PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF ECHAHID CHEIKH LARBI TEBESSI - TEBESSA

FACULTY OF SCIENCE AND TECHNOLOGY

THESIS

Presented for obtaining the diploma of

3rd Cycle LMD DOCTORATE

In: **ELECTROTECHNICAL**

Specialty: Electrical Networks

by: **Nour El Houda BEZZAR**

Title

Energy Forecasting in Smart Grids: Toward an Artificial Intelligence-Driven Approach

Discussed publicly at: 14/10/2025 in front of the assessment committee:

Aziz Boukadoum	***Professor***	*Echahid Cheikh Larbi Tebessi University*	***Chairman***
Abdallah Meraoumia	***Professor***	*Echahid Cheikh Larbi Tebessi University*	***Supervisor***
Lotfi Houam	***MCB***	*Echahid Cheikh Larbi Tebessi University*	***Co-supervisor***
Hocine Labar	***Professor***	*Badji Mokhtar-Annaba University*	***Examiner***
Laatra Yousfi	***MCA***	*Echahid Cheikh Larbi Tebessi University*	***Examiner***
El Bahi Azzag	***Professor***	*Badji Mokhtar-Annaba University*	***Invited***


بسم الله الرحمن الرحيم


وقل رب زدني علما

صدق الله العظيم

Acknowledgements

In the name of Allah, the Most Gracious, the Most Merciful.
All praise is due to Allah, through whose grace good deeds are completed and goals are achieved.

I extend my sincere thanks and deep gratitude to everyone who supported and encouraged me throughout my years of study and in completing this thesis, which represents the culmination of a long and challenging academic journey.

First and foremost, I am especially grateful to my beloved father, who has always been a role model of patience and perseverance, and who instilled in me the value of knowledge and the pursuit of excellence.

To my dearest mother, who has been my primary spiritual support, with her sincere prayers and constant support throughout every stage of my life.

I will never forget my beloved brothers, who have always been a strong support for me, encouraging me and standing by my side in every circumstance. And to all my loved ones who have given me strength and support, standing by my side morally and emotionally during all the difficult times.

I also extend my sincere thanks and gratitude to my supervisor, Professor MERAOUMIA Abdallah, for his academic support and valuable guidance, which were the cornerstone of this work.

I also thank HOUAM Lotfi, my co-supervisor, for his careful follow-up and constructive comments, which contributed to the development of the content of the dissertation.

I would also like to express my gratitude to Professor LAIMECHE Lakhdar for his academic support and valuable advice, which greatly enriched this work.

To all of you, all thanks, appreciation and gratitude. Thanks to you, and with God's help, this achievement was made possible.

Abstract

This thesis focuses on developing an accurate electricity consumption forecasting model within smart grids by leveraging modern artificial intelligence techniques to enhance energy management and improve efficiency in residential environments. The study began with the preparation and preprocessing of detailed electricity consumption data for a single household, including measurements such as active power, reactive power, voltage, current intensity, and consumption from sub-metered appliances.

Initially, a Long Short-Term Memory (LSTM) neural network model was applied for load forecasting. However, after comparative evaluation, the XGBoost model demonstrated superior accuracy and performance in reducing prediction errors. Consequently, efforts were concentrated on optimizing the XGBoost model through hyperparameter tuning and measure the features (using normalization or standardization), resulting in highly accurate forecasts with minimal error margins.

To extend the scope of the research, the model was applied to six different households in a specific area, with hourly and daily predictions performed to provide a detailed and real-time view of electricity consumption patterns. This expansion allowed for a comprehensive understanding of consumption variability and enhanced the model's predictive capability.

In the final phase, an integrated energy management framework was developed based on the forecasting results. This framework optimizes the utilization of multiple energy sources including solar, wind, battery storage, and the electrical grid thus enabling efficient energy distribution and reducing reliance on the grid when possible. The framework represents a significant step towards building intelligent and sustainable energy systems capable of adapting to demand fluctuations and maximizing operational efficiency.

This thesis makes a valuable scientific and practical contribution to the field of smart grids by combining AI techniques with energy management strategies. It opens new avenues for innovative solutions that support energy sustainability and help reduce environmental impacts in the future.

Keywords: Smart Grids, Electricity Consumption Forecasting, Energy Management, XGBoost Model, Artificial Intelligence (AI), Hyperparameter Tuning

ملخص

تُركز هذه الأطروحة على تطوير نموذج تنبؤي دقيق لاستهلاك الكهرباء ضمن الشبكات الذكية، مستفيدًا من تقنيات الذكاء الاصطناعي الحديثة بهدف تحسين إدارة الطاقة وزيادة كفاءتها في البيئات السكنية. بدأ البحث بمرحلة تجهيز وتحليل البيانات الخاصة باستهلاك الكهرباء لمنزل واحد، حيث تم جمع ومعالجة بيانات مفصلة تشمل قياسات متعددة مثل القدرة الفعالة، القدرة غير الفعالة، الجهد، شدة التيار، وكذلك استهلاك الأجهزة الفرعية في المنزل.

تم تطبيق نموذج الشبكات العصبية الذاكرة الطويلة القصيرة المدى (LSTM) كنموذج أولي للتنبؤ بالطلب الكهربائي، لكن بعد المقارنة بين النماذج المختلفة، ظهر تفوق نموذج XGBoost في تقديم تنبؤات أدق وأداء أفضل من حيث تقليل الخطأ. بناءً على ذلك، تم التركيز على تحسين أداء نموذج XGBoost عبر ضبط المعاملات و قياس الميزات (باستخدام التطبيع أو التوحيد القياسي)، مما أسفر عن نتائج دقيقة جداً تقلل من اخطاء التنبؤ.

لتوسيع نطاق الدراسة، تم تطبيق النموذج على ستة منازل مختلفة في منطقة محددة، مع إجراء التنبؤات على أساس الساعة و اليوم، مما أتاح رؤية مفصلة لاستهلاك الكهرباء في الوقت الفعلي. هذه الخطوة سمحت بتقييم نمط الاستهلاك وتنوعه بين المنازل، مما يعزز من دقة التنبؤ ويعطي مؤشرات مهمة لإدارة الطاقة.

في المرحلة النهائية من الدراسة، تم وُضع إطار عمل متكامل لإدارة مصادر الطاقة المختلفة (كالطاقة الشمسية، طاقة الرياح، البطاريات، والشبكة الكهربائية) استنادًا إلى نتائج التنبؤات، مما يساهم في تحقيق توزيع أمثل للطاقة وتقليل الاعتماد على الشبكة الكهربائية. هذا الإطار يُمثل خطوة نحو بناء أنظمة طاقة ذكية ومستدامة قادرة على التكيف مع الطلب وتحقيق كفاءة تشغيلية عالية.

تقدم هذه الأطروحة إسهامًا علميًا وعمليًا مهمًا في مجال الشبكات الذكية، من خلال دمج تقنيات الذكاء الاصطناعي مع إدارة الطاقة، مما يفتح آفاقًا جديدة لتطوير حلول مبتكرة تساعد على تحقيق استدامة الطاقة وتقليل الانبعاثات البيئية مستقبلًا.

الكلمات المفتاحية: الشبكات الذكية، التنبؤ باستهلاك الكهرباء ، إدارة الطاقة، نموذج XGBoost، الذكاء الاصطناعي (AI)، ضبط المعلمات الفائقة

Contents

1. General Introduction **01**
1.1 Motivation and Goal 03
1.2 Contribution 04
1.3 Structure of the Thesis 05
2. Smart Grid **06**
2.1 Electric energy system 06
2.2 Electrical grid operation 07
2.3 Smart grid 12
2.3.1 Definition 13
2.3.2 Smart grid benefits 16
2.3.3 Issues and challenges in smart grids 19
2.3.4 Smart grid component 20
2.4 Conclusion 22
3. Household Energy Management **24**
3.1 Smart Energy Management in Homes: The Role of AI and Forecasting 24
3.1.1 Effects of unregulated energy consumption 24
3.1.2 Role of energy management 25
3.2 Energy consumption prediction 25
3.2.1 Necessity 27
3.2.2 Benefits 27
3.3 Techniques Used 29
3.3.1 Time Series Forecasting 30
3.3.2 Artificial Intelligence and Machine Learning 31
3.4 Literature review 35
3.5 Challenges of home energy management 37
3.5.1 Data accuracy and diverse energy sources 39
3.5.2 Smart systems needed 40
3.6 Impact of home energy management 42
3.6.1 Long-term costs reduction 43
3.6.2 Reducing the carbon footprint 44
3.7 Conclusion 45
4. Efficient Management of Residential Electricity Consumption **47**
4.1 Technical Foundations 48
4.1.1 Time Series Forecasting Principles 48
4.1.2 Long Short-Term Memory Networks 53
4.1.3 Extreme Gradient Boosting 53
4.1.4 Performance measures 55
4.2 Dataset Description 56
4.3 Data Preprocessing 57
4.3.1 Data Cleaning 59
4.3.2 Adopted Decomposition Model 61
4.3.3 Autocorrelation Analysis 64
4.4 Feature Engineering 66
4.4.1 Creation of New Features 66
4.4.2 Normalization and Standardization 67
4.5 Proposed Methodology 70
4.5.1 Forecasting Pipeline 71
4.5.2 Energy Source Management Strategy 72
4.6 Conclusion 74
5. Experimental Results and Interpretation **75**

5.1 Experimental Evaluation 75
5.2 LSTM-Based Electricity Consumption Forecasting 76
5.2.1 Time Series Modeling 78
5.3 Forecasting Electricity Consumption Using XGBoost 80
5.3.1 Hyper parameter Optimization for the XGBoost Model 82
5.3.2 Model Evaluation with Feature Scaling Applied to Dataset 84
5.4 Comparative Analysis with Existing Approaches 87
5.4.1 Hourly Prediction Performance 87
5.4.2 Daily and Second-Level Forecasts 88
5.5 Electricity Consumption Forecasting for Multiple Households Using XGBoost 88
5.5.1 Hyper parameter Optimization for the XGBoost Model 90
5.5.2 Motivation for Multi-Household Forecasting Extension 92
5.6 Energy management system 92
5.6.1 Single-Source Energy Management 93
5.6.2 Multi-Source Energy management 94
5.6.3 Energy Source Switching Control 96
5.6.3 Benefits of Energy Management Strategies 98
5.7 Conclusion 98
6. Conclusion and Future Directions 99
A. Scientific Productions 102
A.1 Publications 102
A.2 International Conferences 102
A.2 National Conferences 102

List of Figures

Figures		Page
II.1	World net electricity generation by energy source, 2010-2040 (trillion kilowatt-hours)	07
II.2	Basic structure of the electric power system	08
II.3	Energy sources	09
II.4	Transmission and Distribution energy lines	10
II.5	Types of electrical consumers	11
II.6	Smart grid	14
II.7	Comparison of traditional grid and smart grid	16
II.8	A conceptual illustration of a microgrid	19
II.9	Overview of Smart Grid design	20
III.1	Smart Home Energy Management System	25
III.2	Conceptual Framework of Energy Consumption Prediction in Smart Home Environments	26
III.3	Comparison of actual vs predicted energy consumption (1 week)	28
III.4	Deep Learning Workflow for Energy Consumption Prediction: Data Sources to Model Evaluation	34
III.5	Challenges in Home Energy Management	38
III.6	Smart Home Energy management system energy sources	39
III.7	Functional Ecosystem of Smart Systems in Home Energy Management	40
III.8	Concepts of multiple HEMS benefits at the household	42
IV.1	Proposed Framework for Electricity Consumption Forecasting	58
IV.2	Percentage of missing values in the dataset	60
IV.3	Household electricity consumption decomposition based on monthly granulate	61
IV.4	Seasonality of household electricity consumption	63
IV.5	Household electricity consumption residual errors	64
IV.6	Autocorrelation of time series data re-sampled	65
IV.7	Probability density function of IHEPC features	68
IV.8	Intelligent energy source selection based on consumption forecast	70

V.1	Features Resampled Over Hour and Day	76
V.2	Hourly and daily prediction results for LSTM model	77
V.3	XGBoost based modeling process	79
V.4	Household prediction vs original observations resampled	81
V.5	XGBoost based modeling process	83
V.6	Household predictions vs. original observations after features normalization	85
V.7	Household predictions vs. original observations after features standardization	85
V.8	Household prediction vs original observations (Hourly granularity) For 6 homes	89
V.9	Household prediction vs original observations (Daily granularity) For 6 homes	89
V.10	Household prediction vs original observations after XGBoost hyper-parameters tuning (Hourly granularity) For 6 homes	90
V.11	Household prediction vs original observations XGBoost hyper-parameters tuning (Daily granularity) For 6 homes	91
V.12	Seasonal variations in energy sourcing	93

List of Tables

Tables		Page
IV.1	Summary statistics of the residual errors	63
V.1	Performance metric (RMSE)	78
V.2	Comparative results of the impact of the two scaling methods	86
V.3	The comparative analysis of the proposed method with state-of-the-art (DL and ML techniques) for hourly data resolution	87
V.4	RMSE after XGBoost hyper-parameters tuning	91

List of Abbreviation

AC	Alternating Current
AEP	Appliances Energy Prediction
AI	Artificial Intelligence
AMI	Advanced Metering Infrastructure
ANN	Artificial Neural Network
API	Application Programming Interface
AR	Autoregressive
ARIMA	AutoRegressive Integrated Moving Average
ARMA	AutoRegressive Moving Average
BACnet	Building Automation and Control Network
BRNN	Bidirectional Recurrent Neural Network
CNN	Convolutional Neural Networks
CO2	Carbon Dioxide
DERs	Distributed Energy Resources
DL	Deep Learning
DR	Demand Response
DSM	Demand Side Management
EDA	AutoRegressive Moving Average
EDA	Exploratory Data Analysis
EI	Energy Information
EKPC	East Kentucky Power Cooperative
EMS	Energy Management Systems
ETS	Exponential Smoothing
EVs	Electric Vehicles
EWO	Earth Worm Optimization
FACTS	Flexible AC Transmission Systems
GBT	Gradient Boosted Trees
GDPR	Independent System Operator

GHG	Greenhouse Gases
GRU	Gated Recurrent Unit
GWO	Gray Wolf Optimization
HEMS	Home Energy Management Systems
HVAC	Heating, Ventilation, and Air Conditioning
Hz	Hertz
IEA	International Energy Agency
IEO	International Energy Outlook
IHEPC	Individual Household Electric Power Consumption
IoT	Internet of Things
ISO	Independent System Operator
k-NN	k-Nearest Neighbors
KPCA	Kernel Principal Component Analysis
KW	Kilowatts
KWh	Kilowatt hours
LightGBM	Light Gradient Boosting Machine
LSTM	Long Short-Term Memory
LWSVR	Locally Weighted Support Vector Regression
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MGOA	Modified Grasshopper Optimization Algorithm
ML	Machine Learning
MLPs	Multilayer Perceptrons
MQTT	Message Queuing Telemetry Transport
MW	Megawatts
NBEATS	Neural Basis Expansion Analysis for Time Series
PV	Photovoltaic
RES	Renewable Energy Sources
RF	Random Forest
RMSE	Root Mean Square Error
RNN	Recurrent neural networks

SARIMA	Seasonal AutoRegressive Integrated Moving Average
SCADA	Supervisory Control and Data Acquisition
SoC	State of Charge
SVM	Support Vector Machines
SVR	Support Vector Regression
T&D	Transmission and Distribution
TCN	Temporal Convolutional Network
UCI	University of California, Irvine
VAR	Vector Autoregression
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

Chapter I

General Introduction

Electricity plays a foundational role in the modern world, serving as the lifeblood of both domestic and industrial operations. It powers homes, fuels economic growth, supports healthcare and transportation, and underpins critical infrastructure. As population growth, urbanization, and technological advancements continue to increase energy demand, ensuring a stable and efficient electricity supply has become a global challenge. Smart grids have emerged in response to this growing demand and complexity. These digitally enhanced power networks integrate sensors, data analytics, and two-way communication to better manage electricity generation, distribution, and consumption. Smart grids facilitate real-time monitoring, support decentralized energy sources, enable demand-side management, and integrate renewable energy efficiently [1].

Accurate electricity consumption forecasting lies at the heart of effective smart grid management. By predicting future energy demand, utility providers can optimize resource allocation, maintain grid stability, and reduce operational waste. Such forecasting also supports the integration of renewable energy sources, like solar and wind, whose variability can introduce significant grid management challenges. Historically, forecasting methods have relied on statistical models such as ARIMA [2], exponential smoothing, and linear regression. While these techniques have offered foundational insights, they often struggle with the nonlinear and dynamic patterns characteristic of modern electricity consumption [3]. Moreover, they typically require substantial manual feature engineering and assume linear relationships, limiting their applicability in complex or rapidly changing environments. Artificial Intelligence (AI) and Machine Learning (ML) techniques offer a robust alternative [4] variables like weather. These data-driven approaches can uncover hidden patterns in large, heterogeneous datasets, adapting dynamically to the underlying data. In particular, supervised ML models are well-suited for energy forecasting due to their ability to capture nonlinear interactions and handle noisy data [5].

Among ML models, ensemble learning, especially gradient boosting techniques like XGBoost, has proven particularly effective. XGBoost is a decision-tree-based model that iteratively optimizes performance by minimizing prediction errors. It supports regularization to prevent overfitting, handles missing data, and allows for efficient parallel processing, making it ideal for large-scale time series forecasting. Applying XGBoost to electricity forecasting enables a more nuanced understanding of consumption trends. This is especially beneficial when working with data from smart meters, which offer high-frequency readings of variables such as voltage, current, and usage across specific household zones (e.g., kitchen, laundry). These granular data help the model detect subtle usage patterns and temporal fluctuations, enhancing forecast accuracy [6]. For consumers, such precision enables better energy management, optimizing usage, identifying peak hours, and lowering costs. For providers, it translates to improved planning, fewer outages, and more efficient infrastructure use. Additionally, policymakers can use these insights to design evidence-based energy regulations and promote sustainable technologies.

However, raw smart meter data often include noise, missing values, and inconsistencies. Therefore, preprocessing steps like data cleaning, normalization, and feature engineering are essential. Transforming timestamps into meaningful time-based features (e.g., day of the week, season, holidays) and incorporating external variables like weather can significantly boost model performance [4]. This research leverages a real-world residential dataset to train and validate the XGBoost model, evaluating its performance across short-term (hourly) and medium-term (daily) forecasts. These temporal levels are chosen for their practical relevance, short-term forecasting aids real-time grid management, while medium-term insights support capacity planning and operational efficiency. To fine-tune the model, hyperparameter optimization techniques such as grid search and Bayesian optimization are employed. These methods systematically identify the best model configurations to minimize forecasting errors. Performance is assessed using key metrics including Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE), which quantify prediction accuracy and model reliability.

Beyond predictive accuracy, the broader value of AI-based forecasting lies in its practical utility. In an era marked by urgent climate goals and volatile energy markets, improved demand forecasting can accelerate the shift toward renewable energy, reduce carbon emissions, and optimize the use of existing infrastructure. More accurate forecasts can also lower operational costs and reduce reliance on fossil fuels, especially during peak demand periods. In residential settings, predictive models can be integrated into smart home systems to guide energy-efficient

behavior. On a larger scale, intelligent energy management systems can automate tasks such as adjusting tariffs or activating distributed energy resources based on demand forecasts [7]. These applications demonstrate how AI models can enhance both user experience and system-wide sustainability.

Nonetheless, several challenges remain. Issues of data governance, privacy, and cybersecurity arise as the energy sector becomes more data-reliant. Ensuring ethical data use, obtaining user consent, and safeguarding digital infrastructure are all critical concerns. Furthermore, while models like XGBoost offer strong predictive power, their complexity can reduce interpretability, a vital factor for stakeholder trust and regulatory compliance. Another challenge is the unpredictable influence of external factors such as weather, economic shifts, or social events. While many of these can be quantified and modeled, their inherent uncertainty limits forecasting precision [8]. Adaptive models capable of learning in real time may help address this variability, though such systems bring their own implementation complexities. Despite these hurdles, the benefits of integrating machine learning into energy systems are substantial. By advancing forecasting capabilities, this research contributes to the modernization of power systems and supports the global transition to sustainable energy. AI-driven models like XGBoost hold the potential to reshape how energy is managed, making electricity systems more flexible, efficient, and environmentally friendly.

This study specifically focuses on the application of XGBoost to residential electricity consumption forecasting. Drawing on expertise from both electrical engineering and data science, it aims to design a robust, practical model capable of handling the intricacies of real-world energy data. Through careful data analysis, modeling, and validation, the research highlights the potential of intelligent systems to improve grid resilience and promote energy sustainability.

I.1 Motivation and Goal

The rapid evolution of energy systems into smart grids has intensified the need for accurate and adaptive electricity consumption forecasting. Growing urbanization, the integration of renewable energy, and the rise of electric vehicles have made electricity demand more variable and complex. Traditional forecasting methods, though historically useful, struggle to capture the nonlinear and dynamic nature of modern consumption patterns, particularly in residential settings where usage is influenced by diverse and unpredictable factors.

In response to these challenges, machine learning, specifically, ensemble methods like Extreme Gradient Boosting (XGBoost), offers a powerful alternative. XGBoost is well-suited

for structured time series data, capable of handling missing values, managing large datasets, and modeling intricate patterns in electricity usage. Its robust performance and scalability make it an ideal tool for forecasting in real-world energy applications.

This research aims to develop and validate an XGBoost-based model to forecast residential electricity consumption at two critical temporal scales: short-term (hourly) and medium-term (daily). These timeframes are essential for tasks ranging from real-time grid control and load balancing to strategic planning and resource allocation. The model is trained using real-world smart meter data, with preprocessing steps to clean and enrich the dataset and hyperparameter tuning to ensure optimal performance.

The overarching goal is to demonstrate how machine learning can enhance energy forecasting accuracy, enabling smarter grid operations and more sustainable energy use. By providing precise demand predictions, the model supports utilities in decision-making, assists consumers in managing consumption, and contributes to broader energy efficiency and climate objectives.

This study is motivated by the dual imperative of improving technical forecasting capabilities and supporting societal efforts toward sustainable, resilient, and intelligent energy systems.

I.2 Contribution

This thesis makes significant contributions to the field of electricity consumption forecasting by developing a robust and accurate model based on advanced machine learning techniques. Unlike traditional methods that often rely on linear assumptions and require extensive manual feature engineering, this research leverages the power of the XGBoost algorithm to capture complex, nonlinear patterns inherent in residential electricity usage. The model is designed to work effectively even when faced with common data challenges such as missing values and noisy observations, which are typical in real-world smart meter datasets. By focusing on both short-term (hourly) and medium-term (daily) forecasting horizons, the study addresses practical needs across different operational and planning scenarios within smart grid environments. Additionally, the integration of supplementary data sources, including weather, socioeconomic factors, and time-related features, enhances the predictive accuracy and contextual relevance of the model. Comprehensive hyperparameter tuning through Bayesian optimization and grid search further refines the model's performance, ensuring optimal balance between bias and variance. The evaluation employs multiple statistical metrics to provide a rigorous assessment of forecasting quality. Beyond theoretical advancements, this research offers valuable insights

and recommendations for implementing AI-driven forecasting models in real-world energy management systems, emphasizing their potential to support sustainable and efficient grid operations. Our main contributions in this context are outlined below:

- Development of an XGBoost-based forecasting model with advanced data preprocessing and feature integration to handle real-world smart grid data complexities.
- Application of sophisticated hyperparameter optimization techniques to improve accuracy across multiple forecasting horizons.
- Comprehensive evaluation and practical recommendations for integrating AI models into energy management, supporting improved grid stability and efficiency.

The findings demonstrate that the proposed XGBoost model is an effective and reliable method for improving electricity consumption forecasting in smart grid contexts.

I.3 Structure of the Thesis

This thesis is systematically organized into six chapters to provide a clear and coherent flow of the research. The first chapter introduces the general background, highlighting the importance of electricity consumption forecasting in smart grid environments and outlining the motivation and scope of the study. Chapter two delves into the concept of smart grids, exploring their architecture, components, and significance in modern energy systems. This sets the foundation for understanding the context in which electricity forecasting is applied.

Chapter three reviews the current state of energy forecasting methodologies, covering both traditional statistical approaches and recent advances in machine learning techniques. The fourth chapter presents the core contributions of this research, detailing the development of an XGBoost-based forecasting model, data preprocessing strategies, and optimization procedures. The fifth chapter provides a comprehensive analysis of the results, evaluating the model's performance across different scenarios and discussing its practical implications. Finally, chapter six concludes the thesis by summarizing the key findings, addressing the limitations, and suggesting future research directions to advance the field further.

Chapter II

Smart Grid

The rapid evolution of energy systems on a global scale has underscored the urgent necessity for a more intelligent, adaptable, and responsive electrical grid infrastructure. Existing traditional power grids, many of which were conceived and constructed several decades ago, are increasingly ill-equipped to cope with the dynamic demands of modern energy usage. These aging systems are under mounting strain due to escalating electricity consumption, stringent environmental regulations, and the growing incorporation of distributed energy resources such as solar panels, wind turbines, and battery storage systems. In light of these emerging challenges, the concept of the Smart Grid has gained prominence as a comprehensive and transformative approach. It represents the convergence of advanced digital communication technologies, real-time data analytics, and automated control mechanisms with conventional power systems. The objective is to enhance operational efficiency, ensure higher levels of reliability, and support long-term sustainability across the entire energy value chain. This innovative energy paradigm introduces a new framework for electricity generation, transmission, distribution, and end-user engagement. By facilitating bidirectional flows of both electricity and information, the Smart Grid enables more precise demand forecasting, better integration of renewable sources, and improved grid stability. As a result, it lays the foundation for a future-oriented energy ecosystem that is not only flexible and interactive but also deeply aligned with environmental stewardship and resilience goals.

II.1 Electric energy system

One of the most significant scientific and technological breakthroughs that has profoundly influenced human civilization and transformed the fabric of modern daily life is the discovery and widespread utilization of electricity. From residential lighting and heating to industrial operations and digital communications, electric energy has become an indispensable foundation for contemporary society. As global population figures continue to rise steadily and the world economy undergoes persistent and accelerated growth, the demand for electrical power is ex-

periencing a corresponding and substantial surge. This trend reflects the deepening interdependence between economic development, urbanization, technological advancement, and energy consumption patterns. According to projections provided by the International Energy Outlook 2013 (IEO2013), global net electricity generation is forecasted to increase significantly, from 20.2 trillion kilowatt-hours (KWh) recorded in 2010 to an estimated 39.0 trillion KWh by the year 2040. This dramatic escalation represents a striking 93% growth over the three-decade period (Fig II.1) [9]. Such a remarkable increase underscores the urgency for nations and energy stakeholders to modernize and expand power infrastructure, enhance generation capacity, and adopt more efficient and sustainable practices in electricity production and distribution to meet future global energy needs responsibly.

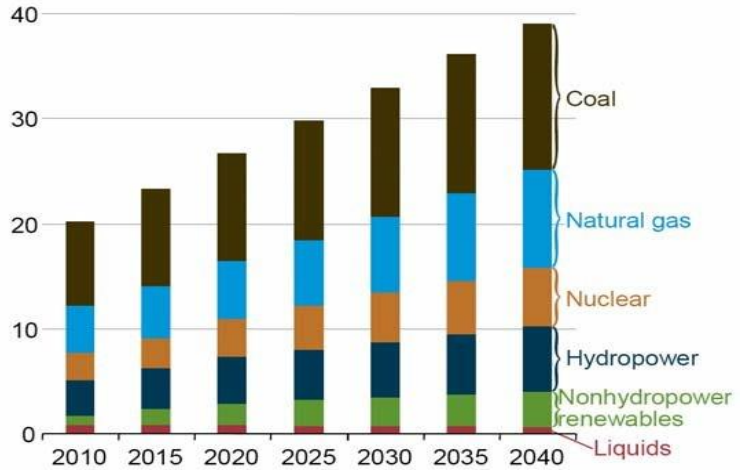


Fig II.1: World net electricity generation by energy source, 2010-2040 (trillion kilowatt-hours) [9]

II.2 Electrical grid operation

The electrical grid functions as the essential transportation infrastructure for electrical energy, facilitating the transfer of power from centralized generation facilities to a vast array of residential, commercial, and industrial consumers. This complex network is the backbone of modern electricity distribution and is designed to ensure the reliable, continuous, and efficient delivery of electric power across extensive geographical areas. Its primary objective is not only to transmit energy from power plants to end-users but also to maintain system stability, optimize overall performance, and minimize operational disruptions. One of the major technical challenges within this network is the occurrence of energy losses, a substantial portion of which is attributed to Joule losses. These losses are a consequence of the inherent electrical resistance in transmission conductors, where energy is dissipated as heat during the flow of current. Minimizing such losses is a priority in grid design and operation; however, reducing resistivity directly, by employing conductors with larger cross-sectional areas or advanced materials with

lower intrinsic resistance, can lead to significant increases in infrastructure costs. Consequently, the most practical and economically advantageous strategy involves increasing the transmission voltage. By raising the voltage level, the same amount of power can be transmitted with a lower current, thereby substantially reducing resistive losses. For this reason, high-voltage transmission lines are deployed at the output terminals of power generation plants to carry large-scale electric power efficiently over long distances.

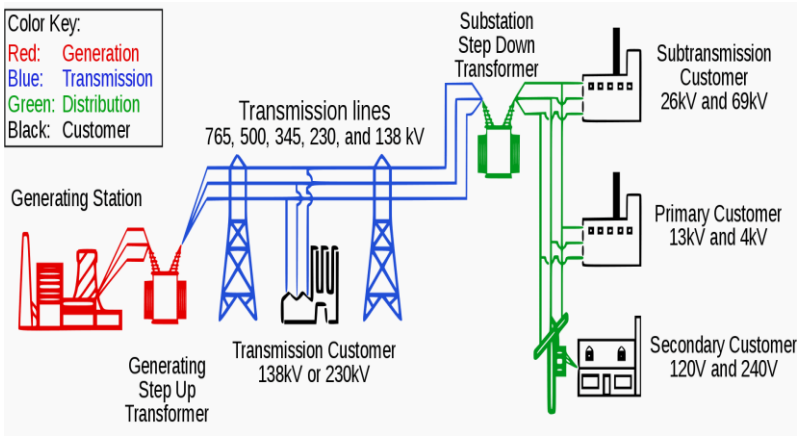


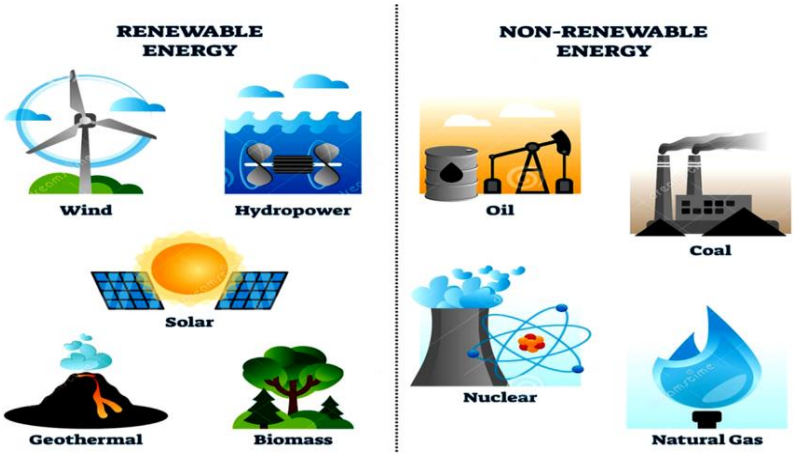
Fig II.2: Basic structure of the electric power system [10]

As electricity approaches end-use locations, the voltage is methodically reduced through a series of step-down transformers and progressively lower-voltage lines. This hierarchical voltage regulation ensures the safe and effective delivery of energy to consumers at appropriate and usable voltage levels. The traditional architecture of this electrical energy system is illustrated in Fig II.2.

The electrical grid consists of four subsystems, each corresponding to different voltage levels: generation, transmission, distribution, and consumption.

• **Generation:** Electrical energy generation is the process of transforming different types of primary energy into electrical energy, typically at medium voltage levels suitable for transmission and distribution. This fundamental step in the power system chain serves as the starting point for delivering electricity to the grid and, ultimately, to the end consumers. The sources used for electrical generation are broadly categorized into two principal groups, as illustrated in Fig II.3: non-renewable and renewable energy sources. Non-renewable sources comprise conventional fuels such as coal, natural gas, fuel oil, and other fossil-based resources, as well as nuclear power derived from uranium fission. These sources have historically formed the backbone of global electricity production due to their high energy density and established infrastructure. However, they are finite in nature and contribute significantly to environmental issues such as greenhouse gas emissions and long-term ecological degradation. On the other

hand, renewable energy sources are derived from natural processes that are continuously replenished. These include solar energy harnessed through photovoltaic systems, wind energy captured via turbines, hydroelectric power generated from moving water, biomass energy from organic materials, and marine-based sources such as wave and tidal energy. Renewable sources are increasingly favored for their sustainability and minimal environmental impact, playing a pivotal role in the transition toward cleaner and more resilient energy systems. The growing integration of these diverse energy resources is reshaping the landscape of electricity generation, necessitating more flexible and adaptive grid technologies to accommodate their variability and distributed nature.



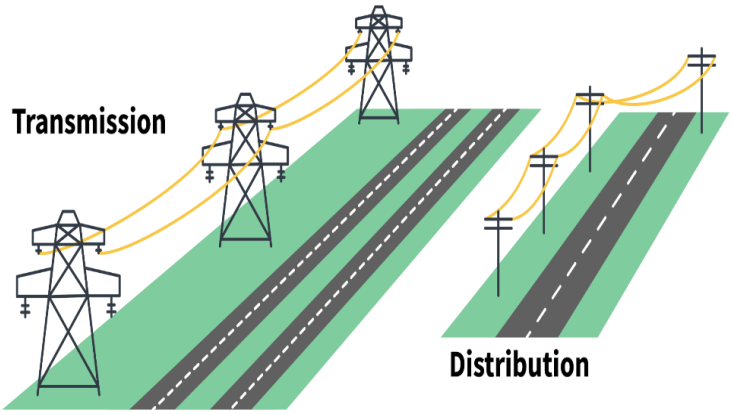
https://edukemy.com/blog/sources-of-energy-renewables-and-energy-conservation-upsc-environment-notes/

Fig II.3: Energy sources

• **Transmission and Distribution:** Transmission and distribution (T&D) networks play a vital role in the electrical power system by ensuring the efficient and reliable delivery of electrical energy from generation facilities to end-users. This process spans multiple voltage levels, very high, high, and medium, which are carefully selected based on distance, load requirements, and system design. As illustrated in Fig II.4, the transmission phase involves the long-distance transport of electricity from power plants to substations using very high and high voltage lines, which help minimize energy losses by reducing current flow. Once the energy reaches regional substations, the distribution phase begins, where the voltage is stepped down to medium levels suitable for residential, commercial, and light industrial consumption. This hierarchical structure ensures that power is delivered safely and efficiently across wide geographical areas, adapting to the varying electrical needs of consumers.

The transmission of energy within these networks is conducted primarily through three-phase alternating current (AC) systems, in which the voltages follow sinusoidal waveforms. These systems are preferred due to their ability to deliver consistent power, support heavy loads,

and allow for more efficient energy conversion. Each phase of the three-phase system is offset by 120 degrees, providing a balanced and stable flow of electricity. The key characteristics defining these systems include frequency, standardized globally at either 50 Hz or 60 Hz depending on the region, as well as specific voltage ratings that align with international and national electrical standards. Additionally, the terminal configurations of these systems typically follow either delta (Δ) or star (Y) arrangements, each offering distinct advantages in terms of grounding, fault tolerance, and load distribution. Together, these design principles enable the T&D network to maintain high performance, system stability, and adaptability across diverse operating conditions.



https://stanwichenergy.com/insights/new-federal-rule-boosts-renewable-energy-transmission-what-it-means-for-electricity-costs

Fig II.4: Transmission and Distribution energy lines

• **Consumption:** Electricity consumption across modern economies is generally categorized into four primary sectors, each reflecting distinct usage patterns and energy demands. The first is the residential sector, which encompasses all forms of housing, from single-family homes to multi-unit apartment buildings. This sector utilizes electricity for a variety of everyday applications, including lighting, space heating and cooling, cooking, refrigeration, and powering a wide array of household appliances and electronic devices. As living standards rise and digital technologies become more embedded in daily life, electricity demand in the residential sector continues to grow. The second sector is the commercial sector, which includes a broad range of non-residential buildings such as offices, retail establishments, schools, hospitals, and government facilities. Electricity in this sector is used to power lighting systems, HVAC (heating, ventilation, and air conditioning), office equipment, elevators, and various specialized machinery needed for service delivery. Because many commercial buildings operate for extended hours or even continuously, their cumulative energy usage is significant. The third category is the industrial sector, which comprises activities related to manufacturing, mining, agriculture,

and construction. This sector is typically the largest consumer of electricity due to the high-power demands of production machinery, process heating and cooling, automated systems, and heavy equipment. Electricity is a critical input for maintaining productivity, ensuring quality control, and supporting continuous operations in industrial settings. Lastly, the transportation sector, while still predominantly reliant on fossil fuels, is experiencing a gradual but notable shift toward electrification. This includes the growing deployment of electric vehicles (EVs), electric buses, and rail systems powered by electricity from the grid. As technological advancements continue and charging infrastructure expands, the share of electricity in transportation energy use is expected to increase steadily. Together, these four sectors provide a comprehensive view of how electrical energy is consumed across different facets of society, shaping policy decisions and infrastructure planning for future energy systems.



https://www.nrel.gov/analysis/electrification-futures.html

Fig II.5: Types of electrical consumers

Traditional energy grids, while once effective in meeting the demands of earlier industrial and residential needs, now face a multitude of significant challenges that render them increasingly inadequate in today's rapidly evolving energy environment. These legacy systems were originally engineered around a linear, one-way model of energy distribution, where electricity flows unidirectionally from centralized power generation plants to end-users. This rigid structure inherently limits operational flexibility and contributes to elevated transmission losses, particularly over long distances. The result is a decrease in overall system efficiency, as a considerable portion of the generated energy is lost before it even reaches the consumer.

One of the major drawbacks of traditional grids is the absence of real-time monitoring and control mechanisms. Without the ability to observe and react to changes in the network as they occur, utility operators are often slow to detect and respond to system faults. This latency in

fault detection can lead to prolonged power outages, decreased service reliability, and higher maintenance expenditures due to the reactive rather than proactive nature of repairs. Furthermore, these grids were designed around large, centralized, fossil fuel-based power stations, such as coal or gas-fired plants. This centralized configuration poses significant challenges when attempting to incorporate decentralized and variable renewable energy sources like solar photovoltaics or wind turbines, which are often distributed and intermittent in nature.

The traditional grid's inability to adapt to these modern energy sources not only hampers the integration of cleaner technologies but also exacerbates environmental degradation through continued dependence on carbon-intensive generation. High operating costs, compounded by aging infrastructure and the need for frequent manual intervention, add to the inefficiencies. Additionally, consumers connected to traditional grids typically have little to no insight or control over their energy consumption, leading to a passive energy use model that lacks engagement or flexibility. These compounding limitations underscore the urgent necessity for a more modern, intelligent, and responsive grid infrastructure that can meet contemporary energy demands while supporting sustainability and resilience objectives.

II.3 Smart grid

To effectively overcome the limitations and inefficiencies inherent in traditional power systems, the transition to smart grid technology has emerged as a critical and transformative necessity. Unlike conventional grids, smart grids are designed to optimize the generation, distribution, and consumption of electrical energy through the use of advanced digital communication, automation, and control technologies. One of their most defining features is the facilitation of two-way energy flow, which allows not only centralized power generation but also decentralized and consumer-based generation, such as rooftop solar panels, to feed electricity back into the grid. This bidirectional capability significantly enhances operational flexibility and responsiveness.

In addition, smart grids are equipped with real-time monitoring and diagnostics, enabling grid operators to detect and respond to faults, load imbalances, and system disturbances almost instantaneously. This capability drastically reduces the duration and frequency of power outages, while also lowering maintenance costs through predictive and condition-based servicing. The intelligent nature of these grids supports a more seamless integration of renewable energy sources, which are inherently variable and geographically distributed. By actively managing fluctuations in generation and demand, smart grids ensure a more stable and reliable power supply, even as the share of renewables continues to rise.

Smart grids also contribute to cost reduction by improving the overall efficiency of energy transmission and distribution, minimizing losses and optimizing asset utilization. Beyond technical benefits, they enhance energy security by making the grid more resilient to cyber threats, physical attacks, and natural disasters through decentralized control and advanced encryption technologies. Importantly, smart grids place consumers at the center of the energy ecosystem by providing access to real-time consumption data, enabling better decision-making and fostering active participation in demand-side management. Smart meters and home energy management systems empower users to monitor, adjust, and even automate their energy usage based on price signals or environmental preferences.

Overall, the adoption of smart grid systems marks a decisive step toward a more sustainable, resilient, and consumer-centric energy landscape. This modern infrastructure not only addresses the shortcomings of traditional grids but also supports long-term energy policy goals related to decarbonization, efficiency, and innovation.

II.3.1 Definition

A smart grid represents a next-generation electricity network that incorporates advanced digital technologies to enhance the generation, delivery, and consumption of electrical energy. Unlike traditional grids, which rely on unidirectional energy flow and limited communication, the smart grid enables two-way digital communication between utilities and consumers. This bidirectional communication is fundamental to the system's ability to operate dynamically and respond to real-time conditions. Through the integration of sensors, automation, and sophisticated software, smart grids facilitate comprehensive monitoring, analysis, control, and communication across all segments of the energy supply chain, from power generation to end-use consumption.

The overarching goal of this intelligent infrastructure is to optimize operational efficiency, reduce overall energy consumption, and minimize system costs, all while enhancing the reliability and transparency of the power supply. By enabling the precise tracking of electricity flow and system performance in real time, smart grids empower utilities to make informed, data-driven decisions that prevent failures, balance supply and demand, and ensure grid stability. In addition, the system supports automated demand response, dynamic pricing models, and distributed generation management, all of which contribute to a more flexible and sustainable energy landscape.

The development of smart grid technologies was driven by the need to overcome the critical limitations of conventional electrical grids, particularly in their inability to integrate variable

renewable energy sources, detect faults quickly, or involve consumers in energy management. One of the cornerstone components of this modernization effort is the smart meter [11], a device that enables accurate, real-time measurement of electricity usage and supports two-way communication between the utility and the user. Smart meters provide detailed consumption data, allow for remote metering, and support automated billing and outage notifications.

This comprehensive transformation influences every stage of the power system, including generation, transmission, distribution, and consumption, by embedding intelligence and adaptability into each layer. As a result, the smart grid forms the foundation of a future-oriented energy infrastructure that is not only more robust and responsive but also better aligned with environmental sustainability and technological innovation. The structure and operation of a modern smart grid system are visually represented in Fig II.6.



Fig II.6: Smart grid [11]

● In the generation part of a smart grid, renewable energy sources like solar and wind are increasingly integrated alongside traditional bulk generation methods, such as coal and natural gas, to create a more sustainable and diversified energy mix. Because solar and wind output can be variable and unpredictable, grid-scale energy storage systems play a vital role by balancing the power flow, acting as a source when discharging energy during peak demand or low generation periods, and as a load when charging during times of excess power. This flexibility helps maintain stability and smooth operation of the grid. On the transmission side, the extensive use of advanced phasor measurement units (PMUs) enables real-time, wide-area monitoring of system conditions, providing operators with detailed, time-synchronized data to quickly detect faults or disturbances. This enhanced visibility significantly improves the overall reliability and

responsiveness of the grid, making it better equipped to handle the complexities of modern electricity networks.

● On the distribution side, microgrids, small, localized energy networks, are becoming increasingly important for integrating distributed energy resources (DERs) such as rooftop solar panels, small wind turbines, and energy storage systems. These microgrids can operate independently or connected to the main grid, providing flexibility, improved reliability, and resilience, especially during outages or peak demand periods. Alongside this, the widespread deployment of smart meters will facilitate real-time, two-way communication between utilities and consumers, enabling more accurate energy usage monitoring, dynamic pricing, and improved demand management. Together, microgrids and smart meters empower both utilities and customers to actively participate in energy management, paving the way for a more efficient, responsive, and decentralized distribution system.

● On the utilization side, the integration of smart appliances equipped with network-based control capabilities is revolutionizing the way consumers interact with the electrical grid. These intelligent devices can communicate with the grid to optimize energy usage based on real-time price signals, demand response programs, or availability of renewable energy, allowing for more efficient and flexible consumption patterns. Additionally, the rapid growth of electric vehicles (EVs) is poised to significantly impact grid dynamics by introducing new, controllable loads that can be managed through smart charging infrastructure. EVs not only represent a shift toward cleaner transportation but also offer potential benefits such as vehicle-to-grid (V2G) services, where parked EV batteries can feed energy back into the grid during peak demand. Together, smart appliances and electric vehicles empower consumers to play an active role in energy management, fostering a more interactive, efficient, and sustainable energy ecosystem.

● With the deployment of a smart grid infrastructure, a wide range of innovative applications and services are rapidly emerging that aim to enhance energy efficiency, convenience, and sustainability. Among these are smart homes and smart buildings, which leverage advanced sensors, automation, and communication technologies to optimize energy use, improve comfort, and reduce costs by intelligently managing lighting, heating, cooling, and appliances based on occupancy and user preferences. Microgrids, which are localized energy systems capable of operating independently or in coordination with the main grid, provide increased resilience and flexibility by integrating distributed energy resources and enabling localized energy management. Additionally, automated demand response (DR) programs are becoming increasingly sophisticated, allowing utilities to dynamically adjust consumer loads during peak periods

through real-time signals and incentives, thereby balancing supply and demand more effectively. These applications collectively demonstrate how smart grid technology is transforming the energy landscape into a more adaptive, efficient, and consumer-centric system.

To better understand the key differences and advancements brought by modern technology, Fig II.7 illustrates a detailed comparison between the traditional electrical grid and the smart grid.

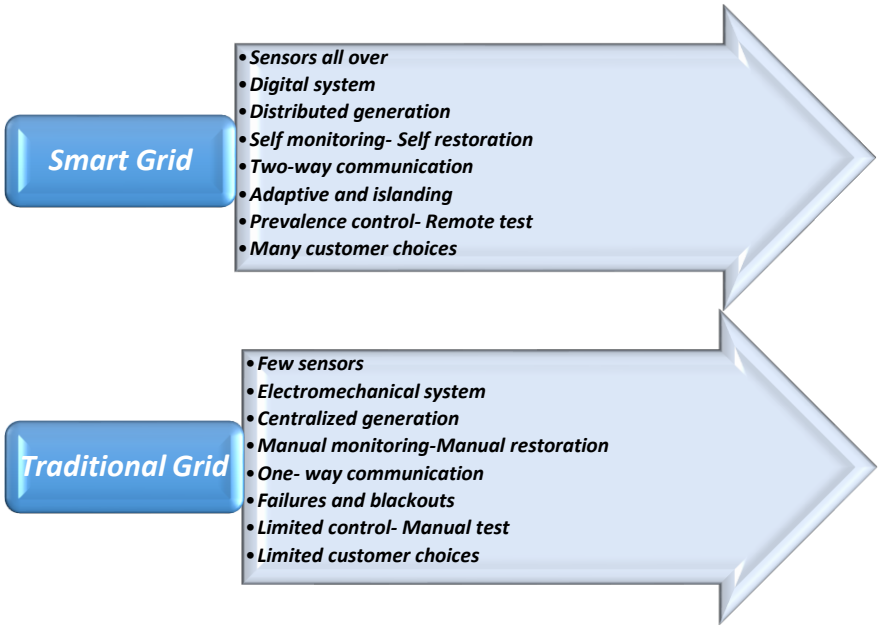


Fig II.7: Comparison of traditional grid and smart grid

II.3.2 Smart grid benefits

Ensuring a secure energy supply within the network brings the issue of energy efficiency to the forefront. Consumers also play a crucial role in this process, starting from their subscription choices to the efficient delivery of electricity. Subscribers can contribute by selecting energy-efficient appliances and using them during off-peak hours. Maintaining system-wide balance, as defined by the network structure, becomes increasingly evident in this context. By adopting more sustainable consumption habits and shifting energy use over time, consumers can enhance the longevity of the network and help reduce technical losses, thereby supporting its overall efficiency and resilience.

- **Reliability:** Smart Grids have the potential to detect faults in real time and enable self-healing within the system. The success of a grid largely depends on meeting customer needs, which are primarily measured in terms of reliability that is, the ability to provide a continuous, error-free supply of electricity [12-13]. Unlike conventional grids, Smart Grids can monitor, record, and analyze data to assess and enhance service reliability. They also support remote

monitoring and management, particularly beneficial in systems integrating hybrid energy generation sources. While traditional grids face challenges in integrating renewable resources, microgrids, and demand response mechanisms, these issues are effectively addressed by Smart Grids. With advances in communication technologies, Smart Grids offer significantly improved reliability and operational efficiency.

- **Security:** With the rise of automation, remote monitoring and control capabilities have become fundamental components of modern grid systems, enabling utilities to manage and optimize the electrical network more efficiently and responsively. These technologies allow for real-time data collection, fault detection, and automated control actions without the need for manual intervention, significantly improving grid performance and reliability. However, the increasing reliance on digital communication and interconnected devices also introduces new vulnerabilities, exposing the grid to potential cyberattacks [13-14]. Malicious actors can exploit weaknesses in the network to disrupt operations, compromise data integrity, or even cause widespread outages. Consequently, cybersecurity has emerged as one of the most critical and complex challenges in the widespread deployment of smart grid technology. Protecting the grid requires robust security protocols, continuous monitoring, and advanced defense mechanisms to safeguard the infrastructure against evolving cyber threats while maintaining the benefits of automation and connectivity.
- **Demand side management system:** Smart Grids enable two-way communication between users and the grid, allowing consumers to actively participate in energy management. This interaction empowers users to utilize electricity more economically and efficiently. Smart Grids enhance efficiency not only on the demand side but also at the distribution level. By enabling load reduction or shifting during peak periods, Smart Grids help alleviate stress on the system [13]. Demand Side Management (DSM) plays a crucial role in this context by allowing consumers to make informed decisions about their energy use. It also assists energy providers in reducing peak load demand and reshaping the load profile [15]. DSM strategies include demand task scheduling, the use of stored electricity, and real-time pricing mechanisms [16]. While DSM increases the operational complexity of the power system and redistributes load, it does not necessarily reduce the total energy consumed by appliances. The effectiveness of DSM is higher when the system operates near its maximum capacity; conversely, it is less impactful when spare capacity is available. Thus, the generation capacity of the grid is a critical factor influencing the success of DSM programs [17,18].

However, DSM and Demand Response (DR) programs face several barriers. On the consumer side, challenges include limited awareness of DR programs, "response fatigue" from tracking

frequently changing prices, and long payback periods for smart meter investments [19,20]. On the producer side, barriers include unclear responsibilities for promoting DR programs, revenue losses from reduced peak usage, and a lack of formal mechanisms to recover initial infrastructure investments [19,21].

Additionally, regional disparities can hinder the development of unified DSM policies in some countries [22]. Moreover, a rebound in energy usage following high price signals may lead to even higher demand peaks, potentially undermining the goals of DSM [23].

- **Metering:** Metering plays a crucial role in enabling two-way communication between consumers and distributors in the Smart Grid framework. Smart meters not only allow distributors to im-plement more accurate billing systems, but also empower consumers to monitor and manage their electricity usage. These devices are equipped with advanced sensors that facilitate au-tomation, power quality monitoring, and real-time outage notifications. Several factors drive the adoption of smart meters, including rising electricity prices following market deregula-tion, consumer dissatisfaction with conventional services, and mandates for monthly meter-ing [13]-[24]. In traditional grid systems, communication was primarily handled by Supervi-sory Control and Data Acquisition (SCADA) systems, which offered centralized monitoring and control capabilities at second-level scan rates. However, SCADA systems are not cost-effective across various levels of power distribution, particularly at the utility level [25].

In contrast, Advanced Metering Infrastructure (AMI) offers a real-time communication solution by collecting consumer usage data and establishing a robust network between the grid and utility. AMI plays a pivotal role in modernizing legacy grid systems by bridging consumers with distribution networks. It supports a wide range of advanced functionalities, including outage management, integration of electric vehicles and smart appliances, monitoring of transformers and feeders, and fault detection and isolation [13][26].

- **Micro-grids and integration of renewable resources:** A microgrid is a small-scale power supply network designed to provide electricity to a localized area, such as a small community. It enables local power generation to serve local loads and typically incorporates various small-scale energy sources, making it highly flexible, efficient, and resilient. Con-nected both to local generating units and the main utility grid, a microgrid helps prevent power outages by offering backup support. Excess power can either be sold back to the utili-ty grid or stored in energy storage systems. The size of a microgrid can vary from a few kil-owatts (kW) for residential estates to several megawatts (MW) for municipal-scale regions.

To ensure voltage stability and reduce power losses in the system, several technical solutions are available. Among them, Flexible AC Transmission System (FACTS) devices are widely

used for reactive power compensation [27]. Deploying renewable Distributed Generation (DG) sources near load centers contributes significantly to system stability, minimizes power losses, and lowers operational costs [28]. Furthermore, renewable energy sources are more environmentally friendly compared to traditional fossil-fuel-based power plants [29].

Optimal planning of DG integration is crucial for enhancing the performance of distribution networks. This includes improving voltage stability, reducing losses, increasing system reliability, and ensuring economic viability [30]. In the future, renewable energy is expected to become the primary driver of electricity generation, helping to conserve fossil fuels and reduce the environmental impact of conventional power generation. Common renewable energy sources include wind, photovoltaic (PV) systems, biomass, nuclear power, and small hydropower plants [31]. Among these, solar and wind energy are the most prominent due to their natural abundance.

In this context, a microgrid is defined as a coordinated system of interconnected loads and distributed energy resources (DERs) capable of supplying power to neighboring areas. Through advanced local power management strategies, microgrids can effectively integrate and coordinate multiple types of renewable and distributed energy technologies. A conceptual diagram illustrating the role of a microgrid within a smart grid framework is shown in Fig II.8 [32].

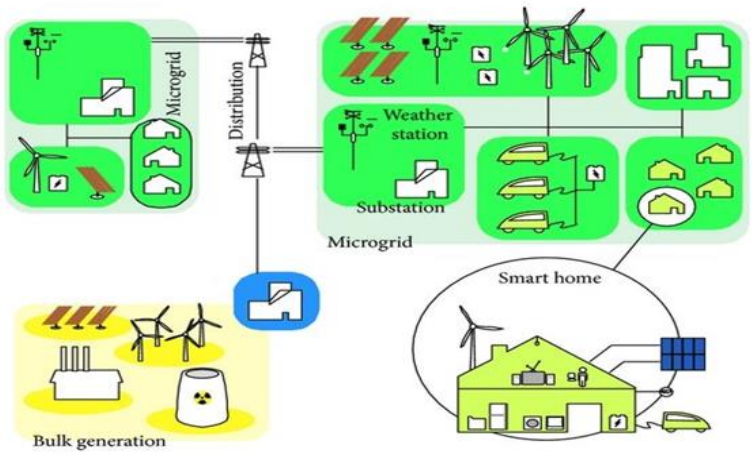


Fig II.8: A conceptual illustration of a microgrid [32]

II.3.3 Issues and challenges in smart grids

The smart grid is expected to revolutionize existing electrical grid by allowing two-way communications to improve efficiency, reliability, economics, and sustainability of the generation, transmission, and distribution of electrical power. However, issues associated with communication and management must be addressed before full benefits of the smart grid can be achieved. Furthermore, how to maximize the use of network resources and available power,

how to ensure reliability and security, and how to provide self-healing capability need to be considered in the design of smart grids.

- battery wearing, low penetration of the Extra Voltages, new battery technologies.
- costs, consumer engagement.
- distributed generation in distribution system and distribution system automation.
- physical security, cyber security, vulnerabilities in metering data.

II.3.4 Smart grid component

There is no specific start of Smart Grid. This concept was start evolving with the start of distribution system of electrical networks. By the time different requirements were needed like control, monitor, prices and services of transmission and distribution of electrical power. Normally, Smart Grid implementation is associated with the installation of smart meter. In 1970s and 80s they were used to send the information of consumer back to the grid [13]. But the most important and fundamental need which is still under consideration even with latest advancement is reliability and efficiency of energy transmission and distribution via electric power grid. But in the latest advancement research is undergo that grids and network systems should not limit to transmission and distribution but also play a vital role in generating clean and sustainable energy in order to reduce greenhouse gases and carbon foot print [13]. It provides its stakeholder an opportunity to maximize the efficiency, reliability, economic performance and security of their electrical network. An overview of its architecture is shown in Fig II.9.

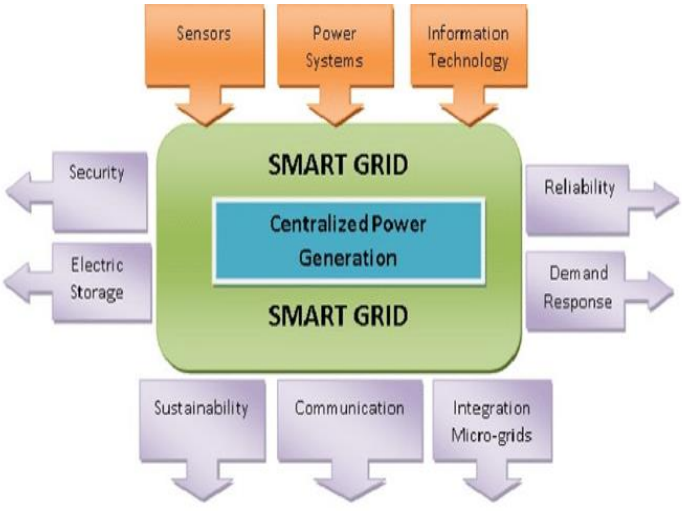


Fig II.9: Overview of Smart Grid design

1. Smart devices interface component

Electronic devices that connect to other devices or networks using a variety of wireless communication protocols and are capable of functioning both interactively with users or other systems and autonomously without direct human intervention are commonly referred to as smart devices. These intelligent devices are essential components within Smart Grid systems, particularly within the generation segment, where they facilitate continuous monitoring and precise control of energy production processes. By enabling real-time data collection, processing, and communication, smart devices allow for rapid detection of anomalies, optimization of generation output, and seamless coordination between distributed energy resources. Their ability to operate autonomously while maintaining interactive capabilities ensures that energy management is both efficient and adaptive, significantly enhancing overall system performance and contributing to the stability and sustainability of modern power grids.

2. Storage component

Because renewable energy sources (RES) such as solar and wind are inherently intermittent and often do not align with peak energy consumption periods, it is crucial to develop and implement effective energy storage solutions to bridge this gap. Energy storage systems play a vital role in improving the reliability and resiliency of both the utility grid and end consumers by storing excess energy generated during periods of high production and releasing it during times of low generation or high demand. This capability helps to smooth out fluctuations in supply and demand, reduce strain on generation resources, and ensure a consistent power supply. Several key energy storage technologies have been developed and deployed to meet these needs, including flow batteries, which offer scalability and long cycle life; ultra-capacitors, known for their rapid charge and discharge capabilities; pumped hydro storage, which leverages gravitational potential energy; superconducting magnetic energy storage, providing high-efficiency energy retention; and compressed air energy storage, which stores energy in the form of pressurized air for later use. Together, these technologies form a critical backbone for integrating renewable energy into modern power systems while maintaining grid stability and performance.

3. Monitoring and control technology component

The monitoring and control technology component encompasses a range of advanced devices and systems designed to enhance the operational intelligence of the power grid by enabling features such as self-monitoring, self-healing, predictability, and adaptability in power generation and distribution. These smart networks and intelligent devices continuously collect

and analyze real-time data, allowing the grid to detect faults, anticipate potential issues, and respond autonomously to maintain stability and reliability. By addressing challenges such as system instability, congestion, and unexpected disruptions, these technologies significantly improve grid resilience and performance. Their ability to adapt dynamically to changing conditions ensures that energy delivery remains consistent and efficient, even in the face of fluctuating demand or supply. Collectively, this monitoring and control infrastructure forms a critical foundation for the development of modern, robust, and responsive power systems within the smart grid framework.

4. Demand side management component

Demand Side Management (DSM) and energy efficiency measures are strategic approaches aimed at actively modifying consumer electricity demand to achieve several important objectives, including reducing overall operating costs by lessening dependence on costly peak generators and postponing the need for new capacity additions. By encouraging consumers to shift or reduce their energy usage during peak periods, DSM improves the reliability of power generation and distribution systems, helping to prevent overloads and outages. Additionally, these strategies contribute significantly to lowering greenhouse gas emissions by decreasing the operation of fossil-fuel-based power plants during high-demand times. One of the key benefits of DSM is its ability to reshape and flatten the utility load curve, smoothing out spikes in electricity consumption and creating a more balanced and manageable demand profile. To maximize the effectiveness of DSM programs, it is essential to implement standardized protocols that enable seamless interaction between utilities and consumers, which is facilitated by two-way communication technologies embedded within smart grid systems. These communication frameworks allow for real-time data exchange and automated demand response, empowering consumers to participate actively in energy management while providing utilities with the tools to optimize grid performance dynamically.

II.4 Conclusion

This chapter has provided a comprehensive overview of the fundamental concepts, key benefits, and notable challenges associated with the development and implementation of smart grid technology. Gaining a solid understanding of these foundational elements is crucial as they form the basis for more advanced discussions on how to effectively manage and forecast energy within modern power systems. Building on this knowledge, the next chapter will focus on critical topics related to energy management and forecasting techniques, exploring innovative

methods and tools designed to enhance operational efficiency, optimize resource utilization, and improve the overall reliability of smart grid operations. This progression aims to equip readers with a deeper insight into how smart grids can meet growing energy demands while supporting sustainability and resilience goals.

Chapter III

Household Energy Management

With the growing global demand for energy and increasing operational costs, energy efficiency has become a key concern for energy and environmental experts. Accurate forecasting and effective management of energy consumption in buildings especially those utilizing multiple energy sources such as the public grid, solar panels, and private generators are essential. Artificial Intelligence (AI) provides advanced capabilities for analyzing both historical and real-time data, enabling precise prediction of consumption patterns and optimal selection of energy sources based on efficiency and cost-effectiveness. This chapter explores the integration of such technologies to forecast building energy consumption, with the objective of achieving dynamic energy source balancing, reducing reliance on traditional grids, improving energy performance, lowering operational costs, and supporting long-term sustainability.

III.1 Smart Energy Management in Homes: The Role of AI and Forecasting

In the face of increasing energy consumption and environmental challenges, effective home energy management and forecasting have become essential for promoting energy efficiency and sustainability. Unregulated consumption results in resource waste and elevated costs, highlighting the need for intelligent solutions grounded in accurate consumption pattern prediction. By utilizing Artificial Intelligence (AI) and time-series forecasting models, households can optimize energy use, dynamically switch between energy sources, and minimize dependence on the public grid during peak demand periods. AI-driven energy management not only enhances household efficiency but also contributes to reducing the burden on national power infrastructures and advancing sustainable development. The integration of predictive technologies into residential energy systems represents a critical advancement toward reducing environmental impact and lowering operational expenses.

III.1.1 Effects of unregulated energy consumption

Home energy management has become a critical necessity in the contemporary era due to its significant implications for environmental sustainability, economic savings, and quality of

life. Efficient and rational use of electricity and household appliances not only reduces energy expenses thereby alleviating financial pressures on families amid rising electricity prices but also limits harmful emissions such as carbon dioxide, contributing to the mitigation of global warming and environmental preservation. Additionally, effective energy management enhances appliance efficiency and longevity, reducing maintenance and replacement costs. It further decreases stress on power grids during peak demand periods, lowering the risk of outages. Beyond these practical benefits, fostering a culture of energy conservation within households promotes environmental awareness and social responsibility, which are essential for achieving sustainable development and preparing society to meet future energy challenges.

III.1.2 Role of energy management

The role of energy management systems in residential settings is to systematically monitor and regulate energy consumption to optimize the balance between daily demand and resource conservation, while minimizing losses and enhancing overall efficiency. These systems analyze consumption patterns with an emphasis on peak load periods, promote the integration of high-efficiency appliances, and implement operational strategies such as deactivating unnecessary lighting and disconnecting idle devices. Through these measures, energy management reduces operational costs, fosters environmental awareness among occupants, and supports national objectives for energy conservation and sustainable development.

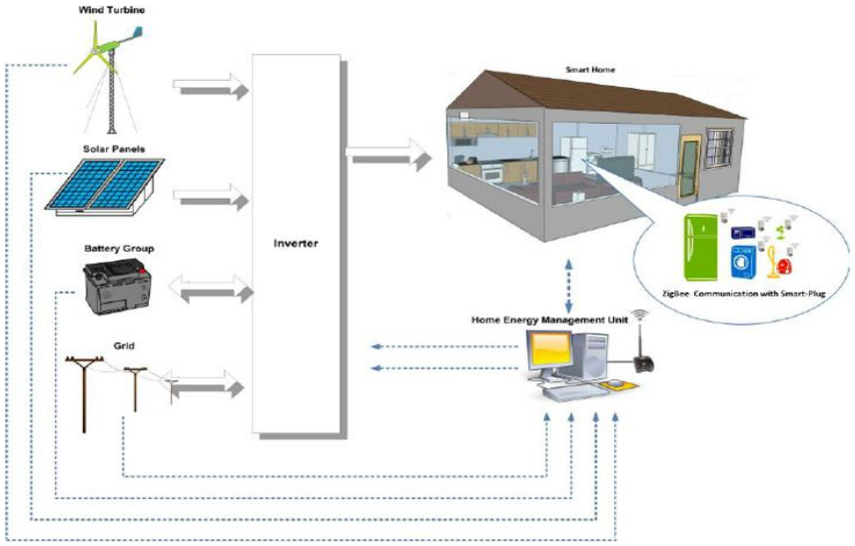


Fig III.1 Smart Home Energy Management System

III.2 Energy consumption prediction

The growing demand for energy, alongside the integration of renewable sources and the advancement of smart grid technologies, has heightened the need for accurate energy

consumption prediction models. Such predictions play a crucial role in enhancing energy efficiency, facilitating demand-side management, and enabling proactive decision-making within energy management systems (EMS) across residential and commercial sectors [33]. As concerns over climate change and resource depletion intensify globally, forecasting energy consumption has become vital for promoting sustainable living and effective infrastructure planning.

Energy consumption prediction refers to estimating the amount of energy a system, building, or household will require over a specific time period by leveraging historical data, usage patterns, environmental factors, and other relevant variables. Depending on the application and operational goals, these predictive models can operate across various time horizons, including short-term (e.g., hourly or daily), medium-term (weekly or monthly), and long-term (seasonal or annual) forecasts [34].

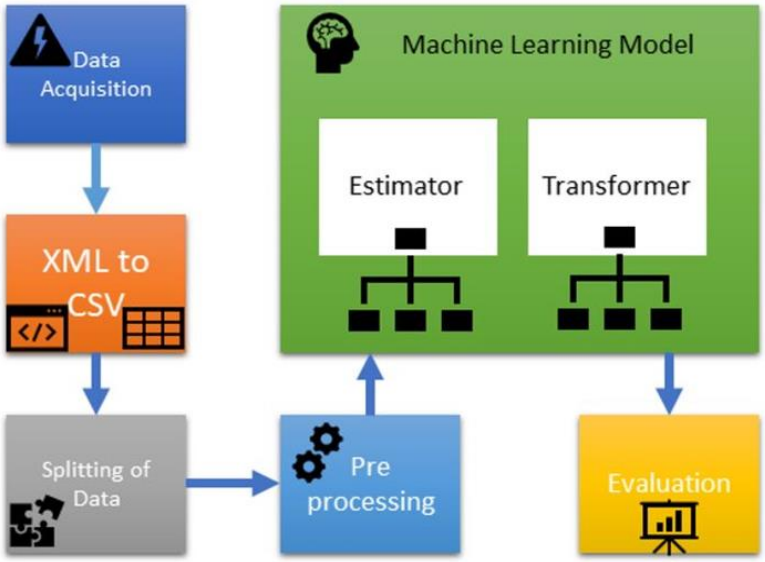


Fig III.2 Conceptual Framework of Energy Consumption Prediction in Smart Home Environments

Advancements in data collection, sensor technology, and computational power have significantly improved the accuracy of energy consumption modeling and forecasting. Today, Internet of Things (IoT) devices and smart meters provide detailed, real-time consumption data that can be utilized to develop mathematical forecasting frameworks or train machine learning models [35]. Consequently, traditional statistical techniques are increasingly being replaced by more adaptable, data-driven approaches such as deep learning, ensemble methods, and hybrid models [36].

Moreover, by aligning energy supply with predicted demand, consumption forecasting supports the integration of renewable energy sources. The intermittent nature of renewables such

as wind and solar can challenge grid stability, but accurate energy prediction mitigates these issues by improving reserve planning and optimizing energy scheduling [37].

III.2.1 Necessity

The increasing complexity of modern energy systems, coupled with the urgent global demand for sustainable development, drives the critical need for accurate energy consumption prediction. As energy demand rises alongside urbanization and population growth, forecasting consumption trends becomes essential for maintaining system efficiency, reliability, and environmental responsibility [38]. Precise consumption forecasts enable utility companies and consumers, particularly in residential and smart home settings, to optimize energy usage and reduce operational costs.

A key factor promoting energy prediction is the transition from conventional power grids to smart grids. These advanced systems use real-time data and predictive analytics to dynamically balance supply and demand. Without accurate consumption forecasts, demand response strategies lose effectiveness, potentially leading to unstable and inefficient grid operations [39]. Energy providers rely on these predictions to plan generation schedules, reduce reserve margins, and minimize energy waste.

The unpredictable and intermittent characteristics of renewable energy sources further emphasize the need for reliable forecasting. Power output from solar and wind sources varies with weather conditions and time of day. By integrating forecasts of renewable generation with consumption predictions, grid operators and prosumers can make better-informed decisions regarding energy scheduling and storage [40]. Accurate demand-side forecasting also supports microgrids and peer-to-peer energy trading models, where decentralized participants depend on real-time matching of production and consumption.

Moreover, decarbonization, efficiency improvements, and clean energy integration are increasingly prioritized in national and international energy policies. Energy consumption prediction underpins essential functions such as modeling carbon emissions, evaluating policy impacts, and supporting environmental reporting frameworks [41]. Governments and energy planners use these forecasts to set realistic climate targets, guide infrastructure investments, and design effective tariff structures.

III.2.2 Benefits

Predicting energy consumption offers numerous benefits, including enhanced environmental sustainability, economic optimization, and improved technical efficiency. As smart homes and

intelligent energy systems become more widespread, and as global energy infrastructures strive to meet climate goals while managing increasing demand, these advantages gain even greater significance [42].

One of the most evident benefits is improved energy efficiency. Predictive models help anticipate usage patterns and enable proactive scheduling of appliances and systems, thereby reducing unnecessary energy consumption. For example, heating and cooling systems can adjust temperatures in advance based on occupancy forecasts, time-of-day patterns, and weather conditions, minimizing energy waste without compromising comfort [43]. Over time, such optimizations can substantially lower the total energy consumption of both residential and commercial buildings.

Another important advantage is the reduction of energy and operational costs. Forecasting enables users to shift consumption to off-peak periods when electricity prices are lower, taking advantage of dynamic pricing schemes. This is particularly relevant in deregulated markets employing time-of-use tariffs. At the utility level, prediction aids in peak shaving and load leveling, reducing reliance on expensive standby generation and excess infrastructure capacity [44].

Prediction models also enhance power system reliability. By anticipating consumption fluctuations and demand spikes, energy providers can better prepare to avoid grid overloads and blackouts [45]. Additionally, forecasting supports more effective local energy management in smart homes, especially when integrated with renewable energy sources or battery storage, fostering greater energy independence and system resilience.

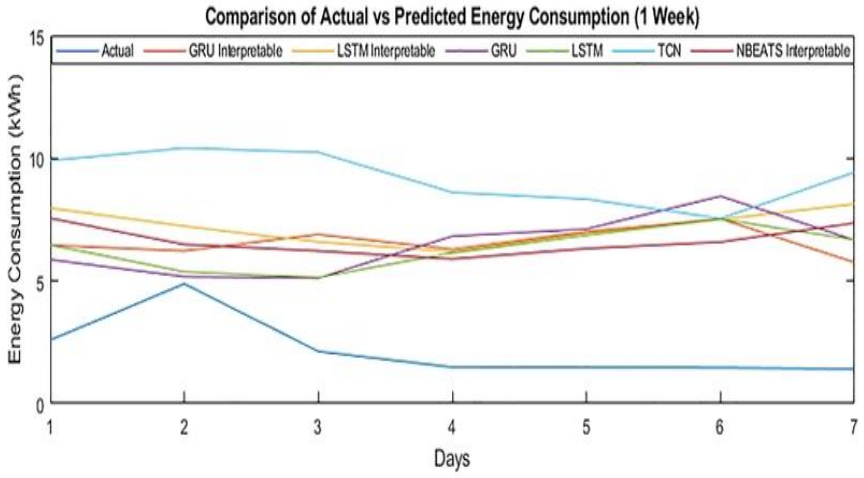


Fig III.3 Comparison of actual vs predicted energy consumption (1 week) [46]

Additionally, predictive analytics provide consumers with personalized energy insights by transforming raw data into actionable information. By analyzing consumption patterns, these systems offer tailored recommendations to reduce usage, set consumption targets, or identify inefficiencies [47]. This not only promotes energy-conscious behavior but also empowers users to make informed decisions. Various prediction models, such as GRU, LSTM, TCN, and NBEATS, demonstrate different levels of accuracy in capturing actual consumption trends, as shown in Fig III.3. While traditional models like LSTM provide a solid predictive foundation, interpretable models enhance transparency and build consumer trust key factors for successful adoption in smart home environments.

III.3 Techniques Used

Recent advancements in data-driven methodologies and computational power have significantly transformed the field of energy consumption prediction. A wide array of techniques ranging from traditional statistical models to advanced machine learning (ML) and deep learning (DL) approaches are now employed to forecast energy usage across different time horizons and levels of detail [48]. These methods are especially vital in connected infrastructures and smart homes, where they support adaptive and responsive energy management systems.

Historically, energy prediction has relied on time series forecasting models such as seasonal decomposition, autoregressive integrated moving average (ARIMA), and exponential smoothing. These traditional approaches perform well when consumption patterns exhibit strong periodicity and limited nonlinearity [49]. However, in more complex, dynamic settings where factors like occupancy, weather variability, and user behavior introduce non-stationary elements the effectiveness of these models declines.

To address the limitations of traditional methods, machine learning (ML) and artificial intelligence (AI) techniques have been widely adopted in energy consumption prediction. These approaches offer greater flexibility and adaptability, especially models such as regression trees, support vector machines (SVM), and ensemble methods, which are effective at uncovering patterns in large, high-dimensional datasets. In recent years, deep learning models have demonstrated superior performance in capturing complex temporal dependencies in energy usage data. Notable examples include recurrent neural networks (RNN), long short-term memory networks (LSTM), gated recurrent units (GRU), temporal convolutional networks (TCN), and Transformer-based architectures [50].

Hybrid models are also gaining traction, combining the strengths of different approaches to deliver more accurate and robust predictions. For instance, deep learning components can capture short-term variability, while statistical methods may be used to model long-term trends [51]. Some hybrid frameworks further enhance interpretability by integrating domain knowledge or incorporating physical system models, which is particularly valuable in regulatory or consumer-facing applications.

The choice of predictive technique often depends on several contextual factors, including data availability, computational constraints, application setting, and forecasting objectives. For example, real-time applications in smart homes may require lightweight, low-latency models, while long-term utility planning might prioritize scalability and interpretability. As energy systems become increasingly decentralized and data-rich, the demand for scalable, transparent, and adaptive prediction solutions will continue to grow.

The following subsections provide a detailed examination of the two main categories of prediction techniques: traditional time series forecasting and advanced machine learning/artificial intelligence methods.

III.3.1 Time Series Forecasting

Time series forecasting has long served as a foundational approach for estimating energy consumption, particularly in structured environments where historical usage exhibits clear temporal regularities. These models predict future values based on sequences of past observations, typically under the assumption that key patterns such as seasonality and autocorrelation persist over time [52].

Among the most widely used classical methods is the Autoregressive Integrated Moving Average (ARIMA) model, renowned for its ability to capture autoregressive (AR) and moving average (MA) components, along with differencing techniques to ensure stationarity [53]. ARIMA is especially effective in short- to medium-term load forecasting scenarios, where linear temporal dependencies are prominent in the historical data. Its seasonal extension, the Seasonal ARIMA (SARIMA) model, further enhances performance in contexts with recurring fluctuations, such as the daily or weekly cycles typical of residential electricity usage.

Another commonly employed technique is Exponential Smoothing (ETS), which assigns exponentially decreasing weights to older observations, thus emphasizing more recent data. This approach is particularly useful in dynamic settings where recent consumption trends are more relevant than distant historical patterns. Hybrid variants like Holt-Winters exponential

smoothing, which incorporates components for level, trend, and seasonality, strike a practical balance between simplicity and accuracy, making them popular in real-world building energy management systems [54].

Despite their computational efficiency and ease of interpretation, classical time series models often struggle to capture nonlinear dynamics, sudden behavioral shifts, and the influence of exogenous variables such as weather conditions, occupancy patterns, and appliance usage. To address these limitations, multivariate time series models such as Vector Autoregression (VAR) and State-Space Models have been proposed, allowing the incorporation of multiple input features [55]. However, these models still face scalability challenges and often rely on assumptions of linearity, making them less suitable for complex, high-dimensional smart home data.

Nonetheless, classical time series models remain valuable, particularly in resource constrained environments where low computational overhead, transparency, and minimal data preprocessing are priorities. Models like ARIMA continue to serve as standard benchmarks in forecasting tasks, and many modern machine learning approaches are evaluated against their performance.

Emerging research trends also highlight efforts to hybridize classical time series methods with machine learning techniques, aiming to leverage the strengths of both linear trend modeling and nonlinear pattern recognition. For instance, hybrid models such as ARIMA-LSTM and SARIMA-ANN have been explored to enhance predictive accuracy while preserving interpretability and domain relevance [56].

III.3.2 Artificial Intelligence and Machine Learning

Artificial Intelligence (AI) and Machine Learning (ML) have revolutionized energy consumption prediction by introducing flexible, data-driven methods capable of learning from high-dimensional datasets, adapting to complex environments, and modeling nonlinear relationships. Unlike traditional statistical approaches, which often depend on assumptions of linearity and stationarity, ML models excel in real-world smart home settings characterized by behavioral variability, irregular usage patterns, and external influences such as weather conditions and dynamic pricing schemes [57].

AI-driven approaches are particularly effective in capturing both short-term fluctuations and long-term consumption trends by leveraging historical data, contextual variables, and real-time sensor inputs. These models employ both supervised and unsupervised learning strategies such as clustering for usage pattern discovery and regression-based methods including decision trees,

support vector machines (SVM), and ensemble techniques like Random Forest and Gradient Boosting to provide robust predictive capabilities [58].

More recently, deep learning architectures have emerged as dominant tools for time series-based forecasting due to their ability to model complex temporal dependencies. Architectures such as Long Short-Term Memory (LSTM), Transformer models, and Convolutional Neural Networks (CNN) have demonstrated superior performance in capturing sequential dynamics at high temporal resolution [59].

In addition to enhancing forecasting accuracy, AI and ML models also offer advantages in terms of scalability, real time adaptability, and seamless integration with Internet of Things (IoT) infrastructures and Home Energy Management Systems (HEMS). The subsequent sections examine the two principal subcategories within this domain classical machine learning and deep learning each with its own strengths and limitations in the context of energy consumption prediction.

1. Classical Machine Learning Techniques

Due to their interpretability, robust performance on structured data, and relatively low computational requirements, classical machine learning (ML) techniques have been extensively adopted for energy consumption prediction. These models typically employ supervised learning approaches to forecast future energy usage based on various input features, including temporal indicators, meteorological data, occupancy status, and appliance level consumption.

Support Vector Machines (SVMs), particularly in their regression form known as Support Vector Regression (SVR), have been widely used in this context. SVR demonstrates strong generalization capabilities, making it especially effective for forecasting tasks involving noisy data and complex nonlinear relationships [60].

Decision tree-based ensemble methods, such as Random Forest (RF) and Gradient Boosted Trees (GBT), are also prominent in energy prediction. These techniques build multiple weak learners and aggregate their outputs to enhance predictive accuracy. While RF is well-suited for handling multicollinearity and missing data, gradient boosting methods like XGBoost and LightGBM have consistently outperformed more complex models in benchmarking studies due to their high accuracy, scalability, and efficient training time [61].

Other classical techniques include Multilayer Perceptrons (MLPs) and k-Nearest Neighbors (k-NN). Although MLPs are a type of shallow neural network, they are often grouped with classical ML methods because of their relatively simple architecture. These methods, however,

typically lack the ability to model temporal dependencies unless augmented with time-series features [62].

A significant advantage of classical ML models lies in their capacity to incorporate domain specific knowledge through feature engineering. Careful feature selection and preprocessing such as encoding time of day, holidays, and weather indicators are critical to improving model accuracy. This flexibility allows for the development of customized, context specific prediction models, making classical ML a practical choice for both residential and commercial energy forecasting applications.

2. Deep Learning Architectures

Deep learning (DL) has emerged as a transformative approach in energy consumption prediction due to its ability to model complex temporal dependencies, nonlinear relationships, and multivariate input sources without extensive manual feature engineering. This capability is particularly valuable in smart home environments, where energy usage data is often sequential, noisy, and influenced by diverse contextual factors such as occupancy patterns, weather variations, and behavioral dynamics [50].

Among the most widely adopted DL models are Recurrent Neural Networks (RNNs) and their advanced variants Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). These architectures are specifically designed to process sequential time-series data by maintaining a memory of past inputs. LSTM networks, in particular, have demonstrated strong performance in short- to medium-term forecasting, owing to their ability to retain long-term dependencies and mitigate issues such as vanishing gradients [51].

Another notable deep learning architecture gaining traction in the field is the Temporal Convolutional Network (TCN). TCNs utilize dilated causal convolutions to model sequential dependencies while enabling parallel computation and reduced training times compared to RNN-based models. Importantly, TCNs preserve the temporal order of input sequences and are capable of capturing both short-term fluctuations and long-range temporal dependencies. These attributes make TCNs a compelling alternative to recurrent models for energy forecasting applications [63].

Recent advancements in energy consumption forecasting have been significantly influenced by the adoption of Transformer-based architectures, originally developed for natural language processing tasks. These models leverage self-attention mechanisms to capture dependencies across all elements in a sequence, irrespective of their relative positions. This allows

Transformers to learn global temporal patterns more effectively than traditional recurrent models. As a result, they have shown superior performance in multi-horizon forecasting and multivariate time series prediction, often outperforming LSTM and GRU architectures in terms of both accuracy and scalability [64].

Furthermore, to simultaneously utilize spatial and temporal information, hybrid deep learning models that blend several architectures, like CNN-LSTM or Transformer-GRU, have been proposed. These hybrids have been incorporated into smart grid and residential HEMS platforms and are especially useful for integrating auxiliary data (such as weather, occupancy, and pricing signals) [65].

DL models necessitate substantial computational resources, sizable labeled datasets, and meticulous hyperparameter tuning despite their high predictive power. Furthermore, energy systems may face difficulties with interpretability and regulatory compliance due to their "black-box" nature. Recent studies have addressed this by concentrating on explainable AI (XAI) strategies that use model-agnostic interpretation techniques, attention mechanisms, or SHAP values to shed light on model decisions.

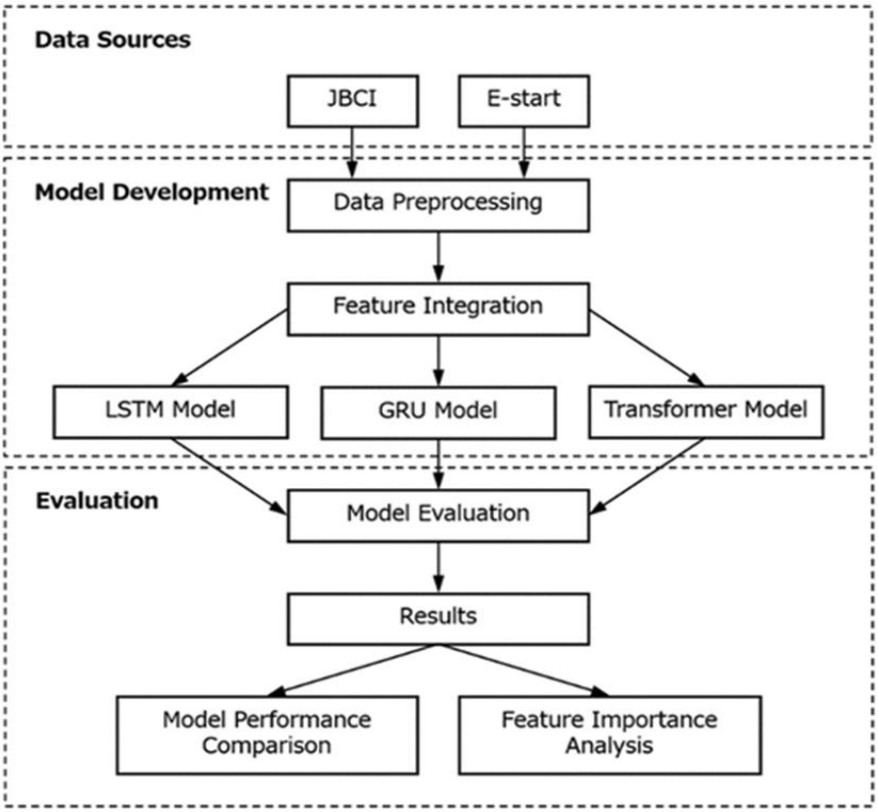


Fig III. 4 Deep Learning Workflow for Energy Consumption Prediction: Data Sources to Model Evaluation [66]

In conclusion, deep learning architectures offer strong instruments for precise and expandable energy consumption prediction. They are positioned as a key component of next generation energy management systems due to their capacity to independently extract features and comprehend complex temporal patterns.

III.4 Literature review

In this section, we will try to highlight the latest work carried out on AI-based electricity consumption prediction using a time series forecasting model. Therefore, due to the many models used, we will try to divide this overview into two parts. The first concerns ML methods, while the second part presents methods based on deep learning (DL). Moreover, we have relied on these referenced studies to compare and validate the results we obtained in our own research.

In [67] developed four statistical models: multiple linear regression (MLP), support vector machine (SVM) with a radial kernel, random forest (RF), and gradient boosting machines (GBMs). The study began with data analysis involving energy consumption aggregation, appliance-level consumption measurement, correlation analysis via pairs plots, hourly heatmaps, and variable importance assessment. All models were trained using tenfold cross-validation to select the best performer. Testing on the passive house planning dataset showed that the GBM model outperformed others, achieving the lowest errors in RMSE (0.665), MAPE (0.389), and MAE (0.352).

In another machine learning study, Wu et al. in [68] proposed a boosting-based framework using multiple kernel learning regression for short-term load forecasting. Their approach first applies boosting to train multiple kernel regressors and then extends this framework to transfer learning, considering both homogeneous and heterogeneous transfer scenarios. Experiments on residential datasets (Open EI and UCI) demonstrated that leveraging knowledge from other households significantly reduces forecasting errors.

The time series forecasting model in [69] combines Kernel Principal Component Analysis (KPCA) and Support Vector Machine (SVM) to predict electricity consumption. KPCA is used for feature selection and dimensionality reduction, while SVM performs the prediction. The process includes data preprocessing steps such as decomposition and normalization, followed by KPCA-based reduction. The model forecasts electricity consumption 30 minutes ahead, achieving strong performance with MAE of 11.48 and RMSE of 13.86.

Moon et al. [70] proposed a hybrid electric load forecasting model for an educational building complex by combining Random Forest (RF) and Multilayer Perceptron (MLP). They used

six years of electrical load data from a college campus and performed preprocessing steps including temperature adjustment and consumption estimation for one week and one year ahead. Various configurations of RF and MLP were evaluated using cross-validation. Experimental results on daily power consumption data from a Korean university showed strong forecasting performance, achieving MAPE of 1.143 and RMSE of 2.946.

With the rise of deep learning (DL) and its success across various fields, researchers have applied these techniques to electricity consumption forecasting. Rajabi and Estebsari [71] proposed a deep learning model based on convolutional neural networks (CNN) to predict individual loads. Their method encodes time series data into recurrence plot images, which are then processed by a CNN to generate predictions. For comparison, support vector machine (SVM) and artificial neural network (ANN) models were also tested on the same dataset. Experiments on the UCI dataset showed that the CNN-based model outperformed SVM and ANN, achieving MAE of 0.59 and RMSE of 0.79.

Tang et al. [72] proposed a multilayer bidirectional recurrent neural network (BRNN) combining long short-term memory (LSTM) and gated recurrent unit (GRU) to improve short-term power load forecasting. Data preprocessing included normalization to scale features uniformly. The BRNN model was tested on the EUNITE and Weather datasets. Results demonstrated improved forecast accuracy and reduced fluctuation, achieving MAPE of 1.90 and RMSE of 16.044.

Sajjad et al. [73] proposed a novel hybrid model combining convolutional neural networks (CNN) and gated recurrent units (GRU) for short-term residential load forecasting. The input data were first normalized using min-max scaling to address nonlinearity. The model extracts spatial features via CNN and captures temporal dependencies through a multilayer GRU. Evaluated on the Appliances Energy Prediction (AEP) and IHEPC datasets, the model achieved strong performance with MAE of 0.31 and RMSE of 0.24 on the AEP dataset.

Authors in [74] proposed a framework combining deep learning and supervised machine learning for electricity consumption prediction. Feature selection was performed using Random Forest (RF) and Extreme Gradient Boosting (XGBoost) to assess feature importance. Two enhanced models were developed: SVM optimized with Gray Wolf Optimization (SVM-GWO) and CNN-GRU optimized with Earth Worm Optimization (CNN-GRU-EWO), where the optimization algorithms tuned the hyperparameters of SVM and CNN-GRU, respectively.

Experimental results showed that the proposed approach improved efficiency by 7% and 3% compared to state-of-the-art methods.

Elattar et al. [75] proposed a method combining locally weighted support vector regression (LWSVR) with a modified grasshopper optimization algorithm (MGOA). The MGOA was enhanced using chaotic initialization and a sigmoid decrement criterion to better tune LWSVR parameters. The method was evaluated on six datasets, including New York, Victoria, EKPC area, CEFCom, ISO New England, and microgrid datasets. Results showed high prediction accuracy on the ISO New England dataset, achieving a MAPE of 1.67 and MAE of 0.89.

Khan et al. [76] developed a hybrid model combining convolutional neural networks (CNN) with an LSTM autoencoder (LSTM-AE) to predict electricity consumption in residential and commercial buildings. The preprocessing step applied various scaling and transformation techniques to handle nulls, redundancies, and outliers. During training, CNN extracted features from the processed data, which were then encoded by an LSTM encoder over four-time steps. The encoded sequences were decoded by another LSTM, followed by a dense layer producing the final prediction. Experimental results demonstrated strong performance, with RMSE of 0.47, MAPE of 0.76, and MAE of 0.31.

III.5 Challenges of home energy management

Home energy management systems (HEMS) face numerous obstacles that prevent their widespread implementation and long-term efficacy, even in the face of the increasing use of smart technologies and sophisticated forecasting models. The nature of data on energy consumption, system interoperability issues, unpredictable user behavior, and the difficulty of integrating various energy sources are some of the causes of these difficulties.

The quality and heterogeneity of input data are among the most basic challenges, as they can have a big influence on how well energy prediction models work. Large amounts of data are produced by smart meters and sensors, but because of hardware constraints or communication issues, these data are frequently noisy, insufficient, or inconsistently sampled. Additionally, contextual metadata, such as occupancy patterns or appliance level disaggregation, that are essential for precise demand forecasting are frequently absent from datasets [77].

The absence of standardized frameworks for system integration and data sharing is another serious problem. Numerous platforms and devices, frequently from various vendors with proprietary protocols, are found in residential settings. The smooth communication between energy storage devices, smart appliances, and renewable energy sources like home batteries or rooftop

solar is hampered by this fragmentation [78]. Consequently, it becomes challenging to coordinate across the energy ecosystem in an optimal manner.

Security and privacy issues also pose a significant obstacle. Data on energy use can provide personal insights into lifestyle trends and household behavior. It is still difficult to guarantee that HEMS protect data privacy and adhere to privacy laws (like the GDPR) without sacrificing functionality [79].

From the standpoint of the user, HEMS's effectiveness is greatly influenced by behavior variability and engagement. Strategies for energy management must adjust to changing user preferences, lifestyle modifications, and unanticipated activities that affect consumption trends. Furthermore, users might be hesitant to embrace or regularly use these systems in the absence of adequate incentives or user-friendly interfaces [80].

Finally, the intermittent nature of distributed and renewable energy resources makes their integration more difficult. Real-time supply and demand predictions are essential for systems, necessitating sophisticated optimization and control techniques in the face of uncertainty. This necessitates scalable, resilient algorithms that can function in decentralized, dynamic environments [81].

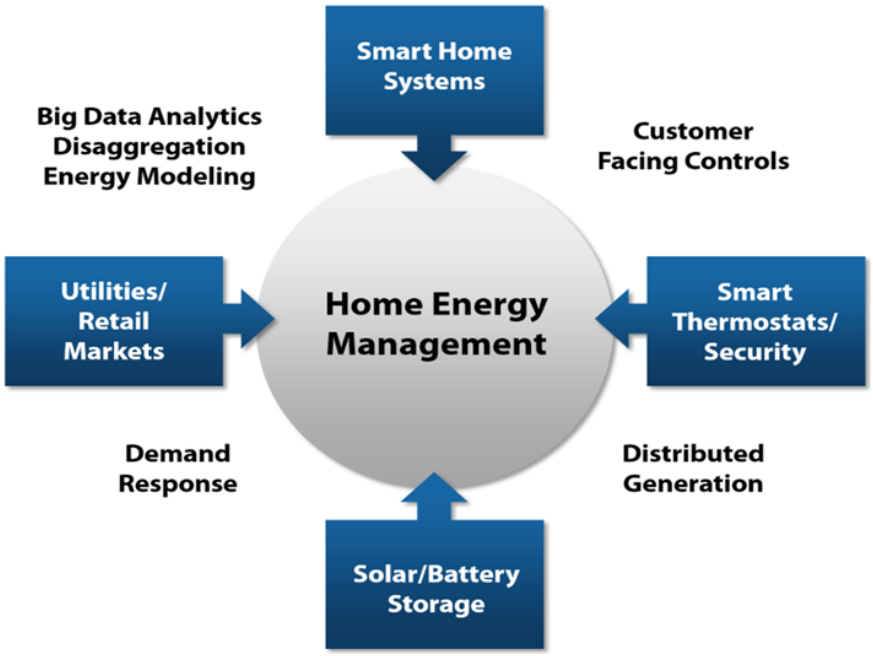


Fig III.5 Challenges in Home Energy Management

These difficulties highlight how crucial it is to create secure, flexible, and intelligent energy management systems that can operate in a wide range of unpredictable and dispersed residential

settings. Two particularly important technical obstacles are examined in the ensuing subsections: the need for smarter systems and data-related constraints.

III.5.1 Data accuracy and diverse energy sources

Since prediction models, control strategies, and optimization algorithms all rely on clean, timely, and granular data, data accuracy is essential to dependable home energy management systems (HEMS). However, the inherent variability in energy consumption and the technical limitations of the existing data acquisition infrastructure make it difficult to achieve high data fidelity in residential settings.

Despite being widely used, smart meters frequently have low sampling frequency, communication delays, or missing values, which restrict the energy management system's responsiveness and resolution. Poor network connectivity, data buffering, or storage limitations in edge devices could be the cause of these problems. For example, smart meters usually record interval data every 15 to 60 minutes, which is not enough to detect abrupt changes in consumption patterns or events at the appliance level [82].

Furthermore, the modeling process is made more difficult by data heterogeneity. Different households may use a combination of battery storage, solar photovoltaic (PV) systems, conventional grid power, and even electric cars. Complex load profiles are produced by these various and frequently sporadic energy sources, necessitating highly adaptive and context-aware prediction models. For instance, weather, shading, and panel orientation all affect solar generation, adding even more unpredictability to energy balancing [83].

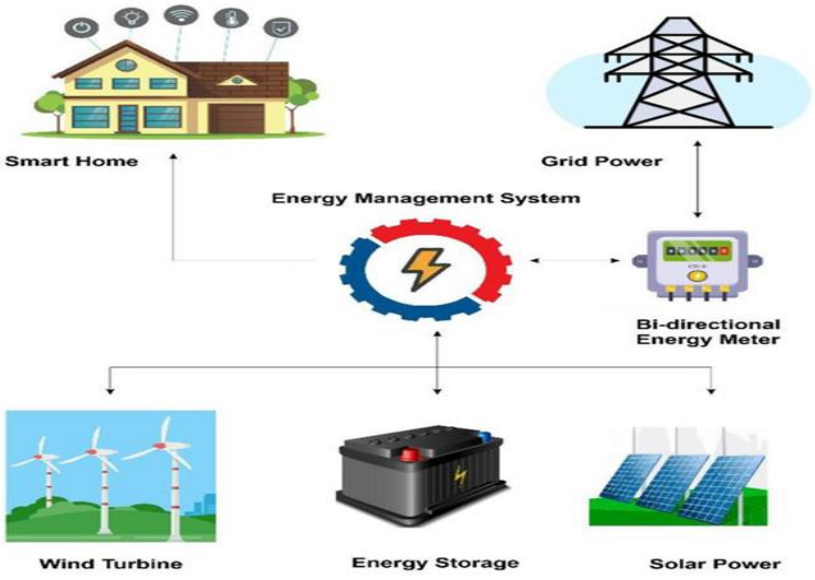


Fig III.6 Smart Home Energy management system energy sources

The problem of inconsistent data is made worse by the incorporation of renewable energy sources into HEMS. Renewable energy systems don't offer a steady or predictable power supply like conventional sources do. As a result, precise generation and consumption forecasting becomes crucial. Due to insufficient temporal granularity or incomplete datasets, many systems are unable to effectively handle this dual prediction requirement [79].

Furthermore, fragmented data pipelines result from a lack of standardized data formats and interfaces across various platforms and devices. Implementing unified monitoring and control strategies is challenging due to interoperability restrictions imposed by proprietary protocols and closed-source APIs. Machine learning and artificial intelligence models are unable to generalize across households or geographical areas as a result of this fragmentation [78].

Lastly, one of the biggest obstacles to training high-performance models is still labeling and annotating energy data, particularly at the device or appliance level. Detailed disaggregated datasets are uncommon and usually produced in controlled settings, which restricts their use in practical implementations.

III.5.2 Smart systems needed

Smarter, adaptive systems that transcend static rule-based controls are desperately needed to handle the complexity and dynamism of residential energy environments. For modern Home Energy Management Systems (HEMS) to balance energy consumption, user comfort, and sustainability goals, they must integrate real-time analytics, self-learning capabilities, and autonomous decision-making frameworks [84]. Context-aware automation, AI algorithms, and strong communication frameworks that can handle diverse devices and decentralized energy sources must all be integrated to achieve this.

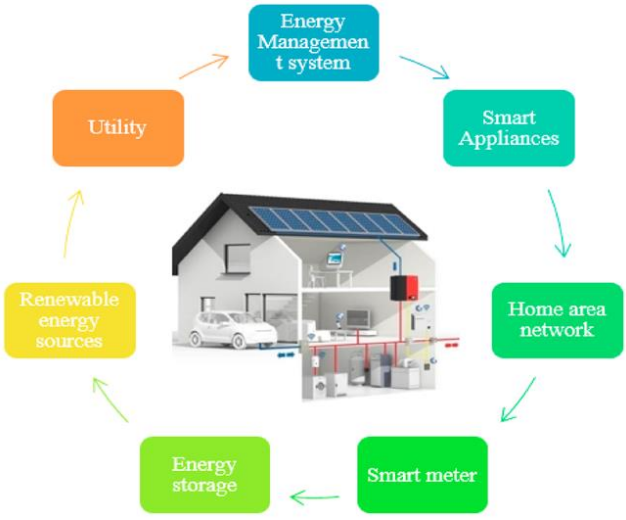


Fig III.7 Functional Ecosystem of Smart Systems in Home Energy Management

Furthermore, the quick adoption of electric vehicles (EVs) and renewable energy systems necessitates that HEMS become pro-active rather than reactive, capable of anticipating shifts in generation or consumption and making the necessary adjustments.

Smart systems need to be user-centric, interoperable, scalable across various home configurations, and technically advanced in order to meet these requirements [85]. To effectively realize such systems, several core features and capabilities are essential, including the following:

● **Real-Time Data Processing:** Energy generation and consumption patterns must be continuously monitored by smart systems. By processing data closer to the source, edge or fog computing lowers latency and makes it possible to instantly modify load balancing or device operation in response to changes in the environment or user behavior.

● **Artificial Intelligence Integration:** By combining machine learning (ML) and deep learning (DL) approaches, HEMS can predict energy consumption, identify irregularities like equipment malfunctions or sudden spikes in load, and automatically optimize appliance schedules. In order to learn temporal patterns and respond adaptively, models such as LSTM, CNNs, and reinforcement learning are especially helpful.

● **Demand Response Capabilities:** In response to pricing signals or grid demands, smart systems should take part in utility-driven demand response (DR) programs, which temporarily move or reduce electricity use during peak hours. In addition to lowering consumer electricity bills, this stabilizes the grid.

● **Interoperability:** It is crucial to ensure smooth integration across HVAC, lighting, renewable energy systems, and storage units due to the large number of device manufacturers and communication standards available on the market. Widely used protocols like ZigBee, Z-Wave, MQTT, and BACnet must be supported by smart HEMS in order to facilitate cross-platform communication and flexible expansion.

● **User Interaction & Personalization:** User engagement is a key component of long-term success. User-friendly dashboards and smartphone apps that provide energy usage data, alerts, and tailored recommendations based on comfort and lifestyle choices should be a feature of smart systems. Changes in occupancy or routines should also be accommodated by adaptive control settings.

● **Resilience and Scalability:** Future-proof HEMS designs ought to be able to grow from a single home to whole residential complexes. They must also use secure, redundant

communication channels and self-healing mechanisms to continue operating even in the event of data loss, cyberattacks, or hardware malfunctions.

III.6 Impact of home energy management

The way residential energy is used, managed, and optimized has completely changed with the introduction of Home Energy Management Systems (HEMS). Intelligent, automated management at the home level becomes critical as energy demands rise as a result of the widespread use of smart devices, electric vehicles, and decentralized energy generation. To optimize energy use in real-time, HEMS integrates a number of technologies, including energy storage systems, smart meters, Internet of Things devices, and predictive analytics [86]. There are significant advantages in terms of the economy, environment, and infrastructure as a result.

Economically speaking, HEMS give users comprehensive information about their patterns of energy use and enable them to modify their usage habits in response to demand projections, appliance schedules, and dynamic pricing signals. When combined with demand-side management and involvement in utility demand response programs, this degree of optimization results in lower electricity costs [87]. Additionally, HEMS serve as centralized control units that coordinate photovoltaics, batteries, and electric vehicles (EVs) for optimal efficiency and financial return as residential energy systems grow more complex.

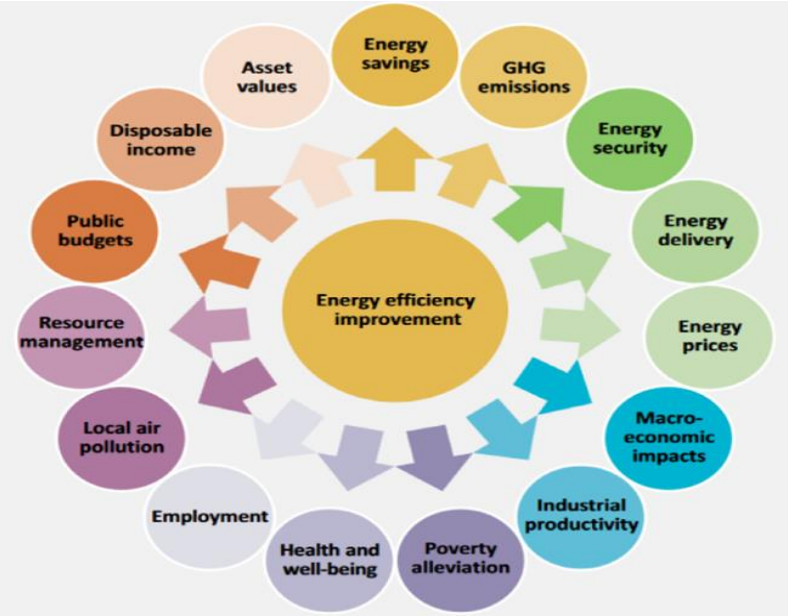


Fig III.8 Concepts of multiple HEMS benefits at the household, community, and utility levels

By increasing energy efficiency and promoting the use of renewable energy sources, HEMS considerably lowers greenhouse gas emissions. These systems reduce dependence on fossil fuel-powered grids during peak hours by strategically scheduling loads and balancing local

generation with storage. Moreover, HEMS improves grid stability, permits improved peak load management, and lessens the need for expensive infrastructure upgrades when deployed on a large scale.

HEMS support the creation of energy-resilient communities and smart cities on a societal level. In addition to promoting sustainable consumption among users, the data they produce aids in the development of energy policies and urban planning [88]. Households can actively participate in a cleaner and more intelligent energy ecosystem by using user-centric interfaces and real-time feedback.

To delve deeper into these advantages, the following subsections explore two of the most measurable impacts: long-term cost reduction and carbon footprint mitigation.

III.6.1 Long-term costs reduction

The possibility of significant long-term cost savings is one of the most alluring benefits of installing Home Energy Management Systems (HEMS). With the help of these systems, homeowners can track, evaluate, and optimize their energy use in real time, which eventually results in financial gains. In both regulated and deregulated energy markets, HEMS enables users to minimize electricity bills by reducing overall demand, shifting energy usage to off-peak hours, and preventing needless energy waste through both passive and active control strategies.

One of HEMS's primary cost-cutting strategies is load shifting, in which the system plans energy-intensive tasks (like running water heaters, dishwashers, or EV chargers) for when electricity prices are at their lowest. This is especially advantageous in smart grids that use real-time pricing models or in time-of-use (TOU) pricing environments. HEMS can also avoid expensive peak-period tariffs by using predictive algorithms to predict periods of high demand and modify device operation accordingly [89].

Additionally, HEMS can control the best possible use of locally produced power when homes incorporate renewable energy systems like energy storage devices and rooftop solar panels. This greatly lessens dependency on utility-supplied electricity by storing extra solar energy for later use or giving self-consumption priority during periods of peak grid pricing. HEMS can even identify the most economical times to export energy back to the grid in are-as with feed-in tariffs or net metering regulations, giving homeowners access to an extra source of income [90].

The automated identification of inefficiencies and malfunctioning equipment also contributes to long-term savings. In order to detect unusual energy use or appliance malfunctions, smart HEMS platforms employ pattern recognition and historical consumption data. To avoid

prolonged inefficiency or equipment failure, the system can, for instance, automatically initiate maintenance procedures or notify the homeowner if a refrigerator starts using more power than usual.

Even modest daily optimizations made possible by HEMS add up to significant financial results over time. Depending on the degree of behavioral adaptation and renewable integration, households utilizing AI-based HEMS saw yearly energy bill reductions of 15% to 30% [90]. Households with flexible loads, like heat pumps, HVAC systems, and electric vehicle chargers, or those with high energy consumption will see the biggest savings.

Furthermore, long-term cost savings extend beyond individual users. More generally, the use of HEMS lowers peak loads and overall grid demand, which can postpone costly grid infrastructure upgrades and lower the cost of electricity supply for all customers. HEMS will be essential to enabling economically viable energy ecosystems as energy markets continue to shift toward decentralized, consumer-driven models.

III.6.2 Reducing the carbon footprint

Apart from their financial benefits, one of Home Energy Management Systems' (HEMS) most significant contributions is their ability to drastically lower the carbon footprint of residential energy use. HEMS provide a comprehensive strategy for reducing greenhouse gas (GHG) emissions at the home and community levels by integrating renewable energy sources, optimizing energy usage patterns, and facilitating demand-side flexibility.

HEMS primarily reduces emissions by improving energy efficiency. These systems reduce wasteful energy use and encourage more environmentally friendly user behavior through automation, intelligent scheduling, and real-time monitoring. HEMS, for example, can detect and respond to energy waste, lower standby power losses, and modify operations according to occupancy or daylight availability, all of which contribute to more thoughtful and effective electricity use [65]. When combined, these enhancements significantly reduce the amount of energy needed by fossil fuel-powered plants, which in turn reduces emissions in cities and neighborhoods.

Additionally, especially at the residential level, HEMS are essential in enabling the integration of renewable energy technologies. Energy storage systems, electric vehicle (EV) chargers, and rooftop solar photovoltaics (PV) can all be seamlessly coordinated with smart energy management. With HEMS, households can become prosumers, consumers and producers of electricity with less of an impact on the environment, by prioritizing the use of locally generated

clean energy and reducing reliance on centralized, carbon-intensive power sources [48]. Additionally, by reducing strain on transmission infrastructure, this decentralized generation reduces losses related to long-distance electricity transportation.

HEMS-enabled strategic load shifting can also match consumption trends with periods of increased renewable generation on the grid, like midday solar surpluses or wind peaks at night. HEMS enhances the efficiency of using renewable resources and indirectly contributes to the grid's overall decarbonization by dynamically moving flexible loads to these times [86]. Furthermore, lower system-wide emissions intensity can result from decreased house-hold demand during peak fossil-fuel generation windows in areas with carbon pricing or emissions trading systems.

Along with operational plans, HEMS's data collection and analysis offer useful information to utilities, researchers, and policymakers who want to lower carbon emissions. The creation of more successful energy efficiency initiatives and climate policies is influenced by behavioral patterns, appliance-level usage profiles, and longitudinal consumption data. Therefore, HEMS's environmental benefits go beyond individual homes to national and international sustainability objectives [40].

Recent studies quantify the carbon mitigation potential of HEMS in various deployment scenarios. For example, a simulation-based analysis demonstrated that intelligent HEMS could reduce annual household CO_2 emissions by up to 35% when combined with PV and battery storage systems [42]. In smart community pilots across Europe and East Asia, emission reductions of 20-30% were reported through coordinated energy management strategies. These outcomes underscore the critical role of HEMS in advancing climate resilience and achieving net-zero carbon objectives in residential sectors.

In summary, HEMS offer a multi-dimensional pathway to carbon footprint mitigation through enhanced efficiency, renewable energy integration, and demand-side flexibility. As climate policies increasingly prioritize decarbonization across all sectors, the adoption of HEMS will be indispensable to achieving sustainable energy transitions at both the micro and macro scales.

III.7 Conclusion

A key component of the shift to smarter, more sustainable residential energy ecosystems is the development of energy consumption prediction and home energy management. The various ways that prediction models and intelligent control techniques can improve the sustainability,

dependability, and efficiency of home energy systems have been covered in this chap-ter. Predictive analytics is clearly not only a technical improvement but also a facilitator of proactive energy stewardship, as evidenced by the implementation of cutting-edge AI and machine learning techniques and the growing need for precise consumption forecasting.

Demand estimation is greatly enhanced by the combination of intelligent algorithms and time series forecasting, which promotes better planning, cost savings, and a smaller environmental impact. Nevertheless, there are drawbacks to these developments, including the need to preserve data accuracy in the face of diverse energy sources and guarantee interoperability across ever-more intricate smart home systems. To fully utilize energy prediction systems and Home Energy Management Systems (HEMS), these problems must be resolved.

The analysis of HEMS demonstrated their transformative impact on the economy, environment, and infrastructure. By promoting behavioral adaptation, enabling demand-side flexibility, and optimizing the use of renewable resources, HEMS dramatically reduces long-term costs and carbon emissions. Empirical evidence and simulation-based assessments support the efficiency of these systems in reducing energy waste, reducing greenhouse gas emissions, and fostering grid resilience.

In the end, the foundation of energy networks prepared for the future is the convergence of predictive technologies and intelligent home energy systems. Adoption of them supports utility operations, empowers consumers, and advances global decarbonization objectives. The creation and implementation of reliable, user-focused, and data-driven home energy solutions will be essential as energy markets shift toward decentralization and sustainability. To guarantee fair, scalable, and climate-resilient energy systems, future research and policy must keep encouraging innovation in this field.

Chapter IV

Efficient Management of Residential Electricity Consumption

Residential energy consumption has experienced a steady increase in recent decades, driven by population growth, urbanization, and the widespread adoption of electrical appliances and smart home technologies. According to the International Energy Agency (IEA), the residential sector accounts for over 25% of global electricity consumption, with this demand expected to continue rising. This growing energy needs places considerable stress on electrical grids, particularly during peak load periods, highlighting the critical necessity for smarter and more adaptive energy management systems.

Accurate electricity consumption forecasting plays a vital role in ensuring grid stability, optimizing resource allocation, and minimizing energy waste. Within smart grid environments characterized by decentralized energy sources such as solar panels, wind turbines, and home battery storage systems forecasting facilitates predictive control and enhances decision-making processes. It also supports demand response strategies by enabling real-time load shifting and reduction.

This chapter proposes a methodology for forecasting residential electricity consumption over the short and medium term. It begins with the deep learning technique (LSTM), followed by machine learning using the XGBoost algorithm. The focus is placed on XGBoost due to its exceptional predictive accuracy and computational efficiency. Based on the forecasting results, an intelligent energy management framework was developed to dynamically select the optimal energy sources. Collectively, these contributions aim to improve prediction accuracy and enhance the operational efficiency of smart residential energy systems, contributing to more sustainable and effective energy management.

IV.1 Technical Foundations

The development of accurate energy forecasting models and intelligent energy management systems relies heavily on a deep understanding of machine learning and time series analysis techniques. In the context of smart grids, where energy consumption patterns are dynamic and influenced by multiple interrelated factors, such as weather, occupancy, time of day, and seasonal trends, traditional linear models often fail to capture the complexity and non-linearity of the data. This necessitates the use of advanced methods capable of learning from historical data and generalizing to future conditions.

This section provides the theoretical and technical underpinnings for the forecasting and decision-making strategies proposed in this chapter. First, it outlines the core principles of time series forecasting, emphasizing key challenges such as trend extraction, seasonality, and autocorrelation. Then, it introduces two powerful predictive modeling techniques: Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN) particularly well-suited for sequential data; and Extreme Gradient Boosting (XGBoost), a highly efficient and scalable ensemble learning algorithm based on decision trees.

These models form the backbone of the forecasting pipeline described in subsequent sections and serve as key components in enabling intelligent, data-driven energy source management.

IV.1.1 Time Series Forecasting Principles

Time series forecasting is the process of using historical data, typically collected over consistent time intervals (e.g., hourly, daily, monthly), to predict future values. It is widely used in various domains, such as energy consumption, finance, economics, weather forecasting, and more. Below are the key principles and concepts that underpin effective time series forecasting:

1. Time series components

In general, a time series is defined as a set of observations recorded sequentially over time. So, a discrete time series $(X_t)_{1 \leq t \leq n}$ is a series of n numerical data points in a chronological order (chronological series), indexed in time with a constant time interval between the indices, where time is represented in seconds, minutes, days, years, and so on. One of the primary objectives in time series analysis is to forecast future values based on previously observed data. Time series forecasting plays a crucial role in numerous fields, including economics, finance, demography, biology, medicine, meteorology, and environmental science. While no

predictive model can perfectly replicate reality or forecast future values with complete accuracy, appropriate models can provide reliable forecasts along with measures of uncertainty, helping to assess the accuracy of predictions. To achieve effective forecasting, a wide range of models can be employed, including regression models, exponential smoothing techniques, and ARMA-type models [91].

A time series generally comprises three key components: trend, seasonality, and residual (or irregular) variations. The trend component reflects the long-term progression or general direction of the series, whether increasing, decreasing, or stable. Identifying the trend allows for the removal of random fluctuations caused by seasonal or cyclical patterns. The seasonal component captures patterns that repeat at regular intervals, such as daily, monthly, or yearly cycles. Lastly, the residual or irregular component accounts for random, unpredictable variations that are not explained by the trend or seasonality. While not all time series exhibit all three components, a proper decomposition is essential to isolate and evaluate the influence of each component on the overall behavior of the series [92].

2. Time series models

Mathematically, the interaction between the three main components of a time series trend, seasonality, and residual gives rise to two fundamental types of models: the additive model and the multiplicative model [93].

- **Additive model:** In this model, the time series is assumed to be the sum of its individual components. It can be mathematically represented as:

$$X_i = g_i S_i W_i \qquad (IV.1)$$

Where $X_i \setminus i$=1…n is the observed series, g_i is the trend, S_i is the seasonal factor and w_i is the residual fluctuation component. If this phenomenon is observed at regular times, the i-order value is observed in time $i \times t_e$, where t_i is the sampling period or the observation period. By convention, S_i is required to have a zero mean, to unambiguously separate the trend and the seasonal factor:

$$\sum_{k=1}^{P} S_{i+k} = 0, \qquad \forall 1 \leq t \leq n-p \qquad (IV.2)$$

Here, p denotes the period of the seasonal fluctuation. To ensure that the overall mean of the observed series x_i is captured within the trend component g_i, it is also assumed that the residual component w_i has a zero mean. In this additive model, the amplitudes of the seasonal

component S_i and the residual component w_i at any given time are assumed to be independent of the value of the trend component g_i at the same time. This implies that the magnitude of seasonal and irregular fluctuations remains constant regardless of the trend level.

- **Multiplicative model:** In this model, the components of the time series interact in a multiplicative manner. The time series can be mathematically expressed as:

$$X_i = g_i(1 + S_i)(1 + w_i) = g_i + g_i S_i + g_i(1 + S_i)w_i = g_i + \hat{S}_i + \hat{w}_i \qquad (IV.3)$$

Where $X_i \setminus i=1\ldots N$ is the observed series, g_i is the trend, $1+S_i$ is the seasonal factor, and $1+w_i$ is the residual fluctuation component. In the multiplicative model, the seasonal factor $\hat{S}_i$ is proportional to the value of the trend component, and the residual fluctuation $1+\hat{w}_i$ is proportional to the sum of the trend and seasonal components. This reflects the idea that both seasonal and irregular variations scale with the level of the series. Furthermore, it is important to note that these two basic models the additive and multiplicative can be combined to form hybrid or mixed models. Various hybrid models have been proposed in the literature to better capture complex behaviors in time series data. An example of such a hybrid model is presented in Equation (IV.4):

$$X_i = (g_i + S_i)(1 + w_i) \qquad (IV.4)$$

The additive model is most appropriate when the seasonal variation remains relatively constant over time. In contrast, the multiplicative model is better suited for cases where the seasonal variation changes proportionally with the level of the series, typically increasing over time. In situations where neither model alone adequately captures the data characteristics, a hybrid model may be employed [94].

3. Time series decomposition

Time series decomposition into trend, seasonal, and residual components is a fundamental step in time series analysis. It serves as an analytical tool that provides an abstract model useful for estimation, forecasting, and simulation processes. The importance of this step lies in its ability to determine whether the time series exhibits consistent or recurring patterns that can be effectively described and directly modeled.

Moreover, each component can guide essential tasks such as data preprocessing, model selection, and feature engineering. The decomposition process typically involves first applying a convolution filter to the data to estimate the trend component. This trend is then removed

from the original series to obtain a detrended series. Finally, the seasonal component is extracted by averaging the detrended series over each time period [95].

In general, the initial step in time series analysis involves a graphical representation of the data for visual inspection. In some cases, the presence of a trend component and a seasonal cycle may be clearly observed. However, these components are not always easily identifiable, and some time series may lack one or both of them.

In fact, the trend component may become apparent only after applying a suitable transformation to the data, such as a logarithmic transformation. For the purpose of this study, we assume that the time series x_t can be represented in the decomposed form described in Equation (IV.1).

- **Trend Estimation:** Trend estimation involves characterizing the fundamental or long-term evolution of a time series. Various methods exist for estimating the trend component; among the most commonly used are Moving Average (MA) smoothing and differencing techniques. These methods help to isolate and analyze the underlying pattern by reducing noise and short-term fluctuations.

(i) Smoothing by Moving Average (MA): The Moving Average (MA) smoothing method is a non-parametric approach used to estimate the trend component of a time series. One of the key advantages of this method is that it does not require any assumptions about the underlying data distribution and can effectively accommodate non-constant trends, even in the presence of seasonal variation [96]. The trend is estimated by applying a symmetric moving average over a window of length p, which smooths out short-term fluctuations and highlights the underlying pattern in the data.

If the period is odd ($p = 2q+ 1$), then

$$(g_i)_e = \frac{1}{2q+1}\sum_{j=-q}^{q} X_{i+j} = \frac{1}{p}\sum_{j=-q}^{q} X_{i+j} \qquad i = q+1, \dots, n-q \qquad (IV.5)$$

If the period is even ($p = 2q$), then

$$(g_i)_e = \frac{1}{p}\left(0.5X_{i-q} + \sum_{j=-q+1}^{-q} X_{i+j} + 0.5X_{i+q}\right) \quad i = q+1, \dots, n-q \qquad (IV.6)$$

The choice of the window length q in the moving average smoothing method involves a trade-off. A value of q that is too small may fail to adequately separate the trend from the re-

sidual noise, while a value that is too large may over smooth the data and obscure important changes in the trend.

(ii) Differencing: To handle time series that exhibit seasonality with a known period p, the differencing method can be applied at order p. This approach helps to eliminate seasonal effects and stabilize the mean of the series. The seasonal differencing operator, denoted by ∇_p, is defined by the following formula:

$$\nabla_p X_i = X_i - X_{i-p} = (1 - B^p)X_i \qquad (IV.7)$$

Whereb B is the backward operator defined as follows:

$$B^p X_i = X_{i-p} \qquad (IV.8)$$

By applying the ∇p operator on the data, we obtain:

$$\nabla_p X_i = g_i - g_{i-p} + w_i + w_{i-p} \qquad (IV.9)$$

Unlike the moving average method, the differencing approach does not compute the trend component g_i explicitly. Instead, it estimates the difference between trend values, (g_i-g_{i-p}), thereby capturing the changes in the trend rather than the trend itself.

- **Seasonal Factor Estimation:** After estimating and subtracting the trend component g_i from the original time series data x_i for each value of i, the next step is to estimate the seasonal component s_i. This is typically done by analyzing the detrended series and averaging the values corresponding to the same positions within each seasonal cycle.

$$\forall 1 \le m \le n, \quad \exists k = \left\lfloor \frac{n}{p} \right\rfloor, \quad (S_m)_e = \frac{1}{k}\sum_{k=1}^{k}\left[(X_{kp+m} - (g_{kp+m})_e)\right] - \frac{1}{p}\sum_{m=1}^{p}(X_{kp+m} - (g_{kp+m})_e \qquad (IV.10)$$

Where $\lfloor \circ \rfloor$ denotes the integer part operator.

- **Residual Fluctuation Estimation:**

Once the seasonal component s_i has been estimated and subtracted from the detrended data X_i-$(g_i)_e$ for each value of i, the remaining values represent an estimate of the residual fluctuations w_i at each point in time. This residual component captures the irregular or random variations in the time series that are not explained by the trend or seasonal factors.

$$w_i = X_i - (g_i)_e - (S_i)_e \qquad (IV.11)$$

IV.1.2 Long Short-Term Memory Networks (LSTM)

Long short-term memory is an artificial neural network used in the fields of artificial intelligence and deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections, it able to address long-term dependency problem. LSTM has a structure known as a memory cell that is able to store an input value where it can handle problems of long-term dependencies and LSTM performs well on jobs with long data sets [97].

The memory cell consists of one principal layer and three gate controllers: input, forget and output. The principal layer analyzes the present entry x_t and the preceding short-term state h_{t-1}.The entry gate belongs to organizing the flow of new memories. Oblivion gate controls the memories that Long-term past state will be discarded $C_{t-1,}$ new memories obtained from the new long-term state C_t. The output gate is established what memories will be considered as the new LSTM cell output of a file specific time step, i.e., the y(t), which at some point during the process is equal to the new short-run state h(t) [98].

The process is expressed mathematically as such:

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * h_{t-1} + W_{ci}C_{t-1} + b_i) \quad (IV.12)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * h_{t-1} + W_{cf}C_{t-1+}b_f \quad (IV.13)$$

$$C_t = f_tC_{t-1} + i_t \tanh(W_{xc} * X_t + W_{hc} * h_{t-1} + b_c) \quad (IV.14)$$

$$O_t = \sigma(W_{xo} * X_t + W_{ho} * h_{t-1} + W_{co}C_cb_o \quad (IV.15)$$

$$h_t = O_t tanh(C_t) \quad (IV.16)$$

here:

$W_{xi,}$ $W_{xf,}$ $W_{xo,}W_{xc}$: The weight matrices for their connection to the input vector x_t.

W_{hi}, W_{hf}, $W_{ho,}$ W_{hc}: The weight matrices for their connection to the previous short-term state h_{t-1}.

b_f, b_c, $b_{i,}$ b_o: The bias terms of each layer.

IV.1.3 Extreme Gradient Boosting (XGBoost)

XGBoost is a powerful implementation of gradient-boosted decision trees, where the trees are built sequentially [99]. In this approach, weights play a crucial role each independent variable is assigned a weight before being passed through the decision tree to make predictions. If a variable is incorrectly predicted, its weight is increased, giving it more importance in the

next tree. This process continues across multiple trees, and the combination of these individual models forms a strong, accurate predictor. XGBoost can be applied to various tasks such as regression, classification, ranking, and prediction. The core steps of the algorithm are as follows:

Step 1: Gradient boosting starts by computing an initial constant prediction, F_0, using the observed data x. Mathematically, this step is expressed as:

$$F_i(x) = argmin \sum_{i=1}^{n} L(y_i, \hat{y}_i) \qquad (IV.17)$$

where y_i is the predicted value, L is the loss function, and argmin indicates that we must find a predicted value, $\hat{y}_i$, for which the loss function is least. Additionally, this phase entails utilizing the loss function provided by equation (IV.18) to determine the minimal value of the anticipated value. On this loss function's differentiation using the given equation (IV.19).

$$L(y_i, \hat{y}_i) = (y_i, \hat{y}_i)^2 \qquad (IV.18)$$

To find *yi* that minimizes ΣL, we are taking a derivative of ΣL with respect to y_i.

$$\frac{\partial}{\partial \hat{y}_i} \sum_{i=1}^{n} L(y_i, \hat{y}_i) = \frac{\partial}{\partial \hat{y}_i} \sum_{i=1}^{n} L(y_i - \hat{y}_i)^2 = -2 \sum_{i=1}^{n} (y_i - \hat{y}_i) = -2 \sum_{i=1}^{n} y_i + 2n\hat{y} / \forall \hat{y}_i = \hat{y} \forall i \quad (IV.19)$$

We are finding $\hat{y}_i$ that makes $\partial \Sigma L / \partial \hat{y}_i$ equal to 0.

$$-2 \sum_{i=1}^{n} y_i + 2n\hat{y} = 0 \Rightarrow \sum_{i=1}^{n} y_i = n\hat{y} \Rightarrow \frac{1}{n} \sum_{i=1}^{n} y_i = \hat{y} \qquad (IV.20)$$

It turned out that the value $\hat{y}_i$ that minimizes ΣL is the mean of y_i.

Step 2: The next step is to compute the pseudo-residual values, which can be obtained by dividing the observed value by the predicted value. This step can be written like this:

$$R_{im} = -\left[\frac{\partial L\big(y_i, F(x_i)\big)}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)} \qquad (IV.21)$$

For i=1, ..., n

$F(x_i)$ is the preceding model in this case, and the decision tree's (DT) number is m. The derivative of the loss function is calculated in this stage in relation to the predicted value.

$$R_{im} = -\frac{\partial L(y_i - F_{m-1})^2}{\partial F_{m-1}} \quad (IV.22)$$

$$= 2(y_i - F_{m-1}) \quad (IV.23)$$

Step 3: This step looks for the output values for each decision tree leaf. It may be important to determine the final production of fall leaves in cases where one leaf yields more than one remnant. The output can be found by simply taking the average of the numbers. This step can be expressed mathematically by:

$$y_m = argmin \sum_{i=1}^{n} L(y_m, F_{m-1}(x_i + yh_m(x_i))) \quad (IV.24)$$

where $h_m(x_i)$ is the DT made on residuals and m is the number of DTs.

Step 4: The final step is to revise the earlier models' predictions. Here's how it can be updated:

$$F_m = F_{m-1}(x) + V_m h_m(x) \quad (IV.25)$$

Where mis the number of performed decision trees, $F_{m\text{-}1(x)}$ is the base model prediction (previous prediction), and v_m denotes the learning rate.

IV.1.4 Performance measures

Statistical metrics, also known as cost functions, serve as key performance indicators for various models [100]. They enable us to assess how well a specific model performs. When evaluating time series models, RMSE (Root Mean Squared Error), MAPE (Mean Absolute Percentage Error), and MAE (Mean Absolute Error) are among the most widely used metrics [101]. These metrics measure the dispersion of observations around the mean, providing insights into model accuracy. While they share similarities, they operate on different scales, meaning their values are not directly comparable.

- **RMSE (Root Mean Squared Error):** Represents the standard deviation of the residuals, or prediction errors. It measures the average distance between observed data points and the values predicted by the regression line [102]. Mathematically, RMSE is expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (IV.26)$$

Here, $\hat{y}_i$ represents the predicted values, y_i denotes the observed values, and n is the total number of observations available for analysis. RMSE (Root Mean Square Error) is widely used and regarded as a versatile and reliable error metric for evaluating predictions.

- **MAPE (Mean Absolute Percentage Error):** Is the second most commonly used metric for evaluating time series models. It is mathematically represented by the following equation:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \qquad (IV.27)$$

This metric calculates the difference between prediction errors and divides it by the actual observed value. As a result, it is scale-independent, meaning it can be used to compare models across different datasets.

- **MAE (Mean Absolute Error):** quantifie the difference between estimated or forecasted values and the actual observed values. While it is commonly used in time series analysis, it can also be applied to various types of statistical estimations. MAE is defined as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (IV.28)$$

Here, y_i represents the actual values, and $\hat{y}_i$ denotes the predicted values.

IV.2 Dataset Description

The dataset utilized in this study originates from the publicly available database known as IHEPC (Individual Household Electric Power Consumption) [103]. It comprises a collection of electrical measurements recorded at regular intervals from a single-family household located in Sceaux, a suburban town in the Île-de-France region of France. This comprehensive dataset contains 2,075,259 records, spanning a period of 47 months**,** from December 2006 to November 2010**.**

The dataset captures several key features that characterize the household's electricity consumption profile. These features include:

- **Global Active Power (Pa)**: Measured in kilowatts, this variable represents the actual amount of electrical energy consumed by appliances to perform useful tasks such as heat-

ing, lighting, or driving a motor. It reflects the *real power* that is effectively used to do work.

- **Global Reactive Power (Pr)**: Expressed in kilovolt-amperes reactive**,** this parameter measures the energy temporarily stored and released by inductive or capacitive components such as motors, transformers, or compressors, and exchanged with the power grid. While not used directly to perform work, it affects the overall power quality and efficiency.
- **Voltage (V)**: Recorded in volts, this is the electrical potential difference between two points in the circuit, specifically measured between phases. It represents the driving force that enables the flow of electric current through conductors.
- **Current Intensity (I)**: Measured in amperes, this variable indicates the rate of flow of electric charge through the circuit, i.e., the amount of electricity passing through a specific point per unit of time.
- **Sub-Metering features (M_1, M_2, M_3):** In addition to the global measures, the dataset includes three specific sub-metering features, which break down the total household electricity consumption into distinct zones or equipment categories:

M_1: This sub-metering value captures the power usage of kitchen appliances, including the microwave, electric oven, stove, refrigerator, and others. These devices are characterized by frequent and rapid spikes in power consumption, especially during meal preparation periods.

M_2: This component includes the washing machine, tumble dryer, iron, and related appliances. Such devices typically operate in cycles, with varying power demands during different phases such as heating, washing, spinning, or drying.

M_3: This sub-meter measures the power consumed by high-load systems such as the electric water heater or the air conditioning unit. These appliances represent significant and often continuous energy demands, generally operating at high power levels and following cyclical usage patterns. Their influence is substantial on the daily load profile of the household.

IV.3 Data Preprocessing

Data analysis represents one of the fundamental pillars of modern statistics, playing a crucial role in both theoretical frameworks and practical applications. It encompasses a comprehensive set of methodologies designed to identify, summarize, model, and interpret the relationships among variables within a dataset. In real-world scenarios, data analysis is central to tasks such as forecasting and decision support, as it enables practitioners to uncover meaning-

ful patterns, reduce noise and anomalies, and predict future behavior based on observed historical trends.

Within the scope of this study, we propose a structured and systematic methodology for forecasting household electricity consumption. Our approach is grounded in the detailed analysis of time series data collected from actual measurements, with the ultimate goal of building a robust and accurate predictive system.

The proposed methodology unfolds in three key stages:

1. **Exploratory Data Analysis (EDA):** This initial phase involves a thorough examination of the raw data in order to understand its structure, distribution, and underlying patterns. Statistical summaries, visualizations, and correlation analyses are employed to detect trends, periodicities, seasonality, and potential outliers. This step provides the foundational insights necessary for subsequent stages of processing and modeling.
2. **Feature Engineering:** In this phase, relevant features are extracted or constructed from the original data to enhance the performance of predictive models. These features may include lagged values, rolling statistics, time-based indicators (such as hour of day, day of week), or domain-specific transformations. Effective feature engineering is essential to capture the temporal dependencies and contextual factors influencing electricity consumption.
3. **Temporal Modeling with Machine Learning Algorithms:** The final stage focuses on building predictive models using advanced machine learning techniques tailored to time series data. Various algorithms such as decision trees, gradient boosting, recurrent neural networks (RNNs), or long short-term memory (LSTM) networks are employed to learn the complex, nonlinear relationships within the temporal data. The objective is to produce accurate forecasts that can generalize well to unseen future data.

The complete architecture of the proposed forecasting framework is schematically illustrated in Fig IV.1, providing an overview of the sequential stages and the data flow between them.

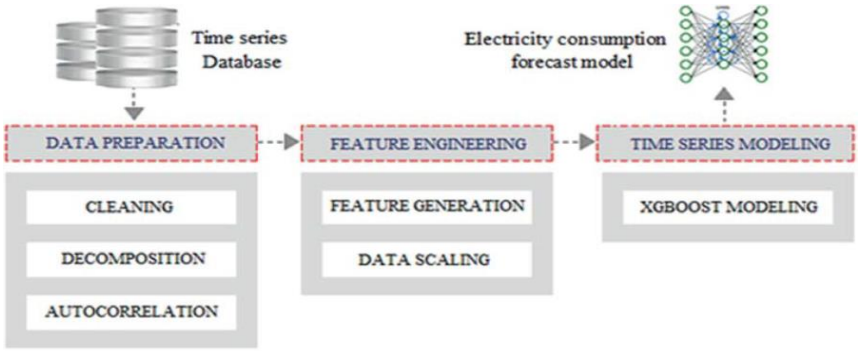


Fig IV.1: Proposed Framework for Electricity Consumption Forecasting

IV.3.1 Data Cleaning

Exploratory Data Analysis (EDA) is a foundational component in the process of data preparation and interpretation, particularly when dealing with time series. This preliminary step is designed to deeply investigate the nature and behavior of raw data prior to any formal modeling or predictive analysis. Specifically, EDA serves to:

1. Explore the statistical distribution of variables, including measures such as the mean, standard deviation, median, skewness, and kurtosis, which help characterize the shape and variability of the data.

2. Detect the presence of anomalies, including outliers, missing values, or inconsistent measurements that may distort the reliability of statistical inferences or affect model performance.

3. Highlight inter-variable relationships, through correlation analysis, scatter plots, or multivariate visualizations that may reveal patterns or dependencies across variables.

4. In the specific context of time series data, which are indexed chronologically, exploratory analysis enables dynamic visualization of how observed phenomena evolve over time. This temporal inspection provides several key methodological benefits:

- It facilitates the identification of underlying trends, which may indicate long-term directional shifts in the data.
- It reveals seasonal patterns or periodic fluctuations that recur at regular intervals (e.g., daily, weekly, or yearly).
- It exposes structural breaks or abrupt changes in behavior that might correspond to external events or modifications in usage patterns.
- It assists in diagnosing non-stationarity, which is a fundamental property that influences the selection and performance of forecasting models.
- Identification des tendances globales : on peut détecter une augmentation ou diminution progressive des valeurs, révélatrice d'un phénomène de fond (par exemple, l'augmentation de la consommation électrique au fil des années).

A. Handling of Missing Values in Sub-Metering M_3

During the exploratory analysis of the dataset, we observed the presence of missing values affecting certain variables, as illustrated in Fig IV.2. More specifically, the M_3 sub-feature, which records electricity consumption associated with either the water heater or the air conditioner**,** exhibited a missing value rate estimated at 1.25% over the entire time series.

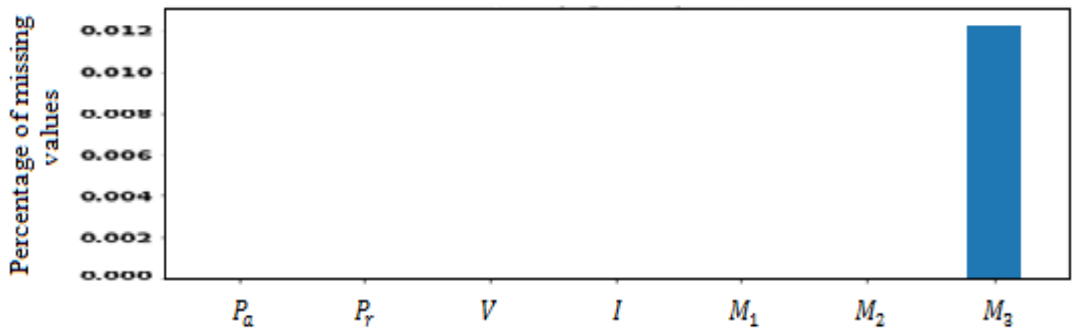


Fig IV.2 Percentage of missing values in the dataset

Although this rate may appear relatively moderate, appropriate handling of these gaps is crucial to ensure the reliability of subsequent analyses and to prevent bias during machine learning model training. Missing or inconsistent values, if left untreated, can lead to misleading statistical interpretations and degraded model performance.

It is important to note that missing values in this dataset are encoded in two distinct ways:

- With the **"?"** character in the timestamp column, indicating a loss or corruption in the temporal index.
- With the NaN (Not a Number) marker in the column corresponding to Sub-metering M_3, signaling an undefined or unavailable measurement at that time point.

Addressing these inconsistencies through a suitable imputation strategy or removal process is a necessary prerequisite before proceeding to model training or time series forecasting.

B. Choice of Imputation Method

In the context of time series data, missing values cannot be imputed arbitrarily or without consideration for the inherent temporal structure of the data. Indeed, successive observations are often temporally correlated, and their evolution may follow recurring patterns driven by usage habits or seasonal effects.

To preserve this temporal structure and avoid distorting the dynamics of the series, we adopted a time-lag imputation strategy. This method involves replacing each missing value with the value observed exactly 24 hours earlier, as defined by the following relation:

$$x_t = x_{t-1440} \qquad (IV.29)$$

where *t* represents the measurement time (in minutes), and 1440 corresponds to the number of minutes in a day. This method relies on the assumption that energy consumption patterns are relatively stable from one day to the next, particularly in the case of appliances with cycli-

cal or regular usage profiles, such as programmed water heaters or air conditioning systems operating during fixed time intervals. By exploiting this daily recurrence, time-lag imputation provides a coherent estimate of missing values without introducing artificial variability into the time series.

IV.3.2 Adopted Decomposition Model

In this study, we adopted an additive time series model, in which the observed series Y(t)Y(t)Y(t) is assumed to be the sum of three distinct components:

$$Y(t) = T(t) + S(t) + R(t) \qquad (IV.30)$$

where:

- $T(t)$ represents the long-term trend,
- $S(t)$ captures regular seasonal variations,
- $R(t)$corresponds to the residual component, interpreted as random noise.

Accordingly, we performed a detailed analysis of the time series derived from the IHEPC dataset. The results of this decomposition are presented in Fig IV.3.

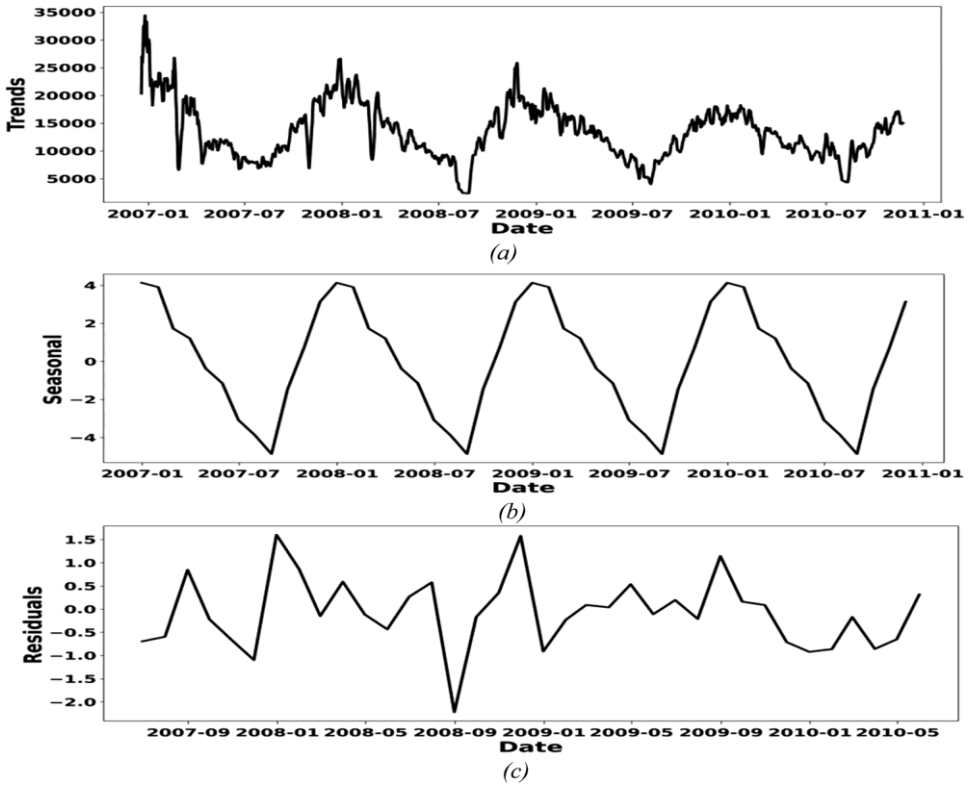


Fig IV.3 Household electricity consumption decomposition based on monthly granulate: *(a)* Trend, *(b)* Seasonality, and *(c)* Stationary (Residuals)

For this analysis, we employed a statistical method based on a convolution process between the observed values of the target variable and a filter defined using a Moving Average (MA) model. The additive assumption implies that the series can be decomposed into a simple sum of its structural components, trend, seasonality, and residuals, each of which is analyzed separately.

To implement this decomposition, the window size and frequency parameters were respectively set to 2 and 12, corresponding to a monthly moving average. This configuration allows for effective smoothing of the data, enabling clearer identification of underlying patterns and periodic structures.

The trend graph obtained from the IHEPC dataset, illustrated in Fig 3(a), offers an overview of the evolution of household electricity consumption over time. This visualization highlights several important observations:

- A significant increase in electricity consumption during the warmer months, particularly between June and August, likely due to the intensified use of air conditioning systems or other cooling appliances.
- A notable decrease during the colder months, from December to March, which may be attributed to reduced household activity or lower usage of electrical devices.
- The presence of an extreme value reaching 97.76 kW, recorded on July 31, 2008, at 16:21, representing a highly unusual peak in consumption within the entire time series.

From a statistical perspective, the trend graph derived from the IHEPC dataset serves as a valuable tool for extracting meaningful features that can be leveraged during the modeling phase of energy consumption forecasting.

By examining Fig 3(a), it becomes evident that the overall long-term trend of the time series is non-linear, yet it exhibits a quasi-linear behavior in the short term, particularly during the first and last six months of the study period. This local linearity is explained by a significant autocorrelation between observations over short time windows, suggesting that from a forecasting standpoint, a model trained on hourly or daily granularity is more effective than one based on monthly aggregation. Finer granularity allows the model to better capture short-term dynamics and local variations in consumption patterns.

Additionally, this figure reveals the presence of outliers, especially in April and October 2007, as well as in May 2008, which underscores the importance of applying data normalization or adjustment techniques before proceeding with model training.

Fig 3(b) illustrates the seasonal component, which captures recurring variations occurring at regular intervals. In this analysis, the frequency was set to 12, corresponding to monthly sampling. A clear periodicity can be observed in the electricity consumption pattern, with a higher spectral density at the beginning of each year. This indicates that long-term electricity consumption seasonality follows an annual cycle, a finding further confirmed by Fig IV.4, which shows a seasonal period of one year.

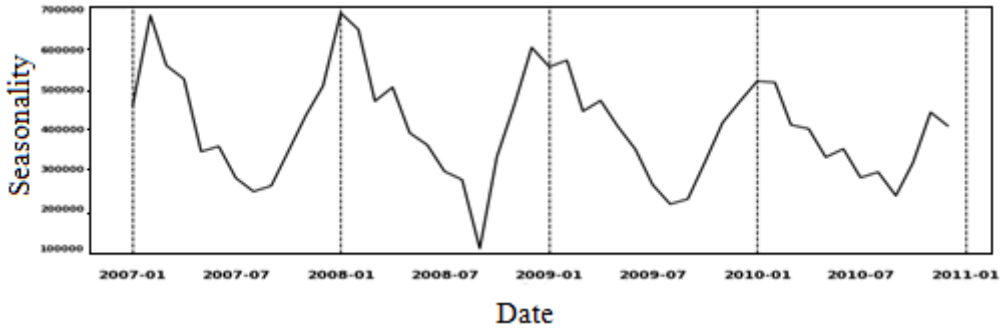


Fig IV. 4 Seasonality of household electricity consumption

Finally, Fig 3(c) displays the residual component of the time series, i.e., the portion not explained by either the trend or the seasonal components. This residual component plays a crucial role in assessing the predictability of the data. By analyzing the residual error statistics, we can evaluate the impact of anomalies on the model's forecasting performance.

In particular, if the mean of the residuals is close to zero and their distribution approximates a normal (Gaussian) law, it indicates that the seasonal component has been accurately modeled, and there is no need for additional seasonal adjustment to smooth out irregularities.

For this reason, our study includes in Table IV.1 a statistical summary of the residual errors, encompassing the mean, standard deviation, percentiles, and observed minimum and maximum values. This summary serves to assess the quality and reliability of the decomposition process applied to the IHEPC time series.

Table IV.1: Summary statistics of the residual errors

Mean	Standard deviation	Minimum Error	Quartile (Q1)	Quartile (Q2)	Quartile (Q3)	Maximum error
0.0042	-2.3193	-5.2181	-1.0298	-0.0092	1.1090	7.0728

The statistics presented in Table IV.1 indicate that the seasonal component does not require any further seasonal adjustment. This finding suggests that the residual variations, that is, the

differences between the observed values and those predicted by the seasonal model, are sufficiently balanced and do not necessitate corrective intervention.

This conclusion is further supported by Fig IV.5, which illustrates the distribution of residual errors. The figure shows that these errors closely follow a Gaussian (normal) distribution, characterized by a symmetrical shape centered around zero and a high concentration of values near the mean. This observation confirms that the residuals are primarily random, devoid of any systematic bias, and that the seasonal model accurately captures the periodic fluctuations in the data.

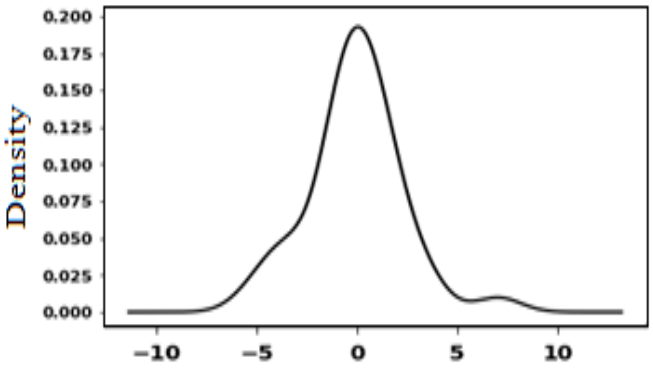


Fig IV.5 Household electricity consumption residual errors

IV.3.3 Autocorrelation Analysis

It is also essential to account for the persistence of the relationship between the current observations of a time series and its past values, commonly referred to as time lag. To assess this relationship, we employed the autocorrelation function applied to the electric power consumption time series.

Mathematically, observations denoted as y_t and y_{t-k} are separated by a lag k, which can be expressed in days, quarters, or years depending on the temporal resolution of the data. Based on the decomposition of the IHEPC dataset, we found that, in contrast to long-term consumption trends, there exists a significant autocorrelation among short-term observations (hourly or daily intervals).

In this study, to compute autocorrelation values at both hourly and daily granularities, we used the Pearson correlation coefficient, which ranges from −1 to 1.

To facilitate a comparative analysis of autocorrelation values across different lag sizes, Fig IV.6 presents the computed autocorrelation coefficients under two levels of granularity:

- **Daily granularity**:
  - *Fig 6(a)*: Autocorrelation with a lag of 7 days.
  - *Fig 6(b)*: Autocorrelation with a lag of 30 days.
- **Hourly granularity**:
  - *Fig 6(c)*: Autocorrelation with a lag of 60 hours.
  - *Fig 6(d):* Autocorrelation with a lag of 12 hours.

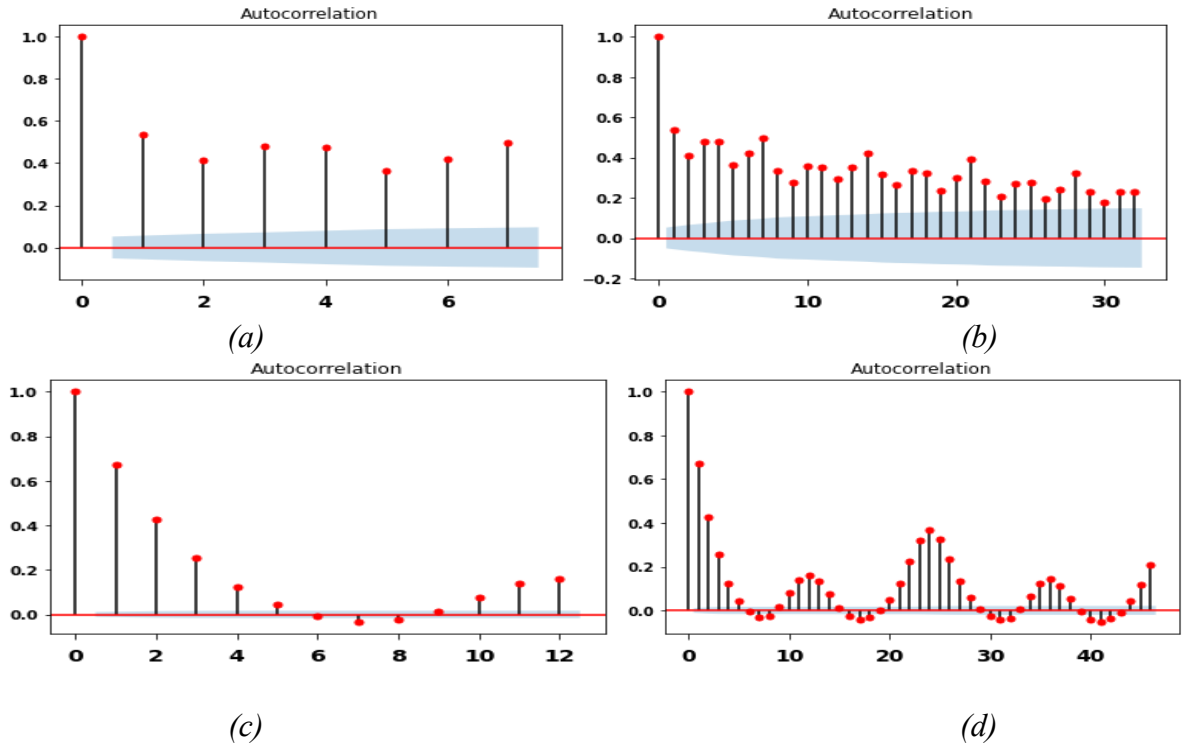


Fig IV.6 Autocorrelation of time series data re-sampled to*: (a-b)* daily mean values, *(c-d)* Hourly mean values

Based on this figure, several key observations can be made:

☞ **For the daily granularity**:

- As illustrated in Fig 6(a), there is a strong correlation between the 1st and 7th day, with an autocorrelation coefficient approaching 0.6, as well as a moderate correlation of approximately 0.4 between the 2nd and 6th day.

- Fig 6(b), which displays autocorrelation at a 30-day lag, also reveals a notable correlation between the 4th and 8th day, suggesting the presence of weekly seasonal patterns in the power consumption data.

☞ **For the hourly granularity**:

- According to Fig 6(c), a moderate autocorrelation is observed between the 13th and 36th hour, with a coefficient value around 0.2.
- In contrast, *Fig* 6(d) shows autocorrelation values very close to zero, indicating no significant correlation at this specific lag interval.

These findings highlight the importance of choosing an appropriate temporal resolution and lag structure when modeling the dynamics of electricity consumption, particularly when designing predictive models that exploit temporal dependencies.

IV.4 Feature Engineering

Feature engineering is a critical phase in the development of a reliable and high-performing predictive model. Its primary objective is to optimize the data representation by transforming, generating, or selecting relevant explanatory variables, thereby enhancing the learning capacity of artificial intelligence or machine learning algorithms. High-quality features can significantly reveal underlying structures within the data, thus facilitating the modeling process.

This engineering process consists of several essential sub-steps. Among them, feature construction is one of the most important. This step involves creating new variables from existing ones by combining, transforming, or cross-referencing them according to predefined rules, often inspired by domain-specific knowledge. For instance, in the context of energy consumption forecasting, it may be useful to derive, from a time-related variable, a new feature that indicates peak hours or off-peak periods. This type of construction allows the model to incorporate relevant latent information, which may be difficult to exploit in its raw form but is critical for enhancing predictive performance.

IV.4.1 Creation of New Features

This step is fundamental in the field of data analysis, as it involves selecting, manipulating, and transforming raw data into variables that can be effectively used by machine learning (ML) techniques. In the context of this work, rather than treating the IHEPC dataset as a multivariate time series, i.e., attempting to predict the electricity consumption associated with each sub-metering variable, we adopted a univariate approach. The goal here is to generate new features derived from our primary time series.

To this end, we designed an additional variable, referred to as M_4 (or *sub-metering 4*), calculated using the following formula:

$$M_4 = (P_a \times 1000/60) - \sum_{i=1}^{3} M_i \qquad (IV.31)$$

Where:

- **M_4** denotes the newly generated feature,
- **pa** refers to the global active power,
- **M_i** represents the values of the three existing sub-meterings, for i = 1, 2, 3.

The resulting variable M_4 thus represents the amount of active power consumed, per minute, by household electrical devices that are not covered by the existing measurements M_1, M_2, and M_3. This construction enables an estimation of the residual power consumption, i.e., the portion attributable to devices that are not specifically instrumented in the original measurement system.

IV.4.2 Normalization and Standardization

It is well established that most machine learning algorithms base their decision-making processes on the features extracted from data. These algorithms often rely on measures of similarity or distance between feature vectors to identify structures, clusters, or trends within the dataset. In this context, when feature values exhibit a certain level of homogeneity, meaning they are close to each other in terms of scale or distribution, the algorithms tend to converge more quickly toward an optimal solution. This homogeneity facilitates learning and generally enhances both the stability and accuracy of the trained model.

Conversely, when features show significant disparities in values, for instance, when two variables differ greatly in order of magnitude, this can lead to imbalances in the internal computations of the algorithm (such as Euclidean distances), potentially degrading learning quality. In such cases, models not only require more time to adapt to the data but may also yield inferior overall performance.

In this study, and as previously highlighted in the exploratory data analysis section, the IHEPC dataset contains several extreme values, representing atypical spikes in electricity consumption. These outliers can negatively affect the learning process if not properly addressed.

To assess the extent of variability across different variables in the dataset, we estimated the Probability Density Function (PDF) for each feature, as illustrated in Fig IV.7. This analysis provides insight into the statistical distribution of the data, helping to identify areas of high concentration as well as rare or extreme occurrences.

It is worth noting that the variable representing household voltage typically follows a normal (Gaussian) distribution, centered around 240 volts, with a typical range between 230 V and 250 V. Due to this regularity and the absence of significant outliers, this variable does not offer particularly relevant information for graphical visualization in this context.

As shown in Fig IV.7, we observed substantial variability both within individual features (intra-feature variation) and across different features (inter-feature variation) in the IHEPC dataset, both in terms of numerical values and units of measurement. This heterogeneity is especially pronounced for the voltage variable, whose scale differs markedly from that of other features.

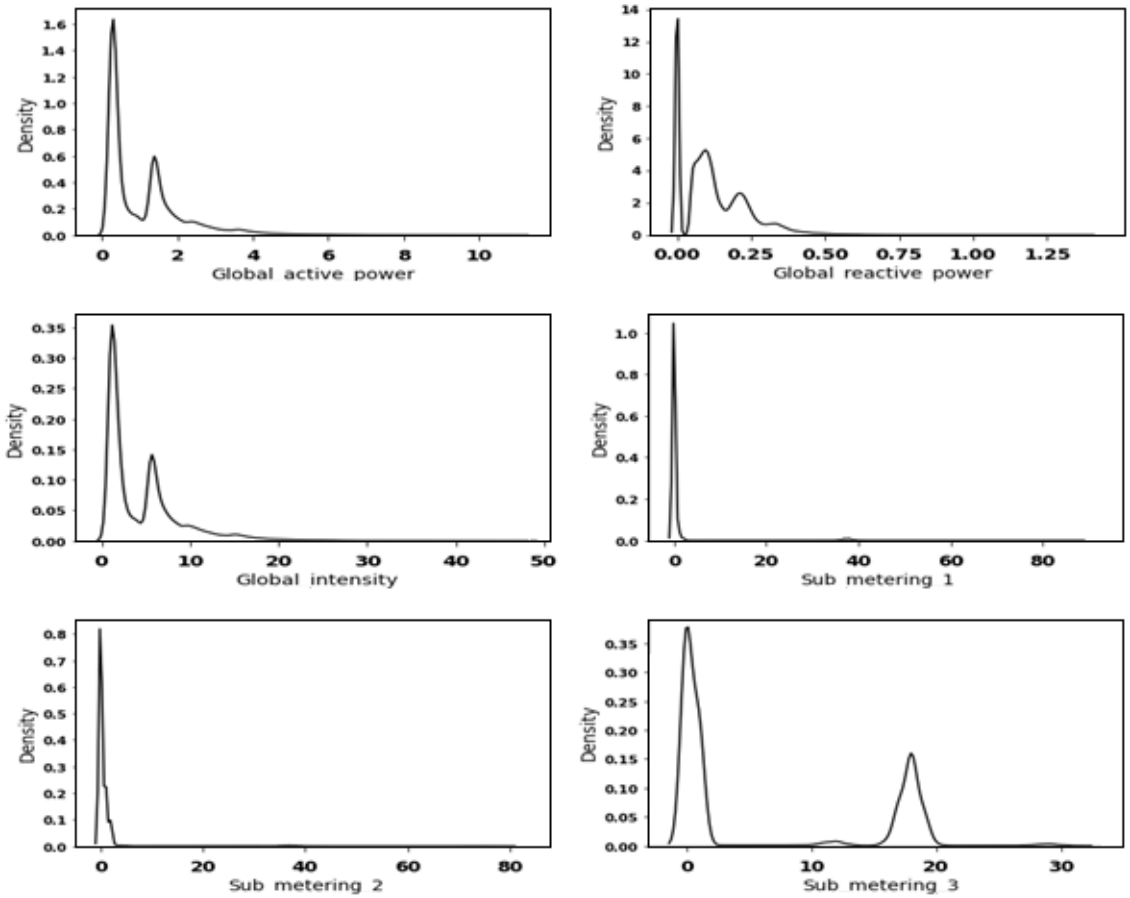


Fig IV.7 Probability density function of IHEPC features

To address these disparities in scale, magnitude, and units present in the data, it is crucial to harmonize feature ranges in order to ensure consistent learning by machine learning algo-

rithms. To this end, two of the most commonly used scaling techniques were applied in this work: normalization and standardization.

A. Normalization: The Min-Max normalization technique is particularly suitable when data features exhibit significant differences in scale across variables. This method involves rescaling all values of a given feature to fit within a fixed interval, typically between 0 and 1.

The primary objective of this transformation is to reduce the range of variation of a feature while minimizing the impact of extreme values (outliers) on machine learning algorithms. By constraining the data to a bounded range, normalization ensures a more balanced contribution of each feature during the learning process, especially in algorithms that are sensitive to distance metrics or gradient-based optimization.

Min-Max normalization is performed using the following formula:

$$\hat{X} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (IV.32)$$

Where:

- $\hat{X}$ is the normalized value of the feature,
- X is the original (raw) value of the feature,
- X_{min}, X_{max} are the minimum and maximum observed values for that feature, respectively.

This transformation preserves the relative proportions among values while rescaling them into a uniform range. It is particularly advantageous in models where relative amplitude comparison plays a central role, such as neural networks or clustering algorithms.

B. Standardization:

Standardization is another widely used data scaling technique, particularly well-suited when the input features follow approximately normal (Gaussian) distributions but exhibit different means and standard deviations. The aim of this transformation is to convert each feature into a standardized normal distribution, denoted as $X \sim N(0,1)$, with zero mean and unit variance. This approach is especially recommended when the dataset's features are expressed in heterogeneous units (e.g., volts, kilowatts, minutes), which could otherwise introduce bias or imbalance in the model's learning process due to unequal feature magnitudes.

Standardization transforms a given feature into a new variable that follows a standard normal distribution, using the formula:

$$Z = \frac{X - \mu}{\sigma} \qquad (IV.33)$$

Where:

- X is the original value of the feature,
- μ is the empirical mean of the feature values,
- σ is the standard deviation of the feature.

This transformation centers the distribution around zero, eliminating any bias associated with a non-zero mean, and reduces its spread by normalizing the variance. As a result, it facilitates the convergence of learning algorithms and often enhances the numerical stability of optimization processes.

IV.5 Proposed Methodology

One of the primary goals of the proposed system is to reduce the electricity bills incurred by residential users while simultaneously contributing to national energy optimization. The framework aims to shift away from static grid reliance by forecasting demand accurately and dynamically selecting the most cost-effective and available energy sources. This dual advantage supports users in minimizing financial charges and supports grid operators and policymakers in alleviating peak demand stress, reducing fuel imports, and progressing toward more sustainable, decentralized energy strategies.

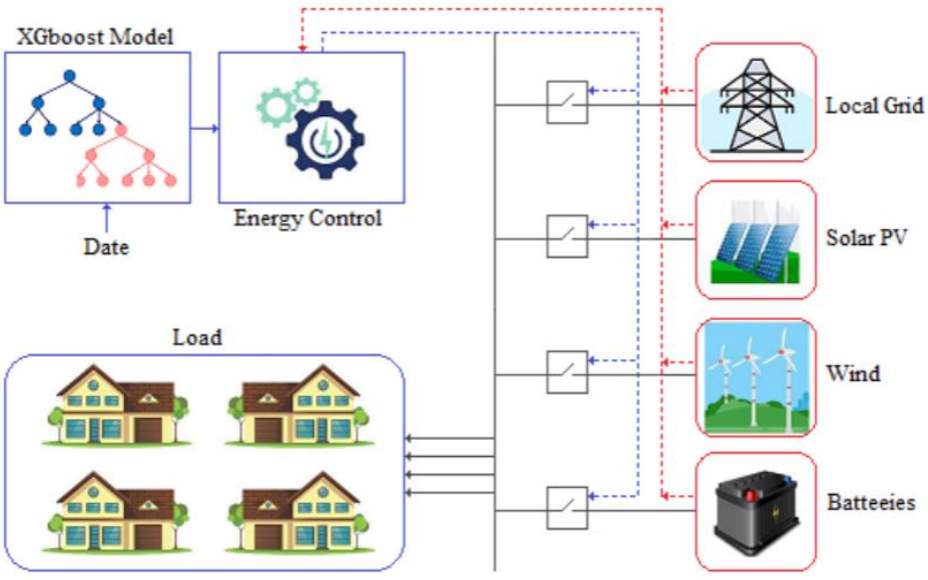


Fig IV.8 Intelligent energy source selection based on consumption forecast

The system is particularly tailored to residential environments, where demand forecasting and renewable integration present both high variability and untapped optimization potential.

The system architecture illustrated in Fig IV.8 integrates a forecasting and control mechanism with multiple energy sources to serve a cluster of residential loads. The forecasting unit, built using the XGBoost model, is responsible for predicting the expected energy demand using date-time and historical consumption features. This forecast acts as the primary input to the energy control logic, which determines the optimal energy source allocation.

The load represents a residential block composed of several houses, each with varying and unpredictable consumption behavior. The energy control block takes the predicted demand and evaluates it against the availability and cost of energy from four potential sources: the local utility grid, solar photovoltaic (PV) arrays, wind turbines, and battery storage systems.

Each source is interfaced with a switch that allows the controller to dynamically enable or disable its contribution to the overall supply. The system operates in a hierarchical logic where renewable energy (solar and wind) is prioritized when available, followed by battery discharging if the state of charge is sufficient. Only when renewable and stored energy are insufficient does the controller fall back on the grid supply. This logic ensures both cost reduction and increased reliance on clean energy sources.

Dashed lines in the figure represent control signals: the red paths indicate forecast-informed source activation commands, while the blue paths show real-time status feedback to the controller. The entire system operates under the assumption of smart meter integration and two-way communication, enabling real-time adaptability and data-driven control decisions.

IV.5.1 Forecasting Pipeline

The forecasting pipeline involves two distinct approaches: Long Short-Term Memory (LSTM) networks, which capture temporal dependencies in the data, and Extreme Gradient Boosting (XGBoost), which extracts decision rules from tabular features. Each model was trained independently and evaluated based on its predictive performance, using customized feature engineering workflows and training procedures specific to each method.

1. Feature Engineering Tailored to Enhance Predictive Model Performance

To ensure accurate forecasting, feature engineering was tailored to the strengths of each prediction model. For time series models like LSTM, constructing meaningful temporal sequences was essential to capturing patterns in electricity consumption. For the XGBoost mod-

el, which excels at interpreting structured data, we focused on generating a wide range of descriptive features, including lagged values and statistical summaries over recent time intervals. This allowed the model to make informed decisions without explicit sequence modeling. By aligning feature extraction with the internal logic of each algorithm, we maximized their respective predictive capabilities and ensured robust performance under varying consumption patterns. The Role of feature selection was crucial in improving performance by removing irrelevant variables through correlation analysis and feature importance derived from XGBoost. This process reduced overfitting and computational load. Domain knowledge guided the retention of key features—such as day of the week, hour, and seasonal indicators—resulting in higher accuracy and stability in forecasting outcomes.

2. XGBoost in Enhancing Energy Forecasting and Control

The XGBoost model complements the LSTM by offering a more interpretable and decision-oriented approach to forecasting. Unlike neural networks, XGBoost is capable of identifying complex feature interactions in structured datasets without needing sequential inputs. This makes it particularly effective at detecting threshold effects and non-obvious triggers in consumption, such as specific combinations of weather conditions and calendar attributes. Its structured output is particularly useful for downstream logic in the energy control module, allowing fine-tuned source selection based on expected load. The inclusion of XGBoost ensures that the forecasting framework is not only accurate but also robust across different data conditions and consumption behaviors.

IV.5.2 Energy Source Management Strategy

An accurate forecast of future energy consumption is only the first step in building a truly intelligent energy system. To translate predictive insights into tangible operational benefits, a robust energy source management strategy is essential. This strategy governs how the system responds to expected demand by orchestrating the use of multiple energy sources, including renewables, battery storage, and the utility grid. Instead of relying on static control rules, the proposed framework uses forecast-driven logic to dynamically determine the most efficient energy mix at any given time. By considering forecasted demand in conjunction with source availability and operational priorities, the system ensures energy is supplied reliably, economically, and with minimal environmental impact. The following sections detail the core mechanisms enabling this strategy.

1. Energy Forecasting: A Crucial Step for Effective Management

Integrating energy forecasts into the energy management strategy provides a predictive backbone for long-term planning and operational scheduling. Rather than reacting to real-time consumption alone, the system benefits from foresight into aggregated demand trends. This allows for better alignment between expected load and renewable energy availability, particularly for solar and wind sources, which often follow seasonal cycles. Forecasting consumption helps the system anticipate peak periods and strategize energy storage or grid usage accordingly. This integration enhances the system's ability to shift from reactive control to anticipatory decision-making, resulting in improved cost-efficiency and greater alignment with both user patterns and renewable generation dynamics.

2. Source Selection Logic

The energy source selection mechanism operates based on a hierarchical logic that prioritizes availability, cost-effectiveness, and sustainability. Renewable sources solar and wind are evaluated first, as they provide clean and virtually free energy when available. If the renewable supply is insufficient to meet the demand, the system then assesses the state of charge (SoC) of the battery to determine its suitability as a supplementary source. Only when both renewables and battery storage are inadequate does the system draw power from the main utility grid.

This tiered selection strategy ensures that energy consumption is optimized in real time, reducing both operational costs and environmental impact. Additionally, the control logic integrates forecasted energy production and demand to proactively adjust source priorities, enabling seamless and efficient energy allocation based on anticipated needs

3. Decision Criteria Based on Forecast and Availability

Effective energy source management relies on precise decision criteria that consider both forecasted consumption and the real-time status of each energy source. The control system continuously monitors forecast trends, battery state of charge, current production from solar and wind sources, and grid tariffs.

Decisions are made via a rules-based algorithm that compares forecasted demand against the projected supply from available sources. Thresholds are established to prevent premature battery discharge or unnecessary grid usage during periods when renewable energy availability is expected.

This logic enables the system to prioritize low-cost and sustainable sources without compromising the reliability of power supply to residential loads. By integrating forecasting into source selection, the system balances economic savings, supply security, and environmental responsibility.

IV.6 Conclusion

The methodology outlined in this chapter proposes a comprehensive and forward-looking framework for intelligent energy forecasting and source management in residential contexts. Although the initial evaluation phase began with the implementation of LSTM networks to explore the capabilities of deep learning in modeling sequential energy data, the primary focus was ultimately placed on the XGBoost model. This decision was made due to XGBoost's superior performance in terms of prediction accuracy, computational efficiency, and its ability to handle missing values and non-linear relationships effectively advantages that made it more suitable for the structured nature of the available electricity consumption data. Instead of relying on other models, XGBoost was selected as the central forecasting tool for its proven robustness and interpretability in real-world energy forecasting tasks. The system incorporates a flexible, forecast-informed control strategy aimed at optimizing energy source usage while minimizing both cost and grid dependence. Its multi-layered design utilizes short- and long-term consumption patterns, real-time data streams, and contextual inputs to guide proactive energy decisions. Feature engineering was adapted to suit model strengths, and the control logic integrates forecasting outputs with practical constraints such as renewable availability and battery capacity. This synergy between accurate forecasting and intelligent control underpins the proposed residential energy management system. The next chapter presents a detailed evaluation of the system's performance across various experimental scenarios and metrics.

Chapter V

Experimental Results and Interpretation

This chapter presents the results obtained from the different stages of the study, along with a comprehensive evaluation of the models used. It outlines the performance of the forecasting approaches and highlights the effectiveness of the proposed methods. The findings are analyzed in terms of accuracy and relevance, serving as a foundation for further developments, including the formulation of energy management strategies. The results are discussed in comparison with existing works to demonstrate the improvements achieved.

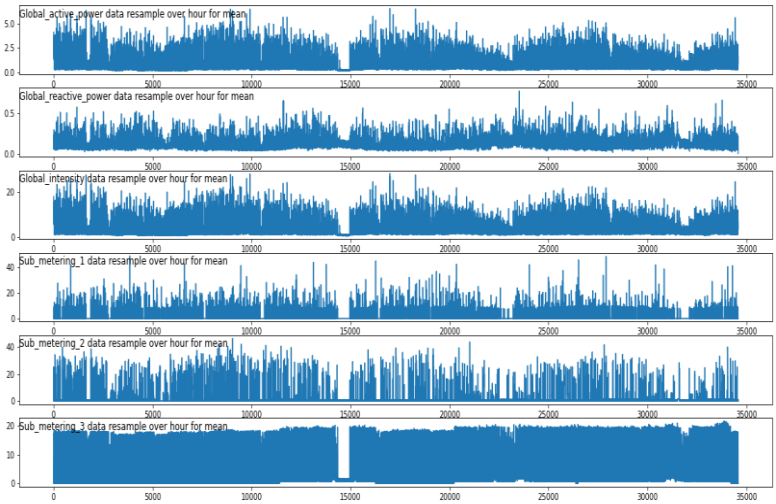
V.1 Experimental Evaluation

In this section, we present the experimental evaluation of the forecasting models applied to residential electricity consumption. The evaluation begins with an overview of results obtained from a Long Short-Term Memory (LSTM) model, as explored in related research, to provide a baseline for comparison. This serves as a reference point for performance assessment and highlights the capabilities of deep learning in time series forecasting. Following this, we focus on the proposed time series forecasting approach using the XGBoost algorithm. The primary objective is to assess the effectiveness and robustness of the XGBoost-based model when applied to electrical consumption prediction, considering both accuracy and practical applicability within smart grid environments.

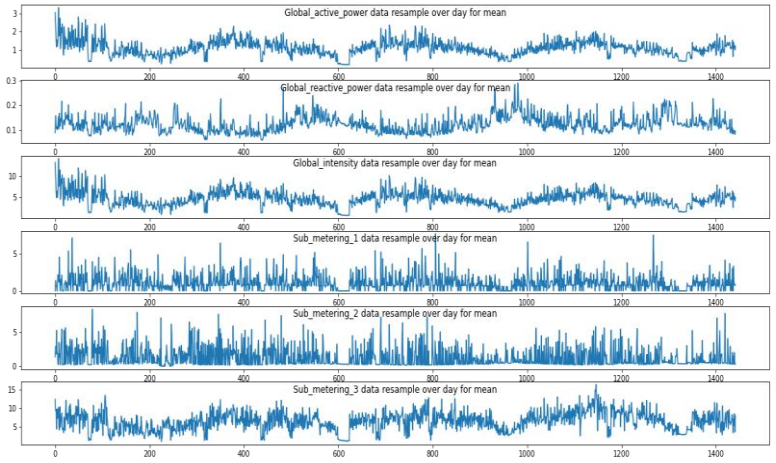
This section provides a comprehensive analysis of the experimental results obtained during the evaluation of the proposed time series forecasting approach. The primary objective is to assess the effectiveness and robustness of the XGBoost-based model applied to electrical consumption prediction.

V.2. LSTM-Based Electricity Consumption Forecasting

We first studied the performance of Long Short-Term Memory (LSTM) networks in predicting household electricity consumption when applied to real-world energy data, demonstrating their effectiveness in monitoring temporal patterns of household electricity use. We performed hourly, daily resampling as shown in Fig V.1. This significantly reduced the data size while maintaining periodic consumption patterns, which is critical for efficient LSTM training.



(a)



(b)

Fig V.1 Features Resampled, (a) Over Hour, (b) Over Day

The model was trained and tested separately on data resampled to hourly, daily intervals. Fig V.2 illustrates the hourly, daily original and the predicted electricity consumption values obtained by LSTM model. In case of hourly and daily granularities, we observe that the prediction curves follow the fluctuation of the original electricity consumption during the day

and night on an hourly basis, except for several peak usage hours. This occurs due to the fact that there are random variations for an individual home user.

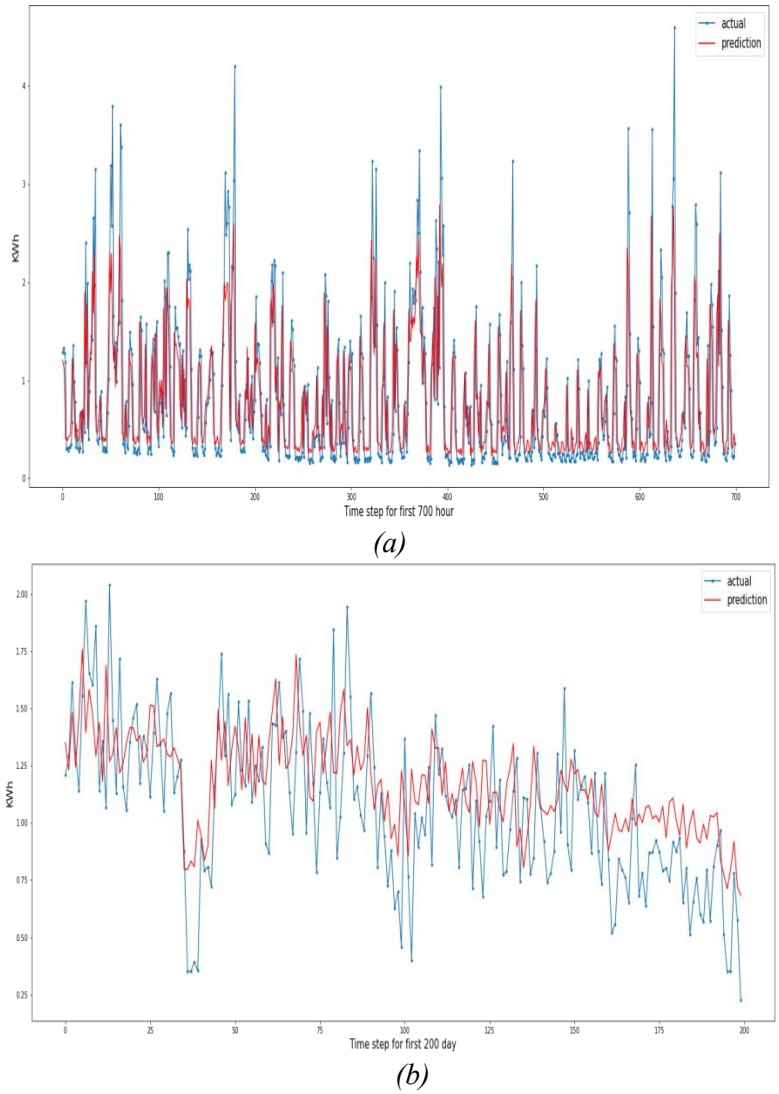


Fig V.2 Hourly and daily prediction results for LSTM model: *(a)* Hourly granularity, and *(b)* Daily granularity

As presented in the table V.1, the quantitative performance metrics support this observation. For the hourly dataset, the model attained a Root Mean Squared Error (RMSE) of 0.614. Using the daily aggregated data, the error values improved, with the RMSE decreasing to 0.333.

The LSTM model demonstrated better performance in predicting daily electricity consumption compared to hourly predictions. This can be attributed to the fact that daily values represent a more stable and aggregated view of consumption patterns, which reduces the effect of short-term fluctuations and noise present in high-frequency data. LSTM networks are designed to learn temporal dependencies over sequences, and they perform better when

the underlying patterns are relatively consistent. In contrast, hourly consumption data tends to be more volatile and influenced by sudden behavioral changes (e.g., turning on appliances), making it more difficult for the LSTM model to accurately learn and predict these rapid variations. Additionally, the task of predicting a single daily value is inherently less complex than predicting 24 separate hourly values, each with its own local dynamics and variability.

Table V.1 performance metric (RMSE)

RMSE	
Hourly	**Daily**
0.614	0.333

Following the implementation and evaluation of the LSTM model for forecasting electricity consumption, it was noted that the model yielded satisfactory results for daily aggregated data but exhibited limitations in capturing the high-frequency, nonlinear temporal dependencies present in hourly data. Given these challenges, the methodology was extended to include the XGBoost algorithm, a gradient boosting framework that excels in handling structured tabular datasets with engineered temporal features. XGBoost's ability to efficiently model complex nonlinear relationships and its robustness against overfitting make it a suitable alternative to recurrent neural networks for this forecasting task, particularly when feature engineering can explicitly encode temporal patterns.

V.2.1 Time Series Modeling

The methodological framework adopted in this study is primarily based on the XGBoost model an advanced and versatile supervised learning algorithm widely acclaimed for its high predictive performance and computational efficiency. As discussed previously, XGBoost (Extreme Gradient Boosting) is particularly effective in handling structured data and capturing complex patterns, making it well-suited for time series forecasting tasks.

To construct the predictive model, we first defined a target variable, denoted by y_i, which, in the context of our research, corresponds to the global active power. The remaining features in the IHEPC dataset, represented as x_i, are employed as predictor variables to estimate this target. The learning process begins with the creation of an initial decision tree, which is trained on the residuals derived from prior predictions. At each iteration, the algorithm computes a similarity score for the residuals, which assesses the homogeneity of the data points grouped within each leaf node of the tree, as well as the gain in similarity achieved

through successive splits. The output value assigned to each leaf is then determined based on the residuals it contains. The prediction generated by a tree is used to compute a new set of residuals, which in turn guide the construction of the next tree. This process continues in an iterative manner, where each new tree corrects the errors of the previous ensemble, thereby progressively refining the model. The iterations proceed either until the residuals cease to decrease significantly, indicating convergence, or until a predefined maximum number of iterations (or epochs) is reached. In the final prediction phase, the outputs from all trees are aggregated by weighting each tree's contribution using a learning rate. This ensemble output is then added to the initial prediction to yield the final, optimized estimation of the target variable. To provide a clearer understanding of this mechanism, Fig V.3 illustrates the internal workflow of the XGBoost algorithm during the training phase.

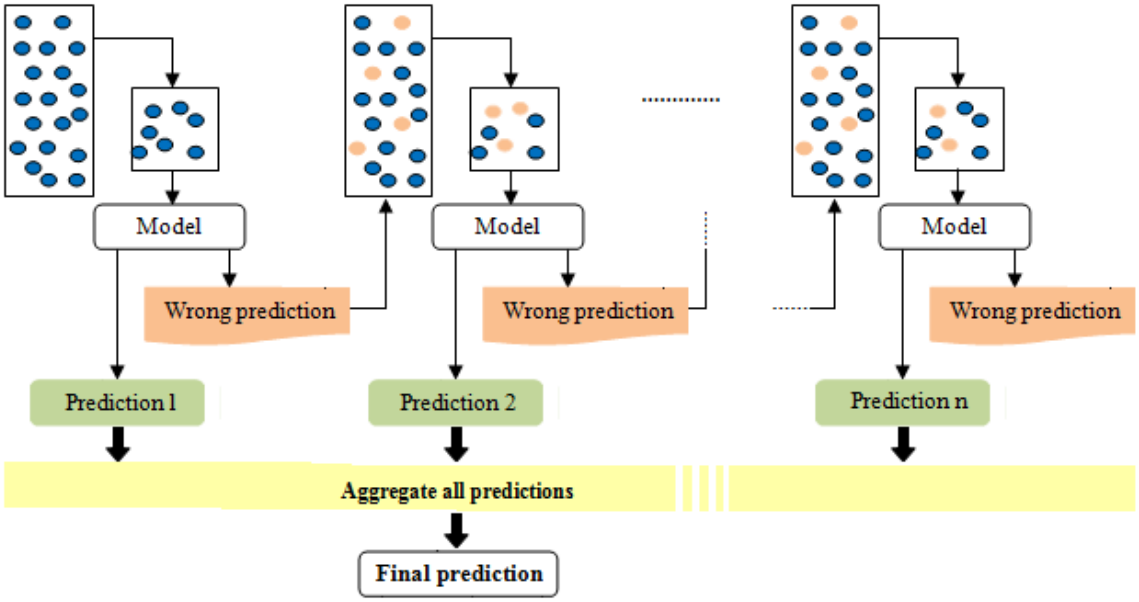


Fig V.3 XGBoost based modeling process

To conduct this evaluation, we utilized the Individual Household Electric Power Consumption (IHEPC) dataset, as previously described. For methodological consistency, the dataset was partitioned into two distinct subsets:

- The training set, comprising data collected from December 26, 2006, to December 31, 2009, was used to train the forecasting models and calibrate their internal parameters.
- The testing set, encompassing the period from January 1, 2010, to November 26, 2010, served to validate the forecasting performance and determine the most suitable prediction horizon.

The experimental protocol was structured into three sequential phases, each designed to address a specific aspect of model performance:

Phase 1: Baseline Evaluation with Default Parameters

In this initial phase, we assessed the predictive capabilities of the XGBoost model using its default hyperparameters. The raw dataset, without any form of preprocessing or scaling, was used as input. This baseline evaluation establishes a reference point against which subsequent optimizations can be compared.

Phase 2: Hyper parameter Optimization

Building on the baseline results, the second phase focuses on fine-tuning the hyperparameters of the XGBoost model to enhance its forecasting accuracy. Parameters such as the number of trees, maximum tree depth, learning rate, and sub-sampling ratio were adjusted systematically using grid search and cross-validation techniques. The aim of this optimization process is to identify a configuration that minimizes forecasting error and maximizes generalization to unseen data.

Phase 3: Evaluation with Feature Scaling

In the final phase, the best-performing XGBoost configuration obtained in Phase 2 was re-evaluated on a version of the dataset that underwent feature scaling (using normalization or standardization). This step aims to assess the impact of data preprocessing on the model's learning efficiency and stability. It also helps mitigate potential issues related to feature magnitude disparities that could adversely influence the model's convergence behavior or skew its predictions.

Through these three phases, the experimental design ensures a systematic and rigorous assessment of the model's behavior under different training conditions. The outcomes of each phase are analyzed in terms of various performance metrics, which will be discussed in detail in the subsequent sections.

V.3. Forecasting Electricity Consumption Using XGBoost

In the initial phase of our experimental study, the time series data was processed to generate two versions of the dataset, corresponding to hourly and daily levels of temporal resolution. This transformation was motivated by the temporal patterns identified during the exploratory data analysis, where recurring dependencies between current and historical values were highlighted.

Following this resampling, we employed the XGBoost regression model to predict future electricity consumption values. For this preliminary experiment, the model was executed using its default configuration settings, namely: Number of estimators: 100, Maximum Tree depth: 6, Learning rate: 0.1, and Booster type: gbtree.

These default parameters served as a baseline to assess the model's initial ability to learn consumption patterns from the IHEPC dataset without any fine-tuning or advanced preprocessing techniques.

The forecasting task was then carried out separately for the two granularities. The resulting predictions are presented in Fig V.4, which displays the estimated electricity consumption over time for both the hourly and daily intervals. These graphical results offer a first visual insight into the model's predictive behavior and set the stage for further refinement in subsequent phases of experimentation.

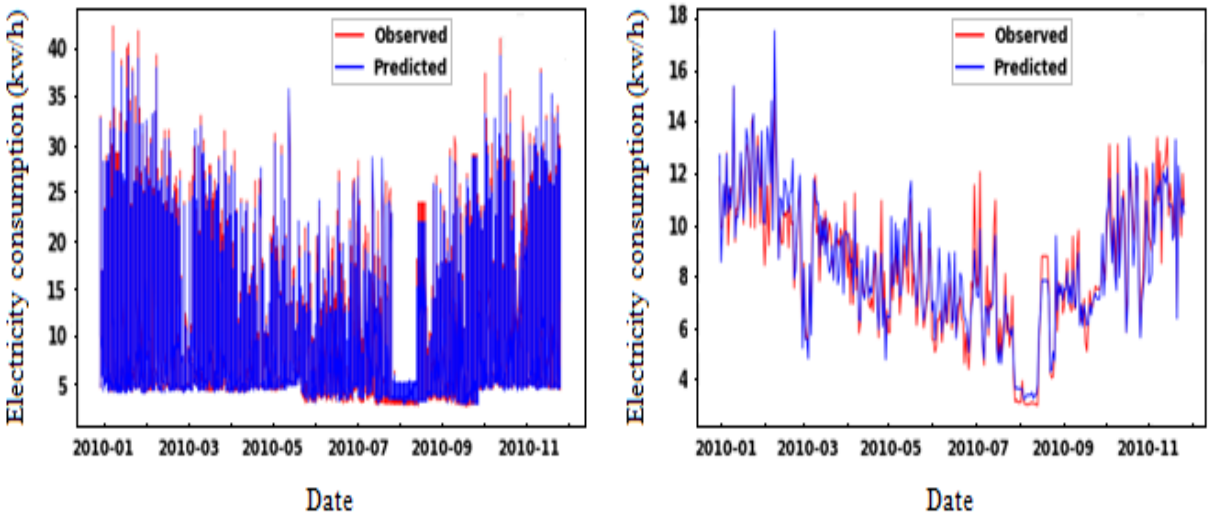


Fig V.4 Household prediction vs original observations re-sampled to: *(a)* Hourly granularity, and (*b)* Daily granularity

A careful visual inspection of the forecasted results reveals that the model effectively captures the overall trends and fluctuations of the electricity consumption across both hourly and daily time scales. The predicted curves closely follow the actual consumption patterns, demonstrating strong alignment with the observed data. However, some deviations are noticeable during specific peak hours and days, which can be attributed to the inherent randomness and unpredictability of individual household usage behaviors.

Quantitatively, the performance metrics reinforce this observation. For the hourly dataset, the model achieved a Root Mean Squared Error (RMSE**)** of **1.120** and a Mean Absolute Percentage Error (MAPE) of **0.108**. When using the daily aggregated data, these error values improved, with an RMSE of **0.955** and a MAPE of **0.090**. These results suggest that while the model performs well overall, its predictive accuracy is somewhat higher when considering longer time intervals, likely due to the smoothing effect of daily aggregation reducing noise in the data.

V.3.1 Hyper parameter Optimization for the XGBoost Model

While the initial results demonstrated promising predictive performance, there remained room for improvement, particularly in accurately forecasting consumption during peak periods, where variability tends to be higher. To address this, we undertook a systematic tuning of the XGBoost model's hyperparameters with the goal of enhancing its overall accuracy and robustness.

Our approach to hyper parameter optimization employed a grid search strategy, which exhaustively tests combinations of key parameters to identify the best-performing configuration. Specifically, we focused on three critical hyperparameters known to strongly influence model behavior:

- **Number of estimators (trees):** tested values were 500, 1000, and 1500.
- **Maximum tree depth:** tested values included 2, 4, 6, and 8.
- **Learning rate:** evaluated at 0.05, 0.10, 0.15, and 0.20.

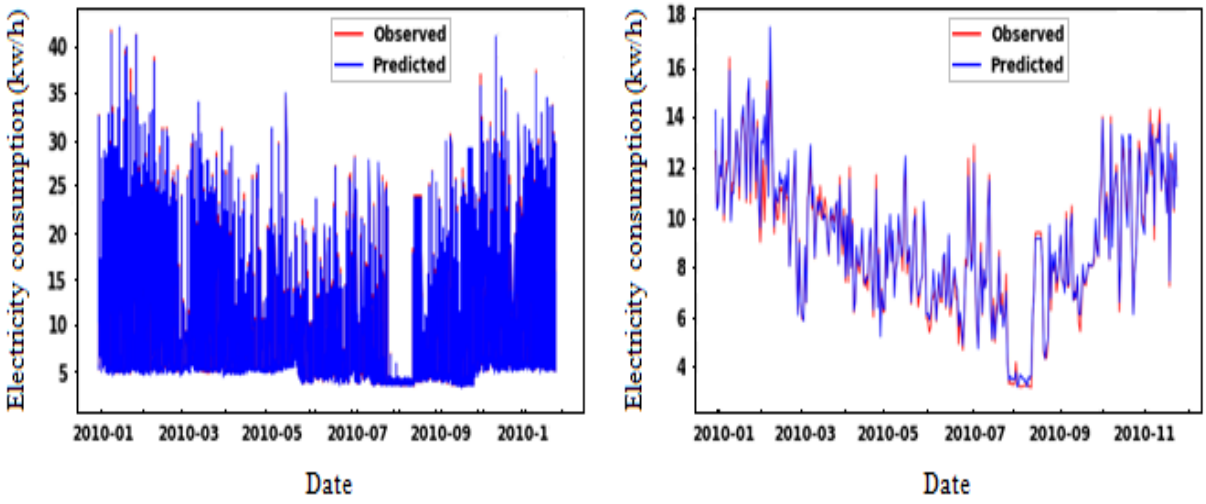
Additionally, two types of boosters were considered to build the trees: **gbtree** (gradient-boosted trees) and **linear booster** (linear functions). Combining all these options resulted in a comprehensive search space comprising **468** unique model configurations.

After training and evaluating all candidate models on both hourly and daily resampled datasets, the optimal hyper parameter set was identified as follows:

- **Number of estimators:** 1500
- **Maximum depth:** 2
- **Learning rate:** 0.15
- **Booster type:** gbtree

This configuration yielded a significant reduction in forecasting errors compared to the baseline model. For the hourly data, the model achieved an RMSE of **0.326** and a MAPE of **0.0271**, demonstrating much tighter alignment with actual consumption values. On the daily aggregated data, the tuned model produced an RMSE of **0.412** and a MAPE of **0.034**, indicating strong predictive performance over longer time intervals as well.

Fig V.5 presents the refined electricity consumption forecasts on both hourly and daily scales after the hyper parameter tuning, highlighting the enhanced precision of the optimized XGBoost model.



(a) *(b)*

Fig V.5 Household prediction vs original observations after XGBoost hyper-parameters tuning resampled to*: (a)* Hourly granularity, and *(b)* Daily granularity

By examining the prediction curves presented in this figure, it is evident that the model's forecasting accuracy has markedly improved following the hyper parameter tuning process. The adjusted curves now more closely follow the actual fluctuations observed in the original electricity consumption data, capturing subtle variations and peak consumption periods with greater fidelity than the initial XGBoost model that used default parameters. This improvement underscores the effectiveness of systematically optimizing model settings to better adapt to the intrinsic patterns within the data.

Furthermore, based on established benchmarks in the literature [99, 100], a model is generally considered reliable and efficient if the Mean Absolute Percentage Error (MAPE) remains below **0.30**. Our experimental results clearly place the tuned XGBoost model well within this threshold, indicating a high level of predictive precision. This strong performance can be attributed to two main factors: first, the thorough exploratory data analysis, which

allowed us to understand and preprocess the dataset effectively; and second, the meticulous tuning of the model's hyperparameters, which enhanced its capacity to generalize and predict future consumption accurately.

Despite this significant advancement, it is important to acknowledge that there is still potential for further enhancement. In the subsequent phase of our experiments, we will explore the impact of applying feature scaling techniques, such as normalization and standardization, to the dataset. These methods aim to harmonize the scales of input variables, potentially enabling the model to learn more effectively and yield even better forecasting results.

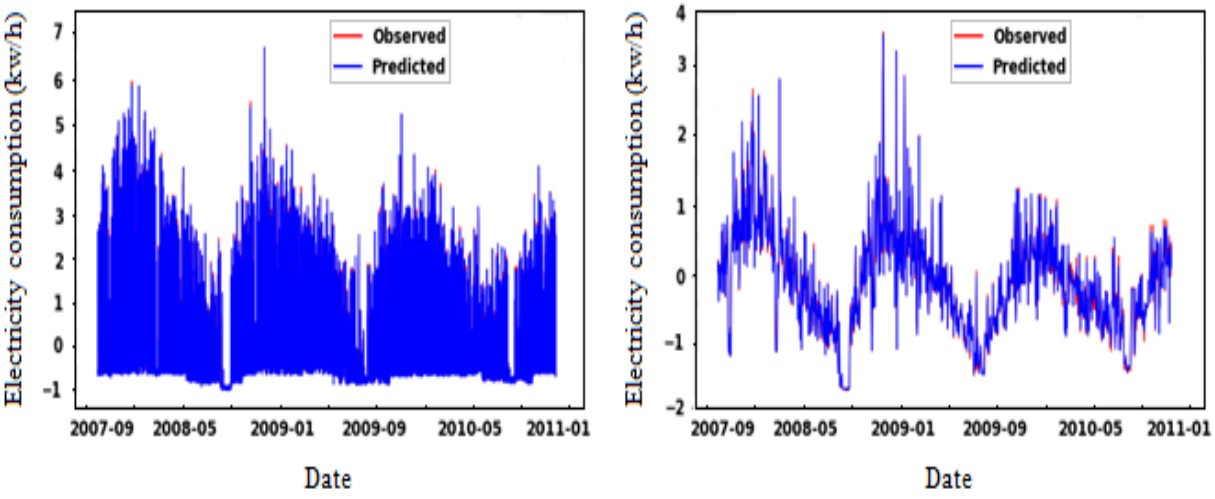
V.3.2 Model Evaluation with Feature Scaling Applied to Dataset

As highlighted during the exploratory data analysis, the IHEPC dataset exhibits extreme values, often referred to as outliers that correspond to unusually high electricity consumption instances. These irregularities create significant variability in the data, resulting in differences in the scales of the input features. Such discrepancies in magnitude can negatively impact the predictive accuracy of the XGBoost model by causing it to focus disproportionately on larger values, thereby reducing its ability to generalize effectively across all data points.

To address this issue and enhance the robustness of our forecasting model, this phase involves applying data scaling techniques to the input features. Specifically, the two scaling methods discussed earlier, normalization and standardization, are utilized to transform the dataset. Normalization rescales the feature values to fit within a fixed range, typically between 0 and 1, while standardization adjusts the features to have a mean of zero and a standard deviation of one, thus standardizing the variance.

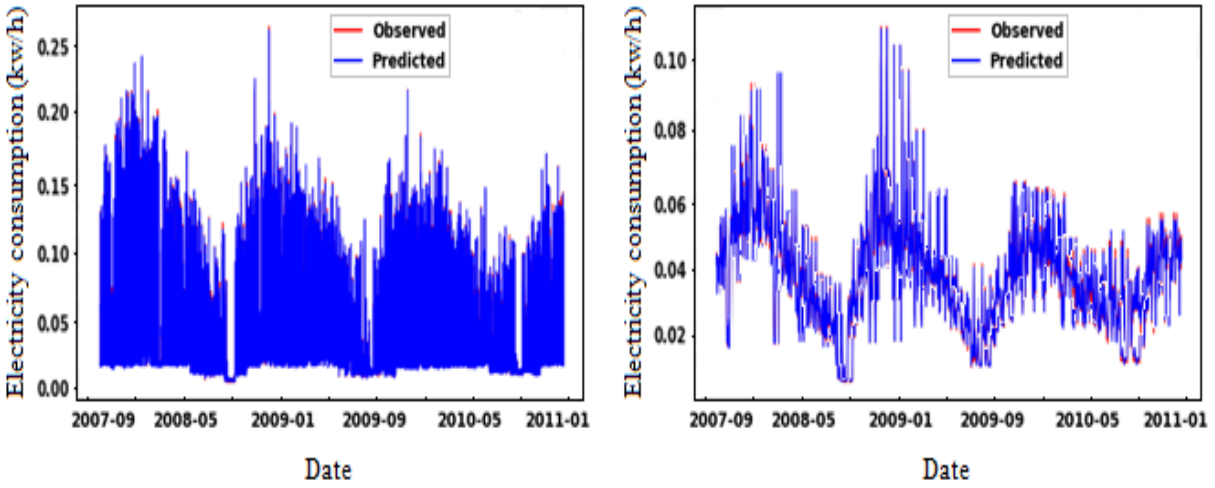
Once the datasets are scaled using these techniques, each transformed version is individually used to train the XGBoost regression model configured with the optimized hyperparameters identified in the previous experiments. This approach allows us to evaluate how the different scaling strategies influence the model's ability to forecast electricity consumption with higher precision.

The results of this analysis are presented in Figures V.6 and V.7, which display the predicted electricity consumption at hourly and daily intervals after the scaling transformations. These figures illustrate the impact of normalization and standardization on the forecasting accuracy, highlighting potential improvements in the model's capacity to track consumption patterns more accurately.



(a) *(b)*

Fig V.6 Household predictions vs. original observations after features normalization: *(a)* Hourly granularity, *(b)* Daily granularity.



(a) *(b)*

Fig V.7 Household predictions vs. original observations after features standardization: *(a)* Hourly granularity, *(b)* Daily granularity

From a careful examination of Figures V.6 and V.7, it is evident that the prediction curves exhibit significant improvement compared to the previous models that did not incorporate feature scaling. Specifically, the models trained on the normalized datasets show a particularly strong ability to closely mirror the fluctuations in the actual electricity consumption data. This enhanced tracking indicates that normalization effectively adjusts the feature values in a way that allows the model to capture subtle variations and patterns that were less well represented before scaling.

Moreover, when comparing the two-scaling methods normalization and standardization both approaches contribute to reducing the volatility observed in the forecasted curves. This means the predicted consumption values vary less erratically over time, producing smoother

and more stable prediction trends. This stability is important because it suggests the model is less sensitive to noise and extreme values, leading to more reliable forecasting outcomes.

To quantitatively support these visual observations, key performance metrics, namely the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE), were calculated for each scaling method and are compiled in Table V.2. These metrics provide an objective evaluation of the prediction accuracy, confirming the visual improvements seen in the forecast curves. The reduced RMSE and MAPE values demonstrate that the scaled feature datasets enhance the model's precision, with normalization showing a particularly noteworthy impact on improving prediction quality.

Table V.2 Comparative results of the impact of the two scaling methods

Standardization				**Normalization**			
Hourly		**Daily**		**Hourly**		**Daily**	
RMSE	**MAPE**	**RMSE**	**MAPE**	**RMSE**	**MAPE**	**RMSE**	**MAPE**
0.346	0.317	0.390	0.233	0.229	0.026	**0.129**	**0.020**

Table V.2 presents a comparative overview of the model's performance when applying two different feature scaling techniques, standardization and normalization, across two-time granularities: hourly and daily. The evaluation metrics used to assess the prediction accuracy are the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE), both widely recognized for measuring forecast precision and error magnitude.

Analyzing the results for standardization, we observe that the RMSE values are **0.346** for the hourly granularity and **0.390** for the daily granularity, while the corresponding MAPE values are **0.317** and **0.233**, respectively. These figures indicate a moderate level of prediction error, with performance slightly better for the hourly data compared to daily data, but overall higher errors relative to normalization.

In contrast, the normalization method demonstrates a marked improvement in both metrics. For hourly data, the RMSE drops significantly to **0.229** and the MAPE decreases to **0.026**. This reduction means the model is far more precise and consistent in its hourly electricity consumption predictions after normalization is applied. Even more impressive are the results for the daily granularity, where the RMSE falls further to **0.129** and the MAPE reaches an exceptionally low value of **0.020**. These figures highlight that the model can very accurately forecast daily electricity consumption with minimal percentage error and deviation from observed values.

The outstanding performance of normalization over standardization may be attributed to its ability to rescale all feature values to a uniform range (typically between 0 and 1), which effectively mitigates the influence of extreme values and disparate measurement units in the dataset. This consistent scaling facilitates the learning process of XGBoost, leading to faster convergence and better generalization.

V.4. Comparative Analysis with Existing Approaches

After rigorously assessing the performance of the proposed forecasting model, this section is dedicated to benchmarking its effectiveness against a selection of state-of-the-art approaches available in the literature. To ensure the integrity and fairness of this comparative analysis, we have restricted our scope to studies that utilized the same dataset, the Individual Household Electric Power Consumption (IHEPC) dataset. This criterion ensures that the observed performance differences arise from model architecture and processing techniques, not from dataset discrepancies.

The comparison spans three temporal granularities of the IHEPC dataset: second-level (every 15 seconds), hourly**,** and daily aggregations. These varying resolutions introduce distinct prediction challenges, such as high-frequency noise at the second-level and smoother long-term trends at the daily level.

Table V.3 The comparative analysis of the proposed method with state-of-the-art (DL and ML techniques) for hourly data resolution

Methods	**Papers**	**Model**	**15 s resampled**		**Hourly down sampled**			**Daily down sampled**		
			RSME	**MAE**	**RMSE**	**MAPE**	**MAE**	**RSME**	**MAPE**	**MAE**
Deep learning	Khan et al., 2020 [76]	CNN +LSTM	**-**	**-**	**0.31**	**0.19**	**0.47**	**0.11**	**0.08**	**0.69**
	Sajjad et al 2020 [73]	CNN-GRU	**-**	**-**	**-**	**-**	**-**	**0.47**	**-**	**0.33**
Machine learning	Puspita et al 2019 [69]	SVM	**-**	**-**	**13.86**	**-**	**11.48**	**-**	**-**	**-**
	Rajabi et al. 2019 [71]	CNN-2D-RP	**0.79**	**0.59**	**-**	**-**	**-**	**-**	**-**	**-**
		SVM	**1.25**	**1.12**	**-**	**-**	**-**	**-**	**-**	**-**
		ANN	**1.15**	**1.08**	**-**	**-**	**-**	**-**	**-**	**-**
	Proposed	**XGBoost**	**-**	**-**	**0.229**	**0.026**	**-**	**0.129**	**0.020**	**-**

V.4.1 Hourly Prediction Performance

In the context of hourly forecasts, as shown in table V.3, the proposed XGBoost-based approach demonstrates significant improvements over existing models cited in works such as

[69, 71, 73, 76]. Particularly, the results highlight that our model consistently produces the lowest RMSE and MAPE values compared to these references. Among these works, while [74, 72] delivered competitive RMSE values **(~0.129)**, the proposed model surpasses them, especially in terms of MAPE, where it achieves a significantly lower value of **0.020**. This indicates a more stable and reliable prediction accuracy across different consumption patterns, including periods with erratic or extreme energy usage.

V.4.2 Daily and Second-Level Forecasts

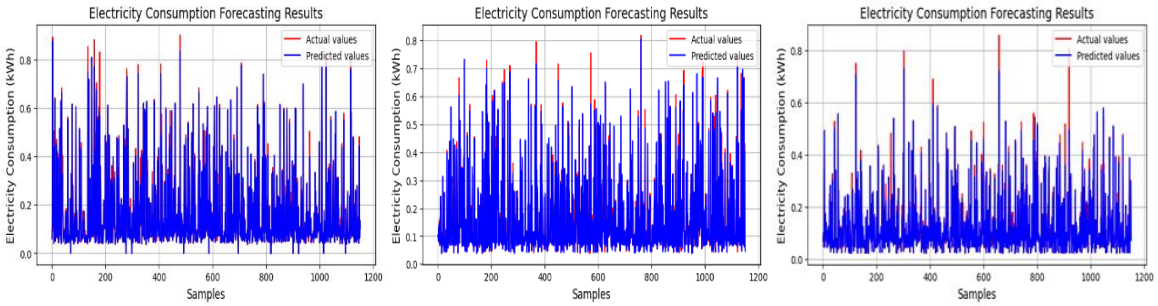
For both daily and second-level (15-second) predictions, the model maintains its advantage. When compared to methodologies in [70, 72, 77], our XGBoost model delivers the smallest error margins across all tested scenarios. Specifically, it achieves an RMSE of **0.229** and MAPE of **0.026** for the second-level aggregated data, and an RMSE of **0.129** and MAPE of **0.020** for daily predictions. These results reaffirm the model's capability to adapt effectively to both short-term and long-term temporal structures in household electricity consumption data.

The consistently lower error rates in both hourly and daily resolutions highlight the strength of the proposed method. The combination of thorough data preprocessing, effective hyper-parameter tuning, and careful feature scaling likely contributed to this enhanced performance.

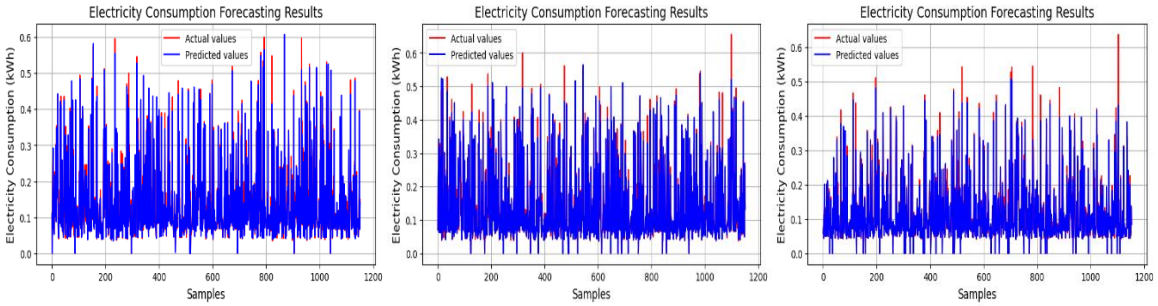
These results suggest that the model can predict power consumption with very small deviations from observed values, thus offering high precision for both short-term and medium-term forecasting horizons.

V.5. Electricity Consumption Forecasting for Multiple Households Using XGBoost

To further evaluate the robustness and scalability of the forecasting approach, the model was extended from a single-household scenario to include six different homes within the same geographic area. This transition aimed to simulate a more realistic smart grid environment, where energy forecasting must accommodate diverse consumption behaviors.



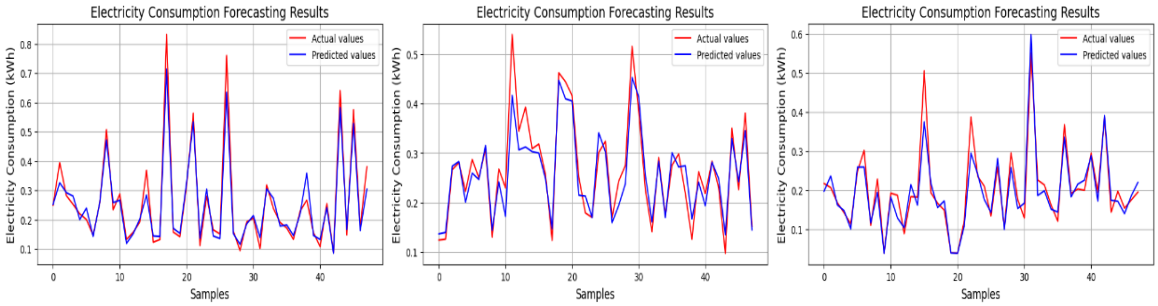
(a) *(b)* *(c)*



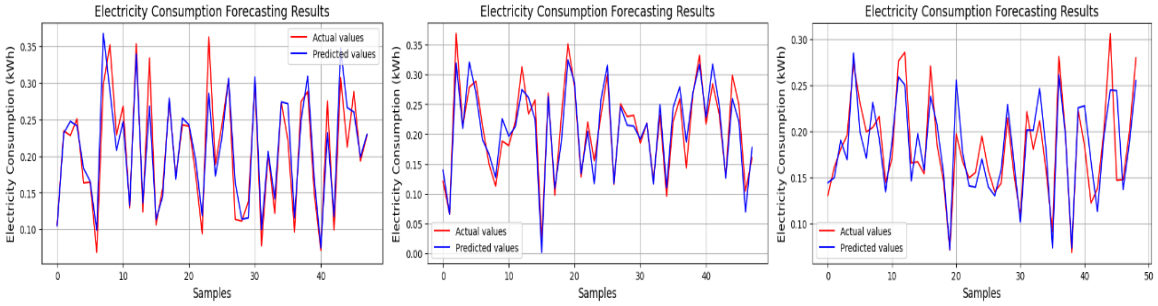
(d) *(e)* *(f)*

Fig V.8 Household prediction vs original observations (Hourly granularity) For 6 homes:

(a) Home 1, *(b)* Home 2, *(c)* Home 3, *(d)* Home 4, *(e)* Home 5, *(f)* Home 6



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



(d) *(e)* *(f)*

Fig V.9 Household prediction vs original observations (Daily granularity) For 6 homes:

(a) Home 1, *(b)* Home 2, *(c)* Home 3, *(d)* Home 4, *(e)* Home 5, *(f)* Home 6

Figures (V.8, V.9) represent the graphs extracted from the hourly and daily forecasting model for each of the six households illustrate a comparison between the actual electricity consumption values and the predicted ones. These plots show that the model is able to reasonably follow the overall trend of electricity usage throughout the day. The predicted values generally capture the fluctuations and peak periods of consumption, reflecting a good alignment with the real consumption patterns. Although there are some deviations between the predicted and actual values at certain time intervals, the model demonstrates an initial ability to adapt to the specific consumption behavior of each household. This indicates that even before tuning, the model has the potential to track consumption trends effectively. Further improvements and reduction of prediction errors are expected after applying hyperparameter optimization.

V.5.1 Hyper parameter Optimization for the XGBoost Model

After hyperparameter tuning, the updated hourly prediction graphs for each of the six households (Fig V.10, V.11) show a clear improvement in the model's ability to capture actual electricity consumption patterns. The predicted values are now more closely aligned with the real data, with less pronounced deviations and improved tracking of both peaks and troughs in consumption. This demonstrates the effectiveness of the tuning process in enhancing model accuracy.

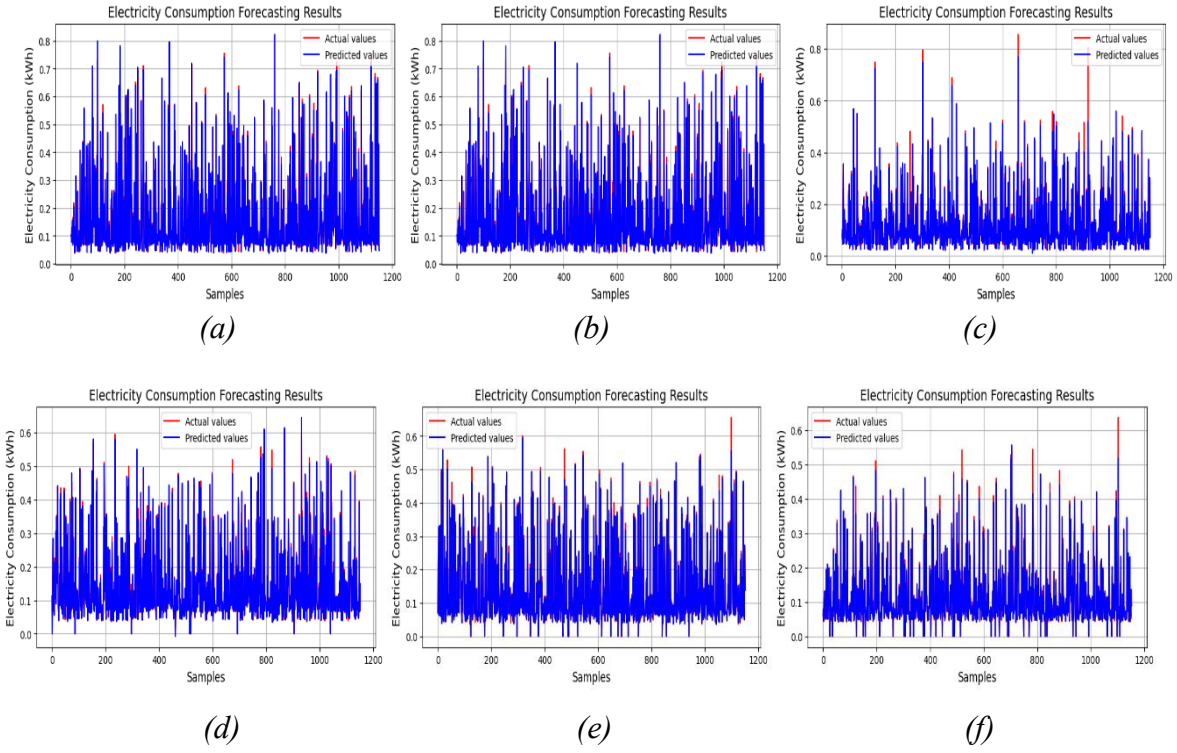


Fig V.10 Household prediction vs original observations after XGBoost hyper-parameters tuning (Hourly granularity) For 6 homes: *(a)* Home 1, *(b)* Home 2, *(c)* Home 3, *(d)* Home 4, *(e)* Home 5, *(f)* Home 6

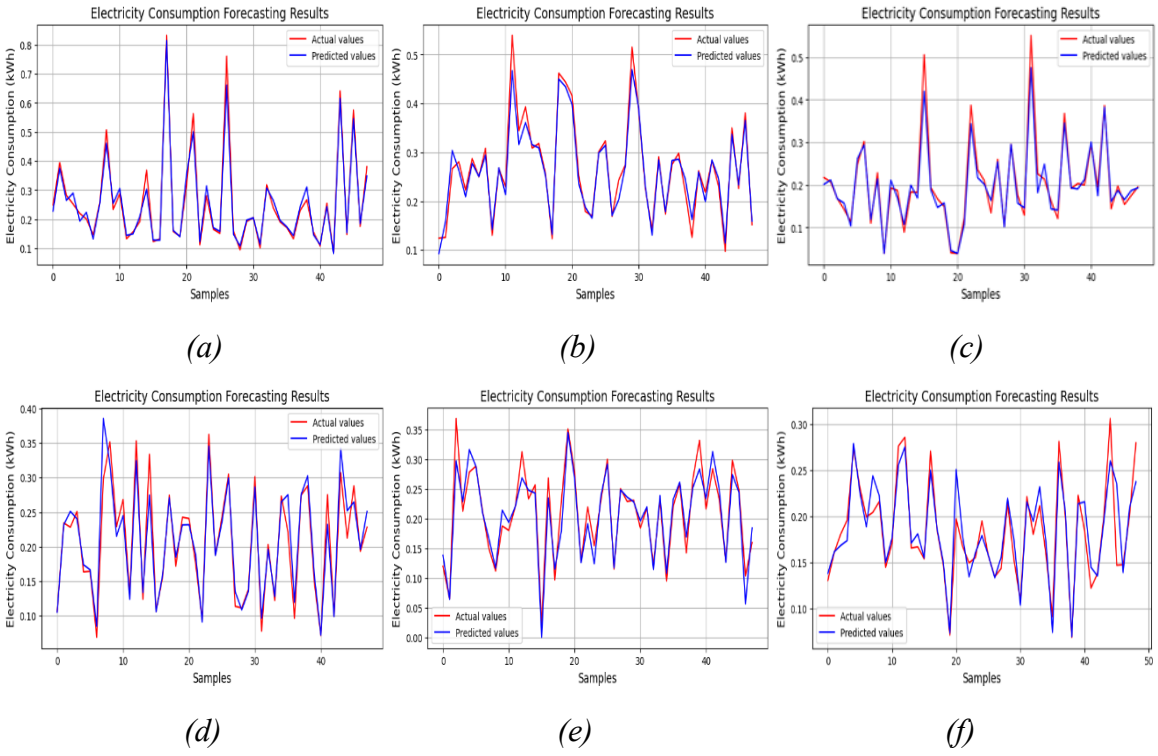


Fig V.11 Household prediction vs original observations XGBoost hyper-parameters tuning (Daily granularity) For 6 homes: *(a)* Home 1, *(b)* Home 2, *(c)* Home 3, *(d)* Home 4, *(e)* Home 5, *(f)* Home 6

To quantify this improvement, Table V.4 summarizes the prediction errors after tuning, using standard evaluation metrics such as Root Mean Squared Error (RMSE) As shown in the table, all six households experienced a noticeable reduction in error values after optimization, further confirming the enhanced performance observed in the visual graphs.

Table V.4 RMSE after XGBoost hyper-parameters tuning

RMSE	Home 1	Home 2	Home 3	Home 4	Home 5	Home 6
Hourly	0.0104	0.0098	0.0163	0.0075	0.0096	0.0096
Daily	0.0270	0.0207	0.0233	0.0235	0.0231	0.0217

Table V.4 presents the Root Mean Squared Error (RMSE) values obtained after hyperparameter tuning for electricity consumption forecasting across six households. The table includes the RMSE results for both hourly and daily predictions, computed separately for each household. These values reflect the forecasting model's accuracy in capturing electricity consumption patterns at different temporal resolutions. The results provide insight into how the model performs when applied individually to multiple households and allow for a quantitative assessment of prediction quality under varying conditions.

V.5.2 Motivation for Multi-Household Forecasting Extension

The transition from forecasting electricity consumption for a single household to modeling six individual households aims to expand the scope of the study toward a more generalized and practical application within smart grid environments. While analyzing a single home allows for detailed modeling and understanding of consumption behavior, it does not fully represent the diversity of usage patterns found across different households. By extending the analysis to six separate homes, the objective is to assess the model's robustness, adaptability, and predictive performance in the face of varying consumption profiles, appliance usage habits, and load dynamics. This broader perspective not only enhances the reliability of the forecasting model but also aligns more closely with real-world scenarios where energy management systems must operate across multiple users. Ultimately, this step supports the development of more scalable and intelligent solutions for household-level demand forecasting in smart grid systems.

V.6 Energy management system

Following the forecasting phase, the study was extended to the design of an energy management system (EMS) for a household connected to four energy sources: solar power, wind energy, a battery storage unit, and the main electrical grid. This setup reflects a typical smart home scenario within a smart grid environment, where local generation and storage capabilities must be optimally coordinated to meet the household's electricity demand. The objective of the EMS is to minimize reliance on the main grid, reduce energy costs, and maximize the use of renewable sources while ensuring supply reliability. The forecasting results serve as a key input to guide the energy allocation strategy across these sources, enabling intelligent and data-driven decision-making in real time.

The core objective of the EMS is to determine, at each time step (daily or hourly), the optimal energy contribution from each source to meet the household demand while minimizing operational costs and maximizing the use of renewable energy. This involves solving a constrained optimization problem where the total energy supplied by solar, wind, battery discharge, and grid purchase must equal the predicted demand, subject to the availability of resources (e.g., solar irradiance, wind speed, battery state of charge) and system constraints (e.g., battery capacity, grid limits, cost tariffs).

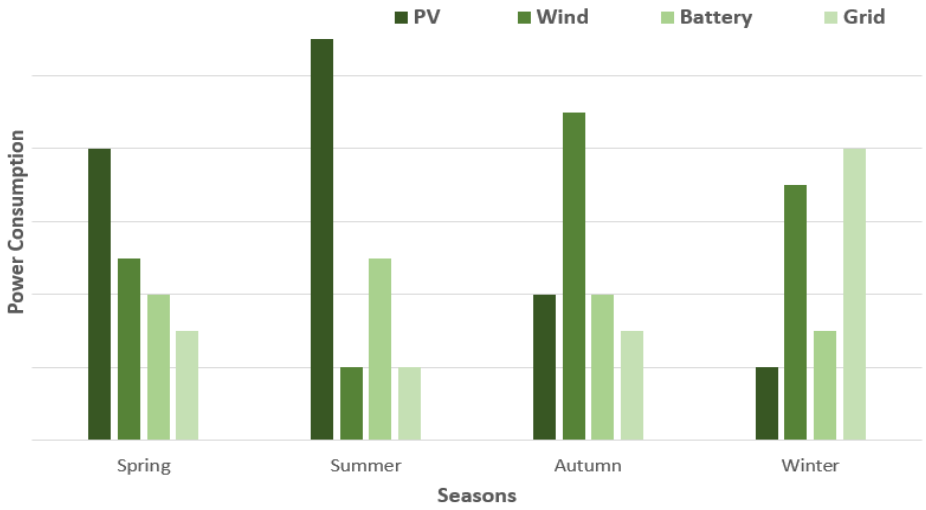


Fig V.12 Seasonal variations in energy sourcing

Fig V.12 illustrates the seasonal variation in energy sourcing, based on typical environmental conditions. As shown, solar energy (PV) dominates in summer, while wind energy is more prevalent in autumn and winter. The battery serves as a buffer in all seasons, storing excess energy and discharging when primary sources are insufficient. Grid energy is used as a last resort, particularly in winter, when renewable sources are limited.

Based on the accurate forecasting results obtained using XGBoost model, several energy management strategies were proposed to optimize the utilization of available resources. Each strategy serves as a high-level plan for intelligently distributing loads and selecting energy sources based on predicted demand. These strategies were then translated into precise and executable algorithms that make automatic decisions regarding which energy source to use at any given time, depending on the time of day, season, and available resources

V.6.1 Single-Source Energy Management

In this energy management approach, the household is connected to only one energy source at a time. This design simplifies the control strategy, minimizes synchronization issues between sources, and reduces hardware complexity. Although multiple energy sources may be available simultaneously (e.g., solar and wind), the system selects the most appropriate single source based on availability, cost, time of day, season, and battery status. The decision-making algorithm ensures that the selected source alone is sufficient to meet the predicted energy demand, thereby avoiding the need for real-time source blending or load balancing.

Algorithm 1: Single-Source Energy Management

Inputs E_pred, E_solar, E_wind, E_battery, E_grid,
season, time_of_day, is_grid_available

Outputs Connected_source

```
If (E_solar ≥ E_pred) and (time_of_day = 'Day') then
   connected_source ← 'Solar'
Else if (E_wind ≥ E_pred) then
   connected_source ← 'Wind'
Else if (E_solar + E_wind ≥ E_pred) then
   connected_source ← 'Hybrid (Solar + Wind)'
Else if (E_battery ≥ E_pred) then
   connected_source ← 'Battery'
Else if is_grid_available and (E_solar + E_wind + E_battery < E_pred) then
   connected_source ← 'Grid'
Else
   connected_source ← 'Load Shedding or Emergency Mode'
End if
```

- This algorithm employs a priority-based strategy to manage energy resources efficiently, giving precedence to renewable and cost-free sources before relying on stored or expensive alternatives. Solar energy is prioritized during daytime hours due to its immediate availability and sustainability. Wind energy is considered a valuable option, particularly at night or during colder seasons such as winter and autumn, when wind activity tends to be stronger. Battery storage is utilized next, but only if sufficient charge remains, in order to preserve battery longevity and avoid deep discharges that could degrade performance. If local renewable and stored sources are insufficient, grid energy serves as a dependable but expensive fallback, ensuring energy demand is still met. Finally, if all other sources fail to cover the required load, the system enters emergency mode, which may involve load shedding or prioritization of critical services to maintain stability.

V.6.2 Multi-Source Energy management

To ensure reliable and sustainable energy supply in a smart grid context, it is important to manage multiple energy sources efficiently. The following algorithm presents a step-by-step logic for powering a household using a hybrid energy management strategy. It considers solar, wind, battery, and grid power sources, and selects the combination of available sources needed to meet the predicted energy demand. Priority is given to renewable sources, while the grid is used only when absolutely necessary. This approach allows the system to dynamically adapt based on time of day, season, and energy availability.

Algorithm 2 : Multi-Source Energy management

Inputs E_pred, E_solar, E_wind, E_battery, E_grid,
season, time_of_day, is_grid_available

Outputs Connected_source

```
Initialize:
   E_needed ← E_pred
   Connected_Sources ← ∅

Step 1: Use Solar Energy
If (time_of_day = 'Day') and (E_solar > 0) then
   If E_solar ≥ E_needed then
      Connected_Sources.Add('Solar')
      E_needed ← 0
   Else
      Connected_Sources.Add('Solar')
      E_needed ← E_needed - E_solar
End if

Step 2: Use Wind Energy
If ((time_of_day = 'Night') or (season ∈ {'Winter', 'Autumn'})) and (E_wind > 0) and
(E_needed > 0) then
   If E_wind ≥ E_needed then
      Connected_Sources.Add('Wind')
      E_needed ← 0
   Else
      Connected_Sources.Add('Wind')
      E_needed ← E_needed - E_wind
End if

Step 3: Use Battery Energy
If (E_battery > 0) and (E_needed > 0) then
   If E_battery ≥ E_needed then
      Connected_Sources.Add('Battery')
      E_needed ← 0
   Else
      Connected_Sources.Add('Battery')
      E_needed ← E_needed - E_battery
End if

Step 4: Use Grid Energy
If is_grid_available and (E_needed > 0) then
   Connected_Sources.Add('Grid')
   E_needed ← 0
End if

Step 5: Emergency Mode
If E_needed > 0 then
   Connected_Sources.Add('Load Shedding or Emergency Mode')
End if
```

The algorithm demonstrates a well-structured and intelligent approach to energy source selection by implementing a clear and logical priority order. It first opts for solar energy during daylight hours, followed by wind energy during the night or in windy seasons, then battery storage, the grid, and finally, an emergency mode as a last resort. This hierarchy reflects a sustainability-driven strategy typical of smart grid systems, emphasizing the use of renewable and cost-effective resources before relying on stored or purchased energy. Moreover, the algorithm supports partial contributions from multiple sources, allowing for flexible energy balancing where shortfalls from one source can be supplemented by others. This ensures realistic and optimized energy utilization. It also exhibits context-aware adaptability by factoring in the time of day, seasonal variations, the real-time availability of each source, and the operational status of the grid. These considerations make the algorithm robust and versatile across varying environmental and system conditions. Lastly, the inclusion of an emergency handling mechanism provides a critical layer of resilience, ensuring continuity of service even in cases where energy demand cannot be fully met, thereby enhancing the system's overall reliability.

V.6.3 Energy Source Switching Control

Effective energy source switching control is a cornerstone in modern hybrid energy systems, where multiple generation and storage options coexist. This algorithm is designed to dynamically allocate energy sources solar, wind, battery, and grid to satisfy the predicted household demand efficiently. It systematically prioritizes the use of renewable energy based on time of day and seasonal availability while preserving battery reserves and minimizing costly grid consumption. By continuously monitoring energy availability and adjusting source connections accordingly, this control strategy enhances energy reliability, reduces operational costs, and supports environmental sustainability. Such intelligent management is crucial for optimizing smart home energy systems in the face of variable renewable energy production and fluctuating consumption patterns**.**

This algorithm exhibits several important strengths that make it well-suited for energy management in a smart grid context. It adopts a clear and logical order for utilizing available energy sources, starting with solar energy during daylight hours, followed by wind energy at night or during colder seasons, then drawing from battery storage, and finally resorting to grid power only if necessary. This priority structure ensures that the most sustainable and cost-effective options are considered first. Additionally, the approach allows for flexible and gradual allocation of energy by continuously updating the remaining demand, enabling

multiple sources to contribute as needed. The inclusion of straightforward control signals to switch each source on or off enhances clarity and simplifies implementation, making the algorithm both practical and adaptable for real-time applications.

Algorithm 3: Energy Source Switching Control

Inputs E_pred, E_solar, E_wind, E_battery, E_grid,
season, time_of_day, is_grid_available

Outputs switch_solar, switch_wind, switch_battery, switch_grid

```
Initialize:
   E_needed ← E_pred
   switch_solar ← 0
   switch_wind ← 0
   switch_battery ← 0
   switch_grid ← 0

Step 1: Use Solar Energy
If (time_of_day = 'Day') and (E_solar > 0) then
   switch_solar ← 1
   If E_solar ≥ E_needed then
      E_needed ← 0
   Else
      E_needed ← E_needed - E_solar
End if

Step 2: Use Wind Energy
If ((time_of_day = 'Night') or (season ∈ {'Winter', 'Autumn'})) and (E_wind > 0) and
(E_needed > 0) then
   switch_wind ← 1
   If E_wind ≥ E_needed then
      E_needed ← 0
   Else
      E_needed ← E_needed - E_wind
End if

Step 3: Use Battery Energy
If (E_battery > 0) and (E_needed > 0) then
   switch_battery ← 1
   If E_battery ≥ E_needed then
      E_needed ← 0
   Else
      E_needed ← E_needed - E_battery
End if

Step 4: Use Grid Energy
If (E_needed > 0) and (is_grid_available = True) then
   switch_grid ← 1
   E_needed ← 0
End if
```

V.6.4 Benefits of Energy Management Strategies

The three proposed algorithms in this study address the efficient management of diverse energy sources to meet the predicted household electrical demand. The first algorithm implements a simple single-source selection based on availability conditions such as time of day and season, prioritizing the use of renewable energy when available. The second algorithm extends this approach to enable simultaneous connection of multiple energy sources, enhancing system flexibility and maximizing renewable energy utilization while reducing dependence on the electrical grid. Finally, the third algorithm introduces a dynamic control mechanism using switching signals to activate or deactivate different energy sources (solar, wind, battery, and grid), optimizing energy management to ensure continuous power supply with minimal cost and maximum reliability. Collectively, these algorithms aim to develop an intelligent home energy management system capable of adapting to demand fluctuations and variable renewable generation, while ensuring both economic and technical sustainability.

V.7 Conclusion

This chapter presents an in-depth evaluation of the proposed electricity consumption forecasting models, beginning with the implementation of an LSTM (Long Short-Term Memory) neural network as a baseline deep learning approach. Although the LSTM model demonstrated reasonable forecasting capabilities, the XGBoost-based model showed superior performance in terms of both accuracy and computational efficiency. Using the IHEPC dataset, several experiments were carried out to assess the model performance across different time resolutions: hourly, daily, and 15-second intervals. The dataset was divided into training and testing sets, and the forecasting experiments were conducted in three progressive phases: 1) Baseline XGBoost Model and 2) Hyper-parameter Tuning.

To address data variability caused by extreme consumption values, normalization and standardization techniques were applied. Normalization, in particular, led to the best results, achieving the lowest RMSE and MAPE (0.129 and 0.020, respectively) for daily predictions.

Finally, the performance of the proposed XGBoost model was compared with several state-of-the-art machine learning and deep learning models from the literature, all of which used the same IHEPC dataset. The proposed model outperformed all referenced approaches across all time resolutions, especially in terms of MAPE and RMSE, highlighting its robustness and reliability.

Building upon these results, the forecasting approach was extended from a single household to six individual households in order to assess the model's scalability and adaptability to diverse consumption patterns.

Moreover, accurate consumption forecasting represents a key enabling component for the design of intelligent energy management systems. The predicted demand values have been effectively used to develop dynamic control algorithms for managing energy sources such as solar, wind, battery, and grid ensuring optimal resource allocation, reduced reliance on the grid, and enhanced overall system efficiency.

Chapter VI

Conclusion and Future Directions

This thesis has addressed one of the most vital and timely topics in the field of smart energy systems: the accurate forecasting of electricity consumption using artificial intelligence. As the global energy landscape continues to evolve driven by sustainability goals, the integration of renewable energy sources, and the digitalization of power systems reliable forecasting models are becoming indispensable for maintaining efficiency, stability, and intelligent energy management.

The research began with the acquisition, preparation, and in-depth preprocessing of real-world household electricity consumption data. This initial step was fundamental in laying the groundwork for a successful prediction model, as it involved careful analysis of multiple energy metrics such as global active power, reactive power, voltage, current intensity, and disaggregated sub-metering for specific household zones (kitchen, laundry, and water heating/cooling systems).

Two advanced machine learning techniques were evaluated Long Short-Term Memory (LSTM) networks, which are widely recognized for their ability to capture temporal dependencies in sequential data, and the XGBoost algorithm, a powerful gradient boosting framework. Although LSTM showed promising results, XGBoost proved to be significantly more efficient, both in terms of training time and prediction accuracy. Thus, the research proceeded with XGBoost as the primary forecasting engine.

Through careful tuning of hyperparameters and model optimization, the XGBoost-based forecasting system delivered highly accurate hourly and daily electricity consumption predictions with remarkably low error rates. The robustness and reliability of the model were further validated by extending the experimental setup from one household to a group of six homes, capturing a broader range of consumption behaviors and usage patterns. This transition

from a single-user to a multi-household scenario not only demonstrated the model's scalability but also enhanced its practical relevance in real-world applications.

Importantly, this thesis did not stop at prediction. In its final stage, it bridged the gap between forecasting and action by proposing a complete energy management framework built upon the model's output. This framework incorporates various energy sources solar panels, wind turbines, battery storage systems, and the conventional electricity grid and is designed to dynamically allocate energy in an optimal manner based on predicted demand. It provides a path toward proactive and data-driven energy distribution, with the potential to minimize energy waste, reduce costs, and improve the reliability and sustainability of energy systems at the household level.

The contributions of this thesis are therefore multifaceted. On the technical side, it introduces a refined and well-validated AI-based forecasting approach. On the practical side, it offers a blueprint for integrating predictive analytics into everyday energy management. From an academic perspective, it lays the groundwork for future research at the intersection of artificial intelligence, renewable energy integration, and smart grid optimization.

In a world where energy challenges are becoming increasingly complex, this research contributes valuable insights and tools to support the transition toward smarter, more sustainable, and more autonomous energy infrastructures. The successful application of AI in this context illustrates not only the power of machine learning in data-rich environments but also its critical role in driving the future of intelligent energy systems.

Building on the outcomes of this thesis, several promising research directions can be pursued to further enhance energy forecasting and management in smart grids using artificial intelligence :

☞ **Real-Time Forecasting and Adaptive Models :** Developing models that can update dynamically with real-time data streams from smart meters and IoT sensors will improve responsiveness to sudden changes in consumption patterns and grid conditions.

☞ **Multi-Scale and Large-Scale Forecasting :** Extending the forecasting framework to cover larger geographical areas, such as neighborhoods, cities, or regions, to support more comprehensive smart grid management and planning.

☞ **Hybrid and Ensemble Learning Approaches :** Investigating hybrid models that combine the strengths of different algorithms (e.g., XGBoost with deep learning architectures

like LSTM or Transformers) to improve prediction accuracy and robustness under diverse conditions.

☞ **Integration with Demand Response and Energy Optimization :** Linking forecasting models with demand response programs and optimization algorithms to enable automated load shifting, peak shaving, and real-time control of energy consumption and generation.

☞ **Incorporation of External Factors :** Including weather data, socio-economic variables, occupancy patterns, and tariff structures into forecasting models to increase prediction accuracy and provide personalized energy management solutions.

☞ **Enhanced Renewable Energy and Storage Management :** Developing advanced strategies for managing renewable energy generation and battery storage, optimizing charge/discharge cycles to maximize efficiency, lifespan, and economic benefits.

☞ **Environmental and Economic Impact Assessment :** Quantifying the environmental benefits (e.g., CO_2 emission reductions) and economic savings of AI-driven energy management systems to support policy-making and incentivize adoption.

☞ **Interoperability and Standardization :** Researching ways to integrate AI-based forecasting systems with existing smart grid standards and platforms to ensure compatibility, scalability, and ease of deployment.

These directions represent a roadmap for evolving the current research into practical, scalable, and impactful smart grid solutions. By addressing these challenges, future work will accelerate the transition toward intelligent, sustainable, and resilient energy systems worldwide.

Appendix A

Scientific Productions

A.1 Publications

1. **Bezzar, N. E-H**., Laimeche, L., Meraoumia, A., & Houam, L. (2022). Data analysis-based time series forecast for managing household electricity consumption. Demonstratio Mathematica, 55(1), 900–921. https://doi.org/10.1515/dema-2022-0176

A.2 International Conferences

1. **Nour E-H, B.,** Lakhdar, L & Abdallah, M. (2022). Time Series Analysis of Household Electric Consumption with XGBoost Model. 4th International Conference on Pattern Analysis and Intelligent Systems (PAIS'2022), Oum El Bouaghi, Algeria. 1–6. https://doi.org/10.1109/pais56586.2022.9946913
2. **Nour E-H, B.,** Abdallah, M & Lotfi, H. (2023). Forecasting Household Energy Consumption Using LSTM Model. 1st International Conference on Advances In Electrical and Computer Engineering (ICAECE'2023), Tebessa, Algeria.
3. **Nour E-H, B.,** Abdallah, M & Lotfi, H. (2023). Energy Management of Hybrid Power System with Storage System (Battery / Supercapacitor). 1st International Conference on Materials for Energy Storage (IC-MES'2023), Algiers, Algeria.

A.3 National Conferences

1. **Nour E-H, B.,** Abdallah, M & Lotfi, H. (2023). Artificial Intelligence Techniques in the Smart Grid: A Review. 1st National Conference on Artificial Intelligence and Its Applications (NCAIA'2023), Telemcen, Algeria.
2. **Nour E-H, B.,** Abdallah, M & Lotfi, H. (2023). A Survey of Challenges and Applications of Artificial Intelligence Techniques in Smart Grid . 1st National Conference on Maintenance and Safety in the Service of The Environment (CNMSSE'2023), Tebessa, Algeria.

3. **Nour E-H, B.,** Abdallah, M & Lotfi, H. (2023). Fuzzy Logic Energy Management Strategy of a Renewable Hybrid Energy System. 2nd National Conference of Materials Sciences and Renewable Energy (Cmsre24), Relizane, Algeria.

Bibliography

[1] A. Ahmad, N. Javaid, A. Mateen, M. Awais, and Z. A. Khan, "Short-term load forecasting in smart grids: An intelligent modular approach," Energies, vol. 12, no. 1, 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/1/164

[2] V. G. Tran, V. DeBusschere, and S. Bacha, one week hourly electricity load forecasting using neuro-fuzzy and seasonal ARIMA models, IFAC Proceeding, 45 (2012), no. 21, 97–102, DOI: https://doi.org/10.3182/20120902-4-FR-2032.00019.

[3] S. M. Molaei, and M. R. Keyvanpour, An analytical review for event prediction system on time series, International Conference on Pattern Recognition and Image Analysis, Rasht, Iran, 2015, pp. 1–6, DOI: https://doi.org/10.1109/PRIA.2015.7161635

[4] R. Wang, S. Lu, and W. Feng, A novel improved model for building energy consumption prediction based on model integration, Appl. Energy 262 (2020), no. 114561, 1–14, DOI: https://doi.org/10.1016/j.apenergy.2020.114561.

[5] G. Shmueli and O. Koppius, "The challenge of prediction in information systems research," SSRN Electronic Journal, 04 2009.

[6] T. Y. Kim and S. B. Cho, predicting residential energy consumption using CNN-LSTM neural networks, Energy J. 182 (2019), 72–81, DOI: https://doi.org/10.1016/j.energy.2019.05.230.

[7] T. Liu, Z. Tan, C. Xu, H Chen, and Z. Li, Study on deep reinforcement learning techniques for building energy consumption forecasting, Energy Buildings 208 (2020), 109675, DOI: https://doi.org/10.1016/j.enbuild.2019.109675.

[8] M. A., Dozdar, M. H. Masoud, and J. M. Ramadhan, A Review on Deep Sequential Models for Forecasting Time Series Data, Appl. Comput. Intell. Soft Comput. 2022 (2022), 1–19, DOI: https://doi.org/10.1155/2022/6596397.

[9] International Energy Outlook 2013. Available: http://www.eia.gov/forecasts/archive/aeo13/

[10] United States Department of Energy. Final Report on The August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations. Report, U.S.-Canada Power System Outage Task Force, April 2014. URL https://www.ferc.gov/industries/electric/indus-act/reliability/blackout/ch1-3.pdf.

[11] Km Ravi Eswar. Smart Grid-Future for Electrical Systems. International Journal of Electrical and Electronics Research Vol, 3, 2015.

[12] Xia S, Luo X, Chan KW. A Framework for Self-Healing Smart Grid with Incorporation of Multi-Agents. December 2014, Energy Procedia 61 :2123-2126, DOI: 10.1016/J.EGYPRO.2014.12.090

[13] Osama Majeed Butt, Et Al. Recent Advancement in Smart Grid Technology: Future Prospects in The Electrical Power Network. Ain Shams Engineering Journal 12 (2021) 687–695, https://doi.org/10.1016/j.asej.2020.05.004.

[14] Von Dollen D. Report to Nist on the Smart Grid Interoperability Standards Roadmap. Electric Power Res Inst (EPRI) Nat Inst Stand Technol 2009.

[15] Logenthiran T, Srinivasan D, Shun TZ. Demand Side Management in Smart Grid Using Heuristic Optimisation. IEEE Trans Smart Grid 2012;3(3):1244–52.

[16] Koutsopoulos I, Tassiulas L. Challenges in Demand Load Control for the Smart Grid. IEEE Netw 2011;25(25):16–21.

[17] Strbac G. Demand Side Management: Benefits and Challenges. Energy Policy 2008;36(12):4419–26.

[18] Khodayar Me, Wu H. Demand Forecasting in The Smart Grid Paradigm: Features and Challenges. Electr J 2015;28(6):51–62.

[19] Saad W, Han Z, Poor HV, Basar T. Game-Theoretic Methods for the Smart Grid: An Overview of Microgrid Systems, Demand-Side Management, And Smart Grid Communications. IEEE Signal Process Mag 2012;29(5):86–105.

[20] Kim JH, Shcherbakova A. Common Failures of Demand Response. Energy 2011;36(2):873–80.

[21] Spees K, Lave LB. Demand Response and Electricity Market Efficiency. Electr J 2007;20(3):69–85.

[22] Wang J, Bloyd CN, Hu Z, Tan Z. Demand Response in China. Energy 2010;35 (4):1592–7.

[23] Ming Z, Li S, Yanying H. Status, Challenges and Countermeasures of Demand side Management Development in China. Renew Sustain Energy Rev 2015; 47:284–94

[24] Leysen R. An Analysis of Smart Meter Deployment in Sweden with Applicability to the Case of India; 2018.

[25] Hart DG. Using Ami to Realize the Smart Grid. Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century IEEE 2008; 2008:1–2.

[26] Hart DG. How Advanced Metering Can Contribute to Distribution Automation. IEEE Smart Grid 2012.

[27] Eltamaly A.M., Mohamed Y.S., El-Sayed AH.M., Elghaffar A.N.A., Abo-Khalil A.G. D-STATCOM for Distribution Network Compensation Linked with Wind Generation. In: Eltamaly A.M., Abdelaziz A.Y., Abo-Khalil A.G. (eds) Control and Operation of Grid-Connected Wind Energy Systems. Green Energy and Technology. Springer, Cham 2021. https://doi.org/10.1007/978-3-030-64336-2_5

[28] M. Eltamaly, A., Sayed Mohamed, Y., M. El-Sayed, AH. Et Al. Power Quality and Reliability Considerations of Photovoltaic Distributed Generation. Technol Econ Smart Grids Sustain Energy 5, 25 (2020). https://Doi.Org/10.1007/S40866020-00096-2

[29] Eltamaly A, Elghaffar ANA. Techno-Economical Study of Using Nuclear Power Plants For Supporting Electrical Grid in Arabian Gulf. Technol Econ Smart Grids Sustain Energy J. https://Doi.Org/10.1007/S40866-017-0031-8

[30] Abou-Hashema M. El-Sayed, Et Al. Analyzing of Wind Distributed Generation Configuration in Active Distribution Network. The 8th (Icmsao'2019) Conference, Bahrain. 04/2019. DOI: 10.1109/ICMSAO.2019.8880291

[31] Eltamaly A, Amer Nasr A. Elghaffar, Et Al: AC Microgrid Protection Coordination. Chapter (8), Wiley Book (Microgrid Technologies), eBOOK 2021, https://doi.org/10.1002/9781119710905.ch8

[32] Muttaqi KM, Aghaei J, Ganapathy V, Esmaeel A. Technical challenges for electric power industries with implementation of distribution system automation in smart grids. Renew Sustain http://dx.doi.org/10.1016/j.rser.2015.02.013.

[33] Negin Alemazkoor, Mazdak Tootkaboni, Roshanak Nateghi, & Arghavan Louhghalam. (2022). Smart-Meter Big Data for Load Forecasting: An Alternative Approach to Clustering. IEEE Access, 10, 8377–8387. https://doi.org/10.1109/access.2022.3142680

[34] Hany Habbak, Mahmoud, M. E., Khaled Metwally, Fouda, M. M., & Ibrahem, M. I. (2023). Load Forecasting Techniques and Their Applications in Smart Grids. 16(3), 1480–1480. https://doi.org/10.3390/en16031480

[35] Franco, P., Martinez, J. M., Kim, Y.-C., & Ahmed, M. A. (2021). A Framework for IoT Based Appliance Recognition in Smart Homes. IEEE Access, 9, 133940–133960. https://doi.org/10.1109/access.2021.3116148

[36] Lee, M., Ser, Y., Selvachandran, G., Thong, P., Cuong, L., Son, L., Tuan, N., & Gerogiannis, V. (2022). A Comparative Study of Forecasting Electricity Consumption Using Machine Learning Models. Mathematics, 10(8), 1329. https://doi.org/10.3390/math10081329

[37] Syed, D., Abu-Rub, H., Ghrayeb, A., & Refaat, S. S. (2021). Household-Level Energy Forecasting in Smart Buildings Using a Novel Hybrid Deep Learning Model. IEEE Access, 9, 33498–33511. https://doi.org/10.1109/access.2021.3061370

[38] Alduailij, M. A., Petri, I., Rana, O., & Aldawood, A. S. (2020). Forecasting peak energy demand for smart buildings. The Journal of Supercomputing, 77(6), 6356–6380. https://doi.org/10.1007/s11227-020-03540-3

[39] Alquthami, T., Zulfiqar, M., Kamran, M., Milyani, A. H., & Rasheed, M. B. (2022). A Performance Comparison of Machine Learning Algorithms for Load Forecasting in Smart Grid. IEEE Access, 10, 48419–48433. https://doi.org/10.1109/access.2022.3171270

[40] Chen, J., Deng, G., Zhang, L., & Ahmadpour, A. (2022). Demand side energy management for smart homes using a novel learning technique–economic analysis aspects. Sustainable Energy Technologies and Assessments, 52, 102023. https://doi.org/10.1016/j.seta.2022.102023

[41] Klyuev, R. V., Morgoev, I. D., Morgoeva, A. D., Gavrina, O. A., Martyushev, N. V., Efremenkov, E. A., & Mengxu, Q. (2022). Methods of Forecasting Electric Energy Consumption: A Literature Review. Energies, 15(23), 8919. https://doi.org/10.3390/en15238919

[42] Hernández, J. L., de Miguel, I., Vélez, F., & Vasallo, A. (2024). Challenges and opportunities in European smart buildings energy management: A critical review. Renewable and Sustainable Energy Reviews, 199, 114472. https://doi.org/10.1016/j.rser.2024.114472

[43] Taheri, S., Hosseini, P., & Razban, A. (2022). Model predictive control of heating, ventilation, and air conditioning (HVAC) systems: A state-of-the-art review. Journal of Building Engineering, 60, 105067. https://doi.org/10.1016/j.jobe.2022.105067

[44] Zhao, X., Gao, W., Qian, F., & Ge, J. (2021). Electricity cost comparison of dynamic pricing model based on load forecasting in home energy management system. Energy, 229, 120538. https://doi.org/10.1016/j.energy.2021.120538

[45] Saxena, A., Shankar, R., El-Saadany, E., Kumar, M., Zaabi, O. A., Hosani, K. A., & Muduli, U. R. (2024). Intelligent Load Forecasting and Renewable Energy Integration for Enhanced Grid Reliability. IEEE Transactions on Industry Applications, 1–15. https://doi.org/10.1109/tia.2024.3436471

[46] Abdul Khalique Shaikh, Nazir, A., Khan, I., & Abdul Salam Shah. (2022). Short term energy consumption forecasting using neural basis expansion analysis for interpretable time series. Scientific reports, 12(1). https://doi.org/10.1038/s41598-022-26499-y

[47] Chataut, R., Phoummalayvane, A., & Akl, R. (2023). Unleashing the Power of IoT: A Comprehensive Review of IoT Applications and Future Prospects in Healthcare, Agriculture, Smart Homes, Smart Cities, and Industry 4.0. Sensors, 23(16), 7194. https://doi.org/10.3390/s23167194

[48] Runge, J., & Zmeureanu, R. (2021). A Review of Deep Learning Techniques for Forecasting Energy Use in Buildings. Energies, 14(3), 608. https://doi.org/10.3390/en14030608

[49] Montero-Manso, P., & Hyndman, R. J. (2021). Principles and algorithms for forecasting groups of time series: Locality and globality. International Journal of Forecasting. https://doi.org/10.1016/j.ijforecast.2021.03.004

[50] Wang, J., Chen, X., Zhang, F., Chen, F., & Xin, Y. (2021). Building Load Forecasting Using Deep Neural Network with Efficient Feature Fusion. Journal of Modern Power Systems and Clean Energy, 9(1), 160–169. https://doi.org/10.35833/mpce.2020.000321

[51] Jiang, Y., Gao, T., Dai, Y., Si, R., Hao, J., Zhang, J., & Gao, D. W. (2022). Very short-term residential load forecasting based on deep-autoformer. Applied Energy, 328, 120120. https://doi.org/10.1016/j.apenergy.2022.120120

[52] Barrera-Animas, A. Y., Oyedele, L. O., Bilal, M., Akinosho, T. D., Delgado, J. M. D., & Akanbi, L. A. (2021). Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting. Machine Learning with Applications, 7, 100204. https://doi.org/10.1016/j.mlwa.2021.100204

[53] Schaffer, A. L., Dobbins, T. A., & Pearson, S.-A. (2021). Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. BMC Medical Research Methodology, 21(1). https://doi.org/10.1186/s12874-021-01235-8

[54] Hewamalage, H., Ackermann, K., & Bergmeir, C. (2022). Forecast evaluation for data scientists: common pitfalls and best practices. Data Mining and Knowledge Discovery. https://doi.org/10.1007/s10618-022-00894-5

[55] Chen, Z., Chen, D., Zhang, X., Yuan, Z., & Cheng, X. (2021). Learning Graph Structures with Transformer for Multivariate Time Series Anomaly Detection in IoT. IEEE Internet of Things Journal, 1–1. https://doi.org/10.1109/jiot.2021.3100509

[56] Ngo, N.-T., Pham, A.-D., Truong, T. T. H., Truong, N.-S., & Huynh, N.-T. (2022). Developing a hybrid time-series artificial intelligence model to forecast energy use in buildings. Scientific Reports, 12(1), 15775. https://doi.org/10.1038/s41598-022-19935-6

[57] Fatehijananloo, M., Stopps, H., & McArthur, J. J. (2024). Exploring artificial intelligence methods for energy prediction in healthcare Facilities: An In-Depth extended systematic review. Energy and Buildings, 320, 114598. https://doi.org/10.1016/j.enbuild.2024.114598

[58] Liu, Z., Wang, Y., Wang, Q., & Hu, M. (2025). Vision Transformer-Based Anomaly Detection in Smart Grid Phasor Measurement Units Using Deep Learning Models. IEEE Access, 13, 44565–44576. https://doi.org/10.1109/access.2025.3549679

[59] Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2020). Deep Learning for Time Series Forecasting: A Survey. Big Data, 9(1). https://doi.org/10.1089/big.2020.0159

[60] Shapi, M. K. M., Ramli, N. A., & Awalin, L. J. (2020). Energy Consumption Prediction by Using Machine Learning for Smart Building: Case Study in Malaysia. Developments in the Built Environment, 5, 100037. https://doi.org/10.1016/j.dibe.2020.100037

[61] Magalhães, B., Bento, P., Pombo, J., Calado, M. do R., & Mariano, S. (2024). Short-Term Load Forecasting Based on Optimized Random Forest and Optimal Feature Selection. Energies, 17(8), 1926. https://doi.org/10.3390/en17081926

[62] Tien, P. W., Wei, S., Darkwa, J., Wood, C., & Calautit, J. K. (2022). Machine Learning and Deep Learning Methods for Enhancing Building Energy Efficiency and Indoor Environmental Quality – A Review. Energy and AI, 10, 100198. https://doi.org/10.1016/j.egyai.2022.100198

[63] Bai, S., J. Zico Kolter, & Koltun, V. (2018). An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.1803.01271

[64] Su, L., Zuo, X., Li, R., Wang, X., Zhao, H., & Huang, B. (2025). A systematic review for transformer-based long-term series forecasting. Artificial Intelligence Review, 58(3). https://doi.org/10.1007/s10462-024-11044-2

[65] Rodrigues, F., Cardeira, C., Calado, J. M. F., & Melicio, R. (2023). Short-Term Load Forecasting of Electricity Demand for the Residential Sector Based on Modelling Techniques: A Systematic Review. Energies, 16(10), 4098. https://doi.org/10.3390/en16104098

[66] Pablo-Andrés Buestán-Andrade, Santos, M., Jesús-Enrique Sierra-García, & Juan-Pablo Pazmiño-Piedra. (2023). Comparison of LSTM, GRU and Transformer Neural Network Architecture for Prediction of Wind Turbine Variables. Lecture Notes in Networks and Systems, 334–343. https://doi.org/10.1007/978-3-031-42536-3_32

[67] L.M.Candanedo,V.Feldheim,andD.Deramaix,Datadrivenpredictionmodelsofenergyuseofappliances inalow-energy house, Energy Buildings 140 (2017),81–97, DOI: https://doi.org/10.1016/j.enbuild.2017.01.083.

[68] D. Wu, B. Wang, D. Precup, and B. J. Boulet, Multiple Kernel learning based transfer regression for electric load forecasting, IEEE Trans Smart Grid 11 (2019), no. 2, 1183–1192, DOI: https://doi.org/10.1109/TSG.2019.2933413.

[69] V. Puspita and Ermatita, Time series forecasting for electricity consumption using kernel principal component analysis (KPCA) and support vector machine (SVM), J. Phys. Conf. Series 1196 (2019), no. 012073, 1–8, DOI: https://doi.org/10.1088/1742-6596/1196/1/012073.

[70] J. Moon, Y. Kim, M. Son, E. Hwang, Hybrid short-term load forecasting scheme using random forest and multilayer perceptron, Energies 11 (2018), no. 12, 3283, 1–20, DOI: https://doi.org/10.3390/en11123283.

[71] R Rajabi and A. Estebsari, Deep learning based forecasting of individual residential loads using recurrence plots, IEEE Milan Power Tech. 2019, pp. 1–5, DOI: https://doi.org/10.1109/PTC.2019.8810899.

[72] X.Tang,Y. Dai, T. Wang,andY.Chen,Short-termpowerloadforecasting basedonmultilayer bidirectional recurrent neural network, IET Gener. Transm. Distrib. 13 (2019), no. 17, 3847–3854, DOI: https://doi.org/10.1049/iet-gtd.2018.6687.

[73] M. Sajjad, Z. A. Khan, A. Ullah, T. Hussain, W. Ullah, M. Y. Lee, et al., A novel CNN-GRU-based hybrid approach for short term residential load forecasting, IEEE Access, 8 (2020), 143759–143768, DOI: https://doi.org/10.1109/ACCESS.2020. 3009537.

[74] N. Ayub, M. Irfan, M. Awais, U. Ali, T. Ali, M. Hamdi, et al., Big data analytics for short and medium-term electricity load forecasting using an AI techniques ensembler, Energies 13 (2020), no. 19, 5193, 1–21, DOI: https://doi.org/10.3390/en13195193.

[75] E. E. Elattar, N. A. Sabiha, M. Alsharef, M. K. Metwaly, A. M. Abd-Elhady, and I. B. M. Taha, Short term electric load forecasting using hybrid algorithm for smart cities, Appl. Intell. 50 (2020), 3379–3399, DOI: https://doi.org/10.1007/ s10489-020-01728-x.

[76] Z. A. Khan, T. Hussain, A. Ullah, S. Rho, M. Lee, and S. W. Baik, Towards efficient electricity forecasting in residential and commercial buildings: A novel hybrid CNN with a LSTM-AE based framework, Sensors, 20 (2020), no. 5, 1399, 1–16, DOI: https://doi.org/10.3390/s20051399.

[77] Czétány, L., Vámos, V., Horváth, M., Szalay, Z., Mota-Babiloni, A., Deme-Bélafi, Z., & Csoknyai, T. (2021). Development of electricity consumption profiles of residential buildings based on smart meter data clustering. Energy and Buildings, 252, 111376. https://doi.org/10.1016/j.enbuild.2021.111376

[78] Albouq, S. S., Sen, A. A. A., Almashf, N., Yamin, M., Alshanqiti, A., & Bahbouh, N. M. (2022). A Survey of Interoperability Challenges and Solutions for Dealing With Them in IoT Environment. IEEE Access, 10, 36416–36428. https://doi.org/10.1109/ACCESS.2022.3162219

[79] Douha, N. Y.-R., Bhuyan, M., Kashihara, S., Fall, D., Taenaka, Y., & Kadobayashi, Y. (2022). A survey on blockchain, SDN and NFV for the smart-home security. Internet of Things, 100588. https://doi.org/10.1016/j.iot.2022.100588

[80] Maghsoudi Nia, E., Qian, Q. K., & Visscher, H. J. (2022). Analysis of Occupant Behaviours in Energy Efficiency Retrofitting Projects. Land, 11(11), 1944. https://doi.org/10.3390/land11111944

[81] Pandiyan, P., Saravanan, S., Usha, K., Kannadasan, R., Alsharif, M. H., & Kim, M.-K. (2023). Technological advancements toward smart energy management in smart cities. Energy Reports, 10, 648–677. https://doi.org/10.1016/j.egyr.2023.07.021

[82] Ammari, C., Belatrache, D., Touhami, B., & Makhloufi, S. (2021). Sizing, optimization, control and energy management of hybrid renewable energy system- a review. Energy and Built Environment, 3(4). https://doi.org/10.1016/j.enbenv.2021.04.002

[83] Hu, Y., Cheng, X., Wang, S., Chen, J., Zhao, T., & Dai, E. (2022). Times series forecasting for urban building energy consumption based on graph convolutional network. Applied Energy, 307, 118231. https://doi.org/10.1016/j.apenergy.2021.118231

[84] Huang, X., Zhang, D., & Zhang, X. (2020). Energy management of intelligent building based on deep reinforced learning. Alexandria Engineering Journal. https://doi.org/10.1016/j.aej.2020.11.005

[85] Rehman, A. U., Hafeez, G., Albogamy, F. R., Wadud, Z., Ali, F., Khan, I., Rukh, G., & Khan, S. (2021). An Efficient Energy Management in Smart Grid Considering Demand Response Program and Renewable Energy Sources. IEEE Access, 9, 148821–148844. https://doi.org/10.1109/access.2021.3124557

[86] Jasim, A. M., Jasim, B. H., Flah, A., Bolshev, V., & Mihet-Popa, L. (2023). A new optimized demand management system for smart grid-based residential buildings adopting renewable and storage energies. Energy Reports, 9, 4018–4035. https://doi.org/10.1016/j.egyr.2023.03.038

[87] Biswas, P., Rashid, A., Al Masum, A., Nasim, M. A. A., Anas Ferdous, A. S. M., Gupta, K. D., & Biswas, A. (2025). An Extensive and Methodical Review of Smart Grids for Sustainable Energy Management-Addressing Challenges with AI, Renewable Energy Integration and Leading-edge Technologies. IEEE Access, 1–1. https://doi.org/10.1109/access.2025.3537651

[88] Shah, S. F. A., Iqbal, M., Aziz, Z., Rana, T. A., Khalid, A., Cheah, Y.-N., & Arif, M. (2022). The Role of Machine Learning and the Internet of Things in Smart Buildings for Energy Efficiency. Applied Sciences, 12(15), 7882. https://doi.org/10.3390/app12157882

[89] Chen, Y., Guo, M., Chen, Z., Chen, Z., & Ji, Y. (2022). Physical energy and data-driven models in building energy prediction: A review. Energy Reports, 8, 2656–2671. https://doi.org/10.1016/j.egyr.2022.01.162

[90] Panda, S., Mohanty, S., Rout, P. K., Sahu, B. K., Bajaj, M., Zawbaa, H. M., & Kamel, S. (2022). Residential Demand Side Management model, optimization and future perspective: A review. Energy Reports, 8, 3727–3766. https://doi.org/10.1016/j.egyr.2022.02.300

[91] G. Mahalakshmi, S. Sridevi and S. Rajaram, A survey on forecasting of time series data, International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), Kovilpatti, India, 2016, pp. 1–8, DOI: https://doi. org/10.1109/ICCTIDE.2016.7725358.

[92] G. Omkar and S. V. Kumar, Time series decomposition model for traffic flow forecasting in urban midblock sections, International Conference On Smart Technologies for Smart Nation (SmartTechCon), Bengaluru, India, 2017, pp. 720–723, DOI: https://doi.org/10.1109/SmartTechCon.2017.8358465.

[93] J. S. Armstrong, Principles of Forecasting: A Handbook for Researchers and Practitioners, Kluwer Academic Publishers, Boston, USA, 2001.

[94] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, John Wiley and Sons, Hoboken, New Jersey, U.S., 2015.

[95] E. B. Dagum and S. Bianconcini, Seasonal Adjustment Methods and Real Time Trend-cycle Estimation, Springer, Cham, 2016.

[96] P. Montero-Manso, G. Athanasopoulos, R. J. Hyndman, and T. S. Talagala, FFORMA: Feature-based forecast model averaging, Int J Forecasting 36 (2020), no. 1, 86–92, DOI: https://doi.org/10.1016/j.ijforecast.2019.02.011.

[97] Hochreiter, S.; Schmidhuber, J. Long short-term memory. Neural Comput. 1997, 9, 1735–1780.

[98] G´eron, A.: Hands-On Machine Learning with Scikit-Learn and TensorFlow, vol. 1. OReilly Media, 1005 Gravenstein Highway North, Sebastopol, CA 95472, 1 edn. (7 2017).

[99] J. Contreras, R. Espinola, F. J. Nogales and A. J. Conejo, ARIMA models to predict next-day electricity prices, IEEE Trans Power Syst. 18 (2003), no. 3, 1014–1020, DOI: https://doi.org/10.1109/TPWRS.2002.804943.

[100] A. Botchkarev, A new typology design of performance metrics to measure errors in machine learning regression algorithms, Interdiscip. J. Inf. Knowl. Manag. 14 (2019), 045–076, DOI: https://doi.org/10.28945/4184.

[101] B. Mathieu, Q. Z. Xiao, N. Elyes, G. Xiaofeng, and C. Patrice, Modeling and forecasting building energy consumption: A review of data-driven techniques, Sustainable Cities Society 48 (2019), no. 101533, 1–27, DOI: https://doi.org/10.1016/ j.scs.2019.101533.

[102] C. Reyes, T. Hilaire, S. Paul and C. F. Mecklenbräuker, Evaluation of the root mean square error performance of the PAST Consensus algorithm, International ITG Workshop on Smart Antennas (WSA), Bremen, Germany, 2010, pp. 156–160, DOI: https://doi.org/10.1109/WSA.2010.5456452.

[103] UCI repository of machine learning database [Online], Available: http://archive.ics.uci.edu/ml/datasets/Individual +household+electric+power+consumption.