages, 2 figures.

[143] Ullah, K., Ahsan, M., Hasanat, S. M., Haris, M., Yousaf, H., Raza, S. F., Tandon, R., Abid, S., and Ullah, Z. (2024). Short-term load forecasting: A comprehensive review and simulation study with cnn-lstm hybrids approach. *IEEE Access*, 12:111858–111881.

[144] Vanting, N. B., Ma, Z., and Jørgensen, B. N. (2021). A scoping review of deep neural networks for electric load forecasting. *Energy Informatics*, 4(Suppl 2):49.

[145] Veldman, E., Gibescu, M., Slootweg, H. J., and Kling, W. L. (2013). Scenario-based modelling of future residential electricity demands and assessing their impact on distribution grids. *Energy Policy*, 56:233–247.

[146] Verdejo, H., Awerkin, A., Becker, C., and Olguin, G. (2017). Statistic linear parametric techniques for residential electric energy demand forecasting. a review and an implementation to chile. *Renewable and Sustainable Energy Reviews*, 74:512–521.

[147] Vogt, Y. (2003). Top-down energy modeling. *Strategic Planning for Energy and the Environment*, 22(4):64–79.

[148] Waheed, W. and Xu, Q. (2024). Data-driven short term load forecasting with deep neural networks: Unlocking insights for sustainable energy management. *Electric Power Systems Research*, 232:110376.

[149] Wang, Z. and Hong, T. (2020). Generating realistic building electrical load profiles through the generative adversarial network (gan). *Energy and Buildings*, 224:110299.

[150] Warren, K., Ambrosio, R., Chen, B., Fu, Y., Ghosh, S., Phan, D., Sinn, M., Tian, C., and Visweswariah, C. (2016). Managing uncertainty in electricity generation and demand forecasting. *IBM Journal of Research and Development*, 60:8:1–8:13.

[151] Wazirali, R., Yaghoubi, E., Abujazar, M. S. S., Ahmad, R., and Vakili, A. H. (2023). State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques. *Electric Power Systems Research*, 225:109792.

[152] Wen, Q. and Liu, Y. (2025). Feature engineering and selection for prosumer electricity consumption and production forecasting: A comprehensive framework. *Applied Energy*, 381:125176.

[153] Widén, J., Lundh, M., Vassileva, I., Dahlquist, E., Ellegård, K., and Wäckelgård, E. (2009). Constructing load profiles for household electricity and hot water from time-use data—modelling approach and validation. *Energy and Buildings*, 41(7):753–768.

[154] Wiese, F., Schlecht, I., Bunke, W.-D., Gerbaulet, C., Hirth, L., Jahn, M., Kunz, F., Lorenz, C., Mühlenpfordt, J., Reimann, J., and Schill, W.-P. (2019). Open power system data – frictionless data for electricity system modelling. *Applied Energy*, 236:401–409.

[155] Xia, Z., Ma, H., Saha, T. K., and Zhang, R. (2022). Consumption scenario-based probabilistic load forecasting of single household. *IEEE Transactions on Smart Grid*, 13(2):1075–1087.

[156] Yang, J. and Johansson, D. J. (2024). Adapting to uncertainty: Modeling adaptive investment decisions in the electricity system. *Applied Energy*, 358:122603.

[157] Yu, Z.-X., Li, M.-S., Xu, Y.-P., Aslam, S., and Li, Y.-K. (2021). Techno-economic planning and operation of the microgrid considering real-time pricing demand response program. *Energies*, 14(15).

[158] Zeiler, M. D. (2012). ADADELTA: An adaptive learning rate method. *arXiv preprint arXiv:1212.5701.* 6 pages.

[159] Zhang, T., Siebers, P.-O., and Aickelin, U. (2011). Modelling electricity consumption in office buildings: An agent based approach. *Energy and Buildings*, 43(10):2882–2892.

[160] Zhang, Y., Cui, L., and Yan, W. (2025). Integrating kolmogorov–arnold networks with time series prediction framework in electricity demand forecasting. *Energies*, 18(6):1365.

[161] Ángela Herraiz-Cañete, Ribó-Pérez, D., Bastida-Molina, P., and Gómez-Navarro, T. (2022). Forecasting energy demand in isolated rural communities: A comparison between deterministic and stochastic approaches. *Energy for Sustainable Development*, 66:101–116.

Acronyms

AB Agent Based. 19, 23, 35, 38, 54–57, 60, 62

ADADELTA Adaptive Delta. 103, 119, 121

ADAGRAD Adaptive Gradient. 103, 119, 121

ADAM Adaptive Moment Estimation. 103, 119, 121, 124

AI Artificial Intelligence. 14, 15, 17–19, 23, 26, 29, 38, 50, 57, 63, 66–68, 86, 88, 89, 98, 101, 138, 146, 149

ANN Artificial Neural Network. xiii, 26, 27, 64, 66, 124, 127, 129, 130

ANOVA Analysis of Variance. 118, 119

ARIMA Autoregressive Integrated Moving Average. 25, 26, 39, 40, 67

ARIMAX ARIMA with exogenous variables. 26, 27, 40

BUM Bottom-Up Method. 29, 31, 32, 35, 48

CAPEX Capital Expenditure. 108

CL Continual Learning. 78, 98, 131–136, 138, 141–143, 148

CNN Convolutional Neural Network. 68

DFNN Deep Feedforward Neural Network. xii, xiii, 94, 119, 121–123, 125, 127, 129, 130, 135, 139

DoA Degree of Adoption. iv, vii, xiii, 125–128, 130, 131, 133, 135, 139, 142, 143, 147

DSM Demand-Side Management. 30, 33, 35, 49, 50

EIFER European Institute for Energy Research. 6, 18

FT Full Training. 132

GB Gradient Boosting. 26, 27

GDP Gross Domestic Product. 27, 28, 30, 36, 42, 46, 61, 66

GIS Geographic Information System. 19, 23, 38, 50–53

HF High-Frequency. 26, 31, 33, 63, 99, 102

HVAC Heating, Ventilation, and Air Conditioning. 48, 49, 58

KAN Kolmogorov–Arnold Network. xiii, 14, 78, 94, 123–125, 127–134, 136, 139, 140, 142–144

LCOE Levelized Cost Of Electricity. 6, 8–10

LSTM Long Short-Term Memory. 64, 68

LTF Long-Term Electricity Demand Forecasting. 25, 28, 29, 40, 46, 63

MAE Mean Absolute Error. xiii, 94, 104, 105, 121, 125, 129, 130, 132

MG microgrid. xiv, 4–10, 13–20, 23, 33, 35, 40, 43, 52, 67, 68, 71, 72, 75, 76, 78–81, 83, 84, 87–89, 92–94, 96–102, 104, 107–109, 112, 113, 123, 125, 132, 134–136, 138, 139, 142–145, 148

ML Machine Learning. 10, 19, 23–27, 29, 33, 38, 40, 43, 47, 50, 53, 63, 65–68, 72, 75, 77, 86, 111

MLP Multilayer Perceptron. xii, xiii, 94, 119, 121–123, 125, 127–130, 135, 139

MPEPAD Maximum Percentage of Electricity Production Accepted from Diesel engine. 6, 9, 10

MSE Mean Squared Error. 94, 104, 105, 121

MTF Medium-Term Electricity Demand Forecasting. 24, 27, 28, 40, 46, 63, 66

NGO Non-governmental Organization. 109

OLP Overestimated Load Profile. xii, xiv, 6–10

PED Positive Energy District. 19, 37, 145

PV Photovoltaic. xiv, 4, 6–10, 18, 59, 67

REC Renewable Energy Community. 19, 145

ReLU Rectified Linear Unit. 64, 121, 124

RFS Random Forest. 27, 65–67

RLP Reference Load Profile. xii, xiv, 5–9

RNN Recurrent Neural Network. 64

SARIMA Seasonal ARIMA. 25, 27, 40, 67

SDG Sustainable Development Goals. 98

SGD Stochastic Gradient Descent. 103, 119, 121

STAM STAM S.r.l.. 19

STF Short-Term Electricity Demand Forecasting. 24–26, 32, 40, 46, 50, 63, 64, 67

SVM Support Vector Machine. 65

SVR Support Vector Regression. 26, 65–67

TDM Top-Down Method. 29, 30, 32, 36, 56

ULP Underestimated Load Profile. xii, xiv, 6–10

ULPGC Universidad de Las Palmas de Gran Canaria. 145